CAUTION: THESE RECORDS WILL BE USED FOR OFFICIAL PURPOSES ONLY, DO NOT REMOVE PAPERS NOR REVEAL CONTENTS TO UNAUTHORIZED PERSON(S)

RECORDS CHARGE-OUT

10211

DATE OF REQUEST	SUSPENSE DATE
25 Jan 61	10 Feb61

FILE OR SERIAL NUMBER AND SUBJECT	From File of Special Consultant (Friedman) Military Cryptanalysis, Part IV Serial No. 316.1		
TO	NAME AND EXTENSION OF PERSON REQUESTING FILE Mr. William Friedman LI6-8520	ORGANIZATION, BUILDING, AND ROOM NUMBER 310 2nd Str,SE,Wash,D.C.	
RETURN TO	Mrs. Christian,AG-24,NSA,Ft.Geo.G.Meade,Md.	DATE RET'ND.	INITIAL HERE
INSTRUCTIONS	WHEN TRANSFERRING FILE TO ANOTHER PERSON, COMPLETE SELF-ADDRESSED TRANSFER COUPON BELOW, DETACH, STITCH TO BLANK LETTER-SIZE PAPER AND PLACE IN OUT-GOING MAIL SERVICE.		

10211

2ND TRANSFER COUPON

TO:

FILE (serial number and subject)

TRANSFERRED TO: (name and extension)

ORGANIZATION, BUILDING, AND ROOM NUMBER

DATE	(sig)	(ext.)

Declassified and approved for release by NSA on 01-29-2014 pursuant to E.O. 13526

REF ID:A59439

WAR DEPARTMENT

OFFICE OF THE CHIEF SIGNAL OFFICER
WASHINGTON

MILITARY CRYPTANALYSIS

PART IV

316.1

record taken from WFF's home

REF ID:A59439

30 April 1959

This document is re-graded "~~CONFIDENTIAL~~" UP of DOD Directive 5200.1 dated 8 July 1957, and by authority of the Director, National Security Agency.

Paul S. Willard
Colonel, AGC
Adjutant General

REF ID:A694439

~~CONFIDENTIAL~~

~~*Restricted*~~

WAR DEPARTMENT
OFFICE OF THE CHIEF SIGNAL OFFICER
WASHINGTON

MILITARY CRYPTANALYSIS

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

CONFIDENTIAL

REF ID:A59439

The Golden Guess
Is Morning-Star to the full round of Truth.
—*Tennyson.*

REF ID:A6194382

MILITARY CRYPTANALYSIS, PART IV.

TRANSPOSITION AND FRACTIONATING SYSTEMS

CONTENTS

Section		Paragraphs	Pages
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	40–79
VI.	Principles of matrix reconstruction	30–31	80–84
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

REF ID:A594382

SECTION I

GENERAL

	Paragraph
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 reassembling 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 reconstructing 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,[1] 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.

RET ID: A59439

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.

e. Although transposition ciphers, as a general rule, are much less complex in their mechanics than are 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 mere 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.

2. 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) removing the 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 inscribing 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 that of *transcription.* Either or both processes may be repetitive, by prearrangement of course, in which case the intermediate steps may be referred to as processes of *rescription,* or *rescriptive* 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.

3. 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 *monophase transposition,* commonly called *single transposition.* When one or more rescriptive 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, especially when the successive transpositions are theoretically correct in principle.

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

SECTION II

SOLUTION OF SIMPLE TRANSPOSITION CIPHERS

	Paragraph
Simple types of transposition	4
The principles of solution of uniliteral route-transposition ciphers	5
Keyed columnar transposition with completely-filled rectangles	6
Example of solution	7
The probable-word method of solution	8
General remarks on solution	9
Reconstruction of literal key	10
Column and row transposition	11

4. Simple types of transposition.—*a.* The simple cases of reversed writing, vertical writing, or rail fence writing hardly require serious attention, since they may be solved almost by inspection. These methods are included here only because they may be encountered in censorship operations.

b. The low degree of cryptographic security afforded by these methods may be increased to a slight degree by adding nulls or by disguising the original word lengths, and regrouping into false words or into groups of regular length.

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

(2) Reversing only the words and regrouping into false word lengths:
Cipher... E G D I R B D E Y O R T S E D T A N E V E L E M P

(3) Reversing the whole text and regrouping into fives:
Cipher... M P N E V E L E T A D E Y O R T S E D E G D I R B

(4) Reversing the whole text, regrouping into fives, and inserting a null in every fifth position:
Cipher... T R I M M P N E V P E L E T A A D E Y R O R T S L
E D E G U D I R B M

(5) 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	B R
R T	I D
I R	G E
D O	D E
G Y	S T
E E	R O
D D	Y E
E	D

B S R T I R D O G Y E E D D E , *or*
B I G D S R Y D R D E E T O E

These simple types can be solved merely by inspection.

5. The principles of solution of uniliteral route-transposition ciphers.—*a.* The so-called uniliteral 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-

REF ID:A59439

angles, of various dimensions suggested by the total number of letters in the message, then nspecting these rectangles, searching for whole words or the fragments of words by reading horizontally, diagonally, vertically, spirally, and so on.[1] (See Special Text No. 165, *Elementary Military Cryptography*, 1935, pars. 20, 21.)

b. The amount of experimentation that must be performed in the solution of ciphers of this type may be materially shortened by means of formulae and tables constructed for the purpose. But because ciphers of this type are of infrequent occurrence today, these formulae and tables are only occasionally useful and hence they have not been included in this text.[2]

6. Keyed columnar transposition with completely-filled rectangles.—*a.* 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 less than 5 letters or more than 25 letters will not commonly be encountered. If the width, that is, the number of columns, 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 without too many errors, it is a common 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*.

b. 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.[3] Consequently, this will the constitute 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 Belshazzar and his dynasty (Daniel V: 1–28). The following partial account of the episode is not as enlightening as one might wish, but it is probably the best explanation available. It is taken from Dr. Max Seligsohn's article on the subject in *The Jewish Encyclopedia*, vol. 8, pp. 490–491 (1925): "MENE, MENE, TEKEL, UPHARSIN (מנא מנא תקל ופרסין) Words written by a mysterious hand on the walls of Belshazzar's palace, and interpreted by Daniel as predicting the doom of the King and his dynasty. The incident is described as follows: Once when King Belshazzar was banqueting with his lords and drinking wine from the golden vessels of the temple of YHWH, a man's hand was seen writing on the wall certain mysterious words. Frightened by the apparition, the King ordered his astrologers to explain the inscription; but they were unable to read it. Daniel was then summoned to the Royal Palace, and the King promised him costly presents if he could decipher the inscription. Daniel read it "Mene, mene, tekel, upharsin," and explained it to mean that God had "numbered" the Kingdom of Belshazzar and brought it to an end; that the King had been weighed and found wanting; and that his Kingdom was divided and given to the Medes and Persians.

The first question which presents itself to the critic, namely, why could the inscription be deciphered by Daniel only—engaged the attention of the Talmudists, who advanced various answers. Certain of them concluded that the Hebrew writing had been changed in the time of Ezra, so that even the Jews that were found in the royal court could not read an inscription written in archaic characters. But those who followed R. Simeon in maintaining that the writing had not changed found other solutions for the problem; *e. g.* it was written in the cryptographic combination את בש, each letter of each pair being substituted by its companion, *e. g.* ימת ימת אדך פרג חמם; or the words were written thus: ממתנס נעקפי אאלרן, one above the other, having to be read vertically; or again, אנם אנם לקח ניטרפו, each word backward; or again, נמא נמא קתל פורסין the first two letters of each word being transposed (Sanh. 22a). It is evident that the author of the Book of Daniel meant that the inscription was written in characters familiar to the King and wise men of Babylon, but that, as often happens with ancient inscriptions, the transposition of certain letters baffled every attempt to decipher them. . . ."

[2] See Lohr, Lenox R. and Friedman, William F., *Formulae for the solution of transposition ciphers*. Riverbank Publication No. 19, Geneva, Illinois, 1918.

[3] See Special Text No. 165, *Elementary Military Cryptography*, 1935, Sec. V. In this text the term "transposition rectangle" will be used to designate the matrix, frame, cage, or design regardless of whether the latter is completely filled or not.

c. In solving a cryptogram of this type the first step taken by the cryptanalyst is to ascertain the dimensions of the rectangle. Clues for this are usually afforded by finding the factors of the total number of letters in the cryptogram. Suppose the cryptogram contains 152 letters. The dimensions of the transposition rectangle may be 4×38 or 8×19, by which is meant that four hypotheses may be made with respect to its dimensions. The rectangle may consist of:

(1) 4 columns with 38 rows, *or*
(2) 38 columns with 4 rows, *or*
(3) 8 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 hypothesis (1) or (2), and for the present these may be discarded. As to choosing between hypotheses (3) and (4), a rather simple test to be described presently will disclose which is the more probable.

d. It is obvious that if the cryptogram is transcribed within a rectangle of the correct dimensions, the letters in each row will be the ones which actually were in those rows in the original transposition rectangle and formed good plain text therein. *In fact, the rows of letters in the correctly-dimensioned rectangle would read plain text were it 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 rectangles of various probable dimensions, that rectangle in which the rows show the best proportions of vowels and consonants is most likely to be correct, and the one that should be tried first.

e. Having ascertained the correct dimensions of the rectangle by the foregoing procedure, the next step is to experiment with the columns of the rectangle, trying to bring together several columns which will then show "good" digraphs, trigraphs, or polygraphs in the *rows* formed by juxtaposing the columns. This process of combining or matching columns in order to build up these fragments of plain text will herein be referred to as *anagramming*.[4]

f. 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, attempt is made to add another column before or after this pair to build up high-frequency trigraphs, and so on, gradually building up longer and longer polygraphs until entire words begin to appear in the respective rows of the rectangle. In this process of anagramming, advantage may be taken of simple mathematical considerations such as adding the normal plain-text frequency values of the digraphs in the rows to assist in discarding 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

[4] The Standard Dictionary defines the word *anagram* as follows: "(noun) 1. The letters of a word or phrase so transposed as to make a different word or phrase; as, 'time' and 'mite' are *anagrams* of 'emit'. 2. A transposition; interchange." As a verb, it is defined as "to anagrammatize; to make an anagram of; make anagrams." (The construction of anagrams was a very widespread pastime in previous centuries. See Wheatley's *Of Anagrams*, London, 1862.) A strict interpretation of the word would therefore confine it to cases wherein the letters to be rearranged already form bonafide words or intelligible phrases. However, this would hardly be broad enough for cryptanalytic purposes. As used in cryptanalysis 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.

REF ID:A59439

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.[5] 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.

g. 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.—*a*. Given the following cryptogram, the steps in solution will be set forth in detail:

CRYPTOGRAM (126 letters)

I L H H D T I E O E U D H T S O N S O O E E E E I O E F T R
R H N E A T N N V U T L B F A E D F O Y C A P D T R R I I A
R I V N L R N R W E T U T C U V R A U O O O F D A O N A J I
U P O L R S O M T N F R A N F M N D M A S A F A T Y E C F X
R T G E T A

b. The cryptogram contains 126 letters (factors of 126: 2, 3, 6, 7, 9, 14, 18, 21), suggesting rectangles of 7×18 or 9×14. 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 6×21, 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 (7×18 or 9×14), then one would proceed to test rectangles 6×21. 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 6*d*, 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 inscribing 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 (7×40 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, row 16, 6 vowels. It is concluded at once that this arrangement is highly improbable. If the plain text had been inscribed vertically in this same rectangle, and then the rows had been transposed in forming the cipher text, then in each column (18 letters) there should be (18×40 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 inscribed vertically and the cryptogram formed by transposing the rows. But when the arrangement at

[5] A suggestion for which the author is indebted to Mr. A. W. Small, junior cryptanalyst in this office. The principle makes practicable the use of tabulating machinery for the purpose of speeding up and facilitating the matching of columns in the anagramming process.

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 inscribed vertically in the rectangle 18×7 and the rows subjected to transposition, there should be (7×40 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:

REF ID:A59439

7 × 18

Row No.	1	2	3	4	5	6	7	Number of vowels
1	I	O	N	T	T	U	M	3
2	L	O	N	R	C	P	A	2
3	H	E	V	R	U	O	S	3
4	H	E	U	I	V	L	A	4
5	D	E	T	I	R	R	F	2
6	T	E	L	A	A	S	A	4
7	I	I	B	R	U	O	T	4
8	E	O	F	I	O	M	Y	5
9	O	E	A	V	O	T	E	5
10	E	F	E	N	O	N	C	3
11	U	T	D	L	F	F	F	1
12	D	R	F	R	D	R	X	0
13	H	R	O	N	A	A	R	3
14	T	H	Y	R	O	N	T	2
15	S	N	C	W	N	F	G	0
16	O	E	A	E	A	M	E	6
17	N	A	P	T	J	N	T	1
18	S	T	D	U	I	D	A	3
Number of vowels	7	11	6	6	10	4	7	

A

18 × 7

Row No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Number of vowels
1	I	E	S	E	T	T	B	Y	R	N	T	A	A	P	T	M	F	X	6
2	L	O	O	E	R	N	F	C	I	L	U	U	O	O	N	N	A	R	9
3	H	E	N	E	R	N	A	A	I	R	T	O	N	L	F	D	T	T	6
4	H	U	S	I	H	V	E	P	A	N	C	O	A	R	R	M	Y	G	7
5	D	D	O	O	N	U	D	D	R	R	U	O	J	S	A	A	E	E	9
6	T	H	O	E	E	T	F	T	I	W	V	F	I	O	N	S	C	T	6
7	I	T	E	F	A	L	O	R	V	E	R	D	U	M	F	A	F	A	8
Number of vowels	2	4	4	6	2	1	3	2	4	1	2	5	5	2	1	2	3	2	

B

9 × 14

Row No.	1	2	3	4	5	6	7	8	9	Number of vowels
1	I	S	T	B	R	T	A	T	F	2
2	L	O	R	F	I	U	O	N	A	5
3	H	N	R	A	I	T	N	F	T	2
4	H	S	H	E	A	C	A	R	Y	4
5	D	O	N	D	R	U	J	A	E	4
6	T	O	E	F	I	V	I	N	C	4
7	I	E	A	O	V	R	U	F	F	5
8	E	E	T	Y	N	A	P	M	X	4
9	O	E	N	C	L	U	O	N	R	4
10	E	E	N	A	R	O	L	D	T	4
11	U	I	V	P	N	O	R	M	G	3
12	D	O	U	D	R	O	S	A	E	5
13	H	E	T	T	W	F	O	S	T	2
14	T	F	L	R	E	D	M	A	A	3
Number of vowels	6	10	3	5	5	7	7	3	5	

C

14 × 9

Row No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Number of vowels
1	I	E	O	F	N	E	T	N	T	O	U	N	M	C	6
2	L	U	O	T	N	D	R	L	C	F	P	F	A	F	3
3	H	D	E	R	V	F	R	R	U	D	O	R	S	X	3
4	H	H	E	R	U	O	I	N	V	A	L	A	A	R	7
5	D	T	E	H	T	Y	I	R	R	O	R	N	F	T	4
6	T	S	E	N	L	C	A	W	A	N	S	F	A	G	4
7	I	O	I	E	B	A	R	E	U	A	O	M	T	E	10
8	E	N	O	A	F	P	I	T	O	J	M	N	Y	T	6
9	O	S	E	T	A	D	V	U	O	I	T	D	E	A	8
Number of vowels	4	3	9	2	2	4	4	2	5	5	3	1	5	2	

D

FIGURE 1.

Rectangle 7 × 18

7 columns and 18 rows:

(1) Horizontal inscription, columnar transcription ..... Very improbable.
(2) Vertical inscription, horizontal transcription ..... Very improbable.

18 columns and 7 rows:

(3) Horizontal inscription, columnar transcription ..... Possible.
(4) Vertical inscription, horizontal transcription ..... Improbable.

Rectangle 9 × 14

9 columns and 14 rows:

(5) Horizontal inscription, columnar transcription ..... Possible.
(6) Vertical inscription, horizontal transcription ..... Improbable.

14 columns and 9 rows:

(7) Horizontal inscription, columnar transcription ..... Improbable.
(8) Vertical inscription, horizontal transcription ..... Very improbable.

c. Discarding all assumptions except (3) and (5), the latter are subjected to further scrutiny. 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 these differences (neglecting signs), and dividing by the total number of rows. For assumptions (3) and (5) the results are as follows:

18 × 7

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Number of vowels	Deviation from expected number
1	I	E	S	E	T	T	B	Y	R	N	T	A	A	P	T	M	F	X	6	1.2
2	L	O	O	E	R	N	F	C	I	L	U	U	O	O	N	N	A	R	9	1.8
3	H	E	N	E	R	N	A	A	I	R	T	O	N	L	F	D	T	T	6	1.2
4	H	U	S	I	H	V	E	P	A	N	C	O	A	R	R	M	Y	G	7	.2
5	D	D	O	O	N	U	D	D	R	R	U	O	J	S	A	A	E	E	9	1.8
6	T	H	O	E	E	T	F	T	I	W	V	F	I	O	N	S	C	T	6	1.2
7	I	T	E	F	A	L	O	R	V	E	R	D	U	M	F	A	F	A	8	.8

Total deviation=8.2
Average deviation=1.2

FIGURE 1*c*.

REF ID:A59439

9×14

	1	2	3	4	5	6	7	8	9	Number of vowels	Deviation from expected number
1	I	S	T	B	R	T	A	T	F	2	1.6
2	L	O	R	F	I	U	O	N	A	5	1.4
3	H	N	R	A	I	T	N	F	T	2	1.6
4	H	S	H	E	A	C	A	R	Y	4	.4
5	D	O	N	D	R	U	J	A	E	4	.4
6	T	O	E	F	I	V	I	N	C	4	.4
7	I	E	A	O	V	R	U	F	F	5	1.4
8	E	E	T	Y	N	A	P	M	X	4	.4
9	O	E	N	C	L	U	O	N	R	4	.4
10	E	E	N	A	R	O	L	D	T	4	.4
11	U	I	V	P	N	O	R	M	G	3	.6
12	D	O	U	D	R	O	S	A	E	5	1.4
13	H	E	T	T	W	F	O	S	T	2	1.6
14	T	F	L	R	E	D	M	A	A	3	.6

Total deviation=12.6
Average deviation= .9

FIGURE 1*f*.

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 1*f* 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 ½-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:

(The frequencies shown are as given in table 6, appendix to *Military Cryptanalysis, Part I*.)

(1)		(2)		(3)		(4)	
7 2	Frequency value	7 6	Frequency value	7 8	Frequency value	7 9	Frequency value
A S	41	A T	47	A T	47	A F	4
O O	6	O U	37	O N	77	O A	7
N N	8	N T	82	N F	9	N T	82
A S	41	A C	14	A R	44	A Y	12
J O	2	J U	2	J A	1	J E	2
I O	41	I V	25	I N	75	I C	22
U E	11	U R	31	U F	1	U F	1
P E	23	P A	14	P M	4	P X	0
O E	3	O U	37	O N	77	O R	64
L E	37	L O	13	L D	9	L T	8
R I	30	R O	28	R M	9	R G	7
S O	15	S O	15	S A	24	S E	49
O E	3	O F	25	O S	14	O T	19
M F	1	M D	1	M A	36	M A	36
Total	262	Total	371	Total	427	Total	313

FIGURE 2.

Combination (3) gives the highest frequency value for the digraphs and an attempt is made to add a column to it. Here are some of the combinations tried:

7 8 1	7 8 2	7 8 3	7 8 9
A T I	A T S	A T T	A T F
O N L	O N O	O N R	O N A
N F H	N F N	N F R	N F T
A R H	A R S	A R H	A R Y
J A D	J A O	J A N	J A E
I N T	I N O	I N E	I N C
U F I	U F E	U F A	U F F
P M E	P M E	P M T	P M X
O N O	O N E	O N N	O N R
L D E	L D E	L D N	L D T
R M U	R M I	R M V	R M G
S A D	S A O	S A U	S A E
O S H	O S E	O S T	O S T
M A T	M A F	M A L	M A A

FIGURE 3.

e. Each of these combinations shows at least one "impossible" trigraph and several "poor" ones.[6] After more or less work along these lines, the cryptanalyst begins to get the feeling that "something is wrong," for, as a rule, once a correct start has been made in cases of this kind, solution comes rather quickly. Hence, the cryptanalyst decides here that possibly his first

[6] Following the steps taken in subpar. *d*, frequency weights may be given the various trigraphs in fig. 3 and the sums obtained taken as indications of the relative probability of each of the four trials. These steps are here omitted, for they are obvious.

REF ID:A59439

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 3d. Placing column 3 after columns 7–6 gives the trigraphs shown in figure 4–A. All of these trigraphs 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 3.

7 6 3
A T T
O U R
N T R
A C H
J U N
I V E
U R A
P A T
O U N
L O N
R O V
S O U
O F T
M D L

FIGURE 4–A.

9 ? ? ? ? ? 7 6 3
F . . . . . A T T
A . . . . . O U R
T . . . . . N T R
Y . . . . . A C H
E . . . . . J U N
C . . . . . I V E
F . . . . . U R A
X . . . . . P A T
R . . . . . O U N
T . . . . . L O N
G . . . . . R O V
E . . . . . S O U
T . . . . . O F T
A . . . . . M D L

FIGURE 4–B.

9 1 5 2 8 4 7 6 3
F I R S T B A T T
A L I O N F O U R
T H I N F A N T R
Y H A S R E A C H
E D R O A D J U N
C T I O N F I V E
F I V E F O U R A
X E N E M Y P A T
R O L E N C O U N
T E R E D A L O N
G U N I M P R O V
E D R O A D S O U
T H W E S T O F T
A T E F A R M D L

FIGURE 5.

f. 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.

g. It will be interesting to see if a calculation based upon the sum of the 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.

	(1)			(2)			(3)			(4)	
7 2	Frequency value	Logarithms	7 6	Frequency value	Logarithms	7 8	Frequency value	Logarithms	7 9	Frequency value	Logarithms
A S	41	1.61	A T	47	1.67	A T	47	1.67	A F	4	0.60
O O	6	.78	O U	37	1.57	O N	77	1.89	O A	7	.85
N N	8	.90	N T	82	1.91	N F	9	.95	N T	82	1.91
A S	41	1.61	A C	14	1.15	A R	44	1.64	A Y	12	1.08
J O	2	.30	J U	2	.30	J A	1	.00	J E	2	.30
I O	41	1.61	I V	25	1.40	I N	75	1.88	I C	22	1.34
U E	11	1.04	U R	31	1.49	U F	1	.00	U F	1	.00
P E	23	1.36	P A	14	1.15	P M	4	.60	P X	0	−1.00
O E	3	.48	O U	37	1.57	O N	77	1.89	O R	64	1.81
L E	37	1.57	L O	13	1.11	L D	9	.95	L T	8	.90
R I	30	.48	R O	28	1.45	R M	9	.95	R G	7	.85
S O	15	1.18	S O	15	1.18	S A	24	1.38	S E	49	1.69
O E	3	.48	O F	25	1.40	O S	14	1.15	O T	19	1.28
M F	1	.00	M D	1	.00	M A	36	1.56	M A	36	1.56
Total	262	13.40	Total	371	17.35	Total	427	16.51	Total	313	13.17

FIGURE 6.

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 .84, which is the logarithm of 7. Likewise, combination 7–6 is roughly 15,000 times more probable than combination 7–9, since 17.35—13.17=4.18.

8. The probable-word method of solution.—*a.* 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.

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

9 1 5 2 8 4 7 6 3
F I R S T B A T T
A L I O N

It will be noted that this assumption requires that there be a column headed by F A, another headed by I L, 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.

c. 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 solution.—*a.* In solving transposition ciphers advantage should be taken of all the characteristics and idiosyncrasies which are applicable to the language of the

324155—41——2

REF ID:A59439

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 T H, which is of highest frequency in literary text; it also combines with C, a letter of medium frequency, to form the fairly frequent digraph C H. The letter V is almost in the low-frequency category yet it combines with E to form the digraph V E, which in military text is the 14th in frequency. The low-frequency letter K often combines with C to form the digraph C K. 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.

b. There is usually in every language at least one letter which can be followed by only a certain other letter, forming what may be termed an *obligatory sequence*, or an *invariable digraph*. In all languages having the letter Q, the combination Q U constitutes such an invariable digraph.[7] 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 C K is employed. In English, the letter J 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 (C 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*.

c. 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 Y, which are all of medium frequency. The important word ATTACK suggests itself if the cryptogram has a K, a letter of low frequency, and a C, one of medium frequency; and so on.

d. The mechanics of simple columnar transposition make 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 O E E E E I O 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 much more frequent than consonant-consonant or vowel-vowel digraphs,[8] 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.

e. 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 C H), that astute cryptographers who realize the clues which such letters afford often replace invariable digraphs by single 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 a factor of importance.

[7] 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, QU may be considered to be an invariable digraph.

[8] The CV and VC digraphs constitute about 62 percent of all digraphs.

f. 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.

g. 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 uniliteral 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.

10. Reconstruction of literal key.—*a.* The reconstruction or 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 if 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 locale of operations, this fact may be of highest importance in subsequent work.

b. 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.

c. 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.

	9	1	5	2	8	4	7	6	3
1		1		2					3
2						4			
3			5					6	
4							7		
5					8				
6	9								

FIGURE 7–A.

	9	1	5	2	8	4	7	6	3
1		ABC DE 1		ABC DE 2					ABC DE 3
2						FGH IJ 4			
3			KLM NO 5					KLM NO 6	
4							LMN OP 7		
5					MNO PR 8				
6	R–Z 9								

FIGURE 7–B.

d. 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

REF ID:A59439

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 2 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.

e. 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.

9	1	5	2	8	4	7	6	3
R	A	K	A	M	F	L	K	A
S	B	L	B	N	G	M	L	B
T	C	M	C	O	H	N	M	C
U	D	N	D	P	I	O	N	D
V	E	O	E	R	J	P	O	E
W								
X								
Y								
Z								

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: [9]

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 Z 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 that 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, O L E, O M E, O N E as possible terminal trigraphs, the best of which from a frequency standpoint is ONE. Seeing the letters P and H in columns 8–4, 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

[9] Taken from Tables 2–D (2) and 2–E (2), p. 111, *Military Cryptanalysis, Part I.*

g. In future studies, cases will be encountered wherein the reconstruction of the numerical key is an essential or, at least, a useful element after the solution of one or more cryptograms has been achieved by cryptanalysis. This is done in order that subsequent cryptograms in the same key can be read directly without cryptanalysis. The reconstruction of the numerical key is, however, a different process than the one illustrated in this paragraph, wherein the problem is solely one of building up a literal key from its numerical equivalent. The purpose in reconstructing the literal key is to give clues as to the *source* from which keys are derived or taken. Sometimes this may lead to ascertaining a book which is used for this purpose and which may be available by purchase at bookshops; or it may be a well-known document, a telephone directory, etc. Obviously, if the source document or book can be located the solution of future cryptograms in the same system becomes merely a matter of decipherment and such cryptograms no longer form the material for cryptanalytic efforts. The method of reconstructing the literal key is, obviously, easier to apply in the case of long numerical keys than in the case of short ones; in general, the longer the numerical key the easier is the recovery of the literal key.

11. Column and row transposition.—It should be obvious that when the rows as well as the columns of a completely-filled rectangle undergo transposition the increase in security is hardly worth mention, since the underlying procedure in solution aims simply at assembling a few columns on the basis of "good" digraphs and trigraphs brought to light by juxtaposing *columns*. After three or four columns have been properly juxtaposed, the placement of additional columns becomes easier and easier, merely by continuing to build upon the fragments of words *in the rows*. Hence, the cryptanalyst is, during a large part of the process, not particularly interested in the intelligibility of the text he is building up; only at the end of the process does this become a factor. When all of the columns have been assembled in proper order, then the text will read continuously in the normal manner (left to right, top to bottom). If it does not, then it is usually a very simple matter to rearrange the rows of the rectangle to bring this about, since the letters at the ends and beginnings of the rows give the necessary clues for continuity.

REF ID:A59439

SECTION III

INCOMPLETELY—FILLED RECTANGLES

	Paragraph
General principles underlying solution	12
Delimiting the lengths of the columns of the rectangle; constructing the "hat" diagram	13
Solution of example	14
Alternative method of solution	15
Concluding remarks on simple columnar transposition	16

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 cryptanalytic point of view, quite profound, and the cryptanalytic effect of this small change in cryptographic procedure is seemingly all out of proportion with the simplicity of the difference.

b. Cryptograms 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 this materially reduces 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 all 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 to form digraphs and trigraphs really come from the same row in the plain-text rectangle. The latter statement calls for a bit more explanation.

c. The *columns* of an incompletely-filled rectangle are of two sorts which may conveniently be 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. This follows, of course, 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 the 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 columns; 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: [(15×20)−13=287] the

long columns will have 20 letters, the short ones, 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.

e. 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.[1] 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.[2]

f. However, this question may be asked: Given a cryptogram of t letters and a rectangle of n columns 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 n elements? 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, then of course there would be no use in forming such digraphs, and thus the number of trials 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 will 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 keyed, columnar-transposition cryptogram.

[1] Factorial 15, or 15×14×13×. . .×1, equals 1,369,944,576,000 different transposition keys.

[2] It is nevertheless pertinent to indicate that machinery for facilitating the "matching" or anagramming of columns has been devised and found to be quite practical in the solution of problems involving columnar transposition.

REF ID:A59439

13. 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

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

DISTRIBUTION

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
19	4	6	6	31	6	3	6	16	1	2	9	7	20	15	1	1	18	12	25	4	4	1	2	5	0	(224)

b. A cryptogram of 224 letters and a rectangle of 12 columns [(12×19)−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	2	3	4	5	6	7	8	9	10	11	12
O	N	M	C	M	H	Y	T	O	A	F	A
D	I	A	T	I	B	O	N	O	I	I	T
N	E	N	R	A	V	I	O	T	I	T	S
N	T	L	M	I	T	K	T	H	E	A	O
P	G	N	S	T	B	N	R	E	A	N	A
T	N	O	C	E	E	K	A	A	T	L	L
I	W	E	L	A	S	T	A	C	E	N	O
R	T	B	C	I	R	N	V	N	E	N	D
N	T	O	S	J	S	D	R	S	A	U	I
T	M	U	O	N	Y	A	I	H	L	A	R
D	E	H	E	S	L	H	R	E	A	C	O
T	D	E	F	O	X	I	S	E	E	L	R
U	S	N	C	I	R	R	T	V	N	R	Y
R	T	L	F	V	O	H	G	T	E	E	N
O	E	E	F	F	R	Q	S	R	E	N	N
E	O	S	T	M	U	I	E	E	M	R	R
X	I	E	E	B	M	E	F	S	Y	T	G
A	T	A	E	I	E	T	E	R	T	R	Y
L	D	A	E	E	T	E	A				

FIGURE 8.

c. The rectangle shown in figure 8 is the same as though it had been assumed that the key numbers 9, 10, 11, and 12 happened to fall at the extreme right in the numerical transposition key. Columns 1 to 8, inclusive, would then be long columns, and columns 9, 10, 11, and 12 would be short columns. But suppose that the key numbers on the extreme right happened to be 1, 2, 3, and 4, instead of 9, 10, 11, and 12. Then columns 1, 2, 3, and 4 would be the short columns, 5 to 12 the long ones. In this case, making reference to figure 8, the final letter of column 1 would pass to the top of column 2; the final 2 letters of column 2 would pass to the top of column 3; the final 3 letters of column 3 would pass to the top of column 4; the final 4 letters of columns 4, 5, 6, 7, and 8 would pass to the top of columns 5, 6, 7, 8, and 9; the final 3 letters of column 9 would pass to the top of column 10; the final 2 letters of column 10 would pass to the top of column 11; and the final letter of column 11 would pass to the top of column 12. The results of the foregoing reasoning are embodied in the matrix or diagram shown in figure 9.

	1	2	3	4	5	6	7	8	9	10	11	12
1					t	m	u	i	e			
2				e	e	b	m	e	f	e		
3			t	a	e	i	e	t	e	s	y	
4		l	d	a	e	e	t	e	a	r	t	r
5	O	N	M	C	M	H	Y	T	O	A	F	A
6	D	I	A	T	I	B	O	N	O	I	I	T
7	N	E	N	R	A	V	I	O	T	I	T	S
8	N	T	L	M	I	T	K	T	H	E	A	O
9	P	G	N	S	T	B	N	R	E	A	N	A
10	T	N	O	C	E	E	K	A	A	T	L	L
11	I	W	E	L	A	S	T	A	C	E	N	O
12	R	T	B	C	I	R	N	V	N	E	N	D
13	N	T	O	S	J	S	D	R	S	A	U	I
14	T	M	U	O	N	Y	A	I	H	L	A	R
15	D	E	H	E	S	L	H	R	E	A	C	O
16	T	D	E	F	O	X	I	S	E	E	L	R
17	U	S	N	C	I	R	R	T	V	N	R	Y
18	R	T	L	F	V	O	H	G	T	E	E	N
19	O	E	E	F	F	R	Q	S	R	E	N	N
20	E	O	S	T	M	U	I	E	E	M	R	R
21	X	I	E	E	B	M	E	F	S	Y	T	G
22	A	T	A	E	I	E	T	E	R	T	R	Y
23	L	D	A	E	E	T	E	A				

FIGURE 9.

d. Now the capital letters in this matrix or diagram, which is often called a *crown* or *hat* diagram,[3] figure 9, represent the latters which are in the columns in case the first hypothesis (key numbers 9, 10, 11, 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 (key 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

[3] Because the lower-case letters at the top form what is usually called the *crown* or *hat*.

REF ID:A59439

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 anagramming 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 enciphering 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.

e. 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 23 of figure 9.

f. 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. Now the question is, can any of these 4 U's follow the U, 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 U is located, it follows that neither one of these U's can be the one that succeeds the U. Thus two candidates are automatically eliminated from consideration. The U in row 17 and the U in row 20 are both possible candidates.

14. 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 anagramming. 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 O 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 10*a*. There are only 17 digraphs in all, whereas there should be at least 18.

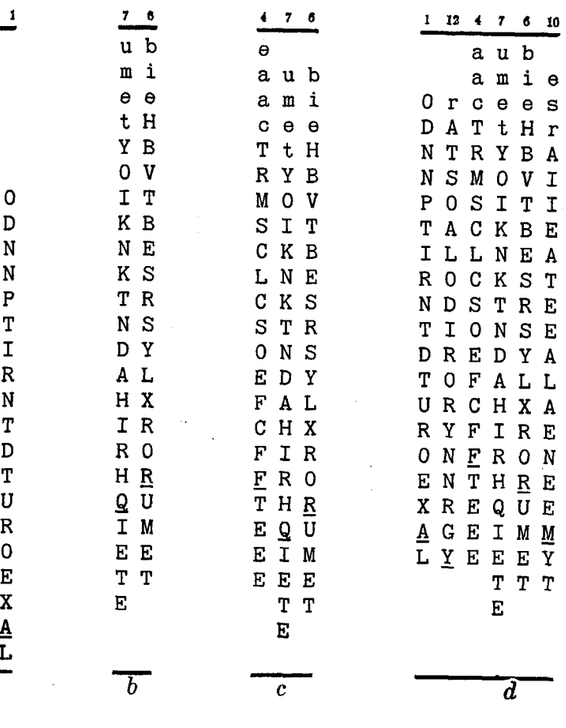


FIGURE 10.

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 H X. Seeing the digraphs K B and K S in these columns, a good assumption to make is that the K's are preceded by the letter C. Is there a column with 2 C's in approximately the correct region? Column 4 meets this requirement. Note the excellent trigraphs it yields, as shown in figure 10*c*. 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 O V suggest 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 10*d*, 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 E E 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 10*e*. 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.

REF ID:A59439

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

FIGURE 10*c*.

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

FIGURE 11.

b. 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.

8	2	5	3	11	9	1	12	4	7	6	10
						ABC 1					
	DEF GHI 2		DEF GHI 3					DEF GHI 4			
		JKL MN 5								JKL MN 6	
									MNO PQ 7		
NOP RST 8					NOP RST 9						NOP RST 10
				R-Z 11			R-Z 12				

FIGURE 12.

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

8	2	5	3	11	9	1	12	4	7	6	10
						A 1					
	DEF GHI 2		DEF GHI 3					I 4			
		JKL MN 5								N 6	
									O 7		
PRS 8					PRS 9						S 10
				T 11			T 12				

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

REF ID:A59439

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 find 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 correspondence, which may afford clues as to minimum and maximum widths of rectangles.

b. 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 or 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 in a single unbroken string at least one complete column of letters from the transposition rectangle. Then a section of consecutive letters of text 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 U 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 found. Then the same thing is done with the Q and all other U's in the text, to insure 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 unlucky and makes several false starts, the process is likely to be a slow one. In choosing from among several 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 7*g*. In the absence of any Q's in the text, recourse must be had to the formation of 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, suppose that a text shows a sequence . . . V E H H K . . .. A sequence such as . . . A R T C C . . . would be excellent to try, for it will yield the digrahhs AV, RE, TH, CH, CK. Or if there is a long sequence of consonants, the cryptanalyst should look for a correspondingly long sequence of vowels, since these 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 Q's, 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; or, the last 20 or 25 letters of the cryptogram may be used as the starting point, since these letters come from the last-numbered column of the rectangle and therefore there is no uncertainty at least as to the letter which is at the bottom of that column.

c. 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" trigraphs and tetragraphs. For example, suppose that the initial juxtaposition has yielded what is shown in figure 14. The digraph P R suggests that it must be followed by a vowel, preferably E, A, or O; the digraph A V might be part of the word CAVALRY, in which case it will be followed by A; the digraph C R 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 letters here desired, *and, of course, at the proper intervals.* Suppose such a sequence is found and yields what is shown in figure 15. The skeletons of words are now beginning to appear. Assuming that A V A is indeed part of the word CAVALRY, there should be an L to follow it; the trigraph T I N suggests the termination G; the trigraph Z E R suggests the word ZERO. A section of text is therefore sought, which will have the letters L, G, and O in the order L . . . . . G . . O. Enough has been shown to demonstrate the procedure. In the course of the work it soon becomes evident where the ends 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 likewise the trigraph at the bottom, E A O. 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 then the "good" portion is composed of the 13 letters between these two extremities 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 this becomes evident the sequence of columns has also become evident.

1 2 3 4 5	1 2
. .	. . .
. .	. . .
R R	R R S
N A	N A T
P R	P R E
T O	T O R
A V	A V A
R E	R E D
T H	T H R
C H	C H U
C K	C K A
I L	I L L
T I	T I N
C R	C R A
B E	B E S
Z E	Z E R
E A	E A O
. .	. . .
. .	. . .

FIGURE 14. FIGURE 15.

d. 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 then 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 length closer to the maximum than to the minimum; consequently a length of 20 letters will be tentatively assumed. Now the Q may be at the top of a column, at the middle,

REF ID:A59439

O 1	D 31	M 61	F 91	T 121	E 151	A 181	A 211
D	S	S	M	N	A	E	L
N	T	C	B	D	O	N	O
N	E	L	I	A	O	E	D
P 5	O 35	C 65	E 95	H 125	T 155	E 185	I 215
T	I	S	H	I	H	M	R
I	T	O	B	R	E	Y	O
R	D	E	V	H	A	T	R
N	M	F	T	Q	C	F	Y
T 10	A 40	C 70	B 100	I 130	N 160	I 190	N 220
D	N	F	E	E	S	T	N
T	L	F	S	T	H	A	R
U	N	T	R	E	E	N	G
R	O	E	S	T	E	L	Y
O 15	E 45	E 75	Y 105	N 135	V 165	N 195	
E	B	E	L	O	T	N	
X	O	M	X	T	R	U	
A	U	I	R	R	E	A	
L	H	A	O	A	S	C	
N 20	E 50	I 80	R 110	A 140	R 170	L 200	
I	N	T	U	V	A	R	
E	L	E	M	R	I	E	
T	E	A	E	I	I	N	
G	S	I	T	R	E	R	
N 25	E 55	J 85	Y 115	S 145	A 175	T 205	
W	A	N	O	T	T	R	
T	A	S	I	G	E	A	
T	C	O	K	S	E	T	
M	T	I	N	E	A	S	
E 30	R 60	V 90	K 120	F 150	L 180	O 210	
D	M	F	T	E	A	A	
S	S	M	N	A	E	L	
T	C	B	D	O	N	O	
E	L	I	A	O	E	D	
O 35	C 65	E 95	H 125	T 155	E 185	I 215	
I	S	H	I	H	M	R	
T	O	B	R	E	Y	O	
D	E	V	H	A	T	R	
M	F	T	Q	C	F	Y	
A 40	C 70	B 100	I 130	N 160	I 190	N 220	

FIGURE 16.

1	2	3	4
O	OT 28	OF 91	OE 177
110 R	110 RM	110 RM	110 RE
U	UE 30	UB	UA
M	MD	MI	ML 180
E	ES	EE 95	EA
T	TT	TH	TE
115 Y	115 YE	115 YB	115 YN
O	OO 35	OV	OE
IO 1	II	IT	IE 185
KD	KT	KB 100	KM
NN	ND	NE	NY
120 KN	120 KM	120 KS	120 KT
TP 5	TA 40	TR	TF
NT	NN	NS	NI 190
DI	DL	DY 105	DT
AR	AN	AL	AA
125 HN	125 HO	125 HX	125 HN
IT 10	IE 45	IR	IL
RD	RB	RO	RN 195
HT	HO	HR 110	HN
→ QU	QU	QU	QU ←
130 IR	130 IH	130 IM	130 IA
EO 15	EE 50	EE	EC
TE	TN	TT	TL 200
EX	EL	EY 115	ER
TA	TE	TO	TE
135 NL	135 NS	135 NI	135 NN
ON 20	OE 55	OK	OR
TI	TA	TN	TT 205
RE	RA	RK 120	RR
AT	AC	AT	AA
140 AG	140 AT	140 AN	140 AT
VN 25	VR 60	VD	VS
RW	RM	RA	RO 210
IT	IS	IH 125	IA
RT	RC	RI	RL
145 SM	145 SL	145 SR	145 SO
TE 30	TC 65	TH	TD
GD	GS	GQ	GI 215
SS	SO	SI 130	SR
ET	EE	ET	EO
1	2	3	4

FIGURE 17-A.

324155—41——3

REF ID:A59439

or at the bottom—there is no way of telling at this point. Hence, to make sure that nothing is overlooked, suppose a section of 41 letters is taken, with the Q at the center. There are 4 U's in the message, and 4 trials are to be made. The results are as indicated in figure 17–A. Examining combination 1 in figure 17–A, the digraphs formed both above and below the Q U are not at all bad. In fact, not one of those above the Q U is impossible and the same is true of those below the Q U until the digraph V N is reached. Hence, combination 1 is possible. As for combination 2, this at once appears to be bad. Digraphs such as I I, and I 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 U, although the digraph H X looks very bad. However, the X might be a sentence separator, so that this combination cannot be discarded. Combination 4 looks very improbable, with the digraph H N occurring twice, 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 possibilities. 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.*

	SOF	
	ERM	
	AUB	
	AMI	
	CEE	
	TTH	
	RYB	
	MOV	
SIO	SIT	
CKD	CKB	
LNN	LNE	
CKN	CKS	
STP	STR	
ONT	ONS	
EDI	EDY	
FAR	FAL	
CHN	CHX	
FIT	FIR	
FRD	FRO	
THT	THR	
→ EQU	EQU	
EIR	EIM	
EEO	EEE	
MTE	MTT	
IEX	IEY	
ATA	ATO	
INL	INI	
TON	TOK	
ETI	ETN	
ARE	ARK	
IAT	IAT	
JAG	JAN	
1	3	

FIGURE 17–B.

Noting the 2 K's (in the digraphs K B and K S) among the combinations before the Q, assume that these 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 63d letter. Suppose the corresponding section is placed in front of the combinations 1 and 3 of figure 17, as shown in figure 17–B. It immediately becomes evident that combination 3 is the correct one, for note the excellent trigraphs it gives, as compared with those in combination 1. Also note that the second trigraph below the E Q U in combination 3 consists of 3 E's, indicating that the end of the columns has been reached just before this trigraph. 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 above the E Q U suggests either THREE or THROUGH; the F R O above the T H R suggests FROM or FRONT. Suppose the word REQUEST is assumed for the E Q U, and the word THREE is assumed for the T H R above it. This requires a section with 2 E's in succession.

e. There are several such places in the text, and further limitation is advisable. The 8th trigraph from the top is certainly suggestive of the word MOVING, which requires an I to follow the V. Is there a place in the text where an I occurs 12 letters before a succession of two E's? There is one such place, and the corresponding section is juxtaposed at the proper place, yielding what is shown in figure 17–C.

SOFV
ERMT
AUBR
AMIE
CEES
TTHR
RYBA
MOVI
SITI
CKBE
LNEA
CKST
STRE
ONSE
EDYA
FALL
CHXA
FIRE
FRON
THRE
EQUE ←
EIMM
EEEY

FIGURE 17–C.

figure 17–C. The upper and lower limits of the columns are now fairly definite and are marked by the horizontal bars; tetragraphs E E E Y at the bottom and A M I 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.

f. When two or three operators are assigned to work upon a cryptogram by this method, solution 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 assimilated into one message which is then completed very speedily.

g. This and the next four subparagraphs will be devoted to some remarks of a general nature concerning columnar transposition of the foregoing type. The degree of cryptographic security afforded by simple columnar transposition methods, especially when incompletely-filled 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. Obviously, the use of low-frequency 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.

h. 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 unoccupied by letters of the text. If only a few cells distributed irregularly within the columns of the transposition matrix are designated as blanks, the disturbing effect upon the anagramming process is quite marked. This 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.

i. In order to fix definitely the positions of the nulls or of the blanks in the transposition matrix, definite prearrangements 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 relatively few different forms, the use of varying transposition keys serves to increase cryptographic security to a rather marked degree.

j. 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 assymmetrically the matter becomes hopeless, as far as a single message is concerned.

k. Of course, if several messages of identical length and in the same key are available for superimposition, the presence of the nulls or blanks then makes little difference, because the

REF ID:A59439

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 hereinafter 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 hereinafter be designated the *cipher→plain sequence*, or *C→P sequence*.

b. These two sequences are obviously related, one being the *inverse* or indexed version of the other. Given one of the sequences, the other can be derived from it by the simple operation of indexing, 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.

T	R	A	N	S	P	O	S	I	T	I	O	N
12	9	1	4	10	8	6	11	2	13	3	7	5
T	H	E	Q	U	I	C	K	B	R	O	W	N
F	O	X	J	U	M	P	S	O	V	E	R	T
H	E	L	A	Z	Y	D	O	G				

Term No.	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
Cipher	E	X	L	B	O	G	O	E	Q	J	A	N	T	C	P	D	W	R	I	M	Y	H	O	E	U	U	Z	K	S	O	T	F	H	R	V

FIGURE 18–A.

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	1	4	10	8	6	11	2	13	3	7	5
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				

Term number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
P→C equences	03	16	29	09	22	35	11	24	04	17	30	13	26	07	20	33	12	25	06	19

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
32	02	15	28	05	18	31	08	21	34	01	14	27	10	23

FIGURE 18–B.

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, T_p, 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_c; the 2d letter of the plain text, H_p, becomes the 22d letter of the cipher text and that in the cryptogram the 22d letter is H_c; and so on. In connection with the P→C sequence, it is to be noted that successive terms in the sequence,

in the case of single transposition, show a constant difference except when passing from a greater to a smaller number, which happens every time a transition is made from a term applying to the bottom element of one column to a term applying to the top element of the next column. For example, in the case of the 1st three terms in the sequence: 16−03=13; 29−16=13. However, in the case of the 3d term of the sequence (29) and the 4th (09) the passage is from a greater to a smaller number and the constant difference, 13, no longer is evident. The cause of the constant difference is, of course, obvious and follows directly from the mechanics of the transposition system itself. The point to be specially noted is that the existence of such a constant difference (with the exceptions noted above) may be taken as one of the identifying characteristics of single columnar transposition; double columnar transposition or other types of complex transposition will show no such constant difference throughout the P→C sequence.

d. Given the P→C sequence in subparagraph *c*, its inverse, the C→P sequence is established merely by preparing an indexed version of the former. Thus:

Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
P→C sequence	03	16	29	09	22	35	11	24	04	17	30	13	26	07	20	33	12	25	06	19

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
32	02	15	28	05	18	31	08	21	34	01	14	27	10	23

Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
C→P sequence	31	22	01	09	25	19	14	28	04	34	07	17	12	32	23	02	10	26	20	15

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
29	05	35	08	18	13	33	24	03	11	27	21	16	30	06

e. The C→P sequence can also be produced in another way. Suppose that numbers are inscribed in the cells of the transposition matrix, not in the normal manner of writing from left to right and form the top downward, but according to the route followed in *transcribing* the numbers to form the "cipher text," that is, in key-number order in the columns of the matrix. Thus:

12	9	1	4	10	8	6	11	2	13	3	7	5
31	22	01	09	25	19	14	28	04	34	07	17	12
32	23	02	10	26	20	15	29	05	35	08	18	13
33	24	03	11	27	21	16	30	06				

FIGURE 18–C.

If these numbers are now transcribed according to the normal manner of writing (from left to right and from the top downward), the sequence produced is 31 22 01 09 25 . . ., which coincides with the C→P sequence shown in subparagraph *d* above, which in turn was derived from the P→C sequence.

Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
C→P sequence	31	22	01	09	25	19	14	28	04	34	07	17	12	32	23	02	10	26	20	15

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
29	05	35	08	18	13	33	24	03	11	27	21	16	30	06

f. The C→P sequence may also be called the *anagram sequence* because it can be established as a result of a solution accomplished by anagramming superimposed messages of identical length. It is clear that what is accomplished in such a solution is to rearrange the letters of the cipher text to bring them back into their original order in the cipher text, that is, the solution involves a C→P conversion.

g. The P→C sequence is called by a recent French author the *κρ* sequence (from the Greek word *kryptos*) because it gives the order of the plain-text letters as they occur in the cryptogram. The P→C sequence is also termed the *encipher sequence* by another writer, and still another has called it the *transposition sequence*. The present author believes that the terminology adopted herein, *viz*, P→C sequence and C→P sequence, is less confusing and serves more accurately to identify or characterize these sequences than the other designations herein indicated.

REF ID:A59439

h. 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 plain-text 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.

i. In single, keyed-columnar transposition an interesting relationship exists between sections of the C→P sequence. Consider the C→P sequence given in subparagraph *d* 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:

Term number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
C→P sequence	31	22	01	09	25	19	14	28	04	34	07	17	12	32	23	02	10	26	20	15	29	05
C→P sequence +1	32	23	02	10	26	20	15	29	05	35	08	18	13	33	24	03	11	27	21	16	30	06

23	24	25	26	27	28	29	30	31	32	33	34	35
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 sequent numbers in *columns*, the following interesting diagram is obtained:

	31	22	01	09	25	19	14	28	04	34	07	17	12
+1	32	23	02	10	26	20	15	29	05	35	08	18	13
+1	33	24	03	11	27	21	16	30	06				

FIGURE 18–D.

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 *c* above.

j. 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 term numbers in the cryptogram) are as follows:

9	25 26	19	14	28
Q	U	I	C	K

The sequence 9–$\frac{25}{26}$–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:

Base	9	25 26	19	14	28	=	Q	U	I	C	K
Derivative 1	10	26 27	20	15	29	=	J	U Z	M	P	S
Derivative 2	11	27 28	21	16	30	=	A	Z K	Y	D	O
Derivative 3	12	28 29	22	17	31	=	N	K S	H	W	T

Here it will be seen that portions of "good" plain text become manifest, viz, JUMPS and AZYDO. The 3d derivative no longer is "good" because the rectangle has but 3 rows and consequently only the 1st and 2d 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.

k. 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 is subparagraphs *i* and *j* in 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+1 sequence. Thus:

P→C sequence	03 16 29	09 22 35	11 24	04 17 30	13 26 07	20 33	12 25
	1	2	3	4	5	6	7
	06 19 32	02 15 28	05 18 31	08 21 34	01	14 27	10 23
	8	9	10	11		12	13

P→C sequence +1	04 17 30	10 23	36	12 25	05 18 31	14 27	08 21 34	13 26 07
	1=4	2=13		3=7	4=10	5=12	6=11	7=5
	20 33	03 16 29	06 19 32	09 22 35	02 15 28	11 24		
	8=6	9=1	10=8	11=2	12=9	13=3		

The equivalencies between identities, as indicated above, indicate not only that the enciphering matrix has 13 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=3, 3=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.

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

m. 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 ascertaining 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 studies cases of this kind will be encountered.

n. 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 a

REF ID:A59439

multiple of the length of the short columns and a multiple of the length of the long columns. For example, consider the adjacent letters C K in the plain-text rectangle in figure 18–A. In the cipher, C_c is the 14th letter, K_c 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: (4×3)+(1×2)=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:

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				

		5	4
0	multiple	0	0
1st	multiple	5	4
2d	multiple	10	8
3d	multiple	—	12
4th	multiple	—	16

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 sums of the combinations. Thus:

No. of long columns	Length in letters	0	1	2	3	4 ---No. of short columns
		0	4	8	12	16 ---Length in letters
0	0	0	4	8	12	16
1	5	5	9	13	17	21
2	10	10	14	18	22	26

If, now, diagonal lines are drawn from the lower left-hand corner to the upper right-hand corner of the diagram, the locations of all possible breaks are given. Thus, there can be a break between the 4th and 5th letters (passing from a short column to the next column, which may be long or short, of course); there cannot be a break between the 5th and 6th letters, nor between the 6th and 7th, nor between the 7th and 8th; there can be a break between the 8th and 9th, as well as between the 9th and 10th, but not between the 10th and 11th, and so on. Suppose that for one reason or another the cryptanalyst has good reason to suspect that a break occurs immediately after the 13th letter. This means that there are 2 short columns (of 4 letters) and 1 long column (of 5 letters) up to that break. The diagram shows that there remain only 2 short columns and 1 long one, and the only breaks that are possible beyond the one at the 13th letter are: Between the 17th and 18th, or between the 21st and 22d letters.

o. The importance of the various principles set forth in this paragraph will become evident as the student progresses in his studies of transposition ciphers.

SECTION IV

OPPORTUNITIES AFFORDED BY STUDYING ERRORS AND BLUNDERS MADE BY ENEMY CRYPTOGRAPHERS

	Paragraph
Importance of the study of errors and blunders in early work upon an unknown system	17
Significance of terms "special solution" and "general solution"	18
Examples to be studied	19

17. Importance of the study of errors and blunders in early work upon an unknown system.—*a*. Blunders and mistakes made by cryptographic clerks in the execution of 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 instances take place and these afford golden opportunities for the enemy cryptanalytic service. This situation exists in respect to all types of cryptographic systems and 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 ineptitude of the adversary's cryptographers.

b. 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.

c. After the cryptanalyst has by painstaking and more or less arduous labors 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 sorts 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.

d. 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 resist all efforts to break it down for a very long time.

18. Significance of terms "special solution" and "general solution."—*a*. 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

REF ID:A59439

the system itself are such that unavoidable or unpredictable circumstances arise so that special solutions 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 is either impossible or improbable, 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 which this may be true are, of course, cryptographically weak but it may be observed that it is perhaps impossible to devise a system which may be considered to be absolutely free from this source of weakness.

c. The advantages of a special solution for any type of cryptographic system are, as a rule, two in number. First, it often makes a solution possible when otherwise this might not be the case. Secondly, if often affords a method of achieving a very rapid solution in the case of a problem which otherwise might require a long time. But a special solution presents one basic disadvantage: It is by its very nature dependent upon the existence of unusual circumstances; in other words, upon chance or good fortune bringing about a set of circumstances favorable for a solution. When these unusual conditions or circumstances do not obtain, then solution may be impossible. Therefore, it is desirable to have, if possible, for every type of system a more or less *general solution* which may be applied in the absence of the unusual conditions necessary for the application of a special solution. In other words, a general solution in cryptanalysis implies a method or procedure which if 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 partakes 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 the latter can be applied 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 it may involve quite large changes or modifications in the general procedure.

19. Examples to be studied.—*a.* 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 circumstances 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 enciphering clerk reexamines his original work and finds that he has made a mistake. He then commits the grave blunder of reenciphering the identical message (without paraphrasing) and transmitting what to the enemy cryptanalysts 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 cryptanalysts are 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 the many opportunities for quick cryptanalytic success afforded by enemy addiction to the use of stereotypic phraseology, especially at the beginnings and endings of messages. Stereotypic phraseology affords even more golden opportunities for cryptanalytic success in the case of transposition systems than it does in the case of substitution systems.

c. In the next few paragraphs some specific examples of the consequences of cryptographic blunders and ineptitude in the case of transposition systems will be studied. These are intended to give the student some idea of the far-reaching effects such studies may have. It is important that he grasp the fundamental principles, for 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.

REF ID:A59439

Section V

SPECIAL SOLUTIONS FOR TRANSPOSITION CIPHERS

	Paragraph
Solution when the beginning or end of the plain text is known	20
The case of an omitted column	21
The case of an interchanged pair of columns	22
Messages with similar beginnings	23
Messages with similar endings	24
The solution of a single message containing a long repetition	25
Solution when several cryptograms of identical length and in the same key are available	26
Reconstruction of the keys in double transposition	27
Special cases of solution of double transposition ciphers	28

20. Solution when the beginning or end of the plain text is known.—*a.* It often happens, when correspondents have fallen into the bad habit of sending stereotyped communications, that the beginnings or the endings of messages become so fixed in their form and content that the enemy can with a fair degree of certainty guess what these will be in specific cases. If so, a quick solution can be reached, the key reconstructed for one message, and this will, of course, enable him to read all other messages in the same key. This is particularly true of simple, keyed-columnar-transposition ciphers. It is only necessary that the cryptanalyst cut the text up in such a manner as to bring the letters composing the assumed text all within the same row or rows of the transposition rectangle.

b. Suppose that the enemy frequently uses the introductory expression REFERRING TO YOUR NUMBER. Here is a cryptogram assumed to begin with this phrase:

Cryptogram

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

c. Assuming that previous experience has indicated that the enemy uses keys varying from 10 to 20 letters in length, the arrangement of the letters in the tops of columns under a key length of 10 would be as shown in Fig. 20.

1	2	3	4	5	6	7	8	9	10
R	E	F	E	R	R	I	N	G	T
O	Y	O	U	R	N	U	M	B	E
R									

Figure 20.

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.

1	2	3	4	5	6	7	8	9	10	11
R	E	F	E	R	R	I	N	G	T	O
Y	O	U	R	N	U	M	B	E	R	

Figure 21.

Here a column is headed by I M, so that this is a possible arrangement. If the width of the rectangle is 11, its outlines are as shown in figure 22. There are 5 columns of 11 letters and 6

R	E	F	E	R	R	I	N	G	T	O
Y	O	U	R	N	U	M	B	E	R	
						A				
						O				
						D				
						R				
						M				
						G				
						R				
						N				

Figure 22.

columns of 10 letters. The text can now be marked off into sections of proper lengths and, moreover, guided by the letters which must be at the heads of columns, the text can be inscribed in the rectangle in key order. For example, column 1 must end with the second group, R M G R N; column 2 therefore begins with E R. There is only one possibility, viz, the fourth column. This is a long column, and must therefore have 11 letters, making column 3 begin with R Y. This definitely fixes the position of the number 3 in the key, and so on. The solution is reached after only a very few moments and is as shown in figure 23.

REF ID:A59439

3	9	6	2	4	7	1	11	5	10	8
R	E	F	E	R	R	I	N	G	T	O
Y	O	U	R	N	U	M	B	E	R	S
E	V	E	N	W	H	A	T	D	I	S
P	O	S	I	T	I	O	N	H	A	S
B	E	E	N	M	A	D	E	O	F	C
R	Y	P	T	O	G	R	A	P	H	I
C	E	Q	U	I	P	M	E	N	T	O
F	M	E	S	S	A	G	E	C	E	N
T	E	R	F	O	U	R	T	H	P	R
O	V	I	S	I	O	N	A	L	B	R
I	G	A	D	E						

FIGURE 23.

d. 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.—*a.* 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. Obviously such a cryptogram will be difficult if not impossible to decipher at the other end, and a repetition is 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 20*b* had first been transmitted as follows:

CRYPTOGRAM

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

b. The column which was omitted is E R N I N T U S F 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 numerical 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 transcription process, and likely as not a repetition will be requested by the addressee. 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.

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

FIRST VERSION

```
O D N I L   N T T H D   G S O H A   O O Q S G   T E R P S
I N E N E   N F U E H   R W R R I   R A T P E   D E T A N
O O C O O   R O G I O   S
```

SECOND VERSION

```
O D N I L   N T T H D   G S O H A   O O Q S G   T E R N F
U E H R W   R P S I N   E N E R I   R A T P E   D E T A N
O O C O O   R O G I O   S
```

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

```
First version... O D N I L N T T H D G S O H A O O Q S G T E R[P S I N E N E]
Second version. O D N I L N T T H D G S O H A O O Q S G T E R[N F U E H R W
                [N F U E H R W R]R I R A T P E D E T A N O O C O O R O G I O S
                 R]P S I N E N E]R I R A T P E D E T A N O O C O O R O G I O S
```

FIGURE 24

d. The two versions are alike except for a pair of interchanged sequences; the bracketed sequence P S I N E N E 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 N F U E H R W R 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.

e. 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 [(2×8)+(1×7)=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

REF ID:A59439

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.

f. The text of the correct version is written out as seen in figure 25*a*. 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 25*b*. In the latter, column 7 has also been placed at the second interval to the right of column 5, because the latter yields good trigraphs with columns 3–4. Seeing the trigraph T R O near the bottom of columns 3–4–5 and the letters O and P in the same row, suggests the word TROOP. The columns are to be rearranged to make this word TROOP. There are

1	2	3	4	5	6	7	8
		a					c
	t	o				d	o
O	H	O	N	P	R	E	O
D	D	Q	F	S	I	T	R
N	G	S	U	I	R	A	O
I	S	G	E	N	A	N	G
L	O	T	H	E	T	O	I
N	H	E	R	N	P	O	O
T	A	R	W	E	E	C	S
T	O		R		D	O	

a

1	2	3	4	5	6	7	8
							c
			a		d		o
O	T	O	N	E	R	P	O
D	H	O	F	T	I	S	R
N	D	Q	U	A	R	I	O
I	G	S	E	N	A	N	G
L	S	G	H	O	T	E	I
N	O	T	R	O	P	N	O
T	H	E	W	C	E	E	S
		A	R	R	O	D	

b

3	4	5	2	6
O	N	E	T	R
O	F	T	H	I
Q	U	A	D	R
S	E	N	G	A
G	H	O	S	T
T	R	O	O	P
E	W	C	H	E
R	R	O		D

c

3	4	7	2	6	8	1	5
O	N	E	T	R	O	O	P
O	F	T	H	I	R	D	S
Q	U	A	D	R	O	N	I
S	E	N	G	A	G	I	N
G	H	O	S	T	I	L	E
T	R	O	O	P	O	N	N
E	W	C	H	E	S	T	E
R	R	O	A	D			

d

FIGURE 25.

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

23. Messages with similar beginnings.—*a.* In military correspondence it is often the case that somewhat similar instructions or information must be conveyed by a superior commander to several subordinate commanders simultaneously. Such a situation frequently results in the circumstance 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 encipherment, then it will result, in such cases as the foregoing, that the first two or more rows of the transposition 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 studying 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.

b. Here are two cryptograms which are assumed to have been intercepted within a few minutes of each other, the messages being addressed to two battalion commanders by the regimental commander.

CRYPTOGRAM 1

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

CRYPTOGRAM 2

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

c. The cryptanalyst, noting the similarities in the first few letters of the two messages, carefully compares the two texts, looking for additional identical sequences of letters between the cryptograms. For example, No. 1 begins with B N T S E and so does No. 2; after an interval of 4 letters in No. 1 and 5 letters in No. 2 he notes the identical sequences L C E T; after an interval of 5 letters in No. 1 and 5 letters in No. 2 he notes the identical sequences E R R O, and so on. The identities are underlined or marked in some distinctive manner throughout the texts, as shown in figure 26.

CRYPTOGRAM 1

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

CRYPTOGRAM 2

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

FIGURE 26.

d. Now it is obvious that these identities exist because the two messages begin alike, and by taking advantage of the identical portions in the cryptograms it will be possible to transcribe the texts of the latter into transposition rectangles which will not only have the identical portions in homologous positions, but also will show which are long columns, which are short. All that is necessary is to begin transcribing the texts on cross-section paper, in columns, arranging matters so that the identical sequences will fall at the tops of the columns. Thus, the first column of No. 1 will contain the letters B N T S E A R K C and the first column of No. 2 will contain the letters B N T S E I N D O T; the second column of No. 1 will contain the letters L C E T T N B I T and the second column of No. 2 will contain the letters L C E T S A F P L,

REF ID:A59439

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 27*a* 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 A E F T F S N P G N P in No. 1, and A E F T F S T O N T N 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 portions 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 27*b*.

1	2
B L E N T	B L E N T
N C R N O	N C R N O
T E R O K	T E R O K
S T O N M	S T O N M
E T T N S	E S M S F
A N A E Z	I A O T E
R B E N T	N F I I Y
K I L O G	D P S I K
C T T O N	O L O U P
	T E C

FIGURE 27*a*.

e. 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,

1														2													
1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9	10	11	12	13	14
B	L	E	N	T	Y	E	A	T	N	I	M	O	A	B	L	E	N	T	Y	E	A	T	N	I	M	O	A
N	C	R	N	O	I	F	R	F	O	O	M	U	G	N	C	R	N	O	I	F	R	F	O	O	M	U	G
T	E	R	O	K	T	T	W	C	E	S	A	T	A	T	E	R	O	K	T	T	W	C	E	S	A	T	A
S	T	O	N	M	D	F	O	T	W	T	O	I	A	S	T	O	N	M	D	F	O	T	W	T	O	I	A
E	T	T	N	S	K	S	I	O	X	I	A	O	P	E	S	M	S	F	V	S	A	T	X	T	O	E	A
A	N	A	E	Z	L	N	A	T	E	D	R	B	N	I	A	O	T	E	S	T	R	L	P	T	S	E	B
R	B	E	N	T	A	P	O	D	R	R	P	I	E	N	F	I	I	Y	I	O	O	T	V	F	C	L	I
K	I	L	O	G	N	G	F	N	F	R	A	E	I	D	P	S	I	K	N	N	E	A	T	O	A	S	T
C	T	T	O	N	A	N	G	I	A	R	T	O	K	O	L	O	U	P	T	T	E	E	I	C	N	O	R
						P		S						T		E		C	A	N	K	A	T		R		T

FIGURE 27*b*.

and columns 2, 4, 11, and 13 as another group, the exact positions occupied by these 6 columns are easily ascertained, as shown in figure 27*c*.

1														2													
7	10								2	11	13	4		7	10								2	11	13	4	
E	N								L	I	O	N		E	N								L	I	O	N	
F	O								C	O	U	N		F	O								C	O	U	N	
T	E								E	S	T	O		T	E								E	S	T	O	
F	W								T	T	I	N		F	W								T	T	I	N	
S	X								T	I	O	N		S	X								S	T	E	S	
N	E								N	D	B	E		T	P								A	T	E	T	
P	R								B	R	I	N		O	V								F	F	L	I	
G	F								I	R	E	O		N	T								P	O	S	I	
N	A								T	R	O	O		T	I								L	C	O	U	
P	S													N	T												

FIGURE 27*c*.

f. 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 27*d*.

1														2													
7	10	3	12	6	11	4	9	5	8	2	11	13	4	7	10	3	12	6	11	4	9	5	8	2	11	13	4
E	N	E	M	Y	B	A	T	T	A	L	I	O	N	E	N	E	M	Y	B	A	T	T	A	L	I	O	N
F	O	R	M	I	N	G	F	O	R	C	O	U	N	F	O	R	M	I	N	G	F	O	R	C	O	U	N
T	E	R	A	T	T	A	C	K	W	E	S	T	O	T	E	R	A	T	T	A	C	K	W	E	S	T	O
F	W	O	O	D	S	A	T	M	O	T	T	I	N	F	W	O	O	D	S	A	T	M	O	T	T	I	N
S	X	T	A	K	E	P	O	S	I	T	I	O	N	S	X	M	O	V	E	A	T	F	A	S	T	E	S
N	E	A	R	L	A	N	T	Z	A	N	D	B	E	T	P	O	S	S	I	B	L	E	R	A	T	E	T
P	R	E	P	A	R	E	D	T	O	B	R	I	N	O	V	I	C	I	N	I	T	Y	O	F	F	L	I
G	F	L	A	N	K	I	N	G	F	I	R	E	O	N	T	S	A	N	D	T	A	K	E	P	O	S	I
N	A	T	T	A	C	K	I	N	G	T	R	O	O	T	I	O	N	T	O	R	E	P	E	L	C	O	U
P	S													N	T	E	R	A	T	T	A	C	K				

FIGURE 27*d*.

24. Messages with similar endings.—*a*. What has been said at the beginning at 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

E T R T E E E S O A A E U N I V A F L N I A M N D R Y H R V M E N R I
E E T R O U D C C C O H T C Y M R R E A R H I T N D E Y E N R N E R V
S R B E N I G S K A I L N R A N F N A D A L O L T X O M A H H R R E I

No. 2

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

REF ID:A59439

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 T R O forming the final trigraph of the 8th group of No. 1 and finds a similar sequence forming the initial trigraph 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

E T R T E E E S 0 A A E U N I V A F L N I A M N D R T H R V M E N R I
(1) (2) (3) (4)
E E T R O U D C C C O H T C Y M R R E A R H I T N D E Y E N R N E R V
(5) (6) (7) (8)
S R B E N I G S K A I L N R A N F N A D A L O L T X O M A H H R R E I
(9) (10) (11) (12) (13)

No. 2

T L V S X O P N R E M E F D S K Y E N R U E E R B T S R E H T I A N T
(1) (2) (3) (4) (5)
I V Y M R V E S I R E E N E I N O L T M N N E D E T R O O P U N A R A
(6) (7) (8) (9) (10) (11)
C I A A I N S C W N A
(12) (13)

c. The identities between 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 crytogram 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 position 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, *viz*, 2, a cyclic sequence is generated from the series of equivalencies 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, ascertained 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 3 12 9 1 4 7 2 11 6 13 5. The completely deciphered messages are shown in figure 28.

No. 1

8	10	3	12	9	1	4	7	2	11	6	13	5
H	E	A	D	R	E	D	C	O	L	U	M	N
I	N	F	A	N	T	R	Y	A	N	D	A	R
T	I	L	L	E	R	Y	M	A	R	C	H	I
N	G	N	O	R	T	H	R	E	A	C	H	E
D	S	I	L	V	E	R	R	U	N	C	R	E
E	K	A	T	S	E	V	E	N	F	O	R	T
Y	A	M	X	R	E	M	A	I	N	H	E	R
E	I	N	O	B	S	E	R	V	A	T	I	O
N												

No. 2

8	10	3	12	9	1	4	7	2	11	6	13
I	N	F	A	N	T	R	Y	P	O	I	N	T
R	E	D	C	O	L	U	M	N	P	A	S	S
E	D	S	I	L	V	E	R	R	U	N	C	R
E	E	K	A	T	S	E	V	E	N	T	W	E
N	T	Y	A	M	X	R	E	M	A	I	N	H
E	R	E	I	N	O	B	S	E	R	V	A	T
I	O	N										

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.

25. Solution of a single message containing a long repetition.—*a.* Sometimes a lengthy phrase or a series of numbers (spelled out in letters) is repeated within a message and if the message is enciphered 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 (169 letters)

O E A E L T R S E D H N U F F R N R Y F N T A E D I L S M Y
(1a) (1b) (2a)
N C E T S L S T O C A W I A O T S L S S L E D H N O R I I S
(2b) (3a) (3b)
F E B N N U U P W E S S M Y E R C N N O R V T T A O G N U G
(4a) (4b) (5a)
G T I F E R S E O M S W E R N R A S T B O S A A A O S N O O
(5b) (6a) (6b)
I B O S D C A Y H L H O N E M S E T F Y H L A U X T A O G G
(7a) (7b) (8a) (8b)
P R S V L Y E E G G T I S S O U U P V
(9a) (9b)

REF ID:A59439

c. 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 19 letters and 2 columns of 18 letters. The first two segments may therefore be assigned the numbers 1*a* and 1*b*, since they come from column 1; the next two segments may be assigned the numbers 2*a* and 2*b*, since they come from column 2, and so on, as shown above. Identical segments may now be equated. Thus:

1*a*	2*a*	3*a*	4*a*	5*a*	6*a*	7*a*	8*a*	9*a*
3*b*	4*b*	2*b*	9*b*	8*b*	1*b*	6*b*	7*b*	5*b*

This gives rise to the cycle 1–3–2–4–9–5–8–7–6, which is a cyclic 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.

1	2	3	4	5	6	7	8	9
O	F	T	R	R	E	A	O	P
E	N	O	I	C	R	A	N	R
A	T	C	I	N	S	O	E	S
E	A	A	S	N	E	S	M	V
L	E	W	F	O	O	N	S	L
T	D	I	E	R	M	O	E	Y
R	I	A	B	V	S	O	T	E
S	L	O	N	T	W	I	F	E
E	S	T	N	T	E	B	Y	G
D	M	S	U	A	R	O	H	G
H	Y	L	U	O	N	S	L	T
N	N	S	P	G	R	D	A	I
U	C	S	W	N	A	C	U	S
F	E	L	E	U	S	A	X	S
F	T	E	S	G	T	Y	T	O
R	S	D	S	G	B	H	A	U
N	L	H	M	T	O	L	O	U
R	S	N	Y	I	S	H	G	P
Y		O	E	F	A		G	V

FIGURE 29.

4	6	9	1	5	3	8	2	7
R	E	P	O	R	T	O	F	A
I	R	R	E	C	O	N	N	A
I	S	S	A	N	C	E	T	O
S	E	V	E	N	A	M	A	S
F	O	L	L	O	W	S	E	N
E	M	Y	T	R	I	E	D	O
B	S	E	R	V	A	T	I	O
N	W	E	S	T	O	F	L	I
N	E	G	E	T	T	Y	S	B
U	R	G	D	A	S	H	M	O
U	N	T	H	O	L	L	Y	S
P	R	I	N	G	S	A	N	D
W	A	S	U	N	S	U	C	C
E	S	S	F	U	L	X	E	A
S	T	O	F	G	E	T	T	Y
S	B	U	R	G	D	A	S	H
M	O	U	N	T	H	O	L	L
Y	S	P	R	I	N	G	S	H
E	A	V	Y	F	O	G		

FIGURE 30.

e. 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 or 7–2. The cyclic permutation of the transposition key obtained in subparagraph *c* is 1–3–2–4–9–5–8–7–6. In order to bring the 2 and 7 adjacent in a sequence 2–7 or 7–2 one must take intervals of 5 and 4, respectively, and "decimate" the cycle, giving the following: 1–5–3–8–2–7–4–6–9 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 or 8–3–5–1–9–6–4–7–2. Only a few moments are necessary to establish the correctness of the former alternative and the solution is at hand. It is as shown in figure 30.

f. 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 paraphrasing the original text; or that 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.

26. Solution when several cryptograms of identical length and in the same key are available.—*a.* 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 mechanics of transposition as a cryptographic scheme, *viz*, that the essential feature of the transposition method consists merely in the alterations 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 lengths*, when transposed according to the same key, will undergo identical alterations in position in the course of encipherment, 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 is 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 funcrion of the number of letters or elements in the plain text itself.* For example, suppose that a plain-text message 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 2d plain-text letter, as the 68th, and so on; in another message of exactly 100 letters, enciphered by the same general system and specific key, it is obvious that the 1st plain-text letter must also appear as the 17th cipher-text letter, the 2d plain-text letter, as the 68th, and so on. In short, all correspondingly numbered plain-text letters in both messages will appear in identical positions in the cryptograms.

b. 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 superimposing the several cryptograms and *anagramming the columns* formed by the superimposition, for it is obvious that any circumstances which can be used as a guide for rearranging the letters in 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.

c. An example of the method involving the application of this general solution will now be given, using as a basis five messages assumed to have been enciphered 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.

REF ID:A59439

d. 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 all 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	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
Message No. 1	I	A	A	L	N	E	O	F	S	G	T	O	G	V	E	R	A	N	O	L	N	D	U	O	D	E	I	H	I	S	A	T
Message No. 2	T	D	N	M	R	G	R	E	O	N	A	R	I	E	U	E	T	N	Y	I	T	C	O	F	E	A	I	E	U	T	T	A
Message No. 3	A	N	E	L	N	E	X	E	H	G	I	L	A	C	E	M	E	E	N	L	F	X	T	E	E	E	I	S	I	G	A	O
Message No. 4	E	E	N	E	T	S	L	N	N	F	T	C	O	I	D	O	S	E	A	I	L	F	I	G	D	W	I	A	A	R	N	O
Message No. 5	R	A	M	E	T	M	I	O	N	O	D	I	U	M	A	L	L	I	N	X	O	A	T	G	T	N	N	A	I	B	T	N

Column No.	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
Message No. 1	F	T	D	N	R	L	V	O	R	O	D	S	W	E	E	R	O	R	Q
Message No. 2	R	D	T	E	D	N	S	O	E	I	P	E	C	M	F	E	A	R	N
Message No. 3	R	W	L	L	D	L	V	V	O	R	D	E	L	O	C	H	O	T	H
Message No. 4	I	H	N	L	L	N	R	F	V	W	L	R	E	M	R	A	I	E	A
Message No. 5	H	I	T	N	I	A	S	D	R	M	S	E	C	U	I	O	V	S	A

FIGURE 31.

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

51 23
Q U
N O
H T
A I
A T

FIGURE 32*a*.

10 51 23
G Q U
N N O
G H T
F A I
O A T

FIGURE 32*b*.

30 51 23
S Q U
T N O
G H T
R A I
B A T

FIGURE 33*a*.

30 51 23 31
S Q U A
T N O T
G H T A
R A I N
B A T T

FIGURE 33*b*.

30 51 23 31 22
S Q U A D
T N O T C
G H T A X
R A I N F
B A T T A

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 T 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 33*a*. Looking for a column with a D at the top (for SQUAD) and either an A (for BATTALION) or an L (for BATTLE), there is only one candidate, column 22, yielding the sequences shown in figure 33*b*. 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 picks 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.

f. 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.

Term number	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
C→P sequence	28	3	14	46	19	37	25	47	48	26	35	41	2	34	27	12	36	45	17	13	40	18	9	24	33	8	1	50	44	11
Message No. 1	H	A	V	E	O	R	D	E	R	E	D	R	A	T	I	O	N	W	A	G	O	N	S	O	F	F	I	R	S	T
Message No. 2	E	N	E	M	Y	D	E	F	E	A	T	E	D	D	I	R	E	C	T	I	O	N	O	F	R	E	T	R	E	A
Message No. 3	S	E	C	O	N	D	E	C	H	E	L	O	N	W	I	L	L	L	E	A	V	E	H	E	R	E	A	T	E	I
Message No. 4	A	N	I	M	A	L	D	R	A	W	N	V	E	H	I	C	L	E	S	O	F	E	N	G	I	N	E	E	R	T
Message No. 5	A	M	M	U	N	I	T	I	O	N	T	R	A	I	N	I	N	C	L	U	D	I	N	G	H	O	R	S	E	D

Term number	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
C→P sequence	30	51	23	31	22	16	7	21	32	42	10	49	4	43	15	5	39	29	20	38	6
Message No. 1	S	Q	U	A	D	R	O	N	T	O	G	O	L	D	E	N	V	I	L	L	E
Message No. 2	T	N	O	T	C	E	R	T	A	I	N	A	M	P	U	R	S	U	I	N	G
Message No. 3	G	H	T	A	X	M	X	F	O	R	G	O	L	D	E	N	V	I	L	L	E
Message No. 4	R	A	I	N	F	O	L	L	O	W	F	I	E	L	D	T	R	A	I	N	S
Message No. 5	B	A	T	T	A	L	I	O	N	M	O	V	E	S	A	T	S	I	X	A	M

FIGURE 34.

g. 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 expertness. 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 lengths, 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.

h. 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:

6 2 5 7 4 1 3
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
A

6 2 5 7 4 1 3
01 02 03 04 05 06 07
08 09 10 11 12 13 14
15 16 17 18 19 20 21
B

P→C sequence for A.... 06 13 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

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

FIGURE 35*a*

REF ID:A59439

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

A_____ 06 13 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_____ 06 13 20 □ 02 09 16 □ 07 14 21 □ 05 12 19 □ 03 10 17 □ 01 08 15 □ 04 11 18 □

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 encipherment. 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 is obvious.

i. 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:

6	2	5	7	4	1	3
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	

A

6	2	5	7	4	1	3
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			

B

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

FIGURE 35*b*

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 experimentation would 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.

j. The foregoing principles will naturally not be applicable to cases where two messages differ in length by but one letter and this small difference brings about a profound difference in the P→C sequences applicable to the messages. This is what happens often in the case of true double transposition,[1] 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 herein 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 16*c*, 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 36*a*.

Term number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
C→P sequence	28	3	14	46	19	37	25	47	48	26	35	41	2	34	27	12	36	45	17	13	40
	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
	18	9	24	33	8	1	50	44	11	30	51	23	31	22	16	7	21	32	42	10	49
	43	44	45	46	47	48	49	50	51												
	4	43	15	5	39	29	20	38	6												

Term number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
P→C sequence	27	13	2	43	46	51	37	26	23	41	30	16	20	3	45	36	19	22	5	49	38
	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
	35	33	24	7	10	15	1	48	31	34	39	25	14	11	17	6	50	47	21	12	40
	43	44	45	46	47	48	49	50	51												
	44	29	18	4	8	9	42	28	32												

FIGURE 36*a*

c. (1) Since there appears to be no constant difference between successive terms in the P → C sequence in figure 36*a*, single columnar transposition is ruled out and double transposition is assumed to have been employed. In passing, it is worthwhile noting that the reconstruction of the keys employed in the case of true double transposition is quite important, because it is often the case that concentrated effort directed toward the cryptanalysis of one or more messages and the subsequent recovery of the transposition keys will, of course, greatly facilitate the reading of all other messages in the same keys.

(2) There are at least four methods suited to the purpose and they will be dealt with in an order most conducive to their comprehension by the student.

(3) A preliminary to the reconstruction of the keys in the case of each of the four methods to be studied consists in establishing or ascertaining the width of either the T–1 or the T–2 matrix, usually the former, because it is easier to do.

[1] See Special Text No. 166, *Advanced Military Cryptography*, sec. IV.

REF ID:A59439

(4) As in paragraph 26*h*, 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.

d. 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 37*a:*

6	2	7	1	5	3	8	4
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
49	50	51					

T–1

3	9	1	7	4	2	11	8	10	6	5
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	01	09
17	25	33	41	49	03	11	19	27	35	43
51	07	15	23	31	39	47				

T–2

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

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

FIGURE 37*a.*

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 this P → C sequence (20 06 48 33 15), set them under the 29th to 34th terms (28 14 05 41 23), and find the difference between superimposed numbers. (When the minuend is less than the subtrahend the superimposed terms are disregarded.) Thus:

29th to 34th terms	28	14	05	41	23
1st to 6th terms	20	06	48	33	15
Differences	8	8		8	8

There is a constant difference between the superimposed terms. The reason for its appearance is not hard to understand if reference is made to figure 37*a* and the matter is studied in the light of the mechanics of the method of encipherment. As for the two terms 28 and the 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 in this case 8 because the superimposed terms happen to be *sequent* in the *columns* in which they fall in the T–1 matrix. If the superimposed 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:

6th to 11th terms	44	30	21	03	39
29th to 34th terms	28	14	05	41	23
	16	16	16		16

Here the difference, 16, is a multiple of 8 because the superimposed terms are separated by one row in the T–1 matrix, as can be seen by referring to figure 37*a*.

e. The foregoing phenomena afford a method of ascertaining the width of the T–1 matrix in an unknown case, and, as noted above, this constitutes the first step in recovering the transposition key or keys. For if a study be made of the terms of the P → C sequence in figure 36*a*, based upon finding sections thereof which show a constant difference, the latter will correspond to either the width of the T–1 matrix or a multiple of the width. An easy way to make this study is to take a section of the P → C sequence in figure 36*a*, add 5, 6, 7, . . . (in successive steps) to each term of the selected section, and then look for repetitions between the original P → C sequence and the P → C sequence plus the additive. A beginning will be made with an assumption of a T–1 matrix of 5 columns. Since the cryptograms contain only 51 letters, all totals greater than 51 will be disregarded. Hence it is best to take a section which has a long series of low numbers, so that when the additive is applied the majority of the totals will not exceed 51. Such a series is the following (only one term in it, the 29th, is close to the maximum):

Term number	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
P → C sequence	38	35	33	24	07	10	15	01	48	31	34	39	25	14	11	17	06
P → C sequence +5	43	40	38	29	12	15	20	06		36	39	44	30	19	16	22	11

Searching for repetitions between the P→C sequence and the P→C sequence +5, the results are negative. Trial is then made of additives 6 to 11, inclusive, with similar negative results. When an additive of 12 is applied, however, the results obtained give positive indication that the T–1 matrix is 12 columns in width. Thus:

Term number	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
P→C sequence	38	35	33	24	07	10	15	01	48	31	34	39	25	14	11	17	06
P→C sequence +12	50	47	45	36	19	22	27	13		43	46	51	37	26	23	29	18

It will be seen, on referring to figure 36*a*, that the following repetitions (with the term numbers in each of the sequences indicated) are present:

Term no. in P→C sequence + 12	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
Repetitions	50	47	45	36	19	22	27	13	02	43	46	51	37	26	23	29	18
Term no. in P→C sequence	38	39	15	16	17	18	01	02	03	04	05	06	07	08	09	44	45

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.

REF ID:A59439

f. (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 differences between successive terms in the P→C sequence of figure 37*a* are calculated and set down as shown below, and then repetitions are sought in the series of differences, the latter constituting what will hereinafter be termed the *P→C interval sequence.* Thus:

Term number	01	20	03	04	05	06	07	08	09	10	11	12	13	14	15
P→C sequence	20	06	48	33	15	44	30	21	03	39	04	42	32	17	51
P→C interval sequence		−14	+42	−15	−18	+29	−14	−9	−18	+36	−35	+38	−10	−15	+34	−15

Term number	16	17	18	19	20	21	22	23	24
P→C sequence	36	22	13	49	31	34	24	09	43
P→C interval sequence		−14	−9	+36	−18	+3	−10	−15	+34	−17

Term number	25	26	27	28	29	30	31	32	33	34	35	36	37		
P→C sequence	26	16	01	35	28	14	05	41	23	10	46	37	19		
P→C interval sequence		−10	−15	+34	−7	−14	−9	+36	−18	−13	+36	−9	−12	−7	

Term number	38	39	40	41	42	43	44	45	46	47	48	49	50	51	
P→C sequence	12	50	40	25	07	18	08	45	27	02	38	29	11	47	
P→C interval sequence		+38	−10	−15	−18	+11	−10	+37	−18	−25	+36	−9	−18	+36	

FIGURE 37*b*

Several repetitions are noted and underscored, in the same manner that ordinary repetitions are indicated in analogous cryptanalytic procedure. Now take the longest repetition, −14−9+36 −18, and find the terms from which it originates in the P→C sequence; a constant difference of 8 will be found. Thus:

(Term numbers 16–20)	36	22	13	49	31
(Term numbers 29–33)	28	14	05	41	23
Differences	8	8	8	8	8

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 37*a*, 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 encipherment.

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

Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
P→C sequence	27	13	02	43	46	51	37	26	23	41	30	16	20	03	45
P→C interval sequence		−14	−11	+41	+3	+5	−14	−11	−3	+18	−11	−14	+4	−17	+42	−9

Term number	16	17	18	19	20	21	22	23	24
P→C sequence	36	19	22	05	49	38	35	33	24
P→C interval sequence		−17	+3	−17	+44	−11	−3	−2	−9	−17

Term number	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
P→C sequence	07	10	15	01	48	31	34	39	25	14	11	17	06	50	47
P→C interval sequence		+3	+5	−14	+47	−17	+3	+5	−14	−11	−3	+8	−11	+44	−3	−26

Term number	40	41	42	43	44	45	46	47	48	49	50	51
P→C sequence	21	12	40	44	29	18	04	08	09	42	28	32
P→C interval sequence		−9	+28	+4	−15	−11	−14	+4	+1	+33	−16	+4

FIGURE 36*b*

Taking the sections of the P→C sequence from which the longest repetition arises and finding the constant difference between the terms involved, a width of 12 for the T–1 matrix is indicated. Thus:

(Term numbers 04–09)	43	46	51	37	26	23
(Term numbers 30–35)	31	34	39	25	14	11
Differences	12	12	12	12	12	12

This is identical with the results found by the other method. The T–1 matrix for the messages of paragraph 26*d* is therefore 12 columns in width.

g. Having ascertained the width of the T–1 matrix, the next step is to ascertain whether the width of the T–2 matrix is the same as that for the T–1, or different. If the same, the suspicion is warranted that the transposition keys for both matrices may be identical, in which case it is necessary to recover but one key. If the widths of the two matrices are different, then it is obvious that two different transposition keys are involved. Having ascertained the widths of both matrices, one can proceed to reconstruct the transposition key or keys which apply thereto. There are, as stated once before, at least four methods suitable for this purpose. They will now be taken up in turn, and each method will be explained in detail.

h. (1) In explaining the first method the discussion will be initiated with a reconsideration of figure 37*a*. If the C→P sequence established in that figure is treated as though it were plain text and enciphered by the double transposition method, using the same two transposition keys as before, an interesting phenomenon is observed. Not the following (fig. 37*c*):

6	2	7	1	5	3	8	4
27	47	09	11	31	02	42	44
23	34	50	38	18	30	05	26
14	43	37	01	08	17	33	22
41	25	46	29	49	07	20	13
04	21	28	16	36	48	10	40
32	12	24	06	45	35	51	03
19	39	15					

T–1

3	9	1	7	4	2	11	8	10	6	5
11	38	01	29	16	06	47	34	43	25	21
12	39	02	30	17	07	48	35	44	26	22
13	40	03	31	18	08	49	36	45	27	23
14	41	04	32	19	09	50	37	46	28	24
15	42	05	33	20	10	51				

T–2

FIGURE 37*c*.

Here it is seen that the numbers in the *columns* of the T–2 matrix are *consecutive*. Obviously, if the columns of this 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, . . . , inscribed within a T–1 matrix in the normal fashion. Thus (fig. 37*d*):

6	2	7	1	5	3	8	4
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
49	50	51					

FIGURE 37*d*.

REF ID:A59439

The result of these three steps has been to effect a return to the original T–1 matrix containing plain text in normal sequence. The principal point to be noted here is the appearance of the T–2 matrix in figure 37*c*, 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 37*a* again will be employed. Let the C→P sequence be inscribed within a matrix of 8 columns (the width having been established by one of the methods set forth in subpars. *d*–*f*). Thus:

1	2	3	4	5	6	7	8
27	47	09	11	31	02	42	44
23	34	50	38	18	30	05	26
14	43	37	01	08	17	33	22
41	25	46	29	49	07	20	13
04	21	28	16	36	48	10	40
32	12	24	06	45	35	51	03
19	39	15					

FIGURE 37*e*

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) | 11 38 01 29 16 06 |

FIGURE 37*f* (1).

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

Step (2)
| 11 38 01 29 16 06 |
| 02 30 17 07 48 35 |

FIGURE 37*f* (2).

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)
| 11 38 01 29 16 06 |
| 02 30 17 07 48 35 |
| 44 26 22 13 40 03 |

Step (4)
| 11 38 01 29 16 06 |
| 02 30 17 07 48 35 |
| 44 26 22 13 40 03 |
| 27 23 14 41 04 32 19 |

Step (5)
| 11 38 01 29 16 06 |
| 02 30 17 07 48 35 |
| 44 26 22 13 40 03 |
| 27 23 14 41 04 32 19 |
| 42 05 33 20 10 51 |

FIGURE 37*f* (3) (4) (5).

The numbers 01 to 05, inclusive, here have 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 06 and 07 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 |
| 02 30 17 07 48 35 |
Step (6) | 44 26 22 13 40 03 | 31 18 08 49 36 45 |
| 27 23 14 41 04 32 19 |
| 42 05 33 20 10 51 |

| 11 38 01 29 16 06 |
| 02 30 17 07 48 35 |
Step (7) | 44 26 22 13 40 03 | 31 18 08 49 36 45 |
| 27 23 14 41 04 32 19 | 09 50 37 46 28 24 15 |
| 42 05 33 20 10 51 |

| 11 38 01 29 16 06 |
| 47 34 43 25 21 12 39 | 02 30 17 07 48 35 |
Step (8) | 44 26 22 13 40 03 | 31 18 08 49 36 45 |
| 27 23 14 41 04 32 19 | 09 50 37 46 28 24 15 |
| 42 05 33 20 10 51 |

FIGURE 37*f* (6) (7) (8).

The process is continued in this manner until, as shown in figure 37*f*(9), all the numbers of the C→P sequence have been placed. (Here the last number is 51.)

| 11 38 01 29 16 06 | 47 34 43 25 21 12 39 |
| 47 34 43 25 21 12 39 | 02 30 17 07 48 35 | 44 26 22 13 40 03 |
Step (9) | 44 26 22 13 40 03 | 31 18 08 49 36 45 | 27 23 14 41 04 32 19 |
| 27 23 14 41 04 32 19 | 09 50 37 46 28 24 15 |
| 09 50 37 46 28 24 15 | 42 05 33 20 10 51 |

FIGURE 37*f* (9).

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 exact outlines of the matrix may now readily be found by referring to the matrix in figure 37*e*. 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. Thus:

3	9	1	7	4	2	11	8	10	6	5	
						11	38	01	29	16	06	47	34	43	25	21	12	39
47	34	43	25	21		12	39	02	30	17	07	48	35	44	26	22	13	40	03
44	26	22				13	40	03	31	18	08	49	36	45	27	23	14	41	04	32	19
27	23					14	41	04	32	19	09	50	37	46	28	24	15				
09	50	37	46	28	24	15	42	05	33	20	10	51									

FIGURE 37*g*.

324155—41——5

REF ID:A59439

Reference to figure 37*c* will show the exact correspondence between the T–2 transposition key and the T–2 matrix of figure 37*g* with these same elements indicated in figure 37*c*. The transposition key for the T–1 matrix in figure 37*e* can now easily be derived from figure 37*g*. It must be as follows:

6	2	7	1	5	3	8	4
27	47	09	11	31	02	42	44
23	34	50	38	18	30	05	26
14	43	37	01	08	17	33	22
41	25	46	29	49	07	20	13
04	21	28	16	36	48	10	40
32	12	24	06	45	35	51	03
19	39	15					

FIGURE 37*h*.

This transposition key and T–1 matrix are identical with the key and T–1 matrix of figure 37*c*.

i. Note the application of the foregoing method to the C→P sequence shown in figure 36*a* in connection with the messages solved in paragraph 26*d*–*e*. It has already been found that the width of the T–1 matrix is 12. The C→P sequence of figure 36*a* is therefore inscribed within a matrix of 12 columns.

1	2	3	4	5	6	7	8	9	10	11	12
28	03	14	46	19	37	25	47	48	26	35	41
02	34	27	12	36	45	17	13	40	18	09	24
33	08	01	50	44	11	30	51	23	31	22	16
07	21	32	42	10	49	04	43	15	05	39	29
20	38	06									

FIGURE 36*c*.

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	
48	40	23	15	28	02	33	07	20	37	45	11	49
	41	24	16	29	03	34	08	21	38	46	12	50	42
		25	17	30	04	35	09	22	39	47	13	51	43
		26	18	31	05								

FIGURE 36*d*.

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:

			4	7	1	8	2	5	9	11	3	12	10	6			
			14	27	01	32	06	19	36	44	10	48	40	23	15		
48	40	23	15	28	02	33	07	20	37	45	11	49	41	24	16	29	
	41	24	16	29	03	34	08	21	38	46	12	50	42	25	17	30	04
		25	17	30	04	35	09	22	39	47	13	51	43	26	18	31	05
		26	18	31	05												

FIGURE 36*e*.

The transposition key for the T–1 matrix is now found to be as indicated at the top of figure 36*e*.

j. A second method for reconstructing the keys will now be explained. To demonstrate this method the data afforded by figure 37*a* 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 37*b*, which gave rise to the thrice-appearing repetition in the interval sequence (−10 −15 +34) are found to be as follows:

1st appearance (term numbers 12–16)	42	32	17	51
2d appearance (term numbers 21–25)	34	24	09	43
3d appearance (term numbers 25–29)	26	16	01	35

FIGURE 37*i*.

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

26	16	01	35
34	24	09	43
42	32	17	51

FIGURE 37*j*.

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 37*a*. The columns of the structure in figure 37*j* are seen to be portions of the T–1 matrix, lying in the following positions in that matrix:

01	.	.	.	.	.	.	.	.	.
09	.	.	.	.	.	.	.	.	16
17	.	.	.	.	.	.	.	.	24
.	26	.	.	.	.	.	.	.	32
.	34	35	.	.	.	.	.	.	.
.	42	43	.	.	.	.	.	.	.
.	.	51							

FIGURE 37*k*.

In the T–2 matrix these numbers fall in the following positions:

.	.	.	.	.	.	.	.	.	26	34
42	.	.	.	.	.	.	.	.	16	24
32	.	.	.	.	.	.	.	.	01	09
17	.	.	.	.	.	.	.	.	35	43
51	.	.	.	.	.	.				

FIGURE 37*l*.

Now if the dimensions of the T–2 matrix were *unknown*, these numbers could nevertheless be placed in a skeletonized T–2 matrix as follows:

26	34	42
16	24	32
01	09	17
35	43	51

FIGURE 37*m*(1).

REF ID:A59439

and the block of numbers could be extended on both sides by referring to the T–1 matrix in figure 37*a*. Thus:

			02	10	18	26	34	42	50									
					08	16	24	32	40	48								
						01	09	17	25	33	41	49						
		03	11	19	27	35	43	51										

FIGURE 37*m*(2).

This structure may next be extended by referring to the P→C sequence in figure 37*a*, 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	01	09	17	25	33	41	49							
	03	11	19	27	35	43	51	07	15	23	31							

FIGURE 37*m*(3).

Again the 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
		06	14	22	30	38	46	08	16	24	32	40	48	05	13	21	29	37	45
			05	13	21	29	37	45	01	09	17	25	33	41	49				
					03	11	19	27	35	43	51	07	15	23	31	39	47		

FIGURE 37*m*(4).

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
42	50	06	14	22	30	38	46	08	16	24	32	40	48	05	13	21	29	37	45
32	40	48	05	13	21	29	37	45	01	09	17	25	33	41	49				
17	25	33	41	49	03	11	19	27	35	43	51	07	15	23	31	39	47		
51	07	15	23	31	39	47													

FIGURE 37*m* (5).

By referring to the T–1 matrix of figure 37*a* and the foregoing structure, the key for T–1 can be recovered. It is 6–2–7–1–5–3–8–4. By referring to the P→C sequence in figure 37*a* 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.

k. Applying this method to the messages solved in paragraph 26*d–e*, the steps are as follows: The width of the T–1 matrix has been established as being 12. The P→C interval sequence in figure 36*b* 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		
		27	13	02	
31	34	39	25	14	11
43	46	51	37	26	23

FIGURE 36*f*.

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

1	2	3	4	5	6	7	8	9	10	11	12
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
49	50	51									

FIGURE 36*g*.

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

	07	31	43
	10	34	46
15	27	39	51
01	13	25	37
	02	14	26
		11	23

FIGURE 36*h*(1).

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

	1	2	3	4	5	6	7
1		07	19	31	43		
2		10	22	34	46		
3	03	15	27	39	51		
4		01	13	25	37	49	
5			02	14	26	38	50
6				11	23	35	47

FIGURE 36*h*(2).

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

	03	15	27	39	51		
	01	13	25	37	49		
		02	14	26	38	50	
07	19	31	43	11	23	35	47
10	22	34	46				

FIGURE 36*h*(3).

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

35	03	15	27	39	51	05	17
33	45	01	13	25	37	49	06
24	36	48	02	14	26	38	50
07	19	31	43	11	23	35	47
10	22	34	46				

FIGURE 36*h*(4).

REF ID:A59439

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	
33	45	01	13	25	37	49	06	18	30	42	12	24	36	48
24	36	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
10	22	34	46											

FIGURE 36*h*(5).

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: 4–7–1–8–2–5–9–11–3–12–10–6. 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.

l. (1) A third method for reconstructing the transposition keys will now be set forth. It will first be explained in connection with the artificial example in figure 37*a*. 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 37*a* and subparagraph *e* give an indicated width of 8 for the T–1 matrix in this case.

(2) Let the additive 8 (found in subpars. *d* and *f*) be applied to the entire P→C sequence of figure 37*a*, and then let the identities between the two sequences be underscored and numbered, as shown in figure 37*n*:

(A) P→C sequence

20 06 48 33 15 | 44 30 21 03 39 | 04 | 42 32 17 51 |
(1) (2) (3)

36 22 13 49 31 | 34 24 09 | 43 | 26 16 01 35 |
(4) (5) (6)

28 14 05 41 23 | 10 46 37 19 | 12 | 50 40 25 07 |
(7) (8) (9)

18 08 45 27 | 02 38 29 11 47 |
(10) (11)

(B) P→C sequence+8

28 14 56 41 23 | 52 38 29 11 47 | 12 | 50 40 25 59 |
(1) (2) (3)

44 30 21 57 39 | 42 32 17 51 | 34 24 09 | 47
(4) (5) (6)

36 22 13 49 31 | 18 54 45 27 | 20 58 48 33 15 |
(7) (8) (9)

26 16 53 35 | 10 46 37 19 55 |
(10) (11)

FIGURE 37*n*.

If now the procedure explained in paragraph 16*k*, 24*c* to *f*, and 25*c* to *e* is applied to the repetitions noted in figure 37*n*, 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: [2]

B	1	2	3	4	5	6	7	8	9	10	11
A	7	11	9	2	3	5	4	10	1	6	8
Chain	1	7	4	2	11	8	10	6	5	3	9

[2] It is to be noted that the B sequence (that is, the P→C sequence plus the additive) must be used as the base, otherwise the chain of equivalents will be a *reversal* of the correct chain.

This is a cyclic 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 transcription 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. This will be noted in the following demonstration of the recovery of the matrices and keys for the messages solved in paragraph 26*d* and *e* by anagramming. It has already been shown how the width of the T–1 matrix was ascertained as being 12 columns (subpar. *f*). The additive 12 is then applied to the entire P→C sequence, identities 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 cyclic permutation of the transposition key for the T–2 matrix. Thus (fig. 36*i*):

	1					2				3			
A. (P→C sequence)	27	13	02	43	46	51	37	26	23	41	30	16	20
B. (P→C sequence+12)	39	25	14	55	58	63	49	38	35	53	42	28	32
	1					2				3			

	4					5				6			
	03	45	36	19	22	05	49	38	35	33	24	07	10
	15	57	48	31	34	17	61	50	47	45	36	19	22
	4					5				6			

	7					8				9			
	15	01	48	31	34	39	25	14	11	17	06	50	47
	27	13	60	43	46	51	37	26	23	29	18	62	59
	7					8				9			

	10				11				12			
	21	12	40	44	29	18	04	08	09	42	28	32
	33	24	52	56	41	30	16	20	21	54	40	44
	10				11				12			

FIGURE 36*i*.

B	1	2	3	4	5	6	7	8	9	10	11	12
A	8	5	12	7	9	4	1	2	11	6	3	10
Chain	1	8	2	5	9	11	3	12	10	6	4	7

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 4–7–1 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 26*d* may be seen in the following (figure 36*j*):

REF ID:A59439

4	7	1	8	2	5	9	11	3	12	10	6
1 H	2 A	3 V	4 E	5 O	6 R	7 D	8 E	9 R	10 E	11 D	12 R
13 A	14 T	15 I	16 O	17 N	18 W	19 A	20 G	21 O	22 N	23 S	24 O
25 F	26 F	27 I	28 R	29 S	30 T	31 S	32 Q	33 U	34 A	35 D	36 R
37 O	38 N	39 T	40 O	41 G	42 O	43 L	44 D	45 E	46 N	47 V	48 I
49 L	50 L	51 E									

T–1

4	7	1	8	2	5	9	11	3	12	10	6
3 V	16 I	27 I	39 T	51 E	5 O	17 N	29 S	41 G	9 R	21 O	33 U
45 E	1 H	13 A	25 F	37 O	49 L	6 R	18 W	30 T	42 O	12 R	24 O
36 R	48 I	2 A	14 T	26 F	38 N	50 L	4 E	16 O	28 R	40 O	7 D
19 A	31 S	43 L	11 D	23 S	35 D	47 V	8 E	20 G	32 Q	44 D	10 E
28 N	34 A	46 N									

T–2

Cryptogram	I	A	A	L	N		E	O	F	S	G	etc.
P→C sequence	27	13	2	43	46		51	37	26	23	41	etc.

FIGURE 36*j*.

m. (1) A fourth and possibly the most elegant solution to the problem of reconstructing the keys for double transposition will now be presented.[3] Reference will be made to the two matrices and keys shown in figure 36*j*. Let the P→C_1 and P→C_2 sequences resulting from the first and the second transpositions, respectively, be shown, as seen below:

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
2. P→C_1 sequence	03	15	27	39	51	05	17	29	41	09	21	33	45	01	13	25	37
P→C_2 sequence	27	13	02	43	46	51	37	26	23	41	30	16	20	03	45	36	19

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
49	06	18	30	42	12	24	36	48	02	14	26	38	50	04	16	28
22	05	49	38	35	33	24	07	10	15	01	48	31	34	39	25	14

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
40	07	19	31	43	11	23	35	47	08	20	32	44	10	22	34	46
11	17	06	50	47	21	12	40	44	29	18	04	08	09	42	28	32

FIGURE 36*k*

[3] The basic principles underlying this fourth and most important method were discovered and first presented 1934 by Solomon Kullback, Ph. D., then Junior Cryptanalyst, S. I. S.

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, $\frac{01}{03}$ 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.

(2) Given only the third row of numbers in figure 36*k*, that is, the P→C_2 sequence (which has heretofore 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_1 sequence. The width of the T–1 matrix can be ascertained by either of the two methods indicated in subparagraphs *e* 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 36*j*, but without the transposition key or the letters in the cells. Thus:

1	2	3	4	5	6	7	8	9	10	11	12
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
49	50	51									

FIGURE 36*l*

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_1 sequence would be as indicated herewith:

1. Term number	01	02	03	04	05
2. P→C_1 sequence	01	13	25	37	49
3. P→C_2 sequence	27	13	02	43	46

Two "conflicts" or contradictions are at once manifested: $\frac{01}{01}$ in rows 1 and 2, $\frac{01}{27}$ in rows 2 and 3; also, $\frac{02}{13}$ in rows 1 and 2, $\frac{13}{13}$ 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:

1. Term number	01	02	03	04	05
2. P→C_1 sequence	02	14	26	38	50
3. P→C_2 sequence	27	13	02	43	46

Here again two conflicts are noted: $\frac{01}{02}$ in rows 1 and 2, $\frac{26}{02}$ in rows 2 and 3; $\frac{02}{14}$ in rows 1 and 2, $\frac{02}{27}$ 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_1 sequence	03	15	27	39	51
3. P→C_2 sequence	27	13	02	43	46

REF ID:A59439

Here there are no contradictions and one check or corroboration: $\frac{03}{27}$ in rows 1 and 2, $\frac{03}{27}$ in rows 2 and 3. If key number 1 really occupies the 3d position in the key, then the superimposition data given in the last set of rows of superimposed numbers may be employed, by transferring the data to the proper positions in the skeletonized figure 36*m*(1):

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
2. P→C_1 sequence	03	15	27	39	51	05								01			
3. P→C_2 sequence	27	13	02	43	46	51	37	26	23	41	30	16	20	03	45	36	19

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
									02					04		
22	05	49	38	35	33	24	07	10	15	01	48	31	34	39	25	14

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
11	17	06	50	47	21	12	40	44	29	18	04	08	09	42	28	32

FIGURE 36*m* (1)

It then becomes at once possible, by referring to the T–1 matrix, to insert more numbers in the P→C_1 sequence. Thus:

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
2. P→C_1 sequence	03	15	27	39	51	05	17	29	41					01	13	25	37
3. P→C_2 sequence	27	13	02	43	46	51	37	26	23	41	30	16	20	03	45	36	19

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
49									02	14	26	38	50	04	16	28
22	05	49	38	35	33	24	07	10	15	01	48	31	34	39	25	14

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
40																
11	17	06	50	47	21	12	40	44	29	18	04	08	09	42	28	32

FIGURE 36*m* (2)

The new placements now permit of placing numbers in the P→C_1 sequence. For example, $\frac{07}{17}$ in rows 1 and 2 permit of placing the number 07 above the number 17 in the P→C_2 sequence; $\frac{08}{29}$ in rows 1 and 2 permit of placing the number 08 above the number 29 in the P→C_2 sequence, and so on. In only a few moments the entire P→C_1 sequence can be established. Thus:

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
2. P→C_1 sequence	03	15	27	39	51	05	17	29	41	09	21	33	45	01	13	25	37
3. P→C_2 sequence	27	13	02	43	46	51	37	26	23	41	30	16	20	03	45	36	19

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
49	06	18	30	42	12	24	36	48	02	14	26	38	50	04	16	28
22	05	49	38	35	33	24	07	10	15	01	48	31	34	39	25	14

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
40	07	19	31	43	11	23	35	47	08	20	32	44	10	22	34	46
11	17	06	50	47	21	12	40	44	29	18	04	08	09	42	28	32

FIGURE 36*n*.

(3) The determination of the T–1 key is now a very simple matter. Since it is known that the key has 12 numbers, it is only necessary to note in the P→C_1 sequence the relative order of the numbers 1 to 12. It is as follows:

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

This is merely the inverse of the actual key; the latter may be obtained by inversion. Thus:

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

Comparison of this key with the T–1 key shown in figure 36*j* will establish the identity of the two. The determination of the T–2 key is obvious, having the T–1 at hand. In this case both matrices and keys are identical.

n. Attention will be directed to a further interesting phenomenon in this case. Referring to figure 36*n*, if chains of equivalents are constructed between elements of the 1st and 3d rows only, the following two chains are obtained:

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

FIGURE 36*o*.

All the terms of the P→C_2 sequence are represented, except the number 24, which stands by itself. If now each of these chains is slid against itself, when properly juxtaposed, the superimposed pairs are identical with those in rows 1 and 2 in figure 36*n*. Note the following:

(1)
01 27 15 45 18 22 35 11 30 31 34 14 03 02 13 20 49 42 40 21 38 50 28
03 02 13 20 49 42 40 21 38 50 28 01 27 15 45 18 22 35 11 30 31 34 14

(2)
04 43 44 29 48 09 23 33 25 07 37 06 51 32 39 47 08 26 10 41 12 16 36 17 19 05 46
39 47 08 26 10 41 12 16 36 17 19 05 46 04 43 44 29 48 09 23 33 25 07 37 06 51 32

FIGURE 36*p*.

The application of the foregoing phenomena in the case under study is obvious. Here it is not even necessary to ascertain the width of the T–1 matrix before proceeding to try to establish the T–1 key. Of course, the number of chains which may be established will vary with the

6	2	7	1	5	3	8	4
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
49	50	51					

A (=T–1)

3	9	1	7	4	2	11	8	10	6	5
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	49	50	51				

B

3	9	1	7	4	2	11	8	10	6	5
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	01	09
17	25	33	41	49	03	11	19	27	35	43
51	07	15	23	31	39	47				

C (=T–2)

FIGURE 37*o*.

REF ID:A59439

specific matrices and keys, but the general principles herein 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.

o. (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 37*a* 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 37*o* and then prepare the T–2 matrix, shown at C.

(3) Write the P→C_1 sequence for T–1, under it write the P→C_1 for T–2, and under the latter write the P→C_2 sequence for the final cryptogram. Thus:

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14
2. P→C_1 sequence for T–1	04	12	20	28	36	44	02	10	18	26	34	42	50	06
3. P→C_1 sequence for T–2	03	14	25	36	47	06	17	28	39	50	01	12	23	34
4. P→C_2 sequence	20	06	48	33	15	44	30	21	03	39	04	42	32	17

15	16	17	18	19	20	21	22	23	24	25	26	27	28
14	22	30	38	46	08	16	24	32	40	48	05	13	21
45	05	16	27	38	49	11	22	33	44	10	21	32	43
51	36	22	13	49	31	34	24	09	43	26	16	01	35

29	30	31	32	33	34	35	36	37	38	39	40	41	42
29	37	45	01	09	17	25	33	41	49	03	11	19	27
04	15	26	37	48	08	19	30	41	02	13	24	35	46
28	14	05	41	23	10	46	37	19	12	50	40	25	07

43	44	45	46	47	48	49	50	51
35	43	51	07	15	23	31	39	47
09	20	31	42	07	18	29	40	51
18	08	45	27	02	38	29	11	47

FIGURE 37*p*.

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 were identical. It follows, therefore, that the principles elucidated under subparagraph *m* 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; an 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.

p. (1) The procedure will be illustrated by employing the P→C sequence in figure 37*b* (which is the same as that labelled P→C_2 sequence in figure 37*p*), it being assumed that nothing is known about the matrices, and that the sequence was obtained from a solution by superimposing and anagramming several messages of identical length.

(2) The width of the T–1 matrix is established as 8 and the T–1 matrix set down:

1	2	3	4	5	6	7	8
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
49	50	51					

T–1

FIGURE 37*q*.

(3) Assuming that key number 1 occupies the first position in the T–1 key, the numbers are inserted in row 2, representing the beginning of the P→C_1 sequence for T–1. The superimposed pairs in rows 1 and 2 are distributed in rows 3 and 4, with the results shown in figure 37*r*(1).

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14
2. P→C_1 sequence for T–1	01	09	17	25	33	41	49							
3. P→C_1 sequence for T–2				05										03
4. P→C_1 sequence	20	06	48	33	15	44	30	21	03	39	04	42	32	17

15	16	17	18	19	20	21	22	23	24	25	26	27	28
				07				02				01	
51	36	22	13	49	31	34	24	09	43	26	16	01	35

29	30	31	32	33	34	35	36	37	38	39	40	41	42
			06									04	
28	14	05	41	23	10	46	37	19	12	50	40	25	07

43	44	45	46	47	48	49	50	51
18	08	45	27	02	38	29	11	47

FIGURE 37*r* (1).

(4) An attempt is now made to construct a T–2 matrix which will produce the distribution and spacing of the numbers in row 3. For example, from the position of the number 05 the matrix would have to be of such dimensions that there are short columns of 2 letters and long columns of 3 letters; or short columns of 3 letters and long columns of 4 letters. The former hypothesis can be discarded at once, for the intervals between the numbers 03, 07, 02, 01, and 06 in row 3 make it untenable. The latter hypothesis may also be discarded, for the intervals between 03 and 07 and between 01 and 06 make it impossible. Hence key number 1 cannot occupy the first position in the T–1 key. Position 2 is assumed for key number 1 and the procedure repeated, also without good results. Note what happens when position 4 is assumed for key number 1 in the T–1 key:

REF ID:A59439

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14
2. P→C_1 sequence for T–1	04	12	20	28	36	44								
3. P→C_2 sequence for T–2	03					06					01			
4. P→C_2 sequence	20	06	48	33	15	44	30	21	03	39	04	42	32	17

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
		05												
	51	36	22	13	49	31	34	24	09	43	26	16	01	35

	29	30	31	32	33	34	35	36	37	38	39	40	41	42
	04									02				
	28	14	05	41	23	10	46	37	19	12	50	40	25	07

	43	44	45	46	47	48	49	50	51
	18	08	45	27	02	38	29	11	47

FIGURE 37r (2).

(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 03 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 { 7 (long) ×5=35
4 (short)×4=16 51

Key of 12 { 3 (long) ×5=15
9 (short)×4=36 51

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

1	2	3	4	5	6	7	8	9	10	11
01	02	03	04	05	06	07	08	09	10	11
12	13	14	15	16	17	18	19	10	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				

A

1	2	3	4	5	6	7	8	19	10	11	12
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
49	50	51									

B

FIGURE 37r (3).

(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 37r (2); these will permit of inserting numbers in row 2. No contradictions and many checks are found. Here is the diagram:

1. Term number	01	02	03	04	05	06	07	08	09	10	11	12	13	14
2. P→C_1 sequence for T–1	04	12	20	28	36	44						42	50	
3. P→C_1 sequence for T–2	03	14	25	36	47	06	17	28	39	50	01	12	23	34
4. P→C_1 sequence	20	06	48	33	15	44	30	21	03	39	04	42	32	17

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
	14	22	30						32	40		05	13	21
	45	05	16	27	38	49								
	51	36	22	13	49	31	34	24	09	43	26	16	01	35

	29	30	31	32	33	34	35	36	37	38	39	40	41	42
						17	25		41	49	03			
	04	15	26	37	48					02	13	24	35	46
	28	14	05	41	23	10	46	37	19	12	50	40	25	07

	43	44	45	46	47	48	49	50	51
			51	07		23	31	39	
	18	08	45	27	02	38	29	11	47

FIGURE 37r (4).

(7) In the first place note, in row 2, the constant difference 8, giving many corroborations that the width of the T–1 matrix is 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 rest follows quite readily, with the final result that figure 37r 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 stereotypic beginnings, so that the first few letters of a message or of 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.—*a*. When the double transposition system is employed in the field and is used 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 cryptographic 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 this 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 columnar transposition is not difficult and goes quite

REF ID:A59439

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.

c. (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 completely-filled rectangle has been employed, and moreover there is no need even to assume various widths.

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

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

1 2 3 4 5 6 7	1 2 3 4 5 6 7	2 6 1 5 3 7 4	2 6 1 5 3 7 4
U S R U O R E	U C T R N O E	C O U N T E R	H O S T I L E
C H G W E F E	S H I E T O L	H O S T I L E	F O R C E E N
T I A D R E M	R G A S O E D	G E R O A D S	C O U N T E R
R E S D D N N	U W D D N O E	W O U N D E D	E D O N R I D
N T O N N C N	O E R D N D I	E D O N R I D	G E R O A D S
O O E O D O V	R F E N C O E	F O R C E E N	E V E N M E N
E L D E I E E	E E M N N V E	E V E N M E N	W O U N D E D
FIGURE 39*a*.	FIGURE 39*b*.	FIGURE 39*c*.	FIGURE 39*d*.

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

d. 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 columns, and second, when the depth is a multiple of the width.

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

6	2	10	1	7	4	9	8	3	5
W 1	H 2	E 3	N 4	W 5	I 6	L 7	L 8	F 9	I 10
R 11	S 12	T 13	S 14	Q 15	U 16	A 17	D 18	R 19	O 20
N 21	R 22	E 23	A 24	C 25	H 26	G 27	O 28	L 29	D 30
E 31	N 32	V 33	I 34	L 35	L 36	E 37	T 38	O 39	N 40
I 41	G 42	H 43	T 44	A 45	D 46	V 47	I 48	S 49	E 50

T–1 Rectangle

6	2	10	1	7	4	9	8	3	5
N 4	S 14	A 24	I 34	T 44	H 2	S 12	R 22	N 32	G 42
F 9	R 19	L 29	O 39	S 49	I 6	U 16	H 26	L 36	D 46
I 10	O 20	D 30	N 40	E 50	W 1	R 11	N 21	E 31	I 41
W 5	Q 15	C 25	L 35	A 45	L 8	D 18	O 28	T 38	I 48
L 7	A 17	G 27	E 37	V 47	E 3	T 13	E 23	V 33	H 43

T–2 Rectangle

P→C sequence..... 34 39 40 35 37 14 19 20 15 17 32 36 31 38 33 2 6 1 8 3 42 46 41 48 43
Cryptogram..... I O N L E S R O Q A N L E T V H I W L E G D I I H

4 9 10 5 7 44 49 50 45 47 22 26 21 28 23 12 16 11 18 13 24 29 30 25 27
N F I W L T S E A V R H N O E S U R D T A L D C G

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

	Section					
Category A	1 —	4	9	10	5	7
	2 —	14	19	20	15	17
	3 —	24	29	30	25	27
	4 —	34	39	40	35	37
	5 —	44	49	50	45	47

	Section					
Category B	6 —	2	6	1	8	3
	7 —	12	16	11	18	13
	8 —	22	26	21	28	23
	9 —	32	36	31	38	33
	10 —	42	46	41	48	43

(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: [4]

1 2 3 4 5		6 7 8 9 10
S R O Q A	and	S U R D T

2 or 8	1 or 6	10	6 or 1	4	7	5	9	8 or 2	3
R	S	T	S	Q	U	A	D	R	O

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

1 2 3 4 5		6 7 8 9 10
I O N L E	and	N L E T V

2 or 8	1 or 6	10	6 or 1	4	7	5	9	8 or 2	3
O	I		N					E	

E N V I L L E T O N

Obviously the proper key for rearrangement is 8-6-10-1-4-7-5-9-2-3. By continuing this procedure the following additional rows of the T–1 rectangle are reconstructed.

1 2 3 4 5		6 7 8 9 10		8 6 10 1 4		7 5 9 2 3	
N F I W L	and	H I W L E	yields	W H E N W		I L L F I	
T S E A V	and	G D I I H	yields	I G H T A		D V I S E	
A L D C G	and	R H N O E	yields	N R E A C		H G O L D	

[4] The fact that the length of the sections corresponds to 5-letter groups has, of course, no bearing on the validity of the method. In this case it just happens that the rectangle contains 5 letters per column.

REF ID: A59439

The various rows are now assembled in sequence, giving the following:

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

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 10×5 or 5×10 suggest themselves. The process of finding cipher groups which form pairs on the same row is 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.

f. 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:

2	5	1	4	3
1 W	2 H	3 E	4 N	5 W
6 I	7 L	8 L	9 F	10 I
11 R	12 S	13 T	14 S	15 Q
16 U	17 A	18 D	19 R	20 O
21 N	22 R	23 E	24 A	25 C
26 H	27 G	28 O	29 L	30 D
31 E	32 N	33 V	34 I	35 L
36 L	37 E	38 T	39 O	40 N
41 I	42 G	43 H	44 T	45 A
46 D	47 V	48 I	49 S	50 E

T–1

2	5	1	4	3
3 E	8 L	13 T	18 D	23 E
28 O	33 V	38 T	43 H	48 I
1 W	6 I	11 R	16 U	21 N
26 H	31 E	36 L	41 I	46 D
5 W	10 I	15 Q	20 O	25 C
30 D	35 L	40 N	45 A	50 E
4 N	9 F	14 S	19 R	24 A
29 L	34 I	39 O	44 T	49 S
2 H	7 L	12 S	17 A	22 R
27 G	32 N	37 E	42 G	47 V

T–2

P→C sequence___13 38 11 36 15 40 14 39 12 37 3 28 1 26 5 30 4 29 2 27 23 48 21 46 25 50 24 49 22 47
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
18 43 16 41 20 45 19 44 17 42 8 33 6 31 10 35 9 34 7 32
D H U I O A R T A G L V I E I L F I L N

Taking the numbers of the P→C sequence and arranging them in sections of 10, the results are as follows:

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

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:

1	2	3	4	5	6	7	8	9	10
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
D	H	U	I	O	A	R	T	A	G
L	V	I	E	I	L	F	I	L	N

→

3	9	1	7	5
R	S	T	S	Q
W	H	E	N	W
N	R	E	A	C
U	A	D	R	O
I	L	L	F	I

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

g. 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 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.

REF ID:A59439

SECTION VI

PRINCIPLES OF MATRIX RECONSTRUCTION

	Paragraph
Special designs or geometric figures	29
Reconstruction of transposition matrix	30

29. Special designs or geometric figures.—*a.* 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 many 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 impractical for military correspondence in any case, so that no attention need be given them in this text.

b. 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 lengths 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.* Assume 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. The final section will be dealt with later.

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

	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
1	A	N	N	C	T	S	Y	O	A	A	C	E	M	E	H	I	E	I	B	I	H	A	D	E	X	T	C	T	U	R
2	W	D	H	E	B	R	N	D	T	T	D	I	Y	A	F	A	D	A	G	R	D	O	E	O	A	A	J	T	R	E
3	A	T	U	A	C	O	D	P	O	B	I	M	N	R	T	I	N	E	S	H	O	Y	N	F	L	I	H	N	R	O
4	M	O	O	C	E	O	I	B	R	S	E	P	Y	C	S	S	S	S	F											

	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
1	R	F	V	S	T	N	N	R	U	I	N	O	U	R	T	F	F	E	N	V	E	L	N	O	E
2	M	A	I	O	N	V	O	T	O	T	T	R	N	O	I	E	U	A	N	H	R	O	C	T	A
3	L	T	Y	E	X	A	E	U	O	A	E	F	R	T	E	X	Y	R	V	R	A	U	I	N	T

b. 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:

40	34	45	85	2	61	20	28	53	10	69	79	41	35	46	84	3	62	21	29	54	11
A	C	H	E	D	R	O	A	D	J	U	N	C	T	I	O	N	F	I	V	E	F
T	E	F	A	R	M	X	S	E	C	O	N	D	B	A	T	T	A	L	I	O	N
B	A	T	T	A	L	I	O	N	T	O	V	I	C	I	N	I	T	Y	O	F	H

c. Examining the numbers forming this partial C→P sequence, note the following sections of the sequence:

40	34	45	85	2	61	20	28	53	10	. . .
41	35	46	84	3	62	21	29	54	11	

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 cross-section paper so that consecutive numbers appear on the same level, as shown in figure 40–A:

C→P sequence		40	34	45	85	2	61	20	28	53	10	69	79	41	35	46	84	3	62	21	29	54	11
Text		A	C	H	E	D	R	O	A	D	J	U	N	C	T	I	O	N	F	I	V	E	F
		T	E	F	A	R	M	X	S	E	C	O	N	D	B	A	T	T	A	L	I	O	N
		B	A	T	T	A	L	I	O	N	T	O	V	I	C	I	N	I	T	Y	O	F	H
Matrix reconstruction diagram	1										10												11
	2									53		69										54	
	3								28				79								29		
	4	40						20						41						21			
	5		34				61								35				62				
	6			45		2										46		3					
	7				85												84						

FIGURE 40–A.

d. 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:

9	68	78	40	34	45	85	2	. . .
S	R	E	A	C	H	E	D	. . .
F	T	A	T	E	F	A	R	. . .
O	U	R	B	A	T	T	A	. . .

To the right of column 11 should come columns 70 80 42 36 47. Thus:

29	54	11	70	80	42	36	47	. . .
V	E	F	I	V	E	S	E	. . .
I	O	N	T	H	I	R	D	. . .
O	F	H	A	R	M	O	N	. . .

REF ID:A59439

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.

e. 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:

C→P sequence number	4	5	6	7	8	9	10	11	12	13	14	15	16
Term number	1	3	7	13	21	31	43	55	65	73	79	83	85

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:

C→P sequence number	48	49	50	51	52	53	54	55	56	57	58
Term number	2	6	12	20	30	42	54	64	72	78	82

C→P sequence number	64	65	66	67	68	69	70	71	72	73	74
Term number	4	8	14	22	32	44	56	66	74	80	84

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:

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

FIGURE 41.

f. 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.

REF ID:A59439

Term number	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
C→P sequence	4	48	5	64	24	49	6	65	75	17	25	50	7	66	76	38	59	18	26	51	8	67	77	39	33	1	60	19	27	52	9	68	78	40	34	45	85	2	61	20	28	53	10	69
Plain text	F	I	R	S	T	B	A	T	T	A	L	I	O	N	F	O	U	R	T	H	I	N	F	A	N	T	R	Y	H	A	S	R	E	A	C	H	E	D	R	O	A	D	J	U
Plain text	D	A	L	O	N	G	U	N	I	M	P	R	O	V	E	D	R	O	A	D	S	O	U	T	H	W	E	S	T	O	F	T	A	T	E	F	A	R	M	X	S	E	C	O
Plain text	S	E	V	E	N	S	I	X	E	I	G	H	T	A	X	P	R	O	P	O	S	E	Y	O	U	M	O	V	E	Y	O	U	R	B	A	T	T	A	L	I	O	N	T	O

Matrix reconstruction diagram

1	1 4		3 5				7 6						13 7								21 8										31 9										43 10		
2		2 48		4 64		6 49		8 65				12 50		14 66						20 51		22 67							30 52		32 68									42 53		44 69	
3					5 24				9 75		11 25				15 76				19 26				23 77					29 27				33 78								41 28				46
4										10 17						16 38		18 18						24 39				28 19					34 40						40 20				
5																	17 59								25 33		27 60								35 34				39 61				
6																										26 1										36 45		38 2						
7																																					37 85							

Term number	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
C→P sequence	26	38	50	1	3	7	13	21	31	43	55	65	73	79	83	85	10	18	28	40	52	62	70	5	11	19	29	41	53	63	71	77	25	35	47	59	69	16	24	34	46	58	68	76

FIGURE 40–B.

REF ID:A59439

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
79	41	35	46	84	3	62	21	29	54	11	70	80	42	36	47	63	22	30	55	12	71	81	43	37	23	31	56	13	72	82	44	32	57	14	73	83	58	15	74	16
N	C	T	I	O	N	F	I	V	E	F	I	V	E	S	E	V	E	N	X	E	N	E	M	Y	P	A	T	R	O	L	E	N	C	O	U	N	T	E	R	E
N	D	B	A	T	T	A	L	I	O	N	T	H	I	R	D	I	N	F	A	N	T	R	Y	N	O	W	A	T	R	O	A	D	J	U	N	C	T	I	O	N
V	I	C	I	N	I	T	Y	O	F	H	A	R	M	O	N	Y	V	I	L	L	E	A	N	D	W	A	I	T	F	U	R	T	H	E	R	I	N	S	T	R
										55 11										65 12								73 13							79 14				83 15		85 16
									54 54		56 70								64 55		66 71						72 56		74 72				78 57		80 73		82 58		84 74	
79								53 29				57 80						63 30				67 81				71 31				75 82		77 32				81 83				
	46 41						52 21						58 42				62 22						68 43		70 23					76 44										
		47 35				51 62								59 36		61 63								69 37																
			48 46		50 3										60 47																										
				49 84																																				
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
36	48	60	2	6	12	20	30	42	54	64	72	78	82	17	27	39	51	61	4	8	14	22	32	44	56	66	74	80	84	9	15	23	33	45	57	67	75	81	49	37

REF ID:A59439

				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-A.

g. 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.

REF ID:A59439

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.

h. 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.

i. 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 *levels* 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.

SECTION VII

SOLUTION OF GRILLES

	Paragraph
Revolving grilles	31
Solution of example	32
Concluding remarks on the solution of revolving grilles	33
Indefinite or continuous grilles	34

31. Revolving grilles.—*a.* In this type of grille [1] 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 undisclosed grid cells (or all but the central grid 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.[2]

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 [3] 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 cryptogram 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 at a time, the message is divided up into 5 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.

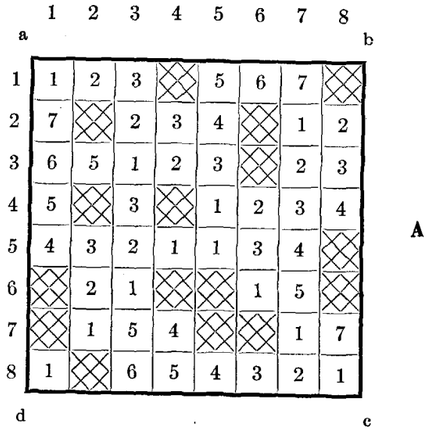
[3] Detailed instructions for the construction of revolving grilles will be found in Special Text No. 166, *Advanced Military Cryptography*, sec. V.

REF ID:A59439

MESSAGE

YOUR LINES TO THIS COMMAND POST CUT BY SHELL FIRE REQUEST YOU CHANGE THE ROUTE.

Grille: 8×8.



	1	2	3	4	5	6	7	8
1	1	2	3	■	5	6	7	■
2	7	■	2	3	4	■	1	2
3	6	5	1	2	3	■	2	3
4	5	■	3	■	1	2	3	4
5	4	3	2	1	1	3	4	■
6	■	2	1	■	■	1	5	■
7	■	1	5	4	■	■	1	7
8	1	■	6	5	4	3	2	1

a b c d

A

POSITION 1

1	2	3	Y 4	5	6	7	O 8
9	U 10	11	12	13	R 14	15	16
17	18	19	20	21	L 22	23	24
25	I 26	27	N 28	29	30	31	32
33	34	35	36	37	38	39	E 40
S 41	42	43	T 44	O 45	46	47	T 48
H 49	50	51	52	I 53	S 54	55	56
57	C 58	59	60	61	62	63	64

a b c d

B

POSITION 2

1	O 2	M 3	4	5	6	7	8
M 9	10	11	12	A 13	14	N 15	16
17	18	19	20	21	22	23	24
25	26	D 27	28	P 29	30	31	O 32
33	S 34	T 35	36	37	38	39	40
41	C 42	43	44	45	U 46	T 47	48
49	50	51	52	53	54	55	56
57	58	B 59	Y 60	61	62	63	S 64

p q r s

C

FIGURE 43

POSITION 3

1	2	3	4	5	6	H 7	8
9	10	E 11	L 12	13	14	15	L 16
F 17	18	19	I 20	R 21	22	23	E 24
R 25	26	27	28	29	30	31	32
33	34	35	36	E 37	38	Q 39	40
41	42	U 43	44	45	46	47	48
49	50	E 51	52	53	54	S 55	56
T 57	58	59	60	Y 61	62	63	64

o p q r

D

POSITION 4

O 1	2	3	4	U 5	C 6	7	8
9	10	11	12	13	14	15	16
17	H 18	A 19	20	21	22	N 23	24
25	26	27	28	29	G 30	E 31	32
T 33	34	35	H 36	37	E 38	39	40
41	42	43	44	45	46	47	48
49	R 50	51	O 52	53	54	55	U 56
57	58	59	60	61	T 62	E 63	64

a b c d

E

FINAL GRID

O 1	O 2	M 3	Y 4	U 5	C 6	H 7	O 8
M 9	U 10	E 11	L 12	A 13	R 14	N 15	L 16
F 17	H 18	A 19	I 20	R 21	L 22	N 23	E 24
R 25	I 26	D 27	N 28	P 29	G 30	E 31	O 32
T 33	S 34	T 35	H 36	E 37	E 38	Q 39	E 40
S 41	C 42	U 43	T 44	O 45	U 46	T 47	T 48
H 49	R 50	E 51	O 52	I 53	S 54	S 55	U 56
T 57	C 58	B 59	Y 60	Y 61	T 62	E 63	S 64

a b c d

F

CRYPTOGRAM

O O M Y U C H O M U E L A R N L F H A I R L N E R I D N P G
E O T S T H E E Q E S C U T O U T T H R E O I S S U T C B Y
Y T E S

FIGURE 43—Continued.

d. 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

REF ID:A59439

apertures concerned will disclose two cells which will also be occupied by a plain-text digraph, *but the letters composing 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 message) in cells 4 and 8. Now note that the symmetrically-corresponding cells disclosed when the grille is in position 3 are cells 57 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 reversed order in the plain text. For example, if the cryptogram of figure 43 is written out as explained above, the result is as shown at figure 44. Now the presence of the Q in position 39 suggests that it be combined with

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
O	O	M	Y	U	C	H	O	M	U	E	L	A	R	N	L	F	H	A	I	R	L	N	E	R	I	D	N	P	G	E	O
S	E	T	Y	Y	B	C	T	U	S	S	I	O	E	R	H	T	T	U	O	T	U	C	S	E	Q	E	E	H	T	S	T

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
T	S	T	H	E	E	Q	E	S	C	U	T	O	U	T	T	H	R	E	O	I	S	S	U	T	C	B	Y	Y	T	E	S
O	E	G	P	N	D	I	R	E	N	L	R	I	A	H	F	L	N	R	A	L	E	U	M	O	H	C	U	Y	M	O	O

FIGURE 44.

a U. If the U in position 43 is taken, then the symmetrical digraph corresponding to Q U would be L I; if the U in position 56 is taken, the symmetrically-corresponding digraph would be M 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 2–6 and 3–6 at A in figure 43. When the grille is turned to positions 2, 3, and 4 they will disclose two sequent 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 2, grid-cell 54 when the grille takes position 3, and grid-cell 42 when the grille takes position 4.

Positions:	1–3	2–4	1–3	2–4	1–3	2–4	1–3	2–4
	1	8	5	25	11	23	19	22
	64	57	60	40	54	42	46	43
	2	16	6	17	12	31	20	30
	63	49	59	48	53	34	45	35
	3	24	7	9	13	26	21	27
	62	41	58	56	52	39	44	38
	4	32	10	15	14	18	28	29
	61	33	55	50	51	47	37	36

FIGURE 45.

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, 8–5, 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.

f. 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 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.

g. In solving a grille, it will be found advisable to prepare a piece of cross-section paper of proper size for the grille and to cut each aperture as soon as its 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.

32. Solution of example.—*a*. 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 4–2 of the grille shown in figure 45–A. At the same time the other 3 cells numbered 4 in the second

	1	2	3	4	5	6	7	8
1	1	2	3	4	5	6	7	1
2	7	1	2	3	4	5	1	2
3	6	5	1	2	3	1	2	3
4	5	4	3	1	1	2	3	4
5	4	3	2	1	1	3	4	5
6	3	2	1	3	2	1	5	6
7	2	1	5	4	3	2	1	7
8	1	7	6	5	4	3	2	1

FIGURE 45–A.

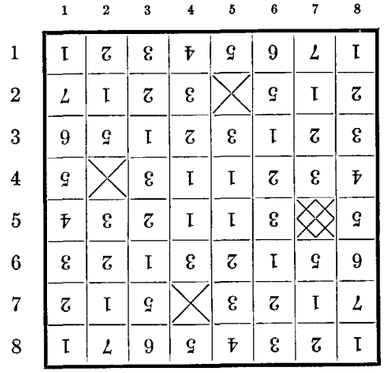


FIGURE 45–B.

band of the grille are marked so that they cannot become apertures. The result is shown in figure 45–B. Conforming to the principle of sequence, the U to be combined with the Q is sought to the right of the Q in figure 43–F. There are three candidates, in positions 43, 46, and 56. They yield:

REF ID:A59439

(Grille in position 3)

39	43		39	46		39	56	
Q	U		Q	U		Q	U	
I	L(=L	I_p)	I	A(=A	I_p)	I	M(=M	I_p)

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 43 would place an aperture at coordinates 6–3 in figure 45–B; the U in position 46 would place an aperture at coordinates 6–6; and the U in position 56 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 51 and E in cell 63. The following possibilities are presented:

39	43	51		39	43	63		39	46	51		39	46	63		39	56	63
Q	U	E		Q	U	E		Q	U	E		Q	U	E		Q	U	E
I	L	R		I	L	O		I	A	R		I	A	O		I	M	O
(=R	L	I)		(=O	L	I)		(=R	A	I)		(=O	A	I)		(=O	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 AI 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 QUEST in this case, as shown in figure 45–C. In position 1 this yields OUR LI . . .; in position 2 it yields two digraphs AN and UT; in position 4 it yields two digraphs H A and R O. 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.

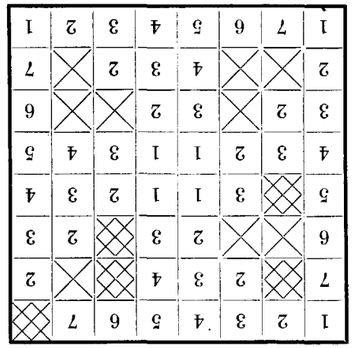


FIGURE 45–C.
(Grille in position 3)

b. Enough has been shown of the procedure to make further demonstration unnecessary. Given the sequence OUR 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.

c. If attempts at solution on the basis of the alpha method of using a grille have failed, the obvious modifications in procedure on the basis of the beta method can readily be made.

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. 31*b*), the several sections each representing the set of letters enciphered on one grid may be superimposed 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 regards 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 canceled, 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 unfilled. 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-filled 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 *vertically*, from left to right, using only as many columns as may be necessary to inscribe the complete message. A 25-letter message would require

REF ID:A59439

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 decryptographing, 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 encipherment 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

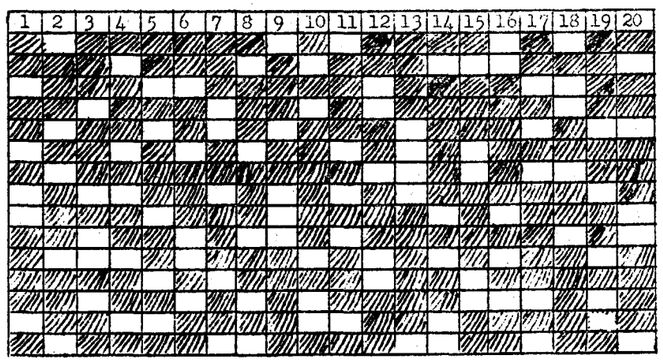


FIGURE 46.

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.[4]

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 31 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.

[4] 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.

324155—41——7

REF ID:A59439

Section VIII

COMBINED SUBSTITUTION-TRANSPOSITION SYSTEMS

	Paragraph
Reasons for combining transposition with substitution	35
Other types of combined substitution-transposition systems	36

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" go, 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 complex 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 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, yet 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

cryptograms can yield no results, because there is nothing of the nature of plain text to guide the cryptanalyst in his juxtaposition of columns.[1]

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

[1] 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. 40.

REF ID:A59439

not be wholly in despair if he is confronted with a problem of this nature in war time, when many cryptograms become available for study. For there are special methods of solution suitable to the occasion, created by special circumstances attendant upon the interception of a voluminous traffic. In subsequent paragraphs the student will come to understand what is here meant by the special circumstances and will learn of these special solutions.

36. Other types of combined substitution-transposition systems.—*a.* There is no technical obstacle to 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 substitution involving four standard alphabets and a keyed columnar transposition with incompletely-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, substitution 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 employed be different from the number of elements in the transposition key, if columnar transposition 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 elimination of the transposition phase by taking advantage of special circumstances or by rearranging the next 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 then these components are subjected to transposition in a second phase. This latter may be followed by a third phase, recombination of distributed components, and a fourth phase, the replacement of the recombined components by letters. Thus, such a system comprises a first substitution, a transposition, a recombination, and a second substitution.[2] In the subsequent paragraphs certain systems of this sort will be dealt with in detail. They are 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 interesting 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.

[2] See Special Text No. 166, *Advanced Military Cryptography*, sec. XI.

SECTION IX

SOLUTION OF THE ADFGVX SYSTEM

	Paragraph
Introductory remarks	37
Special solution by means of identical endings	38
Special solution by means of identical beginnings	39
Special solution by the exact factor method	40
General solution for the ADFGVX system	41
Basic principles of the general solution	42
Illustration of solution	43

37. Introductory remarks.—*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.[1] In this system a 36-character bipartite substitution 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 ADFGVX, and, taken in pairs, the latter are used as substitutes for the letters of the plain text. These substitutive pairs are then inscribed within a rectangle 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 headquarters of divisions and corps. When first instituted on March 1, 1918, the checkerboard consisted of 25 cells, for a 25-letter German alphabet (J was omitted), and the 5 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 26-letter alphabet plus the 10 digits. Transposition keys ranged from 15 to 22 numbers, inclusive, and both the checkerboard and the transposition key were changed daily. The number of messages in this system varied from 25 a day upon the inception of the system to as many as 150 per day, during the last days of May 1918. The first solution was made on April 6 by the French. The cipher continued in use rather extensively until late in June but from that time until the Armistice the volume of messages diminished very considerably. Although only 10 keys, covering a period of as many days were ever solved, the proportion of solved messages in the whole intercepted traffic was about 50 percent. This was true because of the fact that the keys solved were those for days on which the greatest number of messages was intercepted. The same system was employed on the southeastern front from July 1918 to the end of the war. Keys were in effect at first for a period of 2 days and beginning on September 1, for a period of 3 days. In all 17 keys, covering a total of 44 days, were solved.

c. At the time that the Allied cryptanalytic offices were working with cryptograms in this system only three methods were known for their solution and all three of them are classifiable under the heading of *special solutions*, because certain conditions had to obtain before they could be applied. No general solution had been developed until after hostilities had ceased.[2]

[1] Special Text No. 166, *Advanced Military Cryptography*, sec. XI.

[2] 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.

REF ID:A594382

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.

38. 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 the 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 monoalphabetic 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 these 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 monoalphabet 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 intercepted on the same date, the 20th:

No. 1

To CG 22d Brigade:

```
                5                  10                  15                  20                  25                  30
X V A A X  V D D A G  D A D V F  A D A D A  F X G F V  X F A X A

               35                  40                  45                  50                  55                  60
X V A V F  A V X A D  G F F X F  F G A G F  D G D G D  D G A F D

               65                  70                  75                  80                  85                  90
A A D D D  X D A V G  G A A D X  A D F V F  F D F X F  G F G A V

               95                 100                 105                 110                 115                 120
A F A F X  F F X F X  F V D V X  A F F G X  A A A V A  V A F A G

              125                 130                 135                 140                 145                 150
D D F A G  V F A D V  F A V V X  G V A A A  F D F A X  X F A A G

D X
```

REF ID:A59439

No. 2

To CG 23d Brigade:

				5				10				15				20				25				30									
F	D	F	F	F	F	V	F	A	D	D	V	F	V	D	G	A	F	D	F	D	A	G	A	D	F	D	F	A	F		
				35				40				45				50				55				60									
V	A	X	G	D	V	X	G	F	X	V	X	D	X	V	A	A	A	A	D	G	X	F	F	D	V	F	A	A	G		
				65				70				75				80				85				90									
V	G	V	F	F	F	D	A	F	F	F	X	D	A	F	X	G	A	F	D	V	F	V	X	V	D	D	F	A	D		
				95				100				105				110				115				120									
D	A	A	A	X	A	A	F	F	A	F	V	F	X	F	F	A	X	X	A	A	D	V	X	A	V	D	A	V	F		
				125				130				135				140				145				150									
D	F	A	V	X	V	A	D	X	F	A	X	F	F	X	X	A	A	V	X	X	A	D	X	A	A	A	V	V	G		
				155				160				165				170				175				180									
A	G	D	X	X	F	D	F	A	X	F	D	G	D	F	F	X	D	G	X	F	A	G	D	F	F	D	D	V	D		
				185				190																							
D	X	D	A	F	A	G	X	X	A	F	G	A	V																		

d. The delimitation and marking of identities between these two cryptograms is a procedure similar to that explained in paragraph 24*b*, 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 24*b* 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 equivalents), 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 151–54. 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 to 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 D**. Searching through No. 2 for either of these a sequence **X V D D** is found as letters 84–7. Column 1 is probably a long column in No. 1. The word *probably* is used because the identity may extend only over the letters **X V 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

REF ID:A194382

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 V F A D or V F A. Searching through No. 1, a sequence V F A D is found as letters 14–17; a sequence V F A is found as letters 34–6; a sequence V F F D is found as letters 79–82; a sequence V F A D is also found as letters 126–130; a sequence V F A is found as letters 130–2. Here are several possibilities; which is the one to choose? Two of these possibilities coincide exactly with the full sequence being sought, V F A D. 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*n*). "Break tables" are therefore constructed for the messages on the basis of rectangles of 20 columns, and are as shown below.

	0	8	16	24	32	40	48	56	64	72	80	88	96
0	0	8	16	24	32	40	48	56	64	72	80	88	96
7	7	15	23	31	39	47	55	63	71	79	87	95	103
14	14	22	30	38	46	54	62	70	78	86	94	102	110
21	21	29	37	45	53	61	69	77	85	93	101	109	117
28	28	36	44	52	60	68	76	84	92	100	108	116	124
35	35	43	51	59	67	75	83	91	99	107	115	123	131
42	42	50	58	66	74	82	90	98	106	114	122	130	138
49	49	57	65	73	81	89	97	105	113	121	129	137	145
56	56	64	72	80	88	96	104	112	120	128	136	144	152

"Break" table for No. 1 (152 letters)

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140
0	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140
9	9	19	29	39	49	59	69	79	89	99	109	119	129	139	149
18	18	28	38	48	58	68	78	88	98	108	118	128	138	148	158
27	27	37	47	57	67	77	87	97	107	117	127	137	147	157	167
36	36	46	56	66	76	86	96	106	116	126	136	146	156	166	176
45	45	55	65	75	85	95	105	115	125	135	145	155	165	175	185
54	54	64	74	84	94	104	114	124	134	144	154	164	174	184	194

"Break" table for No. 2 (194 letters)

REF ID:A59439

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, 18th, 19th, 20th, . . . 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. 1

X	V	A	A	X	V	D	D	A	G	D	A	D	V	F	A	D	A	D	A	F	X	G	F	V	X	F	A	X	A
				1									2						3									4			
X	V	A	V	F	A	V	X	A	D	G	F	F	X	F	F	G	A	G	F	D	G	D	G	D	D	G	A	F	D
				5									6						7									8			
A	A	D	D	D	X	D	A	V	G	G	A	A	D	X	A	D	F	V	F	F	D	F	X	F	G	F	G	A	V
				9									10						11									12			
A	F	A	F	X	F	F	X	F	X	F	V	D	V	X	A	F	F	G	X	A	A	A	V	A	V	A	F	A	G
				13									14						15									16			
D	D	F	A	G	V	F	A	D	V	F	A	V	V	X	G	V	A	A	A	F	D	F	A	X	X	F	A	A	G
						17							18						19											
D	X																												
20																													

FIGURE 47.

REF ID:A594319

No. 2

F D F F F F V F A D D V F V D G A F D F D A G A D F D F A F
1 2 3

V A X G D V X G F X V X D X V A A A A D G X F F D V F A A G
4 5 6

V G V F F F D A F F F X D A F X G A F D V F V X V D D F A D
7 8 9

D A A A X A A F F A F V F X F F A X X A A D V X A V D A V F
10 11 12

D F A V X V A D X F A X F F X X A A V X X A D X A A A V V G
13 14 15

A G D X X F D F A X F D G D F F X D G X F A G D F F D D V D
16 17 18

D X D A F A G X X A F G A V
19 20

FIGURE 47—Continued

f. A table of equivalencies [3] 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	6	8	10	13	11	17	2	19	15	7	20	14	12	5	18	1	4	3	16

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

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1		9		19		3		8		2		6		11		7		17	
4		10		15		5		13		14		12		20		16		18	

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, 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:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
7	5	17	13	1	14	9	12	19	20	3	16	8	18	2	4	6	10	11	15

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

[3] It is necessary to remark that in setting up the table of equivalencies, 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.

REF ID:A59439

No. 1

7	5	17	13	1	14	9	12	19	20	3	16	8	18	2	4	6	10	11	15
F	X	D	A	X	F	A	F	V	X	A	V	G	V	A	F	A	V	A	F
G	V	F	F	V	X	A	X	A	X	D	A	D	F	G	V	D	G	D	F
A	A	A	A	A	F	D	F	A	F	A	V	D	A	D	X	G	G	F	G
G	V	G	F	A	V	D	G	A	A	D	A	G	V	A	F	F	A	V	X
F	F	V	X	X	D	D	F	F	A	A	F	A	V	D	A	F	A	F	A
D	A	F	F	V	V	X	G	D	G	F	A	F	X	V	X	X	D	F	A
G	V	A	F	D	X	D	A	F	D	X	G	D	G	F	A	F	X	D	A
D	X	D	X	D	A	A	V	A	X	G	D								

No. 2

7	5	17	13	1	14	9	12	19	20	3	16	8	18	2	4	6	10	11	15
A	F	X	V	F	V	A	F	F	F	F	A	A	F	D	F	A	F	A	X
A	X	F	D	D	A	F	A	F	A	D	A	F	F	V	V	D	A	F	A
G	V	D	A	F	D	D	X	D	G	A	A	F	X	F	A	G	D	F	A
V	X	F	V	F	X	V	X	D	X	G	V	F	D	V	X	X	D	A	V
G	D	A	F	F	F	F	A	V	X	A	V	X	G	D	G	F	A	F	X
V	X	X	D	F	A	V	A	D	A	D	G	D	X	G	D	F	A	V	X
F	V	F	F	V	X	X	D	D	F	F	A	A	F	A	V	D	A	F	A
F	A	D	A	F	F	V	V	X	G	D	G	F	A	F	X	V	X	X	D
F	A	G	V	A	F	D	X	D	A	F	D	X	G	D	G	F	A	F	X
D	A	D	X	D	X	D	A	A	V	A	X	G	D						

FIGURE 48.

h. A frequency distribution is now made of all the bipartite pairs, so as to solve the enciphering checkerboard. There is no necessity for going through this part of the solution, for it falls along quite normal lines of monoalphabetic substitution. The two plain-text rectangles are shown in figure 49. The checkerboard [4] is found to be as shown in figure 50*a*.

No. 1

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

FIGURE 49.

[4] Since the second cryptogram is addressed to the CG 23d Brigade and the first cryptogram mentions that the commander of that brigade has been ordered to do so and so, the solution of the groups GG (=2) and FG (=3) is made by inference. This gives the placement of these two digits in the cipher square.

REF ID:A594319

No. 2

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

FIGURE 49—Continued.

2nd component

1st component	A	D	F	G	V	X
A	G		E		R	M
D	A		N	I		L
F	T	Y	C	3	P	H
G		S	B	2	D	F
V				K		O
X		U	V	W	X	

FIGURE 50a.

	A	D	F	G	V	X
A	G		E		R	M
D	A		N	I		L
F	T	Y	C	3	P	H
G		S	B	2	D	
V	F				K	O
X	(Q)	U	V	W	X	(Z)

FIGURE 50b.

	A	D	F	G	V	X
A	G	6	E	5	R	M
D	A	1	N	I	8	L
F	T	Y	C	3	P	H
G	7	S	B	2	D	4
V	F	6	J	Ø	K	O
X	Q	U	V	W	X	Z

FIGURE 50c.

i. Speculating upon the disposition of the letters within the enciphering checkerboard, it becomes evident that the key phrase upon which it is based is GERMAN 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

REF ID:A59439

(fig. 50*b*). The complete checkerboard may therefore be taken almost certainly to be as shown in figure 50*c*. 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:

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

39. Special solution by means of identical beginnings.—*a*. 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 ADFGVX 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 *sine qua non* for the solution.

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

No. 1

V D D F A X F A A X D X G G F F V F X F G X D X G D G A G F
A G D A D V G G D A A A D X X D X A F F A A D A F D F F D A

No. 2

G X D D A D D G D F V G X A X X X G X G A A A A D F A D D X
A V D X F X A D

No. 3

X D A A A G X D D X V F F V D G A D F D X A A A G D F A D G
A F D A D G V G D V F D F X A G F X A F A F A X D D D D F D
X A X V A D X F X F D G A G F G G A D D A G D G X A V G D G
A D A F A X F A A G V A A G A F D V D V D X F D A X F D F F
G D X D V D A D A V D A D D D G A D A G A A A F G G D X A X
F G V X D D G D D F A F A G V A F G X G V D D A X X D V F F
F F D X G V G D F G A V A D A X D A F A A F D G F V F X X X
A A G A G A F D G X A F A F X X G G A G A A F F A A F D G A
G A F V X D G G F G D A A A F D A D A D X V V A X F V A D D
G A F F F G X A X D F D D F X A A A A A

No. 4

A F G F X A G X A G X D D A F A A X A V G D D D D F A F G V
D G D X A F D X A X G F G D D V A D X A X G F A X F D A D D
G D

No. 5

X A A A D D G A A G D D D X F F A V G A X D G G D F F A V A
D A A X A G D X D X X X X D G V F A D A D F F F F V V G F D
X F D G G D A X D G A D F D

REF ID:A594339

No. 6

X D A A V D X D G F X V G D D A V G X A D X A A D X G G A A
G D F D A A A G A X D V F D F D F F D D F D D F X F X X F D
F D X A X G A X F F V D V A F G V D V D D D A G D G G D A A
G G F D D D V F F V V A G V A X A A G G X G X D D D A D X F
A D F F G D G F D A A F G A X F F D V D D D A G A F A D A V
D D D A V G A V A D F G D D F F D G D V D G G X A X A X D A
D X D V F F X V A X G F D A G X F F F F A A X D A F V D X G
X F D A G A G A V D V A G A F D G D A V V D D D D D F X G V
A F F A A F F F D V D F F A F D A G D G G A A A F D X A X A
V A X D A G A D X D V F A F F F G D D A D D D F A G D F A X
D G

No. 7

A G F G V D D D D F D D F X F D D G D F A X V D D V D V X A
D D A X X A A D D F A G G F F A X D D G X D F A D D F D G D
D V A X A X F X D A F X D D G F X G D V G F F G X D A D F A
D D A F F V D G X A A D X F X G V A D A X G X A G A G D G V
X D D V

No. 8

D F G F X D F A F F X D X A G A D G G G D D F G A X G V D F
V V F D A A A X G D A V D V A D D G V D A F A G

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 (290 letters) has 5 long columns of 20 letters and 10 short columns of 19 letters [(15×20)−10=290]. No. 6 (302 letters) has 2 long columns of 21 letters and 13 short columns of 20 letters [(15×21)−13=302]. The identical sections in No. 3 and No. 6 having been marked off as shown in figure 51, the next step is to transcribe the texts into their correct column lengths as given by the study of identical sections, writing them merely in their serial order, 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 sequences, the number 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, viz, in columns 5 and 12.

REF ID:A59439

No. 3

X D A A A G X D D X V F F V D G A D F D X A A A G D F A D G
1 2
A F D A D G V G D V F D F X A G F X A F A F A X D D D D F D
3 4
X A X V A D X F X F D G A G F G G A D D A G D G X A V G D G
5
A D A F A X F A A G V A A G A F D V D V D X F D A X F D F F
6 7
G D X D V D A D A V D A D D D G A D A G A A A F G G D X A X
8
F G V X D D G D D F A F A G V A F G X G V D D A X X D V F F
9 10
F F D X G V G D F G A V A D A X D A F A A F D G F V F X X X
11
A A G A G A F D G X A F A F X X G G A G A A F F A A F D G A
12 13
G A F V X D G G F G D A A A F D A D A D X V V A X F V A D D
14
G A F F F G X A X D F D D F X A A A A A
15

No. 6

X D A A V D X D G F X V G D D A V G X A D X A A D X G G A A
1 2
G D F D A A A G A X D V F D F D F F D D F D D F X F X X F D
3
F D X A X G A X F F V D V A F G V D V D D D A G D G G D A A
4 5
G G F D D D V F F V V A G V A X A A G G X G X D D D A D X F
6
A D F F G D G F D A A F G A X F F D V D D D A G A F A D A V
7 8
D D D A V G A V A D F G D D F F D G D V D G G X A X A X D A
9
D X D V F F X V A X G F D A G X F F F F A A X D A F V D X G
10 11
X F D A G A G A V D V A G A F D G D A V V D D D D D F X G V
12
A F F A A F F F D V D F F A F D A G D G G A A A F D X A X A
13 14
V A X D A G A D X D V F A F F F G D D A D D D F A G D F A X
15
D G

FIGURE 51.

REF ID:A6194382

No. 3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	D	D	F	D	A	D	D	G	X	A	A	F	A	A
D	X	V	D	D	G	F	A	D	D	X	G	F	A	F
A	A	F	X	A	V	F	G	D	V	D	A	A	A	F
A	A	D	A	G	A	G	A	F	F	A	F	A	F	F
A	A	F	X	D	A	D	A	A	F	F	D	F	D	G
G	G	X	V	G	G	X	A	F	F	A	G	D	A	X
X	D	A	A	X	A	D	F	A	F	A	X	G	D	A
D	F	G	D	A	F	V	G	G	D	F	A	A	A	X
D	A	F	X	V	D	D	G	V	X	D	F	G	D	D
X	D	X	F	G	V	A	D	A	G	G	A	A	X	F
V	G	A	X	D	D	D	X	F	V	F	F	F	V	D
F	A	F	F	G	V	A	A	G	G	V	X	V	V	D
F	F	A	D	A	D	V	X	X	D	F	X	X	A	F
V	D	F	G	D	X	D	F	G	F	X	G	D	X	X
D	A	A	A	A	F	A	G	V	G	X	G	G	F	A
G	D	X	G	F	D	D	V	D	A	X	A	G	V	A
A	G	D	F	A	A	D	X	D	V	A	G	F	A	A
D	V	D	G	X	X	D	D	A	A	A	A	G	D	A
F	G	D	G	F	F	G	D	X	D	G	A	D	D	A
		D	A	A		A							G	

FIGURE 52.

REF ID:A59439

No. 6

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	D	D	F	D	A	D	D	G	X	A	A	F	A	A
D	X	V	D	D	G	F	A	D	D	X	G	F	A	F
A	A	F	X	A	V	F	G	D	V	D	A	A	A	F
A	A	D	A	G	A	G	A	F	F	A	F	A	F	F
V	D	F	X	D	X	D	F	F	F	F	D	F	D	G
D	X	D	G	G	A	G	A	D	X	V	G	F	X	D
X	G	F	A	G	A	F	D	G	V	D	D	F	A	D
D	G	F	X	D	G	D	A	D	A	X	A	D	X	A
G	A	D	F	A	G	A	V	V	X	G	V	V	A	D
F	A	D	F	A	X	A	D	D	G	X	V	D	V	D
X	G	F	V	G	G	F	D	G	F	F	D	F	A	D
V	D	D	D	G	X	G	D	G	D	D	D	F	X	F
G	F	D	V	F	D	A	A	X	A	A	D	A	D	A
D	D	F	A	D	D	X	V	A	G	G	D	F	A	G
D	A	X	F	D	D	F	G	X	X	A	D	D	G	D
A	A	F	G	D	A	F	A	A	F	G	F	A	A	F
V	A	X	V	V	D	D	V	X	F	A	X	G	D	A
G	G	X	D	F	X	V	A	D	F	V	G	D	X	X
X	A	F	V	F	F	D	D	A	F	D	V	G	D	D
A	X	D	D	V	A	D	F	D	A	V	A	G	V	G
				V									F	

FIGURE 52—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.

324155—41——8

REF ID:A594339

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:

Cryptogram No.	Total number of letters	Lengths of columns	Number of columns	
			Long	Short
1	60	4	All same length	
2	38	3 and 2	8	7
4	62	5 and 4	2	13
5	74	5 and 4	14	1
7	124	9 and 8	4	11
8	54	4 and 3	9	6

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 with nonfractionated letters.

f. But now consider a plain-text rectangle in the ADFGVX 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, 3d, 5th, . . . pairs down the columns can be distributed in a frequency square. The results are shown in Fig. 53.

REF ID:A59439

	No. 3			No. 6	
	Col. 5	Col. 14		Col. 5	Col. 14
1	D	A	1	D	A
	D	A		D	A
3	A	A	3	A	A
	G	F		G	F
5	D	D	5	D	D
	G	A		G	X
7	X	D	7	G	A
	A	A		D	X
9	V	D	9	A	A
	G	X		A	V
11	D	V	11	G	A
	G	V		G	X
13	A	A	13	F	D
	D	X		D	A
15	A	F	15	D	G
	F	V		D	A
17	A	A	17	V	D
	X	D		F	X
19	F	D	19	F	D
	A	G		V	V
			21	V	F

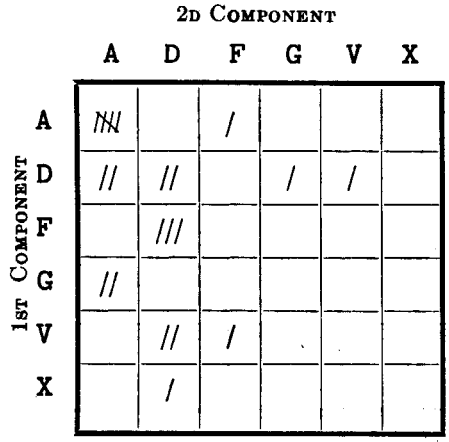


FIGURE 53.

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

Expected value of ϕ for plain text $=.0667\times21\times20=28.01$
Expected value of ϕ for random text $=.0385\times21\times20=16.17$
Observed value of ϕ in this case $=(5\times4)+(2\times1)+(2\times1)+(3\times2)+(2\times1)+(2\times1)=34$

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.

REF ID:A594339

No. 3

	(1)			(2)			(3)			(4)			(5)			(6)		
	5	14	3	5	14	4	5	14	7	14	5	3	14	5	4	14	5	7
1	D	A	D	D	A	F	D	A	D	A	D	D	A	D	F	A	D	D
2	D	A	V	D	A	D	D	A	F	A	D	V	A	D	D	A	D	F
3	A	A	F	A	A	X	A	A	F	A	A	F	A	A	X	A	A	F
4	G	F	D	G	F	A	G	F	G	F	G	D	F	G	A	F	G	G
5	D	D	F	D	D	X	D	D	D	D	D	F	D	D	X	D	D	D
6	G	A	X	G	A	V	G	A	X	A	G	X	A	G	V	A	G	X
7	X	D	A	X	D	A	X	D	D	D	X	A	D	X	A	D	X	D
8	A	A	G	A	A	D	A	A	V	A	A	G	A	A	D	A	A	V
9	V	D	F	V	D	X	V	D	D	D	V	F	D	V	X	D	V	D
10	G	X	X	G	X	F	G	X	A	X	G	X	X	G	F	X	G	A
11	D	V	A	D	V	X	D	V	D	V	D	A	V	D	X	V	D	D
12	G	V	F	G	V	F	G	V	A	V	G	F	V	G	F	V	G	A
13	A	A	A	A	A	D	A	A	V	A	A	A	A	A	D	A	A	V
14	D	X	F	D	X	G	D	X	D	X	D	F	X	D	G	X	D	D
15	A	F	A	A	F	A	A	F	A	F	A	A	F	A	A	F	A	A
16	F	V	X	F	V	G	F	V	D	V	F	X	V	F	G	V	F	D
17	A	A	D	A	A	F	A	A	D	A	A	D	A	A	F	A	A	D
18	X	D	D	X	D	G	X	D	D	D	X	D	D	X	G	D	X	D
19	F	D	D	F	D	G	F	D	G	D	F	D	D	F	G	D	F	G
20	A	G	D	A	G	A	A	G	A	G	A	D	G	A	A	G	A	A

No. 6

	(1)			(2)			(3)			(4)			(5)			(6)		
	5	14	3	5	14	4	5	14	7	14	5	3	14	5	4	14	5	7
1	D	A	D	D	A	F	D	A	D	A	D	D	A	D	F	A	D	D
2	D	A	V	D	A	D	D	A	F	A	D	V	A	D	D	A	D	F
3	A	A	F	A	A	X	A	A	F	A	A	F	A	A	X	A	A	F
4	G	F	D	G	F	A	G	F	G	F	G	D	F	G	A	F	G	G
5	D	D	F	D	D	X	D	D	D	D	D	F	D	D	X	D	D	D
6	G	X	D	G	X	G	G	X	G	X	G	D	X	G	G	X	G	G
7	G	A	F	G	A	A	G	A	F	A	G	F	A	G	A	A	G	F
8	D	X	F	D	X	X	D	X	D	X	D	F	X	D	X	X	D	D
9	A	A	D	A	A	F	A	A	A	A	A	D	A	A	F	A	A	A
10	A	V	D	A	V	F	A	V	A	V	A	D	V	A	F	V	A	A
11	G	A	F	G	A	V	G	A	F	A	G	F	A	G	V	A	G	F
12	G	X	D	G	X	D	G	X	G	X	G	D	X	G	D	X	G	G
13	F	D	D	F	D	V	F	D	A	D	F	D	D	F	V	D	F	A
14	D	A	F	D	A	A	D	A	X	A	D	F	A	D	A	A	D	X
15	D	G	X	D	G	F	D	G	F	G	D	X	G	D	F	G	D	F
16	D	A	F	D	A	G	D	A	F	A	D	F	A	D	G	A	D	F
17	V	D	X	V	D	V	V	D	D	D	V	X	D	V	V	D	V	D
18	F	X	X	F	X	D	F	X	V	X	F	X	X	F	D	X	F	V
19	F	D	F	F	D	V	F	D	D	D	F	F	D	F	V	D	F	D
20	V	V	D	V	V	D	V	V	D	V	V	D	V	V	D	V	V	D
21	V	F		V	F		V	F		F	V		F	V		F	V	

FIGURE 54.

REF ID:A59439

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.

(1) [5–14–3]

	A	D	F	G	V	X
A	𝍸		𝍷	𝍵	𝍶	𝍵
D	𝍶	𝍷		𝍵	𝍵	
F		𝍸				
G	𝍶	𝍵				
V		𝍹	𝍶			𝍵
X		𝍷	𝍶			𝍶

(2) [5–14–4]

	A	D	F	G	V	X
A	𝍸	𝍷	𝍵	𝍵	𝍵	
D	𝍶	𝍶		𝍶	𝍵	
F	𝍶	𝍷				
G	𝍷					
V		𝍷	𝍷	𝍵		
X		𝍷	𝍵	𝍶		𝍵

(3) [5–14–7]

	A	D	F	G	V	X
A	𝍸		𝍹		𝍵	𝍶
D	𝍶	𝍷		𝍵	𝍵	
F		𝍷		𝍶		
G	𝍷					
V	𝍶	𝍹	𝍵			
X	𝍵	𝍷		𝍶	𝍵	

(4) [14–5–3]

	A	D	F	G	V	X
A	𝍸	𝍹		𝍷		
D		𝍶	𝍸		𝍹	𝍵
F	𝍵				𝍵	𝍶
G		𝍸	𝍵			𝍶
V		𝍶				
X		𝍵				

(5) [14–5–4]

	A	D	F	G	V	X
A	𝍸	𝍷	𝍵	𝍶		
D	𝍵	𝍹	𝍷	𝍶	𝍶	𝍶
F	𝍵	𝍵		𝍵	𝍵	
G	𝍶	𝍶	𝍶	𝍵	𝍵	
V		𝍶				
X				𝍵		

(6) [14–5–7]

	A	D	F	G	V	X
A	𝍸	𝍶		𝍶	𝍵	
D		𝍹	𝍸		𝍶	𝍶
F	𝍵	𝍵			𝍶	
G	𝍶	𝍵		𝍹		𝍵
V		𝍶				
X		𝍵				

FIGURE 55.

j. These distributions are now tested for monoalphabeticity, by applying the ϕ test. The number of occurrences in each distribution is 41. Then $41\times40\times.0667=109.4$ is the expected value of ϕ for plain text; $41\times40\times.0385=63.1$ is the expected value of ϕ for random text. Here are the calculations for the first distribution (combination 5–14–3) yielding the observed value of ϕ as 82:

$$(5\times4)+(3\times2)+(1\times0)+(2\times1)+(1\times0)+$$
$$(2\times1)+(3\times2)+(1\times0)+(1\times0)+(5\times4)+(2\times1)+$$
$$(1\times0)+(4\times3)+(2\times1)+(1\times0)+(3\times2)+(2\times2)+(2\times1)=82.$$

The observed values of ϕ for all 6 frequency distributions are shown herewith:

(1)=82	(4)=120
(2)=76	(5)= 70
(3)=78	(6)=110

REF ID:A594389

Only two of these distributions give close approximations to 109, the expected value of ϕ, 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, 5 and 7 form the group of 5 columns at the left of the transposition rectangle, the following combinations are possible:

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

	(1)					(2)					(3)					(4)				
	14	5	3	4	7	14	5	3	7	4	14	5	7	3	4	14	5	7	4	3
1	A	D	D	F	D	A	D	D	D	F	A	D	D	D	F	A	D	D	F	D
2	A	D	V	D	F	A	D	V	F	D	A	D	F	V	D	A	D	F	D	V
3	A	A	F	X	F	A	A	F	F	X	A	A	F	F	X	A	A	F	X	F
4	F	G	D	A	G	F	G	D	G	A	F	G	G	D	A	F	G	G	A	D
5	D	D	F	X	D	D	D	F	D	X	D	D	D	F	X	D	D	D	X	F
6	A	G	X	V	X	A	G	X	X	V	A	G	X	X	V	A	G	X	V	X
7	D	X	A	A	D	D	X	A	D	A	D	X	D	A	A	D	X	D	A	A
8	A	A	G	D	V	A	A	G	V	D	A	A	V	G	D	A	A	V	D	G
9	D	V	F	X	D	D	V	F	D	X	D	V	D	F	X	D	V	D	X	F
10	X	G	X	F	A	X	G	X	A	F	X	G	A	X	F	X	G	A	F	X
11	V	D	A	X	D	V	D	A	D	X	V	D	D	A	X	V	D	D	X	A
12	V	G	F	F	A	V	G	F	A	F	V	G	A	F	F	V	G	A	F	F
13	A	A	A	D	V	A	A	A	V	D	A	A	V	A	D	A	A	V	D	A
14	X	D	F	G	D	X	D	F	D	G	X	D	D	F	G	X	D	D	G	F
15	F	A	A	A	A	F	A	A	A	A	F	A	A	A	A	F	A	A	A	A
16	V	F	X	G	D	V	F	X	D	G	V	F	D	X	G	V	F	D	G	X
17	A	A	D	F	D	A	A	D	D	F	A	A	D	D	F	A	A	D	F	D
18	D	X	D	G	D	D	X	D	D	G	D	X	D	D	G	D	X	D	G	D
19	D	F	D	G	G	D	F	D	G	G	D	F	G	D	G	D	F	G	G	D
20	G	A	D	A	A	G	A	D	A	A	G	A	A	D	A	G	A	A	A	D

FIGURE 56.

REF ID:A59439

No. 6

	(1)					(2)					(3)					(4)				
	14	5	3	4	7	14	5	3	7	4	14	5	7	3	4	14	5	7	4	3
1	A	D	D	F	D	A	D	D	D	F	A	D	D	D	F	A	D	D	F	D
2	A	D	V	D	F	A	D	V	F	D	A	D	F	V	D	A	D	F	D	V
3	A	A	F	X	F	A	A	F	F	X	A	A	F	F	X	A	A	F	X	F
4	F	G	D	A	G	F	G	D	G	A	F	G	G	D	A	F	G	G	A	D
5	D	D	F	X	D	D	D	F	D	X	D	D	D	F	X	D	D	D	X	F
6	X	G	D	G	G	X	G	D	G	G	X	G	G	D	G	X	G	G	G	D
7	A	G	F	A	F	A	G	F	F	A	A	G	F	F	A	A	G	F	A	F
8	X	D	F	X	D	X	D	F	D	X	X	D	D	F	X	X	D	D	X	F
9	A	A	D	F	A	A	A	D	A	F	A	A	A	D	F	A	A	A	F	D
10	V	A	D	F	A	V	A	D	A	F	V	A	A	D	F	V	A	A	F	D
11	A	G	F	V	F	A	G	F	F	V	A	G	F	F	V	A	G	F	V	F
12	X	G	D	D	G	X	G	D	G	D	X	G	G	D	D	X	G	G	D	D
13	D	F	D	V	A	D	F	D	A	V	D	F	A	D	V	D	F	A	V	D
14	A	D	F	A	X	A	D	F	X	A	A	D	X	F	A	A	D	X	A	F
15	G	D	X	F	F	G	D	X	F	F	G	D	F	X	F	G	D	F	F	X
16	A	D	F	G	F	A	D	F	F	G	A	D	F	F	G	A	D	F	G	F
17	D	V	X	V	D	D	V	X	D	V	D	V	D	X	V	D	V	D	V	X
18	X	F	X	D	V	X	F	X	V	D	X	F	V	X	D	X	F	V	D	X
19	D	F	F	V	D	D	F	F	D	V	D	F	D	F	V	D	F	D	V	F
20	V	V	D	D	D	V	V	D	D	D	V	V	D	D	D	V	V	D	D	D
21	F	V				F	V				F	V				F	V			

FIGURE 56—Continued.

m. 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 57. The other squares in figure 57 are constructed in the same way, for the other combinations of figure 56.

(1) [14–5–3–4–7]

	A	D	F	G	V	X
A	𝍸	𝍸		𝍸		//
D		///	𝍸	//	𝍸	/
F	𝍸				///	𝍸
G		𝍸	//	/		//
V		//				/
X		//	/		/	

(2) [14–5–3–7–4]

	A	D	F	G	V	X
A	𝍸	𝍸	///	///	/	
D	//	𝍸	𝍸	////	////	//
F	/	𝍸	////	/	/	//
G	//	𝍸	/	/		//
V		////				
X	/	//	/		/	

FIGURE 57.

REF ID:A594339

(3) [14–5–7–3–4]

	A	D	F	G	V	X
A	𝍸 ///	////		//	/	
D	𝍸	𝍸 ////	𝍸 //// /	//	//	///
F	//	/	𝍸	//	//	//
G	//	///		////		/
V	/	////				
X		//	/	/	/	

(4) [14–5–7–4–3]

	A	D	F	G	V	X
A	𝍸 ///	𝍸	//	//	//	
D	/	𝍸 /	𝍸 ////	/	𝍸 /	𝍸 //
F	//	//	//		///	///
G	//	///	//	𝍸		//
V		///				/
X		/	/			

FIGURE 57—Continued.

n. Again applying the ϕ-test, the expected value of ϕ is $81\times80\times.0667=432$. The observed values for the four combinations of figure 57 are as follows:

(1) For combination 14–5–3–4–7, $\phi=436$
(2) For combination 14–5–3–7–4, $\phi=276$
(3) For combination 14–5–7–3–4, $\phi=344$
(4) For combination 14–5–7–4–3, $\phi=318$

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.

o. Examining 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 [(15×9)—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

14	5	3	4	7	1	2	6	8	9	10	11	12	13	15
A	X	D	V	D	A	D	F	D	X	F	G	D	A	G
D	A	G	D	F	G	F	F	D	D	X	X	A	A	D
A	A	D	V	A	F	D	A	V	A	G	D	F	D	G
X	D	F	X	D	G	D	X	A	F	D	A	F	X	V
G	D	A	A	D	V	F	D	X	X	V	D	V	F	X
X	F	X	D	F	D	X	D	A	D	G	F	D	X	D
A	A	V	D	D	D	F	G	X	D	F	A	G	G	D
G	G	D	A	G	D	D	X	F	G	F	D	X	V	V
A	G	D	X											

FIGURE 58.

REF ID:A59439

10 columns which are to be anagrammed to ascertain their correct sequence. The column to follow column 7 is ascertained 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.

No. 3

14	5	3	4	7	10	15	12	13	1	2	8	6	9	11
A	D	D	F	D	X	A	A	F	X	D	D	A	G	A
A	D	V	D	F	D	F	G	F	D	X	A	G	D	X
A	A	F	X	F	V	F	A	A	A	A	G	V	D	D
F	G	D	A	G	F	F	F	A	A	A	A	A	F	A
D	D	F	X	D	F	G	D	F	A	A	A	A	A	F
A	G	X	V	X	F	X	G	D	G	G	A	G	F	A
D	X	A	A	D	F	A	X	G	X	D	F	A	A	A
A	A	G	D	V	D	X	A	A	D	F	G	F	G	F
D	V	F	X	D	X	D	F	G	D	A	G	D	V	D
X	G	X	F	A	G	F	A	A	X	D	D	V	A	G
V	D	A	X	D	V	D	F	F	V	G	X	D	F	F
V	G	F	F	A	G	D	X	V	F	A	A	V	G	V
A	A	A	D	V	D	F	X	X	F	F	X	D	X	F
X	D	F	G	D	F	X	G	D	V	D	F	X	G	X
F	A	A	A	A	G	A	G	G	D	A	G	F	V	X
V	F	X	G	D	A	A	A	G	G	D	V	D	D	X
A	A	D	F	D	V	A	G	F	A	G	X	A	D	A
D	X	D	G	D	A	A	A	G	D	V	D	X	A	A
D	F	D	G	G	D	A	A	D	F	G	D	F	X	G
G	A	D	A	A										

FIGURE 59.

REF ID:A594319

No. 6

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
A D	D F	D X	A A	F X	D D	A G	A
A D	V D	F D	F G	F D	X A	G D	X
A A	F X	F V	F A	A A	A G	V D	D
F G	D A	G F	F F	A A	A A	A F	A
D D	F X	D F	G D	F V	D F	X F	F
X G	D G	G X	D G	F D	X A	A D	V
A G	F A	F V	D D	F X	G D	A G	D
X D	F X	D A	A A	D D	G A	G D	X
A A	D F	A X	D V	V G	A V	G V	G
V A	D F	A G	D V	D F	A D	X D	X
A G	F V	F F	D D	F X	G D	G G	F
X G	D D	G D	F D	F V	D D	X G	D
D F	D V	A A	A D	A G	F A	D X	A
A D	F A	X G	G D	F D	D V	D A	G
G D	X F	F X	D D	D D	A G	D X	A
A D	F G	F F	F F	A A	A A	A A	G
D V	X V	D F	A X	G V	A V	D X	A
X F	X D	V F	X G	D G	G A	X D	V
D F	F V	D F	D V	G X	A D	F A	D
V V	D D	D A	G A	G A	X F	A D	V
F V							

No. 7

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
A X	D V	D F	G D	A A	D D	F X	G
D A	G D	F X	D A	A G	F D	F D	X
A A	D V	A G	G F	D F	D V	A A	D
X D	F X	D D	V F	X G	D A	X F	A
G D	A A	D V	X V	F V	F X	D X	D
X F	X D	F G	D D	X D	X A	D D	F
A A	V D	D F	D G	G D	F X	G D	A
G G	D A	G F	V X	V D	D F	X G	D
A G	D X						

FIGURE 59—Continued.

p. 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. 38*i*). The results are shown in figure 61. The completely solved messages are shown in figure 62.

REF ID:A59439

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15										
Literal key.....	T	H	E	F	L	O	W	R	S	A	B	M	I	N	P	G	C	D	J	K	Q	U	V	X	Y	Z
Numerical key.....	14	5	3	4	7	10	15	12	13	1	2	8	6	9	11										

2D COMPONENT

1st COMPONENT	A	D	F	G	V	X
A	T	H		E	5	F
D		L	O	W	R	S
F	A		B		M	I
G		N	P	G	7	C
V		D		J		K
X		U	V		Y	

FIGURE 60.

2D COMPONENT

1st COMPONENT	A	D	F	G	V	X
A	T	H	8	E	5	F
D	6	L	O	W	R	S
F	A	1	B	2	M	I
G	9	N	P	G	7	C
V	3	D	4	J	Ø	K
X	Q	U	V	X	Y	Z

FIGURE 61.

No. 1

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
R	E	G	I	M	E	N	T
D V	A G	G G	F X	F V	A G	G D	A

I	N	P	O	S	I	T
A F	X G	D G	F D	F D	X F	X A	A

I	O	N	S	H	A	L	L
F X	D F	G D	D X	A D	F A	D D	D

I	A	T	T	A	C	K
D F	X F	A A	A A	A F	A G	X V	X

No. 2

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
R	E	Q	U	E	S	T	I
D V	A G	X A	X D	A G	D X	A A	F

N	S	T	R	U	C	T
X G	D D	X A	A D	V X	D G	X A	A

I	O	N	S
F X	D F	G D	D X

FIGURE 62.

REF ID:A594319

No. 3

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

FIGURE 62—Continued.

REF ID:A59439

No. 4

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
T	H	I	R	T	Y	S	I
A A	A D	F X	D V	A A	X V	D X	F
X	T	H	F	A	L	E	
X X	G A	A A	D A	X F	A D	D A	G
A	V	I	N	G	G	O	L
F A	X F	F X	G D	G G	G G	D F	D
D	E	N	V	I	L	L	
D V	D A	G G	D X	F F	X D	D D	D
E							
A G							

No. 5

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
C	O	R	P	S	W	I	L
G X	D F	D V	G F	D X	D G	F X	D
L	T	A	K	E	O	V	
D D	D A	A F	A V	X A	G D	F X	F
E	R	T	R	A	F	F	I
A G	D V	A A	D V	F A	A X	A X	F
C	C	O	N	T	R	O	
X G	X G	X D	F G	D A	A D	V D	F
L	A	T	O	N	C	E	
D D	F A	A A	D F	G D	G X	A G	

FIGURE 62—Continued.

REF ID:A594439

No. 6

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
H	O	S	T	I	L	E	T
A D	D F	D X	A A	F X	D D	A G	A
R	O	O	P	S	E	S	
A D	V D	F D	F G	F D	X A	G D	X
T	I	M	A	T	E	D	O
A A	F X	F V	F A	A A	A G	V D	D
N	E	B	A	T	T	A	
F G	D A	G F	F F	A A	A A	A F	A
L	I	O	N	M	O	V	I
D D	F X	D F	G D	F V	D F	X F	F
N	G	U	P	S	T	R	
X G	D G	G X	D G	F D	X A	A D	V
E	A	M	L	I	N	E	S
A G	F A	F V	D D	F X	G D	A G	D
O	U	T	H	W	E	S	
X D	F X	D A	A A	D D	G A	G D	X
T	O	F	R	J	5	7	7
A A	D F	A X	D V	V G	A V	G V	G
H	A	N	D	A	S	S	
V A	D F	A G	D V	D F	A D	X D	X
E	M	B	L	I	N	G	I
A G	F V	F F	D D	F X	G D	G G	F
N	W	O	O	D	S	N	
X G	D D	G D	F D	F V	D D	X G	D
O	R	T	H	E	A	S	T
D F	D V	A A	A D	A G	F A	D X	A
O	F	G	O	L	D	E	
A D	F A	X G	G D	F D	D V	D A	G
N	V	I	L	L	E	S	T
G D	X F	F X	D D	D D	A G	D X	A
O	P	B	A	T	T	E	
A D	F G	F F	F F	A A	A A	A A	G
R	Y	O	F	7	5	S	F
D V	X V	D F	A X	G V	A V	D X	A
I	R	I	N	G	F	R	
X F	X D	V F	X G	D G	G A	X D	V
O	M	O	R	C	H	A	R
D F	F V	D F	D V	G X	A D	F A	D
D	L	E	E	F	A	R	
V V	D D	D A	G A	G A	X F	A D	V
M							
F V							

FIGURE 62—Continued.

REF ID:A59439

No. 7

14 5	3 4	7 10	15 12	13 1	2 8	6 9	11
F	R	O	N	T	L	I	N
A X	D V	D F	G D	A A	D D	F X	G
E	O	U	T	P	O	S	
D A	G D	F X	D A	A G	F D	F D	X
T	R	E	P	O	R	T	S
A A	D V	A G	G F	D F	D V	A A	D
O	U	R	I	N	F	A	
X D	F X	D D	V F	X G	D A	X F	A
N	T	R	Y	M	I	S	S
G D	A A	D V	X V	F V	F X	D X	D
I	O	N	S	S	H	O	
X F	X D	F G	D D	X D	X A	D D	F
T	D	O	W	N	I	N	E
A A	V D	D F	D G	G D	F X	G D	A
N	E	M	Y	L	I	N	
G G	D A	G F	V X	V D	D F	X G	D
E	S						
A G	D X						

No. 8

4 5	3 4	7 10	15 12	13 1	2 8	6 9	11
W	I	R	E	L	I	N	E
D G	F X	D V	A G	D D	F X	G D	A
T	O	B	R	I	G	A	
G A	A D	F F	F D	V F	X G	G F	A
D	I	N	T	E	R	R	U
V D	F X	G D	A A	A G	D V	D V	X
P	T	E	D				
D G	F A	A A	G V	D			

FIGURE 62—Continued.

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 39*e*, is applied in order to produce the best monoalphabetic distribution of bipartite

REF ID:A194382

digraphs down the juxtaposed columns. There is nothing to prevent the simultaneous use of all cryptograms that have been enciphered by completely filled rectangles, for it is clear that if, for example, columns 15 and 4 are to be paired in one cryptogram, the same columns will be paired in all the 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 into play, beyond those already possessed by the student.

b. 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 columns. 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 76 (19×4), 152 (19×8), 190 (19×10) 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.

c. 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 shiftless or careless clerk must go over his work to correct errors is reduced to a minimum. Hence, it often happens that in such cases an enciphering clerk adds one, two, or three letters to complete the last row, thus leading to the transmission of not a few cryptograms enciphered by completely-filled rectangles. Space forbids giving an example of such a solution.

41. General solution for the ADFGVX system.—*a.* All three of the foregoing methods of solving cryptograms in the ADFGVX system fall in the category of special solutions and therefore are dependent upon the fortuitous existence of the special conditions required under each case. What is really desired in the practical situation is a method of solution which is not so dependent upon chance or good fortune for success. A search for a general solution was, of course, made during the time that the system was under minute study by the cryptanalytic agencies of the Allies, but no general solution was devised. All the solutions made during actual hostilities and for a number of weeks thereafter were of the special types described in the preceding paragraphs. The first published description of a general solution is to be found in Givierge's *Cours de Cryptographie*, 1925, but only in broad outlines. A complete general solution was independently conceived by a group of cryptanalysts in the office of the Chief Signal Officer [5] and will be described in paragraphs 42 and 43.

b. 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

[5] See footnote 7 of this section.

REF ID:A59439

not to expect that in practical work a general solution will, in the cryptanalytic as in the mathematical field, *invariably* 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.

42. 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 39*e*. 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 intermixture or alternation 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 Θ_1 and Θ_2 are used to indicate the two classes of components, 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.

EVEN NUMBER OF COLUMNS

+	−	+	−	+	−	+	−	+	−
Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2
Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2
Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2

a

ODD NUMBER OF COLUMNS

+	−	+	−	+	−	+	−	+
Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1
Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2
Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1	Θ_2	Θ_1

b

FIGURE 63.

c. In what follows, the term "odd column" will mean merely that the column in question occupies an odd position (1st, 3d, 5th, . . .) in the transposition rectangle; the term "even column," that it occupies an even position (2d, 4th, 6th, . . .) in the rectangle. The odd or even designation has no reference whatever to the nature of the transposition key number applicable to that column, whether it is odd or even. Now when the transposition is applied to the even-width rectangle *a*, figure 63, the cryptographic text will consist of a number of sections of letters, each section corresponding to a column of the rectangle, and therefore the number of

324155—41——9

REF ID:A59439

sections in this case will be even. Moreover, all the components in a section corresponding to an odd column in rectangle *a* will be Θ_1 or initial components, all those in a section corresponding to an even column, Θ_2 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 $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1 \ldots$; in a section corresponding to an even column the order is $\Theta_2 \rightarrow \Theta_1 \rightarrow \Theta_2 \ldots$. 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 follows along the same lines as explained in paragraph 40, with the modification that in the pairing of columns the number of possibilities is greatly reduced, since it is useless to pair two columns both containing initial components or final components.

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 the normal frequencies of the letters as they appear in English telegraphic plain text [6] are inserted in the cells and totals are obtained vertically and horizontally, these totals will permit of assigning frequency weights to the letters ADFGVX as initial and as final letters of the bipartite cipher equivalents of the plain-text letters. This is shown below in figure 64. The bipartite letter A

2D COMPONENT

1ST COMPONENT	A	D	F	G	V	X	Sums
A	T 92	H 34		E 130		F 28	284
D		L 36	O 75	W 16	R 76	S 61	264
F	A 74		B 10		M 25	I 74	183
G		N 79	P 27	G 16		C 31	153
V		D 42		J 2		K 3	47
X	Q 3	U 26	V 15	X 5	Y 19	Z 1	69
Sums	169	217	127	169	120	198	1.000

FIGURE 64.

[6] As given in fig. 3, p. 13, *Military Cryptanalysis, Part 1.*

REF ID:A59439

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 120 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 ADFGVX 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. For the purposes of this study the set of 12 cryptograms given below will be used.

I

V D D G G G V F D F V D V V F V D G A D D A F F F
V D X F D D X D V X A D V D V F X G D F V A D D G
D G D G V G D D D F X F A D A V D V G D G A D X V
D A D A D F X A V F V D D A A V D F F D F V G D F
V D D G V D D D D A V A D A F A D D X A D D G A D
F V G F V D G A D V F X V X D G D D A G G D D X F
F D D X A D F G D A G X D D A V F D A F G V F V F
A F F V F A F X G F X D G V A D F V D G G A V G G
D D G D V X A X F D D X (212 letters)

II

V D A A V D D F X F X D D A X G X F X D D F X A D
V A G D D F A X D V A V D V D D F V F V F F G D G
F V A X V X A V G D V D X F D X D G A X G F G G F
V F G D F V D X A V X D D V G D D V G V A G F X F
A A A X D D X G (108 letters)

III

D A G A A F G A G V D A F G G X F D X D F V V X G
F X F D X D D A G A D D G V A D D V D D G A F G A
V G D G X D D D A V F V D D F D A A A A D X A G D
X A G G D D A V G V F G D V F V D G G X G G A F F
V F D A X G D D D G D A F D A D G G A D D G D X A
F V D F D X F V G D D V A V F D D D V F A G D F F
F X A A D F A D G G V F D A V D G X F V D A A V G
D X F G G D D X G D A (186 letters)

REF ID:A594319

IV

A D X V F X V G G V F D D V A F G A A V F D G V D
D D G D G F D V V A F G X F X F D D D D V G D A X
D A X D D D A G V F F A A D V G D F X G X G V G D
D D D A D V X V F A V D A X X D F A A F A V D V G
V D V D D A X D A A (110 letters)

V

D F X F D D V V V D X F X F X F F F V A G F D X A
V D A G F D V D G F A D A A D F D V F G D A D F V
F V F X G X D D A G D V G V F D G X X D F F G D G
X G V D D V D D F G F V G D D V F V A G X X D F V
D X A V F G A G A G A X D V D F X G V G D A D D X
A G X D A D F D G X F D G G F V G X V V G D D D A
G X V D G V D V G X D D F D D V A G A A D G D D F
D G A G D F D D D D X G V G V G G G D G X D F G F
A D (202 letters)

VI

G D G F X A G V F V D D X G X D V D D A X D A A X
F A G V G D X F F V X F A D G F F D X A A F V X F
D F X F V G D G F X F D V V X V G D F V D D V F D
F V V D V D G G V F X F G V X F F V G V D D G D D
D D G D D A V G V X G A F F X F V D D D (120 letters)

VII

G A F G F F X F V F G F X A V A G G X D X X D D F
A G V D D V D V F F A D A V A V F V G G A D A A F
V F D F V D X F X X G D X D D F V D F F X D V F X
V A D X V A X D V X A F F V D F D G X F D G F D D
F V D V V A A F V F F V X D G F D D V A D D F D D
D X F F A G F X F X A A G V D G G V D F G G G X D
F D F V A F F G F X G D A X D G D G G D D A V D X
A D F A F V F X D D X V A G D V V D D F X D G X X
D V F V F D D D D A A F D F X D X G D A A F V D F
D V D D V A D D V D V A V D G A F V F X F A A V D
D F V D (254 letters)

VIII

D G V V G F X G G G A D F A F V V V A X A V G G V
V D V G V V D A V G D G D G A V F D D A D D D X X
D X F V F X G V G G D G D F G G D A D F D D X A V
F D D V F A D X G D A D G V A F F X A D F A D X D
G F A D F D D G V D V X A V A D D X F F A G D X F
F V F G F G F D F D V D X X D D G G D (144 letters)

REF ID:A59439

IX

G D D D D X V G V D V D A V G F G D F V D V A V D
G F A G X A V F F G V A D D D A X X A X D G A D G
X A V V D G X X A A A V A D A D G X D V G D D D D
G V F X A A V G G V F X D A F D G V G A F G D D F
A V V G D D V D F X D V D G F V A A G D X F D V A
A D A G D A X F V G D D D A G V A V F G X X F D D
G X F V D G G D A V D A G G F D A X D X F F V G F
A X X A D D F (182 letters)

X

D G D D F V F A V D V F D A D G F V G V G G D F V
D V V X D D F D D V G X G V D X G V G D X D G D X
F X F D X V D A A D D F X D D A F F A A F V F A G
D A A G G F A X G V X X F X A D G D F D G X G D A
D A X G V V V D A A G G V F G V A V F V A A G A X
G X D G A (130 letters)

XI

V F D D V A X G D A D F G G G G F G D D F X X D A
F D D X G G A V G A G D V D F D F D D D G A F A F
D A A A G V A V F G G V A D D G D D F G F V D D A
D F G A F D F V D D F V V V A D A G D X F X X X F
F D X G D F D G F D D F G D A G F A A G G A D X D
G V D G A V G V D F D D F X G A G X F G V F V V D
G V D X D F F F X G X G X A G A G V G D V V X G F
V D X D D X F V D D X (186 letters)

XII

X F D F X V V D V D A V D A D V F A G D G V A D D
F D A A D X A D F V G V D G F X F G D V F V D D D
D G D V V A V V V F A D D A X A V F V A D A X D V
G D D F A X D D G X G V F X A V X V F D G D X D F
D V X A D V A V A V G V D D D A F D F A D V F F V
V G D A G F X D D F A D V X V D F X F F V V G F X
X G F V A V F A G G D A V V D X D X G D D V V A D
D D A G A A G X F G D D D G V F G F V G V X G V F
D F F D A A D V D D X G D F D D V D D G A F G D
(224 letters)

43. Illustration of solution.[7]—*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

[7] This illustration uses 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 ADFGVX Cipher*, prepared by Messrs. Rowlett, Kullback, and Sinkov, in 1934.

REF ID:A594382

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 1). 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. Assuming it is known that the enemy is using transposition keys of not less than 15 nor more than 22 numbers, the latter 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 [(25×9)−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. The sections taken in the case of the 12 cryptograms are shown in figure 65.

Cryptogram	Length	Letters taken	Cryptogram	Length	Letters taken
I	212	VDDGGGVF	VII	254	GAFGFFXFVF
II	108	VDAA	VIII	144	DGVVG
III	186	DAGAAFG	IX	182	GDDDDXV
IV	110	ADXV	X	130	DGDDF
V	202	DFXFDDVV	XI	186	VFDDVAX
VI	120	GDGF	XII	224	XFDFXVVD

FIGURE 65.

b. The odd and the even letters of these 12 sections are then distributed separately, the results being shown in figures 66 and 67. A consideration of the mechanics of this system leads to the expectation that *if the transposition rectangle has an even number of columns the two distributions will be similar; if it has an odd number, they will be different.* The similarity or difference between the two distributions is usually discernible with as few as 20 or 25 letters.

Odd (1st, 3d, . . .) letters

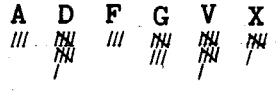


FIGURE 66.

Even (2d, 4th, . . .) letters

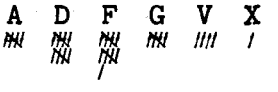


FIGURE 67.

c. 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 V's and X's (over their frequencies in fig. 67); class 2 components (fig. 67) are characterized by a predominance of F's (over its frequency in fig. 66).

REF ID:A59439

e. 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 this column, the V's and X's fall predominantly in the odd positions, the F's fall predominantly in the even positions. Therefore, beginning with position 1, the components in this column show an alternation of the type $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$. By referring to figure 63 it will become clear that if class 1 components are initial components, then it must follow that column 1 occupies an odd position in the transposition rectangle; but if class 1 components are final components, then it must follow that column 1 occupies an even position in the transposition rectangle. Which of these alternatives is true cannot be ascertained at the moment. *But the important point to be noted is that a definite reversal in the type of alternation of class 1 and class 2 components indicates the transit, in the transposition, from the end of one column to the beginning of the next column.* That is, if it is found that from the beginning of the cryptogram the alternation of components is $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$ and after a number of letters this alternation changes to $\Theta_2 \rightarrow \Theta_1 \rightarrow \Theta_2$, the point where this change occurs marks the end of column 1 and the beginning of the column 2. For the sake of brevity in reference, in the subsequent paragraphs the type of alternation $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$ will be designated as the "$\dot{+}$ type," and this type of alternation characterizes columns which fall in the odd positions in the transposition rectangle *i. e.*, in the 1st, 3d, 5th, 7th, . . . positions from the left. The other type, $\Theta_2 \rightarrow \Theta_1 \rightarrow \Theta_2$ will be designated as the "$-$ type," and this type of alternation characterizes columns which fall in the even positions in the transposition rectangle *i. e.*, in the 2d, 4th, 6th, 8th, . . . positions from the left.

f. With these principles in mind, let cryptograms III and XI, each containing 186 letters, be studied. They may be superimposed, since they have identical numbers of letters and therefore the columns end at exactly the same points in both cryptograms.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
III.	D	A	G	A	A	F	G	A	G	V	D	A	F	G	G	X	F	D	X	D	F	V	V
XI.	V	F	D	D	V	A	X	G	D	A	D	F	G	G	G	G	F	G	D	D	F	X	X
	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
III.	X	G	F	X	F	D	X	D	D	A	G	A	D	D	G	V	A	D	D	V	D	D	G
XI.	D	A	F	D	D	X	G	G	A	V	G	A	G	D	V	D	F	D	F	D	D	D	G
	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
III.	A	F	G	A	V	G	D	G	X	D	D	D	A	V	F	V	D	D	F	D	A	A	A
XI.	A	F	A	F	D	A	A	A	G	V	A	V	F	G	G	V	A	D	D	G	D	D	F
	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92
III.	A	D	X	A	G	D	X	A	G	G	D	D	A	V	G	V	F	G	D	V	F	V	D
XI.	G	F	V	D	D	A	D	F	G	A	F	D	F	V	D	D	F	V	V	V	A	D	A
	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115
III.	G	G	X	G	G	A	F	F	V	F	D	A	X	G	D	D	D	G	D	A	F	D	A
XI.	G	D	X	F	X	X	X	F	F	D	X	G	D	F	D	G	F	D	D	F	G	D	A
	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138
III.	D	G	G	A	D	D	G	D	X	A	F	V	D	F	D	X	F	V	G	D	D	V	A
XI.	G	F	A	A	G	G	A	D	X	D	G	V	D	G	A	V	G	V	D	F	D	D	F
	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161
III.	V	F	D	D	D	V	F	A	G	D	F	F	F	X	A	A	D	F	A	D	G	G	V
XI.	X	G	A	G	X	F	G	V	F	V	V	D	G	V	D	X	D	F	F	F	X	G	X
	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184
III.	F	D	A	V	D	G	X	F	V	D	A	A	V	G	D	X	F	G	G	D	D	X	G
XI.	G	X	A	G	A	G	V	G	D	V	V	X	G	F	V	D	X	D	D	X	F	V	D
	185	186																					
III.	D	A																					
XI.	D	X																					

FIGURE 68

RECD 10 DEC 1934

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:

	1	2	3	4	5	6	7	8	9	10
III.	D	A	G	A	A	F	G	A	G	V
XI.	V	F	D	D	V	A	X	G	D	A

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 +.

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

	11	12	13	14	15	16	17	18	19	20
III.	D	A	F	G	G	X	F	D	X	D
XI.	D	F	G	G	G	G	F	G	D	D

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 (19), one in an even position (16). There are no V's among these letters. 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, 3d, . . .) letters

A	D	F	G	V	X
///	𝍸 /	///	𝍸 ///	𝍸 𝍸 /	𝍸 /

Total = 3 + 11 + 3 + 8 + 11 + 6 = 42 letters

Even (2d, 4th, . . .) letters

A	D	F	G	V	X
////	𝍸 𝍸	𝍸 𝍸 /	𝍸	////	/

Total = 4 + 10 + 11 + 5 + 4 + 1 = 35 letters

FIGURE 69.

Since the odd letters have a total frequency of 42, the even, a total frequency of 35, 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.

i. 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.

	11	12	13	14	15	16	17	18	19	20
III.	D	A	F	G	G	X	F	D	X	D
XI.	D	F	G	G	G	G	F	G	D	D

Distribution of odd letters

	A	D	F	G	V	X
		///	///	///		/
f =	0	3	3	3	0	

Distribution of even letters

	A	D	F	G	V	X
	/	///	/	////		/
f =	1	3	1	4	0	1

FIGURE 70.

REF ID:A59439

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 Θ_1's, even letters as Θ_2's):

$0(3)+3(11)+3(3)+3(8)+0(11)+1(6)=72$

$72-\frac{72}{6}=60$

$1(4)+3(10)+1(11)+4(5)+0(4)+1(1)=66 \quad =66$

Total=126

On the basis of figure 69 reversed (even letters as Θ_1's, odd letters as Θ_2's):

$1(3)+3(11)+1(3)\ +4(8)+0(11)+1(6)=77$

$77-\frac{77}{6}=64$

$0(4)+3(10)+3(11)+3(5)+\ 0(4)+1(1)=79 \quad =79$

Total=143

j. Now the frequency sums here obtained (126 *vs.* 143) indicate that an alternation of the type $\Theta_2 \rightarrow \Theta_1 \rightarrow \Theta_2$ 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.

	21	22	23	24	25	26	27	28	29	30
III.	F	V	V	X	G	F	X	F	D	X
XI.	F	X	X	D	A	F	D	D	X	G

Distribution of odd letters

	A	D	F	G	V	X
	—	//	//	—	—	///
f=	1	2	2	1	1	3

Distribution of even letters

	A	D	F	G	V	X
		//	///	—	—	///
f=	0	2	3	1	1	3

On the basis of figure 69 normal (odd letters as Θ_1's, even letters as Θ_2's):

$1(3)+2(11)+\ 2(3)+1(8)+1(11)+3(6)=68$

$68-\frac{68}{6}=57$

$0(4)+2(10)+3(11)+1(5)+\ 1(4)+3(1)=65 \quad =65$

Total=122

On the basis of figure 69 reversed (even letters as Θ_1's, odd letters as Θ_2's):

$0(3)+2(11)+\ 3(3)+1(8)+1(11)+3(6)=68$

$68-\frac{68}{6}=57$

$1(4)+2(10)+2(11)+1(5)+\ 1(4)+3(1)=58 \quad =58$

Total=115

FIGURE 71.

REF ID:A594339

Beginning with the 21st position, the alternation is of type $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$; hence it is of the "+" type. Again the reversal in type of alternation occurs in passing from the 2d set of 10 letters to the 3d set, and this indicates that column 2 of the transposition rectangle ends somewhere near the 20th 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)

	1	2	3	4	5	6	7	8	9	10	11
III.	D	A	G	A	A	F	G	A	G	V	D
XI.	V	F	D	D	V	A	X	G	D	A	D

Distribution of odd letters

	A	D	F	G	V	X
	/	≋		///	//	/
f=	1	5	0	3	2	1

Distribution of even letters

	A	D	F	G	V	X
	≋	/	//	/	/	
f=	5	1	2	1	1	0

Weighted values of distributions:

On the basis of figure 69 normal (odd letters as Θ_1's, even letters as Θ_2's):

$1(3)+5(11)+\ 0(3)+3(8)+2(11)+1(6)=110$

$110-\frac{110}{6}=92$

$5(4)+1(10)+2(11)+1(5)+\ 1(4)+0(1)=61$ $\quad =61$

Total=153

On the basis of figure 69 reversed (even letters as Θ_1's, odd letters as Θ_2's):

$5(3)+1(11)+\ 2(3)+1(8)+1(11)+0(6)=51$

$51-\frac{51}{6}=\ 42$

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

Total=120

The type of alternation is $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$, or "+".

Second Section (letters 12–22)

	12	13	14	15	16	17	18	19	20	21	22
III.	A	F	G	G	X	F	D	X	D	F	V
XI.	F	G	G	G	G	F	G	D	D	F	X

Distribution of odd letters

	A	D	F	G	V	X
		/	≋	///		
f=	0	1	5	3	0	1

Distribution of even letters

	A	D	F	G	V	X
	/	///	/	////	/	//
f=	1	3	1	4	1	2

REF ID:A59439

Weighted values of distributions:

On the basis of figure 69 normal (odd letters as θ_1's, even letters as θ_2's):

$$0(3)+1(11)+5\ (3)+3(8)+0(11)+1(6)=56$$
$$56-\frac{56}{6}=47$$
$$1(4)+3(10)+1(11)+4(5)+\ 1(4)+2(1)=71 \qquad =71$$
$$\text{Total}=118$$

On the basis of figure 69 reversed (even letters as θ_1's, odd letters as θ_2's):

$$1(3)+3(11)+\ 1(3)+4(8)+1(11)+2(6)=94$$
$$94-\frac{94}{6}=78$$
$$0(4)+1(10)+5(11)+3(5)+\ 0(4)+1(1)=81 \qquad =81$$
$$\text{Total}=159$$

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 $\theta_1 \rightarrow \theta_2 \rightarrow \theta_1$, or "+".

THIRD SECTION (letters 23–33)

	23	24	25	26	27	28	29	30	31	32	33
III.	V	X	G	F	X	F	D	X	D	D	A
XI.	X	D	A	F	D	D	X	G	G	A	V

Distribution of odd letters

	A	D	F	G	V	X
	=	≡		=	=	≡
f=	2	3	0	2	2	3

Distribution of even letters

	A	D	F	G	V	X
	~	≡	≡	~		=
f=	1	3	3	1	0	2

Weighted values of distributions:

On the basis of figure 69 normal (odd letters as θ_1's, even letters as θ_2's):

$$2(3)+3(11)+\ 0(3)+2(8)+2(11)+3(6)=95$$
$$95-\frac{95}{6}=79$$
$$1(4)+3(10)+3(11)+1(5)+\ 0(4)+2(1)=74 \qquad =74$$
$$\text{Total}=153$$

On the basis of figure 69 reversed (even letters as θ_1's, odd letters as θ_2's):

$$1(3)+3(11)+\ 3(3)+1(8)+0(11)+2(6)=65$$
$$65-\frac{65}{6}=54$$
$$2(4)+3(10)+0(11)+2(5)+\ 2(4)+3(1)=59 \qquad =59$$
$$\text{Total}=113$$

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 $\theta_1 \rightarrow \theta_2 \rightarrow \theta_1$, or "+".

REF ID:A594439

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 [(17×11)−1=186].

l. There is another cryptogram which also contains but one short column, viz, VII, of 254 letters [17×15)−1=254]. The columns of this cryptogram contain 4 more letters than the corresponding columns of III and XI. Assuming, momentarily, that the last column is the short one, cryptogram VII may be added to the superposition of III and XI, provided these sets of 4 additional letters are accounted for. This has been done in figure 72. In that figure the 4 extra letters pertaining to cryptogram VII are shown as falling under the last letters of the columns of cryptograms III 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.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
III.	D	A	G	A	A	F	G	A	G	V	D	A	F	G	G	X	F	D	X	D	F	V
XI.	V	F	D	D	V	A	X	G	D	A	D	F	G	G	G	G	F	G	D	D	F	X
VII.	G	A	F	G	F	F	X	F	V	F	G	A	G	G	X	D	X	X	D	D	F	A
								F	X	A	V								G	V	D	D

	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
III.	V	X	G	F	X	F	D	X	D	D	A	G	A	D	D	G	V	A	D	D	V	D
XI.	X	D	A	F	D	D	X	G	G	A	V	G	A	G	D	V	D	F	D	F	D	D
VII.	V	D	V	F	F	A	D	A	V	A	V	A	D	A	A	F	V	F	D	F	V	D
								F	V	G	G								X	F	X	X

	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
III.	D	G	A	F	G	A	V	G	D	G	X	D	D	D	A	V	F	V	D	D	F	D
XI.	D	G	A	F	A	F	D	A	A	A	G	V	A	V	F	G	G	V	A	D	D	G
VII.	G	D	X	D	D	F	V	D	F	F	X	V	A	D	X	V	A	X	D	V	X	A
								D	V	F	X								F	F	V	D

	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
III.	A	A	A	A	D	X	A	G	D	X	A	G	G	D	D	A	V	G	V	F	G	D
XI.	D	D	F	G	F	V	D	D	A	D	F	G	A	F	D	F	V	D	D	F	V	V
VII.	F	D	G	X	F	D	G	F	D	D	F	A	A	F	V	F	F	V	X	D	G	F
								V	D	V	V								D	D	V	A

	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
III.	V	F	V	D	G	G	X	G	G	A	F	F	V	F	D	A	X	G	D	D	D	G
XI.	V	A	D	A	G	D	X	F	X	X	X	F	F	D	X	G	D	F	D	G	F	D
VII.	D	D	F	D	D	D	X	F	F	A	G	A	A	G	V	D	G	G	V	D	F	G
								F	X	F	X								G	G	X	D

	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132
III.	D	A	F	D	A	D	G	G	A	D	D	G	D	X	A	F	V	D	F	D	X	F
XI.	D	F	G	D	A	G	F	A	A	G	G	A	D	X	D	G	V	D	G	A	V	G
VII.	F	D	F	V	A	F	F	G	F	X	G	G	D	G	G	D	D	A	V	D	X	A
								D	A	X	D								D	F	A	F

FIGURE 72.

REF ID:A59439

	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154
III.	V	G	D	D	V	A	V	F	D	D	D	V	F	A	G	D	F	F	F	X	A	A
XI.	V	D	F	D	D	F	X	G	A	G	X	F	G	V	F	V	V	D	G	V	D	X
VII.	V	F	X	D	D	X	V	A	G	D	V	X	D	G	X	X	D	V	F	V	F	D
								V	D	D	F								D	D	D	A

	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176
III.	D	F	A	D	G	G	V	F	D	A	V	D	G	X	F	V	D	A	A	V	G	D
XI.	D	F	F	F	X	G	X	G	X	A	G	A	G	V	G	D	V	V	X	G	F	V
VII.	A	F	D	F	X	D	X	G	D	A	A	D	V	D	D	V	A	D	D	V	D	V
								F	V	D	F								A	V	D	G

	177	178	179	180	181	182	183	184	185	186
III.	X	F	G	G	D	D	X	G	D	A
XI.	D	X	D	D	X	F	V	D	D	X
VII.	A	F	V	F	X	F	A	A	V	D
							D	F	V	D

FIGURE 72.—Continued.

Cryptogram	Length	Letters taken	Cryptogram	Length	Letters taken
I	212	VDDGGGVFDFVD	VII	254	GAFGFFXFVFGFXA
II	108	VDAAVD	VIII	144	DGVVGFXG
III	186	DAGAAFGAGV	IX	182	GDDDDXVGVD
IV	110	ADXVFX	X	130	DGDDFVF
V	202	DFXFDDVVVDX	XI	186	VFDDVAXGDA
VI	120	GDGFXAG	XII	224	XFDFXVVDVDAVD

FIGURE 73.

The new frequency weights are therefore as follows:

Odd (1st, 3d, . . .) letters

A	D	F	G	V	X
4	14	5	11	15	10

Total = 4 + 14 + 5 + 11 + 15 + 10 = 59

Even (2d, 4th, . . .) letters

A	D	F	G	V	X
9	15	14	8	7	2

Total = 9 + 15 + 14 + 8 + 7 + 2 = 55

FIGURE 74.

Since the two totals are quite close together, no correction need be made of the nature of that made in preceding calculations, where one-sixth was deducted from the total values of odd letters.

n. Beginning with position 23, in the case of cryptograms III and XI, the next 11 letters, and, in the case of cryptogram VII, the next 15 letters are clearly of the "+" type of alternation. The data are as follows:

	23	24	25	26	27	28	29	30	31	32	33
III.	V	X	G	F	X	F	D	X	D	D	A
XI.	X	D	A	F	D	D	X	G	G	A	V
VII.	V	D	V	F	F	A	D	A	V	A	V
								F	V	G	G

REF ID:A594339

Distribution of odd letters

A	D	F	G	V	X
//	////	/	///	//// //	///
f= 2	4	1	3	7	3

Distribution of even letters

A	D	F	G	V	X
////	////	////	//		//
f= 4	4	5	2	0	2

Weighted values of distributions:

On the basis of figure 74 normal (odd letters as θ_1's, even letters as θ_2's):

$$2(4)+4(14)+1(5)+3(11)+7(15)+3(10)=237$$
$$4(9)+4(15)+5(14)+2(8)+0(7)+2(2)=186$$
$$\text{Total}=423$$

On the basis of figure 74 reversed (even letters as θ_1's, odd letters as θ'_2's):

$$4(4)+4(14)+5(5)+2(11)+0(15)+2(10)=139$$
$$2(9)+4(15)+1(14)+3(8)+7(7)+3(2)=171$$
$$\text{Total}=310$$

Since the greatest total is obtained on the basis of figure 74 normal, the type of alternation for the third section of letters is $\theta_1 \rightarrow \theta_2 \rightarrow \theta_1$, or "+".

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

	34	35	36	37	38	39	40	41	42	43	44
III.	G	A	D	D	G	V	A	D	D	V	D
XI.	G	A	G	D	V	D	F	D	F	D	D
VII.	A	D	A	A	F	V	F	D	V	F	D
								X	F	X	X

Distribution of odd letters

A	D	F	G	V	X
///	//// ///	/		///	//
f= 3	8	1	0	3	2

Distribution of even letters

A	D	F	G	V	X
///	////	////	////	//	/
f= 3	5	5	4	2	1

Weighted values of distributions:

On the basis of figure 74 normal (odd letters as θ_1's, even letters as θ_2's):

$$3(4)+8(14)+1(5)+0(11)+3(15)+2(10)=194$$
$$3(9)+5(15)+5(14)+4(8)+2(7)+1(2)=220$$
$$\text{Total}=414$$

On the basis of figure 74 reversed (even letters as θ_1's, odd letters as θ_2's):

$$3(4)+5(14)+5(5)+4(11)+2(15)+1(10)=191$$
$$3(9)+8(15)+1(14)+0(8)+3(7)+2(2)=186$$
$$\text{Total}=377$$

REF ID:A59439

Since the distribution begins here with an even-numbered position (34), and the greatest total is obtained on the basis of figure 74 normal, hence the alternation for the fourth section or column is of the type $\Theta_2 \rightarrow \Theta_1 \rightarrow \Theta_2$, or "$-$".

p. (1) The data for the letters beyond position 44 are as follows:

	45	46	47	48	49	50	51	52	53	54	55
III.	D	G	A	F	G	A	V	G	D	G	X
XI.	D	G	A	F	A	F	D	A	A	A	G
VII.	G	D	X	D	D	F	V	D	F	F	X
								D	V	F	X

Distribution of odd letters

A	D	F	G	V	X
///	𝍸	/	///	///	///
f= 4	5	1	3	3	4

Distribution of even letters

A	D	F	G	V	X
///	///	𝍸 /	///		
f= 3	4	6	4	0	0

Weighted values of distributions:

On the basis of figure 74 normal (odd letters as Θ_1's, even letters as Θ_2's):

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

On the basis of figure 74 reversed (even letters as Θ_1's, odd letters as Θ_2's):

$$3(4)+4(14)+6(5)+4(11)+0(15)+0(10)=142$$
$$4(9)+5(15)+1(14)+3(8)+3(7)+4(2)=178$$
$$\text{Total}=320$$

Since the distribution starts with an odd position (45) and the greatest total is obtained on the basis of figure 74 normal, the type of alternation for the fifth section or column is $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$, or "$+$".

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

	56	57	58	59	60	61	62	63	64	65	66
III.	D	D	D	A	V	F	V	D	D	F	D
XI.	V	A	V	F	G	G	V	A	D	D	G
VII.	V	A	D	X	V	A	X	D	V	X	A
								F	F	V	D

Distribution of odd letters

A	D	F	G	V	X
𝍸	///	///	/	/	//
f= 5	4	4	1	1	2

Distribution of even letters

A	D	F	G	V	X
/	𝍸 //	/	//	𝍸 ///	/
f= 1	7	1	2	8	1

REF ID:A594389

Weighted values of distributions:

On the basis of figure 74 normal (odd letters as θ_1's, even letters as θ_2's):

$$5(4)+4(14)+\ 4(5)+1(11)+1(15)+2(10)=142$$
$$1(9)+7(15)+1(14)+\ 2(8)+\ 8(7)+1(2)\ =202$$
$$\text{Total}=344$$

On the basis of figure 74 reversed (even letters as θ_1's, odd letters as θ_2's):

$$1(4)+7(14)+\ 1(5)+2(11)+8(15)+1(10)=259$$
$$5(9)+4(15)+4(14)+\ 1(8)+\ 1(7)+\ 2(2)=180$$
$$\text{Total}=439$$

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 $\theta_1 \rightarrow \theta_2 \rightarrow \theta_1$, or "+".

r. But perhaps advantage should be taken of the availability of additional cryptograms. For example, cryptogram V, of 202 letters, has 2 short columns [(17×12)−2=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?

s. Suppose column 1 of cryptogram V is the additional short column. Then the letters of column 2 would be F X F X F F F V A G F D. These letters when evaluated on the basis of figure 74 normal yield a total of 77; when weighted on the basis of figure 74 reversed, a total of 144. The calculation is as follows:

Distribution of odd letters

	A	D	F	G	V	X
	\		≋			
f=	1	0	5	0	0	0

Distribution of even letters

	A	D	F	G	V	X
		\	\	\	\	//
f=	0	1	1	1	1	2

On the basis of figure 74 normal (odd letters as θ_1's, even letters as θ_2's):

$$1(4)+0(14)+\ 5(5)+0(11)+0(15)+0(10)=29$$
$$0(9)+1(15)+1(14)+\ 1(8)+\ 1(7)+\ 2(2)=48$$
$$\text{Total}=77$$

On the basis of figure 74 reversed (even letters as θ_1's, odd letters as θ_2's):

$$1(9)+0(15)+5(14)+\ 0(8)+\ 0(7)+\ 0(2)=79$$
$$0(4)+1(14)+\ 1(5)+1(11)+1(15)+2(10)=65$$
$$\text{Total}=144$$

According to this calculation, column 2 of cryptogram V seems to correspond to the type of alternation $\theta_2 \rightarrow \theta_1 \rightarrow \theta_2$, 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:

REF ID:A59439

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

Distribution of odd letters

	A	D	F	G	V	X
	—		—	—	=	—
f=	1	0	1	1	2	1

Distribution of even letters

	A	D	F	G	V	X
	—	≡	=	—		
f=	1	3	1	1	0	0

Weighted values of distributions:

On the basis of figure 74 normal (odd letters as Θ_1's, even letters as Θ_2's):

$$1(4)+0(14)+1(5)+1(11)+2(15)+1(10) = 60$$
$$1(9)+ 3(15)+1(14)+ 1(8)+90(7)+0(2)= 76$$
$$\text{Total}=136$$

On the basis of figures 74 reversed (even letters as Θ_1's, odd letters as Θ_2'2):

$$1(9)+0(15)+1(14)+ 1(8)+ 2(7)+1\ (2)= 47$$
$$1(4)+3(14)+ 1(5)+1(11)+0(15)+0(10)= 62$$
$$\text{Total}=109$$

Since the greatest total is obtained on the basis of figure 74 normal, the type of alternation is $\Theta_1 \rightarrow \Theta_2 \rightarrow \Theta_1$ 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 "−" 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.

t. 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 "−" shown in figure 75.

Assumption	Column																	Summation of +'s and −'s
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
(1) 6th short	+	+	+	−	+	+	+	+	−	+	+	+	−	−	−	−	−	10+, 7−
(2) 7th short	+	+	+	−	+	+	−	+	−	+	+	+	−	−	−	−	−	9+, 8−
(3) 8th short	+	+	+	−	+	+	−	−	−	+	+	+	−	−	−	−	−	8+, 9−
(4) 9th short	+	+	+	−	+	+	−	−	+	+	+	+	−	−	−	−	−	9+, 8−
(5) 10th short	+	+	+	−	+	+	−	−	+	−	+	+	−	−	−	−	−	8+, 9−
(6) 11th short	+	+	+	−	+	+	−	−	+	−	−	+	−	−	−	−	−	7+, 10−
(7) 12th short	+	+	+	−	+	+	−	−	+	−	−	−	−	−	−	−	−	6+, 11−
(8) 13th short	+	+	+	−	+	+	−	−	+	−	−	−	+	−	−	−	−	7+, 10−
(9) 14th short	+	+	+	−	+	+	−	−	+	−	−	−	+	+	−	−	−	8+, 9−
(10) 15th short	+	+	+	−	+	+	−	−	+	−	−	−	+	+	+	−	−	9+, 8−
(11) 16th short	+	+	+	−	+	+	−	−	+	−	−	−	+	+	+	+	−	10+, 7−
(12) 17th short	+	+	+	−	+	+	−	−	+	−	−	−	+	+	+	+	+	11+, 6−

FIGURE 75.

u. The correct assumption must satisfy the following conditions:

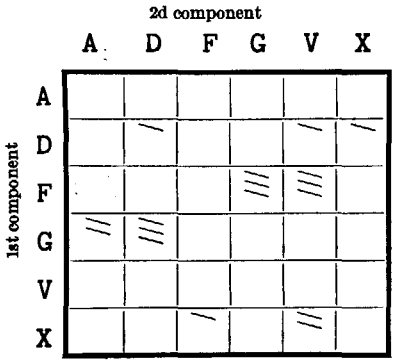
(*a*) There must be 9 "−" and 8 "+" columns.

(*b*) The short column must be "−".

324155—41——10

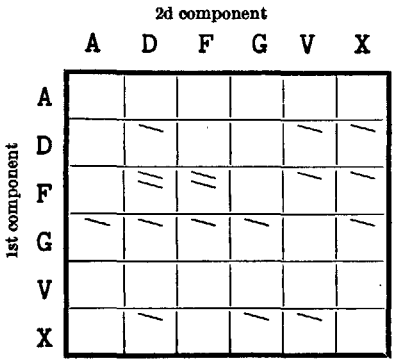
REF ID:A594389

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 monoalphabeticity 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.



$E(\phi) = .0667 \times 17 \times 16 = 18.14$
$\phi = 22$

FIGURE 76.



$E(\phi) = .0667 \times 17 \times 16 = 18.14$
$\phi = 4$

FIGURE 77.

v. 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.

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

16	5	7	6	9	3	14	1	13	11	17	10	4	12	15	2	8
V	I	K	I	N	G	S	C	R	O	W	N	H	O	T	E	L

2d component

1st component	A	D	F	G	V	X
A	V	I	9	K	N	G
D	7	S	C	3	R	O
F	W	H	8	T	E	5
G	L	A	1	B	2	D
V	4	F	6	J	Ø	M
X	P	Q	U	X	Y	Z

FIGURE 78.

x. 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-

REF ID:A59439

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 (θ_1 or θ_2) 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 39*e–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.[8]

1st component \ 2d component	A	D	F	G	V	X	Sums
A	I 74	Q 3	S 61		U 26		164
D		T 92		W 16	C 31	P 27	166
F	G 16		A 74			N 79	169
G	X 5	V 15	J 2	E 130	B 10	K 3	165
V	R 76	M 25	F 28		L 36		165
X		D 42	Z 1	Y 19	O 75	H 34	169
Sums	171	177	166	165	158	163	1,000

FIGURE 79.

z. 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.

[8] The frequencies indicated as those given in fig. 3, p. 13, *Military Cryptanalysis*, Part I.

REF ID:A194382

SECTION X

SOLUTION OF BIFID FRACTIONATING SYSTEMS

	Paragraph
Review of principles underlying the cryptographic method	44
Example of a simple bifid cipher	45
Principles of solution	46
Example of solution	47
Special solution	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.[1] 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 plain-text letter is replaced by two components, Θ_c^1 and Θ_c^2, of a bifid or bipartite alphabet; (2) a process of separation (transposition), in which the $\Theta^1\Theta_c^2$ components originally paired together are separated; (3) a process of recombination (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 plain-text 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.

[1] See Special Text No. 166, *Advanced Military Cryptography*, sec. XI and *Military Cryptanalysis, Part I*, sec. IX.

REF ID:A59439

1st component \ 2d component	1	2	3	4	5
1	M	A	N	U	F
2	C	T	R	I	G
3	B	D	E	H	K
4	L	O	P	Q	S
5	V	W	X	Y	Z

FIGURE 80.

Plain text	O	N	E		P	L	A	N	E		R	E	P	O	R	T	E	D		L	O	S	T		A	T		S	E	A
Components	4	1	3		4	4	1	1	3		2	3	4	4	2	2	3	3		4	4	4	2		1	2		4	3	1
	2	3	3		3	1	2	3	3		3	3	3	2	3	2	3	2		1	2	5	2		2	2		5	3	2

The third process, that of recombination or recomposition, 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:

O	N	E		P	L	A	N	E		R	E	P	O	R	T	E	D		L	O	S	T		A	T		S	E	A
4	1	3		4	4	1	1	3		2	3	4	4	2	2	3	3		4	4	4	2		1	2		4	3	1
2	3	3		3	1	2	3	3		3	3	3	2	3	2	3	2		1	2	5	2		2	2		5	3	2

giving the pairs 21, 33, 34, 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 preagreed. 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	O	N	E		P	L	A	N	E		R	E	P	O	R	T	E	D		L	O	S	T		A	T		S	E	A
Components	4	1	3		4	4	1	1	3		2	3	4	4	2	2	3	3		4	4	4	2		1	2		4	3	1(4)
	2	3	3		3	1	2	3	3		3	3	3	2	3	2	3	2		1	2	5	2		2	2		5	3	2
Cipher	C	E	H		H	M	C	E	D		E	H	H	T	D	R	E	I		U	I	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:

O	N	E	P	L	A	N	E	
42	13	33	43	41	12	13	33	. .
C	E	H	H	M	C	E		

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.

Rec'd 10 Oct 1938

46. Principles of solution.—*a.* Note the following skeleton encipherments, using the checkerboard shown in figure 80:

C E N O	S E N D	R E N C	H E N A	T E N Y
2 3 1 4	4 3 1 3	2 3 1 2	3 3 1 1	2 3 1 5
1 3 3 2	5 3 3 2	3 3 3 1	4 3 3 2	2 3 3 4
N B H	X B E	E B D	P B B	R B K
(1)	(2)	(3)	(4)	(5)

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 trigraphic 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 these five encipherments have in common. They are as follows:

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

Reference to the checkerboard discloses the very interesting and important fact that the letters preceding the cipher repetition (B_c) *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_p produces N B H_c. The N_c is the result of combining the second component of the bipartite equivalent of C_p with the first component bipartite equivalent of E_p, yielding the combination 13, which is N. No matter what the other three letters in the plain-text tetragraph may be, if the second letter is E_p, the second component bified equivalent of the first letter of the cipher trigraph must be a 3. This means that this first letter of the cipher trigraph 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 trigraphs 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 trigraphs: these all must 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.*[2] 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.

[2] The principle involved in such reconstruction was, to my knowledge, first pointed out and successfully employed early in 1938 by Associate Cryptanalysts S. Kullback and A. Sinkov.

REF ID:A59439

d. Note the following encipherments:

Plain	L	P	U	R		O	R	M	I
	41	43	14	23		42	23	11	24
Cipher	U	B	O			T	B	A	
	(6)					(7)			

Here the B_c is preceded by letters (U and T) which not only are not in the same column as those in the corresponding position in the case of the first five encipherments, but also these two letters are themselves in different columns. The cause of this is not difficult to see. It is merely that the second component of the P_p and the second component of the R_p *happen* to be identical, the first component of the U_p and the first component of the M_p also *happen* to be identical, thus producing the same cipher letter in both cases. This is a phenomenon which must happen by chance a certain number of times, a number which is dependent not only upon the mechanics of the system but also upon the exact composition of the checkerboard. Disregarding for the moment the latter factor, it is obvious that if the checkerboard is perfectly balanced, the bifid element 3, for example, should occur 20 percent of the time as the first or as the second element of a bifid pair, since there are 5 elements and each can theoretically appear an equal number of times. However, since the checkerboard is not perfectly balanced, the bifid element 5 can, in the case of figure 80, appear as a second component of the bipartite equivalent of a cipher letter only very rarely, since it corresponds to the first component of the bifid equivalents of the letters V, W, X, Y, and Z, all of which are of low frequency. On the other hand, the bifid element 3, in the case of figure 80, can appear very frequently as a first component of the bifid equivalent of a cipher letter because it is the second component of the bifid equivalents for the high-frequency plain-text letters N, R, and E, which are all in column 3. However, since the exact composition of the checkerboard is unknown when cryptograms of this sort are to be solved, frequency weights can, of course, not be assigned to any of the components or bipartite elements and it will have to be assumed that each one has an equal probability of occurrence, that is, one-fifth.

e. From the foregoing discussion it is obvious that it would be unwise merely to study the prefixes and suffixes to identical single letters of the cipher text in an attempt to solve cryptograms of this sort, for the disturbing effect of the accidental identities of certain cipher letters would be sufficient to retard solution. A few detailed examples of the type of study that must be made in connection with repetitions in such systems as this will now be given.

f. It was stated in subparagraph *a* that the mechanics of the system tend to reduce by one character the lengths of the repetitions in the cipher text. The expression "tend to reduce" aptly describes the situation, for not only can it happen that a 3-letter repetition in the plain text may appear to remain a 3-letter repetition in the cipher text, but also it can happen that a 3-letter repetition in the plain text may even appear as a pseudo 4-letter repetition in the cipher text. Study the following examples (based on fig. 80) and note what happens in each case:

(*A*)	F	T	H	E	C		(*A*)	N	T	H	E	D	A
	15	22	34	33	21			13	22	34	33	32	12
	W	R	P	D				D	R	P	E	C	
(*B*)	Y	T	H	E	Y		(*B*)	T	T	H	E	H	A
	54	22	34	33	54			22	22	34	33	34	12
	O	R	P	K				T	R	P	E	L	
	(1)							(2)					

A 3-letter plain text repetition appears as a 2-letter cipher text repetition.

A 3-letter plain-text repetition appears as a 3-letter cipher-text repetition because the 1st components of the D_p and H_p happen to be identical (D and H are in same row in checkerboard).

REF ID:A594319

(A)	I	N	T	H	E	G	
	24	13	22	34	33	25	
		L	D	R	P	D	
(B)	O	R	T	H	E	U	
	42	23	22	34	33	14	
		T	D	R	P	B	

(3)

A 3-letter plain-text repetition appears as a 3-letter cipher-text repetition because the second component of the N_p and R_p happen to be identical (R and N are in same column in checkerboard).

(A)	T	O	T	H	E	M	A
	22	42	22	34	33	11	12
		I	T	R	P	B	M
(B)	N	D	T	H	E	N	E
	13	32	22	34	33	13	33
		E	T	R	P	B	E

(4)

A 3-letter plain-text repetition appears as a 4-letter cipher-text repetition because the phenomena of case 2 and case 3 occur simultaneously (O and D are in same column; M and N are in the same row in the checkerboard).

g. From a study of these phenomena the rule may be deduced that an n-letter repetition in the plain text is really reduced to an $(n-1)$-letter repetition in the cipher text, but it can happen fortuitously [8] 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.

h. 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. 80 as a basis:

F	O	U	R	D	A	(YS)	
15	42	14	23	32	12	. . .	
Y	C	O	D	C			

S	O	M	E	D	U	(TY)	
45	42	11	33	32	14	. . .	
Y	C	N	D	C			

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 cryptanalyst must be very careful in respect to the deductions he may make in the case of short repetitions. In the example cited it happens that the accidental repetitions are such as to make the sequences as a whole almost appear to be identical.

i. It is these pseudo-repetitious elements which complicate the solution of what would otherwise be a simple system. To illustrate what is meant, note that in case (1) of subparagraph *f* the letters W_c and O_c, the prefixes to the repetition RP_c, do actually come from the same column of the checkerboard; the letters D_c and K_c, the suffixes, do actually come from the same row. But now note in case (2) that while the prefixes D_c and T_c come from the same column, the suffixes C_c and L_c do not come from the same row. Note also in case (3) that while the prefixes L_c and T_c do not come from the same column, the suffixes, D_c and B_c, do come from the same row; while in case (4) the prefixes turn out to be the same letter, T_c (which constitutes an example where the two prefixes come from the same column) but the suffixes, M_c and E_c, come from different rows. Since the exact length of the real repetition, without its pseudo-repetitious elements, does not readily manifest itself in the cipher text (although in favorable cases it may be deduced by a careful detailed analysis and comparison with nearly similar repetitions) the nature of the difficulties confronting the cryptanalyst become apparent.

j. The nature of the detailed analysis and comparison of repetitions referred to above may require a few words of explanation. Suppose that a cryptogram shows many occurrences of RP_c (=THE_p in the foregoing examples). It would indicate a high-frequency plain-text trigraph. A few repetitions of such cipher trigraphs as RPE_c, DRP_c, $TRPB_c$, would lead to the surmise that the latter are of the type where pseudo-repetitious elements have crept into the picture and there-

[8] Strictly speaking, of course, not really fortuitously. It depends upon the exact letters which precede or follow the plain-text repetition and the exact positions these letters occupy in the checkerboard.

REF ID:A59439

fore the cryptanalyst should be very hesitant to assume that the adventitious prefixed letters are in the same columns, or that the adventitious suffixed letters are in the same row. In fact, he would be warranted in tentatively assuming the very opposite condition, that they are not in the same columns or rows, respectively. The conclusions derivable from a study of short repetitions can be carried over to the longer ones. Note the following four cases from which several conclusions may be reached:

(1)	U	R	P	O	S	I	T	I	O	N	S	A	
	14	23	43	42	45	24	22	24	42	13	45	12	
	O	H	H	I	W	O	T	Q	C	H	V		
(2)	I	S	P	O	S	I	T	I	O	N	O	F	
	24	42	43	42	45	24	22	24	42	13	42	15	
	Q	Y	H	I	W	O	T	Q	C	H	C		
(3)	H	E	P	O	S	I	T	I	O	N	D	E	
	34	33	43	42	45	24	22	24	42	13	32	33	
	P	H	H	I	W	O	T	Q	C	E	R		
(4)	O	R	P	O	S	I	T	I	O	N	L	U	
	42	23	43	42	45	24	22	24	42	13	41	14	
	T	H	H	I	W	O	T	Q	C	H	M		

First, the 7-letter cipher sequence H I W O T Q C is common to all four cases; if only the cipher 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 O, P, and T, they may or may not come from the same column, most probably not. (Actually, O 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 V, C, and M, they may or may not come from the same row, most probably not. (Actually, all 3 letters come from different rows.)

k. Note the following cases of encipherment: The fact that the 7-letter cipher sequence is common to all four cases means that the plain-text repetition consists of 8 letters.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
	N	T	O	P	O	S	I	T	I	O	N	S	I	F	T	H	E
(1)	13	22	42	43	42	45	24	22	24	42	13	45	24	15	22	34	33
	D	I	I	H	I	W	O	T	Q	C	H	W	L	W	R	P	
	O	N	T	P	O	S	I	T	I	O	N	F	I	F	T	H	R
(2)	42	13	22	43	42	45	24	22	24	42	13	15	24	15	22	34	23
	C	D	I	H	I	W	O	T	Q	C	B	W	L	W	R	O	
	W	A	S	P	O	S	I	T	I	O	N	L	I	F	T	I	N
(3)	52	12	45	43	42	45	24	22	24	42	13	41	24	15	22	24	13
	C	I	Y	H	I	W	O	T	Q	C	H	A	L	W	T	L	
	T	A	L	P	O	S	I	T	I	O	N	T	I	F	L	I	S
(4)	22	12	41	43	42	45	24	22	24	42	13	22	24	15	41	24	45
	C	I	U	H	I	W	O	T	Q	C	D	T	L	Y	A	Q	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

REF ID:A59439

The pseudo-repetitious letters, I in the third position in cases (1) and (2), and the letters Y_c and U_c in corresponding positions in cases (3) and (4) mean that I, U, and Y, come from the same column of the checkerboard. The I_c in position 2, in cases (1), (3), and (4) and the D_c in the corresponding position in case (2) indicates that I_c and D_c are probably in different columns in the checkerboard. In position 11, the H_c, B_c, and D_c give indications of being in the same row of the checkerboard. In position 14, W_c and Y_c 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:

X A B Q
N A B R
Z A B T
N A B Q
L A B I
Z A B T

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 keyword 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.

REF ID:A594389

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	Location
E	W	V	R	V	S	D	D	T	W																																	2I
D	W	V	R	V	S	D	D	T	I																																	2B
R	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	B																											1B
													G	U	H	K	I	R	X	Y	I	R	N	G	T	I	W	A	U	U	I	A	A	Q	C	W	N			1D
P	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	H	K	I	R	X	Y	I	R	N	G	T	I	W	A	U	U	I	A	A	Q	C	Q	D	T	Q	1L
X	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	H	K	I	R	X	Y	I	R	N	B																	2D
															L	K	I	R	X	Y	I	R	N	G	T	I	W	A	U	U	I	A	A	Q	C	Q	D	T	V	1I
																				K	I	R	N	G	S	T	L	B	D	D	T	I								2F
																				N	I	R	N	G	S	T	L	B	D	D	T	W								1G
																										T	W	A	U	U	I	A	A	Q	D					1H

																					L	Q	D	T	X	D	W	Y	I	W	X	N	I	N	B	D	R	Q	1H
													W	S	A	V	I																								1A
													Q	S	A	V	N	I	I	U	H	Q	D	T	X	W	I															1E
														T	A	V	N	N	I																							1F
														Q	A	V	N	I	I	U	H	Q	D	T	X	D	W	Y	I	W	X	N	I	N	B	D	R	W	2G
																					L	Q	D	T	X	C	W	S	C	I	V	D	T	N	I	L	V	N	2H
																				D	L	Q	D	T	X	W	S	C	L	P										1K
																					C	Q	D	T	Z															1J
												N	I	L	V	N	O	D	N	U	L	Q	D	U	B	G															2I
											R	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	B	I															1B
											X	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	H	K															2E
																					L	L	Q	D	U	K															2C

																					T	V	T	I	N	I	N	D	C	U	W	W	D	C							1C
																					W	V	T	I	N	I	N	D	C	U	W	W	D	S	N	I				1K

																					D	N	B	A	X																1A
																					R	N	B	A	Y	I	I	I													2E
																					T	E	W	A	Y	I	I	L													2E
																					I	E	W	A	Y	U	L	Q													1E

														Y	I	I	I	G	T	E	W	A	Y	I																	2E
																	B	I	E	W	A	Y	U																		1E
																	Q	V	E	W	C	W	S																		1I
												I	S	D	C	Y	I	W	V	E	W	D	N	B																	1A
																	B	G	T	E	W	V	R																		2I
												T	X	D	W	Y	I	W	X	N	I	N																			1H

																					A	I	N	B																	1B
																					K	I	N	D																	2A
																					W	A	X	L																	2A
																					B	A	X	W																	2A

FIGURE 81.

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
E	W	V	R	V	S	D	D	T	W																																
D	W	V	R	V	S	D	D	T	I																																
R	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	B																										
													G	U	H	K	I	R	X	Y	I	R	N	G	T	I	W	A	U	U	I	A	A	Q	C	W	N		
P	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	H	K	I	R	X	Y	I	R	N	G	T	I	W	A	U	U	I	A	A	Q	C	Q	D	T	Q
X	W	V	R	V	S	D	D	T	N	U	L	Q	D	U	H	K	I	R	X	Y	I	R	N	B																
															L	K	I	R	X	Y	I	R	N	G	T	I	W	A	U	U	I	A	A	Q	C	Q	D	T	V
																				K	I	R	N	G	S	T	L	B	D	D	T	I							
																				N	I	R	N	G	S	T	L	B	D	D	T	W							
																										T	W	A	U	U	I	A	A	Q	D			

FIGURE 82.

REF ID:A59439

47. Example of solution.—*a.* Suppose the two following cryptograms suspected of being in the same key are at hand:

No. 1

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

No. 2

Line																				
A	I	W	I	L	T	G	S	I	H	W	W	A	W	K	I	N	D	C	U	W	W	A	X	L	X
B	D	I	W	I	R	C	V	N	O	D	N	G	S	L	N	G	I	G	W	L	V	F	D	W	V
C	R	V	S	D	D	T	I	L	L	Q	D	U	K	D	W	S	S	H	X	S	E	N	C	Q	D
D	T	Q	G	T	E	U	D	V	Q	C	O	I	W	T	X	W	V	R	V	S	D	D	T	N	U
E	L	Q	D	U	H	K	I	R	X	Y	I	R	N	B	A	Y	I	I	I	G	T	E	W	A	Y
F	I	I	L	W	N	K	I	R	N	G	S	T	L	B	D	D	T	I	T	I	L	U	D	V	L
G	V	T	T	A	Q	A	V	N	I	I	U	H	Q	D	T	X	D	W	Y	I	W	X	N	I	N
H	B	D	R	W	S	C	L	P	L	W	A	H	N	T	L	Q	D	T	X	C	W	S	C	I	V
I	D	T	N	I	L	V	N	O	D	N	U	L	Q	D	U	B	G	T	E	W	V	R	V	S	D
J	D	T	W																						

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.

REF ID:A594339

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.
(6) From column 21: Y, K, N 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.

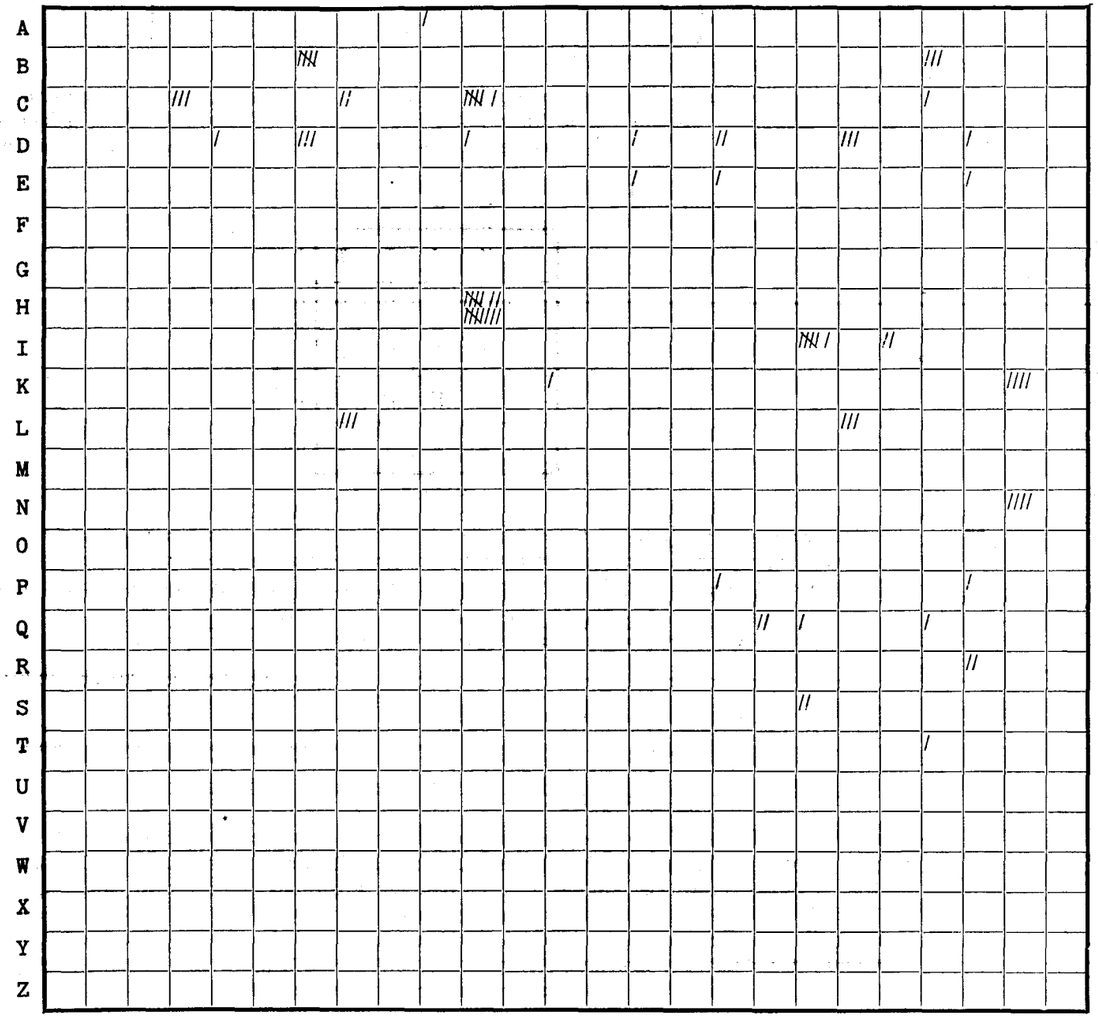
d. Tentative deduction (1) indicates that E, D, R, P, and X belong in the same column of the recomposition checkerboard. If correct, the complete set of 5 letters of one column is at hand. But tentative deduction (3) indicates that D and G belong in the same column and this would mean that the column 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.

REF ID:A59439

PREFIXES

(Letters in same column)



SUFFIXES

(Letters in same row)

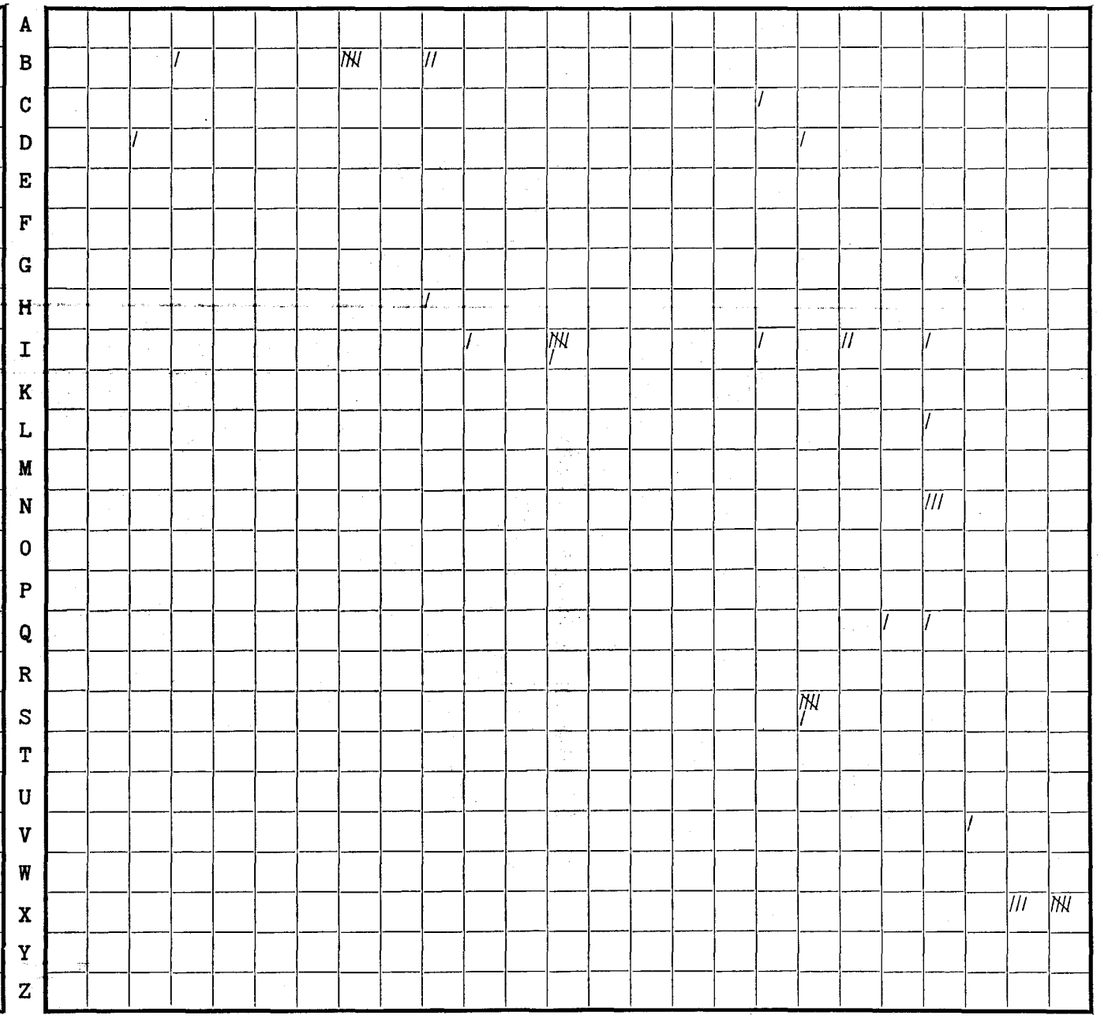


FIGURE 83.

REF ID:A59439

e. Suppose the data have been reduced to the following:

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

1st component \ 2d component	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	L	T	G	S	I	H	W	W	A	W	K	I	N	D
12	11	12	13	21	42	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.
.	T	W	T	D	N	?	W	?	S	W	L	B	N	B	T	H	.

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:

REF ID:A594339

1st component \ 2d component	1	2	3	4	5
1	N	U	T	Y	P
2	E	O	I	F	L
3	A	M	B	C	S
4	R	W	G	D	H
5	K	Q	V	X	Z

A
(Decomposition)

1st component \ 2d component	1	2	3	4	5
1	W	I	L	N	O
2	T	U	R	E	M
3	S	D	C	A	F
4	B	G	H	K	P
5	Q	V	X	Y	Z

B
(Recomposition)

FIGURE 85.

g. Speculating upon the key words used to produce the mixed sequences in these checkerboards, the trade-name N U T Y P E (a typewriter cleaning fluid) suggests looking at the label on the box containing a bottle of this chemical. It reads: N U T Y P E N O N – I N F L A M M A B L E T Y P E C L E A N S E R , prepared by W A L T E R G. G I E S. This yields the sequence for the decomposition checkerboard. The legend on the box also reads: W I L L N O T I N J U R E M O S T D E L I C A T E F A B R I C S A N D F U R N I T U R E, which yields the sequence for the recomposition checkerboard.

REF ID:A59439

h. The two cryptograms may now be deciphered directly from the checkerboards. The plain-texts are as follows:

No. 1

I S D C Y I W V E W D N B A X W S A V I W I L T K
12 31 32 33 54 12 11 52 24 11 32 14 41 34 53 11 31 34 52 12 11 12 13 21 44
I T I S R E P O R T E D T H A T T H E E N E M Y H

X L Y I W D A I N B D R W V R V S D D T N U L Q D
53 13 54 12 11 32 34 12 14 41 32 23 11 52 23 52 31 32 32 21 14 22 13 51 32
A S R E T I R E D T O A P O S I T I O N W E S T O

U B I N R K V L Q D N L Q D T V T I N I N D C U W
22 41 12 14 23 44 52 13 51 32 14 13 51 32 21 52 21 12 14 12 14 32 33 22 11
F N E W C H E S T E R S T O P O N E R E G I M E N

W D C V S O D C G S I S S G U H K I R X Y I R N G
11 32 33 52 31 15 32 33 42 31 12 31 31 42 22 43 44 12 23 53 54 12 23 14 42
T I S I N V I C I N I T Y O F C R O S S R O A D O

T I W A U U I A A Q C W N B I E W A Y U L Q S A V
21 12 11 34 22 22 12 34 34 51 33 11 14 41 12 24 11 34 54 22 13 51 31 34 52
N E T W O E I G H T A N D N O R T H W E S T T H E

N I I U H Q D T X W I T A V N N I N D C U W W C G
14 12 12 22 43 51 32 21 53 11 12 21 34 52 14 14 12 14 32 33 22 11 11 33 42
R E O F S T O P A N O T H E R R E G I M E N T C O

T L G W W A L W D S N S I L N N I R N G S T L B D
21 13 42 11 11 34 13 11 32 31 14 31 12 13 14 14 12 23 14 42 31 21 13 41 32
N C E N T R A T I N G N E A R R O A D J U N C T I

D T W A U U I A A Q D G T E L Q D T X D W Y I W X
32 21 11 34 22 22 12 34 34 51 32 42 21 24 13 51 32 21 53 32 11 54 12 11 53
O N T W O E I G H T F O U R S T O P B E P R E P A

N I N B D R Q T Q V E W C W S C L P I T L K I R X
14 12 14 41 32 23 51 21 51 52 24 11 33 11 31 33 13 45 12 21 13 44 12 23 53
R E D T O S U P P O R T A T T A C K O N C R O S S

Y I R N G T I W A U U I A A Q C Q D T Z I I W V T
54 12 23 14 42 21 12 11 34 22 22 12 34 34 51 33 51 32 21 55 12 12 11 52 21
R O A D O N E T W O E I G H T S T O P K E E P O N

I N I N D C U W W D S N I L V N O D L Q D T X W S
12 14 12 14 32 33 22 11 11 32 31 14 12 13 52 14 15 32 13 51 32 21 53 11 31
E R E G I M E N T I N R E S E R V E S T O P A T T

C L P W V R V S D D T N U L Q D U H K I R X Y I R
33 13 45 11 52 23 52 31 32 32 21 14 22 13 51 32 22 43 44 12 23 53 54 12 23
A C K P O S I T I O N W E S T O F C R O S S R O A

N G T I W A U U I A A Q C Q D T V
14 42 21 12 11 34 22 22 12 34 34 51 33 51 32 21 52
D O N E T W O E I G H T S T O P

REF ID:A594439

No. 2

I W I L T G S I H W W A W K I N D C U W W A X L X
12 11 12 13 21 42 31 12 43 11 11 34 11 44 12 14 32 33 22 11 11 34 53 13 53
E N E M Y I N F A N T R Y R E G I M E N T H A S B

D I W I R C V N O D N G S L N G I G W L V F D W V
32 12 11 12 23 33 52 14 15 32 14 42 31 13 14 42 12 42 11 13 52 35 32 11 52
E E N O B S E R V E D I N A D E F E N S I V E P O

R V S D D T I L L Q D U K D W S S H X S E N C Q D
23 52 31 32 32 21 12 13 13 51 32 22 44 32 11 31 31 43 53 31 24 14 33 51 32
S I T I O N E A S T O F G E T T Y S B U R G S T O

T Q G T E U D V Q C O I W T X W V R V S D D T N U
21 51 42 21 24 22 32 52 51 33 15 12 11 21 53 11 52 23 52 31 32 32 21 14 22
P Y O U W I L L T A K E U P A P O S I T I O N W E

L Q D U H K I R X Y I R N B A Y I I I G T E W A Y
13 51 32 22 43 44 12 23 53 54 12 23 14 41 34 54 12 12 12 42 21 24 11 34 54
S T O F C R O S S R O A D T H R E E F O U R T H R

I I L W N K I R N G S T L B D D T I T I L U D V L
12 12 13 11 14 44 12 23 14 42 31 21 13 41 32 32 21 12 21 12 13 22 32 52 13
E E A N D R O A D J U N C T I O N O N E M I L E S

V T T A Q A V N I I U H Q D T X D W Y I W X N I N
52 21 21 34 51 34 52 14 12 12 22 43 51 32 21 53 32 11 54 12 11 53 14 12 14
O U T H T H E R E O F S T O P B E P R E P A R E D

B D R W S C L P L W A H N T L Q D T X C W S C I V
41 32 23 11 31 33 13 45 13 11 34 13 24 11 53 51 32 21 53 33 11 31 33 12 52
T O A T T A C K A T D A W N S T O P B A T T A L I

D T N I L V N O D N U L Q D U B G T E W V R V S D
32 21 14 12 13 52 14 15 32 14 22 13 51 32 22 41 42 21 24 11 52 23 52 31 32
O N R E S E R V E W E S T O F Y O U R P O S I T I

D T W
32 21 11
O N

48. Special solution.—*a.* The preceding example of solution constitutes the general solution for this system, since no special conditions are prerequisite 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. 1. G C O D M G C E G B W I L W G M O N G B S X O P C N G E S F L N I
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 D I W I R C V N O

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

No. 1. Q P F U A Q M S A Z P H Z G N L M S O W O V X G X H M G A T A K N
No. 2. H X S E N C Q D T Q G T E U D V Q C O I W T X W V R V S D D T N U

No. 1. O M S W C U S H Q L S S T M S U W N N E H U A S U W N T G E L S S
No. 2. L Q D U H K I R X Y I R N B A Y I I I G T E W A Y I I L W N K I R

No. 1. T M O V C V A T A E A C O G N L O P H V S V S U N W T W F M S A X
No. 2. N G S T L B D D T I T I L U D V L V T Q A Q A V N I I U H Q D T X

FIRST COMPONENTS OF CRYPTOGRAM No. 1

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			T_2I_1 T_2N_1		T_2I_1				S_2S_1	T_2N_1						N_2C_1		W_2A_1 S_2C_1	D_2D_1				T_2X_1 W_2V_1		T_2Q_1
B											K_2D_1							W_2A_1				W_2A_1			
C					I_2H_1		G_2W_1 D_2W_1						D_2I_1	I_2L_1						H_2K_1	L_2B_1	N_2I_1			
D							N_2T_1					T_2G_1	N_2U_1									F_2D_1			
E	I_2T_1						H_2W_1 C_2W_1	G_2T_1			N_2K_1							I_2R_1	L_2N_1						
F											C_2V_1	L_2V_1 H_2Q_1 L_2Q_1								S_2E_1					
G	S_2D_1 W_2S_1	W_2W_1 W_2W_1	W_2I_1 S_2I_1		W_2I_1 S_2L_1 W_2N_1	W_2L_1 T_2L_1						N_2D_1	U_2D_1					W_2A_1		W_2Y_1			W_2V_1 W_2X_1		
H												R_2V_1				R_2X_1				T_2E_1	T_2Q_1				T_2E_1
I																S_2H_1						O_2D_1			
K													N_2U_1		B_2G_1										
L							D_2W_1					V_2Q_1	V_2N_1 G_2I_1	V_2L_1				Y_2I_1 K_2I_1				K_2I_1			
M				V_2F_1			G_2S_1 G_2S_1 V_2S_1							D_2C_1 G_2S_1				Q_2D_1 Q_2C_1 B_2A_1 B_2D_1	V_2D_1							
N			I_2G_1		I_2G_1		U_2W_1 I_2W_1		N_2O_1		D_2V_1	I_2V_1	I_2I_1	U_2L_1					I_2L_1			N_2I_1		L_2P_1	
O				L_2T_1			L_2U_1				L_2V_1	L_2Q_1	C_2U_1		L_2X_1 L_2V_1						W_2T_1 S_2T_1	O_2I_1			
P			X_2D_1	H_2N_1		X_2S_1		G_2T_1 V_2T_1																	
Q											X_2Y_1	C_2Q_1			H_2X_1										
R																									
S	D_2T_1					R_2C_1	R_2W_1 L_2W_1	I_2R_1					C_2L_1 C_2I_1	C_2O_1	A_2H_1			I_2R_1 D_2R_1	R_2N_1	A_2Y_1 A_2V_1	A_2Q_1	D_2U_1	A_2X_1		
T	D_2T_1						L_2W_1				N_2G_1	N_2G_1 N_2B_1		L_2L_1								I_2U_1			
U	E_2N_1 E_2W_1												V_2N_1					K_2I_1				Y_2I_1			
V	B_2D_1		T_2L_1															Q_2A_1					T_2X_1		
W		U_2K_1	U_2H_1	D_2N_1		U_2H_1	I_2N_1 D_2W_1 I_2W_1		A_2W_1	U_2B_1									D_2N_1 I_2I_1 I_2N_1						
X			X_2D_1		X_2C_1		X_2W_1	V_2R_1																	
Y																		P_2L_1							
Z							E_2U_1								Q_2G_1										

FIGURE 86.

REF ID:A594139

REF ID:A59439

No. 1. C G U W G X N W T M S S G A S N Y S G S P D G F M S A X E G A S N
No. 2. D W Y I W X N I N B D R W S C L P L W A H N T L Q D T X C W S C I

No. 1. M T A C W O L N I W D N O M S W K P H U A X H M G A T A C
No. 2. V D T N I L V N O D N U L Q D U B G T E W V R V S D D T I

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_c, the second determines the second half of G_c. Therefore, for cryptanalytic purposes, the actual, but unknown, numerical components may be represented by the symbols G_1 and G_2, 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----	C_1C_2	G_1G_2	C_1C_2	O_1O_2	D_1D_2	M_1M_2	G_1G_2	C_1C_2	E_1E_2
No. 2----------	I	W	I	L	T	G	S	I	H
Components----	I_1I_2	W_1W_2	I_1I_2	L_1L_2	T_1T_2	G_1G_2	S_1S_2	I_1I_2	H_1H_2

Now from the method of encipherment it is clear that C_2G_1 and I_2W_1 represent the same plain-text letter, since both messages are assumed to contain identical plain texts. That is, C_2G_1 of cryptogram No. $1 = I_2W_1$ of cryptogram No. 2. Likewise $G_2C_1 = W_2I_2$; $C_2O_1 = I_2L_1$; $O_2D_1 = L_2T_1$; 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, A_2C_1 of cryptogram No. $1 = T_2I_1$ of cryptogram No. 2. This means that I_1 and N_1 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 $C_2G_1 = G_2W_1$ and D_2W_1, which means that G and D must be in the same column in that checkerboard. Again, $A_2S_1 = W_2A_1 = S_2C_1$; 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:

In same row:

(1) I, N, L, W
(2) A, C, S, D
(3) Q, V, X, Y
(4) B, G

In same column:

(1)	(2)	(3)	(4)	(5)
T	H	Y	Q	U
W	C	K	B	I
S	L			D
	R			G
				V

324155—41——11

REF ID:A594339

An attempt is now made to bring together these results to reconstruct the recomposition checkerboard for message No. 2. This yields the following:

1st Component \ 2d Component					
	U	K?	R	T	K?
	I	N	L	W	K?
	D	A	C	S	K?
	G	K?	H	B	K?
	V	Y?	X	Q	Y?

FIGURE 85a.

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.

f. 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 47*f*.

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

No. 1

1st Component \ 2d Component	1	2	3	4	5
1	R	E	F	L	C
2	T	I	N	G	P
3	O	A	B	D	H
4	K	M	Q	S	U
5	V	W	X	Y	Z

A
(Decomposition)

1st Component \ 2d Component	1	2	3	4	5
1	W	A	S	H	I
2	N	G	T	O	D
3	C	B	E	F	K
4	L	M	P	Q	R
5	U	V	X	Y	Z

B
(Recomposition)

G C O D M G C E G B W I L W G M O N G B S X O P C N G E S F
22 31 24 25 42 22 31 33 22 32 11 15 41 11 22 42 24 21 22 32 13 53 24 43 31 21 22 33 13 34
E N E M Y I N F A N T R Y R E G I M E N T H A S B E E N O B

L N I W T M G E T L N C G F M D W G X H M G A T A C T O M S
41 21 15 11 23 42 22 33 23 41 21 31 22 34 42 25 11 22 53 14 42 22 12 23 12 31 23 24 42 13
S E R V E D I N A D E F E N S I V E P O S I T I O N E A S T

W B L G A I Q P F U A Q M S A Z P H Z G N L M S O W O V X G
11 32 41 22 12 15 44 43 34 51 12 44 42 13 12 55 43 14 55 22 21 41 42 13 24 11 24 52 53 22
O F G E T T Y S B U R G S T O P Y O U W I L L T A K E U P A

X H M G A T A K N O M S W C U S H Q L S S T M S U W N N E H
53 14 42 21 12 23 12 35 21 24 42 13 11 31 51 13 14 44 41 13 13 23 42 13 51 11 21 21 33 14
P O S I T I O N W E S T O F C R O S S R O A D T H R E E F O

REF ID:A59439

U A S U W N T G E L S S T M O V C V A T A E A C O G N L O P
51 12 13 51 11 21 23 22 33 41 13 13 23 42 24 52 31 52 12 23 12 33 12 31 24 22 21 41 24 43
U R T H R E E A N D R O A D J U N C T I O N O N E M I L E S

H V S V S U N W T W F M S A X C G U W G X N W T M S S G A S
14 52 13 52 13 51 21 11 23 11 34 42 13 12 53 31 22 51 11 22 53 21 11 23 42 13 13 22 12 13
O U T H T H E R E O F S T O P B E P R E P A R E D T O A T T

N Y S G S P D G F M S A X E G A S N M T A C W O L N I W D N
21 54 13 22 13 43 25 22 34 42 13 12 53 33 22 12 13 21 42 23 12 31 11 24 41 21 15 11 25 21
A C K A T D A W N S T O P B A T T A L I O N R E S E R V E W

O M S W K P H U A X H M G A T A C
24 42 13 11 35 43 14 51 12 53 14 42 22 12 23 12 31
E S T O F Y O U R P O S I T I O N

No. 2

1st Component \ 2d Component	1	2	3	4	5
1	N	U	T	Y	P
2	E	O	I	F	L
3	A	M	B	C	S
4	R	W	G	D	H
5	K	Q	V	X	Z

1st Component \ 2d Component	1	2	3	4	5
1	W	I	L	N	O
2	T	U	R	E	M
3	S	D	C	A	F
4	B	G	H	K	P
5	Q	V	X	Y	Z

W I L T G S I H W W A W K I N D C U W W A X L X D I W I R C
11 12 13 21 42 31 12 43 11 11 34 11 44 12 14 32 33 22 11 11 34 53 13 53 32 12 11 12 23 33
E N E M Y I N F A N T R Y R E G I M E N T H A S B E E N O B

V N O D N G S L N G I G W L V F D W V R V S D D T I L L Q D
52 14 15 32 14 42 31 13 14 42 12 42 11 13 52 35 32 11 52 23 52 31 32 32 21 12 13 13 51 32
S E R V E D I N A D E F E N S I V E P O S I T I O N E A S T

U K D W S S H X S E N C Q D T Q C T E U D V Q C O I W T X W
22 44 32 11 41 31 43 53 31 24 14 33 51 32 21 51 42 21 24 22 32 52 51 33 15 12 11 21 53 11
O F G E T T Y S B U R G S T O P Y O U W I L L T A K E U P A

V R V S D D T N U L Q D U H K I R X Y I R N B A Y I I I G T
52 23 52 31 32 32 21 14 22 13 51 32 22 43 44 12 23 53 54 12 23 14 41 34 54 12 12 12 42 21
P O S I T I O N W E S T O F C R O S S R O A D T H R E E F O

E W A Y I I L W N K I R N G S T L B D D T I T I L U D V L V
24 11 34 54 12 12 13 11 14 44 12 23 14 42 31 21 13 41 32 32 21 12 21 12 13 22 32 52 13 52
U R T H R E E A N D R O A D J U N C T I O N O N E M I L E S

T T A Q A V N I I U H Q D T X D W Y I W X N I N B D R W S C
21 21 34 51 34 52 14 12 12 22 43 51 32 21 53 32 11 54 12 11 53 14 12 14 41 32 23 11 31 33
O U T H T H E R E O F S T O P B E P R E P A R E D T O A T T

L P L W A H N T L Q D T X C W S C I V D T N I L V N O D N U
13 45 13 11 34 13 24 11 53 51 32 21 53 33 11 31 33 12 52 32 21 14 12 13 52 14 15 33 14 22
A C K A T D A W N S T O P B A T T A L I O N R E S E R V E W

L Q D U B G T E W V R V S D D T W
13 51 32 22 41 42 21 24 11 52 23 52 31 32 32 21 11
E S T O F Y O U R P O S I T I O N

REF ID:A594319

h. It is seen that the principles elucidated permit of solving this fairly good cipher system without recourse to frequency studies and detailed, difficult analytical research. What can be done with complete messages of identical texts will give the student a clue to what might be done when fairly lengthy sequences of identical plain texts (but not complete messages) are available for study. Messages with similar beginnings, or similar endings will afford data for such reconstruction.

49. 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 preagreement 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:

1st component \ 2d component	1	2	3	4	5
1	M	A	N	U	F
2	C	T	R	I	G
3	B	D	E	H	K
4	L	O	P	Q	S
5	V	W	X	Y	Z

FIGURE 87.

O	N	E	P	L		A	N	E	R	E		P	O	R	T	E		D	L	O	S	T		A	T	S	E	A
4	1	3	4	4		1	1	3	2	3		4	4	2	2	3		3	4	4	4	2		1	2	4	3	1
2	3	3	3	1		2	3	3	3	3		3	2	3	2	3		2	1	2	5	2		2	2	5	3	2

Recombinations are effected horizontally within the periods, by joining components in pairs, the first period yielding the pairs 41, 34, 42, 33, 31. These pairs are then replaced by letters from the original checkerboard, yielding the following:

O	N	E	P	L		A	N	E	R	E		P	O	R	T	E		D	L	O	S	T		A	T	S	E	A
4	1	3	4	4		1	1	3	2	3		4	4	2	2	3		3	4	4	4	2		1	2	4	3	1
2	3	3	3	1		2	3	3	3	3		3	2	3	2	3		2	1	2	5	2		2	2	5	3	2
L	H	O	E	B		M	D	D	E	E		Q	T	E	R	R		H	Q	T	A	W		A	P	A	G	D

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.

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 while elaborating upon to make its cryptanalytic significance perfectly clear. Note what happens when an even period is employed:

O	N	E	P	L	A		N	E	R	E	P	O		R	T	E	D	L	O	
4	1	3	4	4	1		1	3	2	3	4	4		2	2	3	3	4	4	. . .
2	3	3	3	1	2		3	3	3	3	3	2		3	2	3	2	1	2	. . .
L	H	L	R	E	A		N	R	Q	E	E	D		T	E	Q	D	D	A	. . .

REF ID:A59439

Now if each 6-letter cipher group is split in the middle into two sections 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, $ON_p = LR_c$; $EP_p = HE_{c1}$ 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.[3] 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.

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

O N E P L A N E
L R H E L A N E

FIGURE 88.

b. But the latter statement is no longer true in the case of odd periods. Note, in the example under paragraph 49*a*, that the cipher equivalent of the first plain-text digraph of the first group, ON, is composed of the initial and final components of the letter L_c, the final component of the letter O_c, and the initial component of the letter L_c. 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 ON_p.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
12 34 5	12 34 5	12 34 5	12 34 5	12 34 5	12 34 5	12 34 5
ON EP L	ON TH E	ON CR U	PR ON G	CO NT I	PO NG I	AT IO N
41 34 4	41 23 3	41 22 1	42 41 2	24 12 2	44 12 2	12 24 1
23 33 1	23 24 3	23 13 4	33 23 5	12 32 4	32 35 4	22 42 3
LH OE B	LR DD P	LT AB H	OL RD K	IA CR I	QA RR Y	AI AI R

c. The foregoing examples fall into two classes. In the first, where the O of ON_p falls in an odd position in the period, the first letter of the trigraphic cipher equivalent must be an L_c, 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_c may combine with 5×5 or 25 pairs of letters to form the second and third letters of the 3-letter equivalent of ON_p. In the other class, where the O of ON_p 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 R_c. Therefore, R_c may combine with 5×5 or 25 pairs of letters to form the first and second letters of the 3-letter equivalent of ON_p in this position in the period. Hence, ON_p 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.

[3] An example of the solution of a cryptogram of this type was given in *Military Cryptanalysis, Part I*, sec. IX.

REF ID:A594439

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 49 *a:*

O	N	E	P	L		A	N	E	R	E		P	O	R	T	E		D	L	O	S	T		A	T	S	E	A
4	1	3	4	4		1	1	3	2	3		4	4	2	2	3		3	4	4	4	2		1	2	4	3	1
2	3	3	3	1		2	3	3	3	3		3	2	3	2	3		2	1	2	5	2		2	2	5	3	2
L	H	O	E	B		M	D	D	E	E		Q	T	E	R	R		H	Q	T	A	W		A	P	A	G	D

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 bifid components that come from the first and second positions in the upper row of components in the period. These components are not known, but whatever they are the first of them is the first component of L, the second of them is the second component of L. Therefore, just as in paragraph 48*c*, the actual but unknown, numerical components may be represented by the symbols L_1 and L_2, 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:

Cipher	L	H	O	E	B
Components	L_1	L_2	H_1	H_2	O_1
	O_2	E_1	E_2	B_1	B_2

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

Plain text	O	N	E	P	L
Components	4	1	3	4	4
	2	3	3	3	1
Cipher	L	H	O	E	B

O	N	E	P	L
L_1	L_2	H_1	H_2	O_1
O_2	E_1	E_2	B_1	B_2
L	H	O	E	B

By comparing the two diagrams it becomes obvious that L_1, H_2, and O_1 all represent the coordinate 4; H_1, E_1, E_2, and B_1 all represent the coordinate 3, and so on. If this equivalency 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 monoalphabetic encipherment. But this equivalency is not known in the case of a cryptogram that is to be solved; basically the problem is to establish the equivalency.

e. It is obvious that the vertical pair of components $\begin{matrix} L_1 \\ O_2 \end{matrix}$ represents O_p, the vertical pair $\begin{matrix} L_2 \\ E_1 \end{matrix}$ represents N_p, and so on. The complete example therefore becomes:

Plain	O N E P L	A N E R E	P O R T E	D L O S T	A T S E A
Components	L_1 L_2 H_1 H_2 O_1	M_1 M_2 D_1 D_2 D_1	Q_1 Q_2 T_1 T_2 E_1	H_1 H_2 Q_1 Q_2 T_1	A_1 A_2 P_1 P_2 A_1
	O_2 E_1 E_2 B_1 B_2	D_2 E_1 E_2 E_1 E_2	E_2 R_1 R_2 R_1 R_2	T_2 A_1 A_2 W_1 W_2	A_2 G_1 G_2 D_1 D_2
Cipher	L H O E B	M D D E E	Q T E R R	H Q T A W	A P A G D

f. Note that a plain-text letter in an odd position in the period has its components in the order $\Theta_1\Theta_2$; in an even position in the period the components of a plain-text letter are in the order $\Theta_2\Theta_1$. For example, note the O_p in the first period $\left(=\begin{matrix} L_1 \\ O_2 \end{matrix}\right)$ and the O_p in the third period $\left(=\begin{matrix} Q_2 \\ R_1 \end{matrix}\right)$. This distinction must be retained since the component indicators for rows and columns are not interchangeable in this system. From this it follows that the vertical pairs of components represent-

REF ID:A59439

ing a given plain-text letter are of two classes: $\Theta_1\Theta_2$ and $\Theta_2\Theta_1$, and the two must be kept separate in cryptanalysis.

g. Now consider the equivalent of O_p in the first period. It is composed of $\begin{matrix} L_1 \cdot \\ O_2 \end{matrix}$ This is only one of a number of equivalents for O_p in an odd position in the period. The row of the substitution checkerboard indicated by L_1 may be represented by 4 other components, since that row contains 5 letters. Therefore the upper component of the $\begin{matrix} \Theta_1 \\ \Theta_2 \end{matrix}$ equivalent of O_p may be any one of 5 letters. The same is true of the lower component. Hence, O_p in an odd position in the period may be represented by any one of $5\times5=25$ combinations of vertical components in the sequence $\Theta_1\rightarrow\Theta_2$. O_p 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, $\Theta_2\rightarrow\Theta_1$. Thus, disregarding the position in the period, this system may be described as a monoalphabetic 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) occupied by a cipher letter in the period is obvious after the length of the period has been established, a proper segregation of the cipher letters will permit of handling the cipher letters in the two classes referred to above, in which case one has to deal with only 25 variants for each plain-text letter. Obviously, the 25 variants are related to one another by virtue of their having been produced from a single enciphering matrix of but 25 letters. This relationship can be used to good advantage in reconstructing the matrix in the course of the solution and the relationship will be discussed in its proper place.

h. Now if the foregoing encipherment is studied intently several important phenomena may be observed. Note, for instance, how many times either the Θ_1 or the Θ_2 component coincides with the plain-text letter of which it is a part. In the very first period the O_p has an O_2 under it; in the same period the E_p has an E_2 under it. The same phenomenon is observed in columns 3 and 5 of the second period, in column 3 of the third period, and in column 1 of the fifth period. In column 5 of the third, fourth, and fifth periods the Θ_1 components coincide with the respective plain-text letters involved. There are, in this short example, 9 cases of this sort, giving rise to instances of what seems to be a sort of self-encipherment of plain-text letters. How does this come about? And is it an accident that all these cases involve plain-text letters in odd positions in the periods?

i. If the periods in the foregoing example in subparagraph *e* are studied closely, the following observations 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: In these cases the two 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 first 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 *h* 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 Θ_1 component standing directly under the plain-text letter is identical with the latter in each case.

j. But 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 Θ_2 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.

k. Since the probability of the occurrence of the event in question is 1/5 for Θ_1 components and 1/5 for Θ_2 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.

l. As regards the plain-text letters in even positions, analysis will show why only in a 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 this order will give wholly different letters, except in those 5 cases in which both components are identical. (The letters involved are those which occupy the 5 cells along the diagonal from the upper left-hand corner to the lower right-hand corner of the checkerboard.) Now in every case of a letter in an odd position in a period the two vertical components are in the $\Theta_1\Theta_2$ order, corresponding to the order in which they are normally taken in finding letter equivalents in the checkerboard. But in every case of a letter in an even position in a period, the two vertical components are in the order $\Theta_2\Theta_1$, which is a reversal of the normal order. It has been seen that in the case of letters in odd positions in the periods the probability that one of the components will coincide with the plain-text letter is 40 percent. The reasoning which led to this determination in the case of the odd letters is exactly the same as that in the case of letters in even positions, except that in the final recomposition process, since the components in the even positions are in the $\Theta_2\Theta_1$ order, which is the reverse of the normal order, identity between one of the components and the plain-text letter can occur in only 1/5 of the 40 = 8 percent of the cases. It may be said then that in this system 48 percent of all the letters of the plain text will be "self-enciphered" and represented by one or the other of the two components; in the case of the letters in odd positions, the amount is 40 percent, in the case of letters in even positions, it is 8 percent.

m. 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 *e*? 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_c'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 cipher letter is repeated at an interval of 2, for this will bring about the superimposition of a Θ_1 and Θ_2 with the same principal 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.

REF ID:A59439

n. All of the foregoing phenomena will be useful when the solution of an example is undertaken. But before coming to such an example it is necessary to explain how to ascertain the period of a cryptogram to be solved.

51. Ascertaining the length of the period.—*a*. There are several methods available for ascertaining the length of the period. The simplest, of course, is to look for repetitions of the ordinary sort. If the period is a short one, say 3, 5, 7 letters, and if the message is fairly long, the chances are good that a polygraph which occurs several times within the message will fall in homologous positions within two different periods and therefore will be identically enciphered both times. There will not be many such repetitions, it is true, but factoring the intervals between such as do occur will at least give some clue, if it will not actually disclose the length of the period. For example, suppose that a 7-letter repetition is found, the two occurrences being separated by an interval of 119. The factors of 119 are 7 and 17; the latter is unlikely to be the length of the period, the former, quite likely.

b. If a polygraph is repeated but its two occurrences do not fall in homologous positions in two periods, there will still be manifestations of the presence of repetition but the repeated letters will be separated by one or more intervals in the periods involved. The number of repeated letters will be a function of the length of the polygraph and 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. Note what happens in the following example:

S	E	N	D	T	H	R		E	E	M	E	N	D	O		W	N	T	O	E	N	D		O	F	E	N	D	I	C		O	T	T	R	O	A	D
4	3	1	3	2	3	2		3	3	1	3	1	3	4		5	1	2	4	3	1	3		4	1	3	1	3	2	2		4	2	2	2	4	1	3
5	3	3	2	2	4	3		3	3	1	3	3	2	2		2	3	2	2	3	3	2		2	5	3	3	2	4	1		2	2	2	3	2	2	2
P	N	R	G	E	T	P		E	N	N	P	B	E	T		V	I	B	D	D	R	D		L	B	D	T	X	D	L		O	T	L	D	T	D	T

CRYPTOGRAM

P N R G E T P E N N P B E T V I B D D R D L B D T X D L O T L D T D T

Here the plain text contains the trigraph END 4 times. The END_p in the first period gives rise to the cipher letters $\overset{1\ 2\ 3\ 4\ 5\ 6\ 7}{.\ N\ .\ .\ E\ .\ .}$; in the second period this trigraph also produces $\overset{1\ 2\ 3\ 4\ 5\ 6\ 7}{.\ .\ N\ .\ .\ E\ .}$ The interval between the N_c and the E_c is 3 in both cases. Two times this interval plus one gives the length of the period. In this case the initial letter of the repeated trigraph falls in an even position in the period in both occurrences. The END_p in the third period gives rise to the cipher letters $\overset{1\ 2\ 3\ 4\ 5\ 6\ 7}{.\ .\ B\ .\ .\ .\ D}$; in the fourth period it also produces $\overset{1\ 2\ 3\ 4\ 5\ 6\ 7}{.\ B\ .\ .\ .\ D\ .}$ The interval between the B_c and the D_c is 4 in both cases. Two times this interval minus one gives the length of the period. In this case the initial letter of the repeated trigraph falls in an odd position in the period in both occurrences.

c. The foregoing properties of repetitions in this system afford a means of ascertaining the length of the period in an unknown example. First, it is evident that a repeated trigraph in the plain text produces two different pairs of cipher equivalents according to whether the initial letter of the trigraph occurs in an odd or an even position in the period. The two letters constituting the repetition in the cryptogram will not be sequent but will be separated by an interval of 1, 2, 3, . . . letters depending upon the length of the period. This interval, however, is half of the period plus or minus one.[4] Conversely, if in a cryptogram there are repetitions of pairs of

[4] The student must remember that the text is here concerned only with cases in which the period is odd. In the case of even periods the interval separating the 2 letters is always exactly half of the length of the period.

REF ID:A594319

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: $p(\text{length of period}) = 2x \pm 1$.

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_c, B_c, C_c, . . . Z_c 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_c one looks to see if there are 2 or more cases in which the same letter appears 2, 3, 4, . . . 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_c, C_c, and so on. The tallies representing coincidences may be amalgamated for all the letters A, B, C, . . . 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, . . . Z, and noting the coincidences occurring 1, 2, 3, . . . up to say 10 letters beyond the juxtaposed letters. For example, beginning with A_c, 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 . . . Z's. The record made of the coincidences may consist merely of a tally stroke written under the intervals 1, 2, 3, . . . 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 Θ_1 and Θ_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.

REF ID:A59439

SECOND (θ_2) COMPONENTS

FIRST (θ_1) 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	
A		///	/	//			//	////		/			A			/	/			/				//	///		A 21
B	/		/	//	/	/				/	/	/	B		/			//				/		/	/		B 15
C			/	///			/			///	/		C			/	//							/	/		C 14
D		////			//		/		/		////		D	/			/	/	/	/				///	/	/	D 23
E	/	/	/				/	//			//		E	/		/		/				/			/	/	E 14
F									//	/			F	/						//							F 6
G						//							G		///			/					//				G 8
H			/	/			/		/		//		H									/	/				H 8
I	/	/		/	/						//		I	/	/												I 8
K		///	/	/	/	//							K			/		/	/					/		/	K 13
L	//								/	/	/		L		///		//	///	/			/			/			L 17
M		/			//	/	/	/		/	////		M	/		/	/						/	/			M 16
	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	
N	/			/					/	////	///		N		/		/	/				/					N 14
O												/	O			/		//	/				/		/		O 7
P	/		/							/		/	P					/					//				P 7
Q		//	/			/							Q		//								//				Q 8
R										//		/	R					//	/					/		/	R 8
S		///	/										S		/	//		/	/								S 9
T	//				/		/	/					T							////				/			T 11
U	//		//	/	////		/	//					U								/	/	/			U 15
V	/				/			/	/		///		V							/							V 8
W								/					W					//		/				///			W 7
X		/	//	/						/	/		X		///	/		////	/					/			X 16
Y			//	/						//		////	Y		//		/							//			Y 14
Z					////					/			Z		/					//					/			Z 9
	A 12	B 19	C 15	D 14	E 17	F 7	G 9	H 12	I 7	K 18	L 26	M 8		N 5	O 18	P 9	Q 9	R 23	S 7	T 13	U 0	V 5	W 13	X 17	Y 9	Z 4	(Total=296)	296

FIGURE 90.

FIRST (θ_1) COMPONENTS

SECOND (θ_2) 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	
A	/	/	//	///		/				//	//	//	A		/					/		/	/	//			A 20
B		/							/		///		B		//	/								/	/		B 10
C							/				//		C		/		/	/		///					/		C 10
D	///	//		//		//	/	///		//		/	D	//													D 18
E	/	/											E				/	///		/		/	/				E 9
F					//						/		F										/	/			F 5
G					/			/		/			G				//	/					.				G 6
H										//			H				/		/								H 4
I	/				/						/		I		/										/		I 5
K	/			/	//		/						K		/								/				K 7
L		/			//								L			///	/	//		/			/	/			L 12
M								/		/			M				//	///		///			/			/	M 12
	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	
N		//	//	/	/	/	/	/	/		/		N					/						/			N 13
O							//						O									/					O 3
P			/					/					P		/			/									P 4
Q					/	/					///		Q					/							/		Q 7
R													R					/		//			//				R 5
S											/		S		///			/						/	/		S 7
T		/	/							//			T					//	///								T 9
U	/		/	//	/	/			/		//	/	U		//	//								/			U 15
V				/	/								V	/										//	/		V 6
W										/			W					/					/				W 3
X	/									//			X	/				/	//				//		/		X 10
Y		///	///					/			/		Y	/	/			/									Y 13
Z			/				/		/		////		Z		/												Z 8
	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	221
	9	14	11	10	12	6	7	8	4	13	21	4		5	14	6	8	20	6	11	0	3	9	10	7	3	(Total=221)

FIGURE 91.

324155—41 (Face p. 167) No. 2

REF ID:A594439

REF ID:A59439

52. Illustration of solution.—*a.* With the foregoing principles in mind, the following cryptogram will be studied:

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 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 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
F D D T E B A M F D E T E N A K G D F O Q D U B N D C L Y D V W B A X
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 R 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 E 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 X K U P S E M C T N K T K E B O E E P G V Q T G W E R 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 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 Θ_1 and Θ_2 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 Θ_1 and Θ_2 components according to a period of 7 (see fig. 89) and complete distributions of $\Theta_1\Theta_2$ and $\Theta_2\Theta_1$ 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.

1	2	3	4	5
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 C E P Q L
K_1 K_2 Z_1 Z_2 F_1 F_2 B_1	Y_1 Y_2 Y_1 Y_2 M_1 M_2 O_1	B_1 B_2 L_1 L_2 Z_1 Z_2 D_1	B_1 B_2 L_1 L_2 P_1 P_2 K_1	U_1 U_2 C_1 C_2 C_1 C_2 E_1
B_2 E_1 E_2 I_1 I_2 L_1 L_2	O_2 C_1 C_2 B_1 B_2 R_1 R_2	D_2 O_1 O_2 T_1 T_2 G_1 G_2	K_2 Y_1 Y_2 W_1 W_2 C_1 C_2	E_2 P_1 P_2 Q_1 Q_2 L_1 L_2

6	7	8	9	10
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
A_1 A_2 M_1 M_2 E_1 E_2 Y_1	X_1 X_2 W_1 W_2 H_1 H_2 L_1	Y_1 Y_2 D_1 D_2 R_1 R_2 W_1	I_1 I_2 Z_1 Z_2 E_1 E_2 B_1	Y_1 Y_2 E_1 E_2 S_1 S_2 O_1
Y_2 L_1 L_2 Z_1 Z_2 Q_1 Q_2	L_2 R_1 R_2 W_1 W_2 Q_1 Q_2	W_2 B_1 B_2 M_1 M_2 T_1 T_2	B_2 E_1 E_2 L_1 L_2 A_1 A_2	O_2 B_1 B_2 R_1 R_2 Y_1 Y_2

11	12	13	14	15
Q V B B L Y 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
Q_1 Q_2 V_1 V_2 B_1 B_2 B_1	N_1 N_2 A_1 A_2 B_1 B_2 Q_1	Y_1 Y_2 M_1 M_2 Q_1 Q_2 D_1	N_1 N_2 A_1 A_2 C_1 C_2 O_1	R_1 R_2 G_1 G_2 A_1 A_2 S_1
B_2 L_1 L_2 Y_1 Y_2 X_1 X_2	Q_2 B_1 B_2 D_1 D_2 O_1 O_2	D_2 L_1 L_2 W_1 W_2 L_1 L_2	O_2 X_1 X_2 C_1 C_2 R_1 R_2	S_2 W_1 W_2 Q_1 Q_2 B_1 B_2

16	17	18	19	20
F D D T E B A	M F D E T E N	A K G D F O Q	D U B N D C L	Y D V W B A X
F_1 F_2 D_1 D_2 D_1 D_2 T_1	M_1 M_2 F_1 F_2 D_1 D_2 E_1	A_1 A_2 K_1 K_2 G_1 G_2 D_1	D_1 D_2 U_1 U_2 B_1 B_2 N_1	Y_1 Y_2 D_1 D_2 V_1 V_2 W_1
T_2 E_1 E_2 B_1 B_2 A_1 A_2	E_2 T_1 T_2 E_1 E_2 N_1 N_2	D_2 F_1 F_2 O_1 O_2 Q_1 Q_2	N_2 D_1 D_2 C_1 C_2 L_1 L_2	W_2 B_1 B_2 A_1 A_2 X_1 X_2

FIGURE 89.

REF ID:A594382

21

C	A	U	G	G	X	O
C_1	C_2	A_1	A_2	U_1	U_2	G_1
G_2	G_1	G_2	X_1	X_2	O_1	O_2

22

A	R	T	X	X	T	S
A_1	A_2	R_1	R_2	T_1	T_2	X_1
X_2	X_1	X_2	T_1	T_2	S_1	S_2

23

D	A	Y	X	H	K	O
D_1	D_2	A_1	A_2	Y_1	Y_2	X_1
X_2	H_1	H_2	K_1	K_2	O_1	O_2

24

L	S	X	A	B	R	K
L_1	L_2	S_1	S_2	X_1	X_2	A_1
A_2	B_1	B_2	R_1	R_2	K_1	K_2

25

R	P	U	Z	W	H	O
R_1	R_2	P_1	P_2	U_1	U_2	Z_1
Z_2	W_1	W_2	H_1	H_2	O_1	O_2

26

M	T	D	H	T	S	G
M_1	M_2	T_1	T_2	D_1	D_2	H_1
H_2	T_1	T_2	S_1	S_2	G_1	G_2

27

M	L	S	L	Q	P	O
M_1	M_2	L_1	L_2	S_1	S_2	L_1
L_2	Q_1	Q_2	P_1	P_2	O_1	O_2

28

U	N	H	C	I	C	K
U_1	U_2	N_1	N_2	H_1	H_2	C_1
C_2	I_1	I_2	C_1	C_2	K_1	K_2

29

K	A	Q	B	D	O	F
K_1	K_2	A_1	A_2	Q_1	Q_2	B_1
B_2	D_1	D_2	O_1	O_2	F_1	F_2

30

L	E	K	A	P	R	G
L_1	L_2	E_1	E_2	K_1	K_2	A_1
A_2	P_1	P_2	R_1	R_2	G_1	G_2

31

S	X	U	P	O	W	A
S_1	S_2	X_1	X_2	U_1	U_2	P_1
P_2	O_1	O_2	W_1	W_2	A_1	A_2

32

L	M	A	V	Q	H	L
L_1	L_2	M_1	M_2	A_1	A_2	V_1
V_2	Q_1	Q_2	H_1	H_2	L_1	L_2

33

M	L	A	X	K	P	W
M_1	M_2	L_1	L_2	A_1	A_2	X_1
X_2	K_1	K_2	P_1	P_2	W_1	W_2

34

S	T	M	C	X	K	Q
S_1	S_2	T_1	T_2	M_1	M_2	C_1
C_2	X_1	X_2	K_1	K_2	Q_1	Q_2

35

V	H	S	I	X	S	L
V_1	V_2	H_1	H_2	S_1	S_2	I_1
I_2	X_1	X_2	S_1	S_2	L_1	L_2

36

L	W	X	L	X	R	S
L_1	L_2	W_1	W_2	X_1	X_2	L_1
L_2	X_1	X_2	R_1	R_2	S_1	S_2

37

G	Z	D	F	K	L	N
G_1	G_2	Z_1	Z_2	D_1	D_2	F_1
F_2	K_1	K_2	L_1	L_2	N_1	N_2

38

Y	B	X	M	R	B	N
Y_1	Y_2	B_1	B_2	X_1	X_2	M_1
M_2	R_1	R_2	B_1	B_2	N_1	N_2

39

A	D	K	T	T	B	A
A_1	A_2	D_1	D_2	K_1	K_2	T_1
T_2	T_1	T_2	B_1	B_2	A_1	A_2

40

E	O	B	H	W	V	L
E_1	E_2	O_1	O_2	B_1	B_2	H_1
H_2	W_1	W_2	V_1	V_2	L_1	L_2

41

Y	S	X	M	B	O	W
Y_1	Y_2	S_1	S_2	X_1	X_2	M_1
M_2	B_1	B_2	O_1	O_2	W_1	W_2

42

P	G	X	K	O	R	Z
P_1	P_2	G_1	G_2	X_1	X_2	K_1
K_2	O_1	O_2	R_1	R_2	Z_1	Z_2

43

I	U	C	E	A	D	Y
I_1	I_2	U_1	U_2	C_1	C_2	E_1
E_2	A_1	A_2	D_1	D_2	Y_1	Y_2

44

I	D	B	L	Z	M	I
I_1	I_2	D_1	D_2	B_1	B_2	L_1
L_2	Z_1	Z_2	M_1	M_2	I_1	I_2

45

T	A	N	H	C	A	I
T_1	T_2	A_1	A_2	N_1	N_2	H_1
H_2	C_1	C_2	A_1	A_2	I_1	I_2

46

D	N	C	I	D	D	O
D_1	D_2	N_1	N_2	C_1	C_2	I_1
I_2	D_1	D_2	D_1	D_2	O_1	O_2

47

Y	I	B	C	N	O	L
Y_1	Y_2	I_1	I_2	B_1	B_2	C_1
C_2	N_1	N_2	O_1	O_2	L_1	L_2

48

Y	U	U	M	C	E	P
Y_1	Y_2	U_1	U_2	U_1	U_2	M_1
M_2	C_1	C_2	E_1	E_2	P_1	P_2

49

O	T	D	M	G	B	F
O_1	O_2	T_1	T_2	D_1	D_2	M_1
M_2	G_1	G_2	B_1	B_2	F_1	F_2

50

U	N	A	H	L	B	D
U_1	U_2	N_1	N_2	A_1	A_2	H_1
H_2	L_1	L_2	B_1	B_2	D_1	D_2

51

W	X	N	X	K	K	C
W_1	W_2	X_1	X_2	N_1	N_2	X_1
X_2	K_1	K_2	K_1	K_2	C_1	C_2

52

S	C	T	O	X	T	S
S_1	S_2	C_1	C_2	T_1	T_2	O_1
O_2	X_1	X_2	T_1	T_2	S_1	S_2

53

D	A	Y	X	H	K	C
D_1	D_2	A_1	A_2	Y_1	Y_2	X_1
X_2	H_1	H_2	K_1	K_2	C_1	C_2

54

N	L	D	K	R	R	F
N_1	N_2	L_1	L_2	D_1	D_2	K_1
K_2	R_1	R_2	R_1	R_2	F_1	F_2

55

K	Y	A	P	M	H	C
K_1	K_2	Y_1	Y_2	A_1	A_2	P_1
P_2	M_1	M_2	H_1	H_2	C_1	C_2

56

A	N	M	B	V	G	R
A_1	A_2	N_1	N_2	M_1	M_2	B_1
B_2	V_1	V_2	G_1	G_2	R_1	R_2

57

E	Z	Q	A	T	C	Y
E_1	E_2	Z_1	Z_2	Q_1	Q_2	A_1
A_2	T_1	T_2	C_1	C_2	Y_1	Y_2

58

I	M	N	D	L	R	L
I_1	I_2	M_1	M_2	N_1	N_2	D_1
D_2	L_1	L_2	R_1	R_2	L_1	L_2

59

G	M	T	W	E	T	R
G_1	G_2	M_1	M_2	T_1	T_2	W_1
W_2	E_1	E_2	T_1	T_2	R_1	R_2

60

C	V	V	K	T	E	D
C_1	C_2	V_1	V_2	V_1	V_2	K_1
K_2	T_1	T_2	E_1	E_2	D_1	D_2

61

U	F	D	E	L	X	H
U_1	U_2	F_1	F_2	D_1	D_2	E_1
E_2	L_1	L_2	X_1	X_2	H_1	H_2

62

E	Q	V	C	B	L	Y
E_1	E_2	Q_1	Q_2	V_1	V_2	C_1
C_2	B_1	B_2	L_1	L_2	Y_1	Y_2

63

U	D	U	G	Y	A	F
U_1	U_2	D_1	D_2	U_1	U_2	G_1
G_2	Y_1	Y_2	A_1	A_2	F_1	F_2

64

H	N	Q	L	K	F	R
H_1	H_2	N_1	N_2	Q_1	Q_2	L_1
L_2	K_1	K_2	F_1	F_2	R_1	R_2

65

U	C	N	V	D	L	H
U_1	U_2	C_1	C_2	N_1	N_2	V_1
V_2	D_1	D_2	L_1	L_2	H_1	H_2

66

L	Z	D	R	E	L	K
L_1	L_2	Z_1	Z_2	D_1	D_2	R_1
R_2	E_1	E_2	L_1	L_2	K_1	K_2

67

X	K	U	P	S	E	M
X_1	X_2	K_1	K_2	U_1	U_2	P_1
P_2	S_1	S_2	E_1	E_2	M_1	M_2

68

C	T	N	K	T	K	E
C_1	C_2	T_1	T_2	N_1	N_2	K_1
K_2	T_1	T_2	K_1	K_2	E_1	E_2

69

B	O	E	E	P	G	V
B_1	B_2	O_1	O_2	E_1	E_2	E_1
E_2	P_1	P_2	G_1	G_2	V_1	V_2

70

Q	T	G	W	E	R	H
Q_1	Q_2	T_1	T_2	G_1	G_2	W_1
W_2	E_1	E_2	R_1	R_2	H_1	H_2

71

L	Z	D	R	E	L	K
L_1	L_2	Z_1	Z_2	D_1	D_2	R_1
R_2	E_1	E_2	L_1	L_2	K_1	K_2

72

F	A	X	I	Y	D	A
F_1	F_2	A_1	A_2	X_1	X_2	I_1
I_2	Y_1	Y_2	D_1	D_2	A_1	A_2

73

K	Z	L	X	X	O	R
K_1	K_2	Z_1	Z_2	L_1	L_2	X_1
X_2	X_1	X_2	O_1	O_2	R_1	R_2

74

R	P	E	R	R	R	R
R_1	R_2	P_1	P_2	E_1	E_2	R_1
R_2	R_1	R_2	R_1	R_2	R_1	R_2

75

N	C	I	E
N_1	N_2	C_1	C_2
I_1	I_2	E_1	E_2

FIGURE 89—Continued.

REF ID:A59439

c. The text now being transcribed into periods of 7, with the Θ_1 and Θ_2 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 Θ_1 and Θ_2 superimposed components are identical, whereupon 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 equivalencies can be established between a few of the Θ_1 components, or between a few of the Θ_2 components, or between a few Θ_1 and Θ_2 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 25 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:

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

FIGURE 92.

REF ID:A59439

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 Θ_1 coordinates applying to these two rows, A_1 and F_1, 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.

f. 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 N F F U U F M M N A N U N F A U M A N U F
F..... A M M A N F F U U F M M N A N U N F A U M A N U F

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 N F F U U F M M N A N U N F A U M A N U F
C..... T C C T R G G I I G C C R T R I R G T I C T R I G

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 C, F 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.

g. In applying the usual χ-test for matching two distributions use is made of the important 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

REF ID:A59439

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:

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

Adding the squares of the probabilities for separate occurrence [5] of the letters in each category:

A	.0054	B	.0001	K	.0000
E	.0169	C	.0009	Q	.0000
I	.0054	D	.0018	V	.0002
N	.0063	F	.0008	W	.0002
O	.0057	G	.0003	X	.0000
R	.0057	H	.0012	Z	.0000
S	.0037	L	.0013		
T	.0084	M	.0006	Total=	.0004
		P	.0007	Average=	.00007
Total=	.0575	U	.0007		
Average=	.0072	Y	.0004		
		Total=	.0088		
		Average=	.0008		

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:

$$5\times.0072=.0360$$
$$15\times.0008=.0120$$
$$5\times.00007=.0004$$
$$\text{Total}=.0484$$

This, then, is the constant that should be applied in the χ-test for the problem under consideration. Suppose, for convenience, the approximation .05 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 apply to identical rows of letters and indicate two letters in the same row in the checkerboard, since 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.

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

[5] As given in the table on p. 114 of *Military Cryptanalysis*, Part 1, Appendix 2, par. 2*e* (1).

REF ID:A594439

	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	
D_1		4			2		1		1		3		1			1		1					3	1	1	N=23
M_1		1			2	1	1	1		1	4		1		1	1						1	2			N=16

FIGURE 93.

Applying the χ-test, the observed value of $\chi=34$, the expected value is $.05\ (23\times16)=18$. 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?

i. 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 equivalency between some of the variant forms of $\begin{matrix}\Theta_1\\ \Theta_2\end{matrix}$ vertical pairs in the cryptogram. If the student will think over the matter he will quickly see that this manifestation of equivalency is but a reflection of the principle elucidated in paragraph 46, expressed in a little different way. In other words, establishing equivalence between two Θ_1 components means that the two base letters involved belong in the same row of the checkerboard; establishing equivalence betweeen two Θ_2 components means that the two base letters involved belong in the same column of the checkerboard. Note the following instances of apparent equivalency between D_1 and M_1:

Period			
	16	$D_2\ B_1$	$D_1\ B_2$
	20	$Y_2\ B_1$	$D_1\ B_2$
	49	$T_2\ B_1$	$D_1\ B_2$
	2	$Y_2\ B_1$	$M_1\ B_2$
	3	$Z_2\ G_1$	$D_1\ G_2$
	56	$N_2\ G_1$	$M_1\ G_2$
	13	$Q_2\ L_1$	$D_1\ L_2$
	37	$Z_2\ L_1$	$D_1\ L_2$
	58	$N_2\ L_1$	$D_1\ L_2$
	66	$Z_2\ L_1$	$D_1\ L_2$
	71	$Z_2\ L_1$	$D_1\ L_2$
	6	$A_2\ L_1$	$M_1\ L_2$
	13	$Y_2\ L_1$	$M_1\ L_2$
	58	$I_2\ L_1$	$M_1\ L_2$

Period				
	18	$G_2\ Q_1$	$D_1\ Q_2$	
	32	$L_2\ Q_1$	$M_1\ Q_2$	
	16	$F_2\ E_1$	$D_1\ E_2$	
	17	$F_2\ E_1$	$D_1\ E_2$	
	59	$G_2\ E_1$	$M_1\ E_2$	
	12	$A_1\ B_2$	$A_2\ D_1$	$B_1\ D_2$
	50	$A_1\ B_2$	$A_2\ D_1$	$H_1\ D_2$
	8	$D_1\ B_2$	$D_2\ M_1$	$R_1\ M_2$
	19	$D_1\ N_2$	$D_2\ D_1$	$U_1\ D_2$
	46	$D_1\ I_2$	$D_2\ D_1$	$N_1\ D_2$
	44	$D_1\ Z_2$	$D_2\ M_1$	$B_1\ M_2$
	43	$U_1\ A_2$	$U_2\ D_1$	$C_1\ D_2$
	67	$U_1\ E_2$	$U_2\ M_1$	$P_1\ M_2$

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

D_1+M_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
	2	5			4		2		1		7		2		1	2		1				1	5	1	1

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

	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	
D_1		4			2		1		1		3		1			1		1					3	1	1	N=23
A_1		3	1		2						4					1							3			N=24

REF ID:A59439

Expected for plain text: .05 (23×21)=24
Expected for random text: .038 (23×21)=18
Observed =26

	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	
M_1		~			=	~	~	~		~	≡		~		~	~						~	~			N=16
A_1	≡	~	=			=	≡		~						~	~			~				=	≡		N=21

Expected for plain text: .05 (16×21)=17
Expected for random text: .038 (16×21)=13
Observed =14

	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	
$D_1 M_1$		≋			≡	~	=	~	~	~	≋		=		~	=	~	~	~			~	≡	~	~	N=39
A_1	≡	~	=			=	≡		~						~	~			~				=	≡		N=21

Expected for plain text: .05 (39×21)=41
Expected for random text: .038 (39×21)=31
Observed =39

From the foregoing calculations it appears that A_1 may be similar to the D_1 and the M_1 distributions, for the observed values, while not as great as expected for plain text, are higher than those expected for random text. Perhaps more conclusive evidence may be found if a search were made through the text to see if any equivalences between A_1, D_1, and M_1, appear.

Note the following cases:

Period				
	8	Y_2 B_1	A_1 B_2	A_2 M_1
	12	N_2 B_1	A_1 B_2	A_2 D_1
	50	N_2 B_1	A_1 B_2	A_2 D_1
	16	D_2 B_1	D_1 B_2	D_2 A_1
	20	Y_2 B_1	D_1 B_2	D_2 A_1
	49	T_2 B_1	D_1 B_2	D_2 F_1
	2	Y_2 B_1	M_1 B_2	M_2 R_1
	21	C_2 G_1	A_1 G_2	A_2 X_1
	30	K_2 G_1	A_1 G_2	S_1 P_2
	3	Z_2 G_1	D_1 G_2	B_1 K_2
	56	N_2 G_1	M_1 G_2	M_2 R_1
	24	X_2 K_1	A_1 K_2	R_1 Z_2
	34	T_2 K_1	M_1 K_2	M_2 Q_1

Period				
	33	L_2 P_1	A_1 P_2	A_2 W_1
	48	U_2 P_1	M_1 P_2	O_1 M_2
	15	G_2 Q_1	A_1 Q_2	A_2 B_1
	18	G_2 Q_1	D_1 Q_2	D_1 N_2
	32	L_2 Q_1	M_1 Q_2	M_2 H_1
	14	N_2 X_1	A_1 X_2	A_2 C_1
	61	F_2 X_1	D_1 X_2	D_2 H_1
	57	Q_2 Y_1	A_1 Y_2	I_1 D_2
	72	F_2 Y_1	A_1 Y_2	A_2 D_1
	63	U_2 Y_1	D_1 Y_2	D_2 A_1

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.

324155—41——12

REF ID:A594382

j. Among the most frequent combinations is the pair $Y_2 B_1$, appearing in the following sequences:

Period					
	2	$Y_2 C_1$	$Y_1 C_2$	$Y_2 B_1$	$M_1 B_2$	$M_2 R_1$
	8	$L_1 Q_2$	$Y_1 W_2$	$Y_2 B_1$	$A_1 B_2$	$A_2 M_1$
	10	$B_1 A_2$	$Y_1 O_2$	$Y_2 B_1$	$E_1 B_2$	$E_2 R_1$
	20	$N_1 L_2$	$Y_1 W_2$	$Y_2 B_1$	$D_1 B_2$	$D_2 A_1$
	41	$H_1 L_2$	$Y_1 M_2$	$Y_2 B_1$	$S_1 B_2$	$S_2 O_1$

Note how M_1, A_1, E_1, D_1, and S_1 all appear to be interchangeable. Are these the 5 letters which belong in the same row? The probable equivalence among A_1, D_1, and M_1 has been established by noting cases of equivalency in the text. A further search will be made to see if E_1 and S_1 also show equivalencies with A_1, D_1, and M_1.

Note the following:

Period				
	21	$C_2 G_1$	$A_1 G_2$	
	30	$K_2 G_1$	$A_1 G_2$	
	3	$Z_2 G_1$	$D_1 G_2$	
	69	$O_2 G_1$	$E_1 G_2$	
	56	$N_2 G_1$	$M_1 G_2$	
	23	$D_1 X_2$	$D_2 H_1$	$A_1 H_2$
	32	$M_1 Q_2$	$M_2 H_1$	$A_1 H_2$
	61	$D_1 X_2$	$D_2 H_1$	$E_1 H_2$

Period				
	12	$N_2 B_1$	$A_1 B_2$	$A_2 D_1$
	10	$Y_2 B_1$	$E_1 B_2$	$E_2 R_1$
	30	$L_2 P_1$	$E_1 P_2$	$E_2 R_1$
	33	$L_2 P_1$	$A_1 P_2$	$A_2 W_1$
	43	$C_2 Y_1$	$E_1 Y_2$	$I_1 L_2$
	57	$Q_2 Y_1$	$A_1 Y_2$	$I_1 D_2$

Here are indications that E_1 belongs to the same series, but not enough cases where S_1 is interchangeable with A, D, E, or M can be found to be convincing. But perhaps it is best not to go too fast in these early stages. Let it be assumed for the present that A, D, E, and M are in the same row of the substitution checkerboard. In period 16 there is the pair of vertical components D_1 E_2. Since $D_1 = E_1$ this pair may be written $E_1 E_2$, 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.

k. 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 Θ_2 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.

l. 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 90 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.

m. 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 involved is indicated by either its Θ_1 or Θ_2 component. This property affords a fair basis for assuming a probable word. For example, the cryptogram here studied shows the following two periods:

REF ID:A59439

Period	35							36						
Plain text					S			L						
Components	V_1	V_2	H_1	H_2	S_1	S_2	I_1	L_1	L_2	W_1	W_2	X_1	X_2	L_1
	I_2	X_1	X_2	S_1	S_2	L_1	L_2	L_2	X_1	X_2	R_1	R_2	S_1	S_2
Cipher text	V	H	S	I	X	S	L	L	W	X	L	X	R	S

Two letters are quite definite, S_s and L_s. Suppose the possible plain-text letters be indicated.

Period	35							36						
Possible plain-text letters	V	.	H	.	S	.	I	L	.	W	.	X	.	L
	I		X				L			X		R		S
Components	V_1	V_2	H_1	H_2	S_1	S_2	I_1	L_1	L_2	W_1	W_2	X_1	X_2	L_1
	I_2	X_1	X_2	S_1	S_2	L_1	L_2	L_2	X_1	X_2	R_1	R_2	S_1	S_2
Cipher text	V	H	S	I	X	S	L	L	W	X	L	X	R	S

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:

Period	35					36	
Plain text	H	O	S	T	I	L	E
Cipher-text components	V_1	V_2	H_1	H_2	S_1	S_2	I_1	L_1	L_2
	I_2	X_1	X_2	S_1	S_2	L_1	L_2	L_2	X_1
Plain-text components		H_1	O_1	S_1	T_1	I_1	L_1	E_1
		H_2	O_2	S_2	T_2	I_2	L_2	E_2

This word, if correct, yields the following equivalencies: $H_2=X_2=O_1$; $S_1=O_2$; $T_1=S_2$; $L_1=T_2$; $I_2=L_2=E_1$; $X_1=E_2$. 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 $I_1 L_2=I_1 I_2=I_p$. This is unfortunate, so that additional words will have to be assumed. The 14th period shows a C_p and the components after it suggest that the word CROSSROADS may be present. Thus:

Period	14							15						
Plain text					C	R	O	S	S	R	O	A	D	S
Components	N_1	N_2	A_1	A_2	C_1	C_2	O_1	R_1	R_2	G_1	G_2	A_1	A_2	S_1
	O_2	X_1	X_2	C_1	C_2	R_1	R_2	S_2	W_1	W_2	Q_1	Q_2	B_1	B_2
Cipher text	N	A	C	O	X	C	R	R	G	A	S	W	Q	B

Take the first letter R_p, represented by $C_2 R_1$.

Since $R_p=C_2 R_1$,

Therefore, $R_1 R_2=C_2 R_1$

Hence $R_1=C_2$ and $R_2=R_1$

Therefore, $R_1=R_2=C_2$

Again, in the case of the first O_p,

$O_p=O_1 R_2$

But $O_p=O_1 O_2=O_1 R_2$

$O_2=R_2$

Therefore, $R_1=R_2=O_2=C_2$

REF ID: A594319

The various equivalencies yielded are as follows:

$C_2=R_1=O_2=S_1=R_2=G_1=W_2=Q_1$
$S_2=W_1=B_2=T_1$
$H_2=X_2=O_1=G_2$
$O_2=Q_1$
$Q_2=A_2=D_1$

$B_1=D_2$
$L_1=T_2$
$L_2=I_2=E_1$
$X_1=E_2$

n. Let all the equivalencies found thus far from subparagraphs *e*, *f*, and *h* be collected in two tables, as shown in figure 94, one for $\Theta_1\ \Theta_2$ combinations, the other for $\Theta_2\ \Theta_1$ combinations.

$\Theta_1\Theta_2$ COMBINATIONS

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
D_1	H_1		A_1	A_1		O_2	B_1			T_2	A_1	B_1	H_2		G_1	G_1	O_2	S_2			T_1	E_2		
E_1	N_1		E_1	D_1		C_2	N_1				D_1	H_1	X_2		O_2	C_2	C_2	W_1			S_2			
M_1	D_2		M_1	M_1		M_2	D_2				E_1	D_2	G_2		C_2	O_2	M_2	B_2			B_2			
I_2			I_2	I_2		W_2					I_2				M_2	M_2	W_2							
L_2			L_2	L_2		R_1					L_2				W_2	W_2	R_1							
A_2			A_2	A_2		R_2					A_2				R_1	R_2	R_2							
Q_2			Q_2	Q_2		S_1					Q_2				R_2	S_1	G_1							
						Q_1									S_1	Q_1	Q_1							

$\Theta_2\Theta_1$ COMBINATIONS

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
A_1	T_1	M_2	B_1	X_1		O_1	X_2	L_2		I_2	C_2		C_2		L_2	C_2	T_1	L_1			C_2	H_2		
D_1	S_2	O_2	H_1			H_2	O_1	A_1		A_1	O_2		M_2		I_2	M_2	W_1				M_2	O_1		
E_1	W_1	W_2	N_1			X_2	G_2	D_1		D_1	W_2		W_2		A_1	O_2	B_2				O_2	G_2		
M_1		R_1						E_1		E_1	R_1		S_1		D_1	W_2					S_1			
I_2		S_1						M_1		M_1	S_1		R_1		E_1	S_1					R_1			
L_2		R_2						A_2		A_2	R_2		R_2		M_1	R_1					R_2			
Q_2		G_1						Q_2		Q_2	G_1		G_1		A_2	G_1					G_1			
		Q_1									Q_1		Q_1			Q_1					Q_1			

FIGURE 94

A study of the equivalencies indicates that—

(1) A, D, E, M 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) I, L, A, Q belong in the same column.
(6) X, H, G belong in the same column.
(7) The 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.

o. 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, Q, S, fall in 3 different rows; C, O, M, W, R and I, 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 $R_1=O_2=C_2=$

REF ID:A59439

$M_2=W_2$, 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

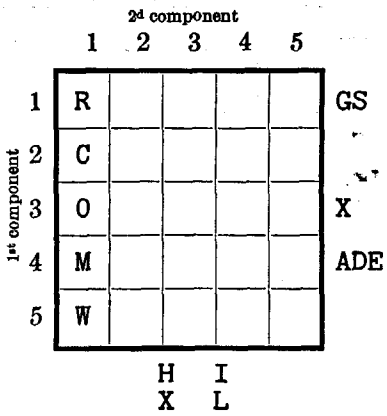


FIGURE 95–A.

must be in row 4, H and X must be in column 3. This means that θ_1 for A, D, and E must be 4, and that θ_2 for H and X must be 3. And since $M_1=I_2=L_2$, θ_2 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	36
Plain text	H O S T I	L E
Components	$V_1\ V_2\ H_1\ H_2\ S_1\ S_2\ I_1$	$L_1\ L_2\ W_1\ W_2\ X_1\ X_2\ L_1$
	$I_2\ X_1\ X_2\ S_1\ S_2\ L_1\ L_2$	$L_2\ X_1\ X_2\ R_1\ R_2\ S_1\ S_2$
Cipher text	V H S I X S L	L W X L X R S

the letters R_p and O_p as the second and fourth letters after E_p, suggesting that the word after HOSTILE is TROOP. This gives $W_1\ X_2=T_p$, which permits of placing T in position 5–3. Since T in HOSTILE$=S_2\ L_1$, therefore $S_2=5$ and $L_1=3$. Since S is in row 1, and $S_2=5$, S must go in position 1–5. Since $L_2=4$ and $L_1=3$, L must go in position 3–4. Since O_p (the 1st O in TROOP)$=X_1R_2$ and it is known that $O_p=3$–1, therefore X must be in position 3–3. The checkerboard is now as shown in figure 95–B. From figure 94, $X_1=E_2$. Now $X_1=3$, and since the E must be in row 4,

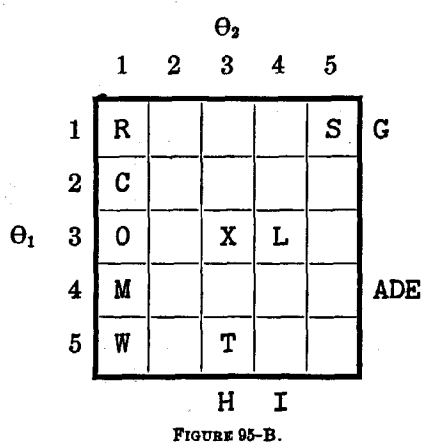


FIGURE 95–B.

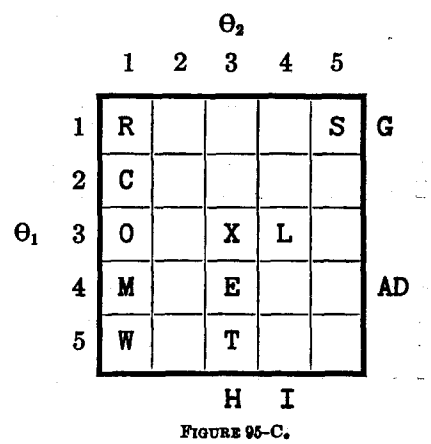


FIGURE 95–C.

REF ID:A594339

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 as shown in figure 95–D. It will decipher the entire cryptogram as it stands, but speculating upon the presence of W U T V Z in the last row, and assuming a key-word mixed sequence has brought this about, a rearrangement of the columns of the checkerboard is made to give T U V W Z, 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.

θ_2

θ_1	1	2	3	4	5
1	R	K	G	Q	S
2	C	N	H	I	B
3	O	Y	X	L	P
4	M	D	E	A	F
5	W	U	T	V	Z

FIGURE 95–D.

θ_2

θ_1	3	2	4	1	5
1	G	K	Q	R	S
2	H	N	I	C	B
3	X	Y	L	O	P
4	E	D	A	M	F
5	T	U	V	W	Z

FIGURE 95–E.

θ_2

θ_1	1	2	3	4	5
1	X	Y	L	O	P
2	H	N	I	C	B
3	E	D	A	M	F
4	G	K	Q	R	S
5	T	U	V	W	Z

FIGURE 95–F.

p. The completely deciphered cryptogram is as follows:

1

S I T U A T I
4 2 5 5 3 5 2
5 3 1 2 3 1 3
K Z F B E I L
7

2

O N O N F R O
1 2 1 2 3 4 1
4 2 4 2 5 4 4
Y Y M O C B R
14

3

N T O F T W E
2 5 1 3 5 5 3
2 1 4 5 1 4 1
B L Z D O T G
21

4

N T Y F O U R
2 5 1 3 1 5 4
2 1 2 5 4 2 4
B L P K Y W C
28

5

T H B R I G A
5 2 2 4 2 4 3
1 1 5 4 3 1 3
U C C E P Q L
35

6

D E A S F O L
3 3 3 4 3 1 1
2 1 3 5 5 4 3
A M E Y L Z Q
42

7

L O W S C O L
1 1 5 4 2 1 1
3 4 4 5 4 4 3
X W H L R W Q
49

8

O N F I R S T
1 2 3 2 4 4 5
4 2 5 3 4 5 1
Y D R W B M T
56

9

B A T T A L I
2 3 5 1 3 1 2
5 3 1 1 3 3 3
I Z E B E L A
63

10

O N F O R T Y
1 2 3 1 4 5 1
4 2 5 4 4 1 2
Y E S O B R Y
70

11

S E V E N T H
4 3 5 3 2 5 2
5 1 3 1 2 1 1
Q V B B L Y X
77

12

I N F A N T R
2 2 3 3 2 5 4
3 2 5 3 2 1 4
N A B Q B D O
84

13

Y H A S R E A
1 2 3 4 4 3 3
2 1 3 5 4 1 3
Y M Q D L W L
91

14

C H E D C R O
2 2 3 3 2 4 1
4 1 1 2 4 4 4
N A C O X C R
98

15

S S R O A D S
4 4 4 1 3 3 4
5 5 4 4 3 2 5
R G A S W Q B
105

16

E V E N F I V
3 5 3 2 3 2 5
1 3 1 2 5 3 3
F D D T E B A
112

REF ID:A59439

17

E S E V E N D
3 4 3 5 3 2 3
1 5 1 3 1 2 2
M F D E T E N

119

18

D A S H R O A
3 3 4 2 4 1 3
2 3 5 1 4 4 3
A K G D F O Q

126

19

D J U N C T I
3 2 5 2 2 5 2
2 3 2 2 4 1 3
D U B N D C L

133

20

O N F I V E T
1 2 3 2 5 3 5
4 2 5 3 3 1 1
Y D V W B A X

140

21

H R E E T H R
2 4 3 3 5 2 4
1 4 1 1 1 1 4
C A U G G X O

147

22

E E G S T O P
3 3 4 4 5 1 1
1 1 1 5 1 4 5
A R T X X T S

154

23

E N E M Y H O
3 2 3 3 1 2 1
1 2 1 4 2 1 4
D A Y X H K O

161

24

L D S W O O D
1 3 4 5 1 1 3
3 2 5 4 4 4 2
L S X A B R K

168

25

S S O U T H W
4 4 1 5 5 2 5
5 5 4 2 1 1 4
R P U Z W H O

175

26

E S T O F C H
3 4 5 1 3 2 2
1 5 1 4 5 4 1
M T D H T S G

182

27

A R L E S T O
3 4 1 3 4 5 1
3 4 3 1 5 1 4
M L S L Q P O

189

28

W N I N C O N
5 2 2 2 2 1 2
4 2 3 2 4 4 2
U N H C I C K

196

29

S I D E R A B
4 2 3 3 4 3 2
5 3 2 1 4 3 5
K A Q B D O F

203

30

L E F O R C E
1 3 3 1 4 2 3
3 1 5 4 4 4 1
L E K A P R G

210

31

S T O P W I L
4 5 1 1 5 2 1
5 1 4 5 4 3 3
S X U P O W A

217

32

L M A K E E V
1 3 3 4 3 3 5
3 4 3 2 1 1 3
L M A V Q H L

224

33

E R Y E F F O
3 4 1 3 3 3 1
1 4 2 1 5 5 4
M L A X K P W

231

34

R T T O D R I
4 5 5 1 3 4 2
4 1 1 4 2 4 3
S T M C X K Q

238

35

V E H O S T I
5 3 2 1 4 5 2
3 1 1 4 5 1 3
V H S I X S L

245

36

L E T R O O P
1 3 5 4 1 1 1
3 1 1 4 4 4 5
L W X L X R S

252

37

S O U T A N D
4 1 5 5 3 2 3
5 4 2 1 3 2 2
G Z D F K L N

259

38

O C C U P Y D
1 2 2 5 1 1 3
4 4 4 2 5 2 2
Y B X M R B N

266

39

E F E N S I V
3 3 3 2 4 2 5
1 5 1 2 5 3 3
A D K T T B A

273

40

E P O S I T I
3 1 1 4 2 5 2
1 5 4 5 3 1 3
E O B H W V L

280

41

O N S T O P M
1 2 4 5 1 1 3
4 2 5 1 4 5 4
Y S X M B O W

287

42

Y T R O O P S
1 5 4 1 1 1 4
2 1 4 4 4 5 5
P G X K O R Z

294

43

H A V I N G D
2 3 5 2 2 4 3
1 3 3 3 2 1 2
I U C E A D Y

301

44

I F F I C U L
2 3 3 2 2 5 1
3 5 5 3 4 2 3
I D B L Z M I

308

45

T Y M A I N T
5 1 3 3 2 2 2
1 2 4 3 3 2 3
T A N H C A I

315

46

A I N I N G C
3 2 2 2 2 4 2
3 3 2 3 2 1 4
D N C I D D O

322

47

O N N E C T I
1 2 2 3 2 5 2
4 2 2 1 4 1 3
Y I B C N O L

329

48

O N W I T H F
1 2 5 2 5 2 3
4 2 4 3 1 1 5
Y U U M C E P

336

49							50							51							52						
O	R	T	Y	F	I	F	T	H	I	N	F	A	N	T	R	Y	O	N	N	O	R	T	H	S	T	O	P
1	4	5	1	3	2	3	5	2	2	2	3	3	2	5	4	1	1	2	2	1	4	5	2	4	5	1	1
4	4	1	2	5	3	5	1	1	3	2	5	3	2	1	4	2	4	2	2	4	4	1	1	5	1	4	5
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
						343							350							357							364

53							54							55							56						
E	N	E	M	Y	N	O	N	C	O	M	M	I	S	S	I	O	N	E	D	O	F	F	I	C	E	R	C
3	2	3	3	1	2	1	2	2	1	3	3	2	4	4	2	1	2	3	3	1	3	3	2	2	3	4	2
1	2	1	4	2	2	4	2	4	4	4	4	3	5	5	3	4	2	1	2	4	5	5	3	4	1	4	4
D	A	Y	X	H	K	C	N	L	D	K	R	R	F	K	Y	A	P	M	H	C	A	N	M	B	V	G	R
						371							378							385							392

57							58							59							60						
A	P	T	U	R	E	D	N	E	A	R	C	H	A	R	L	E	S	T	O	W	N	S	T	A	T	E	S
3	1	5	5	4	3	3	2	3	3	4	2	2	3	4	1	3	4	5	1	5	2	4	5	3	5	3	4
3	5	1	2	4	1	2	2	1	3	4	4	1	3	4	3	1	5	1	4	4	2	5	1	3	1	1	5
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	P
						399							406							413							420

61							62							63							64						
T	H	A	T	E	N	E	M	Y	S	E	V	E	N	T	H	D	I	V	I	S	I	O	N	I	S	M	O
5	2	3	5	3	2	3	3	1	4	3	5	3	2	5	2	3	2	5	2	4	2	1	2	2	4	3	1
1	1	3	1	1	2	1	4	2	5	1	3	1	2	1	1	2	3	3	3	5	3	4	2	3	5	4	4
U	F	D	E	L	X	H	E	Q	V	C	B	L	Y	U	D	U	G	Y	A	F	H	N	Q	L	K	F	R
						427							434							441							448

65							66							67							68						
V	I	N	G	I	N	T	O	A	T	T	A	C	K	P	O	S	I	T	I	O	N	S	T	O	N	I	G
5	2	2	4	2	2	5	1	3	5	5	3	2	4	1	1	4	2	5	2	1	2	4	5	1	2	2	4
3	3	2	1	3	2	1	4	3	1	1	3	4	2	5	4	5	3	1	3	4	2	5	1	4	2	3	1
U	C	N	V	D	L	H	L	Z	D	R	E	L	K	X	K	U	P	S	E	M	C	T	N	K	T	K	E
						455							462							469							476

69							70							71							72						
H	T	P	R	E	P	A	R	A	T	O	R	Y	T	O	A	T	T	A	C	K	A	T	D	A	Y	L	I
2	5	1	4	3	1	3	4	3	5	1	4	1	5	1	3	5	5	3	2	4	3	5	3	3	1	1	2
1	1	5	4	1	5	3	4	3	1	4	4	2	1	4	3	1	1	3	4	2	3	1	2	3	2	3	3
B	O	E	E	P	G	V	Q	T	G	W	E	R	H	L	Z	D	R	E	L	K	F	A	X	I	Y	D	A
						483							490							497							504

73							74							75			
G	H	T	T	O	M	O	R	R	O	W	M	O	R	N	I	N	G
4	2	5	5	1	3	1	4	4	1	5	3	1	4	2	2	2	4
1	1	1	1	4	4	4	4	4	4	4	4	4	4	2	3	3	1
K	Z	L	X	X	O	R	R	P	E	R	R	R	R	N	C	I	E
						511							518				522

q. The steps taken in recovering the original substitution checkerboard demonstrate that cyclic permutations of a correct checkerboard will serve to decipher such a cryptogram just as well as the original checkerboard. In other words, a cryptogram prepared according to this method is decipherable by factorial 5 ($5\times4\times3\times2\times1=120$) checkerboards, all of which are cyclically equivalent. Even though the identities of the components will be different if the same message is enciphered by two different cyclically-equivalent checkerboards, when these components are recombined, they will yield identical cipher texts, and therefore so far as external appearances are concerned different checkerboards yield identical cryptograms. The reason

REF ID:A59439

that there are only factorial 5 cyclically-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 aforesaid relationship hold true. If two checkerboards have identical 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 it 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 about 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 can be found almost from the context alone. And in establishing these 6 values, the letters involved are not so interrelated that all 6 have to be ascertained 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*, not one by one. 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 security of the bifid system as compared with that afforded by the simple monoalphabetic system. It is a principle well worth remembering and speculating upon.

53. Special solutions for bifid systems.—*a.* The security of the bifid system is very considerably reduced if the situation in which it is employed happens to be such that two or more messages with identical beginnings, endings, or internal portions can often be expected to occur. For in this case it is possible to establish equivalencies between components and quickly reconstruct the substitution checkerboard. An example will be given to illustrate the steps in a specific case.

b. Here are two cryptograms transmitted by two coordinate units to a superior headquarters at about the same time. They show certain identities, which have been underlined.

REF ID:A594439

No. 1. QVBBL YXNAB QBDOY HONDW VUYTE MHQZD QTLKE EWAPK QSLIP QDWC
No. 2. VBNHY XDABG BDOIH OBNWV LYTFW HQXDQ VLKEW WAXDQ SABCA NXGX

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 Θ_1 and Θ_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, $\begin{matrix} N_1 & N_2 \\ Q_2 & D_1 \end{matrix}$; hence $N_1=N_2$ and $Q_2=D_1$. $\begin{matrix} N_2 & H_1 \\ B_1 & D_2 \end{matrix}$; hence $N_2=H_1$ 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:

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

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*, wherein ZA stands for Z and ZB for the 27th character.

REF ID:A59439

	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
	Q_1	Q_2	V_1	V_2	B_1	B_2	B_1	N_1	N_2	A_1	A_2	B_1	B_2	Q_1	Y_1	Y_2	H_1	H_2	O_1	O_2	N_1	U_1	U_2	Y_1	Y_2	T_1	T_2	E_1	Z_1	Z_2	D_1	D_2	Q_1	Q_2	T_1	E_1	E_2	W_1	W_2	A_1	A_2	P_1	L_1	L_2	I_1	I_2	P_1	P_2	Q_1
	B_2	L_1	L_2	Y_1	Y_2	X_1	X_2	Q_2	B_1	B_2	D_1	D_2	O_1	O_2	N_2	D_1	D_2	W_1	W_2	V_1	V_2	N_2	E_2	M_1	M_2	H_1	H_2	Q_1	Q_2	T_2	L_1	L_2	K_1	K_2	E_1	E_2	P_2	K_1	K_2	Q_1	Q_2	S_1	S	Q_2	D_1	D_2	W_1	W_2	C_1	C_2
No. 1	Q	V	B	B	L	Y	X	N	A	B	Q	B	D	O	Y	H	O	N	D	W	V	U	Y	T	E	M	H	Q	Z	D	Q	T	L	K	E	E	W	A	P	K	Q	S	L	I	P	Q	D	W	C
	V_1	V_2	B_1	B_2	N_1	N_2	H_1	A_1	A_2	B_1	B_2	G_1	G_2	B_1	H_1	H_2	O_1	O_2	B_1	B_2	N_1	Y_1	Y_2	T_1	T_2	F_1	F_2	W_1	D_1	D_2	Q_1	Q_2	V_1	V_2	L_1	W_1	W_2	A_1	A_2	X_1	X_2	D_1	B_1	B_2	C_1	C_2	A_1	A_2	N_1
	H_2	Y_1	Y_2	X_1	X_2	D_1	D_2	B_2	D_1	D_2	O_1	O_2	I_1	I_2	N_2	W_1	W_2	V_1	V_2	L_1	L_2	W_2	H_1	H_2	Q_1	Q_2	X_1	X_2	W_2	L_2	K_1	K_2	E_1	E_2	W_1	W_2	D_2	Q_1	Q_2	S_1	S_2	A_1	A_2	N_2	X_1	X_2	G_1	G_2	X_1	X_2
No. 2	V	B	N	H	Y	X	D	A	B	G	B	D	O	I	H	O	B	N	W	V	L	Y	T	F	W	H	Q	X	D	Q	V	L	K	E	W	W	A	X	D	Q	S	A	B	C	A	N	X	G	X

FIGURE 96.

Θ_1 Θ_2 . . .	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	D_2		Q_2	F_1	E_1	Q_1	N_2	N_2	W_2	E_2	L_2	N_2		X_1	G_1			V_1	B_2	T_1	Z_2	T_2	G_2	F_2
	Q_2	Y_2		I_2	V_2	V_2		D_2	D_2	M_2	X_2	Q_2	B_1		T_2							P_2	P_1		
	D_1	N_1		A_1				Y_2	Y_2			D_1	U_2												
	L_2	N_2		L_2				N_1	N_1			A_1	D_2												
	I_2	H_1		A_2				B_1	B_1			A_2	Y_2												
	M_1	K_2		M_1				K_2	K_2			I_2	K_2												
		I_1						I_1	H_1				I_1												
		U_2						U_2	U_2				H_1												

Θ_2 Θ_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
	A_1	U_1		B_1	L_1	Z_1	Y_1		D_1	D_2	M_1	W_2	N_1		W_1	D_1			X_1	N_1	E_1	M_2	L_1	B_1	W_1
	Q_2			N_1	X_2				Q_2	B_1	Q_2	K_1	H_1		Z_2	I_2			P_1	N_2	F_1	K_1	E_2	D_2	P_2
	D_1			Y_2					A_1	Y_2	D_1		I_1			A_1				B_1				N_1	
	L_2			N_2					L_2	N_1	A_1		B_1			L_2				D_2				N_2	
	I_2			H_1					A_2	N_2	A_2		U_2			A_2				H_1				U_2	
	M_1			K_2					M_1	H_1	I_2		D_2			M_1				K_2				H_1	
				I_1						I_1			K_2							I_1				K_2	
				U_2						U_2			Y_2							Y_2				I_1	

FIGURE 97.

REF ID:A59439

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	R	E	L	I	E	F	O	F	Y	O	U	R	R	E	G	I	M	E	N	T	T	O	M	O	R	R	O	W
Components	2	1	2	1	1	1	2	1	3	2	3	2	2	1	1	1	2	1	2	3	3	2	2	2	2	2	2	3
	3	2	1	3	2	2	2	2	3	2	1	3	3	2	3	3	2	2	2	1	1	2	2	2	3	3	2	2
	3	2	3	3	2	3	3	3	1	3	3	3	3	2	1	3	1	3	2	2	2	3	1	3	3	3	3	2
Cipher text	K	A	Q	H	O	R	R	H	W	F	L	X	I	Z	B	F	?	N	A	T	N	N	N	W	R	O	I	Z

CRYPTOGRAM

K A Q H O R R H W F L X I Z A B F Z B N A T N N N W R O I Z

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

REF ID:A194382

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 the chances of finding two or more cryptograms of identical length become quite good and the general solution may be applied.

b. Contrary to the situation in the case of substitution, in that of transposition wherein the letters of the plain-text itself are transposed (not code) the shorter the cryptogram the greater the possibility of solution. For, in the case of a message of say only 25 or 30 letters, one might shift the letters about and actually reconstruct the plain text as one does in the case of the game called "anagrams." Of course, several different "solutions" may thus be obtained, but having such "solutions" it may be possible to reconstruct the system upon which the transposition was based and thus "prove" one of the solutions.

c. The text has confined itself almost entirely to cases of uniliteral transposition, in order to demonstrate basic principles. But there is inherently no reason why transposition may not be applied to digraphs, trigraphs, or tetragraphs. If longer sequences are used as the units of transposition the security decreases very sharply, as in the case of the ordinary route ciphers of the Civil War period.

d. Transposition designs, diagrams, or patterns are susceptible of yielding cryptograms of good security, if they are at 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 "superencipher" 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.

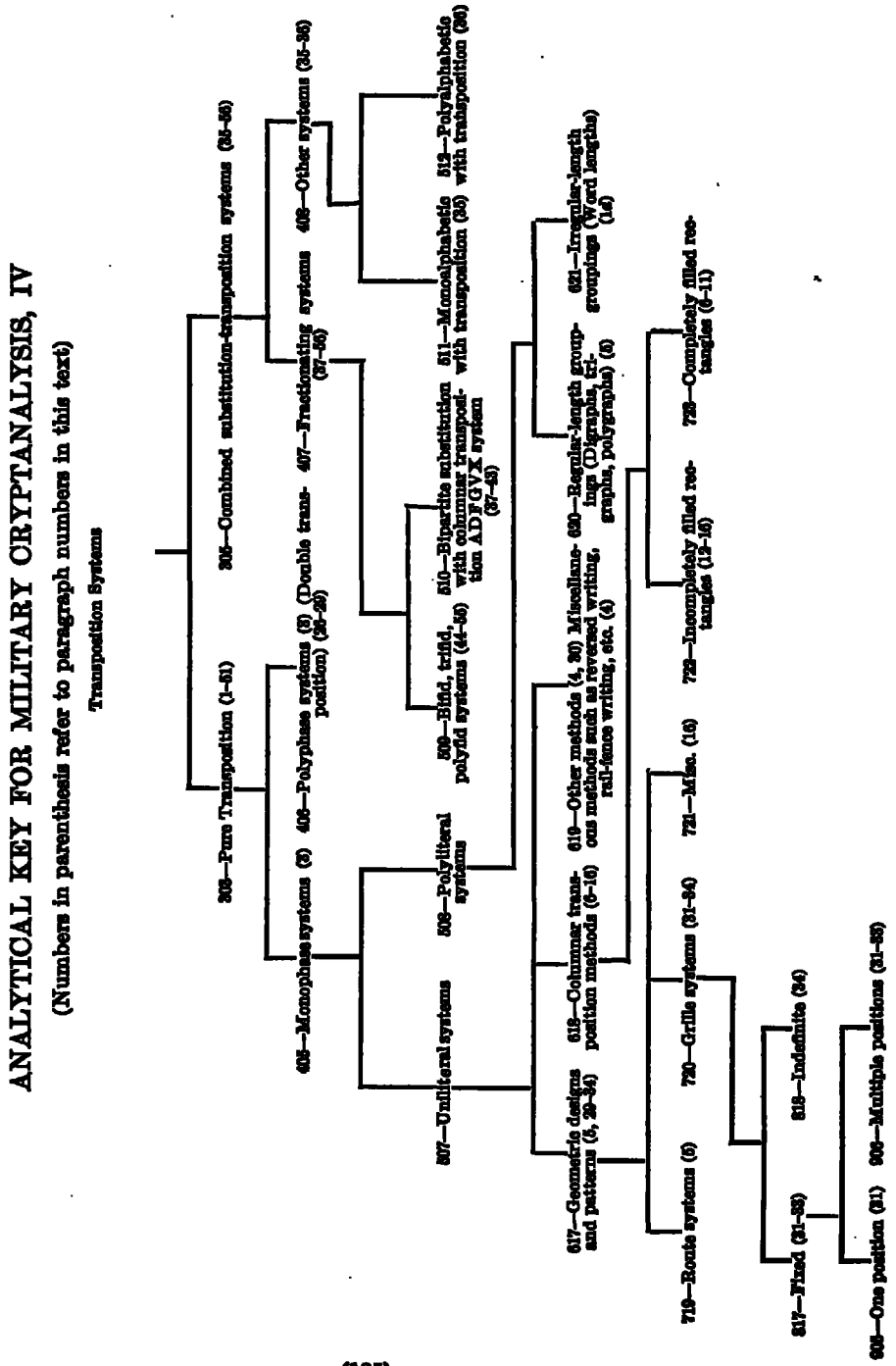
REF ID:A59439

SECTION XI

ANALYTICAL KEY

Analytical key ..... Paragraph 57

57. Analytical key.—Continuing the scheme initiated in the first text of this series, the analytical key applicable to the subject matter and systems embraced in this text is given below.



REF ID:A594439

INDEX

	Page
ADFGVX system	97–143
General solution	124–143
Alternation of components	125–126
Basic principles	125–127, 130–131
Final components	125–126
Illustration of solution	127–143
Initial components	125–126
Minus alternation	131
Plus alternation	131
Special solution	98–124
Exact-factor method of solution	123–124
Solution by means of identical beginnings	105–123
Solution by means of identical endings	98–105
Alternation of components	110, 131
Anagram sequence	33
Anagramming	5, 51–53
Analytical key	185
Ascertaining period in bifid system	165–166
Bifid fractionating systems	144–183
Ascertaining period	165–166
Basic steps of	144–146–160–161
Bipartite equivalents of	144–147
Column coordinate	146
Even-length periods	161
General principles underlying solution	146–150, 160–165
Illustration of solution	150–160, 166–181
Matching of distributions	170–173
Odd-length periods	161
Periods of fixed length	160
Preparation of index	151, 166
Probability of occurrence of even-positioned letters	170–171
Probability of occurrence of odd-positioned letters	170–171
Probable-word method of attack	174–176
Reconstruction of original substitution checkerboard	152–154, 175–178
Row coordinate	146
Security of	181, 183
Solution of	146–183
Special solution of	156–160, 181–182
Vertical pairs of components	144–145
Bipartite equivalents	144, 162
Blanks (in matrix)	31, 184
"Breaks"	35–36
Break table	36, 100
"Cage"	2, 4
Cipher→Plain sequence	32, 55–75

REF ID:A59439

	Page
Columnar transposition ciphers	3–79
Completely-filled rectangles	4–17
Column and row transposition	17
Consonants and vowels, deviation of	6–10
Invariable digraph	14
Keyword reconstruction diagram	15
Limited affinity	14
Matrix	2
Matrix reconstruction	80–84
Obligatory sequences	14
Pilot letters	14
Probable-word method of solution	13–14, 37–39
Reconstruction of literal key	15–17, 25
General solution	18, 51–53
Incompletely-filled rectangles	18–36
Alternative method of solution	25–31
Formula for calculating length and number of long and short columns	18
General principles underlying solution	18–24
General solution	18, 51–53
Keyword reconstruction diagram	15, 25
Long columns of	18
Short columns of	18
Special solution of	37–38, 40–55
Width	18
Special solutions	37–38, 40–55
Cryptograms of identical length in same key	51–53
Interchanged pair of columns	43–44
Messages with similar beginnings	44–47
Messages with similar endings	47–49
Omitted column	42
Single message containing a long repetition	49–50
Stereotyped phraseology	37–39
Combined substitution-transposition systems	94–96
Using digraphic substitution	96
Using fractionating systems	96
Using known alphabets	94–95
Using monoalphabetic substitution	94–95
Using polyalphabetic or polygraphic substitution	96
Completely-filled rectangles	4
C→P sequence	32, 55–75
"Crown" diagram	20–21
Cyclic permutation of transposition key	35, 180
Double transposition ciphers	51–79
Depth of rectangle a multiple of width	78–79
Enciphering rectangle a perfect square	76
Failure to execute double transposition properly	75–76
Reconstructing keys	55–75
Special cases of solution	75–79
Width of rectangle a multiple of depth	76–78
Encipher sequence	33
Exact-factor method of solving ADFGVX cipher	123–124

REF ID: A594382

	Page
Fractionating systems	97–184
ADFGVX	97–143
Bifid	144–182
Trifid	182–183
"Frame"	2, 4
General solution	32, 37–38, 51–53, 124–125
Geometric designs	80–84
Grilles, indefinite or continuous	91–93
Grilles, revolving	85–91
Alpha method	85–91
Beta method	85
Principle of exclusion	89
Principle of sequence	89
Principle of symmetry	87
"Hat" diagram	20–21
Inscription	2
Interchanged pair of columns	43–44
Interval sequence	58
Invariable digraph	14
Invariant relationship	69
Inverse sequence	32–33
Keyword reconstruction diagram	15
$\kappa\rho$ sequence	33
Limited affinity	14
Literal key, reconstruction of	15–17, 25
Logarithms of probabilities, use of	6, 12–13, 143
Matching distributions	170–173
Matrix	2, 4
Matrix reconstruction	80–84
Monophase transposition	2
Nulls (in matrix)	1, 31, 184
Obligatory sequences	14
Omitted column	42
Partial C→P sequence	75
P→C interval sequence	58
P→C sequence	32, 55–75
Pilot letters	14
Polyphase transposition	2
Processes, rescriptive	2
Plain→Cipher sequence	32, 55–75
Rail-fence writing	3
Rescription, process of	2
Reversed writing	3, 95
Route transposition	3–4

REF ID:A59439

	Page
Single transposition	2
Solution by superimposition	51–53
Special solution	37–38, 40–55, 75–79
Superimposition, solution by	51–53
Term number	34
Transcription	2
Transposition:	
Columnar	4
Double	55–79
Monophase	2
Polyphase	2
Sequence	33
Simple types of	3–17
Single	2
Uniliteral route	3–4
Vertical writing	3
Trifid fractionating system, solution of	182–183
Uniliteral transposition	4
Vertical writing	3
$\kappa\rho$ sequence	33

324155—41——18