~~CONFIDENTIAL~~

312

~~CONFIDENTIAL~~

312

record taken from WFF's home

~~RESTRICTED~~ ~~CONFIDENTIAL~~

MILITARY CRYPTANALYSIS

PART II

SIMPLER VARIETIES OF POLYALPHABETIC SUBSTITUTION SYSTEMS

by

WILLIAM F. FRIEDMAN

Principal Cryptanalyst

Prepared under the direction of the Chief Signal Officer.

1937

~~CONFIDENTIAL~~

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
Paul S. Willard
Colonel, AGC
Adjutant General

REF ID:A4146445

~~CONFIDENTIAL~~

MILITARY CRYPTANALYSIS. PART II

Simpler Varieties of Polyalphabetic Substitution Systems

ERRATA

Page	Paragraph	Line	Now reads	Correction
1	1	3	thier	their
13		4	2, 3, 4, 8, 12	2, 3, 4, 6, 8, 12
17	Table	3	Interval Factors 315 3,5,7,9	Interval Factors 215 5
43		last	Fig. 19 (continued)	Fig. 19
55	27	2	Paragraph 3	Paragraph 6
56	28b	last	Cipher	(2) Cipher
58	31a	2	27 b	28 b
59	31b	2	27 b	28 b
61	31k	5		Delete one CD digraph
61	31k	last		Delete AB after questionably.
62	31 1	6	dispite	despite
62	31 1	12 & 13		Delete AB after questionably
65	31q (2)	7	IM	MI
66	Col. 22	4	T	Y
67	32b	2		Delete principles of the
68	Col. 2	CC	G X K W D V B L S E	G C K W D V B L S E
69	Col. S, Table II		GW	GU
69	Col. Z, " II		RH	RX
70	Col. D, " V		VT	YT
70	Col. E, " V		HA	HQ
72	Col. 2	CC	G X K W D V B L S E	G C K W D V B L S E
73	"	"	"	"
74	"	"	"	"
75	"	"	"	"
85	36	14	let	led
86	Fig. 35			Blank under H should be Q
87	36d (4)	3	First O	P
90	Fig. 41	group 6	T R O R O R E	T H O R O R E
90	38	3	Paragraph 3	Paragraph 6
91	39	2	Paragraph 3	Paragraph 6
92	Fig. 44, Col.1,	1	O Q W W Q W M I O P	O Q E W Q W M I O P
96	Group 2,	7	V N X J K	V V X J K
97	44c	7	First D	F
102	45b, group 3,	5	E X T Z L	E Z T Z L

LESSON SHEETS

Lesson	Problem	Line	Group	Page	Now Reads	Correction
2	1	5	2		W D B X N	W D B S N
2	1	4	3	2	Y B K A O	Y B K V O

SECTION I

INTRODUCTORY REMARKS

1. The essential difference between monoalphabetic and polyalphabetic substitution. - a. In the substitution methods thus far discussed it has been pointed out that thier basic feature is that of monoalphabeticity. From the cryptanalytic standpoint, neither the nature of the cipher symbols, nor their method of production is an essential feature, although these may be differentiating characteristics from the cryptographic standpoint. It is true that in those cases designated as monoalphabetic substitution with variants or multiple equivalents, there is a departure, more or less considerable, from strict monoalphabeticity. In some of those cases, indeed, there may be available two or more wholly independent sets of equivalents, which, moreover, may even be arranged in the form of completely separate alphabets. Thus, while a loose terminology might permit one to designate such systems as polyalphabetic, it is better to reserve this nomenclature for those cases wherein polyalphabeticity is the essence of the method, specifically introduced with the purpose of imparting a positional variation in the substitutive equivalents for plain-text letters, in accordance with some rule directly or indirectly connected with the absolute positions the plain-text letters occupy in the message. This point calls for amplification.

b. In monoalphabetic substitution with variants the object of having different or multiple equivalents is to suppress, so far as possible by simple methods, the characteristic frequencies of the letters occurring in plain text. As has been noted, it is by means of these characteristic frequencies that the cipher equivalents can usually be identified. In these systems the varying equivalents for plain-text letters are subject to the free choice and caprice of the enciphering clerk; if he is careful and conscientious in the work, he will really make use of all the different equivalents afforded by the system; but if he is slip-shod and hurried in his work, he will use the same equivalents repeatedly rather than take pains and time to refer to the charts, tables, or diagrams to find the variants. Moreover, and this is a crucial point, even if the individual enciphering clerks are extremely careful, when many of them employ the same system it is entirely impossible to insure a complete diversity in the encipherments produced by two or more clerks working at different message centers. The result is inevitably to produce plenty of repetitions in the texts emanating from several stations, and when texts such as these are all available for study they are open to solution, by a comparison of their similarities and differences.

c. In true polyalphabetic systems, on the other hand, there is established a rather definite procedure which automatically determines the shifts or changes in equivalents or in the manner in which they are introduced, so that these changes are beyond the momentary whim or choice of the enciphering clerk. When the method of shifting or changing the equivalents is scientifically sound and sufficiently complex the research necessary to establish the values of the cipher characters is much more prolonged and difficult than is the case even in complicated monoalphabetic substitution with variants, as will later be seen. These are the objects of true polyalphabetic substitution systems. The number of such systems is quite large, and it will be possible to describe in detail the cryptanalysis of only a few of the more common or typical examples of methods encountered in practical military cryptanalysis.

d. The three methods, (1) mono-equivalent monoalphabetic substitution, (2) monoalphabetic substitution with variants, and (3) true polyalphabetic substitution show the following relationships as regards the equivalency between plain-text and cipher-text units:

A. In method (1), there is a set of 26 symbols; a plain-text letter is always represented by one and only one of these symbols; conversely, a symbol always represents the same plain-text letter. The equivalence between the plain-text and the cipher letters is constant in both encipherment and decipherment.

B. In method (2), there is a set of 26+n symbols, where n may be any number; a plain-text letter may be represented by 1, 2, 3, ... different symbols; conversely, a symbol always represents the same plain-text letter, the same as is the case in method (1). The equivalence between the plain-text and the cipher letters is variable in encipherment but constant in decipherment.[1]

C. In method (3), there is, as in the first method a set of 26 symbols; a plain-text letter may be represented by 1, 2, 3, ... 26 different symbols; conversely, a symbol may represent 1, 2, 3, ... 26 different plain text letters, depending upon the system and the specific key. The equivalence between the plain-text and the cipher letters is variable in both encipherment and decipherment.

2. Primary classification of polyalphabetic systems. - a. A primary classification of polyalphabetic systems into two rather distinct types may be made: (1) periodic systems and (2) aperiodic systems. When the enciphering process involves a cryptographic treatment which is repetitive in character, and which results in the production of cyclic phenomena in the cryptographic text, the system is termed periodic. When the enciphering process is not of the type described in the foregoing general terms, the system is termed aperiodic. The substitution in both cases involves the use of two or more cipher alphabets.

b. The cyclic phenomena inherent in a periodic system may be exhibited externally, in which case they are said to be patent, or they may not be exhibited externally, and must be uncovered by a preliminary step in the analysis, in which case they are said to be latent. The

---

[1] There is a monoalphabetic method in which the inverse result obtains, the correspondence being constant in encipherment but variable in decipherment; this is a method not found in the usual books on cryptography but in an essay on that subject by Edgar Allan Poe, entitled, in some editions of his works, "A few words on secret writing" and, in other editions, "Cryptography". The method is to draw up an enciphering alphabet such as the following (using Poe's example):

Plain -	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
Cipher -	S	U	A	V	I	T	E	R	I	N	M	O	D	O	F	O	R	T	I	T	E	R	I	N	R	E

In such an alphabet, because of repetitions in the cipher component, the plain-text equivalents are subject to a considerable degree of variability, as will be seen in the deciphering alphabet:

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

This type of variability gives rise to ambiguities in decipherment. A cipher group such as TIE$_c$ would yield such plain-text sequences as REG, FIG, TEU, REU, etc., which could be read only by context. No system of such a character would be practical for serious usage.

periodicity may be quite definite in nature, and therefore determinable with mathematical exactitude allowing for no variability, in which case the periodicity is said to be <u>fixed</u>. In other instances the periodicity is more or less flexible in character and even though it may be determinable mathematically, allowance must be made for a degree of variability subject to limits controlled by the specific system under investigation. The periodicity is in this case said to be <u>flexible</u>, or <u>variable within limits</u>.

3. Primary classification of periodic systems. - a. Periodic polyalphabetic substitution systems may primarily be classified into two kinds:

(1) Those in which only a few of a whole set of cipher alphabets are used in enciphering individual messages, these alphabets being employed repeatedly in a fixed sequence throughout each message. Because it is usual to employ a secret word, phrase, or number as a key to determine the number, identity, and sequence with which the cipher alphabets are employed, and this key is used over and over again in encipherment, this method is often called the <u>repeating-alphabet system</u>. It is also sometimes referred to as the <u>multiple-alphabet system</u> because if the keying of the entire message be considered as a whole it is composed of multiples of a short key used repetitively.[1] In this text the designation "repeating-key system" will be used.

(2) Those in which all the cipher alphabets comprising the complete set for the system are employed one after the other progressively in the encipherment of a message, and when the last alphabet of the series has been used, the encipherer begins over again with the first alphabet. This is commonly referred to as a <u>progressive-alphabet system</u> because the cipher alphabets are used in progression.

4. Sequence of study of polyalphabetic systems. - a. In the studies to be followed in connection with polyalphabetic systems, the order in which the work will proceed conforms very closely to the classifications made in paragraphs 2 and 3. Periodic polyalphabetic substitution ciphers will come first, because they are, as a rule, the simpler and because a thorough understanding of the principles of their analysis is prerequisite to a comprehension of how aperiodic systems are solved. But in the final analysis the solution of examples of both types rests upon the conversion or reduction of polyalphabeticity into monoalphabeticity. If this is possible, solution can always be achieved, granted there are sufficient data in

---

[1] French terminology calls this the "double-key method", but there is no logic in such nomenclature.

the final monoalphabetic distributions to permit of solution by recourse to the ordinary principles of frequency.

b. First in the order of study of periodic systems will come the analysis of repeating-key systems. Some of the more simple varieties will be discussed in detail, with examples. Subsequently, ciphers of the progressive type will be discussed. There will then follow a more or less detailed treatment of aperiodic systems.

SECTION II

CIPHER ALPHABETS FOR POLYALPHABETIC SUBSTITUTION

5. Classification of cipher alphabets upon the basis of their derivation. - a. The substitution processes in polyalphabetic methods involve the use of a plurality of cipher alphabets. The latter may be derived by various schemes, the exact nature of which determines the principal characteristics of the cipher alphabets and plays a very important role in the preparation and solution of polyalphabetic cryptograms. For these reasons it is advisable, before proceeding to a discussion of the principles and methods of analysis, to point out these various types of cipher alphabets, show how they are produced, and how the method of their production or derivation may be made to yield important clues and short-cuts in analysis.

b. A primary classification of cipher alphabets for polyalphabetic substitution may be made into the two following types:

(1) Independent or unrelated cipher alphabets.

(2) Derived or interrelated cipher alphabets.

c. Independent cipher alphabets may be disposed of in a very few words. They are merely separate and distinct alphabets showing no relationship to one another in any way. They may be compiled by the various methods discussed in Pars. 44 - 48, inclusive, Section IX of Special Text No. 165, Elementary Military Cryptography. The solution of cryptograms written by means of such alphabets is rendered more difficult by reason of the absence of any relationship between the equivalents of one cipher alphabet and those of any of the other alphabets of the same cryptogram. On the other hand, from the point of view of practicability in their production and their handling in cryptographing and decryptographing, they present some difficulties which make them less favored by cryptographers than cipher alphabets of the second type.

d. Derived or interrelated alphabets, as their name indicates, are most commonly produced by the interaction of two primary components,[1] which when juxtaposed at the various points of coincidence can be made to yield secondary alphabets.[2]

6. Primary components and secondary alphabets. - Two basic, slidable sequences or components of *n* characters each will yield *n* secondary alphabets. The components may be classified according to various schemes. For cryptanalytic purposes the following classification will be found useful:

CASE A. The primary components are both normal sequences.

- (1) The sequences proceed in the same direction. (The secondary alphabets are direct standard alphabets.)
- (2) The sequences proceed in opposite directions. (The secondary alphabets are reversed standard alphabets and are reciprocal.)

CASE B. The primary components are not both normal sequences.

- (1) The plain component is normal, the cipher component is a mixed sequence. (The secondary alphabets are mixed alphabets.)
- (2) The plain component is a mixed sequence, the cipher component is normal. (The secondary alphabets are mixed alphabets.)
- (3) Both components are mixed sequences.
  - (a) Components are identical mixed sequences.
    - I. Sequences proceed in the same direction. (The secondary alphabets are mixed alphabets.) (Par. 23)

---

[1] See Par. 37, Special Text No. 165.

[2] See Pars. 49 and 59, Special Text No. 165.

II. Sequences proceed in opposite directions. (The secondary alphabets are reciprocal mixed alphabets.) (Par. 38)

(b) Components are different mixed sequences. (The secondary alphabets are mixed alphabets.) (Par. 39)

7. Cipher disks and cipher squares. - a. Reference is now made to Pars. 60 - 62, Section XII, Special Text No. 165, wherein was shown the equivalency that subsists between the results produced by sliding primary components and cipher disks and square tables of the Vigenere type. In all cases the results produced by the successive juxtapositions of two sliding components may be duplicated by using a cipher square; the converse relationship is true only when the columns or rows of the cipher square show symmetry; that is, the sequences in the columns or rows are identical but merely displaced 1, 2, 3, ... intervals successively.

b. In cryptanalytic studies it is usually more convenient and useful, wherever possible, to consider the problem from the point of view of sliding components rather than cipher squares.

SECTION III

THEORY OF SOLUTION OF REPEATING-KEY SYSTEMS

8. The three steps in the analysis of repeating-key systems. - a. The method of enciphering according to the principle of the repeating-key, or repeating alphabets is adequately explained in Pars. 57 and 58 of Special Text No. 165, Elementary Military Cryptography, and no further reference need be made at this time. The analysis of a cryptogram of this type, regardless of the kind of cipher alphabets employed, or their method of production, resolves itself into three distinct and successive steps.

(1) Determination of the length of the repeating key, which is the same as the determination of the exact number of alphabets involved in the cryptogram;

(2) Allocation or distribution of the letters of the cipher text into the respective cipher alphabets to which they belong, which reduces the polyalphabetic text to monoalphabetic terms;

(3) Analysis of the individual monoalphabetic distributions to determine plain-text values of the cipher letters in each distribution or alphabet.

b. The foregoing steps will be treated in the order in which mentioned. The first step may be described briefly as that of determining the period. The second step may be described briefly as that of reduction to monoalphabetic terms. The third step may be designated as identification of cipher-text values.

9. First step: finding the length of the period. - a. The determination of the period, that is, the length of the key or the number of cipher alphabets involved in a cryptogram enciphered by the repeating-key method is, as a rule, a relatively simple matter. The cryptogram itself usually manifests externally certain phenomena which are the direct result of the use of a repeating key. The principles involved are, however, so fundamental in cryptanalysis that their elucidation warrants a somewhat detailed treatment. This will be done in connection with a short example of encipherment, shown below in Fig. 1.

b. Regardless of what system is used, identical plain-text letters enciphered by the same cipher alphabet[1] must yield identical cipher letters. Referring to Fig. 1, such a condition is brought about every time that identical plain-text letters happen to be enciphered with the same key-letter, or every time identical plain-text letters fall into the same column in the encipherment.[2] Now since the number of columns or positions with respect to the key is very limited (except in the case of very long key words), and since the repetition of letters is an inevitable condition in plain text, it follows that there will be in a message of fair length many cases where identical plain-text letters **must** fall into the same column. They will thus be enciphered by the same cipher alphabet, resulting, therefore, in the production of many identical letters in the cipher text. When identical plain-text polygraphs fall into identical

---

[1] It is to be understood, of course, that cipher alphabets with single equivalents are meant in this case.

---

[2] The frequency with which this condition may be expected to occur can be definitely calculated. A discussion of this point falls beyond the scope of the present text.

MESSAGE

THE ARTILLERY BATTALION MARCHING IN THE REAR OF THE ADVANCE GUARD KEEPS ITS COMBAT TRAIN WITH IT INSOFAR AS PRACTICABLE.

(Key: BLUE, using direct standard alphabets)

CIPHER ALPHABETS

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

BLUE	BLUE	BLUE	BLUE
THEA	ARDK	THEA USYE	ARDK BCXO
RTIL	EEPS	RTIL SECP	EEPS FPJW
LERY	ITSC	LERY MPLC	ITSC JEMG
BATT	OMBA	BATT CLNX	OMBA PXVE
ALIO	TTRA	ALIO BWCS	TTRA UELE
NMAR	INWI	NMAR OXUV	INWI JYQM
CHIN	THIT	CHIN DSCR	THIT USCX
GINT	INSO	GINT HTHX	INSO JYMS
HERE	FARA	HERE IPLI	FARA GLLE
AROF	SPRA	AROF BCIJ	SPRA TALE
THEA	CTIC	THEA USYE	CTIC DECG
DVAN	ABLE	DVAN EGUR	ABLE BMFI
CEGU		CEGU DPAY	
a	a	b	b

CRYPTOGRAM

USYES ECPMP LCCLN XBWCS OXUVD SCRHT
HXIPL IBCIJ USYEE GURDP AYBCX OFPJW
JEMGP XVEUE LEJYQ MUSCX JYMSG LLETA
LEDEC GBMFI

Figure 1.

columns the result is the formation of identical cipher-text polygraphs, that is, repetitions of groups of 2, 3, 4, . . . letters are exhibited in the cryptogram. Repetitions of this type will hereafter be called causal repetitions, because they are produced by a definite, traceable cause, viz., the encipherment of identical letters by the same cipher alphabets.

c. It will also happen, however, that different plain-text letters falling in different columns will, by mere accident, produce identical cipher letters. Note, for example, in Fig. 1 that in Column 1, R_p becomes S_c and that in Column 2, H_p also becomes S_c. The production of an identical cipher-text letter in these two cases (that is, a repetition where the plain-text letters are different and enciphered by different alphabets) is merely fortuitous. It is, in every day language, "a mere coincidence", or "an accident." For this reason repetitions of this type will hereafter be called accidental repetitions. Such repetitions will, of course, happen fairly frequently with individual letters, but less frequently with digraphs, because in this case the same kind of an "accident" must take place twice in succession. Intuitively one feels that the chances that such a purely fortuitous coincidence will happen two times in succession must be much less than that it will happen every once in a while in the case of single letters. Similarly, intuition makes one feel that the chances of such accidents happening in the case of three or more consecutive letters are still less than in the case of digraphs, decreasing very rapidly as the repetition increases in length. This phenomenon may, however, be dealt with statistically, thus taking the matter outside the realm of intuition.

d. (1) Imagine a box containing an infinite number of the 26 letters of the alphabet, all in equal proportions, so that there are exactly the same numbers of A's, B's, C's, . . . Z's. The box is thoroughly shaken so that the letters are thoroughly mixed and a single letter is now drawn at random. What are the chances that it is an A? Obviously the chances are 1 in 26. The chances that the letter is B, C, D, . . . Z are also the same. In mathematical language, the probability[1] of drawing any specified letter is $\frac{1}{26}$. Suppose an A has been drawn.

(2) Now suppose that this letter is replaced in the box, the latter again shaken, and a second drawing is made. What are the chances that the second drawing will also be an A? Another way of asking the same question, which will perhaps make it clearer is this: suppose two letters are drawn simultaneously from the box, what is the probability of drawing two A's? Since the probability of occurrence of two events which are independent is the product of the probability of their separate occurrence, the probability of drawing two A's is $\frac{1}{26} \times \frac{1}{26}$, or $\frac{1}{676}$. This is also the probability

---

[1] The definition of probability implies philosophical questions which are beyond this discussion. However, for the purposes of this text, the following definition of a priori probability will be found sufficient. The probability that an event will occur is the ratio of the number of favorable cases to the number of total possible cases, all cases being equally likely to occur, where by a favorable case is meant one which will produce the event in question.

For example he has a message of a certain length and it shows repetitions. Is the number of repetitions higher or lower than ~~what~~ he should expect by pure accident? How ~~many~~ repetitions and of what length should he expect?

of drawing two B's, an A and a B, or any two specified letters in a specified order: that is, the probability of drawing any specified digraph is $\frac{1}{26^2}$. Similarly, the probability of drawing a specified trigraph is $\frac{1}{26^3}$, a specified tetragraph $\frac{1}{26^4}$, and so on. In general, for any specified polygraph the probability is $\frac{1}{26^n}$ where n is the number of letters in the polygraph.

(3) However, in ~~studying the phenomena of~~ cryptanalysis questions concerning the phenomena of repetition, chiefly ~~repetition~~, the student is concerned not so much or so often with questions concerning ~~with~~ the probability of occurrence of a specified single letter, digraph, trigraph, or polygraph, ~~but~~ as with the repetition probability of the (recurrence) of these elements. The problem is now different.

e. Consider the cryptogram in subparagraph b, which contains exactly 100 letters, and assume that these letters constitute a perfectly random assortment; that is, assume that the cryptographic system which produced the cryptogram is of such a nature that the text may be considered to be the same as though one had made 100 drawings, with replacements, of the letters from the box of letters referred to above. What is the probability that a specified single letter will not be repeated in the cryptogram? What is the probability that a specified single letter will be repeated exactly (that is, no more and no less than) 1, 2, 3, . . . times? What is the probability that a specified single letter will appear at least once; that is, including all cases in which that letter will appear 1, 2, 3, 4, . . . times? What is the probability that a specified single letter will appear at least 2 times, at least 3 times, and so on? What are the answers to the same questions as regards digraphs, trigraphs, and longer polygraphs? Another, and possibly more concrete, way of putting these questions is this: In the 100-letter cryptogram being studied, assuming it were perfectly random text, how many letters should not occur at all? How many should occur exactly 1, 2, 3, . . . times? How many should occur at least 1, 2, 3, . . . times? What are the answers to these same questions as regards digraphs and trigraphs? What is the probability that the ~~same~~ digraph, trigraph, tetragraph will appear at least twice in this cryptogram?

f. It may be stated at once that the answers to ~~the latter~~ questions of this kind are by no means easy to find, and a complete discussion would fall quite outside the scope of this text. However, it will be sufficient for the purpose if the mathematics involved are converted into a form that will be of practical use to the student. With this in view Chart 1 has been prepared and its use will now be explained.[1]

---

[1] This chart was constructed from calculations based upon Poisson's exponential expansion, or the "law of small probabilities." Students without a thorough grounding in the mathematical theory of probability and statistics "will have to take the chart on faith." Those interested in ~~its~~ the theory and the derivation of the Poisson distribution are referred to the following texts:

Fisher, R. A., Statistical Methods for Research Workers, London, 1937.
Fry, T. C., Probability and its Engineering Uses, New York, 1928.

What is the prob. that a specified single letter, say A, will not appear at all?

g. (1) Suppose a cryptogram of 100 letters is being studied. Assuming that the 100 letters had been drawn out of the box, so that they constitute a perfectly random assortment of letters, what is the probability that a specified single letter will not appear at all in this assortment? It has been seen that in a perfectly random assortment, the probability for selecting ~~a specified single~~ an A letter is $\frac{1}{26}$, or, in mathematical language,

$P_1 = .0385.$

There being 100 letters in the cryptogram in question, P_1 is to be multiplied by 100, giving 3.85. Referring now to Chart 1, find the point corresponding to the value 3.85 on the x axis of the chart, that is, the horizontal scale at the bottom; select the curve marked zero; find the point where this curve intersects the vertical ordinate corresponding to the value 3.85 on the horizontal scale; follow this point straight over to the left and read the value on the y axis of the chart, that is the vertical scale. It is approximately .021. This means that the probability that a single specified letter will not appear at all in the cryptogram, if it were a perfectly random assortment of letters, is .021. That is, according to the theory of probability, in 1000 cases of random text messages of 100 letters each there should be 21 messages in which a single specified letter will not appear at all. This, of course, is merely a theoretical expectancy; it indicates only what probably will happen in the long run.

The various frequency [illegible] the [illegible] may be regarded as [illegible] the [illegible] two [illegible] times the event frequencies may be expected to recur

(2) What is the probability that a single specified letter will appear exactly once, no more and no less? To answer this question, find the point of intersection of the vertical ordinate corresponding to 3.85 with the curve marked "1". Its value on the vertical scale is 0.082; that is, in 1000 cases of random text messages of 100 letters each the theoretical expectancy is that there will be 82 messages in which a single specified letter will appear exactly once, no more and no less.

(3) In exactly the same way, the probability that a single specified letter will appear exactly twice, is found to be 0.158. That is, the probability that a single specified letter will be repeated exactly once (two occurrences) is .158; the probability that it will be repeated exactly twice (three occurrences) is found to be .202 and so on. The following table gives the probabilities for exact numbers of occurrences from 0 to 10, inclusive:

100 letters of random text

Frequency	Probability that a specified single letter will occur exactly . . . times.	By table III interp.
0	0.021	.0213
1	.082	.0820
2	.15~~5~~8	.1577
3	.20~~0~~2	.2023
4	.195	.1947
5	.150	.1500
6	.096	.0962
7	.053	.0530
8	.02~~5~~6	.0255
9	.011	.0108
10	.00~~3~~4	.0042

which took into account occurrences only up to 10.

The chart enables one to find the probability for occurrences up to 16 and recurrences up to 15, both inclusive, for various numbers of letters in random assortments.

(4) To find the probability that a specified single letter will occur at least 1, 2, 3, . . . times in a series of letters constituting random text, one reasons as follows: Since the concept "at least 1" implies that the number specified is to be considered only as the minimum, with no limit indicated as to maximum, occurrences of 2, 3, 4, . . . are also "favorable" cases; the probabilities for exactly 1, 2, 3, 4, . . . occurrences should therefore be added and this will give the probability for "at least 1." Thus, in the case of 100 letters, the sum of the probabilities for exactly 1 to 10 occurrences, as set forth in the table directly above, is .977, and the latter value approximates the probability for at least 1 occurrence.

(5) A more accurate result will be obtained by the following reasoning. The probability for zero occurrences is .021. Since it is certain that a specified letter will occur either zero times or 1, 2, 3, . . . times, to find the probability for at least one time it is merely necessary to subtract the probability for zero occurrences from 1. That is, 1 - .021 = .979, which is somewhat greater than the result obtained by the other method. The reason it is greater is that the value .979 includes all the occurrences beyond 10, which were excluded from the previous calculation. Of course, the probabilities for these occurrences beyond 10 are very small, but taken all together they add up to .002, the difference between the results obtained by the two methods. The probability for at least 2 occurrences is the difference between unity and the sum of the probability for zero and exactly 1 occurrences; that is, $1 - (P_0 + P_1) = 1 - (.021 + ~~.158~~) = 1 - .179$ $= ~~.821~~$ .897.

.082 = 1 - .103

(6) The foregoing calculations refer to random text containing 100 letters. For other numbers of letters, it is merely necessary to multiply the probability for drawing a single specified letter out of the box, which is $\frac{1}{26}$ or .0385, by the number of letters in the assortment, and refer to the chart. For example, for a random assortment of 200 letters, the product of 200 x .0385 or 7.7 gives the value of the point to be sought along the horizontal or x axis of the chart; the intersections of the vertical line corresponding to this point with the various curves for 0, 1, 2, 3, . . . occurrences give the probabilities for these occurrences, the reading being taken on the vertical or y axis of the chart.

(7) The discussion thus far has dealt with random assortments of letters. What about other types of texts, for example, normal plain text? What is the probability that E will occur 0, 1, 2, 3, . . . times in 50 letters of normal English? The relative frequency value or probability that a letter selected at random from a large volume of normal English text will be E is .12604. For 50 letters this value must be multiplied by 50, giving 6.3 as the point along the x axis of the chart. The probabilities for 0, 1, 2, 3, . . . occurrences are tabulated below:

Frequency	Probability that an "E" will be drawn exactly . . . times	Probability that an "E" will be drawn at least . . . times
0	0.002	1.000
1	.011	.998
2	.036	.987
3	.076	.951
4	.120	.875
5	.151	.755
6	.159	.604
7	.143	.445
8	.113	.302
9	.079	.223
10	.050	.173
11	.028	.145
12	.015	.130
13	.007	.123
14	.003	.120
15	.001	.119
16	.000	.119

h. (1) The discussion thus far has dealt with the probabilities for 0, 1, 2, 3, . . . occurrences of specified single letters. It may be of more practical advantage to the student if he could be shown how to find the answer to these questions: Given a random assortment of 100 letters how many letters may be expected to occur exactly 0, 1, 2, 3, . . . times? How many may be expected to occur at least 1, 2, 3, . . . times? Chart 1 may also be used to answer these questions, by a very simple calculation: multiply the probability value as obtained above for a specified single letter by the number of different elements being considered. For example, the probability that a specified single letter will occur exactly twice in a perfectly random assortment of 100 letters is .15~~5~~; since the number of different letters is 26, the absolute number of single letters that may be expected to occur exactly 2 times in this assortment is .15~~5~~ x 26 = ~~4.03~~. That is, in 100 letters of random text there should be about four letters which occur exactly 2 times. The following table gives the data for various numbers of occurrences:

8

4.108

158
26
948
316
4108

What is the prob that a specified will not occur? What is the prob that a spec. dig will appear exactly one time, etc.

100 letters of random text

Frequency	Probability that a specified single letter will occur exactly . . . times	Probability that a specified single letter will occur at least . . . times	Expected number of letters appearing exactly . . . times	Expected number of letters appearing at least . . . times
0	0.020	1.000	0.52	26.00
1	.080	.980	2.08	25.48
2	.157	.900	4.09	23.39
3	.206	.743	5.37	19.30
4	.197	.537	5.07	13.93
5	.152	.340	3.97	8.86
6	.098	.188	2.55	4.89
7	.051	.090	1.33	2.34
8	.026	.039	.68	1.01
9	.010	.023	.26	.33
10	.003	.003	.08	.08

1.000
.02
.98

(2) Thus far the discussion has been restricted to single letters, but the chart may also be used for calculations referring to digraphs, trigraphs, and longer polygraphs. The method of using the chart is exactly the same as before, but the points selected on the x axis are now determined by the value of the probability of selecting a specified pair of letters or a set of 3 or more letters from the box of letters referred to above.

(3) Taking up the case of digraphs, and assuming the box now to contain an unlimited number of all 676 digraphs in equal proportions, the probability of selecting a specified digraph from the box is $\frac{1}{676}$ = .00148. Given a random assortment of 100 digraphs ~~letters~~, the value along the x axis is now 100 x .00148 = .148. The following values are obtained from the chart:

*

100 Digraphs

Frequency	Probability for exact number of . . . occurrences of digraphs	Probability for at least . . . occurrences of a specified digraphs	Number of digraphs expected to occur exactly . . . times ~~(approximately)~~	Number of digraphs expected to occur at least . . . times ~~(approximately)~~
0	0.86	1.00	581 580	676
1	.13	~~.15~~ 14	88. 88	95 93
2	.01	.02 01	7 7	7
3	.001	.01 00	~~1~~ 1	0
4	.000	.009	0 0	0

576

(4) The student may have some good practice by making the same sort of calculations for trigraphs. ~~to be expected in 100, 1,000, and 10,000 letters.~~

676
.14
2704
676
9464

(5) Referring again to Chart 1, and specifically to the tabulated results set forth under subparagraph g (4) above, it will be seen that the probability that there will be exactly one repetition of a specified single letter in 100 letters of random text is less than the probability that there will be exactly two repetitions; in other words, the chances that a letter will be repeated exactly twice are better by about 25% than that it will be repeated only once. If this sounds absurd to the student, let him cogitate upon the implications of the word "exactly" in the foregoing statements and the reasoning on which the whole argument is based. He will find assistance from studying the shape of the various curves in Chart 1, especially those for 1, 2, and 3 occurrences, wherein the curves approximate the shape of the bell-shaped normal probability curve.

i. (1) The message in subparagraph a is now to be studied from the viewpoint of the number of repetitions it contains as compared with the number theoretically to be expected. First, the repetitions of two or more letters are underscored.

A. USYES ECPMP LCCLN XBWCS OXUVD
B. SCRHT HXIPL IBCIJ USYEE GURDP
C. AYBCX OFPJW JEMGP XVEUE LEJYQ
D. MUSCX JYMSG LLETA LEDEC GBMFI

(2) Here are the repetitions, listed for convenience:

Group	Number of occurrences
BC	2
CX	2
EC	2
LE	3
JY	2
PL	2
SC	2
SY	2
US	3
YE	2
USY	2
USYE	2

While the number of digraphs appearing twice (8) is not much more than expected (7); the number appearing three times (2) is twice as great as expectancy (1).

digraphs drawn at random

has 99 digraphs, or practically 100. Yet within these 100 digraphs there are 8

j. Referring to the table set forth under subparagraph h (3), it will be seen that in 100 ~~letters of random text~~ the expectancy is that 7 digraphs will appear 2 times, and 1 will appear 3 times. The message being studied ~~has 8 di~~graphs occurring 2 times, and 2 digraphs occurring 3 times. ~~In~~ other words, the numbers of digraphs that occur 2 and 3 times in the message are greater than expected if the message were random text.

k. ~~Moreover,~~ The message contains a trigraphic ~~tetragraphic~~ repetition. (USY) Chart 1 shows that the probability of 2 occurrences of a specified tri~~tetra~~graph in 100 ~~letters is~~ approximately .005; that is, this may be expected to happen only about 5 times in a thousand. Yet it has happened here.

tetragraphs drawn at random

The message has 95 trigraphs ...

l. A consideration of the facts set forth in the subparagraphs i-k leads to but one conclusion, viz, that the repetitions exhibited by the cryptogram under investigation are not accidental but are causal in their origin; and the cause is in this case not difficult to find: repeated letters in the plain text were actually enciphered by identical alphabets. In order for this to occur, it was necessary that the tetragraph USYE, for example, fall both times in exactly the same relative position with respect to the key. Note, for example, that USYE in Fig. 1 represents in both cases the plain-text polygraph THEA. The first time it occurred it fell in positions 1-2-3-4 with respect to the key; the second time it occurred it happened to fall in the very same relative positions, although it might just as well have happened to fall in any of the other three possible relative positions with respect to the key, viz, 2-3-4-1, 3-4-1-2, or 4-1-2-3.

m. Lest the student be misled, however, a few more words are necessary on this subject. In the preceding subparagraph the word "happened" was used; this word correctly expresses the idea in mind, because the insertion or deletion of a single plain-text letter between the two occurrences would have thrown the second occurrence one letter forward or backward, respectively, and thus caused the polygraph to be enciphered by a sequence of alphabets such as can no longer produce the cipher polygraph USYE from the plain-text polygraph THEA. On the other hand, the insertion or deletion of this one letter might bring the letters of some other polygraph into similar columns so that some other repetition would be exhibited in case the USYE repetition had thus been suppressed.

n. The encipherment of similar letters by similar cipher alphabets is therefore the cause of the production of repetitions in the cipher text in the case of repeating-key ciphers. What principles can be derived from this fact, and how can they be employed in the solution of cryptograms of this type?

o. If a count is made of the number of letters from and including the first USYE to, but not including, the second occurrence of USYE, a total of 40 letters is found to intervene between the two occurrences. This number, 40, must, of course, be an exact multiple of the length of the key. Having the plain-text before one, it is easily seen that it is the 10th multiple; that is, the 4-letter key has repeated itself 10 times between the first and the second occurrence of USYE. It follows, therefore, that if the length of the key were not known, the number 40 could safely be taken to be an exact multiple of the length of the key; in other words, one of the factors of the number 40 would be equal to the length of the key. The word "safely" is used in the preceding sentence to mean that the interval 40 applies to a repetition of 4 letters and it has been shown that the chances that this repetition is accidental are small. The factors of 40 are 2, 4, 5, 8, 10, and 20. So far as this single repetition of USYE is concerned, if the length of the key were not known, all that could be said about the latter would be that it is equal to one of these factors. The repetition by itself gives no further indications. How can the exact factor be selected from among a list of several possible factors?

p. Let the intervals between all the repetitions in the cryptogram be listed. They are as follows:

Repetition	Interval	Factors
1st USYE to 2d USYE	40	2, 4, 5, 8, 10, 20.
1st BC to 2d BC	16	2, 4, 8.
1st CX to 2d CX	25	5.
1st EC to 2d EC	88	2, 4, 11, 22, 44
1st LE to 2d LE	16	2, 4, 8.
2d LE to 3d LE	4	2, 4.
1st LE to 3d LE	20	2, 4, 5, 10.
1st JY to 2d JY	8	2, 4.
1st PL to 2d PL	24	2, 3, 4, 6, 8, 10,12.
1st SC to 2d SC	52	2, 4, 13, 26.
(1st SY to 2d SY, already included in USYE.)		
(1st US to 2d US, already included in USYE.)		
2d US to 3d US	36	2, 3, 4, 6, 9, 18.
(1st US to 3d US, already included in USYE.)		
(1st YE to 2d YE, already included in USYE.)		

q. Are all these repetitions <u>causal</u> repetitions? It has been seen that the odds against a theory that the USYE repetition is accidental are about 995 to 5 (since the probability for its occurrence is .005), or 199 to 1. It has also been seen that the odds against a theory that the eight digraphs which occur twice are accidental repetitions are about 99 to 1 (since the probability for 2 occurrences of a specified digraph is .01); the odds against a theory that the two digraphs which occur 3 times are accidental repetitions are 999 to 1. The chances are very great, therefore, that all or nearly all these repetitions are causal. Certainly the chances against the two occurrences of the tetragraph USYE and the three occurrences of the two different digraphs (LE and US) being accidental are quite high, and it is therefore not astonishing that the intervals between all the various repetitions, except in one case, contain the factors 2 and 4.

r. This means that if the cipher is written out in either 2 columns or 4 columns, all these repetitions (except the CX repetition) would fall into the same columns. From this it follows that the length of the key is either 2 or 4, the latter, on practical grounds, being more probable than the former. Doubts concerning the matter of choosing between a 2-letter and a 4-letter key will be dissolved when the cipher text is distributed into its component uniliteral frequency distributions.

s. The repeated digraph CX in the foregoing message is an accidental repetition, as will be apparent by referring to Fig. 1. Had the message been longer there would have been more such accidental repetitions, but, on the other hand, there would be a proportionately greater number of causal repetitions. This is because the phenomenon of repetition in plain text is so all-pervading.

t. Sometimes it happens that the cryptanalyst quickly notes a repetition of a polygraph of four or more letters, the interval between the first and second occurrences of which has only two factors, of which one is a relatively small number, the other a relatively high incommensurable number. He may therefore assume at once that the length of the key is equal to the smaller factor without searching for additional recurrences upon which to corroborate his assumption. Suppose, for example, that in a relatively short cryptogram the interval between the first and second occurrences of a polygraph of five letters happens to be a number such as 203, the factors of which are 7 and 29. Evidently the number of alphabets may at once be assumed to be 7, unless one is dealing with messages exchanged among correspondents known to use long keys. In the latter case one could assume the number of alphabets to be 29.

u. The foregoing method of determining the period in a polyalphabetic cipher is commonly referred to in the literature as "factoring the intervals between repetitions"; or more often it is simply called "factoring." Because the latter is an apt term and is brief, it will be employed hereafter in this text to designate the process.

10. General remarks on factoring. - a. The statement made in Par. 2 with respect to the cyclic phenomena said to be exhibited in cryptograms of the periodic type now becomes clear. The use of a short repeating key produces a periodicity of recurrences or repetitions collectively termed "cyclic phenomena", an analysis of which leads to a determination of the length of the period or cycle, and this gives the length of the key. Only in the case of relatively short cryptograms enciphered by a relatively long key does factoring fail to lead to the correct determination of the number of cipher alphabets in a repeating-key cipher; and of course, the fact that a cryptogram contains repetitions whose factors show constancy is in itself an indication and test of its periodic nature. It also follows that if the cryptogram is not a repeating-key cipher, then factoring will show no definite results, and conversely the fact that it does not yield definite results at once indicates that the cryptogram is not a periodic, repeating-key cipher.

b. There are two cases in which factoring leads to no definite results. One is in the case of monoalphabetic substitution ciphers. Here recurrences are very plentiful as a rule, and the intervals separating these recurrences may be factored, but the factors will show no constancy; there will be several factors common to many or most of the recurrences. This in itself is an indication of a monoalphabetic substitution cipher, if the very fact of the presence of many recurrences fails to impress itself upon the inexperienced cryptanalyst. The other case in which the process of factoring is nonsignificant involves certain types of nonperiodic, polyalphabetic ciphers. In certain of these ciphers recurrences of digraphs, trigraphs, and even polygraphs may be plentiful in a long message, but the intervals between such recurrences bear no definite multiple relation to the length of the key, such as in the case of the true periodic, repeating-key cipher, in which the alphabets change with successive letters and repeat themselves over and over again.

c. Factoring is not the only method of determining the length of the period of a periodic, polyalphabetic substitution cipher, although it is by far the most common and easily applied. At this point it will merely be noted that when the message under study is relatively short in comparison with the length of the key, so that there are only a few cycles of cipher text and no long repetitions affording a basis for factoring, there are several other methods available. However, it being deemed inadvisable to interject the data concerning those other methods at this point, they will be explained subsequently. It is desirable at this juncture merely to indicate that methods other than factoring do exist and are used in practical work.

11. Second step: distributing the cipher text into the component monoalphabets. - a. After the number of cipher alphabets involved in the cryptogram has been ascertained, the next step is to rewrite the message in groups corresponding to the length of the key, or in columnar fashion, whichever is more convenient, and this automatically divides up the text so that the letters belonging to the same cipher alphabet occupy similar positions in the groups, or, if the columnar method is used, fall in the same column. The letters are thus allocated or distributed into the respective cipher alphabets to which they belong. This reduces the polyalphabetic text to monoalphabetic terms.

b. Then separate monoliteral frequency distributions for the thus isolated individual alphabets are compiled. For example, in the case of the cipher on page 9, having determined that four alphabets are involved, and having rewritten the message in four columns, a frequency distribution is made of the letters in Column 1, another is made of the letters in Column 2, and so on for the rest of the columns. Each of the resulting distributions is therefore a monoalphabetic frequency distribution. If these distributions do not give the irregular crest and trough appearance of single frequency distributions, then the analysis which led to the hypothesis as regards the number of alphabets involved is fallacious. In fact, the appearance of these individual distributions may be considered to be an index of the correctness of the factoring process; for theoretically, and practically, the individual distributions constructed upon the correct hypothesis will tend to conform more closely to the irregular crest and trough appearance of a single alphabet frequency distribution than will the graphic tables constructed upon an incorrect hypothesis.

12. Third step: solving the monoalphabetic distributions. The difficulty experienced in analyzing the individual or isolated frequency distributions depends mostly upon the type of cipher alphabets that is used. It is apparent that mixed alphabets may be used just as easily as standard alphabets, and, of course, the cipher letters themselves give no indication as to which is the case. However, just as it was found that in the case of monoalphabetic, substitution ciphers a monoliteral frequency distribution will give clear indications whether the cipher alphabet is a standard or a mixed alphabet, by the relative positions and extensions of the crests and troughs in the table, so it is found that in the case of repeating-key ciphers, monoliteral frequency distributions for the isolated or individual alphabets will also give clear indications as to whether these alphabets are standard alphabets or mixed alphabets. Only one or two such frequency distributions are necessary for this determination; if they appear to be standard alphabets, similar distributions can be made for the rest of the alphabets; but if they appear to be mixed alphabets, then it is best to compile triliteral frequency distributions for all the alphabets. The

analysis of the values of the cipher letters in each table proceeds along the same lines as in the case of monoalphabetic ciphers. The analysis is more difficult only because of the reduced size of the tables, but if the message be very long, then each frequency distribution will contain a sufficient number of elements to enable a speedy solution to be achieved.

SECTION IV

REPEATING-KEY SYSTEMS WITH STANDARD CIPHER ALPHABETS

13. Solution by applying principles of frequency. - a. In the light of the foregoing principles, let the following cryptogram be studied:

MESSAGE

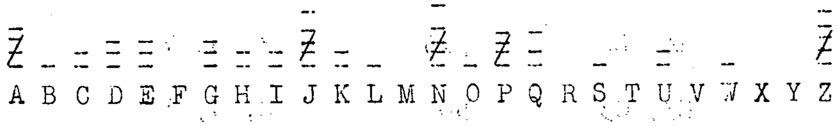
	1	2	3	4	5
A	A U K H Y	J A M K I	Z Y M W M	J M I G X	N F M L X
B	E T I M I	Z H B H R	A Y M Z M	I L V M E	J K U T G
C	D P V X K	Q U K H Q	L H V R M	J A Z N G	G Z V X E
D	N L U F M	P Z J N V	C H U A S	H K Q G K	I P L W P
E	A J Z X I	G U M T V	D P T E J	E C M Y S	Q Y B A V
F	A L A H Y	P O E X W	P V N Y E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

A search for repetitions discloses the following short list of most of the longer repetitions, with the intervals and factors below 11 listed (for previous experience may lead to the conclusion that it is unlikely that the cryptogram involves more than 10 alphabets, showing the number of recurrences which it does):

Repetition	Location	Interval	Factors
LUFMPZJNVC	D1, K3	160	2, 4, 5, 8, 10
JZXIG	E1, H4	90	2, 3, 5, 6, 9, 10
EJK	B4, L2	315	3, 5, 7, 9
PTE	E3, G3	50	2, 5, 10
QGK	D4, H1	85	5
UKH	A1, C2	55	5
ZLA	J1, L4	65	5
AS	D3, L3	175	3, 5, 7
EJ	B4, L2	115	5
FM	A5, D1	57	3
FM	A5, J2	185	5
FM	J2, J4	12	2, 3, 4, 6
FM	J4, K3	20	2, 4, 5, 10
FM	K3, L4	30	2, 3, 5, 6, 10
JA	A2, C4	60	2, 3, 4, 5, 6, 10
LA	F1, J1	75	3, 5
LA	J1, L4	65	5
LL	G5, H2	10	2, 5
NL	D1, H2	105	3, 5, 7
NL	H2, K1	45	3, 5, 9
VX	C1, C5	20	2, 4, 5, 10
YM	A3, B3	25	5

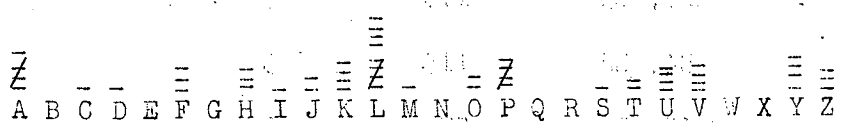
b. The factor 5 appears in all but two cases, each of which involves only a digraph. It seems almost certain that the number of alphabets is five. Since the text already appears in groups of five letters, it is unnecessary to rewrite the message. The next step is to make a monoliteral frequency distribution for Alphabet 1 to see if it can be determined whether or not standard alphabets are involved. It is as follows:

Alphabet 1.



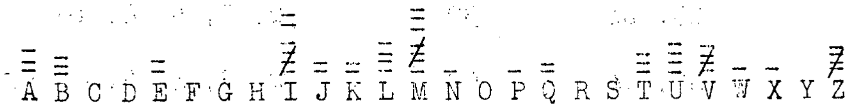
c. Although the indications are not very clear cut, yet if one takes into consideration the small amount of data the assumption of a direct standard alphabet with $W_c = A_p$, is worth further test. Accordingly a similar distribution is made for Alphabet 2.

Alphabet 2.

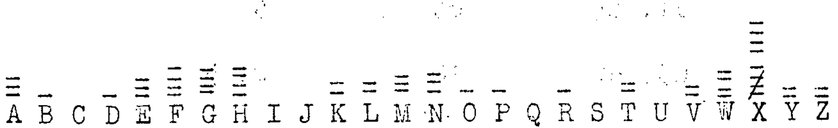


d. There is every indication of a direct standard alphabet, with $H_c = A_p$. Let similar distribution be made for the last three alphabets. They are as follows:

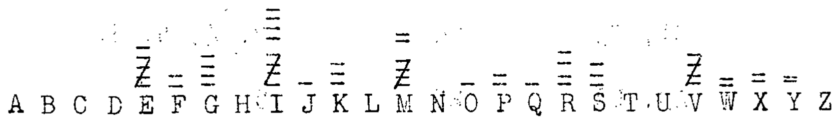
Alphabet 3.



Alphabet 4.



Alphabet 5.



e. After but little experiment it is found that the distributions can best be made to fit the normal when the following values are assumed:

Alphabet 1 -- $A_p = W_c$

Alphabet 2 -- $A_p = H_c$

Alphabet 3 -- $A_p = I_c$

Alphabet 4 -- $A_p = T_c$

Alphabet 5 -- $A_p = E_c$

f. Note the key word given by the successive equivalents of A_p: WHITE. The real proof of the correctness of the analysis is, of course, to test the values of the solved alphabets on the cryptogram. The five complete cipher alphabets are as follows:

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

Fig. 2

g. Applying these values to the first few groups of our message, the following is found:

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	
Cipher -	A	U	K	H	Y	J	A	M	K	I	Z	Y	M	W	M	J	M	I	G	X	N	F	M	L	X	. . .
Plain -	E	N	C	O	U	N	T	E	R	E	D	R	E	D	I	N	F	A	N	T	R	Y	E	S	T	. . .

h. Intelligible text at once results, and the solution can now be completed very quickly. The complete message is as follows:

> ENCOUNTERED RED INFANTRY ESTIMATED AT ONE REGIMENT AND MACHINE GUN COMPANY IN TRUCKS NEAR EMMITSBURG. AM HOLDING MIDDLE CREEK NEAR HILL 543 SOUTHWEST OF FAIRPLAY. WHEN FORCED BACK WILL CONTINUE DELAYING REDS AT MARSH CREEK. HAVE DESTROYED BRIDGES ON MIDDLE CREEK BETWEEN EMMITSBURG-TANEYTOWN ROAD AND RHODES MILL.

i. It is obvious that reversed standard alphabets may be used. The solution is accomplished in the same manner. In fact, the now obsolete cipher disk used by the United States Army for a number of years yields exactly this type of cipher and may just as readily be solved. In fitting the isolated frequency distributions to the normal direction of "reading" the crests and troughs is merely reversed.

14. Solution by completing the plain-component sequence. - a. There is another method of solving this type of cipher, which is worthwhile explaining, because the underlying principles will be found useful in many cases. It is a modification of the method of solution by completing the plain-component sequence, already explained in Par. 20 of Part I.

b. After all, the individual alphabets of a cipher such as the one just solved are merely standard direct alphabets. It has been seen that monoalphabetic ciphers in which standard cipher alphabets are employed may be solved almost mechanically by completing the plain-component sequence. The plain text reappears on only one generatrix and this generatrix is the same for the whole message. It is easy to pick this generatrix out of all the other generatrices because it is the only one which yields intelligible text. Is it not apparent that if the same process is applied to the cipher letters of the individual alphabets of the cipher just solved that the plain-text equivalents of these letters must all reappear on one and the same generatrix? But how will the generatrix which actually contains the plain-text letters be distinguishable from the other generatrices, since these plain-text letters are not consecutive letters in the plain text but only letters separated from one another by a constant interval? The answer is simple. The plain-text generatrix should be distinguishable from the others because it will show more and a better assortment of high-frequency letters, and can thus be selected by the eye from the whole set of generatrices. If this is done with all the alphabets in the cryptogram, it will merely be necessary to assemble the letters of the thus selected generatrices in proper order, and the result should be consecutive letters forming intelligible text.

c. An example will serve to make the process clear. Let the same message be used as before. Factoring showed that it involves five alphabets. Let the first ten cipher letters in each alphabet be set down in a horizontal line and let the normal alphabet sequences be completed. Thus:

	Alphabet 1	Alphabet 2	Alphabet 3	Alphabet 4	Alphabet 5
1	AJZJNEZAIJ	UAYMFTHYLK	KMMIMIBMVU	HKWGLMHZMT	YIMXXIRMEG
2	BKAKOFABJK	VBZNGUIZML	LNNJNJCNWV	ILXHMNIANU	ZJNYYJSNFH
3	CLBLPGBCKL	WCAOHVJANM	MOOKOKDOXW	JMYINOJBOV	AKOZZKTOGI
4	DMCMQHCDLM	XDBPIWKBON	NPPLPLEPYX	KNZJOPKCPW	BLPAALUPHJ
5	ENDNRIDEMN	YECQJXLCPO	OQQMQMFQZY	LOAKPQLDQX	CMQBBMVQIK
6	FOEOSJEFNO	ZFDRKYMDQP	PRRNRNGRAZ	MPBLQRMERY	DNRCCNWRJL
7	GPFPTKFGOP	AGESLZNERQ	QSSOSOHSBA	NQCMRSNFSZ	EOSDDOXSKM
8	HQGQULGHPQ	BHFTMAOFSR	RTTPTPITCB	ORDNSTOGTA	FPTEEPYTLN
9	IRHRVMHIQR	CIGUNBPGTS	SUUQUQJUDC	PSEOTUPHUB	GQUFFQZUMO
10	JSISWNIJRS	DJHVOCQHUT	TVVRVRKVED	QTFPUVQIVC	HRVGGRAVNP
11	KTJTXOJKST	EKIWPDRIVU	UWWSWSLWFE	RUGQVWRJWD	ISWHHSBWOQ
12	LUKUYPKLTU	FLJXQESJWV	VXXTXTMXGF	SVHRWXSKXE	JTXIITCXPR
13	MVLVZQLMUV	GMKYRFTKXW	WYYUYUNYHG	TWISXYTLYF	KUYJJUDYQS
14	NWMWARMNVW	HNLZSGULYX	XZZVZVOZIH	UXJTYZUMZG	LVZKKVEZRT
15	OXNXBSNOWX	IOMATHVMZY	YAAWAWPAJI	VYKUZAVNAH	MWALLWFASU
16	PYOYCTOPXY	JPNBUIWNAZ	ZBBXBXQBKJ	WZLVABWOBI	NXBMMXGBTV
17	QZPZDUPQYZ	KQOCVJXOBA	ACCYCYRCLK	XAMWBCXPCJ	OYCNNYHCUW
18	RAQAEVQRZA	LRPDWKYPCB	BDDZDZSDML	YBNXCDYQDK	PZDOOZIDVX
19	SBRBFWRSAB	MSQEXLZQDC	CEEAEATENM	ZCOYDEZREL	QAEPPAJEWY
20	TCSCGXSTBC	NTRFYMARED	DFFBFBUFON	ADPZEFASFM	RBFQQBKFXZ
21	UDTDHYTUCD	OUSGZNBSFE	EGGCGCVGPO	BEQAFGBTGN	SCGRRCLGYA
22	VEUEIZUVDE	PVTHAOCTGF	FHHDHDWHQP	CFRBGHCUHO	TDHSSDMHZB
23	WFVFJAVWEF	QWUIBPDUHG	GIIEIEXIRQ	DGSCHIDVIP	UEITTENIAC
24	XGWGKBWXFG	RXVJCQEVIH	HJJFJFYJSR	EHTDIJEWJQ	VFJUUFOJBD
25	YHXHLCXYGH	SYWKDRFWJI	IKKGKGZKTS	FIUEJKFXKR	WGKVVGPKCE
26	ZIYIMDYZHI	TZXLESGXKJ	JLLHLHALUT	GJVFKLGYLS	XHLWWHQLDF

Fig. 3

d. If now high-frequency generatrices underlined in Fig. 3 are selected and their letters are juxtaposed in columns, the consecutive letters of intelligible plain text immediately present themselves. Thus:

Selected Genera- trices	For Alphabet 1, generatrix	5	E N D N R I D E M N
	For Alphabet 2, generatrix	20	N T R F Y M A R E D
	For Alphabet 3, generatrix	19	C E E A E A T E N M
	For Alphabet 4, generatrix	8	O R D N S T O G T A
	For Alphabet 5, generatrix	23	U E I T T E N I A C

Columnar juxtaposition of letters from selected generatrices

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

Fig. 4

Plain text: ENCOUNTERED RED INFANTRY ESTIMATED AT ONE
REGIMENT AND MAC . . . . .

e. Solution by this method can thus be achieved without the compilation of any frequency tables whatever and is very quickly attained. The inexperienced cryptanalyst may have difficulty at first in selecting the generatrices which contain the most and the best assortment of high-frequency letters, but with increased practice, a high degree of proficiency is attained. After all it is only a matter of experiment, trial, and error to select and assemble the proper generatrices so as to produce intelligible text.

f. If the letters on the sliding strips were accompanied by numbers representing their relative frequencies in plain text, and these numbers were added across each generatrix then that generatrix with the highest total frequency would theoretically always be the plain-text generatrix. Practically it will be among the generatrices which show the first three or four greatest totals. Thus, an entirely mathematical solution for this type of cipher may be applied.

g. If the cipher alphabets are reversed standard alphabets, it is only necessary to convert the cipher letters of each isolated alphabet into their normal plain component equivalents and then proceed as in the case of direct standard alphabets.

h. It has been seen how the key word may be discovered in this type of cryptogram. Usually the key is made up of those letters in the successive alphabets whose equivalents are A_p. Sometimes a key number is used, such as 8-4-7-1-12, which means merely that A_p is represented by the eight letter from A (in the normal alphabet) in the first cipher alphabet, by the fourth letter from A in the second cipher alphabet, and so on. However, the method of solution as illustrated above, being independent of the nature of the key, is the same as before.

15. Solution by the "probable word method". - a. The common use of key words in cryptograms such as the foregoing makes possible a method of solution that is simple and can be used where the more detailed method of analysis using frequency distributions or by completing the plain-component sequence is of no avail, so that in the case of a very short message which may show no recurrences and give no indications as to the number of alphabets involved, this modified method will be found useful.

b. Briefly, the method consists in assuming the presence of a probable word in the message, and referring to the alphabets to find the key letters applicable when this hypothetical word is assumed to be present in various positions in the cipher text. If the assumed word happens to be correct, and is placed in the correct location in the message, the key letters produced by referring to the alphabets will yield the key word. In the following example it is assumed that reversed standard alphabets are known to be used by the enemy.

MESSAGE

M D S T J L Q C X C K Z A S A N Y Y K O L P

c. Extraneous circumstances lead to the assumption of the presence of the word AMMUNITION. One may assume that this word begins the message. Using sliding normal alphabets, one reversed, the other direct, one proceeds to find the key letters by noting what the successive equivalents of A_p are. Thus:

If M D S T J L Q C X C equals
A M M U N I T I O N, then the key letters (= A_p) are
M P E N W T J K L P.

The "key" does not spell any intelligible word. One therefore shifts the assumed word one letter forward and another trial is made.

If D S T J L Q C X C K equals
A M M U N I T I O N, then the key letters (= A_p) are
D E F D Y Y V F Q X.

This also yields no intelligible key word. One continues to shift the assumed word forward one space at a time until the following point is reached:

If L Q C X C K Z A S A equals
A M M U N I T I O N, then the key letters (= A_p) are
L C O R P S S I G N.

The key stands out: It is a cyclic permutation of the name SIGNAL CORPS.[1]

d. If the assumption of reversed standard alphabets yields no good results, then direct standard alphabets are assumed and the test made exactly in the same manner. Solution by this method is inevitable when the correct word has been assumed and its correct position ascertained. Here again is an example of the efficacy of the "probable word" method. Furthermore, as will be shown subsequently, it can also be used as a last resort when mixed alphabets are employed.

---

[1] It should be clear that since the key word or key phrase repeats itself during the encipherment of such a message, the plain-text word upon whose assumed presence in the message this test is being based may begin to be enciphered at any point in the key, and continue over into its next repetition if it is longer than the key. When this is the case it is merely necessary to shift the latter part of the sequence of determined key letters to the first part, as in the case noted: LCORPSSIGN is transposed into SIGN...LCORPS, and thus SIGNAL CORPS.

e. It will be seen in the foregoing method of solution that the length of the key is of no particular interest or consequence in the steps taken in effecting the solution. The determination of the length and elements of the key comes after the solution rather than before it. In this case the length of the period is seen to be eleven (SIGNAL CORPS).

f. The foregoing method is one of the other methods of determining the length of the key (besides factoring), referred to in Par. 10 c.

SECTION V

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, I.

16. Reason for the use of mixed alphabets. - a. It has been seen in the examples considered thus far that the use of several alphabets in the same message does not greatly complicate the analysis of such a cryptogram. There are three reasons why this is so: Firstly, only relatively few alphabets were employed; secondly, these alphabets were employed in a periodic or repeating manner, giving rise to cyclic phenomena in the cryptogram, by means of which the number of alphabets could be determined; and, thirdly, the cipher alphabets were <u>known</u> alphabets, by which is meant merely that the sequences of letters in both components of the cipher alphabets were known sequences.

b. In the case of monoalphabetic ciphers it was found that the use of a mixed alphabet delayed the solution to a considerable degree, and it will now be seen that the use of mixed alphabets in polyalphabetic ciphers renders the analysis much more difficult than the use of standard alphabets, but the solution is still fairly easy to achieve.

17. Interrelated mixed alphabets. - a. It was stated in Par. 2 that the method of producing the mixed alphabets in a polyalphabetic cipher often affords clues which are of great assistance in the analysis of the cipher alphabets. This is so, of course, only when the cipher alphabets are interrelated secondary alphabets produced by sliding components. Reference is now made to the classification set forth in Par. 6, in connection with the types of alphabets which may be employed in polyalphabetic substitution. It will be seen that thus far only Cases A(1) and (2) have been treated. Case B(1) will now be discussed.

b. Here one of the components, the plain component, is the normal sequence, while the cipher component is a mixed sequence, the sliding of the two components yielding mixed alphabets. The mixed component may be a systematically-mixed or a random-mixed sequence. If the successive alphabets produced by the sliding of two such components are set down as in the case of the Vigenère Square, a symmetrical square such as that shown in Fig. 5 results therefrom.

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

Fig. 5

c. Such a table may be used in exactly the same manner as the Vigenere Table. With the key word BLUE the following secondary alphabets would be used:

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

Fig. 6

18. Principles of direct symmetry of position. - a. It was stated directly above that Fig. 5 is a symmetrical cipher square, by which is meant that the letters in its successive horizontal lines show a direct symmetry of position with respect to one another. They constitute, really, one and only one sequence or series of letters, the sequences being merely displaced successively 1, 2, 3, ... intervals. The symmetry exhibited is obvious and is said to be patent, or "direct". This fact can be used to good advantage.

b. Consider, for example, the pair of letters G and V in the B, or 1st, cipher alphabet directly above; the letter V is the 15th letter to the right of G. In the L, or 2d, cipher alphabet, V is also the 15th letter to the right of G, as is the case in every one of these secondary alphabets, since the relative positions they occupy are the same in each horizontal line, that is, in each cipher alphabet. If, therefore, the relative positions occupied by a given pair of letters in one of these cipher alphabets is known, and one of the members of this same pair has been located in another of these cipher alphabets, one may at once place the other member of this pair in its proper position in the second of the cipher alphabets. Suppose, for example, that as the result of an analysis based upon considerations of frequency, the following values in a given cryptogram have been tentatively determined:

	Plain --	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
Cipher	1 --					G										Y					V						
	2 --					N										G					P						
	3 --					L										B					I						
	4 --					W										I					Q						

Fig. 7

The letter G is common to Alphabets 1 and 2. In Alphabet 2 it is noted that N occupies the 10th position to the left of G, and the letter P occupies the 5th position to the right of G. One may therefore place these letters, N and P, in their proper positions in Alphabet 1, the letter N being placed 10 letters before G, and the letter P, 5 letters after G. Thus:

Plain --	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
1 --					G					P					Y					V	N					

Thus, the values of two new letters in Alphabet 1, viz., $P_c = J_p$, and $N_c = U_p$ have been automatically determined; these values were obtained without any analysis based upon the <u>frequency</u> of P_c and N_c. Likewise, in Alphabet 2, the letters Y and V may be inserted in these positions:

Plain --	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
2 --				V	N										G					P					Y	

This gives the new values $V_c = D_p$ and $Y_c = Y_p$ in Alphabet 2. Alphabets 3 and 4 have a common letter I, which permits of the placement of Q and W in Alphabet 3, and of B and L in Alphabet 4.

c. The new values thus found are of course immediately inserted throughout the cryptogram, thus leading to the assumption of further values in the cipher text. This process, the reconstruction of the primary components by the application of the principles of direct symmetry of position, thus facilitates and hastens solution.

d. It must be clearly understood that before the principles of direct symmetry of position can be applied in cases such as the foregoing, it is necessary that the plain component be a <u>known</u> sequence. Whether it is the normal sequence or not is immaterial, so long as the sequence is known. Obviously, if the sequence is unknown, symmetry even if present, cannot be detected by the cryptanalyst because he has no <u>base</u> upon which to try out his assumptions for symmetry. In other words, direct symmetry of position is manifested in the illustrative example because the plain component was a known sequence, and not because it was the normal alphabet. The significance of this point will become apparent later on in connection with the problem discussed in Par. 26b.

19. Initial steps in the solution of a typical example. -- a. In the light of the foregoing principles let a typical message now be studied.

MESSAGE

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

b. The principal repetitions of three or more letters have been underlined in the message and the factors (up to 20 only) of the intervals between them are as follows:

QWBRIVWY -- 45 = 3, 5, 9, 15
CGXGB -- 60 = 2, 3, 4, 5, 6, 10, 12, 15, 20
PJEL -- 95 = 5, 19
ZZGI -- 145 = 5
BRIV -- 285 = 3, 5, 15, 19
BRI -- 45 = 3, 5, 9, 15
KAG -- 75 = 3, 5, 15
QRD -- 165 = 3, 5, 15
QWB -- 45 = 3, 5, 9, 15
QWB -- 275 = 5, 11
WIC -- 130 = 2, 5, 10, 13
XNF -- 45 = 3, 5, 9, 15
YZT -- 225 = 3, 5, 15
ZTC -- 145 = 3, 5

The factor 5 is common to all of these repetitions, and there seems to be every indication that five alphabets are involved. Since the message already appears in groups of five letters, it is unnecessary in this case to rewrite it in groups corresponding to the length of the key. The monoliteral frequency distribution for Alphabet 1 is as follows:

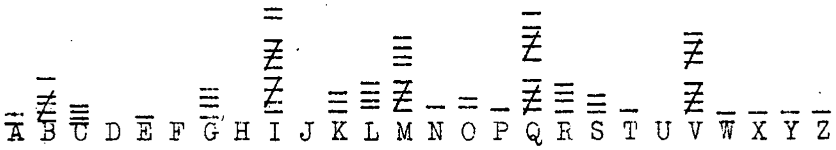


Fig. 8.

c. Attempts to fit this distribution to the normal on the basis of a direct or reversed standard alphabet do not give positive results, and it is assumed that mixed alphabets are involved. Individual <u>trigraphic</u> frequency distributions are then compiled and are shown in Fig. 9. These tables are similar to those made for single mixed alphabet ciphers, and are made in the same way except that instead of taking the letters one after the other, we now must assemble in separate tables the letters which belong to the separate alphabets. For example, in Alphabet 1, the trigraph QAC means that A occurs in Alphabet 1; Q, its prefix, occurs in Alphabet 5, and C, its suffix, occurs in Alphabet 2. We may avoid all confusion by placing numbers indicating the alphabets in which they belong above the letters, thus: $\overset{512}{\text{QAC}}$.

Alphabet 1.

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
QC	GW	NT		TV		AE		AS		UD	UW	IT	UT	QP	NX	-W	LB	LA	LA		IW	NN	QI	UX	QR
PT	OP	TC				AD		WC		FI	QX	II		UP		YW	YW	DE			IW				
	GK	TT				LX		HW		FW	LV	OT				NW	QD	RB			UE				
	OW	WB				LW		ND			LR	SY				QC	QD				LC				
	GL							GV				WC				GI					GP				
	GX							WC				GP				QL					QB				
								XD				AB				RI					NW				
								GB				JF				YV					QE				
								IV				DI				NY					IP				
								NR								SW					UP				
								AK								QW									
								QB																	

Fig. 9.

Alphabet 2.

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
SN	RZ	IJ	IM	GG	MD			MB		IW	QF		WB		BD		ZH	IP	MZ		IX	QB	GN	MJ	
TG	VG	QG	GG	VZ				QG		BZ	BG				OD		IG		CG		QF	VY	BD	QA	
	IE	VG	ID	SZ				QI							VW		LZ		NZ		LV	QY	PF		
	MJ	CB	RG	VD				KL							OJ				MY		IJ	LM	YN		
	SG	IG	KH					MY							MJ				CJ		EG	QB	LN		
	CY	MJ	RZ					XN							VJ				AY			VY			
	IW	AJ													VY							BN			
																						IJ			
																						BF			
																						RN			
																						VD			
																						QB			
																						KF			
																						GF			
																						QJ			

Alphabet 3.

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
YH	WR		PB	BY	WE	CQ	RC	IE	CC		IC	WG	WB		SJ						VC	PM	VC	WC	BE
	IK		PK		LC	EX	DC		WK			DR	WF									KJ		WE	TE
	WR		DR		VW	IV			YJ				XF									BR		WI	EY
	CY		WY		XP	TY			CK				XF											TZ	KZ
	WI		XB		WZ	CX			PQ				AC											IZ	TA
	NR		FZ		WJ	DI			PE				XC											TE	EZ
			EC			CX			BJ				IB											BZ	RN
						LQ			TR															PH	DY
						BR			CR																
						DD			VR																
						BZ			PE																
						AD			WY																
						RQ																			
						VQ																			

Alphabet 4.

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
ZO	NQ	YA	GG	ZY	NL	MW	AQ	YG	PL	BN		WR	ZQ		FU	GH	BI				GN	FY	GN	ZG	ZG
	DL	JI	GN	YU	NW		YL	GG	JY	JA						GQ	BI						GG	GO	YT
	DN	XU		ZI	NG			BI	JF	DA						JQ	MU						GG	BQ	ZG
	NA	FO		FQ					WQ	JX						GP	GS							DQ	DT
		HN		IW					FQ							GU	DU							EU	YQ
		ND		JL													JD							ZF	GN
		HA		JL													JR							JQ	YT
		LJ		YW													JN								FL
		DQ															BI								
		NL															WX								
		VS																							

Fig. 9 (continued)

Alphabet 5.

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
CI			CS		JK	IB	QI	RV	CM		JR		KQ	YB	QA	BQ	MQ	RM	ZC	EL		GI	KI	EQ	
KG			RM		YK	YQ		CM			BV		XI	AB		EQ	RS	CQ	ZC	RV		EI	R-	JQ	
KG						XB		EM			FG		VC	CM		YO			ZE	CN		FM		WR	
CM						ZI		RV			ES		CV			QV				RO		EC			
BI						IV		II			CL		BP			QZ				PY					
						XB		RV			ET		ZQ			YR				YK					
						DB					HL		RW			ZA				QV					
						FM					ZG		DI			HV									
						ZI										CL									
																NX									
																JR									
																JQ									
																YI									

Condensed table of repetitions.

5-1-2
I V W-2
Q R D-2
W I C-2

1-2-3
Q W B-3
V W Y-2

2-3-4
C G X-2
P J E-2
W B R-2
X N F-2

3-4-5
B R I-3
G X G-2
J E L-2
Y Z T-2
Z Z G-2

4-5-1
K A G-2
X G B-2
Z G I-2
Z T C-2
R I V-3

1-2
Q W-5
V P-3
V W-3

2-3
C G-3
C J-3
P J-3
W B-3
W F-3
W Y-3
X N-3

3-4
B R-3
G Q-4
G X-3
J R-3
N F-3
Y Z-3

4-5
R I-3
Y Q-3
Z T-3

5-1
G B-4
I V-3
Q Q-3

Fig. 9 (continued)

d. One now proceeds to analyze each alphabet distribution, in an endeavor to establish identifications of cipher equivalents. First, of course, attempts should be made to separate the vowels from the consonants in each alphabet, using the same test as in the case of a simple mixed alphabet cipher. There seems to be no doubt about the equivalent of E_p in each alphabet: $E = \overset{1}{I}_c \quad \overset{2}{W}_c \quad \overset{3}{G}_c \quad \overset{4}{C}_c \quad \overset{5}{Q}_c$.

e. The letters of greatest frequency in Alphabet 1 are I, M, Q, V, B, G, L, R, S, and C. I_c has already been assumed to be E_p. If $\overset{2}{W}_c$ and $\overset{5}{Q}_c = E_p$, then one should be able to distinguish the vowels from the consonants among the letters M, Q, V, B, G, L, R, S, and C by examining the prefixes of $\overset{2}{W}_c$, and the suffixes of $\overset{5}{Q}_c$. The prefixes and suffixes of these letters, as shown by the trigraphic frequency tables, are these:

Prefixes of $\overset{2}{W}_c$ (= $\overset{2}{E}_p$)	Suffixes of $\overset{5}{Q}_c$ (= $\overset{5}{E}_p$)
$\underline{\underline{\underline{\underline{Q}}}}$ G K $\underline{\underline{V}}$ R $\underline{B}$ I L	I $\overline{\overline{Q}}$ $\overline{R}$ X L $\overline{V}$ A Z O

f. Consider now the letter $\overset{1}{M}_c$; it does not occur either as a prefix of $\overset{2}{W}_c$, or as a suffix of $\overset{5}{Q}_c$. Hence it is most probably a vowel, and on account of its high frequency it may be assumed to be O_p. On the other hand, note that $\overset{1}{Q}_c$ occurs five times[1] as a prefix of $\overset{2}{W}_c$ and three times as a suffix of $\overset{5}{Q}$. It is therefore a consonant, most probably R, for it would give the digraph ER (= $\overset{51}{QQ}_c$) as occurring three times and RE (= $\overset{12}{QW}_c$) as occurring five times.

g. The letter $\overset{1}{V}_c$ occurs three times as a prefix of $\overset{2}{W}_c$ and twice as a suffix of $\overset{5}{Q}_c$. It is therefore a consonant, and on account of its frequency, let it be assumed to be T_p. The letter $\overset{1}{B}_c$ occurs twice as a prefix of $\overset{2}{W}_c$ but not as a suffix of $\overset{5}{Q}_c$. Its frequency is only medium, and it is probably a consonant. In fact, the twice repeated digraph $\overset{12}{BW}_c$ is once a part of the trigraph $\overset{512}{GBW}$, and $\overset{5}{G}_c$, the letter of second highest frequency in Alphabet 5, looks excellent for T_p. Might not the trigraph $\overset{512}{GBW}$ be THE? It will be well to keep this possibility in mind.

h. The letter $\overset{1}{G}_c$ occurs only once as a prefix of $\overset{2}{W}_c$ and does not occur as a suffix of $\overset{5}{Q}_c$. It may be a vowel, but one can not be sure.

---

[1] The letter Q has four tallies under it, plus one occurrence indicated by the presence of the letter itself among the prefixes, equals five occurrences. The same applies to the other letters.

The letter $\overset{1}{L}_c$ occurs once as a prefix of $\overset{2}{W}_c$ and once as a suffix of $\overset{5}{Q}_c$. It may be considered to be a consonant. $\overset{1}{R}_c$ occurs once as a prefix of $\overset{2}{W}_c$, and twice as a suffix of $\overset{5}{Q}_c$, and is certainly a consonant. Neither the letter $\overset{1}{S}_c$ nor the letter $\overset{1}{C}_c$ occurs as a prefix of $\overset{2}{W}_c$ or as a suffix of $\overset{5}{Q}_c$; both would seem to be vowels, but a study of the prefixes and suffixes of these letters lends more weight to the assumption that $\overset{1}{C}_c$ is a vowel than that $\overset{1}{S}_c$ is a vowel. For all the prefixes of C, viz, $\overset{5}{N}$, $\overset{5}{T}$, and $\overset{5}{W}$, are in subsequent analysis of Alphabet 5 classified as consonants, as are likewise its suffixes, viz, T, C, and B in Alphabet 2. On the other hand, only one prefix, $\overset{5}{L}_c$, and one suffix, $\overset{2}{B}_c$, of $\overset{1}{S}_c$ are later classified as consonants. Since vowels are more often associated with consonants than with other vowels, it would seem that $\overset{1}{C}_c$ is more likely to be a vowel than $\overset{1}{S}_c$. At any rate $\overset{1}{C}_c$ is assumed to be a vowel, for the present, leaving $\overset{1}{S}_c$ unclassified.

i. Going through the same steps with the remaining alphabets, the following results are obtained:

Alphabet	Consonants	Vowels
1	Q, V, B, L, R, G?	I, M, C
2	B, C, D, T	W, P, I
3	J, N, D, Y, F	G, Z
4	Y, Z, J, Q	C, E?, R?, B?
5	G, N, A, I, W, L, T	Q, U

20. Application of principles of direct symmetry of position. - a. The next step is to try to determine a few values in each alphabet. In Alphabet 1, from the analysis above, the following data are on hand:

Plain --	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
Cipher --	C?				I				C?						M			Q		V						

Let the values of E_p already assumed in the remaining alphabets, be set down, as follows:

		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
Plain	--	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
Cipher 1	--	C?				I				C?						M			Q		V						
Cipher 2	--					W																					
Cipher 3	--					G																					
Cipher 4	--					C																					
Cipher 5	--					Q																					

Fig. 10

b. It is seen that by good fortune the letter Q is common to Alphabets 1 and 5, and the letter C is common to Alphabets 1 and 4. If it is assumed that one is dealing with a case in which a mixed component is sliding against the normal component, one can apply the principles of direct symmetry of position to these alphabets, as outlined in Par. 18. For example, one may insert the following values in Alphabet 5:

		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
Plain	--	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
Cipher 1	--	C?				I				C?						M			Q		V						
Cipher 5	--		M			Q		V							C?				I				C?				

Fig. 11

c. The process at once gives three definite values: $\overset{5}{M}_c = B_p$, $\overset{5}{V}_c = G_p$, $\overset{5}{I}_c = R_p$. Let these deduced values be substantiated by referring to the frequency distribution. Since B and G are normally low or medium frequency letters in plain text, one should find that M_c and V_c, their hypothetical equivalents in Alphabet 5, should have low frequencies. As a matter of fact, they do not appear in this alphabet, which thus far corroborates the assumption. On the other hand, since $\overset{5}{I}_c = R_p$, if the values derived from symmetry of position are correct, $\overset{5}{I}_c$ should be of high frequency, and it is. The position of C is doubtful; it belongs either under N_p or V_p. If the former is correct, then the frequency of $\overset{5}{C}_c$ should be high, for it would equal N_p; if the latter is correct, then its frequency should be low, for it would equal V_c. As a matter of fact $\overset{5}{C}_c$ does not occur, and it must be concluded that it belongs under V_p. This in turn settles the value of $\overset{1}{C}_c$, for it must now be placed definitely under I_p and removed from beneath A_p.

d. The definite placement of C now permits the insertion of new values in Alphabet 4, and one now has the following:

	Plain --	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
Cipher	1 --					I				C						M			Q		V						
	2 --					W																					
	3 --					G																					
	4 --	I				C						M			Q		V										
	5 --		M			Q		V											I				C				

Fig. 12.

21. Subsequent steps in solution. - a. It is high time that the thus far deduced values be inserted in the cipher text, for by this time it must seem that one has certainly gone too far with work based upon unproved hypotheses. The following results:

MESSAGE

```
QWBRI  VWYCA  ISPJL  RBZEY  QWYEU  LWMGW  ICJCI  MTZEI  MIBKN  QWBRI
RE  R  TE E   E             RE      E     E  ER  O   R  O      RE  R

VWYIG  BWNBQ  QCGQH  IWJKA  GEGXN  IDMRU  VEZYG  QIGVN  CTGYO  BPDBL
TE A    E  E  R EN   EE       E    E      T      R EP   I E

VCGXG  BKZZG  IVXCU  NTZAO  BWFEQ  QLFCO  MTYZT  CCBYQ  OPDKA  GDGIG
T E           E  E           E  E  R  E   O      I   E           EA

VPWMR  QIIEW  ICGXG  BLGQQ  VBGRS  MYJJY  QVFWY  RWNFL  GXNFW  MCJKX
T  K   R      E E      ENE  T E    O      R       E            O

IDDRU  OPJQQ  ZRHCN  VWDYQ  RDGDG  BXDBN  PXFPU  YXNFG  MPJEL  SANCD
E         NE     E   TE  E    E                         O         E

SEZZG  IBEYU  KDHCA  MBJJF  KILCJ  MFDZT  CTJRD  MIYZQ  ACJRR  SBGZN
       E         E   O         E   O      I      O   E           E

QYAHQ  VEDCQ  LXNCL  LVVCS  QWBII  IVJRN  WNBRI  VPJEL  TAGDN  IRGQP
R   E  T  EE     E      E   RE AR  E          R  T        E    E EN

ATYEW  CBYZT  EVGQU  VPYHL  LRZNQ  XINBA  IKWJQ  RDZYF  KWFZL  GWFJQ
       I        EN   T          E         E   E          E      E  E

QWJYQ  IBWRX
RE  E  E
```

b. The combinations given are excellent throughout and no inconsistencies appear. Note the trigraph $\overset{123}{QWB}$, which is repeated in the following polygraphs (underlined in the foregoing text):

```
1 2 3 4 5 1         5 1 2 3 4 5 1
Q W B R I V . . . . S Q W B I I I
R E     R T . . .   R E   A R E
```

c. The letter $\overset{3}{B}_c$ is common to both polygraphs, and a little imagination will lead to the assumption of the value $\overset{3}{B}_c = P_p$, yielding the following:

```
1 2 3 4 5 1         5 1 2 3 4 5 1
Q W B R I V . . . . S Q W B I I I
R E P O R T . . . . P R E P A R E
```

d. Note also the following polygraph: $\underset{A}{\overset{4}{I}}\ \overset{5}{G}\ \underset{T}{\overset{1}{V}}\ \overset{2}{P}\ \overset{3}{W}\ \underset{K}{\overset{4}{M}}$, which looks like the word ATTACK. The frequency distributions are consulted to see whether the frequencies given for $\overset{5}{G}_c$ and $\overset{2}{P}_c$ are high enough for T_p and A_p, respectively, and also whether the frequency of $\overset{3}{W}_c$ is good enough for C_p; it is noted that they are excellent. Moreover, the digraph $\overset{51}{GB}_c$, which occurs four times, looks like TH, thus making $\overset{1}{B}_c = H_p$. Does the insertion of these four new values in our diagram of alphabets bring forth any inconsistencies? The insertion of the value $\overset{2}{P}_c = A_p$ and $\overset{1}{B}_c = H_p$ gives no indications either way, since neither letter has yet been located in any of the other alphabets. The insertion of the value $\overset{5}{G}_c = T_p$ gives a value common to Alphabets 3 and 5, for the value $\overset{3}{G}_c = E_p$ was assumed long ago. Unfortunately an inconsistency is found here. The letter I has been placed two letters to the left of G in the mixed component, and has given good results in Alphabets 1 and 5; if the value $\overset{3}{W}_c = C_p$, as obtained above from the assumption of the word ATTACK, is correct, then W, and not I, should be the second letter to the left of G. Which shall be retained? There has been so far nothing to establish the value of $\overset{3}{G}_c = E_p$; this value was assumed from frequency considerations solely. Perhaps it is wrong. It certainly behaves like a vowel, and one may see what happens when one changes its value to O_p. The following placements result from the analysis when only two or three new values have been added as a result of the clues afforded by the deductions:

		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
Cipher	1 --			S		I		G	B	C						M		P	Q	R	V	W					
	2 --	P	Q	R	V	W								S		I		G	B	C						M	
	3 --	R	V	W								S		I		G	B	C						M		P	Q
	4 --	I		G	B	C						M		P	Q	R	V	W								S	
	5 --		M		P	Q	R	V	W								S		I		G	B	C				

Plain -- in the first row of column heads; "Cipher" spans rows 1 to 5.

Fig. 13 a.

e. Many new values are produced, and these are inserted throughout the message, yielding the following:

```
QWBRI  VWYCA  ISPJL  RBZEY  QWYEU  LWMGW  ICJCI  MTZEI  MIBKN  QWBRI
REPOR  TE E   EMY    SR     RE      EWCH  ES ER  O   R  OOP    REPOR

VWYIG  BWNBQ  QCGQH  IWJKA  GEGXN  IDMRU  VEZYG  QIGVN  CTGYO  BPDBL
TE AT  HE DE  RSON   EE     G O    E WO   T   T  ROOP   I O    HA D

VCGXG  BKZZG  IVXCU  NTZAO  BWFEQ  QLFCO  MTYZT  CCBYQ  OPDKA  GDGIG
TSO T  H   T  ED E          HE  E  R  E   O      ISP E   A     G OAT

VPWMR  QIIEW  ICGXG  BLGQQ  VBGRS  MYJJY  QVFWY  RWNFL  GXNFW  MCJKX
TACKF  ROM H  ESO T  H ONE  TROOP  O      RD Q   SE     G   H  OS

IDDRU  OPJQQ  ZRHCN  VWDYQ  RDGDG  BXDBN  PXFPU  YXNFG  MPJEL  SANCD
E  O    A NE   C E   TE  E  S O T  H  D   Q  M      T  OA     C  E

SEZZG  IBEYU  KDHCA  MBJJF  KILCJ  MFDZT  CTJRD  MIYZQ  ACJRR  SBGZN
C   T  ER         E  OR      O E   O      I  O   OO  E   S OF  CRO

QYAHQ  VEDCQ  LXNCL  LVVCS  QWBII  IVJRN  WNBRI  VPJEL  TAGDN  IRGQP
R   E  T  EE      E   DBEP  REPAR  ED O   U POR  TA       O    ECOND

ATYEW  CBYZT  EVGQU  VPYHL  LRZNQ  XINBA  IKWJQ  RDZYF  KWFZL  GWFJQ
    H  IR     DON    TA       C E   O D   E   E  S        E    GE  E

QWJYQ  IBWRX
RE  E  ER O
```

22. Completing the solution. - a. Completion of solution is now a very easy matter. The mixed component is finally found to be the following sequence, based upon the word EXHAUSTING:

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

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

Fig. 13 b.

b. Note that the successive equivalents of A_p spell the word APRIL, which is the key for the message. The plain-text message is as follows:

> REPORTED ENEMY HAS RETIRED TO NEWCHESTER. ONE TROOP IS REPORTED AT HENDERSON MEETING HOUSE: TWO OTHER TROOPS IN ORCHARD AT SOUTHWEST EDGE OF NEWCHESTER. 2D SQ IS PREPARING TO ATTACK FROM THE SOUTH. ONE TROOP OF 3D SQ IS ENGAGING HOSTILE TROOP AT NEWCHESTER. REST OF 3D SQ IS MOVING TO ATTACK NEWCHESTER FROM THE NORTH. MOVE YOUR SQ INTO WOODS EAST OF CROSSROAD 539 AND BE PREPARED TO SUPPORT ATTACK OF 2D AND 3D SQ. DO NOT ADVANCE BEYOND NEWCHESTER. MESSAGES HERE.
>
> TREER,
> COL.

c. The preceding case is a good example of the value of the principles of direct symmetry of position when applied properly to a cryptogram enciphered by the sliding of a mixed component against the normal. The cryptanalyst starts off with only a very limited number of assumptions and builds up many new values as a result of the placement of the few original values in the diagram of the alphabets.

23. Solution of subsequent messages enciphered by the same cipher component. - a. Preliminary remarks. Let it be supposed that the correspondents are using the same basic or primary component but with different key words for other messages. Can the knowledge of the sequence of letters in the reconstructed primary component be used to solve the subsequent messages? It has been shown that in the case of a monoalphabetic cipher in which a mixed alphabet was used, the process of completing the plain component could be applied to solve subsequent messages in which the same cipher component was used even though the cipher component was set at a different key letter. A modification of the procedure used in that case can be used in this case, where a plurality of cipher alphabets based upon a sliding primary component is used.

b. The message. Let it be supposed that the following message passing between the same two correspondents as in the preceding message has been intercepted:

MESSAGE

SFDZR YRRKX MIWLL AQRLU RQFRT IJQKF XUWBS MDJZK
MICQC UDPTV TYRNH TRORV BQLTI QBNPR RTUHD PTIVE
RMGQN LRATQ PLUKR KGRZF JCMGP IHSMR GQRFX BCABA
OEMTL PCXJM RGQSZ VB

c. Factoring and conversion into plain component equivalents. The presence of a repetition of a four-letter polygraph whose interval is 21 letters suggests a key word of seven letters. There are very few other repetitions, and this is to be expected in a short message with a key of such length.

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

Fig. 14

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

Fig. 15

d. Transcription into periods. Let the message be written in groups of seven letters, in columnar fashion, as shown in Fig. 14. The letters in each column belong to a single alphabet. Let the letters in each column be converted into their plain component equivalents by setting the reconstructed cipher component against the normal alphabet at any arbitrarily selected point, for example, that shown below:

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

The columns of equivalents are now as shown in Fig. 15.

e. Examination and selection of generatrices. It has been shown that in the case of a monoalphabetic cipher it was merely necessary to complete the normal alphabet sequence beneath the plain-component equivalents and the plain text all reappeared on one generatrix. It was also found that in the case of a multiple-alphabet cipher involving standard alphabets, the plain-text equivalents of each alphabet reappeared on the same generatrix, and it was necessary only to combine the proper generatrices in order to produce the plain text of the message. In the case at hand both processes are combined: the normal alphabet sequence is continued beneath the letters of each column and then the generatrices are combined to produce the plain text. The completion diagrams for the first two columns are as follows (Fig. 16):

Column 1.		Column 2.
FVQUPRLWVGVHIQZHVDLF		NPDNNMUGSHGWQENCNSBZ
GWRVQSMXWHWIJRAIWEMG	1	OQEOONVHTIHXRFODOTCA
HXSWRTNYXIXJKSBJXFNH	2	PRFPPOWIUJIYSGPEPUDB
IYTXSUOZYJYKLTCKYGOI	3	QSGQQPXJVKJZTHQFQVEC
JZUYTVPAZKZLMUDLZHPJ	4	RTHRRQYKWLKAUIRGRWFD
KAVZUWQBALAMNVEMAIQK	5	SUISSRZLXMLBVJSHSXGE
LBWAVXRCBMBNOWFNBJRL	6	TVJTTSAMYNMCWKTITYHF
MCXBWYSDCNCOPXGOCKSM	7	UWKUUTBNZONDXLUJUZIG
NDYCXZTEDODPQYHPDLTN	8	VXLVVUCOAPOEYMVKVAJH
OEZDYAUFEPEQRZIQEMUO	9	WYMWWVDPBQPFZNWLWBKI
PFAEZBVGFQFRSAJRFNVP	10	XZNXXWEQCRQGAOXMXCLJ
QGBFACWHGRGSTBKSGOWQ	11	YAOYYXFRDSRHBPYNYDMK
RHCGBDXIHSHTUCLTHPXR	12	ZBPZZYGSETSICQZOZENL
SIDHCEYJITIUVDMUIQYS	13	ACQAAZHTFUTJDRAPAFOM
TJEIDFZKJUJVWENVJRZT	14	BDRBBAIUGVUKESBQBGPN
UKFJEGALKVKWXFOWKSAU	15	CESCCBJVHWVLFTCRCHQO
VLGKFHBMLWLXYGPXLTBV	16	DFTDDCKWIXWMGUDSDIRP
WMHLGICNMXMYZHQYMUCW	17	EGUEEDLXJYXNHVETEJSQ
XNIMHJDONYNZAIRZNVDX	18	FHVFFEMYKZYOIWFUFKTR
YOJNIKEPOZOABJSAOWEY	19	GIWGGFNZLAZPJXGVGLUS
ZPKOJLFQPAPBCKTBPXFZ	20	HJXHHGOAMBAQKYHWHMVT
AQLPKMGRQBQCDLUCQYGA	21	IKYIIHPBNCBRLZIXINWU
BRMQLNHSRCRDEMVDRZHB	22	JLZJJIQCODCSMAJYJOXV
CSNRMOITSDSEFNWESAIC	23	KMAKKJRDPEDTNBKZKPYW
DTOSNPJUTETFGOXFTBJD	24	LNBLLKSEQFEUOCLALQZX
EUPTOQKVUFUGHPYGUCKE	25	MOCMMLTFRGFVPDMBMRAY

Fig. 16.

f. Combining the selected generatrices. After some experimenting with these generatrices the 23d generatrix of Column 1 and the 1st of Column 2, which yields the digraphs shown in Fig. 17, are combined. The generatrices of the subsequent columns are examined to select those which may be added to these already selected in order to build up the plain text. The results are shown in Fig. 18. This process is a very valuable aid in the solution of messages after the primary component has been recovered as a result of the longer and more detailed analysis of the frequency tables of the first message intercepted. Very often a short message can be solved in no other way than the one shown, when the primary alphabet is completely known.

1	2
C	O
S	Q
N	E
R	Q
M	O
O	N
I	V
T	H
S	T
D	I
S	H
E	X
F	R
N	F
W	O
E	D
S	O
A	T
I	C
C	A

Fig. 17.

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

Fig. 18.

g. Recovery of the key. It may be of interest to find the key word for the message. All that is necessary is to set the mixed component of the cipher alphabet underneath the plain component so as to produce the cipher letter indicated as the equivalent of any given plain-text letter in each of the alphabets. For example, in the first alphabet it is noted that $C_p = S_c$. Setting the two components under each other so as to bring S of the cipher component beneath C of the plain component, thus:

```
Plain:  ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher:                        EXHAUSTINGBCDFJKLMOPQRVWYZ
```

It is noted that $A_p = A_c$. Hence, the first letter of the key word to the message is A. The 2d, 3d, 4th, ... 7th key letters are found in exactly the same manner, and the following is obtained:

When C O F I R S T equals
S F D Z R Y R then A_p successively equals
A Z I M U T H

24. Summation of relative frequencies as an aid to the selection of the correct generatrices. - a. In the foregoing example, under subparagraph f, there occurs this phrase: "After some experimenting with these generatrices..." By this was meant, of course, that the selection of the correct initial pair of generatrices of plain-text equivalents is in this process a matter of trial and error. The test of "correctness" is whether, when juxtaposed, the two generatrices so selected yield "good" digraphs, that is, high-frequency digraphs such as occur in normal plain text. In his early efforts the student may have some difficulty in selecting, merely with his eyes, the most likely generatrices to try. There may be in each diagram several generatrices which contain good assortments of high-frequency letters, and the number of trials of combinations of generatrices may be quite large. Perhaps a simple mathematical method may be of assistance in the process.

b. Suppose, in Fig. 16, that each letter were accompanied by a number which corresponds to its relative frequency. Then, by adding the numbers along each <u>horizontal</u> line, the totals thus found will give a numerical measure of the frequency value of each generatrix. Theoretically, the generatrix with the greatest value will be the correct generatrix because its total will represent the sum of the individual values of the actual plain-text letters. In actual practice, of course, the generatrix with the greatest value may not be the correct one, but the correct one will certainly be among the three or four generatrices with the largest values. Thus, the number of trials may be greatly reduced, in the attempt to put together the correct generatrices.

c. Using the preceding message as an example, note the respective generatrix values in Fig. 19.

Generatrix	Column 1.																				Frequency Value
0	F	V	Q	U	P	R	L	W	V	G	V	H	I	Q	Z	H	V	D	L	F	
	3	2	0	3	3	8	4	2	2	2	2	3	7	0	0	3	2	4	4	3	57
1	G	W	R	V	Q	S	M	X	W	H	W	I	J	R	A	I	W	E	M	G	
	2	2	8	2	0	6	2	0	2	3	2	7	0	8	7	7	2	13	2	2	77
2	H	X	S	W	R	T	N	Y	X	I	X	J	K	S	B	J	X	F	N	H	
	3	0	6	2	8	9	8	2	0	7	0	0	0	6	1	0	0	3	8	3	66
3	I	Y	T	X	S	U	O	Z	Y	J	Y	K	L	T	C	K	Y	G	O	I	
	7	2	9	0	6	3	8	0	2	0	2	0	4	9	3	0	2	2	8	7	74
4	J	Z	U	Y	T	V	P	A	Z	K	Z	L	M	U	D	L	Z	H	P	J	
	0	0	3	2	9	2	3	7	0	0	0	4	2	3	4	4	0	3	3	0	49
5	K	A	V	Z	U	W	Q	B	A	L	A	M	N	V	E	M	A	I	Q	K	
	0	7	2	0	3	2	0	1	7	4	7	2	8	2	13	2	7	7	0	0	74
6	L	B	W	A	V	X	R	C	B	M	B	N	O	W	F	N	B	J	R	L	
	4	1	2	7	2	0	8	3	1	2	1	8	8	2	3	8	1	0	8	4	73
7	M	C	X	B	W	Y	S	D	C	N	C	O	P	X	G	O	C	K	S	M	
	2	3	0	1	2	2	6	4	3	8	3	8	3	0	2	8	3	0	6	2	66
8	N	D	Y	C	X	Z	T	E	D	O	D	P	Q	Y	H	P	D	L	T	N	
	8	4	2	3	0	0	9	13	4	8	4	3	0	2	3	3	4	4	9	8	91

Fig. 19 (continued)

Fig. 19 (continued)

Column 1 (continued)

Generatrix																					Frequency Value
9	O	E	Z	D	Y	A	U	F	E	P	E	Q	R	Z	I	Q	E	M	U	O	
	8	13	0	4	2	7	3	3	13	3	13	0	8	0	7	0	13	2	3	8	110
10	P	F	A	E	Z	B	V	G	F	Q	F	R	S	A	J	R	F	N	V	P	
	3	3	7	13	0	1	2	2	3	0	3	8	6	7	0	8	3	8	2	3	82
11	Q	G	B	F	A	C	W	H	G	R	G	S	T	B	K	S	G	O	W	Q	
	0	2	1	3	7	3	2	3	2	8	2	6	9	1	0	6	2	8	2	0	67
12	R	H	C	G	B	D	X	I	H	S	H	T	U	C	L	T	H	P	X	R	
	8	3	3	2	1	4	0	7	3	6	3	9	3	3	4	9	3	3	0	8	82
13	S	I	D	H	C	E	Y	J	I	T	I	U	V	D	M	U	I	Q	Y	S	
	6	7	4	3	3	13	2	0	7	9	7	3	2	4	2	3	7	0	2	6	90
14	T	J	E	I	D	F	Z	K	J	U	J	V	W	E	N	V	J	R	Z	T	
	9	0	13	7	4	3	0	0	0	3	0	2	2	13	8	2	0	8	0	9	83
15	U	K	F	J	E	G	A	L	K	V	K	W	X	F	O	W	K	S	A	U	
	3	0	3	0	13	2	7	4	0	2	0	2	0	3	8	2	0	6	7	3	65
16	V	L	G	K	F	H	B	M	L	W	L	X	Y	G	P	X	L	T	B	V	
	2	4	2	0	3	3	1	2	4	2	4	0	2	2	3	0	4	9	1	2	50
17	W	M	H	L	G	I	C	N	M	X	M	Y	Z	H	Q	Y	M	U	C	W	
	2	2	3	4	2	7	3	8	2	0	2	2	0	3	0	2	2	3	3	2	52
18	X	N	I	M	H	J	D	O	N	Y	N	Z	A	I	R	Z	N	V	D	X	
	0	8	7	2	3	0	4	8	8	2	8	0	7	7	8	0	8	2	4	0	86
19	Y	O	J	N	I	K	E	P	O	Z	O	A	B	J	S	A	O	W	E	Y	
	2	8	0	8	7	0	13	3	8	0	8	7	1	0	6	7	8	2	13	2	103
20	Z	P	K	O	J	L	F	Q	P	A	P	B	C	K	T	B	P	X	F	Z	
	0	3	0	8	0	4	3	0	3	7	3	1	3	0	9	1	3	0	3	0	51
21	A	Q	L	P	K	M	G	R	Q	B	Q	C	D	L	U	C	Q	Y	G	A	
	7	0	4	3	0	2	2	8	0	1	0	3	4	4	3	3	0	2	2	7	55
22	B	R	M	Q	L	N	H	S	R	C	R	D	E	M	V	D	R	Z	H	B	
	1	8	2	0	4	8	3	6	8	3	8	4	13	2	2	4	8	0	3	1	88

Fig. 19 (continued)

Column 1 (continued)

Generatrix																					Frequency Value
23	C	S	N	R	M	O	I	T	S	D	S	E	F	N	W	E	S	A	I	C	
	3	6	8	8	2	8	7	9	6	4	6	13	3	8	2	13	6	7	7	3	129
24	D	T	O	S	N	P	J	U	T	E	T	F	G	O	X	F	T	B	J	D	
	4	9	8	6	8	3	0	3	9	13	9	3	2	8	0	3	9	1	0	4	102
25	E	U	P	T	O	Q	K	V	U	F	U	G	H	P	Y	G	U	C	K	E	
	13	3	3	9	8	0	0	2	3	3	3	2	3	3	2	2	3	3	0	13	78

Column 2

Generatrix																					Frequency Value
0	N	P	D	N	N	M	U	G	S	H	G	W	Q	E	N	C	N	S	B	Z	
	8	3	4	8	8	2	3	2	6	3	2	2	0	13	8	3	8	6	1	0	90
1	O	Q	E	O	O	N	V	H	T	I	H	X	R	F	O	D	O	T	C	A	
	8	0	13	8	8	8	2	3	9	7	3	0	8	3	8	4	8	9	3	7	119
2	P	R	F	P	P	O	W	I	U	J	I	Y	S	G	P	E	P	U	D	B	
	3	8	3	3	3	8	2	7	3	0	7	2	6	2	3	13	3	3	4	1	84
3	Q	S	G	Q	Q	P	X	J	V	K	J	Z	T	H	Q	F	Q	V	E	C	
	0	6	2	0	0	3	0	0	2	0	0	0	9	3	0	3	0	2	13	3	46
4	R	T	H	R	R	Q	Y	K	W	L	K	A	U	I	R	G	R	W	F	D	
	8	9	3	8	8	0	2	0	2	4	0	7	3	7	8	2	8	2	3	4	88
5	S	U	I	S	S	R	Z	L	X	M	L	B	V	J	S	H	S	X	G	E	
	6	3	7	6	6	8	0	4	0	2	4	1	2	0	6	3	6	0	2	13	79
6	T	V	J	T	T	S	A	M	Y	N	M	C	W	K	T	I	T	Y	H	F	
	9	2	0	9	9	6	7	2	2	8	2	3	2	0	9	7	9	2	3	3	94
7	U	W	K	U	U	T	B	N	Z	O	N	D	X	L	U	J	U	Z	I	G	
	3	2	0	3	3	9	1	8	0	8	8	4	0	4	3	0	3	0	7	2	68

Fig. 19 (continued)

Column 2 (continued)

Generatrix																					Frequency Value
8	V	X	L	V	V	U	G	O	A	P	O	E	Y	M	V	K	V	A	J	H	
	2	0	4	2	2	3	3	8	7	3	8	13	2	2	2	0	2	7	0	3	73
9	W	Y	M	W	W	V	D	P	B	Q	P	F	Z	N	W	L	W	B	K	I	
	2	2	2	2	2	2	4	3	1	0	3	3	0	8	2	4	2	1	0	7	50
10	X	Z	N	X	X	W	E	Q	C	R	Q	G	A	O	X	M	X	C	L	J	
	0	0	8	0	0	2	13	0	3	8	0	2	7	8	0	2	0	3	4	0	60
11	Y	A	O	Y	Y	X	F	R	D	S	R	H	B	P	Y	N	Y	D	M	K	
	2	7	8	2	2	0	3	8	4	6	8	3	1	3	2	8	2	4	2	0	75
12	Z	B	P	Z	Z	Y	G	S	E	T	S	I	C	Q	Z	O	Z	E	N	L	
	0	1	3	0	0	2	2	6	13	9	6	7	3	0	0	8	0	13	8	4	85
13	A	C	Q	A	A	Z	H	T	F	U	T	J	D	R	A	P	A	F	O	M	
	7	3	0	7	7	0	3	9	3	3	9	0	4	8	7	3	7	3	8	2	93
14	B	D	R	B	B	A	I	U	G	V	U	K	E	S	B	Q	B	G	P	N	
	1	4	8	1	1	7	7	3	2	2	3	0	13	6	1	0	1	2	3	8	73
15	C	E	S	C	C	B	J	V	H	W	V	L	F	T	C	R	C	H	Q	O	
	3	13	6	3	3	1	0	2	3	2	2	4	3	9	3	8	3	3	0	8	79
16	D	F	T	D	D	C	K	W	I	X	W	M	G	U	D	S	D	I	R	P	
	4	3	9	4	4	3	0	2	7	0	2	2	2	3	4	6	4	7	8	3	77
17	E	G	U	E	E	D	L	X	J	Y	X	N	H	V	E	T	E	J	S	Q	
	13	2	3	13	13	4	4	0	0	2	0	8	3	2	13	9	13	0	6	0	108
18	F	H	V	F	F	E	M	Y	K	Z	Y	O	I	W	F	U	F	K	T	R	
	3	3	2	3	3	13	2	2	0	0	2	8	7	2	3	3	3	0	9	8	76
19	G	I	W	G	G	F	N	Z	L	A	Z	P	J	X	G	V	G	L	U	S	
	2	7	2	2	2	3	8	0	4	7	0	3	0	0	2	2	2	4	3	6	59
20	H	J	X	H	H	G	O	A	M	B	A	Q	K	Y	H	W	H	M	V	T	
	3	0	0	3	3	2	8	7	2	1	7	0	0	2	3	2	3	2	2	9	59
21	I	K	Y	I	I	H	P	B	N	C	B	R	L	Z	I	X	I	N	W	U	
	7	0	2	7	7	3	3	1	8	3	1	8	4	0	7	0	7	8	2	3	81

Fig. 19 (continued)

Column 2 (continued)

Generatrix																					Frequency Value
22	J	L	Z	J	J	I	Q	C	O	D	C	S	M	A	J	Y	J	O	X	V	
	0	4	0	0	0	7	0	3	8	4	3	6	2	7	0	2	0	8	0	2	56
23	K	M	A	K	K	J	R	D	P	E	D	T	N	B	K	Z	K	P	Y	W	
	0	2	7	0	0	0	8	4	3	13	4	9	8	1	0	0	0	3	2	2	66
24	L	N	B	L	L	K	S	E	Q	F	E	U	O	C	L	A	L	Q	Z	X	
	4	8	1	4	4	0	6	13	0	3	13	3	8	3	4	7	4	0	0	0	85
25	M	O	C	M	M	L	T	F	R	G	F	V	P	D	M	B	M	R	A	Y	
	2	8	3	2	2	4	9	3	8	2	3	2	3	4	2	1	2	8	7	2	77

d. It will be noted that the frequency value of the 23d generatrix for the first column of cipher letters is the greatest value; that of the first generatrix for the second column is the greatest. In both cases these are the correct generatrices. Thus the selection of the correct generatrices in such cases has been reduced to a purely mathematical basis which is at times of much assistance in effecting a quick solution. Moreover, an understanding of the principles involved will be of considerable value in subsequent work.

25. Solution by the probable-word method. - a. Occasionally one may encounter a cryptogram which is so short that it contains no recurrences of even digraphs, and thus gives no indications of the number of alphabets involved. If the sliding mixed component is known one may apply the method illustrated in Par. 15, assuming the presence of a probable word, and checking it against the text and the sliding components to establish a key, if the correspondents are using key words.

b. For example, suppose that the presence of the word ENEMY is assumed in the message in Par. 23b above. One proceeds to check it against an unknown key word, using the already reconstructed mixed component sliding against the normal and starting with the first letter of the cryptogram in this manner:

If SFDZR equals ENEMY, then the successive equivalents of A_p equal XENFW.

The sequence XENFW spells no intelligible word. Therefore one shifts the location of the assumed word ENEMY one letter forward in the cipher text, and the test is made again, just as was explained on page 23.

When the group AQRLU is tried one obtains as the key letters ZIMUT, which, taken as a part of a word, suggests the word AZIMUTH. The method must yield solution when a correct word is assumed and correctly placed.

c. The danger to cryptographic security resulting from the inclusion of cryptographed addresses and signatures in cryptographic messages is directly connected with the principles of solution by the probable-word method. To illustrate, reference is made to the message employed in Pars. 19-22. It will be noted in Par. 22 b that the message carried a signature (Treer, Col.) and that the latter was enciphered. Suppose that this were an authorized practice, and that every message could be assumed to conclude with a cryptographed signature. The signature "TREER COL" would at once afford a very good basis for the quick solution of subsequent messages emanating from the same headquarters as did the first message, because presumably this same signature would appear in other messages. It is for this reason that addresses and signatures must not be cryptographed; if they must be included they should be cryptographed in a totally different system or by a wholly different method, perhaps by means of a special address and signature code. It would be best, however, to omit all addresses and signatures, and to let the call signs of the headquarters concerned also convey these parts of the message, leaving the distribution or delivery to the offices concerned a matter for local action.

26. Solution when the plain component is a mixed sequence, the cipher component, the normal. - a. This falls under Case B (2) outlined in Par. 6. It is not the usual method of employing a single mixed component, but may be encountered occasionally in cipher devices.

b. The preliminary steps, as regards factoring to determine the length of the period, are the same as usual. The message is then transcribed into its periods. Frequency distributions are then made, as usual, and these are attacked by the principles of frequency and recurrence. An attempt is made to apply the principles of direct symmetry of position, but this attempt will be futile, for the reason that the plain component is in this case an unknown mixed sequence. (See Par. 18 d.) Any attempt to find symmetry in the secondary alphabets based upon the normal sequence can therefore disclose no symmetry because the symmetry which exists is based upon a wholly different sequence.

c. However, if the principles of direct symmetry of position are of no avail in this case, there are certain other principles of symmetry which may be employed to great advantage. To explain them an actual example will be used. Let it be assumed that it is known to the cryptanalyst that the enemy is using the general system under discussion, viz, a mixed sequence variable from day to day is used as plain component, the normal sequence is used as cipher component, and a repeating key, variable from message to message, is used in the ordinary manner.

The following message has been intercepted:

QEOVK LRMLZ JVGTG NDLVK EVNTY ERMUE

VRZMO YAAMP DKEIJ SFMYO YHMME GQAMB

UQAXR HUFBU KQYMU NELVT KQILE KZBUE

ULIBK NDAXB XUDGL LADVK POAYO DKKYK

LADHY BVNFV UEEME FFMTE GVWBY TVDZL

SPBHB XVAZC UDYUE LKMMA EUDDK NCFSH

HSAHY TMGUJ HQXPP DKOUE XUQVB FVWBX

NXALB TCDLM IVAAA NSZIL OVWVP YAGZL

SHMME GQDHO YHIVP NCRRE XKDQZ GKNCG

NQGUY JIWYY TMAHW XRLBL OADLG NQGUY

JUUGB JHRVX ERFLE GWGUO XEDTP DKEIZ

VXNWA FAANE MKGHB SSNLO KJCBZ TGGLO

PKMBX HGERY TMWLZ NQCYY TMWIP DKATE

FLNUJ NDTVX JRZTL OPAHC DFZYY DEYCL

GPGTY TECXB HQEBR KVWMU NINGJ IQDLP

JKATE GUWBR HUQWM VRQBW YRFBF KMWMB

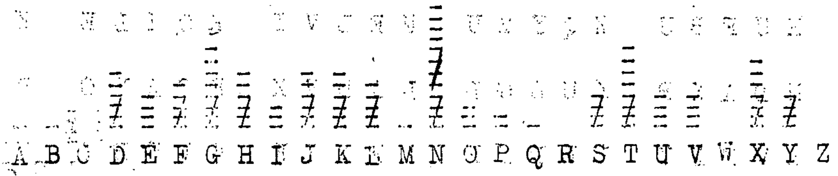
TMULZ LAAHY JGDVK LKRRE XKNAO NDSBX

XCGZA HDGTL VKMBW ISAUE FDNWP NLZIJ

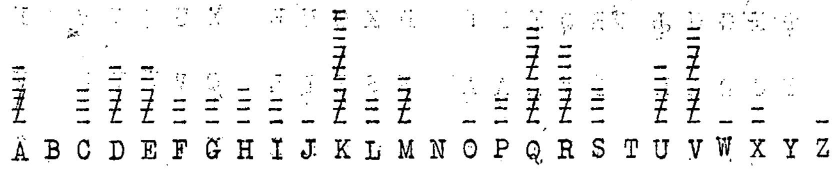
SRQZL AVNHL GVWVK FIGHP GECZU KQAP

d. A study of the recurrences and factoring their intervals discloses that five alphabets are involved. Monoliteral bar frequency distributions are made and are as follows:

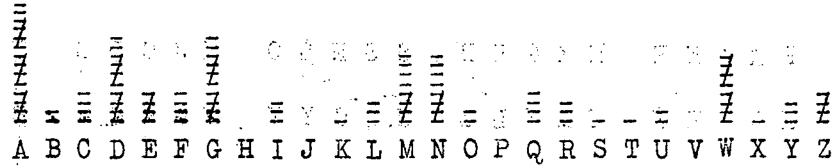
Alphabet 1.



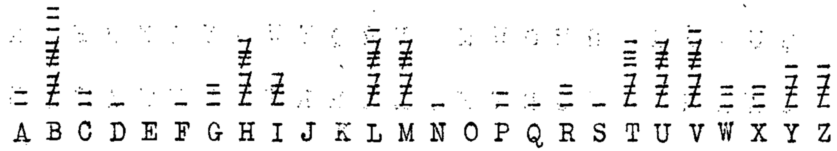
Alphabet 2.



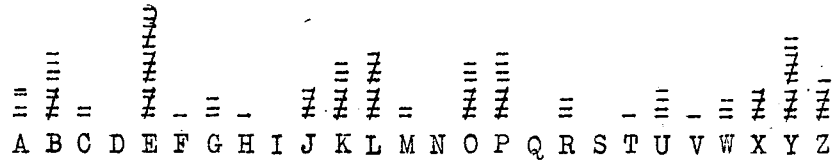
Alphabet 3.



Alphabet 4.



Alphabet 5.



e. Since the cipher component in this case is the normal alphabet, it follows that the five frequency distributions are based upon a sequence which is known. Therefore the five frequency distributions should manifest a direct symmetry of distribution of crests and troughs. By shifting the five distributions relative to one another, all five can be matched as regards the positions of the crests and troughs, thus reducing the five distributions to a single equivalent monoalphabetic distribution. Note how this has been done in the case of the five illustrative distributions:

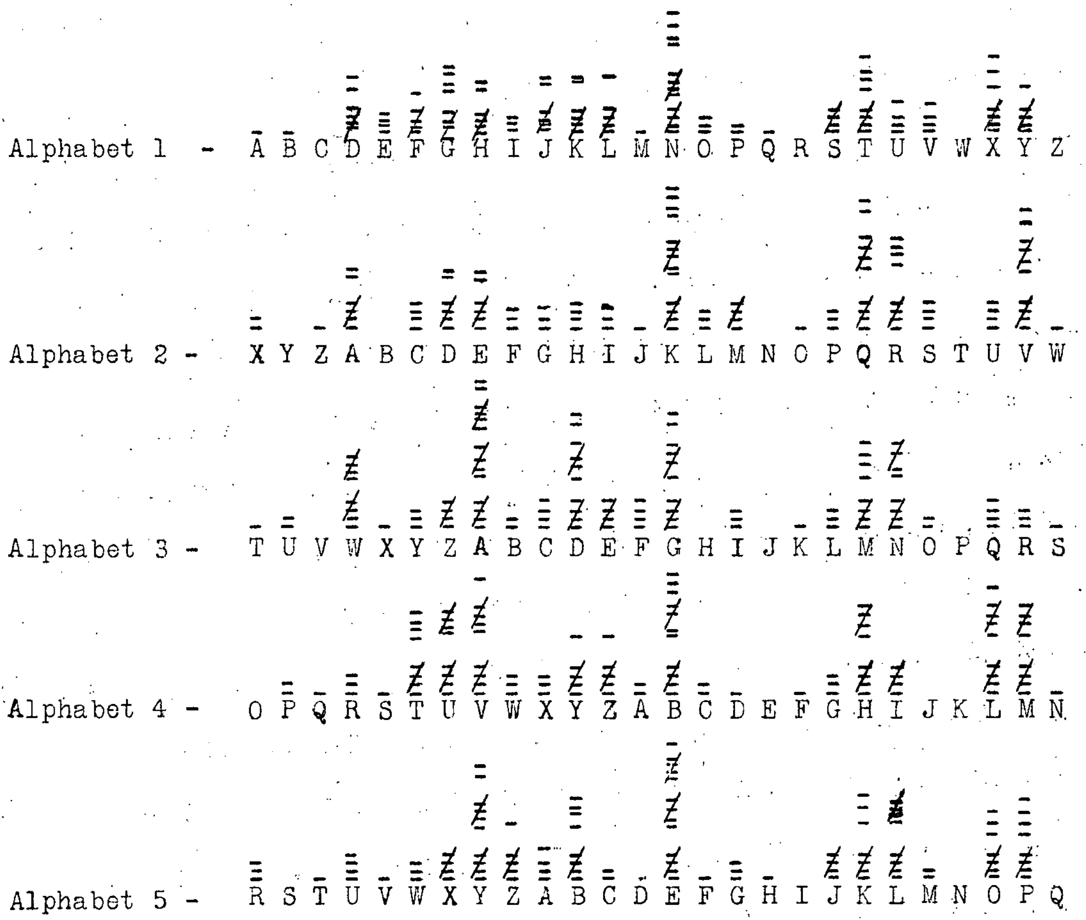
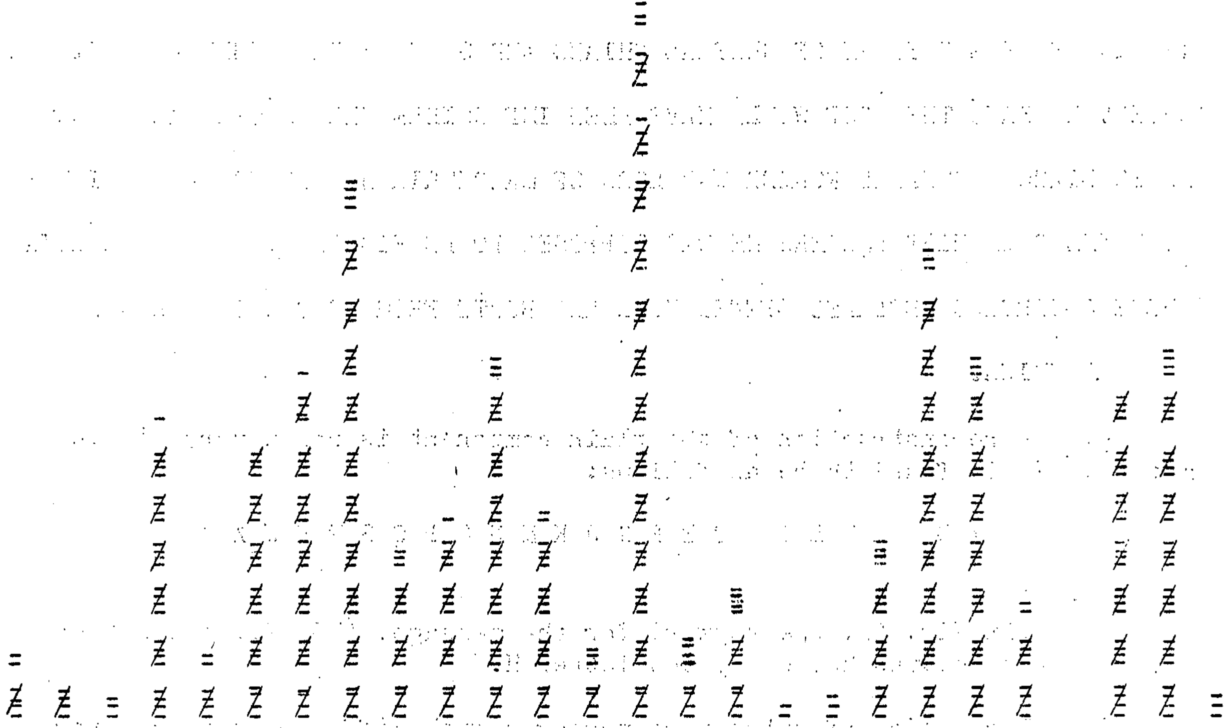


Fig. 21.

f. The superimposition of the respective distributions enables one to convert the cipher letters of the five alphabets into one alphabet. Suppose it is decided to convert alphabets 2, 3, 4, and 5 into alphabet 1. It is merely necessary to substitute for the respective letters in the four alphabets those which stand above them in Alphabet 1. For example, in Fig. 21, X_c in Alphabet 2 is directly under A_c in Alphabet 1; hence, if the superimposition is correct then $\underset{X_c}{2} = \underset{A_c}{1}$. Therefore, in the cryptogram it is merely necessary to replace every X_c in the second position by A_c. Again T_c in Alphabet 3 = A_c in Alphabet 1; therefore, in the cryptogram one replaces every T_c in the third position by A_c. The entire process gives the following converted message:

```
Q H V H T  L U T X L  J Y N F P  N G S H T  E Y U F H  E U T G N  V U G Y X

Y D H Y Y  D N L U S  S I T K X  Y K T Y N  G T H Y K  U T H J A  H X M N D

K T F Y D  N H S H C  K T P X N  K C I G N  U O P N T  N G H J K  X X K S U

L D K H T  P R H K X  D N R K T  L D K T H  B Y U R E  U H L Y N  F I T F N

G Y D N H  T Y K L U  S S I T K  X Y H L L  U G F G N  L N T Y J  E X K P T

N F M E Q  H V H T H  T P N G S  H T E B Y  D N V G N  X X X H K  F Y D N G

N A H X K  T F K X V  I Y H M J  N V G U U  O Y D H Y  Y D N L U  S K T Y N

G T K T X  Y K P H Y  N F Y D N  X N K C I  G N U O P  N T N G H  J L D K H

T P H T F  X U S N U  O D K X P  N T N G H  J X B S K  J K Y H G  E U M X N

G Z N G X  X H K F Y  D N L U I  V A U I J  F D H Z N  M N N T K  S V U X X

K M J N I  T J N X X  P N T N G  H J L D H  T P D X I  N T J K H  T P D U Y

D N H F N  F O U G S  N G A H G  J U G F U  O S H T L  D I G K H  D H F O U

G S N F H  T H J J K  H T L N A  K Y D Y D  N L U S S  I T K X Y  J N H F N

G X D N A  H X X I V  V U X N F  Y U M N O  K P D Y K  T P B X I  L D H T H

J J K H T  L N Y D N  X N U M X  N G Z N G  X F N L J  H G N F U  V N T N F

I V H G N  F G U I Y  N O G U S  S U X L U  A Y U T U  G Y D H T  F L N T Y

G H J L D  K T H B
```

The monoalphabetic frequency for this follows. Note that the frequency of each letter is the sum of the five frequencies in the corresponding columns of Fig. 21.



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
7	5	3	31	7	30	36	54	18	21	38	22	3	69	9	14	2	3	19	48	38	12	-	35	39	3

Fig. 21a.

g. The problem having been reduced to monoalphabetic terms, a trigraphic frequency distribution can now be made and solution readily attained by simple principles. It yields the following:

JAPAN CONSULTED GERMANY TODAY ON REPORTS THAT THE COMMUNIST INTERNATIONAL WAS BEHIND THE AMAZING SEIZURE OF GENERALISSIMO CHIANG KAI SHEK IN CHINA. TOKYO ACTED UNDER THE ANTICOMMUNIST ACCORD RECENTLY SIGNED BY JAPAN AND GERMANY. THE PRESS SAID THERE WAS INDISPUTABLE PROOF THAT THE COMINTERN INSTIGATED THE SEIZURE OF GENERAL CHIANG AND SOME OF HIS GENERALS. MILITARY OBSERVERS SAID THE COUP WOULD HAVE BEEN IMPOSSIBLE UNLESS GENERAL CHANG HSUEN LIANG HOTHEADED FORMER WAR LORD OF MANCHURIA HAD FORMED AN ALLIANCE WITH THE COMMUNIST LEADERS HE WAS SUPPOSED TO BE FIGHTING. SUCH AN ALLIANCE THESE OBSERVERS DECLARED OPENED UP A RED ROUTE FROM MOSCOW TO NORTH AND CENTRAL CHINA.

h. The reconstruction of the plain component is now a very simple matter. It is found to be as follows:

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

Note also, in Fig. 21, the keyword for the message, (HEAVY), the letters being in the columns headed by the letter H.

i. The solution of subsequent messages with different keys can now be reached directly, by a simple modification of the principles explained in Paragraph 18. This modification consists in using for the completion the mixed plain component (now known) instead of the normal alphabet, after the cipher letters have been converted into their plain component equivalents. Let the student confirm this by an experiment.

j. The probable-word method of solution discussed under Paragraph 20 is also applicable here, in case of very short cryptograms. This method presupposes of course, possession of the mixed component; the procedure is essentially the same as that in Paragraph 20. In the example discussed in the present paragraph, the letter A on the plain component was successfully set against the key letters HEAVY; but this is not the only possible procedure.

k. The student should go over carefully the principle of "conversion into monoalphabetic terms" explained in subparagraph f above until he thoroughly understands it. Later on he will encounter cases in which this principle is of very great assistance in the cryptanalysis of more complex problems.

SECTION VI.

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, II.

27. Further cases to be considered. - a. Thus far Cases B (1) and (2), mentioned in Paragraph 3 have been treated. There remains Case B (3) to be studied. This case has been further subdivided as follows:

Case B (3). Both components are mixed sequences.

- (a) Components are identical mixed sequences.
    - (1) Sequences proceed in the same direction. (The secondary alphabets are mixed alphabets).
    - (2) Sequences proceed in opposite directions. (The secondary alphabets are reciprocal mixed alphabets).
- (b) Components are different mixed sequences. (The secondary alphabets are mixed alphabets).

b. The first of the foregoing subcases will now be examined.

28. Identical, primary mixed components proceeding in the same direction. - a. It is often the case that the mixed components are derived from an easily remembered word or phrase, so that they can be reproduced at any time from memory. Thus, for example, given the key word QUESTIONABLY, the following mixed sequence is derived:

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

b. By using this sequence as both plain and cipher component, that is, sliding this sequence against itself, a series of 26 secondary mixed alphabets may be produced. For example, by setting the two sliding strips against each other in the two positions shown below, the cipher alphabets labeled (1) and (2) given by the two settings are seen to be different.

Key letter = A (i. e., $Q_p = A_c$).

Plain component.
↓
QUESTIONABLYCDFGHJKMPRVWXZQUESTIONABLYCDFGHJKMPRVWXZ
QUESTIONABLYCDFGHJKMPRVWXZ
↑
Cipher component.

Secondary alphabet:

Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ
(1) Cipher - HJPRLVWXDZQKUGFEASYCBTIOMN

Key letter = B (i. e., $Q_p = B_c$).

Plain component.
↓
QUESTIONABLYCDFGHJKMPRVWXZQUESTIONABLYCDFGHJKMPRVWXZ.
QUESTIONABLYCDFGHJKMPRVWXZ
↑
Cipher component.

Secondary alphabet:

Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ
(2) Cipher - JKRVYWXZFQUMEHGSBTCDLIONPA

c. In enciphering a message by such sliding strips, a key word is used to designate the particular positions in which the strips are to be set, the same as was the case in previous examples of the use of sliding components. The method of designating the positions is, however, slightly different, the reasons for which will appear in the succeeding paragraph. In the methods heretofore given, the key letter, as located on the cipher component, was set opposite A, as located on the plain component; in other words, if A was the key letter, then the two sliding strips were set so that $A_p = A_c$. In this case, however, where identical mixed sliding components are used, the key letter is set opposite the first letter of the sequence upon which the primary components are based; that is, if A is the key letter, then the sliding strips are set so that $Q_p = A_c$ in the case of the mixed components shown above. Hence, in the first of the two examples above, the key letter for the first example being A, then A_c is set opposite Q_p; in the second of these examples, the key letter being B, then B_c is set opposite Q_p.

d. Very frequently a quadricular or square table is employed by the correspondents, instead of sliding strips, but the results are the same. The square table based upon the word QUESTIONABLY is shown in Table 6. It will be noted that the table does nothing more than set forth the successive positions of the two primary sliding components, and the top line of the table is the plain component, the successive horizontal lines below it, the cipher component in its various juxtapositions. The usual method of employing such a table is to take as the cipher equivalent of a plain-text letter that letter which lies at the intersection of the vertical column headed by the plain-text letter and the horizontal row begun by the key letter. For example, the cipher equivalent of E_p with key letter T is the letter O_c or E_p (T_k) = O_c. The method given in paragraph b, for determining the cipher equivalents by means of the two sliding strips yields the same results as does the square table.

TABLE 6.

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

29. Cryptographing and decryptographing by identical, primary mixed components. - There is nothing of special interest to be noted in connection with the use either of identical mixed components of an equivalent quadricular table such as that shown in Table 6, in enciphering or deciphering a message. The basic principles are the same as in the case of the sliding of one mixed component against the normal changeable keywords of varying lengths. The components may be changed at will and so on. All this has been demonstrated adequately enough in Special Text No. 165, Elementary Military Cryptography.

30. Principles of solution. - a. Basically the principles of solution in the case of a cryptogram enciphered by two identical mixed sliding components are the same as in the preceding case. Primary recourse is had to the principles of frequency and repetition of single letters, digraphs, trigraphs, and polygraphs. Once an entering wedge has been forced into the problem, the subsequent steps may consist merely in continuing along the same lines as before, building up the solution bit by bit.

b. Doutless the question has already arisen in the student's mind as to whether any principles of symmetry of position can be used to assist in the solution and in the reconstruction of the cipher alphabets in cases of this kind under consideration. This phase of the subject will be taken up in the next section and will be treated in a somewhat detailed manner, because the theory and principles involved are of very wide application in cryptanalytics.

SECTION VII.

THEORY OF INDIRECT SYMMETRY OF POSITION IN SECONDARY ALPHABETS.

	Par.
Reconstruction of primary components from secondary alphabets	31

31. Reconstruction of primary components from secondary alphabets.

a. Note the two secondary alphabets (1) and (2) given in paragraph 27b. Externally they show no resemblance or symmetry despite the fact that they were produced from the same primary components. Nevertheless, when the matter is studied with care, a symmetry of position is discoverable. Because it is a hidden or latent phenomenon, it may be termed latent symmetry of position. However, in previous texts the phenomenon has been designated as an indirect symmetry of position and this terminology has grown into usage so that a change is perhaps now inadvisable. Indirect symmetry of position is a very interesting and exceedingly useful phenomenon in cryptanalytics.

b. Consider the following secondary alphabet (the one labeled (2) in paragraph 27b):

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

c. Assuming it to be known that this is a secondary alphabet produced by two primary identical mixed components, it is desired to reconstruct the latter. Construct a chain of alternate $\theta_p - \theta_c - \theta_p$ values, beginning at any point and continuing until the chain has been completed. Thus, for example, beginning with $A_p = J_c$, $J_p = Q_c$, $Q_p = B_c$, and dropping out the letters common to successive pairs, there results the sequence A J Q B . . . By completing the chain the following sequence of letters is established:

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

d. This sequence consists of 26 letters, and when slid against itself will produce exactly the same secondary alphabets as do the primary components based upon the word QUESTIONABLY. To demonstrate that this is the case, compare the secondary alphabets given by the two settings of the externally different components shown below:

```
                      Plain component.
                           ↓
QUESTIONABLYCDFGHJKMPRVWXZQUESTIONABLYCDFGHJKMPRVWXZ
                          QUESTIONABLYCDFGHJKMPRVWXZ
                             ↑
                       Cipher component.
```

Secondary alphabet:

```
         Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ
(1)     Cipher - JKRVYWXZFQUMEHGSBTCDLIONPA
```

```
                      Plain component.
                           ↓
AJQBKULMEYPSCRTDVIFWOGXNHZAJQBKULMEYPSCRTDVIFWOGXNHZ
                          AJQBKULMEYPSCRTDVIFWOGXNHZ
                                    ↑
                              Cipher component.
```

Secondary alphabet:

```
         Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ
(2)     Cipher - JKRVYWXZFQUMEHGSBTCDLIONPA
```

e. Since the sequence A J Q B K . . . gives exactly the same equivalents in the secondary alphabets as the sequence Q U E S T . . . gives, it is termed an equivalent primary component. If the real or original primary component is a key-word mixed sequence, it is hidden or latent within the equivalent primary sequence; it can be made patent by decimation of the equivalent primary component. Find three letters in the equivalent primary component such as are likely to have formed an unbroken sequence in the original primary component, and see if the interval between the first and second is the same as that between the second and third. Such a case is presented by the letters W, X, and Z in the equivalent primary component above; the distance or interval between them is two letters. Continuing the chain by adding letters two intervals removed, the latent original primary component is made patent.

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

f. It is possible to perform the steps given in c and e in a combined single operation when it is suspected that the original primary component is a key-word mixed sequence. Starting with any pair of letters (in the cipher component of the secondary alphabet) likely to be sequent in the key-word mixed sequence, such as JK_c in the secondary alphabet labeled (2), the following chain of digraphs may be set up. Thus, J, K, in the plain component stand over Q, U, respectively, in the cipher component; Q, U, in the plain component stand over B, L, respectively, in the cipher component, and so on. Connecting the pairs in a series, the following results are obtained:

JK - QU - BL - KM - UE - LY - MP - ES - YC - PR - ST - CD - RV -

TI - DF - VW - IO - FG - WX - ON - GH - XZ - NA - HJ - ZQ - AB -

These may now be united by means of their common letters:

JK - KM - MP - PR - RV - etc. = JKMPRVWXZQUESTIONABLYCDFGH

The original primary component is thus completely reconstructed.

g. Not all of the 26 secondary alphabets of the series yielded by two sliding primary components may be used to develop a complete equivalent primary component. If examination be made, it will be found that only 13 of these secondary alphabets will yield complete equivalent primary components when the method of reconstruction shown in subparagraph c above is followed. For example, the following secondary alphabet, which is also derived from the primary components based upon the word QUESTIONABLY will not yield a complete chain of 26 plain text-cipher-plain text equivalents:

Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher - CDHJOKMPBRVFWYLXTZNAIQUEGS

Equivalent primary component:

A C H P X E O L F K V Q T A C H . . (The A C H sequence begins again).

h. It is seen that only 13 letters of the chain have been established before the sequence begins to repeat itself. It is evident that exactly one-half of the chain has been established. The other half may be established by beginning with a letter not in the first half. Thus:

B D J R Z S N Y G M W U I B D J . . . (The B D J sequence begins again).

i. It is not necessary to distribute the letters of each half-sequence within 26 spaces, to correspond with their placements in a complete alphabet. This can only be done by allowing between the letters of one of the half-sequences a constant odd number of spaces. Distributions are therefore made upon the basis of 3, 5, 7, 9, . . . spaces. Select that distribution which most nearly coincides with the distribution to be expected in a key-word component. Thus, for example, with the first half-sequence the distribution selected is the one made by leaving three spaces between the letters; it is as follows:

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
A	-	L	-	C	-	F	-	H	-	K	-	P	-	V	-	X	-	Q	-	E	-	T	-	O	-

j. Now interpolate, by the same constant interval (three in this case), the letters of the other half-sequence. Noting that the group F - H appears in the foregoing distribution, it is apparent that G of the second half-sequence should be inserted between F and H. The letter which immediately follows G in the second half-sequence, viz, M, is next inserted in the position three spaces to the right of G, and so on, until the interpolation has been completed. This yields the original primary component, which is as follows:

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

k. Another method of handling cases such as the foregoing is indicated in subparagraph f. By extending the principles set forth in that subparagraph, one may reconstruct the following chain of 13 pairs from the secondary alphabet given in subparagraph g:

CD - HJ - PR - XZ - ES - ON - LY - FG - KM - VW - QU - TI - AB - CD

Now find, in the foregoing chain, two pairs likely to be sequent, for example HJ and KM and count the interval between them in the chain. It is 7 (counting by pairs). If this decimation interval is now applied to the chain of pairs, the following is established:

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

l. The reason why a complete chain of 26 letters cannot be constructed from the secondary alphabet given under subparagraph g is that it represents a case in which two primary components of 26 letters were slid an even number of intervals apart. There are in all 12 such cases, none of which will admit of the construction of a complete chain of 26 letters. In addition, there is one case wherein, dispite the fact that the primary components are an odd number of intervals apart, the secondary alphabet cannot be made to yield a complete chain of 26 letters for an equivalent primary component. This is the case in which the displacement is 13 intervals. Note the following secondary alphabet based upon the primary components shown in subparagraph d:

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

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

m. If an attempt is made to construct a chain of letters from this secondary alphabet alone, no progress can be made because the alphabet is completely reciprocal. However, the cryptanalyst need not at all be baffled by this case. The attack will follow along the lines shown below in subparagraphs n and o.

n. If the original primary component is a key-word mixed sequence, the cryptanalyst may reconstruct it by attempting to "dovetail" the 13 reciprocal pairs (AR, BV, CZ, DQ, EG, FU, HS, IK, JT, LW, MO, NP, and XY) into one sequence. The members of these pairs are all 13 intervals apart. Thus:

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

Fig.22.

Write out the series of numbers from 1 to 26 and insert as many pairs into position as possible, being guided by considerations of probable sequence in the key-word mixed sequence. Thus:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	B	C	D	.	.	.	.	.	.	.	.	R	V	Z	Q

It begins to look as though the key-word commences with the letter Q, in which case it should be followed by U. This means that the next pair to be inserted is FU. Thus:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	B	C	D	F	.	.	.	.	.	.	.	R	V	Z	Q	U

The sequence A B C D F means that E is in the key. Perhaps the sequence is A B C D F G H. Upon trial, using the pairs EG and HS, the following placements are obtained:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
A	B	C	D	F	G	H	.	.	.	.	.	R	V	Z	Q	U	E	S

This suggests the word QUEST or QUESTION. The pair JT is added:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	B	C	D	F	G	H	J	.	.	.	.	R	V	Z	Q	U	E	S	T

The sequence G H J suggests G H J K, which places an I after T. Enough of the process has been shown to make the steps clear.

o. Another method of circumventing the difficulties introduced by the 14th secondary alphabet (displacement interval, 13) is to use it in conjunction with another secondary alphabet which is produced by an even-interval displacement. For example, suppose the following two secondary alphabets are available.

0 - 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
1 - R V Z Q G U E S K T I W O P M N D A H J F B L Y X C
2 - X Z E S K T I O R N A Q B W V L H Y M P J C D F U G

Fig. 23.

The first of these secondaries is the 13 interval secondary; the second is one of the even-interval secondaries, from which only half-chain sequences can be constructed. But if the construction be based upon the two sequences, 1 and 2 in the foregoing diagram, the following is obtained:

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

This is a complete equivalent primary component. The original key-word mixed component can be recovered from it by decimination based upon the 9th interval:

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

p. (1) When the primary components are identical mixed sequences proceeding in opposite directions, all the secondary alphabets will be reciprocal alphabets. Reconstruction of the primary component can be accomplished by the procedure indicated under subparagraph o above. Note the following three reciprocal secondary alphabets:

0 - 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
1 - P M H G Q F D C W Y L K B R V A E N Z X U O I T J S
2 - W V M K S J H G Q F D R C X Z Y I L E U T B A N P O
3 - T S S Z L X W V N R P E M I O K C J B A Y H G F U D

Fig. 24.

(2) Using lines 1 and 2 the following chain can be constructed (equivalent primary component):

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

Or, using lines 2 and 3:

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

The original key-word mixed primary component (based on the word QUESTIONABLY) can be recovered from either of the two foregoing equivalent primary components. But if lines 1 and 3 are used, only half-chains can be constructed:

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

This is because 1 and 3 are both odd-interval secondary alphabets, whereas 2 is an even-interval secondary. It may be added that odd-interval secondaries are characterized by having two cases in which $\theta_p = \theta_c$. (Note that in secondary number 1 above, $F_p = F_c$ and $U_p = U_c$; in secondary number 3 above, $M_p = M_c$ and $\theta_p = \theta_c$). This characteristic will enable the cryptanalyst to select at once the proper two secondaries to work with in case several are available; one should show two cases where $\theta_p = \theta_c$; the other should show none.

q. (1) When the primary components are different mixed sequences, their reconstruction from secondary cipher alphabets follows along the same lines as set forth under b to j inclusive, above, with the exception that the selection of letters for building up the chain of equivalents for the primary cipher component is restricted to those below the zero line. Having reconstructed the primary cipher component, the plain component can be readily reconstructed. This will become clear if the student will study the following example.

0 -	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
1 -	T	V	A	B	U	L	I	Q	X	Y	C	W	S	N	D	P	F	E	Z	G	R	H	J	K	M	O
2 -	Z	J	S	T	V	I	Q	R	M	O	N	K	X	E	A	G	B	W	P	L	H	Y	C	D	F	U

Fig. 25.

(2) Using only lines 1 and 2, the following chain is constructed:

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

This is an equivalent primary cipher component. By finding the values of the successive letters of this chain in terms of the plain component of the first secondary alphabet (the zero line), the following is obtained:

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

The sequence A S P T . . . is an equivalent primary plain component. The original key-word mixed components may be recovered from each of the equivalent primary components. That for the primary plain component is based upon the key PUBLISHERS MAGAZINE; that for the primary cipher component is based upon the key QUESTIONABLY.

(3) Another method of accomplishing the process indicated above can be illustrated graphically by the following two chains, based upon the two secondary alphabets set forth in subparagraph q (1):

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
0 -	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
1 -	T	V	A	B	U	L	I	Q	X	Y	C	W	S	N	D	P	F	E	Z	G	R	H	J	K	M	O
2 -	Z	J	S	T	V	I	Q	R	M	O	N	K	X	E	A	G	B	W	P	L	H	T	C	D	F	U

<u>Col. 1.</u> <u>Col. 2.</u>

A(0-1) → T(1-1); T(2-4) → D(0-4);

D(0-4) → B(1-4); B(2-17) → Q(0-17);

Q(0-17) → F(1-17); F(2-25) → Y(0-25);

Y(0-25) → M(1-25); M(2-9) → I(0-9);

I(0-9) → X(1-9); X(2-13) → M(0-13);

M(0-13) → S(1-13); S(2-3) → C(0-3);

etc. etc.

Fig. 26.

(4) By joining the letters in Column 1, the following chain is obtained: A D Q Y I M, etc. If this be examined, it will be found to be an equivalent primary of the sequence based upon P U B L I S H E R S M A G A Z I N E. By joining the letters in Column 2, the following chain is obtained: T B F M X S. This is an equivalent primary of the sequence based upon Q U E S T I O N A B L Y.

SECTION VIII.

APPLICATION OF PRINCIPLES OF INDIRECT SYMMETRY OF POSITION.

32. Applying the principles to a specific example. - a. The preceding section, with the many details covered, now forms a sufficient base for proceeding with an exposition of how the principles of indirect symmetry of position can be applied very early in the solution of a polyalphabetic substitution cipher in which sliding primary components were employed to produce the secondary cipher alphabets for the enciphering of the cryptogram.

b. The case described below will serve not only to explain the principles of the method of applying these principles but will at the same time show how the solution of a single, rather difficult, polyalphabetic substitution cipher can be greatly facilitated by applying these principles. It is realized, of course, that the cryptogram could be solved by the usual methods of frequency and long, patient experimentation. However, the method to be described was actually applied and very materially reduced the amount of time and labor that would otherwise have been required for solution.

33. The cryptogram employed in the exposition. - a. The problem that will be used in this exposition involves an actual cryptogram submitted for solution in connection with a cipher device having two concentric disks upon which the same random mixed alphabet appears, both alphabets progressing in the same direction. This was obtained from a study of the descriptive circular accompanying the cryptogram. By the usual process of factoring, it was determined that the cryptogram involved 10 alphabets. The message as arranged according to its period is shown in Figure 27, in which all repetitions of two or more letters are indicated.

b. The trigraphic frequence distributions are given in Figure 28. It will be seen that on account of the brevity of the message, considering the number of alphabets involved, the frequency distributions do not yield many clues. By a very careful study of the repetitions, tentative individual determinations of values of cipher letters, as illustrated in Figures 29, 30, 31, and 32, were made. These are given in sequence and in detail in order to show that there is nothing artificial or arbitrary in the preliminary stages of analysis here set forth.

THE CRYPTOGRAM

(Repetitions underlined)

	1	2	3	4	5	6	7	8	9	0
A	W	F	U	P	C	F	O	C	J	Y
B	G	B	Z	D	P	F	B	O	U	O
C	G	R	F	T	Z	M	Q	M	A	V
D	K	Z	U	G	D	Y	F	T	R	W
E	G	J	X	N	L	W	Y	O	U	X
F	I	K	W	E	P	Q	Z	O	K	Z
G	P	R	X	D	W	L	Z	I	C	W
H	G	K	Q	H	O	L	O	D	V	M
I	G	O	X	S	N	Z	H	A	S	E
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
M	S	R	Q	E	W	M	L	N	A	E
N	G	S	X	E	R	O	Z	J	S	E
O	G	V	Q	W	E	J	M	K	G	H
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
S	C	F	B	S	C	V	X	C	H	Q
T	Z	T	Z	S	D	M	X	W	C	M
U	R	K	U	H	E	Q	E	D	G	X
V	F	K	V	H	P	J	J	K	J	Y
W	Y	Q	D	P	C	J	X	L	L	L
X	G	H	X	E	R	O	Q	P	S	E
Y	G	K	B	W	T	L	F	D	U	Z
Z	O	C	D	H	W	M	Z	T	U	Z
A A	K	L	B	P	C	J	O	T	X	E
B B	H	S	P	O	P	N	M	D	L	M
C C	G	X	K	W	D	V	B	L	S	E
D D	G	S	U	G	D	P	O	T	H	X
E E	B	K	D	Z	F	M	T	G	Q	J
F F	L	F	U	Y	D	T	Z	V	H	Q
G G	Z	G	W	N	K	X	J	T	R	N
H H	Y	T	X	C	D	P	M	V	L	W
I I	B	G	B	W	W	O	Q	R	G	N
J J	H	H	V	L	A	Q	Q	V	A	V
K K	J	O	W	O	O	T	T	N	V	Q
L L	B	K	X	D	S	O	Z	R	S	N
M M	Y	U	X	O	P	P	Y	O	X	Z
N N	H	O	Z	O	W	M	X	C	G	Q
O O	J	J	U	G	D	W	Q	R	V	M
P P	U	K	W	P	E	F	X	E	N	F
Q Q	C	C	U	G	D	W	P	E	U	H
R R	Y	B	W	E	W	V	M	D	Y	J
S S	R	Z	X							

Fig. 27.

FIGURE 28.

Trigraphic Frequency Distributions.

I.

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
	EB	FF			XK	YB	ES	XK	ZC	VZ	WQ			ZC	ZR	DC	HC	HR		MK		-F		YQ	QT
	HZ	FC				OR	NH		VQ	ZL	JF						MK							NT	QG
	XK					WJ	ZO		QJ								JZ							NU	
	WG					WK																		HB	
	QK					MO																			
						ES																			
						EV																			
						LH																			
						EK																			
						MC																			
						ES																			

II.

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
	GZ	QB			WU	ZW	GX		GX	IW	KB			GX		LZ	GF	GX	ZZ	YX	GQ				KU
	BJ	JQ			CB	BB	HV		JU	GQ				HZ		YD	PX	HP	YX						BZ
	YW	RV			LU					RU						JW	SQ	GW							RH
		OD								FV															
		GK								GB															
		CU								BD															
										BX															
										UW															

III.

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
	CZ		QP		RT				BI	CW					SO	KH				FP	CO	KE	JN		BD
	FS		CH													CR				ZG	KH	GN	RD		QA
	KW		KZ													RE				KH	HL	QO	OS		ZC
	LP															VW				SG		KP	SE		TS
	GW																			FY		BE	HE		OO
																				JG			TC		
																				CG			KD		
																							UO		
																							Z-		

FIGURE 28 (Cont).

IV.

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
ZA		ZK	ZP	WP		UD	QO	JP			VA		XL	VP	UC		QQ	XN	FZ			QE		UD	BE
		XD	XW	QW		UD	UE						WK	PP	DC			BC				BT			DF
			XS	XR		UD	VP							WO	BC			ZD				KD			
				XR		UD	DW							XP	WE							BW			
				WW										ZW											

V.

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
AA		PF	GY	ZX	ZM					CQ	NW		SZ	HL	DF	RF	EO	DO	WL			DL			TM
LQ		SV	SM	WJ						NX				OT	EQ		EO					EM			
		PJ	WV	HA											IQ							HM			
		PJ	GP	PF											ON							WO			
			VT												HJ							OM			
			CP												ON							EV			
			GW												OP										
			GW																						

VI.

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
AM					CO				EM		WZ	ZQ	PB	RZ	DO	PZ			DZ		CX	LY	EQ	DF	NH
					PB				PJ		OO	WL	PM	RQ	DM	PF			OT		DB	DQ	KJ		
					QV				CX		TF	DX		WQ	PY	KO					WM	DP			
					EX				CO			WZ		SZ		EE									
												FT				AQ									
												WX													

FIGURE 28 (Cont).

VII.

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
	FO			QD	YT		ZA		JK		MN	JK		FC	WE	MM			MG		FM		VC	WO	QO
	NL				QJ				XT			AD		LD		XT			TN				MW	PO	LI
	VL				LD							ND		QI		OP							JL		OJ
												PV		JT		OR							MC		MT
												VD		PT		QV							FE		TV
																WR									OR

VIII.

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
HS		OJ	OV	XN		TQ		ZC	FH	MG	BC	QA	LA	BU	QS		QG		FR		ZH	XC			
		XH	MC	PU				OK	ZS	JJ	XL	VL	TV	YU			ZS		QX		ML				
		XG	EG								BS			ZK			QV		ZU		QA				
			FU											YX					OX						
			ML																OH						
			MY																JR						

IX.

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
MV		IW				KH	JD		CY	OZ	MH		EF			GJ	TW	AE		OO	DM		TZ	DJ	
NE		LW				DX	CQ		KY	IF	LL						TN	JE		OX	NQ		TE		
VV		DH				RN	TX				DM							PE		DZ	RM		OZ		
		WM				CQ	VQ				VW							LE		TZ					
																		RN		EH					

X.

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
			HQ	SB	KC		LS		QL		LG	VG	RY	UG		HZ					AK	RG	UI	JG	KP
				AG	NC		GR		YR			CR	GH			HZ					AJ	GG	GF	JY	XJ
				SG			CB					LG	SY			VB						CL	HB		UO
				SG			UY					VU				GJ						LB			UK
				XH																					XH
				SG																					

INITIAL VALUES FROM ASSUMPTIONS.

$\overset{1}{G}_c = E_p$; $\overset{2}{K}_c = E_p$; $\overset{3}{X}_c = E_p$; and $\overset{4}{D}_c = E_p$, from frequency considerations.

$\overset{345}{UGD} =$ THE; $\overset{456}{PCJ} =$ THE; and $\overset{901}{SEG} =$ THE, from study of repetitions.

	1	2	3	4	5	6	7	8	9	0
A	W	F	U T	P T	C H	F	O	C	J	Y
B	G E	B	Z	D	P	F	B	O	U	O
C	G E	R	F	T	Z	M	Q	M	A	V
D	K	Z	U T	G H	D E	Y	F	T	R	W
E	G E	J	X E	N	L	W	Y	O	U	X
F	I	K E	W	E	P	Q	Z	O	K	Z
G	P	R	X E	D	W	L	Z	I	C	W
H	G E	K E	Q	H	O	L	O	D	V	M
I	G E	O	X E	S	N	Z	H	A	S T	E H
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
M	S	R	Q	E	W	M	L	N	A	E H
N	G E	S	X E	E	R	O	Z	J	S T	E H
O	G E	V	Q	W	E	J E	M	K	G	H
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
S	C	F	B	S	C H	V	X	C	H	Q
T	Z	T	Z	S	D E	M	X	W	C	M
U	R	K E	U T	H	E	Q	E	D	G	X
V	F	K E	V	H	P	J E	J	K	J	Y
W	Y	Q	D	P T	C H	J E	X	L	L	L
X	G E	H	X E	E	R	O	Q	P	S T	E H
Y	G E	K E	B	W	T	L	F	D	U	Z
Z	O	C	D	H	W	M	Z	T	U	Z
A A	K	L	B	P T	C H	J E	O	T	X	E H
B B	H	S	P	O	P	N	M	D	L	M
C C	G E	X	K	W	D E	V	B	L	S T	E H
D D	G E	S	U T	G H	D E	P	O	T	H	X
E E	B	K E	D	Z	F	M	T	G	Q	J
F F	L	F	U T	Y	D E	T	Z	V	H	Q
G G	Z	G	W	N	K	X	J	T	R	N
H H	Y	T	X E	C	D E	P	M	V	L	W
I I	B	G	B	W	W	O	Q	R	G	N
J J	H	H	V	L	A	Q	Q	V	A	V
K K	J	Q	W	O	O	T	T	N	V	Q
L L	B	K	X	D	S	O	Z	R	S	N
M M	Y	U E	X E	O	P	P	Y	O	X T	Z
N N	H	O	Z	O	W	M	X	C	G	Q
O O	J	J	U T	G H	D E	W	Q	R	V	M
P P	U	K E	W	P T	E	F	X	E	N	F
Q Q	C	C	U T	G H	D E	W	P	E	U	H
R R	Y	B	W	E	W	V	M	D	Y	J
S S	R	Z	X E							

Fig. 29.

ADDITIONAL VALUES FROM ASSUMPTIONS (I)

Refer to line DD in Figure 29; ${}^{2}_{9}S_c$ assumed to be N_p.

Refer to line M in figure 29; A_c assumed to be W_p.

Then in lines C-D, $\overset{9}{A}\ \overset{10}{V}\ \overset{1}{K}\ \overset{2}{Z}\ \overset{3}{U}\ \overset{4}{G}\ \overset{5}{D}$ is assumed to be WITH THE.

	1	2	3	4	5	6	7	8	9	0
A	W	F	U	P	C	F	O	C	J	Y
			T	T	H					
B	G	B	Z	D	P	F	B	O	U	O
	E									
C	G	R	F	T	Z	M	Q	M	A	V
	E								W	I
D	K	Z	U	G	D	Y	F	T	R	W
	T	H	T	H	E					
E	G	J	X	N	L	W	Y	O	U	X
	E		E							
F	I	K	W	E	P	Q	Z	O	K	Z
		E								
G	P	R	X	D	W	L	Z	I	C	W
			E							
H	G	K	Q	H	O	L	O	D	V	M
	E	E								
I	G	O	X	S	N	Z	H	A	S	E
	E		E						T	H
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
M	S	R	Q	E	W	M	L	N	A	E
									W	H
N	G	S	X	E	R	O	Z	J	S	E
	E	N	E						T	H
O	G	V	Q	W	E	J	M	K	G	H
	E					E				
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
		H								
S	C	F	B	S	C	V	X	C	H	Q
					H					
T	Z	T	Z	S	D	M	X	W	C	M
					E					
U	R	K	U	H	E	Q	E	D	G	X
		E	T							
V	F	K	V	H	P	J	J	K	J	Y
		E				E				
W	Y	Q	D	P	C	J	X	L	L	L
				T	H	E				

	1	2	3	4	5	6	7	8	9	0
X	G	H	X	E	R	O	Q	P	S	E
	E		E						T	H
Y	G	K	B	W	T	L	F	D	U	Z
	E	E								
Z	O	C	D	H	W	M	Z	T	U	Z
A A	K	L	B	P	C	J	O	T	X	E
	T			T	H	E				
B B	H	S	P	O	P	N	M	D	L	M
		N								
C C	G	X	K	W	D	V	B	L	S	E
	E				E				T	H
D D	G	S	U	G	D	P	O	T	H	X
	E	N	T	H	E					
E E	B	K	D	Z	F	M	T	G	Q	J
		E								
F F	L	F	U	Y	D	T	Z	V	H	Q
			T		E					
G G	Z	G	W	N	K	X	J	T	R	N
H H	Y	T	X	C	D	P	M	V	L	W
			E		E					
I I	B	G	B	W	W	O	Q	R	G	N
J J	H	H	V	L	A	Q	Q	V	A	V
									W	I
K K	J	Q	W	O	O	T	T	N	V	Q
L L	B	K	X	D	S	O	Z	R	S	N
		E	E						T	
M M	Y	U	X	O	P	P	Y	O	X	Z
N N	H	O	Z	O	W	M	X	C	G	Q
O O	J	J	U	G	D	W	Q	R	V	M
			T	H	E					
P P	U	K	W	P	E	F	X	E	N	F
		E		T						
Q Q	C	C	U	G	D	W	P	E	U	H
			T	H	E					
R R	Y	B	W	E	W	V	M	D	Y	J
S S	R	Z	X							
		H	E							

Fig. 30.

ADDITIONAL VALUES FROM ASSUMPTIONS (II)

1 2 3 4 5 6 7 8 910
Refer to Figure 30, line A; W F U P C F O C J Y; assume to be BUT THOUGH.
- - T T H - - - - -

3456
Refer to Figure 30, lines N and X, where repetition XERO occurs; assume EACH
E---

	1	2	3	4	5	6	7	8	9	0
A	W	F	U	P	C	F	O	C	J	Y
	B	U	T	T	H	O	U	G	H	
B	G	B	Z	D	P	F	B	O	U	O
	E					O				
C	G	R	F	T	Z	M	Q	M	A	V
	E								W	I
D	K	Z	U	G	D	Y	F	T	R	W
	T	H	T	H	E					
E	G	J	X	N	L	W	Y	O	U	X
	E		E							
F	I	K	W	E	P	Q	Z	O	K	Z
		E		A						
G	P	R	X	D	W	L	Z	I	C	W
			E							
H	G	K	Q	H	O	L	O	D	V	M
	E	E					U			
I	G	O	X	S	N	Z	H	A	S	E
	E		E						T	H
J	B	B	J	I	P	Q	F	J	H	D
K	Q	C	B	Z	E	X	Q	T	X	Z
L	J	C	Q	R	Q	F	V	M	L	H
						O				
M	S	R	Q	E	W	M	L	N	A	E
				A					W	H
N	G	S	X	E	R	O	Z	J	S	E
	E	N	E	A	C	H			T	H
O	G	V	Q	W	E	J	M	K	G	H
	E					E				
P	R	C	V	O	P	N	B	L	C	W
Q	L	Q	Z	A	A	A	M	D	C	H
R	B	Z	Z	C	K	Q	O	I	K	F
		H					U			
S	C	F	B	S	C	V	X	C	H	Q
		U			H			G		
T	Z	T	Z	S	D	M	X	W	C	M
					E					
U	R	K	U	H	E	Q	E	D	G	X
		E	T							
V	F	K	V	H	P	J	J	K	J	Y
		E					E		H	
W	Y	Q	D	P	C	J	X	L	L	L
				T	H	E				

	1	2	3	4	5	6	7	8	9	0
X	G	H	X	E	R	O	Q	P	S	E
	E		E	A	C	H			T	H
Y	G	K	B	W	T	L	F	D	U	Z
	E	E								
Z	O	C	D	H	W	M	Z	T	U	Z
A A	K	L	B	P	C	J	O	T	X	E
	T			T	H	E	U			H
B B	H	S	P	U	P	N	M	D	L	M
		N								
C C	G	X	K	W	D	V	B	L	S	E
	E				E				T	H
D D	G	S	U	G	D	P	O	T	H	X
	E	N	T	H	E		U			
E E	B	K	D	Z	F	M	T	G	Q	J
		E								
F F	L	F	U	Y	D	T	Z	V	H	Q
		U	T		E					
G G	Z	G	W	N	K	X	J	T	R	N
H H	Y	T	X	C	D	P	M	V	L	W
			E		E					
I I	B	G	B	W	W	O	Q	R	G	N
						H				
J J	H	H	V	L	A	Q	Q	V	A	V
									W	I
K K	J	Q	W	U	O	T	T	N	V	Q
L L	B	K	X	D	S	O	Z	R	S	N
		E	E			H			T	
M M	Y	U	X	O	P	P	Y	O	X	Z
N N	H	O	Z	U	W	M	X	C	G	Q
								G		
O O	J	J	U	G	D	W	Q	R	V	M
			T	H	E					
P P	U	K	W	P	E	F	X	E	N	F
		E		T		O				
Q Q	C	C	U	G	D	W	P	E	U	H
			T	H	E					
R R	Y	B	W	E	W	V	M	D	Y	J
				A						
S S	R	Z	X							
		H	E							

Fig. 31.

ADDITIONAL VALUES FROM ASSUMPTIONS (III).

456
OPN - assume ING from repetition and frequency.

901
HQZ - assume ING from repetition and frequency.

```
     1 2 3 4 5 6 7 8 9 0
A    W F U P C F O C J Y
     B U T T H O U G H
B    G B Z D P F B O U O
     E       N C
C    G R F T Z M Q M A V
     E               W I
D    K Z U G D Y F T R W
     T H T H E
E    G J X N L W Y O U X
     E   E
F    I K W E P Q Z O K Z
       E   A N
G    P R X D W L Z I C W
         E
H    G K Q H O L O D V M
     E E         U
I    G O X S N Z H A S E
     E   E           T H
J    B B J I P Q F J H D
             N       I
K    Q C B Z E X Q T X Z
L    J C O R Q F V M L H
               C
M    S R Q E W M L N A E
           A         W H
N    G S X E R O Z J S E
     E N E A C H     T H
O    G V Q W E J M K G H
     E         E
P    R C V O P N B L C W
           I N G
Q    L Q Z A A A M D C H
R    B Z Z C K Q O I K F
       H         U
S    C F B S C V X C H Q
       U     H     G I N
T    Z T Z S D M X W C M
     G       E
U    R K U H E Q E D G X
       E T
V    F K V H P J J K J Y
       E     N E     H
W    Y Q D P C J X L L L
           T H E
```

```
     1 2 3 4 5 6 7 8 9 0
X    G H X E R O Q P S E
     E   E A C H     T H
Y    G K B W T L F D U Z
     E E
Z    O C D H W M Z T U Z
A A  K L B P C J O T X E
     T     T H E U     H
B B  H S P O P N M D L M
       N   I N G
C C  G X K W D V B L S E
     E       E       T H
D D  G S U G D P O T H X
     E N T H E   U   I
E E  B K D Z F M T G Q J
       E
F F  L F U Y D T Z V H Q
       U T   E       I N
G G  Z G W N K X J T R N
     G
H H  Y T X C D P M V L W
         E   E
I I  B G B W W O Q R G N
               H
J J  H H V L A Q Q V A V
                     W I
K K  J Q W O O T T N V Q
           I           N
L L  B K X D S O Z R S N
       E E     H     T
M M  Y U X O P P Y O X Z
           I N
N N  H O Z O W M X C G Q
           I       G   N
O O  J J U G D W Q R V M
         T H E
P P  U K W P E F X E N F
       E   T   O
Q Q  C C U G D W P E U H
         T H E
R R  Y B W E W V M D Y J
           A
S S  R Z X
       H E
```

Fig. 32.

c. From the initial and subsequent tentative identifications shown in Figures 29, 30, 31, and 32, the values obtained were arranged in the form of the secondary alphabets shown in Figure 33.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
0	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
1		W			G		Z													K						
2					K			Z						S							F					
3					X															U						
4	E							G	O											P						
5			R		D			C						P												
6					J		N	O							F											
7																					O					
8							C																			
9								J	H											S			A			
10								E	V																	

Fig. 33.

34. Fundamental theory. - a. In paragraph 31 methods of reconstructing primary components from secondary alphabets were given in detail. It is necessary that those methods be fully understood before the following steps be studied. It was there shown that the primary component can be one of a series of 26 equivalent primary sequences, all of which will give exactly similar results so far as the secondary alphabets and the cryptographic text are concerned. It is not necessary that the identical or original primary component employed in the cryptographing be reconstructed; any equivalent primary sequence will serve. The whole question is one of establishing a sequence of letters the interval between which is either identical with that in the original primary component or else is an exact constant multiple of the interval separating the letters in the original primary component. For example, suppose K P X N Q forma a sequence in the original primary component. Here the interval between K and P, P and X, X and N, N and Q is one; in an equivalent primary component, say the sequence I. . P. . X. . N. . Q, the interval between K and P is three, that between P and X also three, and so on; and the two sequences will yield the same secondary alphabets. So long as the interval between K and P, P and X, X and N, N and Q is a constant one, the sequence will yield the same secondary alphabets as do those of the original primary sequence. However, it is necessary that this interval be an odd number other than 13, as these are the only cases which will yield one unbroken sequence of 26 letters. Suppose a secondary alphabet to be as follows:

Plain -	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
Cipher -														X		K	N							P		

It can be said that the primary component contains the following sequences:

XN   KP   NQ   PX

These, when united by means of their common letters, yield K P X N Q.

Suppose also the following secondary alphabet is at hand:

Plain -	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
Cipher -														P			X							K		N

Here the sequences PN, XQ, KX, and NZ can be obtained, which when united yield the two sequences KXQ and PNZ.

By a comparison of the sequences K P X N Q, K X Q, and P N Z, one can establish the following:

K P X N Q

K . X . Q

P . N . Z

It follows that one can now add the letter Z to the sequence, making it K P X N Q Z.

b. The reconstruction of a primary alphabet from one of the secondaries by the process given in paragraph 31 requires a complete or nearly complete secondary alphabet. This is at hand only <u>after</u> a cryptogram has been completely solved. But if one could employ several very scant or skeletonized secondary alphabets simultaneously with the analysis of the cryptogram, one could then possibly build up a primary component from fewer data and thus solve the cryptogram much more rapidly than would otherwise be the case.

c. Suppose only the cipher components of the two secondary alphabets given above be placed into juxtaposition. Thus:

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
.	.	.	.	.	.	.	.	.	.	.	.	.	X	.	K	N	.	.	.	.	.	.	P	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	.	X	.	.	.	.	.	.	K	.	N

The sequences PX, XN, and KP result, which, united, yield KPXN as part of the primary sequence. It follows, therefore, that one can employ the <u>cipher components of secondary alphabets</u> as sources of independent data to assist in building up the primary sequences. The usefulness of this point will become clearer subsequently.

35. Application of principles. - a. Refer now to Figure 33. Hereafter, in order to avoid all ambiguity and for ease in reference, the position of a letter in Figure 33 will be indicated by coordinates in parentheses. Thus, N (6 - 7) refers to the letter N in line 6 and in column 7 of Figure 33.

b. (1) Now, consider the following pairs of letters:

E (0 - 5) J (6 - 5)

G (0 - 7) N (6 - 7)

(H (0 - 8) C (6 - 8))
( ) HO, CF = HOF
(O (0 -15) F (6 - 15))

(One is able to use the line marked zero in Figure 33 since this is a mixed sequence sliding against itself.)

(2) The immediate results of this set of values will now be given. Having HOF as a sequence, with EJ as belonging to the same interval set, suppose HOF and EJ are placed into juxaposition as portions of sliding alphabets. Thus:

Plain - . . . H O F . . .
Cipher - . . . E J . . . .

When $H_p = E_c$, then $O_p = J_c$.

(3) Refer now to alphabet 10, Figure 33, where it is seen that $H_p = E_c$. The derived value, $O_p = J_c$, can immediately be inserted in the same alphabet and substituted in the cryptogram.

c. (1) Again, CN belongs to the same set of interval values as do EJ and HOF. Hence, by superimposition:

Plain - . . . H O F . . .
Cipher - . . . G N . . . .

(2) When $H_p = G_c$, then $O_p = N_c$. Therefore, the value $O_p = N_c$ can be inserted and also substituted in the cryptogram.

(3) Furthermore, note the corroboration we find from this particular superimposition.

H (0 - 8) G (0 - 7)
O (6 - 8) N (6 - 7)

This checks up the value in alphabet 6, $G_p = N_c$.

d. (1) Again superimpose HOF and GN:

H O F
 G N

(2) Note this corroboration:

O	(6 - 8)	G	(4 - 8)
F	(6 -15)	N	(4 -15)

which has just been inserted in Figure 7, as stated above.

e. (1) Again using HOF and EJ, but in a different superimposition, we have:

. . . H O F . .
. . E J . . . .

(2) Refer now to H (9 - 9) J (9 - 8). Directly under these letters is found V (10 - 9) E (10 - 8). Therefore, the V can be added immediately before H O F, making the sequence V H O F.

f. (1) Now take V H O F and juxapose it with E J, thus:

V H O F
E J

(2) Refer now to Figure 33, and find the following:

V	(10 - 9)	E	(10 - 8)
H	(9 - 9)	J	(9 - 8)
O	(4 - 9)	G	(4 - 8)
I	(0 - 9)	H	(0 - 8)

(3) From the value O G it follows that G can be set next to J in E J. Thus:

V H O F
E J G

(4) But G N is already a member of the same interval as E J. Therefore, it is now possible to combine E J, J G, and G N into one sequence, E J G N, yielding:

V H O F
E J G N

g. (1) Refer now to Figure 33.

V (0 - 22)	E (0 - 5)
? (1 - 22)	G (1 - 5)
? (2 - 22)	K (2 - 5)
? (3 - 22)	X (3 - 5)
? (5 - 22)	D (5 - 5)
? (6 - 22)	J (6 - 5)

(2) The only values which can be inserted are:

O (1 - 22)	G (1 - 5)
H (6 - 66)	J (6 - 5)

(3) This means that $V_p = O_c$ in alphabet 1 and that $V_p = H_c$ in alphabet 6. There is one O_c in the frequency distribution for alphabet 1, and no H_c in that for alphabet 6. The frequency distribution is, therefore, corroborative insofar as these values are concerned:

h. (1) Further, taking E J G N and V H O F, superimpose them thus:

```
  E J G N

V H O F
```

(2) Refer now to Figure 33.

E (0 - 5)	H (0 - 8)
G (1 - 5)	? (1 - 8)

(3) From the diagram of superimposition the value G (1 - 5) F (1 - 8) can be inserted, which gives $H_p = F_c$ in alphabet 1.

i. (1) Again, V H O F and E J G N are juxtaposed:

```
  V H O F

E J G N
```

(2) Refer to Figure 33 and find the following:

H (0 - 8)	G (4 - 8)
A (0 - 1)	E (4 - 1)

This means that it is possible to add A, thus:

```
A V H O F
E J G N
```

(3) In the set there are also:

E (0 - 5)   G (1 - 5)
G (0 - 7)   Z (1 - 7)

Then in the superimposition

```
  E J G N

E J G N
```

It is possible to add Z under G, making the sequence E J G N Z.

(4) Then taking

```
  A V H O F
E J G N Z
```

and referring to Figure 33:

H (0 - 8)   N (0 - 14)
O (6 - 8)   ? (6 - 14)

It will be seen that O = Z from superimposition, and hence in alphabet 6 $N_p = Z_c$, an important new value, but occurring only once in the cryptogram. Has an error been made? The work so far seems too corroborative in interlocking details to think so.

j. (1) The possibilities of the superimposition and sliding of the AVHOF and the EJGNZ sequences have by no means been exhausted as yet, but a little different trail this time may be advisable.

E (0 - 5)   T (0 - 20)
G (1 - 5)   K (1 - 20)
K (3 - 5)   U (3 - 20)

(2) Then:

```
E J G N X
T . K
```

(3) Now refer to the following:

E (0 - 5)   K (2 - 5)
N (0 -14)   S (2 -14)

whereupon the value S can be inserted:

E J G N Z
T . K . . S

k. (1) Consider all the values based upon the interval corresponding to JG:

J (6 - 5)	G (1 - 5) →	J (9 - 8)	G (4 - 8)		
N (6 - 7)	Z (1 - 7)	H (9 - 9)	O (4 - 9)		
		S (9 -20)	P (4 -20) →	S (2 - 14)	P (5 - 14)
				Z (2 - 8)	C (5 - 8)
				K (2 - 5)	D (5 - 5)

(2) Since J and G are sequent in the EJGNZ sequence, it can be said that all the letters of the foregoing pairs are also sequent. Hence Z C, S P, and K D are available as new data. These give E J G N Z C and T . K D . S P.

T (0 - 20)	P (4 - 20)
A (0 - 1)	E (4 - 1)
H (0 - 8)	G (4 - 8)
I (0 - 9)	O (4 - 9)

(3) Now in the T . K D . S P sequence the interval between T and P is $\overset{\;1\;2\;3\;4\;5\;6}{T\;.\;.\;.\;.\;.\;P}$. Hence the interval between A and E is 6 also. It follows therefore that the sequences A V H O F and E J G N Z C should be united thus:

1 2 3 4 5 6
A V H O F . E J G N Z C

(4) Corroboration is found in the interval between H and G, which is six. The letter I can be placed into position, from the relation I (0 - 9) O (4 - 9), thus:

1 2 3 4 5 6
I . . A V H O F . E J G N Z C

l. (1) From Figure 33:

H (0 - 8)	Z (2 - 8)
E (0 - 5)	K (2 - 5)
N (0 -14)	S (2 -14)
U (0 -21)	F (2 -21)

(2) From the I . . A V H O F . E J G N Z C sequence one can write:

1 2 3 4 5 6 7 8
H . . . . Z
E . . . . K
N . . . . S
U . . . . F

(3) Hence one can make the sequence

										1	2	3	4	5	6	7	8
I	.	.	A	V	H	O	F	.	$\underline{E}$	J	G	N	Z	C	.	.	$\underline{K}$

Then I . . A V H O F . E J G N Z C T . K D . S P

and	$\underline{U}$	I	.	.	A	V	H	O	$\underline{F}$	.	E	J	G	$\underline{N}$	Z	C	T	.	K	D	.	$\underline{S}$	P
		1	2	3	4	5	6	7	8						1	2	3	4	5	6	7	8	

m. (1) Subsequent derivations can be indicated very briefly as follows:

E (0 - 5) C (0 - 3)
D (5 - 5) R (5 - 3)

	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
From	U	I	.	.	A	V	H	O	F	.	E	J	G	N	Z	C	T	.	K	D	.	S	P	.	.	.
one can write											E	.	.	.	.	C										
												1	2	3	4	5										
and																				D	.	.	.	.	R	
																					1	2	3	4	5	

making the sequence

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
U	I	.	.	A	V	H	O	F	.	E	J	G	N	Z	C	T	.	K	D	.	S	P	.	R	.

(2) Another derivation:

U (3 - 20) T (0 -20)
X (3 - 5) E (0 - 5)

	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)	
From	U	I	.	.	A	V	H	O	F	.	E	J	G	N	Z	C	T	.	K	D	.	S	P	.	R	.)	$\underline{U}$
one can write	U	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T										
and	.	.	.	.	.	.	.	.	.	.	E	.	.	.	.	.	.	.	.	.	X	.	.	.	.	.	

making the sequence

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
U	I	.	.	A	V	H	O	F	.	E	J	G	N	Z	C	T	.	K	D	X	S	P	.	R	.

(3) Another derivation:

E (0 - 5) G (1 - 5)
B (0 - 2) W (1 - 2)

From E J G
one can write E . G
and then B . W

There is only one place where B . W can fit, viz, at the end:

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
U	I	.	.	A	V	H	O	F	.	E	J	G	N	Z	C	T	.	K	D	X	S	P	$\underline{B}$	R	$\underline{W}$

n. Only four letters remain to be placed into the sequence: L, M, Q, and Y. They were easily found by application of the primary component to the message. Having the primary component almost fully constructed, decipherment of the cryptogram can be completed with speed and precision. The text is as follows:

```
WFUPCFOCJY  RCVOPNBLCW  BKDZFMTGQJ
BUTTHOUGHW  POSINGTHES  SELFWILLGO

GBZDPFBOUO  LQZAAAMDCH  LFUYDTZVHQ
ECANNOTASY  OLARSYSTEM  OUTBECOMIN

GRFTZMQMAV  BZZCKQOIKF  ZGWNKXJTRN
ETREVIEWWI  SHALLTURNA  GACOLDANDL

KZUGDYFTRW  CFBSCVXCHQ  YTXCDPMVLW
THTHEMINDS  NUNCHANGIN  IFELESSMAS

GJXNLWYOUX  ZTZSDMXWCM  BGBWWOQRGN
EYEOURPAST  GFACEINPER  SANDTHESOL

ITWEPQZOKZ  RKUHEQEDGX  HHVLAQQVAV
WECANTOANE  PETUITYTOT  ARSYSTEMWI

PRXCWLZICW  FKVHPJJKJY  JQWOOTTNVQ
XTENTFORES  HESUNEACHW  LLCIRCLEUN

GKQHOLODVM  YQDPCJXLLL  BKXDSOZRSN
EEOURFUTUR  ILLTHENHAV  SEENGHOSTL

GOXSNZHASE  GHXEROQPSE  YUXOPPYOXZ
EWECANWITH  EREACHEDTH  IKEINSPACE

BBJIPQFJHD  GKBWTLFDUZ  HOZOWMXCGQ
SCIENTIFIC  EENDOFITSE  AWAITINGON

QCBZEXQTXZ  OCDHWMZTUZ  JJUGJWQRVM
CONFIDENCE  VOLUTIONSE  LYTHERESUR

JCQRQFVMLH  KLBPCJOTXE  UKWPEFXENF
LOOKFORWAR  TINTHEUNCH  RECTIONOFA

SRQEWMLNAE  HSPOPNMDLM  CCUGDWPEUH
DTOATIMEWH  ANGINGSTAR  NOTHERCOSM

GSXEROZJSE  GCKWDVBLSE  YBWEWVMDYJ
ENEACHOFTH  EOFDEATHTH  ICCATASTRO

GVQWEJMKGH  GSUGDPOTHX  RZX
EBODIESCOM  ENTHESUNIT  PHE
```

Fig. 34.

36. General remarks. - a. It is to be stated that the sequence of steps described in the preceding paragraphs corresponds quite closely with that actually followed in solving the problem. It is also to be pointed out that this method can be used as a control in the early stages of analysis because it will allow the cryptanalyst to check assumptions for values. For example, the very first value derived in applying the principles of indirect symmetry to the problem herein described was $H_c = A_p$ in alphabet 1. As a matter of fact the writer had been inclined toward this value, from a study of the frequency and combinations which H_c showed; when the indirect-symmetry method actually substantiated his tentative hypothesis he immediately proceeded to substitute the value given. If he had assigned a different value to H_c, or if he had assumed a letter other than H_c for A_p in that alphabet, the conclusion would immediately follow that either the assumed value for H_c was erroneous, or that one of the values which let to the derivation of $H_c = A_p$ by indirect symmetry was wrong. Thus, these principles aid not only in the systematic and nearly automatic derivation of new values (with only occasional, or incidental references to the actual frequencies of letters), but they also assist very materially in serving as corroborative checks upon the validity of the assumptions already made.

b. Furthermore, while the writer has set forth, in Figure 33, a set of 30 values apparently obtained before he began to reconstruct the primary component, this was done for purposes of clarity and brevity in exposition of the principles herein described. As a matter of fact, what he did was to watch very carefully, when inserting values in Figure 33, to find the very first chance to employ the principles of indirect symmetry; and just as soon as a value could be derived, he substituted the value in the cryptographic text. This is good procedure for two reasons. Not only will it disclose impossible combinations but also it gives opportunity for making further assumptions for values by the addition of the derived values to those previously assumed. Thus, the processes of reconstructing the primary component and finding additional data for the reconstruction proceed simultaneously in an ever-widening circle.

c. It is worth noting that the careful analysis of only a sum total of 30 values in Figure 33 results in the derivation of the entire table of secondary alphabets, 676 values in all. And while the elucidation of the method seems long and tedious, in its actual application the results are speedy, accurate, and gratifying in their corroborative effect upon the mental activity of the cryptanalyst.

d. (1) The problem here used as an illustrative case is by no means one that most favorably presents the application and the value of the method, for it has been applied in other cases with much speedier success. For example, suppose that in a cryptogram of 6 alphabets the equivalents of only THE in all 6 alphabets are fairly certain. As in the previous case, it is supposed that the secondary alphabets are obtained by sliding a mixed alphabet against itself. Suppose the secondary alphabets to be as follows:

Ø	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
1					B															E						
2					C			L												X						
3					I			V												C						
4					N			P												B						
5					X			O												P						
6					T			Z												V						

Fig. 35.

(2) Consider the following chain of derivatives arranged diagrammatically:

```
H  (Ø - 8)    O  (5 - 8)
T  (Ø -20)    P  (5 -20)
E  (Ø - 5)    X  (5 - 5) ---> E  (1 - 20)  X (2 -20)
                              Q  (1 -  8)  L (2 - 8)
                              B  (1 -  5)  C (2 - 5) -> B (4 -20)  C (3 -20)
                                                        N (4 - 5)  I (3 - 5)
                                                        P (4 - 8)  V (3 - 8) ->

->P (5 -20)    V  (6 -20)
  O (5 - 8)    Z  (6 - 8)
  X (5 - 5)    T  (6 - 5) ---> X  (2 -20)  T (Ø -20)
                               L  (2 - 8)  H (Ø - 8)
                               C  (2 - 5)  E (Ø - 5) -> C (3 -20)  E (1 -20)
                                                        V (3 - 8)  Q (1 - 8)
                                                        I (3 - 5)  B (1 - 5)
```

Fig. 36.

(3) These pairs are manifestly all of the same interval, and therefore unions can be made immediately. The complete list is as follows:

EX QL NI LH

HO BC OZ CE

TP PV XT VQ

IB

(4) Joining pairs by their common letters, the following sequence is obtained:

N I B C E X T O V Q L H O Z

e. With this as a nucleus the cryptogram can be solved speedily and accurately. When it is realized that the cryptanalyst can assume THE's rather readily in some cases, the value of this principle becomes apparent. When it is further realized that if a cryptogram has sufficient text to enable the THE's to be found easily, it is usually also not at all difficult to make correct assumptions for values for two or three other high-frequency letters, it is clear that the principles of indirect symmetry of position may often be used with gratifyingly quick success to reconstruct the complete primary component.

SECTION IX.

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, III.

37. Solution of subsequent messages enciphered by the same primary components. - a. In the discussion of the methods of solving repeating-key ciphers using secondary alphabets derived from the sliding of a mixed component against the normal component, (Section V), it was shown how subsequent messages enciphered by the same pair of primary components but with different keys could be solved by application of principles involving the completion of the plain-component sequence (paragraphs 23, 24). The present paragraph deals with the application of these same principles to the case where the primary components are identical mixed sequences.

b. Suppose that the following primary component has been reconstructed from the analysis of a lengthy cryptogram:

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

A new message exchanged between the same correspondents is intercepted and is suspected of having been enciphered by the same primary components but with a different key. The message is as follows:

NFWWP NOMKI WPIDS CAAET QVZSE
YOJSC AAAFG RVNHD WDSCA EGNFP
FOEMT HXLJW PNOMK IQDBJ IVNHL
TFNCS BGCRP

c. Factoring discloses that the period is 7 letters. The text is transcribed accordingly, and is as follows:

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

Fig. 37.

d. The letters belonging to the same alphabet are then employed as the initial letters of completion sequences, in the manner shown in paragraph 23e, using the already reconstructed primary component. The completion diagrams for the first five letters of the first three alphabets are as follows:

Alphabet 1.

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

Alphabet 2.

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

Alphabet 3.

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

Fig. 38.

e. Examining the successive generatives to select the ones showing the best assortment of high-frequency letters, those marked in Figure 38 by asterisks are chosen. These are then assembled in columnar fashion and yield the following plain text:

1	2	3	4	5	6	7
H	A	V				
E	C	T				
C	O	N				
I	M	E				
C	O	N				

Fig. 39.

f. The corresponding key-letters are sought and are found to be JOU, which suggests the keyword JOURNEY. Testing the key-letters RNEY for alphabets 4, 5, 6, and 7, the following results are obtained:

```
1 2 3 4 5 6 7
J O U R N E Y

N F W W P N O
H A V E D I R

S C A A E T Q
E C T E D S E
```

Fig. 40.

The message may now be completed with ease. It is as follows:

```
J O U R N E Y          J O U R N E Y

H A V E D I R          S A I N C E I
N F W W P N O          P F O E M T H

E C T E D S E          N T H E D I R
M K I W P I D          X L J W P N O

C O N D R E G          E C T I O N O
S C A A E T Q          M K I Q D B J

I M E N T T O          F H O R S E S
V Z S E Y O J          I V N H L T F

C O N D U C T          H O E F A L L
S C A A A F G          N C S B G C R

T R O R O R E          S
R V N H D W D          P

C O N N A I S
S C A E G N F
```

Fig. 41.

38. Solution of repeating-key ciphers in which the identical mixed components proceed in opposite directions. - The secondary alphabets in this case (paragraph 3, Case B (3) (a) (II) are reciprocal. The steps in solution are essentially the same as in the preceding case (paragraph 28).

the principles of indirect symmetry of position can also be applied with the necessary modifications introduced by virtue of the reciprocity existing within the respective secondary alphabets (paragraph 31 p).

39. Solution of repeating-key ciphers in which the primary components are different mixed sequences. - This is Case B (3) (b) of paragraph 3. The steps in solution are essentially the same as in paragraphs 28 and 31, except that in applying the principles of indirect symmetry of position it is necessary to take cognizance of the fact that the primary components are different mixed sequences (paragraph 31 q).

40. Solution of subsequent messages after the primary components have been recovered. - a. In the case in which the primary components are identical mixed sequences proceeding in opposite directions, as well as in that in which the primary components are different mixed sequences, the solution of subsequent messages[1] is a relatively easy matter. In both cases, however, the student must remember that before the method illustrated in paragraph 37 can be applied it is necessary to convert the cipher letters into their plain-component equivalents before completing the plain-component sequence. From there on, the process of selecting and assembling the proper generatrices is the same as usual.

b. Perhaps an example may be advisable. Suppose the enemy has been found to be using primary components based upon the keyword QUESTIONABLY, the plain component running from left to right, the cipher component in the reverse direction. The following new message has arrived from the intercept station:

MVXOX BZIYZ NLWZH OXIEO OOEPZ

FXSRX EJBSH BONAU RAPZI NRAMV-

XOXAI JYXWF KNDOW JERCU RALVB

ZAQUW JWXYI DGRKD QBDRM QECYV

QW

---

[1] That is, messages intercepted after the primary components have been reconstructed, and enciphered by keys different from those used in the messages upon which the reconstruction of the primary components was accomplished.

c. Factoring discloses that the period is 6 and the message is accordingly transcribed into 6 columns, Fig. 42. The letters of these columns are then converted into their plain component equivalents by juxtaposing the two primary components at any point of coincidence, for example $Q_p = Z_c$. The converted letters are shown in Fig.43. The letters of the individual columns are then used as the initial letters of completion sequences, using the QUESTIONABLY primary sequence. The final step is the selection and assembling of the selected generatrices. The results for the first ten letters of the first three columns are shown below:

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

Fig.42

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

Fig.43.

Column 1.

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

Column 2.

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

Column 3.

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

Fig. 44.

Columnar assembling of selected generatrices gives what is shown in Fig. 45.

1	2	3	4	5	6
F	I	R	.	.	.
A	V	A	.	.	.
L	E	S	.	.	.
I	R	D	.	.	.
A	D	R	.	.	.
I	L	L	.	.	.
U	P	Y	.	.	.
D	E	F	.	.	.
F	I	R	.	.	.
E	L	A	.	.	.

Fig. 45.

d. The key letters are sought, and found to be NUM, which suggests NUMBER. The entire message may now be read with ease. It is as follows:

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

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

Fig. 46.

e. If the primary components are different mixed sequences, the procedure is identical with that just indicated. The important point to note is that one must not fail to convert the letters into their plain-component equivalents before the completion-sequence method is applied.

SECTION X.

REPEATING-KEY SYSTEMS WITH MIXED CIPHER ALPHABETS, IV.

41. General remarks. - The preceding three sections have been devoted to an elucidation of the general principles and procedure in the solution of typical cases of repeating-key ciphers. This section will be devoted to a consideration of the variations in cryptanalytic procedure arising from special circumstances. It may be well to add that by the designation special circumstances it is not meant to imply that the latter are necessarily <u>unusual</u> circumstances. <u>The student should always be on the alert to seize upon any opportunities that may appear in which he may apply the methods to be described.</u> In practical work such opportunities are by no means rare and are seldom overlooked by competent cryptanalysts.

42. Deriving the secondary alphabets, the primary components, and the key, given a cryptogram with its plain text. - a. It may happen that a cryptogram and its equivalent plain text may be at hand, as the result of capture, pilferage, compromise, etc. This as a general rule affords a very easy attack upon the whole system.

b. Taking first the case where the plain component is the normal alphabet, the cipher component a mixed sequence, the first thing to do is to write out the cipher text with its letter-for-letter decipherment. From this, by a slight modification of the principles of "factoring", one discovers the length of the key. It is obvious that when a word of three or four letters is enciphered by the same cipher text, the interval between the two occurrences is almost certainly a multiple of the length of the key. By noting a few recurrences of plain text and cipher letters, one can quickly determine the length of the key (assuming of course that the message is long enough to afford sufficient data). Having determined the length of the key, the message is rewritten according to its periods, with the plain text likewise in periods under the cipher letters. From this arrangement one can now reconstruct complete or partial secondary alphabets. If the secondary alphabets are complete, they will show direct symmetry of position; if they are but fragmentary in several alphabets, then the primary component can be reconstructed by the application of the principles of direct symmetry of position.

c. If the plain component is a mixed sequence, the cipher component the normal (direct or reversed sequence), the secondary alphabets will show no direct symmetry unless they are converted into their reciprocals (deciphering alphabets). The student should be on the lookout for such cases.

d. (1) If the plain and cipher primary components are identical mixed sequences proceeding in the same direction, the secondary alphabets will show indirect symmetry of position, and they can be used for the speedy reconstruction of the primary components (Paragraph 31 a. to o.).

(2) If the plain and the cipher primary components are identical mixed sequences proceeding in opposite directions, the secondary alphabets will be completely reciprocal secondary alphabets and the primary component may be reconstructed by applying the principles outlined in paragraph 31 p.

(3) If the plain and the cipher **primary** components are different mixed sequences, the secondary alphabets will show indirect symmetry of position and the primary components may be reconstructed by applying the principles outlined in paragraph 31 q.

e. In all the foregoing cases, after the primary components have been reconstructed, the keys can be readily recovered.

43. Deriving the secondary alphabets, the primary components, and the keywords for messages, given two or more cryptograms in different keys and suspected to contain identical plain text. - a. The simplest case of this kind is that involving two monoalphabetic substitution ciphers with mixed alphabets derived from the same pair of sliding components. An understanding of this case is necessary to that of the case involving repeating-key ciphers.

b. (1) A message is transmitted from station A to station B. B. sends A some operating signals which indicate that B cannot decipher the message, and soon thereafter A sends a second message, identical in length with the first. This leads to the suspicion that the plain text of both messages is the same. The intercepted messages are superimposed. Thus:

```
1.  NXGRV  MPUOF  ZQVCP  VWERX  QDZVX  WXZQE
2.  EMLHJ  FGVUB  PRJNG  JKWHM  RAPJM  KMPRW

1.  TBDSP  VNXJK  RFZWH  ZUWLU  IYVZQ  FXOAR
2.  ZTAXG  JJMCD  HBPKY  PVKIV  QOJPR  BMUSH
```

(2) In initiating a chain of cipher-text equivalents from message 1 to message 2, the following complete sequence 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
N	E	W	K	D	A	S	X	M	F	B	T	Z	P	G	L	I	Q	R	H	Y	O	U	V	J	C

(3) Experimentation along already-indicated lines soon discloses the fact that the foregoing component is an equivalent primary component of the original primary based upon the keyword QUESTIONABLY, decimated on the 21st interval. Let the student decipher the cryptogram.

(4) The foregoing example is somewhat artificial in that the plain text was consciously selected with a view to making it contain every letter of the alphabet. The purpose in doing this was to permit the construction of a complete chain of equivalents from only two short messages, in order to give a simple illustration of the principles involved. If not every letter of the alphabet is present in the plain-text message, then only partial chains of equivalents can be constructed. These may be united, if circumstances will permit, by recourse to the various principles elucidated in paragraph 31.

(5) The student should carefully study the foregoing example in order to obtain a thorough comprehension of the <u>reason</u> why it was possible to reconstruct the primary component from the two cipher messages without having any plain-text to begin with at all. Since the plain text of both messages is the same, the relative displacement of the primary components in the case of message 1 differs from the relative displacement of the same primary components in the case of message 2 by a <u>fixed</u> interval. Therefore, the distance on the primary component, between N and E (the first letters of the two messages), regardless of what plain-text letter these two cipher letters represent, is the same as the distance between E and W (the 18th letters), W and K (the 17th letters), and so on. Thus this fixed interval permits of establishing a complete chain of letters separated by constant intervals and this chain becomes an equivalent primary component.

44. The case of repeating-key systems. - a. With the foregoing basic principles in mind the student is ready to note the procedure in the case of two repeating-key ciphers having identical plain texts. First, the case in which both messages have keywords of identical length but different compositions will be studied.

b. Given the following two cryptograms suspected to contain the same plain text:

Message 1.

YHYEX UBUKA PVLLT ABUVV DYSAB
PCQTU NGKFA ZEFIZ BDJEZ ALVID
TROQS UHAFK

Message 2.

CGSLZ QUBMN CTYBV HLQFT FLRHL
MTAIQ ZWMDQ NSDWN LCBLQ NETOC
VSNZR BJNOQ

The first step is to try to determine the length of the period. The usual method of factoring cannot be employed because there are no long repetitions and not enough repetitions even of digraphs to give any convincing indications. However, a subterfuge will be employed, based upon the theory of factoring.

c. Let the two messages be superimposed.

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

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1.	D	Y	S	A	B	P	C	Q	T	U	N	G	K	F	A	Z	E	F	I	Z
2.	D	L	R	H	L	M	T	A	I	Q	Z	W	M	D	Q	N	S	D	W	N

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1.	B	D	J	E	Z	A	L	V	I	D	T	R	O	Q	S	U	H	A	F	K
2.	L	C	B	L	Q	N	E	T	O	C	V	S	N	Z	R	B	J	N	O	Q

Now let a search be made of cases of identical superimposition. For example, $\begin{matrix}4\\E\\L\end{matrix}$ and $\begin{matrix}44\\E\\L\end{matrix}$ are separated by 40 letters, $\begin{matrix}6\\U\\Q\end{matrix}$, $\begin{matrix}18\\U\\Q\end{matrix}$, and $\begin{matrix}30\\U\\Q\end{matrix}$ are separated by 12 letters. Let these intervals between identical superimpositions be factored, just as though they were ordinary repetitions. That factor which is the most frequent should correspond with the length of the period for the following reason. If the period is the same and the plain text is the same in both messages, then the condition of identity of superimposition can only be the result of identity of encipherments by identical cipher alphabets. This is only another way of saying that the same relative position in the keying cycle has been reached in both cases of identity. Therefore, the distance between identical superimpositions must be either equal to or else a multiple of the length of the period. Hence, factoring the intervals must yield the length of the period. The complete list of intervals and factors applicable to cases of identical superimposed pairs is as follows (factors above 12 are omitted):

1st EL to 2d EL - 40 = 2, 4, 5, 8, 10

1st UQ to 2d UQ - 12 = 2, 3, 4, 6, 12

2d UQ to 3d UQ - 12 = 2, 3, 4, 6, 12

1st UB to 2d UB - 48 = 2, 3, 4, 6, 8, 12

1st KM to 2d KM - 24 = 2, 3, 4, 6, 8, 12

1st AN to 2d AN - 36 = 2, 3, 4, 6, 9, 12

2d AN to 3d AN - 12 = 2, 3, 4, 6, 12

1st VT to 2d VT - 8 = 2, 4, 8

2d VT to 3d VT - 28 = 2, 4, 7

1st TV to 2d TV - 36 = 2, 3, 4, 6, 9, 12

1st AH to 2d AH - 8 = 2, 4, 8

1st BL to 2d BL - 8 = 2, 4, 8

2d BL to 3d BL - 16 = 2, 4, 8

1st SR to 2d SR - 32 = 2, 4, 8

1st FD to 2d FD - 4 = 2, 4

1st ZN to 2d ZN - 4 = 2, 4

1st DC to 2d DC - 8 = 2, 4, 8

The factor 4 is the only one common to every one of these intervals and it may be taken as beyond question that the length of the period is 4.

d. Let the messages now be superimposed according to their periods:

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

e. Now distribute the superimposed letters into "secondary alphabets". Thus:

0.	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
1.		L		F	S			J	O		M	Y			N					I				Z	C	Q
2.	N			C		D		G				B				M	Z				Q				L	
3.	Q	U	T			O			W	B		E		Z		C		R		V		F			S	
4.	H				L		W				Q						A	S			B	T				N

by the usual methods, construct the primary or an equivalent primary component. Taking lines 0 and 1, the following sequences are noted:

BL, DF, ES, HJ, IO, KM, LY, ON, TI, XZ, YC, ZQ,

which, when united by means of common letters and study of other sequences, yield the complete original primary component based upon the keyword QUESTIONABLY:

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

The fact that the pair of lines with which the process was commenced yield the original primary sequence is purely accidental; it might have just as well yielded an equivalent primary sequence.

f. Having the primary component, the solution of the messages is now a relatively simple matter. An application of the method elucidated in paragraph 37 is made, involving the completion of the plain-component sequence for each alphabet and selecting those generatrices which contain the best assortments of high-frequency letters. Thus, using Message 1:

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

Fig. 48.

The selected generatrices (those marked by asterisks in Fig. 48) are assembled in columnar manner:

A	L	L	A
R	R	A	N
G	E	M	E
N	T	S	F
O	R	R	E

Fig. 49.

The key letters are sought and give the keyword SOUP. The plain text for the second message is now known, and by reference to the cipher text and the primary components, the keyword for this message is found to be TIME. The complete textx are as follows:

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

Fig. 50.

45. The case of identical messages enciphered by keywords of different lengths. - a. In the foregoing case the keywords for the two messages, although different, were identical in length. When this is not true and the keywords are of different lengths, the procedure need be only slightly modified.

b. Given the following two cryptograms suspected of containing the same plain-text enciphered by the same primary components but with different keywords of different lengths.

Message 1.

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

Message 2.

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

c. The messages are long enough to show a few short repetitions which permit factoring. The latter discloses that Message 1 has a period of 4, Message 2 a period of 6 letters. The messages are superimposed, with numbers marking the position of each letter in the corresponding period, as shown below:

```
      1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2
No.1-I Y L F F P H X G C E Z T Z L A M B K I B Y L Z E L F E I L
No.2-A M T U K M F G F H U N N N T R W A H V A G B N S K A G B B
      1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

      3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
No.1-B H N Z F U W N X S Z O R V K B G S L J P S L P F I H K F H
No.2-N N O S D B Q G K H S I M D J D F Y D Z F H F M C V G V D X
      1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

      1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2
No.1-Y Y X U T Z F H W L Y X A D K O D L G L I Z S W S I L X N Z
No.2-F M K F A X C N V F L O Y R C M J B D U T S E I O D T Y Y X
      1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6

      3 4 1 2 3 4 1 2 3 4 1 2
No.1-L W L K F H G O U W L A
No.2-A F B V D X K F R L F N
      1 2 3 4 5 6 1 2 3 4 5 6
```

d. A table of "secondary alphabets" is now constructed by distributing the letters in respective lines corresponding to the 12 different superimposed pairs of numbers. For example, all pairs corresponding to the superimposition of position 1 of Message 1 with position 1 of Message 2 all distributed in lines 0 and 1 of the Table. Thus, the very first superimposed pair is $\begin{matrix}1\\I\\A\\1\end{matrix}$; the letter A is inserted in line 1 under the letter I. The next $^{1}_{1}$pair is the 13th superimposition, with $^{T}_{N}$; the letter N is inserted in line 1 under the letter T, and so on. The completed diagram is as follows:

Superimposed pairs.

0	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
1-1	Y				S		K		A			D							O	N			Q		F	
2-2				R					D	Z		K		G	F										M	N
3-3						A					C	T				F					R			K		
4-4	R				G	U									M				H		F		L	Y		
1-5				J		K			B			F	W	Y						A						S
2-6	N	A										B			I	M										X
3-1		N				C	D	F			H	A						M								
4-2								N	V			U										D	F	G		
1-3		A					F	G	T		J	B		O									V			
2-4		D	H								V	F													G	S
3-5					U	D	F					B							E						L	
4-6								X											Y		B		I	O		N

Fig. 51.

e. There are more than sufficient data here to permit of a complete reconstruction of the primary component, which is found to be that based upon the keyword QUESTIONABLY.

f. The plain text and the keywords for both messages may now be found very easily. They are shown below:

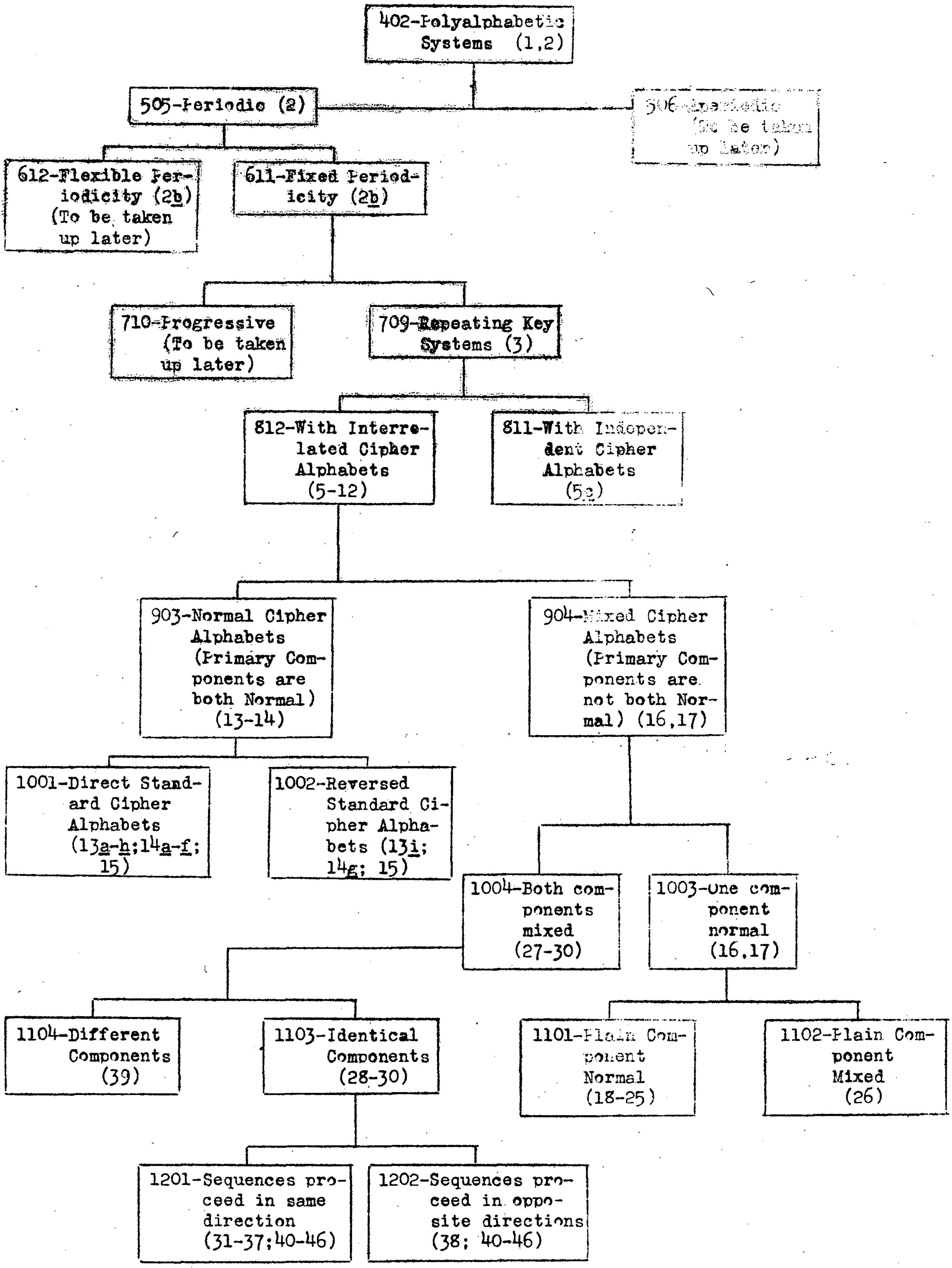
STAR	STAR	STAR	OCEANS	OCEANS
I Y L F	W N X S	A D K O	A M T U K M	C V G V D X
E N E M	P S H A	I B L Y	E N E M Y H	O L D F O R
F P H X	Z O R V	D L G L	F G F H U N	F M K F A X
Y H A S	V E D U	L O N G	A S C A P T	A N H O U R
G C E Z	K B G S	I Z S W	N N T R W A	C N V F L O
C A P T	G I N A	E R R E	U R E D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

Fig. 52.

46. Concluding remarks. - The observant student will have noted that a large part of this text is devoted to the elucidation and application of a very few basic principles. These principles are, however, extremely important and their proper usage in the hands of a skilled cryptanalyst makes them practically indispensable tools of his art. The student should therefore drill himself in the application of these tools by having someone make up problem after problem for him to practice upon, until he acquires facility in their use and feels competent to apply them in practice whenever the least opportunity presents itself. This will save him much time and effort.

Analytical Key for Military Cryptanalysts, Part II *

(Numbers in parentheses refer to Paragraph Numbers in this text)



* For explanation of the use of this chart see Par. 50 of Military Cryptanalysis, Part I.