



NATIONAL SECURITY AGENCY
FORT GEORGE G. MEADE, MARYLAND 20755-6000

FOIA Case: 68177B
9 December 2020

This responds to your Freedom of Information Act (FOIA) request of 7 July 2012 for "a copy of Military Cryptanalytics, Part III, by Lambros D. Callamahos. Please review the sections marked as classified for possible declassification and release." A copy of your request is enclosed. Your request has been processed under the FOIA and the document you requested is on the enclosed CD. Certain information, however, has been deleted from the enclosure.

Some of the withheld information has been found to be currently and properly classified in accordance with Executive Order 13526. The information meets the criteria for classification as set forth in Subparagraph (c) of Section 1.4 and remains classified SECRET as provided in Section 1.2 of Executive Order 13526. The information is classified because its disclosure could reasonably be expected to cause serious damage to the national security. Because the information is currently and properly classified, it is exempt from disclosure pursuant to the first exemption of the FOIA (5 U.S.C. Section 552(b)(1)). The information is exempt from automatic declassification in accordance with Section 3.3(b)(1) of E.O. 13526.

In addition, this Agency is authorized by various statutes to protect certain information concerning its activities. We have determined that such information exists in this document. Accordingly, those portions are exempt from disclosure pursuant to the third exemption of the FOIA, which provides for the withholding of information specifically protected from disclosure by statute. The specific statutes applicable in this case are Title 18 U.S. Code 798; Title 50 U.S. Code 3024(i); and Section 6, Public Law 86-36 (50 U.S. Code 3605).

Since these deletions may be construed as a partial denial of your request, you are hereby advised of this Agency's appeal procedures.

You may appeal this decision. If you decide to appeal, you should do so in the manner outlined below. NSA will endeavor to respond within 20 working days of receiving any appeal, absent any unusual circumstances.

- The appeal must be sent via U.S. postal mail, fax, or electronic delivery (e-mail) and addressed to:

NSA FOIA/PA Appeal Authority (P132)
National Security Agency
9800 Savage Road STE 6932
Fort George G. Meade, MD 20755-6932

The facsimile number is 443-479-3612.

The appropriate email address to submit an appeal is
FOIARSC@nsa.gov.

- It must be postmarked or delivered electronically no later than 90 calendar days from the date of this letter. Decisions appealed after 90 days will not be addressed.
- Please include the case number provided above.
- Please describe with sufficient detail why you believe the denial of requested information was unwarranted.

You may also contact our FOIA Public Liaison at foialo@nsa.gov for any further assistance and to discuss any aspect of your request. Additionally, you may contact the Office of Government Information Services (OGIS) at the National Archives and Records Administration to inquire about the FOIA mediation services they offer. The contact information for OGIS is as follows:

Office of Government Information Services
National Archives and Records Administration
8601 Adelphi Rd. - OGIS
College Park, MD 20740
ogis@nara.gov
877-684-6448
(Fax) 202-741-5769

Sincerely,

A handwritten signature in black ink, appearing to read "Ronald Mapp", with a stylized "for" written to the left of the main signature.

RONALD MAPP
Chief, FOIA/PA Office
NSA Initial Denial Authority

Encls:
a/s

~~SECRET~~

NATIONAL SECURITY AGENCY

MILITARY CRYPTANALYTICS

Part III

By

LAMBROS D. CALLIMACHOS

October 1977

Classified by DIRNSA/CHCSS (NSA/CSSM 123-2)
Exempt from GDS, EO 11652, Cat 2
Declassify Upon Notification by the Originator

National Security Agency
Fort George G. Meade, Maryland

~~SECRET~~

~~SECRET~~

NATIONAL SECURITY AGENCY

MILITARY CRYPTANALYTICS

Part III

By

LAMBROS D. CALLIMAHOS

October 1977

Classified by DIRNSA/CHCSS (NSA/CSSM 123-2)
Exempt from GDS, EO 11652, Cat 2
Declassify Upon Notification by the Originator

National Security Agency
Fort George G. Meade, Maryland

~~SECRET~~

Give me an ounce of civet, good apothecary, to sweeten my imagination.

—Shakespeare.

(King Lear, Act IV, Sc. 6)

Preface

1. I wish to acknowledge my indebtedness to William F. Friedman in drawing upon portions of his early work, "Military Cryptanalysis, Part III," for much of the material treated in Chapters I–V. Chapters IV–XI are revisions of seven of my monographs in the *NSA Technical Literature Series*, viz.: Monograph No. 19, "The Cryptanalysis of Ciphertext and Plaintext Autokey Systems"; Monograph No. 20, "The Analysis of Systems Employing Long or Continuous Keys"; Monograph No. 21, "The Analysis of Cylindrical Cipher Devices and Strip Cipher Systems"; Monograph No. 22, "The Analysis of Systems Employing Geared Disk Cryptomechanisms"; Monograph No. 23, "Fundamentals of Key Analysis"; Monograph No. 15, "An Introduction to Teleprinter Key Analysis"; and Monograph No. 18, "Ars Conjectandi: The Fundamentals of Cryptodiagnosis."

2. I also wish to acknowledge my indebtedness to Francis T. Leahy for keeping me out of statistical mischief, and to Bruce W. Fletcher for his expert assistance in the final proofreading, and for checking the cryptograms and the various diagrams.

—L. D. C.

TABLE OF CONTENTS

MILITARY CRYPTANALYTICS, PART III

Aperiodic Substitution Systems

(b) (1)
 (b) (3) -18 USC 798
 (b) (3) -50 USC 3024 (i)
 (b) (3) -P.L. 86-36

Chapter	Page
I. Introduction	1
1. Preliminary remarks. 2. General remarks on cryptographic periodicity. 3. Effects of varying the length of plaintext groupings. 4. Primary and secondary periods; resultant periods. 5. Cryptographic principles of aperiodic systems. 6. Fundamental cryptanalytic considerations in the solution of aperiodic systems.	
II. Systems using constant-length keying units to encipher variable-length plaintext groupings	7
7. General remarks. 8. Aperiodic encipherment produced by plaintext sequences grouped according to word lengths. 9. Solution when known cipher alphabets are involved. 10. Solution when unknown cipher alphabets are involved. 11. Solution by means of idiomorphs and the probable-word method. 12. Solution by means of isomorphs. 13. Additional remarks.	
III. Systems using variable-length keying units to encipher constant-length plaintext groupings	25
14. General. 15. Plaintext interruptor systems. 16. Ciphertext interruptor systems. 17. Systems employing externally generated or determined keys. 18. Solution when known cipher alphabets are employed. 19. Solution when unknown cipher alphabets are employed. 20. Additional remarks.	
IV. Ciphertext autokey systems	41
21. The cryptography of autokey encipherment. 22. Solution of ciphertext autokeyed cryptograms when known cipher alphabets are employed. 23. Principles of solution by frequency analysis. 24. Example of solution by frequency analysis. 25. Solution by means of isomorphs. 26. Solution of isologs involving the same pair of unknown primary components. [REDACTED] 28. Further remarks on ciphertext autokey systems.	
V. Plaintext autokey systems	75
29. Preliminary remarks on plaintext autokeying. 30. Solution of plaintext autokey systems when known cipher alphabets are employed and the introductory key consists of a single letter. 31. Solution of plaintext autokey systems involving known cipher alphabets when the introductory key consists of several letters. 32. Analysis of a case involving unknown components. [REDACTED] 34. Analysis of digital plaintext autokey systems. 35. Concluding remarks on autokey systems.	
VI. Systems employing long or continuous keys	121
36. Preliminary remarks. 37. Depth and its exploitation. 38. Solution of a single cryptogram involving known primary components and an unknown plaintext running key. [REDACTED] 41. Recovery of plain texts and the unknown primary components from a number of messages in flush depth. [REDACTED]	

VII. Cylindrical cipher devices and strip cipher systems	151
45. General. 46. Reconstruction of unknown cipher alphabets. 47. Analysis of cryptograms involving known alphabets but with unknown keys. 48. Further remarks.	
VIII. Systems employing geared disk cryptomechanisms	173
49. Introduction. 50. The Wheatstone cipher device. 51. Analysis of the Wheatstone cipher device. 52. The Kryha cipher machine. 53. Analysis of the original Kryha machine.	
IX. Fundamentals of key analysis	227
56. Convenient sources of key. 57. Manual key generation methods. 58. Mechanical and electronic key generators. 59. General analytic approaches. 60. Analysis of key in a double transposition cipher.	
62. Concluding remarks.	
X. Teleprinter key analysis	263
63. General. 64. Teleprinter key generation methods. 65. Analysis of combination streams.	
XI. Principles of cryptodiagnosis	323
71. General. 72. The basic steps in diagnosis. 73. The diagnostician and his attributes. 74. Embarking on the unknown cryptosystem. 75. Preliminary actions in attacking the unknown cryptosystem. 76. First step: manipulating the data. 77. Second step: recognizing the phenomena. 78. Third step: interpreting the phenomena. 79.	
Post mortem.	
XII. Concluding remarks	415
81. Special cases of aperiodic encipherment. 82. Analysis and solution of a first case. 83. Analysis and solution of a second case. 84. Final remarks.	
APPENDICES	
1. De Profundis; or the ABC of Depth Reading.....	437
2. Synoptic Tables, Cipher Device M-94.....	447
3. Tables of the Poisson distribution.....	463
4. Table of the Binomial distribution for $p = \frac{1}{10}$	537
5. Plaintext and random material for sampling purposes.....	553
6. Basic letter frequency data, 24 foreign languages.....	561
7. Problems—Military Cryptanalytics, Part III.....	611
INDEX	653

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024(i)
(b) (3)-P.L. 86-36

(b) (1)
 (b) (3) -18 USC 798
 (b) (3) -50 USC 3024 (i)
 (b) (3) -P.L. 86-36

CHAPTER I

INTRODUCTION

	Paragraph
Preliminary remarks.....	1
General remarks on cryptographic periodicity.....	2
Effects of varying the length of plaintext groupings.....	3
Primary and secondary periods; resultant periods.....	4
Cryptographic principles of aperiodic systems.....	5

1. **Preliminary remarks.**—*a.* This text constitutes the third in the series of six basic texts on the science of cryptanalytics.¹ The first two texts together have covered most of the necessary fundamentals of cryptanalytics; this and the remaining three texts will be devoted to more specialized and more advanced aspects of the science.

b. It is assumed that the cryptanalyst reader has studied *Military Cryptanalytics, Parts I and II*, and is familiar with the cryptologic terminology, concepts, principles, and techniques of solution of the various cryptosystems treated in those texts. *This general background is a necessary prerequisite to the thorough understanding of the principles expounded in this and the succeeding volumes.* Where appropriate, however, reference will be made to particular portions of the first two volumes; the reader would be wise to have these volumes handy when undertaking the study of this present text.

c. The text immediately preceding this one dealt with various types of periodic polyalphabetic substitution, commonly called *repeating-key systems*. It was seen in these repeating-key systems how a regularity in the employment of a limited number of alphabets, or even the employment of a complete set of alphabets in succession as in a progressive alphabet system, results in the manifestation of periodicity or cyclic phenomena in the cryptogram, by means of which the latter may be solved. The difficulty of solution is directly correlated with the type and number of cipher alphabets employed in specific examples.

d. Two procedures might suggest themselves to an enemy cryptographer for consideration if he realizes the foregoing circumstances and he thinks of methods to eliminate the weaknesses inherent in repeating-key ciphers. First, noting that the difficulties in solution increase as the length of the key increases, he might consider employing much longer keys as a means of increasing the security of the messages. Upon second thought, however, if the enemy cryptographer recognizes that, as a general rule, the first step in the solution of these ciphers consists in ascertaining the number of alphabets employed, it might seem to him that the most logical thing to do would be to use a procedure which will avoid periodicity altogether, eliminating the cyclic phenomena that are normally manifested within cryptograms of periodic construction, and thus foil even a first step towards solution. In other words, the cryptographer might progress from the use of rather short repeating keys (of perhaps no more than a dozen letters or so) to the use of key phrases of, let us say, 25-40 letters or thereabouts; subsequently, he might embark upon the use of keying procedures which would have the effect of producing keys of a length approximately equal to that of the average message being enciphered; and finally, he might advance to a stage of keying sophistication wherein the key consists of hundreds or thousands of elements, or even of an infinite number of elements (as, for example, in autokey systems).

¹ Before the echoes of the first sentence of this third volume have died down, the distinction between the *science* of cryptanalytics and the *art* of cryptanalysis should be re-emphasized. The cryptanalyst pursues studies along general and detailed lines, in order to equip himself technically for the duties of the moment or of the future. This parallels quite closely the technical studies of a violinist, who progresses from elementary exercises to the études of Kreutzer and Rode and finally to the Caprices of Paganini; in the meanwhile, the violinist has also studied various solo works and chamber music as a means of enhancing his comprehension and appreciation of music in general. All that a technical background does for the violinist is to give him the means of artistic expression or synthesis of musical thoughts from the coding of clefs, keys, and notes; all that a technical background does for the cryptanalyst is to give him the means for imaginative expression or synthesis of plaintext meanings from the coding of systems, keys, and characters. See also in this connection footnote 5 on p. 3 of *Military Cryptanalytics, Part I*.

e. At this point in our discussion it would be well to examine two terms defined in the previous volume:

- (1) *periodic system*. A system in which the enciphering process is repetitive in character and which usually results in the production of cyclic phenomena in the cryptographic text.
- (2) *aperiodic system*. A system in which the method of keying does not bring about cyclical phenomena in the cryptographic text.

The foregoing are *practical* definitions—nobody in his right mind (and that of course includes all of our readers)² would classify a Hagelin C-38 system³ as periodic just because it really *is* periodic with a finite cycle of 26x25x23x21x19x17 or 101,405,850; nor would the same right-minded individual quibble with the classification of a system as aperiodic if the length of the key is only 1000 letters and messages very rarely exceed that length. In brief, what we are in effect saying is that, even if a system embraces in its principle a fixed cycle or period, unless the period is considerably shorter than the messages being enciphered (thus permitting the manifestation of cyclic phenomena), the system may nevertheless for all practical purposes be considered as aperiodic *since the solution of a message is not predicated on writing the cipher text on several superimposed cycles and then solving the cryptographic depth thus produced*.

f. In this text we shall first examine varieties of aperiodic (as just defined) polyalphabetic substitution systems; then we shall study methods of extending or lengthening short mnemonic keys, followed by systems using lengthy keys (to include digital and teleprinter systems). Subsequently, we shall study methods of solution of some typical cryptomechanisms and cipher machines, and aperiodic combination systems. The text proper will end with a discussion of principles of key analysis as applied in manual and machine cryptosystems, followed by an extensive treatment of cryptodiagnosis. The appendices include useful cryptologic and cryptomathematical reference material, concluding with a course of problems designed to insure comprehension of the principles expounded in this volume.

2. General remarks on cryptographic periodicity.—a. When we consider the nature of periodicity in polyalphabetic substitution systems, we note that it is composed of *two* fundamental factors, because there are in reality *two* elements involved in its production. We have appreciated the fact that periodicity necessitates the use of a keying element employed in a cyclic manner; now we begin to realize that there is also another element involved, *viz.*, that unless the key is applied to constant-length plaintext groupings, no periodicity will be manifested externally in the cipher text, despite the repetitive or cyclic use of a constant-length key. This realization is quickly followed by the idea that possibly all periodicity may be avoided or suppressed by either or both of two ways: (1) by using constant-length keying units to encipher variable-length plaintext groupings, or (2) by using variable-length keying units to encipher constant-length plaintext groupings.

b. In the usual types of polyalphabetic substitution systems, successive letters of the repeating key are applied to successive letters of the text. With respect to the employment of the key, the cryptographic process may be said to be *constant* or *fixed* in character. This is true even if a single keying unit serves to encipher two or more letters at a time, provided only the groupings of plaintext letters are constant in length. In all such cases of encipherment by constant-length groupings, the apparent length of the period (as found by applying the factoring process to the cryptograms) is a multiple of the real length and the multiple corresponds to the length of the groupings, i.e., the number of plaintext letters enciphered by the same key letter. It is to be noted, however, that all these cases are still periodic, because *both* the keying units and the plaintext groupings are constant in length.

3. Effects of varying the length of plaintext groupings.—a. Now let us consider the effects of making either one or the other of these two elements *variable* in length. Suppose that the plaintext groups are made variable in length and that the keying units are kept constant in length. Then, even though the key may be cyclic and may repeat itself many times in the course of encipherment, external periodicity is suppressed, *unless the law governing the variation in plaintext groupings is itself cyclic, and the length of the message is greater than that of the cycle applicable to this variable grouping*.

² To scholars of English who experience a quick intake of breath at this point, the author hastens to clarify that the parenthetical phrase is intended to modify only the four immediately preceding words.

³ Cf. pp. 458-464 of *Military Cryptanalytics, Part II*.

b. As an example illustrating the italicized portion of the preceding sentence, let us suppose the correspondents agree to use reversed standard cipher alphabets with the key word SIGNAL, and that in the encryption the message is divided into groups as shown below:

S	I	G	N	A	L	S	I	G	N	A	L	S	I	G
1	12	123	1234	12345	1	12	123	1234	12345	1	12	123	1234	12345
C	OM	MAN	DING	GENER	A	LF	IRS	TARM	YHASI	S	SU	EDO	RDER	SEFFE
Q	UW	UGT	KFAH	UWNWJ	L	HN	ARQ	NGPU	PGNVF	I	TR	OPE	RFER	OCBBC
N	A	L	S	I	G	N	A	L	S	I	G	N	A	L
1	12	123	1234	12345	1	12	123	1234	12345	1	12	123	1234	12345
C	TI	VET	WENT	YFIRS	T	AT	NOO	NDIR	ECTIN	G	TH	ATT	ELEP	HONES
L	HS	QHS	WOFZ	KDARQ	N	NU	NMM	YIDU	OQZKF	C	NZ	NUU	WPWL	EXYHT
S	I	G	N	A	L	S	I							
1	12	123	1234	12345	1	12	123							
C	OM	MAS	WITC	HBOAR	D	SC	OMM...							
Q	UW	UGO	RFUL	TZMAJ	I	AQ	UWW							

Cryptogram

QUWUG TKFAH UWNWJ LHNAR QNGPU PGNVF ITROP ERFER OCBBC LHSQH
 SWOFZ KDARQ NNUNM MYIDU OQZKF CNZNU UWPWL EXYHT QUWUG ORFUL
 TZMAJ IAQUW W...

The cipher text in this example shows a tetragraphic and a pentagraphic repetition. The two occurrences of QUWUG_c (=COMMA_p) are separated by an interval of 90 letters; the two occurrences of ARQN_c (=IRST_p) by 39 letters. The first repetition (QUWUG_c), it will be noted, is a true periodic repetition, since the plaintext letters, their groupings, and the key letters are identical. The interval in this case, if counted in terms of letters, is the product of the keying cycle, 6, and the grouping cycle, 15. The second repetition (ARQN_c) is not a true periodic repetition in the sense that both cycles have been completed at the same point, as is the case in the first repetition. It is true that ARQN_c, representing IRST_p, both times, is a *causal* repetition produced by the action of the same combination of key letters, I and G, but the enciphering points in the grouping cycle are different in the two occurrences. Repetitions of this type may be termed *partially periodic* repetitions, to distinguish them from those of the *completely periodic* type.

c. When the intervals between the two repetitions noted above are more carefully studied, especially from the point of view of the interacting cycles which brought them about, it will be seen that, counting according to *groupings* and not according to single letters, the two pentagraphs QUWUG_c are separated by an interval of 30 groupings. Or, if one prefers to look at the matter in the light of the keying cycle, the two occurrences of QUWUG_c are separated by 30 key letters. Since the key is but 6 letters long, it has gone through 5 cycles. Thus, the number 30 is the product of the number of letters in the keying cycle (6) and the number of different-length groupings in the grouping cycle (5). The interaction of these two cycles is like that of two gears in mesh, one driven by the other. One of these gears has 6 teeth, the other 5, and the teeth are numbered. If the two gears are adjusted so that the teeth marked "1" are adjacent to each other and the gears are caused to revolve, these two teeth will not come together again until the larger gear has made 5 revolutions and the smaller one 6. During this time, a total of 30 meshings of individual teeth will have occurred. But since one revolution of the smaller gear (=the grouping cycle) represents the encipherment of 15 letters (when translated in terms of letters), the 6 complete revolutions of this gear mean the encipherment of 90 letters. This accounts for the period of 90, when stated in terms of letters.

d. The two occurrences of the other repetition, ARQN_c, are at an interval of 39 letters; but in terms of the number of intervening groupings, the interval is 12, which is obviously two times the length of the keying cycle. In other words, the key has in *this* case passed through two cycles.

e. In a long message enciphered according to such a scheme as the foregoing, there would be many repetitions of both types discussed above (the completely periodic and the partially periodic) so that

the cryptanalyst might encounter some difficulty in his attempts to reach a solution, especially if he had no information as to the basic system. It is to be noted in this connection that if any one of the groupings exceeds, say, 5 letters or so in length, the scheme may give itself away rather easily, since it is clear that *within each grouping the encipherment is strictly monoalphabetic*. Therefore, in the event of groupings of more than 5 or 6 letters, the monoalphabetic equivalents of telltale words such as **ATTACK**, **BATTALION**, **DIVISION**, would stand out. This system is most efficacious, therefore, with short groupings.

f. It should also be noted that there is nothing about the scheme which requires a regularity in the grouping cycle such as that embodied in the example. A lengthy grouping cycle guided by a key of its own may just as easily be employed; for example, the *number* of dots and dashes contained in the International Morse signals⁴ for the letters composing the 25-letter key phrase **DECLARATION OF INDEPENDENCE** might be used. Thus, A (. —) has 2, B (— . . .) has 4, and so on. Hence:

D E C L A R A T I O N O F I N D E P E N D E N C E
3 1 4 4 2 3 2 1 2 3 2 3 4 2 2 3 1 4 1 2 3 1 2 4 1

The grouping cycle is 3+1+4+4+2+ . . . , or 60 letters in length, and if the same phrase is used as a repeating key the total period would of course be 60, since after the encipherment of 60 letters the first key letter would be used again to encipher 3 letters, and so on, repeating the cycle. Suppose, however, that the foregoing 60-element keying pattern were used in conjunction with a different literal sequence for the actual key letters, say the 38-letter phrase **CONSTITUTION OF THE UNITED STATES OF AMERICA**. The period would then be the least common multiple of 38 and 60, or 1140 letters. This system might appear at first glance to yield a fairly high degree of cryptographic security; but this is not the case, as we shall presently see.

4. Primary and secondary periods; resultant periods.—*a.* It has been noted that the length of the complete period in a system such as the foregoing is the least common multiple of the length of the two component or interacting periods. In a way, therefore, since the *component* periods constitute the basic elements of the scheme, they may be designated as the *basic* or *primary* periods. These are also *hidden* or *latent* periods. The *apparent* or *patent* period, that is, the complete period, may be designated as the *resultant* or *secondary* period. In certain types of cipher machines there may be more than two primary periods which interact to produce a resultant period; also, there are cases in which the latter may interact with another primary period to produce a tertiary period, and so on.⁵ The *final*, or *resultant*, or *apparent* period is sometimes the one which is usually ascertained first as a result of the study of the intervals between repetitions. This may or may not be broken down into its component primary periods.

b. Although a solution may often be obtained without breaking down a resultant period into its component primary periods, the reading of many messages pertaining to a widespread system of secret communication is much facilitated when the analysis is pushed to its lowest level, that is, to the point where the final cryptographic scheme has been reduced to its simplest terms. This may involve the discovery of a multiplicity of simple elements which interact in successive cryptographic strata.

5. Cryptographic principles of aperiodic systems.—*a.* A discussion of the methods for avoiding periodicity was contained in the preceding text.⁶ A brief résumé of these methods is given below:

(1) Elements of a fixed or invariable-length key are applied to variable or irregular-length groupings of the plain text.

(2) Elements of irregular-length (variable-length) key are applied to regular and fixed groupings of the plain text.

(3) The principles of (1) and (2) are combined into a single system.

⁴ Cf. p. 23, *Military Cryptanalytics, Part I*.

⁵ An example of a cipher machine with several interacting latent periods is the Hagelin C-38. This machine produces in effect at any given moment six simultaneous reversed-standard-alphabet monoalphabetic substitutions in all 26 combinations of their presence or absence. The activity of each contributing monoalphabetic substitution is strictly periodic, with cycles of 26, 25, 23, 21, 19, or 17, conforming to the six regularly stepping pinwheels of the stated sizes. The total cycle of the machine is the product of the six relatively prime numbers, but the presence of individual subcycles constitutes one of the serious weaknesses of the machine.

⁶ Cf. par. 99, *Military Cryptanalytics, Part II*.

(4) The key does not repeat itself; this is brought about either by constructing a nonrepeating key, or by employing the key in a special manner (such as in plaintext- and ciphertext interruptor systems and plaintext- and ciphertext autokey systems).

b. From the standpoint of cryptographic mechanics, aperiodic systems may be divided into two main classes, viz.:

(1) Systems in which the key elements are not in any way determined or influenced by any elements of the plain or cipher text; and

(2) Systems in which the key elements are generated or governed by the plain text being enciphered or by the resultant cipher text.

⁷ Cf. par. 65 (on p. 157) of *Military Cryptanalytics, Part II*.

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024(i)
(b) (3)-P.L. 86-36

CHAPTER II

SYSTEMS USING CONSTANT-LENGTH KEYING UNITS TO ENCIPHER
VARIABLE-LENGTH PLAINTEXT GROUPINGS

	Paragraph
General remarks.....	7
Aperiodic encipherment produced by plaintext sequences grouped according to word lengths.....	8
Solution when known cipher alphabets are involved.....	9
Solution when unknown alphabets are involved.....	10
Solution by means of idiomorphs and the probable-word method.....	11
Solution by means of isomorphs.....	12
Additional remarks.....	13

7. General remarks.—*a.* The system described in subpar. 3*b* is obviously not to be classified as aperiodic in nature, despite the incorporation into the cryptosystem of a variable which in that case consisted of irregularity in the length of one of the two elements (key text and plain text) involved in polyalphabetic substitution. The variable there was subject to a law which in itself was periodic in character.

b. To make such a system truly aperiodic (under the definition given in subpar. 1*e*), by elaborating upon the basic scheme for producing variable-length plaintext groupings, would be possible, but impractical. For example, using the Morse code method illustrated in subpar. 3*f* for determining the key and simultaneously the lengths of the groupings, one might employ the text of a book; and if the book is longer than the message to be enciphered, the cryptogram would certainly show no periodicity as regards the intervals between any repetitions which might occur. However, as already indicated, such a scheme would not be very practical for regular intercommunication between a large number of correspondents, for reasons which are quite apparent. Encipherment and decipherment would be slow, cumbersome, onerous, and very subject to error; the book would have to be safeguarded as would a code book; and, unless the same key text were used for all messages, methods or indicators would have to be adopted to show exactly where encipherment begins in each message. Therefore a simpler method is desirable for producing constantly changing, aperiodic plaintext groupings.

8. Aperiodic encipherment produced by plaintext sequences grouped according to word lengths.—

a. The simplest method for producing aperiodic plaintext groupings is encipherment according to the actual word lengths of the message being encrypted. Although the *average* number of letters composing the words of any alphabetical language is fairly constant, *successive* words comprising plain text vary a great deal in this respect, and this variation is subject to no law.¹ In telegraphic English, for example, the mean length of words is 5.2 letters; the words may contain from 1 to 15 or more letters, but the successive words vary in length in an extremely irregular manner, no matter how long the text may be.

b. As a consequence, the use of word lengths for determining the number of letters to be enciphered by each key letter of a repetitive key suggests itself to a cryptographer as soon as he comes to understand the way in which repeating-key ciphers are solved. For, he asks, if there is no periodicity in the cryptograms, how can the letters of the cipher text written in 5-letter groups be distributed into their respective monoalphabets? And if this very first step is impossible, how can the cryptograms be solved? We shall see.

9. Solution when known cipher alphabets are involved.—*a.* Despite the foregoing rhetorical questions, the solution is really quite simple when the cipher alphabets involved are standard alphabets or are otherwise composed of known sequences. All that is involved is the completion of the plain-component sequence (preceded by, if the situation so demands, conversion into plain-component equivalents). In monoalphabetic substitution systems, all of the words of the entire message come out on a single gen-

¹ It is true, of course, that the differences between the vocabularies of two writers are often marked and can be measured. These differences may be subject to certain laws, but these laws are psychological rather than mathematical. See Rickert, E., *New Methods for the Study of Literature*, University of Chicago Press, Chicago, 1927.

eratrix in the completion diagram; in the case of the system discussed in subpar. 8b, since the individual, separate words of a message are enciphered by different key letters, *these words will reappear on different generatrices of the diagram*. All the cryptanalyst has to do is to pick them out; he can do this once he has found a good starting point, by using a little imagination and following clues afforded by the context.

b. As an example, let us consider the following intercepted message:

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

In the course of routine study of the message, the plain-component sequence is completed for the first 15 letters of the cryptogram, on the assumptions of direct and reversed standard cipher alphabets, as shown in Figs. 2a and b, respectively, below:²

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

FIGURE 2a

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

FIGURE 2b

c. In the diagram in Fig. 2b we note the word CAN at the beginning of one generatrix, then in the very next six columns the words YOU and GET in two other generatrices. That we should get some three-letter words on various generatrices is not particularly remarkable; (note the short words produced purely by accident in the generatrices of Fig. 2a) but that these words should follow one another in direct sequence in succeeding columns, and that the three words in question should be in excellent *contextual relationship* to form a plausible and *convincing* sentence beginning such as "CAN YOU GET . . ."

² One of the first things, if not the very first, to be done to a cryptogram in an undiagnosed system is the completion of the plain-component sequence on the basis of standard alphabets. In certain cases a solution is sometimes achieved by this means that would be impossible by any other. The completion is painless if accomplished by sliding strips; its probability of success in an isolated case is small, but the ratio of the time expended to its potential value is very large. This is a typical illustration of the application of the maxim of the experienced cryptanalyst: "Try the simplest thing first."

is more than remarkable (=a probability of .01 of random occurrence)—it is astonishing (=random probability of .0001).³

d. From here on the rest of the solution follows easily. If the cryptanalyst comes to a temporary halt (as in the example in Fig. 2b) in recovering further words on the generatrices, he can search in subsequent positions of the generatrix diagram for more words to be disclosed, and then he can fill in the missing portions from context and take another look at the generatrices. Or, it might be simpler if the cryptanalyst recovers a fragment of the specific key for the message, and then expands this key by steps to assist in reading the rest of the plain text. For example, in the case under discussion the cryptanalyst would get U, N, and I as key letters⁴ for the successive words of the plain text CAN YOU GET; these letters suggest the words UNION, UNITED, UNIVERSITY, etc. The complete solution is given below, with the recovered specific key being UNITED NATIONS.⁵

U	N	I	T	E	D	N	A	T	I	O	N	S
CAN	YOU	GET	IN	TOUCH	WITH	SECOND	DETACHMENT	STOP	LINE	OUT	OF	ORDER
SUH	PZT	CEP	LG	LQKCX	HVKW	VJLZAK	XWHAYTOWNH	BAFE	XAVEQ	AUV	ZI	EBPOB

The only minor difficulty of such a solution is that of making the first step and getting a good start on a word. If the words are short it is rather easy to overlook good possibilities and thus spend some time in fruitless searching. However, solution must come; if nothing good appears at the beginning of the message, search should be made in the interior of the cryptogram or at the end.

10. Solution when unknown cipher alphabets are involved.—a. It has been seen from the foregoing that solution of cryptograms involving word-length encipherment by standard alphabets is rather trivial, not because there is any magical quality about standard alphabets, but because the components are *known sequences*. If any other components had been used, say a plain component based upon a HYDRAULIC keyword-mixed sequence and a cipher component based on a QUESTIONABLY keyword-mixed sequence, and if these components were *known*, the problem would have been pursued in exactly the same way, *viz.*, conversion of the cipher letters of the cryptogram into their plain-component (HYDRAULIC . . . XZ) equivalents and completion of the plain-component (HYDRAULIC . . . XZ) sequence.

b. But what if one or both of the components are unknown mixed sequences? The simple procedure of completing the plain-component sequence obviously cannot be used. Since the messages are poly-alphabetic, and since the process of factoring cannot be applied, it would seem that the solution of messages enciphered in different alphabets and according to word lengths would be a difficult matter. Nevertheless, as is about to be demonstrated, the solution, even of a single message, is not nearly so difficult as first impression might lead one to imagine; the *modus operandi* will be explained in pars. 11 and 12.

11. Solution by means of idiomorphs and the probable-word method.—a. The first case to be studied involving unknown alphabets will be one wherein the original word lengths are retained in the cryptogram; this case will be discussed not because it is often encountered in practical military cryptography, but because it affords a good introduction to the usual case in which the original word lengths are no longer in evidence in the cryptogram, the latter appearing in the customary 5-letter groups. If the words

³ We must never forget that probabilities are influenced by the *amount* of material under examination; if we looked at *enough* material, it might not be at all astonishing if we obtained even a 10-letter word by accident. In all the probability considerations in this text, unless otherwise stated it is assumed that we are dealing with a limited amount of traffic, limited enough so that a probability of .01 is remarkable, and a probability of .001 exciting.

⁴ The key letters are assumed to be under A_p as the index letter. Throughout this text, whenever encipherment processes are under discussion, the pair of enciphering equations commonly referred to as characterizing the so-called Vigenère method will be understood unless otherwise indicated. This method involves the pair of enciphering equations $\theta_{k/2} = \theta_{i/1}$; $\theta_{p/1} = \theta_{e/2}$. That is, the index letter, which is usually the initial letter of the plain component, is set opposite the key letter in the cipher component; the plaintext letter to be enciphered is sought in the plain component and its equivalent is the letter opposite it in the cipher component. See in this connection subpar. 13f of *Military Cryptanalytics Part II*.

⁵ This is the specific key as recovered from this single message. It is quite possible that the complete key is UNITED NATIONS ORGANIZATION, UNITED NATIONS SECURITY COUNCIL, etc.; a longer message would prove whether the key is UNITED NATIONS used repetitively, or whether it is a phrase beginning with these two words.

of a message are enciphered monoalphabetically, the true and complete idiomorphs of word patterns will be patent, regardless of the identity of the particular alphabet used in the encryption of each word. These idiomorphs and word lengths can then be used as a basis for the probable-word method of attack.

b. Let us study the following low-echelon ground message in which the actual word lengths have been preserved in the cipher text:⁶

IUITD QHIWE LVCGWPCLZ RP NIV GYPYSYCV NC IXHCXWUJ ORS ZXH
GRPPRVQDOB SE OKYNMMHKV GUJLTN MYIN WZ IVURNI CLSWZVHS

We note some strong idiomorphic sequences, in particular the following:

- (1) IUITD (2) GYPYSYCV (3) GRPPRVQDOB (4) OKYNMMHKV
aba abaca abba abcddea

Looking up these patterns in idiomorph lists,⁷ and guided by the delimitations of the words, we arrive at the following assumptions:

- (1) IUITD (2) GYPYSYCV (3) GRPPRVQDOB (4) OKYNMMHKV
ENEMY DIVISION BATTALIONS ARTILLERY

The cipher values of these plain-cipher equivalencies are entered into a sequence reconstruction matrix of four levels (representing the four word assumptions), as follows:

P:	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)							I									T	U									D
(2)				G					Y							V	C		S		P					
(3)	R	G							Q				V		O	D			B	P						
(4)	O			H				N			M							K	Y						V	

FIGURE 3a

Noting in lines (2) and (3) that the intervals between the letters G, V, and P are the same in both cases, we can assume direct symmetry⁸ of position. In a few moments our reconstruction matrix will look like this:

P:	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)	S	H	B	P	I	N			M		R	G	T	U	K	Y	Q				V	C	O	D		
(2)	M		R	G	T	U	K	Y	Q				V	C	O	D		S	H	B	P	I	N			
(3)	R	G	T	U	K	Y	Q					V	C	O	D		S	H	B	P	I	N			M	
(4)	O	D		S	H	B	P	I	N			M		R	G	T	U	K	Y	Q			V	C		

FIGURE 3b

c. The rest of the plain text can be recovered either by (1) completion of the plain-component sequence, insofar as possible, in order to reveal further plaintext fragments which may be expanded and thus make possible the filling in of additional values in the cipher component, or by (2) recovery and expansion of the partial specific key for the message. An important additional step in solution is the recovery of the missing letters in the cipher component by analysis of the construction of the component in cases of systematic derivation. These points will be taken up in order in the subparagraphs below.

⁶ Foolish as this may be, it has happened in operational practice.

⁷ Cf. Appendix 3, *Military Cryptanalytics, Part I*.

⁸ Cf. par. 28, *Military Cryptanalytics, Part II*.

(1) Let us complete the plain-component sequence on the second and third words of the message, after first converting the cipher letters into their plain-component equivalents (where known), using for this purpose the uppermost cipher alphabet given in Fig. 3b. This is shown in the illustration below:

IUITD	QHIWE	LVCGWPC LZ
ENEMY	SBE	VWL DW
	TCF	WXM EX
	UDG	XYN FY
	VEH	YZO GZ
	WFI	ZAP HA
	XGJ	ABQ IB
	YHK	BCR JC
	ZIL	CDS KD
	AJM	*DET LE
	BKN	EFU MF
	*CLO	FGV NG
	DMP	GHV OH
	ENQ	HIX PI
	*FOR	IJY QJ
	GPS	JKZ RK
	HQT	KLA SL
	IRU	LMB TM
	JSV	MNC UN
	KTW	NOD VO
	LUX	OPE WP
	MVY	PQF XQ
	NWZ	QRG YR
	OXA	RSH ZS
	PYB	*STI AT
	QZC	TUJ BU
	*RAD	UVK CV

FIGURE 4a

LVCGWPC LZ
GVWL DW
HWXM EX
IXYN FY
JYZO GZ
KZAP HA
LABQ IB
MBCR JC
NCDS KD
ODET LE
PEFU MF
QFGV NG
RGHW OH
SHIX PI
TIJY QJ
UJKZ RK
VKLA SL
WLMB TM
XMNC UN
YNOD VO
ZOPE WP
APQF XQ
BQRG YR
CRSH ZS
DSTI AT
ETUJ BU
FUVK CV

FIGURE 4b

LVCGWPC LZ
HVWL DW
IWXM EX
JXYN FY
KYZO GZ
LZAP HA
MABQ IB
NBCR JC
OCDS KD
PDET LE
QEFU MF
RFGV NG
SGHW OH
THIX PI
UIJY QJ
VJKZ RK
WKLA SL
XLMB TM
YMNC UN
ZNOD VO
AOPE WP
BPQF XQ
CQRG YR
DRSH ZS
*ESTI AT
FTUJ BU
GUVK CV

FIGURE 4c

The generatrices with the most plausible possibilities for the continuation of plain text are marked with an asterisk. If the context of the message cannot be gotten from this diagram, what we can do is to take the third word, LVCGWPC LZ, and assume that the letters for which we have no plain-component equivalents in the first cipher alphabet of Fig. 3b represent one of the eight missing plaintext letters, G, H, J, P, R, T, U, or Z. If we assume that the first letter (L_c) of this word represents G_p (on the first or conversion row of the generatrix diagram just beneath the ciphertext letters), we obtain the result shown in Fig. 4b; when we try $L_c = H_p$, as shown in Fig. 4c, we obtain an excellent plaintext tetragraph on the third generatrix from the bottom, and see that the word is **ESTIMATED**. The newly recovered values in the cipher alphabet will aid in establishing the remaining unknown letters in the generatrix diagrams for other words of the message.

(2) For the second method, let us refer again to Fig. 4a. The key letter used to encipher the first word, **ENEMY**, is S_k (assuming A_p to be the index letter in the usual Vigenère equation), since $I_c = E_p$. Now for the second word, if $Q_c = C_p$ (one of the asterisked good generatrices for this word), the key is Y_k ; if $Q_c = F_p$, $\theta_k = U$; and if $Q_c = R_p$, $\theta_k = H$. The first key digraphs thus formed, SY, SU, and SH, are all compatible as English word beginnings. For the third word of the message, considering the two asterisked generatrices in Fig. 4a, if $V_c = D_p$, $\theta_k = Q$; if $V_c = S_p$, $\theta_k = P$. Therefore the first three key letters are now resolved as SYP or SUP; SYP is quickly discarded, and SUP should be followed by an E, I (less likely), P, or R, suggesting words such as SUPERIOR, SUPPORT, or SUPREME. A quick check on the message establishes that, with $\theta_k = R$, the fourth word deciphers to AT_p . Proceeding in this fashion, we are able to recover the key and simultaneously the plain text in record time.

(3) In cases wherein the cipher component has been constructed in some systematic manner, analysis of its derivation will make possible recovery of the component in its entirety after a sufficient number of values has already been placed correctly.⁹ What constitutes "a sufficient number of values" depends upon the type of construction of the component, as well as the vagaries of the particular situation at hand. Taking for example the cipher component as established in Fig. 3b,

S H B P I N . . M . R G T U K . Y . Q . . V C O D . .

we observe the digraphic fragments BP and GT. If these are a part of a transposition-mixed sequence, the mechanics of the system would indicate that the fragments are part of the diagram B . . . G,
P Q R S T

which means that three of the letters CDEF lie in order between B and G, and that directly above them are the letters composing the key word for the transposition matrix. However, since R immediately precedes the G in the sequence, it appears that R is part of the key word and not part of the remaining

H Y D R

alphabetic portion. Thus the fragmentary matrix B . . G can be reconstructed, from which, with but
P Q S T

little imagination, the key word HYDRAULIC may be seen emerging, so that the entire component is derivable from the following diagram:

4 9 3 7 1 8 6 5 2
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

d. By means of the foregoing methods, we can establish that the primary components are the following:

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

The complete message and the specific key are given below:

S U P R E M E C O U
ENEMY FORCE ESTIMATED AT ONE DIVISION OF INFANTRY AND TWO
IUITD QHIWE LVCGWPCLZ RP NIV GYPYSYCV NC IXHCXWUJ ORS ZXH

R T O F T H E U (NITED STATES)
BATTALIONS OF ARTILLERY MOVING WEST OF NEWTON JUNCTION
GRPPRVQDOB SE OKYNMMHKV GUJLTN MYIN WZ IVURNI CLSWZVHS

Now that the components have become known sequences, the solution of subsequent messages enciphered with these components but with different specific keys is a simple matter, involving only a conversion of the cipher letters into their plain-component equivalents and a completion of the plain-component sequence. This point required re-emphasizing because in actual operational problems it is frequently forgotten.

e. The example in subpar. b involved a case of direct symmetry of position. If both the plain and the cipher components had involved mixed sequences, indirect symmetry of position would have

⁹ For a treatment of the cryptographic mechanics of systematically mixed sequences and their cryptanalytic recovery, see par. 51 (on pp. 86-90) of *Military Cryptanalytics, Part I*.

applied.¹⁰ As an example of such a case, let us suppose that the cipher text of the message in question had been different, and that the sequence reconstruction matrix in Fig. 3a had been the following:

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					H								U	L												O
2				J					P				Y	D				U		I						
3	R	C							L		U		M	N				Q	S							
4	O			W				S		Q							N		C							K

FIGURE 5

(1) We observe the proportion AR (Ø-1, 3-1)=ON (Ø-15, 3-15) which is duplicated in AR (Ø-1, Ø-18)=ON (4-1, 4-18);¹¹ this is indicative that symmetry extends to the Ø line, and therefore that the plain and cipher components are identical sequences. Consequently, we are able to chain to the Ø line, deriving the following sets of partial chains:

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

FIGURE 6

(2) We note that the fragmentary chains ONM and TSQ of the Ø-3 set appear to be parts of a keyword-mixed sequence in reverse; so, proceeding with the graphical method¹² of indirect symmetry, we assign to these chains the notation →, and then we arbitrarily assign the notation ↓ to the Ø-2 chains. The four sets of fragmentary chains will then be amalgamated into the diagram shown in Fig. 7a, below.

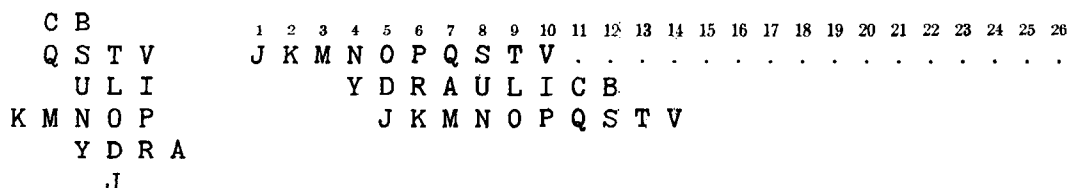


FIGURE 7a

FIGURE 7b

This diagram may then be expanded into that shown in Fig. 7b, consisting of the integration of two major chains tied together by the vertical VIP relationship.

(3) Now noting in Fig. 7b the sequence VCS on a diagonal and the letters S.V in the top row, we realize that the distance V to C when measured on the primary component should be 12, i.e., one-half of the distance (24) between V and S on the top row. Consequently, we may place the C at a position 12 spaces to the right of the V, which permits us to expand our diagram into the following:

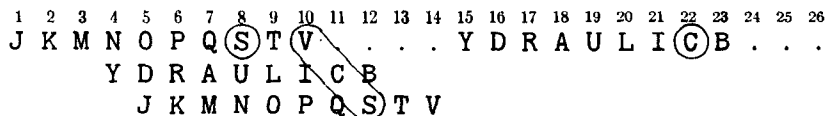


FIGURE 7c

(4) The fragmentary chains EH and EW in Fig. 6 could have been placed in their proper positions earlier in Figs. 7a-c; however, in order to illustrate a point, we have delayed their amalgamation until now. We note that the Ø-1 chains in Fig. 6 are at a decimation of -10 in the sequence in Fig. 7c; there is only one possible placement of the letters E and H at this interval, which then fixes the position of the

¹⁰ Cf. Chapter VI, *Military Cryptanalytics, Part II*.

¹¹ This notation has been discussed in footnote 2 on p. 92, *Military Cryptanalytics, Part II*.

¹² Cf. par. 46, *Military Cryptanalytics, Part II*.

last unused letter, W, the placement of which heretofore could have been ambiguous. These letters fit into the reconstructed sequence as follows:

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

It is a pleasure to use, without encountering a risk of cavilation the word "obvious"¹³ as regards the positions of the missing letters:

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

f. The immediately preceding example treated a case of identical sequences proceeding in the same direction for the plain and cipher components. If the cipher component had run in the reverse direction, or if the components had been two different [unknown] mixed sequences, indirect symmetry would still have applied, with the exception (and a very important exception indeed) that chains to the Ø line would have been excluded, all chaining being done within the matrix. This prohibition would result in the situation that not only would a single short message encrypted in such a system be well-nigh unsolvable, but that even if we had a long message or a small volume of traffic, it would probably be necessary to make a fairly large number of assumptions, all correct, before there would be enough data available to permit their manipulation and exploitation by indirect symmetry.

g. Now that we have understood the details of solution of cases wherein the true word lengths have been preserved, we will take up the situation wherein the cipher text has been transmitted in its usual form of 5-letter groups.

(1) Let us suppose that we have a number of messages, all of which are known to have been enciphered monoalphabetically by word lengths with the same pair of unknown primary mixed components and (although this is not a vital consideration) in the same message key.¹⁴ Five messages have been selected from the aggregate because of the presence of polygraphic repetitions between them; the beginnings of these messages are shown in Fig. 8a, below:

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

FIGURE 8a

¹³ The reader is reminded of the pithy anecdote on the word "obvious" quoted in footnote 11 on p. 6 of *Military Cryptanalytics, Part I*.

¹⁴ If this latter fact had not been known, it could have been conjectured, from an examination of the I.C.'s of groups of columns, that the same message key was used for all the messages. In the particular example in Fig. 8a, the I.C. of the first 5 columns (taken collectively) is 1.56, while that of the first 10 is 1.76, and thereafter the I.C. drops off very rapidly even though we are adding more data to our distribution for evaluation. The grouped I.C.'s for the first N columns are summarized in the diagram below:

N	I.C.	N	I.C.
5	1.56	12	1.53
6	1.55	15	1.33
8	1.73	20	1.37
10	1.76	25	1.24

The reason for the low I.C.'s of the first 5 and the first 6 columns is that the sample was insufficient to portray what we expect of English plain text; on the other hand, the reason for the high I.C.'s of the first 8 and the first 10 columns is that the beginning words of these messages probably exceed the average length (5.2 letters) of all English words.

(2) The 5-letter and 9-letter repetitions have the length and idiomorphic patterns of **ENEMY** and **ARTILLERY**, respectively. Taking into account that the average word length in telegraphic English plain text is 5.2 letters, it appears that both of these words were probably enciphered by the third letter of the message key,¹⁵ although the relative numerical identity of the particular alphabet is really of no concern to us at the moment. On the basis of the idiomorphic beginning, Message No. 3 could start with the word **AMMUNITION**, making the 4-letter repetition **TION** which is cryptolinguistically titillating; the first word of Message No. 1, **LOCATION**, comes immediately thereafter, which is followed by **COUNTERATTACK** at the beginning of Message No. 5, **HOSTILE** at the beginning of Message No. 4, and **THRUST** at the beginning of Message No. 2. From the solution of the first three words of these five messages, and with the concurrent exploitation of the direct symmetry manifested, the primary cipher component is established as

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

(3) The key letters (under **A_p**) of the first three alphabets are S, U, and P. The rest of the solution proceeds either by the generatrix method as outlined in subpar. 11c(1), or by analysis of the key as illustrated in subpar. 11c(2). The complete texts of the message beginnings are shown in Fig. 8b, below:

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

FIGURE 8b

(4) It may be seen from the foregoing example that the general theory of idiomorphic attack and the probable-word method remains the same for 5-letter texts as it is for text divided into bona fide word lengths; only the details of the execution differ. Where a small volume of homogeneous traffic is at hand, and something is known about the correspondents and the nature of the messages, solution should pose no problems (other than usual cryptanalytic headaches concomitant with operational situations of minor systems in which only a few messages are available).

12. Solution by means of isomorphs.—*a.* The phenomenon of isomorphism and an illustration of the exploitation of isomorphs in cipher text were covered in the previous volume.¹⁶ In practical cryptanalysis the phenomena of isomorphism afford a constantly astonishing source of clues and aids in solution. The alert cryptanalyst is always on the lookout for situations in which he can take advantage of these phenomena, for they are among the most interesting and most important in cryptanalytics.

¹⁵ We have already noted that the common plain-cipher equivalencies $E_p = L_c$ and $Y_p = H_c$ in the two words establish the fact that these words were enciphered by the same alphabet.

¹⁶ Cf. par. 71, *Military Cryptanalytics, Part II*.

b. Let us consider the case of word-length encipherment involving an unknown pair of primary components, the cipher text being transmitted in the customary 5-letter groups. The following cryptogram is available for study:

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

c. There are no long polygraphic repetitions in evidence. An isomorphic search,¹⁷ however, uncovers several isomorphic sequences indicated by the dotted lines above; these are grouped into the following two sets of isomorphs:

Set "A"

Set "B"

(α) L H J J T Y Z L D X Z H Y
 (β) P G Z Z I J F P K E F G J
 (γ) D V B B O W T D X S T V W

(δ) D F G O O B D
 (ϵ) T A U E E D T

If these isomorphs are causal isomorphs, i.e., isomorphs produced by the different encryptions of identical plaintext sequences, then the relationships between corresponding letters of the isomorphs reflect the relationships between different juxtapositions or slides of a cipher component against a plain component; these relationships, latent in the isomorphs, may be made patent through the mechanics of indirect symmetry.

d. The partial chains derivable from these isomorphs are given below:

α - β : L P H G Y J Z F T I D K X E
 α - γ : L D X S H V J B Z T O Y W
 β - γ : P D G V Z B I O J W F T K X E S
 δ - ϵ : B D T F A G U O E

Using the graphical method of indirect symmetry, these partial chains may be amalgamated into the diagrams shown in Figs. 9a and b, below. We note in Fig. 9a the $\leftarrow \begin{smallmatrix} 1 \\ 2 \end{smallmatrix} \right$ relationship of the letters XZ, ST,

OP, and EF, and conclude that the cipher component must be a keyword-mixed sequence. We now expand the diagram of Fig. 9a by placing the W in position diagonally ahead of the XZ, and we duplicate the remaining letters in their proper position with respect to the W just placed a moment before; this

Y W
 J B
 L D X S
 Z T O
 P K E
 F I
 A

FIGURE 9a

FIGURE 9b

Y W
 J B Y (W)
 L D (X) S
 (Z) T O J B
 P K E L D X S
 F I Z T O
 A P K E
 F I
 A

FIGURE 9c

Y W
 J B Y (W) H (V)
 L D (X) S G
 (Z) T O J B U
 P K E L D X S
 F I Z T O
 A P K E
 F I
 A

FIGURE 9d

¹⁷ For a systematic method of searching for isomorphs, see footnote 7 on p. 174 of *Military Cryptanalytics, Part II*.

is shown in Fig. 9c. This facilitates placing the diagram of Fig. 9b into the array (on the basis of the VWXZ diagonal), resulting in the final diagram shown in Fig. 9d. From this latter figure, the original cipher component, minus 5 letters, may be chained out:

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
B E F G J K . . O P . S T V W X Z H Y D . A U L I .

Our old friend HYDRAULIC.

e. We have the sequence for the cipher component, but now what? We could assume the plain component to be the normal sequence, direct and then reversed, and we could convert the first few cipher letters into their plain-component equivalents on these hypotheses and then complete the plain-component sequence; if we are correct in our assumption, a plaintext word would be revealed on one generatrix, another word on a different generatrix, etc. We could also assume the plain component to be the same as the cipher, in the same or in the reverse direction, and we could complete the plain-component sequence accordingly. All of these attempts fail, so it means that we are faced with a plain component of unknown composition. We have the cipher component at hand, it is true, but *unless we know or can deduce the motion of the cipher component*,¹⁸ it will be impossible for us to convert the cipher text into monoalphabetic terms; in other words, the original cipher is already reduced as far as it will go. Plaintext assumptions are now an absolute necessity.

f. It can be seen by referring to the two sets of isomorphs in sub-par. c that Set "A" has the complete idiomorphic pattern for COMMUNICATION, and that Set "B" has the idiomorphic pattern contained in ARTILLERY. If the now known cipher component is set down and the plaintext equivalents for the first occurrences of the assumed COMMUNICATION and ARTILLERY are recorded in the rows labelled P₁ and P₂ of the diagram below, direct symmetry of position will of course apply, provided that there

C:	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
P ₁	O	N	A						C						M						U				T	I
P ₂		R								E		T	I			A		L	Y							

is a tie-in letter between the sequences; there happen to be three such letters, so that the plain component may be expanded as follows:

O N A . L Y C M . R U E . T I

Since there are manifested the phenomena of a keyword-mixed sequence in the plain component, we may further expand the sequence into the following:

O N A . L Y C D F G H J K M . R U E . T I

If the key word for the sequence cannot be guessed from this partial sequence, we might finish the solution by a modification of the method indicated in subpar. 11c(1), with the difference that, in this case, not only will some of the cipher letters not have plain-component equivalents, but also that the plain component itself will have gaps in its sequence in the plain-component completion diagram. After the key word (QUESTIONABLY) for the plain component has been recovered, the solution can be completed by the generatrix method, keeping in mind the reconstruction of the message key as a means of quick

¹⁸ See in this connection subpar. 71d on p. 175 of *Military Cryptanalytics, Part II*.

analysis after a few letters of the key have been derived. The complete decipherment of the message is shown below:

[illegible]

FIGURE 10

The key for the message, under Q_p as the index letter, is "STRIKE WHILE THE IRON IS . . . (HOT?)"¹⁹

g. In connection with the solution of the problem in this paragraph, let us take a closer look at the isomorphs listed in subpar. *c*. These are given below, together with their plaintext equivalents:

Set "A"												Set "B"											
<u>C O M M U N I C A T I O N</u>												<u>(A) R T I L L E R (Y)</u>											
(α)	L	H	J	J	T	Y	Z	L	D	X	Z	(δ)	D	F	G	O	O	B	D				
(β)	P	G	Z	Z	I	J	F	P	K	E	F	(ϵ)	T	A	U	E	E	D	T				
(γ)	D	V	B	B	O	W	T	D	X	S	T												
	V	W																					

In case it had escaped attention before, note the ciphertext fragments **XZHY**, **EFGJ**, and **STVW** at the ends of the isomorphs of Set "A". These three tetragraphs, transparent in the cipher text, are actually fragments of the keyword-mixed sequence constituting the cipher component. The reason for their presence is not hard to find: the plaintext equivalent of the isomorphs ended with **TION** and the letters **TION** happened to be a fragment of the keyword-mixed sequence constituting the *plain* component. (Note also, from Set "B", that **AU** must also be in sequence in the cipher component.) This information would have been of assistance to us in the chaining process pursued in subpar. *d*; for pedagogical reasons, however, we delayed drawing attention to this situation until now. Needless to say, this situation or a recognizable variation of it could be of considerable assistance in the solution of a difficult problem of only a few messages in actual operations.

h. One more very important facet of isomorphism should be discussed at this point. Let us suppose that we have recovered the cipher component of the message under study through the exploitation of isomorphs as just demonstrated; but let us suppose that the two plaintext assumptions (**COMMUNICATION** and **ARTILLERY**) were insufficient to disclose enough of the sequence for the plain component to permit its facile recovery.²⁰ Additional plaintext assumptions are necessary, but we seem to have milked the

19 Hot.

²⁰ This would be the case if the plain component were not a keyword-mixed sequence but were, let us say, a transposition-mixed sequence.

cipher text dry with the two cribs we have already placed. The problem confronting us is how to make further "educated guesses" that might display a trace (or more, we hope) of erudition.

i. Since the cipher component has become a known sequence, let us set it down, numbering its elements serially from 1 to 26, 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
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

Let us first replace the letters of the cipher text by their numerical equivalents according to the HYDRAULIC sequence. We will then take a delta or lateral difference stream²¹ from these numerical values, by subtracting each number from the following one;²² however, instead of recording the numerical difference, we will record the literal equivalent of this numerical difference according to the HYDRAULIC sequence above. The result of this process is shown in Fig. 11, below:

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

FIGURE 11

j. We can see, by comparing Fig. 11 with the original plain text as given in Fig. 10, that the delta stream has revealed *all* of the polygraphic repetitions of trigraphs or better in the underlying plain text.²³ Note the IXF repetition in the delta stream, which means that the ciphertext sequences PHZF and IMKH must represent the same plain text (probably the word WITH, since it is a four-letter repetition following COMMUNICATION); the PHZF and IMKH sequences are actually isomorphic, but we were unable to recognize them as such until now. Also note the delta repetition FAESJH, which means that the ciphertext sequences YJPRXGJ and KHUNFZH (an isomorphic pair whose isomorphism we were unable to trust before, because of a lack of sufficient corroborative values in the isomorphic repetition pattern) must represent the same plain text (in this case, the assumption of the word THROUGH would be permitted). Note further the HH digraphic fragment in the delta stream, which means that GJK_e must represent ION_p (from COMMUNICATION); since this is not preceded by T_p, the assumption of S_p and therefore DIVISION is encouraged. With these plaintext values and those which follow as a direct result of our analysis thus far, it would be a simple matter to reconstruct the plain component almost *in toto* and the plain text of the message in its entirety.

13. Additional remarks.—a. One of the practical difficulties in employing systems in which the keying process shifts according to word lengths is that in handling such a message the deciphering clerk is often not exactly certain when the termination of a word has been reached, which results in the loss of time and effort. For instance, in deciphering a word such as INFORM, the clerk would not know whether he now has the complete word and should shift to the next key letter or not; the word might be INFORMS,

²¹ The application of delta stream techniques to the solution of digital cipher systems has been illustrated in Chapters XI and XII of *Military Cryptanalytics, Part II*.

²² In this process, subtraction is performed *mod* 26: i.e., if we are to subtract a large number from a smaller, we add 26 to the smaller before subtraction. For example, $1-7=(1+26)-7=20=Q$ in the scale above. 26 is equivalent to 0 in this modulus, so that $14-14=0=Z$.

²³ The plaintext repetitions are foreshortened by one letter in the delta stream; e.g., COMMUNICATION, a 13-letter word, appears as a 12-letter sequence in the delta stream.

INFORMED, INFORMING, INFORMAL, INFORMATION, etc. The past tense of verbs, the plural of nouns, and terminations of various sorts capable of being added to word roots would give rise to difficulties, and the latter would be especially troublesome if the messages contained a few telegraphic errors to boot. Consequently, *word separators* are often adopted to circumvent this source of trouble.²⁴ These separators usually consist of an infrequent letter, such as X_p or Q_p , which is placed after every word of the plain text and is encrypted along with the rest of the message.²⁵

(1) When word separators are employed and this fact is once suspected or discovered, their presence is of as much aid to the cryptanalyst in his solution as it is to the clerks who are to decipher the messages. As an example, let us study the following cryptogram:

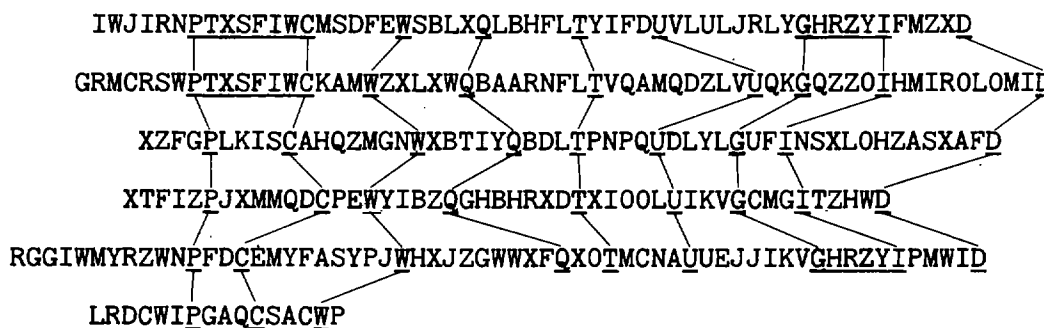
IWJIR	NPTXS	FIWCM	SDFEW	SBLXQ	LBHFL	TYIFD	UVLUL	JRLYG	HRZYI
FMZXD	GRMCR	SWPTX	SFIWC	KAMWZ	XLXWQ	BAARN	FLTVQ	AMQDZ	LVUQK
GQZZO	IHMIR	OLOMI	DXZFG	PLKIS	CAHQZ	MGNWX	BTIYQ	BDLTP	NPQUD
LYLGU	FINSX	LOHZA	SXAFD	XTFIZ	PJXMM	QDCPE	WYIBZ	QGHBH	RXDTX
IOOLU	IKVGC	MGITZ	HWDRG	GIWMY	RZWNP	FDCEM	YFASY	PJWHX	JZGWW
XFQXO	TMCNA	UUEJJ	IKVGH	RZYIP	MWIDL	RDCWI	PGAQC	SACWP	

Collateral information indicates that the cryptosystem involves monoalphabetic encipherment by word lengths, a word separator being used to signal the change to a new key letter; the key letters themselves form a plaintext word as a mnemonic key.

(2) If the encipherment is by word lengths and a word separator is used, the average length of words should be 6.2 letters. Since a key word is used to control the selection of alphabets, if a polygraphic repetition of significant length is present in the cipher text, the interval between the first and second occurrences should give a fair indication of the length of the key, unless there are repeated letters in the key and these polygraphic repetitions happen to be produced by identical key letters in *different* positions in the key word. We note the 8-letter repetition PTXSFIWC at an interval of 56 letters; this would seem to indicate that the key word is $\frac{56}{6.2}$ or 9 letters long, give or take a letter. Since there is another

polygraphic repetition present, GHRZYI at an interval of 224 letters, the division $\frac{224}{6.2} = 36 = 4 \times 9$ furnishes corroboration of the length of the key word, and dispels fears that these repetitions may have been produced by identical key letters in different positions in the key word.

(3) When word separators have been used, the first and last letters of long polygraphic repetitions are most likely to be word separators;²⁶ consequently, in the case of the first repeated sequence, PTXSFIWC (representing either the second or third word of the message), P_c and C_c should represent the word separators. Now if the cipher text of the message is written out in lines of 50-60 letters or so using the repeated sequence PTXSFIWC as a sort of base, we might be able to pick out the successive word separators; this is shown in the diagram below:



Note that the next-to-last separator, I_c , was preferred to a preceding Z_c (also a possibility as a word separator) because of the final letter of the GHRZYI repetition. With the word divisions now known, it is an easy matter to make plaintext assumptions based on word lengths and idiomorphs; the rest of the solution is left to the reader as an exercise.

(4) The foregoing example involved but a single message and a relatively short message key so that during the encryption five complete cycles of the message key were employed; it was this re-use of the keying sequence that permitted a solution. It is obvious that, regardless of the length of the key, if we had five or six messages or so in the same key, we could have written them out over one another and by careful scrutiny we could have determined the word separators much in the same fashion as was illustrated by the diagram of Fig. 12, above.

b. The systems thus far discussed are all based upon word-length encipherment using different cipher alphabets. Words are markedly irregular in regard to this feature of their construction and thus aperiodicity is imparted to such cryptograms. But variations in the method, aimed at making the latter somewhat more secure, are possible. Instead of enciphering according to natural word lengths, the irregular groupings of the text may be regulated by other agreements. For example, suppose that the numerical value (in the normal sequence) of each key letter be used to control the number of letters enciphered by the successive cipher alphabets. Depending then upon the composition of the key word or key phrase, there would be a varying number of letters enciphered in each alphabet. If the key word were PREPARE, for instance, then the first cipher alphabet would be used for 16 ($=P$) letters, the second cipher alphabet, for 18 ($=R$) letters, and so on. Monoalphabetic encipherment would therefore allow plenty of opportunity for telltale word patterns to manifest themselves in the cipher text. Once an entering wedge is found in this manner, solution would be achieved rather rapidly.

c. If the key of the system described in the foregoing subparagraph is short, and the message is long, periodicity will be manifested in the cryptograms, so that it would be possible to ascertain the length of the basic cycle (in this case, the length of the key) from a single message, despite the irregular groupings in encipherment. The determination of the length of the cycle might, however, present difficulties in some cases, since the basic or fundamental period would not be clearly evident because of the presence of repetitions which are not periodic in their origin. For example, suppose the word PREPARE were used as a key, each key letter being employed to encipher a number of letters corresponding to its numerical value in the normal sequence. It is clear that the length of the basic period, in terms of letters, would here be the sum of the numerical values of $P(=16) + R(=18) + E(=5) + \dots$, totalling 79 letters. But because the key itself contains repeated letters and because encipherment by each key letter is monoalphabetic, there would be plenty of cases in which the first letter P would encipher the same or part of the same word as the second letter P, producing repetitions in the cipher text. The same would be true as regards encipherments by the two R's and the two E's in this key word. Consequently, the basic period of 79 would be distorted or masked by aperiodic repetitions, the intervals between which would not be a function of, nor bear any relation to, the length of the key. Cases are frequently encountered in which a fundamental periodicity is masked or obscured by the presence of ciphertext repetitions not attributable to a fundamental cycle; the experienced cryptanalyst is on the lookout for phenomena of this type, when he finds in a polyalphabetic cipher plenty of repetitions but with no factorable constancy which leads to the disclosure of a short period. He may conclude, then, either that the cryptosystem involves several primary periods which interact to produce a long resultant period, or that it involves a fairly long fundamental cycle within which repetitions of a non-periodic origin are present and obscure the phenomena manifested by repetitions of periodic origin.²⁷

d. A logical extension of the principle of polyalphabetic encipherment of variable-length plaintext groupings is the case in which these plaintext groupings rarely exceed 4 letters, so that a given cipher alphabet is in play for only a very short time, thus breaking up what might otherwise appear as fairly long repetitions in the cipher text.

²⁷ See also in this connection footnote 8 on p. 26 of *Military Cryptanalytics, Part II*.

(1) For example, suppose that the letters of the alphabet to be used as key letters are arranged in the order of their relative frequencies in English plain text, and are set off into four groups of 5, 6, 7, and 8 letters, respectively, as follows:

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

Suppose that a key letter in Group 1 means that one letter will be enciphered; a letter in Group 2, that two letters will be enciphered; and so on. Suppose, next, that a rather lengthy phrase were used as a key; for example, I KNOW NOT WHAT COURSE OTHERS MAY TAKE BUT AS FOR ME GIVE ME LIBERTY OR GIVE ME DEATH. Suppose, finally, that each letter of the key were used to control the number of letters to be enciphered by the selected alphabet, according to the scheme outlined above. Such an enciphering scheme, using the HYDRAULIC . . . XZ primary cipher component sliding against a normal plain component, would yield the following groupings:

Grouping:	2	4	1	1	4	1	1	1	4	2	2	1	3	1	3	1	2	1	1	1	2	
Key:	I	K	N	O	W	N	O	T	W	H	A	T	C	O	U	R	S	E	O	T	H	
Plain:	TW	ENTI	E	T	HREG	I	M	E	N	THE	AD	QU	A	RTE	R	SNO	W	LO	C	A	T	ED
Cipher:	HR	PYIV	S	E	AKYR	X	R	Z	ENAY	HR	SX	T	ZYG	C	WPQ	Z	UC	G	O	K	AR	

Grouping:	1	1	2	3	2	3	1	2	4	1	4	3	1	2	2	3	1	1	
Key:	E	R	S	M	A	Y	T	A	K	E	B	U	T	A	S	F	O	R	
Plain:	N	E	AR	HEA	DQ	UAR	T	ER	SOFF	O	RTYS	ECO	N	DR	EG	IME	N	T	. . .
Cipher:	W	I	SF	VQM	IS	TYP	K	CT	LDQQ	X	HDIY	BIQ	C	IT	XH	QWM	A	V	

(2) Here it will be seen that any tendency for the formation of lengthy repetitions would be counteracted by the short groupings and quick shifting of alphabets. Before a long plaintext passage can be enciphered by *exactly* the same sequence of key letters, an interval of exactly 135 letters (the sum of the values of the letters in the key phrase) or a multiple thereof must intervene between the two occurrences of the plaintext passage.²⁸ When, however, a repeated plaintext passage is at an interval of only one or two letters off from 135 or a multiple of 135, there can occur in the system under discussion a phenomenon of intermittent coincidences; i.e., coincidences not among all the ciphertext letters representing the repeated plaintext passage, but among only a few of these ciphertext letters. As an example, let us consider the following message beginnings of two messages in flush depth:

Grouping:	2	4	1	1	4	1	1	1	4	2	2	1	3	1	3	1	2	1	1	1	2	
Key:	<u>I</u>	<u>K</u>	<u>N</u>	<u>O</u>	<u>W</u>	<u>N</u>	<u>O</u>	<u>T</u>	<u>W</u>	<u>H</u>	<u>A</u>	<u>T</u>	<u>C</u>	<u>O</u>	<u>U</u>	<u>R</u>	<u>S</u>	<u>E</u>	<u>O</u>	<u>T</u>	<u>H</u>	
Msg "A":	TW	ENTI	E	T	HREG	I	M	E	N	THE	AD	QU	A	RTE	R	SNO	W	LO	C	A	T	ED . . .
Cipher "A":	HR	PYIV	S	E	AKYR	X	R	Z	ENAY	HR	SX	T	ZYG	C	WPQ	Z	UC	G	O	K	AR	
Msg "B":	FI	FTYF	I	R	STDI	V	I	S	IONH	EA	DQ	U	ART	E	RSI	S	MO	V	I	N	G . . .	
Cipher "A":	GM	QIGQ	X	C	MNHU	F	Z	J	UFEA	AH	IS	M	CZY	T	VWJ	T	LC	U	Z	C	L	

The word HEADQUARTERS is offset one position to the right in Message "B" with respect to its position in Message "A". (This same situation would arise if the second occurrence of HEADQUARTERS in Message "A" were at an interval of 136 letters from the first occurrence, or 271 letters, or any other multiple of 135 plus one more letter.) If we set down the cipher equivalents of the two occurrences of HEADQUARTERS under the plain text, we have the following:

H	E	A	D	Q	U	A	R	T	E	R	S
<u>A</u>	<u>Y</u>	<u>H</u>	<u>R</u>	<u>S</u>	<u>X</u>	<u>T</u>	<u>Z</u>	<u>Y</u>	<u>G</u>	<u>C</u>	<u>W</u>
<u>A</u>	<u>A</u>	<u>H</u>	<u>I</u>	<u>S</u>	<u>M</u>	<u>C</u>	<u>Z</u>	<u>Y</u>	<u>T</u>	<u>V</u>	<u>W</u>

²⁸ Note that, in the case of this particular key, two occurrences of a 14-letter plaintext passage could receive identical encipherments at two different positions of the key at which there are identical fragments (GIVE ME) in the key phrase; the intervals between these repeated ciphertext sequences would have nothing to do with the length of the period.

We notice that the cipher equivalents agree only in the first, third, fifth, eighth, ninth, and twelfth letters. The repetitions here extend only to one or two letters; longer repetitions can occur only exceptionally. The two encipherments yield only occasional coincidences, i.e., places where the cipher letters are identical; moreover, the *distribution* of the coincidences is quite irregular and of an intermittent character. This phenomenon of intermittent coincidences, involving coincidences of single letters, pairs of letters, or short sequences (rarely ever exceeding pentagraphs) is one of the characteristics of this general class of polyalphabetic substitution, wherein the cryptograms commonly manifest what appears to be disturbed or distorted periodicity.

e. As has already been noted, in aperiodic systems wherein the key is determined or generated *apart* from the plain text being enciphered (as is the case of the example in the foregoing subparagraph), cryptographic depth is possible; therefore the analyst may be able, if keying conditions permit, to superimpose messages and solve the resultant superimposition. On the other hand, in systems wherein the plain or cipher text influences or governs in any way the selection of keys, cryptographic depth is usually impossible of establishment, except in very special circumstances.

f. The essence of the systems described in this chapter really comprises *monoalphabetic* substitution in irregular, and usually small, segments; nevertheless, these segments were large enough to permit of their isolation and exploitation. As the size of these segments decreases, ultimately to units of single letters, so does the difficulty of solution increase—but not beyond the potentials of cryptanalysis, as will shortly be demonstrated.

CHAPTER III

SYSTEMS USING VARIABLE-LENGTH KEYING UNITS TO ENCIPHER CONSTANT-LENGTH PLAINTEXT GROUPINGS

	Paragraph
General.....	14
Plaintext interruptor systems.....	15
Ciphertext interruptor systems.....	16
Systems employing externally generated or determined keys.....	17
Solution when known cipher alphabets are employed.....	18
Solution when unknown cipher alphabets are employed.....	19
Additional remarks.....	20

14. General.—a. The systems treated in the preceding chapter incorporated simple methods of eliminating or avoiding periodicity by enciphering variable-length groupings of the plain text, using constant-length keying units; the essence of those systems was really *monoalphabetic* encipherment by sections,¹ the sections comprising irregular-length plaintext groupings. In subpar. 2a, however, it was pointed out that periodicity can also be suppressed by applying variable-length keying units to constant-length plaintext groupings; the essence of such systems is polyalphabetic substitution applied to the plaintext units (usually single letters). One such method consists in *irregularly interrupting* the keying sequence, if the latter is of a limited or fixed length, and recommencing it (from its initial point) after such interruption, so that the keying sequence becomes equivalent to a series of keys of different lengths. Thus, the key phrase BUSINESS MACHINES might be expanded, by a particular keying convention, into a series of irregular-length keying sequences, such as BUSI/BUSINE/BU/BUSINESSM/BUSINESSMAC, etc. Various schemes or prearrangements for determining the type or character of the interruptions may be adopted. Several typical methods will now be described.

b. There are many methods of interrupting a keying sequence which is basically cyclic and which therefore would give rise to periodicity if not interfered with in some way. These methods may, however, be classified into six general cases as regards what happens after the interruption occurs.²

Case I: The keying sequence merely stops and begins again at the initial point of the cycle.

Case II: Certain elements of the keying sequence may "stutter" or be repeated a fixed or a variable number of times.

Case III: One or more of the elements in the keying sequence may be omitted from time to time irregularly.

Case IV: The keying sequence irregularly alternates in its direction of progression.³

Case V: The keying sequence irregularly alternates in its direction of progression, and, in addition, certain elements of the keying sequence may be repeated one or more times.

Case VI: The keying sequence irregularly alternates in its direction of progression, and, in addition, one or more of the elements in the keying sequence are omitted from time to time irregularly.

c. The foregoing methods may, for clarity, be represented graphically as follows. Suppose the key consists of a cyclic sequence of 10 elements represented symbolically by the series of numbers 1, 2, 3, . . . 0. Indicating an interruption by a vertical line,⁴ we show in Fig. 13, below, the relationship between

¹ See in this connection subpar. 84b(1) on p. 220 of *Military Cryptanalytics, Part I*.

² In addition to these cases, a "septimum quid" could be listed, as a catchall for "everything else," which includes progressions so irregular as to defy classification.

³ It is to be noted that this fourth case could be treated as though it were a special form (with irregularly occurring small or large skips) of the third case.

⁴ What *specifically* brings about the interruption is here not stated, nor for that matter does it concern us at the moment. Suffice it to say that, whatever the cause of the interruption, it is not a function of the plain or of the cipher text, but is in this case predetermined by an external convention with steps 3, 1, 6, 7, 5, 8, 2, 4, 9 . . .

the letter at each position of the message and the identity of the element of the keying sequence in the six general cases discussed above.

Letter No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Key elements, Case I:	1	2	3	1	1	2	3	4	5	6	1	2	3	4	5	6	7	1	2	3	4	5	1	2	3
Key elements, Case II:	1	2	3	3	3	4	5	6	7	8	8	9	0	1	2	3	4	4	5	6	7	8	8	9	0
Key elements, Case III:	1	2	3	5	7	8	9	0	1	2	4	5	6	7	8	9	0	2	3	4	5	6	8	9	0
Key elements, Case IV:	1	2	3	2	3	4	5	6	7	8	7	6	5	4	3	2	1	2	3	4	5	6	5	4	3
Key elements, Case V:	1	2	3	3	3	4	5	6	7	8	8	7	6	5	4	3	2	2	3	4	5	6	6	5	4
Key elements, Case VI:	1	2	3	5	7	8	9	0	1	2	4	3	2	1	0	9	8	0	1	2	3	4	6	5	4

Letter No.	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
Key elements, Case I:	4	5	6	7	8	1	2	1	2	3	4	1	2	3	4	5	6	7	8	9	1	2	3	4	5
Key elements, Case II:	1	2	3	4	5	5	6	6	7	8	9	9	0	1	2	3	4	5	6	7	7	8	9	0	1
Key elements, Case III:	1	2	3	4	5	7	8	0	1	2	3	5	6	7	8	9	0	1	2	3	5	6	7	8	9
Key elements, Case IV:	2	1	0	9	8	9	0	9	8	7	6	7	8	9	0	1	2	3	4	5	4	3	2	1	0
Key elements, Case V:	3	2	1	0	9	9	0	0	9	8	7	7	8	9	0	1	2	3	4	5	5	4	3	2	1
Key elements, Case VI:	3	2	1	0	9	1	2	4	3	2	1	3	4	5	6	7	8	9	0	1	3	2	1	0	9

FIGURE 13

Note that in Cases III and VI the amount of skip is here portrayed as being constant. This is not a necessary condition in these keying methods; the amount of skip could have consisted of irregular jumps as established by the keying convention employed.

d. If we *knew* just when the interruptions take place, and if we also knew the exact nature of the effect of each interruption,⁵ then the successive ciphertext sections of encrypted messages in the foregoing six cases could be properly superimposed so as to be in true cryptographic depth. In the diagrams below, the digits in the top line represent the ten keying elements, while the numbers 1-50 underneath this line represent the positional identities of the first 50 ciphertext letters.

	1	2	3	4	5	6	7	8	9	0
	(1	2	3)							
	(4)									
	(5	6	7	8	9	10)				
	(11	12	13	14	15	16	17)			
	(18	19	20	21	22)					
Case I	(23	24	25	26	27	28	29	30)		
	(31	32)								
	(33	34	35	36)						
	(37	38	39	40	41	42	43	44	45)	
	(46	47	48	50	...					

	1	2	3	4	5	6	7	8	9	0
	(1	2	3)							
			(4)							
			(5	6	7	8	9	10)		
	→14	15	16	17)				(11	12	13→
			(18	19	20	21	22)			
Case II	→26	27	28	29	30)			(23	24	25→
			(31	32)						
			(33	34	35	36)				
	→39	40	41	42	43	44	45)	(37	38→	
	→50	...					(46	47	48	49→

⁵ These two points could have been determined either through physical compromise of a cipher machine incorporating such principles, or through the *analytical* compromise of a cryptosystem.

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

Case III

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

Case IV

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

Case V

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

Case VI

Obviously, if we did not know when or how the interruptions take place, then the successive sections of keying elements cannot be superimposed as indicated above.

e. The interruption of the fundamental cyclic keying sequence usually takes place according to some prearranged plan or convention. The *identity* of the plaintext letters being enciphered might be involved in the determination of the interruption (as in plaintext interruptor systems);⁶ or the identity of the ciphertext letters might be a factor (as in ciphertext interruptor systems); or, finally, the interruption of the fundamental cyclic keying sequence might be predicated upon a separate convention,

⁶ In the Wheatstone cipher device the interruption of the keying sequence of the 26 cipher alphabets used in sequential progression is predicated upon the relative *position in the plain component* of a plaintext letter with respect to the *position* (in the plain component) occupied by the next plaintext letter to be encrypted. (See Chapter VIII in this connection.)

mechanism, or prearrangement, without regard to the plain text or the cipher text. Some basic methods of interruption will now be taken up, using a short mnemonic key as an example.

15. **Plaintext interruptor systems.**—a. Suppose the correspondents agree that the interruption in the key will take place after the occurrence of a specified letter in the plain text, after which the key begins anew at its initial position.⁷ Since there is nothing fixed about the time the interruption will occur—it will take place at no fixed intervals—not only does the interruption become quite irregular, following no pattern, but also the method never reverts to one having periodicity. Let us assume that the correspondents have agreed upon R_p as the interruptor letter, and that they are using the normal sequence for the plain component and the HYDRAULIC . . . XZ sequence for the cipher component. If the mnemonic key phrase is BUSINESS MACHINES, this key would be interrupted by the occurrences of R_p as in the following example:

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

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

Key: B U S I B U S B U S I N E B U S I N
 Plain: T H I R D A R T I L L E R Y S T O P . . .
 Cipher: D G D X G U F D J U P S Y I W J T U

The final cipher text, in groups of five letters, would be the following:

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

It will be noted that the two long polygraphic repetitions are at intervals of 44 and 34, respectively, which intervals have nothing in common with 16, the length of the basic, uninterrupted period.

b. Instead of employing an ordinary plaintext letter as the interruptor letter, one might use a 25-letter plain component, combining I with J, and then use the 26th character (J) as a null plaintext letter which is inserted at random by the encipherer to serve as the interruptor letter. Note the following example:

Key: B U S I N E S S M A B U S I N E S S M A C H I N E S B U S B U S I N
 Plain: P R O C E E D T O J R O A D I U N C T I O N S I X T W O J F I V E J . . .

c. It is obvious that repetitions would be plentiful in cryptograms of this construction, regardless of whether a letter of high-, medium-, or low-frequency is selected as the signal for key interruption. If a letter of high frequency is chosen, repetitions will occur quite often, not only because that letter will certainly be a part of many common words, but also because it will be followed by words that are frequently repeated; and since the key starts again with each such interruption, these frequently repeated words will be enciphered by the same sequence of cipher alphabets. This is the case in the first of the two foregoing examples. It is clear, for instance, that every time the word ARTILLERY appears in the cryptogram the cipher equivalents of TILLERY must be the same. If the interruptor letters were A_p, instead of R_p, the repetition would include the cipher equivalents of RTILLERY; if it were T_p, ILLERY, and so on. On the other hand, if a letter of very low frequency were selected as the interruptor letter, then the encipherment would tend to approximate that of normal periodic substitution, and repetitions would be plentiful on that basis alone. Of course, the intervals between the repetitions in any of the

⁷ This is Case I of subpar. 14b.

foregoing cases (except perhaps that in which the plaintext interruptor is a letter of very low frequency) would be markedly irregular, so that periodicity would not be manifested.

16. **Ciphertext interruptor systems.**—*a.* In the systems of the preceding paragraph, a plaintext letter serves as the interruptor letter. But now suppose the correspondents agree that the interruption in the key will take place immediately after a previously agreed-upon letter, say Q_c , occurs in the cipher text. The key would then be interrupted as shown in the following example:

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

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

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

The cipher text in 5-letter groups is as follows:

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

b. In the foregoing example, there are no significant repetitions; such as do occur comprise only digraphs, several of which are purely accidental. But the absence of significant, long repetitions is itself purely accidental, for had the interruptor letter been a letter other than Q_c , then the phrase **AMMUNITION FOR** (which occurs twice) might have been enciphered identically both times. If a short key is employed, repetitions may be plentiful. For example, note the following, in which S_c is the interruptor letter: ⁸

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

c. This last example gives a clue to one method of attacking this type of system. There will be repetitions within short sections, and the interval between them will sometimes permit ascertaining the length of the basic key. In such short sections, the letters which intervene between the repeated sequences may be eliminated as possible interruptor letters. Thus, in the foregoing example, we can deduce that the length of the basic key is 5 letters, and that the cipher letters A, C, B, and N may be eliminated as interruptor letters. By extension of this principle to the letters intervening between other repetitions, one may more-or-less quickly ascertain what ciphertext letter serves as the interruptor. ⁹

d. The ciphertext interruptor might be a letter which is not otherwise used in the cryptographic scheme; for example, the plain component might be a 25-letter sequence (combining I and J) and the cipher component a 25-letter sequence excluding, let us say, Z. This letter Z may then be inserted in appropriate places in the cipher text to signal the interruptions in the keying cycle. In some cases such a special interruptor letter may be used in *addition* to a ciphertext interruptor which arises from the bona fide encryption of a plaintext letter, as a means of insuring that interruption of the keying cycle will take place frequently enough to suit the cryptographer or his procedures-prescribing superiors.

⁸ Note that the periodic repetitive phenomena manifested would also have arisen in a *plaintext* interruptor system, if the interruptor had been, let us say, A_p —or, for that matter, *any* other letter not present in the fragment FOURFI VETOFOURFI.

⁹ The method described in this subparagraph may also be applied in the case of plaintext interruptor systems, with certain modifications.

(For that matter, there is nothing to bar the use of two or more letters as interruptors in the usual manner, in either plaintext interruptor or ciphertext interruptor systems.)

17. **Systems employing externally generated or determined keys.**—*a.* In subpars. 3*b* and *f* we have seen two examples of keying procedures which do not depend upon conventions affiliated with identities of plaintext or ciphertext letters, but which are established by an independent, external keying convention. The keying methods of subpar. 3*b*, if modified to incorporate variable-length *polyalphabetic* keying units (as contrasted with the variable-length monoalphabetic keying units illustrated in that example), could take on a form such as the following:

S	IG	GNA	NALS	ALSIG	L	SI	IGN	GNAL	NALSI	A	LS	SIG	IGNA	GNALS
1	12	123	1234	12345	1	12	123	1234	12345	1	12	123	1234	12345
N	AL	LSI	SIGN	IGNAL	G	NA	ALS	LSIG	SIGNA	I	GN	NAL	ALSI	LSIGN
1	12	123	1234	12345	1	12	123	1234	12345	1	12	123	1234	12345
S	IG	GNA	NALS	ALSIG	L	SI	IGN	etc.						
1	12	123	1234	12345	1	12	123							

Similarly, the keying method of subpar. 3*f*, modified to embrace the aspect of variable-length polyalphabetic keying units, could be transformed into one of the following, among other possibilities:

- (1) D E F/E/C D E F/L M N O/A B/R S T/A B/T/I J/O P Q/N O/O P Q/F G H I/. . .
- (2) D E C/E/C L A R/L A R A/A R/R A T/A T/T/I O/O N O/N O/O F I/F I N D/. . .

b. The foregoing methods have as their purpose the establishment of keys of fair length from a short mnemonic key. There are other simple methods for accomplishing this, as illustrated in the examples which follow. Let us consider the mnemonic key **HYDRAULIC**, and derive from it a numerical key:

H Y D R A U L I C
4 9 3 7 1 8 6 5 2

We may now take the key letters in numerical-key order, and in groupings as determined by the numerical key, so that the original key of only 9 letters is expanded to one of 45 letters. Thus:

- (1) A/C H/D R A/H Y D R/I C H Y D/L I C H Y D/R A U L I C H/
U L I C H Y D R/Y D R A U L I C H/

Two other methods of deriving 45-element key sequences from the basic 9-letter key word are shown below:

- (2) H Y D R/Y D R A U L I C H/D R A/R A U L I C H/A/
U L I C H Y D R/L I C H Y D/I C H Y D/C H/
- (3) H Y D R A/H Y D R A U L I C/H Y D/H/H Y D R A U L I/
H Y D R A U L/H Y D R /H Y D R A U/H Y/

Method (2) is essentially the same as (1), except that the key fragments are taken in the order in which they appear in the key word. Method (3) involves taking the successive sections of the numerical key, these sections terminating with the successive numbers 1, 2, 3, . . . of the numerical key.¹⁰

c. Many other methods exist for the establishment of keys consisting of variable-length keying units. Furthermore, some of these methods merge into the domain of methods of lengthening or extending keys in general, apart from any considerations of variable-length keying units. Several of the most important of these methods will be discussed in subsequent chapters of this text.

18. **Solution when known cipher alphabets are employed.**—*a.* (1) Let us suppose that a particular cryptosystem has been in use for some time, and that the general nature of the system and the cipher

¹⁰ This method is equivalent to an interrupted-key columnar transposition system. See in this connection subpar. 51*h* on p. 89 of *Military Cryptanalytics, Part I*.

alphabets have become known, either through successful cryptanalysis or through light-fingered techniques coming under the formal term of "physical compromise," which includes among its manifold tachydactylurgic aspects that which has been referred to colloquially as "wastebasket cryptanalysis."¹¹ Only the specific key to messages remains unknown. The cipher text is examined for repetitions, and an attack is made on the basis of searching for a probable word. Thus, taking the cryptogram in subpar. 15a as an example (quoted here below for convenience), suppose the presence of the word **ARTILLERY** is suspected.

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

Attempts are made to locate this word, basing the search upon the recognition of an intelligible key; we will assume in this case that the cipher component is the **HYDRAULIC . . . XZ** sequence sliding against the normal sequence for the plain component.

(2) Beginning with the very first letter of the message, we juxtapose the word **ARTILLERY** against the cipher text and ascertain the key letters. Thus:

Key: B H J Q P I B F U
Cipher: B O L Y R P J D R
Plain: A R T I L L E R Y

Since this "key" is certainly not intelligible text, the assumed word is moved one letter to the right and the test repeated, and so on until the 19th position in the text is reached.¹²

Key: S I B U S I N E B
Cipher: S X D J U P S Y I
Plain: A R T I L L E R Y

(3) The sequence **BUSINE** suggests **BUSINESS**; moreover, it is noted that the key appears to be interrupted both times by the letter **R**. The key may now be applied to the beginning of the message, to see whether the whole key or only a portion of it has been recovered. Thus:

Key: B U S I N E S S B U S I N E S
Cipher: B O L Y R P J D R O J K X K J . . .
Plain: A M M U N I T I U M T H I E T

(4) It is obvious that **BUSINESS** is only a part of the key. But the first word of the message is plainly **AMMUNITION**. When this is tried, the key is extended to **BUSINESS MA . . .** This key crib is now slid through the rest of the cipher text and the remainder of the message is quickly deciphered and the entire key recovered.

¹¹ This is really not stealing. For the pure in heart, this should be thought of as conversion of raw data, and that the parties so generously supplying these raw data are, unknowingly, cooperating in government work.

¹² In actual practice, the search for the placement of the probable word would have been accomplished by means of the following diagram (see in this connection subpar. 22d on pp. 41-42 of *Military Cryptanalytics, Part II*):

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

~~SECRET~~

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024 (i)
(b) (3)-P.L. 86-36

~~SECRET~~

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024(i)
(b) (3)-P.L. 86-36

d. (1) Another technique, if we know or can assume the method of key interruption (e.g., a skip over one element of the key after the occurrence of a previously designated ciphertext letter, in this case W_e), involves writing out the modified cipher text of a single message on trial widths in order to see if any cyclic properties are present in the basic, uninterrupted key. We can then determine statistically when the correct cyclic write-out is reached by the application of a technique discussed in the preceding text.¹⁶ As an example, let us assume the following message is at hand:

```

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

(2) If we know the method of interruption and also the identity of the ciphertext interruptor,¹⁷ we would write out the appropriately modified cipher text on various widths, testing each hypothesis in turn, until a satisfactory I.C. is reached for an entire columnar array. For example, if we know that the enemy is using key words and phrases from 11 to 40 letters in length as the basic key sequence, we would begin by writing out the modified cipher text (on the assumption of W_e as the interruptor letter) on a width of 11 as shown below, together with the appropriate ϕ values for the computations:

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

$\Sigma \phi = 64$

(b) (1)
 (b) (3)-18 USC 798
 (b) (3)-50 USC 3024(i)
 (b) (3)-P.L. 86-36

(3) The row of numbers immediately beneath the write-out represents the ϕ values of the columns; the row beneath the " ϕ " row, labelled "N", represents the number of letters in the columns. Since there

¹⁶ Cf. Subpar. 18e on pp. 28-31 of *Military Cryptanalytics, Part II*.

¹⁷ If worst came to worst, we could test each of the 26 letters in turn as the ciphertext interruptor. This testing, coupled with the writing out of the cipher text on the various widths, would be quite laborious and time-consuming by manual methods; data processing machine techniques would here be very useful.

are 3 columns of 13 letters each, 5 columns of 12 letters each, 2 columns of 11 letters, and 1 column of 9 letters, the expected value of ϕ random (ϕ_r) is given by the formula

$$\phi_r = \frac{3(13 \cdot 12) + 5(12 \cdot 11) + 2(11 \cdot 10) + 1(9 \cdot 8)}{26} = \frac{1420}{26} = 54.6$$

The δ I.C. is defined as the ratio of the observed value of ϕ to the expected value of ϕ random, or $\frac{\phi_o}{\phi_r}$. Now since the ϕ_o (which is the sum of all the ϕ values for the columns) is 64, our I.C. formula becomes, by simple algebraic transformation,

$$\delta I.C. = \frac{26.64}{3(13.12) + 5(12.11) + 2(11.10) + 1(9.8)} = \frac{1664}{1420} = 1.17$$

(4) This I.C. of 1.17 is not satisfactory, so we continue testing successively greater widths, until the width of 32 is reached:

[illegible]

At this width, the I.C. calculation becomes

$$\delta I.C. = \frac{26 \cdot 30}{7(5 \cdot 4) + 21(4 \cdot 3) + 3(3 \cdot 2) + 1(2 \cdot 1)} = \frac{780}{412} = 1.89,$$

giving statistical credence to the assumption of 32 as the correct width, since we were looking for an I.C. in the vicinity of 1.73 for the correct case.¹⁸ Knowing the components involved, we may complete the plain-component sequence on the letters of the columns to effect a speedy solution.

19. Solution when unknown cipher alphabets are employed.—*a.* In the first text in this series, it was pointed out that “in the final analysis, the solution of every cryptogram involving a form of substitution depends upon its reduction to monoalphabetic terms, if it is not originally in those terms.”¹⁹ In the preceding volume, it was observed that when in the course of solution of an ordinary repeating-key cipher the text is written out in period-lengths, “another way of looking at the matter is to conceive of the text as having thus been transcribed into *superimposed periods*; in such a case the letters in each

¹⁸ The reader might be interested in the I.C.'s for all the widths from 11 to 40; these are shown in the following table:

w	δ	w	δ	w	δ	w	δ	w	δ	w	δ
11	1. 17	16	1. 11	21	0. 91	26	1. 06	31	1. 08	36	1. 33
12	1. 09	17	0. 95	22	1. 35	27	1. 24	32	1. 89	37	1. 69
13	1. 06	18	1. 03	23	1. 18	28	0. 98	33	0. 79	38	1. 28
14	0. 91	19	1. 02	24	1. 17	29	1. 67	34	0. 68	39	0. 98
15	1. 29	20	0. 86	25	1. 20	30	1. 29	35	1. 14	40	1. 01

Note the I.C. of 1.67 for the width of 29, and the I.C. of 1.69 for the width of 37. These I.C.'s are certainly satisfactory; however, the widths from which they were derived are incorrect, so that they represent the vagaries of the touch not of a Mephistophelian finger, but rather that of a Bernoullian digit when an insufficient number of trials is involved.

¹⁹ *Military Cryptanalytics, Part I*, subpar. 17b.

column have undergone the same kind of treatment by the same elements (plain and cipher components of the cipher alphabet).''²⁰ It follows that, even if the repeating key is very long, if there are many short cryptograms all enciphered by exactly the same key and each message begins at the same point in the key, the distributions applicable to the successive columns of text can be solved.²¹ Even in aperiodic systems, if there is available a number of messages starting out in the same key which then diverges in the course of encipherment according to the nature of the cryptosystem, this *solution by superimposition* may be applicable in particular cases, so long as the key divergence is not too radical for cryptanalytic comfort.

b. Let us study the following beginnings of 30 messages, passed between correspondents known to have used various types of aperiodic keying:

1. Y F W F M R I Q M X X E L M J . . .	16. G O E Q B Q O T L E S A C R B . . .
2. H W W T T E C T D O Z F D O V . . .	17. W T S R G X M Z T V S J Q L X . . .
3. T P Y F K S O V W I H F N C J . . .	18. W T E V F C I B T S P R C A T . . .
4. Y P E P S N L S K Z N V T J B . . .	19. Z C V Y M B V N Y W Q U Z G U . . .
5. E A Q U Z D V E S K C I U P A . . .	20. Z C T T Z W C T T I K H Q U T . . .
6. Z C G M W T N B I M K U S N L . . .	21. Z C C T S N E S K O U B M P T . . .
7. E P D O Z F D O V B I L V L W . . .	22. A F E S J O N K T D V E S K C . . .
8. E P T L E S A C R B M P T P J . . .	23. Z C F F D T N P F D H D T P F . . .
9. W M L S O T O Z E E J Z G V K . . .	24. V Z I E X R X R F F U N T Q E . . .
10. Z C F F D C F R J W H L P D T . . .	25. E P S N L S K L O H W P T R G . . .
11. H C Q E D T P Y I L N R E M V . . .	26. Y T S V W L S T L E S A C R B . . .
12. C L C T Z I K S O E O Z C T T . . .	27. W A Z X Z Q A C H Q U T L S T . . .
13. H C Q E F D K I F Q W O C L M . . .	28. N O F T Z N L H Q U T J H Z A . . .
14. E P T W K S U Z N V V A U C S . . .	29. V R C W K M O L N X W S D O L . . .
15. Z A B M Z H G O F X Q I G M M . . .	30. S P R C P F X E O J C Q F W M . . .

²⁰ *Military Cryptanalytics, Part II*, subpar. 65a.

²¹ Cf. subpar. 65b, *Military Cryptanalytics, Part II*.

The presence of digraphic and polygraphic repetitions in the initial columns could mean that the messages start out in flush depth, and the presence of offset repetitions could be an indication of shifts in the keying sequence. Frequency distributions for the columns are made and are shown in Fig. 16, below, accompanied by their I.C.'s:

1.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	2.63
2.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	3.41
3.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.31
4.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.37
5.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.85
6.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.20
7.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.08
8.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.08
9.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.13
10.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	0.95
11.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.02
12.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	0.72
13.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.43
14.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.02
15.	$\bar{A} \bar{B} \bar{C} \bar{D} \bar{E} \bar{F} \bar{G} \bar{H} \bar{I} \bar{J} \bar{K} \bar{L} \bar{M} \bar{N} \bar{O} \bar{P} \bar{Q} \bar{R} \bar{S} \bar{T} \bar{U} \bar{V} \bar{W} \bar{X} \bar{Y} \bar{Z}$	1.61

FIGURE 16

c. The first two columns are certainly monoalphabetic; after that, there is a rapid falling off in monoalphabeticity, with the exceptions of cols. 5, 13 and 15 which could be due to chance. We note the digraph ZC , which occurs 6 times in cols. 1 and 2; this could well be the equivalent of RE_p , and HC_e in cols. 1 and 2 could stand for SE_p . On this basis, $ZCFFD$ in Messages 10 and 23 could represent $REFER$, and $HCQE$ in Messages 11 and 13 could be $SEND$. We then note

				5			10			15						
23.	Z	C	F	F	D	<u>T</u>	N	P	F	D	H	D	T	P	F	. . .
6.	Z	C	G	M	W	<u>T</u>	N	B	I	M	K	U	S	N	L	. . .

which could represent REFERENCE and REGIMENT. We now turn our attention to the following four message beginnings:

28. N O F ⁵T Z N L H Q U T J H Z A . . .
 F R

20. Z C T ¹⁰T Z W C T T I K H Q U T . . .
 R E R

21. Z C C T S N E S K O U B M P T . . .
 R E

12. C L C ¹⁵T Z I K S O E O Z C T T . . .
 R

If we assume that T_e in col. 4 represents O_p, then in No. 28—FOR . . . becomes INFORM(ATION), in No. 20 RE-OR— becomes REPORT, in No. 21 RE-O . . . becomes RECOMMEND, and in No. 12—COR—N— becomes ACCORDING.

d. The plain-cipher equivalencies from the foregoing assumptions are entered into a sequence reconstruction matrix, as shown below:

	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	C										N										Z	H				
2		L		C											O											
3		C			F	G									Q		T									
4			(E)	F						(M)																
5														w _s						D _Z						
6				I	T									N									W			
7	L				E					K					N											
8		P													S								B _H			
9			(K)	F		O				(Q)																
10																U										
11															T											

Conflicts are noted in lines 5 and 8, and between lines 4 and 9; however, possibility of direct symmetry is noted in the top four lines, which indicates that the recoveries in these lines could well be homogeneous, not having been affected by vagaries of the keying. Transferring values among these four lines, we will develop the reconstruction matrix into the following, in which the cipher components are slides of what is patently a keyword-mixed sequence (derived values in lower case):

P:	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	C		e	f	g			m	N	o		q		t				Z	H							l
2			L		C		e	f	g			m	n	O		q		t					z	h		
3	l		C		e	F	G			m	n	o		Q		T						z	h			
4		c		E	F	g				M	n	o		q		T						z	h			l

Our work sheet will now look as illustrated in Fig. 17. which includes the values in the first four columns obtained by direct symmetry shown in lower case:

1. Y F W F M R I Q M X X E L M J . . . h e	16. G O E Q B Q V O C F X E S N H . . . e N e m
2. H W W T T E C T D O Z F D O V . . . S O	17. W T S R G X M Z T V S J Q L X . . . r
3. T P Y F K S O V W I H F N C J . . . n E	18. W T E V F C I B T S P R C A T . . . r e T
4. Y P E P S N L S K Z N V T J B . . . e M M A N D	19. Z C V Y M B V N Y W Q U Z G U . . . R E
5. E A Q U Z D V E S K C I U P A . . . c N R	20. Z C T T Z W C T T I K H Q U T . . . R E P O R T
6. Z C G M W T N B I M K U S N L . . . R E G I M E N T	21. Z C C T S N E S K O U B M P T . . . R E C O M M E N D
7. E P D O Z F D O V B I L V L W . . . c k R	22. A F E S J O N K T D V E S K C . . . h e N
8. E P T L E S A C R B M P T P F . . . c P z	23. Z C F F D T N P F D H D T P F . . . R E F E R E N C E
9. W M L S O T O Z E E J Z G V K . . . l a E	24. V Z I E X R X R F F U N T Q E . . . v D E
10. Z C F F D C F R J W H L P H X . . . R E F E R	25. E P S N L S K L O H W P T R G . . . c j I G
11. H C Q E D T P Y I L N R E M V . . . S E N D R E	26. Y T S V W L S T L E S A C R B . . . r M
12. C L C T Z I K S O E O Z C T T . . . A C C O R D I N G	27. W A Z X Z Q A C H Q U T L S T . . . t R
13. H C Q E F D K I F Q W O C L M . . . S E N D I E	28. N O F T Z N L H Q U T J H Z A . . . I N F O R M A T I O N
14. E P T W K S U Z N V V A U C S . . . c P	29. V R C W K M O L N X W S D O L . . . C
15. Z A B M Z H G O F X Q I G M M . . . R I R E	30. S P R C P F X E O J C Q F W M . . . b G

FIGURE 17

e. At this point more plain text could be assumed in the messages from the fragments already present; the cipher component would be recovered in its entirety, the basic key determined, and the cause of the key interruptions (as manifested by the apparent garbles) ascertained. Or, as another

approach, we might take an introspective look at the first 5 letters of matched plain and cipher of the following three message beginnings:

```

10. Z C F F D . . .
    R E F E R
20. Z C T T Z . . .
    R E P O R.
28. N O F T Z . . .
    I N F O R

```

We note the $D_c=R_p$ and $Z_c=R_p$ in position 5 of Messages 10 and 20, and observe that, if the system employs a ciphertext interruptor, it may be either F_c or T_c ; but if $TZ_c=OR_p$ in Messages 20 and 28 is causal, F_c and T_c are eliminated and therefore there is *no* ciphertext interruptor in the cryptosystem. We then note the common $F_c=F_p$ between Messages 10 and 28 and the fact that in position 5 of Message 28 $Z_c=R_p$, and we may conclude that, if a plaintext interruptor is present, it must be O_p . We find this to be true, and when we finish the solution of the problem we find the cipher component to be our perennial friend, the HYDRAULIC . . . XZ sequence, and the basic key to be CALIFORNIAGOLDR(USH).

20. Additional remarks.—*a.* We have seen in the preceding paragraph a demonstration of solution of only one irregularly keyed system involving unknown cipher alphabets. The solution involved a set of very fortunate circumstances indeed, all of which were happily present awaiting rapid exploitation by the cheerful cryptanalyst. Modern cryptanalysis is quite often contingent upon miracles—minor miracles for minor systems, and healthy miracles for some of the complex systems encountered in present-day operations. When we come right down to it, all cryptanalysis is astonishing; it certainly is so to a layman, and it is so even to an expert—if he pauses long enough from his breaking of one system after another to marvel at the phenomenal luck he has had, shuddering at the thought of what would have happened *if* (i.e., *if* the enemy had done this instead of that, if he had used this instead of that, and *if* . . .). On the other hand, all cryptanalysis is quite commonplace:²² after all, messages have been encrypted with certain invariant mathematico-philosophical-procedural elements, and all the cryptanalyst does is to discover and exploit these elements. And, in retrospect, after a problem has been solved, we often shrug our shoulders and say “Well, how else would one have done it?” Many systems of the types treated in this volume could be virtually unsolvable, or might appear to be so, if only a small amount of traffic is available for study, and if little or nothing is known about the nature of the cryptosystem. However, as happens time and again in actual operations, Fortuna smiles and the incredible is shorn of its prefix.

b. Operationally, cryptodilemmas are resolved by the exploitation of contingencies which are by now well-known to the reader: (1) messages in the same or nearly the same keys; (2) depths and partial depths; (3) polygraphic repetitions; (4) cribs; (5) various kinds of cryptographic errors; (6) isologs; (7) matched plain and cipher; (8) isomorphs; (9) indicators. Each problem presents a very special case, and therefore demands its own special requirements for solution.

c. Most of the types of aperiodic substitution discussed in this chapter are rather unsuitable for practical military usage. Encipherment is slow and subject to error. In some cases encipherment can be accomplished only by a single-letter operation. For, in interruptor systems, if the interruptor is a cipher letter the key is interrupted by a letter which cannot be known in advance; if the interruptor is a plaintext letter, while the interruptions can be indicated before encipherment is begun, the irregularities occasioned by the interruptions in keying cause confusion and quite materially retard the enciphering process. In deciphering, the rate of speed would be just as slow in either method. It is obvious that one of the principal disadvantages in all these methods is that if an error in transmission is made, if some letters are omitted, or if anything happens to the interruptor letter, the message becomes difficult or impossible to decipher by the ordinary cipher clerk. In spite of all these objections, plus the fact that the degree of cryptosecurity attainable by most of these methods is not sufficient for military purposes, these systems have been and are still occasionally encountered—which is what makes the cryptologic world go round.

²² It must have been a deep thinker who first uttered the statement that “all problems in cryptanalysis, like mathematics, are either trivial or impossible.”

(b) (1)
(b) (3) -18 USC 798
(b) (3) -50 USC 3024 (i)
(b) (3) -P.L. 86-36

CHAPTER IV

CIPHERTEXT AUTOKEY SYSTEMS

	Paragraph
The cryptography of autokey encipherment.....	21
Solution of ciphertext autokeyed cryptograms when known cipher alphabets are employed.....	22
Principles of solution by frequency analysis.....	23
Example of solution by frequency analysis.....	24
Solution by means of isomorphs.....	25
Solution of isologs involving the same pair of unknown primary components.....	26
	27
Further remarks on ciphertext autokey systems.....	28

21. The cryptography of autokey encipherment.—*a.* The mechanics of autokey encipherment were treated briefly in the preceding volume.¹ In autokey systems there are two possible sources for successive key letters: the cipher text or the plain text of the message itself. In either case, the *initial* key letter or key letters are supplied by prearrangement between the correspondents, or are designated by means of an indicator; after that, the text letters that are to serve as the key are displaced 1, 2, 3 . . . intervals to the right, depending upon the length of the prearranged key.

b. An example of ciphertext autokey encipherment is shown below, wherein the cipher alphabets are direct standard alphabets, and the single letter X is the prearranged initial key:

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

FIGURE 18a

Instead of having a single letter serve as the initial key, a word or even a long phrase may be used, as in the example below wherein the word **FORTUNE** is used as the initial key:

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

FIGURE 18b

Sometimes only the last cipher letter resulting from the use of the prearranged key word is used as the key letter for enciphering the autokeyed portion of the text. Thus, in the preceding example, the plain text beginning **GIMENT** . . . would be enciphered differently as follows:

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

FIGURE 18c

c. In plaintext autokey encipherment the procedure is quite similar, as is shown in the following example wherein the prearranged initial key is the letter X:

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

FIGURE 19a

¹ Cf. subpar. 99e on pp. 310-311 of *Military Cryptanalytics, Part II*.

If the word FORTUNE were used as the initial key, the plain text would be enciphered as follows:

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

FIGURE 19b

d. In the foregoing examples, direct standard alphabets were used; however, mixed alphabets, either interrelated or independent,² may be used just as readily. Furthermore, instead of the ordinary type of cipher alphabets, the cryptographic process may employ a mathematical process of addition, but the difference between the latter process and the ordinary one using sliding alphabets is more apparent than real. For example, let us consider the following numerical sequence for the 26 letters

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
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 0

and let the plaintext message be the same as before. Let us assume that the cryptographic rules prescribe that the first plaintext letter will be self-enciphered,³ and that each cipher letter from that point on is produced in turn by finding the sum (mod 26) of the numerical equivalents of the preceding cipher letter and the plaintext letter to be enciphered; in other words, a type of numerical ciphertext autokey system. This is shown in the diagram below, wherein P' denotes the numerical equivalents of the plain text, C' the sum of the key and the numerical (i.e., intermediate) plain text, and C the conversion into letters of the intermediate cipher text C'.

K : 0 22 23 5 9 12 16 1 14 22 12 23 14 10 19 11 1 17 22 13 16 9 1 22 18 8 0 23 5 22
P : T H I R D R E G I M E N T C O M M A N D P O S T M O V I N G . . .
P' : 22 1 8 4 3 4 11 13 8 16 11 17 22 9 18 16 16 5 17 3 19 18 21 22 16 18 23 8 17 13
C' : 22 23 5 9 12 16 1 14 22 12 23 14 10 19 11 1 17 22 13 16 9 1 22 18 8 0 23 5 22 9
C : T V A C F M H J T F V J B P E H N T G M C H T O I Z V A T C

FIGURE 20a

e. That the difference between the types of encipherment in the preceding subparagraph and the ordinary method of ciphertext autokey encipherment is illusory is demonstrated by the example in Fig. 20b, below:

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

FIGURE 20b

In this example, the plain and cipher components are keyword-mixed sequences based upon HYDRAULIC, and Z_p is the index letter against which the key letters in the cipher component are set; the cryptographic results are identical to those obtained in Fig. 20a, above.

f. Since the analysis of ciphertext autokey systems is usually easier than the analysis of plaintext autokey systems, the former will be the first to be discussed.

22. Solution of ciphertext autokeyed cryptograms when known cipher alphabets are employed.—a. First of all, it is to be noted that if the cryptanalyst knows the cipher alphabets which were employed in encipherment, the solution hardly presents any problem. It is only necessary to decipher the message beyond the key-letter or key-word portion and the initial part of the plain text enciphered by this key letter or key word can be filled in from the context.

² For instance, an autokey system might incorporate independent, random alphabets such as those illustrated in Fig. 33 on p. 70.

³ This, on a numerical scale, is tantamount to the effect of a key of 0.

(1) For example, let us consider the following beginning of an intercepted message:

Q X F W Z Q U A I U Y L E G U G S S F I . . .

On the assumption of ciphertext autokey involving direct standard alphabets, if we write the cipher text as key letters, displaced one interval to the right, we obtain the following decipherment:

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

The introductory key letter required to make $Q_c = T_p$ is found to be X_k .

(2) As a second example, let us consider the following beginning of a cryptogram suspected to have been enciphered by ciphertext autokey with direct standard alphabets:

B P A U V N L F J A L Y M L Q N A E L R . . .

Assuming an introductory key of one letter, we obtain the following decipherment:

K: B P A U V N L F J A L Y M L Q N A E L
C: B P A U V N L F J A L Y M L Q N A E L R . . .
P: O L U B S Y U E R L N O Z F X N E H G

Nothing. We now assume that the introductory key consisted of two letters, and we get the following results:

K: B P A U V N L F J A L Y M L Q N A E
C: B P A U V N L F J A L Y M L Q N A E L R . . .
P: Z J V T Q S Y V C Y B N E C K R L N

Still nothing. We make several more trials, and finally, on the assumption of an introductory key of 8 letters, the following is obtained.

K: B P A U V N L F J A L Y
C: B P A U V N L F J A L Y M L Q N A E L R . . .
P: I L L E R Y F I R E A T

It is clear that the introductory key is 8 letters in length. Doing what comes naturally,

K: N U M B P A U V N L F J A L Y
C: B P A U V N L F J A L Y M L Q N A E L R . . .
P: A R T I L L E R Y F I R E A T

shows that the introductory key ends with NUM; now with but little experimentation, either with the letters of the key or with the beginning of the message plain text, we obtain the complete solution:

K: A L U M I N U M B P A U V N L F J A L Y
C: B P A U V N L F J A L Y M L Q N A E L R . . .
P: B E G I N A R T I L L E R Y F I R E A T

(3) In a third case, we will assume that the following is the beginning of an intercepted message:

D I T G C M G T Z B P C V D Q K Y S K P . . .

Again assuming direct standard alphabets, writing the cipher text as key letters displaced one interval to the right and deciphering, we obtain the following:

K: D I T G C M G T Z B P C V D Q K Y S K
 C: D I T G C M G T Z B P C V D Q K Y S K P . . .
 P: F L N W K U N G C O N T I N U O U S F

We note the plain text "CONTINUOUS F . . ." emerging, preceded by the NG which is probably ING; this indicates that ciphertext autokey is involved with an initial key of 7 letters, and that the last letter of the initial key is used to start the autokeyed portion. After a little experimentation with the initial portion of the message text and the key,⁴ we recover the key word and the first word of the message, as follows:⁵

K: M E R C U R Y | G T Z B P C V D Q K Y S K
 C: D I T G C M G T Z B P C V D Q K Y S K P . . .
 P: R E C E I V I N G C O N T I N U O U S F

b. A mechanical method of solution for ciphertext autokeyed cryptograms when the components are known sequences may be of interest. The method involves the use of sliding alphabet strips aligned in such a manner that, as one progresses from left to right across the strips, each key letter is set opposite the letter k on the preceding strip:⁶ the plain text will appear to the left of the pertinent cipher letter on each strip. In other words, what we have is a mechanical method of correlating the letters of the key, cipher, and plain text; the method is best understood by examples.

(1) In Fig. 21 is illustrated the arrangement of standard-alphabet strips for the first 10 letters of putative key, QXFWZQUAIU, for the message beginning given in subpar. a(1), above. If we assume that a one-letter introductory key has been used, the key letters just named were used to key the 2d through 11th cipher letters, XFWZQUAIUY; therefore we search for these cipher letters consecutively across the strips and we note the letters to their immediate left. In this case the plain text HIRDREGIME is manifested and the problem is solved.

⁴ The I of ING gives a key of Y, which should be preceded by a T, A, L, or very few other letters.

⁵ The key word and the first word of the message may be recovered by working backwards from ING_p, or by assuming various initial digraphs for the plain text or the key; a trial of RE_p for the message beginning would yield ME_x, and we could go on from here to read this "depth of one."

⁶ This is under the assumption that A_p is the index letter in the cryptographic equations.

(K): (Q X F W Z Q U A I U) (B P A U V N L F J A) (B P A U V N L F J A)

(C): (X F W Z Q U A I U Y) (P A U V N L F J A L) (J A L Y M L Q N A E)

<u>A</u> QNSONDXXFZ	<u>A</u> BQQKFSDIRRR	<u>A</u> BQQKFSDIRRR
BROTPOEYYGA	BCRRRLGTEJSS	BCRRRLGTEJSS
CSPUQPFFZZHB	CDSSSMHUFKTT	CDSSSMHUFKTT
DTQVR <u>Q</u> GA <u>A</u> IC	DETTNIVGLUU	DETTNIVGLUU
EURW <u>S</u> R <u>H</u> B <u>B</u> JD	EFUUOJWHMVV	EFUUOJWHMVV
FVSXTSICCKE	FGVVVPKXINWW	FGVVVPKXINWW
GWTYUTJDDLF	GHWWQLYJOXX	GHWWQLYJOXX
H <u>X</u> UZVUKEEMG	HIXXRMZKPY	HIXXRMZKPY
IYV <u>A</u> WVLFNNH	IJYY <u>S</u> N <u>A</u> LQZZ	IJYY <u>S</u> N <u>A</u> LQZZ
JZWBXWMMGOI	JKZZTOBMRA <u>A</u>	JKZZTOBMRA <u>A</u>
K <u>A</u> XCYXNHHPJ	KL <u>A</u> AUPCNS <u>B</u> B	KL <u>A</u> AUPCNS <u>B</u> B
LB <u>Y</u> DZYOI <u>I</u> QK	LM <u>B</u> B <u>V</u> QDOTCC	LM <u>B</u> B <u>V</u> QDOTCC
MCZEA <u>A</u> ZPJRL	MNCCWREPUDD	MNCCWREPUDD
ND <u>A</u> FBA <u>Q</u> KKSM	NODDXSFQVEE	NODDXSFQVEE
OEBGCBRLLTN	OPEEYTGRRWF	OPEEYTGRRWF
PFCHDCSMMUO	PQFFZUHSXGG	PQFFZUHSXGG
QGDIEDTNNVP	QRGG <u>A</u> VITYHH	QRGG <u>A</u> VITYHH
RHEJFEUOOWQ	RSHHBWJUZZII	RSHHBWJUZZII
S <u>I</u> FKGFVPPXR	STII <u>C</u> XKV <u>A</u> JJ	STII <u>C</u> XKV <u>A</u> JJ
TJGLHGWQQYS	TUJJ <u>D</u> YLWBKK	TUJJ <u>D</u> YLWBKK
UKHMIHXRRZT	UVKKEZMXCLL	UVKKEZMXCLL
VLINJIIYSS <u>A</u> U	VWLLF <u>A</u> NYDMM	VWLLF <u>A</u> NYDMM
WMJOKJZTTBV	WXMMSGBOZEN	WXMMSGBOZEN
XNKPLKA <u>U</u> UCW	XYNNHC <u>P</u> AFOO	XYNNHC <u>P</u> AFOO
YOLQMLBVVDX	YZOOIDQBGP	YZOOIDQBGP
ZPMRNM CWWEY	Z <u>A</u> PPJERCHQQ	Z <u>A</u> PPJERCHQQ

(P): (H I R D R E G I M E) (O L U B S Y U E R L) (I L L E R Y F I R E)

FIGURE 21

FIGURE 22a

FIGURE 22b

(2) The next example in Fig. 22a illustrates the strip arrangement for the first 10 letters of key, BPAUVNLFJA, for the message beginning given in subpar. a(2). If a one-letter introductory key has been used, these key letters apply to the 2d through 11th cipher letters, PAUVNLFJAL; the decipherment of these letters is found to their immediate left, which is OLUBSYUERL, obviously not plain text. On the same diagram we then search for the decipherment of the 3d through 12th letters, assuming that a two-letter introductory key was employed; again this yields no valid plain text. Finally, on the 8th trial, on the assumption that an 8-letter introductory key is involved, we obtain the plain text ILLERY FIRE; this is shown in Fig. 22b.

(3) For the third and final example, there is illustrated in Fig. 23a the strip arrangement for the first 10 letters of assumed key, DITGCMGTZB, for the message beginning given in subpar. a(3).

(K): (D I T G C M G T Z B)
(C): (I T G C M G T Z B P)

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

(P): (F L N W K U N G C O)

FIGURE 23a

(Z B P C V D Q K Y S)
(B P C V D Q K Y S K)

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

(C O N T I N U O U S)

FIGURE 23b

Nothing is seen here, so a number of additional trials is made, sliding the assumed key over successive 10-letter segments of the cipher text, all without success. We could now assume that an introductory key word was used, and that the autokeyed portion began with the last letter of cipher text after the end of the introductory key. With this in mind, we take as hypothetical key some text after the beginning of the cryptogram, say the 9th through 18th letters, ZBPCVDQKYS; trying this as key for the 10th through 19th letters, BPCVDQKYSK, we are successful on the first trial as shown in Fig. 23b with the emergence of the plain text CONTINUOUS.⁷

c. The foregoing mechanical method serves in helping to understand the mechanics of solution of ciphertext autokey encipherment involving known components. A simpler approach, however, is the use of the method of searching for the location of a probable word, as illustrated in the previous volume.⁸

⁷ Had we been a little more observant, we could have noticed what appears to be a good plaintext fragment NGCO in the very first trial in Fig. 23a; this was a contrived lapse of observation, the better to illustrate a pedagogical point in Fig. 23b.

⁸ Cf. Subpar. 22d on pp. 41-42 of *Military Cryptanalytics, Part II*.

If we had only the cipher text, and assumed that it was the result of ciphertext autokey encipherment with direct standard alphabets and a one-letter introductory key, we would perform the following decipherment on the basis of A_p as the index letter:

K : A R J K X Y M C U Q E B E Q O K G Q N
 C : A R J K X Y M C U Q E B E Q O K G Q N A . . .
 "P": R S B N B O Q S W O X D M Y W W K X N

The "decipherment" does not yield plain text; but if we *complete the plain-component sequence on the result of this decipherment*, we will obtain the true plain text on one of the generatrices, as shown in the diagram below:

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

23. Principles of solution by frequency analysis.—*a.* It is apparent that repetitions in ciphertext autokey systems will not be nearly as plentiful in the cipher text as they are in the plain text, because in these systems before a repetition can appear two things must happen simultaneously. First, of course, the plaintext sequence must be repeated, and second, one or more ciphertext letters (depending upon the length of the introductory key) immediately before the second appearance of the plaintext sequence must be identical with one or more ciphertext letters immediately before the first appearance of the plaintext sequence. This can happen only as the result of chance. In the following example the introductory key is the single letter X, and the components are direct standard alphabets employed in the usual Vigenère manner:

K: X C K B T M D H N V H L Y . . . K D K S J M D H N V H L Y
 P: F I R S T R E G I M E N T . . . T H I R D R E G I M E N T
 C: C K B T M D H N V H L Y R . . . K D K S J M D H N V H L Y R

The repeated plaintext word, REGIMENT, has only 8 letters but the repeated ciphertext sequence contains 9, of which only the last 8 letters actually represent the plaintext repetition. In order that the word REGIMENT be enciphered by DHNVHLYR the second time this word appeared in the text, it was necessary that the key letter for its first plaintext letter, R, be M *both* times; no other key letter will produce the same cipher sequence for the word REGIMENT in this case. Each different key letter for enciphering the first letter of REGIMENT will produce a different encipherment for the word, so that the chance for a repetition in this case is 1 in 26. This is the principal cause for the reduction in repetitions in this system. If an introductory key of two letters were used, it would be necessary that the two cipher letters immediately before the second appearance of the repeated word REGIMENT be identical with the two cipher letters immediately before the first appearance of the word; therefore the chance for a repetition in this case is 1 in 26^2 . In general, then, an n -letter repetition in the cipher text, represents an $(n-k)$ letter repetition in the plain text, where n is the length of the ciphertext repetition and k is the length of the introductory key.

b. There is a second phenomenon of interest in connection with ciphertext autokey systems. Let the letter opposite which the key letter is placed (when using sliding components for encipherment) be termed, for convenience in reference, the "base letter." Normally the base letter is the initial letter of the plain component, but it has been pointed out in the preceding volume that this is only a convention. Now when the introductory key is a single letter, if the base letter occurs as a plaintext letter its cipher equivalent is identical with the immediately preceding cipher letter; that is, there is produced a double letter in the cipher text, no matter what the cipher component is and no matter what the key letter happens to be for that encipherment. For example, using the HYDRAULIC . . . XZ sequence for both primary components, with H (the initial letter of the plain component) as the base letter, and using the introductory key letter X, the following encipherment is produced:

```

K: X|U N F F T T V K U H H M B N
P: I F T H E H Y P O T H E S I S . . .
C: U N F F T T V K U H H M B N E

```

Note the doublets FF, TT, and HH. Each time such a doublet occurs it means that the second letter represents H_p , which is the base letter in this case (the initial letter of the plain component). Now if the base letter happens to be a high-frequency letter in normal plain text, for example the letter E or T, then the cipher text will show a large number of doublets; if it happens to be a low-frequency letter then the cipher text will show very few doublets. In fact, the number of doublets will be directly proportional to the frequency of the base letter in normal plain text. Thus, if the cryptogram contains 1,000 letters there should be about 72 occurrences of doublets if the base letter is A, since in 1,000 letters of plain text there should be about 72 A's. Conversely, if a cryptogram of 1,000 letters shows about 72 doublets, the base letter is likely to be A; if it shows about 90, it is likely to be T, and so on. Furthermore, when a clue to the identity of the base letter has been obtained in this manner, it is possible immediately to insert the corresponding plaintext letter throughout the text of the message. The distribution of this letter may not only serve as a check (if no inconsistencies develop) but may also lead to the assumption of values for other cipher letters.

c. When the introductory key is two letters, then this same phenomenon will produce groups of the formula ABA, where A and B may be any letters, but the first and third must be identical. The occurrence of patterns of this type in this case indicates the encipherment of the base letter.

d. The phenomenon noted above can be used to considerable advantage in the solution of ciphertext autokey cryptograms. For instance, if it is known that the ordinary Vigenère method of encipherment ($\theta_{x/2} = \theta_{1/1}$; $\theta_{p/1} = \theta_{c/2}$) is used, then the initial letter of the plain component is the base letter. If, further, it is known that the plain component is the normal A-Z sequence, then the base letter is A and a word such as BATTALION will be enciphered by a group having the pattern AABCCDEFG. If the plain component is a mixed sequence and happens to start with the letter E, then a word such as ENEMY would be enciphered by a sequence having the pattern AABBCD.¹⁰ Sequences such as these are, of course, idiomorphic and if words yielding such idiomorphisms are frequent in the text there will be produced in the latter several

¹⁰ Six letters are shown because the idiomorphism in this case extends over that many letters.

or many cases of isomorphism. When these are analyzed by the principles of indirect symmetry of position, a quick solution may follow.

e. A final principle underlying the solution of ciphertext autokeyed cryptograms remains to be discussed; it concerns the nature of the frequency distribution required for the analysis of such cryptograms. Consider the message beginning illustrated in Fig. 18a in subpar. 21b. It happens that the letter W_c occurs three times in this short message beginning and, because of the nature of the ciphertext autokeying method, this letter must also appear three times in the key. Now it is obvious that all plaintext letters enciphered by key letter W_k will be in the same cipher alphabet; in other words, if the key text is cipher text offset one letter to the right of the cipher text, *then every cipher letter which immediately follows a W_c in the cryptogram will belong to the same cipher alphabet*, and this alphabet may be designated conveniently as the W cipher alphabet. Now if there were sufficient text, so that there were, say, 30 to 40 W_c 's in it, then a frequency distribution of the letters immediately following the W_c 's will exhibit monoalphabeticity. What has been said of the letters following the W_c 's applies equally well to the letters following all the other letters of the cipher text, the A_c 's, B_c 's, C_c 's, and so on. In short, if 26 distributions are made, one for each letter of the alphabet, showing the cipher letter immediately succeeding each different letter of the cipher text, the text of the cryptograms can be allocated into 26 uniliteral, monoalphabetic frequency distributions which can be solved by frequency analysis, provided that there are sufficient data for this purpose.

(1) The foregoing principle has been described as pertaining to the case when the introductory key is a single letter; that is, when the key text is offset or displaced but one interval to the right of the cipher text. But it applies equally to cases wherein the key text is offset more than one interval, provided that the frequency distributions are based upon the proper interval, as determined by the displacement due to the length of the introductory key. For instance, suppose the introductory key consists of two letters, as in the following example.

```

K: X Z M R H F H G F N Q R X O M R M V W E E
P: R E L I A B L E I N F O R M A T I O N . . .
C: M R H F H G F N Q R X O M R M V W E E . . .

```

The key text in this case is offset two intervals to the right of the cipher text; therefore if we made frequency distributions by taking the cipher letters one interval to the right of a given cipher letter (each time that letter occurs), these distributions will not be monoalphabetic because some letter not related at all to the given cipher letter is the key letter for enciphering the letter one interval to the right of the letter. For example, note the three R_c 's in the foregoing illustration. The first R_c is followed by H_c , representing the encipherment of L_p by M_k ; the second R_c is followed by X_c , representing the encipherment of F_p by Q_k ; the third R_c is followed by M_c , representing the encipherment of A_p by M_k . The three cipher letters H , X , and M are here entirely unrelated and do not belong to the same cipher alphabet because they represent encipherments by three different key letters. On the other hand, the cipher letters *two* intervals to the right of the R_c 's, *viz.*, F , O , and V , are in the same cipher alphabet because these cipher letters are the results of enciphering plaintext letters I , O , and T , respectively, by the *same* key letter, R . It is obvious, then, that when the introductory key consists of two letters and the key text is displaced two intervals to the right of the cipher text, the proper frequency distributions for monoalphabeticity will be based upon the letter at the second interval to the right of each cipher letter. Likewise, if the introductory key consists of three letters and the key text is displaced three intervals to the right of the cipher text, the distributions must be based upon the third interval, and so on, in each case the interval used corresponding to the amount of displacement between key text and cipher text.

(2) Conversely, in solving a problem of this type, when the length of the introductory key and therefore the amount of displacement are not known, the appearance of the frequency distributions based upon various intervals after each different cipher letter will disclose this unknown factor, since only one set of distributions will exhibit monoalphabeticity and the interval corresponding to that set will be the correct interval.

24. Example of solution by frequency analysis.—a. It will be assumed that previous studies have disclosed that the enemy is using ciphertext autokey systems; it will be further assumed that these studies have also disclosed that (1) the introductory key is usually a single letter, (2) the usual Vigenère method of employing sliding primary components is used, and (3) the plain component is usually the normal A-Z sequence, the cipher component a mixed sequence which changes daily. The following cryptograms, all of the same date, have been intercepted:

Message No. 1

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

Message No. 2

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

Message No. 3

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

Message No. 4

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

Message No. 5

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

Message No. 6

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

Message No. 7

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

b. A distribution table is now compiled, the results of which are shown in Fig. 26, below; in originally making the distribution, tallies had been recorded in the appropriate cell in the pertinent horizontal line of the table to indicate the cipher letter which immediately followed each occurrence of the letter to which that line applies. Obviously, the best method of compiling the data is to treat the text biliterally, taking the first and second letters, the second and third letters, and so on, distributing the digraphs as tallies in a digraphic distribution.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	N	ϕ	I.C.
A	3	1	2				1			3	1	1	1	4		6	3						1		2	29	64	2.05	
B		1	1	2				1		3		1	1	4		1	1	2	2	1		2	2		1	24	26	1.22	
C	2		2	5	3	3		1		2	3			1	1				2	2	1	1		2	4	35	62	1.35	
D	2	1		2			5	2	10		1	1				1		2		1					2	30	120	3.59	
E			4		1	1		3	2	1	3					2		2	5		1	1		2		28	52	1.79	
F	1			1	1	4		2	1	4	1		1	2					1	1				1		21	28	1.73	
G		2	1	1	1		3		3	1	2		1			1	3	1	3				1	3	1	29	32	1.02	
H	1		1			1		2	3		2	2		2		3		1			1	1	1		2	23	22	1.13	
I	1	4		5	5	1				1	2					1	1	2	1	1		2	3	3		33	70	1.72	
J	1		3		1		2	1	2	4		1	2	3	4	1					4		1	2		32	56	1.47	
K	4	1	1						5	3	1	4	2	1			2	1	1		1		2	2		31	58	1.62	
L	2	1		1			1					5	3							4		1	2	2		22	44	2.48	
M		1			3	2			1	1	4	4	2			2		2	2	2	5	3		2		37	70	1.37	
N		1	2		1				4	3	1	2	2			2			2					1		21	28	1.73	
O		3	2				2	2	1	2		1	3			1	1				5	2		3		28	48	1.66	
P	2	1	1	3		1	1				1			3	1		1		2	2						19	18	1.37	
Q	2	1			2		1	1	1	1	2		1	1	1		1	1		1				5	4	26	38	1.52	
R					1	2	1		2	1	2		2	1	2	2	2	1		2	1	2	1			25	18	0.78	
S	1	1	1		2	4	1			1	1				1		1	2	5			1				22	36	2.03	
T	1	4	6	1		2			2			1	1	2	3	1		6	4		1	2		4		41	110	1.74	
U	3		3		1		2	3	3		1	1			1	4		1	2		2	2	1			30	44	1.31	
V			1	2	1		3	1	1			1	2			1			4		1		1	1		20	22	1.51	
W			1							6	1	1	1	1	4			1	2		1	1	1	1		22	44	2.48	
X	2		3	4	2	1	4	1	2			3	2							2		2				28	50	1.72	
Y		1	1		1	2	2				2	1	1	2			2	1	2					2	2	22	16	0.90	
Z	1		1	1	4		1	2		2	2	5	1	1			2	1	2			1				27	42	1.56	
																										705	1218	42.85	

FIGURE 26

c. The individual frequency distributions give every appearance of being monoalphabetic, which confirms the assumption that the enemy is using ciphertext autokey with a *single-letter* introductory key. The average I.C. of the rows of the matrix is $\frac{42.85}{26} = 1.65$, which is fine;¹¹ or, as a better approach, we could calculate the *digraphic* I.C. of the matrix by considering the sum (1218) of the ϕ values of Fig. 26 as the observed value of ϕ and substituting in the formula $\delta = \frac{676 \sum f(f-1)}{N(N-1)} = \frac{676(1218)}{705 \times 704} = 1.66$, again substantiating the same assumption.¹² (This discrepancy between the two figures lies in the round-off errors introduced in obtaining an average I.C.)

¹¹ The arithmetic mean here suffices because the values of N involved are fairly close to one another; since, as has been previously stated, in ciphertext autokey systems the cipher letters are equiprobable (the over-all I.C. of the cipher text in this example is 1.005), a weighted mean is unnecessary.

¹² Ciphertext autokey systems may therefore be identified statistically from the appropriate digraphic distribution (i.e., on the assumption of the correct length of the introductory key) by the fact that the digraphic I.C. will reflect the *monographic* I.C. of the language.

d. The total number of letters of text is 712, comprising 705 digraphs. If the base letter is A, then there should be approximately $705 \times 7.4\% = 52$ cases of doubled letters in the text. There are actually 53 doublets, which checks very well with the expectancy. The letter A is substituted throughout the text for the second letter of each doublet.

e. The following sequence is noted at the beginning of Message No. 5:

G I S U H W Z H S T T Z O I D D H O O V N B T J G X C T B S
 A A A

Assume that the sequence DDHOOVNBT represents the word BATTALION, in which case we will have the following key-cipher-plain relationships:

K: I D D H O O V N B T
 C: D D H O O V N B T
 P: B A T T A L I O N

If this assumption is correct, the frequency of H_c in the D alphabet should be high, since H_c=T_p; the H_c has only two occurrences. Likewise, the frequency of O_c (=T_p) in the H alphabet should be high; it is also only two. The frequency of V_c in the O alphabet should be medium or low, since it would equal L_p; it is five, which is too high. The rest of the letters of the assumed word are similarly checked against the appropriate frequency distributions, with the result that, on the whole, the assumption that the DDHOOVNBT sequence represents BATTALION does not appear to be warranted. Similar attempts are made at other points in the text, with the same or other probable words. Some of these attempts may have to be carried to the point where the placement of values in the tentative cipher component leads to serious inconsistencies. Finally, attention is fixed upon the following sequence in the second line of Message No. 6:

M M T T F U U N M G
 A A A

If we assume that this skeleton represents the word AVAILABLE, the following fragment of key, cipher, and plain should be true:

K: M M T T F U U N M G
 C: M T T F U U N M G
 P: A V A I L A B L E

Reference is now made to the appropriate frequency distributions to see how well the actual individual frequencies correspond to the expected ones; these data are tabulated in the diagram below:

Alphabet	Assumed		Frequency		Approximation
	θ _c	θ _p	Expected	Actual	
M	T	V	Low	2	Fair
T	F	I	High	2	Fair
F	U	L	Medium	1	Good
U	N	B	Low	1	Good
N	M	L	Medium	2	Fair
M	G	E	High	2	Poor

~~SECRET~~

This assumption of AVAILABLE cannot be discarded just yet. Let the values derivable from the assumption be inserted in their proper places in a cipher component, and, using the latter in conjunction with a normal A-Z sequence as the plain component, let an attempt be made to find corroboration for these values. The following placements may be made:¹³

P: 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
C: M F G U N T

The letter M_c appears twice in the cipher sequence and when this partially reconstructed cipher component is tested it is found that the value L_p(N_k) = M_c is corroborated. Having the letters M, F, G, U, N, and T tentatively placed in the cipher component, it is possible to insert certain plaintext values in the text. For example, in the M alphabet, F_c=D_p, G_c=E_p, U_c=O_p, N_c=P_p, and T_c=V_p. In the F alphabet, G_c=B_p, U_c=L_p, N_c=M_p, T_c=S_p, and M_c=X_p. The other letters yield additional values in the appropriate alphabets. The plaintext values thus obtainable are inserted in the cipher text. No inconsistencies appear and, moreover, certain good digraphs are brought to light. For instance, note what is manifested at the end of the third line of Message No. 5:

K: U Q Z H Z M T F H
C: U Q Z H Z M T F H Z
P: V I

Now if the letter H can be placed in the cipher component, several values might be added to this partial decipherment. We note that F and G are sequent in the cipher component; now let us suppose that H follows G therein, and we obtain the following:

K: U Q Z H Z M T F H
C: U Q Z H Z M T F H Z
P: V I C

Suppose the VIC is the beginning of VICINITY. This assumption permits the placement of A, C, L, and Z in the cipher component, as follows:

P: 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
C: M A F G H L Z U N T C

These additional values check in very nicely and presently the entire cipher component is reconstructed. It is found to be as follows:

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

The key phrase is clearly based upon UNDERWOOD TYPEWRITER COMPANY. All the messages may now be deciphered with ease. The following gives a letter-for-letter decipherment of the first three groups of each message:¹⁴

Message No. 1 K: KI J X W X E E C D A C N Q E
C: I J X W X E E C D A C N Q E T . . .
P: R I G H T F A I R L Y Q U I E
Message No. 2 K: EG R V R M Z W K X G W P C K
C: G R V R M Z W K X G W P C K K . . .
P: N O T H I N G O F S P E C I A

¹³ Note that, had we not known (or assumed) the plain-component sequence, we would first have entered these values in a 26×26 square rather than in a single strip for the cipher component, and then we would exploit any manifestations of direct or indirect symmetry present.

¹⁴ The introductory keys for these messages are presumed to have been specified by prearrangement, or indicated by the message number, file time, or some other element of the message externals.

~~SECRET~~

Message No. 3 K: R|R W K A O L T C J M Z D K V
 C: R W K A O L T C J M Z D K V U . . .
 P: A B O U T O N E H U N D R E D

Message No. 4 K: J|X J J P M L T Z K X E C A Q
 C: X J J P M L T Z K X E C A Q Z . . .
 P: G U A R D I N S U F F I C I E

Message No. 5 K: E|G I S U H W Z H S T T Z O I
 C: G I S U H W Z H S T T Z O I D . . .
 P: N U M E R O U S F L A S H E S

Message No. 6 K: B|T B J P A A R Y Y P V H I D
 C: T B J P A A R Y Y P V H I D I . . .
 P: T H E R E A R E A B O U T S I

Message No. 7 K: B|T B J P A Q A A Z T R X A L
 C: T B J P A Q A A Z T R X A L X . . .
 P: T H E R E I S A M I X U P H E

f. In the foregoing example the plain component was the normal sequence, so that with the Vigenère method of encipherment the base letter is A. If the plain component is a mixed sequence, the base letter may no longer be A, but in accordance with the principle set forth in subpar. 23b, the frequency of doublets in the cipher text will correspond with the frequency of the base letter as a letter of normal plain text.¹⁵ If a good clue is afforded by the frequency of doublets in the cipher text, the insertion of the corresponding base letter in the plain text will lead to further clues. The solution from there on can be handled along the lines indicated above.

25. **Solution by means of isomorphs.**—a. It was stated in subpar. 23d that in ciphertext autokey systems the production of isomorphs is a frequent phenomenon and that analysis of these isomorphs may yield a quick solution. An example of this sort will now be studied, using as an illustration the following three messages which are suspected of being in a ciphertext autokey system:

Message No. 1

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

Message No. 2

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

Message No. 3

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

¹⁵ If the plain and cipher components had been identical sequences, this fact together with the identity of the base letter could have been determined from the digraphic distribution: one of the rows of the distribution (the row corresponding to the base letter) would reflect an approximation of the normal frequency distribution, i.e., peaks for the letters AEINORST, and blanks or near-blanks for JKQXZ. Furthermore, the reconstruction matrix would have displayed symmetry about the main diagonal (from upper left to lower right); see in this connection subpar. 33e in the next chapter.

b. Frequency distributions are made, based upon the second letters of pairs, as in the preceding example. These distributions are shown in the table in Fig. 27, below. The digraphic I.C. is $\frac{676(500)}{451 \times 450} = 1.67$, confirming ciphertext autokey with a single-letter introductory key. Nevertheless, the data in each distribution are relatively scanty and it would appear that the solution is going to be a rather difficult matter.

		2d letter																												
		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	N	φ	
1st letter	A	3	1	1	2							2	2	2	1					3					1		18	20		
	B	2		1	1	1		1	2	1			2	4									3	2		3	23	32		
	C	1		1		1		1							2			1	1	1	2	1	3			1	1	17	8	
	D	2		2	1				2			1	4	1			2		1									16	20	
	E	1			1	1	1				4			1	2	1					1		2	2	1	2	20	20		
	F	1	1		2			1	1		1	1			1			1	2						2		1	15	6	
	G								1						2				1	3			1		1	1	3	13	14	
	H	4		3								1		1		2							1					12	20	
	I		1	1			2	1		2	1			1			1	1			2					5		18	26	
	J										1		4		1	4	1		1		2						2	16	28	
	K				1					2		2	1		1	1		1		1		1	4	4			2	20	30	
	L	1	1			2	1	1			1	3	1		2		2	1							1	2	3	22	20	
	M	1		1					2				4	3		1	1	1	2		1					1		18	22	
	N	2	3		1	1				2		2			3			1	1	1						2	1	20	20	
	O	1	2	3	1	3	1		1				1		1						1	4	2	3			2	26	36	
	P		2			1		1	1		1		2	1									2	2				12	8	
	Q					1	2	2		2			1			1		2						1	1			13	8	
	R			1		4				2		1	1		1	2	1		2	3			2			2		22	28	
	S			3		1	1				1	2								1		1				1	1	12	8	
	T		1	3				1	1		2	3		1	2				1		2					2	1	1	21	20
	U		1	1					1	1					2		2	1	2	1			1		1		1	15	6	
	V	2	3		2	3	1					3	3		1	1				1				2			1	23	30	
	W					1			2				1						1		2	2						9	6	
	X	1	1		5			2					3	1		1		1	3	1	1					2	1	23	36	
	Y		2					1			1	1	1	1			3						2				1	13	10	
	Z				3					1	2				3	2	1		2									14	18	

451 500

FIGURE 27

c. Before becoming discouraged too quickly, however, we make a search throughout the text to see if any isomorphs are present. Fortunately, there appear to be several of them. Note the following:

Message No. 1 (1) . . . D B D D Q G S U N S F B O B E K . . .
 (2) . . . N E V V J L K Z E K U R C N I F . . .
 (3) . . . T N K K U X L O D L W T H C Z R]
 Message No. 2 (4) . . . C R I I X O Z E T Z N K]
 Message No. 3 (5) . . . C Q V V J L K Z E K U R F R F X . . .

First, it is necessary to delimit the length of the isomorphs. Isomorph (2) shows that the isomorphism begins with the doubled letters; for there is an E before VV in that case and also an E within the isomorph. If the phenomenon included the E, then the letter immediately before the DD in the case of isomorph (1) would have to be an N, to match its homolog, E, in isomorph (2), which it is not. Corroborating data are

given by isomorphs (3), (4), and (5) in this respect. Hence, we may take it as established that the isomorphism begins with the doubled letters. As for the end of the isomorphism, the fact that isomorphs (2) and (5) consist of the same set of 10 letters seems to indicate that this number defines the length of the isomorphism. The fact that Message No. 2 ends 2 letters after the last tie-in letter, Z, corroborates this assumption. It is at least certain that the isomorphism does not extend beyond 11 letters because the recurrence of R in isomorph (5) is not matched by the recurrence of R in isomorph (2), nor by the recurrence of T in isomorph (3). Hence it may be assumed that the isomorphic sequence is probably 10 letters in length, possibly 11. But to be on safe ground it is best to proceed on the 10-letter basis.

d. By applying the principles of indirect symmetry to the superimposed isomorphs, partial chains may be constructed, as shown below:

```
(1-2) D V   Q J   G L   S K   F U Z   N E   B R
(1-3) N D K   Q U O   G X   S L   F W   B T
(1-4) D I   Q X   G O   S Z   U E   F N T   B K
(2-3) V K L X   J U W   Z O   E D   R T
(2-4) V I   J X   L O   R K Z E T   U N
(3-4) D T K I   U X O E   L Z   W N
```

These partial chains may be amalgamated into the following sequence:

L O D J X B S U N . G W . Q . . . F V I . R K Z E T

Noting the J K at an interval of -7, and also W X Z at the same interval, we conclude that a keyword-mixed sequence is involved, and we derive the original sequence as

W X Z . . D R . U L I . B E F G J K . N O . Q S T V,

whereupon we recognize our perennial friend HYDRAULIC and fill in the missing six letters.

e. We now have the cipher component, and the plain component remains to be reconstructed. The simplest and most foolproof solution ordinarily is a reduction to monoalphabetic terms, using the recovered cipher component and the known offset of the cipher text against itself as key.¹⁶ However, the probable word method, if the probable words are at all probable, may be used to good advantage. A good crib to assume for the 10-letter repetition found in Message Nos. 1 and 3 is ARTILLERY (especially since the doublet rate of the distribution in Fig. 27 is $\frac{33}{451} = 7.3\%$, which is just right for a base letter of A_p to represent the doublet in the repetition). This single assumption is sufficient to place 7 letters in the plain component, thus:

```
K:   V V J L K Z E K U R
C:   V V J L K Z E K U R
P:   A R T I L L E R Y
```

A . . . E . . . I . . L R . T Y .

These few letters (few, but how beautifully spaced!) are sufficient to suggest that the plain component is in all probability the normal sequence. A few moments' testing proves this to be true. The two components are therefore:

```
P: 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
C: 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
```

¹⁶ Note that if the LODJX . . . sequence in subpar. d, above, had not been of systematic construction to enable us to analyze its derivation and thus fill in the missing 6 letters, we still could have converted most of the cipher text to monoalphabetic terms, solved the text, and recovered both components.

f. With the two components at hand, the decipherment of the message is a low-order triviality. Since a single-letter introductory key is known to have been used,¹⁷ we decipher the first five groups of Message No. 1 as follows:

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

The mangled beginning is the result either of garbles, or of specialized keying procedure wherein the *last* letter of an introductory key was used as the introductory key letter for enciphering the subsequent autokeyed portion of the text (see Fig. 18c in subpar. 21b). If we assume that the IVE before the word FIRE is the ending of the first word of the plain text, and that this word is INTENSIVE, the introductory key word is found to be WICKER. Thus:

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

The beginnings of the other two messages are recoverable in the same way and are found to be as follows:

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

g. The example solved in the foregoing subparagraphs offers an important lesson to the student, insofar as it teaches him that *he should not immediately feel discouraged when confronted with a problem presenting only a small quantity of text and therefore affording what seems at first glance to be an insufficient quantity of data for solution*. For in this example, while it is true that there are insufficient data for analysis by simple principles of frequency, it turned out that solution was achieved *without any recourse to the principles of frequency of occurrence*. Here, then, is one of those interesting cases of substitution ciphers of rather complex construction which are solvable without any study whatsoever of frequency distributions. Indeed, it will be found to be true that in more than a few instances the solution of quite complicated cipher systems may be accomplished not by the application of the principles of frequency, but by recourse to inductive and deductive reasoning based upon other considerations, even though the latter may often appear to be very tenuous and to rest upon quite flimsy supports.

26. Solution of isologs involving the same pair of unknown primary components.—a. Two messages containing identical plain text encrypted in a ciphertext autokey system with two different single-letter introductory keys may be solved in a manner identical to that described in the last paragraph, since what we really have is a pair of long isomorphs one letter shorter than the length of the messages. Even if the introductory keys are words of different lengths and compositions, if the key usage is similar to that illustrated in Fig. 18c in subpar. 21b the message can be solved very rapidly by reconstructing the primary components, since the cryptographic texts of such messages will be isomorphic after the initial keyword portions.

¹⁷ We know this from (a) statistical evidence of the digraphic distribution at an offset of 1, (b) the indications of the correct plain component emerging from a tentative decipherment of the 10-letter repetition, and (c) the unlikelihood that with three rather short messages a long ciphertext repetition would have manifested itself if the offset were more than 1 letter. We knew that the three messages were autokeyed at the same offset, otherwise the isomorphs would have not appeared among all three messages.

(1) Note the two following superimposed messages, in which isomorphism between the two cryptograms is both obvious and consistent after their 6th letters:

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

Starting with any pair of superimposed letters (after the 6th pair), the following chains are derived:

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

The foregoing fragments either are part of two 13-letter chains, or they are parts of a complete 26-letter sequence. If the former is the case, then the two 13-letter chains must be (Z O B Y Q F R T J U W M I) and (L X N C H P E D S G A V K); and, a few moments later, noting phenomena associated with keyword-mixed sequences in the two chains, we superimpose them in the diagram ¹⁸

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

from which we speedily obtain the HYDRAULIC . . . XZ sequence.¹⁹

(2) Only the cipher component has been recovered thus far. If we assume that the plain component is the same as the cipher, the initial key words and the message plain texts are at once deciphered; it will be found that the initial key word for Message "A" is PENCE, and that for Message "B" is LATERAL.²⁰ If the plain component had not been guessed in this case, we could have "deciphered" the message text using an arbitrary plain component (say, the A-Z sequence), resulting in a conversion of the complex cipher text into monoalphabetic terms which can then speedily be solved.

(3) The foregoing solution affords a clue to the solution of cases in which the texts of two or more messages are not completely identical but are in part identical because they happen to have similar beginnings or endings, or contain nearly similar information or instructions. The progress in such cases is not so rapid as in the case of messages with wholly identical texts because much care must be exercised in blocking out the isomorphic sequences upon which the reconstruction of the primary components will be based.

b. In the preceding example the autokeyed portions of the texts started with the last letters of the introductory keys. If full autokeying (i.e., the method shown in Fig. 18b), had been employed the solution would hardly be more difficult.

¹⁸ See subpar. 44i on pp. 89-90 of *Military Cryptanalytics, Part II*.

¹⁹ If the four fragments (1), (2), (3), and (4) had been parts of a complete 26-letter sequence, there would have been only 6 ways to permute them, viz., 1-2-3-4, 1-2-4-3, 1-3-2-4, 1-3-4-2, 1-4-2-3, and 1-4-3-2; therefore the problem would still be solvable without too much effort, even if the cipher component has been a random sequence.

²⁰ The reason that the cryptographic texts are isomorphic after the initial keyword portions is, of course, that since the text beyond the key word is enciphered autokey fashion by the preceding cipher letter, the letters before the last letter of the key have no effect upon the encipherment at all. Hence two messages having identical plain text cannot be other than isomorphic after the initial keyword portions.

(1) In order to illustrate such a case, let the same plain texts used in the preceding example be enciphered by introductory key words of the same lengths but different compositions: PENCE and LATER. Thus:

Message "A"

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

Message "B"

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

(2) Now let the cipher texts be superimposed and isomorphisms be sought. They are shown underlined below:

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

It will be noted that the intervals between identical vertical pairs show a constant factor of 5, indicating that the messages have been enciphered with 5-letter introductory key words.

(3) The vertical pairs beyond the first five letters of the messages are now distributed in a reconstruction matrix according to their position based upon this interval of 5, similar to the treatment of vertical pairs in periodic-cipher isologs arising from the use of repeating keys of the same lengths.²¹ This is shown below:

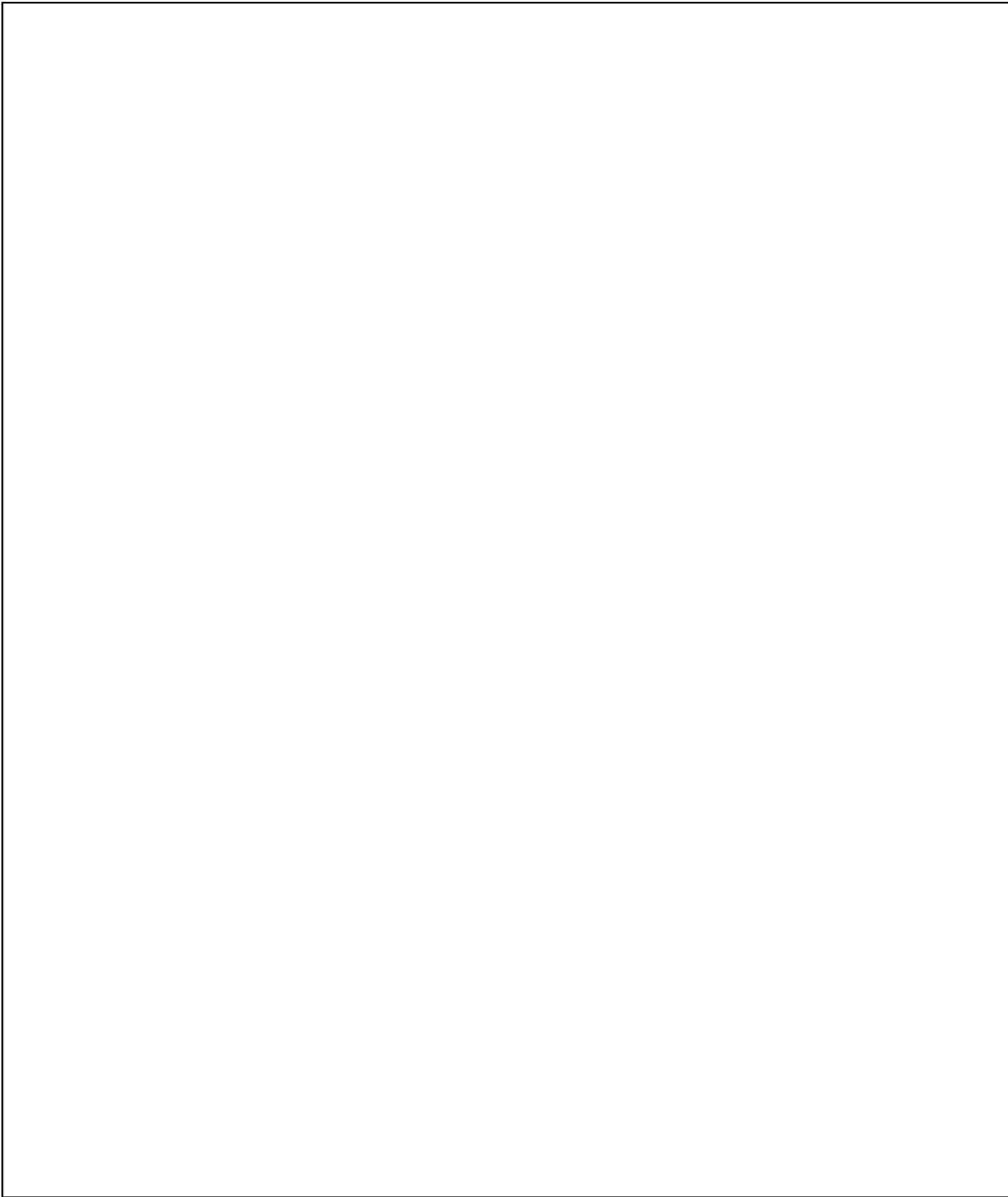
Ø	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	P	W		N			H		T	Y		S	R		L											
2		R	D			U					H	B	E		G									O		
3	B								G		Q		S	T						H	E		D			
4	L		E					D	B							T	U					Z	H		Y	
5					A		Q	H				C		E						K	X	M				

From the values in this matrix the original cipher component, the HYDRAULIC . . . XZ sequence, may quickly be recovered, because the Ø line may be included in the chaining.

²¹ Cf. subpar. 60f, *Military Cryptanalytics, Part II*.

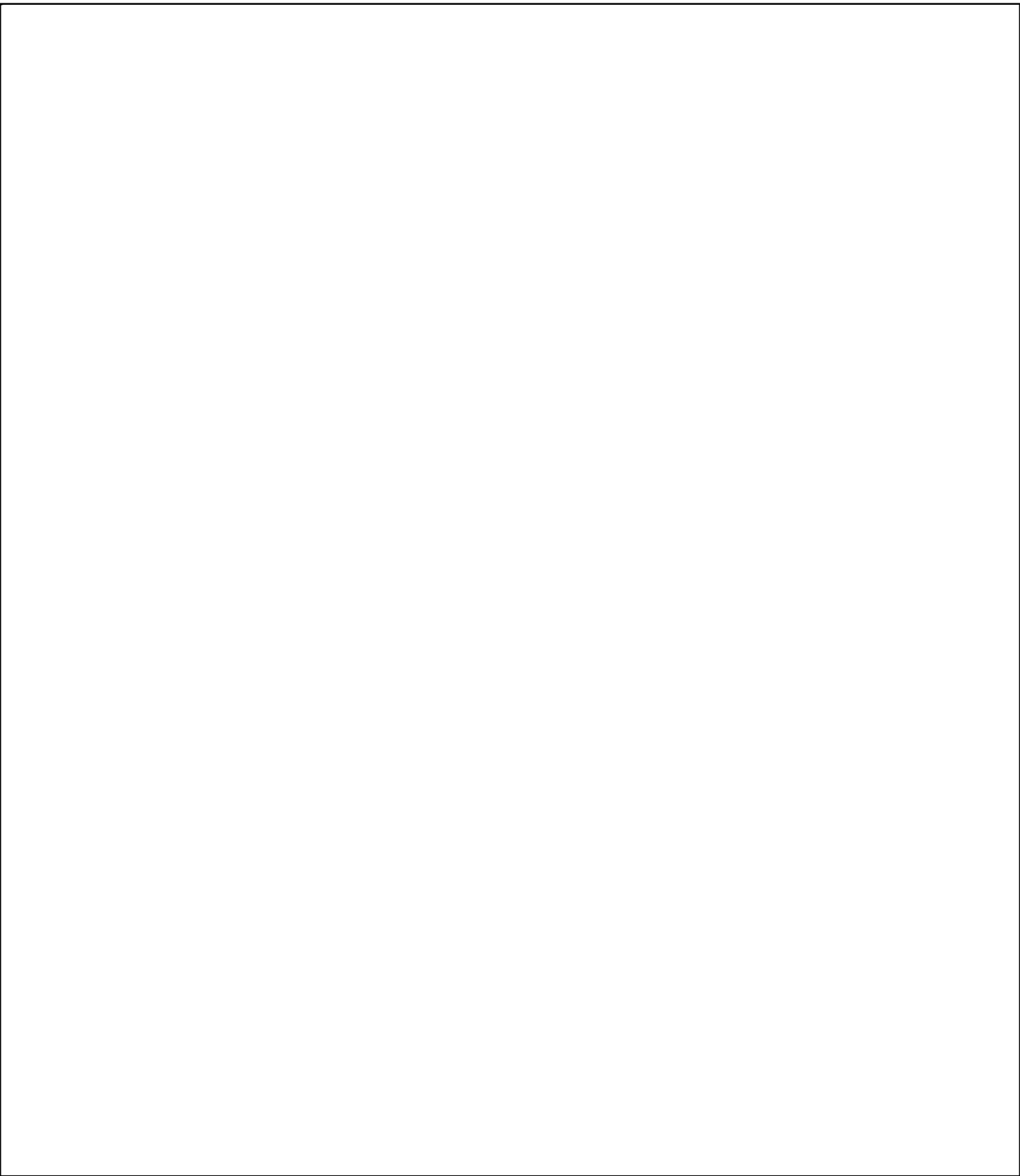
(b) (1)
(b) (3) -18 USC 798
(b) (3) -50 USC 3024(i)
(b) (3) -P.L. 86-36

(b) (1)
(b) (3) -18 USC 798
(b) (3) -50 USC 3024 (i)
(b) (3) -P.L. 86-36



(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024 (i)
(b) (3)-P.L. 86-36

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024 (i)
(b) (3)-P.L. 86-36

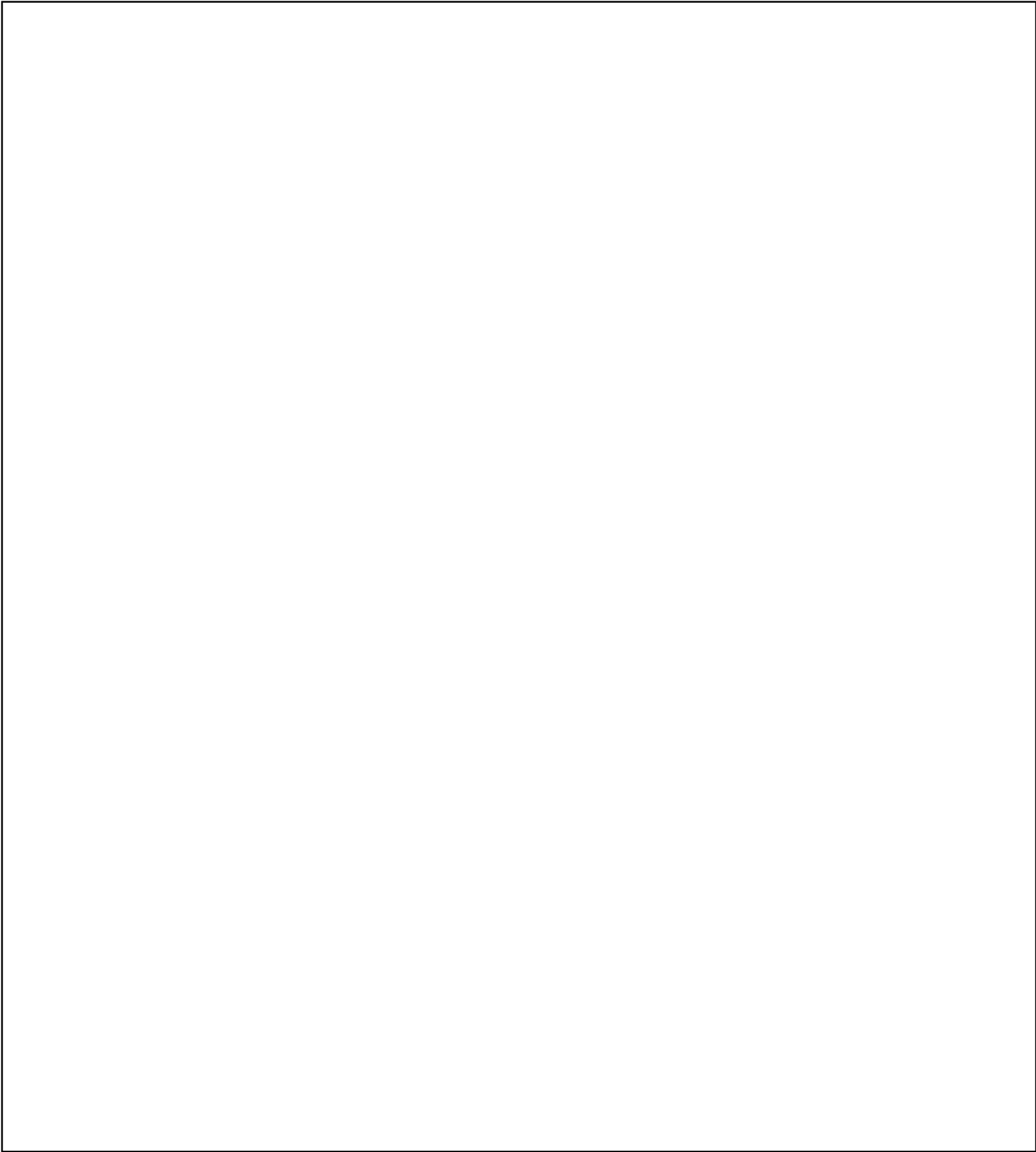


(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024 (i)
(b) (3)-P.L. 86-36

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024(i)
(b) (3)-P.L. 86-36

(b) (1)
(b) (3) -18 USC 798
(b) (3) -50 USC 3024 (i)
(b) (3) -P.L. 86-36

(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024(i)
(b) (3)-P.L. 86-36



(b) (1)
(b) (3)-18 USC 798
(b) (3)-50 USC 3024(i)
(b) (3)-P.L. 86-36

(b) (1)
(b) (3) -18 USC 798
(b) (3) -50 USC 3024 (i)
(b) (3) -P.L. 86-36

28. Further remarks on ciphertext autokey systems.—a. All of the discussion on ciphertext autokey systems thus far has been limited to alphabetical systems employing sliding primary components (or the equivalent form of a square table). There is no reason, of course, why a set of 26 unrelated random sequences in a table such as that in Fig. 33, below, could not be used for the cipher alphabets. In such

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

FIGURE 33

cases, the general methods treated in par. 22 still apply, with necessary modifications, as also do the methods in pars. 23 and 24, except that it is obvious that (a) there will be no determinable base letter, and (b) there will be no causal isomorphs. For that matter, even with a matrix such as that of Fig. 34 below, in which the key letters are designated by *arbitrary* letters to the left of the square (instead of

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

FIGURE 34

the letters in a column under a particular plaintext letter—the base letter), there is likewise no determinable base letter and no causal isomorphs can be produced, as can be shown by the following isologous message beginnings:

Message "A"

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

Message "B"

K: YQNSTTVULPAESJOUBNOADTEXPAOEFT
P: THIRDRREGIMENTCOMMANDPOSTMOVING...
C: QNSTTVULPAESJOUBNOADTEXPAOEFTU

The ciphertext doublets are the result of chance. Note in this example that, since the plain and cipher components are identical sequences, one of the distributions of the letters immediately following a particular cipher letter (in this case, A_c) will fit the normal, since the plaintext letters enciphered by this key letter will be self-enciphered; the distributions of the letters following the other cipher letters will of course be monoalphabetic.

b. The general principles of ciphertext autokeying apply equally well to digital systems, and for that matter to Baudot systems. The modulus in digital systems is the usual mod-10 arithmetic,²⁶ so that in effect the cipher component is a *known* sequence—thus a reduction to monoalphabetic terms suggests itself at once, if not sooner.

c. As an example, let us study the following message, suspected of having been enciphered in a ciphertext autokey system with the additive method of encipherment ($P+K=C$).²⁷

```

1 5 6 3 5 9 2 0 0 1 1 3 7 5 6 9 4 0 9 3 0 5 1 5 1 4 0 0 9 4
2 2 1 2 3 3 0 4 7 6 6 9 7 2 0 3 9 4 9 8 7 9 9 2 9 5 8 1 1 2
7 6 1 1 1 4 9 3 6 9 2 8 0 7 2 4 8 3 0 1 9 0 0 7 4 4 5 5 4 8
4 0 5 4 4 3 1 5 8 9 1 2 9 4 3 5 0 5 2 9 5 2 8 1 4 7 4 5 1 6
7 1 9 3 2 4 0 5 0 6 1 1 6 8 2 0 5 7 9 3 1 0 6 4 9 2 5 6 9 5
1 9 6 8 9 9 9 6 6 1

```

If a single-digit introductory key has been used, the cipher text is offset against itself at an interval of 1 and a decipherment obtained, the beginning of which is shown below:²⁸

```

K:  1 5 6 3 5 9 2 0 0 1 1 3 7 5 6 9 4 0 9 3 0 5 1 5
C: 1 5 6 3 5 9 2 0 0 1 1 3 7 5 6 9 4 0 9 3 0 5 1 5 1 . . .
P:  4 1 7 2 4 3 8 0 1 0 2 4 8 1 3 5 6 9 4 7 5 6 4 6

```

The I.C. of the entire deciphered message is 0.99, so the length of the introductory key was probably not 1. The cipher text is then offset against itself at intervals of 2, 3, 4 . . . , up to an interval of 9: the I.C.'s obtained from the resulting decipherments are all unsatisfactory, as may be seen from the following table:

Offset	I.C.	Offset	I.C.	Offset	I.C.
1	0.99	4	0.97	7	0.99
2	1.02	5	0.96	8	0.98
3	1.02	6	0.96	9	0.97

²⁶ If other than mod-10 arithmetic is used, say an arbitrary conversion-square encipherment with a square such as the following,

		Plain											
		0	1	2	3	4	5	6	7	8	9		
Key	0	3	1	6	7	5	8	2	4	9	0		
	1	1	2	3	5	4	6	8	9	0	7		
	2	6	4	0	2	8	1	5	3	7	9		
	3	7	3	2	9	0	5	6	1	8	4		
	4	5	8	4	0	9	2	3	7	1	6		
	5	8	5	1	6	7	9	0	2	4	3		
	6	2	9	8	1	3	7	4	0	6	5		
	7	4	6	9	8	2	0	7	5	3	1		
	8	9	0	7	3	6	4	1	8	5	2		
	9	0	7	5	4	1	3	9	6	2	8		

the problem must be solved as a general case of ciphertext autokey as treated earlier in this chapter. (The conversion square shown here is a *Latin square*, with ten unique digits in each of the rows and in each of the columns; other conversion squares may have columns containing repeated digits.)

²⁷ As regards the effect of the use of different encipherment conventions, see subpars. 35g and h (on pp. 117-120).

²⁸ We have assumed that the enciphering procedure was the *additive* method (i.e., wherein $P+K=C$); see also the remarks in subpar. 35h in the next chapter if subtractive or minuend methods had been involved.

d. When an offset of 10 is tried, the I.C. of the deciphered text jumps to 1.52 and a long repetition is in evidence, revealing that 10 is the length of the introductory key. The decipherment and its appertaining frequency distribution are shown below:

K: 1 5 6 3 5 9 2 0 0 1 1 3 7 5 6 9 4 0 9 3
 C: 1 5 6 3 5 9 2 0 0 1 1 3 7 5 6 9 4 0 9 3 0 5 1 5 1 4 0 0 9 4
 P: 0 8 1 2 1 0 2 0 9 2 9 2 4 0 5 5 6 0 0 1

K: 0 5 1 5 1 4 0 0 9 4 2 2 1 2 3 3 0 4 7 6 6 9 7 2 0 3 9 4 9 8
 C: 2 2 1 2 3 3 0 4 7 6 6 9 7 2 0 3 9 4 9 8 7 9 9 2 9 5 8 1 1 2
 P: 2 7 0 7 2 9 0 4 8 2 4 7 6 0 7 0 9 0 2 2 1 0 2 0 9 2 9 7 2 4

K: 7 9 9 2 9 5 8 1 1 2 7 6 1 1 1 4 9 3 6 9 2 8 0 7 2 4 8 3 0 1
 C: 7 6 1 1 1 4 9 3 6 9 2 8 0 7 2 4 8 3 0 1 9 0 0 7 4 4 5 5 4 8
 P: 0 7 2 9 2 9 1 2 5 7 5 2 9 6 1 0 9 0 4 2 7 2 0 0 2 0 7 2 4 7

K: 9 0 0 7 4 4 5 5 4 8 4 0 5 4 4 3 1 5 2 9 1 2 9 4 3 5 0 5 6 9
 C: 4 0 5 4 4 3 1 5 2 9 1 2 9 4 3 5 0 5 6 9 5 2 8 1 4 7 4 5 5 6
 P: 5 0 5 7 0 9 6 0 8 1 7 2 4 0 9 2 9 0 4 0 4 0 9 7 1 2 4 0 9 7

K: 5 2 8 1 4 7 4 5 5 6 7 1 9 3 2 4 0 5 4 6 1 1 6 8 2 0 5 7 3 3
 C: 7 1 9 3 2 4 0 5 4 6 1 1 6 8 2 0 5 7 3 3 1 0 6 4 9 2 5 6 3 5
 P: 2 9 1 2 8 7 6 0 9 0 4 0 7 5 0 6 5 2 9 7 0 9 0 6 7 2 0 9 0 2

K: 1 0 6 4 9 2 5 6 3 5
 C: 1 9 6 8 9 9 9 6 0 1
 P: 0 9 0 4 0 7 4 0 7 6

≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡
 ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡
 ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡
 ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡
 ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡
 ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡ ≡
 1 2 3 4 5 6 7 8 9 0

From here on the solution of the intermediate text is a simple matter; monome-dinome characteristics are observed in the preliminary examination, and recovery of the plain text and of the enciphering matrix quickly follow.²⁹

²⁹ The analysis of the intermediate text is given in par. 77 of *Military Cryptanalytics, Part I*.