~~CONFIDENT~~

THE DISTRIBUTION OF THIS SPECIAL TEXT WILL BE RESTRICTED TO REGULARLY ENROLLED EXTENSION COURSE STUDENTS, TO MILITARY PERSONNEL, AND TO OTHER PERSONS COMING WITHIN THE MEANING OF THE PHRASE "FOR OFFICIAL USE ONLY"

ARMY EXTENSION COURSES

SPECIAL TEXT No. 165

ELEMENTARY MILITARY CRYPTOGRAPHY

1935 EDITION

PREPARED UNDER THE DIRECTION OF THE CHIEF SIGNAL OFFICER
FOR USE WITH THE ARMY EXTENSION COURSES

William F. Friedman
Washington
1935
7

Records taken from
WFF's home
Box 25

UNITED STATES
GOVERNMENT PRINTING OFFICE
WASHINGTON : 1935

558

~~CONFIDENT~~

WAR DEPARTMENT,
WASHINGTON, *March 1, 1935.*

The following Special Text No. 165, Elementary Military Cryptograpy, for use with the Army Extension Courses, is published for the information and guidance of all concerned.

[A. G. 352.6 (12-26-34).]

BY ORDER OF THE SECRETARY OF WAR:

DOUGLAS MACARTHUR,
General,
Chief of Staff.

OFFICIAL:

JAMES F. McKINLEY,
Major General,
The Adjutant General.

~~CONFIDENTIAL~~

SPECIAL TEXT NO. 165

ELEMENTARY MILITARY CRYPTOGRAPHY

PART ONE

GENERAL

SECTION I

PRELIMINARY DEFINITIONS

1. Cryptology, secret writing, cryptography.—*a.* That branch of knowledge which treats of all the means and methods of secret intercommunication is called *cryptology*. The importance of cryptology as an adjunct to intercommunication between military commanders in the field, and between them and their home government, has been recognized from the earliest days of organized warfare.

b. Intercommunication may be conducted by any means susceptible of ultimate interpretation by one of the five senses, but those most commonly employed are visual or auditory in character. Aside from the use of simple visual and auditory signals for intercommunication over relatively short distances and the use of the telephone for direct intercommunication over greater distances, the usual method of intercommunication at a distance involves, at one stage or another, the act of *writing*.

c. Writing may be either *visible* or *invisible*. In the former, the characters are inscribed with ordinary writing materials and can be seen with the naked eye; in the latter, they are inscribed with certain chemicals called *invisible*, *sympathetic*, or *secret inks*, which have the property of either being initially invisible to the naked eye or of becoming so after a short time. In order to make writing which has been inscribed by the use of secret inks visible, special processes must usually be applied. There are, in addition to the foregoing, certain methods of producing writing which is invisible because its characters are microscopic in size. These methods usually employ either special photographic apparatus or very delicate mechanical instruments called micropantographs, by means of which ordinary writing may be copied in extremely reduced size. In order to become visible to the naked eye, and hence before such writing can be read, magnifying lenses of high power must be used.

d. Invisible writing, and visible writing which has been prepared in such a form as to be unintelligible in the language in which it is

written, constitute *secret writing*. Both of these forms of secret writing have their uses in military communications, but this text will deal only with visible secret writing.

c. That branch of cryptology which treats of visible secret writing is called *cryptography*. A more specific definition will be given later. (Par. 4*d*.)

2. Plain language and secret language.—*a.* Visible writing which conveys an intelligible meaning in the language in which it is written is said to be in *plain language*. The text of such writing is referred to as *plain text*, *clear text*, or *open text*.

b. Visible writing may apparently convey an intelligible meaning but the latter may not be the real meaning intended to be conveyed by the writer to the person to whom he is writing. Thus, to quote but one of thousands of examples of communications containing a secret or hidden meaning, prepared with the intention of escaping suppression by censors in war times, the sentence "Package sent today" may mean "Three transports left today." Communications of this type, although occasionally useful in espionage and counter-espionage, are entirely impractical for field military use, and for this reason they will not be dealt with further in this text.

c. Visible writing which conveys no intelligible meaning in any language, or which apparently conveys an intelligible meaning that is not the real meaning intended to be conveyed, is said to be in *secret language*. The text of such writing constitutes *secret text* or *cryptographic text*.

3. Message, cryptogram, correspondents, enemy.—*a.* The term "*message,*" as used in this text, applies to all communications in visible writing, whether in plain language or in secret language, transmitted by any of the agencies of signal communication, or carried by postal agencies.

b. A message in plain language will herein be called a *plain-language message*, or a *plain-text message*, or a *message in clear*.

c. The term "*cryptogram*" as used in this text applies only to communications in visible writing in secret language. These also may be transmitted by any of the agencies of signal communication, or carried by postal agencies, but the agencies of signal communication principally employed to transmit cryptograms are electrical in character, viz., radio, telegraph, and telephone.

d. As employed herein the term "*correspondents*" will be applied to designate persons who communicate with one another by the exchange of messages. The *originator* of the message is the person who drafts the plain text; the *addressee* is the person for whom the plain text is intended. Between the originator and the addressee there may be persons who handle the message, who convert its plain text into a cryptogram, or who reconvert the cryptogram into plain text.

The originator and the addressee may also serve in this capacity but in our Army this is not usually the case, this work being performed by special personnel who act as agents of the correspondents.

e. The term "*enemy*" will be applied in this text to designate all persons who obtain messages or copies of messages not intended for them.

4. Cryptographing, decryptographing, and cryptography.—*a.* To *cryptograph* is to convert a plain-text message into a cryptogram by following certain rules mutually agreed upon in advance by the correspondents, or furnished them or their agents by higher authority.

b. To *decryptograph* is to reconvert a cryptogram into the equivalent plain-text message *by a direct reversal of the cryptographing process.*

c. A person skilled in the art of cryptographing and decryptographing is called a *cryptographer,* and a clerk who cryptographs and decryptographs, or who assists in such work, is called a *cryptographic clerk.*

d. Cryptography is that branch of cryptology which treats of the various means, methods, and devices for converting plain-text messages into cryptograms and reconverting the so-produced cryptograms into their plain-text form by a direct reversal of the steps or processes employed in the original conversion.

e. The noun *cryptograph* is restricted in its usage in this text to apply only to a mechanical (usually hand-operated) device or instrument employed in cryptographing or decryptographing. In other words, a cryptograph is a cipher device. (Compare the terms "*cryptography,*" "*cryptogram*" and "*cryptograph*" with the terms "*telegraphy,*" "*telegram,*" "*telegraph.*")

5. Codes, ciphers, and enciphered code.—*a.* Cryptographing and decryptographing are accomplished by means collectively designated as *codes and ciphers.* They are used for a two-fold purpose: (1) secrecy and (2) economy. The former is far more important than the latter in military cryptography. The specialized meanings of the terms code and cipher will be explained in detail later on; it will be sufficient at this point to indicate that, broadly speaking, in ciphers or *cipher systems* cryptograms are produced by applying the cryptographic treatment to *individual letters* of the plain-text messages, whereas in codes or *code systems* cryptograms are produced by applying the cryptographic treatment to entire words, phrases, and sentences of the plain-text messages.

b. A cryptogram which has been produced by means of a cipher system is said to be in cipher and is called a cipher message, or sometimes, simply a *cipher.* The text of the cryptogram is referred to as *cipher text.* The cryptographing process in this case is called encipher-

ing, and the enciphered version of the plain text is often referred to as its *encipherment*. The cryptographic clerk who performs the process serves as an *encipherer*. The corresponding terms applicable to the decryptographing process in the case of cipher systems are "*deciphering*," "*decipherment*," and "*decipherer*." When a cryptographic clerk serves both as an encipherer and decipherer, he is called a *cipher clerk*.

c. A cryptogram which has been produced by means of a code system is said to be *in code* and is called a *code message*, or sometimes simply a *code*. The text of the cryptogram is referred to as *code text*. The cryptographing process in this case is called *encoding*, and the encoded version of the plain-text message is often referred to as its *encodement*. The cryptographic clerk who performs the process serves as an *encoder*. The corresponding terms applicable to the decryptographing process in the case of code systems are "*decoding*, *decodement*," and "*decoder*." When a cryptographic clerk serves both as an encoder and decoder, he is called a *code clerk*. In our Army, code methods of secret communication predominate, and for this reason Tables of Organization refer to the cryptographic clerks as code clerks, though they may occasionally be called upon to encipher and decipher messages.

d. Sometimes, for special purposes, the code text of a cryptogram subsequently undergoes encipherment, producing what is called a cryptogram in *enciphered code*, or an *enciphered-code message*. *Encoded cipher*, that is, where the cipher text of a cryptogram subsequently undergoes encodement, is also possible but rare.

6. General system and specific key.—*a*. The sum total of all the basic, invariable rules to be followed in cryptographing a message according to a given method, together with all the agreements, conventions or private understandings drawn up between the correspondents or their authorized agents, or furnished them by higher authority, constitutes what is herein termed the "*general cryptographic system*."

b. Usually the general cryptographic system operates in connection with a number, word, phrase, or sentence which controls or directs the details of the steps to be followed in cryptographing a message. This element, usually of a variable nature and easily changeable at the will of the correspondents, or prearranged for them or for their agents by higher authority, constitutes the *specific key*. The specific key may also consist of a set of specially prepared tables, a special document, or it may even be a book; a single letter may be used.

c. Hereafter, when an ambiguity can arise, the general cryptographic system will be referred to more briefly as the *system*; the specific key, as the *key*.

7. Cryptanalytics and cryptanalysis.—*a.* It may be stated that, as a general rule, all or nearly all cryptographic systems suitable for *practical* use can be *broken down*, or *solved*,—that is, properly prepared cryptograms can be "translated" or read without a knowledge or possession of the general cryptographic system and the specific key applying to the cryptograms.

b. That branch of cryptology which deals with the principles, methods, and means employed in the solution or *analysis* of cryptograms is called *cryptanalytics*.

c. The steps and operations performed in applying the principles of cryptanalytics constitute *cryptanalysis*. To cryptanalyze or to *decrypt* a cryptogram is to solve it by cryptanalysis.

d. A person skilled in the art of cryptanalysis is called a *cryptanalyst*, and a clerk who assists in such work is called a *cryptanalytic clerk*.

SECTION II

RELATED INFORMATION OF AN INTRODUCTORY NATURE

8. Interception, radiogoniometry, systems of secret signaling.—*a.* It is a well-known fact that communications transmitted by electrical means can, under certain circumstances, also be heard and copied by persons who are not the correspondents or their authorized agents. Communications transmitted by radio can be manually copied or automatically recorded by suitably adjusted radio apparatus within listening range. Except in special cases, communications transmitted over wire lines can likewise be manually copied or automatically recorded by special apparatus suited for the purpose. In the case of radio traffic, the correspondents have no way of knowing whether or not their traffic is being copied by the enemy, since the unauthorized copying does not interfere in the slightest degree with the signals being transmitted. In the case of wire traffic, there are methods of listening-in and copying in which there is either no interference or so little interference with the transmitted signals that the correspondents can never be certain whether or not their traffic is being copied by the enemy. The general term applied to the act of

listening-in and copying or recording electrically transmitted communications by persons other than the correspondents or their authorized agents is called *"interception."* In time of war, it must be assumed that the enemy will make every effort to intercept all traffic possible, and that he will usually be in a position to intercept all communications transmitted by agencies susceptible of interception.

b. A less well-known fact concerning radio transmission is that it is possible to determine with a fair degree of accuracy the geographic location of a station that is emitting radio waves. The science which deals with the means and methods of locating a radio transmitting station by taking bearings on the waves emitted by the station is called *radiogoniometry.* In war time, much valuable information about the enemy can often be obtained merely by determining the locations of his radio transmitting stations.

c. Information may be conveyed from one person to another by signals which are made by means of special apparatus constructed with a view to distorting, disguising, or completely hiding the signals themselves so that third parties will have great difficulty in intercepting and recording the signals or may not even be aware of their existence. All such methods of transmitting intelligence fall in the class herein designated as *systems* of *secret signaling.*

d. It may be stated that, as a general rule, the signals of all or practically all systems of secret signaling can be intercepted and recorded in a form suitable to making the signals apprehensible by one of the senses, usually visual or auditory, without possession of the specific apparatus employed in their formation or emission, or of the "key" employed in their distortion, or disguise.

9. Information derivable from intercept, radiogoniometric, and cryptanalytic activities.—*a.* In addition to (1) the intercept and radiogoniometric means of obtaining enemy communications and information relating to enemy communications, further data may be obtained by (2) the employment of secret agents engaged in espionage activities; (3) the capture of prisoners, messengers, and homing pigeons; (4) the capture of headquarters or command posts in the theater of operations; and (5) treason or carelessness of personnel entrusted with the handling of communications. These five sources are listed in their relative order of importance. The first is by far the most prolific source of enemy communications or of information relating to them.

b. The amount of important information that can be obtained from the efficient employment of intercept, radiogoniometric, and cryptanalytic activities cannot be accurately estimated. It fluctuates with time, place, and personnel. It is obvious that if all enemy transmitting stations can be located quickly and if all his communications

113477°—35——2

can be intercepted and solved, extremely valuable information concerning his strength, disposition of forces, state of morale of his troops, and his intentions may be continually available to our own forces.

10. Allocation of code and cipher work to different agencies of the Military Establishment.—*a.* The coordination, control, and supervision of all work connected with codes and ciphers in our Army is a function of the Intelligence (G–2) Division of the General Staff of the headquarters at which such work is conducted.

b. As regards operations, code and cipher work in the Army is of two sorts: (1) That connected with our own communications, and (2) that connected with enemy communications.

c. In the former there are two agencies concerned:

(1) The Signal Corps is charged with preparing, publishing, revising, storing, accounting for and distributing all codes and ciphers, cipher tables, and instructions pertaining to them, and with the development, production, storage, and issue of cipher devices.

(2) The actual handling of codes and ciphers in the organizations to which they are distributed for use in cryptographing and decryptographing messages is one of the functions of certain agencies called *message centers*. In organizations down to and including division, the message centers are operated by Signals Corps personnel; in those below division, they are operated by personnel assigned by the commanders of the organizations to serve under the unit signal officer.

d. In addition to the duties named in paragraph *c* (1) above, and as regards enemy communications, the Signal Corps is charged in war time with—

(1) Location, by radiogoniometric means, enemy transmitting stations.

(2) Intercepting by electrical means, enemy communications.

(3) Solving all enemy code and cipher messages thus intercepted or forwarded to it for solution by authorized agencies.

(4) Establishing and operating laboratories for the employment and detection of secret inks.

e. The results obtained by the Signal Corps from the sources given under *d* above are forwarded directly to the G–2 Division of the staff served by the cryptanalytic section concerned. The latter evaluates this information and distributes it to those concerned or interested.

11. Discussion with reference to the time required for cryptanalysis; the factors upon which the latter depends.—*a.* Assuming that certain essential conditions to be set forth later are complied with, the cryptanalysis of any given cryptographic system is merely a matter of time. In military operations it is especially true that "time is of the essence", and therefore a brief discussion of the factors upon which the length of time required for the complete

analysis of a given cryptographic system is essential as a preliminary to a study of systems suitable for employment in military communications.

b. The influence or effect that the analysis of military cryptograms will have on the tactical situation depends upon the sum total of several factors, of which the following are most important:

(1) The length of time necessary to transmit the enemy cryptograms to the solving headquarters. This is usually not a negligible factor unless signal communication agencies are properly organized to perform this function.

(2) The length of time required to solve the cryptograms, including that required in making copies, tabulating, and recording the data, etc.

(3) The nature of the information disclosed by the solved cryptograms, whether it is of immediate importance in connection with an impending attack or an action in the near future, or whether it is only of historical interest in connection with an action in the past.

(4) The length of time necessary to transmit the thus derived information to the G–2 Division of the General Staff, and for the latter to evaluate the information.

(5) The length of time necessary for G–2 to transmit the resulting assimilated and correlated information to the Operations (G–3) Division of the General Staff, and for the latter to prepare the orders for the action determined by the information thus obtained and to transmit them to the combat units concerned. The final sentence under subparagraph (1) applies here also.

c. Of the five factors mentioned in the preceding subparagraph, the only one which is of direct interest to us in our study is the second—namely, the length of time required to solve the cryptograms. It is subject to great variation and is dependent upon several factors, of which the following are the most important:

(1) The *degree of cryptographic security* inherent in the cryptographic system itself, by which is meant the resistance it offers to, or the obstacles it places in the way of cryptanalysis. Cryptographic systems vary more widely in this respect than is commonly supposed, but a demonstration of this variation by means of actual examples of cryptanalysis falls outside the scope of this text.

(2) The amount or volume of cryptographic text available for study. Other things being equal, the greater the volume of text, the more easily and speedily can it be solved. A single cryptogram in a given system may present an almost hopeless task for the cryptanalyst, but if several or many cryptograms belonging to the same cryptographic system are available for study, the solution may be reached in an astonishingly short time.

(3) The number, skill, and efficiency of organization and cooperation of the cryptanalytic personnel assigned to the work. It is

sufficient here merely to indicate that in order to avoid duplication of effort and, especially in forward areas where the solved information is most useful, to make possible the quick interpretation of cryptograms in already solved systems, cryptanalytic headquarters are organized in units of ascending size, the forward echelons consisting perhaps of only four or five persons, while the rear echelons may consist of as many as 100 persons. In all these units proper organization of highly skilled workers is absolutely essential for efficient operation.

(4) The amount and character of collateral intelligence made available to the cryptanalytic headquarters. Isolated cryptograms exchanged between a very restricted, small group of correspondents, about whom and about whose business no information is available, may resist the efforts of even a highly organized, skilled cryptanalytic office for a long time, or even indefinitely. If however a certain amount of information in these respects is obtained, the situation may be entirely changed. In military operations usually a great deal of general information is available. In many cases a fair amount of definite information concerning the contents of specific cryptograms is at hand, such as proper names of persons and places, events in the immediate past or future, etc. In such cases, the solution of the cryptograms is greatly facilitated.

12. Degree of cryptographic security required of a system for military use.—The ideal cryptographic system for military purposes would be one which is practicable for use not only in the zone of the interior and in the zone of communications but also in the theater of operations, and which presents such a great degree of cryptographic security that no matter how much traffic becomes available, all in the same key, the cryptograms composing this traffic should resist solution indefinitely. This, however, lies outside the realm of possibility so far as our present methods are concerned. *The best that can be expected is that the degree of security should be great enough to delay solution by the enemy for such a length of time that when the solution is finally accomplished the information thus obtained has lost its immediate value.*

13. Fundamental practical requirements of a cryptographic system for military use.—*a.* Before proceeding to an exposition of various types of cryptographic systems, it is necessary to indicate briefly certain fundamental requirements of a practical nature which military cryptograms must meet because of definite limiting conditions in the present state of the art of military signal communication. Some of these requirements are obvious but they are often overlooked by persons who set out to devise new cryptographic systems for military use.

(1) Cryptograms must be in a form suitable for the most economical transmission by Morse telegraphy. In the first place it automatically

eliminates all cryptograms containing symbols or characters for which there exist no equivalents or signals in the two most important Morse telegraphic alphabets: American Morse, and International Morse. Although there are signals in these alphabets for certain signs of punctuation, such as . , : ; etc., if these are inserted in telegrams they are either spelled out in words by the operator or, if their transmission as symbols is insisted upon, they are heavily taxed. Thus, this requirement eliminates all systems except those which produce cryptograms composed *exclusively* of the 26 letters of the alphabet, or exclusively of the 10 arabic digits. Cryptographic systems employing arabic numerals are not as desirable as those employing letters because the Morse signals for numbers are longer and are handled with greater difficulty by the average telegraph or radio operator than are letters. When the cryptograms must be transmitted by telephone, as is sometimes the case in the theater of operations, they are however better than systems employing letters. Systems which produce cryptograms composed of mixtures of letters and figures, or of letters, figures, punctuation signs, and the like, are entirely unsuited for practical usage. Further, in order to be suitable for economical transmission, the cryptographic text must be capable of being arranged in regular sets of characters for the reason: first, that it promotes accuracy in telegraphic transmission (since an operator knows that he must receive a definite number of characters in each group, no more and no less); and second, for the reason that cryptanalysis is usually made more difficult when the lengths of the words, phrases, and sentences of the plain text are no longer apparent. The usual grouping is in sets of five characters, though occasionally other groupings may be encountered under special circumstances.

(2) The regular channels of signal communication can only carry a limited volume of traffic. Their most efficient operation demands that, other things being equal, the smallest number of characters actually necessary to convey a given amount of intelligence be transmitted. Therefore, in the case of a system for field use, the cryptographic text should be no longer than the equivalent clear text. In the case of systems for other than field use, the cryptographic text may be somewhat longer than the equivalent clear text, but a system in which the cryptographic text is twice the length of the equivalent clear text can be deemed useful only if it is of outstanding merit otherwise and is suitable for certain restricted or special employment. No system in which the cryptographic text is more than twice the length of the equivalent clear text can be considered practicable for military usage.

(3) The operations of cryptographing and decryptographing must be relatively simple and rapid. In our Army these are performed for the most part by enlisted men, and under difficult conditions. Therefore, they must not require the remembering and application of a long

series of steps or rules. They must be such as to reduce to a minimum the mental strain upon the operator. As a rule, complex processes involving several distinct steps are not suited to the conditions in the combat zone, but occasionally systems involving only two steps, provided each is simple and rapid, may be practicable for military usage. In the case of a system for field use, it should require the services of only one operator either to cryptograph or to decryptograph messages.

(4) Cipher devices or instruments for field use must be light in weight, rugged in construction, and simple in operation, requiring the services of only one operator.

(5) The system must be such that errors, which invariably occur in cryptographic communication, can be corrected easily and rapidly by cryptograph clerks without the necessity of calling for a repetition of the whole transmission, or for a rechecking of the original cryptographing.

b. There are other requirements in connection with the degree of cryptographic security that must be fulfilled. These will be discussed later but only very briefly.

c. In the systems to be set forth in this text, only such as fulfill the foregoing practical requirements can be included.

SECTION III

THE TWO CLASSES OF CRYPTOGRAPHIC SYSTEMS

14. Transposition and substitution.—*a.* Technically there are two and only two distinctly different types of treatment which may be applied to plain text to convert it into secret text, yielding two different *classes* of cryptograms. In the first, called *transposition*, the *elements* or *units* of the plain text, whether one is dealing with individual letters or groups of letters, syllables, whole words, phrases and sentences, retain their original identities but merely undergo some change in their relative positions or sequences so that the message becomes unintelligible. In the second, called *substitution*, the elements of the plain text retain their original positions or sequences but are replaced by other elements with different values or meanings.

b. It is possible to cryptograph a message by a substitution method and then to apply a transposition method to the substitution text, or vice versa. Such combined transposition-substitution methods do not form a third category of methods. They are occasionally encountered in military cryptography, but the types of combinations that are sufficiently simple to be practicable for field use are very restricted.

15. Letter, syllable, and word methods.—Under each of the two principal classes of cryptograms as outlined in the preceding paragraph, a further classification can be made with respect to the nature of the *textual elements* or *units* with which the cryptographic process deals. Broadly speaking these textual units are (1) individual letters, or groups of letters in regular sets, and (2) complete words. Methods which deal with the first type of units are called *letter methods*, including, when such is the case, *syllable methods;* those which deal with the second type of units are called *word methods*.

16. Cipher systems and code systems.—It is necessary to indicate that the latter classification into methods is more or less abitrary or artificial in nature, and is established for purpose of convenience only. No sharp line of demarcation can be drawn in every case, for occasionally a given system may combine methods of treating single letters, groups of letters, syllables, whole words, phrases and sentences. Whenever in a single system the general method is such that the cryptographic treatment is as a general rule applied to textual units of regular length, usually single letters or pairs, and is only exceptionally applied to textual units of irregular length, the system will be designated as a *cipher system*. Likewise, whenever in a single system the general method is such that the cryptographic treatment is as a general rule applied to textual units of irregular length, usually whole words, phrases and sentences, and is only exceptionally applied to single letters, pairs, or groups of letters and syllables, the method will be designated as a *code system*.

17. Scope of subject matter to be covered.—In this text a few typical examples of cipher systems and code systems will be presented; the procedure in cryptographing and decryptographing by their means will be shown in detail; methods of preparing keys suitable for use in connection with them will be illustrated; errors and their correction will be discussed; and, finally, a few of the most important precautions that should be observed in order to safeguard the systems and the cryptograms from enemy cryptanalysts will be set forth. In all this only such considerations as apply to military cryptography will be included.

PART TWO

CIPHER SYSTEMS

A. Transportation systems

SECTION IV

SIMPLE MONOLITERAL TRANSPOSITION METHODS

18. Transposition ciphers in general.—Transposition ciphers are roughly analogous to "jig-saw puzzles" in that all the pieces of which the whole original is composed are present but are merely disarranged. The pieces into which the picture forming the basis of a jig-saw puzzle may be divided are irregular in size and shape, but the pieces into which the plain text forming the basis of a transposition cipher may be divided must be much more regular in these respects, for the sake of practicability. They must be either single letters or pairs of letters or sets of letters in regular groupings or finally, in an exceptional case, whole words. The majority of transposition methods however, deal with individual letters and are therefore termed "*monoliteral methods.*" The other methods are termed "*polyliteral methods.*"

19. Geometric designs.—*a.* Practically all monoliteral or polyliteral transposition ciphers involve the use of a design or geometric figure, such as a square, rectangle, triangle, trapezoid, etc., in which the letters of the plain text are first *inscribed* or written according to a previously agreed-upon direction of writing and then *transcribed* or rewritten according to another and different, previously agreed-upon direction to form the text of the cryptogram. In nearly all cases the specific key consists in (1) employing designs of a specific nature and dimensions, and (2) varying the direction or manner of inscription or transcription, or both.

b. In working with transposition ciphers or, for that matter, most types of ciphers, cross-section paper will be found very convenient. Cross-section paper with ¼-inch squares is most suitable. For brevity

in reference, the individual small squares of such cross-section paper will hereafter be called *cells*.

20. Route transpositions.—*a*. Suppose the correspondents agree to use the method of monoliteral transposition known as *route transposition*. The message is inscribed within a rectangle in the usual manner of writing, i. e., from left to right and in consecutive lines from top to bottom. If one or more cells remain vacant at the end, *nulls* or *dummy* letters — letters having no significance—are inserted as "fillers" so as to complete the rectangle. Then, to form the cipher text, the letters in the design are taken out of the design and rewritten or transcribed by following or tracing one of many different routes. It is possible for each route to have a different starting point, and normally it is one of the four corners of the rectangle. A few typical routes are illustrated in Figure 1 where, for the sake of ease in following the route, the plain-text message is assumed to be merely the sequence of letters A B C ... X.

(A) Simple horizontal:

(1)	(2)	(3)	(4)
ABCDEF	FEDCBA	STUVWX	XWVUTS
GHIJKL	LKJIHG	MNOPQR	RQPONM
MNOPQR	RQPONM	GHIJKL	LKJIHG
STUVWX	XWVUTS	ABCDEF	FEDCBA

(B) Simple vertical:

(1)	(2)	(3)	(4)
AEIMQU	UQMIEA	DHLPTX	XTPLHD
BFJNRV	VRNJFB	CGKOSW	WSOKGC
CGKOSW	WSOKGC	BFJNRV	VRNJFB
DHLPTX	XTPLHD	AEIMQU	UQMIEA

(C) Alternate horizontal:

(1)	(2)	(3)	(4)
ABCDEF	FEDCBA	XWVUTS	STUVWX
LKJIHG	GHIJKL	MNOPQR	RQPONM
MNOPQR	RQPONM	LKJIHG	GHIJKL
XWVUTS	STUVWX	ABCDEF	FEDCBA

(D) Alternate vertical:

(1)	(2)	(3)	(4)
AHIPQX	XQPIHA	DELMTU	UTMLED
BGJORW	WROJGB	CFKNSV	VSNKFC
CFKNSV	VSNKFC	BGJORW	WROJGB
DELMTU	UTMLED	AHIPQX	XQPIHA

(E) Simple diagonal:

(1)	(2)	(3)	(4)
ABDGKO	OKGDBA	GKOSVX	XVSOKG
CEHLPS	SPLHEC	DHLPTW	WTPLHD
FIMQTV	VTQMIF	BEIMQU	UQMIEB
JNRUWX	XWURNJ	ACFJNR	RNJFCA

(5)	(6)	(7)	(8)
ACFJNR	RNJFCA	JNRUWX	XWURNJ
BEIMQU	UQMIEB	FIMQTV	VTQMIF
DHLPTW	WTPLHD	CEHLPS	SPLHEC
GKOSVX	XVSOKG	ABDGKO	OKGDBA

(F) Alternate diagonal:

(1)	(2)	(3)	(4)
ABFGNO	ONGFBA	GNOUVX	XVUONG
CEHMPU	UPMHEC	FHMPTW	WTPMHF
DILQTV	VTQLID	BEILQS	SQLIEB
JKRSWX	XWSRKJ	ACDJKR	RKJDCA

(5)	(6)	(7)	(8)
ACDJKR	RKJDCA	JKRSWX	XWSRKJ
BEILQS	SQLIEB	DILQTV	VTQLID
FHMPTW	WTPMHF	CEHMPU	UPMHEC
GNOUVX	XVUONG	ABFGNO	ONGFBA

(G) Spiral, clockwise:

(1)	(2)	(3)	(4)
ABCDEF	LMNOPA	DEFGHI	IJKLMN
PQRSTG	KVWXQB	CRSTUJ	HUVWXO
OXWVUH	JUTSRC	BQXWVK	GTSRQP
NMLKJI	IHGFED	APONML	FEDCBA

(H) Spiral, counterclockwise:

(1)	(2)	(3)	(4)
APONML	FEDCBA	NMLKJI	IHGFED
BQXWVK	GTSRQP	OXWVUH	JUTSRC
CRSTUJ	HUVWXO	PQRSTG	KVWXQB
DEFGHI	IJKLMN	ABCDEF	LMNOPA

FIGURE 1.

b. It is apparent that instead of following the normal direction of writing, i. e., from left to right and from the top downwards, the letters of the plain text may be inscribed according to any one of the routes agreed upon, and then transcribed to form the cipher text by taking the letters from the rectangle in the normal manner, i. e., in this case from left to right, and from the top downwards, or *by following any other route of transposition.*

21. Example of encipherment and decipherment by monoliteral route transposition.—*a.* Let us now take a special example of encipherment by monoliteral route transposition. We will use the message ATTACK HAS BEEN POSTPONED UNTIL TOMORROW TWO AM, and employ a relatively complicated method. Suppose that the general system agreed upon is the one being described, and that the specific key consists of the following elements:

(1) Using a completely filled rectangle of seven columns;

(2) Inscribing the letters of the plain text within the rectangle by following route (F) (3) of Figure 1;

(3) Transcribing the thus inscribed letters (to form the cipher text) by following route (E) (6) of Figure 1.

Since the message contains a total of 40 letters, and it has been agreed to use a completely filled rectangle of seven columns, it is necessary to add two nulls to make the total number of letters a multiple of seven. A rectangle of seven columns of cells and six lines of cells is therefore prepared. The design is then filled in as shown in Figure 2.

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

Cryptogram:

TMLAO WROWT ROMOT DUNTI LENOP
TSEEN POBSA HACKT TA

FIGURE 2.

b. To decryptograph such a cryptogram the process is merely reversed. First, the total number of letters in the cipher text must be found. Since it is 42, and since a completely filled rectangle of seven columns has been agreed upon, a design consisting of seven columns and six rows is outlined on cross-section paper. The cipher text is then inscribed according to route (E) (6) of Figure 1, and after this has been completed the plain-text letters are read according to route (F) (3) of Figure 1. It is apparent that it is necessary to remember a relatively long series of rules, and even when the cryptographing

has all been accomplished correctly the degree of security is very low. Note how obviously the whole word UNTIL manifests itself in the cipher text. Parts of other words can also be seen. Despite the rather extended variability that this system affords as regards the dimensions of the rectangle, the method of inscription and transcription and their starting points, the degree of security remains very low.

22. Use of nulls in transposition.—*a.* It will be noted that the two nulls selected as fillers to complete the rectangle in the preceding example were the letters L and T. These were chosen rather than such letters as J, K, Q, X, or Z, for a reason which is important to note. Since transposition ciphers of this type involve merely a rearrangement of the letters, without any change whatever in their identities, it follows that the natural or normal frequencies of letters of plain text remain unchanged. Now, the letters of every alphabetic language have characteristic frequencies, as a result of which certain clues are afforded in cryptanalysis. The presence, in transposition ciphers, of letters of very low frequency (in English), such as J, K, Q, X, or Z, is very unusual and therefore if these are employed merely as fillers they may afford clues as to the real number of letters in the plain text, the starting or finishing points of the real text, etc. For this reason it is best to insert as fillers in transposition ciphers letters of medium or high frequency, such as E, T, R, I, N, O, A, S, D, L, or C, for these will not afford any clues to solution. Nulls, when employed for the purpose of making cryptanalysis more difficult, may also be inserted in specific positions as prearranged, or they may be inserted at random if the system permits. This is true of other cryptographic systems, but as a general rule the use of nulls, especially in cipher systems, is to be discouraged. Very often they add little if any security, and thus merely increase the length of the cryptographic text without any compensating advantages.

b. Whenever it is necessary to add nulls in order to complete a transposition message in any respect, or for any reason whatsoever, they must be added *before* the transposition process is applied and not afterward, otherwise the decryptographing clerk will have great difficulty in reading the message, if the possibility is not wholly destroyed so far as he is concerned. This applies especially to the case where the service regulations require that the final group in a cryptogram be a complete group, containing exactly as many letters as all other groups in the message.

23. Special cases of route transposition.—*a.* The oldest and simplest transposition method known, that called reversed writing, is a special case of one of the routes shown in Figure 1. Here the text is written in the opposite direction from the normal; for example, BRIDGE DESTROYED is written EGDIRB DEYORTSED.

The variability of the scheme, that is, the *specific key*, consists in the fact that the reversal may be applied to groups of fixed length, to whole words, to sentences, or to the whole text. The security of simple reversed writing may be somewhat increased by disguising the original word lengths, by which is meant a destruction of the normal, or natural word limits by combining a part of one word with a part of the next to form either *false words* or groups of regular length.

b. Some examples of reversed writing follow. Let the message be: BRIDGE DESTROYED AT ELEVEN PM.

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

Cipher: EGDIRB DEYORTSED TA NEVELE MP

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

Cipher: EG DIRB DEYORT SEDTA NEVE LEMP

(3) Reversing the whole text and regrouping into fives:

Cipher: MPNEV ELETA DEYOR TSEDE GDIRB

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

Cipher: MPNER VELEO TADEB YORTH SEDEA GDIRB

c. A second very simple type of transposition, that known as *vertical writing*, is a special case of another of the routes shown in Figure 1.

The message BRIDGE DESTROYED is written in two vertical columns, as shown in Figure 3, and the cipher text is taken from the horizontal pairs thus formed. The message becomes:

BS
RT
IR
DO
GY
EE
DD
E

FIGURE 3.

BSRTI RDOGY EEDDE

When the plain text is inscribed in pairs of letters in vertical writing and then the cipher text is taken by transcribing the columns, a slightly different result is obtained. This is shown in Figure 4, using the plain text message BRIDGE DESTROYED. The cipher becomes:

BR
ID
GE
DE
ST
RO
YE
D

FIGURE 4.

BIGDS RYDRD EETOE

This type of transposition is sometimes called the *rail-fence* cipher because it can be produced by writing the message in the following form:

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

which yields the same cipher result as before.

24. Remarks on monoliteral route transposition.—Reversed writing and vertical writing of the types indicated yield extremely simple cryptograms. In practice they are sometimes used in connection with other more or less simple cryptographing methods to increase their security. The cryptographic security of the other methods thus far indicated is also very low, despite the apparently large degree of variability they afford. The reason is that the route to be followed in the inscription or transcription process is definitely fixed under each type of route. In other types of transposition soon to be discussed, a much wider latitude for variation in the route is afforded by the use of key words to control or to guide these processes. Geometric designs are also used in these types of transposition, and key words determine the dimensions of the design, or else, in case only one key word is used, it determines one dimension, the other being determined by the length of the text. Examples to be given in their proper place will serve to illustrate the processes.

25. Key words and numerical keys.—*a.* It is often necessary, in performing certain cryptographic operations, to employ a *numerical key*, which may consist of a relatively long sequence of numbers difficult or impossible for the average cipher clerk to memorize. In order to avoid making it necessary that such sequences of numbers be carried about on the person in the form of written memoranda, a procedure which would often be dangerous, cryptographers have devised very simple methods of deriving such sequences from words, phrases, or sentences, which can usually be remembered much more easily than can unintelligible, relatively long sequences of numbers. One of the simplest methods is to assign numerical values to the letters of the key in accordance with their relative positions in the ordinary alphabet. Such a key is called a *derived numerical key*. This method of assigning the numbers is very flexible and varies with different uses to which numerical keys are put. For purposes of transposition, the method shown below is very satisfactory.

b. Let the prearranged key word be the word CARBUNCLE. Since the word contains the letter A, which comes first in the alphabet, the number 1 is written under this letter in the key word. Thus:

C A R B U N C L E
1

The next letter of the normal alphabet that occurs in the key word is B, which is assigned the number 2. The letter C, which occurs twice in the key word, is assigned the number 3 for its first occurrence, the number 4 for its second occurrence, and so on. The final result is:

Basic key word:—C–A–R–B–U–N–C–L–E
Derived numerical key:—3–1–8–2–9–7–4–6–5

c. The method may, of course, be applied to phrases or to sentences, so that a very long numerical key, impossible ordinarily to remember, may be derived at will from an easily remembered *key text*.

d. It is advisable to make note of a few points valuable in connection with the choice of a key text:

(1) It should be such as can be easily remembered. Often a key composed of two or more short words is better than one consisting of a single long word. Thus, the whole sentence WHEN DO WE EAT would be better than the single word EXTRAORDINARY.

(2) It should consist of one or more *simple*, familiar words admitting of but *one* spelling. A word such as REINFORCEMENT is inadvisable because the spelling REENFORCEMENT is also admissible. It goes almost without saying that words such as form good material for "spelling bees", even though they may be familiar, everyday words as, for example, DEFINITELY, SEPARATELY, REPETITION, etc., are likewise inadvisable.

(3) It should contain, as a rule, as many different letters as possible, in no systematic sequence. Words with several repeated letters, such as ELEMENT, BANANA, MISSISSIPPI, etc., form poor key words.

(4) It should present no associations with the special situation under which it is employed, so as not to be easily guessed by the enemy. For example, to use personal or geographic names associated with a region in the theater of operations is bad practice. The key word GETTYSBURG employed in a cryptogram originating in the vicinity of Gettysburg would be bad practice. Or to use for this purpose words of common military usage, such as BATTALION, REGIMENT, ARTILLERY, SIGNAL CORPS, MACHINE GUN, etc., is likewise bad practice.

e. It is convenient to designate key text in letters as a *literal key*. As noted above, a literal key may consist of a single letter, a single word, a phrase, a sentence, a whole paragraph or even a book. The method of deriving a numerical key from a literal key given in subparagraph *b* above is only one of a number of methods, but it is the most commonly employed. It is also subject to variation in detail. But, so far as the cryptanalyst is concerned, just how the numerical key is derived from a specific literal key is usually of interest to him only if this knowledge will assist in subsequent solutions of cryptograms prepared according to the same basic system. Often the cryptanalyst is wholly unconcerned as to whether a literal or a numerical key has been employed in connection with the cryptographing of the messages, and he may frequently be unaware of the fact that a literal key has been employed as the basis for deriving a numerical key.

SECTION V

COLUMNAR TRANSPOSITION METHODS

26. Columnar transposition with completely filled rectangles.—*a.* One of the most common types of transposition involving the use of a key word or a derived numerical key is that known as *keyed* or *variable-key columnar transposition.* In this type the letters are usually written in a geometric design, most often a rectangle, by inscribing them in the ordinary manner, i. e., in horizontal lines from left to right and from the top downwards, and then the letters are transcribed by "reading" the columns in the sequence determined by the numerical key. If the text does not contain a sufficient number of letters to fill the last line completely, as many nulls as are necessary to do so are added at the end. Figure 5 shown below is an example of cryptographing by this method.

Key word—

L–I–B–E–R–T–Y

Numerical key—

4	3	1	2	5	6	7
R	E	P	O	R	T	L
O	C	A	T	I	O	N
O	F	S	E	C	O	N
D	B	A	T	T	A	L
I	O	N	C	O	M	M
A	N	D	P	O	S	T
T	O	D	A	Y	D	N

NOTE.—The letters D and N in the final two cells are nulls, inserted to complete the rectangle.

Cryptogram:

PASAN DDOTE TCPAE CFBON OROOD IATRI CTOOY
TOOAM SDLNN LMTN

FIGURE 5.

b. To decryptograph such a cryptogram, a rectangle with the proper number of cells, as determined by the length of the message and the length of the key, must first be prepared. In the foregoing example, since the cipher text consists of 49 letters and the key consists of 7 letters or numbers, the rectangle shown in (a) of Figure 6 is prepared and then the columns (of cells) are filled in numerical

order. An early stage in the decrytographing is represented in (b) of the figure. It is only after the process has been finished that the complete message reappears, as shown in (c) of the figure.

Cryptogram:

PASAN DDOTE TCPAE CFBON OROOD IATRI CTOOY TOOAM
SDLNN LMTN

4 – 3 – 1 – 2 – 5 – 6 – 7

(a)

4 – 3 – 1 – 2 – 5 – 6 – 7

		P				
		A				
		S				
		A				
		N				
		D				
		D				

(b)

4 – 3 – 1 – 2 – 5 – 6 – 7

R	E	P	O	R	T	L
O	C	A	T	I	O	N
O	F	S	E	C	O	N
D	B	A	T	T	A	L
I	O	N	C	O	M	M
A	N	D	P	O	S	T
T	O	D	A	Y	D*	N*

(c)

FIGURE 6.

*The letters D and N are recognized as nulls.

c. The method indicated above is susceptible of considerable variation, consisting in (1) changing the key word, and (2) changing the direction of inscribing the letters of the plain text, or in transcribing them to form the cipher text. As to the first factor, little need be said. A daily change in key, or oftener, is possible; or, by drawing up a whole list of daily keys for a given period, automatic change in key can be provided for without the necessity of giving any indication in the cryptograms as to the key applicable. It is also possible to prepare a long list of suitable keys and to designate each key by an *indicator* which is inserted in the cryptogram in a prearranged position. Indicators may consist of words, numbers, groups of letters, or single letters. For example, each key in a list of 500 may be indicated by a single pair of letters which may be inserted at the beginning of the cryptogram, at the end, or at any prearranged position. This procedure has the disadvantage, however, that if an error occurs at the particular position of the cryptogram containing the indicator, the decryptographing is made difficult if not impossible. For this reason indicators, if used, are often inserted in at least two different positions in the cryptogram, usually at or near the beginning and end.

d. As to the second factor, the letters of the plain text may be inscribed in the rectangle according to any one of the routes indicated in Figure 1. So long as the transcribing process is accomplished by reading whole columns or whole rows, according to a prearranged plan that follows a route perpendicular to the inscribing route (except in the case of spiral inscription) the decryptographing process is possible. It is obvious that only certain of the more simple combinations of inscription and transcription are suitable for military use, the most practicable being that illustrated in Figures 5 and 6.

27. Columnar transposition with incompletely filled rectangles.—*a.* The degree of cryptographic security of columnar transposition is much increased if the rectangle is *not completely filled.* It is impossible to go into the reasons for this increased security without actually demonstrating solutions; suffice it to say that difficulties placed in the way of the solution are more than would be suspected as a result of so simple a change in method as that which merely involves leaving one or more cells vacant in the last row of cells in the rectangle. An example of cryptographing and decryptographing follows:

Message:

REQUEST IMMEDIATE REENFORCEMENTS

Key word: P – R – O – D – U – C – T
Numerical key: 4 – 5 – 3 – 2 – 7 – 1 – 6

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

FIGURE 7.

Cryptogram:

SINEU EEEQM RCRIT OTEME RSTAF NEDEM

b. To decryptograph such a cryptogram one must first count the number of letters in the text and then outline on cross-section paper a rectangle which will exactly contain the message, crossing off the cells which must remain vacant. In the foregoing example, the text contains 30 letters and, since the key contains 7 letters, the outlined rectangle is as shown in Figure 8 (a). From the complete rectangle of $7 \times 5 = 35$ cells, 5 cells must remain vacant at the end.

Cryptogram:

SINEU EEEQM RCRIT OTEME RSTAF NEDEM

4 – 5 – 3 – 2 – 7 – 1 – 6

(a)

4 – 5 – 3 – 2 – 7 – 1 – 6

		Q	U		S	
		M	E		I	
			E		N	
			E		E	

(b)

FIGURE 8.

c. The cipher text is then inserted in key-number order, the result of inserting the first two groups of the text being shown in Figure

8 (b). It is only after the process has been finished that the complete message becomes apparent.

28. Modification of columnar method.—A variation of the foregoing procedure as regards columns, but one that produces exactly the same results as columnar transposition, may be found useful. It will now be described briefly. First, write the message out in groups corresponding to the length of the key, underneath the latter. Thus, using the same key and message as in paragraph 27, the following is obtained:

```
4-5-3-2-7-1-6 4-5-3-2-7-1-6 4-5-3-2-7-1-6 4-5-3-2-7-1
R E Q U E S T I M M E D I A T E R E E N F O R C E M E
    6 4-5
    N T S
```

The letters are then taken from the groups and are transcribed in groups of five, all letters marked 1 being taken first, then all those marked 2, and so on. Thus, the first two cipher-text groups are SINEU EEEQM, and the complete text is identical with that produced in Figure 7.

29. Addition of nulls to complete a final group.—It will be noted that the example given in the preceding case happened to contain 30 letters, a number that is an exact multiple of five. Thus, the final group in the cryptogram automatically became a complete group. If it is required that the final group of every message be a complete group, a procedure which is conducive to accuracy in transmission, the number of letters comprising the original text of a message must be counted and if not an exact multiple of five must be made so by the addition of nulls, *before* the transposition process is applied (see par. 22 *b*).

SECTION VI

MISCELLANEOUS TRANSPOSITION METHODS

30. Transposition systems employing special designs.—*a.* It is impossible here to demonstrate by example all the different types of designs employed for producing transposition ciphers and to show how they are used. Mention can be made of such designs as triangles, trapezoids, and polygons of various symmetrical shapes. Most of them are impractical for general military use but may occasionally serve as special devices for the use of secret agents.

b. A common transposition device of some practical importance is that known under the general name of grille. This is usually made of a square sheet of cross-section paper from which cells have been cut in definite but apparently irregular positions. The grille is superimposed on another sheet of cross-section paper of the same dimensions and the letters of the message are written in the cells exposed by the perforations. Usually the grille is then given a 90° turn clockwise or counterclockwise, as agreed, and the fresh cells exposed by the perforations are filled with the next few letters of the text. If the grille has been prepared properly it is possible to give it four turns of 90° each, at the end of which all the cells on the under sheet of cross-section paper are occupied by letters. The grille is then removed and the letters of the sheet underneath it are transcribed in accordance with some prearranged route to form the cipher text. Naturally, the correspondents must have identical grilles and every step must be definitely prearranged. Although it is possible to construct grilles with many different arrangements of perforations, the necessity for carrying the device on the person, and the many agreements and understandings necessary for its successful operation make the method hardly suitable for field military use. Furthermore, practical difficulties connected with the preparation and distribution of many grilles would make it almost inevitable that several messages would be enciphered by the same grille. The degree of cryptographic security afforded by them is not so great as may be suspected; sometimes single messages of fair length may be solved.

31. Polyliteral and word transposition.—*a.* The methods indicated thus far employ individual letters as the units for the transposition process. It is of course possible to employ pairs of letters, sets of three or more letters, or entire words as units for the process, and the same methods as have been described in connection with monoliteral methods may be applied in polyliteral transpositions. Sometimes more complicated routes may be followed in the transposition process; for example, a route composed of a prearranged succession of the moves made by the knight in playing chess. It is usually necessary to have at hand a printed form showing the complete route, and this makes these methods impractical for field use. They may, however, be employed in special cases.

b. The cipher system used by the Federal Army in our Civil War represents a good example of word transposition. In the earliest form in which this cipher was used by the Federals only one route was employed, which consisted in writing the text in six columns, going up the sixth, down the first, up the fifth, down the second, up the fourth, and down the third. Arbitrary words were substituted for proper names, nulls were introduced at regular positions, and it was

usual to misspell words in order further to obscure the meaning. For example, the word "operation" was often spelled as two words: "opera", and "shun". Later, many additional routes were provided, relatively long lists of arbitrary equivalents for names, numbers, dates, common military terms, etc., were added and the system made considerably more complicated as a whole. While the degree of cryptographic security afforded by this system was probably great enough for those days, it would hardly be enough today to permit of its use even in cases where a delay of only a few hours is required. Furthermore, if long lists of arbitrary equivalents must be handled, the system presents all the disadvantages of a poor cipher system with but few of the advantages offered by a good code system.

32. Single and double transposition methods.—In all the methods described thus far the letters go through a single transposition from plain text to cipher text. It is, however, possible to take the letters resulting from a first transposition and apply a second transposition to them, resulting often in cryptograms presenting a very great degree of security. Triple and quadruple transposition is likewise possible but wholly impracticable for common use. Only a very limited number of double transposition methods are practicable for military use but the degree of security afforded by certain of them is much greater than that afforded by certain much more complicated substitution methods.

33. Concluding remarks on transposition.—*a.* The various transposition methods described in the preceding explanation differ quite markedly as regards cryptographic security; in some it is almost nil, in others it is of a very high degree. As a general rule, all transposition systems present important advantages as regards speed and simplicity. These advantages have led to attempts to increase the degree of cryptographic security in some manner or other, and hence arise double transposition schemes, rotating grilles, and the like more complicated methods. In only certain types are written memoranda or devices required. Very often the entire cryptographing process in even very complex methods is susceptible of being easily memorized by persons of very good intelligence, such as secret agents. It is for these reasons that transposition systems are often useful in espionage and counterespionage activities.

b. But transposition ciphers for military usage present three very serious disadvantages. In the first place, the methods are in general of such a nature that they do not allow any latitude for the occurrence of errors in handling. Often if a single letter is omitted or added, as not infrequently happens in telegraphic transmission, the whole message becomes difficult if not impossible to decryptograph. In the second place, the fundamental nature of transposition processes is such that if two or more messages prepared in the same key and

containing exactly the same number of letters are available for study, no matter how complicated the method employed, the cryptograms can be solved, the key can be recovered, and applied to other cryptograms in the same key but with different numbers of letters. Since in military cryptography it is not unusual, in cases of heavy traffic, to have as many as 100 or 200 messages transmitted on the same day, all in the same key, and since it would obviously be impractical to try to control from a central headquarters the exact length of messages to be transmitted by many subordinate headquarters, the chances that the enemy may actually intercept and find several messages of identical length are not negligible. Thus, a transposition method presenting an extremely high degree of cryptographic security when only a few messages are to be cryptographed, fails quite markedly when it must be employed for heavy traffic. Finally, in certain cases, where the great degree of security afforded by the system is due to a double process of transposition, it is almost inevitable that a poorly trained or careless cryptographic clerk will fail to perform both steps correctly. This results in not only laying the messages prepared by the one poor or careless operator open to easy solution but also in laying all other messages even though correctly prepared by other careful operators open to solution.

B. Substitution systems

Section VII

GENERAL INTRODUCTORY REMARKS AND DEFINITIONS

34. Fundamental nature of substitution methods, cipher systems and code systems.—The methods now to be described differ from those already described in that the elements or textual units composing the original plain text retain their relative positions but do not retain their identities, being replaced by other elements or textual units so that the external form of the writing is cryptographic in nature. It is for this reason that these methods are termed *substitution methods*. They may deal with individual letters, pairs of letters, sets of letters in regular groups, syllables, whole words, phrases, and sentences. On the basis of the units treated in the cryptographic process, one may, broadly speaking, classify the methods into *letter methods*, *syllable methods*, and *word methods*, just

as in the case under transposition methods; but it should be recognized that such a classification is a rather arbitrary one and is not based on the nature, form, or external appearance of the cryptographic text. For example, a substitution method dealing with single letters of the plain text may not involve their replacement by other single letters. In some cases whole words may be used to replace single letters. In outward form such a cryptogram gives the appearance of dealing with words, but its internal nature is quite clear. Single-letter substitution has been effected. The classification indicated is, nevertheless, a useful one from a practical point of view. Broadly speaking, when the cryptographic process involves, as the general rule, the treatment of individual letters or pairs of letters, and only exceptionally the treatment of syllables or whole words, the method will be referred to as a *substitution cipher system;* and when the process involves, as the general rule, the treatment of whole words, phrases, or sentences, and only exceptionally the treatment of wholly individual letters, groups of letters, or syllables, the method will be referred to as a *code system,* because it usually necessitates the use of a *code book.*

35. Nature of alphabets.—*a.* The simplest kind of substitution cipher is that which is known in the literature as Julius Caeser's Cipher, but which, as a matter of fact, was a favorite long before his day. In this cipher each letter of the text of a message is replaced by the letter standing the third to the right of it in the ordinary alphabet; the letter A is replaced by D, the letter B by E, and so on. The word CAB becomes converted into "FDE" which, we now say, is cipher. The answer to the question, "Why is the 'word' FDE cipher?" is not so simple as one might think at first hand, and involves a consideration of the mechanics of written language, a subject which is of considerable interest to cryptographers.

b. The English language is written by means of a set of 26 simple characters called *letters* which, taken together and considered as a *sequence of symbols,* constitute the *alphabet* of the language. Not all written languages are of this nature. The Chinese language is composed of about 44,000 more or less complex characters, each representing one sense of a word. Whereas English words are composite or polysyllabic and may consist of one to seven or eight syllables, Chinese words are all monosyllables and each monosyllable is a word. The written languages of the majority of other civilized peoples of today are, however, alphabetic and polysyllabic in construction, so that the principles discussed herein apply in general to all of them.

c. The letters composing the alphabet we have today are the results of a long period of evolution, the complete history of which may rest forever unknown. They are merely conventional symbols representing *elementary sounds,* and any other simple symbols, so long as the sounds which they represent are agreed upon by those concerned, will

serve the purpose equally well. If we were taught from early childhood that the symbols $, *, and @ represent the sounds "Ay", "Bee", and "See", respectively, the combination @$* would still be pronounced CAB, and would, of course, have exactly the same meaning as before. Or let us suppose that two persons have agreed to change the sound values of the letters, F, G, and H, and after long practice have become accustomed to pronouncing them as "Ay", "Bee", and "See", respectively. They would then write the "word" HFG, pronounce it CAB, and see nothing strange whatever in the matter. But to us and to others not party to their agreements HFG constitutes cipher. The combination of sounds called for by this combination of symbols is perfectly intelligible to the two who have adopted the new sound values for those symbols, and therefore pronounce HFG, as CAB, but HFG is utterly unpronounceable and wholly unintelligible to us who are reading it according to our own and long established sound-symbol basis. We should say, even if we could pronounce it, that there is no such word as HFG, by which we should mean merely that the particular combination of sounds represented by this combination of letters has not been adopted by convention to represent a thing or an idea in our language. Thus we see that in order for the written words of a language to be pronounceable and intelligible to all who speak that language, it is necessary, first, that the sound values of the letters or symbols be universally understood and agreed upon and, second, that the particular combination of sounds denoted by the letters should have been adopted to represent a thing or an idea. Spoken plain language consists of vocables, that is, combinations and permutations of elementary speech-sounds which have by long usage come to be adopted and recognized as representing definite things and ideas; written plain language consists of *words;* that is, combinations and permutations of simple symbols, called letters, which represent visually and call forth vocally the elementary speech-sounds of which the spoken language is composed.

d. It is clear also that in order to write a polysyllabic language with facility it is necessary to establish and to maintain, by common agreement or convention, equivalency between *two* sets of elements, first, a set of elementary sounds and, second, a set of elementary symbols to represent the sounds. When this is done we have what we call an *alphabet*, a word derived from the names of the first two letters of the Greek alphabet, "alpha" and "beta."

e. Theoretically, in an ideal alphabet each symbol or letter would denote only one elementary sound, and each elementary sound would invariably be represented by the same symbol. But such an alphabet would be far too difficult for the average person to use. It has been conservatively estimated that a minimum of 100 characters would be

necessary for English alone. Attempts toward producing and introducing into usage a practical, scientific alphabet have been made, the most recent being that of the Simplified Spelling Board in 1928, which advocated a revised alphabet of 42 characters. Were such an alphabet adopted into current usage, in books, letters, telegrams, etc., the flexibility of cryptographic systems would be infinitely extended and the difficulties set in the path of the enemy cryptanalysts vastly increased. The chances for its adoption in our day are, however, quite small. On account of the continually changing nature of every living language, it is doubtful whether an original perfect alphabet could, over any long period of time, remain so and serve to indicate with great precision the exact sounds which it was originally designed to represent.

36. Normal alphabets and cipher alphabets.—*a.* In the study of cryptography the dual nature of the alphabet becomes apparent. It consists of two parts or components, (1) an arbitrarily arranged sequence of sounds, and (2) an arbitrarily arranged sequence of symbols.

b. The *normal alphabet* for any language is one in which these two components are the ordinary sequences that have been definitely fixed by long usage or convention. The dual nature of our normal or everyday alphabet is often lost sight of. When we write A, B, C, . . . we really mean:

Sequence of sounds: "Ay" "Bee" "See" . . . .
Sequence of symbols: A B C . . . .

The normal alphabets of the different languages vary considerably as regards the number of characters composing them and the arrangement or sequence of the characters. The English, Dutch, and German alphabets each have 26, the French 25, the Italian 21, Spanish 27 (including the digraphs *ch* and *ll*), Russian 35. The Japanese language has a syllabary consisting of 72 syllabic sounds, to express which 48 characters are employed.

c. A *cipher alphabet* or, as it is sometimes called, a *substitution alphabet* is one in which the elementary speech-sounds are represented by characters other than those representing them in the normal alphabet. These characters may be letters, figures, signs, symbols, or combinations of them.

d. We may now give a more technical definition of a familiar cipher: When the plain text of a message is converted into secret text by the use of one or more cipher alphabets, the resultant cryptogram constitutes a *substitution cipher.*

37. The two components of an alphabet.—It will be convenient to designate that component of a cipher alphabet constituting the sequence of speech-sounds the *plain component*, and the component constituting the sequence of symbols the *cipher component.* In writing a cipher alphabet, if the plain component is omitted, the lat-

ter is understood to be the normal sequence. For the sake of brevity and in order to avoid ambiguity or confusion, a letter of the plain text, or of the plain component of a cipher alphabet, will be designated by suffixing a small letter "p" to it: A_p means A of the plain text, or of the plain component of a cipher alphabet. Similarly, a letter of the cipher text, or of the cipher component of a cipher alphabet, will be designated by suffixing a small letter "c" to it: X_c means X of the cipher text, or of the cipher component of a cipher alphabet. The expression $A_p=X_c$ means that A of the plain text, or A of the plain component of a cipher alphabet, is represented by X in the cipher text, or by X in the cipher component of a cipher alphabet.

38. Standard and mixed cipher alphabets.—As regards the arrangement or sequence of the letters forming its cipher component, cipher alphabets are of two kinds:

a. Standard cipher alphabets, in which the sequence of letters in the cipher component is the same as the normal, but (a) merely reversed in direction or (b) shifted from its normal point of coincidence with the plain component.

b. Mixed cipher alphabets, in which the sequence of letters or characters in the cipher component is no longer the same as the normal in its entirety.

39. Enciphering and deciphering alphabets.—There are various sorts of standard and mixed cipher alphabets, and they will be taken up in their proper places later on, but all cipher alphabets may be classified on the basis of their arrangement as *enciphering* or *deciphering* alphabets. An enciphering alphabet is one in which the sequence of letters in the plain component coincides with the normal sequence, and is arranged in that manner simply for convenience in encipherment. In a deciphering alphabet the sequence of letters in the cipher component coincides with the normal, for convenience in deciphering. For example, in Figure 9 (a) there is shown a mixed cipher alphabet arranged as an enciphering alphabet; (b) shows the deciphering alphabet corresponding thereto. An enciphering alphabet and its corresponding deciphering alphabet present a *verse and inverse* relationship to each other. To invert a deciphering alphabet is to write the corresponding enciphering alphabet; to invert an enciphering alphabet is to write the corresponding deciphering alphabet.

Enciphering Alphabet

(a) Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: JKQVXZWESTRNUIOLGAPHCMYBDF

Deciphering Alphabet

(b) Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: RXUYHZQTNABPVLOSCKIJMDGEWF

FIGURE 9.

SECTION VIII

MONOALPHABETIC SUBSTITUTION SYSTEMS

40. Single-alphabet substitution.—If a cipher alphabet is drawn up and a message is enciphered by its means, letter-for-letter consistently throughout the message, it is said that the cryptogram has been enciphered by a single alphabet, and that it is a single-alphabet substitution cipher. We shall see later that such a cipher represents the most simple type of substitution and that there are more complex ciphers in which several or many alphabets are employed in the encipherment of a single message. The former type, in which a single cipher alphabet is employed, is technically called *monoalphabetic substitution;* the latter type, in which two or more cipher alphabets are employed, is called *polyalphabetic substitution.* Cases of the first type will now be described.

41. Standard alphabet ciphers.—*a.* Standard cipher alphabets are of two sorts:

(1) *Direct standard,* in which the cipher component is the normal sequence but shifted to the right or left of its point of coincidence in the normal alphabet. Example:

```
          ——→
Plain:  ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: QRSTUVWXYZABCDEFGHIJKLMNOP
                  ——→
```

It is obvious that the cipher component can be applied to the plain component at any one of 26 points of coincidence, but since the alphabet that results from one of these applications coincides exactly with the normal alphabet, a series of only 25 (direct standard) cipher alphabets results from the shifting of the cipher component.

(2) *Reversed standard,* in which the cipher component is also the normal sequence but runs in the opposite direction from the normal. Example:

```
          ——→
Plain:  ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: QPONMLKJIHGFEDCBAZYXWVUTSR
              ←——
```

Here the cipher component can be applied to the plain component at any one of 26 points of coincidence, each yielding a different cipher alphabet. There is in this case, therefore, a series of 26 (reversed standard) cipher alphabets.

b. It is often convenient to be able to refer to or designate one of a series of cipher alphabets without ambiguity or circumlocution.

The usual method is to indicate the particular alphabet to which reference is being made by citing a pair of equivalents in that alphabet. For example, the reversed alphabet above, one of a series of 26 related alphabets, may be designated as that in which $L_p=F_c$, or $W_p=U_c$. But the most common basis of reference is the letter which represents the first or initial letter of the plain component, usually A_p. Thus, the key for the cipher alphabet just referred to, as well as that preceding it, is $A_p=Q_c$, and it is said that the key letter for the cipher alphabet is Q_c.

42. Reciprocal alphabets.—Attention may be directed to the fact that the cipher alphabet directly above is also a *reciprocal alphabet*, that is, the equivalents are reversible or reciprocal in pairs. For example, in the alphabet referred to $A_p=Q_c$, and $Q_p=A_c$; $B_p=P_c$ and $P_p=B_c$, etc. This reciprocity holds throughout the whole alphabet and is a result of the manner in which it is formed; the two components are identical sequences, but run in opposite directions. Reciprocal alphabets may be produced (1) by juxtaposing two identical mixed sequences, one running in the opposite direction from the other; or (2) by building up a complete reciprocal alphabet by random assignment of values in pairs.[1]

43. Procedure in encipherment and decipherment.—*a.* The process of enciphering a message by means of a single cipher alphabet is simple. The letters of the text are consistently replaced by their equivalents as noted in the cipher alphabet selected or agreed upon. For example, if the correspondents agreed to employ reversed standard alphabets, and to indicate the particular alphabet used in a given message by writing as the first letter of the final cryptogram the equivalent of A_p, a message may be enciphered as shown in Figure 10 below:

Message: THREE MACHINE GUNS CAPTURED.
Enciphering Alphabet: Reversed Standard, $A_p=D_c$
Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: DCBAZYXWVUTSRQPONMLKJIHGFE
Letter-for-letter encipherment:
THREE MACHINE GUNS CAPTURED
KWMZZ RDBWVQZ XJQL BDOKJMZA

The cipher text is then grouped in fives, the indicator letter D being inserted as the initial letter of the first group (or any other prearranged group).

Cryptogram:

DKWMZ ZRDBW VQZXJ QLBDO KJMZA

FIGURE 10.

b. The procedure in decipherment is merely the reverse of that in encipherment. The initial letter of the message, D, serving as the

[1] A reciprocal alphabet is inverse to or with respect to itself, since it may serve indifferently either as an enciphering or deciphering alphabet.

indicator for the cipher alphabet, shows the latter to be $A_p=D_c$. The deciphering alphabet is therefore as follows:

Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: DCBAZYXWVUTSRQPONMLKJIHGFE

The message deciphers thus:

Cipher: (D) KWMZ ZRDBW VQZXJ QLBDO KJMZA
Plain: THRE EMACH INEGU NSCAP TURED

The deciphering clerk rewrites the text in word lengths:

THREE MACHINE GUNS CAPTURED.

c. When a mixed alphabet is used, the enciphering and deciphering processes are the same in nature as those already described under subparagraphs *a* and *b*. For speed in cryptographing, the cipher alphabet is prepared in the form of an enciphering alphabet, and for speed in decryptographing, in the form of a deciphering alphabet. A brief discussion of various types of mixed alphabets will be useful.

SECTION IX

TYPES OF MIXED CIPHER ALPHABETS

44. Systematically mixed cipher alphabets.—It will be recalled that in a mixed cipher alphabet the sequence of letters or characters in the cipher component does not correspond to the normal sequence. There are various methods of mixing up the letters of the cipher component, and those which are based upon a scheme that is systematic in its nature are very useful because they make possible the derivation of one or more mixed sequences from any easily remembered word or phrase, and thus do not necessitate the carrying of written memoranda. They are called *systematically mixed cipher alphabets.*

45. Key-word mixed alphabets.—*a.* One of the simplest types of systematically mixed cipher alphabets is that known as the *key-word mixed alphabet.* In this type one merely writes down the prearranged key word or key phrase, repeated letters, if present, being omitted after their first occurrence, and then one completes the sequence with the rest of the letters of the alphabet in their normal sequence. Such letters as already occur in the key are, of course, omitted.

b. Mixed alphabets formed by including all repeated letters of the key word or key phrase were common in Edgar Allan Poe's day but are impractical because they make decipherment difficult. An example of such an alphabet is the following:

Enciphering alphabet—
Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: NOWISTHETIMEFORALLGOODMENT

Deciphering alphabet—
Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: P VHMSGD QKAB OEF C
L J RWYN I
X T Z
U

The average cipher clerk would have considerable difficulty in decryptographing a cipher group such as TOOET, each letter of which has three or more equivalents, and from which the plain-text words (N)INTH, . . FT THI(S), IT THI . . . , etc., can be formed on decipherment.

c. An example of a key-word mixed alphabet is shown in Figure 9, where, in its enciphering form, the cipher component presents to the experienced eye the skeleton of the key upon which the alphabet is based: WESTERN UNION TELEGRAPH COMPANY. Any easily remembered word, phrase, or sentence may be employed. The starting point of the sequence, when used as a cipher component, may be indicated in the usual manner. For example, in the alphabet referred to, the alphabet key is $A_p=J_c$. Two or more correspondents using the prearranged key, WESTERN UNION TELEGRAPH COMPANY, would obtain the same disarranged sequence; when this sequence is to form the cipher component of a cipher alphabet, the prearranged key letter, $A_p=J_c$, would result in giving each correspondent exactly the same cipher alphabet. The key words or key phrases need not consist of any definite number of letters, but it is advisable to use for keys such words or phrases as will most thoroughly disarrange the normal sequence. (See, in this connection, par. 25.) It should be noted that a key-word mixed alphabet will manifest the key word or parts of it only when the alphabet is in the form of an enciphering alphabet. Note that alphabet (b) of Figure 9 no longer gives any external evidence of having been derived from the phrase WESTERN UNION TELEGRAPH COMPANY.

46. Transposition-mixed alphabets.—*a.* It is possible to disarrange the sequence even more thoroughly by taking the key-word sequence and applying a simple method of transposition to it just as though it were a message. An example is that illustrated in Figure 11(a), wherein a simple columnar transposition, taking the columns in regular order from left to right, is effected with the key-word sequence based upon the word TELEPHONY. In (b) of the same figure, a columnar transposition based upon the numerical

key derived from the key word itself is applied, yielding a different result.

Key word: TELEPHONY

(a) Simple columnar transposition:

```
TELPHONY
ABCDFGIJ
KMQRSUVW
XZ
```

Mixed sequence:

TAKXEBMZLCQPDRHFSOGUNIVYJW

(b) Numerical key, columnar transposition:

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

Mixed sequence:

EBMZHFSLCQNIVOGUPDRTAKXYJW

FIGURE 11.

b. The last two systematically mixed cipher alphabets may be designated as *transposition-mixed alphabets*, and it is obvious that almost any of the methods of transposition described in Sections IV and V may be applied in their production.

47. Decimation method of producing mixed alphabets.— Another simple method of producing a mixed alphabet is that called the *decimation method*, which consists in agreeing upon a number and then "counting off" the letters in the normal alphabet, or in a key-word mixed alphabet, according to the interval selected. For this purpose the basic sequence to be decimated is regarded as a circle, and as each letter is decimated it is written down in a separate list and the letter is eliminated in the continued decimation of the basic sequence. Thus, suppose the number 7 is agreed upon, the decimation to be applied to the key-word mixed alphabet based on the key phrase SING A SONG OF SIX PENCE:

Key word sequence

SINGAOFXPECBDHJKLMQRTUVWXZ

Decimated Alphabet

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

Mixed Sequence

FHTIEMZPQNDWCVBSLXAGOKYJRU

48. Random-mixed alphabets.—There are, of course, practical considerations which set a limit to the complexities that may be introduced in constructing systematically mixed alphabets, and beyond a certain point there is no object in further mixing. The greatest amount of mixing by systematic processes will give no more security than that resulting from mixing the alphabet by random selection, such as by putting the 26 letters in a box, thoroughly shaking them up, and then drawing the letters out one at a time. Whenever the laws of chance operate in the construction of a mixed alphabet, a thorough disarrangement is bound to be produced. Random-mixed alphabets give more cryptographic security than do the various less complicated types of systematically mixed alphabets because they afford no clues with regard to the positions of any letters, given the positions of a few of them, as is the case with the latter type. Their chief disadvantage is that they must be reduced to writing, since they can not readily be remembered, nor can they be reproduced at will from an easily remembered key word.

49. Number of single alphabets available from a basic alphabet.—It is obvious that the cipher component of a cipher alphabet may be shifted or slid against the plain component at 26 points of contact or coincidence so as to produce a series of different enciphering alphabets. For example, the mixed sequence given under Figure 11 (b), when used as a cipher component, yields the following two of a series of 26 cipher alphabets:

Enciphering Alphabets

(1) Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: EBMZHFSLCQNIVOGUPDRTAKXYJW
(2) Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: WEBMZHFSLCQNIVOGUPDRTAKXYJ

The message DAILY REPORT NOT RECEIVED YET would be enciphered by the first alphabet as:

Plain: DAILY REPOR TNOTR ECEIV EDYET
Cipher: ZECIJ DHUGD TOGTD HMHCK HZJHT

and by the second alphabet as:

Plain: DAILY REPOR TNOTR ECEIV EDYET
Cipher: MWLNY PZGOP RVORP ZBZLA ZMYZR

Externally the two cryptograms seem different, except as regards length. The two enciphering alphabets present the same sequence in the cipher component, but this entirely disappears in the corresponding deciphering alphabets, which are as follows:

113477°—35——6

Deciphering Alphabets

(1) Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: UBIRAFOELYVHCKNQJSGTPMZWXD
(2) Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: VCJSBGPFMZWIDLORKTHUQNAXYE

It is possible to write the same message in 26 different external forms, each employing a different cipher alphabet of a series derivable from a basic sequence. The basic sequence in such a case is often called a *primary sequence* or a *primary alphabet;* the derived alphabets are called *secondary alphabets.* In producing the secondary alphabets the basic sequence must be juxtaposed and slid against itself, or against the normal sequence, or against another mixed sequence. In all cases the secondary alphabets form a series of alphabets that are interrelated and that either directly or indirectly manifest relationships which are important from a cryptanalytic point of view. It should be clear now that by means of a single, prearranged, secret word it is possible for two correspondents to send a whole set of messages all in different mixed alphabets, or to use a different alphabet for each of 26 consecutive days.

50. Miscellaneous types of cipher alphabets.—*a.* The cipher alphabets shown thus far have employed only letters, but alphabets in which the cipher component consists of figures, or groups of figures, are not uncommon in military cryptography. Cipher alphabets employing signs and symbols are not suitable for military cryptography because they can neither be telegraphed nor telephoned with any degree of accuracy, speed, or facility even if at all possible. As to figure ciphers, since there are but ten digits it is obvious that, in order to represent a complete alphabet, combinations of at least two digits are necessary. The simplest kind of such an alphabet is that in which $A_p=01, B_p=02, \ldots Z_p=26$.

b. Instead of a simple alphabet of the preceding type, it is possible to have a simple diagram of the type shown in Figure 12. Here the

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

FIGURE 12.

digits at the side and top of the rectangle are used to designate, according to the coordinate system, the cell occupied by each letter and punctuation marks within the rectangle. When employed for

such purposes, the figures (or letters) constituting coordinate elements are referred to as *row and column indicators*. It is usually necessary to agree beforehand upon which indicator will be given as the first half of the equivalent for a letter, the row indicator or the column indicator, in order to avoid ambiguity or error. In all the systems to be described herein, the row indicator will always form the first half of an equivalent. Accordingly in Figure 12 the letter $A_p=11$, $B_p=12$, etc.

c. A variation of the foregoing diagram is exemplified in Figure 13.

(2)

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

FIGURE 13.

Here, letters of the alphabet are inserted in the 25 cells of a large square, I and J being written together in one cell. Then a key word of five letters is applied to the top of the large square and the same or a different key word is applied to the side of the square to form column and row indicators. In Figure 13, for example, S_p=TI; W_p=EH; etc.

The message RAIDERS HAVE GONE is enciphered thus:

Plain: R A I D E R S H A V E G O N E
Cipher: TH WW HT WT WE TH TI HI WW EW WE HH IT II WE

The cryptogram is then transmitted in groups of five letters:

Cryptogram: THWWH TWTWE THTIH IWWEW WEHHI TIIWE

d. It is obvious that in these two systems just described—

(1) The letters of the alphabet within the square or rectangle may be fixed in a mixed sequence, either systematically or random-mixed sequences being possible.

(2) The column and row indicators may be the same, or different; when letters are used they may form a key word or they may not; the key words, if formed, may be identical or nonidentical.

e. When letters are used as column and row indicators they may be selected so as to result in producing cipher text that resembles

"made-up words", that is, words composed of regular alternations of vowels and consonants. For example, if in Figure 13, the row indicators consisted of the vowels A E I O U and in this sequence from the top downwards, and the column indicators consisted of the consonants B C D F G, in this sequence, from left to right, the word RAIDS would be enciphered as OCABE FAFOD, which very closely resembles code of the type formerly called artificial code language. Such a system as the one just described may be designated as a *false*, or *pseudo-code system*.[1]

SECTION X

MONOALPHABETIC SUBSTITUTION WITH VARIANTS

51. Purpose of providing variant values.—It is well known that the individual letters composing ordinary intelligible plain text are employed with varying frequencies; some, such as (in English) E, T, R, I, and N, are used much more often than others, such as J, K, Q, X, and Z. In fact, each letter has a *characteristic frequency* by means of which definite clues are afforded in the solution of simple substitution ciphers. This has led cryptographers to devise methods for disguising, suppressing, or eliminating the characteristic frequencies manifested by the letters of cryptograms produced by simple monoalphabetic substitution. One of such methods is that in which the letters of the plain component of the cipher alphabet are assigned two or more equivalents in the cipher component and they are, for this reason, called *variant values*. In some cases the letters of the plain component receive numbers of variant values, or variants, in proportion to their normal frequencies; in other cases, all the letters receive equal numbers of variant values, as determined by the total number available. We shall now proceed to note a few examples.

52. Figure ciphers with variant values.—*a.* The use of figures as substitution equivalents, when employed in pairs, makes available a total of 100 different pairs, those from 00 to 99. They may all be used in a complete system or only certain ones may be selected, as prearranged.

[1] Prior to 1934, International Telegraph Regulations required code words of five letters to contain at least one vowel and code words of ten letters to contain at least three vowels. The Madrid Conference held in 1932 amended these regulations to permit the use of code groups containing any combination of letters. These unrestricted code groups were authorized for use after January 1, 1934.

b. One of the most common varieties of ciphers using all the pairs of digits is that in which the alphabet is reduced to 25 letters (by making I and J interchangeable or by eliminating a letter such as Q), and each letter is assigned four values which may be used indifferently or at random. The assignment of values may be based upon a key word of four letters, each of which designates the starting points of a *normal sequence* of 25 numbers. An example is shown in Figure 14, wherein the key word is TRIP. This means that in the first set of numbers, 01 to 25, the first number, 01, is assigned to the letter T; in the second set, from 26 to 50, the first number, 26, is assigned to the letter R; in the third set, from 51 to 75, the first number, 51, is assigned to the letter I; finally, in the last set, from 76 to 00, the first number, 76, is assigned to the letter P.

A—	08	35	68	87
B—	09	36	69	88
C—	10	37	70	89
D—	11	38	71	90
E—	12	39	72	91
F—	13	40	73	92
G—	14	41	74	93
H—	15	42	75	94
I–J—	16	43	51	95
K—	17	44	52	96
L—	18	45	53	97
M—	19	46	54	98
N—	20	47	55	99
O—	21	48	56	00
P—	22	49	57	76
Q—	23	50	58	77
R—	24	26	59	78
S—	25	27	60	79
T—	01	28	61	80
U—	02	29	62	81
V—	03	30	63	82
W—	04	31	64	83
X—	05	32	65	84
Y—	06	33	66	85
Z—	07	34	67	86

FIGURE 14.

The letter A_p may be represented by any one of four equivalents, 08, 35, 68, and 87; the letter B_p by 09, 36, 69, 88; and so on. The equivalent used in any particular instance is merely selected at random, so that the word CAB may be represented in cipher by any one of a total of 64 combinations, such as 10–08–09, 70–35–09, 37–08–69, etc. In the final cryptogram the figures may be run together in groups of five. The cipher group 10080, on deciphering, would be split up into 10–08–0.

c. Within each set of 25 in this case, the numbers progress serially, each set being treated as a ring or circle. It is of course possible to mix the sequence to destroy this serial progression, thus giving four mixed alphabets which can be used at random.

d. Another variation is to assign each letter a set of numbers in accordance with its relative frequency in ordinary English, so that each of the most frequently used letters such as E, T, R, I, and N will have perhaps seven or eight different equivalents, whereas letters of low frequency such as J, K, Q, X, and Z will each have but one equivalent.

53. Use of rectangles as a basis for providing variant values.—*a.* Instead of having alphabets drawn up as shown in Figure 14, it is possible to use the diagram shown in Figure 12, but with several variant digits as the row indicators instead of a single

digit for each row. For example, the row indicators may be of the following arrangements:

1–6–7	1–2–3	1–2–3	5–4–3
2–5–8	4–5–6	8–9–4	6–9–2
3–4–9	7–8–9	7–6–5	7–8–1, etc.

Thus, if the first arrangement is used, A_p would have the equivalents 11, 61, 71; B_p, 12, 62, 72; etc. The word RUN might be represented by any one of 27 different combinations, such as 28–31–24, 28–91–54, etc.

b. A variation of the foregoing scheme is that in which, using a diagram of the type shown in Figure 13, a number of different letters are applied to each row and column, or 2-figure numbers may be used for this purpose, in which case a series of as many as 50 pairs of digits may be employed as the row indicators and another series of 50 pairs, as the column indicators.

c. The use of variants lends itself quite well to application in a pseudo-code system such as described in paragraph 50*e*. It presents many possibilities for variation, with or without key words, with one or more alphabets distributed within the square or rectangle, with alphabets extended to include figures, punctuation signs, common syllables and words, etc. Sometimes pseudo code is encountered when the groups of a numerical cipher system (or a figure-code system) are converted into letters, in order to make the cryptographic text conform to certain telegraph regulations and thus have the message accorded a more favorable rate of charge (see Section XV). Thus, a group such as 0125784256 might be converted into the group BAFOSULAFE. If the conversion table is irregular in its construction and is kept secret, this adds an encipherment step to the system.

54. Remarks on monoalphabetic substitution with variants.—The obvious disadvantage of all such methods as are discussed under *a*, *b*, *c*, and *d* of the preceding paragraph is that the cryptographic text is exactly twice as long as the original plain text. Furthermore, this important disadvantage from the point of view of practicability is not compensated by any really worthwhile advantage from the point of view of cryptographic security. When the methods are such that the cipher equivalents are passed through another process which returns the cipher text to a length identical with that of the equivalent plain text, they are usually too complicated, too slow, and too subject to error to be practical. They are often the result of combining substitution with transposition processes into one system. Methods which substitute three or more characters for one letter of the original text are not at all practical for military cryptography.

SECTION XI

POLYALPHABETIC SUBSTITUTION SYSTEMS

55. Monoalphabetic and polyalphabetic substitution.—*a.* In the substitution methods thus far discussed it has been noted that only one cipher alphabet is employed in the encipherment of a message, and that as a class they constitute the type of system designated as monoalphabetic substitution. It is true that in certain of the systems set forth, namely, those in which monoalphabetic substitution with variant equivalents takes place, there are two or more complete alphabets involved and that these systems may, therefore, with apparently good reason be designated as polyalphabetic substitution, but this designation will be seen to be somewhat inaccurate when cases of *true polyalphabetic substitution* come to be studied. The real or essential difference between the two systems may best be made clear by setting forth the primary object in each case.

b. In monoalphabetic substitution with variant values, the object of having different sets of equivalents is to suppress so far as possible by *simple* methods the characteristic frequencies of letters. One such method consists in merely providing one or more different values as cipher equivalents of the same plain-text letter, or a few different values as equivalents of some of the high-frequency letters. Now there are certain conditions inherent in the method itself, conditions which can not here be indicated, that result in producing in the cryptograms certain definite clues leading to the more or less rapid establishment, in cryptanalysis, of the equivalent of different variant values. Furthermore, in these systems the varying or alternative equivalents for plain-text letters are subject to the free choice and caprice of the encipherer; if he is careful and conscientious in the work he will actually make use of all the variant values afforded by the system; but if he is slip-shod and hurried in his work, he will use the same equivalent repeatedly rather than take pains and time to refer to his charts, tables, or diagrams to find variants. The result of all this is that the cryptograms based upon these methods are open to rather easy solution as a result of carelessness, even when the basic methods are such as would make a solution difficult without the interception of carelessly enciphered messages. What is necessary is a system in which there is established a rather definite procedure for more or less automatically shifting or changing the cipher alphabets employed in the

encipherment of a single message; some method which within certain limits is beyond the momentary whims of cipher clerks, and which to a higher degree makes difficult the establishment of the equivalency of different cipher values. These are the objects of true polyalphabetic substitution systems. The number of such systems is quite large, and it will be possible to describe only a few of the more common or typical examples of methods practicable for military use.

c. The three methods, (1) simple monoalphabetic substitution, (2) monoalphabetic substitution with variants, and (3) true polyalphabetic substitution, are attended by the following consequences in the plain text cipher relationship, a careful study of which will help toward an understanding of the similarities and differences existing among them:

A. Encipherment—

In method (1) each plain-text letter is represented by one and always the same cipher equivalent.

In method (2) as well as in method (3) each plain-text letter is represented by two or more different cipher equivalents, the identities of which are not determined by the positions they occupy in the text.

B. Decipherment—

In method (1) as well as in method (2) each cipher equivalent represents one and always the same plain-text letter.

In method (3) one and the same cipher equivalent represents two or more different plain-text letters, the identities of which are determined by the positions they occupy in the text.

56. Example of polyalphabetic substitution.—*a.* We may illustrate what is meant by true polyalphabetic substitution by a simple example. Suppose that two correspondents agree upon a numerical key, for example, 74030274, each digit of which means that the plain-text letter to which the digit applies as a key number is to be replaced by the letter that stands a corresponding number of places to the right of it in the normal alphabet. For example, if R is to be enciphered by key number 7, it is to be replaced by Y. The numerical key is written under the letters of the plain-text letter for letter, and is repeated until the whole text is covered. Let the message be REENFORCEMENTS BEING RUSHED. The encipherment of a message is shown in Figure 15. For convenience in counting forward (to the right) to find cipher equivalents, a normal alphabet is given at the top of the figure.

```
Normal alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: REENFORCEMENTS BEING RUSHED
Key:  74030274740302 74740 302747
Cipher:YIEQFQYGLQEQTU IIPRG UUUOIK
```

The text is then transmitted in *five-letter groups.*
Cryptogram: YIEQF QYGLQ EQTUI IPRGU UUOIK

FIGURE 15.

b. To decipher such a cryptogram, the clerk writes the numerical key over the cipher letters and then counts backward (to the left) in the normal alphabet as many places as indicated by the key number standing over each letter. Thus:

Normal alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Key: 74030 27474 03027 47403 02747
Cipher: YIEQF QYGLQ EQTUI IPRGU UUOIK
Plain: REENF ORCEM ENTSB EINGR USHED
Message: REENFORCEMENTS BEING RUSHED.

FIGURE 16.

57. Systematizing the work.—The work of encipherment may be materially shortened by systematizing the procedure. Instead of having to write the key over and over again in order to cover the text completely, the text may be written in sets of letters corresponding in length to the length of the key. Thus the text may be written underneath a single appearance of the key in successive short horizontal lines, leaving space between the lines for the insertion of cipher equivalents, as shown in Figure 17 a.

7 4 0 3 0 2 7 4
R E E N F O R C
E M E N T S B E
I N G R U S H E
D

FIGURE 17 a.

Instead of enciphering the letters by individual, repeated countings, two strips of paper bearing normal alphabets may be juxtaposed in the proper relative positions to encipher a whole *column* of letters at one setting of the strips. Thus, for the first column, with the key number 7, the strips are juxtaposed so that the first letter in the column, viz., R (which is to be represented by the seventh letter to the right of it, and is therefore to be enciphered by Y of the lower strip) is directly above Y. Thus:

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRS
Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ

The equivalents for the rest of the letters of the first column may now be rewritten down in their proper places, reference being made to the alphabet strips to see what the cipher letters should be: $E_p=I_c$; $I_p=P_c$; $D_p=K_c$. For the second column the two alphabet strips are in these relative positions:

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUV
Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ

The cipher equivalents for the second column are: $E_p=I_c$; $M_p=Q_c$; $N_p=R_c$. The process is continued in this manner until all the columns have been enciphered, as shown in Figure 17 b.

7 4 0 3 0 2 7 4
R E E N F O R C Y I E Q F Q Y G
E M E N T S B E L Q E Q T U I I
I N G R U S H E P R G U U U O I
D K

FIGURE 17 b.

The cipher text is then transcribed in groups of five letters, reading the successive lines in the normal manner, i. e., from left to right and from the top downwards, yielding the groups YIEQF, QYGLQ, etc.

It is no more difficult to encipher a message by this systematized procedure than by the longer and slower method of writing the text out in long lines and repeating the key over and over again. What is more important, however, is that the shortened procedure promotes accuracy in encipherment. A few seconds careful checking of the relative positions in which the two alphabet strips are set is all that is required but this checking is very necessary, for if that is wrong all the cipher letters in that column to which this setting applies will be in error.

58. Using key words to indicate the number, indentity and sequence of the cipher alphabets employed.—*a.* If reference is made to the two settings of alphabet strips in paragraph 57, it will be noted that in the first setting $A_p=H_c$, in the second $A_p=E_c$. If the eight settings of the strips are studied it will be found thathe letters which A_p represents successively are H, E, A, D, A, C, H, and E, giving the word HEADACHE. These settings, when first presented in the foregoing description, correspond merely to the numerical key 74030274, but now it is noted that this numerical key is expressible in terms of letters, and that the letters when put together properly spell a word. This is only another way of showing that key words may be employed in this type of substitution as in those previously described. Key words of various lengths and composition may be used, consisting of single words, long phrases, or sentences. In general, the longer the key the greater is the degree of cryptographic security. The method as a whole is often referred to as the *repeating key method*, or the *multiple alphabet system.*

b. The number of elements in the key—that is, the number of letters or figures composing it—determines the number of alphabets to be employed; the identity of each element of the key—that is, the specific letter or figure it happens to be—determines specifically which of a set of cipher alphabets pertaining to the whole system will be used; and the specific sequence of the elements of the key—that is, their relative order—determines specifically the sequence with which the cipher alphabets are employed within the encipherment.

The total number of cipher alphabets pertaining to or composing the system may be limited or unlimited. When they are produced as a result of the sliding of two basic or primary alphabets against each other, the number is limited to 26 in the case of the English alphabet.

c. A brief notation for indicating or designating a specific key letter is to suffix the subscript "k" to it, just as the subscripts "p" and "c" are suffixed to letters to indicate letters of the plain text or cipher text, respectively. When the key letter occurs in an equation, it can be enclosed within parentheses to avoid ambiguity. Thus, $B_p\ (D_k)=E_c$ means that plain-text letter B when enciphered by key letter D (in a certain alphabet system) yields the cipher letter E.

59. Use of other types of alphabets.—*a.* It has been noted that in the case of monoalphabetic ciphers, alphabets of various types may be employed. This is likewise true of polyalphabetic ciphers. It is obvious that instead of using two alphabet strips bearing the normal alphabetic sequence to determine the cipher equivalent of a letter enciphered by a given key number or key letter, one may use a pair of strips, one of which bears the normal direct, the other the normal reversed sequence. In the former case we are dealing with direct standard, in the latter, with reversed standard alphabets.

b. Without going into further detail it may be stated that polyalphabetic substitution with direct or reversed standard alphabets does not result in nearly so great a degree of cryptographic security as that resulting from the simple artifice of providing mixed alphabets for the strips. All sorts of mixed alphabets may be used. One of the strips may bear the normal direct or reversed sequence; the other a mixed sequence. Both strips may bear identical mixed sequences proceeding (a) in the same direction, or (b) in opposite directions. Finally, both strips may bear different mixed sequences.

c. In all cases, except in those where reciprocal alphabets are produced, it is essential that the correspondents agree upon the sequence or strip from which the plain and the cipher letters respectively will be taken; i. e., it is necessary to indicate which sequence constitutes the plain component, which the cipher component. If this is not done, two correspondents will have difficulty in deciphering one another's messages. Also, as noted above, it is necessary to agree as to which letter the key letter is to be set against. The usual method is to agree that the initial letter of the plain component, usually A_p, will be set opposite the key letter, though other conventions are possible.

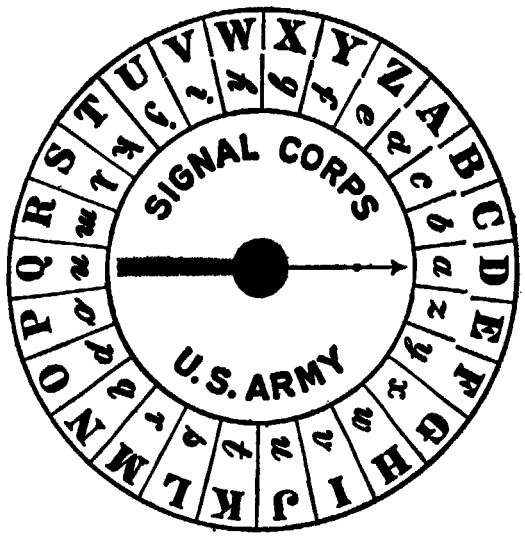
d. The sequences on the strips may be permanent or invariable, but naturally the degree of cryptographic security in this case is considerably lower than if they can be changed easily at the will of the correspondents and by prearrangment. It is possible that a secret word may serve as the basis not only for the key for shifting the

strips, but also for the mixing of the alphabetic sequences. For example, two correspondents may agree to use the key CENTRAL AMERICA; to use the first part as the basis for constructing the mixed plain component; the second part, for constructing the cipher component; and to use the whole phrase as the key for enciphering the message. All the methods of constructing systematically mixed alphabets as described in Section IX are applicable.

SECTION XII

CIPHER DISKS AND SQUARE TABLES

60. Cipher disks.—*a.* In the foregoing remarks it was noted that the separate alphabets employed in the encipherment are produced by the use of only two strips of paper bearing the normal alphabet. Such strips are often referred to as *sliding alphabets* because they can be shifted or slid against each other in any one of 26 points of contact or coincidence. Exactly the same results, so far as cipher equivalents are concerned, can be obtained by the use of other devices. First, there are the so-called cipher wheels or cipher disks in which an alphabet is written on the periphery of a *rotating* disk, the circumference of which is divided into 26 equal segments, and this disk is made to revolve concentrically upon a similar but slightly larger *fixed disk*. In Figure 18 there is shown the now obsolete U. S. Army Cipher Disk, which is of this simple type. Here the alphabetic sequences are printed on glossy celluloid, are permanent, and



To encipher a message, the key letter or the first letter of the key word or phrase is set opposite "a." Let us assume it to be "E." The cipher letters to be written are those opposite the text letter when "a" on the circle is set opposite "E" on the card. For example, "send powder" would be written "MARBPQIBAN." To use a key word or phrase, each letter is used in turn to encipher one letter only. When the last letter of the key word is used, repeat until all letters of the message are enciphered. Numbers when enciphered with the disk must be spelled out.

FIGURE 18.

admit of no variation. The use of unglazed celluloid upon which blank segments appear would permit of writing letters and erasing them as often as desirable. Thus, quick and easy change of alphabets would be possible.

b. The cipher alphabets produced by the cipher disk shown in the figure are merely reversed standard alphabets, the same as are produced by the use of sliding strips of paper, and, as shall soon be seen, by the use of certain tables. The method of employing the disk needs no discussion. It may serve in monoalphabetic or polyalphabetic substitution with a key word or key number.

61. Square tables.—*a*. Tables known in the literature of cryptography under various names, such as "Vigenère Table", "Square Table", "Quadricular Table", "Pythagorean Table", "Cipher Square", "Cipher Chart", etc., are often employed in polyalphabetic substitution. All the results produced by their use can be duplicated by the employment of sliding alphabets or revolving disks. The Vigenère Table is shown in Figure 19.

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

FIGURE 19.—The Original Vigenère Table.

Such a table may be used in various ways, differing from one another in minor details. The most common method is to consider the top line of the table as containing the plain-text letters, the first column at the left as containing the key letters. Then each successive horizontal line contains the cipher equivalents for the plain-text sequence ABC . . . Z enciphered by the key letter which stands at its left in the first column. Thus, the cipher alphabet corresponding to key letter D is the sequence of letters in the fourth horizontal line under the plain-text line, where $A_p=D_c$, $B_p=E_c$, etc. It will be easy to remember, in using such a table, that the equivalent of a given plain-text letter, T_p, for example, enciphered by a given key letter, O_k, lies at the intersection of the vertical column headed by T, and the horizontal row begun by O. In this case $T_p(O_k)=H_c$. The same result will be found on referring to sliding, direct standard alphabets.

b. Minor modifications of the Vigenère Table are encountered. If the top line is made a reversed normal sequence, leaving the interior of the table unchanged, or if the successive horizontal rows are made to contain the reversed normal sequence, leaving the top row (plain text) unchanged, then the results given by using the table are the same as those given by using the obsolete cipher disk shown in Figure 18. Again, the same general results can be obtained by using a set of alphabets in tabular form known under the names of Porta's Table and Napoleon's Table, which is shown in Figure 20.

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

etc.

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

FIGURE 20.

In this table the alphabets are all reciprocal, for example, $G_p(W_k)=V_c$, $V_p(W_k)=G_c$. Reciprocal alphabets when arranged in this form are sometimes called *complementary alphabets*. Note that in each alphabet either of two letters may serve as key letter indifferently: $G_p(W_k)$ or $G_p(X_k)=V_c$.

c. Another modification of the basic table, and one that employs numbers instead of letters as cipher equivalents is shown in Figure 21. Since many more than 26 different equivalents are available (100 pairs of digits from 00 to 99), it is possible to insert many plain-text elements in the top line of the table in addition to the 26 letters. For example, one could have the 10 digits; a few common double-letter combinations, such as DD, LL, RR, SS; a few of the most frequently used pairs of letters, such as TH, ER, IN, or even such common syllables as ENT, ING, and ION.

*	a	b	c	d	e	f	g	h	i	*
a	10	11	12	13	14	15	16	17	18	a
b	11	12	13	14	15	16	17	18	19	b
c	12	13	14	15	16	17	18	19	20	c
d	13	14	15	16	17	18	19	20	21	d
e	14	15	16	17	18	19	20	21	22	e
f	15	16	17	18	19	20	21	22	23	f
g	16	17	18	19	20	21	22	23	24	g
h	17	18	19	20	21	22	23	24	25	h
i	18	19	20	21	22	23	24	25	26	i
j	19	20	21	22	23	24	25	26	27	j
k	20	21	22	23	24	25	26	27	28	k
l	21	22	23	24	25	26	27	28	29	l
m	22	23	24	25	26	27	28	29	30	m
n	23	24	25	26	27	28	29	30	31	n
o	24	25	26	27	28	29	30	31	32	n
p	25	26	27	28	29	30	31	32	33	p
q	26	27	28	29	30	31	32	33	34	q
r	27	28	29	30	31	32	33	34	35	r
s	28	29	30	31	32	33	34	35	36	s
t	29	30	31	32	33	34	35	36	37	t
u	30	31	32	33	34	35	36	37	38	u
v	31	32	33	34	35	36	37	38	39	v
w	32	33	34	35	36	37	38	39	40	w
x	33	34	35	36	37	38	39	40	41	x
y	34	35	36	37	38	39	40	41	42	y
z	35	36	37	38	39	40	41	42	43	z
*	a	b	c	d	e	f	g	h	i	*

*	j	k	l	m	n	o	p	q	r	*
a	19	20	21	22	23	24	25	26	27	a
b	20	21	22	23	24	25	26	27	28	b
c	21	22	23	24	25	26	27	28	29	c
d	22	23	24	25	26	27	28	29	30	d
e	23	24	25	26	27	28	29	30	31	e
f	24	25	26	27	28	29	30	31	32	f
g	25	26	27	28	29	30	31	32	33	g
h	26	27	28	29	30	31	32	33	34	h
i	27	28	29	30	31	32	33	34	35	i
j	28	29	30	31	32	33	34	35	36	j
k	29	30	31	32	33	34	35	36	37	k
l	30	31	32	33	34	35	36	37	38	l
m	31	32	33	34	35	36	37	38	39	m
n	32	33	34	35	36	37	38	39	40	n
o	33	34	35	36	37	38	39	40	41	o
p	34	35	36	37	38	39	40	41	42	p
q	35	36	37	38	39	40	41	42	43	q
r	36	37	38	39	40	41	42	43	44	r
s	37	38	39	40	41	42	43	44	45	s
t	38	39	40	41	42	43	44	45	10	t
u	39	40	41	42	43	44	45	10	11	u
v	40	41	42	43	44	45	10	11	12	v
w	41	42	43	44	45	10	11	12	13	w
x	42	43	44	45	10	11	12	13	14	x
y	43	44	45	10	11	12	13	14	15	y
z	44	45	10	11	12	13	14	15	16	z
*	j	k	l	m	n	o	p	q	r	*

*	s	t	u	v	w	x	y	z	*
a	28	29	30	31	32	33	34	35	a
b	29	30	31	32	33	34	35	36	b
c	30	31	32	33	34	35	36	37	c
d	31	32	33	34	35	36	37	38	d
e	32	33	34	35	36	37	38	39	e
f	33	34	35	36	37	38	39	40	f
g	34	35	36	37	38	39	40	41	g
h	35	36	37	38	39	40	41	42	h
i	36	37	38	39	40	41	42	43	i
j	37	38	39	40	41	42	43	44	j
k	38	39	40	41	42	43	44	45	k
l	39	40	41	42	43	44	45	10	l
m	40	41	42	43	44	45	10	11	m
n	41	42	43	44	45	10	11	12	n
o	42	43	44	45	10	11	12	13	o
p	43	44	45	10	11	12	13	14	p
q	44	45	10	11	12	13	14	15	q
r	45	10	11	12	13	14	15	16	r
s	10	11	12	13	14	15	16	17	s
t	11	12	13	14	15	16	17	18	t
u	12	13	14	15	16	17	18	19	u
v	13	14	15	16	17	18	19	20	v
w	14	15	16	17	18	19	20	21	w
x	15	16	17	18	19	20	21	22	x
y	16	17	18	19	20	21	22	23	y
z	17	18	19	20	21	22	23	24	z
*	s	t	u	v	w	x	y	z	*

*	0	1	2	3	4	5	6	7	8	9	*
a	36	37	38	39	40	41	42	43	44	45	a
b	37	38	39	40	41	42	43	44	45	10	b
c	38	39	40	41	42	43	44	45	10	11	c
d	39	40	41	42	43	44	45	10	11	12	d
e	40	41	42	43	44	45	10	11	12	13	e
f	41	42	43	44	45	10	11	12	13	14	f
g	42	43	44	45	10	11	12	13	14	15	g
h	43	44	45	10	11	12	13	14	15	16	h
i	44	45	10	11	12	13	14	15	16	17	i
j	45	10	11	12	13	14	15	16	17	18	j
k	10	11	12	13	14	15	16	17	18	19	k
l	11	12	13	14	15	16	17	18	19	20	l
m	12	13	14	15	16	17	18	19	20	21	m
n	13	14	15	16	17	18	19	20	21	22	n
o	14	15	16	17	18	19	20	21	22	23	o
p	15	16	17	18	19	20	21	22	23	24	p
q	16	17	18	19	20	21	22	23	24	25	q
r	17	18	19	20	21	22	23	24	25	26	r
s	18	19	20	21	22	23	24	25	26	27	s
t	19	20	21	22	23	24	25	26	27	28	t
u	20	21	22	23	24	25	26	27	28	29	u
v	21	22	23	24	25	26	27	28	29	30	v
w	22	23	24	25	26	27	28	29	30	31	w
x	23	24	25	26	27	28	29	30	31	32	x
y	24	25	26	27	28	29	30	31	32	33	y
z	25	26	27	28	29	30	31	32	33	34	z
*	0	1	2	3	4	5	6	7	8	9	*

FIGURE 21.

62. Square tables employing mixed alphabets.—*a.* In the tables thus far shown the alphabets have been direct or reversed standard sequences, but just as mixed sequences may be written upon sliding strips and revolving disks, so can mixed alphabets appear in tabular form. The following table, based upon the key word sequence derived from the word LEAVENWORTH, is an example that is equivalent to the use of a strip bearing the same key word sequence sliding against another strip bearing the normal alphabet.

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

FIGURE 22.

The usual method of employing such a table is the same as that in the preceding cases. The only difference is that the key letters must now be sought in a mixed sequence, whereas in the preceding tables they were located in normal direct or reversed sequences. Example, using Figure 22: $C_p(S_k)=X_c$.

b. In the table shown in Figure 23, there is a case wherein a mixed alphabet is sliding against itself.

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

FIGURE 23.

The usual method of employing such a table is exactly the same as before. The only difference is that both the plain-text letters and the key letters must be looked for in mixed sequences. Example, using Figure 23: $U_p(R_k)=V_c$.

c. There is one item to which it is necessary to call attention in connection with the last table. It has been indicated that the basis of reference in most cryptographic operations involving key words is the letter A_p. In employing sliding alphabets it has been usual to set the key letter as located in the cipher component opposite the letter A as located in the plain component. But it was also indicated (pars. 41 *b*, 59 *c*) that the key letter as located in the cipher component is usually set opposite the initial letter of the plain component. In all previous examples, this initial letter has been A, but in the case of Figure 23, since the plain component is also a mixed sequence, and since its initial letter is Q, the sliding alphabets are set against each other so that the given key letter in the cipher component is opposite Q in the plain component. Thus, to duplicate the results given by the use of Figure 23 in finding the value of $U_p(R_k)$, it is necessary to set the sliding strips in the following relative positions:

```
Plain: QUESTIONABLYCDFGHJKMNPRVWXZQUESTIONABLYCDEFGHJKM
Cipher:       QUESTIONABLYCDFGHJKMNPRVWXZ
```

Here we see that $U_p(R_x)=V_c$, which is identical with the result obtained from the use of the table. It must be remembered, however, that there are other ways of using the table, each having a correspondingly modified method of employing sliding strips in order to obtain identical results.

SECTION XIII

CONCLUDING REMARKS ON CIPHER SYSTEMS

63. More complex substitution systems.—*a.* The substitution systems thus far described represent relatively simple methods. They can all be solved, and their solution can usually be reached in time for the information to be of great value. For that reason more complicated methods have been devised and are used, but since they do not fall within the scope of this subcourse only brief reference to them can be made herein.

b. Practically all systems based upon the principle of a repeating key can be solved on account of certain phenomena of a cyclic or periodic nature which the use of such a key causes to be exhibited externally or internally in the cryptograms. There are methods for preventing the external manifestation in the cryptograms of these phenomena, or their suppression and disguise if present internally. In some, the principle is to make the elements of a fixed or invariable-length key apply to variable or irregular-length groupings of the plain text so that no cyclic phenomena are exhibited by the cryptograms. In others, the principle is to apply irregular lengths of the key, or a variable-length key to regular and fixed groupings of the plain text, with the same object in view. In still other methods, both principles are combined, and in still others the key itself is of such a nature that it does not repeat itself. This may be brought about either by constructing or establishing a nonrepeating key, or by employing the key in a special manner. Systems in which the successive letters of the cipher text or successive letters of the plain text after the initial letter serve as successive key-letters are also used with the object of avoiding or eliminating periodicity.

c. In the majority of the methods herein described the encipherment deals with single letters, and is therefore *monographic* in nature. There are, however, certain methods in which encipherment is by pairs of letters, called *digraphic substitution*, or by sets of three letters, called *trigraphic substitution*. *Polygraphic substitution methods*, as they are called, have for their object the suppression, so far as

possible, of the characteristic frequencies of individual letters, by means of which solution may be reached. The methods may employ extensive tables, small squares, rectangles, and other designs, or sets of sliding or rotating alphabets. The *Playfair Cipher*, which was for many years a standard field cipher in the British Army and was for a short time during the World War employed by our own Army, is an example of digraphic substitution.

64. Combined substitution-transposition systems.—In paragraph 14 *b*, reference was made to the possibility of combining within a single system both transposition and substitution methods; that is, of first enciphering by a method of one type and then taking the resulting cipher text and passing it through an encipherment of the other type. The usual order is first to substitute and then to transpose, but the reverse of this order of procedure is also possible. In some methods, quite complex, there may be a first substitution, then a transposition, and finally a substitution again. Despite the fact that three steps are involved, certain of these systems may be practical for military use under special conditions where speed is not so important as security. These cannot be described herein.

65. Cipher devices and cipher machines.—*a*. Only a little practical experience with any of the methods herein described is necessary to convince one that on the whole they are slow, more or less cumbersome, and subject to errors that often delay or make impossible the decryptographing of messages. Furthermore, from the point of view of cryptographic security when employed in regular voluminous traffic, they leave much to be desired. Consequently, cryptographers, both experienced and inexperienced, have been led to attempt to devise apparatus which will not only facilitate cryptographing and decryptographing, but will also increase the degree of cryptographic security. Small devices constructed for this purpose, operated by hand, are often called *cryptographs*. Scores of them have been devised, of which only a very few are sufficiently practicable for field use, and still fewer are of such construction that they produce cryptograms of unusual security. Among the better examples of such cryptographs is that employed in our Army under the name of *Cipher Device, Type M–94*. It is normally used for messages the origin or destination of which is forward of battalion headquarters. The degree of cryptographic security of the device, however, is not especially great, particularly when employed under circumstances where the enemy is in a good position to assume with a fair degree of probability the presence in messages of such words as ENEMY, BATTALION, ARTILLERY, etc.

b. There are larger cryptographic machines which are much more automatic in nature and can therefore be operated at a much greater rate of speed. These are usually equipped with typewriter keyboards

which can be manipulated with considerable speed; the machine may also print the enciphered and deciphered text. Sometimes they are equipped with electrical transmitters and can thus serve not only to encipher and decipher messages but also to transmit them automatically. A mechanism of the latter nature is usually in the form of a modified printing telegraph apparatus. One of the cipher systems adopted for war time use in the Army is of this type, and is known as the *printing telegraph cipher system.* It can only be used between the larger headquarters where traffic is very great.

66. Disadvantages and limitations of cipher systems.—Aside from certain cipher machines that are operated electrically or mechanically in conjunction with a typewriter keyboard, all cryptographic methods employing cipher systems are comparatively slow, cumbersome, and subject to error. Practically all of them are open to solution by enemy cryptanalysts, and such as are suitable for use in the theater of operations can by the very nature of the limitations imposed by such use offer fewer obstacles to solution than can systems suitable for use in the rear areas. Furthermore, cipher systems are not economical as regards the number of time units required in electrical transmission, since the best that they can do in this respect is to produce cryptograms no longer than the original plain text. When it is considered that there are other cryptographic methods which offer more advantages in respect to speed, simplicity, and economy, and at the same time afford as great or even greater degrees of cryptographic security, it is not surprising to find that the latter methods predominate over cipher methods in those fields in which these factors are essential.

Part Three

CODE SYSTEMS

Section XIV

INTRODUCTORY REMARKS ON CODE SYSTEMS

67. Difference between code and cipher systems as methods of cryptography.—In the final analysis, a code system is only a more or less highly specialized form of substitution. The basic principle underlying substitution cipher systems is the replacement of the individual letter constituting the plain text of a message by other letters, figures, symbols, and the like. It is only occasionally or exceptionally that the replacement or substitution process is applied to groups of letters, and when this is the case the groups are usually of

definite, or regular length. Broadly speaking, in cipher systems the units with which the cryptographic treatment deals are fundamentally the smallest ones of which plain text can be composed. They are roughly analogous, let us say, to the atoms with which 19th century chemistry concerned itself. The basic principle underlying code systems, on the other hand, is the replacement of the entire words, long phrases, or complete sentences constituting the plain text of a message by arbitrarily selected equivalents having little or no relation or connection with the elements they replace. These equivalents may be other words, groups of letters, groups of figures, or both. It is only occasionally or exceptionally that the replacement or substitution process is applied to elements smaller than whole words, and when this is the case it is immaterial whether these elements are single letters, groups of letters, or syllables. Broadly speaking, in code systems the units with which the cryptographic treatment deals are aggregates of smaller units—individual letters combined in various groups of irregular length; that is, words, phrases, sentences. If the units with which cipher systems are concerned are roughly analogous to the atoms of the old atomic chemistry, then the units with which code systems are concerned are roughly analogous to the molecules of molecular chemistry, and to the more complex compound substances with which organic chemistry deals.

68. Code books and codes.—*a.* If it were possible to memorize a long list of words, phrases, and sentences together with the arbitrary equivalents called *code groups* assigned to represent them, there would be no need of writing them down. Indeed, at least one code based on a mnemonic principle has been proposed and an example constructed, but its practical use is naturally restricted to a very limited field. There is hardly any practical method of dealing with code other than to have a *written document* in which the words, phrases, and sentences are listed in some systematic manner and are accompanied by their arbitrarily assigned code equivalents. It is obvious that correspondents must possess identical copies of the document in order to communicate with one another. Although an ordinary dictionary may, and often does serve the purpose of code communication, so far as single words are concerned, as a rule a specially prepared document containing the words, phrases, sentences, etc., suited to particular types of correspondence, is employed. Such documents are called, in this country and in Great Britain, *code books* or simply *codes*. In other countries they are called *repertories*, *word books*, *cipher dictionaries*, *enciphering and deciphering tables*, etc., although the term "code" is becoming more and more prevalent throughout the world.

b. There are various types of codes depending on the particular types of correspondence for which they are adapted. Some are fairly

large books suitable for general business or social correspondence; others are much more specialized for particular industries—for example, the rubber, sugar, steel, automobile industries—and therefore contain highly specialized technical vocabularies. Most large commercial firms have their *private codes*, constructed especially for their own use, as those of the foremost banking houses. We, however, are more concerned here with codes such as are more suitable for military communication, and while the resemblances between the ordinary commercial codes and the usual military codes are quite marked, the primary purposes are different in the two cases. The principal purpose of code in commercial communications is to effect economy in cost of communicating, secrecy being usually of secondary importance. The principal purpose of code in military communications is to effect secrecy; and economy, while a very important additional feature, is of secondary importance.

69. Operations of encoding and decoding.—These two terms have been referred to before and apply to the cryptographing and decryptographing, respectively, of messages by means of the code concerned. In encoding it is merely necessary to replace the various words, phrases, sentences, numbers, etc., by the code groups given as their equivalents in the code. The code text is built up from individual code units each representing the longest possible plain-text unit the code book affords. For example, if the sentence ENEMY FORCE ESTIMATED AT ONE BATTALION ENCOUNTERED ONE MILE SOUTH-EAST OF ROCK CREEK CHURCH is to be encoded, and the code book lists the phrase ENEMY FORCE ESTIMATED AT, the code group representing the latter phrase would be used rather than separate code groups representing the individual words ENEMY, FORCE, ESTIMATED, and AT, all of which might also be present in the code. In the case of words or proper names which are not listed, provision is usually made for building up the word by means of code groups representing individual letters, groups of letters, and syllables; these are often contained in a special table called a *spelling table* or a *syllabary*. The process of decoding is, of course, merely the reverse of that of encoding. Each code group is looked up in the code book, its meaning found and written down. Where the errors in transmission are few, the process is quite rapid; but it is obvious that even a small number of errors in a message may obscure the meaning or render a message unintelligible.

70. Economies afforded by code systems.—*a.* It is obvious that messages cryptographed by means of a code book are secret only when the code book is kept secret. There are however code systems the purpose of which does not include the factor of secrecy. They are intended merely for economy. This possibility arises from the fact that code books afford a means for abbreviating or condensing the

writing necessary to convey a given amount of information, since a single comparatively short group of code characters may represent a whole word of as many as 15 or more letters, a long phrase, or a complete sentence. Thus, as a rule, the code text of a message is much shorter than the plain text, a result which is very conducive to economy in communication. Naturally, the *condensing power* of a code book varies with its size,— that is, with the extensiveness of its *vocabulary* or contents, since in a small book there can be listed only the most common words and only a few phrases and sentences; whereas, in a large book practically all the words likely to be used in telegraphic communication and many common phrases and sentences may be included. When a code book is intended only to condense the text for purposes of economy, it is called a *nonsecret code*. Examples of such codes may be found in the ordinary commercial codes already referred to, which are purchasable from book dealers. A code book may combine the features of economy and secrecy, in which case the book itself must be safeguarded from the enemy as a *secret code*.

b. Economy expressible in terms of money is not the only form of economy that code systems afford, so far as the user of codes is concerned. There is in addition a very noticeable saving in time and labor in cryptographing, copying, and general handling of code as compared with cipher. These are extremely important in military cryptography. Furthermore, the simplicity of the processes involved in cryptographing and decryptographing code messages as compared with those involved in the case of cipher results in an economy in time and labor—two additional important advantages of code.

SECTION XV

CODE GROUPS

71. Composition of code groups.—*a*. As regards the elements of which code groups are composed, they may be of one or more of the following types:

(1) *Bona fide words*, that is, real words taken from the dictionaries of one or more languages. The usual languages employed as sources for code words of this type are Dutch, English, French, German, Italian, Latin, Portuguese, and Spanish.

(2) *Artificial words*, that is, groups of letters having no intrinsic meanings, and constructed more or less systematically of arrangements of vowels and consonants so as to impart to these groupings the appearance and pronounceability of bona fide words.

(3) Groups of letters presenting no appearance of bona fide or artificial words and resembling cipher groups.

(4) Groups of arabic figures.

b. For special purposes code groups composed of intermixtures of letters and figures within groups may be used, and then only in military or naval communications. Radio call signs for amateur stations, such as W2KA and W5AZZ, are examples of such intermixtures often used in radio call-sign codes. These are not examples of code groups used in message codes. A code may, however, contain two or more parallel sets of code groups of different types. For example, in many commercial codes and in some military and naval codes, there is one series of code groups of the bona fide or artificial word type and another series of the figure group type, both applying to the same series of words, phrases, and sentences of the code. There are several reasons for this. In most parts of the world where italic or roman letters are used for writing, letters possess greater advantages as regards accuracy in reading and handling by telegraph personnel, this being the prerequisite to correct transmission and reception of messages. However, in some parts of the world—for example, Turkey, Russia, China—telegraph personnel, except in the larger cities, are unfamiliar with our alphabet and hence many errors arise. But the arabic digits are almost universally recognized and used, so that for communications between obscure ports and small cities in foreign countries, figure groups are preferred above letter groups. Again, in telephoning code messages, letters are harder to understand and are received with less accuracy than are the ten digits, on account of certain technical limitations in this means of communication. Again, there are certain methods of condensing code groups composed of figures into still smaller groups composed of letters by means of *condensers*, so that many firms use figure groups for such purposes in expensive transmissions. Finally, in certain methods of enciphering code messages for the sake of more secrecy, figure groups form the basis for the encipherment more readily than do letter groups.

c. Prior to January 1, 1934, in practically all modern codes constructed by experts, the letter code groups were of the artificial-word type. On that date new rules in international communication became effective,[1] permitting the use of letter code groups without restriction as to their formation, i. e., class (3) in *a* above. It is probable that almost all of the codes published in the future will contain letter code groups of the unrestricted type.[2]

d. The greatest advantage possessed by letter groups over figure groups lies in the availability of a far greater number of permutations

[1] See Telegraph Regulations, International Telecommunication Convention, Madrid, 1932.

[2] For a treatise of the development of codes see "The History of Codes and Code Language, the International Telegraph Regulations pertaining thereto, and the bearing of this history on the Cortina Report", by Major William F. Friedman, Sig.-Res., Government Printing Office, 1928.

of letter groups, because there are 26 letters which may be permuted to form letter code groups, whereas there are only 10 digits which may be permuted to form figure groups. If code groups of five elements are used, then there are available 26^5, or 11,881,376 groups of five letters, and only 10^5, or 100,000 groups of five figures. Now since the number of permutations of 26 letters taken in groups of five is so great, only permutations conforming to special types may be selected for use, and there will still remain a sufficient number of code groups for even the largest codes. The selection of certain types of code groups is done with a view to reducing to a minimum the inevitable errors in telegraphic transmission. Furthermore, if the code groups have been constructed scientifically it is possible to provide a quick and effective means of correcting such errors as do creep in, without having to call for a repetition of the message.

72. Length of code groups.—The length of code groups used, i. e., whether they are groups consisting of two, three, four, or five elements, depends upon the size of the code. This however applies almost exclusively to field military or naval codes, where transmission is through a governmental agency; for in commercial messages or in governmental communications transmitted over privately owned and operated lines, five-letter or five-figure groups are used almost exclusively on account of the regulations adopted by the International Telegraph Conferences, and by the commercial telegraph and cable companies. Throughout the world in transmission of code and cipher messages, each group of five letters is charged for as one word regardless of the number and arrangement of vowels. In all countries, except the United States and Canada, each group of five figures is likewise charged for as one word. In the United States and Canada each group of five figures is charged for as *five* words.

73. Permutation tables and the two-letter differential.—*a.* The code groups of modern codes are constructed by the use of tables which permit of the more or less automatic and systematic construction of code groups of the form desired as regards their length, similarities and differences. These are called *permutation* tables. Because they may be used to correct the majority of the errors made in transmission or writing, such tables are usually incorporated in the code book and are often called *mutilation tables, garble tables, error-detector charts*, etc. When scientifically constructed, such tables include a feature that has greatly increased the reliability of code as a system of communication, a matter in which it was rather deficient before the invention of permutation tables. This feature is discussed in *b* below.

b. To make an error in a group of five letters is not at all unusual on the part of the average telegraph or radio operator. If a difference in only one letter distinguishes one code group from another in the

same code, e. g., ABABA and ABABE, then serious errors may be introduced in the meaning of a message, or the message may be rendered unintelligible or obscure by the presence of only a few transmission errors. If, however, every code group in the code book is distinguished from all other code groups in the same code by a difference in at least *two* letters, then there would have to be two errors in a single group and these two errors would have to be such as to produce a code group actually present in the code before a wrong meaning could be conveyed by the so mutilated group. This principle of having code groups of the same code differ from each other by a minimum of two letters is called the *two-letter differential.* It is most easily incorporated in code groups by constructing the permutation table with this end in view. The differential may consist in absolute difference in the identities of the two letters, or in the relative positions occupied by them. For example, BACOF and BACUG differ from each other in the identities of the final pair of letters; considered as a *combination* of letters, the two groups present a two-letter difference. The two groups BACOF and BOCAF, however, differ from each other in the relative positions occupied by two of their letters, but considered as a permutation of letters, these two groups as well as the two groups BACOF and BACUG present a two-letter difference. In short, when at least two *homologous* letters in a pair of code groups differ in their identities, the two code groups are said to present a 2-letter difference. Errors arising from the exchange of position of two letters, without a change in their identities, are referred to as errors due to *transposition.* They are not at all unusual but fortunately, as a rule, they involve only letters which are either adjacent or alternate. For example, in the pair of groups BACOF and BOCAF there is a transposition of the alternate-letter type. In some of the most recent codes, attempts have been made to devise permutation tables which will eliminate, avoid, or suppress one of the two members of every pair of groups differing from each other by the mere transposition of two adjacent or alternate letters. Inasmuch as codes employing groups based upon a permutation table show the table and explain how to use it in correcting the usual mutilations of groups, it is unnecessary to go further into this subject herein.

Section XVI

ONE-PART AND TWO-PART CODES

74. Arrangement of contents of codes.—As regards their construction or arrangement, codes may be of two types:

(1) *One-part*, or alphabetical codes, in which the plain-text groups are arranged in alphabetical order accompanied by their code groups

which are also arranged in alphabetical order, or numerical order. Such a code serves for decoding as well as for encoding.

(2) *Two-part*, or randomized codes, in which the plain-text groups are arranged in alphabetical order accompanied by their code groups arranged in a nonalphabetical or random order, the code groups being assigned to the plain-text groups in an absolutely arbitrary and random manner, by drawing the code groups out of a box in which they have been thoroughly mixed up, or by some other manner in which the element of chance operates in assigning the code groups to the plain-text groups. It follows, therefore, that such a list can serve only for encoding and that, for decoding, another list must be provided in which the code groups are arranged in alphabetical or numerical order and are accompanied by their meanings as given in the encoding section. For this reason a two-part code is often called a *cross-reference code*.

The following brief extracts from typical one-part and two-part codes will serve to illustrate the difference between them:

One-part code.	
ABABD	A
ABACF	Abaft
ABAHK	Abandon
ABAJL	. . . . it
ABALN	Abandoned
ABAMP	. . . . by
ABAWZ	Abandoning
ABBAD	Abandonment
----------	----------
----------	----------
ZYZYZ	Zero

Two-part code.			
Encoding Section		Decoding Section	
GAJVY	A	ABABD	Obstructed
TOGTY	Abaft	ABACF	Term
FEHIL	Abandon	ABAHK	Zero
BAYLT	. . . . it	ABAJL	If it has not
ZYZYZ	Abandoned	ABALN	To be sent by
NYSYZ	. . . . by	ABAMP	Acceding
IFWUZ	Abandoning	ABAWZ	Building
RUMGO	Abandonment	ABBAD	Do not attempt
----------	----------	----------	----------
----------	----------	----------	----------
ABAHK	Zero	ZYZYZ	Abandoned

75. Purposes of the two-part type of code.—*a*. The two-part code is a comparatively recent development in code systems. Its purposes are two-fold: (1) Greater secrecy, and (2) greater accuracy. These two features will now be explained.

b. In a one-part code the plain-text groups progress from A to Z in a regular alphabetical sequence, accompanied by their code groups, also in a regular alphabetical or numerical sequence. If the word ABAFT is represented by a code group whose initial letter is A, or whose initial number is 1, then the word ABANDON will be represented by a group whose initial letter is also A, or whose initial number is also 1. In other words, the enemy cryptanalysts have definite clues to follow in breaking down the code as a direct result of the parallelism existing between the two sequences; the determination of the value of one code group affords definite clues to the value of many other code groups. In a two-part code, however, the word ABAFT

might be represented by a group whose initial letter is T, or whose initial number is 8, and the word ABANDON might be represented by a code group whose initial letter is F, or whose initial number is 3. In other words, the two sequences show no parallelism in the progression and hence the determination of the value of one code group affords no clues to the value of any other group.

c. With regard to the greater accuracy of a two-part code over a one-part code, consider the following pair of phrases which appear in a hypothetical one-part code:

WOVAM Will be ready to attack
WOVEN Will not be ready to attack

Such an arrangement is subject to two sources of error. A code clerk working under great difficulties, in a hurry, may accidentally write down WOVAM instead of WOVEN, as a result of the contiguity of the two sets of letters which are nearly similar in appearance and are so close together on the page that his eye may take the group from the wrong line. Again, on account of the similarity in sound, his ear may deceive him into writing WOVEN when he should have written WOVAM. Now the meaning of the one group is the exact opposite of the meaning of the other and, since either meaning may fit in correctly with the context of the message, the error may remain undiscovered for some time, thus causing serious inconvenience or, in the case of combat, actual loss of life. Furthermore, although the making of two errors in a single group is rather unusual in transmission or reception, yet it does happen and, in such a case as the above, the error would not be detected. The phenomenon referred to is especially true in connection with tabular material such as lists of numbers, dates, names, etc., in which the context often fails to yield clues to the correction of garbles or errors, or to give conclusive evidence of the presence of an error. But in a two-part code such errors are impossible. In the case of the first source of error mentioned above, the code clerk would be very much less likely to confuse two entirely different groups of letters; in the second case, if two errors are made in the transmission or reception and if these errors involve two letters, producing a group which actually has a meaning in the code, this meaning is so unlikely to be such as to fit in correctly with the context that its probability of occurrence may be altogether neglected. Thus, if this sort of error does happen, the meaning of the group fails to fit in with the context and at once indicates that an error is involved. Knowledge of the existence of such an error, even if it is impossible to correct, is a much more preferable condition than ignorance of its existence, with a possible action based upon an erroneous decodement.

d. Two-part codes are used by large governments for their secret diplomatic, military, and naval communications because the two advantages they offer over one-part codes, as explained above, are more than sufficient to compensate for their two disadvantages, now to be explained. A two-part code is, physically, at least twice as large in content as a one-part code, since each code group and each plain-text element must appear twice in such a code. The cost of printing such a code is therefore approximately double that for a one-part code. This, however, by no means constitutes the most important item in the increased cost. The amount of labor involved in compiling a two-part code is much more than double that involved in compiling a one-part code, on account of the necessity for preparing the extremely accurate cross reference arrangement which forms its basic principle. It is perhaps no exaggeration to say that this item is nearly four times greater than in the case of a one-part code. Finally, from the point of view of practicability in handling, a two-part code is twice as bulky as a one-part code of the same vocabulary content, and is therefore not so easily manipulated.

e. As to the sequence, progression, or arrangement of the *phrases* included in the vocabulary of a code, when an absolutely strict alphabetical arrangement is adhered to, the code is said to be a *strictly alphabetical code;* when the phrases are listed under separate headings based upon the principal word or idea in the whole expression, the code is said to be a *caption code.* The following extracts will serve to illustrate the two types:

Caption code	Strictly alphabetical
Assistance	Assistance
Give *assistance*	Assistance for
Require *assistance*	Assistance from
No *assistance* required	Assistance has been sent
Assistance has been sent	Assistance to
Assistance for	Assistant
Assistance from	Assisted
Assistance to	---
Assistant	Give
Assisted	Give assistance
etc.	---
	No
	No assistance required

	Require
	Require assistance

f. A caption code permits, perhaps, of more precise and more economical encoding than does a strictly alphabetical code, because it is easier under the former type of arrangement to assemble under

each specific principal heading a rather extended variety of expressions and different shades of expressions than under the latter type of arrangement. But on the other hand, the use of a caption code involves more time and labor in encoding, especially by untrained or unskilled personnel, than does the use of a strictly alphabetical code. Where the phraseology of communication is quite standardized or stereotypic, all the most common expressions regardless of their length may be listed in a strictly alphabetical code as readily as in a caption. In both types of codes there may be tabulated material of various sorts, such as tables of numbers, dates, equipment, geographical or personal designations, etc., either forming isolated sections in the code or inserted in the vocabulary under appropriate headings.

SECTION XVII

ENCIPHERED CODE

76. Purposes of enciphered code.—*a.* Sometimes the code groups of a code message undergo a further process of encipherment, in which case the resulting cryptogram constitutes an *enciphered code message*. There are two circumstances in which enciphered code is employed. First, if the code book is not secret and it is desirable to transmit a secret message in this code, it becomes necessary to encipher the code groups. Secondly, even if the code book is kept secret, it is desirable in the case of highly secret communications to encipher the messages, in order to increase the degree of security by delaying as long as possible the reduction of the code by the enemy cryptanalysts.

b. An example of a situation in which encipherment of code text is resorted to because the code book itself is not secret is found in the case of commercial codes that can be purchased in book stores. It often happens that it is desirable to impart some secrecy to the plain code message to obviate the possibility of a translation of the message by unauthorized persons. In military cryptography it may also be desirable in special cases to add this factor of safety to messages exchanged between commanders who must employ a code that has been given a wide distribution. Two commanders may find it necessary to communicate with each other secretly in connection with matters affecting subordinate personnel who may also be in possession of the code, and a system of encipherment is agreed upon between them.

c. It has already been indicated that code messages may be solved by cryptanalytic principles without possession of the code. The length of time required for the process varies widely in different cases

and is dependent upon the special conditions surrounding the work, as explained in Section II. In order to increase the length of time required for solution, in the case of secret codes, the code text of the messages resulting from the use of the code is passed through a cipher process so that the messages will be in different keys, thus delaying the assembling and study of the data, which is a prerequisite to the solution.

77. Types of encipherment.—*a.* Both of the two general classes of cipher methods, transposition and substitution, may be employed in enciphering code. The augmented degree of secrecy due to the encipherment depends entirely upon the nature of the system applied.

b. Transposition systems involving a rearrangement of complete groups may be employed where the degree of increased security does not have to be of a high order, and where the original form of the groups must be retained even after encipherment. Transposition systems in which the order of the letters within groups is changed may also be employed. For example, a numerical key such as 3–2–1–5–4, derived from a key word, may serve to indicate the transposed order of the letters of the code groups, so that the group BACRA is transposed into CABAR.

c. Substitution systems of many sorts may be employed, ranging from simple monoalphabetic to the most complex types of substitution, even with cipher machinery. Tables of alphabets are often used. In some systems, a simple transposition process may be combined with a simple substitution process.

d. A favorite method in codes with numbered code groups is that in which the code group which stands 1, 2, 3, . . . *n* places before or after the code group representing the word or phrase intended to be conveyed is substituted for the latter. When the method is one in which a given number is added to the number of the code group to be conveyed and the code group designated by the sum of the two values is transmitted, it is referred to as the *additive method;* when subtraction is involved, it is referred to as the *subtractive method.* Both additive and subtractive methods may be combined into a single system, operated by means of a key word, so that addition and subtraction take place alternately, or at regular or irregular intervals, as controlled by the key.

PART FOUR

CONCLUDING REMARKS

SECTION XVIII

COMPARISON OF CODE AND CIPHER SYSTEMS

78. Advantages and disadvantages of each type of system.—*a.* Each of these two general methods of secret communication has its place in the military service, and both are at present indispensable adjuncts to any real system of signal communication. When and if cipher machines that will meet every requirement necessary in a cryptographic system for use in governmental affairs are finally invented and constructed, it may be that cipher will entirely supersede code in military, naval, and diplomatic correspondence. But so far, no single machine has yet been constructed which will meet all the requirements of simplicity, secrecy, and portability, so that it can be used for all forms of secret communication necessary in the military service. Hence, in the comparisons which follow, only cipher methods operated without machines, or in other words "hand methods," will be considered.

b. The principal factors to be taken into account in comparing code and cipher methods as systems of secret communication are—

(1) Simplicity, rapidity, practicability.
(2) Secrecy.
(3) Accuracy.
(4) Economy.

c. In general it may be said that code is a more rapid, simple, and practicable method of secret communication than cipher, both as regards encoding and decoding. The processes of enciphering and deciphering require very close mental attention to avoid errors, and are usually much slower than those of encoding and decoding which more nearly approach automatic processes and thus require less concentrated mental effort. This is of greatest importance in the combat zone where time is most pressing, and the mental strain and excitement of battle are apt to lead to many errors. What has been said here applies only to cipher methods operated by hand, and not to ciphers produced by an automatic machine or device. There are very small cipher devices which tend to reduce the mental strain to a minimum, but in general the cryptograms they yield are not secure, especially when many messages are available for interception by the enemy.

d. Code systems are, on the whole, more secret than cipher systems, depending upon (1) the extent of the vocabulary and its arrangement; (2) the extent to which the code is used—that is, the number of messages transmitted. *Furthermore, in the case of a code system, the solution of one message does not entail the immediate break-down of the whole system, with the consequent solution of all cipher messages in the same key, as is usually the case in a cipher system.* On the other hand, in the case of a code system it is absolutely necessary to guard at all times the code book, so that it does not fall into the possession of the enemy. Actual possession is not always necessary, for unauthorized sight of the book, with opportunity to copy or memorize certain portions of it, is often sufficient to compromise the whole code. Small codes may be carried about very easily, but then they are all the more likely to fall into the hands of the enemy. Large codes cannot, of course, be carried about so easily and are sometimes inaccessible or unavailable at the most inopportune moments.

e. On the whole, it may be said that code systems are less accurate than cipher systems and are more subject to the necessity for repetition of messages than are cipher systems. This is because a mistake in one or two code groups may obscure, alter, or render unintelligible the meaning of a whole message whereas, in the case of ciphers, the meaning of a few letters which are in error may be supplied by the context. On the other hand, it must be remembered that in some cipher systems a single error of a fundamental type may prevent the deciphering of the message altogether, but this circumstance must be regarded as an unusual occurrence and not as a regular phenomenon inherent in the method.

f. Since code text is usually shorter than the equivalent plain text, on account of the abbreviating features of code, the latter is more economical than cipher. This is of great importance where the amount of traffic is very heavy, and each unnecessary character transmitted requires the time and labor of a large personnel and the uneconomical use of a great amount of equipment. On the other hand, it is true that codes must be prepared, printed, and distributed, processes which take much time and labor and are often attended with considerable difficulties. A continuously operative code compilation section must be maintained to replace codes as fast as they become compromised by continued use, or by capture. The handling of the manuscript, proofs, etc., in printing entails the necessity of ever watchful secrecy; and finally, the difficulties of a prompt and thorough distribution of the codes to all who must use them are sometimes very great, especially where this distribution must be made over an extensive territory. In the long run, therefore, ciphers are possibly more economical than codes, but this increase in degree of economy is probably very low.

g. A thorough understanding of the foregoing items upon which the comparison of code and cipher systems has been made will show clearly why it is that at the present time code methods predominate over cipher methods in military cryptography. Were it not for the fact that code books involve much time, labor, and money in their production and distribution, and that they must at all times be carefully safeguarded, ciphers of the hand-operated type would rarely be employed in military communications. In Section II the principal requirements as regards the practicability of a cipher system for military use were discussed; in the description of various types of cipher systems the requirements as regards secrecy were barely touched upon, and it is necessary to summarize them.

h. It has been seen that every good cipher system combines two more or less separate and distinct elements: (1) a basic or unchangeable method or process, which is termed the general system, and (2) a specific or variable factor which governs or controls the steps under the general system and is termed the specific key. The secrecy of any cipher system for military use must be entirely dependent upon the second of these two elements because it *must be assumed that the enemy is in full possession of all the details concerning the general system.* This assumption is not only warranted by the whole history of military cryptography but is also based upon the two following considerations which all experienced cryptanalysts regard as valid. In the first place, the general circumstances under which cryptography is employed in military operations are such that the enemy has far more prolific sources from which information concerning cryptographic methods may be obtained without his engaging in laborious efforts to solve messages, than is the case in cryptographic methods employed between private individuals in isolated instances. In short, he can sooner or later come into possession of full information regarding the general cryptographic system by one means or another. In the second place, within a very short time the number of messages available for study becomes so great, and the instances where the unforeseen and the perhaps inevitable "blunders" in the handling of communications have become so numerous that a solution by detailed study can more or less readily be accomplished by the enemy, with a consequent disclosure of the general system. In this connection it may be definitely stated that an instance in military cryptography wherein the circumstances just indicated have not been true has yet to be encountered. Now if the cryptographic systems adopted for military use were such that, once the general methods underlying them become known, messages could easily be solved when, with or without a knowledge of the specific keys applicable to them, it is obvious that every time this happened it would entail a change in the

entire system, with the consequent loss of much past training and the necessity for devising a new system, distributing information concerning it to thousands of persons in the military service, and training them in its manipulation, which would naturally be utterly impracticable. This assumption with regard to the enemy's knowledge of the general cryptographic system must also be extended to cover those cases in which cipher devices, instruments or machines are employed. The only limitation in this respect can apply solely to code books and cipher tables of various sorts, which are given a limited distribution and can more or less readily be kept secret from the enemy; but here again it must be admitted that they can be kept so only for a variable length of time before they must undergo internal changes in composition. These changes, however, do not as a rule affect their method of usage. In the case of cipher systems, the specific key must be susceptible of easy and rapid change by prearrangement between correspondents. In the case of systems for use in the theater of operations the key may consist of an easily remembered word, phrase, sentence or number; it must not require the carrying about of written notes on the person. In the case of systems for use in the larger headquarters in the rear, the specific key may be in the form of written memoranda, paper tapes, and the like. Generally speaking, the specific key must be the same throughout a given period of time for all the members of an intercommunicating network, or at least only a very limited number of specific keys must be in simultaneous effect, otherwise confusion and delay are inevitable. This requirement has as a consequence that the enemy may intercept as many as 100 or 200 messages all in the same specific key. Conforming to all the requirements as to practicability as set forth in Section II, and to the foregoing one with respect to the specific key, a cipher system for military use must be of such a nature that despite the enemy's full knowledge of all the details of the general method (or his possession of the cipher device or apparatus, if used) and despite the fact that he may have available for study as many as 100 or more cryptograms all in the same specific key, it must nevertheless be practically impossible for him to decrypt any of the messages within a sufficiently short period of time for the information obtained in this manner to be of any real or immediate value to him in the tactical situation. It goes almost without saying that there is no cipher system yet known which fully meets all these requirements, and that a good code system, employing a well constructed code book, more nearly meets them than does any cipher system thus far known. It is for this reason that code systems predominate in our Army at the present time.

79. Limitations and disadvantages of all methods.—It may be stated that the necessity for cryptographing messages and the dis-

advantages entailed by all the present "hand methods" of cryptography, including both code and cipher systems, constitute the "neck of the bottle" of military signal communication. This seems to be true not only in this country and in all the military, naval and diplomatic services thereof, but also in foreign countries. More and more is the efficiency of combatants becoming dependent upon efficient signal communication, which involves quick, accurate and economical service in the transmission of intelligence. In only the last-named respect can modern cryptography be said to be effective, and then only when large codes are employed. All methods are subject to inaccuracies and many of them are hopelessly slow when action must be speedy in order to be fruitful. The increasingly widespread use of radio, a means of communication that lends itself readily to enemy interception, and the consequent indispensability of cryptography demand the invention of systems involving the elimination of the disadvantages of present known methods. The path along which progress will be made in this regard seems to lie in the direction pointing to highly efficient, automatic, mechanical and electrical cipher machines.

SECTION XIX

CORRECTION OF ERRORS

80. Sources of error in cryptography.— Errors, mutilations and garbles are some of the names applied to the inaccuracies that occur in the execution of all the operations involved in or connected with cryptographic communication. They are so common and so troublesome that commanders who, for the most part, are already prone to regard cryptographic processes as hopelessly slow and cumbersome, often become much prejudiced against their employment in active operations. It is therefore one of the essential parts of the training of personnel assigned to cryptographic work that they receive instruction in the correction of errors. Training and experience will greatly reduce the time necessary to correct the most common types of errors, which may be traced to the following sources:

a. Those made in cryptographing and decryptographing, including the simple process of copying by hand or by typewriter.

b. Those made in transmission and reception by all means of signal communication other than those in which the cryptograms are physically carried from origin to destination.

81. Practical suggestions for eliminating or avoiding errors.—*a.* Errors in cryptographing and decryptographing can be

much reduced, though not wholly eliminated, by systematizing the work so far as possible and *invariably* checking it. Great care in the formation of letters in writing must be exercised, and roman capitals should always be used. Reference is made in this connection to paragraph 227, Basic Field Manual, Vol. IV. If messages have been copied and are to be checked as to correctness of the work by two operators, one reading the letters to the other, a phonetic alphabet must be used in order to prevent misunderstandings. (See par. 48, B. F. M., Vol. IV.) In forward areas it is naturally impossible to provide as suitable or convenient quarters for the personnel engaged in cryptographic work as is desirable, but in rear areas and at the larger headquarters this personnel will be found to work much more efficiently in a quiet, well-ventilated office. In checking cryptographic work it is always advisable, when possible, to check the accuracy by having some operator other than the original cryptographic clerk decryptograph the message. If an operator checks his own work, he should actually decryptograph it—not merely check his cryptographing—because it is a well-known psychological fact that persons have a tendency to repeat an error unconsciously. In this connection it may be stated that the most serious errors in cryptographic work leading to difficulties and delays in decryptographing are not those involving mere mistakes in the writing down of letters, but are of the type in which an error of a fundamental nature is involved and of which the operator says, when it comes back to him, "I don't see how I could have made it." Checking by actually decryptographing will usually eliminate such errors. At the destination, the *final copy* of a decryptographed message should likewise be invariably checked against the *original* work sheets before being turned over to the addressee, and again, preferably, by another operator. It is easy to omit the word NOT from a decoded message and to fail to note the omission, if the same operator merely reads over the decryptographed message. Here, again, psychological factors are involved. Especially careful must those clerks be who are predisposed to transpose letters and words, an unconscious habit of a peculiar psychological origin.

b. Errors in transmission and reception are harder to avoid, especially in transmission by radio, on account of interference, atmospheric disturbances, and the like. For this reason cryptographic clerks should be familiar with the Morse telegraph alphabets and the most common errors of wire and radio transmission methods, so as to be able to refer an error to its probable origin or to find clues for the correction of badly garbled groups when all other means fail. The following table will be found useful:

Most common errors in telegraphic transmission

Continental Morse alphabet (used in radio, cables, and outside United States)			American Morse alphabet (used in the United States, except for radio)		
Letters and figures	Alphabet	Frequent errors	Letters and figures	Alphabet	Frequent errors
A	•—	i, m, t, et	A	•—	i, t, et
B	—•••	d, ts	B	—•••	d, h, ts
C	—•—•	f, k, j, r, nn	C	•••	s, z, ie
D	—••	b, s, l, ti	D	—••	b, ti
E	•	t, a, i	E	•	t
F	••—•	q, r, in	F	•—•	r, q, en
G	——•	m, n, o, q, me	G	——•	n, c, me
H	••••	s, v, b, se	H	••••	s, p, z, y, es
I	••	a, n, s	I	••	a, o, e
J	•———	w, o, eo, am	J	—•—•	c, k, ke
K	—•—	a, n, d, o, ta	K	—•—	j, n, ta
L	•—••	x, r, d, ed	L	——	t, n
M	——	a, n, i, tt	M	——	n, a, tt
N	—•	i, m, t, te	N	—•	o, t, te
O	———	g, k, m, w, mt	O	••	n, i, ee
P	•——•	j, w, g, l, r, an	P	•••••	h, s
Q	——•—	g, k, o, x, z, ma	Q	••—•	f, g, u, in
R	•—•	a, n, f, g, s, l, w	R	•••	s, i, ei
S	•••	h, d, i, r, u, v	S	•••	h, r, i
T	—	a, e, n	T	—	l, e, n
U	••—	a, s, v, it	U	••—	v, a, w, it
V	•••—	h, u, x, st	V	•••—	u, st
W	•——	a, m, o, r, u, at	W	•——	f, a, u, m, at
X	—••—	d, v, u, k, y, tu	X	•—••	l, y, f, ai
Y	—•——	x, w, k, c, nm	Y	•• ••	h, ii
Z	——••	b, d, g, q, mi	Z	••• •	h, c, se
1	•————	0, 2	1	•—•—	p
2	••———	1, 3	2	••—••	3
3	•••——	2, 4	3	•••—•	4
4	••••—	3, 5	4	••••—	3
5	•••••	4, 6	5	———	
6	—••••	5, 7	6	••••••	p
7	——•••	6, 8	7	——••	
8	———••	7, 9	8	—••••	
9	————•	8, 0	9	—••—	x
0	—————	9	0	——	L

c. Every message to be decryptographed should be examined as to its correspondence with the word count or *check* carried as a part of the preamble to a telegram or radiogram; and each group should be examined to see that it has its proper quota of letters—no more and no less. In this connection reference is made to paragraph 228, Basic Field Manual, Vol. IV, which the student will find valuable. *Except in the case of transposition ciphers*, if the final group of a cryptogram lacks 1, 2, 3, . . . letters to make it a complete group, equal in length to all the other groups in the message, it is advisable to make it so by adding the necessary number of nulls. In the case of transposition ciphers, the total number of letters in the plain text should be counted and the nulls, if necessary, added *before cryptographing*, for if

added *after* cryptographing, the message will not yield to quick decryptographing, if it yields at all.

d. Efficiency in cryptographic work requires, in addition to the usual qualities of carefulness, accuracy, and attention to detail, the possession of certain psychological characteristics peculiar to the work. These characteristics can, as a rule, not be developed if initially absent, but they can be intensified and made more efficient by constant practice and experience. It is therefore advisable to select personnel for cryptographic work as for any other specialized work, to train them carefully, and retain them as long as possible, for the longer they remain in this work the less likely are they to repeat the errors with which it abounds and the more likely are they to render highly efficient service.

SECTION XX

FUNDAMENTAL RULES FOR SAFEGUARDING CRYPTOGRAMS

82. Fundamental rules for safeguarding cryptograms.— *a.* There are a few fundamental rules which must be observed in all cryptographic work. Failure to observe such rules will inevitably lead to a more rapid solution by the enemy than would otherwise be the case. *Much of the success which attends the efforts of the cryptanalyst is due to ignorance and carelessness on the part of the clerks who are entrusted with the work of cryptographing and decryptographing messages.* The following general rules would seem to be self-evident, but they are violated every day.

b. A cryptographed message once transmitted must never be repeated in any other key or in any other code or cipher whatsoever. It is only permissible to correct errors of a very minor nature, such as mistakes involving individual letters that have been garbled in transmission or in copying; but any of the following blunders is absolutely fatal, and is invariably discovered by the enemy: (1) to send identical plain-text messages to two different correspondents in two different keys; (2) to reencipher the same message in a different key; (3) to reencipher a message in the same key in case an error involving one letter affects all or many subsequent letters; (4) to encode the same message in a different code; (5) to encipher the same code message in a different key; (6) to send identical or practically identical plain-text messages to two correspondents in two different codes. It is sometimes discovered that a message that has just been transmitted has been cryptographed in the wrong key or in the wrong code. It is fatal to recall or to cancel the message and to transmit the same plain text in the correctly cryptographed form.

What must invariably be done in all cases where a repetition in a different cryptographic form is absolutely unavoidable is to *paraphrase* the message; that is, rewrite it so as to change its original wording as much as possible without changing the meaning of the message. This is done by altering the positions of sentences in the message, by altering the positions of subject, predicate, and modifying phrases or clauses in the sentence, and by altering as much as possible the diction by the use of synonyms and synonymous expressions. In this process, deletion rather than expansion of the wording of the message is preferable because if an ordinary message is paraphrased simply by expanding it along its original lines, an expert can easily reduce the so-paraphrased message to its lowest terms, and the resultant wording will be practically the original message. It is very important to eliminate repeated words or proper names, if at all possible, by the use of carefully selected pronouns; by the use of the words FORMER, LATTER, FIRST-MENTIONED, SECOND-MENTIONED; or by the use of special means provided in the code book. After carefully paraphrasing, the message can be sent in the other key or code.

c. A message once sent in cryptographed form must never be repeated in clear under any circumstances. Vice versa, a message once sent in clear must never be repeated in cryptographed form and, of course, a cryptographed message must never be answered in clear. So far as possible, no information of any kind should ever be given in letters, in plain-text messages, or in a cryptographed message which may be connected in any way directly by verbiage with a cryptographed message previously sent.

d. Never insert or leave plain text of any sort in cryptographed messages. This includes punctuation and abbreviations of any description. They afford valuable clues to the enemy. If a message is to be cryptographed at all, it should be *completely* cryptographed.

e. Plain text and its equivalent cryptographed form must never appear on the same sheet of paper for final copy or for filing purposes. *Work sheets should be destroyed by burning.* Preferably, no verbatim plain-language copy should be retained, but if a copy must be retained, it should be only a paraphrased version.

f. All rules and precautions set forth in the instructions to the various codes and ciphers issued for use must be observed very carefully. These rules have been adopted as a result of experience gained in solving enemy messages during the late war and are intended to delay the solution of our own messages as long as possible. Practice in the preparation of code messages is especially recommended because of the familiarity that is soon gained with the particular words and phrases contained within the book. With familiarity of contents, and speed in operation, the length of the messages may be

reduced very considerably, as well as the time necessary to prepare them. If every cryptogram were checked carefully, and preferably by an operator other than the one who cryptographed the message originally, many errors would be avoided. The best method would be to decryptograph the cryptogram (rather than merely check the cryptographing) and compare the decryptographed version with the original plain text.

g. The more messages sent, and the longer the messages are, the sooner will the enemy be able to solve them. Messages can be materially shortened by the deletion of unnecessary words, punctuation, etc.

h. The formation and adoption of fixed habits as regards the phraseology of messages, arrangement of their contents, use of punctuation, etc., is a most dangerous practice, and will assist the enemy cryptanalysts very greatly. Routine reports of all kinds should be sent by means and agencies not susceptible of interception.

i. The beginnings and endings of all secret messages are cryptographically their weakest spots and are usually the points first to be attacked and solved by the enemy cryptanalysts. If the address and signature of a message must be cryptographed, this should be done according to a different system than is applied to the interior of the message. Sometimes a special address and signature code must be provided for this purpose, which code must not be employed for any other purpose.

j. It may be stated that one of the most important sources of cryptographic information consists in the study of the material furnished newspaper men or given out for publicity in the form of communiques or as information in connection with negotiations conducted by cable or radio. Anything which will enable an alert enemy to compare a given piece of plain text with a cryptogram that supposedly contains this plain text is highly dangerous to the safety of the cryptographic system or code book. Where information must be given out for publicity, or where it is unavoidable that this information be handled by many individuals, *the plain text version should be very carefully paraphrased before distribution*, in order to minimize so far as possible the data an enemy might obtain from an accurate comparison of the cryptographic text with the equivalent, original plain text.

k. Finally, the utmost care should be taken to prevent the loss or unauthorized sight of the codes or lists of cipher keys in use. It is possible to photograph an entire code in two or three hours. Mere continued possession of the code is, therefore, no absolute guaranty that it has not been compromised by photography or some other method of reproduction. The only absolute assurance of its not having been compromised is that it has never left the possession of the person into whose care it has been entrusted or the safe in which it

is kept when not in use. Even if knowledge that a code has been compromised follows immediately after such compromise, the time and difficulties attendant upon the notification of the fact to all concerned and the distribution of a new code are so great that much serious damage is caused by the delay and interruption in communication, not to speak of the danger resulting from the decoding of the most recent messages in the compromised code.

O