Copy No. ________

NATIONAL SECURITY AGENCY

MILITARY CRYPTANALYSIS

Part I

4th Edition

By

WILLIAM F. FRIEDMAN

Revised and enlarged by
LAMBROS D. CALLIMAHOS

National Security Agency
Washington 25, D. C.

December 1952

SECTION VII

MULTILITERAL SUBSTITUTION WITH SINGLE-EQUIVALENT CIPHER ALPHABETS

(In preparation)

(BLANK)

~~RESTRICTED~~

SECTION VIII

MULTILITERAL SUBSTITUTION WITH VARIANTS

(In preparation)

(BLANK)

SECTION IX

POLYGRAPHIC SUBSTITUTION SYSTEMS

(In preparation)

(BLANK)

SECTION X

CONCLUDING REMARKS

(In preparation)

(BLANK)

APPENDIX 1

GLOSSARY

(In preparation)

(BLANK)

APPENDIX 5

LETTER FREQUENCY DATA - FOREIGN LANGUAGES

APPENDIX 6

LIST OF FREQUENT WORDS - ENGLISH AND FOREIGN LANGUAGES

(In preparation)

~~RESTRICTED~~

(BLANK)

~~RESTRICTED~~

APPENDIX 7

CRYPTOGRAPHIC SUPPLEMENT

(In preparation)

(BLANK)

APPENDIX 8

LESTER S. HILL ALGEBRAIC ENCIPHERMENT

(In preparation)

(BLANK)

APPENDIX 9

OPEN CODES AND CONCEALMENT SYSTEMS

(In preparation)

~~RESTRICTED~~

(BLANK)

~~RESTRICTED~~

~~RESTRICTED~~

APPENDIX 10

COMMUNICATION INTELLIGENCE OPERATIONS

APPENDIX 11

PRINCIPLES OF COMMUNICATION SECURITY

(In preparation)

~~RESTRICTED~~

(BLANK)

APPENDIX 12

BIBLIOGRAPHY; RECOMMENDED READING

(In preparation)

~~RESTRICTED~~

(BLANK)

RESTRICTED

APPENDIX 13

PROBLEMS - MILITARY CRYPTANALYSIS, PART I

(In preparation)

RESTRICTED

(BLANK)

APPENDIX 14

FOREIGN LANGUAGE PROBLEMS

(In preparation)

(BLANK)

REF ID:A56895

~~RESTRICTED~~
~~Security Information~~

NATIONAL SECURITY AGENCY

MILITARY CRYPTANALYSIS

Part I

4th Edition

By

WILLIAM F. FRIEDMAN

Revised and enlarged by
LAMBROS D. CALLIMAHOS

National Security Agency
Washington 25, D. C.

December 1952

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

-- Tennyson.

Preface to the 4th Edition

This edition represents an extensive expansion and revision of the original text, in both scope and content, necessitated by the considerable advancement made in the art since the publication of the previous editions.

I wish to express grateful acknowledgment for Mr. Friedman's generous assistance and invaluable collaboration in the preparation of this edition.

-- *L. D. C.*

TABLE OF CONTENTS

MILITARY CRYPTANALYSIS, PART I

Monoalphabetic Substitution Systems

(BLANK)

SECTION I

INTRODUCTORY REMARKS

1. Scope of this text.--a. This text constitutes the first of a series of six basic texts[1] on the art of cryptanalysis. Although most of the information contained herein is applicable to cryptograms of various types and sources, special emphasis will be laid upon the principles and methods of solving military[2] cryptograms. Except for an introductory discussion of fundamental principles underlying the science of cryptanalytics, this first text in the series will deal solely with the principles and methods for the analysis of monoalphabetic substitution ciphers. Even with this limitation it will be possible to discuss only a few of the many variations of this one type that are met in practice; but with a firm grasp upon the general principles few difficulties should be experienced with any modifications or variations that may be encountered.

b. This and some of the succeeding texts will deal only with basic types of cryptosystems not because they may be encountered unmodified in military operations but because their study is essential to an understanding of the principles underlying the solution of the modern, very much more complex types of codes, ciphers, and certain encrypted transmission systems that are likely to be employed by the larger governments of today in the conduct of their military affairs in time of war.

c. It is presupposed that the student has no prior background in the field of cryptology; therefore cryptography is presented concurrently with cryptanalysis. Basic terminology and preliminary cryptologic considerations are treated in Section II; other terms are usually defined upon their first occurrence, or they may be found in the Glossary (Appendix 1).

d. The cryptograms presented in the examples embrace messages from hypothetical air, ground, and naval traffic; thus, the student will have the opportunity to familiarize himself with the language and phraseology of all three Services comprising the Armed Forces of the United States.

---

[1] Each text has its accompanying course in cryptanalysis, so that the student may test his learning and develop his skill in the solution of the types of cryptograms treated in the respective texts. The problems which pertain to this text constitute Appendix 13.

[2] The word "military" is here used in its broadest sense. In this connection see subpar. d, below.

2. Mental equipment necessary for cryptanalytic work.--a. Captain Parker Hitt, in the first United States Army manual[3] dealing with cryptology, opens the first chapter of his valuable treatise with the following sentence:

"Success in dealing with unknown ciphers is measured by these four things in the order named: perseverance, careful methods of analysis, intuition, luck."

These words are as true today as they were then. There is no royal road to success in the solution of cryptograms. Hitt goes on to say:

"Cipher work will have little permanent attraction for one who expects results at once, without labor, for there is a vast amount of purely routine labor in the preparation of frequency tables, the rearrangement of ciphers for examination, and the trial and fitting of letter to letter before the message begins to appear."

The present author deems it advisable to add that the kind of work involved in solving cryptograms is not at all similar to that involved in solving crossword puzzles, for example. The wide vogue the latter have had and continue to have is due to the appeal they make to the quite common interest in mysteries of one sort or another; but in solving a crossword puzzle there is usually no necessity for performing any preliminary labor, and palpable results become evident after the first minute or two of attention. This successful start spurs the crossword "addict" on to complete the solution, which rarely requires more than an hour's time. Furthermore, crossword puzzles are all alike in basic principles and once understood, there is no more to learn. Skill comes largely from the embellishment of one's vocabulary, though, to be sure, constant practice and exercise of the imagination contribute to the ease and rapidity with which solutions are generally reached. In solving cryptograms, however, many principles must be learned, for there are many different systems of varying degrees of complexity. Even some of the simpler varieties require the preparation of tabulations of one sort or another, which many people find irksome; moreover, it is only toward the very close of the solution that results in the form of intelligible text become evident. Often, indeed, the student will not even know whether he is on the right track until he has performed a large amount of preliminary "spade work" involving many hours of labor. Thus, without at least a willingness to pursue a fair amount of theoretical study, and a more than average amount of patience and perseverance, little skill and experience can be gained in the rather difficult art of cryptanalysis. General Givierge, the author of an excellent treatise on cryptanalysis, remarks in this connection:[4]

"The cryptanalyst's attitude must be that of William the Silent: No need to hope in order to undertake, nor to succeed in order to persevere."

---

[3] Hitt, Capt. Parker, Manual for the Solution of Military Ciphers. Army Service Schools Press, Fort Leavenworth, Kansas, 1916. 2d Edition, 1918. (Both out of print.)

[4] Givierge, Général Marcel, Cours de Cryptographie, Paris, 1925, p. 301.

b. As regards Hitt's reference to careful methods of analysis, before one can be said to be a cryptanalyst worthy of the name it is necessary that one should have, firstly, a sound knowledge of the basic principles of cryptanalysis, and secondly, a long, varied, and active practical experience in the successful application of those principles. It is not sufficient to have read treatises on this subject. One month's actual practice in solution is worth a whole year's mere reading of theoretical principles. An exceedingly important element of success in solving the more intricate cryptosystems is the possession of the rather unusual mental faculty designated in general terms as the power of inductive and deductive reasoning. Probably this is an inherited rather than an acquired faculty; the best sort of training for its emergence, if latent in the individual, and for its development is the study of the natural sciences, such as chemistry, physics, biology, geology, and the like. Other sciences such as linguistics, archaeology, and philology are also excellent.

c. Aptitude in mathematics is quite important, more especially in the solution of ciphers and enciphered codes than in codebook reconstruction, which latter is purely and simply a linguistic problem. Although in the early days of the emergence of the science of cryptanalytics little thought was given to the applications of mathematics in this field, many branches of mathematics and, in particular, probability and statistics, have now found cryptologic applications. Those portions of mathematics and those mathematical methods which have cryptologic applications[5] are known collectively as cryptomathematics.

[5] It is quite important to stress at this point that in professional cryptologic work the science of cryptanalytics is subordinated to the art of cryptanalysis, just as in the world of music the technical virtuosity of a great violinist is adjuvant to the expression of music, that is, the virtuosity is a "tool" for the recovery of the complete musical "plain text" conceived by the composer. Since the practice of cryptanalysis is an art, mathematical approaches cannot always be expected to yield a solution in cryptology, because art can and must transcend the cold logic of scientific method. By way of example, an experienced Indian guide can usually find his way out of a dense forest more readily than a surveyor equipped with all the refined apparatus and techniques of his profession. Likewise, an experienced cryptanalyst can generally find his way through a cryptosystem more readily than a pure mathematician equipped merely with the techniques of his field no matter how abstruse or refined they may be. A cryptomathematician of repute once stated that "the only effect of [refined mathematical techniques] is frequently to discourage one so much that one does nothing at all and some unmathematical ignoramus then gets the problem out in some very unethical way. This is intensely irritating." See also in this connection the remarks made in subpar. 27e in reference to the validity of statistical tests in cryptanalysis.

d. An active imagination, or perhaps what Hitt and other writers call intuition, is essential, but mere imagination uncontrolled by a judicious spirit will be more often a hindrance than a help. In practical cryptanalysis the imaginative or intuitive faculties must, in other words, be guided by good judgment, by practical experience, and by as thorough a knowledge of the general situation or extraneous circumstances that led to the sending of the cryptogram as is possible to obtain. In this respect the many cryptograms exchanged between correspondents whose identities and general affairs, commercial, social, or political, are known are far more readily solved[6] than are isolated cryptograms exchanged between unknown correspondents, dealing with unknown subjects. It is obvious that in the former case there are good data upon which the intuitive powers of the cryptanalyst can be brought to bear, whereas in the latter case no such data are available. Consequently, in the absence of such data, no matter how good the imagination and intuition of the cryptanalyst, these powers are of no particular service to him. Some writers, however, regard the intuitive spirit as valuable from still another viewpoint, as may be noted in the following:[7]

"Intuition, like a flash of lightning, lasts only for a second. It generally comes when one is tormented by a difficult decipherment and when one reviews in his mind the fruitless experiments already tried. Suddenly the light breaks through and one finds after a few minutes what previous days of labor were unable to reveal."

This, too, is true, but unfortunately there is no way in which the intuition may be summoned at will, when it is most needed.[8] There are certain authors who regard as indispensable the possession of a somewhat

[6] The application in practical, operational cryptanalysis of "probable words" or "cribs", i.e., plain text assumed or known to be present in a cryptogram, is developed in time of war into a refinement the extent and usefulness of which cannot be appreciated by the uninitiated. Even as great a thinker as Voltaire found the subject of cryptanalysis stretching his credulity to the point that he said:

"Those who boast that they can decipher a letter without knowing its subject matter, and without preliminary aid, are greater charlatans than those who would boast of understanding a language which they have never learned."--Dictionnaire Philosophique, under the article "Poste".

[7] Lange et Soudart, Traité de Cryptographie, Libraire Félix Alcan, Paris, 1925, p. 104.

[8] The following extracts are of interest in this connection:

"The fact that the scientific investigator works 50 per cent of his time by non-rational means is, it seems, quite insufficiently recognized. There is without the least doubt an instinct for research, and often the most successful investigators of nature are quite unable to give an account of their reasons for doing such and such an experiment, or for placing side by side two apparently unrelated facts. Again, one of the most salient traits in the character of the successful scientific worker is the capacity for knowing that a point is proved when it would not appear to be proved to an outside intelligence functioning in a purely rational manner; thus the investigator feels that some proposition is true, and proceeds at once to the next set of experiments without waiting and wasting time in the elaboration of the formal proof of the point which heavier minds would need. Questionless such a scientific intuition may and does sometimes lead investigators astray, but it is quite certain that if they did not widely make use of it, they would not get a quarter as far as they do. Experiments confirm each other, and a rare, rather mysterious faculty that they designate by the word "flair", or by the expression "cipher brains". Even so excellent an authority as General Givierge,[9] in referring to this mental faculty, uses the following words:

"Over and above perseverance and this aptitude of mind which some authors consider a special gift, and which they call intuition, or even, in its highest manifestation, clairvoyance, cryptographic studies will continue more and more to demand the qualities of orderliness and memory."

Although the present author believes a special aptitude for the work is essential to cryptanalytic success, he is sure there is nothing mysterious about the matter at all. Special aptitude is prerequisite to success in all fields of endeavor. There are, for example, thousands of physicists, hundreds of excellent ones, but only a handful of world-wide fame. Should it be said, then, that a physicist who has achieved very notable success in his field has done so because he is the fortunate possessor of a mysterious faculty? That he is fortunate in possessing a special aptitude for his subject is granted, but that there is anything mysterious about it, partaking of the nature of clairvoyance (if, indeed, the latter is a reality) is not granted. While the ultimate nature of any mental process seems to be as complete a mystery today as it has ever been, the present author would like to see the superficial veil of mystery removed from a subject that has been shrouded in mystery from even before the Middle Ages down to our own times. (The principal and readily understandable reason for this is that governments have always closely guarded cryptographic secrets and anything so guarded soon becomes "mysterious".) He would, rather, have the student approach the subject as he might approach any other science that can stand on its own merits with other sciences, because cryptanalytics, like other sciences, has a practical importance in human affairs. It presents to the inquiring mind an interest in its own right as a branch of knowledge; it, too, holds forth many difficulties and disappointments, and these are all the more

---

false step is usually soon discovered. And not only by this partial replacement of reason by intuition does the work of science go on, but also to the born scientific worker—and emphatically they cannot be made—the structure of the method of research is as it were given, he cannot explain it to you, though he may be brought to agree *a posteriori* to a formal logical presentation of the way the method works".—Excerpt from Needham, Joseph, *The Sceptical Biologist*, London, 1929, p. 79.

"The essence of scientific method, quite simply, is to try to see how data arrange themselves into causal configurations. Scientific problems are solved by collecting data and by "thinking about them all the time." We need to look at strange things until, by the appearance of known configurations, they seem familiar, and to look at familiar things until we see novel configurations which make them appear strange. We must look at events until they become luminous. That is scientific method . . . Insight is the touchstone . . . The application of insight as the touchstone of method enables us to evaluate properly the role of imagination in scientific method. The scientific process is akin to the artistic process: it is a process of selecting out those elements of experience which fit together and recombining them in the mind. Much of this kind of research is simply a ceaseless mulling over, and even the physical scientist has considerable need of an armchair . . . Our view of scientific method as a struggle to obtain insight forces the admission that science is half art . . . Insight is the unknown quantity which has eluded students of scientific method".—Excerpts from an article entitled *Insight and Scientific Method*, by Willard Waller, in *The American Journal of Sociology*, Vol. XL, 1934.

[9] Op. cit., p. 302.

keenly felt when the nature of these difficulties is not understood by those unfamiliar with the special circumstances that very often are the real factors that led to success in other cases. Finally, just as in the other sciences wherein men labor long and earnestly for the true satisfaction and pleasure that comes from work well done, so the mental pleasure that the successful cryptanalyst derives from his accomplishments is very often the only reward for much of the drudgery that he must do in his daily work. General Givierge's words in this connection are well worth quoting:[10]

"Some studies will last for years before bearing fruit. In the case of others, cryptanalysts undertaking them never get any result. But, for a cryptanalyst who likes the work, the joy of discoveries effaces the memory of his hours of doubt and impatience."

e. With his usual deft touch, Hitt says of the element of luck, as regards the role it plays in analysis:

"As to luck, there is the old miners' proverb: 'Gold is where you find it.'"

The cryptanalyst is lucky when one of the correspondents whose cryptograms he is studying makes a blunder that gives the necessary clue; or when he finds two cryptograms identical in text but in different keys in the same system; or when he finds two cryptograms identical in text but in different systems, and so on. The element of luck is there, to be sure, but the cryptanalyst must be on the alert if he is to profit by these lucky "breaks".

f. If the present author were asked to state, in view of the progress in the field since 1916, what elements might be added to the four ingredients Hitt thought essential to cryptanalytic success, he would be inclined to mention the following:

(1) A broad, general education, embodying interests covering as many fields of practical knowledge as possible. This is useful because the cryptanalyst is often called upon to solve messages dealing with the most varied of human activities, and the more he knows about these activities, the easier his task.

(2) Access to a large library of current literature, and wide and direct contacts with sources of collateral information. These often afford clues as to the contents of specific messages. For example, to be able instantly to have at his disposal a newspaper report or a personal report of events described or referred to in a message under investigation goes a long way toward simplifying or facilitating solution. Government cryptanalysts are sometimes fortunately situated in this respect, especially where various agencies work in harmony.

(3) Proper coordination of effort. This includes the organization of cryptanalytic personnel into harmonious, efficient teams of cooperating individuals.

---

[10] Op. cit., p. 301.

(4) Under mental equipment he would also include the faculty of being able to concentrate on a problem for rather long periods of time, without distraction, nervous irritability, and impatience. The strain under which cryptanalytic studies are necessarily conducted is quite severe and too long-continued application has the effect of draining nervous energy to an unwholesome degree, so that a word or two of caution may not here be out of place. One should continue at work only so long as a peaceful, calm spirit prevails, whether the work is fruitful or not. But just as soon as the mind becomes wearied with the exertion, or just as soon as a feeling of hopelessness or mental fatigue intervenes, it is better to stop completely and turn to other activities, rest, or play. It is essential to remark that systematization and orderliness of work are aids in reducing nervous tension and irritability. On this account it is better to take the time to prepare the data carefully, rewrite the text if necessary, and so on, rather than work with slipshod, incomplete, or improperly arranged material.

(5) A retentive memory is an important asset to cryptanalytic skill, especially in the solution of codes. The ability to remember individual groups, their approximate locations in other messages, the associations they form with other groups, their peculiarities and similarities, saves much wear and tear of the mental machinery, as well as much time in looking up these groups in indexes.

(6) The assistance of machine aids in cryptanalysis. The importance and value of these aids cannot be overemphasized in their bearing on practical, operational cryptanalysis, especially in the large-scale effort that would be made in time of war on complex, high-grade cryptosystems at a theater headquarters or in the zone of the interior. These aids, under the general category of rapid analytical machines, comprise both punched-card tabulating machinery and certain other general- and special-purpose high-speed electrical and electronic devices. Some of the more compact equipment may be employed by lower echelons within a theater of operations to facilitate the cryptanalysis of medium-grade cryptosystems found in tactical communications.

g. It may be advisable to add a word or two at this point to prepare the student to expect slight mental jars and tensions which will almost inevitably come to him in the conscientious study of this and the subsequent texts. The present author is well aware of the complaint of students that authors of texts on cryptanalysis base much of their explanation upon their foreknowledge of the "answer"--which the student does not know while he is attempting to follow the solution with an unbiased mind. They complain, too, that these authors use such expressions as "it is obvious that", "naturally", "of course", "it is evident that", and so on, when the circumstances seem not at all to warrant their use. There is no question that this sort of treatment is apt to discourage the student, especially when the point elucidated becomes clear to <u>him</u> only after many hours' labor, whereas, according to the book, the author noted the weak spot at the first moment's inspection. The present author can only promise to try to avoid making the steps appear to be much more simple than they really are, and to suppress glaring instances of unjustifiable "jumping at conclusions". At the same time he must indicate that for pedagogical reasons in many cases a message has been consciously "manipulated" so as to allow certain principles to become more obvious in the illustrative examples than they ever are in practical work. During the course of some of the explanations attention will even be directed to cases of unjustified inferences. Furthermore, of the student who is quick in observation and deduction, the author will only ask that he bear in mind that if the elucidation of certain principles seems prolix and occupies more space than necessary, this is occasioned by the author's desire to carry the explanation forward in very short, easily-comprehended, and plainly-described steps, for the benefit of students who are perhaps a bit slower to grasp but who, once they understand, are able to retain and apply principles slowly learned just as well, if not better than the students who learn more quickly.[11]

3. Validity of results of cryptanalysis.--Valid or authentic cryptanalytic solutions cannot and do not represent "opinions" of the cryptanalyst. They are valid only so far as they are wholly objective, and are susceptible of demonstration and proof, employing authentic, objective methods. It should hardly be necessary (but an attitude frequently encountered among laymen makes it advisable) to indicate that the validity of the results achieved by any serious cryptanalytic studies on authentic material rests upon the same sure foundations and are reached by the same general steps as the results achieved by any other scientific studies; viz., observation, hypothesis, deduction and induction, and confirmatory experiment. Implied in the latter is the possibility that two or more qualified investigators, each working independently upon the same material, will achieve identical (or practically identical) results--there is one and only one (valid) solution to a cryptogram. Occasionally a "would-be" or pseudo-cryptanalyst offers "solutions" which cannot withstand such tests; a second, unbiased, investigator working independently either cannot consistently apply the methods alleged to have been applied by the pseudo-cryptanalyst, or else, if he can apply

---

[11] In connection with the use of the word "obvious", the following extract is of interest:

"Now the word 'obvious' is a rather dangerous one. There is an incident, which has become something of a legend in mathematical circles, that illustrates this danger. A certain famous mathematician was lecturing to a group of students and had occasion to use a formula which he wrote down with the remark, 'This statement is obvious.' Then he paused and looked rather hesitantly at the formula. 'Wait a moment,' he said. 'Is it obvious? I think it's obvious.' More hesitation, and then, 'Pardon me, gentlemen, I shall return.' Then he left the room. Thirty-five minutes later he returned; in his hands was a sheaf of papers covered with calculations, on his face a look of quiet satisfaction. 'I was right, gentlemen. It is obvious,' he said, and proceeded with his lecture."--Excerpt from The Anatomy of Mathematics by Kershner and Wilcox. New York, 1950.

them at all, the results (plaintext translations) are far different in the two cases. The reason for this is that in such cases it is generally found that the "methods" are not clear-cut, straightforward or mathematical in character. Instead, they often involve the making of judgments on matters too tenuous to measure, weigh, or otherwise subject to careful scrutiny. Often, too, they involve the "correction" of an inordinate number of "errors" which the pseudo-cryptanalyst assumes to be present and which he "corrects" in order to make his "solution" intelligible. And sometimes the pseudo-cryptanalyst offers as a "solution" plain text which is intelligible only to him or which he makes intelligible by expanding what he alleges to be abbreviations, and so on. In all such cases, the conclusion to which the unprejudiced observer is forced to come is that the alleged "solution" obtained by the pseudo-cryptanalyst is purely subjective.[12] In nearly all cases where this has happened (and they occur from time to time) there has been uncovered nothing which can in any way

[12] A mathematician is often unable to grasp the concept behind the expression "subjective solution" as used in the cryptanalytic field, since the idea is foreign to the basic philosophy of mathematics and thus the expression appears to him to represent a contradiction in terms. As an illustration, let us consider a situation in which a would-be cryptanalyst offers a solution to a cryptogram he alleges to be a simple monoalphabetic substitution cipher. His so-called solution, however, requires that he assume the presence of, let us say, approximately 50% garbles (which he claims to have been introduced by cipher clerks' errors, faulty radio reception because of adverse weather conditions, etc.). That is, the "plain text" he offers as the "solution" involves his making helter-skelter many "corrections and emendations", which, one may be sure, will be based on what his subconscious mind expects or desires to find in the cleartext message. Unfortunately, another would-be cryptanalyst working upon the same cryptogram and hypothesis independently might conceivably "degarble" the cryptogram in different spots and produce an entirely dissimilar "plain text" as his "solution". Both "solutions" would be invalid because they are based upon an erroneous hypothesis--the cryptogram actually happens to be a polyalphabetic substitution cipher which when correctly analyzed requires on the part of unbiased observers no assumption of garbles to a degree that strains their credulity. The last phrase is added here because in professional cryptanalytic work it is very often necessary to make a few corrections for errors but it is rarely the case that the garble rate exceeds more than a few percent of the characters of the cryptogram, say 5 to 10% at the outside. It is to be noted, however, that occasionally the solution to a cryptogram may involve the correction of more than this percentage of errors, but the solution would be regarded as valid only if the errors can be shown to be <u>systematic</u> in some significant respect, or can otherwise be explained by <u>objective</u> rationalization.

be used to impugn the integrity of the pseudo-cryptanalyst. The worst that can be said of him is that he has become a victim of a special or peculiar form of self-delusion, and that his desire to solve the problem, usually in accord with some previously-formed opinion, or notion, has over-balanced, or undermined, his judgment and good sense.[13]

---

[13] Specific reference can be made to the following typical "case histories":

Donnelly, Ignatius, The Great Cryptogram. Chicago, 1888.
Owen, Orville W., Sir Francis Bacon's Cipher Story. Detroit, 1895.
Gallup, Elizabeth Wells, Francis Bacon's Biliteral Cipher. Detroit, 1900.
Arensberg, Walter Conrad, The Cryptography of Shakespeare. Los Angeles, 1922.
The Shakespearean Mystery. Pittsburgh, 1928.
The Baconian Keys. Pittsburgh, 1928.
Margoliouth, D. S., The Homer of Aristotle. Oxford, 1923.
Newbold, William Romaine, The Cipher of Roger Bacon. Philadelphia, 1928. (For a scholarly and complete demolition of Professor Newbold's work, see an article entitled Roger Bacon and the Voynich MS, by John M. Manly, in Speculum, Vol. VI, No. 3, July 1931.)
Feely, Joseph Martin, The Shakespearean Cypher. Rochester, N. Y., 1931.
Deciphering Shakespeare. Rochester, N. Y., 1934.
Roger Bacon's Cypher: the right key found. Rochester, N. Y., 1943.
Wolff, Werner, Déchiffrement de l'Ecriture Maya. Paris, 1938.
Strong, Leonell C., Anthony Askham, the author of the Voynich manuscript, in Science, Vol. 101, June 15, 1945, pp. 608-9.

SECTION II

BASIC CRYPTOLOGIC CONSIDERATIONS

4. Cryptology, communication intelligence, and communication security. The occasional or frequent need for secrecy in the conduct of important affairs has been recognized from time immemorial. In the case of diplomacy and organized warfare this need is especially important in regard to communications. However, when such communications are transmitted by electrical means, they can be heard and copied by unauthorized persons. The protection resulting from all measures designed to deny to unauthorized persons information of value which may be derived from such communications is called communication security. The evaluated information concerning the enemy, derived principally from a study of his electrical communications, is called communication intelligence. The collective term including all phases of communication intelligence and communication security is cryptology.[1] Or, stated in broad terms, cryptology is that branch of knowledge which treats of hidden, disguised, or secret[2] communications.

---

[1] From the Greek kryptos (hidden) + logos (learning). The prefix "crypto-" in compound words pertains to "cryptologic", "cryptographic", or "cryptanalytic", depending upon the use of the particular word as defined.

[2] In this text the term "secret" will be used in its ordinary sense as given in the dictionary. Whenever the designation is used in the more restricted sense of the security classification as defined in official regulations, it will be capitalized. There are in current use the four classifications Restricted, Confidential, Secret, and Top Secret, listed in ascending order of degree.

5. Secret communication.--a. Communication may be conducted by any means susceptible of ultimate interpretation by one of the five senses, but those most commonly used are sight and hearing. Aside from the use of simple visual and auditory signals for communication over relatively short distances, the usual method of communication between or among individuals separated from one another by relatively long distances involves, at one stage or another, the act of writing or of speaking over a telephone.

b. Privacy or secrecy in communication by telephone can be obtained by using equipment which affects the electrical currents involved in telephony so that the conversations can be understood only by persons provided with suitable equipment properly arranged for the purpose. The same thing is true in the case of electrical transmission of pictures, drawings, maps, and television images. However, this text will not treat of these aspects[3] of cryptology.

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, the characters are inscribed by means or methods which make the writing invisible to the naked eye. Invisible writing can be prepared with certain chemicals called invisible, sympathetic, or secret inks, and in order to "develop" such writing, that is, make it visible, special processes must usually be applied. There are also methods of producing writing which is invisible to the naked eye because the characters are of microscopic size, thus requiring special photographic or microscopic apparatus to make such writing visible to the naked eye.

d. Invisible writing and unintelligible visible writing constitute secret writing.

6. Plain text and encrypted text.--a. Visible writing which is intelligible, that is, conveys a more or less understandable or sensible meaning (in the language in which written) and which is not intended to convey a hidden meaning, is said to be in plain text.[4] A message in plain text is termed a plaintext message, a cleartext message, or a message in clear.

[3] These aspects of cryptology are now known as ciphony (from cipher + telephony); cifax (from cipher + facsimile); and civision (from cipher + television).

[4] Visible writing may be intelligible but the meaning it obviously conveys may not be its real meaning, that is, the meaning intended to be conveyed. To quote a simple example of an apparently innocent message containing a secret or hidden meaning, prepared with the intention of escaping censorship, the sentence "Son born today" may mean "Three transports left today." Messages of this type are said to be in open code. Secret communication methods or artifices of this sort (concealment systems) are impractical for field military use but are often encountered in espionage and counter-espionage activities.

b. Visible writing which conveys no intelligible meaning in any recognized language[5] is said to be in encrypted text and such writing is termed a cryptogram.[6]

7. Cryptography, encrypting, and decrypting.--a. Cryptography is that branch of cryptology which treats of various means, methods, and apparatus for converting or transforming plaintext messages into cryptograms and for reconverting the cryptograms into their original plaintext forms by a simple reversal of the steps used in their transformation.

b. To encrypt is to convert or transform a plaintext message into a cryptogram by following certain rules, steps, or processes constituting the key or keys and agreed upon in advance by correspondents, or furnished them by higher authority.

c. To decrypt is to reconvert or to transform a cryptogram into the original equivalent plaintext message by a direct reversal of the encrypting process, that is, by applying to the cryptogram the key or keys (usually in a reverse order) used in producing the cryptogram.

d. A person skilled in the art of encrypting and decrypting, or one who has a part in devising a cryptographic system is called a cryptographer; a clerk who encrypts and decrypts, or who assists in such work, is called a cryptographic clerk.

8. Codes, ciphers, and enciphered code.--a. Encrypting and decrypting are accomplished by means collectively designated as codes and ciphers. Such means are used for either or both of two purposes: (1) secrecy, and (2) economy or brevity. Secrecy usually is far more important in military cryptography than economy or brevity. In ciphers or cipher systems, cryptograms are produced by applying the cryptographic treatment to individual letters of the plaintext messages, whereas, in codes or code systems, cryptograms are produced by applying the cryptographic treatment to entire words, phrases, and sentences of the plaintext messages. The specialized meanings of the terms code and cipher are explained in detail later.

b. A cryptogram 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 act or operation of encrypting a cipher message is called enciphering,

---

[5] There is a certain type of writing which is considered by its authors to be intelligible, but which is either completely unintelligible to the wide variety of readers or else requires considerable mental struggle on their part to make it intelligible. Reference is here made to so-called "modern literature" and "modern verse", products of such writers as E. E. Cummings, Gertrude Stein, James Joyce, et al.

[6] From kryptos + gramma (that which is written). Analogous terminology would call a plaintext message a phanerogram (phaneros = visible, manifest, open).

and the enciphered version of the plain text, as well as the act or process itself, is often referred to as the encipherment. The cryptographic clerk who performs the process serves as an encipherer. The corresponding terms applicable to the decrypting of cipher messages are deciphering, decipherment, and decipherer. A clerk who serves as both an encipherer and decipherer of messages is called a cipher clerk.

c. A cipher device is a relatively simple mechanical contrivance for encipherment and decipherment, usually "hand-operated" or manipulated by the fingers, as for example a device with concentric rings of alphabets, manually powered; a cipher machine is a relatively complex apparatus or mechanism for encipherment and decipherment, usually equipped with a typewriter key board and often requiring an external power source.

d. A cryptogram produced by means of a code system is said to be in code and is called a code message. The text of the cryptogram is referred to as code text. This act or operation of encrypting is called encoding, and the encoded version of the plain text, as well as the act or process itself, is referred to as the encodement. The clerk who performs the process serves as an encoder. The corresponding terms applicable to the decrypting of code messages are decoding, decodement, and decoder. A cryptographic clerk who serves as both an encoder and decoder of messages is called a code clerk.

e. Sometimes, for special purposes (usually increased security), the code text of a cryptogram undergoes a further step in concealment involving superencryption, that is, encipherment of the characters comprising the code text, thus producing what is called an enciphered-code message, or enciphered code. Encoded cipher, that is, the case where the final cryptogram is produced by enciphering the plain text and then encoding the cipher text obtained from the first operation, is also possible, but rare.

9. General system, specific key, and cryptosystem.--a. There are a great many different methods of encrypting messages, so that correspondents must first of all be in complete agreement as to which of them will be used in their secret communications, or in different types or classes of such communications. Furthermore, it is to be understood that all the detailed rules, processes, or steps comprising the cryptography agreed upon will be invariant, that is, constant or unvarying in their use in a given set of communications. The totality of these basic, invariable rules, processes, or steps to be followed in encrypting a message according to the agreed method constitutes the general cryptographic system or, more briefly, the general system.

b. It is usually the case that the general system operates in connection with or under the control of a number, a group of letters, a word, a phrase, or sentence which is used as a key, that is, the element which specifically governs the manner in which the general system will be applied in a specific message, or the exact setting of a cipher device or a cipher machine at the initial point of encipherment or decipherment of a specific message. This element--usually of a variable nature or changeable at the will of the correspondents, or prearranged for them by higher authority--is called the specific key. The specific key may also involve the use of a set of specially prepared tables, a special document, or even a book.

c. The term cryptosystem[7] is used when it is desired to designate or refer to all the cryptomaterial (device, machine, instructions for use, key lists, etc.) as a unit to provide a single, complete system and means for secret communication.

10. Cryptanalytics and cryptanalysis.--a. In theory any cryptosystem (except one[8]) can be "broken", i.e., solved, if enough time, labor, and skill are devoted to it, and if the volume of traffic in that system is large enough. This can be done even if the general system and the specific key are unknown at the start. In military operations theoretical rules must usually give way to practical considerations. How the theoretical rule in this case is affected by practical considerations will be discussed in Appendix 11.

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

c. The steps and operations performed in applying the principles of cryptanalytics constitute cryptanalysis. To cryptanalyze 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.

11. Transposition and substitution.--a. Technically there are only two distinct types of treatment which may be applied to written 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 retain their original identities and merely undergo some change in their relative positions, with the result that the original text becomes unintelligible. In the second, called substitution, the elements of the plain text retain their original relative positions but are replaced by other elements with different values or meanings, with the result that the original text becomes unintelligible. Thus, in the case of transposition ciphers, the unintelligibility is brought about merely by a change in the original sequence of the elements or units of

---

7 The term cryptosystem is used in preference to cryptographic system so as to permit its use in designating secret communication systems involving means other than writing, such as ciphony and cifax.

8 The exception is the "one-time" system in which the key is used only once and in itself must have no systematic construction, derivation, or meaning.

the plain text; in the case of substitution ciphers, the unintelligibility is brought about by a change in the elements or units themselves, without a change in their relative order.

b. It is possible to encrypt 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 class 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 limited.[9]

c. Under each of the two principal classes of cryptograms as outlined above, a further classification can be made based upon the number of characters composing the textual elements or units undergoing cryptographic treatment. These textual units are composed of (1) individual letters, (2) combinations of letters in regular groupings, (3) combinations of letters in irregular, more or less euphonious groupings called syllables, and (4) complete words, phrases, and sentences. Methods which deal with the first type of units are called monographic methods; those which deal with the second type are called polygraphic (digraphic, trigraphic, etc.); those which deal with the third type, or syllables, are called syllabic; and, finally, those which deal with the fourth type are called lexical (of or pertaining to words).

d. It is necessary to indicate that the foregoing classification of cryptographic methods is more or less 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, regular or irregular-length groupings of letters, syllables, words, phrases, and complete sentences. When in a single system the cryptographic treatment is applied to textual units of regular length, usually monographic or digraphic (and seldom longer, or intermixed monographic and digraphic), the system is called a cipher system. Likewise, when in a single system the cryptographic treatment is applied to textual units of irregular length, usually syllables, whole words, phrases, and sentences, and is only exceptionally applied to single letters or regular groupings of letters, the system is called a code system and generally involves the use of a code book.[10]

12. Nature of alphabets.--a. One of the simplest kinds of substitution ciphers is that which is known in cryptologic literature as Julius Caesar'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

---

[9] One notable exception is the ADFGVX system, used extensively by the Germans in World War I. See in this connection the Cryptographic Supplement (Appendix 7).

[10] A list of single letters, frequent digraphs, trigraphs, syllables, and words is often called a syllabary; cryptographic treatment of the units of such syllabaries places them in the category of code systems.

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 is cipher.

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

c. The letters comprising the English alphabet used today are the results of a long period of evolution, the complete history of which may never fully be known.[11] They are 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 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. Again, 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 we pronounce the letters A, B, and C, respectively; they would then write the "word" HFG, pronounce it cab, and see nothing strange whatever in the matter. But to others no party to their arrangements, 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 others who are reading it according to their own long-established system of sound and symbol equivalents. It would be stated that there is no such word as HFG, which would 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 the English language. Thus, it is seen 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, secondly, 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.

---

[11] An excellent and most authoritative book on this subject is The Alphabet; a key to the history of Mankind by David Diringer. London, 1949.

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 the result is what is called 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, one 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 considerably extended and the difficulties set in the path of the enemy cryptanalysts greatly increased. The chances for its adoption in the near future are, however, quite small. Because of the continually changing nature of every living language, it is doubtful whether an initially "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.

13. Types of 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	....

Normal alphabets of different languages vary considerably in 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; the Spanish, 27 (including the digraphs CH and LL); and the Russian, 31.[12] The Japanese language has a syllabary consisting of 72 syllabic sounds which require 48 characters for their representation.

---

[12] In contrast to the foregoing alphabets, it is of interest to note that in the Hawaiian language the alphabet consists of only 12 letters, viz, the five vowels A, E, I, O, U, and the seven consonants H, K, L, M, N, P, W.

c. A cipher alphabet, or substitution alphabet as it is sometimes called, 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. When the plain text of a message is converted into encrypted text by the use of one or more cipher alphabets, the resultant cryptogram constitutes a substitution cipher. If only one cipher alphabet is involved, it is called a monoalphabetic substitution cipher; if two or more cipher alphabets are involved, it is called a polyalphabetic substitution cipher.

e. It is convenient to designate that component of a cipher alphabet constituting the sequence of speech-sounds as the plain component and the component constituting the sequence of symbols as the cipher component. If omitted in a cipher alphabet, the plain component is understood to be the normal sequence. For brevity and clarity, a letter of the plain text, or of the plain component of a cipher alphabet, is 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.

f. With reference to the arrangement or sequence of letters forming their components, cipher alphabets are of two types:

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

(2) Mixed cipher alphabets, in which the sequence of letters or characters in one or both of the components is no longer the same as the normal in its entirety.

g. Although the basic considerations of the preceding paragraphs place the student in a position to undertake the study of certain fundamental principles of cryptanalysis, this may be a good point at which to pause and to make a few remarks with regard to the role that cryptanalysis plays in the whole chain of more or less complex operations involved in deriving communication intelligence, after which these fundamental cryptanalytic principles will be treated.

(BLANK)

SECTION III

FUNDAMENTAL CRYPTANALYTIC OPERATIONS

14. The role of cryptanalysis in communication intelligence operations.--a. Through the medium of communication intelligence an attempt is made to answer three questions concerning enemy communications: "Who?" "Where?" "What?"--Who are their originators and addressees? Where are these originators and addressees located? What do the messages say?

b. All of the foregoing questions are very important in the military application of communication intelligence. Hence, even though this text deals almost exclusively with the principles and operations involved in deriving the answer to the third question--"What do the messages say?"--a few words on the importance of the first and second questions may be useful. It is a serious mistake to think that one can necessarily and always correctly interpret the mere text of a message without identifying and locating the originator and the addressee or, on many occasions, without having a background against which to interpret the message in order to appreciate its real import or significance.

c. The very first step in the series of activities involved in deriving communication intelligence is the collection of the raw material, that is, the interception[1] and copying of the transmissions constituting the messages to be studied and analyzed.

d. Then, with the raw material in hand, studies are made in order to answer the first two questions--"Who?" and "Where?" The answers to these questions are not always obvious in modern military communications, especially in the case of messages exchanged by units in the combat zone, since messages of this sort rarely indicate in plain language who the

[1] To intercept means, in its cryptologic sense, to gain possession of communications which are intended for other recipients, without obtaining the consent of these addressees and without preventing or ordinarily delaying the transmission of the communications to them.

originator and the addressee are or where they are located. Consequently, certain apparatus and techniques specifically developed for finding the answers to these questions must be employed. These apparatus and techniques are embraced by that part of communication intelligence theory and practice which is known as traffic analysis. This latter subject and interception are treated briefly in Appendix 10, "Communication intelligence operations". (The serious student will derive much practical benefit from a careful reading of this appendix.)

e. The foregoing operations, interception and traffic analysis, along with cryptanalysis constitute the first three operations of communication intelligence. But generally there must follow at least one additional operation. If the plain texts recovered through cryptanalysis are in a foreign language, they must usually be translated, and translation constitutes this fourth operation. In the course of translating, it may be found that, because of errors in transmission or reception, corrections and emendations must be made in these plain texts; however, although this often requires skill and experience of a high order, it does not constitute another communication intelligence operation, since it is but an auxiliary step to the process of translation.

f. In a large-scale communication intelligence effort these four steps, interception, traffic analysis, cryptanalysis, and translation, must be properly organized and coordinated in order to gain the most benefit from the potentialities of communication intelligence, that is, the production of the maximum quantity of information from the raw traffic. This information must then be evaluated by properly trained intelligence specialists, collated with intelligence derived from other sources, and, finally, disseminated to the commanders who need the intelligence in time to be of operational use to them, rather than of mere historical interest. The foregoing operations and especially the first three--interception, traffic analysis, and cryptanalysis--usually complement one another. This, however, is not the place for elaboration on the interrelationships which exist and which when properly integrated make the operations as a whole an efficient, unified complex geared to the fulfillment of its principal goal, namely, the production of timely communication intelligence.

g. With the foregoing general background, the student is prepared to proceed to the technical considerations and principles of cryptanalysis.

15. The four basic operations in cryptanalysis.--a. The solution of practically every cryptogram involves four fundamental operations or steps:

(1) The determination of the language employed in the plaintext version.

(2) The determination of the general system of cryptography employed.

(3) The reconstruction of the specific key in the case of a cipher system, or the reconstruction, partial or complete, of the code book, in the case of a code system; or both, in the case of an enciphered code system.

(4) The reconstruction or establishment of the plain text.

b. These operations will be taken up in the order in which they are given above and in which they usually are performed in the solution of cryptograms, although occasionally the second step may precede the first.[1]

16. The determination of the language employed.--a. There is not much that need be said with respect to this operation except that the determination of the language employed seldom comes into question in the case of studies made of the cryptograms of an organized enemy. By this is meant that during wartime the enemy is of course known, and it follows, therefore, that the language he employs in his messages will almost certainly be his native or mother tongue. Only occasionally nowadays is this rule broken. Formerly it often happened, or it might have indeed been the general rule, that the language used in diplomatic correspondence was not the mother tongue, but French. In isolated instances during World War I the Germans used English when their own language could for one reason or another not be employed. For example, for a year or two before the entry of the United States into that war, during the time America was neutral and the German Government maintained its embassy in Washington, some of the messages exchanged between the Foreign Office in Berlin and the Embassy in Washington were encrypted in English, and a copy of the code used was deposited with the Department of State and our censor. Another instance is found in the case of certain Hindu conspirators who were associated with and partially financed by the German Government in 1915 and 1916; they employed English as the language of their cryptographic messages. Occasionally the cryptograms of enemy agents may be in a language different from that of the enemy. But in general these

---

[1] Although the foregoing four steps represent the classical or ideal approach to cryptanalysis, the art may be reduced to the following:

	Procedures in cryptanalysis	Requirements
1.	Arrangement and rearrangement of data to disclose non-random characteristics or manifestations (i.e., in frequency counts, repetitions, patterns, symmetrical phenomena, etc.).	Experience or ingenuity, and time (which latter may be appreciably lowered by the use of machine aids in cryptanalysis).
2.	Recognition of the non-random characteristics or manifestations when disclosed.	Experience or statistics.
3.	Explanation of the non-random characteristics when recognized.	Experience or imagination, and intelligence.

In all of the foregoing, the element of luck plays a very important part, as it is possible to side-step a large amount of labor and effort, in many cases, if "hunches" or intuition lead the analyst forthwith to the right path. Therefore, the phrase "or luck" should be added to each of the requirements above.

In fact, it all boils down to the simple statement: "Find something significant, and attach some significance thereto."

are, as has been said, isolated instances; as a rule, the language used in cryptograms exchanged between members of large organizations is the mother tongue of the correspondents. Where this is not the case, that is, when cryptograms of unknown origin must be studied, the cryptanalyst looks for any indications on the cryptograms themselves which may lead to a conclusion as to the language employed. Address, signature, and other data, if in plain text in the preamble, in the body, or at the end of the cryptogram, all come under careful scrutiny, as well as all extraneous circumstances connected with the manner in which the cryptograms were obtained, the person on whom they were found, or the locale of their origin and destination.

b. In special cases, or under special circumstances a clue to the language employed is found in the nature and composition of the cryptographic text itself. For example, if the letters K and W are entirely absent or appear very rarely in messages, it may indicate that the language is Spanish, for these letters are absent in the alphabet of that language and are used only to spell foreign words or names. The presence of accented letters or letters marked with special signs of one sort or another, peculiar to certain languages, will sometimes indicate the language used. The Japanese Morse telegraph alphabet and the Russian Morse telegraph alphabet contain combinations of dots and dashes which are peculiar to those alphabets and thus the interception of messages containing these special Morse combinations at once indicates the language involved. Finally, there are certain peculiarities of alphabetic languages which, in certain types of cryptograms, viz., pure transposition, give clues as to the language used. For example, the frequent digraph CH, in German, leads to the presence, in cryptograms of the type mentioned, of many isolated C's and H's; if this is noted, the cryptogram may be assumed to be in German.

c. In some cases it is perfectly possible to perform certain steps in cryptanalysis before the language of the cryptogram has been definitely determined. Frequency studies, for example, may be made and analytic processes performed without this knowledge, and by a cryptanalyst wholly unfamiliar with the language even if it has been identified, or who knows only enough about the language to enable him to recognize valid combinations of letters, syllables, or a few common words in that language. He may, after this, call to his assistance a translator who may not be a cryptanalyst but who can materially aid in making necessary assumptions based upon his special knowledge of the characteristics of the language in question. Thus, cooperation between cryptanalyst and translator results in solution.[2]

---

[2] The writer has seen in print statements that "during World War I . . . . decoded messages in Japanese and Russian without knowing a word of either language." The extent to which such statements are exaggerated will soon become obvious to the student. Of course, there are occasional instances in which a mere clerk with quite limited experience may be able to "solve" a message in an extremely simple system in a language of which he has no knowledge at all; but such a "solution" calls for nothing more arduous than the ability to recognize pronounceable combinations of vowels and consonants—an ability that hardly deserves to be rated as "cryptanalytic" in any real sense. To say that it is possible to solve a cryptogram in a foreign language "without knowing a word of that language" is not quite the same as to say that it is possible to do so with only a slight knowledge of the language; and it may be stated without cavil that the better the cryptanalyst's knowledge of the language, the greater are the chances for his success and, in any case, the easier is his work.

17. The determination of the general system.--a. Except in the case of the more simple types of cryptograms, the step referred to as diagnosis, that is, ascertaining the general system according to which a given cryptogram has been produced is usually a difficult, if not the most difficult, step in its solution. The reason for this is not hard to find.

b. As will become apparent to the student as he proceeds with his study, in the final analysis, the solution of every cryptogram involving a form of substitution depends upon its reduction to monoalphabetic terms, if it is not originally in those terms. This is true not only of ordinary substitution ciphers, but also of combined substitution-transposition ciphers, and of enciphered code. If the cryptogram must be reduced to monoalphabetic terms, the manner of its accomplishment is usually indicated by the cryptogram itself, by external or internal phenomena which become apparent to the cryptanalyst as he studies the cryptogram. If this is impossible, or too difficult, the cryptanalyst must, by one means or another, discover how to accomplish this reduction, by bringing to bear all the special or collateral information he can get from all the sources at his command. If both these possibilities fail him, there is little left but the long, tedious, and often fruitless process of elimination. In the case of transposition ciphers of the more complex type, the discovery of the basic method is often simply a matter of long and tedious elimination of possibilities. For cryptanalysis has unfortunately not yet attained, and may indeed never attain, the precision found today in qualitative analysis in chemistry, for example, where the analytic process is absolutely clear-cut and exact in its dichotomy. A few words in explanation of what is meant may not be amiss. When a chemist seeks to determine the identity of an unknown substance, he applies certain specific reagents to the substance and in a specific sequence. The first reagent tells him definitely into which of two primary classes the unknown substance falls. He then applies a second test with another specific reagent, which tells him again quite definitely into which of two secondary classes the unknown substance falls, and so on, until finally he has reduced the unknown substance to its simplest terms and has found out what it is. In striking contrast to this situation, cryptanalysis affords exceedingly few "reagents" or tests that may be applied to determine positively that a given cipher belongs to one or the other of two systems yielding externally similar results. And this is what makes the analysis of an isolated, complex cryptogram so difficult. Note the limiting adjective "isolated" in the foregoing sentence, for it is used advisedly. It is not often that the general system fails to disclose itself or cannot be discovered by painstaking investigation when there is a great volume of text accumulating from a regular traffic between numerous correspondents in a large organization. Sooner or later the system becomes known, either because of blunders and carelessness on the part of the personnel entrusted with the encrypting of the messages, or because the accumulation of text itself makes possible the determination of the general system by cryptanalytic, including statistical, studies. But in the case of a single or even a few isolated cryptograms concerning which little or no information can be gained by the cryptanalyst, he is often unable, without a knowledge of, or a shrewd guess as to the general system employed, to decompose the heterogeneous text of the cryptogram into homogeneous, monoalphabetic text, which is the ultimate and essential step in analysis. The only knowledge that the cryptanalyst can bring to his aid in this most difficult step is that gained by long experience and practice in the analysis of many different types of systems. In this respect the practice of cryptanalysis is analogous to the practice of medicine: correct diagnosis is the most important and often the most difficult first step toward success.

c. On account of the complexities surrounding this particular phase of cryptanalysis, and because in any scheme of analysis based upon successive eliminations of alternatives the cryptanalyst can only progress as far as the extent of his own knowledge of all the possible alternatives will permit, it is necessary that detailed discussion of the eliminative process be postponed until the student has covered most of the field. For example, the student will perhaps want to know at once how he can distinguish between a cryptogram that is in code or enciphered code from one that is in cipher. It is at this stage of his studies impracticable to give him any helpful indications on his question. In return it may be asked of him why he should expect to be able to do this in the early stages of his studies when often the experienced expert cryptanalyst is baffled on the same score!

d. Nevertheless, in lieu of more precise diagnostic tests not yet discovered, a general guide that may be useful in cryptanalysis will be built up, step by step as the student progresses, in the form of a series of charts comprising what may be designated An Analytical Key for Cryptanalysis. (See Section X.) It may be of assistance to the student if, as he proceeds, he will carefully study the charts and note the place which the particular cipher he is solving occupies in the general cryptanalytic panorama. These charts admittedly constitute only very brief outlines, and can therefore be of but little direct assistance to him in the analysis of the more complex types of cryptosystems he may encounter later on. So far as they go, however, they may be found to be quite useful in the study of elementary cryptanalysis. For the experienced cryptanalyst they can serve only as a means of assuring that no possible step or process is inadvertently overlooked in attempts to solve a difficult cryptosystem.

e. Much of the labor involved in cryptanalytic work, as referred to in par. 2, is connected with this determination of the general system. The preparation of the text, its rewriting in different forms, sometimes being rewritten in dozens of ways, the recording of letters, the establishment of frequencies of occurrences of letters, comparisons and experiments made with known material of similar character, and so on, constitute much labor that is most often indispensable, but which sometimes turns out to have been wholly unnecessary, or in vain. In one treatise[3] it is stated quite boldly that "this work once done, the determination of the system is often relatively easy." This statement can certainly apply only to the simpler types of cryptosystems; it is entirely misleading as regards the much more frequently encountered complex cryptograms of modern times.

18. The reconstruction of the specific key.--a. Nearly all practical cryptographic methods require the use of a specific key to guide, control, or modify the various steps under the general system. Once the latter has been disclosed, discovered, or has otherwise come into the possession of the cryptanalyst, the next step in solution is to determine, if necessary and if possible, the specific key that was employed to encrypt the message or messages under examination. This determination may not be in complete detail; it may go only so far as to lead to a knowledge of the number of alphabets involved in a substitution cipher, or the number of columns involved in a transposition cipher, or that a one-part code has been used, in the case of a code system. But it is often desirable to determine the specific key in as complete a form and with as much detail as possible, for this information will very frequently be useful in the solution of subsequent cryptograms exchanged between the same correspondents, since the nature or source of the specific key in a solved case may be expected to give clues to the specific key in an unsolved case.

b. Frequently, however, the reconstruction of the key is not a prerequisite to, and does not constitute an absolutely necessary preliminary step in, the fourth basic operation, viz., the reconstruction or establishment of the plain text. In many cases, indeed, the two processes are carried along simultaneously, the one assisting the other, until in the final stages both have been completed in their entireties. In still other cases the reconstruction of the specific key may follow the reconstruction of the plain text instead of preceding it and is accomplished purely as a matter of academic interest; or the specific key may, in unusual cases, never be reconstructed.

19. The reconstruction of the plain text.--a. Little need be said at this point on this phase of cryptanalysis. The process usually consists, in the case of substitution ciphers, in the establishment of equivalency between specific letters of the cipher text and the plain text, letter by letter, pair by pair, and so on, depending upon the particular type of substitution system involved. In the case of transposition ciphers, the process consists in rearranging the elements of the cipher text, letter by letter, pair by pair, or occasionally word by word, depending upon the particular type of transposition system involved, until the letters or words have been returned to their original plaintext order. In the case of code, the process consists in determining the meaning of each code group and inserting this meaning in the code text to reestablish the original plain text.

---

[3] Lange et Soudart, op. cit., p. 106.

b. The foregoing processes do not, as a rule, begin at the beginning of a message and continue letter by letter, or group by group in sequence up to the very end of the message. The establishment of values of cipher letters in substitution methods, or of the positions to which cipher letters should be transferred to form the plain text in the case of transposition methods, comes at very irregular intervals in the process. At first only one or two values scattered here and there throughout the text may appear; these then form the "skeletons" of words, upon which further work, by a continuation of the reconstruction process, is made possible; in the end the complete or nearly complete[4] text is established.

c. In the case of cryptograms in a foreign language, the translation of the solved messages is a final and necessary step, but is not to be considered as a cryptanalytic process. However, it is commonly the case that the translation process will be carried on simultaneously with the cryptanalytic, and will aid the latter, especially when there are lacunae which may be filled in from the context. (See also subpar. 16c in this connection.)

20. The utilization of traffic intercepts.[5]--a. There are, of course, other operations which are not as basic in nature as those just outlined but which must generally be performed as preliminary steps in practical cryptanalytic work (as distinguished from academic cryptanalysis). Before a military cryptanalyst can begin the analysis of an enemy cryptosystem, it is necessary for him to study the intercept material that is available to him, isolate the messages that hve been encrypted by means of the cryptosystem to be exploited, and to arrange the latter in a systematic order for analysis. This work, although apparently very simple, may require a great deal of time and effort.

b. Since, whenever practicable, two or more intercept stations are assigned to copy traffic emanating from the stations of one enemy radio net, it is natural that there should be a certain amount of duplication in the work of the several stations. This is desirable since it provides the cryptanalysts with two or more sets of the same messages, so that when one intercept station fails to receive all the messages completely and correctly, because of radio difficulties, local static, or poor operation, it is possible by studying the other sets to reconstruct accurately the entire traffic of the enemy net.

---

[4] Sometimes in the case of code, the meaning of a small percentage of the code groups occurring in the traffic may be lacking, because there is insufficient text to establish their meaning.

[5] A traffic intercept is a copy of a communication gained through interception.

c. In all intercept activities where operators are used for copying the traffic, one of the most likely errors to be found is caused by the human element in reception. For this reason cryptanalysts and their

Ltrs. and Figs.	Morse equivalent	Frequent Errors	Ltrs. and Figs.	Morse equivalent	Frequent Errors
A	•-	i, m, t, et	S	•••	h, d, i, r, u
B	-•••	d, ts	T	-	a, e, n
C	-•-•	f, k, r, nn	U	••-	a, s, v, it
D	-••	b, s, l, ti	V	•••-	h, u, x, st
E	•	t, i	W	•--	a, m, o, r, u, at
F	••-•	r, in	X	-••-	v, k, y, tu
G	--•	m, o, z, me	Y	-•--	x, c, nm
H	••••	s, v, b, ii, se	Z	--••	b, g, q, mi
I	••	a, n, s	1	•----	∅, 2
J	•---	w, o, am, eo	2	••---	1, 3
K	-•-	d, o, ta	3	•••--	2, 4
L	•-••	r, d, ed	4	••••-	3, 5
M	--	a, n, tt	5	•••••	4, 6
N	-•	i, m, t, te	6	-••••	5, 7
O	---	g, k, w, mt	7	--•••	6, 8
P	•--•	j, g, l, w, an	8	---••	7, 9
Q	--•-	o, x, z, ma	9	----•	8, ∅
R	•-•	a, f, g, l, n, s, w	∅	-----	9, 1

Chart 1. Most common errors in telegraphic transmission.

assistants should be familiar with the international Morse alphabet and the most common errors in wire and radio transmission methods so as to be able to correct garbled groups when they occur. In this connection, Chart 1, above, will be found useful.

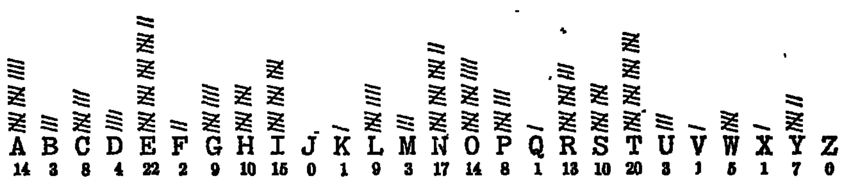
(BLANK)

~~RESTRICTED~~

SECTION IV

FREQUENCY DISTRIBUTIONS AND THEIR FUNDAMENTAL USES

21. The simple or uniliteral frequency distribution.--a. It has long been known to cryptographers and typographers that the letters composing the words of any intelligible written text composed in any language which is alphabetic in construction are employed with greatly varying frequencies. For example, if on cross-section paper a simple tabulation, shown in Fig. 1, called a uniliteral frequency distribution, is made of the letters composing the words of the preceding sentence, the variation in frequency is strikingly demonstrated. It is seen that whereas certain letters, such as A, E, I, N, O, R, and T, are employed very frequently, other letters, such as C, G, H, L, P, and S are employed not nearly so frequently, while still other letters, such as F, J, K, Q, V, X, and Z are employed either seldom or not at all.



(Total=200 letters)

Figure 1.

b. If a similar tabulation is now made of the letters comprising the words of the second sentence in the preceding subparagraph, the distribution shown in Fig. 2 is obtained. Both sentences have exactly the same number of letters (200).

~~RESTRICTED~~

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
12	2	8	7	25	7	4	5	20	0	1	9	5	17	14	6	2	13	14	17	5	1	2	1	3	0

(Total=200 letters)

Figure 2.

c. Although each of these two distributions exhibits great variation in the relative frequencies with which different letters are employed in the respective sentences to which they apply, no marked differences are exhibited between the frequencies of the same letter in the two distributions. Compare, for example, the frequencies of A, B, C . . . Z in Fig. 1 with those of A, B, C . . . Z in Fig. 2. Aside from one or two exceptions, as in the case of the letter F, these two distributions agree rather strikingly.

d. This agreement, or similarity, would be practically complete if the two texts were much longer, for example, five times as long. In fact, when two texts of similar character, each containing more than 1,000 letters, are compared, it would be found that the respective frequencies of the 26 letters composing the two distributions show only very slight differences. This means, in other words, that in normal plain text each letter of the alphabet occurs with a rather constant or characteristic frequency which it tends to approximate, depending upon the length of the text analyzed. The longer the text (within certain limits), the closer will be the approximation to the characteristic frequencies of letters in the language involved. However, when the amount of text being analyzed has reached a substantial volume (roughly, 1,000 letters), the practical gain in accuracy does not warrant further increase in the amount of text.[1]

e. An experiment along these lines will be convincing. A series of 260 official telegrams[2] passing through the Department of the Army Message Center was examined statistically. The messages were divided into five sets, each totaling 10,000 letters, and the five distributions shown in Table 1-A, were obtained.

---

1 See footnote 5, page 38.

2 These comprised messages from several official sources in addition to the Department of the Army and were all of an administrative character.

TABLE 1-A.—*Absolute frequencies of letters appearing in five sets of Governmental plain-text telegrams, each set containing 10,000 letters, arranged alphabetically*

Set No. 1		Set No. 2		Set No. 3		Set No. 4		Set No. 5	
Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency
A	738	A	788	A	681	A	740	A	741
B	104	B	103	B	98	B	83	B	99
C	319	C	300	C	288	C	326	C	301
D	387	D	413	D	423	D	451	D	448
E	1,367	E	1,294	E	1,292	E	1,270	E	1,275
F	253	F	287	F	308	F	287	F	281
G	166	G	175	G	161	G	167	G	150
H	310	H	351	H	335	H	349	H	349
I	742	I	750	I	787	I	700	I	697
J	18	J	17	J	10	J	21	J	16
K	36	K	38	K	22	K	21	K	31
L	365	L	393	L	333	L	386	L	344
M	242	M	240	M	238	M	249	M	268
N	786	N	794	N	815	N	800	N	780
O	685	O	770	O	791	O	756	O	762
P	241	P	272	P	317	P	245	P	260
Q	40	Q	22	Q	45	Q	38	Q	30
R	760	R	745	R	762	R	785	R	736
S	658	S	583	S	585	S	628	S	604
T	936	T	879	T	894	T	958	T	928
U	270	U	233	U	312	U	247	U	238
V	163	V	173	V	142	V	183	V	155
W	166	W	163	W	186	W	183	W	182
X	43	X	50	X	44	X	53	X	41
Y	191	Y	155	Y	179	Y	213	Y	229
Z	14	Z	17	Z	2	Z	11	Z	5
Total	10,000		10,000		10,000		10,000		10,000

f. If the five distributions in Table 1-A are summed, the results are as shown in Table 2-A.

TABLE 2-A.—*Absolute frequencies of letters appearing in the combined five sets of messages totaling 50,000 letters, arranged alphabetically*

A	3,683	G	819	L	1,821	Q	175	V	766
B	487	H	1,694	M	1,237	R	3,788	W	780
C	1,534	I	3,676	N	3,975	S	3,058	X	231
D	2,122	J	82	O	3,764	T	4,595	Y	967
E	6,498	K	148	P	1,335	U	1,300	Z	49
F	1,416								

g. The frequencies noted in Table 2-A above, when reduced to the basis of 1,000 letters and then used as a basis for constructing a simple chart that will exhibit the variations in frequency in a striking manner, yield the following distribution which is hereafter designated as the normal or standard uniliteral frequency distribution for English telegraphic plain text:

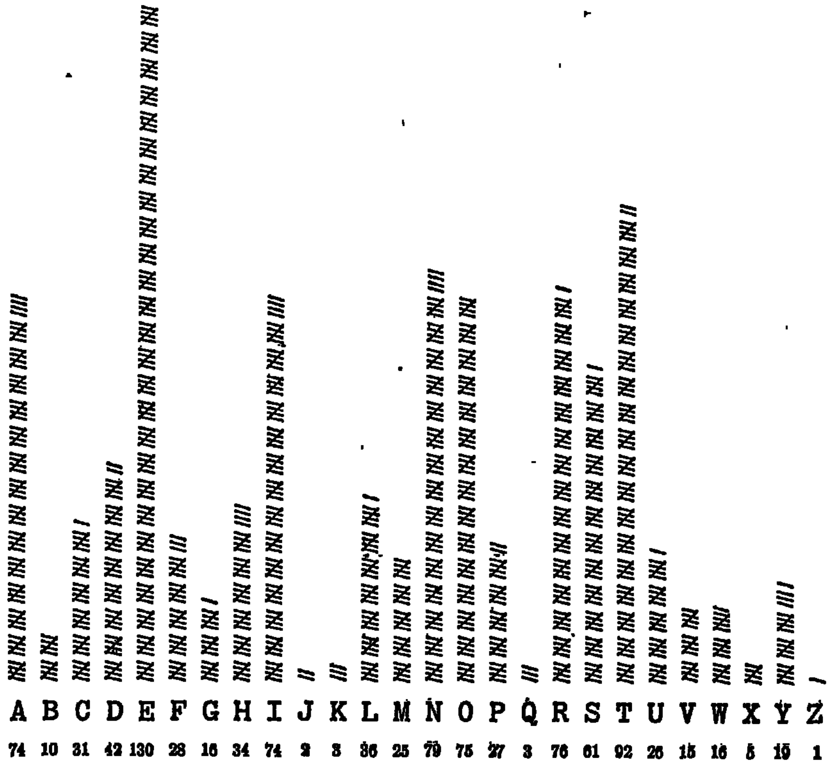


Figure 3.

22. Important features of the normal uniliteral frequency distribution.--a. When the distribution shown in Fig. 3 is studied in detail, the following features are apparent:

(1) It is quite irregular in appearance. This is because the letters are used with greatly varying frequencies, as discussed in the preceding paragraph. This irregular appearance is often described by saying that the distribution shows marked crests and troughs, that is, points of high frequency and low frequency.

(2) The relative positions in which the crests and troughs fall within the distribution, that is, the spatial relations of the crests and troughs, are rather definitely fixed and are determined by circumstances which have been explained in subpar. 13b.

(3) The relative heights and depths of the crests and troughs within the distribution, that is, the linear extensions of the lines marking the respective frequencies, are also rather definitely fixed, as would be found if an equal volume of similar text were analyzed.

(4) The most prominent crests are marked by the vowels A, E, I, O, and the consonants N, R, S, T; the most prominent troughs are marked by the consonants J, K, Q, X, and Z.

(5) The important data are summarized in tabular form in Table 3.

TABLE 3

	Frequency	Percent of total	Percent of total in round numbers
6 Vowels: A E I O U Y	398	39.8	40
20 Consonants:			
5 High Frequency (D N R S T)	350	35.0	35
10 Medium Frequency (B C F G H L M P V W)	238	23.8	24
5 Low Frequency (J K Q X Z)	14	1.4	1
Total	1,000	100.0	100

(6) The frequencies of the letters of the alphabet, reduced to a base of 1000, are as follows:

A	74	G	16	L	36	Q	3	V	15
B	10	H	34	M	25	R	76	W	16
C	31	I	74	N	79	S	61	X	5
D	42	J	2	O	75	T	92	Y	19
E	130	K	3	P	27	U	26	Z	1
F	28								

(7) The relative order of frequency of the letters is as follows:

E	130	I	74	C	31	Y	19	X	5
T	92	S	61	F	28	G	16	Q	3
N	79	D	42	P	27	W	16	K	3
R	76	L	36	U	26	V	15	J	2
O	75	H	34	M	25	B	10	Z	1
A	74								

(8) The four vowels A, E, I, O (combined frequency 353) and the four consonants N, R, S, T (combined frequency 308) form 661 out of every 1,000 letters of plain text; in other words, <u>less than one-third of the alphabet is employed in writing two-thirds of normal plain text</u>.

b. The data given in Fig. 3 and Table 3 represent the relative frequencies found in a large volume of English telegraphic text of a governmental, administrative character.[3] These frequencies will vary somewhat with the nature of the text analyzed. For example, if an equal number of telegrams dealing solely with commercial transactions in the leather industry were studied statistically, the frequencies would be slightly different because of the repeated occurrence of words peculiar to that industry. Again, if an equal number of telegrams dealing solely with military messages of a tactical character were studied statistically, the frequencies would differ slightly from those found above for general governmental messages of an administrative character.

c. If ordinary English literary text (such as may be found in any book, newspaper, or printed document) were analyzed, the frequencies of certain letters would be changed to an appreciable degree. This is because in telegraphic text words which are not strictly essential for intelligibility (such as the definite and indefinite articles, certain prepositions, conjunctions, and pronouns) are omitted. In addition, certain essential words, such as "stop", "period", "comma", and the like, which are usually indicated in written or printed matter by symbols not easy to transmit telegraphically and which must, therefore, be spelled out in telegrams, occur very frequently. Furthermore, telegraphic text often employs longer and more uncommon words than does ordinary newspaper or book text.

d. As a matter of fact, other tables compiled from Army sources gave slightly different results, depending upon the source of the text. For example, three tables based upon 75,000, 100,000, and 136,257 letters taken from various sources (telegrams, newspapers, magazine articles, books of fiction) gave as the relative order of frequency for the first 10 letters the following:

For 75,000 letters................ E T R N I O A S D L
For 100,000 letters............... E T R I N O A S D L
For 136,257 letters............... E T R N A O I S L D

[3] Just as the individual letters constituting a large volume of plain text have more or less characteristic or fixed frequencies, so it is found that digraphs and trigraphs (two- and three-letter combinations, respectively) have characteristic frequencies, when a large volume of text is studied statistically. In Table 6 of Appendix 2, "Letter frequency data - English", are shown the relative frequencies of all digraphs appearing in the 260 telegrams referred to in subpar. 21e. This appendix also includes several other kinds of tables and lists of frequency data which will be useful to the student in his work. It is suggested that the student refer to this appendix now, to gain an idea of the data available for his future reference.

Other languages, of course, each have their own individual characteristic plaintext frequencies of single letters, digraphs, trigraphs, etc. A brief summary of the letter frequency data for German, French, Italian, Spanish, Portuguese, and Russian constitute Appendix 5, "Letter frequency data - foreign languages".

e. Frequency data applicable purely to English military text were compiled by Hitt,[4] from a study of 10,000 letters taken from orders and reports. The frequencies found by him are given in Tables 4 and 5.

TABLE 4.—*Frequency table for 10,000 letters of literary English, as compiled by Hitt*

ALPHABETICALLY ARRANGED

A	778	G	174	L	372	Q	8	V	112
B	141	H	595	M	288	R	651	W	176
C	296	I	667	N	686	S	622	X	27
D	402	J	51	O	807	T	855	Y	196
E	1,277	K	74	P	223	U	308	Z	17
F	197								

ARRANGED ACCORDING TO FREQUENCY

E	1,277	R	651	U	308	Y	196	K	74
T	855	S	622	C	296	W	176	J	51
O	807	H	595	M	288	G	174	X	27
A	778	D	402	P	223	B	141	Z	17
N	686	L	372	F	197	V	112	Q	8
I	667								

TABLE 5.—*Frequency table for 10,000 letters of telegraphic English, as compiled by Hitt*

ALPHABETICALLY ARRANGED

A	813	G	201	L	392	Q	38	V	136
B	149	H	386	M	273	R	677	W	166
C	306	I	711	N	718	S	656	X	51
D	417	J	42	O	844	T	634	Y	208
E	1,319	K	88	P	243	U	321	Z	6
F	205								

ARRANGED ACCORDING TO FREQUENCY

E	1,319	S	656	U	321	F	205	K	88
O	844	T	634	C	306	G	201	X	51
A	813	D	417	M	273	W	166	J	42
N	718	L	392	P	243	B	149	Q	38
I	711	H	386	Y	208	V	136	Z	6
R	677								

23. Constancy of the standard or normal unciliteral frequency distribution.--a. The relative frequencies disclosed by the statistical study of large volumes of text may be considered to be the standard or normal frequencies of the letters of written English. Counts made of smaller volumes of text will tend to approximate these normal frequencies,

[4] Op. cit., pp. 6-7.

and, within certain limits,[5] the smaller the volume, the lower will be the degree of approximation to the normal, until, in the case of a very short message, the normal proportions may not manifest themselves at all. It is advisable that the student fix this fact firmly in mind, for the sooner he realizes the true nature of any data relative to the frequency of occurrence of letters in text, the less often will his labors toward the solution of specific ciphers be thwarted and retarded by too strict an adherence to these generalized principles of frequency. He should constantly bear in mind that such data are merely statistical generalizations, that they will be found to hold strictly true only in large volumes of text, and that they may not even be approximated in short messages.

b. Nevertheless the normal frequency distribution or the "normal expectation" for any alphabetic language is, in the last analysis, the best guide to, and the usual basis for, the solution of cryptograms of a certain type. It is useful, therefore, to reduce the normal, uniliteral frequency distribution to a basis that more or less closely approximates the volume of text which the cryptanalyst most often encounters in individual cryptograms. As regards length of messages, counting only the letters in the body, and excluding address and signature, a study of the 260 telegrams referred to in par. 21 shows that the arithmetical average is 217 letters; the statistical mean, or weighted average[6], however, is 191 letters. These two results are, however, close enough together to warrant the statement that the average length of telegrams is approximately 200 letters. The frequencies given in par. 21 have therefore been reduced to a basis of 200 letters, and the following uniliteral frequency distribution may be taken as showing the most typical distribution to be expected in 200 letters of English telegraphic text:

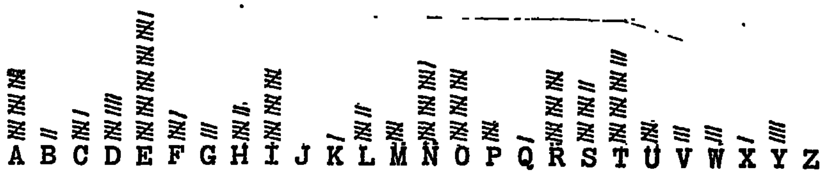


Figure 4.

---

[5] It is useless to go beyond a certain limit in establishing the normal-frequency distribution for a given language. As a striking instance of this fact, witness the frequency study made by an indefatigable German, Kaeding, who in 1898 made a count of the letters in about 11,000,000 words, totaling about 62,000,000 letters in German text. When reduced to a percentage basis, and when the relative order of frequency was determined, the results he obtained differed very little from the results obtained by Kasiski, a German cryptographer, from a count of only 1,000 letters. See Kaeding, *Haeufigkeitswoerterbuch*, Steglitz, 1898; Kasiski, *Die Geheimschriften und die Dechiffrir-Kunst*, Berlin, 1863.

[6] The arithmetical average is obtained by adding each different length and dividing by the number of different-length messages; the mean is obtained by multiplying each different length by the number of messages of that length, adding all products, and dividing by the total number of messages.

c. The student should take careful note of the appearance of the distribution[7] shown in Fig. 4, for it will be of much assistance to him in the early stages of his study. The manner of setting down the tallies should be followed by him in making his own distributions, indicating every fifth occurrence of a letter by an oblique tally. This procedure almost automatically shows the total number of occurrences for each letter, and yet does not destroy the graphical appearance of the distribution, especially if care is taken to use approximately the same amount of space for each set of five tallies. Cross-section paper is very useful for this purpose.

d. The word "uniliteral" in the designation "uniliteral frequency distribution" means "single letter", and it is to be inferred that other types of frequency distributions may be encountered. For example, a distribution of pairs of letters, constituting a biliteral frequency distribution, is very often used in the study of certain cryptograms in which it is desired that pairs made by combining successive letters be listed. A biliteral distribution of A B C D E F would take these pairs: AB, BC, CD, DE, EF. The distribution could be made in the form of a large square divided up into 676 cells. When distributions beyond biliteral are required (triliteral, quadriliteral, etc.) they can only be made by listing them in some order, for example, alphabetically based on the 1st, 2d, 3d, . . . letter.

---

[7] The use of the terms "distribution" and "frequency distribution", instead of "table" and "frequency table", respectively, is considered advisable from the point of view of consistency with the usual statistical nomenclature. When data are given in tabular form, with frequencies indicated by numbers, then they may properly be said to be set out in the form of a table. When, however, the same data are distributed in a chart which partakes of the nature of a graph, with the data indicated by horizontal or vertical linear extensions, or by a curve connecting points corresponding to quantities, then it is more proper to call such a graphic representation of the data a distribution.

24. The three facts which can be determined from a study of the unilitoral frequency distribution for a cryptogram.--a. The following three facts (to be explained subsequently) can usually be determined from an inspection of the uniliteral frequency distribution for a given cipher message of average length, composed of letters:

(1) Whether the cipher belongs to the substitution or the transposition class;

(2) If to the former, whether it is monoalphabetic[8] or non-monoalphabetic[9] in character;

(3) If monoalphabetic, whether the cipher alphabet is standard (direct or reversed) or mixed.

b. For immediate purposes the first two of the foregoing determinations are quite important and will be discussed in detail in the next two paragraphs; the other determination will be touched upon very briefly, leaving its detailed discussion for subsequent sections of the text.

25. Determining the class to which a cipher belongs.--a. The determination of the class to which a cipher belongs is usually a relatively easy matter because of the fundamental difference between transposition and substitution as cryptographic processes. In a transposition cipher the original letters of the plain text have merely been rearranged, without any change whatsoever in their identities, that is, in the conventional values they have in the normal alphabet. Hence, the numbers of vowels (A, E, I, O, U, Y), high-frequency consonants (D, N, R, S, T), medium-frequency consonants (B, C, F, G, H, L, M, P, V, W), and low-frequency consonants (J, K, Q, X, Z) are exactly the same in the cryptogram as they are in the plaintext message. Therefore, the percentages of vowels, high-, medium-, and low-frequency consonants are the same in the transposed text as in the equivalent plain text. In a

[8] In connection with uniliteral frequency distributions, the term monoalphabetic is considered to embrace the concept of monoalphabetic-monographic-uniliteral systems only, thus excluding polygraphic and multiliteral systems, both of which, however, usually fall into the monoalphabetic category.

[9] The term non-monoalphabetic as applied in this instance is considered to embrace all deviations from the characteristic appearance of monoalphabetic distributions. These deviations include the phenomena inherent in polyalphabetic, polygraphic, and multiliteral cryptograms, as well as in random text, i.e., text which appears to have been produced by chance or accident, having no discernible patterns or limitations.

substitution cipher, on the other hand, the identities of the original letters of the plain text have been changed, that is, the conventional values they have in the normal alphabet have been altered. Consequently, if a count is made of the various letters present in such a cryptogram, it will be found that the number of vowels, high-, medium-, and low-frequency consonants will usually be quite different in the cryptogram from what they are in the original plaintext message. Therefore, the percentages of vowels, high-, medium-, and low-frequency consonants are usually quite different in the substitution text from what they are in the equivalent plain text. From these considerations it follows that if in a specific cryptogram the percentages of vowels, high-, medium-, and low-frequency consonants are approximately the same as would be expected in normal plain text, the cryptogram probably belongs to the transposition class; if these percentages are quite different from those to be expected in normal plain text the cryptogram probably belongs to the substitution class.

b. In the preceding subparagraph the word "probably" was emphasized by italicizing it, for there can be no certainty in every case of this determination. Usually these percentages in a transposition cipher are close to the normal percentages for plain text; usually, in a substitution cipher, they are far different from the normal percentages for plain text. But occasionally a cipher message is encountered which is difficult to classify with a reasonable degree of certainty because the message is too short for the general principles of frequency to manifest themselves. It is clear that if in actual messages there were no variation whatever from the normal vowel and consonant percentages given in Table 3, the determination of the class to which a specific cryptogram belongs would be an extremely simple matter. But unfortunately there is always some variation or deviation from the normal. Intuition suggests that as messages decrease in length there may be a greater and greater departure from the normal proportions of vowels, high-, medium-, and low-frequency consonants, until in very short messages the normal proportions may not hold at all. Similarly, as messages increase in length there may be a lesser and lesser departure from the normal proportions, until in messages totalling a thousand or more letters there may be no difference at all between the actual and the theoretical proportions. But intuition is not enough, for in dealing with specific messages of the length of those commonly encountered in practical work the question sometimes arises as to exactly how much deviation (from the normal proportions) may be allowed for in a cryptogram which shows a considerable amount of deviation from the normal and which might still belong to the transposition rather than to the substitution class.

c. Statistical studies have been made on this matter and some graphs have been constructed thereon. These are shown in Charts 2 - 5 in the form of simple curves, the use of which will now be explained. Each chart contains two curves marking the lower and upper limits, respectively, of the theoretical amount of deviation (from the normal percentages) of vowels or consonants which may be allowable in a cipher believed to belong to the transposition class.

d. In Chart 2, curve V_1 marks the lower limit of the theoretical amount of deviation[10] from the number of vowels theoretically expected to appear[11] in a message of given length; curve V_2 marks the upper limit of the same statistic. Thus, for example, in a message of 100 letters in plain English there should be between 33 and 47 vowels (A E I O U Y). Likewise, in Chart 3 curves H_1 and H_2 mark the lower and upper limits as regards the high-frequency consonants. In a message of 100 letters there should be between 28 and 42 high-frequency consonants (D N R S T). In Chart 4 curves M_1 and M_2 mark the lower and upper limits as regards the medium-frequency consonants. In a message of 100 letters there should be between 17 and 31 medium-frequency consonants (B C F G H L M P V W). Finally, in Chart 5, curves L_1 and L_2 mark the lower and upper limits as regards the low-frequency consonants. In a message of 100 letters there should be between 0 and 3 low-frequency consonants (J K Q X Z). In using the charts, therefore, one finds the point of intersection of the vertical coordinate corresponding to the length of the message, with the horizontal coordinate corresponding to (1) the number of vowels, (2) the number of high-frequency consonants, (3) the number of medium-frequency consonants, and (4) the number of low-frequency consonants actually counted in the message. If all four points of intersection fall within the area delimited by the respective curves, then the numbers of vowels and high-, medium-, and low-frequency consonants correspond with the numbers theoretically expected in a normal plaintext message of the same length; since the message under investigation is not plain text, it follows that the cryptogram may certainly be classified as a transposition cipher. On the other hand, if one or more of these points of intersection fall outside the area delimited by the respective curves, it follows that the cryptogram is probably a substitution cipher. The distance that the point of intersection falls outside the area delimited by these curves is a more or less rough measure of the improbability of the cryptogram's being a transposition cipher.

e. Sometimes a cryptogram is encountered which is hard to classify with certainty even with the foregoing aids, because it has been consciously prepared with a view to making the classification difficult. This can be done either by selecting peculiar words (as in "trick cryptograms") or by employing a cipher alphabet in which letters of approximately similar normal frequencies have been interchanged. For example, E may be replaced by O, T by R, and so on, thus yielding a cryptogram giving external indications of being a transposition cipher but which is really a substitution cipher. If the cryptogram is not too short, a close study will usually disclose what has been done, as well as the futility of so simple a subterfuge.

---

[10] In Charts 2 - 5, inclusive, the limits of the upper and lower curves have been calculated to include approximately 70 percent of messages of the various lengths.

[11] The expression "the number of ... theoretically expected to appear" is often condensed to "the theoretical expectation of ..." or "the normal expectation of ..."

f. In the majority of cases, in practical work, the determination of the class to which a cipher of average length belongs can be made from a mere inspection of the message, after the cryptanalyst has acquired a familiarity with the normal appearance of transposition and of substitu-

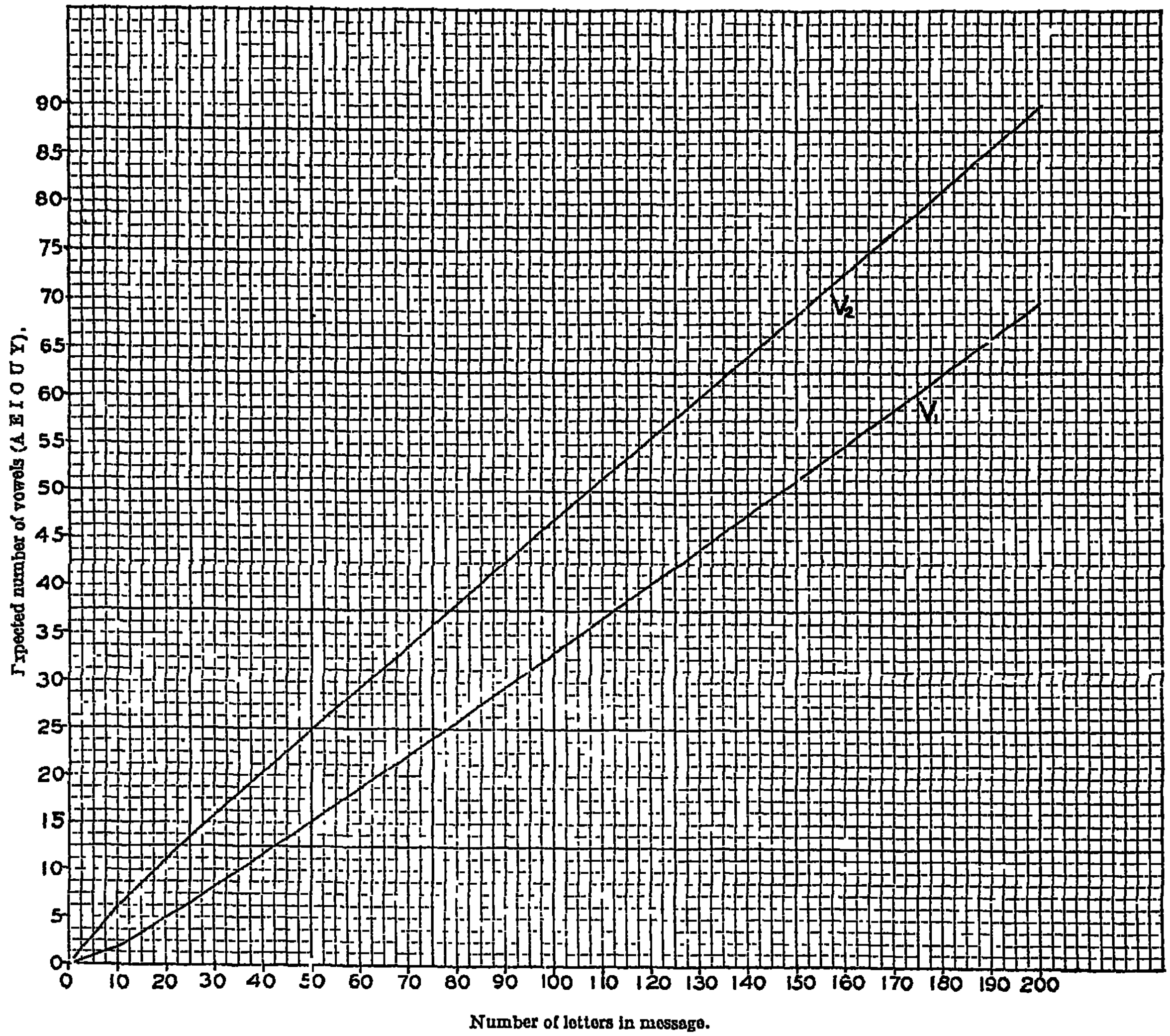


Chart 2. Curves marking the lower and upper limits of the theoretical amount of deviation from the number of vowels theoretically expected in messages of various lengths. (See subpar. 25d.)

tion ciphers. In the former case, his eyes very speedily note many high-frequency letters, such as E, T, N, R, O, and S, with the absence of low-frequency letters, such as J, K, Q, X, and Z; in the latter case, his eyes just as quickly note the presence of many low-frequency letters, and a corresponding absence of some of the high-frequency letters.

g. Another rather quickly completed test, in the case of the simpler varieties of ciphers, is to look for repetitions of groups of letters. As will become apparent very soon, recurrences of syllables, entire words and short phrases constitute a characteristic of all normal plain text. Since a transposition cipher involves a change in the sequence of the letters

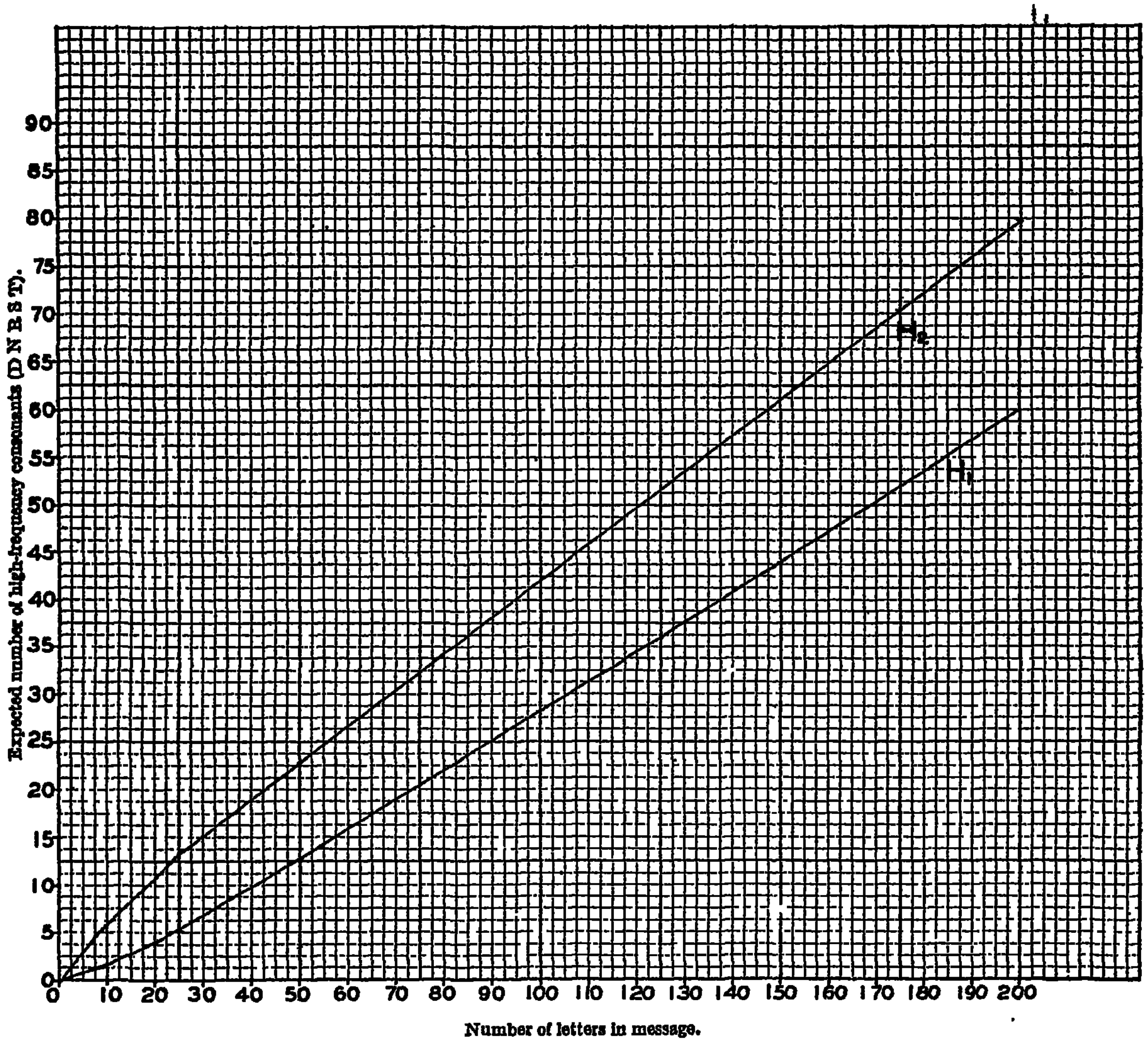


Chart 3. Curves marking the lower and upper limits of the theoretical amount of deviation from the number of high-frequency consonants theoretically expected in messages of various lengths. (See subpar. 25d.)

composing a plaintext message, such recurrences are broken up so that the cipher text no longer will show repetitions of more or less lengthy sequences of letters. But if a cipher message does show many repetitions and these are of several letters in length, say over four or five, the conclusion is at once warranted that the cryptogram is most probably a substitution and not a tranposition cipher. However, for the beginner in cryptanalysis, it will be advisable to make the unilateral frequency distribution, and note the frequencies of the vowels and of the high-,

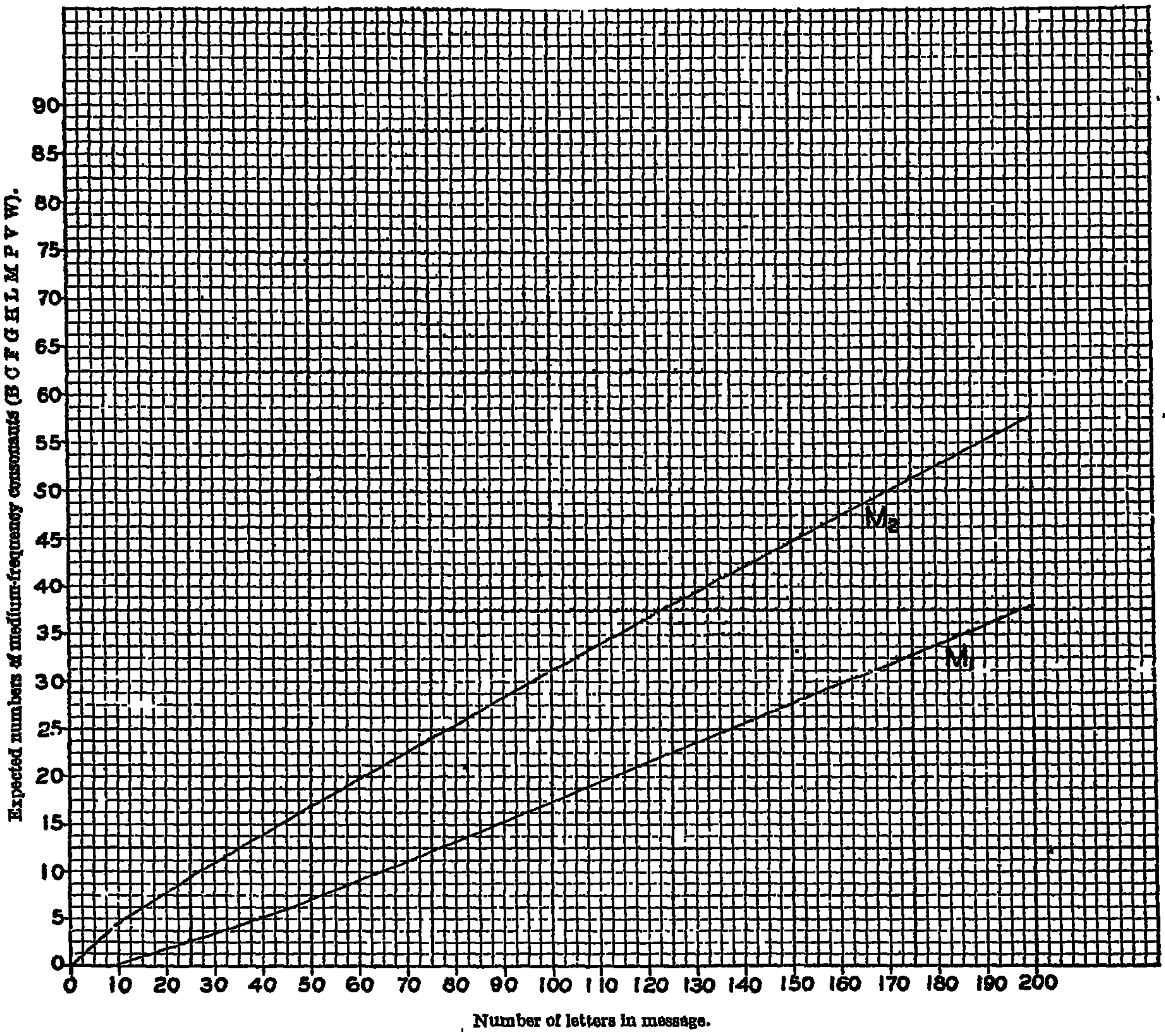


Chart 4. Curves marking the lower and upper limits of the theoretical amount of deviation from the number of medium-frequency consonants theoretically expected in messages of various lengths. (See subpar. 25d.)

medium-, and low-frequency consonants. Then, referring to Charts 2 to 5, he should carefully note whether or not the observed frequencies for these categories of letters fall within the limits of the theoretical frequencies for a normal plaintext message of the same length, and be guided accordingly.

h. It is obvious that the foregoing rule applies only to ciphers composed wholly of letters. If a message is composed entirely of figures, or of arbitrary signs and symbols, or of intermixtures of letters, figures and other symbols, it is immediately apparent that the cryptogram is a substitution cipher.

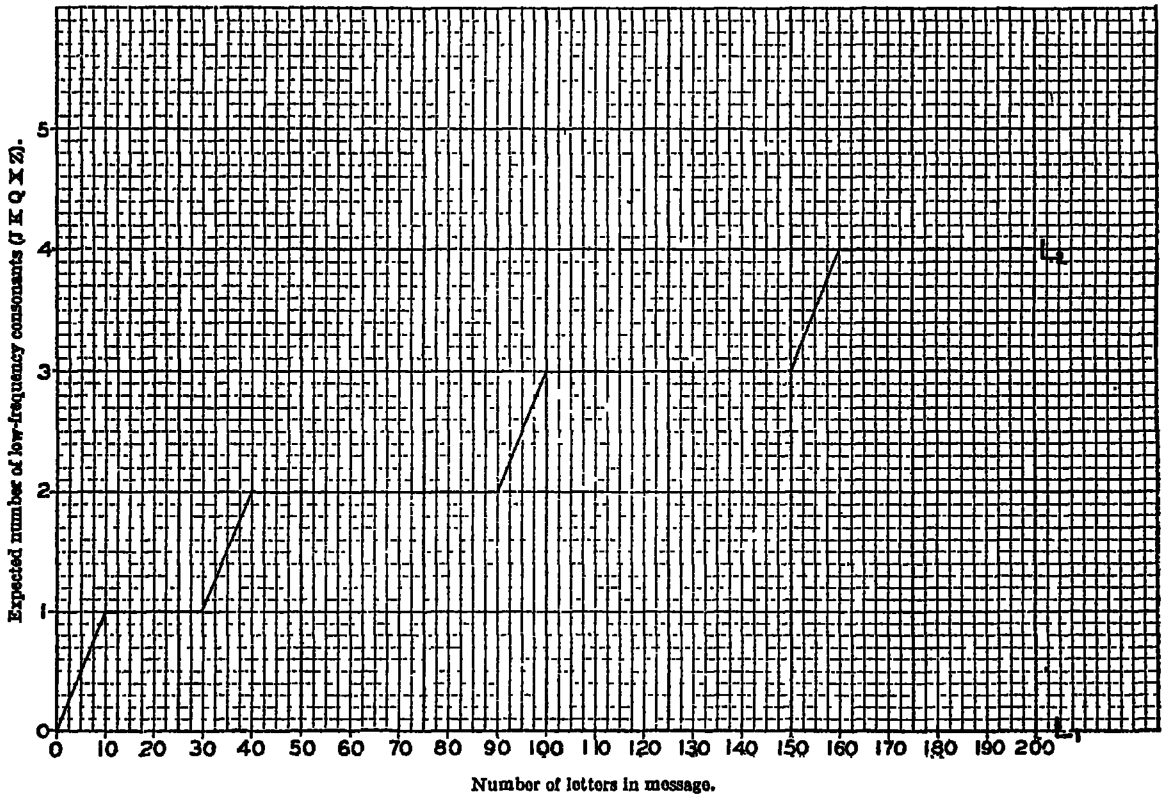


Chart 5. Curves marking the lower and upper limits of the theoretical amount of deviation from the number of low-frequency consonants theoretically expected in messages of various lengths. (See subpar. 25d.)

i. Finally, it should be mentioned that there are certain kinds of cryptograms whose class cannot be determined by the method set forth in subparagraph d above. These exceptions will be discussed in a subsequent section of this text.[12]

---

12 Section X.

26. Determining whether a substitution cipher is monoalphabetic or non-monoalphabetic.--a. It will be remembered that a monoalphabetic substitution cipher is one in which a single cipher alphabet is employed throughout the whole message; that is, a given plaintext letter is invariably represented throughout the message by one and the same letter in the cipher text. On the other hand, a polyalphabetic substitution cipher is one in which two or more cipher alphabets are employed within the same message; that is, a given plaintext letter may be represented by two or more different letters in the cipher text, according to some rule governing the selection of the equivalent to be used in each case. From this it follows that a single cipher letter may represent two or more different plaintext letters. A similar situation prevails in the case of multiliteral substitution, in which a particular cipher letter may constitute a part of the equivalents for several plaintext letters, giving rise to phenomena resembling those of polyalphabeticity.

b. It is easy to see why and how the appearance of the uniliteral frequency distribution for a substitution cipher may be used to determine whether the cryptogram is monoalphabetic or non-monoalphabetic in character. The normal distribution presents marked crests and troughs by virtue of two circumstances. First, the elementary sounds which the symbols represent are used with greatly varying frequencies, it being one of the striking characteristics of every alphabetic language that its elementary sounds are used with greatly varying frequencies.[13] In the second place, except for orthographic aberrations peculiar to certain languages (conspicuously, English and French), each such sound is represented by the same symbol. It follows, therefore, that since in a monoalphabetic substitution cipher each different plaintext letter (=elementary sound) is represented by one and only one cipher letter (=elementary symbol), the uniliteral frequency distribution for such a cipher message must also exhibit the irregular crest-and-trough appearance of the normal distribution, but with this important modification--the absolute positions of the crests and troughs will not be the same as in the normal. That is, the letters accompanying the crests and the troughs in the distribution for the cryptogram will be different from those accompanying the crests and the troughs in the normal distribution. But the marked irregularity or "roughness" of the distribution, that is, the presence of accentuated crests and troughs, is in itself an indication that each symbol or cipher letter always represents the same plaintext letter in that cryptogram. Hence the general rule: A marked crest-and-trough appearance in the uniliteral frequency distribution for a given cryptogram indicates that a single cipher alphabet is involved and constitutes one of the tests for a monoalphabetic substitution cipher.

c. On the other hand, suppose that in a cryptogram each cipher letter represents several different plaintext letters. Some of them are of high frequency, others of low frequency. The net result of such a

---

[13] The student who is interested in this phase of the subject may find the following reference of value: Zipf G.K., Selected Studies of the Principle of Relative Frequency in Language, Cambridge, Mass., 1932.

situation, so far as the uniliteral frequency distribution for the cryptogram is concerned, is to prevent the appearance of any marked crests and troughs and to tend to reduce the elements of the distribution to a more or less common level. This imparts a "flattened out" appearance to the distribution. For example, in a certain cryptogram of polyalphabetic construction, $K_c=E_p$, G_p, and J_p; $R_c=A_p$, D_p, and B_p; $X_c=O_p$, L_p, and F_p. The frequencies of K_c, R_c, and X_c will be approximately equal because the summations of the frequencies of the several plaintext letters each of these cipher letters represents at different times will be about equal. If this same phenomenon were true of all the letters of the cryptogram, it is clear that the frequencies of the 26 letters, when shown by means of the ordinary uniliteral frequency distribution, would show no striking differences and the distribution would have the flat appearance of a typical polyalphabetic substitution cipher. Hence, the general rule: The absence of marked crests and troughs in the uniliteral frequency distribution indicates that a complex form of substitution is involved. The flattened-out appearance of the distribution, then, is one of the criteria for the rejection of a hypothesis of monoalphabetic[14] substitution.

d. The foregoing test based upon the appearance of the frequency distribution is only one of several means of determing whether a substitution cipher is monoalphabetic o. non-monoalphabetic in composition. It can be employed in cases yielding frequency distributions from which definite conclusions can be drawn with more or less certainty by mere ocular examination. In those cases in which the frequency distributions contain insufficient data to permit drawing definite conclusions by such examination, certain statistical tests can be applied. One of these tests, called the ϕ (phi) test, warrants detailed treatment and is discussed in paragraph 27 below.

e. At this point, however, one additional test will be given because of its simplicity of application. This ~~test~~ test, the Λ (lambda) or blank-expectation test, may be employed in testing messages up to 200 letters in length, it being assumed that in messages of greater length ocular examination of the frequency distribution offers little or no difficulty. This test concerns the number of blanks in the frequency distribution, that is, the number of letters of the alphabet which are entirely absent from the message. It has been found from statistical studies that rather definite "laws" govern the theoretically expected number of blanks in normal plaintext messages and in frequency distributions for cryptograms of different natures and of various sizes. The results of certain of these studies have been embodied in Chart 6.

f. This chart contains two curves. The one labeled P applies to the average number of blanks theoretically expected in frequency distributions based upon normal plaintext messages of the indicated lengths. The other curve, labeled R, applies to the average number of blanks theoretically expected in frequency distributions based upon perfectly random assortments of letters; that is, assortments such as would be found by random

[14] Cf., footnote 8 on page 40.

selection of letters out of a hat containing thousands of letters, all of the 26 letters of the alphabet being present in equal proportions, each letter being replaced after a record of its selection has been made. Such random assortments correspond to polyalphabetic cipher messages in which the number of cipher alphabets is so large that if uniliteral frequency distributions are made of the letters, the distributions are practically identical with those which are obtained by random selections of letters out of a hat.

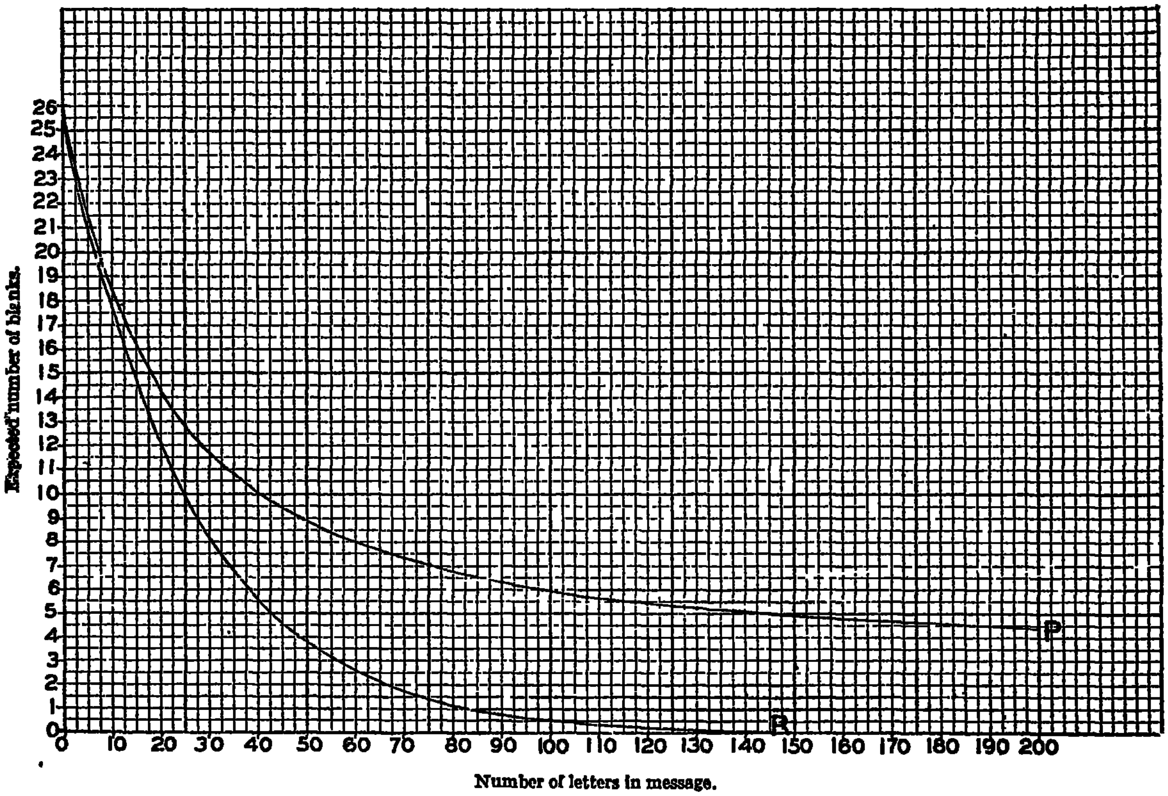


Chart 6. Curves showing the average number of blanks theoretically expected in distributions for plain text (P) and for random text (R) for messages of various lengths. (See subpar. 26f.)

g. In using this chart, one finds the point of intersection of the vertical coordinate corresponding to the length of the message, with the horizontal coordinate corresponding to the observed number of blanks in the distribution for the message. If this point of intersection falls closer to curve P than it does to curve R, the number of blanks in the message approximates or corresponds more closely to the number theoretically expected in a plaintext message than it does to a random (ciphertext) message of the same length; therefore, this is evidence that the cryptogram is monoalphabetic. Conversely, if this point of intersection falls closer to curve R than to curve P, the number of blanks in the message approximates or corresponds more closely to the number theoretically expected in a random text than it does to a plaintext message of the same length; therefore, this is evidence that the cryptogram is non-monoalphabetic.

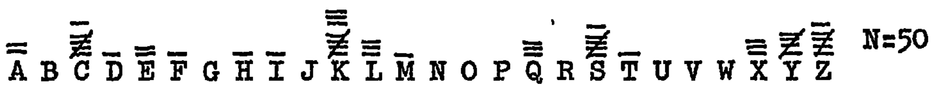
27. The ϕ (phi) test for determining monoalphabeticity.--a. The student has seen in the preceding paragraph how it is possible to determine by ocular examination whether or not a substitution cipher is monoalphabetic. This tentative determination is based on the presence of a marked crest-and-trough appearance in the uniliteral frequency distribution, and also on the number of blanks in the distribution. However, when the distribution contains a small number of elements, ocular examination and evaluation becomes increasingly difficult and uncertain. In such cases, recourse may be had to a mathematical test, known as the ϕ test, to determine the relative monoalphabeticity or non-monoalphabeticity of a distribution.

b. Without going into the theory of probability at this time, or into the derivation of the formulas involved, let it suffice for the present to state that with this test the "observed value of ϕ" (symbolized by ϕ_o) is compared with the "expected value of ϕ random" (ϕ_r) and the "expected value of ϕ plain" (ϕ_p). The formulas are $\phi_r = .0385N(N-1)$ and, for English military text, $\phi_p = .0667N(N-1)$, where N is the total number of elements in the distribution.[15] The use of these formulas is best illustrated by an example.

c. The following short cryptogram with its accompanying uniliteral frequency distribution is at hand:

QCYCH ADSKS YZZQE CYKYK QZYSK

LSZAC TKFCX LKLKC ESZMX KISZX



[15] The constant .0385 is the decimal equivalent of 1/26, i.e., the reciprocal of the number of elements in the alphabet. The constant .0667 is the sum of the squares of the probabilities of occurrence of the individual letters in English plain text. These constants are treated in detail in Military Cryptanalysis, Part II.

ϕ_o is calculated by applying the formula f(f-1) to the frequency (f) of each letter and totaling the result; or, expressed in mathematical notation,[16] $\phi_o = \Sigma f(f-1)$. Thus,

	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	
Σf =	2		6	1	2	1		1	1		8	3	1				3		6	1				3	5	6	= 50
$\Sigma f(f-1)$ =	2		30	0	2	0		0	0		56	6	0				6		30	0				6	20	30	= 188

For this distribution, $\phi_r = .0385N(N-1) = .0385 \times 50 \times 49 = 94$, and
$\phi_p = .0667N(N-1) = .0667 \times 50 \times 49 = 163$.

Now since ϕ_o, 188, is in fact greater than ϕ_p, we have a mathematical corroboration of the hypothesis that the cryptogram is a monoalphabetic substitution cipher. If ϕ_o were nearer to ϕ_r, then the assumption would be that the cryptogram is not a monoalphabetic cipher. If ϕ_o were just half way between ϕ_r and ϕ_p, then decision would have to be suspended, since no further statistical proof in the matter is possible with this particular test.[17]

d. Two further examples may be illustrated:

(1) N=25

	$\bar{A}$	B	C	$\bar{D}$	$\bar{\bar{E}}$	$\bar{\bar{\bar{F}}}$	$\bar{\bar{\bar{\bar{G}}}}$	$\bar{\bar{H}}$	I	J	$\bar{K}$	L	M	N	$\bar{\bar{\bar{\bar{O}}}}$	$\bar{\bar{P}}$	Q	$\bar{R}$	S	T	U	V	W	$\bar{X}$	Y	$\bar{\bar{\bar{Z}}}$	
	0			0	2	6	12	2			0				12	2		0						0		6	$\Sigma f(f-1) = 42$

---

[16] The more usual mathematical notation for expressing ϕ_o would be $\sum_{i=A}^{Z} f_i(f_i-1)$, which is read as "the sum of all the terms for all integral values of f from A to Z inclusive. In turn, $\sum_{i=A}^{Z} f_i(f_i-1)$ would be expanded as $f_A(f_A-1) + f_B(f_B-1) + f_C(f_C-1) + \ldots + f_Z(f_Z-1)$. However, in the interest of simplicity the notation $\Sigma f(f-1)$ is employed; likewise, the notations ϕ_r and ϕ_p are employed in lieu of the more usual $E(\phi_r)$ and $E(\phi_p)$.

[17] Another method of determining the relative monoalphabeticity of a cryptogram is based upon comparing the <u>index of coincidence</u> (abbr. <u>I.C.</u>) of the cryptogram under examination with the theoretical I.C. of plain text. The I.C. of a message is defined as the ratio of ϕ_o to ϕ_r; thus, in the example above, the I.C. is $\frac{188}{94}$, which equals 2. The theoretical I.C. of English plain text is 1.73, which is the decimal equivalent of $\frac{.0667}{.0385}$, the ratio of the "<u>plain</u> constant" to the "<u>random</u> constant". The I.C. of random text is 1, i.e., $\frac{.0385}{.0385}$.

(2) $A\ B\ \overline{C}\ D\ E\ F\ \overline{G}\ \overline{H}\ \overline{\overline{I}}\ \overline{J}\ \overline{K}\ \overline{L}\ \overline{\overline{\overline{M}}}\ \overline{N}\ \overline{O}\ P\ \overline{Q}\ \overline{\overline{R}}\ S\ \overline{T}\ \overline{U}\ V\ \overline{W}\ \overline{X}\ \overline{\overline{Y}}\ \overline{\overline{\overline{Z}}}$ N=25

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	
		0				0	0	2	0	0	0	6	0	0		0	2		0	0		0	0	2	6	$\Sigma f(f-1)=18$

Since both distributions have 25 elements, then for both

$$\phi_r = .0385 \times 25 \times 24 = 21, \text{ and}$$

$$\phi_p = .0667 \times 25 \times 24 = 40.$$

Hence distribution (1) is monoalphabetic, while (2) is not.

e. The student must not assume that statistical tests in cryptanalysis are infallible or absolute in themselves[18]; statistical approaches serve only as a means to the end, in guiding the analyst to the most probably fruitful sources of attack. Since no one test in cryptanalysis gives definite proof of a hypothesis (in fact, not even a battery of tests gives absolute proof), all applicable statistical means at the disposal of the cryptanalyst should be used; thus, in examination for monoalphabeticity, the ϕ test, Λ test, and even other tests[19] could profitably be employed. To illustrate this point, if the ϕ test is taken on the distribution of the plaintext letters of the phrase

A QUICK BROWN FOX JUMPS OVER THE LAZY DOG

$\overline{\overline{A}}\ \overline{B}\ \overline{C}\ \overline{D}\ \overline{\overline{E}}\ \overline{F}\ \overline{G}\ \overline{H}\ \overline{I}\ \overline{J}\ \overline{K}\ \overline{L}\ \overline{M}\ \overline{N}\ \overline{\overline{\overline{\overline{O}}}}\ \overline{P}\ \overline{Q}\ \overline{\overline{R}}\ \overline{S}\ \overline{T}\ \overline{\overline{U}}\ \overline{V}\ \overline{W}\ \overline{X}\ \overline{Y}\ \overline{Z}$ N=33

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
2				2										12			2			2						$\Sigma f(f-1)=20$

$$\phi_r = 41; \ \phi_p = 70$$

it will be noticed that ϕ_o is less than half of ϕ_r, thus conclusively "proving" that the letters of this phrase could not possibly constitute plain text nor a monoalphabetic encipherment of plain text in any language! The student should be able to understand the cause of this cryptologic curiosity.

---

[18] The following quotation from the Indian mathematician P. C. Mahalanobis, concerning the fallibility of statistics, is particularly appropriate in this connection: "If statistical theory is right, predictions must sometimes come out wrong; on the other hand, if predictions are always right, then the statistical theory must be wrong."--Sankhyā, Vol. 10, Part 3, p. 203. Calcutta, 1950.

[19] One of these, the chi-square test, will be treated in a subsequent text.

28. Determining whether a cipher alphabet is standard (direct or reversed) or mixed.--a. Assuming that the uniliteral frequency distribution for a given cryptogram has been made, and that it shows clearly that the cryptogram is a substitution cipher and is monoalphabetic in character, a consideration of the nature of standard cipher alphabets[20] almost makes it obvious how an inspection of the distribution will disclose whether the cipher alphabet involved is a standard cipher alphabet or a mixed cipher alphabet. If the crests and troughs of the distribution occupy positions which correspond to the relative positions they occupy in the normal frequency distribution, then the cipher alphabet is a standard cipher alphabet. If this is not the case, then it is highly probable that the cryptogram has been prepared by the use of a mixed cipher alphabet. A mechanical test may be applied in doubtful cases arising from lack of material available for study; just what this test involves, and an illustration of its application will be given in the next section, using specific examples.

b. Of course, if it has been determined that a standard cipher alphabet is involved in a particular instance, it goes without saying that at the same time it must have been found whether the alphabet is a direct standard or reversed standard cipher alphabet. The difference between the distribution of a direct standard alphabet cipher and one of a reversed standard alphabet cipher is merely a matter of the direction in which the sequence of crests and troughs progresses--to the right, as is done in normally reading or writing the alphabet (A B C → Z), or to the left, that is, in the reversed direction (Z ← C B A). With a direct standard cipher alphabet the direction in which the crests and troughs of the distribution progress is the normal direction, from left to right; with a reversed standard cipher alphabet this direction is reversed, from right to left.

[20] See par. 12.

(BLANK)

SECTION V

UNILITERAL SUBSTITUTION WITH STANDARD CIPHER ALPHABETS

29. Types of standard cipher alphabets.--a. Standard cipher alphabets are of two types:

(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:

$\longrightarrow$

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: QRSTUVWXYZABCDEFGHIJKLMNOP

$\longrightarrow$

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:

$\longrightarrow$

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: QPONMLKJIHGFEDCBAZYXWVUTSR

$\longleftarrow$

Here the cipher component can be applied to the plain component at any 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 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 made by citing a pair of equivalents in that alphabet, such as, in the example above, $A_p=Q_c$. 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.

c. The cipher alphabet in subpar. a(2), above, is also a reciprocal alphabet; that is, the cipher alphabet contains 13 distinct pairs of equivalents which are reversible. 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. The reciprocity exists throughout the alphabet and is a result of the method by which it was formed. (Reciprocal alphabets may be produced by juxtaposing any two components which are identical but progress in opposite directions.)

30. Procedure in encipherment and decipherment by means of uniliteral substitution.--a. When a message is enciphered by means of uniliteral substitution, or simple substitution (as it is often called), the individual letters of the message text are replaced by the single-letter equivalents taken from the cipher alphabet selected by prearrangement. Example:

Message: EIGHTEEN PRISONERS CAPTURED

Enciphering alphabet: Direct standard, $A_p=T_c$

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: TUVWXYZABCDEFGHIJKLMNOPQRS

Letter-for-letter encipherment:

EIGHTEEN PRISONERS CAPTURED
XBZAMXXG IKBLHGXKL VTIMNKXW

The cipher text is then regrouped, for transmission, into groups of five.

Cryptogram:

XBZAM XXGIK BLHGX KLVTI MNKXW

b. The procedure in decipherment is merely the reverse of that in encipherment. The cipher alphabet selected by prearrangement is set up with the cipher component arranged in the normal sequence and placed above the plain component for ease in deciphering. The letters of the cryptogram are then replaced by their plaintext equivalents, as shown below.

Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: HIJKLMNOPQRSTUVWXYZABCDEFG

The message deciphers thus:

Cipher: XBZAM XXGIK BLHGX KLVTI MNKXW
Plain: EIGHT EENPR ISONE RSCAP TURED

The deciphering clerk rewrites the text in word lengths:

EIGHTEEN PRISONERS CAPTURED

c. In subpar. a, above, the cryptogram was prepared in final form for transmission by dividing the cryptographic text into groups of five. This is generally the case in military communications involving cipher systems. It promotes accuracy in telegraphic transmission since an operator knows he must receive a definite number of characters in each group, no more and no less. Also, it usually makes solution of the messages by unauthorized persons more difficult because the length of the words, phrases, and sentences of the plain text is hidden. If the last group of the cipher text in subpar. 30a had not been a complete group of five letters, it might have been completed by adding a sufficient number of meaningless letters (called nulls).

31. Principles of solution by construction and analysis of the uniliteral frequency distribution.--a. The analysis of monoalphabetic cryptograms prepared by the use of standard cipher alphabets follows almost directly from a consideration of the nature of such alphabets. Since the cipher component of a standard cipher alphabet consists either of the normal sequence merely displaced 1, 2, 3, . . . intervals from the normal point of coincidence, or of the normal sequence proceeding in a reversed-normal direction, it is obvious that the uniliteral frequency distribution for a cryptogram prepared by means of such a cipher alphabet employed monoalphabetically will show crests and troughs whose relative positions and frequencies will be exactly the same as in the uniliteral frequency distribution for the plain text of that cryptogram. The only thing that has happened is that the whole set of crests and troughs of the distribution has been displaced to the right or left of the position it occupies in the distribution for the plain text; or else the successive elements of the whole set progress in the opposite direction. Hence, it follows that the correct determination of the plaintext value of the cipher letter marking any crest or trough of the uniliteral frequency distribution, coupled with the correct determination of the relative direction in which the plain component sequence progresses, will result at one stroke in the correct determination of the plaintext values of all the remaining 25 letters respectively marking the other crests and troughs in that distribution. The problem thus resolves itself into a matter of selecting that point of attack which will most quickly or most easily lead to the determination of the value of one cipher letter. The single word identification will hereafter be used for the phrase "determination of the value of a cipher letter"; to identify a cipher letter is to find its plaintext value.

b. It is obvious that the easiest point of attack is to assume that the letter marking the crest of greatest frequency in the frequency distribution for the cryptogram represents E_p. Proceeding from this initial point, the identifications of the remaining cipher letters marking the other crests and troughs are tentatively made on the basis that the letters of the cipher component proceed in accordance with the normal alphabetic sequence, either direct or reversed. If the actual frequency of each letter marking a crest or a trough approximates to a fairly close degree the normal or theoretical frequency of the assumed plaintext equivalent, then the initial identification $\theta_c = E_p$ may be assumed to be correct and therefore the derived identifications of the other cipher letters also may be assumed to be correct.[1] If the original starting point for assignment of plaintext values is not correct, or if the direction of "reading" the successive crests and troughs of the distribution is not correct, then the frequencies of the other 25 cipher letters will not correspond to or even approximate the normal or theoretical frequencies of their hypothetical plaintext equivalents on the basis of the initial identification. A new initial point, that is, a different cipher equivalent, must then be selected to represent E_p; or else the direction of "reading" the crests and troughs must be reversed. This procedure, that is, the attempt to make the actual frequency relations exhibited by the uniliteral frequency distribution for a given cryptogram conform to the theoretical frequency relations of the normal frequency distribution in an effort to solve the cryptogram, is referred to technically as "fitting the actual uniliteral frequency distribution for a cryptogram to the theoretical uniliteral frequency distribution for normal plain text", or, more briefly, as "fitting the frequency distribution for the cryptogram to the normal frequency distribution", or, still more briefly, "fitting the distribution to the normal." In statistical work the expression commonly employed in connection with this process of fitting an actual distribution to a theoretical one is "testing the goodness of fit." The goodness of fit may be stated in various ways, mathematical in character.[2]

c. In fitting the actual distribution to the normal, it is necessary to regard the cipher component (that is, the letters A . . . Z marking the successive crests and troughs of the distribution) as partaking of the nature of a circle, that is, a sequence closing in upon itself, so that no matter with what crest or trough one starts, the spatial and frequency relations of the crests and troughs are constant. This manner of regarding the cipher component as being cyclic in nature is valid because it is obvious that the relative positions and frequencies of the crests and troughs of any uniliteral frequency distribution must remain the same regardless of what letter is employed as the initial point of the distribution. Fig. 5 gives a clear picture of what is meant in this connection, as applied to the normal frequency distribution.

---

[1] The Greek letter θ (theta) is used to represent a character or letter without indicating its identity. Thus, instead of the circumlocution "any letter of the plain text", the symbol θ_p is used; and for the expression "any letter of the cipher text", the symbol θ_c is used.

[2] One of these tests for expressing the goodness of fit, the χ (chi) test, will be treated in Military Cryptanalysis, Part II.

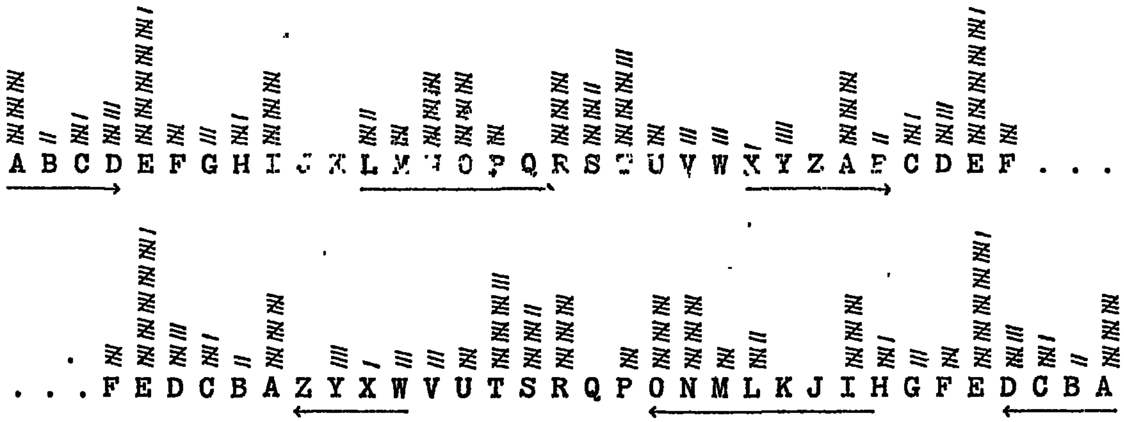


Figure 5.

d. In the third sentence of subparagraph b, the phrase "assumed to be correct" was advisedly employed in describing the results of the attempt to fit the distribution to the normal, because the final test of the goodness of fit in this connection (that is, of the correctness of the assignment of values to the crests and troughs of the distribution) is whether the consistent substitution of the plaintext values of the cipher characters in the cryptogram will yield intelligible plain text. If this is not the case, then no matter how close the approximation between actual and theoretical frequencies is, no matter how well the actual frequency distribution fits the normal, the only possible inferences are that (1) either the closeness of the fit is a pure coincidence in this case and that another equally good fit may be obtained from the same data, or else (2) the cryptogram involves something more than simple monoalphabetic substitution by means of a single standard cipher alphabet. For example, suppose a transposition has been applied in addition to the substitution. Then, although an excellent correspondence between the uniliteral frequency distribution and the normal frequency distribution has been obtained, the substitution of the cipher letters by their assumed equivalents will still not yield plain text. However, aside from such cases of double encipherment, instances in which the uniliteral frequency distribution may be easily fitted to the normal frequency distribution and in which at the same time an attempted simple substitution fails to yield intelligible text are rare. It may be said that, in practical operations whenever the uniliteral frequency distribution can be made to fit the normal frequency distribution, substitution of values will result in solution; and, as a corollary, whenever the uniliteral frequency distribution cannot be made to fit the normal frequency distribution, the cryptogram does not represent a case of simple, monoalphabetic substitution by means of a standard alphabet.

32. Theoretical example of solution.--a. The foregoing principles will become clearer by noting the encryption and solution of a theoretical example. The following message is to be encrypted.

HOSTILE FORCE ESTIMATED AT ONE REGIMENT INFANTRY AND TWO PLATOONS CAVALRY MOVING SOUTH ON QUINNIMONT PIKE STOP HEAD OF COLUMN NEARING ROAD JUNCTION SEVEN THREE SEVEN COMMA EAST OF GREENACRE SCHOOL FIRED UPON BY OUR PATROLS STOP HAVE DESTROYED BRIDGE OVER INDIAN CREEK.

b. First, solely for purposes of demonstrating certain principles, the uniliteral frequency distribution for this plaintext message is presented in Figure 6.

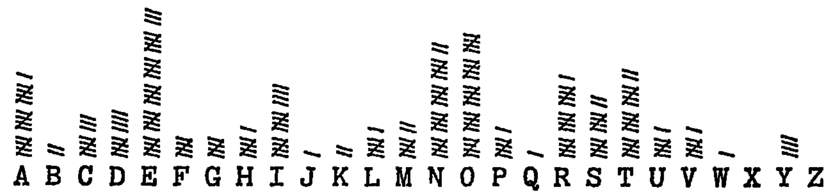


Figure 6.

c. Now let the foregoing message be encrypted monoalphabetically by the following standard cipher alphabet, yielding the cryptogram shown below and the frequency distribution shown in Figure 7.

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

Plain - - - HOSTI LEFOR CEEST IMATE DATON EREGI MENTI NFANT RYAND
Cipher - - NUYZO RKLUX IKKYZ OSGZK JGZUT KXKMO SKTZO TLGTZ XEGTJ

Plain - - - TWOPL ATOON SCAVA LRYMO VINGS OUTHO NQUIN NIMON TPIKE
Cipher - - ZCUVR GZUUT YIGBG RXESU BOTMY UAZNU TWAOT TOSUT ZVOQK

Plain - - - STOPH EADOF COLUM NNEAR INGRO ADJUN CTION SEVEN THREE
Cipher - - YZUVN KGJUL IURAS TTKGX OTMXU GJPAT IZOUT YKBKT ZNXKK

Plain - - - SEVEN COMMA EASTO FGREE NACRE SCHOO LFIRE DUPON BYOUR
Cipher - - YKBKT IUSSG KGYZU LMXKK TGIXK YINUU RLOXK JAVUT HEUAX

Plain - - - PATRO LSSTO PHAVE DESTR OYEDB RIDGE OVERI NDIAN CREEK
Cipher - - VGZXU RYYZU VNGBK JKYZX UEKJH XOJMK UBKXO TJOGT IXKKQ

Cryptogram

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

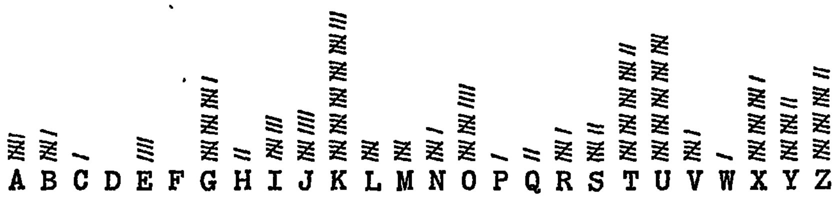


Figure 7.

d. Let the student now compare Figs. 6 and 7, which have been superimposed in Fig. 8 for convenience in examination. Crests and troughs are present in both distributions; moreover their relative positions and frequencies have not been changed in the slightest particular. Only the absolute position of the sequence as a whole has been displaced six places to the right in Fig. 7, as compared with the absolute position of the sequence in Fig. 6.

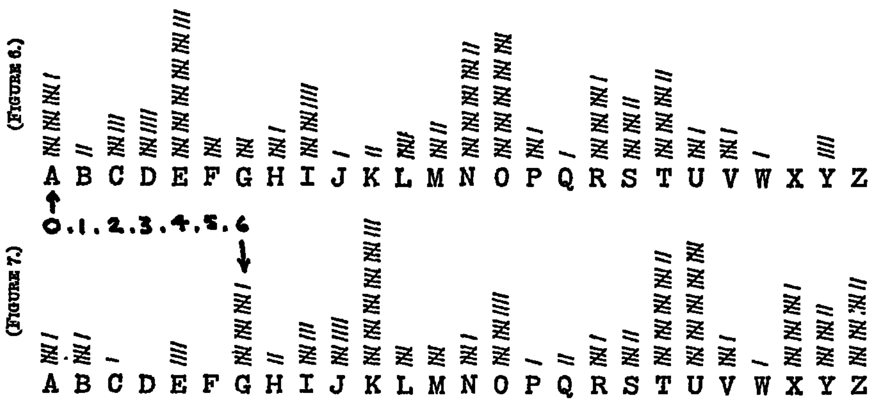


Figure 8.

e. If the two distributions are compared in detail the student will clearly understand how easy the solution of the cryptogram would be to one who knew nothing about how it was prepared. For example, the frequency of the highest crest, representing E_p in Fig. 6 is 28; at an interval of four letters before E_p there is another crest representing A_p with frequency 16. Between A and E there is a trough, representing the medium-frequency letters B, C, D. On the other side of E, at an interval of four letters, comes another crest, representing I with frequency 14. Between E and I there is another trough, representing the medium-frequency letters F, G, H. Compare these crests and troughs with their homologous crests and troughs in Fig. 7. In the latter, the letter K marks the highest crest in the distribution with a frequency of 28; four letters before K there is another crest, frequency 16, and four letters on the other side of K there is another crest, frequency 14. Troughs corresponding to B, C, D and F, G, H are seen at H, I, J and L, M, N in Fig. 7. In fact, the two distributions may be made to coincide exactly, by shifting the frequency distribution for the cryptogram six places to the left with respect to the distribution for the equivalent plaintext message, as shown herewith.

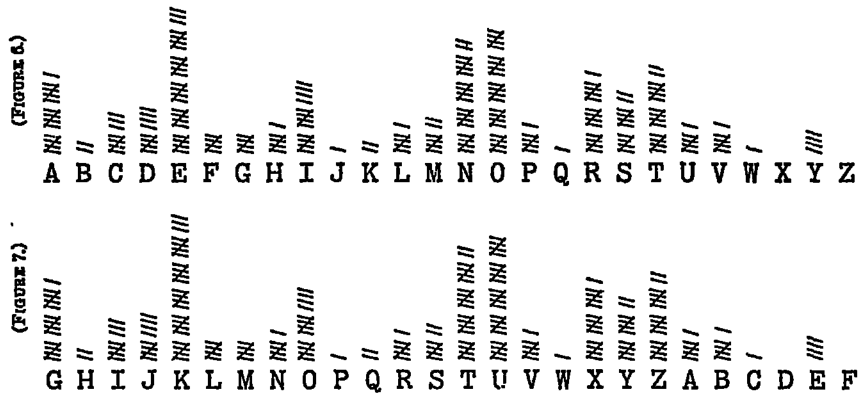


Figure 9.

f. Let us suppose now that nothing is known about the process of encryption, and that only the cryptogram and its uniliteral frequency distribution is at hand. It is clear that simply bearing in mind the spatial relations of the crests and troughs in a normal frequency distribution would enable the cryptanalyst to fit the distribution to the normal in this case. He would naturally first assume that $K_c=E_p$, from which it would follow that if a direct standard alphabet is involved, $L_c=F_p$, $M_c=G_p$, and so on, yielding the following (tentative) deciphering alphabet:

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

g. Now comes the final test! If these assumed values are substituted in the cipher text, the plain text immediately appears. Thus:

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

h. It should be clear, therefore, that the initial selection of G_c as the specific key (that is, to represent A_p) in the process of encryption has absolutely no effect upon the relative spatial and frequency relations of the crests and troughs of the frequency distribution for the cryptogram. If Q_c had been selected to represent A_p, these relations would still remain the same, the whole series of crests and troughs being merely displaced further to the right of the positions they occupy when $G_c=A_p$.

33. Practical example of solution by the frequency method.--
a. The case of direct standard alphabet ciphers.--(1) The following cryptogram is to be solved by applying the foregoing principles:

N W N V H   C A X X Y   B J C C J   L T R W P   X D A Y X   B R C R X

W B N J B   C X O W N   F C X W B   C X Y Y N   C H A B L   X U R W O

(2) From the presence of so many low-frequency letters such as B, W, and X it is at once suspected that this is a substitution cipher. But to illustrate the steps that must be taken in difficult cases in order to be certain in this respect, a uniliteral frequency distribution is constructed, and then reference is made to Charts 2 to 5 to note whether the actual numbers of vowels, high-, medium-, and low-frequency consonants fall inside or outside the areas delimited by the respective curves.

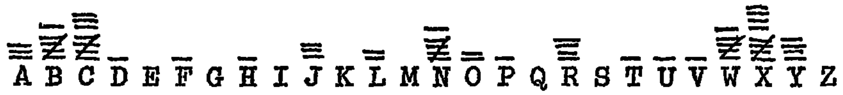


Figure 10 a.

Letters	Frequency	Position with respect to areas delimited by curves
Vowels (A E I O U Y)	10	Outside, chart 1.
High-frequency Consonants (D N R S T)	12	Outside, chart 2.
Medium-frequency Consonants (B C F G H L M P V W)	26	Outside, chart 3.
Low-frequency Consonants (J K Q X Z)	12	Outside, chart 4.
Total	60	

(3) All four points falling completely outside the areas delimited by the curves applicable to these four classes of letters, the cryptogram is clearly a substitution cipher.

(4) The appearance of the frequency distribution, with marked crests and troughs, indicates that the cryptogram is probably monoalphabetic. At this point the ϕ test is applied to the distribution. The observed value of ϕ is found to be 258, while the expected value of ϕ plain and ϕ random are calculated to be 236 and 136, respectively. The fact that the observed value is not only closer to but greater than ϕ_p is taken as statistical evidence that the cryptogram is monoalphabetic. Furthermore, reference being made to Chart 6, the point of intersection of the message length (60 letters) and the number of blanks (8) falls directly on curve P; this is additional evidence that the message is probably monoalphabetic.

(5) The next step is to determine whether a standard or a mixed cipher alphabet is involved. This is done by studying the positions and the sequence of crests and troughs in the frequency distribution, and trying to fit the distribution to the normal.

(6) The first assumption to be made is that a direct standard cipher alphabet is involved. The highest crest in the distribution occurs over X_c. Let it be assumed that $X_c=E_p$. Then Y_c, Z_c, A_c, ... $=F_p$, G_p, H_p, ..., respectively; thus:

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

Figure 10b.

It may be seen quickly that the approximation to the expected frequencies is very poor. There are too many occurrences of J_p, Q_p, U_p and F_p and too few occurrences of N_p, O_p, R_p, S_p, T_p and A_p. Moreover, if a substitution is attempted on this basis, the following is obtained for the first two cipher groups:

```
Cipher.....N W N V H   C A X X Y
"Plain text"U D U C O   J H E E F
```

This is certainly not plain text and it seems clear that X_c is not E_p, if the hypothesis of a direct standard alphabet cipher is correct. A different assumption will have to be made.

(7) Suppose $C_c=E_p$. Going through the same steps as before, again no satisfactory results are obtained. Further trials[3] are made along the same lines, until the assumption $N_c=E_p$ is tested:

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

Figure 10c.

(8) The fit in this case is quite good; possibly there are too few occurrences of A_p, D_p, and R_p. But the final test remains: trial of the substitution alphabet on the cryptogram itself. This is done and the results are as follows:

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

C: W B N J B   C X O W N   F C X W B   C X Y Y N   C N A B L   X U R W O
P: N S E A S   T O F N E   W T O N S   T O P P E   T E R S C   O L I N F
```

ENEMY TROOPS ATTACKING OUR POSITIONS EAST OF NEWTON. PETERS COL INF.

---

[3] It is unnecessary, of course, to write out all the alphabets and pseudo-decipherments, as shown above, when testing assumptions. This is usually done mentally.

(9) It is always advisable to note the specific key. In this case the correspondence between any plaintext letter and its cipher equivalent will indicate the key. Although other conventions are possible, and equally valid, it is usual, however, to indicate the key by noting the cipher equivalent of A_p. In this case $A_p=J_c$.

b. The case of reversed standard alphabet ciphers.--(1) Let the following cryptogram and its frequency distribution be studied.

```
F W F X L   Q S V V U   R J Q Q J   H Z B W D   V P S U V   R B Q B V
W R F J R   Q V E W F   N Q V W R   Q V U U F   Q F S R H   V Y B W E
```

(2) The preliminary steps illustrated above, under subpar. a (1) to (4) inclusive, in connection with the test for class and monoalphabeticity, will here be omitted, since they are exactly the same in nature. The result is that the cryptogram is obviously a substitution cipher and is monoalphabetic.

(3) Assuming that it is not known whether a direct or a reversed standard alphabet is involved, attempts are at once made to fit the frequency distribution to the normal direct sequence. If the student will try them he will soon find out that these are unsuccessful. All this takes but a few minutes.

(4) The next logical assumption is now made, viz., that the cipher alphabet is a reversed standard alphabet. When on this basis F_c is assumed to be E_p, the distribution can readily be fitted to the normal, practically every crest and trough in the actual distribution corresponding to a crest or trough in the expected distribution.

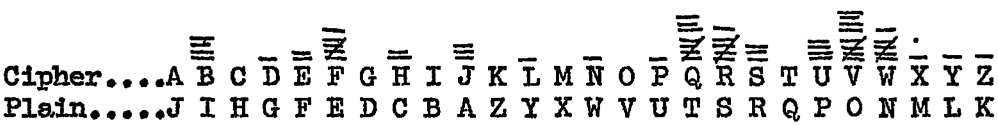


Figure 10d.

(5) When the substitution is made in the cryptogram, the following is obtained.

```
Cryptogram...F W F X L   Q S V V U   R J Q Q J
Plain text...E N E M Y   T R O O P   S A T T A
```

(6) The plaintext message is identical with that in subpar. a. The specific key in this case is also $A_p=J_c$. If the student will compare the frequency distributions in the two cases, he will note that the relative positions and extents of the crests and troughs are identical; they merely progress in opposite directions.

c. General note on solution by the frequency method.--In actual practice, the procedure of subpars. a and b are given a more rapid treatment than that just described, the practical treatment being based, not on the initial finding of some single crest or trough, but rather on locating the more readily-discernible clusters of crests which usually appear in a distribution, such as the distinctive crest-patterns representing "A...E...I" and "RST". These crest-patterns are searched for, with a quick scanning of the distribution, and then the relative placement with respect to each other is tested to see if it conforms to the expectation for a direct standard cipher alphabet, and, if not, then for a reversed standard cipher alphabet. During this latter step, which consists of little more than counting in one direction and then (when necessary) in the other, the blank (or nearly-blank) expectation of "JK" followed by the characteristic curve for "LMNOP" and the blank "Q" are considered, as a means of either substantiating or invalidating the original "identification" of the crests.

34. Solution by completing the plain-component sequence.--
a. The case of direct standard alphabet ciphers.--(1) The foregoing method of analysis, involving as it does the construction of a uniliteral frequency distribution, was termed a solution by the frequency method because it involves the construction of a frequency distribution and its study. There is, however, another method which is much more rapid, almost wholly mechanical, and which, moreover, does not necessitate the construction or study of any frequency distribution whatever. An understanding of the method follows from a consideration of the method of encipherment of a message by the use of a single, direct standard cipher alphabet.

(2) Note the following encipherment:

Message----- TWO CRUISERS SUNK

Enciphering Alphabet

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

Encipherment

```
Plain text----- T W O   C R U I S E R S   S U N K
Cryptogram----- Z C U   I X A O Y K X Y   Y A T Q
```

Cryptogram

```
Z C U I X   A O Y K X   Y Y A T Q
```

(3) The enciphering alphabet shown above represents a case wherein the sequence of letters of both components of the cipher alphabet is the normal sequence, with the sequence forming the cipher component merely shifted six places to the left (or 20 positions to the right) of the position it occupies in the normal alphabet. If, therefore, two strips

of paper bearing the letters of the normal sequence, equally spaced, are regarded as the two components of the cipher alphabet and are juxtaposed at all of the 25 possible points of coincidence, it is obvious that one of these 25 juxtapositions must correspond to the actual juxtaposition shown in the enciphering alphabet directly above.[4] It is equally obvious that if a record were kept of the results obtained by applying the values given at each juxtaposition to the letters of the cryptogram, one of these results would yield the plain text of the cryptogram.

(4) Let the work be systematized and the results set down in an orderly manner for examination. It is obviously unnecessary to juxtapose the two components so that $A_c=A_p$, for on the assumption of a direct standard alphabet, juxtaposing two direct normal components at their normal point of coincidence merely yields plain text. The next possible juxtaposition, therefore, is $A_c=B_p$. Let the juxtaposition of the two sliding strips therefore be $A_c=B_p$, as shown here:

```
Plain-------------                            ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher----------- ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
```

The values given by this juxtaposition are substituted for the letters of the cryptogram and the following results are obtained.

Cryptogram------------	Z C U I X	A O Y K X	Y Y A T Q
1st Test--"Plain text"	A D V J Y	B P Z L Y	Z Z B U R

This certainly is not intelligible text; obviously, the two components were not in the position indicated in this first test. The plain component is therefore slid one interval to the left, making $A_c=C_p$, and a second test is made. Thus

```
Plain-------------                           ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher----------- ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
```

Cryptogram------------	Z C U I X	A O Y K X	Y Y A T Q
2d Test---"Plain text"	B E W K Z	C Q A M Z	A A C V S

Neither does the second test result in disclosing any plain text. But, if the results of the two tests are studied a phenomenon that at first seems quite puzzling comes to light. Thus, suppose the results of the two tests are superimposed in this fashion.

Cryptogram------------	Z C U I X	A O Y K X	Y Y A T Q
1st Test--"Plain text"	A D V J Y	B P Z L Y	Z Z B U R
2d Test---"Plain text"	B E W K Z	C Q A M Z	A A C V S

---

[4] One of the strips should bear the sequence repeated. This permits juxtaposing the two sequences at all 26 possible points of coincidence so as to have a complete cipher alphabet showing at all times.

(5) Note what has happened. The net result of the two experiments was merely to continue the normal sequence begun by the cipher letters at the heads of the <u>columns</u> of letters. It is obvious that if the normal sequence is completed in each column <u>the results will be exactly the same as though the whole set of 25 possible tests had actually been performed</u>. Let the columns therefore be completed, as shown in Fig. 11.

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

Figure 11.

An examination of the successive horizontal lines of the diagram discloses <u>one and only one</u> line of plain text, that marked by the asterisk and reading T W O C R U I S E R S S U N K.

(6) Since each column in Fig. 11 is nothing but a normal sequence, it is obvious that instead of laboriously writing down these columns of letters every time a cryptogram is to be examined, it would be more convenient to prepare a set of strips each bearing the normal sequence doubled (to permit complete coincidence for an entire alphabet at any setting), and have them available for examining any future cryptograms. In using such a set of sliding strips in order to solve a cryptogram prepared by means of a single direct standard cipher alphabet, or to make a test to determine whether a cryptogram has been so prepared, it is only necessary to "set up" the letters of the cryptogram on the strips, that is, align them in a single row across the strips (by sliding the individual strips

up or down). The successive horizontal lines, called generatrices (singular, generatrix)[5], are then examined in a search for intelligible text. If the cryptogram really belongs to this simple type of cipher, one of the generatrices will exhibit intelligible text all the way across; this text will practically invariably be the plain text of the message. This method of analysis may be termed a solution by completing the plain-component sequence. Sometimes it is referred to as "running down" the sequence. The principle upon which the method is based constitutes one of the cryptanalyst's most valuable tools.[6]

b. The case of reversed standard alphabets.--(1) The method described under subpar. a may also be applied, in slightly modified form, in the case of a cryptogram enciphered by a single reversed standard alphabet. The basic principles are identical in the two cases, as will now be demonstrated.

(2) Let two sliding components be prepared as before, except that in this case one of the components must be a reversed normal sequence, the other, a direct normal sequence.

(3) Let the two components be juxtaposed A to A, as shown below, and then let the resultant values be substituted for the letters of the cryptogram. Thus:

CRYPTOGRAM

N K S E P   M Y O C P   O O M T W

```
Plain-----------                           ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher---------- ZYXWVUTSRQPONMLKJIHGFEDCBAZYXWVUTSRQPONMLKJIHGFEDCBA
    Cryptogram------------ N K S E P   M Y O C P   O O M T W
    1st Test--"Plain text" N Q I W L   O C M Y L   M M O H E
```

(4) This does not yield intelligible text, and therefore the reversed component is slid one space forward and a second test is made. Thus:

```
Plain-----------                          ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher---------- ZYXWVUTSRQPONMLKJIHGFEDCBAZYXWVUTSRQPONMLKJIHGFEDCBA
    Cryptogram------------ N K S E P   M Y O C P   O O M T W
    2d Test---"Plain text" O R J X M   P D N Z M   N N P I F
```

(5) Neither does the second test yield intelligible text. But let the results of the two tests be superimposed. Thus:

```
    Cryptogram------------ N K S E P   M Y O C P   O O M T W
    1st Test--"Plain text" N Q I W L   O C M Y L   M M O H E
    2d Test---"Plain text" O R J X M   P D N Z M   N N P I F
```

---

[5] Pronounced: jĕn′ĕr-ā-trī′sēz and jĕn′ĕr-ā′trĭks, respectively.

[6] A set of heavy paper strips, suitable for use in completing the plain-component sequence, has been prepared for use as a training aid in connection with the courses in Military Cryptanalysis.

(6) It is seen that the letters of the "plain text" given by the <u>second</u> trial are merely the continuants of the normal sequences initiated by the letters of the "plain text" given by the first trial. If these sequences are "run down"--that is, completed within the columns--the results must obviously be the same as though successive tests exactly similar to the first two were applied to the cryptogram, using one reversed normal and one direct normal component. If the cryptogram has really been prepared by means of a single reversed standard alphabet, one of the generatrices of the diagram that results from completing the sequences <u>must</u> yield intelligible text.

(7) Let the diagram be made, or better yet, if the student has already at hand the set of sliding strips referred to in footnote 6 to page 69, let him "set up" the letters given by the <u>first</u> trial. Fig. 12 shows the diagram and indicates the plaintext generatrix.

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

Figure 12.

(8) The only difference in procedure between this case and the preceding one (where the cipher alphabet was a direct standard alphabet) is that the letters of the cipher text are first "deciphered" by means of <u>any</u> reversed standard alphabet and then the columns are "run down", according to the normal A B C . . . Z sequence. For reasons which will

become apparent very soon, the first step in this method is technically termed converting the cipher letters into their plain-component equivalents; the second step is the same as before, viz., completing the plain-component sequence.

35. Special remarks on the method of solution by completing the plain-component sequence.--a. The terms employed to designate the steps in the solution set forth in par. 34b(8), viz., "converting the cipher letters into their plain-component equivalents" and "completing the plain-component sequence", accurately describe the process. Their meaning will become more clear as the student progresses with the work. It may be said that whenever the components of a cipher alphabet are known sequences, no matter how they are composed, the difficulty and time required to solve any cryptogram involving the use of those components is considerably reduced. In some cases this knowledge facilitates, and in other cases is the only thing that makes possible, the solution of a very short cryptogram that might otherwise defy solution. Later on an example will be given to illustrate what is meant in this regard.

b. The student should take note, however, of two qualifying expressions that were employed in a preceding paragraph to describe the results of the application of the method. It was stated that "one of the generatrices will exhibit intelligible text all the way across; this text will practically invariably be the plain text." Will there ever be a case in which more than one generatrix will yield intelligible text through its extent? That obviously depends almost entirely on the number of letters that are aligned to form a generatrix. If a generatrix contains but a very few letters, only five, for example, it may happen as a result of pure chance that there will be two or more generatrices showing what might be "intelligible text." Note in Fig. 11, for example, that there are several cases in which 3-letter and 4-letter English words (LAD, COB, MESH, MAPS, etc.) appear on generatrices that are not correct, these words being formed by pure chance. But there is not a single case, in this diagram, of a 5-letter or longer word appearing fortuitously, because obviously the longer the word the smaller the probability of its appearance purely by chance; and the probability that two generatrices of 15 letters each will both yield intelligible text along their entire length is exceedingly remote, so remote, in fact, that in practical cryptology such a case may be considered nonexistent.[7]

c. The student should observe that in reality there is no difference whatsoever in principle between the two methods presented in subpars. a and b of par. 34. In the former the preliminary step of converting the cipher letters into their plain-component equivalents is apparently not present but in reality it is there. The reason for its apparent absence is that in that case the plain component of the cipher alphabet is identical in all respects with the cipher component, so that the cipher letters

[7] A person with patience and an inclination toward the curiosities of the science might construct a text of 15 or more letters which would yield two "intelligible" texts on the plain-component completion diagram.

require no conversion, or, rather, they are identical with the equivalents that would result if they were converted on the basis $A_c=A_p$. In fact, if the solution process had been arbitrarily initiated by converting the cipher letters into their plain-component equivalents at the setting $A_c=O_p$, for example, and the cipher component slid one interval to the right thereafter, the results of the first and second tests of par. 34a would be as follows:

Cryptogram---------------	Z C U I X A O Y K X Y Y A T Q
1st Test--"Plain text"---	N Q I W L O C M Y L M M O H E
2d Test---"Plain text"---	O R J X M P D N Z M N N P I F

Thus, the foregoing diagram duplicates in every particular the diagram resulting from the first two tests under par. 34b: a first line of cipher letters, a second line of letters derived from them but showing externally no relationship with the first line, and a third line derived immediately from the second line by continuing the direct normal sequence. This point is brought to attention only for the purpose of showing that a single, broad principle is the basis of the general method of solution by completing the plain-component sequence, and once the student has this firmly in mind he will have no difficulty whatsoever in realizing when the principle is applicable, what a powerful cryptanalytic tool it can be, and what results he may expect from its application in specific instances.

d. In the two foregoing examples of the application of the principle, the components were normal sequences; but it should be clear to the student, if he has grasped what has been said in the preceding subparagraph, that these components may be mixed sequences which, if known (that is, if the sequence of letters comprising the sequences is known to the cryptanalyst), can be handled just as readily as can components that are normal sequences.

e. It is entirely immaterial at what points the plain and the cipher components are juxtaposed in the preliminary step of converting the cipher letters into their plain-component equivalents. For example, in the case of the reversed alphabet cipher solved in par. 34b, the two components were arbitrarily juxtaposed to give the value $A_p=A_c$, but they might have been juxtaposed at any of the other 25 possible points of coincidence without in any way affecting the final result, *viz.*, the production of one plaintext generatrix in the completion diagram.

<u>36. Value of mechanical solution as a short cut.</u>--a. It is evident that <u>the very first step the student should take in his attempts to solve an unknown cryptogram that is obviously a substitution cipher is to try the mechanical method of solution by completing the plain-component sequence, using the normal alphabet, first direct, then reversed. This takes only a very few minutes and is conclusive in its results.</u> It saves the labor and trouble of constructing a frequency distribution in case the cipher is of this simple type. Later on it will be seen how certain variations of this simple type may also be solved by the application of this method. Thus, a very easy short cut to solution is afforded, which even the experienced cryptanalyst never overlooks in his first attack on an unknown cipher.

b. It is important now to note that if neither of the two foregoing attempts is successful in bringing plain text to light and the cryptogram is quite obviously monoalphabetic in character, the cryptanalyst is warranted in assuming that the cryptogram involves a mixed cipher alphabet.[8]

37. Basic reason for the low degree of cryptosecurity afforded by monoalphabetic cryptograms involving standard cipher alphabets.--The student has seen that the solution of monoalphabetic cryptograms involving standard cipher alphabets is a very easy matter. Two methods of analysis were described, one involving the construction of a frequency distribution, the other not requiring this kind of tabulation, being almost mechanical in nature and correspondingly rapid. In the first of these two methods it was necessary to make a correct assumption as to the value of but one of the 26 letters of the cipher alphabet and the values of the remaining 25 letters at once became known; in the second method it was not necessary to assume a value for even a single cipher letter. The student should understand what constitutes the basis of this situation, viz., the fact that the two components of the cipher alphabet are composed of known sequences. What if one or both of these components are, for the cryptanalyst, unknown sequences? In other words, what difficulties will confront the cryptanalyst if the cipher component of the cipher alphabet is a mixed sequence? Will such an alphabet be solvable as a whole at one stroke, or will it be necessary to solve its values individually? Since the determination of the value of one cipher letter in this case gives no direct clues to the value of any other letter, it would seem that the solution of such a cipher should involve considerably more analysis and experiment than has the solution of either of the two types of ciphers so far examined. The steps to be taken in the cryptanalysis of a mixed-alphabet cipher will be discussed in the next section.

---

[8] There is but one other possibility, already referred to under subpar. 31d which involves the case where transposition and monoalphabetic substitution processes have been applied in successive steps. This is unusual, however, and will be discussed in its proper place.

(BLANK)

SECTION VI

UNILITERAL SUBSTITUTION WITH MIXED CIPHER ALPHABETS

38. Literal keys and numerical keys.--a. As has been previously mentioned, most cryptosystems involve the use of a specific key to control the steps followed in encrypting or decrypting a specific message (see subpar. 9b). Such a key may be in literal form or in numerical form.

b. It is convenient to designate a key which is composed of letters as a literal key. As already mentioned, a literal key may consist of a single letter, a single word, a phrase, a sentence, a whole paragraph, or even a book; and, of course, it may consist merely of a sequence of letters chosen at random.

c. Certain cryptosystems involve the use of a numerical key, which may consist of a relatively long sequence of numbers difficult or impossible for the average cipher clerk to memorize. Several simple methods for deriving such sequences from words, phrases, or sentences have been devised, and a numerical key produced by any of these methods is called a derived numerical key (as opposed to a key consisting of randomly-selected numbers). One of the commonly-used methods consists of assigning numerical values to the letters of a selected literal key in accordance with their relative positions in the ordinary alphabet, as exemplified in the following subparagraph.

d. Let the prearranged key word be the word LOGISTICS. Since C, the penultimate letter of the key word, appears in the normal alphabet before any other letter of the key word, it is assigned the number 1:

L	O	G	I	S	T	I	C	S
							1	

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

L	O	G	I	S	T	I	C	S
5	6	2	3	7	9	4	1	8

This method of assigning the numbers is very flexible and varies with different uses to which numerical keys are put. It may, of course, be applied to phrases or to sentences, so that a very long numerical key, ordinarily impossible to remember, may be thus derived at will from an easily-remembered key text.

e. As far as the cryptanalyst is concerned, the derivation of a numerical key from a specific literal key is of interest to him because this knowledge may assist in subsequent solutions of cryptograms prepared according to the same basic system, or in identifying the source from which the literal key was selected - perhaps an ordinary book, a magazine, etc. However, it should be pointed out that in some instances the cryptanalyst may be unaware that a literal key has in fact been used as the basis for deriving a numerical key.

39. Types of mixed cipher alphabets.--a. It will be recalled that in a mixed cipher alphabet the sequence of letters or characters in one of the components (usually the cipher component) does not correspond to the normal sequence. There are various methods of composing the sequence of letters or elements of this mixed 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. Alphabets involving a systematic method of mixing are called systematically-mixed cipher alphabets.

b. One of the simplest types of systematically-mixed cipher alphabets is the keyword-mixed alphabet. The cipher component consists of a key word or phrase (with repeated letters, if present, omitted after their first occurrence)[1], followed by the letters of the alphabet in their normal sequence (with letters already occurring in the key omitted of course). Example, with GOVERNMENT as the key word:

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: GOVERNMTABCDFHIJKLPQSUWXYZ

c. It is possible to disarrange the sequence constituting the cipher component even more thoroughly by applying a simple method of transposition to the keyword-mixed sequence. Two common methods are illustrated below, using the key word TELEPHONY.

(1) Simple columnar transposition:

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 (formed by transcribing the successive columns from left to right):

TAKXEBMZLCQPDRHFSOGUNIVYJW

(2) Numerically-keyed 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 (formed by transcribing the columns in a sequence determined by the numerical key derived from the key word itself):

EBMZHFSLCQNIVOGUPDRTAKXYJW

---

[1] Mixed alphabets formed by including all repeated letters of the key word or key phrase in the cipher component were common in Edgar Allan Poe's day but are impractical because they are ambiguous, making decipherment difficult; an example:

(a) Alphabet for enciphering.-- Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: NOWISTHETIMEFORALLGOODMENT

(b) Inverse form of (a), for deciphering.---------- 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 decrypting a cipher group such as TOOET, each letter of which has three or more equivalents, and from which the plaintext fragments (N)INTH., ..FT THI(S), IT THI..., etc. can be formed on decipherment.

d. The last two systematically-mixed sequences are examples of transposition-mixed sequences. Almost any method of transposition may be used to produce such sequences.

e. Another simple method of forming a mixed sequence is the decimation method. In this method, letters in the normal alphabet, or in a keyword-mixed sequence, are "counted off" according to any selected interval. As each letter is decimated--that is, eliminated from the basic sequence by counting off--it is entered in a separate list to form the new mixed sequence. For example, to form a mixed sequence by this method from a keyword-mixed sequence based on the key phrase SING A SONG OF SIXPENCE with 7 the interval selected, proceed as follows:

Keyword-mixed (or basic) sequence:

SINGAOFXPECBDHJKLMQRTUVWYZ

When the letters are counted off by 7's from left to right, F will be the first letter arrived at, H the second, T the third:

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

These letters are entered in a separate list (F first, H second, T third, and so on) and eliminated from the keyword-mixed sequence. When the end of the keyword-mixed sequence is reached, return to the beginning, skipping the letters already eliminated:

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

The decimation-mixed sequence:

FHTIEMZPQNDWCVBSLXAGOKYJRU

f. Practical considerations, of course, set a limit to the complexities that may be introduced in constructing systematically-mixed alphabets. 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, the probability of producing a thorough disarrangement of letters is very great. Random-mixed alphabets give more cryptographic security than do the less complicated systematically-mixed alphabets, because they afford no clues to positions of letters, given the position of a few of them. Their chief disadvantage is that they must be reduced to writing, since they cannot readily be remembered, nor can they be reproduced at will from an easily-remembered key word.

~~RESTRICTED~~

40. Additional remarks on cipher alphabets.--a. 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 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, (1), below, shows a mixed cipher alphabet arranged as an enciphering alphabet; (2) shows the corresponding deciphering alphabet. An enciphering alphabet and its corresponding deciphering alphabet present an 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

(1) Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: JKQVXZWESTRNUIOLGAPHCMYBDF

Deciphering Alphabet

(2) Cipher: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Plain: RXUYHZQTNABPVLOSCKIJMDGEWF

b. A series of related reciprocal alphabets may be produced by juxtaposing at all possible points of coincidence two components which are identical but progress in opposite directions. This holds regardless of whether the components are composed of an even or an odd number of elements. The following reciprocal alphabet is one of such a series of 26 alphabets:

Plain: HYDRAULICBEFGJKMNOPQSTVWXZ
Cipher: GFEBCILUARDYHZXWVTSQPONMKJ

A single or isolated reciprocal alphabet may be produced in one of two ways:

(1) By constructing a complete reciprocal alphabet by arbitrary or random assignments of values in pairs. That is, if A_p is made the equivalent of K_c, then K_p is made the equivalent of A_c; if B_p is made R_c, then R_p is made B_c, and so on. If the two components thus constructed are slid against each other no additional reciprocal alphabets will be produced.

(2) By juxtaposing a sequence comprising an even number of elements against the same sequence shifted exactly half way to the right (or left), as seen below:

ABCDEFGHIJKLMNOPQRSTUVWXYZ
ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ

~~RESTRICTED~~

41. Preliminary steps in the analysis of a monoalphabetic, mixed-alphabet cryptogram.--a. The student is now ready to resume his cryptanalytic studies. Note the following cryptogram:

```
SFDZF IOGHL PZFGZ DYSPF HBZDS GVHTF UPLVD FGYVJ VFVHT GADZZ ATTYD ZYFZJ
ZTGPT VTZBD VFHTZ DFXSB GIDZY VTXOI YVTEF VMGZZ THLLV XZDFM HTZAI TYDZY
BDVFH TZDFK ZDZZJ SXISG ZYGAV FSLGZ DTHHT CDZRS VTYZD OZFFH TZATT YDZYG
AVDGZ ZTKHI TYZYS DZGHU ZFZTG UPGDI XWGHX ASRUZ DFUID EGHTV EAGXX
```

b. A casual inspection of the text discloses the presence of several long repetitions as well as of many letters of normally low frequency, such as F, G, V, X, and Z; on the other hand, letters of normally high frequency, such as the vowels, and the consonants N and R, are relatively scarce. The cryptogram is obviously a substitution cipher and the usual mechanical tests for determining whether it is possibly of the monoalphabetic, standard-alphabet type are applied. The results being negative, a uniliteral frequency distribution is immediately constructed, as shown in Figure 13, and the ϕ test is applied to it.

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
8	4	1	23	3	19	19	15	10	3	2	5	2	0	3	5	0	2	10	22	5	16	1	8	14	35

$\phi_p=3668$ $\phi_r=2117$ $\phi_o=3862$

Figure 13.

c. The fact that the frequency distribution shows very marked crests and troughs indicates that the cryptogram is very probably monoalphabetic, and the results of the ϕ test further support this hypothesis. The fact that the cryptogram has already been tested by the method of completing the plain-component sequence and found not to be of the monoalphabetic, standard-alphabet type, indicates with a high degree of probability that it involves a mixed cipher alphabet. A few moments might be devoted to making a careful inspection of the distribution to insure that it cannot be made to fit the normal; the object of this would be to rule out the possibility that the text resulting from substitution by a standard cipher alphabet had not subsequently been transposed. But this inspection in this case is hardly necessary, in view of the presence of long repetitions in the message.[2] (See subpar. 25g.)

[2] This possible step is mentioned here for the purpose of making it clear that the plain-component sequence completion method cannot solve a case in which transposition has followed or preceded monoalphabetic substitution with standard alphabets. Cases of this kind will be discussed in a later text. It is sufficient to indicate at this point that the frequency distribution for such a combined substitution-transposition cipher would present the characteristics of a standard alphabet cipher and yet the method of completing the plain-component sequence would fail to bring out any plain text.

d. One might, of course, attempt to solve the cryptogram by applying the simple principles of frequency. One might, in other words, assume that Z_c (the letter of greatest frequency) represents E_p, D_c (the letter of next greatest frequency) represents T_p, and so on. If the message were long enough this simple procedure might more or less quickly give the solution. But the message is relatively short and many difficulties would be encountered. Much time and effort would be expended unnecessarily, because it is hardly to be expected that in a message of only 235 letters the relative order of frequency of the various cipher letters should exactly coincide with, or even closely approximate the relative order of frequency of letters of normal plain text found in a count of 50,000 letters. It is to be emphasized that the beginner must repress the natural tendency to place too much confidence in the generalized principles of frequency and to rely too much upon them. It is far better to bring into effective use certain other data concerning normal plain text, such as digraphic and trigraphic frequencies.

42. Preparation of the work sheet.--a. The details to be considered in this paragraph may at first appear to be superfluous, but long experience has proved that systematization of the work and preparation of the data in the most utilizable, condensed form is most advisable, even if this seems to take considerable time. In the first place, if it merely serves to avoid interruptions and irritations occasioned by failure to have the data in an instantly available form, it will pay by saving mental wear and tear. In the second place, especially in the case of complicated cryptograms, painstaking care in these details, while it may not always bring about success, is often the factor that is of greatest assistance in ultimate solution. The detailed preparation of the data may be irksome to the student, and he may be tempted to avoid as much of it as possible, but, unfortunately, in the early stages of solving a cryptogram he does not know (nor, for that matter, does the expert always know) just which data are essential and which may be neglected. Even though not all of the data may turn out to have been necessary, as a general rule, time is saved in the end if all the usual data are prepared as a regular preliminary to the solution of most cryptograms.

b. First, the cryptogram is recopied in the form of a work sheet. This sheet should be of a good quality of paper so as to withstand considerable erasure. If the cryptogram is to be copied by hand, cross-section paper of $\frac{1}{4}$-inch squares is extremely useful. The writing should be in ink, and plain, carefully-made roman capital letters should be used in all cases.[3] If the cryptogram is to be copied on a typewriter, the ribbon employed should be impregnated with an ink that will not smear or smudge under the hand.

[3] It is advisable to use, for this purpose, the system of standardized manual printing adopted by Service communications personnel. The use of this system, which is included in Appendix 7, assures that work sheets are completely legible, not only to the person preparing them, but to others as well.

c. The arrangement of the characters of the cryptogram on the work sheet is a matter of considerable importance. If the cryptogram as first obtained is in groups of regular length (usually five characters to a group) and if the uniliteral frequency distribution shows the cryptogram to be monoalphabetic, the characters should be copied without regard to this grouping. It is advisable to allow one space between letters (this is especially true for work sheets prepared on the typewriter), and to write a constant number of letters per line, approximately 25. At least two spaces, preferably three spaces, should be left between horizontal lines, to allow room for multiple assumptions. Care should be taken to avoid crowding the letters in any case, for this is not only confusing to the eye but also mentally irritating when later it is found that not enough space has been left for making various sorts of marks or indications. If the cryptogram is originally in what appears to be word lengths (and this is the case, as a rule, only with the cryptograms of amateurs), naturally it should be copied on the work sheet in the original groupings.[4] If further study of a cryptogram shows that some special grouping is required, it is often best to recopy it on a fresh work sheet rather than to attempt to indicate the new grouping on the old work sheet.

d. In order to be able to locate or refer to specific letters or groups of letters with speed, certainty, and without possibility of confusion, it is advisable to use coordinates applied to the lines and columns of the text as it appears on the work sheet. To minimize possibility of confusion, it is best to apply letters to the horizontal lines of the text, numbers to the vertical columns. In referring to a letter, the horizontal line in which the letter is located is usually given first. Thus, referring to the work sheet shown below, coordinates A17 designate the letter Y, the 17th letter in the first line. The letter I is usually omitted from the series of line indicators so as to avoid confusion with the figure 1. If lines are limited to 25 letters each, then each set of 100 letters of the text is automatically blocked off by remembering that 4 lines constitute 100 letters.

e. Above each character of the cipher text may be some indication of the frequency of that character in the whole cryptogram. This indication may be the actual number of times the character occurs, or, if colored pencils are used, the cipher letters may be divided up into three categories or groups--high-frequency, medium-frequency, and low-frequency. It is perhaps simpler, if clerical help is available, to indicate the actual frequencies. This saves constant reference to the frequency tables, which interrupts the train of thought, and saves considerable time in the end, since it enables the student better to visualize frequency-patterns of words. In any case, it is recommended that the frequencies of the letters comprising the repetitions be inscribed over their

---

[4] In some cryptosystems, certain low-frequency letters are employed as word separators to indicate the end of a word; if the meaning of these letters is discovered, it is tantamount to having the cryptogram in word lengths and thus the work sheet is made accordingly. See also in this connection the treatment on word separators in Section VII.

respective letters; likewise, the frequencies of the first 10 and last 10 letters should also be inscribed, as these positions often lend themselves readily to attack.[5]

f. After the special frequency distribution, explained in Par. 43 below, has been constructed, repetitions of digraphs and trigraphs should be underscored. In so doing, the student should be particularly watchful for trigraphic repetitions which can be further extended into tetragraphs and polygraphs of greater length. Repetitions of more than ten characters should be set off by heavy vertical lines, as they indicate repeated phrases and are of considerable assistance in solution. If a repetition continues from one line to the next, put an arrow at the end of the underscore to signal this fact. Reversible digraphs and trigraphs should also be indicated by an underscore with an arrow pointing in both directions. Anything which strikes the eye as being peculiar, unusual, or significant as regards the distribution or recurrence of the characters should be noted. All these marks should, if convenient, be made with ink so as not to cause smudging. The work sheet will now appear as shown below (not all the repetitions are underscored):

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
	10	19	23	35	19	10	3	19	15	5	5	35	19	19	35	23	14	10	5	19	15	4	35	23	10
A	S	F	D	Z	F	I	O	G	H	L	P	Z	F	G	Z	D	Y	S	P	F	H	B	Z	D	S
	19	16	15	22	19	5	5	5	16	23	19	19	14	16	3	16	19	16	15	22	19	8	23	35	35
B	G	V	H	T	F	U	P	L	V	D	F	G	Y	V	J	V	F	V	H	T	G	A	D	Z	Z
	8	10	22	14	23	35	14	19	35	3	35	22	19	5	22	16	22	35	4	23	16	19	15	22	35
C	A	I	T	Y	D	Z	Y	F	Z	J	Z	T	G	P	T	V	T	Z	B	D	V	F	H	T	Z
	23	19	8	10	4	19	10	23	35	14	16	22	8	3	10	14	16	22	3	19	16	2	19	35	35
D	D	F	X	S	B	G	I	D	Z	Y	V	T	X	O	I	Y	V	T	E	F	V	M	G	Z	Z
	22	15	5	5	16	8	35	23	19	2	15	22	35	8	10	22	14	23	35	14	4	23	16	19	15
E	T	H	L	L	V	X	Z	D	F	M	H	T	Z	A	I	T	Y	D	Z	Y	B	D	V	F	H
	22	35	23	19	2	35	23	35	35	3	10	8	10	10	19	35	14	19	8	16	19	10	5	19	35
F	T	Z	D	F	K	Z	D	Z	Z	J	S	X	I	S	G	Z	Y	G	A	V	F	S	L	G	Z
	23	22	15	15	22	1	23	35	2	10	16	22	14	35	23	3	35	19	19	15	22	35	8	10	22
G	D	T	H	H	T	C	D	Z	R	S	V	T	Y	Z	D	O	Z	F	F	H	T	Z	A	I	T
	14	23	35	14	19	8	16	23	19	35	35	22	2	15	10	22	14	35	14	10	23	35	19	15	5
H	Y	D	Z	Y	G	A	V	D	G	Z	Z	T	K	H	I	T	Y	Z	Y	S	D	Z	G	H	U
	35	19	35	22	19	5	5	19	23	10	8	1	10	15	8	8	10	2	5	35	23	19	5	10	23
J	Z	F	Z	T	G	U	P	G	D	I	X	W	G	H	X	A	S	R	U	Z	D	F	U	I	D
	3	19	15	22	16	3	8	19	8	8															
K	E	G	H	T	V	E	A	G	X	X															

[5] See Appendix 4 in this connection.

43. Triliteral frequency distributions.--a. In what has gone before, a type of frequency distribution known as a uniliteral frequency distribution was used. This, of course, shows only the number of times each individual letter occurs. In order to apply the normal digraphic and trigraphic frequency data (given in Appendix 2) to the solution of a cryptogram of the type now being studied, it is obvious that the data with respect to digraphs and trigraphs occurring in the cryptogram should be compiled and should be compared with the data for normal plain text. In order to accomplish this in suitable manner, it is advisable to construct a more comprehensive form of distribution termed a triliteral frequency distribution.[6]

b. Given a cryptogram of 50 or more letters and the task of determining what trigraphs are present in the cryptogram, there are three ways in which the data may be arranged or assembled. One may require that the data show (1) each letter with its two succeeding letters; (2) each letter with its two preceding letters; (3) each letter with one preceding letter and one succeeding letter.

c. A distribution of the first of the three foregoing types may be designated as a "triliteral frequency distribution showing two suffixes"; the second type may be designated as a "triliteral frequency distribution showing two prefixes"; the third type may be designated as a "triliteral frequency distribution showing one prefix and one suffix." Quadriliteral and pentaliteral frequency distributions may occasionally be found useful.

d. Which of these three arrangements is to be employed at a specific time depends largely upon what the data are intended to show. For present purposes, in connection with the solution of a monoalphabetic substitution cipher employing a mixed alphabet, possibly the third arrangement, that showing one prefix and one suffix, is most satisfactory.

e. It is convenient to use $\frac{1}{4}$-inch cross-section paper for the construction of a triliteral frequency distribution in the form of a distribution showing crests and troughs, such as that in Figure 14. In that figure the prefix to each letter to be recorded is inserted in the left half of the cell directly above the cipher letter being recorded; the suffix to each letter is inserted in the right half of the cell directly above the letter being recorded; and in each case the prefix and the suffix to the letter being recorded occupy the same cell, the prefix being directly to the left of the suffix. The number in parentheses gives the total frequency for each letter.

---

[6] It is felt advisable here to distinguish between two closely related terms. A triliteral distribution of A B C D E F would consider the groups A B C, B C D, C D E, D E F; a trigraphic distribution would consider only the trigraphs A B C and D E F. (See also subpar. 23d.)

CONDENSED TABLE OF REPETITIONS

Digraphs			*Trigraphs*		*Longer Polygraphs*
DZ–9	TZ–5	VF–4	DZY–4	FHT–3	HTZAITYDZY–2
ZD–9	TY–5	VT–4	HTZ–4	TYD–3	BDVFHTZDF–2
HT–8	FH–4	ZF–4	ITY–4	YDZ–3	ZAITYDZY–3
ZY–6	GH–4	ZT–4	ZDF–4	ZAI–3	FHTZ–3
DF–5	IT–4	ZZ–4	AIT–3		
GZ–5					

Letter	Count	Digraphs (top to bottom)
A	(8)	EG XS GV ZI GV ZI ZI GD
B	(4)	YD SG ZD HZ
C	(1)	TD
D	(23)	IE ZF GI SZ VG YZ ZO CZ ZT ZZ ZF BV YZ ZF IZ ZF BV YZ AZ VF ZS ZY FZ
E	(3)	VA DG TF
F	(19)	DU ZZ FH ZF VS DK VH DM EV DX VH YZ VV DG TU PH ZG ZI SD
G	(19)	AX EH WH PD TU ZH DZ YA LZ YA SZ MZ BI TP TA FY SV FZ OH
H	(15)	GT GX GU KI FT HT TH FT MT TL FT VT VT FB GL
I	(10)	UD DX HT AT XS AT OY GD AT FO
J	(3)	ZS ZZ VV
K	(2)	TH FZ
L	(5)	SG LV HL PV HP
M	(2)	FH VG
N	(0)	
O	(3)	DZ XI IG
P	(5)	UG GT UL SF LZ
Q	(0)	
R	(2)	SU ZS
S	(10)	AR YD RV FL IG JX XB DG YP –F
T	(22)	HV ZG IY ZK IY HZ VY HC DH HZ IY HZ ZH VE VX HZ VZ PV ZG IY HG HF
U	(5)	FI RZ GP HZ FP
V	(16)	TE AD ST AF DF LX FM YT YT DF TT FH JF YJ LD GH
W	(1)	XG
X	(8)	X– GX HA IW SI VZ TO FS
Y	(14)	ZS TZ ZG TD TZ ZG ZB TD IV ZV ZF TD GV DS
Z	(35)	UD FT UF DG YY ZT GZ DY TA OF YD DR GD GY ZJ DZ KD TD DY TA XD ZT GZ DY TD TB JT FJ DY ZA DZ BD GD PF DF

FIGURE 14.

RESTRICTED

f. The triliteral frequency distribution is now to be examined with a view to ascertaining what digraphs and trigraphs occur two or more times in the cryptogram. Consider the pair of columns containing the prefixes and suffixes to D_c in the distribution, as shown in Fig. 14. This pair of columns shows that the following digraphs appear in the cryptogram:

Digraphs based on prefixes (arranged as one reads up the column)

FD, ZD, ZD, VD, AD, YD, BD,
ZD, ID, ZD, YD, BD, ZD, ZD,
ZD, CD, ZD, YD, VD, SD, GD,
ZD, ID

Digraphs based on suffixes (arranged as one reads up the column)

DZ, DY, DS, DF, DZ, DZ, DV,
DF, DZ, DF, DZ, DV, DF, DZ,
DT, DZ, DO, DZ, DG, DZ, DI,
DF, DE

The nature of the triliteral frequency distribution is such that in finding what digraphs are present in the cryptogram it is immaterial whether the prefixes or the suffixes to the cipher letters are studied, <u>so long as one is consistent in the study</u>. For example, in the foregoing list of digraphs based on the prefixes to D_c, the digraphs FD, ZD, ZD, VD, etc., are found; if now, the student will refer to the suffixes of F_c, Z_c, V_c, etc., he will find the very same digraphs indicated. This being the case, the question may be raised as to what value there is in listing both the prefixes and the suffixes to the cipher letters. The answer is that by so doing the trigraphs are indicated at the same time. For example, in the case of D_c, the following trigraphs are indicated:

FDZ, ZDY, ZDS, VDF, ADZ, YDZ, BDV, ZDF, IDZ, ZDF, YDZ, BDV, ZDF,
ZDZ, ZDT, CDZ, ZDO, YDZ, VDG, SDZ, GDI, ZDF, IDE.

g. The <u>repeated</u> digraphs and trigraphs can now be found quite readily. Thus, in the case of D_c, examining the list of digraphs based on suffixes, the following repetitions are noted:

DZ appears 9 times; DF appears 5 times; DV appears 2 times

Examining the trigraphs with D_c as central letter, the following repetitions are noted:

ZDF appears 4 times; YDZ appears 3 times; BDV appears 2 times

h. It is unnecessary, of course, to go through the detailed procedure set forth in the preceding subparagraphs in order to find all the repeated digraphs and trigraphs. The repeated trigraphs with D_c as central letter can be found merely from an inspection of the prefixes and suffixes opposite D_c in the distribution. It is necessary only to find those cases in which two or more prefixes are identical at the same time that the suffixes are identical. For example, the distribution shows at once that in four cases the prefix to D_c is Z_c at the same time that the suffix to this letter is F_c. Hence, the trigraph ZDF appears four times. The repeated trigraphs may all be found in this manner.

i. The most frequently repeated digraphs and trigraphs are then assembled in what is termed a condensed table of repetitions, so as to bring this information prominently before the eye. As a rule, in messages of average length, digraphs which occur less than four or five times, and trigraphs which occur less than three or four times may be omitted from the condensed table as being relatively of no importance in the study of repetitions. In the condensed table the frequencies of the individual letters forming the most important digraphs, trigraphs, etc., should be indicated.

44. Classifying the cipher letters into vowels and consonants.--
a. Before proceeding to a detailed analysis of the repeated digraphs and trigraphs, a very important step can be taken which will be of assistance not only in the analysis of the repetitions but also in the final solution of the cryptogram. This step concerns the classification of the high-frequency cipher letters into two groups--(1) those which most probably represent vowels, and (2) those which most probably represent consonants. For if the cryptanalyst can quickly ascertain the equivalents of the four vowels, A, E, I, and O, and of only the four consonants, N, R, S, and T, he will then have the values of approximately two-thirds of all the cipher letters that occur in the cryptogram; the values of the remaining letters can almost be filled in automatically.

b. The basis for the classification will be found to rest upon a comparatively simple phenomenon: the associational or combinatory behavior of vowels is, in general, quite different from that of consonants. If an examination be made of Table 7-B in Appendix 2, showing the relative order of frequency of the 18 digraphs composing 25 percent of English telegraphic text, it will be seen that the letter E enters into the composition of 9 of the 18 digraphs; that is, in exactly half of all the cases the letter E is one of the two letters forming the digraph. The digraphs containing E are as follows:

ED	EN	ER	ES		
	NE	RE	SE	TE	VE

The remaining nine digraphs are as follows:

AN	ND	OR	ST
IN	NT		TH
ON			TO

c. None of the 18 digraphs is a combination of vowels. Note now that of the 9 combinations with E, 7 are with the consonants N, R, S, and T, one is with D, one is with V, and none is with any vowel. In other words, E_p combines most readily with consonants but not with other vowels, or even with itself. Using the terms often employed in the chemical analogy, E shows a great "affinity" for the consonants N, R, S, T, but not for the vowels. Therefore, if the letters of highest frequency occurring in a given cryptogram are listed, together with the number of times each of them combines with the assumed cipher equivalent of E_p, those which show considerable combining power or affinity for the cipher equivalent of E_p may be assumed to be the cipher equivalents of N, R, S, T_p; those which do not show any affinity for the cipher equivalent of E_p may be assumed to be the cipher equivalents of A, I, O, U_p. Applying these principles to the problem in hand, and examining the triliteral frequency distribution, it is quite certain that Z_c=E_p, not only because Z_c is the letter of highest frequency, but also because it combines with several other high-frequency letters, such as D_c, F_c, G_c, etc. The nine letters of next highest frequency are:

23	22	19	19	16	15	14	10	10
D	T	F	G	V	H	Y	S	I

Let the combinations these letters form with Z_c be indicated in the following manner:

Number of times Z_c occurs as prefix---	𝍷𝍷𝍷𝍷𝍸	𝍷𝍷𝍷𝍷	𝍷𝍷𝍷𝍷	𝍷			𝍷𝍸		
Cipher Letter------	D(23)	T(22)	F(19)	G(19)	V(16)	H(15)	Y(14)	S(10)	I(10)
Number of times Z_c occurs as suffix---	𝍸𝍷𝍷𝍷𝍷	𝍸	𝍷𝍷	𝍸			𝍷𝍷		

d. Consider D_c. It occurs 23 times in the message and 18 of those times it is combined with Z_c, 9 times in the form Z_cD_c (=$E\theta_p$), and 9 times in the form D_cZ_c (=θE_p). It is clear that D_c must be a consonant. In the same way, consider T_c, which shows 9 combinations with Z_c, 4 in the form Z_cT_c (=$E\theta_p$) and 5 in the form T_cZ_c (=θE_p). The letter T_c appears to represent a consonant, as do also the letters F_c, G_c, and Y_c. On the other hand, consider V_c, occurring in all 16 times but never in combination with Z_c; it appears to represent a vowel, as do also the letters H_c, S_c, and I_c. So far, then, the following classification would seem logical:

Vowels	Consonants
Z_c(=E_p), V_c, H_c, S_c, I_c	D_c, T_c, F_c, G_c, Y_c

45. Further analysis of the letters representing vowels and consonants.--a. O_p is usually the vowel of second highest frequency. Is it possible to determine which of the letters V, H, S, I_c is the cipher equivalent of O_p? Let reference be made again to Table 6 in Appendix 2, where it is seen that the 10 most frequently occurring diphthongs are:

Diphthong------	IO	OU	EA	EI	AI	IE	AU	EO	AY	UE
Frequency------	41	37	35	27	17	13	13	12	12	11

If V, H, S, I_c are really the cipher equivalents of A, I, O, U_p (not respectively), perhaps it is possible to determine which is which by examining the combinations they make among themselves and with Z_c (=E_p). Let the combinations of V, H, S, I, and Z that occur in the message be listed. There are only the following:

ZZ_c--4 VH_c--2 HH_c--1 HI_c--1 IS_c--1 SV_c--1

ZZ_c is of course EE_p. Note the doublet HH_c; if H_c is a vowel, then the chances are excellent that H_c=O_p because the doublets AA_p, II_p, UU_p, are practically non-existent, whereas the double vowel combination OO_p is of next highest frequency to the double vowel combination EE_p. If $H_c=O_p$, then V_c must be I_p because the digraph VH_c occurring two times in the message could hardly be AO_p, or UO_p, whereas the dipthong IO_p is the one of high frequency in English. So far then, the tentative (because so far unverified) results of the analysis are as follows:

$$Z_c=E_p \quad H_c=O_p \quad V_c=I_p$$

This leaves only two letters, I_c and S_c (already classified as vowels) to be separated into A_p and U_p. Note the digraphs:

$$HI_c=O\theta_p \quad IS_c=\theta\theta_p \quad SV_c=\theta I_p$$

Only two alternatives are open:

(1) Either $I_c=A_p$ and $S_c=U_p$,
(2) Or $I_c=U_p$ and $S_c=A_p$.

If the first alternative is selected, then

$$HI_c=OA_p \quad IS_c=AU_p \quad SV_c=UI_p$$

If the second alternative is selected, then

$$HI_c=OU_p \quad IS_c=UA_p \quad SV_c=AI_p$$

The eye finds it difficult to choose between these alternatives; but suppose the frequency values of the plaintext diphthongs as given in Table 6 of Appendix 2 are added for each of these alternatives, giving the following:

$HI_c=OA_p$, frequency value=	7	$HI_c=OU_p$, frequency value=	37
$SV_c=UI_p$, frequency value=	5	$SV_c=AI_p$, frequency value=	17
$IS_c=AU_p$, frequency value=	13	$IS_c=UA_p$, frequency value=	5
Total--------	25	Total--------	59

Mathematically, the second alternative appears to be more probable than the first.[7] Let it be assumed to be correct and the following (still tentative) values are now at hand:

$$Z_c=E_p \quad H_c=O_p \quad V_c=I_p \quad S_c=A_p \quad I_c=U_p$$

b. Attention is now directed to the letters classified as consonants: How far is it possible to ascertain their values? The letter D_c, from considerations of frequency alone, would seem to be T_p, but its frequency, 23, is not considerably greater than that for T_c. It is not

7 A more accurate guide for choosing between the alternative groups of digraphs could be obtained through a consideration of the logarithmic weights of their assigned probabilities, rather than their plaintext frequency values. These weights are given in Appendix 2, along with an explanation of the method for their derivation; a detailed treatment of their application is presented in Military Cryptanalysis, Part II.

much greater than that for F_c or G_c, with a frequency of 19 each. But perhaps it is possible to ascertain not the value of one letter alone but of two letters at one stroke. To do this one may make use of a tetragraph of considerable importance in English, viz., $TION_p$. For if the analysis pertaining to the vowels is correct, and if $VH_c=IO_p$, then an examination of the letters immediately before and after the digraph VH_c in the cipher text might disclose both T_p and N_p. Reference to the text gives the following:

$$\begin{matrix} GVHT_c & FVHT_c \\ \theta IO\theta_p & \theta IO\theta_p \end{matrix}$$

The letter T_c follows VH_c in both cases and very probably indicates that $T_c=N_p$; but as to whether G_c or F_c equals T_p cannot be decided. However, two conclusions are clear: first, the letter D_c is neither T_p nor N_p, from which it follows that it must be either R_p or S_p; second, the letters G_c and F_c must be either T_p and S_p, respectively, or S_p and T_p, respectively, because the only tetragraphs usually found (in English) containing the diphthong IO_p as central letters are $SION_p$ and $TION_p$. This in turn means that as regards D_c, the latter cannot be either R_p or S_p; it must be R_p, a conclusion which is corroborated by the fact that ZD_c ($=ER_p$) and DZ_c ($=RE_p$) occur 9 times each. Thus far, then, the identifications, when inserted in an enciphering alphabet, are as follows:

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

46. Substituting deduced values in the cryptogram.--a. Thus far the analysis has been almost purely hypothetical, for as yet not a single one of the values deduced from the foregoing analysis has been tried out in the cryptogram. It is high time that this be done, because the final test of the validity of the hypotheses, assumptions, and identifications made in any cryptographic study is, after all, only this: do these hypotheses, assumptions, and identifications ultimately yield verifiable, intelligible plain text when consistently applied to the cipher text?

b. At the present stage in the process, since there are at hand the assumed values of but 9 out of the 25 letters that appear, it is obvious that a continuous "reading" of the cryptogram can certainly not be expected from a mere insertion of the values of the 9 letters. However, the substitution of these values should do two things. First, it should immediately disclose the fragments, outlines, or "skeletons" of "good" words in the text; and second, it should disclose no places in the text where "impossible" sequences of letters are established. By the first is meant that the partially deciphered text should show the outlines or skeletons of words such as may be expected to be found in the communication; this will become quite clear in the next subparagraph. By the second is meant that sequences, such as "AOOEN" or "TNRSENO" or the like, obviously not possible or extremely unusual in normal English text, must not result from the substitution of the tentative identifications resulting from the analysis. The appearance of several such extremely unusual or impossible sequences would at once signify that one or more of the assumed values is incorrect.

~~RESTRICTED~~

c. Here are the results of substituting the nine values which have been deduced by the reasoning based on a classification of the high-frequency letters into vowels and consonants and the study of the members of the two groups:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
A	10 S A	16 F T S	23 L R	35 Z E	19 F T S	10 I	3 O	19 G S T	16 H O	5 L	5 P	35 Z E	19 F T S	19 G S T	35 Z E	23 D R	14 Y	10 S A	5 P	19 F T S	15 H O	4 B	35 Z E	23 D R	10 S A
B	19 G S T	16 V I	15 H O	22 T N	19 F T S	5 U	5 P	5 L	16 V I	23 D R	19 F T S	19 G S T	14 Y	16 V I	3 J	16 V I	19 F T S	16 V I	15 H O	22 T N	19 G S T	8 A	23 D R	35 Z E	35 Z E
C	8 A	10 I	22 T N	14 Y	23 D R	35 Z E	14 Y	19 F T S	35 Z E	3 J	35 Z E	22 T N	19 G S T	5 P	22 T N	16 V I	22 T N	35 Z E	4 B	23 D R	16 V I	19 F T S	15 H O	22 T N	35 Z E
D	23 D R	19 F T S	8 X	10 S A	4 B	19 G S T	10 I	23 D R	35 Z E	14 Y	16 V I	22 T N	8 X	3 O	10 I	14 Y	16 V I	22 T N	3 E	19 F T S	16 V I	2 M	19 G S T	35 Z E	35 Z E
E	22 T N	15 H O	5 L	5 L	16 V I	8 X	35 Z E	23 D R	10 F T S	2 M	15 H O	22 T N	35 Z E	8 A	10 I	22 T N	14 Y	23 D R	35 Z E	14 Y	4 B	23 D R	16 V I	19 F T S	15 H O
F	22 T N	35 Z E	23 D R	19 F T S	2 K	35 Z E	23 D R	35 Z E	35 Z E	3 J	10 S A	8 X	10 I	10 S A	19 G S T	35 Z E	14 Y	19 G S T	8 A	16 V I	19 F T S	10 S A	5 L	19 G S T	35 Z E
G	23 D R	22 T N	15 H O	15 H O	22 T N	1 C	23 D R	35 Z E	2 R	10 S A	16 V I	22 T N	14 Y	35 Z E	23 D R	3 O	35 Z E	19 F T S	19 F T S	15 H O	22 T N	35 Z E	8 A	10 I	22 T N
H	14 Y	23 D R	35 Z E	14 Y	19 G S T	8 A	16 V I	23 D R	19 G S T	35 Z E	35 Z E	22 T N	2 K	15 H O	10 I	22 T N	14 Y	35 Z E	14 Y	10 S A	23 D R	35 Z E	19 G S T	15 H O	5 U
J	35 Z E	19 F T S	35 Z E	22 T N	19 G S T	5 U	5 P	19 G S T	23 D R	10 I	8 X	1 W	24 G S T	15 H O	8 X	8 A	10 S A	2 R	5 U	35 Z E	23 D R	19 F T S	5 U	10 I	23 D R
K	3 E	19 G S T	15 H O	22 T N	16 V I	3 E	8 A	10 G S T	8 X	8 X															

~~RESTRICTED~~

~~RESTRICTED~~

d. No impossible sequences are brought to light, and, moreover, several long words, nearly complete, stand out in the text. Note the following portions:

	A21									
	H	B	Z	D	S	G	V	H	T	F
(1)	O	?	E	R	A	S	I	O	N	T
						T				S

	C15												
	T	V	T	Z	B	D	V	F	H	T	Z	D	F
(2)	N	I	N	E	?	R	I	T	O	N	E	R	T
								S					S

	F22								
	S	L	G	Z	D	T	H	H	T
(3)	A	?	S	E	R	N	O	O	N
			T						

The words are obviously OPERATIONS, NINE PRISONERS, and AFTERNOON. The value G_c is clearly T_p; that of F_c is S_p; and the following additional values are certain:

$$B_c=P_p \qquad L_c=F_p$$

47. Completing the solution.--a. Each time an additional value is obtained, substitution is at once made throughout the cryptogram. This leads to the determination of further values, in an ever-widening circle, until all the identifications are firmly and finally established, and the message is completely solved. In this case the decipherment is as follows:

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

~~RESTRICTED~~

Message: AS RESULT OF YESTERDAYS OPERATIONS BY FIRST DIVISION THREE HUNDRED SEVENTY NINE PRISONERS CAPTURED INCLUDING SIXTEEN OFFICERS ONE HUNDRED PRISONERS WERE EVACUATED THIS AFTERNOON REMAINDER LESS ONE HUNDRED THIRTEEN WOUNDED ARE TO BE SENT BY TRUCK TO CHAMBERSBURG TONIGHT

b. The solution should, as a rule, not be considered complete until an attempt has been made to discover all the elements underlying the general system and the specific key to a message. In this case, there is no need to delve further into the general system, for it is merely one of uniliteral substitution with a mixed cipher alphabet. It is necessary or advisable, however, to reconstruct the cipher alphabet because this may give clues that later may become valuable.

c. Cipher alphabets should, as a rule, be reconstructed by the cryptanalyst in the form of <u>enciphering</u> alphabets because they will then usually be in the form in which the encipherer used them. This is important for two reasons. First, if the sequence in the cipher component gives evidence of system in its construction or if it yields clues pointing toward its derivation from a key word or a key phrase, this may often corroborate the identifications already made and may lead directly to additional identifications. A word or two of explanation is advisable here. For example, refer to the skeletonized enciphering alphabet given at the end of subpar. 45b:

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

Suppose the cryptanalyst, looking at the sequence DGFI or DFGI in the cipher component, suspects the presence of a keyword-mixed alphabet. Then DFGI is certainly a more plausible sequence than DGFI. Examining the skeleton cipher component more carefully, he notes that S . . . Z would allow for insertion of three of the missing letters UWXY, since the letters T and V occur later, probably in the keyword itself; further, he notes that the key word probably begins under F_p and ends in TH, making it probable that the TH is followed by AB or BC. This would mean that either P, Q_p=A, B_c or B, C_c. Assuming that P, Q_p=A, B_c, he refers to the frequency distribution and finds that the assumptions P_p=A_c and Q_p=B_c are not good; on the other hand, assuming that P, Q_p=B, Cc, the frequency distribution gives excellent corroboration. A trial of these values would materially hasten solution because it is often the case in cryptanalysis that if the value of a very low-frequency letter can be surely established it will yield clues to other values very quickly. Thus, if Q_p is definitely identified it almost invariably will identify U_p, and will give clues to the letter following the U_p, since it must be a vowel. In the case under discussion the identification P, Q_p=B, C_c would have turned out to be correct. For the foregoing reason an attempt should always be made in the early stages of the analysis to determine, if possible, the basis of construction or derivation of the cipher alphabet; as a rule this can be done only by means of the enciphering alphabet, and not the deciphering alphabet. For example, the skeletonized deciphering alphabet corresponding to the enciphering alphabet directly above is as follows:

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

Here no evidences of a keyword-mixed alphabet are seen at all. However, if the enciphering alphabet has been examined and shows no evidences of systematic construction, the deciphering alphabet should then be examined with this in view, because occasionally it is the deciphering alphabet which shows the presence of a key or keying element, or which has been systematically derived from a word or phrase. The second reason why it is important to try to discover the basis of construction or derivation of the cipher alphabet is that it affords clues to the general type of key words or keying elements employed by the enemy. This is a psychological factor, of course, and may be of assistance in subsequent studies of his traffic. It merely gives a clue to the general type of thinking indulged in by certain of his cryptographers.

d. In the case of the foregoing solution, the complete enciphering alphabet is found to be as follows:

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

Obviously, the letter Q, which is the only letter not appearing in the cryptogram, should follow P in the cipher component. Note now that the latter is based upon the keyword LEAVENWORTH, and that this particular cipher alphabet has been composed by shifting the mixed sequence based upon this keyword five intervals to the right so that the key for the message is $A_p=S_c$.[8] Note also that the deciphering alphabet fails to give any evidence of keyword construction based upon the word LEAVENWORTH.

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

e. If neither the enciphering nor the deciphering alphabet exhibits characteristics which give indication of derivation from a key word by some form of mixing or disarrangement, the use of such a key word for this purpose is nevertheless not finally excluded as a possibility. For the reconstruction of such mixed alphabets the cryptanalyst must use ingenuity and a knowledge of the more common methods of suppressing the appearance of key words in the mixed alphabets. Several of these methods are given detailed treatment in par. 51 below.

f. It is very important in practical cryptanalytic work to prepare a technical summary of the solution of a system. Step-by-step

---

[8] It is usual practice to employ as the specific key the equivalent of either A_p, or the equivalent of the first letter of the plain component when this component is a mixed sequence.

commentaries should accompany an initial solution; the steps taken should be jotted down as they are made, and at the end they should be combined into a complete résumé of the analysis. The résumé should be brief and concise, yet comprehensive enough that at any future time the solution may be reconstructed following the exact manner in which it was originally accomplished. Assumptions of words, etc., should be referred to with work sheet line- and column-indicators, and should be couched in the proper cryptologic language or symbols. A short exposition of the mechanics of the general system, enciphering alphabets, enciphering diagrams, etc., as well as all key words (together with their derivation) and specific keys should be included. On the work sheet there should be a letter-for-letter decryptment under the cipher text; the final plaintext version should be in word lengths, with any errors or garbles corrected. Nulls or indicators showing sentence separation, change of key, etc., may be enclosed in parentheses. All work sheets and notes should be kept together with the solution.

48. General remarks on the foregoing solution.--a. The example solved above is admittedly a more or less artificial illustration of the steps in analysis, made so in order to demonstrate general principles. It was easy to solve because the frequencies of the various cipher letters corresponded quite well with the normal or expected frequencies. However, all cryptograms of the same monoalphabetical nature can be solved along the same general lines, after a certain amount of experimentation, depending upon the length of the cryptogram, and the skill and experience of the cryptanalyst.[9]

b. It is no cause for discouragement if the student's initial attempts to solve a cryptogram of this type require much more time and effort than were apparently required in solving the foregoing purely illustrative example. It is indeed rarely the case that every assumption made by the cryptanalyst proves in the end to have been correct; more often it is the case that a good many of his initial assumptions are incorrect, and that he loses much time in casting out the erroneous ones. The speed and facility with which this elimination process is conducted is in many cases all that distinguishes the expert from the novice.

---

[9] The use of monoalphabetic substitution in modern military operations is exceedingly rare because of the simplicity of solution. However, such cases have occurred, and one rather illuminating instance may be cited. In an important communication on 5 August 1918, General Kress von Kressenstein used a single mixed alphabet, and the intercepted radio message was solved at American GHQ very speedily. A day later another message, but in a very much more difficult cipher system, was intercepted and solved. When translated, it read as follows:

"GHQ Kress:

The cipher prepared by General von Kress was at once solved here. Its further use and employment is forbidden.

Chief Signal Officer, Berlin."

c. Nor will the student always find that the initial classification into vowels and consonants can be accomplished as easily and quickly as was apparently the case in the illustrative example. The principles indicated are very general in their nature and applicability, and there are, in addition, some other principles that may be brought to bear in case of difficulty. Of these, perhaps the most useful are the following:

(1) In normal English it is unusual to find more than two consonants in succession, each of high frequency. If in a cryptogram a succession of three or four letters of high-frequency appear in succession, it is practically certain that at least one of these represents a vowel.[10]

(2) Successions of three vowels are rather unusual in English.[11] Practically the only time this happens is when a word ends in two vowels and the next word begins with a vowel.[12]

(3) When two letters already classified as vowel-equivalents are separated by a sequence of six or more letters, it is either the case that one of the supposed vowel-equivalents is incorrect, or else that one or more of the intermediate letters is a vowel-equivalent.[13]

(4) Reference to Table 7-B of Appendix 2 discloses the following:

Distribution of first 18 digraphs forming 25 percent of English text

Number of consonant-consonant digraphs	4
Number of consonant-vowel digraphs	6
Number of vowel-consonant digraphs	8
Number of vowel-vowel digraphs	0

Distribution of first 53 digraphs forming 50 percent of English text

Number of consonant-consonant digraphs	8
Number of consonant-vowel digraphs	23
Number of vowel-consonant digraphs	18
Number of vowel-vowel digraphs	4

---

10 Sequences of seven consonants are not impossible, however, as in STRENGTH THROUGH.

11 Note that the word RADIOED, past tense of the verb RADIO, is coming into usage.

12 A sequence of seven vowels is not impossible, however, as in THE WAY YOU EARN.

13 Some cryptanalysts place a good deal of emphasis upon this principle as a method of locating the remaining vowels after the first two or three have been located. They recommend that the latter be underlined throughout the text and then all sequences of five or more letters showing no underlines be studied attentively. Certain letters which occur in several such sequences are sure to be vowels. An arithmetical aid in the study is as follows: Take a letter thought to be a good possibility as the cipher equivalent of a vowel (hereafter termed a *possible vowel-equivalent*) and find the length of each interval from the possible vowel-equivalent to the next *known* (fairly surely determined) vowel-equivalent. Multiply the interval by the number of times this interval is found. Add the products and divide by the total number of intervals considered. This will give the *mean* interval for that possible vowel-equivalent. Do the same for all the other possible vowel-equivalents. The one for which the mean is the greatest is most probably a vowel-equivalent. Underline this letter throughout the text and repeat the process for locating additional vowel-equivalents, if any remain to be located.

The latter tabulation shows that of the first 53 digraphs which form 50 percent of English text, 41 of them, that is, over 75 percent, are combinations of a vowel with a consonant. In short, in normal English the vowels and the high-frequency consonants are in the long run distributed fairly evenly and regularly throughout the text.

(5) As a rule, repetitions of trigraphs in the cipher text are composed of high-frequency letters forming high-frequency combinations. The latter practically always contain at least one vowel; in fact, if reference is made to Table 10-A of Appendix 2 it will be noted that 36 of the 56 trigraphs having a frequency of 100 or more contain one vowel, 17 of them contain two vowels, and only three of them contain no vowel. In the case of tetragraph repetitions, Table 11-A of Appendix 2 shows that no tetragraph listed therein fails to contain at least one vowel; 27 of them contain one vowel, 25 contain two vowels, and 2 contain three vowels.

(6) Quite frequently when two known vowel-equivalents are separated by six or more letters none of which seems to be of sufficiently high frequency to represent one of the vowels A E I O, the chances are good that the cipher-equivalent of the vowel U or Y is present.

d. To recapitulate the general principles, vowels may then be distinguished from consonants in that they are usually represented by:

(1) high-frequency letters;

(2) high-frequency letters which do not readily contact each other;

(3) high-frequency letters which have a great variety of contact;

(4) high-frequency letters which have an affinity for low-frequency letters (i.e., low-frequency plaintext consonants).

e. In the foregoing example the amount of experimentation or "cutting and fitting" was practically nil. (This is not true of real cases as a rule.) Where such experimentation is necessary, the underscoring of all repetitions of several letters is very essential, as it calls attention to peculiarities of structure that often yield clues.

f. After a few basic assumptions of values have been made, if short words or skeletons of words do not become manifest, it is necessary to make further assumptions for unidentified letters. This is accomplished most often by assuming a word.[14] Now there are two places in every message which lend themselves more readily to successful attack by the assumption of words than do any other places--the very beginning and the very end of the message. The reason is quite obvious, for although words may begin or end with almost any letter of the alphabet, they usually begin

---

[14] This process does not involve anything more mysterious than ordinary, logical reasoning; there is nothing of the subnormal or supernormal about it. *If* cryptanalytic success seems to require processes akin to those of medieval magic, if "hocus-pocus" is much to the fore, the student should begin to look for items that the claimant of such success has carefully hidden from view, for the mystification of the uninitiated. If the student were to adopt as his personal motto for all his cryptanalytic ventures the quotation (from Tennyson's poem *Columbus*) appearing on the back of the title page of this text, he will frequently find "short cuts" to his destination and will not too often be led astray!

and end with but a few very common digraphs and trigraphs. Very often the association of letters in peculiar combinations will enable the student to note where one word ends and the next begins. For example suppose, E, N, S, and T have been definitely identified, and a sequence like the following is found in a cryptogram:

. . . E N T S N E . . .

Obviously the break between two words should fall either after the S of E N T S or after the T of E N T, so that two possibilities are offered: . . . E N T S / N E . . ., or . . . E N T / S N E . . .. Since in English there are very few words with the initial trigraph S N E, it is most likely that the proper division is . . . E N T S / N E . . .. Of course, when several word divisions have been found, the solution is more readily achieved because of the greater ease with which assumptions of additional new values may be made.

g. Although a considerable amount of detailed treatment has been devoted to vowel-consonant analysis, it is felt advisable again to caution the student against the natural tendency to accept without question the results of any one cryptanalytic technique exclusively, even one such as vowel-consonant analysis which seems quite scientific in character.

49. The "probable-word" method; its value and applicability.--a. In practically all cryptanalytic studies, short cuts can often be made by assuming the presence of certain words in the message under study. Some writers attach so much value to this kind of an "attack from the rear" that they practically elevate it to the position of a method and call it the "intuitive method" or the "probable-word method." It is, of course, merely a refinement of what in everyday language is called "assuming" or "guessing" a word in the message. The value of making a "good guess" can hardly be overestimated, and the cryptanalyst should never feel that he is accomplishing a solution by an illegitimate subterfuge when he has made a fortunate guess leading to solution. A correct assumption as to plain text will often save hours or days of labor, and sometimes there is no alternative but to try to "guess a word", for occasionally a system is encountered the solution of which is absolutely dependent upon this artifice.

b. The expression "good guess" is used advisedly. For it is "good" in two respects. First, the cryptanalyst must use care in making his assumptions as to plaintext words. In this he must be guided by extraneous circumstances leading to the assumption of probable words--not just any words that come to his mind. Therefore he must use his imagination but he must nevertheless carefully control it by the exercise of good judgement. Second, only if the "guess" is correct and leads to solution, or at least puts him on the road to solution, it is a good guess. But, while realizing the usefulness and the time and labor-saving features of a solution by assuming a probable word, the cryptanalyst should exercise discretion in regard to how long he may continue in his efforts with this method. Sometimes he may actually waste time by adhering to the method too long, if straightforward, methodical analysis will yield results more quickly.

c. Obviously, the "probable-word" method has much more applicability when working upon material the general nature of which is known, than when working upon more or less isolated communications exchanged between correspondents concerning whom or whose activities nothing is known. For in the latter case there is little or nothing that the imagination can seize upon as a background or basis for the assumptions.[15] However, in the case of military cryptanalysis in time of active operations there is, indeed, so great a probability that certain words and expressions are present in certain cryptograms that those words and expressions ("cliches") are often referred to as "cribs" (as defined in Webster's New Collegiate Dictionary: "...a plagiarism; hence, a translation, etc., to aid a student in reciting."). The cryptanalyst is quite sure they are present in the cryptogram under examination--what he must do is to "fit the crib to the text", that is, locate it in the cipher text.

d. Very frequently, the choice of probable words is aided or limited by the number and positions of repeated letters. These repetitions may be patent--that is, externally visible in the cryptographic text as it originally stands--or they may be latent--that is, externally invisible but susceptible of being made patent as a result of the analysis. For example, in a monoalphabetic substitution cipher, such as that discussed in the preceding paragraph, the repeated letters are directly exhibited in the cryptogram; later the student will encounter many cases in which the repetitions are latent, but are made patent by the analytical process. When the repetitions are patent, then the pattern or formula to which the repeated letters conform is of direct use in assuming plaintext words; and when the text is in word-lengths, the pattern is obviously of even greater assistance. Suppose the cryptanalyst is dealing with military text, in which case he may expect such words as DIVISION, BATTALION, etc., to be present in the text. The positions of the repeated letter I in DIVISION, of the reversible digraph AT, TA in BATTALION, and so on, constitute for the experienced cryptanalyst tell-tale indications of the presence of these words, even when the text is not divided up into its original word lengths.

e. The important aid that a study of word patterns can afford in cryptanalysis warrants the use of definite terminology and the establishment of certain data having a bearing thereon. The phenomenon herein under discussion, namely, that many words are of such construction as regards the number and positions of repeated letters as to make them readily identifiable, will be termed idiomorphism (from the Greek "idios"= one's own, individual, peculiar + "morphe"=form). Words which show this phenomenon will be termed idiomorphic. It will be useful to deal with the idiomorphisms symbolically and systematically as described below.

---

[15] General Givierge in his *Cours de Cryptographie* (p. 121) says: "However, expert cryptanalysts often employ such details as are cited above [in connection with assuming the presence of 'probable words'], and the experience of the years 1914 to 1918, to cite only those, prove that in practice one often has at his disposal elements of this nature, permitting assumptions much more audacious than those which served for the analysis of the last example. The reader would therefore be wrong in imagining that such fortuitous elements are encountered only in cryptographic works where the author deciphers a document that he himself enciphered. Cryptographic correspondence, if it is extensive, and if sufficiently numerous working data are at hand, often furnishes elements so complete that an author would not dare use all of them in solving a problem for fear of being accused of obvious exaggeration."

f. When dealing with cryptograms in which the word lengths are determined or specifically shown, it is convenient to indicate their lengths and their repeated letters in some easily recognized manner or by formulas. This is exemplified, in the case of the word DIVISION, by the formula ABCBDBEF; in the case of the word BATTALION, by the formula ABCCBDEFG. If the cryptanalyst, during the course of his studies, makes note of striking formulas he has encountered, with the words which fit them, after some time he will have assembled a quite valuable body of data. And after more or less complete lists of such formulas have been established in some systematic arrangement, a rapid comparison of the idiomorphs in a specific cryptogram with those in his lists will be feasible and will often lead to the assumption of the current word. Such lists can be arranged according to word length, as shown herewith:

3/aba : DID, EVE, EYE, etc.
abb : ADD, ALL, ILL, OFF, etc.
4/abac : ARAB, AWAY, etc.
abbc : ALLY, BEEN, etc.
abca : AREA, BOMB, DEAD, etc.
abcb : ANON, CEDE, etc.
etc. etc.

g. When dealing with cryptographic text in which the lengths of the words are not indicated or otherwise determinable, lists of the foregoing nature are not so useful as lists in which the words (or parts of words) are arranged according to the intervals between identical letters, in the following manner:

1 Interval	2 Intervals	3 Intervals	Repeated digraphs
-DiD-	AbbAcy	AbeyAnce	COCOa
-EvE-	ArAbiA	hAbitAble	-dERER
-EyE-	AblAtive	lAborAtory	ICICle
dIvIsion	AboArd	AbreAst	-ININg
revIsIon	-AciA-	AbroAd	bAGgAGe
etc.	etc.	etc.	etc.

h. The most usual practice, however, in designating idiomorphic patterns and classifying them into systematic lists is to assign a literal nomenclature to that portion of a word (or sequence of plaintext letters) which contains the distinctive pattern, beginning with the first letter which is repeated in the pattern and ending with the last letter which is repeated in the pattern. Thus, the word DIVISION would be termed as an idiomorph of the abaca class (based on the sequence IVISI contained therein), and the word BATTALION as an idiomorph of the abba class (based on the sequence ATTA). In Appendix 3 will be found a compendium of the more frequent military words in English, arranged according to word-lengths in alphabetical order and in rhyming order; in addition, there will be found in this appendix a listing of idiomorphs arranged first according to pattern and then according to the first letter of the idiomorphic sequence.

50. Solution of additional cryptograms produced by the same components.--a. To return, after a rather long digression, to the cryptogram solved in pars. 44 - 47, once the components of a cipher alphabet have been reconstructed, subsequent messages which have been enciphered by means of the same components may be solved very readily, and without recourse to the principles of frequency, or application of the probable-word method. It has been seen that the illustrative cryptogram treated in paragraphs 41 - 47 was enciphered by juxtaposing the cipher component against the normal sequence so that $A_p=S_c$. It is obvious that the cipher component may be set against the plain component at any one of 26 different points of coincidence, each yielding a different cipher alphabet. After the components have been reconstructed, however, they become known sequences and the method of converting the cipher letters into their plain-component equivalents and then completing the plain-component sequence[16] begun by each equivalent can be applied to solve any cryptogram which has been enciphered by these components.

b. An example will serve to make the process clear. Suppose the following message, passing between the same two stations as before, was intercepted shortly after the first message had been solved:

I Y E W K   C E R N W   O F O S E   L F O O H   E A Z X X

It is assumed that the same components were used, but with a different key letter. First the initial two groups are converted into their plain-component equivalents by setting the cipher component against the plain component at any arbitrary point of coincidence. The initial letter of the former may as well be set against A of the latter, with the following result:

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

Cryptogram---- I Y E W K   C E R N W   . . .
Equivalents--- P Y B F R   L B H E F   . . .

The plain component sequence initiated by each of these conversion equivalents is now completed, with the results shown in Fig. 15. Note the plaintext generatrix, CLOSEYOURS, which manifests itself without further analysis. The rest of the message may be read either by continuing the same process, or, what is even more simple, the key letter of the message may now be determined quite readily and the message deciphered by its means.

[16] It must be noted that if the plain component is a mixed sequence, then it is this mixed sequence which must be used to complete the columns.

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

Figure 15.

c. In order that the student may understand without question just what is involved in the latter step, that is, discovering the key letter after the first two or three groups have been deciphered by the conversion-completion process, the foregoing example will be used. It was noted that the first cipher group was finally deciphered as follows:

Cipher---------- I Y E W K
Plain----------- C L O S E

Now set the cipher component against the normal sequence so that $C_p = I_c$. Thus:

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

It is seen here that when $C_p = I_c$ then $A_p = F_c$. This is the key for the entire message. The decipherment may be completed by direct reference to the cipher alphabet. Thus:

Cipher-- I Y E W K C E R N W O F O S E L F O O H E A Z X X
Plain--- C L O S E Y O U R S T A T I O N A T T W O P M X X

Message: CLOSE YOUR STATION AT TWO PM

d. The student should make sure that he understands the fundamental principles involved in this quick solution, for they are among the most important principles in cryptanalytics. How useful they are will become clear as he progresses into more and more complex cryptanalytic studies.

e. It must be kept in mind that there are four ways that two basic sequences may be used to form a cipher alphabet, subject to the instructions guiding the cryptographer in the use of his cryptosystem; this fact must be considered when additional cryptograms appear in a particular cryptosystem for which the primary components have been recovered. Assuming that the sequences just recovered are labelled "A" and "B", then the following contingencies might arise in the encryption of subsequent messages:

(1) "A" direct for the plain component, and "B" direct for the cipher component (as in the original recovery);

(2) "A" direct for the plain, and "B" reversed for the cipher;

(3) "B" direct for the plain, and "A" direct for the cipher; and

(4) "B" direct for the plain, and "A" reversed for the cipher.

51. Derivation of key words.--a. Concurrent with the solution of a cryptogram, there should be a simultaneous effort in the reconstruction of cipher alphabets and recovery of key words. Much labor can thus be saved as recovery of the keys early in the stages of solution may transform the process of cryptanalysis into one of decipherment.

b. A mixed cipher alphabet falls into one of five categories, according to the composition of its components, viz.,

(1) the plain component is the normal sequence and the cipher component is mixed;

(2) the cipher component is the normal sequence and the plain component is mixed;

(3) both components are the same mixed sequence;

(4) both components are the same mixed sequence, but running in reverse; or

(5) the components are different mixed sequences.

c. Let us examine several types of mixed sequences, using the key word HYDRAULIC as an example. The ordinary keyword-mixed sequence produced from this key word is:

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

The two principal transposition-mixed types based on this key word are derived from the diagram:

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

and read:

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

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

Other types may arise from various types of route transpositions such as the following, using the foregoing diagram:

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

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

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

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

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

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

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

Any transposition system may be employed to produce a systematically-mixed sequence; practicability of method is the only determining factor. It must be remembered that the greatest amount of systematic mixing will produce a sequence inherently no more secure than a random-mixed alphabet.

d. The student would do well to construct both enciphering and deciphering versions of cipher alphabets recovered, as has been previously mentioned. For example, in the following case

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

no semblance of a key is apparent; but in the inverse form

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

the key-phrase "NOW IS THE TIME FOR ALL GOOD MEN TO COME TO THE AID OF THEIR PARTY" is quite clear. In other types of mixed sequences, first the one form is attacked, and then if negative results are obtained the inverse form is treated.

e. Let us consider the following cipher alphabet:

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

The section V W X seems to comprise superimposed parts of the non-keyword
J K L
portions of mixed sequences. Adding Y Z to the plain component, we get

V W X Y Z which is certainly consistent as far as alphabetical progres-
J K L N P
sion goes, and indicates that the letters M and O are present in the key word of the cipher component. Continuing in this vein, the section
M N O Q S T V W X Y Z is rapidly established by correlating both se-
B E F G H I J K L N P
quences. It is obvious that the plain component key word begins right after the Z, and that the cipher component key word probably just precedes the B. Going to the right, Z R H suggests key words like RHOMBOID,
P Q R
RHEUMATISM, etc. These trials are quickly repudiated; therefore we go on to Z R E which is acceptable. Z R E K is found wanting, but Z R E P is
P Q S   P Q S T   P Q S U
very satisfactory, and this is soon expanded to Z R E P U B L I C, and in
P Q S U V W X Y Z
a moment or two we recover the complete cipher alphabet:

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

f. In the example below the student will observe that the alphabets are reciprocal: this is an indication of identical sequences at a shift of 13, or that a mixed sequence running against itself in reverse has been employed. In this case the W X Y Z points to the latter hypothesis.
Z Y X W

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

Starting with the V W X Y Z R cluster, we see that the key word begins
R Z Y X W V
with the letter R; therefore the next letter should be a vowel. Z R A
W V H
is not acceptable, but Z R E is fine, showing that the letter U appears
W V T
in the key word. Continuing the same line of reasoning as in the preceding example, and with a little further experimentation, the final alphabet is discovered to be

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

g. In the next example, all efforts to derive key words on the basis of keyword-mixed sequences are fruitless: the conclusion is therefore drawn that this is a case of a transposition.

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

Considering the mechanics of the cryptography involved, and assuming for the time being that Z is at the bottom of the matrix and not in the key word, we start with the letters to the left, or if this fails, to the right of Z in the cipher component, obtaining the column N which is not
K
Z
incompatible if N is in the key word on the top row. If we place Y to

the left of Z and build up <u>its</u> column, we get E N which is excellent.

```
E N
J K
Y Z
```

This is expanded into

```
I M E N
G H J K
W X Y Z
```

which quickly becomes

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

This last example was very easy because none of the letters V W X Y Z appeared in the key word; but other cases should hardly prove more difficult.

<u>h</u>. Two additional methods that have been encountered for deriving mixed sequences may be mentioned. One is a slight modification of the preceding paragraph, when the key word contains repeated letters:

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

which produces the mixed sequence:

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

The other method is an interrupted-key columnar transposition system:[17]

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

which produces the mixed sequence:

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

The first example will succumb to the treatment outlined in subpar. <u>g</u>, whereas the second method is vulnerable owing to the presence of the fragments D J N, F K O, and G M P in the sequence which may be anagrammed. Note the fair-sized fragment B D J N R S, composed of an ascending sequence of letters; this is an outward manifestation of the interrupted-key columnar method.

<u>i</u>. There are still other methods used for the production of mixed sequences, but space does not permit giving further examples. However, the student should by this time be able to devise methods of attack for any special cases that may present themselves, based upon the cryptanalytically exploitable weaknesses or peculiarities inherent in the system of cryptography involved.

---

[17] It is to be noted that in this particular case the numerical key serves two purposes: (1) determining the cut-off point (and therefore the number of letters) in each <u>row</u> of the diagram, after the appearance of the keyword; and (2) determining the order of transcription of the <u>columns</u>.

TABLE OF CONTENTS

MILITARY CRYPTANALYSIS, PART I

Monoalphabetic Substitution Systems

(BLANK)

SECTION VII

MULTILITERAL SUBSTITUTION WITH SINGLE-EQUIVALENT CIPHER ALPHABETS

52. General types of multiliteral cipher alphabets.--a. Monoalphabetic substitution methods in general may be classified into uniliteral and multiliteral systems. In the former there is a strict "one-to-one" correspondence between the length of the units of the plain and those of the cipher text; that is, each letter of the plain text is replaced by a single character in the cipher text. In the latter this correspondence is no longer $1_p:1_c$ but may be $1_p:2_c$, where each letter of the plain text is replaced by a combination of two characters in the cipher text; or $1_p:3_c$, where a three-character combination in the cipher text represents a single letter of the plain text, and so on. A cipher in which the correspondence is of the $1_p:1_c$ type is termed uniliteral in character; one in which it is of the $1_p:2_c$ type, biliteral; $1_p:3_c$, triliteral, and so on. Ciphers in which one plaintext letter is represented by cipher characters of two or more elements are classed as multiliteral.[1]

b. Biliteral alphabets are usually composed of a set of 25 or 26 combinations of a limited number of characters taken in pairs. An example of such an alphabet is the following:

Plain--------	A	B	C	D	E	F	G	H	I	J	K	L	M
Cipher-------	WW	WH	WI	WT	WE	HW	HH	HI	HT	HT	HE	IW	IH
Plain--------	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher-------	II	IT	IE	TW	TH	TI	TT	TE	EW	EH	EI	ET	EE

This alphabet is derived from the cipher square or matrix shown in Fig. 16. The cipher equivalent of each plaintext element is made up of two coordinate letters from outside the cipher matrix, one letter being the coordinate of the row, the other being the coordinate of the column

---

[1] The terms uniliteral and multiliteral, although originally applied only to cipher text composed of letters, are used here in their broader sense to embrace cipher text in letters, digits, and even other symbols. In more precise terminology, these terms would probably be monosymbolic and polysymbolic, respectively, but the terms uniliteral and multiliteral are too well established in literature to be changed at this late time.

in which the plaintext letter is located. In other words, the letters at the side and top of the matrix have been used to designate, according to

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

Figure 16.

a coordinate system, the cell occupied by each letter within the matrix. The letters (or figures) constituting the coordinate elements of such matrices are termed row and column indicators.

c. If a message is enciphered by means of the foregoing biliteral alphabet, the cryptogram is still monoalphabetic in character. A frequency distribution based upon pairs of letters will obviously have all the characteristics of a simple, uniliteral distribution for a monoalphabetic substitution cipher.

d. The cipher alphabets shown thus far in this text have involved only letters, but alphabets in which the cipher component consists of figures, or groups of figures, are not uncommon in military cryptography.[2] Since there are but 10 digits it is obvious that, in order to represent an alphabet of more than 10 characters by means of figure ciphers, 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, . . . Z_p=26; that is, one in which the plaintext letters have as their equivalents two-digit numbers indicating their positions in the normal alphabet.

e. Instead of a simple alphabet of the preceding type, it is possible to use a diagram of the type shown in Fig. 17. In this cipher

	1	2	3	4	5	6	7	8	9	Ø
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 17.

[2] Although, as an extension of this idea, cipher alphabets employing signs and symbols are possible, such alphabets are not suitable for modern cryptography because they can be neither telegraphed nor telephoned with any degree of accuracy, speed, or facility.

the letter A_p is represented by the dinome[3] 11, B_p by the dinome 12, etc. Furthermore, this matrix includes provision for the encipherment of some of the frequently-used punctuation marks in addition to the 26 letters.

f. Other types of biliteral cipher alphabets are illustrated in the examples below:

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

Figure 18.

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

Figure 19.

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

Figure 20.

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

Figure 21.

g. It is to be noted that in alphabets of the foregoing types, the row indicators may be distinct from the column indicators (e.g., Fig. 18), or they may not (e.g., Fig. 19); of course, when there is any duplication between the row and column indicators, it is necessary to agree beforehand upon which indicator will be given as the first half of the equivalent for a letter, in order to avoid ambiguity. (In all of the systems described in this and subsequent sections of this text, the row indicator will always form the first part of an equivalent). When letters are used as row and column indicators they may form a key word (e.g., Fig. 20), or they may not (e.g., Fig. 21); the key words, if formed, may be identical (e.g., Fig. 16) or different (e.g., Fig. 20). Furthermore, the plaintext letters may be arranged within the matrix as a mixed sequence (e.g., Fig. 20), either systematically- or random-mixed; and the matrix may contain, in addition to the letters of the alphabet, punctuation symbols (Fig. 17), numbers (Figs. 20, 21), etc., permitting their encipherment as such, instead of having to be spelled out.

[3] A pair of digits is called a dinome; similarly, a trinome is a set of three digits; a tetranome, a set of four digits; etc. Although a single digit would properly be termed a mononome, for the sake of euphony it is shortened into the term monome.

h. When letters are used as row and column indicators, they may be selected so as to result in producing cipher text that resembles artificial words; that is, words composed of alternate vowels and consonants. For example, if in Figure 16 the row indicators consisted of the vowels A E I O U in this sequence from the top down, 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 may be called a false, or pseudo-code system.[4]

i. As a weak type of subterfuge, biliteral ciphers may involve a third character appended to the basic two-character cipher unit; this is done to "camouflage" the biliteral nature of the cipher text. This third character may be produced through the use of a cipher matrix of the type illustrated in Fig. 22 (wherein A_p=611, B_p=612, etc.); or the third character may be a "sum-checking" digit which is the non-carrying sum (i.e., the sum modulo 10)[5] of the preceding two digits, such as in the trinomes 257, 831, and 662; or it may merely be a randomly-selected character (inserted solely for the purpose of leading the cryptanalyst astray).

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

Figure 22.

j. Another possibility that lends itself to certain multiliteral ciphers is the use of a word spacer or word separator. This word separator might be represented by a value in the matrix; i.e., the separator is enciphered (for instance, the dinome "39" in Fig. 19 might stand for a word separator). The word separator might instead be a single element not otherwise used in the cryptosystem; i.e., unenciphered, and thus not giving rise to any possible ambiguity. Thus, in Fig. 19 the digit Ø and in Fig. 21 the letter J might be used as word separators, since no confusion would arise in decrypting.

---

[4] 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 International Telegraph Conference held in Madrid in 1932 amended these regulations to permit the use of 5-letter code groups containing any combination of letters. These unrestricted code groups were authorized for use after 1 January 1934.

[5] The term modulo (abbreviated mod) pertains to a cyclic scale or basis of arithmetic; thus, in the modulus of 7, the numbers 8 and 15 are equivalent to 1, and 9 and 16 are equivalent to 2, etc.; or expressed differently, 8 mod 7 is 1, 9 mod 7 is 2. In cryptology, many operations are expressed mod 10 and mod 26.

k. The biliteral alphabets yielded by the matrices of Figs. 16-21 may also be termed bipartite, because the cipher units of these alphabets may be divided into two separate parts whose functions are clearly defined, viz., row indicators and column indicators. As will be discussed later, this bipartite nature of most biliteral alphabets produced from cipher matrices constitute one of the weaknesses of these alphabets which make them recognizable as such to a cryptanalyst. However, it is possible to employ a cipher matrix in a manner which will produce a biliteral alphabet not bipartite in character. For example, using the matrix of Fig. 23 one could produce the following biliteral cipher alphabet in

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

Figure 23.

which the equivalent for any letter in the matrix is the sum of the two coordinates which indicate its cell in the matrix:

Plain--------	A	B	C	D	E	F	G	H	I	J	K	L	M
Cipher-------	14	20	19	12	22	23	24	10	18	18	25	17	26

Plain--------	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher-------	28	29	30	31	13	32	34	16	35	36	37	11	38

The cipher units of this alphabet are, of course, biliteral; but they are not bipartite. Note the equivalent of A_p, that is 14--if divided, it yields the digits 1 and 4 which have no meaning per se: plaintext letters whose cipher equivalents begin with 1 may be found in two different rows of the matrix, and those whose equivalents end in 4 appear in three different columns.

53. The Baconian and Trithemian ciphers.--a. An interesting example in which the cipher equivalents are five-letter groups and yet the resulting cipher is strictly monoalphabetic in character is found in the cipher system invented by Sir Francis Bacon (1561-1626) over 300 years ago. Despite its antiquity the system possesses certain features of

merit which are well worth noting.[6] Bacon proposes the following 24-element cipher alphabet, composed of permutations of two elements taken five at a time:[7]

A=aaaaa	I-J=abaaa	R=baaaa
B=aaaab	K=abaab	S=baaab
C=aaaba	L=ababa	T=baaba
D=aaabb	M=ababb	U-V=baabb
E=aabaa	N=abbaa	W=babaa
F=aabab	O=abbab	X=babab
G=aabba	P=abbba	Y=babba
H=aabbb	Q=abbbb	Z=babbb

If this were all there were to Bacon's invention it would be hardly worth bringing to attention. But what he pointed out, with great clarity and simple examples, was how such an alphabet might be used to convey a secret message by enfolding it in an innocent, external message which might easily evade the strictest kind of censorship. As a very crude example, suppose that a message is written in capital and lower-case letters, any capital letter standing for an "a" element of the cipher alphabet, and any small letter, for a "b" element. Then the external sentence "All is well with me today" can be made to contain the secret message "Help." Thus:

```
A L l  i s   W E l L   W I t H   m E   T o d a Y
a a b  b b   a a b a   a a b a   b a   a b b b a
    H              E          L              P
```

Instead of employing a device so obvious as capital and small letters, suppose that an "a" element be indicated by a very slight shading, or a

[6] For a true picture of this cipher, the explanation of which is often distorted beyond recognition even by cryptographers, see Bacon's own description of it as contained in his De Augmentis Scientiarum (The Advancement of Learning), as translated by any first class editor, such as Gilbert Watts (1640) or Ellis, Spedding, and Heath (1857, 1870). The student is cautioned, however, not to accept as true any alleged "decipherments" obtained by the application of Bacon's cipher to literary works of the 16th century. These readings are purely subjective.

[7] Bacon's alphabet was called by him a "biliteral alphabet" because it employs permutations of two letters. But from the cryptanalytic standpoint the significant point is that each plaintext letter is represented by a 5-character equivalent. Hence, present terminology requires that this alphabet be referred to as a quinqueliteral alphabet. Although the quinqueliteral alphabet affords 32 permutations, Bacon used only 24 of them, because in the 16th century the letters I and J, U and V were used interchangeably. Note the regularity of construction of Bacon's biliteral alphabet, a feature which easily permits its reconstruction from memory.

very slightly heavier stroke. Then a secret message might easily be thus enfolded within an external message of exactly opposite meaning. The number of possible variations of this basic scheme is very high. The fact that the characters of the cryptographic text are hidden in some manner or other has, however, no effect upon the strict monoalphabeticity of the scheme.

b. Almost 100 years before Bacon's time, the abbot Trithemius, born Johann von Heydenberg (1462-1516), invented a triliteral alphabet which he evidently intended to use in a fashion similar to Bacon's alphabet; i.e., as a means of disguise or cover for a secret text. This alphabet, modified to include the 26 letters of the present-day English alphabet, is shown in Fig. 23 below; it consists of all the permutations of three things taken three at a time, i.e., 3^3 or 27 in all.

A=111	D=121	G=131	J=211	M=221	P=231	S=311	V=321	Y=331
B=112	E=122	H=132	K=212	N=222	Q=232	T=312	W=322	Z=332
C=113	F=123	I=133	L=213	O=223	R=233	U=313	X=323	*=333

Figure 23.

The cipher text of course does not have to be restricted to digits; any groupings of three things taken three at a time will do.

54. <u>Analysis of multiliteral, monoalphabetic substitution ciphers</u>.--
<u>a</u>. Biliteral ciphers and those of the other multiliteral (triliteral, quadriliteral, . . .) types are often readily detected externally by the fact that the cryptographic text is usually composed of but a very limited number of different characters. They are handled in exactly the same manner as are uniliteral, monoalphabetic substitution ciphers. So long as the same character, or combination of characters, is always used to represent the same plaintext letter, and so long as a given letter of the plain text is always represented by the same character or combination of characters, the substitution is strictly monoalphabetic and can be handled in the simple manner described in the preceding section of this text.

<u>b</u>. In the case of biliteral ciphers in which the row and column indicators are not identical, and the direction of reading the cipher pairs is chosen at will for each succeeding cipher pair, an analysis of the <u>contacts</u> of the letters comprising the cipher pairs will disclose that there are <u>two</u> distinct families of letters, and a cipher pair will never consist of two letters of the same family. With this fact discovered, the cipher may be quickly reduced to uniliteral terms and solved in the manner previously mentioned.

<u>c</u>. If a multiliteral cipher includes provision for the encipherment of a word separator, the cipher equivalent of this word separator may be readily identified because it will have the <u>highest</u> frequency of any cipher unit. On the other hand, if the word separator is a <u>single</u> character (see subpar. 52<u>j</u>. on the use of the digit Ø and the letter J), this

character may be identified throughout the encrypted text by its positional appearance spaced "wordlength-wise" in the cipher text, and by the fact that it never contacts itself. If this single character is used as a null indiscriminately throughout the cipher text, instead of as a word separator, the analysis is a bit more complicated but not as great as might be thought.

d. As a general rule, it is advisable to reduce multiliteral cipher text to uniliteral equivalents, especially if a triliteral frequency distribution is to be made. If not more than 36 different combinations are present in a cryptogram, the extra values over 26 may be represented by digits for the purpose of this reduction. If, however, more than 36 different combinations are found in the encrypted text, it is usually not worth the trouble to attempt any uniliteral reduction, and the cipher text can be attacked in its multiliteral groupings.

e. As one of the first steps in the solution of any multiliteral cipher in letters which appears to involve the use of a cipher matrix, it is generally advisable to anagram the letters comprising the row and column indicators in an attempt to disclose any key words for these indicators. When the anagramming process does disclose such a key word or words, the next step is to make a skeleton reconstruction matrix which is a duplicate of the original enciphering matrix in that the indicators are arranged in the same order as on the original. Then, as plain text is recovered in the cryptogram by any of the methods outlined in the previous section of this text, the recovered plaintext letters should be inserted in the proper cells of the reconstruction matrix, so that any systematic arrangement of the plaintext letters, if present in the original, may be disclosed prior to recovery of the complete plain text. Furthermore, it may in some instances be found worthwhile, immediately after successfully uncovering the key words used as indicators, to make a frequency distribution of the particular cryptogram in the form of tally marks within the properly arranged frame of the reconstruction matrix, because it may be that a few moments' study of the locations of the crests and troughs in the distribution made in that form may, if the plaintext letters are arranged in the normal sequence or in a keyword-mixed sequence (especially if it is related to the key words for the indicators), provide a basis for the derivation of this sequence at one stroke, without recourse to analysis of the cipher text.

55. Historically interesting examples.--a. Two examples of multiliteral ciphers of historical interest will be cited as illustrations. During the campaign for the presidential election of 1876 (Hayes vs. Tilden) many cipher messages were exchanged between the Tilden managers and their agents in several states where the voting was hotly contested. Two years later the New York Tribune[8] exposed many irregularities in the

[8] New York Tribune, Extra No. 44, The Cipher Dispatches, New York, 1879.

campaign by publishing the decipherments of many of these messages. These decipherments were achieved by two investigators employed by the Tribune, and the plain text of the messages seems to show that illegal attempts and measures to carry the election for Tilden were made by his managers. Here is one of the messages:

JACKSONVILLE, Nov. 16 (1876).

GEO. F. RANEY, Tallahassee.

P p y y e m n s n y y y p i m a s h n s y y s s i t e p a a e n s h n s
p e n s s h n s m m p i y y s n p p y e a a p i e i s s y e s h a i n s s s p
e e i y y s h n y n s s s y e p i a a n y i t n s s h y y s p y y p i n s y y
s s i t e m e i p i m m e i s s e i y y e i s s i t e i e p y y p e e i a a s s
i m a a y e s p n s y y i a n s s s e i s s m m p p n s p i n s s n p i n s i m
i m y y i t e m y y s s p e y y m m n s y y s s i t s p y y p e e p p p m a
a a y y p i i t

L'Engle goes up tomorrow.

DANIEL.

Examination of the message discloses that only ten different letters are used. It is probable, therefore, that what one has here is a cipher which employs a multiliteral alphabet. First assuming that the alphabet is one in which combinations of two letters represent single letters of the plain text, the message is rewritten in pairs and substitution of arbitrary letters for the pairs is made, as seen below:

PP	YY	EM	NS	NY	YY	PI	MA	SH	NS	YY	SS	etc.
A	B	C	D	E	B	F	G	H	D	B	I	etc.

A triliteral frequency distribution is then made and analysis of the message along the lines illustrated in the preceding section of this text yields solution, as follows:

Jacksonville, Nov. 16.

GEO. F. RANEY, Tallahassee:

Have Marble and Coyle telegraph for influential men from Delaware and Virginia. Indications of weakening here, Press advantage and watch Board. L'Engle goes up tomorrow.

DANIEL.

b. The other example, using numbers, is as follows:

Jacksonville, Nov. 17.

S. PASCO and E. M. L'ENGLE:

84	55	84	25	93	34	82	31	31	75	93	82	77	33	55	42
93	20	93	66	77	66	33	84	66	31	31	93	20	82	33	66
52	48	44	55	42	82	48	89	42	93	31	82	66	75	31	93

DANIEL.

There were, of course, several messages of like nature, and examination disclosed that only 26 different numbers in all were used. Solution of these ciphers followed very easily, the decipherment of the one given above being as follows:

Jacksonville, Nov. 17.

S. PASCO and E. M. L'ENGLE:

Cocke will be ignored, Eagan called in. Authority reliable.

DANIEL.

c. The Tribune experts gave the following alphabets as the result of their decipherments:

AA=O	EN=Y	IT=D	NS=E	PP=H	SS=N
AI=U	EP=C	MA=B	NY=M	SH=L	YE=F
EI=I	IA=K	MM=G	PE=T	SN=P	YI=X
EM=V	IM=S	NN=J	PI=R	SP=W	YY=A
20=D	33=N	44=H	62=X	77=G	89=Y
25=K	34=W	48=T	66=A	82=I	93=E
27=S	39=P	52=U	68=F	84=C	96=M
31=L	42=R	55=O	75=B	87=V	99=J

They did not attempt to correlate these alphabets, or at least they say nothing about a possible relationship. The present author has, however, reconstructed the rectangle upon which these alphabets are based, and it is given below (Fig. 24).

2d Letter or Number (columns); 1st Letter or Number (rows)

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

Figure 24.

It is amusing to note that the conspirators selected as their key a phrase quite in keeping with their attempted illegalities - HIS PAYMENT - for bribery seems to have played a considerable part in that campaign. The blank cells in the matrix probably contained proper names, numbers, etc.[9]

56. The international (Baudot) teleprinter code.--a. Modern printing telegraph systems,[10] or teleprinter systems as they are more often called, make use of a five-unit code[11] or alphabet which is similar to the Baconian alphabet treated in par. 53. Like the Baconian alphabet, the teleprinter alphabet is composed of permutations of two elements taken five at a time, making it possible to obtain 32 different permutations, 26 of which are assigned to the letters of the alphabet, leaving 1 for an "idle condition" and 5 for certain printer operations called functions, such as "space", "figure shift", "letter shift," etc.

b. During electrical transmission, the two distinct elements of which each character is composed take the form of (1) a timed interval of electrical current and (2) a timed interval of no current, which are commonly referred to as "mark" impulses and "space" impulses, respectively. In certain operations, a paper tape is prepared of the traffic to be transmitted, or a paper tape may be prepared of the incoming traffic at the receiving end; in such tapes, the elements of the Baudot characters take the form of punched holes ("mark" impulses) and imperforate positions ("space" impulses).

---

9 As was mentioned in a previous footnote, a matrix containing such items would be termed a syllabary square; for example of such matrices see the treatment of syllabary squares and code charts in Section X.

10 Such systems are characterized by the transmission and reception-printing of messages by electrical means, incorporating two electrically-connected instruments resembling typewriters. When a key of the keyboard on the transmitting instrument is depressed, an electrical signal is transmitted to the receiving instrument, causing the corresponding character to be printed therein. Usually the message is printed at the local as well as the distant station. The system has been adapted to radio as well as wire and overseas cable transmission.

11 The five-unit code was first applied to teleprinter systems by Jean Maurice Emile Baudot (1845-1903), and is commonly known as the Baudot code. It is worthwhile to point out that Baudot apparently constructed his alphabet to correspond with normal frequencies of characters (with certain exceptions), since the most frequent ones are represented by permutations requiring the least electrical energy on the basis of "marking" and "spacing." In this respect Baudot "took a leaf out of Morse's notebook."

~~RESTRICTED~~

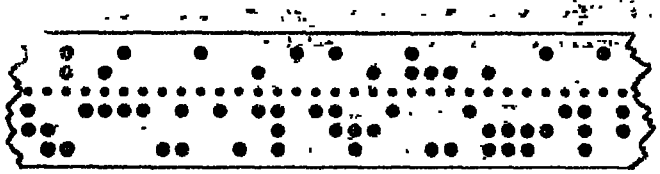
c. The teleprinter code in international use is given in Chart 7, below, wherein the mark and space impulses (known collectively as bauds) are illustrated as the holes (shown as black dots) and "no-holes" of a teleprinter tape. The letter equivalents ("lower case") are self-explanatory. The figure shift is used to change the meaning of a particular character to an "upper case" equivalent, and when it is desired to return to lower case, the letter shift is used; in regular teleprinter usage,

UPPER CASE	WEATHER SYMBOLS	↑	⊕	○	↗	3	→	↘	↓	8	↙	←	↖	.	⊕	9	Ø	1	4	Δ	5	7	ⓞ	2	/	6	+	–	<	≡			
	COMMUNICATIONS	–	?	:	$	3	!	&	£	8	'	(	)	.	,	9	Ø	1	4	Δ	5	7	;	2	/	6	"	»	<	≡			
LOWER CASE		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	BLANK	C.R.	L.F.	SPACE	LTR SHIFT	FIG. SHIFT
	1	●	●		●	●	●				●	●						●		●		●		●	●	●	●					●	●
	2	●		●				●		●	●	●	●				●	●	●			●	●	●						●		●	●
	FEED HOLES	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	3			●			●		●	●		●		●	●		●	●		●		●	●		●	●					●	●	
	4		●	●	●		●	●			●	●		●	●	●			●				●		●				●			●	●
	5		●					●	●				●	●		●	●	●			●		●	●	●	●	●					●	●

Chart 7. International teleprinter code.

the "communications" set of upper-case equivalents are the ones recorded on the typed copy by the teleprinter, whereas the "weather symbols" are the upper-case equivalents which are printed in teleprinter systems designed for the sending and receiving of weather information. The space is used to separate words; the carriage return (C.R.) effects the return of the teleprinter carriage to the right and the line feed (L.F.) rolls the platen to the next line for printing (cf. the corresponding functions of an ordinary typewriter). In addition, when the upper-case equivalent of "S" is used, a bell rings in the receiving teleprinter as a signal to call the operator to his machine, or to indicate that traffic is about to be sent.

d. In Fig. 25 is shown a portion of a teleprinter tape containing the beginning of the phrase "Now is the time for all good men . . ."



NOW IS THE TIME FOR ALL GOOD MEN

Figure 25.

The small holes, one of which appears in every position of the tape between the second and third levels, are sprocket holes used for advancing the tape through the transmitter unit.

~~RESTRICTED~~

e. It is to be emphasized that messages are not made secure from unauthorized reading merely by sending them by means of an ordinary teleprinter system--the teleprinter alphabet is internationally known, just as the English, Russian, etc. alphabets are. In order to provide security for a teleprinter message, it is just as necessary to apply thereto some sort of cryptographic treatment as it is to any other kind of message. The cryptosystems used for teleprinter encryption may involve either, or both, of the two classes of cryptographic treatment, viz., substitution and transposition. A substitution treatment might involve changing certain of the mark impulses of the characters comprising a message to space impulses, and vice versa, according to a prearranged system; a transposition treatment might involve changing the order of the 5 impulses in the Baudot equivalents for the characters comprising a message; and so on. The cryptographic treatment can be accomplished by a special cipher attachment (called an "appliqué unit") to a teleprinter; thus no modification of the teleprinter itself would be necessary. There are, of course, self-contained cipher teleprinters designed as such for engineering or cryptographic reasons, or both.

f. In the analysis of encrypted teleprinter systems, recourse is had to special tables[12] of the frequencies of single Baudot characters, digraphs, trigraphs, etc., as they appear in teleprinter traffic. It is important to note that in teleprinter traffic, as in any other type of traffic involving the use of a word separator, this character has the highest frequency of any plaintext element. Furthermore, one of the highest-frequency plaintext digraphs, in addition to those wherein the word separator constitutes one of the elements, will be the combination "carriage-return/line-feed", since this combination of characters is used in the normal procedure of typing each line of text on the teleprinter.

---

[12] In such tables, as is common in cryptanalytic practice, the mark impulses are designated by a plus symbol (+), and the space impulses are designated by a minus symbol (−). In addition, it is usual in such tables to denote the character representing the figure shift by the digit "2", the space by "3", the letter shift by "4", the line feed by "5", the blank by "6", and the carriage return by "7".

(BLANK)

Mr. Friedman

SECTION VIII

MULTILITERAL SUBSTITUTION WITH VARIANTS

57. Purpose of providing variants in monoalphabetic substitution.--

a. It has been seen that the individual letters composing ordinary intelligible plain text are used 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 which affords definite clues in the solution of simple monoalphabetic ciphers, such as those discussed in the preceding sections of this text. In addition, the associations which individual letters form in combining to make up words, and the peculiarities which certain of them manifest in plain text, afford further direct clues by means of which ordinary monoalphabetic substitution encipherments of such plain text may be more or less speedily solved. This has led cryptographers to devise methods for disguising, suppressing, or eliminating the foregoing characteristics manifested in cryptograms produced by the simpler methods of monoalphabetic substitution. One category of such methods, the one to be discussed in this section, is that in which the letters of the plain component of a cipher alphabet are assigned two or more cipher equivalents, which are called variant values (or, more simply, variants).

b. Basically, systems involving variants are multiliteral[1] and, in such systems, because of the large number of equivalents made available

---

[1] Uniliteral substitution with variants is also possible. Note the following cipher alphabet, illustrated by Captain Roger Baudouin in his excellent treatise, Eléments de Cryptographie, p. 101 (Paris, 1939):

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

Baudouin proposed that J_p and Y_p be replaced by I_p; K_p by C_p or Q_p; and W_p by VV_p--thus four cipher letters would be available as variants for the high-frequency plaintext letters in French.

by the combinations and permutations of a limited number of elements, each letter of the plain text may be represented by several multiliteral cipher equivalents which may be selected at random. For example, if 3-letter combinations are employed as the multiliteral equivalents, there are available 26^3 or 17,576 such equivalents for the 26 letters of the plain text; they may be assigned in equal numbers of different equivalents for the 26 letters, in which case each letter would be representable by 676 different 3-letter equivalents; or they may be assigned on some other basis, for example, proportionately to the relative frequencies of plaintext letters. For this reason this type of system may be more completely described as a monoalphabetic, multiliteral substitution with a multiple-equivalent cipher alphabet.[2] Some authors term such a system "simple substitution with multiple equivalents"; others term it "monoalphabetic substitution with variants", or multiliteral substitution with variants. For sake of brevity and precise terminology, the latter designation will be employed in this text, it being understood without further restatement that only such systems as are monoalphabetic will be discussed.

c. The primary object of monoalphabetic substitution with variants is, as has been mentioned above, to provide several values which may be employed at random in a simple substitution of cipher equivalents for the plaintext letters.

d. A word or two concerning the underlying theory of (monoalphabetic) multiliteral substitution with variants may not be amiss. Whereas in simple or single-equivalent substitution it has been seen that

(1) the same letter of the plain text is invariably represented by but one and always the same character of the cryptogram, and
(2) the same character of the cryptogram invariably represents one and always the same letter of the plain text,

in multiliteral substitution with variants it will be seen that

(1) the same letter of the plain text may be represented by one or more different characters of the cryptogram, but
(2) the same character of the cryptogram nevertheless invariably represents one and always the same letter of the plain text.

58. Simple types of cipher alphabets with variants.--a. The matrices shown on the next page provide some of the simpler means for accomplishing monoalphabetic substitution with variants. The systems incorporating these matrices are extensions of the basic ideas of multiliteral substitution treated in par. 52. The variant equivalents for any plaintext letter may be chosen at will; thus, in Fig. 26, E_p=10, 15, 60, or 65; in Fig. 27, E_p=AU_c, AZ_c, FU_c, FZ_c, LU_c, or LZ_c; etc.

---

[2] Cf. the title of the preceding section, "Multiliteral substitution with single-equivalent cipher alphabets."

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

Figure 26

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

Figure 27

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

Figure 28

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

Figure 29

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

Figure 30

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

Figure 31

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

Figure 32

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

Figure 33

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

Figure 34

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

Figure 35

b. It is to be noted that encipherment by means of the matrices in Figures 27, 28, and 31 is commutative; i.e., he coordinates may be read in either row-column or column-row order without cryptographic ambiguity, since there is no duplication between the row and column coordinates. The remaining matrices above are non-commutative; therefore a convention must be agreed upon as to the order of reading the coordinates. It should also be noted that in Figs. 30 and 31 the letters in the square have been inscribed in such a manner that, coupled with the particular arrangement of the row and column coordinates, the number of variants available for each plaintext letter is roughly proportional to the frequencies of the letters in plain text. A similar idea is found in Fig. 35, wherein the top row of the rectangle contains a word composed of high-frequency letters, and the coordinates are arranged in a manner roughly corresponding to the frequencies of plaintext letters. The matrix in Fig. 28 is a modification of the pseudo-code system described in par. 52h, with the added feature of variants.

c. Other simple ideas for producing variant systems are matrices such as the following:

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
08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	01	02	03	04	05	06	07
35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	26	27	28	29	30	31	32	33	34
68	69	70	71	72	73	74	75	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
87	88	89	90	91	92	93	94	95	96	97	98	99	00	76	77	78	79	80	81	82	83	84	85	86

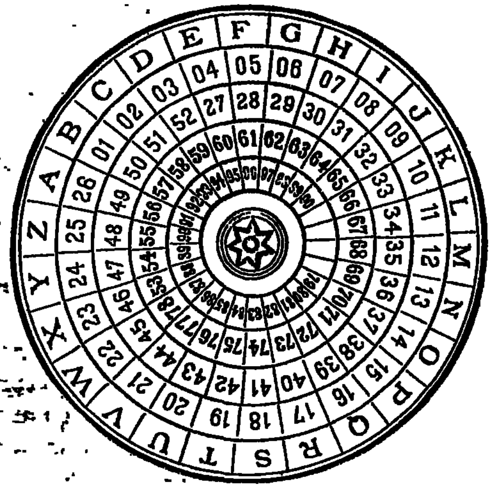
Figure 36

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
14	15	16	17	18	19	20	21	22	23	24	25	26	01	02	03	04	05	06	07	08	09	10	11	12	13
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	53	54	55	56	57
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00					79	80

Figure 37

In these two matrices there has been a regular inscription of the dinomes in the rows. Furthermore, in Fig. 36 the dinomes 01, 26, 51, and 76 (i.e., the lowest number in each of the four sequences) give the key word (TRIP) for that matrix; and in Fig. 37, the dinomes 01, 27, 53, and 79 denote the key word (NAVY) for that matrix. The security of systems involving such matrices would of course be greatly improved if the dinomes were assigned in a random manner; but then the easy mnemonic feature of the four sequences and the key word would be lost.

d. An interesting adaptation in a disc form of the type of matrix illustrated in Fig. 37 is the following device reputedly once used by the Mexican Army:



The device consisted of five concentric discs, the outer disc bearing the 26 letters of the alphabet, and the other four bearing the sequences 01-26, 27-52, 53-78, and 79-00. The rotatable discs made it possible to change the keys at frequent intervals, without the necessity of writing out a new matrix each time.

59. More complicated types of cipher alphabets with variants.--

a. Matrices such as those in Figs. 38, 39, and 40 below are termed __frequential matrices__, since the number of cipher values available for any given plaintext letter closely approximates its relative plaintext frequency.

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

(676 - cell matrix)

Figure 38

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

Figure 39

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

Figure 40

b. In the fragmentary matrix illustrated in Fig. 38, the number of occurrences of a particular letter within the matrix is proportional to its frequency in plain text; the letters are inscribed in a random manner, in order to enhance further the security of the system. In Fig. 39, we have a modification of the idea set forth in Fig. 38, except that the size of the matrix has been reduced from 26x26 to 10x10; in this case, the letters (with appropriate number of repetitions) have been inscribed in a simple diagonal route (lower left to upper right) within the square, and the coordinates have been scrambled, for greater security. In Fig. 40, there is illustrated a type of cipher square which is known in cryptologic literature as the Grandpré cipher; in this square there are inscribed ten 10-letter words containing all the letters of the alphabet in their approximate plaintext frequencies. These ten words are further linked together by a 10-letter word which appears vertically in the first column, as a mnemonic feature for the inscription of the words in the rows.

c. The frequential-type system represented in Fig. 41a (enciphering matrix) and 41b (deciphering matrix) was described by Sacco[3], who proposed that the dinomes inscribed in the enciphering matrix be thoroughly disarranged by applying a double transposition to the dinomes 00-99 as a means of suppressing any patent relationships among the variant values for the various plaintext letters; furthermore, the nulls incorporated in the matrix were to be used occasionally during the encryption of a message, in order to throw a cryptanalyst off the track. In this example the number of variant values for each plaintext letter has been established, of course, from the standpoint of Italian letter frequencies.

---

[3] Sacco, Generale Luigi, Manuale di Crittografia, 3d Ed., Rome, 1947, p. 22.

~~RESTRICTED~~

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

Figure 41a.

	1	2	3	4	5	6	7	8	9	∅
1	S	-	N	-	U	period	H	E	zero	I
2	-	Y	I	F	A	R	four	C	eight	Q
3	nine	comma	T	Z	E	U	E	G	M	O
4	P	-	H	one	-	seven	-	-	L	B
5	O	A	I	-	-	-	F	S	Z	three
6	period	A	-	-	E	one	O	M	A	five
7	E	O	N	-	I	-	Q	E	A	C
8	J	I	-	two	X	V	I	T	O	D
9	five	zero	B	R	W	K	G	P	U	-
∅	U	V	A	six	L	-	O	D	-	U

Figure 41b.

~~RESTRICTED~~

d. The Baconian cipher described in par. 53a may be used as a basis for superimposing additional complexities. For instance, the "a" elements may be represented by any one of the 20 consonants as variants, while the "b" elements may be represented by any one of the six vowels; or the letters A-M may be used to represent the "a" elements and the letters N-Z for the "b" elements; digits may be used for the "a" and "b" elements, either on the basis of the first five and last five digits, or on the basis of the odd and even digits; or the first 10 consonants (B-M) and the last 10 consonants (N-Z) may be used for the "a" and "b" elements, with the vowels used occasionally as nulls--thus the resultant cryptograms will resemble those of a fairly complex cryptosystem. However, once the cryptanalyst assumes the possibility of such a system, its complexity is more apparent than real. Similarly, variations of this genre may be superimposed on triliteral systems such as the Trithemian cipher illustrated in par. 53b; variants for the "1", "2", and "3" elements may be chosen in such a way as to provide a large number of equivalents for each basic triliteral combination.

e. Another scheme for a complex variant system is a summing-trinome system. In this cryptosystem, each plaintext letter is assigned a unique value of 1 to 26; this value is then expressed as a trinome, the digits of which sum to the designated value of the letter. For example, if a letter has been assigned the value "4", it may be represented by any one of the following permutations and combinations[4]:

004	031	112	202	301
013	040	121	211	310
022	103	130	220	400

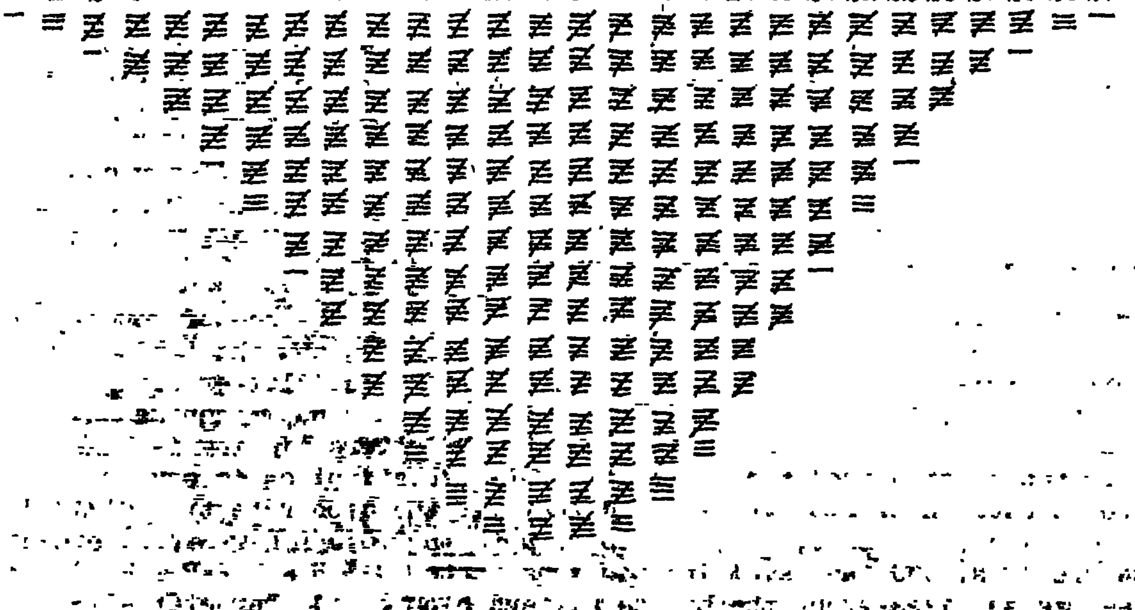
Since the values toward the middle of the range 1-26 may be represented by a very considerable number of summing-trinomes (e.g., for the values 13 and 14 there are 75 variants each), such a system would offer a cryptographer wide latitude in the choice of cipher equivalents in enciphering,

---

[4] The representations of an integer (i.e., a whole number) as the sum of integers in all possible ways are termed the partitions of that number. The partitions in this subparagraph are mod 10 and also include the digit Ø in order to form trinome equivalents out of all the possible permutations.

especially if the basic values of the plaintext letters were chosen to correspond with the scale of their relative frequencies, such as the following:

	J	Q	B	W	Y	U	F	H	D	I	O	N	E	T	R	A	S	L	C	P	M	G	V	X	K	Z	
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27



The tallies beneath each value represent the number of variants possible for the particular value. The unused values for Ø and 27 (uniquely represented by 000 and 999, respectively) may be used for punctuation marks, nulls, or other special-purpose symbols. Since such a system, once suspected, would offer little difficulty[5] to a cryptanalyst, certain modifications would be necessary in order to pose any real obstacles in the way of solution. For instance, if the numerical value of a letter is expressed by permutations of 3 letters (instead of digits) out of a set of the 10 letters A-J wherein the sequence of the letters A-J represents a <u>disarranged</u> sequence of the digits Ø-9, such a system may be among the most complex types of ciphers in the realm of monoalphabetic substitution, requiring the solution of many simultaneous equations. A further refinement would involve the use of all 26 letters as variants, in predetermined groups, to represent the digits Ø-9. Fortunately for the cryptanalyst, such systems are impracticable for field military use; but if they were encountered, a sufficiently large volume of text, coupled with Hitt's four essentials quoted in Section I, would eventually make a solution possible. The actual cryptanalytic complexity of certain apparently exceedingly complex cryptosystems is dependent on their being correctly used at all times, which is not invariably the case with military ciphers.

---

[5] The solution would involve simply dividing the cipher text into groups of 3 digits, summing the trinomes thus produced to yield 28 possible basic values, and solving these basic values as in any simple monoalphabetic substitution cipher.

60. Analysis of simple examples.--a. The following cryptogram is available for study:

```
QMDCV  PLFNF  DHNWJ  WLKDK  NHBPV  RLTVM
BKLWD  WVHVK  SHBCL  PQKJR  VWSML  KGCNR
LRNKV  MGFXW  JRGMV  WGTJH  QKXFN  ZVFDM
LTBPL  PVFLM  DCNWN  HBCVZ  NMLWQ  FDHDW
VZBRV  KLCVC  VRDHL  RVTLF  NCDKG  MXWXM
DTSCB  CLZLR  LMVTS  ZNKBW  VPBRN  CLRXR
DCNKV  PBTNT  GHJZL  FQFVK  BWDZX  PNHSP
GHLKL  FVZLT  VMLKD  PQRNZ  LZDTB  MNTGM
NZVFX  KSFDC  LZVTV  FDFVR  GCLPQ  PNCDW
VRJTN  HLZLM  VWNPV  PDZDW  JPNWL  RJKVM
XMDTS  MGFDR  DKLWJ  FLPJM  SFQWB  FNCBZ
DKVWG  ZSHBH  DHJCX
```

The first thing that strikes the eye is the total absence of vowels, remarkable not only because six letters are missing (cf. the Λ test) in a text of this size, but also because all six of these letters fall into an identical limited category--a significant non-random phenomenon. Since a unilateral substitution alphabet with six letters missing is highly improbable, the conclusion of multiliteral substitution is obvious. Upon closer inspection it is found that, if the cipher text is divided into pairs of letters, only ten consonants (B D G J L N Q S V X) are used as prefixes, and the remaining ten consonants (C F H K M P R T W Z) are used as suffixes--thus the biliteral (and bipartite) characteristics of the cipher text are disclosed. A digraphic[6] distribution is therefore constructed:

	C	F	H	K	M	P	R	T	W	Z
B	≡	—	—	—	—	=	=	—	=	—
D	≣	—	≣	≣	—	—	—	≡	≣	=
G	=	=	=		≡			—		—
J	—	—	—	—	—	—	=	—	—	—
L	—	≣		≣	≡	≣	≣ (struck)	≡	≡	≣
N	≣	—	≣	≡	—	—	—	=	≣	≡
Q		=		=	—	—	—		—	
S	—	=	=		=	—				—
V	—	≣	—	≡	≣	≣	≣	≡	≣	≡
X		—		—	=	—	—		=	

6 If it had not been noticed that the cryptogram should be divided into pairs for analysis, a biliteral distribution (see par. 23d) might have been made, in order to reveal contact affinities of the cipher letters.

RESTRICTED

b. It is possible that the cryptogram under study may involve the use of a small enciphering matrix with variants for the rows and columns. Since there is available an easily-applied special solution which permits the determination of the row indicators which are equivalent (i.e., interchangeable variants) and the column indicators which are equivalent, merely from a study of the digraphic distribution, this possibility is examined. The special solution is based on the following considerations: in a message of moderate length for such a cryptosystem, it may be assumed that the various possible cipher pairs for a given plaintext letter will be used with approximately equal frequency; for this reason, the cipher letters which pair with one of the letters used to indicate any particular row of the enciphering matrix may be expected to pair equally often with any other cipher letter which has been used to indicate the same row (and, of course, the same is true concerning the column-indicator letters). Thus, in the digraphic distribution of such a cryptogram, sets of rows appear which have similar "profiles" and, likewise, sets of similar columns.[7] First a study will be made of the rows of the distribution just compiled, in an attempt to locate and isolate those which match with each other; then, the same will be done with the columns of the distribution.

c. It is noted that the "L" and "V" distributions have pronounced similarities (Fig. 42a)--these rows came under consideration first because of their unique "heaviness" of their frequency characteristics. Likewise, the "D" and "N" rows have homologous attributes in their appearance (Fig. 42b). However, the further grouping of the rows by ocular inspection may present difficulties to the student, since he may not yet trust his eye

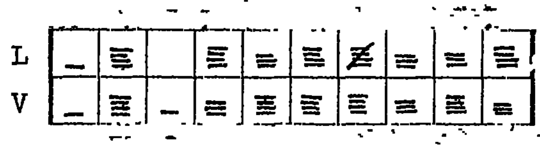


Figure 42a.

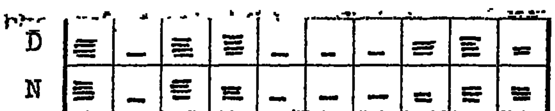


Figure 42b.

in matching distributions; and he may feel the need for some kind of statistical assurance. In the following subparagraphs there is given the technique of a more precise method for matching, mathematical in nature.

[7] These similarities are especially pronounced when the encipherer uses a "check-off" procedure for choosing his variants for each letter, that is, when he systematically "checks off" the variants used during encryption to insure that all possible variants are used in approximately equal proportions.

RESTRICTED

d. This method of matching in an attempt to "equate" interchangeable variants involves computing a separate value for each trial matching of a particular row (or column) against each of a series of other rows (or columns, as appropriate)--such a value is taken as an indication of the "goodness of match" exhibited by the particular trial, the theory being that the correct match will produce the highest value.[8] The value for a particular trial match is computed by multiplying the number of tallies in each cell of one row (or column) by the number of tallies in each corresponding cell in the other row (or column) and then totaling the products thus obtained. Because of the way in which it is produced, such a value is termed a "cross-products sum".

e. In subparagraph c above, it was determined that the "L" and "V" rows were equivalent, and that the "D" and "N" rows also formed an equivalent pair. The next "heavy" row is the "G" row; this is to be tested for match with the five remaining unmatched rows. Let the "G" row be tested first against the "B" row. These two rows are given below, with their cross-products sum. For convenience, the cross-products sum is symbolized by $\chi(\theta^1, \theta^2)$, where θ^1 and θ^2 represent the designators of the distributions to be matched.[9]

"G": 2 2 2 - 3 - - 1 - 1
"B": 3 1 1 1 1 2 2 1 2 1
χ(G,B): 6 2 2 - 3 - - 1 - 1 = 15

The complete table of the comparisons of the "G" row with the five available rows is as follows:

χ(G,B): 6 2 2 - 3 - - 1 - 1 = 15
χ(G,J): 2 2 2 - 3 - - 1 - 1 = 11
χ(G,Q): - 4 - - 3 - - - - - = 7
χ(G,S): 2 4 4 - 6 - - - - 1 = 17
χ(G,X): - 2 - - 6 - - - - - = 8

The results indicate that the most probable match with the "G" row is the "S" row.

f. Since the next "heaviest" row to be tested is the "B" row, its matchings with the three remaining rows are made, and are given below:

χ(B,J): 3 1 1 1 1 2 4 1 2 1 = 17
χ(B,Q): - 2 - 2 1 2 2 - 2 1 = 12
χ(B,X): - 1 - 1 2 2 2 - 4 - = 12

---

[8] In this connection, note the considerations treated in subpar. 60j.

[9] The Greek letter χ (chi) is often used in cryptology to symbolize matching operations.

The correct matching of the "B" and "J" rows is indicated by the results. This leaves only the "Q" and "X" rows, which are presumed to go together, since not only is their cross-products sum satisfactory (when compared to the χ values for some of the other rows which have been matched), but, equally important, their patterns of <u>crests</u> and <u>troughs</u> are similar. Since we have not found more than <u>two</u> rows for any one set of interchangeable values, it appears that the original matrix had only five rows, with two variants for each row. The rows of the distribution diagram are therefore combined in the following diagram:

	C	F	H	K	M	P	R	T	W	Z
BJ	4	2	2	2	2	3	4	2	3	2
DN	8	2	8	7	2	2	2	5	7	5
GS	3	4	4	-	5	1	-	1	-	2
LV	2	8	1	7	7	8	9	6	7	7
QX	-	3	-	3	3	2	2	-	3	-

Figure 43

<u>g</u>. Analysis of the distributions of the columns of Fig. 43 quickly reveals that columns "C" and "H" may be matched as a pair, and likewise columns "F" and "M", and columns "P" and "R". In order to decide the groupings of the remaining columns, the six possible χ values are derived:

χ(K,T): 4 35 - 42 - = 81
χ(K,W): 4 49 - 49 9 = 113
χ(K,Z): 4 35 - 49 - = 88
χ(T,W): 6 35 - 42 - = 83
χ(T,Z): 4 25 2 42 - = 73
χ(W,Z): 6 35 - 49 - = 90

Combinations:
KT, WZ: 81 + 90 = 171
KW, TZ: 113 + 73 = 186
KZ, TW: 88 + 83 = 171

It appears that the proper pairings of the columns are "K" and "W", "T" and "Z".

<u>h</u>. The groupings of the columns having been determined, the frequency diagram is reduced to its basic 5x5 square, and the ϕ test is

	C	F	K	P	T
	H	M	W	R	Z
BJ	6	4	5	7	4
DN	16	4	14	4	10
GS	7	9	-	1	3
LV	3	15	14	17	13
QX	-	6	6	4	-

ϕ_p=1962
ϕ_r=1132
ϕ_o=1670

taken as further statistical assurance of the matchings. Although ϕ_o in this case does not come up to the best expectations, we feel nevertheless that the matching has been carefully and correctly accomplished, and so the next step is continued with a conversion of the multiliteral text into uniliteral equivalents, using the following reduction square containing an arbitrary sequence:

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

The converted cryptogram is now easily solved, using the principles set forth in Section VI. The first fifteen letters of the plaintext message are found to read "WEATHER FORECAST.....", and the original enciphering matrix is recovered, based on the key word ATMOSPHERIC, as follows:

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

i. The method of matching rows and columns just described in the preceding subparagraphs applies equally well to all the matrices in Figs. 26-35, and similar variations. If in the process of equating indicators the cryptanalyst sees that the row indicators are falling into the same groupings as the column indicators, he might be able to accelerate the equating process by taking advantage of this feature alone, as would be the case if he had encountered a cryptogram involving a matrix with indicators arranged in a manner similar to that shown in Figs. 29 and 30. Furthermore, a cryptogram enciphered in a commutative system, wherein the equivalents have been taken in row-column and column-row order indiscriminately, may be recognized as such through a study of the digraphic distribution of the cryptogram since the " α " row of the distribution will have an appearance similar to the " α " column, the " β " row will be similar to the " β " column, etc;[10] this matter is discussed further in subpar. 61d.

[10] It is often convenient to use arbitrary symbols in cryptanalytic work, to prevent confusion with designations of actual elements of plain text, cipher text, or key (see footnote 1 on page 58). For this purpose Greek letters are often used; for reference, the 24 letters of the Greek alphabet and their names are appended in the chart below:

Α α alpha	Ε ε epsilon	Ι ι iota	Ν ν nu	Ρ ρ ro	Φ φ phi
Β β beta	Ζ ζ zeta	Κ κ kappa	Ξ ξ xi	Σ σ sigma	Χ χ chi
Γ γ gamma	Η η eta	Λ λ lambda	Ο ο omicron	Τ τ tau	Ψ ψ psi
Δ δ delta	Θ θ theta	Μ μ mu	Π π pi	Υ υ ypsilon	Ω ω omega

j. It is important to point out that in matching, the cryptanalyst should begin with the "best" rows or columns--best not only from the standpoint of "heaviness" of the distribution, but also best from the point of view of a distinctive pattern of crests and troughs. If insufficient text is available to allow equating all the interchangeable coordinates of a particular enciphering matrix, it may still be possible that a conversion of the cipher text by means of a partially-reduced reconstruction matrix may yield enough idiomorphic patterns and other data to make possible an entry into the text. If the cryptographer has not used a "check-off" process in enciphering, but instead has favored certain equivalents for the various plaintext letters, matching may not be possible; nevertheless, an entry into the text may be facilitated in this case, because some of the resultant peaks in the cipher text may be correctly identified. Furthermore, since no variant system can possibly disguise the letters of low frequency in plain text, their low-frequency equivalents in the cipher text may provide possible approaches to solution. (See also subpar. 61e).

k. In addition to the method of solution by matching and combining rows and columns of a digraphic distribution of a multiliteral cipher, there is also the general approach applicable without exception to any variant system. This method, involving the correlation of cipher elements suspected to be the equivalents of specific but unknown plaintext letters, is treated in detail in paragraphs 61 and 62.

l. Systems such as the 4-level dinome cipher illustrated in Fig. 36 are susceptible to a very easy solution, if the dinomes have been inscribed in numerical order as indicated. Assuming such a case in a specific cryptogram, the first six groups of which are

68321 09022 48057 65111 88648 42036 . .

a four-part frequency distribution of the entire message, is taken as illustrated in Fig. 44 below:

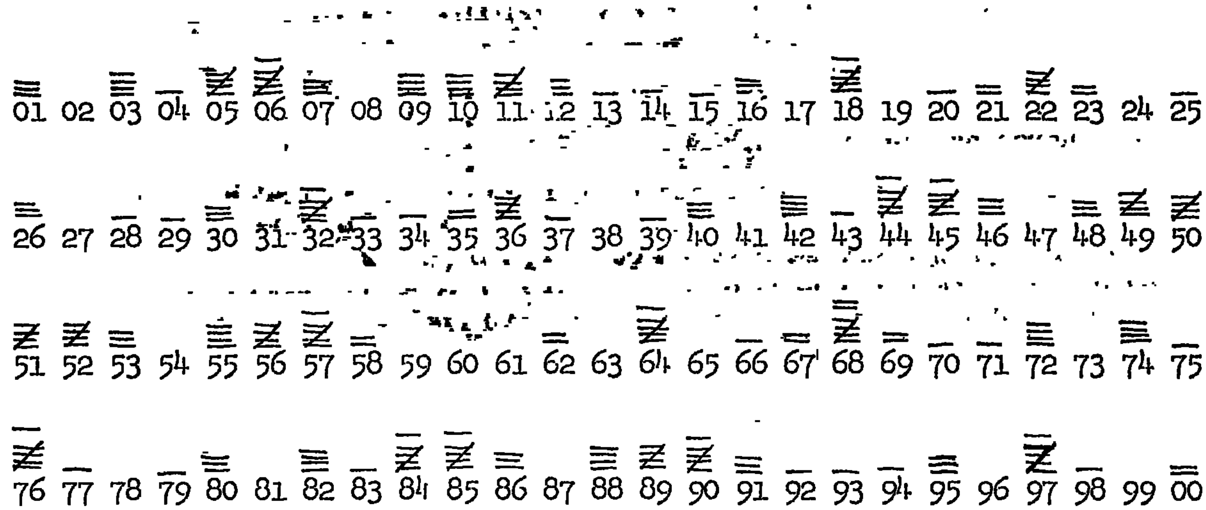


Figure 44.

If the student will bring to bear upon this problem the principles he learned in Section V of this text, he will soon realize that what he now has before him are four simple, monoalphabetic frequency distributions similar to those involved in a monoalphabetic substitution cipher using standard alphabets. The realization of this fact immediately provides the clue to the next step: "fitting each of the distributions to the normal". (See par. 31). This can be done without difficulty in this case (remembering that a 25-letter alphabet is involved and assuming that I and J are combined) and the following alphabets result:

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

The key word is seen to be JUNE and the beginning of the cryptogram is deciphered as "EASTERN ENTRANCE....."

m. If instead of 25-element alphabets, a system such as that in Fig. 37 has been used, only a slight modification of the procedure in subparagraph j would have been necessary, i.e., the distributions would have had to be considered on a basis of 26, and the process of fitting the distributions to the normal would have gone on as in the previous example.

n. One further application of principles learned in Section V deserves to be mentioned here, in connection with the solution of systems such as those of Fig. 36. Let the following short message be considered:

48226 88423 52099 93604 76059 05651
36683 52267 97114 54466 76

If it is known that the correspondents have been using a variant system such as that in Fig. 36, a special solution may be employed in those cases wherein there is insufficient cipher text to permit analysis by the method of fitting the frequency distribution to the normal. Thus, a short cryptogram may be solved by a variation of the plain-component completion method described in par. 34.[11] First, let the cryptogram be copied in dinomes, with an indication of the level (i.e., the "alphabet") the dinome would occupy in the 4-level matrix; thus:

48	22	68	84	23	52	09	99	36	04	76	05	90	56	51	36	68	35	22	67	97	11	45	44	66	76
2	1	3	4	1	3	1	4	2	1	4	1	4	3	3	2	3	2	1	3	4	1	2	2	3	4

The dinomes belonging to the four levels are as follows:

(1) 22 23 09 04 05 22 11
(2) 48 36 36 35 45 44
(3) 68 52 56 51 68 67 66
(4) 84 99 76 90 97 76

These dinomes are converted into terms of the plain component by setting each of the cipher sequences against the plain component at an arbitrary point of coincidence, such as in the following example:

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
01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00

(1) 22=W; 23=X; 09=I; 04=D; 05=E; 22=W; 11=L
(2) 48=X; 36=L; 36=L; 35=K; 45=U; 44=T
(3) 68=S; 52=B; 56=F; 51=A; 68=S; 67=R; 66=Q
(4) 84=I; 99=Y; 76=A; 90=P; 97=W; 76=A

[11] It should be clear to the student that the reason this method can be applied in this instance is that both the plain component (ABC.....Z) and the cipher component (01, 02, 03 ..... 25; 26-50, 51-75, 76-00) are known sequences (or thus assumed).

~~RESTRICTED~~

o. The plain component sequence is now completed on the letters of the four levels, as follows:

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

It is seen that the generatrices with the best assortment[12] of high-frequency letters for the four levels are:

1st level	2d level	3d level	4th level
E F R M N E T	R E E D O N	E O S N E D C	N C E T A E

[12] In evaluating generatrices, the sum of the arithmetical frequencies of the letters in each row may be used as an indication of their relative "goodness". A statistically much more accurate method of evaluating generatrices involves the use of logarithms of the probabilities of the plaintext letters forming the generatrices. (See also footnote 7 on page 89.)

~~RESTRICTED~~

RESTRICTED

If the letters of these generatrices are arranged in the order of appearance of their dinome equivalents, according to the way they fall into the various levels,

48	22	68	84	23	52	09	99	36	04	76	05	90	56	51	36	68	35	22	67	97	11	45	44	66	76
	E			F		R			M		N							E			T				
R								E							E		D					O	N		
		E			O								S	N		E			D					C	
			N				C			E		T								A					E

the plain text "REENFORCEMENTS NEEDED AT ONCE" is clearly seen. Or, more simply, if we examine the equivalents of 01, 26, 51, and 76 after the generatrix determination has been made, the key word JUNE is revealed. If an error had been made in the selection of a generatrix, the error could be resolved by hypothesizing the probable key word, or by deciphering the text on the basis of the assumed diagram and then noting and degarbling the systematic errors (which, it would be noticed, all come from one level).

p. The student should note that no one generatrix will yield plain text all the way across as in the example in par. 34. Instead, the generatrices must be considered separately for the four levels, since it is within each of the four levels that there is a homogeneous relationship of dinomes. Obviously if dinomes from more than one level were used to complete the plain component sequence, the generatrices would not consist of a homogeneous group of letters but instead would represent an assortment of letters from two or more "alphabets".

61. Analysis of more complicated examples.--a. As soon as a beginner in cryptography realizes the consequences of the fact that letters are used with greatly varying frequencies in normal plain text, a brilliant idea very speedily comes to him. Why not disguise the natural frequencies of letters by a system of substitution using many equivalents, and let the numbers of equivalents assigned to the various letters be more or less in direct proportion to the normal frequencies of the letters? Let E, for example, have 13 equivalents; T, 9; N, 8; etc., and thus (he thinks) the enemy cryptanalyst can have nothing in the way of telltale or characteristic frequencies to use as an entering wedge.

b. If the text available for study is small in amount and if the variant values are wholly independent of one another, the problem can become exceedingly difficult. But in practical military communications such methods are rarely encountered, because the volume of text is usually great enough to permit of the establishment of equivalent values. To illustrate what is meant, suppose a number of cryptograms produced by the monoalphabetic-variant method described above show the following two sets of groupings[13] of cipher elements in the text, Set "A" being assumed to be different representations of one particular underlying plain text, and Set "B" assumed to be representations of another underlying plain text:

Set "A"	Set "B"
(12-37-02-79-68-13-03-37-77)	(71-12-02-51-23-05-77)
(82-69-02-79-13-68-23-37-35)	(11-82-51-02-03-05-35)
(82-69-51-16-13-13-78-05-35)	(11-91-02-02-23-37-35)
(91-05-02-01-68-42-78-37-77)	(97-12-51-02-78-69-77)

An examination of these groupings would lead to the following tentative conclusions with regard to probable equivalents:

(12, 82, 91)	(02, 51)	(13, 42, 68)	(35, 77)
(05, 37, 69)	(01, 16, 79)	(03, 23, 78)	(11, 71, 97)

The establishment of these equivalencies would sooner or later lead to the finding of additional sets of equal values. The completeness with which this can be accomplished will determine the ease or difficulty of solution. Of course, if many equivalencies can be established the problem can then be reduced practically to monoalphabetic terms and a speedy solution can be attained.

c. Theoretically, the determination of equivalencies may seem to be quite an easy matter, but practically it may be very difficult, because the cryptanalyst can never be <u>certain</u> that a combination showing what may appear to be a variant value is really such and does not represent a part of a <u>different</u> plaintext sequence. For example, take the groups --

17-82-31-82-14-63, and
27-82-40-82-14-63

Here one might suspect that 17 and 27 represent the same letter, 31 and 40 another letter. But it happens that one group represents the word MANAGE, the other DAMAGE. There are hundreds of such cases in English and in other languages.

d. When reversible combinations are used as variants, the problem is perhaps a bit more simple. For example, using the accompanying Fig. 45

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

Figure 45

---

[13] The alert student might be able to determine the underlying plain text of the two sets of ciphertext groupings.

for encipherment, two messages with the same initial words, REFERENCE YOUR, may be enciphered as follows:

	R E F E R	E N C E Y	O U R			
(1)	N H W D R	X L S H C	D W W Z N	R S L H P	S R B J C	H
(2)	C H D W R	X S L H N	D W Z W N	R L S H P	R W J B N	H

The experienced cryptanalyst, noting the appearance of the very first few cipher groups, assumes that not only have the messages identical beginnings in their plain texts, but also that he is here confronted with a variant system involving biliteral reversible equivalents. One of the manifestations of such a cryptosystem is that in the digraphic distribution of the cipher text the "B" row will have an appearance similar to the "B" column, the "C" row will resemble the "C" column, etc.; thus, the cryptanalyst will almost immediately realize that he has encountered a commutative system involving a matrix smaller than that indicated by the size of matrix necessary for making the digraphic distribution.

e. The probable-word method of solution may be used, but with a slight variation introduced because of the fact that, regardless of the system, letters of low frequency in plain text remain infrequent in the cryptogram. Hence, suppose a word containing low-frequency letters, but in itself a rather common word strikingly idiomorphic in character is sought as a "probable word"; for example, words such as CAVALRY, ATTACK, and PREPARE. Such a word may be written on a slip of paper and slid one interval at a time under the text, which has been marked so that the high- and low-frequency characters are indicated. Each coincidence of a low-frequency letter of the text with a low-frequency letter of the assumed word is examined carefully to see whether the adjacent text letters correspond in frequency with the other letters of the assumed word; or, if the latter presents repetitions, whether there are correspondences between repetitions in the cipher text and those in the word. Many trials are necessary but this method will produce results when the difficulties are otherwise too much for the cryptanalyst to overcome.

62. Analysis involving the use of isologs.--a. In military communications it is not unusual that cryptograms are produced containing identical plain text but which have been subjected to different cryptographic treatment, thus yielding different cipher texts. This difference in cryptographic treatment may be caused by the use of an entirely different general system, or by the use of a different specific key, or merely by the choice of equivalents in a variant system. Messages which present different encrypted texts but which contain identical plain text are called isologs (from the Greek iso = "equal" and logos = "word"). One of the easily-noted indications of the possible presence of isologs is equality or near-equality in the lengths of two (or more) cryptograms. Isologs, no matter how the cryptographic treatment varies, are among the most powerful media available to the cryptanalyst for the successful solution of a difficult cryptosystem--and, in some cases, may provide the

only possible entries into a complex cryptosystem. An inkling of the help afforded by isologs was revealed by the example contained in subpar. 61d above; however, a much more striking illustration is given in the next few subparagraphs.

b. The following two cryptograms, suspected to be isologs, are available for study:

Message "A"

82265	63103	74839	69842	32529	70115
80277	89106	94000	13828	54082	40065
63629	33918	43158	81048	26458	45039
81713	52538	73309	20749	61752	16476
38728	91147	99926	41468	13365	33881
89697	93816	51750	57074	11804	43255
28120	27730	31199	79962	27865	60653
90870	40867	46594	19855	10822	22987
46729	36245				

Message "B"

30150	87497	14511	97360	49676	50106
45647	99181	69672	53889	41563	25203
90628	77536	20351	10570	89277	75011
35199	90138	99974	50232	04115	89216
38463	17547	14648	00646	85864	53898
26121	83878	94889	33728	11272	20504
06484	32103	98715	42662	80760	89880
44105	52900	59728	22855	87300	70893
59682	46253				

On the possibility that some dinome system (or systems) is involved, the messages are written under each other in dinomes to facilitate the examination of the similarities and differences of such a grouping of the cipher texts, as shown on the next page:

					5					10					15
A	82	26	56	31	03	74	83	96	98	42	32	52	97	01	15
A'	30	15	08	74	97	14	51	19	73	60	49	67	65	01	06
B	80	27	78	91	06	94	00	01	38	28	54	08	24	00	65
B'	45	64	79	91	81	69	67	25	38	89	41	56	32	52	03
C	63	62	93	39	18	43	15	88	10	48	26	45	84	50	39
C'	90	62	87	75	36	20	35	11	05	70	89	27	77	50	11
D	81	71	35	25	38	73	30	92	07	49	61	75	21	64	76
D'	35	19	99	01	38	99	97	45	02	32	04	11	58	92	16
E	38	72	89	11	47	99	92	64	14	68	13	36	53	38	81
E'	38	46	31	75	47	14	64	80	06	46	85	86	45	38	98
F	89	69	79	38	16	51	75	05	70	74	11	80	44	32	55
F'	26	12	18	38	78	94	88	93	37	28	11	27	22	05	04
G	28	12	02	77	30	31	19	97	99	62	27	86	56	06	53
G'	06	48	43	21	03	98	71	54	26	62	80	76	08	98	80
H	90	87	04	08	67	46	59	41	98	55	10	82	22	29	87
H'	44	10	55	29	00	59	72	82	28	55	87	30	07	08	93
J	46	72	93	62	45										
J'	59	68	24	62	53										

The dinome distributions for the two messages are as follows:

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

Distribution for Message "A"

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

Distribution for Message "B"

c. Since a general absence of marked crests and troughs is noted in both distributions, if the division of these cryptograms into dinomes is correct, and if they are both monoalphabetic, it is quite probable that some type of variant system (or systems) has been used. With this in mind, the encrypted texts and their distributions are scrutinized further for some indication of the kind of relationship which exists between the methods of encipherment of the two messages. The distributions are seen to be strikingly similar, not only with respect to the location of the one predominant peak in each, but also in the close correlation of the locations of the blanks in each.[14] Furthermore, upon examination of the superimposed messages themselves, it is observed that there are several instances wherein a value in message "A" coincides with the same value in message "B" (e.g., see positions A/A' 14, B/B' 9). This observation, taken in conjunction with the marked similarity of the distributions, strongly indicates that not only has the same general cryptosystem been used for the encryption of both messages, but that the same enciphering matrix has been used for both. Also, in the case of the values 38 and 62, it is noted that wherever either occurs in one message the same value

---

[14] For the benefit of the student with a mathematical background, it might be interesting to point out certain applications of cryptomathematics in connection with these two distributions. First of all, each of the two distributions is much flatter than that which would be expected for a sample of 125 dinomes of random text; i.e., a drawing (with replacement) and recording from an urn containing equal numbers of counters in each of 100 categories labeled 00-99 consecutively. In other words, whereas "random" follows a characteristic distributional appearance, approximated by the normal or binomial distributions, the samples at hand exhibit phenomena even flatter (or "worse") than that expected for random, approaching the theoretical (and fantastically non-random) "equilibrium" of exactly the same number of tallies in each cell of a distribution. The following table gives the observed number of x-fold repetitions in the two distributions, together with the expected number of x-fold repetitions in a sample of like size of random text, which expected number has been computed from tables of the Poisson exponential distribution (see Military Cryptanalysis, Part III):

x	Observed Msg. "A"	Observed Msg. "B"	Expected
Ø	14	17	29
1	51	52	36
2	33	23	22
3	1	6	9
4	-	1	3
5	1	1	1

It is to be noted that in the distribution for Message "A" the observed number of blanks (14) against the expected number of blanks from random text (29) represents a sigmage or standard deviation of 2.78 σ, which occurs in the other message, a phenomenon explainable on the assumption that the plaintext equivalents of these values are of such low frequency that no variant values have been provided for these plaintext letters in the cryptosystem.

d. With the foregoing details determined, it is now realized that it should be possible to form, between the two messages, "chains" of those cipher values which represent identical plaintext letters, as exemplified below. Beginning with the first value in each message, 82 and 30, a partial chain of equivalent variants is started; now locating some other occurrence of either value elsewhere (e.g., 82 at position H'8), and noting the cipher value coinciding with it (in this case, 41), the partial chain may be extended (including now 82, 30, and 41). After this particular chain is extended to include as many values as possible, another chain is formed by starting with any value which has not already been included in the preceding chain, this procedure being repeated until

---

can be translated as odds of 368 to 1 against its occurrence by pure chance. Likewise the other entries besides ∅ (in particular, the x-values of 1 and 2, and the cumulative values of 3-and-better) may be evaluated in terms of sigmages, and the conclusion would be reached that the two distributions have a most remote chance of being as flat as they are through mere chance; for instance, it is 3.05 σ or 877 to 1 against distribution "A", having only two tallies occurring three or more times when 13 such tallies are expected by random--and this sigmage when taken into consideration with that of the number of blanks yields a sigmage of 4 σ or approximately 31,000 to 1 of occurring through sheer chance. The sum total of all the deviations could be collectively evaluated, but this would involve the laborious computation of a multinomial distribution. Since the distributions of the two messages are much worse than would even be expected for random chance, the conclusion is drawn that the dinome grouping is highly significant and therefore must be correct, and furthermore that the cryptosystem involves variants in sufficient numbers for the plaintext letters to permit the encipherer to select the cipher equivalents with a view to suppressing as much of the phenomena of repetition as possible. Secondly, the χ test of the two distributions gives a χ value of 206, as against the χ value of 156 for random samples of this size; this represents a sigmage of 4.02 σ, or a ratio of 33,000 to 1 against its happening by pure chance; i.e., if the cryptograms were not in the same general system and specific keys. Therefore it is a foregone conclusion statistically that not only do the cryptosystems involve dinomes as the ciphertext grouping, but that the identical cryptosystem is involved in the two messages; and that because of the close correlation of the patterns of the two distributions, there is a good probability that the cryptograms contain identical plain text and therefore are isologs. This specific illustration of the potentialities of cryptomathematics indicates the important role that this branch of science may play in the art of cryptanalysis.

all possible chains are completed. It is found that the following chains, arbitrarily arranged here according to length, may be derived from the two messages:

(06 14 15 26 28 31 35 73 74 81 89 98 99)
(02 07 20 22 43 44 63 90)
(12 37 48 51 69 70 83 94)
(03 30 41 54 65 82 97)
(05 10 24 32 49 87 93)
(16 18 36 76 78 79 86)
(27 45 53 64 80 92)
(11 39 75 88)
(21 58 77 84)
(46 59 68 72)
(00 52 67)
(04 55 61)
(08 29 56)
(19 71 96)
(01 25)
(13 85)
(42 60)

Single dinomes:

(38) (47) (50) (62) (91)

If we now make an arbitrary assignment of a different letter to represent each chain (and each single dinome) and convert either of the messages to uniliteral terms by means of these arbitrarily-assigned values, we note the pattern of the opening stereotype "REFERENCE YOUR MESSAGE.....", and quickly recover the plain text.

e. The plaintext values when inserted into a 10x10 matrix having arbitrarily-arranged coordinates yield the following:

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

Manipulating the rows and columns with a view to uncovering some symmetry or systematic phenomena, the latent diagonal pattern of the equivalents for certain of the letters (such as E_p, N_p, O_p, R_p, and S_p) is revealed, and the rows and columns of the reconstruction diagram are permuted to yield the following original enciphering matrix:

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

There are no observable relationships in or between the sequences of digits in the row and column coordinates; therefore for want of any visible phenomena or further information on the derivation (if any) of these digits, it is assumed that they must have been assigned at random. The student will note that the final matrix is identical to that of Figure 39 in paragraph 59.

f. It should be emphasized that in the example of the preceding subparagraphs it was only possible to form chains of values from both messages <u>reciprocally</u> because the same enciphering matrix had been used for both. A non-reciprocal chaining procedure would have been required if only the general system had been the same for both but the enciphering matrices had differed in some respect, or if two completely different variant systems had been used (e.g., one using a frequential matrix and the other involving a less complex type of variant matrix, such as Fig. 29). Specifically, it would have been necessary to maintain two separate groups of chains, one group for each message; otherwise heterogeneous values would have become intermingled.

g. Although an analysis of but one isolated example by means of isologs was presented, the student should be able to appreciate the significance and potentially enormous value of isologs to a cryptanalyst. This value goes far beyond the simple variant encryption in a monoalphabetic substitution system; isologs produced by the use of two different code books, or two different enciphered code versions of the same underlying plain text, or two encryptions of identical plain text by two different "settings" of a cipher machine, may all prove of inestimable value in the attack on a difficult cryptosystem.

63. Further remarks on variant systems.--a. A few words should be added with regard to certain subterfuges which are sometimes encountered in monoalphabetic substitution with variants, and which, if not recognized in time, cause considerable delays. The considerations treated before in subpars. 52i and j on the disguise of the length of the basic multiliteral group apply equally here to multiliteral substitution with variants; thus, in dinome systems, a sum-checking digit or a null might be added in specified positions of the group to form a trinome. In complex variant systems, the presence of a null as one of the digits of a trinome would add greatly to the complexities of cryptanalysis of that system. The most important of the subterfuges have to deal with the use of nulls which are of a different size than the real cryptographic units, inserted occasionally to prevent the cryptanalyst from breaking up the text into its proper units. The student should take careful note of the last phrase; the mere insertion of symbols having the same characteristics as the symbols of the cryptographic text, except that they have no meaning, is not what is meant. This class of nulls rarely achieves the purpose intended. What is really meant can best be explained by an example. Suppose that a 5x5 variant matrix with the row and column indicators shown in Fig. 46 is adopted for encipherment. Normally, the cipher units would consist of 2-letter combinations of the indicators, invariably giving the row indicator first (by agreement).

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

Figure 46

The phrase COMMANDER OF SPECIAL TROOPS might be enciphered thus:

```
 C   O   M   M   A   N   D   E   R   O   F   . . .
VI  EB  PH  IU  FT  IE  AB  TM  WO  PW  GT   . . .
```

These would normally then be arranged in 5-letter groups, thus:

```
VIEBP HIUFT IEABT MWOPW GT . . .
```

b. It will be noted, however, that only 20 of the 26 letters of the alphabet have been employed as row and column indicators, leaving J, K, Q, X, Y, and Z unused. Now, suppose these six letters are used as nulls, <u>not in pairs, but as individual letters inserted at random</u> just before the real text is arranged in 5-letter groups. Occasionally, a <u>pair</u> of letters might be inserted, in order to mask the characteristics of "avoidance" of these letters for each other. Thus, for example:

V I E <u>X</u> B   P H <u>K</u> I U   F <u>J</u> <u>X</u> T I   E A <u>J</u> B T   M W O <u>Q</u> P   W G <u>K</u> T <u>Y</u>

The cryptanalyst, after some study suspecting a biliteral cipher, proceeds to break up the text into pairs:

VI EX BP HK IU FJ XT IE AJ BT MW OQ PW GK TY

Compare this set of 2-letter combinations with the correct set. Only 4 of the 15 pairs are "proper" units. It is easy to see that without a knowledge of the <u>existence</u> of the nulls--and even with a knowledge, if he does not know <u>which</u> letters are nulls--the cryptanalyst would be confronted with a problem for the solution of which a fairly large amount of text might be necessary. The careful employment of the variants also very materially adds to the security of the method because repetitions can be rather effectively suppressed.

<u>c</u>. Similarly in the examples under paragraph 58, the letter J in Figs. 27 and 29 may be used as a null; the letter Y in Fig. 28; and the digit Ø in Figs. 33 and 34. In Fig. 30, any letters in the range of P - Z might be used as nulls, but this usage might be weak because of the extremely low frequency of these letters as compared with the letters A - O; this is an important point to consider in the examination of encrypted text for possible poor usages of nulls.

<u>d</u>. From the cryptographic standpoint, usage of nulls in the manner outlined above results in cryptographic text even more than twice as long as the plain text, thus constituting a serious disadvantage. From the cryptanalytic standpoint, the meaning of the cipher units in the system described in subpar. b above constitutes the most important obstacle to solution; this, coupled with the use of variants, makes this system considerably more difficult to solve, despite its monoalphabeticity.

(BLANK)

~~RESTRICTED~~ ~~Security Information~~

NATIONAL SECURITY AGENCY

COURSE

IN

MILITARY CRYPTANALYSIS, PART I

NOTICE: This document contains information affecting the national defense of the United States within the meaning of the Espionage Laws, Title 18, U.S.C., Sections 793, 794 and Title 50, U.S.C., Sections 46, 46a and 46b. Its transmission or the revelation of its contents in any manner to an unauthorized person is prohibited by law.

National Security Agency
Washington 25, D. C.

December 1952

~~RESTRICTED~~

(BLANK)

COURSE IN MILITARY CRYPTANALYSIS, PART I

Monoalphabetic Substitution Systems

Introduction

This is the first of a series of six basic courses in the art of military cryptanalysis. The purpose of this course is to impart to the student the methods and techniques which form the basis for the cryptanalysis of the simple types of military cipher systems. An understanding of these principles is necessary to grasp the more advanced cryptanalytic techniques employed in the attack on the complex cryptosystems which constitute present-day military cryptography.

The scope of this course is: fundamental principles; uniliteral substitution; multiliteral substitution; polygraphic substitution; and miscellaneous monoalphabetic substitution systems. It consists of ten lessons and an examination as follows:

Lesson 1, Fundamental principles

Lesson 2, Uniliteral substitution with standard and mixed cipher alphabets

Lesson 3, Multiliteral substitution: miscellaneous matrices; Baconian and Trithemius systems; elementary Baudot systems

Lesson 4, Multiliteral substitution with variants

Lesson 5, Polygraphic substitution: small matrices

Lesson 6, Polygraphic substitution: quadricular tables

Lesson 7, Polygraphic substitution: miscellaneous systems

Lesson 8, Miscellaneous monoalphabetic substitution systems; concealment systems

Lesson 9, Monoalphabetic substitution with irregular-length cipher units: monome-dinome systems; miscellaneous systems

Lesson 10, Syllabary squares and code charts

Examination

The text reference for this course is the National Security Agency publication, "Military Cryptanalysis, Part I" (December 1952).

This course has been designed as a self-study or extension-type course; therefore, there is no limit placed on the number of hours that may be spent in the completion of the course, any lesson, or the examination. However, for statistical purposes it is requested that the student indicate the number of hours spent in the completion of each lesson and the examination.

The cryptograms in this course have for the most part been arranged in proper worksheet form, obviating the necessity of recopying; and frequency distributions have been given to reduce the amount of time spent on the purely clerical labor incidental to the solution. The underlying texts of the cryptograms comprise hypothetical ground, naval, air, and general administrative messages. Where necessary for solution, the specific nature of the text of any particular cryptogram is indicated. Otherwise, the text of a message may be assumed to be general administrative or ground text.

The only materials required are cross-section paper of $\frac{1}{4}$-inch squares, and a set of printed and blank alphabet strips. An eraser is of the utmost importance.

Special Instructions

So far as is practicable, detailed work sheets which usually form a part of the solution should be submitted with the solutions. In all the lessons of this course, it is required that the student recover all cipher alphabets, cipher tables, and specific keys used. He will also be required to state the method of operation of each cryptosystem and give the key words upon which each component is based.

~~RESTRICTED~~ ~~Security Information~~

NATIONAL SECURITY AGENCY
Washington 25, D. C.

COURSE — Military Cryptanalysis, Part I

LESSON 2 — Uniliteral substitution with standard and mixed cipher alphabets

TEXT ASSIGNMENT — Sections V and VI

1. a. What is the first step one should take in attempting to solve an unknown cryptogram that is obviously a substitution cipher?

b. If this step is unsuccessful and the cryptogram is obviously monoalphabetic in character, what type of cipher alphabet may be assumed to have been used?

2. a. Name two methods of solving monoalphabetic substitution ciphers involving standard cipher alphabets.

b. In the solution of a substitution cipher by completing the plain component sequence involving reversed standard alphabets, what are the successive steps?

c. Why do monoalphabetic cryptograms involving standard cipher alphabets yield such a low degree of cryptosecurity?

3. What are four characteristics of vowels which permit their classification as such in monoalphabetic substitution ciphers involving mixed cipher alphabets?

4. a. What two places in every message lend themselves more readily to successful attack by the assumption of words than do any other places? Explain.

b. What is meant by the "probable word method" of solution?

5. a. What is meant by the word pattern "A B C B A D B"?

b. For each pattern given below, indicate one good English word that contains the pattern:

(1) A B C B A D B

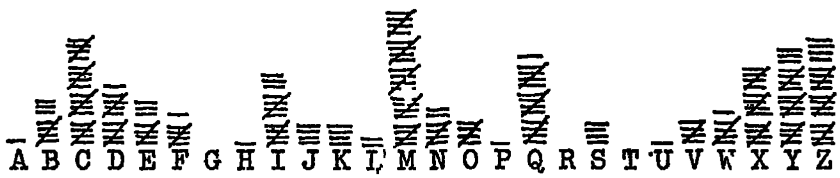
(2) A A B A

(3) A B C D A

~~RESTRICTED~~

6. Solve the following cryptogram and indicate the specific key ($A_p = \theta_c$):

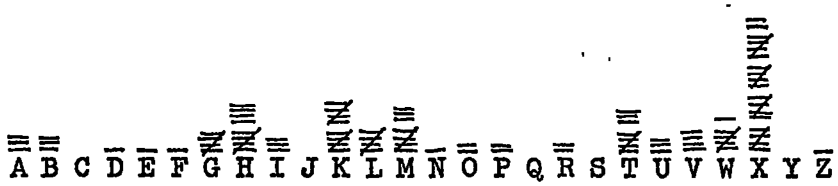
```
JMQVS  QZXIF  FMZSL  IZMLZ  CEMEB
FQOME  MDXYQ  OZCYY  XJMZI  VMZIY
OQWYI  DKYMV  MZMNQ  EQKMX  CCWZB
CYIXI  CDYYX  CBZQI  FZCQN  HWDOX
ICDJQ  YPMMD  YMVMZ  MFSNQ  EQKMN
QDNEW  OJMAW  IBEMD  XNMYX  ZCSMN
YXCBU  MQZME  CVIDK  CWZXZ  CCBYX
CZMQZ  BCYIX  ICDYY  XCBZQ  FYXCD
```



$\phi_p = 2655$ $\phi_r = 1531$ $\phi_o = 2636$

7. Solve the following cryptogram, and indicate the specific key:

```
WXLMK  HRXKL  ATOXU  XXGHK  WXKXW
MHIKH  VXXWT  MHGVX  MHTKX  TPAXK
XLNUF  TKBGX  TVMBO  BMRAT  LUXXG
KXIHK  MXWLM  HITVD  GHPEX  WZXXX
```



$\phi_p = 660$ $\phi_r = 381$ $\phi_o = 848$

~~RESTRICTED~~

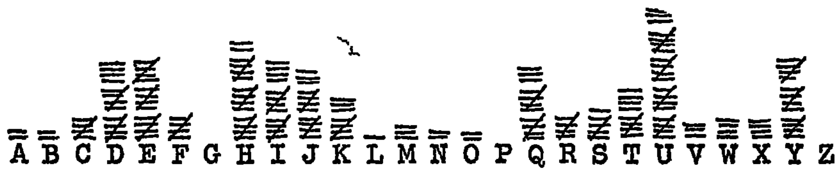
8. Solve the following cryptograms, and indicate the specific keys:

a. Q H H Y L Y D W Q J J M E F C

b. Y X S E D Y F S X U H W X U S

9. The following badly garbled cryptogram was intercepted. Reconstruct the original plaintext message, resolving the errors and omissions, and indicate the specific key:

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



$\phi_p = 2270$ $\phi_r = 1311$ $\phi_o = 2136$

~~RESTRICTED~~

10. a. Construct a triliteral frequency distribution showing one prefix and one suffix of the letters of the cryptogram below. On the work sheet below, indicate by underscoring in black all repetitions of three or more letters. Other significant details may be marked in different colors.

b. Prepare a condensed table of repetitions of digraphs and trigraphs appearing more than twice, and include all repetitions of longer polygraphs.

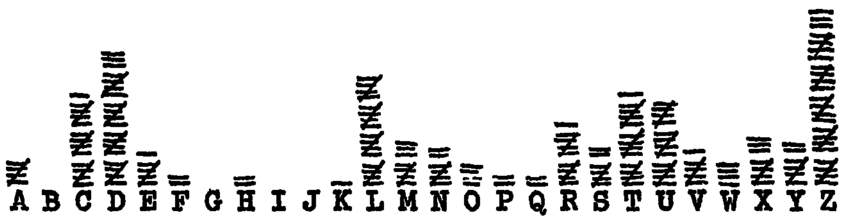
c. Using the data obtained in a and b above, complete the solution of the cryptogram, and recover all keys.

	5	10	15	20	25
A	U B S Y B	V X R P N	C G U M Z	X G P N P	C U B Q P
B	U X X F Z	X B N B M	I G V R P	N V X U Y	R X G N D
C	F B Z H I	Z U X G L	L B U I B	M Q L Z R	B M B N X
D	V G N O P	P A B A Z	U B Z P N	B C G H B	M G L B V
E	N P U X F	B Z V X P	C D U B B	N H G L L	B V X P Q
F	Q F P X P	D U Z Q F	G R U B R	P N N Z G	V V Z N R
G	B M G V V	G P N V N	B D Z X G	H B E B R	Z Y V B P
H	C Z A H B	U V B O B	Z X F B U	R P N A G	X G P N V

~~RESTRICTED~~

11. Solve the cryptogram below, suspected to contain the probable word "BLOCKADE"; recover all keys.

	5	10	15	20	25
A	L C T C E	L U Z O D	U C R E A	W Z U S N	F Z X D Y
B	D R T L D	S D R Z S	D E U C M	U Z Z K Z	U D C D V
C	T Q T X D	A O Y Z C	Z W Y D X	P T V Z D	S C M Z Z
D	R Z A Q L	L D E C M	Z U R X D	T L C M T	L W Z Z R
E	Z S S Z X	C Z V L C	D O U D X	P Z C W T	U U T H Z
F	S U D A D	E U F Z L	L Z Y L X	D R C N R	E Z L C D
G	M T U T L	L M D L C	N Y Z L M	D U Z O D	L N C N D
H	R L T R V	M T L V T	A T H Z V	U T N Y Y	N R Z L X



$\phi_p = 2655$ $\phi_r = 1532$ $\phi_o = 2770$

12. Solve the following cryptogram, and recover all keys:

	5	10	15	20	25
A	J Z D F V	W H E D Z	V H W D S	Y K T W D	O E D Z D
B	E D S E C	C W H H W	E D Z T E	X X W S Z	V N Z V Z
C	S P F J K	V Z T Y P	H J D W O	L J W D P	V P W T I
D	R E D Z E	X E K V F	P J V E Y	H H J E F	E D Z F V →
E	← W H E D Z	V H J P J	Z H J L P	J X E K V	J L T W M
F	W H W E D	W H W D M	W S W D W	J R E X I	Y K Z C E
G	K D J P W	D C E M W	D O N Z H	J J E P J	J P S B E →
H	← K V F E H	W J W E D	H N Z H J	E X X P W	V J E N D
J	H J E F S	E D X W V	C P J W E	D V Z G K	Z H J Z T

$\phi_p=3362$ $\phi_r=1940$ $\phi_o=3560$

13. Using the sequences recovered in Problem 12, solve the following cryptograms and indicate the specific keys:

a. U R J J R X Q U Q X K S A R B B E T O I

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

$\phi_p=25$ $\phi_r=15$ $\phi_o=16$

b. F D L D Y X Z U M U E U F P N D V O F E A L Y R W

U M L J X A F D Y E X E K Q P D O Y C V R E U A X

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

$\phi_p=163$ $\phi_r=94$ $\phi_o=118$

14. The following cryptograms, enciphered with random cipher alphabets, are in bona fide word lengths. Solve them.

a. H Y A R V J Z G H A R O T V K C G K M M G K H Z M L K U G

L K U G O R O E H O Z E M V H F S R M J R O T

J E H Z P U H G V E G M R O M C J K K S J K U M E

b. R G R Q R U T D S P Y U R D P Z F T A V D R C A Y C F O

J O D R Z Y U U F S P P F U Z R T F A D Y G P

c. C D G W D S A L C A U M M D C R B U C D Y V D V D J R

I Y S U A U Y V S L Z C Y S S C U T D C

15. In solving several unrelated monoalphabetic cryptograms, the following cipher alphabets were reconstructed. Recover all key words in each case. To facilitate solution, significant segments have been underlined.

a.

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

b.

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

c.

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

d.

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

e.

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

f.

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

g.

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

h.

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

NATIONAL SECURITY AGENCY
Washington 25, D. C.

COURSE — Military Cryptanalysis, Part I

LESSON 1 — Fundamental principles

TEXT ASSIGNMENT — Sections I-IV, inclusive.

1. a. What four things were thought by Captain Hitt to be essential to cryptanalytic success?

b. What six additional elements are also highly desirable?

2. a. Define the terms "cryptology", "cryptography", and "cryptanalysis."

b. What are the essential differences between substitution and transposition?

c. Differentiate between a code and a cipher system.

d. Explain the difference between the terms "general system" and "specific key".

e. Distinguish between monoalphabetic and polyalphabetic substitution.

3. What four fundamental operations are involved in the solution of practically every cryptogram?

4. In the solution of cryptograms involving a form of substitution, to what simple terms is it necessary to reduce them in order to reach a solution?

5. Is it always necessary to determine the specific key in order to reconstruct the plain text? Explain.

6. Indicate the language in which you would expect the plain text of the encrypted portion of the following message to be written. Give reasons for your answer.

From: João Fialho, São Paulo, Brasil.
To: Gualterio Costa, New York City.

Com referência ao seu telegrama. NSM NRJPN INJ PMVCOCEN VNPSN PMBMPCEN QMT JBCVCJ IJUM DTGAJ LTMCPN KPJUCEMIVCNP PMHMQQN UMIVCHMISJQ SMPVMCPJ SPCHMQSPM.

~~RESTRICTED~~

7. a. The letter E represents what percentage (in round numbers) of the letters in English telegraphic text?

b. What are the four most frequent consonants in English telegraphic text?

c. What are the five letters of lowest frequency in English telegraphic text?

d. What are the four most frequent digraphs in English telegraphic text?

e. Account for the discrepancies between frequencies of letters in English literary text and English telegraphic text.

8. What three facts can be determined from a study of the uniliteral frequency distribution?

9. In the following extract from a speech given during World War II, each dash indicates the omission of a letter. Complete the text by writing the necessary letters over each dash to form appropriate words.

"Washington's Birthday is a most a p _ _ _ _ _ _ _ _ _ occasion for us to talk with each _ _ _ _ _ about things as they are _ _ _ _ _ and things as we _ _ _ _ they shall be in the _ _ _ _ _ _.

"For _ _ _ _ t years, General Washington and his _ _ _ _ _ _ _ _- _ _ _ Army were faced c o _ _ _ _ _ _ _ _ _ with formidable _ _ _ _ and recurring _ _ _ _ _ _ _. _ _ _ _ _ _ _ _ and equipment were lacking. In a _ _ _ _ _, every winter was a Valley Forge. Throughout the _ _ _ _ _ _ _ _ states there existed selfish men, jealous men, _ _ _ _ _ u l men, who _ _ _ _ _ _ _ _ _ _ that Washington's _ _ _ _ _ was hopeless, that he should ask for a n _ _ _ _ _ _ _ _ _ peace.

"Washington's _ _ _ _ _ _ _ in those hard _ _ _ _ _ has provided the _ _ _ _ _ for all Americans ever since--a model of moral _ _ _- _ _ _ a. He held to his _ _ _ _ _ _, as it had been charted in the Declaration of Independence. He and the _ _ _ _ _ men who _ _ _ _ _ _ with him knew that no man's life or _ _ _ _ _ _ _ _ was secure, without freedom and free i _ _ _ _ _ _ _ _ _ n s.

"The present _ _ _ _ _ struggle has _ _ _ _ _ _ us increasingly that _ _ _ _ _ o m of person and _ _ _ _ _ _ _ y of property anywhere in the _ _ _ _ _ depend upon the security of the rights and obligations of liberty and _ _ _ _ _ _ _ everywhere in the world.

"This war is a new _ _ _ _ of war. It is _ _ _ _ _ _ _ _ _ from all other wars of the _ _ _ _, not only in its methods and

~~RESTRICTED~~

_ _ _ _ _ _ _ but also in its geography. It is warfare in terms of every c o n _ _ _ _ _ _, every _ _ _ _ n d, every sea, and every a _ _ - _ _ n e in the world. The _ _ _ _ _ oceans which have been h e r _ _ _ _ _ _ in the past as our _ _ _ _ _ _ _ _ _ _ _ from attack have become _ _ _ _ _ _ s s battlefields on which we are _ _ _ - _ _ _ _ _ _ _ being challenged by our enemies."

10. a. In the following examples the words of sentences have been transposed. Rearrange the words to make plain text.

(1) AT NOTHING REPORT THIS TIME TO

(2) ARTILLERY SECTOR BARRAGE NORTHWEST HEAVY IN

b. In the following examples the letters of several words of each sentence have been transposed. Rearrange the letters to make good words that will give intelligible plain text.

(1) Eight SESTYODRER have DTPADERE to join SAKT REOFC

(2) ABELNU to contact ATTAINBLO on my right AFKLN

c. In the following examples the words of each sentence have been transposed and, in the case of several words, the letters have also been transposed. Reconstruct the plain text.

(1) OLANG RIDGE TANK GIMNOV EHOTISL EAST NOMLCU

(2) DOWN MEYEN OFANERTON SIX THIS OTHS SNEALP

d. In the following examples, the letters of each word of each sentence have been rearranged in the order in which they appear in the normal alphabet. Reconstruct the plain text.

(1) ADELY AACKTT CDDEEHLSU OT CCEEMMNO AT EGHIT HIST GIMNNOR

(2) ADEEIILMMTY NOPU CEEIPRT ADHIRTWW OT AADEEGNPRRR IINOOPST

e. In the following examples the plain text has been broken up into groups of five letters and then in each group of five the letters have been rearranged in the order in which they appear in the normal alphabet. Reconstruct the plain text.

(1) ORSUU ABIMR AEHNS ENSUV ADKOR ADEGM EEINN EMNVY EELSS S

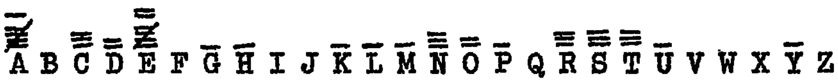
(2) AEIRR ACNNO AINSS ACEPT ELORR OPRST AILRT EELRY ACIMP EEMNT DESST DEORY

11. Using cross-section paper prepare a unilateral frequency bar distribution of the letters of the following paragraph:

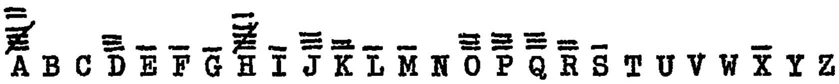
"The shortest and surest way to live with honor in the world is to be in reality what we would appear to be; all human virtues increase and strengthen themselves by the practice and experience of them."

12. Determine the class to which the cipher systems, which were used in enciphering the following messages, belong:

a. ORANA THPNO SKTCD MEEES CERAE
RNUSA ETLGD AYECA



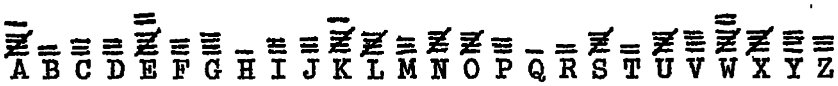
b. DHJJK QOAHR XKSOF HPQGA PPHLA
DIADE HJROA MAHQA



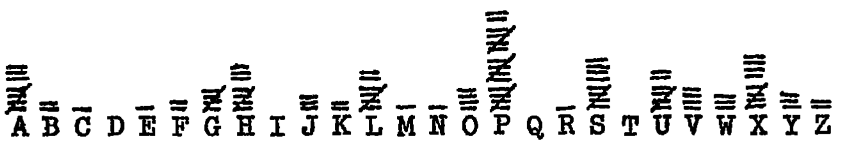
c. ROLEH KBWFZ CQCPZ NVJWZ MIVEQ
EPCIN OJSJU YMWQB

13. Which of the following substitution ciphers are monoalphabetic?

a. UJKLW EUVKL FSPAQ PHTKR DZNGL
SELYN XYXBX JDATU WEUZG WFVXM
MNZAY AOSGU DCLGI OEWJE IFOKM
KNWAP KOIEV AROEV WSCWN SBCYX



b. HUPYP XXAEP AFGZP VGLHA SLXHU
SXXAY PWKAS LHPRH ALOBA XPLVS
WUPJP OBSHU HUPGF XGKPH PVSWU
PJOPZ SVPYS MPOAX ULSLP CGNJX



c. GXYVL ZXMXS LOZGR WEJLX PWTKZ
GMXLW QIVZW QBRXK KTDVL MXAEX
VHMXA LOTLY TKDWX GBQKQ LWZXG
RTYYZ KTOXG AWXLQ LOZGR XVWGQ

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

14. The following messages were enciphered monoalphabetically. Determine in each case whether the cipher alphabet used was a standard or mixed alphabet and if standard, whether direct or reversed.

a. ANVOR LOUNQ RLEZW ZHNEZ WZBOR
ZKYLF AOZSO ONORF PJZPP LDZDN
LRZLB LABWZ HNAPO WQHOO RZIZU

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. ESPAP LVDLY OECZF RSDTY ESTDO
TDECT MFETZ YBFTN VWJTO PYETQ
JTELD OTCPN EDELY OLCON TASPC

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

c. PYHYL XOLWY JJVYX OILYR YQYPJ
KNYLK YHYLC PAYAC LYXIR QYJVO
ZKOXC PCREK UKUPJ IUJUO PRIAS

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

15. Derive the ϕ_p, ϕ_r, ϕ_o, Λ_p, Λ_r, and Λ_o for each of the following distributions, and evaluate the /monoalphabetic/ goodness of ϕ_o and Λ_o of each in terms of "good", "fair", or "poor", entering these data in the attached diagram. On the basis of the foregoing, decide which distributions are most probably monoalphabetic and which are most probably non-monoalphabetic, indicating your decision by a check (✓) in the diagram; in the case of those not clearly belonging in either of these categories, check "decision suspended".

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

c. A 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

d. 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

e. 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

f. 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

g. 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

h. 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

								Goodness of ϕ_o			Goodness of Λ_o			Decision		
	N	ϕ_p	ϕ_r	ϕ_o	Λ_p	Λ_r	Λ_o	G	F	P	G	F	P	mono.	non-mono.	susp.
a.																
b.																
c.																
d.																
e.																
f.																
g.																
h.																

~~RESTRICTED~~

16. From the intercepted traffic of three intercept stations operating in the same sector of the front, the following code messages were selected for study by a member of the cryptanalytic section at GHQ. They are undoubtedly three versions of one enemy message, but there appears to be a number of differences, due no doubt to operating difficulties at the several stations. Study the messages and reconstruct from them the actual code text sent by the enemy station.

I. Time intercepted 1612 by HS   W F F   V   L D C

GR 35 $\overline{BT}$

NR 17 DYBIE DUFTO AMEJA KIBON

SGCOY FOBAK DODLA LUFYD KAWAL

APAYN CODAP KEDUR JOPID JENOX

MEHAZ LOGIS KUTEG EVAUK IPBEM

KEHZA HOBWE AVDUZ FOFA_ EMCOZ

EGBLO DOFYO ENC__ MAWEN _____

_____ _____ _____ _____ _____

II. Time intercepted 1610 by MR   M F F   V   L D C

GR 35 $\overline{BT}$

NR I_ DYBIE BUFTO AMEJA KIBON

IPKO_ F_BAK DODLA LUFYL KAWAL

APAYN _____ __DUA __PID JENOX

NEHAZ LOGIS KUTEG EVAUC IRBW

KEHZA SOBWE VADUZ FOFET EMCOZ

EGBLO DOFYO AECDA MAWEN ___OM

EMCOZ ACFAH LOFIR 0935

III. Time intercepted 1612 by YG   W F F   V   L D K

GR __ $\overline{BT}$

NR 17 DYBIE DUFTO AMEJA KSBON

IPCOY ___A_ DO___ LUFYL KAWAL

APETYN CODAP KEDUR WOPID JENOX

MEHAZ LOGHKUTEG EVAUK IPBEM

KEHZA HOBWE AVDUZ FOFET EMCOZ

EGBLO DOFYO ENCOA MAWEN MAWEN

EXFOM EMCOZ ACFAH LOFIR 0935

~~RESTRICTED~~

(BLANK)

~~RESTRICTED~~

~~Security Information~~

NATIONAL SECURITY AGENCY
Washington 25, D. C.

COURSE — Military Cryptanalysis, Part I

LESSON 3 — Multiliteral substitution with single-equivalent cipher alphabets

TEXT ASSIGNMENT — Section VII

1. Solve the following cryptogram, and recover all keys:

					5					10					15
A	DT	LR	WE	OE	OE	WH	RR	WR	LA	WH	WA	DE	DA	WR	LE
B	LE	OR	RE	WT	OR	WA	OH	WH	OR	LE	LR	WA	RR	RR	WH
C	WA	WH	OE	OR	LE	LE	WR	WA	WH	OH	LR	LE	LR	WA	OH
D	OE	LR	OA	OA	OE	LR	OR	RE	OA	OA	WH	WT	WH	WA	WA
E	WR	WA	WH	DE	RT	OE	WH	WH	RE	OR	OA	RT	OE	LR	OR
F	RE	WR	WE	WA	OH	DE	WR	LR	WA	WA	WR	WA	WH	DE	DA
G	LR	LR	WA	WH	OA	DE	LR	LT	LT	LR	OA	WR	DE	WR	LR
H	WA	OA	LR	RA	RA	LR	WE	OE	DE	RT	OE	WH	RR	WR	LA
J	WH	WA	DE	DA	WR	LE	LE	OT	WH	OE	WH	WH	WA	RA	LR
K	OE	OH	WH	RE	OT	DT	OR	RE	RE	WR	DE	WR	LR	WA	OR
L	LE	OR	OE	DE	WR	LE	LE	WH	OE	DT	OA	WE	LT	LT	LR
M	OE	DE	OA	DE	LR	LT	OH	LR	LE	LR	WA	WH	LE	OT	WH
N	WA	WA	WR	WA	RR										

(For distribution, see next page)

~~RESTRICTED~~

	A	E	H	R	T
D	3	12	–	–	3
L	2	13	–	21	5
O	10	14	6	10	3
R	3	7	–	5	3
W	22	4	22	13	2

ϕ_p=2270 ϕ_r=1362 ϕ_o=2288

(25-element alphabet)

2. This message was sent by the Fifteenth Infantry. Solve it and recover all keys:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	CY	AO	NX	CN	NO	CN	AO	AO	OG	ON	NG	BY	OX	OX	RO
B	CG	NY	RO	AN	RE	AG	RO	OX	AO	AN	AX	AX	AG	AN	AG
C	CN	RO	OX	OX	BY	AN	AG	CN	BE	CX	BN	BX	CG	RO	ON
D	CO	RE	CN	AY	BG	CE	ON	NO	AO	OG	RO	NO	NO	RO	RE
E	OO	NG	BY	OX	OX	RY	AG	AX	BY	AN	OG	CN	AO	OY	OG
F	NO	OX	CY	NX	OG	AO	AN	CN	AG	RE	AG	BY	OG	NO	AO
G	BO	AO	CN	CG	AG	CN	ON	BO	CN	AO	OY	CO	OE	ON	NO
H	AO	OG	RO	NO	NG	RO	NO	AG	CN	RE	AO	OX	RX	AE	BY
J	AN	BO													

	E	G	N	O	X	Y
A	1	9	7	12	3	1
B	1	1	1	3	1	6
C	1	3	11	2	1	2
N	–	3	–	9	2	1
O	1	7	5	1	9	2
R	5	–	–	9	1	1

ϕ_p=960 (approx.) ϕ_r=410 ϕ_o=716

(36-element alphabet)

~~RESTRICTED~~

3. Solve the following cryptogram, and recover all keys:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	RG	GP	EE	GR	RG	GP	ES	GR	RG	PP	GE	PR	GE	RG	GS
B	AS	GR	RR	GS	AE	PP	GP	GA	PP	RA	EA	ES	GR	RG	PP
C	GE	RA	PR	GS	RE	GP	AR	GP	GS	PP	GP	RG	RA	EA	PP
D	PS	PG	AR	PE	GA	RR	RG	GP	RR	RE	PG	PP	RA	EA	RS
E	PG	PE	RG	AR	PE	GA	RR	RG	GP	RR	RP	AE	GS	GA	AP
F	GP	PP	RA	EP	ES	GP	RA	GP	RA	PE	PR	PR	AE	GR	GP
G	RA	GA	GP	GP	RR	GP	RR	GR	AS	AS	GP	RR	GR	GS	PP
H	GP	AE	GE	RS	PG	RG	GS	RE	PP	GR	GG	GS	PP	GR	PG
J	GA	PG	RS	RE	PG	AS	PR	GS	GA	GE	RR	EA	ES	GR	RG
K	RR	RP	GS	PP	PP	GS	AE	GR	PG	GA	EP	RG	GP	EE	GR
L	RA	GR	PP	GR	PG	GA	AR	GS	RA	RP	GP	GP	GA	GS	PE
M	ES	PG	RG	GR	ER	GP	RR	RP	GE	RG	GP	AG	GR	AS	GP
N	GA	PP	GS	AE	AR	PA	EP	RG	GP	PR	AE	GE	RG	GP	EE
P	GP	RA	PP	GP	RR										

	A	E	G	P	R	S
A	–	7	1	1	5	5
E	4	3	–	3	1	5
G	11	7	1	27	16	14
P	1	5	10	16	6	1
R	11	4	16	4	12	3

$\phi_p = 2260$ (approx.) $\phi_r = 1164$ $\phi_o = 2294$

(30-element alphabet)

~~RESTRICTED~~

4. Solve the following cryptogram, and recover all keys:

					5					10
A	AAC	AAB	BBA	AAB	AAC	AAB	ABB	ACC	AAB	CCA
B	ABA	ABC	AAC	CAA	AAB	BAA	BAA	AAA	BBB	AAB
C	ABB	ABC	CAA	BAB	AAB	AAC	BBA	ACB	CBA	AAB
D	BBA	BCC	ACB	BBB	BBC	ACA	BBA	ABA	ABC	AAC
E	ACA	BBC	AAC	AAB	AAB	BBC	AAA	BAA	BAB	AAB
F	AAB	ABB	ACC	AAA	ABB	ACC	AAB	BCC	BCC	AAB
G	BAC	CCC	ABB	AAB	CBC	ACA	ACA	AAC	ACB	CAB
H	AAA	ACA	CCB	AAB	AAC	ABA	BAA	ACB	CBC	CCB
J	AAB	AAC	ABA	CCB	AAB	AAC	ABA			

		A	A	A	B	B	B	C	C	C
	2:	A	A	A	B	B	B	C	C	C
	3:	A	B	C	A	B	C	A	B	C
1:	A	4	18	10	5	5	3	5	4	3
	B	4	2	1	4	2	3	-	-	3
	C	2	1	-	1	-	2	1	3	1

$\phi_p=499$ $\phi_r=277$ $\phi_o=542$

(27-element alphabet)

5. Solve the following naval message, and recover all keys:

11101	10333	12231	03023	33122	31000
06002	60610	15231	40424	24052	33206
03042	61122	33263	12334	11052	33011
00001	12200	20010	02600	06151	62611
13367	89310	62222	26050	41221	04101
30511	24230	52604	22221	21604	10151
10023	14122	30105	00113	50024	11111
33504	10131	42305	03042	60623	10360

RESTRICTED

6. Solve the following cryptogram, and recover all keys:

45264	56282	02523	29276	16145	23820
63216	52729	27212	60652	16729	47694
56529	02146	04161	25424	90692	12143
65026	45672	92325	61272	84543	04182
04221	67262	94523	41252	92945	23820
46272	34506	52921	63023	45646	74565
29082	21670	23456	12582	02947	27650
29210	23472	12543	65000		

7. Solve the following cryptogram, and recover all keys:

05105	23804	91161	38349	22702	74491
16138	33834	92274	27505	31612	74492
16127	14914	92274	38216	12724	91161
27138	10523	84274	05405	23801	61491
16105	22713	80271	05227	44910	51052
05327	14921	60491	05227	10502	74163
38016	11653	85492	27405	20531	61494
49238	42713	82492	27427	20522	71380
49127	02714	91270	49149	12702	72273
05505	30522	74272	16127	13814	93052
49449	24910	52380	05149	23834	91492
27449	23823	82384	38105	23844	91050

RESTRICTED

8. The following is a text in the Baudot teleprinter code enciphered by a simple machine employing five two-position switches which operate polarized relays. Each switch has the function of changing the polarity of its respective baud (a single "mark" or "space" impulse), if the switch is in the 'active' position. If the switch is in the 'inactive' position, the polarity of the baud is unaffected. The switch settings remain constant for each message. As an example, if switches 1 and 4 are active (x), and 2, 3 and 5 are inactive (o), then the word ENEMY is enciphered thus:

Key:	xooxo	xooxo	xooxo	xooxo	xooxo
Plain:	+----	--++-	+----	--+++	+-+-+
Cipher:	---+-	+-+--	---+-	+-+-+	--+++

Solve the message and recover the switch settings.

	1	2	3	4	5	6	7	8	9	10
A	+-+--	+-++-	+---+	++-++	++++-	+---+	-+++-	-+---	-----	-+-+-
B	+-+--	+----	+---+	--+++	++--+	++-+-	+-+++	+++-+	+++-+	--+++
C	++--+	+++++	--+-+	++-++	+-+--	+---+	-++-+	+++++	+-+--	++--+
D	+++--	+++--	--+-+	+++++	-----	+---+	---++	++--+	+++++	--+-+
E	+---+	+-++-	+--++	+---+	+--++	--+-+	-++--	+-+--	+-++-	+--++
F	+---+	-+++-	-+-++	++--+	-++--	++---	-++-+	-+--+	+-++-	-----
G	+++--	+--+-	-+-+-	+++-+	+++-+	+---+	+---+	-++-+	+-+--	-++--
H	-++-+	+-+--	--+-+	+++++	+---+	++-++	+-++-	+++--	+---+	-+++-
J	-+---	+-+--	-+-+-	+---+	++--+	+--++	---++			

	3:	+	+	+	+	-	-	-	-
	4:	+	+	-	-	+	+	-	-
	5:	+	-	+	-	+	-	+	-
1,2:	++	5	1	4	4	3	1	6	1
	+-	1	5	-	8	4	1	13	1
	-+	-	3	4	3	1	3	1	2
	--	2	-	5	-	2	-	-	3

ϕ_p=480 (approx.) ϕ_r=234 ϕ_o=386

(32-element alphabet)

~~RESTRICTED~~ ~~Security Information~~

NATIONAL SECURITY AGENCY
Washington 25, D. C.

COURSE — Military Cryptanalysis, Part I

LESSON 4 — Multiliteral substitution with variants

TEXT ASSIGNMENT — Section VIII

1. Solve the following cryptogram, and recover all keys:

					5					10					15
A	RA	DE	KE	PE	VE	TI	BO	LA	GO	DU	JO	BE	KI	BI	JO
B	BU	JA	VA	ME	LA	BE	KI	RE	FE	DO	VI	JO	SA	DO	JE
C	KI	BA	MO	SA	CU	GE	GE	PI	BO	KI	JU	CE	CI	MI	NE
D	PO	JU	CE	RE	NA	BU	BE	KO	RA	DE	KE	TE	SE	TI	JO
E	FA	GO	DU	DO	JE	KI	DI	JO	BU	JA	CE	BO	FO	BA	BU
F	DA	LE	JO	NI	DO	NA	BO	BE	PI	GI	ME	TE	CO	JO	TI
G	SA	BO	TI	DU	MO	FA	BU	NA	DU	DE	TO	GI	BE	SE	BU
H	GE	CO	PA	TA	KE	CE	NA	VA	MO	LO	ME	NA	DU	DE	CE
J	BO	FO	DA	DU	DA	LE	BO	SI	JO	VA	DO	DE	TI	NI	DO
K	CO	FI	DE	VE	CI	BU	DA	LE	BO	VI	DO	NA	JO	BE	KI
L	VA	DU	DE	KO	GO	RE	MO	PE	SA	RA	JE	KA	DO	PI	RI

(For distribution, see page 5)

2. Solve the following cryptogram, and recover all keys:

					5					10					15
A	DR	DD	SY	DA	RA	RR	SB	YA	BT	TY	AR	HI	DB	TB	AD
B	YY	YB	SA	AA	HI	DA	TD	HR	YB	TD	RB	RI	AI	HH	BT
C	DD	IA	AI	BB	HA	YD	TH	YA	HI	BA	YT	YD	YY	BD	YH
D	SD	DI	SB	AA	ST	YD	RH	SD	SR	YR	DT	SR	RA	RR	YB
E	SA	BT	TY	HR	AI	DB	IB	AD	DY	YB	SA	HA	HI	DA	TD
F	TS	DB	SH	YH	DI	SD	TT	TT	YY	HH	ST	YI	SB	AA	ST
G	DD	AH	DH	YT	RH	HI	ID	AR	SB	BA	RI	HB	AI	HI	RH
H	DB	SH	HA	RI	DA	AI	IB	YB	DI	SI	DD	YA	BB	YT	HH
J	II	YH	TY	BS	DD	YR	SR	RI	HH	TD	DT	TA	AI	RY	ST
K	SH	DH	AB	AI	TI	YT	AH	HY	AR	AI	RH	DI	YD	DD	YA
L	TB	DT	HH	SB	AA	DT	DD	RH	YD	DR	YB	DH	SH	SR	DD
M	DA	SI	RI	ID	ST	BD	SI	SD	TT	BH	SH	RI	AA	HI	BB
N	IS	BI	HI	RH	AY	DB	BA	AI	DH	SH					

(For distribution, see page 5)

~~RESTRICTED~~

3. Solve the following cryptogram, and recover all keys:

					5					10					15
A	99	18	57	82	12	28	78	90	25	04	15	30	04	06	14
B	57	34	64	20	72	15	30	02	57	44	84	52	66	11	81
C	87	58	35	78	31	14	70	90	68	47	30	13	15	21	86
D	92	43	10	30	35	20	31	32	64	18	57	26	84	12	06
E	34	25	69	72	90	78	07	90	31	29	57	50	82	19	53
F	31	72	51	36	10	86	36	47	18	67	26	04	92	82	30
G	08	31	58	90	88	87	91	10	20	82	31	14	56	57	31
H	88	04	31	30	66	47	30	36	18	99	20	06	97	31	21
J	55	99	18	20	10	28	74	68	90	41	69	82	90	78	31
K	86	88	15	91	26	92	72	87	14	43	20	53	28	64	92
L	47	02	58	35	10	96	05	34	37	85	06	26	80	50	92
M	68	10	70	81	92	18	02	86	49	47	07	82	94	06	69
N	15	21	90	56	10	40	01	68	90	15	35	57	52	32	60
P	47	64	36	71	06	55	00	68	78	45	52	12	69	43	

(For distribution, see page 5)

~~RESTRICTED~~

~~RESTRICTED~~

4. This message is suspected of having an ending similar to Problem 3. Solve it and recover all keys:

					5					10					15
A	22	08	71	29	19	83	05	34	76	58	05	56	62	26	22
B	35	48	75	13	78	58	34	65	02	07	71	51	87	35	96
C	10	32	69	45	47	81	46	11	01	14	67	37	75	79	35
D	30	53	29	37	46	60	19	30	94	66	49	68	88	57	98
E	84	93	30	86	28	90	51	04	53	03	84	76	58	31	57
F	42	12	86	49	36	79	54	26	09	38	24	41	86	63	79
G	08	28	67	68	66	94	22	63	71	66	83	56	05	07	58
H	95	60	19	62	26	48	23	59	40	38	15	67	43	92	42
J	62	77	43	79	54	69	38	65	16	82	10	96	67	97	57
K	48	93	24	13	53	29	46	37	32	65	12	94	84	95	68
L	83	93	98	37	75	79	45	12	97	84	53	03	75	76	95
M	31	29	32	21	49	17	25	73	00	69	86	36	79	45	19
N	77	98	38	95	97	93	94	98	72	42	59	00	08	50	44
P	27	26	62	57	06	91	23								

~~RESTRICTED~~

FREQUENCY DISTRIBUTIONS

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

Problem 1

	A	B	D	H	I	R	S	T	Y
A	5	1	2	2	9	3	-	-	1
B	3	3	2	1	1	-	1	3	-
D	5	5	8	4	4	2	-	4	1
H	3	1	-	5	8	2	-	-	1
I	1	3	1	-	1	-	1	-	-
R	2	1	-	6	6	2	-	-	1
S	3	5	4	6	3	4	-	5	1
T	1	2	4	1	1	-	1	3	3
Y	4	6	5	3	1	2	-	4	3

Problem 2

	0	1	2	3	4	5	6	7	8	9
0	1	1	3	-	4	1	6	2	1	-
1	7	1	3	1	4	6	-	-	6	1
2	6	3	-	-	-	2	4	-	3	1
3	7	10	2	-	3	4	4	1	-	-
4	1	1	-	3	1	1	-	6	-	1
5	2	1	3	2	-	2	2	7	3	-
6	1	-	-	-	4	-	2	1	5	4
7	2	1	4	-	1	-	-	-	5	-
8	1	2	6	-	2	1	4	3	3	-
9	9	2	6	-	1	-	1	1	-	3

Problem 3

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

Problem 4

5. Solve the following cryptogram, and recover all keys:

80713	06941	35696	80213	28061	37695
69680	91394	78800	25513	28096	91134
47713	68026	97695	13913	72502	56475
80280	88091	35802	25247	31341	39696
25525	12508	09132	47825	81314	74256
69525	51301	36477	13169	46966	90699
80247	46951	30801	80525	11378	04470
69213	11308	03477			

6. Solve the following cryptogram, and recover all keys:

18905	52131	89011	04414	52131	34022
05518	92022	35156	19005	52240	55145
19020	21561	67189	08815	60110	44190
08801	11900	22055	05514	54044	15460
35832	53583	14303	41532	53474	15459
46035	83813	14280	27946	04603	14448
51628	03143	58404	33637	04044	15291
37031	43036	73730	72971	87296	73684
70757	26957	30572	71872	97075	72550
57261	76847	29729	60661	77186	51572
71871	85385	94572			

7. Solve the following cryptogram, and recover all keys:

72109	19015	41776	04657	89925	96235
70368	62717	67091	83938	99294	88596
52368	62170	37091	22620	80735	96695
04627	17032	53136	77644	22537	12262
47907	38026	22703	88434	30196	04118
66826	27034	15596	84825	35230	46569
16375	84979	74893	10920	85780	73541
97477	67212	08479	35210	91365	78947
39865	97030	28334	15432	54516	59910
04639	82992	26541	09142	43430	28208
75852	33987	03712	25322	67217	58578

~~RESTRICTED~~

8. The following cryptograms are suspected to be isologs. Solve them, and recover all keys:

Message "A"

09728	23144	33987	73514	27769	10677
94418	99479	41948	66432	24374	48499
56758	47636	35546	81176	12242	30777
76194	15272	62644	85211	21361	71687
28759	72459	47047	20204	22145	53570
21377	58467	36166	13037	05358	25876
64403	33524	36847	98975	76679	83637
79946	05777	46243	95667	15086	47920
54391	27284	32060	43178	94367	66414
32190	15429	62648	60975	47915	66679
14422	70281	93894	71368	35325	27686
21707	79439	22000			

Message "B"

87560	77444	35211	41109	33772	89084
55415	78586	41056	35506	15844	48995
20110	23777	58199	19437	57052	62714
37174	88756	25154	11724	98779	72367
61813	38507	47890	68719	65521	08875
68548	81270	33609	17554	83811	72477
85433	50805	37598	60718	37306	17704
06159	62714	46551	69370	50945	58696
19561	70681	86600	83474	55377	71502
16576	41295	65052	00751	47289	33956
59497	38764	66574	72261	08560	73763
68350	48516	25000			

~~RESTRICTED~~

9. The following naval messages are suspected to be isologs, containing the probable word "TASK FORCE". Solve them, and recover all keys.

Message "A"

43022	83524	26060	98448	56175	57368
05544	54713	25748	18995	73211	78809
78230	46746	55566	38971	52835	54310
66179	30225	49705	63605	75310	83452
92351	03132	27998	93539	26288	11095
80473	12200	63369	42108	52097	11477
11306	68721	98883	68453	95650	15184
59749	92076	67000			

Message "B"

77639	32338	96687	32583	16771	36033
25195	21007	61936	37147	94702	74323
91551	84030	23211	74696	15784	34746
34170	59391	35584	17645	65752	24915
07432	64598	99104	17307	66639	31127
90402	53353	77760	84479	75139	10388
02285	42214	80132	62568	27529	42875
07934	45455	20000			

10. The following cryptogram is suspected to begin with the opening stereotype "REFERENCE YOUR MESSAGE....". Solve it, and recover all keys.

40162	42385	52104	83121	44422	37211
99099	42127	37912	77785	80116	44444
13378	77640	12255	50022	48883	78850
22287	84629	99920	06648	91253	20729
01331	81222	90051	99523	19391	41936
61045	48376	88311	15454	00022	05509
60615	57129	18859	20396	66603	14945
35079	88552	82411	08663	05032	28600
07722	55212	00080	00774	72883	40900

APPENDIX 2

LETTER FREQUENCY DATA - ENGLISH

ENGLISH CRYPTANALYTIC DATA
FREQUENCY TABLES

* * * * *

SPECIAL-PURPOSE DATA

TABLE 1-A.—*Absolute frequencies of letters appearing in five sets of Governmental plain-text telegrams, each set containing 10,000 letters, arranged alphabetically*

Set No. 1		Set No. 2		Set No. 3		Set No. 4		Set No. 5	
Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency
A	738	A	783	A	681	A	740	A	741
B	104	B	103	B	98	B	83	B	99
C	319	C	300	C	288	C	326	C	301
D	387	D	413	D	423	D	451	D	448
E	1,367	E	1,294	E	1,292	E	1,270	E	1,275
F	253	F	287	F	308	F	287	F	281
G	166	G	175	G	161	G	167	G	150
H	310	H	351	H	335	H	349	H	349
I	742	I	750	I	787	I	700	I	697
J	18	J	17	J	10	J	21	J	16
K	36	K	38	K	22	K	21	K	31
L	365	L	393	L	333	L	386	L	344
M	242	M	240	M	238	M	249	M	268
N	786	N	794	N	815	N	800	N	780
O	685	O	770	O	791	O	756	O	762
P	241	P	272	P	317	P	245	P	260
Q	40	Q	22	Q	45	Q	38	Q	30
R	760	R	745	R	762	R	735	R	786
S	658	S	583	S	585	S	628	S	604
T	936	T	879	T	894	T	958	T	928
U	270	U	233	U	312	U	247	U	238
V	163	V	173	V	142	V	133	V	155
W	166	W	163	W	136	W	133	W	182
X	43	X	50	X	44	X	53	X	41
Y	191	Y	155	Y	179	Y	213	Y	229
Z	14	Z	17	Z	2	Z	11	Z	5
Total	10,000		10,000		10,000		10,000		10,000

TABLE 1–B.—*Absolute frequencies of letters appearing in five sets of Governmental plain-text telegrams, each set containing 10,000 letters, arranged according to frequency*

Set No. 1		Set No. 2		Set No. 3		Set No. 4		Set No. 5	
Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency	Letter	Absolute Frequency
E	1,367	E	1,294	E	1,292	E	1,270	E	1,275
T	936	T	879	T	894	T	958	T	928
N	786	N	794	N	815	N	800	R	786
R	760	A	783	O	791	O	756	N	780
I	742	O	770	I	787	A	740	O	762
A	738	I	750	R	762	R	735	A	741
O	685	R	745	A	681	I	700	I	697
S	658	S	583	S	585	S	628	S	604
D	387	D	413	D	423	D	451	D	448
L	365	L	393	H	335	L	386	H	349
C	319	H	351	L	333	H	349	L	344
H	310	C	300	P	317	C	326	C	301
U	270	F	287	U	312	F	287	F	281
F	258	P	272	F	308	M	249	M	268
M	242	M	240	C	288	U	247	P	260
P	241	U	233	M	238	P	245	U	238
Y	191	G	175	Y	179	Y	213	Y	229
G	166	V	173	G	161	G	167	W	182
W	166	W	163	V	142	V	138	V	155
V	163	Y	155	W	136	W	133	G	150
B	104	B	103	B	98	B	83	B	99
X	43	X	50	Q	45	X	53	X	41
Q	40	K	38	X	44	Q	38	K	31
K	36	Q	22	K	22	K	21	Q	30
J	18	J	17	J	10	J	21	J	16
Z	14	Z	17	Z	2	Z	11	Z	5
Total	10,000		10,000		10,000		10,000		10,000

TABLE 1–C.—*Absolute frequencies of vowels, high-frequency consonants, medium-frequency consonants, and low-frequency consonants appearing in five sets of Governmental plain-text telegrams, each set containing 10,000 letters*

Set No.	Vowels	High-Frequency Consonants	Medium-Frequency Consonants	Low-Frequency Consonants
1	3,993	3,527	2,329	151
2	3,985	3,414	2,457	144
3	4,042	3,479	2,356	123
4	3,926	3,572	2,358	144
5	3,942	3,546	2,389	123
Total [1]	19,888	17,538	11,889	685

[1] Grand total, 50,000.

TABLE 2-A.—*Absolute frequencies of letters appearing in the combined five sets of messages totalling 50,000 letters, arranged alphabetically*

A	3,683	G	819	L	1,821	Q	175	V	766
B	487	H	1,694	M	1,237	R	3,788	W	780
C	1,534	I	3,676	N	3,975	S	3,058	X	231
D	2,122	J	82	O	3,764	T	4,595	Y	967
E	6,498	K	148	P	1,335	U	1,300	Z	49
F	1,416								

TABLE 2-B.—*Absolute frequencies of letters appearing in the combined five sets of messages totalling 50,000 letters, arranged according to frequency*

E	6,498	I	3,676	C	1,534	Y	967	X	231
T	4,595	S	3,058	F	1,416	G	819	Q	175
N	3,975	D	2,122	P	1,335	W	780	K	148
R	3,788	L	1,821	U	1,300	V	766	J	82
O	3,764	H	1,694	M	1,237	B	487	Z	49
A	3,683								

TABLE 2-C.—*Absolute frequencies of vowels, high-frequency consonants, medium-frequency consonants, and low-frequency consonants appearing in the combined five sets of messages totalling 50,000 letters*

Vowels	19,888
High-frequency consonants (D, N, R, S, and T)	17,538
Medium-frequency consonants (B, C, F, G, H, L, M, P, V, and W)	11,889
Low-frequency consonants (J, K, Q, X, and Z)	685
Total	50,000

TABLE 2-D.—*Absolute frequencies of letters as initial letters of 10,000 words found in Governmental plain-text telegrams*

(1) ARRANGED ALPHABETICALLY

A	905	G	109	L	196	Q	30	V	77
B	287	H	272	M	384	R	611	W	320
C	664	I	344	N	441	S	965	X	4
D	525	J	44	O	646	T	1,253	Y	88
E	390	K	23	P	433	U	122	Z	12
F	855								
								Total	10,000

(2) ARRANGED ACCORDING TO FREQUENCY

T	1,253	R	611	M	384	L	196	J	44
S	965	D	525	I	344	U	122	Q	30
A	905	N	441	W	320	G	109	K	23
F	855	P	433	B	287	Y	88	Z	12
C	664	E	390	H	272	V	77	X	4
O	646								
								Total	10,000

TABLE 2-E.—*Absolute frequencies of letters as final letters of 10,000 words found in Governmental plain-text telegrams*

(1) ARRANGED ALPHABETICALLY

A	269	G	225	L	354	Q	8	V	4
B	22	H	450	M	154	R	769	W	45
C	86	I	22	N	872	S	962	X	116
D	1,002	J	6	O	575	T	1,007	Y	866
E	1,628	K	53	P	213	U	31	Z	9
F	252								
								Total	10,000

(2) ARRANGED ACCORDING TO FREQUENCY

E	1,628	R	769	F	252	C	86	I	22
T	1,007	O	575	G	225	K	53	Z	9
D	1,002	H	450	P	213	W	45	Q	8
S	962	L	354	M	154	U	31	J	6
N	872	A	269	X	116	B	22	V	4
Y	866								
								Total	10,000

TABLE 3.—*Relative frequencies of letters appearing in 1,000 letters based upon Table 2-B*

(1) ARRANGED ALPHABETICALLY

A	73.66	G	16.38	L	36.42	Q	3.50	V	15.32
B	9.74	H	33.88	M	24.74	R	75.76	W	15.60
C	30.68	I	73.52	N	79.50	S	61.16	X	4.62
D	42.44	J	1.64	O	75.28	T	91.90	Y	19.34
E	129.96	K	2.96	P	26.70	U	26.00	Z	.98
F	28.32								
								Total	1,000.00

(2) ARRANGED ACCORDING TO FREQUENCY

E	129.96	I	73.52	C	30.68	Y	19.34	X	4.62
T	91.90	S	61.16	F	28.32	G	16.38	Q	3.50
N	79.50	D	42.44	P	26.70	W	15.60	K	2.96
R	75.76	L	36.42	U	26.00	V	15.32	J	1.64
O	75.28	H	33.88	M	24.74	B	9.74	Z	.98
A	73.66								
								Total	1,000.00

(3) VOWELS

A	73.66
E	129.96
I	73.52
O	75.28
U	26.00
Y	19.34
Total	397.76

(4) HIGH-FREQUENCY CONSONANTS

D	42.44
N	79.50
R	75.76
S	61.16
T	91.90
Total	350.76

TABLE 3, Contd.—*Relative frequencies of letters appearing in 1,000 letters based upon Table 2–B*

(5) MEDIUM-FREQUENCY CONSONANTS

B	9.74
C	30.68
F	28.82
G	16.38
H	33.88
L	36.42
M	24.74
P	26.70
V	15.32
W	15.60
Total	237.78

(6) LOW-FREQUENCY CONSONANTS

X	4.62
Q	3.50
K	2.96
J	1.64
Z	.98
Total	13.70

Total (3), (4), (5), (6) 1,000.00

TABLE 4.—*Frequency distribution for 10,000 letters of literary English, as compiled by Hitt* [1]

(1) ARRANGED ALPHABETICALLY

A	778	G	174	L	372	Q	8	V	112
B	141	H	595	M	288	R	651	W	176
C	296	I	667	N	686	S	622	X	27
D	402	J	51	O	807	T	855	Y	196
E	1,277	K	74	P	223	U	308	Z	17
F	197								

(2) ARRANGED ACCORDING TO FREQUENCY

E	1,277	R	651	U	308	Y	196	K	74
T	855	S	622	C	296	W	176	J	51
O	807	H	595	M	288	G	174	X	27
A	778	D	402	P	223	B	141	Z	17
N	686	L	372	F	197	V	112	Q	8
I	667								

TABLE 5.—*Frequency distribution for 10,000 letters of telegraphic English, as compiled by Hitt* [1]

(1) ARRANGED ALPHABETICALLY

A	813	G	201	L	392	Q	38	V	136
B	149	H	386	M	273	R	677	W	166
C	306	I	711	N	718	S	656	X	51
D	417	J	42	O	844	T	634	Y	208
E	1,319	K	88	P	243	U	321	Z	6
F	205								

(2) ARRANGED ACCORDING TO FREQUENCY

E	1,319	S	656	U	321	F	205	K	88
O	844	T	634	C	306	G	201	X	51
A	813	D	417	M	273	W	166	J	42
N	718	L	392	P	243	B	149	Q	38
I	711	H	386	Y	208	V	136	Z	6
R	677								

[1] Hitt, Capt. Parker. *Manual for the Solution of Military Ciphers.* Army Service Schools Press, Fort Leavenworth, Kansas, 1916.

TABLE 6-A.—*Frequency distribution of digraphs, based on 50,000 letters of Governmental plain-text telegrams; reduced to 5,000 digraphs*

SECOND LETTER

FIRST 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	Total	Blanks
A	3	6	14	27	1	4	6	2	17	1	2	32	14	64	2	12		44	41	47	18	7	3		12		374	3
B	4				18				2	1		6	1		4			2	1	1	2				7		49	14
C	20		3	1	32	1		14	7		4	5	1	1	41			4	1	14	4		1		1		155	8
D	32	4	4	8	33	8	2	2	27	1		3	5	4	16	5	2	12	13	15	5	3	4		1		209	3
E	35	4	32	60	42	18	4	7	27	1		29	14	111	12	20	12	37	54	37	3	20	7	7	4	1	648	1
F	5		2	1	10	11	1		39			2	1		40	1		9	3	11	3		1		1		141	9
G	7		2	1	14	2	1	20	5	1		2	1	3	6	2		5	3	4	2		1				82	7
H	20	1	3	2	20	5			33			1	2	3	20	1	1	17	4	28	8		1		1		171	7
I	8	2	22	6	13	10	19				2	28	9	75	41	7		27	35	27		25		15		2	368	7
J	1				2										2						2						7	22
K	1		1		6				2			1		1					1								13	19
L	28	3	3	9	37	3	1	1	20			27	2	1	13	3		2	6	8	2	2	2		10		183	5
M	36	6	3	1	26	1		1	9				13		10	8		2	4	2	2				2		126	10
N	26	2	19	52	57	9	27	4	30	1	2	5	5	8	18	3	1	4	24	82	7	3	3		5		397	2
O	7	4	8	12	3	25	2	3	5	1	2	19	25	77	6	25		64	14	19	37	7	8	1	2		376	2
P	14	1	1	1	28	2		3	6			13	4	1	17	11		18	6	8	3	1	1		1		135	6
Q													1					1			15						17	23
R	39	2	9	17	98	6	7	3	30	1	1	5	9	7	28	13		11	31	42	5	5	4		9		382	3
S	24	3	13	5	49	12	2	26	34		1	2	3	4	15	10		5	19	68	11	1	4		1		307	4
T	28	3	6	6	71	7	1	78	45			5	6	7	50	2	1	17	19	19	5		36		41	1	454	4
U	5	3	3	3	11	1	8		5			6	5	21	1	2		31	12	12		1					130	9
V	6				57				12						1					1							77	21
W	12				22			4	13			1		2	19			1	1						1		76	16
X	2		2	1	1	1		1	2					1	1	2		1	1	7							23	13
Y	6	2	4	4	9	11	1	1	3			2	2	6	10	3		4	11	15	1		1				96	7
Z	1				2				1																		4	23
Total	370	46	154	217	657	137	82	170	374	8	14	189	128	397	378	130	17	368	304	462	180	75	77	23	99	4	5,000	
Blanks	1	11	6	7	1	7	12	10	3	18	19	6	6	7	3	8	21	4	4	5	7	15	11	23	10	23		248

TABLE 6-B.—*Frequency distribution of digraphs (naval text), based on 20,000 letters of naval text; reduced to 2,000 digraphs*[1]

SECOND LETTER

FIRST 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	Total	Blanks
A	1	4	9	5		2	3	1	8		3	7	2	29		4		16	11	31	1	3		1	5		146	6
B	4			1	8				1			6	2		4						1				2		29	17
C	7		1		10		2	5	1		4				22	1		4	1	4							62	14
D	10	2	2	2	15	3	1	1	12			2	2		4	3		8	6	6	2	1	1		3		86	6
E	9	3	8	24	25	7	1	2	7	1	1	6	6	34	6	10	1	43	23	18	1	7	2	4	1	4	254	0
F	2		1		2	1			13			5	1		12	1		2	1	5	1				1		48	12
G	4		1	1	8	1	1	11	2			2		1	2	1		2	2	6	3					1	49	9
H	6				7	1			6				1		3	1		7	1	11	6		1				51	14
I	2	1	6	2	2	5	11					8	2	42	21	2		10	10	11		9		5			149	9
J															2												2	25
K	1	1	1		3	1			2			1		1													11	18
L	14	1	1		15	1			8			6			7	2		1	2	1	1	2			2		64	11
M	11	1			5				4			1	2		4	2			1						3		34	16
N	10	3	8	22	22	5	22	2	6		2	2	2	3	10	2		2	9	27	3		1				168	6
O	3	3	8	11	4	9	2		6		1	4	9	38	2	8		20	9	7	20	1	4	1	1	1	167	3
P	4				18			1	1			5			7	3		8	8	2	1						58	15
Q																					3						3	25
R	14	2	6	9	34	2	3		19		1	1	3	3	24	2		2	8	10	4		1				148	7
S	8	2	8	1	15	2		4	13			2	1	1	5	6	1	1	6	23	6		3				108	7
T	16	1	4	3	27	4	1	21	28			3	1	2	22	3		10	3	8	4		12		8	4	185	5
U	4	3	1	2	3		1		4			2	2	9		1		1	4	10							47	12
V	3				17				4						1												25	22
W	4				10			1	5						6			1									27	20
X			1			1		1	4			1								2							10	20
Y	3	1	2	1	2	3			1			3	2		2	2		1	2	2			1				28	11
Z					10																	1					11	24
Total	140	28	63	84	262	48	48	50	150	1	12	67	38	163	166	54	2	139	107	184	57	24	26	11	26	10	1,960	
Blanks	4	12	9	18	4	10	15	15	4	25	20	7	11	15	6	8	24	8	8	8	11	19	17	22	17	22		334

[1] Fractional values have been discarded. This accounts for the discrepancy between the indicated total (1,960) and the stated total (2,000).

TABLES 7–11, Inclusive

Absolute frequencies of digraphs, trigraphs, and tetragraphs and the logarithms of their assigned probabilities [1]

1. For each of the following 18 tables, the basic data were first arranged according to their absolute frequencies (F), and then the logarithms—$L_{10}(F)$ of the frequencies found.

2. The tables are designed to facilitate determination of the relative weights or probability of occurrence of sets of digraphs, trigraphs, or tetragraphs, particularly with respect to various "matching" operations. For example, are the matched digraphs RE and ET more probable than the matched digraphs RT and EF? Table 7–A shows the frequencies (F) of the digraphs to be as follows: RE=98, ET=37, RT=42, EF=18. Therefore, 98 times 37 is compared with 42 times 18, or 3,626 with 756. **This arithmetic method of approach is extremely cumbersome for** a large number of comparisons. By using the logarithms of the individual frequencies, the operation is greatly simplified, since the addition of the logarithms of two numbers is equivalent to the multiplication of their equivalent arithmetic values. Thus, the foregoing computation may be expressed as Log 98+Log 37, compared with Log 42+Log 18, or 0.96+0.79 versus 0.81+0.66 (see Table 7–A and explanation below). If more than one occurrence of a particular digraph is involved, it is merely necessary to multiply the logarithmic value by the number of the occurrences, viz., Log X+2(Log Y)+3(Log Z), as compared with Log A+3(Log B)+2(Log C).

3. The logarithm of any given number is the power to which 10 must be raised to equal the given number. Thus, $10^2=100$, or the logarithm of 100=2. Similarly, $10^3=1{,}000$, or the logarithm of 1,000=3. The sum of logarithms is equal to the logarithm of the product of their antilogs (arithmetic numbers they represent). For example, $10^2=100$; $10^3=1000$; $10^{2+3}=100\times1000$; Log 100,000=5. Also, $10^0=1$, or Log 1=0. The Log of 0 is minus infinity ($-\infty$).

4. In the compilation of the logarithms of the elements constituting these tables, frequencies of 1, of course, had a logarithmic value of 0.00. Digraphs which did not occur,[2] i.e., those with 0 occurrences, had a logarithmic value of minus infinity ($-\infty$). For practical use, each of the original frequency occurrences in these tables was doubled; i. e., EN was given a frequency of 222 instead of 111, the frequency of RE became 196 instead of 98, etc. Thus, single occurrences were doubled ($2\times1=2$), and the logarithms of those elements became 0.30 instead of 0. This is equivalent to saying Log 1+Log 2=0.00+0.30=0.30. Those elements which occurred 0 times, now were assumed to have an occurrence of 1, with an equivalent logarithmic value of 0.00.

5. In order to place all the logarithms of the initial frequencies on a comparable logarithmic basis, it was merely necessary to add 0.30 to each of them. While EN had a frequency of 111 in the original compilation, it now had a frequency of 222, or 2(111). The logarithm of 222 is 2.35. This is equivalent to saying Log 111+Log 2=2.05+0.30=2.35.

6. The frequencies as stated in terms of their actual logarithms do not readily indicate their relative size for each distribution. Therefore, the highest frequency in each group was given a value of 0.99, and the lowest a value of 0; frequencies intermediate between these extremes were

---

[1] These frequency distributions are based upon data derived from 50,000 letters of U. S. Governmental plain-text telegrams, reduced to 5,000 digraphs.

[2] While in general it is possible to assign probability values to digraphs in accordance with their observed frequencies, it is not strictly correct to associate the probability "0" with a frequency of zero. This would be equivalent to saying: "Because a specified digraph has not occurred, it cannot occur," and would be reflected in the mathematics: "Log probability zero equals minus infinity." What may be said is: "Since a specified digraph has not occurred in the data its true probability value is unknown, except that it must be below the probability value assigned to a frequency of one." The proper way to assign a probability value to digraphs with frequencies of zero is to continue counting until they have at least one occurrence; then the true relative probability can be found.

A simple practical method of taking this difficulty into account is merely to assume that in twice the amount of data the digraph probably would have occurred at least once; that is, it has a frequency of one-half.

It should be pointed out, however, that since probabilities are multiplied (by summing logarithms) a 10% error in evaluating the digraph ZZ for example, makes the product, wherever ZZ occurs, 10% wrong, and is just as serious as a 10% error in evaluating the high-frequency digraph EN.

In practice, however, results obtained from the logarithmic method are so satisfactory that refinements are not needed.

evaluated in proportion to their respective frequencies. This is equivalent to expressing the frequencies in logarithms with a base other than 10. In other words, this procedure of converting the logarithms to the range from .00 to .99 consists in dividing up the original range of logarithms into 100 equal parts and assigning each one to the proper rank in the range.

7. The new base (C) used to convert each of the digraphic frequencies to the logarithmic range 0 to 0.99 is derived as follows, when 222 is the highest frequency (F):-

$$\text{Let } 222 = C^{0.99}$$
$$\text{Log}_{10}\, 222 = \text{Log}_{10}\, C^{0.99}$$
$$\text{Log}_{10}\, 222 = (0.99)\ (\text{Log}_{10}\, C)$$
$$C = \text{Antilog}\,\frac{\text{Log}_{10}\, 222}{0.99} = \text{Antilog}\,\frac{2.35}{0.99}$$
$$C = 224$$

8. The formula for the computation of the logarithm to the new base (C) of any actual frequency (Y) of a series is:

$$\text{Log}_c\, Y = \frac{\text{Log}_{10}\, Y}{\text{Log}_{10}\, C}$$

It is more expeditious to use reciprocals in the conversion of a whole series of logarithmic values, as in this instance. The formula is: $(\text{Log}_{10}\, C)^{-1} \cdot (\text{Log}_{10}\, Y) = \text{Log}_c\, Y$.

9. The digraphic index chart, Table 15, on page 37, summarizes the logarithmic frequencies of all English plain-text digraphs, computed to a base of 224 so that the logarithm of the highest frequency (EN) is 0.99.

Example:

$$EN = 222$$
$$\text{Log}_{10}\, 222 = 2.35$$
$$(\text{Log}_{10}\, C)^{-1} = (\text{Log}_{10}\, 224)^{-1} = 0.421$$
$$\text{Log}_c\, 222 = 0.421 \times 2.35 = 0.99$$

10. Likewise, the trigraphs and tetragraphs have been computed to the bases L586 and L244, respectively, so that the logarithms of the highest-frequency trigraph (ENT) and tetragraph (TION) are 0.99. Since no use is being made of the trigraphs appearing less than 100 times and tetragraphs appearing less than 50 times, the basic frequencies of the trigraphs and tetragraphs have not been doubled in computing the new bases of the logarithms.

TABLE 7–A.—*The 428 different digraphs of Table 6–A, arranged according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$
EN	111	2.05	.99
RE	98	1.99	.96
ER	87	1.94	.94
NT	82	1.91	.93
TH	78	1.89	.92
ON	77	1.89	.92
IN	75	1.88	.92
TE	71	1.85	.91
AN	64	1.81	.89
OR	64	1.81	.89
ST	63	1.80	.88
ED	60	1.78	.88
NE	57	1.76	.87
VE	57	1.76	.87
ES	54	1.73	.86
ND	52	1.72	.85
TO	50	1.70	.84
SE	49	1.69	.84
	[1]1,249		
AT	47	1.67	.83
TI	45	1.65	.82
AR	44	1.64	.82
EE	42	1.62	.81
RT	42	1.62	.81
AS	41	1.61	.80
CO	41	1.61	.80
IO	41	1.61	.80
TY	41	1.61	.80
FO	40	1.60	.80
FI	39	1.59	.80
RA	39	1.59	.80
ET	37	1.57	.79
LE	37	1.57	.79
OU	37	1.57	.79
MA	36	1.56	.78
TW	36	1.56	.78
EA	35	1.54	.78
IS	35	1.54	.78
SI	34	1.53	.77
DE	33	1.52	.77
HI	33	1.52	.77
AL	32	1.51	.76
CE	32	1.51	.76

	F	$L_{10}(F)$	$L_{224}(2F)$
DA	32	1.51	.76
EC	32	1.51	.76
RS	31	1.49	.75
UR	31	1.49	.75
NI	30	1.48	.75
RI	30	1.48	.75
EL	29	1.46	.74
HT	28	1.45	.74
LA	28	1.45	.74
RO	28	1.45	.74
TA	28	1.45	.74
	[2]2,495		
AD	27	1.43	.73
DI	27	1.43	.73
EI	27	1.43	.73
IR	27	1.43	.73
IT	27	1.43	.73
LL	27	1.43	.73
NG	27	1.43	.73
ME	26	1.41	.72
NA	26	1.41	.72
SH	26	1.41	.72
IV	25	1.40	.72
OF	25	1.40	.72
OM	25	1.40	.72
OP	25	1.40	.72
NS	24	1.38	.71
SA	24	1.38	.71
IL	23	1.36	.70
PE	23	1.36	.70
IC	22	1.34	.69
WE	22	1.34	.69
UN	21	1.32	.68
CA	20	1.30	.67
EP	20	1.30	.67
EV	20	1.30	.67
GH	20	1.30	.67
HA	20	1.30	.67
HE	20	1.30	.67
HO	20	1.30	.67
LI	20	1.30	.67
IG	19	1.28	.67
NC	19	1.28	.67

	F	$L_{10}(F)$	$L_{224}(2F)$
OL	19	1.28	.67
OT	19	1.28	.67
SS	19	1.28	.67
TS	19	1.28	.67
TT	19	1.28	.67
WO	19	1.28	.67
BE	18	1.26	.66
EF	18	1.26	.66
NO	18	1.26	.66
PR	18	1.26	.66
AI	17	1.23	.64
HR	17	1.23	.64
PO	17	1.23	.64
RD	17	1.23	.64
TR	17	1.23	.64
DO	16	1.20	.63
DT	15	1.18	.62
IX	15	1.18	.62
QU	15	1.18	.62
SO	15	1.18	.62
YT	15	1.18	.62
AC	14	1.15	.61
AM	14	1.15	.61
CH	14	1.15	.61
CT	14	1.15	.61
EM	14	1.15	.61
GE	14	1.15	.61
OS	14	1.15	.61
PA	14	1.15	.61
AU	13	1.11	.59
DS	13	1.11	.59
IE	13	1.11	.59
LO	13	1.11	.59
MM	13	1.11	.59
PL	13	1.11	.59
RP	13	1.11	.59
SC	13	1.11	.59
WI	13	1.11	.59
	[3]3,745		
AP	12	1.08	.58
AY	12	1.08	.58
DR	12	1.08	.58
EO	12	1.08	.58

	F	$L_{10}(F)$	$L_{224}(2F)$
EQ	12	1.08	.58
OD	12	1.08	.58
SF	12	1.08	.58
US	12	1.08	.58
UT	12	1.08	.58
VI	12	1.08	.58
WA	12	1.08	.58
FF	11	1.04	.56
FT	11	1.04	.56
PP	11	1.04	.56
RR	11	1.04	.56
SU	11	1.04	.56
UE	11	1.04	.56
YF	11	1.04	.56
YS	11	1.04	.56
FE	10	1.00	.55
IF	10	1.00	.55
LY	10	1.00	.55
MO	10	1.00	.55
SP	10	1.00	.55
YO	10	1.00	.55
FR	9	0.95	.53
IM	9	0.95	.53
LD	9	0.95	.53
MI	9	0.95	.53
NF	9	0.95	.53
RC	9	0.95	.53
RM	9	0.95	.53
RY	9	0.95	.53
YE	9	0.95	.53
DD	8	0.90	.51
DF	8	0.90	.51
HU	8	0.90	.51
IA	8	0.90	.51
LT	8	0.90	.51
MP	8	0.90	.51
NN	8	0.90	.51
OC	8	0.90	.51
OW	8	0.90	.51
PT	8	0.90	.51
UG	8	0.90	.51
AV	7	0.85	.48
BY	7	0.85	.48
CI	7	0.85	.48

[1] The 18 digraphs above this line compose 25% of the total.
[2] The 53 digraphs above this line compose 50% of the total.
[3] The 122 digraphs above this line compose 75% of the total.

TABLE 7-A, Contd.—*The 428 different digraphs of Table 6-A, arranged according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	L_{244} (2F)		F	$L_{10}(F)$	L_{244} (2F)		F	$L_{10}(F)$	L_{244} (2F)		F	$L_{10}(F)$	L_{244} (2F)
EH	7	0.85	.48	RU	5	0.70	.42	GS	3	0.48	.33	JE	2	0.30	.25
EW	7	0.85	.48	RV	5	0.70	.42	HC	3	0.48	.33	JO	2	0.30	.25
EX	7	0.85	.48	SD	5	0.70	.42	HN	3	0.48	.33	JU	2	0.30	.25
GA	7	0.85	.48	SR	5	0.70	.42	LB	3	0.48	.33	KI	2	0.30	.25
IP	7	0.85	.48	TL	5	0.70	.42	LC	3	0.48	.33	LM	2	0.30	.25
NU	7	0.85	.48	TU	5	0.70	.42	LF	3	0.48	.33	LR	2	0.30	.25
OA	7	0.85	.48	UA	5	0.70	.42	LP	3	0.48	.33	LU	2	0.30	.25
OV	7	0.85	.48	UI	5	0.70	.42	MC	3	0.48	.33	LV	2	0.30	.25
RG	7	0.85	.48	UM	5	0.70	.42	NP	3	0.48	.33	LW	2	0.30	.25
RN	7	0.85	.48	AF	4	0.60	.38	NV	3	0.48	.33	MR	2	0.30	.25
TF	7	0.85	.48	BA	4	0.60	.38	NW	3	0.48	.33	MT	2	0.30	.25
TN	7	0.85	.48	BO	4	0.60	.38	OE	3	0.48	.33	MU	2	0.30	.25
XT	7	0.85	.48	CK	4	0.60	.38	OH	3	0.48	.33	MY	2	0.30	.25
AB	6	0.78	.45	CR	4	0.60	.38	PH	3	0.48	.33	NB	2	0.30	.25
AG	6	0.78	.45	CU	4	0.60	.38	PU	3	0.48	.33	NK	2	0.30	.25
BL	6	0.78	.45	DB	4	0.60	.38	RH	3	0.48	.33	OG	2	0.30	.25
GO	6	0.78	.45	DC	4	0:60	.38	SB	3	0.48	.33	OK	2	0.30	.25
ID	6	0.78	.45	DN	4	0.60	.38	SM	3	0.48	.33	OY	2	0.30	.25
KE	6	0.78	.45	DW	4	0.60	.38	TB	3	0.48	.33	PF	2	0.30	.25
LS	6	0.78	.45	EB	4	0.60	.38	UB	3	0.48	.33	RB	2	0.30	.25
MB	6	0.78	.45	EG	4	0.60	.38	UC	3	0.48	.33	SG	2	0.30	.25
OO	6	0.78	.45	EY	4	0.60	.38	UD	3	0.48	.33	SL	2	0.30	.25
PI	6	0.78	.45	GT	4	0.60	.38	YI	3	0.48	.33	TP	2	0.30	.25
PS	6	0.78	.45	HS	4	0.60	.38	YP	3	0.48	.33	UP	2	0.30	.25
RF	6	0.78	.45	MS	4	0.60	.38	AH	2	0.30	.25	WN	2	0.30	.25
TC	6	0.78	.45	NH	4	0.60	.38	AK	2	0.30	.25	XA	2	0.30	.25
TD	6	0.78	.45	NR	4	0.60	.38	AO	2	0.30	.25	XC	2	0.30	.25
TM	6	0.78	.45	OB	4	0.60	.38	BI	2	0.30	.25	XI	2	0.30	.25
UL	6	0.78	.45	PM	4	0.60	.38	BR	2	0.30	.25	XP	2	0.30	.25
VA	6	0.78	.45	RW	4	0.60	.38	BU	2	0.30	.25	YB	2	0.30	.25
YA	6	0.78	.45	SN	4	0.60	.38	DG	2	0.30	.25	YL	2	0.30	.25
YN	6	0.78	.45	SW	4	0.60	.38	DH	2	0.30	.25	YM	2	0.30	.25
CL	5	0.70	.42	WH	4	0.60	.38	DQ	2	0.30	.25	ZE	2	0.30	.25
DM	5	0.70	.42	YC	4	0.60	.38	FC	2	0.30	.25	AE	1	0.00	.13
DP	5	0.70	.42	YD	4	0.60	.33	FL	2	0.30	.25	AJ	1	0.00	.13
DU	5	0.70	.42	YR	4	0.60	.33	GC	2	0.30	.25	BJ	1	0.00	.13
FA	5	0.70	.42	AA	3	0.48	.33	GF	2	0.30	.25	BM	1	0.00	.13
GI	5	0.70	.42	AW	3	0.48	.33	GL	2	0.30	.25	BS	1	0.00	.13
GR	5	0.70	.42	CC	3	0.48	.33	GP	2	0.30	.25	BT	1	0.00	.13
HF	5	0.70	.42	DL	3	0.48	.33	GU	2	0.30	.25	CD	1	0.00	.13
NL	5	0.70	.42	DV	3	0.48	.33	HD	2	0.30	.25	CF	1	0.00	.13
NM	5	0.70	.42	EU	3	0.48	.33	HM	2	0.30	.25	CM	1	0.00	.13
NY	5	0.70	.42	FS	3	0.48	.33	IB	2	0.30	.25	CN	1	0.00	.13
OI	5	0.70	.42	FU	3	0.48	.33	IK	2	0.30	.25	CS	1	0.00	.13
RL	5	0.70	.42	GN	3	0.48	.33	IZ	2	0.30	.25	CW	1	0.00	.13

TABLE 7–A, Concluded.—*The 428 different digraphs of Table 6–A, arranged according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
CY----	1	0.00	.13	HW----	1	0.00	.13	PD----	1	0.00	.13	WL----	1	0.00	.13
DJ----	1	0.00	.13	HY----	1	0.00	.13	PN----	1	0.00	.13	WR----	1	0.00	.13
DY----	1	0.00	.13	JA----	1	0.00	.13	PV----	1	0.00	.13	WS----	1	0.00	.13
EJ----	1	0.00	.13	KA----	1	0.00	.13	PW----	1	0.00	.13	WY----	1	0.00	.13
EZ----	1	0.00	.13	KC----	1	0.00	.13	PY----	1	0.00	.13	XD----	1	0.00	.13
FD----	1	0.00	.13	KL----	1	0.00	.13	QM----	1	0.00	.13	XE----	1	0.00	.13
FG----	1	0.00	.13	KN----	1	0.00	.13	QR----	1	0.00	.13	XF----	1	0.00	.13
FM----	1	0.00	.13	KS----	1	0.00	.13	RJ----	1	0.00	.13	XH----	1	0.00	.13
FP----	1	0.00	.13	LG----	1	0.00	.13	RK----	1	0.00	.13	XN----	1	0.00	.13
FW----	1	0.00	.13	LH----	1	0.00	.13	SK----	1	0.00	.13	XO----	1	0.00	.13
FY----	1	0.00	.13	LN----	1	0.00	.13	SV----	1	0.00	.13	XR----	1	0.00	.13
GD----	1	0.00	.13	MD----	1	0.00	.13	SY----	1	0.00	.13	XS----	1	0.00	.13
GG----	1	0.00	.13	MF----	1	0.00	.13	TG----	1	0.00	.13	YG----	1	0.00	.13
GJ----	1	0.00	.13	MH----	1	0.00	.13	TQ----	1	0.00	.13	YH----	1	0.00	.13
GM----	1	0.00	.13	NJ----	1	0.00	.13	TZ----	1	0.00	.13	YU----	1	0.00	.13
GW----	1	0.00	.13	NQ----	1	0.00	.13	UF----	1	0.00	.13	YW----	1	0.00	.13
HB----	1	0.00	.13	OJ----	1	0.00	.13	UO----	1	0.00	.13	ZA----	1	0.00	.13
HL----	1	0.00	.13	OX----	1	0.00	.13	UV----	1	0.00	.13	ZI----	1	0.00	.13
HP----	1	0.00	.13	PB----	1	0.00	.13	VO----	1	0.00	.13		5,000		
HQ----	1	0.00	.13	PC----	1	0.00	.13	VT----	1	0.00	.13				

TABLE 7–B.—*The 18 digraphs composing 25% of the digraphs in Table 6–A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their initial letters*

(1) AND ACCORDING TO THEIR FINAL LETTERS

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
AN----	64	1.81	.89	ON----	77	1.89	.92
				OR----	64	1.81	.89
ED----	60	1.78	.88	RE----	98	1.99	.96
EN----	111	2.05	.99				
ER----	87	1.94	.94	SE----	49	1.69	.84
ES----	54	1.73	.86	ST----	63	1.80	.88
				TE----	71	1.85	.91
IN----	75	1.88	.92	TH----	78	1.89	.92
				TO----	50	1.70	.84
ND----	52	1.72	.85	VE----	57	1.76	.87
NE----	57	1.76	.87		1,249		
NT----	82	1.91	.93				

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
AN----	64	1.81	.89	ON----	77	1.89	.92
				OR----	64	1.81	.89
EN----	111	2.05	.99	RE----	98	1.99	.96
ER----	87	1.94	.94				
ED----	60	1.78	.88	ST----	63	1.80	.88
ES----	54	1.73	.86	SE----	49	1.69	.84
				TH----	78	1.89	.92
IN----	75	1.88	.92	TE----	71	1.85	.91
				TO----	50	1.70	.84
NT----	82	1.91	.93	VE----	57	1.76	.87
NE----	57	1.76	.87		1,249		
ND----	52	1.72	.85				

~~RESTRICTED~~

TABLE 7-C.—*The 53 digraphs composing 50% of the 5,000 digraphs of Table 6-A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their initial letters*

(1) AND ACCORDING TO THEIR FINAL LETTERS

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
AL	32	1.51	.76	MA	36	1.56	.78
AN	64	1.81	.89				
AR	44	1.64	.82	ND	52	1.72	.85
AS	41	1.61	.80	NE	57	1.76	.87
AT	47	1.67	.83	NI	30	1.48	.75
				NT	82	1.91	.93
CE	32	1.51	.76				
CO	41	1.61	.80	ON	77	1.89	.92
				OR	64	1.81	.89
DA	32	1.51	.76	OU	37	1.57	.79
DE	33	1.52	.77				
				RA	39	1.59	.80
EA	35	1.54	.78	RE	98	1.99	.96
EC	32	1.51	.76	RI	30	1.48	.75
ED	60	1.78	.88	RO	28	1.45	.74
EE	42	1.62	.81	RS	31	1.49	.75
EL	29	1.46	.74	RT	42	1.62	.81
EN	111	2.05	.99				
ER	87	1.94	.94	SE	49	1.69	.84
ES	54	1.73	.86	SI	34	1.53	.77
ET	37	1.57	.79	ST	63	1.80	.88
FI	39	1.59	.80	TA	28	1.45	.74
FO	40	1.60	.80	TE	71	1.85	.91
				TH	78	1.89	.92
HI	33	1.52	.77	TI	45	1.65	.82
HT	28	1.45	.74	TO	50	1.70	.84
				TW	36	1.56	.78
IN	75	1.88	.92	TY	41	1.61	.80
IO	41	1.61	.80				
IS	35	1.54	.78	UR	31	1.49	.75
LA	28	1.45	.74	VE	57	1.76	.87
LE	37	1.57	.79		2,495		

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
AN	64	1.81	.89	MA	36	1.56	.78
AT	47	1.67	.83				
AR	44	1.64	.82	NT	82	1.91	.93
AS	41	1.61	.80	NE	57	1.76	.87
AL	32	1.51	.76	ND	52	1.72	.85
				NI	30	1.48	.75
CO	41	1.61	.80				
CE	32	1.51	.76	ON	77	1.89	.92
				OR	64	1.81	.89
DE	33	1.52	.77	OU	37	1.57	.79
DA	32	1.51	.76				
				RE	98	1.99	.96
EN	111	2.05	.99	RT	42	1.62	.81
ER	87	1.94	.94	RA	39	1.59	.80
ED	60	1.78	.88	RS	31	1.49	.75
ES	54	1.73	.86	RI	30	1.48	.75
EE	42	1.62	.81	RO	28	1.45	.74
ET	37	1.57	.79				
EA	35	1.54	.78	ST	63	1.80	.88
EC	32	1.51	.76	SE	49	1.69	.84
EL	29	1.46	.74	SI	34	1.53	.77
FO	40	1.60	.80	TH	78	1.89	.92
FI	39	1.59	.80	TE	71	1.85	.91
				TO	50	1.70	.84
HI	33	1.52	.77	TI	45	1.65	.82
HT	28	1.45	.74	TY	41	1.61	.80
				TW	36	1.56	.78
IN	75	1.88	.92	TA	28	1.45	.74
IO	41	1.61	.80				
IS	35	1.54	.78	UR	31	1.49	.75
LE	37	1.57	.79	VE	57	1.76	.87
LA	28	1.45	.74		2,495		

~~RESTRICTED~~

TABLE 7-D.—*The 122 digraphs composing 75% of the 5,000 digraphs of Table 6-A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their initial letters*

(1) AND ACCORDING TO THEIR FINAL LETTERS

	F	$L_{10}(F)$	$L_{224}(2F)$
AC	14	1.15	.61
AD	27	1.43	.73
AI	17	1.23	.64
AL	32	1.51	.76
AM	14	1.15	.61
AN	64	1.81	.89
AR	44	1.64	.82
AS	41	1.61	.80
AT	47	1.67	.83
AU	13	1.11	.59
BE	18	1.26	.66
CA	20	1.30	.67
CE	32	1.51	.76
CH	14	1.15	.61
CO	41	1.61	.80
CT	14	1.15	.61
DA	32	1.51	.76
DE	33	1.52	.77
DI	27	1.43	.73
DO	16	1.20	.63
DS	13	1.11	.59
DT	15	1.18	.62
EA	35	1.54	.78
EC	32	1.51	.76
ED	60	1.78	.88
EE	42	1.62	.81
EF	18	1.26	.66
EI	27	1.43	.73
EL	29	1.46	.74
EM	14	1.15	.61
EN	111	2.05	.99
EP	20	1.30	.67

	F	$L_{10}(F)$	$L_{224}(2F)$
ER	87	1.94	.94
ES	54	1.73	.86
ET	37	1.57	.79
EV	20	1.30	.67
FI	39	1.59	.80
FO	40	1.60	.80
GE	14	1.15	.61
GH	20	1.30	.67
HA	20	1.30	.67
HE	20	1.30	.67
HI	33	1.52	.77
HO	20	1.30	.67
HR	17	1.23	.64
HT	28	1.45	.74
IC	22	1.34	.69
IE	13	1.11	.59
IG	19	1.28	.67
IL	23	1.36	.70
IN	75	1.88	.92
IO	41	1.61	.80
IR	27	1.43	.73
IS	35	1.54	.78
IT	27	1.43	.73
IV	25	1.40	.72
IX	15	1.18	.62
LA	28	1.45	.74
LE	37	1.57	.79
LI	20	1.30	.67
LL	27	1.43	.73
LO	13	1.11	.59

	F	$L_{10}(F)$	$L_{224}(2F)$
MA	36	1.56	.78
ME	26	1.41	.72
NA	26	1.41	.72
NC	19	1.28	.67
ND	52	1.72	.85
NE	57	1.76	.87
NG	27	1.43	.73
NI	30	1.48	.75
NO	18	1.26	.66
NS	24	1.38	.71
NT	82	1.91	.93
OF	25	1.40	.72
OL	19	1.28	.67
OM	25	1.40	.72
ON	77	1.89	.92
OP	25	1.40	.72
OR	64	1.81	.89
OS	14	1.15	.61
OT	19	1.28	.67
OU	37	1.57	.79
PA	14	1.15	.61
PE	23	1.36	.70
PO	17	1.23	.64
PR	18	1.26	.66
QU	15	1.18	.62
RA	39	1.59	.80
RD	17	1.23	.64
RE	98	1.99	.96
RI	30	1.48	.75
RO	28	1.45	.74

	F	$L_{10}(F)$	$L_{224}(2F)$
RS	31	1.49	.75
RT	42	1.62	.81
SA	24	1.38	.71
SE	49	1.69	.84
SH	26	1.41	.72
SI	34	1.53	.77
SO	15	1.18	.62
SS	19	1.28	.67
ST	63	1.80	.88
TA	28	1.45	.74
TE	71	1.85	.91
TH	78	1.89	.92
TI	45	1.65	.82
TO	50	1.70	.84
TR	17	1.23	.64
TS	19	1.28	.67
TT	19	1.28	.67
TW	36	1.56	.78
TY	41	1.61	.80
UN	21	1.32	.68
UR	31	1.49	.75
VE	57	1.76	.87
WE	22	1.34	.69
WO	19	1.28	.67
YT	15	1.18	.62
	3,745		

TABLE 7-D, Concluded.—*The 122 digraphs composing 75% of the 5,000 digraphs of Table 6-A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their initial letters*

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES

	F	$L_{10}(F)$	$L_{224}(2F)$
AN----	64	1.81	.89
AT----	47	1.67	.83
AR----	44	1.64	.82
AS----	41	1.61	.80
AL----	32	1.51	.76
AD----	27	1.43	.73
AI----	17	1.23	.64
AC----	14	1.15	.61
AM----	14	1.15	.61
AU----	13	1.11	.59
BE----	18	1.26	.66
CO----	41	1.61	.80
CE----	32	1.51	.76
CA----	20	1.30	.67
CH----	14	1.15	.61
CT----	14	1.15	.61
DE----	33	1.52	.77
DA----	32	1.51	.76
DI----	27	1.43	.73
DO----	16	1.20	.63
DT----	15	1.18	.62
DS----	13	1.11	.59
EN----	111	2.05	.99
ER----	87	1.94	.94
ED----	60	1.78	.88
ES----	54	1.73	.86
EE----	42	1.62	.81
ET----	37	1.57	.79
EA----	35	1.54	.78
EC----	32	1.51	.76
EL----	29	1.46	.74

	F	$L_{10}(F)$	$L_{224}(2F)$
EI----	27	1.43	.73
EP----	20	1.30	.67
EV----	20	1.30	.67
EF----	18	1.26	.66
EM----	14	1.15	.61
FO----	40	1.60	.80
FI----	39	1.59	.80
GH----	20	1.30	.67
GE----	14	1.15	.61
HI----	33	1.52	.77
HT----	28	1.45	.74
HA----	20	1.30	.67
HE----	20	1.30	.67
HO----	20	1.30	.67
HR----	17	1.23	.64
IN----	75	1.88	.92
IO----	41	1.61	.80
IS----	35	1.54	.78
IR----	27	1.43	.73
IT----	27	1.43	.73
IV----	25	1.40	.72
IL----	23	1.36	.70
IC----	22	1.34	.69
IG----	19	1.28	.67
IX----	15	1.18	.62
IE----	13	1.11	.59
LE----	37	1.57	.79
LA----	28	1.45	.74
LL----	27	1.43	.73
LI----	20	1.30	.67
LO----	13	1.11	.59

	F	$L_{10}(F)$	$L_{224}(2F)$
MA----	36	1.56	.78
ME----	26	1.41	.72
NT----	82	1.91	.93
NE----	57	1.76	.87
ND----	52	1.72	.85
NI----	30	1.48	.75
NG----	27	1.43	.73
NA----	26	1.41	.72
NS----	24	1.38	.71
NC----	19	1.28	.67
NO----	18	1.26	.66
ON----	77	1.89	.92
OR----	64	1.81	.89
OU----	37	1.57	.79
OF----	25	1.40	.72
OM----	25	1.40	.72
OP----	25	1.40	.72
OL----	19	1.28	.67
OT----	19	1.28	.67
OS----	14	1.15	.61
PE----	23	1.36	.70
PR----	18	1.26	.66
PO----	17	1.23	.64
PA----	14	1.15	.61
QU----	15	1.18	.62
RE----	98	1.99	.96
RT----	42	1.62	.81
RA----	39	1.59	.80
RS----	31	1.49	.75

	F	$L_{10}(F)$	$L_{224}(2F)$
RI----	30	1.48	.75
RO----	28	1.45	.74
RD----	17	1.23	.64
ST----	63	1.80	.88
SE----	49	1.69	.84
SI----	34	1.53	.77
SH----	26	1.41	.72
SA----	24	1.38	.71
SS----	19	1.28	.67
SO----	15	1.18	.62
TH----	78	1.89	.92
TE----	71	1.85	.91
TO----	50	1.70	.84
TI----	45	1.65	.82
TY----	41	1.61	.80
TW----	36	1.56	.78
TA----	28	1.45	.74
TS----	19	1.28	.67
TT----	19	1.28	.67
TR----	17	1.23	.64
UR----	31	1.49	.75
UN----	21	1.32	.68
VE----	57	1.76	.87
WE----	22	1.34	.69
WO----	19	1.28	.67
YT----	15	1.18	.62
	3,745		

TABLE 7-E.—*All the 428 digraphs of Table 6-A, arranged first alphabetically according to their initial letters and then alphabetically according to their final letters.*

(SEE TABLE 6-A.—READ ACROSS THE ROWS)

TABLE 8.—*The 428 different digraphs of Table 6–A, arranged first alphabetically according to their initial letters and then according to their absolute frequencies under each initial letter,*[1] *accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$
AN	64	1.81	.89
AT	47	1.67	.83
AR	44	1.64	.82
AS	41	1.61	.80
AL	32	1.51	.76
AD	27	1.43	.73
AI	17	1.23	.64
AC	14	1.15	.61
AM	14	1.15	.61
AU	13	1.11	.59
AP	12	1.08	.58
AY	12	1.08	.58
AV	7	0.85	.48
AB	6	0.78	.45
AG	6	0.78	.45
AF	4	0.60	.38
AA	3	0.48	.33
AW	3	0.48	.33
AH	2	0.30	.25
AK	2	0.30	.25
AO	2	0.30	.25
AE	1	0.00	.13
AJ	1	0.00	.13
BE	18	1.26	.66
BY	7	0.85	.48
BL	6	0.78	.45
BA	4	0.60	.38
BO	4	0.60	.38
BI	2	0.30	.25
BR	2	0.30	.25
BU	2	0.30	.25
BJ	1	0.00	.13
BM	1	0.00	.13
BS	1	0.00	.13
BT	1	0.00	.13
CO	41	1.61	.80
CE	32	1.51	.76
CA	20	1.30	.67
CH	14	1.15	.61

	F	$L_{10}(F)$	$L_{224}(2F)$
CT	14	1.15	.61
CI	7	0.85	.48
CL	5	0.70	.42
CK	4	0.60	.38
CR	4	0.60	.38
CU	4	0.60	.38
CC	3	0.48	.33
CD	1	0.00	.13
CF	1	0.00	.13
CM	1	0.00	.13
CN	1	0.00	.13
CS	1	0.00	.13
CW	1	0.00	.13
CY	1	0.00	.13
DE	33	1.52	.77
DA	32	1.51	.76
DI	27	1.43	.73
DO	16	1.20	.63
DT	15	1.18	.62
DS	13	1.11	.59
DR	12	1.08	.58
DD	8	0.90	.51
DF	8	0.90	.51
DM	5	0.70	.42
DP	5	0.70	.42
DU	5	0.70	.42
DB	4	0.60	.38
DC	4	0.60	.38
DN	4	0.60	.38
DW	4	0.60	.38
DL	3	0.48	.33
DV	3	0.48	.33
DG	2	0.30	.25
DH	2	0.30	.25
DQ	2	0.30	.25
DJ	1	0.00	.13
DY	1	0.00	.13
EN	111	2.05	.99
ER	87	1.94	.94

	F	$L_{10}(F)$	$L_{224}(2F)$
ED	60	1.78	.88
ES	54	1.73	.86
EE	42	1.62	.81
ET	37	1.57	.79
EA	35	1.54	.78
EC	32	1.51	.76
EL	29	1.46	.74
EI	27	1.43	.73
EP	20	1.30	.67
EV	20	1.30	.67
EF	18	1.26	.66
EM	14	1.15	.61
EO	12	1.08	.58
EQ	12	1.08	.58
EH	7	0.85	.48
EW	7	0.85	.48
EX	7	0.85	.48
EB	4	0.60	.38
EG	4	0.60	.38
EY	4	0.60	.38
EU	3	0.48	.33
EJ	1	0.00	.13
EZ	1	0.00	.13
FO	40	1.60	.80
FI	39	1.59	.80
FF	11	1.04	.56
FT	11	1.04	.56
FE	10	1.00	.55
FR	9	0.95	.53
FA	5	0.70	.42
FS	3	0.48	.33
FU	3	0.48	.33
FC	2	0.30	.25
FL	2	0.30	.25
FD	1	0.00	.13
FG	1	0.00	.13
FM	1	0.00	.13
FP	1	0.00	.13
FW	1	0.00	.13
FY	1	0.00	.13

	F	$L_{10}(F)$	$L_{224}(2F)$
GH	20	1.30	.67
GE	14	1.15	.61
GA	7	0.85	.48
GO	6	0.78	.45
GI	5	0.70	.42
GR	5	0.70	.42
GT	4	0.60	.38
GN	3	0.48	.33
GS	3	0.48	.33
GC	2	0.30	.25
GF	2	0.30	.25
GL	2	0.30	.25
GP	2	0.30	.25
GU	2	0.30	.25
GD	1	0.00	.13
GG	1	0.00	.13
GJ	1	0.00	.13
GM	1	0.00	.13
GW	1	0.00	.13
HI	33	1.52	.77
HT	28	1.45	.74
HA	20	1.30	.67
HE	20	1.30	.67
HO	20	1.30	.67
HR	17	1.23	.64
HU	8	0.90	.51
HF	5	0.70	.42
HS	4	0.60	.38
HC	3	0.48	.33
HN	3	0.48	.33
HD	2	0.30	.25
HM	2	0.30	.25
HB	1	0.00	.13
HL	1	0.00	.13
HP	1	0.00	.13
HQ	1	0.00	.13
HW	1	0.00	.13
HY	1	0.00	.13

[1] For arrangement alphabetically first under initial letters and then under final letters, see Table 6–A.

TABLE 8, Contd.—*The 428 different digraphs of Table 6-A, arranged first alphabetically according to their initial letters and then according to their absolute frequencies under each initial letter,*[1] *accompanied by the logarithms of their assigned probabilities*

	F	Lg(F)	Lg/(2F)
IN....	75	1.88	.92
IO....	41	1.61	.80
IS....	35	1.54	.78
IR....	27	1.43	.73
IT....	27	1.43	.73
IV....	25	1.40	.72
IL....	23	1.36	.70
IC....	22	1.34	.69
IG....	19	1.28	.67
IX....	15	1.18	.62
IE....	13	1.11	.59
IF....	10	1.00	.55
IM....	9	0.95	.53
IA....	8	0.90	.51
IP....	7	0.85	.48
ID....	6	0.78	.45
IB....	2	0.30	.25
IK....	2	0.30	.25
IZ....	2	0.30	.25
JE....	2	0.30	.25
JO....	2	0.30	.25
JU....	2	0.30	.25
JA....	1	0.00	.13
KE....	6	0.78	.45
KI....	2	0.30	.25
KA....	1	0.00	.13
KC....	1	0.00	.13
KL....	1	0.00	.13
KN....	1	0.00	.13
KS....	1	0.00	.13
LE....	37	1.57	.79
LA....	28	1.45	.74
LL....	27	1.43	.73
LI....	20	1.30	.67
LO....	13	1.11	.59
LY....	10	1.00	.55
LD....	9	0.95	.53
LT....	8	0.90	.51
LS....	6	0.78	.45
LB....	3	0.48	.33
LC....	3	0.48	.33
LF....	3	0.48	.33
LP....	3	0.48	.33
LM....	2	0.30	.25
LR....	2	0.30	.25
LU....	2	0.30	.25
LV....	2	0.30	.25
LW....	2	0.30	.25
LG....	1	0.00	.13
LH....	1	0.00	.13
LN....	1	0.00	.13
MA....	36	1.56	.78
ME....	26	1.41	.72
MM....	13	1.11	.59
MO....	10	1.00	.55
MI....	9	0.95	.53
MP....	8	0.90	.51
MB....	6	0.78	.45
MS....	4	0.60	.38
MC....	3	0.48	.33
MR....	2	0.30	.25
MT....	2	0.30	.25
MU....	2	0.30	.25
MY....	2	0.30	.25
MD....	1	0.00	.13
MF....	1	0.00	.13
MH....	1	0.00	.13
NT....	82	1.91	.93
NE....	57	1.76	.87
ND....	52	1.72	.85
NI....	30	1.48	.75
NG....	27	1.43	.73
NA....	26	1.41	.72
NS....	24	1.38	.71
NC....	19	1.28	.67
NO....	18	1.26	.66
NF....	9	0.95	.53
NN....	8	0.90	.51
NU....	7	0.85	.48
NL....	5	0.70	.42
NM....	5	0.70	.42
NY....	5	0.70	.42
NH....	4	0.60	.38
NR....	4	0.60	.38
NP....	3	0.48	.33
NV....	3	0.48	.33
NW....	3	0.48	.33
NB....	2	0.30	.25
NK....	2	0.30	.25
NJ....	1	0.00	.13
NQ....	1	0.00	.13
ON....	77	1.89	.92
OR....	64	1.81	.89
OU....	37	1.57	.79
OF....	25	1.40	.72
OM....	25	1.40	.72
OP....	25	1.40	.72
OL....	19	1.28	.67
OT....	19	1.28	.67
OS....	14	1.15	.61
OD....	12	1.08	.58
OC....	8	0.90	.51
OW....	8	0.90	.51
OA....	7	0.85	.48
OV....	7	0.85	.48
OO....	6	0.78	.45
OI....	5	0.70	.42
OB....	4	0.60	.38
OE....	3	0.48	.33
OH....	3	0.48	.33
OG....	2	0.30	.25
OK....	2	0.30	.25
OY....	2	0.30	.25
OJ....	1	0.00	.13
OX....	1	0.00	.13
PE....	23	1.36	.70
PR....	18	1.26	.66
PO....	17	1.23	.64
PA....	14	1.15	.61
PL....	13	1.11	.59
PP....	11	1.04	.56
PT....	8	0.90	.51
PI....	6	0.78	.45
PS....	6	0.78	.45
PM....	4	0.60	.38
PH....	3	0.48	.33
PU....	3	0.48	.33
PF....	2	0.30	.25
PB....	1	0.00	.13
PC....	1	0.00	.13
PD....	1	0.00	.13
PN....	1	0.00	.13
PV....	1	0.00	.13
PW....	1	0.00	.13
PY....	1	0.00	.13
QU....	15	1.18	.62
QM....	1	0.00	.13
QR....	1	0.00	.13

[1] For arrangement alphabetically first under initial letters and then under final letters, see Table 6-A.

TABLE 8, Concluded.—*The 428 different digraphs of Table 6–A, arranged first alphabetically according to their initial letters and then according to their absolute frequencies under each initial letter,*[1] *accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
RE	98	1.99	.96	SR	5	0.70	.42	US	12	1.08	.58	XI	2	0.30	.25
RT	42	1.62	.81	SN	4	0.60	.38	UT	12	1.08	.58	XP	2	0.30	.25
RA	39	1.59	.80	SW	4	0.60	.38	UE	11	1.04	.56	XD	1	0.00	.13
RS	31	1.49	.75	SB	3	0.48	.33	UG	8	0.90	.51	XE	1	0.00	.13
RI	30	1.48	.75	SM	3	0.48	.33	UL	6	0.78	.45	XF	1	0.00	.13
RO	28	1.45	.74	SG	2	0.30	.25	UA	5	0.70	.42	XH	1	0.00	.13
RD	17	1.23	.64	SL	2	0.30	.25	UI	5	0.70	.42	XN	1	0.00	.13
RP	13	1.11	.59	SK	1	0.00	.13	UM	5	0.70	.42	XO	1	0.00	.13
RR	11	1.04	.56	SV	1	0.00	.13	UB	3	0.48	.33	XR	1	0.00	.13
RC	9	0.95	.53	SY	1	0.00	.13	UC	3	0.48	.33	XS	1	0.00	.13
RM	9	0.95	.53					UD	3	0.48	.33				
RY	9	0.95	.53	TH	78	1.89	.92	UP	2	0.30	.25	YT	15	1.18	.62
RG	7	0.85	.48	TE	71	1.85	.91	UF	1	0.00	.13	YF	11	1.04	.56
RN	7	0.85	.48	TO	50	1.70	.84	UO	1	0.00	.13	YS	11	1.04	.56
RF	6	0.78	.45	TI	45	1.65	.82	UV	1	0.00	.13	YO	10	1.00	.55
RL	5	0.70	.42	TY	41	1.61	.80					YE	9	0.95	.53
RU	5	0.70	.42	TW	36	1.56	.78	VE	57	1.76	.87	YA	6	0.78	.45
RV	5	0.70	.42	TA	28	1.45	.74	VI	12	1.08	.58	YN	6	0.78	.45
RW	4	0.60	.38	TS	19	1.28	.67	VA	6	0.78	.45	YC	4	0.60	.38
RH	3	0.48	.33	TT	19	1.28	.67	VO	1	0.00	.13	YD	4	0.60	.38
RB	2	0.30	.25	TR	17	1.23	.64	VT	1	0.00	.13	YR	4	0.60	.38
RJ	1	0.00	.13	TF	7	0.85	.48					YI	3	0.48	.33
RK	1	0.00	.13	TN	7	0.85	.48	WE	22	1.34	.69	YP	3	0.48	.33
				TC	6	0.78	.45	WO	19	1.28	.67	YB	2	0.30	.25
ST	63	1.80	.88	TD	6	0.78	.45	WI	13	1.11	.59	YL	2	0.30	.25
SE	49	1.69	.84	TM	6	0.78	.45	WA	12	1.08	.58	YM	2	0.30	.25
SI	34	1.53	.77	TL	5	0.70	.42	WH	4	0.60	.38	YG	1	0.00	.13
SH	26	1.41	.72	TU	5	0.70	.42	WN	2	0.30	.25	YH	1	0.00	.13
SA	24	1.38	.71	TB	3	0.48	.33	WL	1	0.00	.13	YU	1	0.00	.13
SS	19	1.28	.67	TP	2	0.30	.25	WR	1	0.00	.13	YW	1	0.00	.13
SO	15	1.18	.62	TG	1	0.00	.13	WS	1	0.00	.13				
SC	13	1.11	.59	TQ	1	0.00	.13	WY	1	0.00	.13	ZE	2	0.30	.25
SF	12	1.08	.58	TZ	1	0.00	.13					ZA	1	0.00	.13
SU	11	1.04	.56					XT	7	0.85	.48	ZI	1	0.00	.13
SP	10	1.00	.55	UR	31	1.49	.75	XA	2	0.30	.25		5,000		
SD	5	0.70	.42	UN	21	1.32	.68	XC	2	0.30	.25				

[1] For arrangement alphabetically first under initial letters and then under final letters, see Table 6–A.

TABLE 9-A.—*The 428 different digraphs of Table 6-A, arranged first alphabetically according to their final letters and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{424}(2F)$		F	$L_{10}(F)$	$L_{424}(2F)$		F	$L_{10}(F)$	$L_{424}(2F)$		F	$L_{10}(F)$	$L_{424}(2F)$
RA	39	1.59	.80	EC	32	1.51	.76	RE	98	1.99	.96	GF	2	0.30	.25
MA	36	1.56	.78	IC	22	1.34	.69	TE	71	1.85	.91	PF	2	0.30	.25
EA	35	1.54	.78	NC	19	1.28	.67	NE	57	1.76	.87	CF	1	0.00	.13
DA	32	1.51	.76	AC	14	1.15	.61	VE	57	1.76	.87	MF	1	0.00	.13
LA	28	1.45	.74	SC	13	1.11	.59	SE	49	1.69	.84	UF	1	0.00	.13
TA	28	1.45	.74	RC	9	0.95	.53	EE	42	1.62	.81	XF	1	0.00	.13
NA	26	1.41	.72	OC	8	0.90	.51	LE	37	1.57	.79				
SA	24	1.38	.71	TC	6	0.78	.45	DE	33	1.52	.77				
CA	20	1.30	.67	DC	4	0.60	.38	CE	32	1.51	.76	NG	27	1.43	.73
HA	20	1.30	.67	YC	4	0.60	.38	ME	26	1.41	.72	IG	19	1.28	.67
PA	14	1.15	.61	CC	3	0.48	.33	PE	23	1.36	.70	UG	8	0.90	.51
WA	12	1.08	.58	HC	3	0.48	.33	WE	22	1.34	.69	RG	7	0.85	.48
IA	8	0.90	.51	LC	3	0.48	.33	HE	20	1.30	.67	AG	6	0.78	.45
GA	7	0.85	.48	MC	3	0.48	.33	BE	18	1.26	.66	EG	4	0.60	.38
OA	7	0.85	.48	UC	3	0.48	.33	GE	14	1.15	.61	DG	2	0.30	.25
VA	6	0.78	.45	FC	2	0.30	.25	IE	13	1.11	.59	OG	2	0.30	.25
YA	6	0.78	.45	GC	2	0.30	.25	UE	11	1.04	.56	SG	2	0.30	.25
FA	5	0.70	.42	XC	2	0.30	.25	FE	10	1.00	.55	FG	1	0.00	.13
UA	5	0.70	.42	KC	1	0.00	.13	YE	9	0.95	.53	GG	1	0.00	.13
BA	4	0.60	.38	PC	1	0.00	.13	KE	6	0.78	.45	LG	1	0.00	.13
AA	3	0.48	.33					OE	3	0.48	.33	TG	1	0.00	.13
XA	2	0.30	.25					JE	2	0.30	.25	YG	1	0.00	.13
JA	1	0.00	.13	ED	60	1.78	.88	ZE	2	0.30	.25				
KA	1	0.00	.13	ND	52	1.72	.85	AE	1	0.00	.13				
ZA	1	0.00	.13	AD	27	1.43	.73	XE	1	0.00	.13				
				RD	17	1.23	.64					TH	78	1.89	.92
AB	6	0.78	.45	OD	12	1.08	.58					SH	26	1.41	.72
MB	6	0.78	.45	LD	9	0.95	.53					GH	20	1.30	.67
DB	4	0.60	.38	DD	8	0.90	.51	OF	25	1.40	.72	CH	14	1.15	.61
EB	4	0.60	.38	ID	6	0.78	.45	EF	18	1.26	.66	EH	7	0.85	.48
OB	4	0.60	.38	TD	6	0.78	.45	SF	12	1.08	.58	NH	4	0.60	.38
LB	3	0.48	.33	SD	5	0.70	.42	FF	11	1.04	.56	WH	4	0.60	.38
SB	3	0.48	.33	YD	4	0.60	.38	YF	11	1.04	.56	OH	3	0.48	.33
TB	3	0.48	.33	UD	3	0.48	.33	IF	10	1.00	.55	PH	3	0.48	.33
UB	3	0.48	.33	HD	2	0.30	.25	NF	9	0.95	.53	RH	3	0.48	.33
IB	2	0.30	.25	CD	1	0.00	.13	DF	8	0.90	.51	AH	2	0.30	.25
NB	2	0.30	.25	FD	1	0.00	.13	TF	7	0.85	.48	DH	2	0.30	.25
RB	2	0.30	.25	GD	1	0.00	.13	RF	6	0.78	.45	LH	1	0.00	.13
YB	2	0.30	.25	MD	1	0.00	.13	HF	5	0.70	.42	MH	1	0.00	.13
HB	1	0.00	.13	PD	1	0.00	.13	AF	4	0.60	.38	XH	1	0.00	.13
PB	1	0.00	.13	XD	1	0.00	.13	LF	3	0.48	.33	YH	1	0.00	.13

RESTRICTED

TABLE 9-A, Contd.—*The 428 different digraphs of Table 6-A, arranged first alphabetically according to their final letters and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	L_{224} (2F)
TI	45	1.65	.82
FI	39	1.59	.80
SI	34	1.53	.77
HI	33	1.52	.77
NI	30	1.48	.75
RI	30	1.48	.75
DI	27	1.43	.73
EI	27	1.43	.73
LI	20	1.30	.67
AI	17	1.23	.64
WI	13	1.11	.59
VI	12	1.08	.58
MI	9	0.95	.53
CI	7	0.85	.48
PI	6	0.78	.45
GI	5	0.70	.42
OI	5	0.70	.42
UI	5	0.70	.42
YI	3	0.48	.33
BI	2	0.30	.25
KI	2	0.30	.25
XI	2	0.30	.25
ZI	1	0.00	.13
AJ	1	0.00	.13
BJ	1	0.00	.13
DJ	1	0.00	.13
EJ	1	0.00	.13
GJ	1	0.00	.13
NJ	1	0.00	.13
OJ	1	0.00	.13
RJ	1	0.00	.13
CK	4	0.60	.38
AK	2	0.30	.25
IK	2	0.30	.25
NK	2	0.30	.25
OK	2	0.30	.25
RK	1	0.00	.13
SK	1	0.00	.13
AL	32	1.51	.76
EL	29	1.46	.74

	F	$L_{10}(F)$	L_{224} (2F)
LL	27	1.43	.73
IL	23	1.36	.70
OL	19	1.28	.67
PL	13	1.11	.59
BL	6	0.78	.45
UL	6	0.78	.45
CL	5	0.70	.42
NL	5	0.70	.42
RL	5	0.70	.42
TL	5	0.70	.42
DL	3	0.48	.33
FL	2	0.30	.25
GL	2	0.30	.25
SL	2	0.30	.25
YL	2	0.30	.25
HL	1	0.00	.13
KL	1	0.00	.13
WL	1	0.00	.13
OM	25	1.40	.72
AM	14	1.15	.61
EM	14	1.15	.61
MM	13	1.11	.59
IM	9	0.95	.53
RM	9	0.95	.53
TM	6	0.78	.45
DM	5	0.70	.42
NM	5	0.70	.42
UM	5	0.70	.42
PM	4	0.60	.38
SM	3	0.48	.33
HM	2	0.30	.25
LM	2	0.30	.25
YM	2	0.30	.25
BM	1	0.00	.13
CM	1	0.00	.13
FM	1	0.00	.13
GM	1	0.00	.13
QM	1	0.00	.13
EN	111	2.05	.99
ON	77	1.89	.92
IN	75	1.88	.92

	F	$L_{10}(F)$	L_{224} (2F)
AN	64	1.81	.89
UN	21	1.32	.68
NN	8	0.90	.51
RN	7	0.85	.48
TN	7	0.85	.48
YN	6	0.78	.45
DN	4	0.60	.38
SN	4	0.60	.38
GN	3	0.48	.33
HN	3	0.48	.33
WN	2	0.30	.25
CN	1	0.00	.13
KN	1	0.00	.13
LN	1	0.00	.13
PN	1	0.00	.13
XN	1	0.00	.13
TO	50	1.70	.84
CO	41	1.61	.80
IO	41	1.61	.80
FO	40	1.60	.80
RO	28	1.45	.74
HO	20	1.30	.67
WO	19	1.28	.67
NO	18	1.26	.66
PO	17	1.23	.64
DO	16	1.20	.63
SO	15	1.18	.62
LO	13	1.11	.59
EO	12	1.08	.58
MO	10	1.00	.55
YO	10	1.00	.55
GO	6	0.78	.45
OO	6	0.78	.45
BO	4	0.60	.38
AO	2	0.30	.25
JO	2	0.30	.25
UO	1	0.00	.13
VO	1	0.00	.13
XO	1	0.00	.13
OP	25	1.40	.72
EP	20	1.30	.67

	F	$L_{10}(F)$	L_{224} (2F)
RP	13	1.11	.59
AP	12	1.08	.58
PP	11	1.04	.56
SP	10	1.00	.55
MP	8	0.90	.51
IP	7	0.85	.48
DP	5	0.70	.42
LP	3	0.48	.33
NP	3	0.48	.33
YP	3	0.48	.33
GP	2	0.30	.25
TP	2	0.30	.25
UP	2	0.30	.25
XP	2	0.30	.25
FP	1	0.00	.13
HP	1	0.00	.13
EQ	12	1.08	.58
DQ	2	0.30	.25
HQ	1	0.00	.13
NQ	1	0.00	.13
TQ	1	0.00	.13
ER	87	1.94	.94
OR	64	1.81	.89
AR	44	1.64	.82
UR	31	1.49	.75
IR	27	1.43	.73
PR	18	1.26	.66
HR	17	1.23	.64
TR	17	1.23	.64
DR	12	1.08	.58
RR	11	1.04	.56
FR	9	0.95	.53
GR	5	0.70	.42
SR	5	0.70	.42
CR	4	0.60	.38
NR	4	0.60	.38
YR	4	0.60	.38
BR	2	0.30	.25
LR	2	0.30	.25
MR	2	0.30	.25
QR	1	0.00	.13
WR	1	0.00	.13
XR	1	0.00	.13

RESTRICTED

TABLE 9-A, Concluded.—*The 428 different digraphs of Table 6–A, arranged first alphabetically according to their final letters and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	L_{10}(F)	L_{224}(2F)		F	L_{10}(F)	L_{224}(2F)		F	L_{10}(F)	L_{224}(2F)		F	L_{10}(F)	L_{224}(2F)
ES	54	1.73	.86	OT	19	1.28	.67	JU	2	0.30	.25	PW	1	0.00	.13
AS	41	1.61	.80	TT	19	1.28	.67	LU	2	0.30	.25	YW	1	0.00	.13
IS	35	1.54	.78	DT	15	1.18	.62	MU	2	0.30	.25				
RS	31	1.49	.75	YT	15	1.18	.62	YU	1	0.00	.13	IX	15	1.18	.62
NS	24	1.38	.71	CT	14	1.15	.61					EX	7	0.85	.48
SS	19	1.28	.67	UT	12	1.08	.58	IV	25	1.40	.72	OX	1	0.00	.13
TS	19	1.28	.67	FT	11	1.04	.56	EV	20	1.30	.67				
OS	14	1.15	.61	LT	8	0.90	.51	AV	7	0.85	.48	TY	41	1.61	.80
DS	13	1.11	.59	PT	8	0.90	.51	OV	7	0.85	.48	AY	12	1.08	.58
US	12	1.08	.58	XT	7	0.85	.48	RV	5	0.70	.42	LY	10	1.00	.55
YS	11	1.04	.56	GT	4	0.60	.38	DV	3	0.48	.33	RY	9	0.95	.53
LS	6	0.78	.45	MT	2	0.30	.25	NV	3	0.48	.33	BY	7	0.85	.48
PS	6	0.78	.45	BT	1	0.00	.13	LV	2	0.30	.25	NY	5	0.70	.42
HS	4	0.60	.38	VT	1	0.00	.13	PV	1	0.00	.13	EY	4	0.60	.38
MS	4	0.60	.38					SV	1	0.00	.13	MY	2	0.30	.25
FS	3	0.48	.33	OU	37	1.57	.79	UV	1	0.00	.13	OY	2	0.30	.25
GS	3	0.48	.33	QU	15	1.18	.62					CY	1	0.00	.13
BS	1	0.00	.13	AU	13	1.11	.59	TW	36	1.56	.78	DY	1	0.00	.13
CS	1	0.00	.13	SU	11	1.04	.56	OW	8	0.90	.51	FY	1	0.00	.13
KS	1	0.00	.13	HU	8	0.90	.51	EW	7	0.85	.48	HY	1	0.00	.13
WS	1	0.00	.13	NU	7	0.85	.48	DW	4	0.60	.38	PY	1	0.00	.13
XS	1	0.00	.13	DU	5	0.70	.42	RW	4	0.60	.38	SY	1	0.00	.13
				RU	5	0.70	.42	SW	4	0.60	.38	WY	1	0.00	.13
NT	82	1.91	.93	TU	5	0.70	.42	AW	3	0.48	.33				
ST	63	1.80	.88	CU	4	0.60	.38	NW	3	0.48	.33	IZ	2	0.30	.25
AT	47	1.67	.83	EU	3	0.48	.33	LW	2	0.30	.25	EZ	1	0.00	.13
RT	42	1.62	.81	FU	3	0.48	.33	CW	1	0.00	.13	TZ	1	0.00	.13
ET	37	1.57	.79	PU	3	0.48	.33	FW	1	0.00	.13				
HT	28	1.45	.74	BU	2	0.30	.25	GW	1	0.00	.13				
IT	27	1.43	.73	GU	2	0.30	.25	HW	1	0.00	.13		5,000		

~~RESTRICTED~~

TABLE 9-B.—*The 18 digraphs composing 25% of the 5,000 digraphs of Table 6-A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their final letters*

(1) AND ACCORDING TO THEIR INITIAL LETTERS

	F	$L_{10}(F)$	$L_{224}(2F)$
ED	60	1.78	.88
ND	52	1.72	.85
NE	57	1.76	.87
RE	98	1.99	.96
SE	49	1.69	.84
TE	71	1.85	.91
VE	57	1.76	.87
TH	78	1.89	.92
AN	64	1.81	.89
EN	111	2.05	.99

	F	$L_{10}(F)$	$L_{224}(2F)$
IN	75	1.88	.92
ON	77	1.89	.92
TO	50	1.70	.84
ER	87	1.94	.94
OR	64	1.81	.89
ES	54	1.73	.86
NT	82	1.91	.93
ST	63	1.80	.88
	1,249		

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES

	F	$L_{10}(F)$	$L_{224}(2F)$
ED	60	1.78	.88
ND	52	1.72	.85
RE	98	1.99	.96
TE	71	1.85	.91
NE	57	1.76	.87
VE	57	1.76	.87
SE	49	1.69	.84
TH	78	1.89	.92
EN	111	2.05	.99
ON	77	1.89	.92

	F	$L_{10}(F)$	$L_{224}(2F)$
IN	75	1.88	.92
AN	64	1.81	.89
TO	50	1.70	.84
ER	87	1.94	.94
OR	64	1.81	.89
ES	54	1.73	.86
NT	82	1.91	.93
ST	63	1.80	.88
	1,249		

TABLE 9-C.—*The 53 digraphs composing 50% of the 5,000 digraphs of Table 6-A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their final letters*

(1) AND ACCORDING TO THEIR INITIAL LETTERS

	F	$L_{10}(F)$	$L_{224}(2F)$
DA	32	1.51	.76
EA	35	1.54	.78
LA	28	1.45	.74
MA	36	1.56	.78
RA	39	1.59	.80
TA	28	1.45	.74
EC	32	1.51	.76
ED	60	1.78	.88
ND	52	1.72	.85
CE	32	1.51	.76
DE	33	1.52	.77
EE	42	1.62	.81
LE	37	1.57	.79

	F	$L_{10}(F)$	$L_{224}(2F)$
NE	57	1.76	.87
RE	98	1.99	.96
SE	49	1.69	.84
TE	71	1.85	.91
VE	57	1.76	.87
TH	78	1.89	.92
FI	39	1.59	.80
HI	33	1.52	.77
NI	30	1.48	.75
RI	30	1.48	.75
SI	34	1.53	.77
TI	45	1.65	.82
AL	32	1.51	.76
EL	29	1.46	.74

	F	$L_{10}(F)$	$L_{224}(2F)$
AN	64	1.81	.89
EN	111	2.05	.99
IN	75	1.88	.92
ON	77	1.89	.92
CO	41	1.61	.80
FO	40	1.60	.80
IO	41	1.61	.80
RO	28	1.45	.74
TO	50	1.70	.84
AR	44	1.64	.82
ER	87	1.94	.94
OR	64	1.81	.89
UR	31	1.49	.75

	F	$L_{10}(F)$	$L_{224}(2F)$
AS	41	1.61	.80
ES	54	1.73	.86
IS	35	1.54	.78
RS	31	1.49	.75
AT	47	1.67	.83
ET	37	1.57	.79
HT	28	1.45	.74
NT	82	1.91	.93
RT	42	1.62	.81
ST	63	1.80	.88
OU	37	1.57	.79
TW	36	1.56	.78
TY	41	1.61	.80
	2,495		

~~RESTRICTED~~

TABLE 9-C, Concluded.—*The 53 digraphs composing 50% of the 5,000 digraphs of Table 6–A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their final letters*

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES

	F	$L_{10}(F)$	$L_{224}(2F)$
RA	39	1.59	.80
MA	36	1.56	.78
EA	35	1.54	.78
DA	32	1.51	.76
LA	28	1.45	.74
TA	28	1.45	.74
EC	32	1.51	.76
ED	60	1.78	.88
ND	52	1.72	.85
RE	98	1.99	.96
TE	71	1.85	.91
NE	57	1.76	.87
VE	57	1.76	.87
SE	49	1.69	.84

	F	$L_{10}(F)$	$L_{224}(2F)$
EE	42	1.62	.81
LE	37	1.57	.79
DE	33	1.52	.77
CE	32	1.51	.76
TH	78	1.89	.92
TI	45	1.65	.82
FI	39	1.59	.80
SI	34	1.53	.77
HI	33	1.52	.77
NI	30	1.48	.75
RI	30	1.48	.75
AL	32	1.51	.76
EL	29	1.46	.74

	F	$L_{10}(F)$	$L_{224}(2F)$
EN	111	2.05	.99
ON	77	1.89	.92
IN	75	1.88	.92
AN	64	1.81	.89
TO	50	1.70	.84
CO	41	1.61	.80
IO	41	1.61	.80
FO	40	1.60	.80
RO	28	1.45	.74
ER	87	1.94	.94
OR	64	1.81	.89
AR	44	1.64	.82
UR	31	1.49	.75

	F	$L_{10}(F)$	$L_{224}(2F)$
ES	54	1.73	.86
AS	41	1.61	.80
IS	35	1.54	.78
RS	31	1.49	.75
NT	82	1.91	.93
ST	63	1.80	.88
AT	47	1.67	.83
RT	42	1.62	.81
ET	37	1.57	.79
HT	28	1.45	.74
OU	37	1.57	.79
TW	36	1.56	.78
TY	41	1.61	.80
	2,495		

TABLE 9-D.—*The 122 digraphs composing 75% of the 5,000 digraphs of Table 6–A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their final letters*

(1) AND ACCORDING TO THEIR INITIAL LETTERS

	F	$L_{10}(F)$	$L_{224}(2F)$
CA	20	1.30	.67
DA	32	1.51	.76
EA	35	1.54	.78
HA	20	1.30	.67
LA	28	1.45	.74
MA	36	1.56	.78
NA	26	1.41	.72
PA	14	1.15	.61
RA	39	1.59	.80
SA	24	1.38	.71
TA	28	1.45	.74
AC	14	1.15	.61
EC	32	1.51	.76
IC	22	1.34	.69
NC	19	1.28	.67
AD	27	1.43	.73
ED	60	1.78	.88

	F	$L_{10}(F)$	$L_{224}(2F)$
ND	52	1.72	.85
RD	17	1.23	.64
BE	18	1.26	.66
CE	32	1.51	.76
DE	33	1.52	.77
EE	42	1.62	.81
GE	14	1.15	.61
HE	20	1.30	.67
IE	13	1.11	.59
LE	37	1.57	.79
ME	26	1.41	.72
NE	57	1.76	.87
PE	23	1.36	.70
RE	98	1.99	.96
SE	49	1.69	.84
TE	71	1.85	.91
VE	57	1.76	.87
WE	22	1.34	.69

	F	$L_{10}(F)$	$L_{224}(2F)$
EF	18	1.26	.66
OF	25	1.40	.72
IG	19	1.28	.67
NG	27	1.43	.73
CH	14	1.15	.61
GH	20	1.30	.67
SH	26	1.41	.72
TH	78	1.89	.92
AI	17	1.23	.64
DI	27	1.43	.73
EI	27	1.43	.73
FI	39	1.59	.80
HI	33	1.52	.77
LI	20	1.30	.67
NI	30	1.48	.75
RI	30	1.48	.75

	F	$L_{10}(F)$	$L_{224}(2F)$
SI	34	1.53	.77
TI	45	1.65	.82
AL	32	1.51	.76
EL	29	1.46	.74
IL	23	1.36	.70
LL	27	1.43	.73
OL	19	1.28	.67
AM	14	1.15	.61
EM	14	1.15	.61
OM	25	1.40	.72
AN	64	1.81	.89
EN	111	2.05	.99
IN	75	1.88	.92
ON	77	1.89	.92
UN	21	1.32	.68

TABLE 9-D, Contd.—*The 122 digraphs composing 75% of the 5,000 digraphs of Table 6-A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their final letters*

(1) AND ACCORDING TO THEIR INITIAL LETTERS—Concluded

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
CO	41	1.61	.80	AR	44	1.64	.82	RS	31	1.49	.75	TT	19	1.28	.67
DO	16	1.20	.63	ER	87	1.94	.94	SS	19	1.28	.67	YT	15	1.18	.62
FO	40	1.60	.80	HR	17	1.23	.64	TS	19	1.28	.67				
HO	20	1.30	.67	IR	27	1.43	.73					AU	13	1.11	.59
IO	41	1.61	.80	OR	64	1.81	.89					OU	37	1.57	.79
LO	13	1.11	.59	PR	18	1.26	.66	AT	47	1.67	.83	QU	15	1.18	.62
NO	18	1.26	.66	TR	17	1.23	.64	CT	14	1.15	.61	EV	20	1.30	.67
PO	17	1.23	.64	UR	31	1.49	.75	DT	15	1.18	.62	IV	25	1.40	.72
RO	28	1.45	.74					ET	37	1.57	.79				
SO	15	1.18	.62	AS	41	1.61	.80	HT	28	1.45	.74	TW	36	1.56	.78
TO	50	1.70	.84	DS	13	1.11	.59	IT	27	1.43	.73				
WO	19	1.28	.67	ES	54	1.73	.86	NT	82	1.91	.93	IX	15	1.18	.62
				IS	35	1.54	.78	OT	19	1.28	.67				
EP	20	1.30	.67	NS	24	1.38	.71	RT	42	1.62	.81	TY	41	1.61	.80
OP	25	1.40	.72	OS	14	1.15	.61	ST	63	1.80	.88		3,745		

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
RA	39	1.59	.80	RE	98	1.99	.96	TH	78	1.89	.92	OM	25	1.40	.72
MA	36	1.56	.78	TE	71	1.85	.91	SH	26	1.41	.72	AM	14	1.15	.61
EA	35	1.54	.78	NE	57	1.76	.87	GH	20	1.30	.67	EM	14	1.15	.61
DA	32	1.51	.76	VE	57	1.76	.87	CH	14	1.15	.61				
LA	28	1.45	.74	SE	49	1.69	.84					EN	111	2.05	.99
TA	28	1.45	.74	EE	42	1.62	.81	TI	45	1.65	.82	ON	77	1.89	.92
NA	26	1.41	.72	LE	37	1.57	.79	FI	39	1.59	.80	IN	75	1.88	.92
SA	24	1.38	.71	DE	33	1.52	.77	SI	34	1.53	.77	AN	64	1.81	.89
CA	20	1.30	.67	CE	32	1.51	.76	HI	33	1.52	.77	UN	21	1.32	.68
HA	20	1.30	.67	ME	26	1.41	.72	NI	30	1.48	.75				
PA	14	1.15	.61	PE	23	1.36	.70	RI	30	1.48	.75	TO	50	1.70	.84
				WE	22	1.34	.69	DI	27	1.43	.73	CO	41	1.61	.80
EC	32	1.51	.76	HE	20	1.30	.67	EI	27	1.43	.73	IO	41	1.61	.80
IC	22	1.34	.69	BE	18	1.26	.66	LI	20	1.30	.67	FO	40	1.60	.80
NC	19	1.28	.67	GE	14	1.15	.61	AI	17	1.23	.64	RO	28	1.45	.74
AC	14	1.15	.61	IE	13	1.11	.59					HO	20	1.30	.67
												WO	19	1.28	.67
				OF	25	1.40	.72	AL	32	1.51	.76	NO	18	1.26	.66
ED	60	1.78	.88	EF	18	1.26	.66	EL	29	1.46	.74	PO	17	1.23	.64
ND	52	1.72	.85					LL	27	1.43	.73	DO	16	1.20	.63
AD	27	1.43	.73	NG	27	1.43	.73	IL	23	1.36	.70	SO	15	1.18	.62
RD	17	1.23	.64	IG	19	1.28	.67	OL	19	1.28	.67	LO	13	1.11	.59

Table 9–D, Concluded.—*The 122 digraphs composing 75% of the 5,000 digraphs of Table 6–A, accompanied by the logarithms of their assigned probabilities, arranged alphabetically according to their final letters*

(2) AND ACCORDING TO THEIR ABSOLUTE FREQUENCIES—Concluded

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
OP	25	1.40	.72	ES	54	1.73	.86	NT	82	1.91	.93	OU	37	1.57	.79
EP	20	1.30	.67	AS	41	1.61	.80	ST	63	1.80	.88	QU	15	1.18	.62
				IS	35	1.54	.78	AT	47	1.67	.83	AU	13	1.11	.59
				RS	31	1.49	.75	RT	42	1.62	.81				
ER	87	1.94	.94	NS	24	1.38	.71	ET	37	1.57	.79	IV	25	1.40	.72
OR	64	1.81	.89	SS	19	1.28	.67	HT	28	1.45	.74	EV	20	1.30	.67
AR	44	1.64	.82	TS	19	1.28	.67	IT	27	1.43	.73	TW	36	1.56	.78
UR	31	1.49	.75	OS	14	1.15	.61	OT	19	1.28	.67	IX	15	1.18	.62
IR	27	1.43	.73	DS	13	1.11	.59	TT	19	1.28	.67				
PR	18	1.26	.66					DT	15	1.18	.62	TY	41	1.61	.80
HR	17	1.23	.64					YT	15	1.18	.62		3,745		
TR	17	1.23	.64					CT	14	1.15	.61				

Table 9–E.—*All the 428 different digraphs of Table 6–A, arranged alphabetically first according to their final letters and then according to their initial letters*

(SEE TABLE 6–A.—READ DOWN THE COLUMNS)

Table 10–A.—*The 56 trigraphs appearing 100 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
ENT	569	2.76	.99	TOP	174	2.24	.82	EIG	135	2.13	.79
ION	260	2.41	.88	NTH	171	2.23	.82	FIV	135	2.13	.79
AND	228	2.36	.86	TWE	170	2.23	.82	MEN	131	2.12	.78
ING	226	2.35	.86	TWO	163	2.21	.81	SEV	131	2.12	.78
IVE	225	2.35	.86	ATI	160	2.20	.81	ERS	126	2.10	.78
TIO	221	2.34	.85	THR	158	2.20	.81	UND	125	2.10	.78
FOR	218	2.34	.85	NTY	157	2.20	.81	NET	118	2.07	.77
OUR	211	2.32	.85	HRE	153	2.18	.80	PER	115	2.06	.76
THI	211	2.32	.85	WEN	153	2.18	.80	STA	115	2.06	.76
ONE	210	2.32	.85	FOU	152	2.18	.80	TER	115	2.06	.76
NIN	207	2.32	.85	ORT	146	2.16	.80	EQU	11[illegible]	2.06	.76
STO	202	2.31	.84	REE	146	2.16	.80	RED	113	2.05	.76
EEN	196	2.29	.84	SIX	146	2.16	.80	TED	112	2.05	.76
GHT	196	2.29	.84	ASH	143	2.16	.80	ERI	109	2.04	.76
INE	192	2.28	.83	DAS	140	2.15	.79	HIR	106	2.03	.75
VEN	190	2.28	.83	IGH	140	2.15	.79	IRT	105	2.02	.75
EVE	177	2.25	.82	ERE	138	2.14	.79	DER	101	2.00	.74
EST	176	2.25	.82	COM	136	2.13	.79	DRE	100	2.00	.74
TEE	174	2.24	.82	ATE	135	2.13	.79				

~~RESTRICTED~~

TABLE 10-B.—*The 56 trigraphs appearing 100 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their initial letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
AND	228	2.36	.86	GHT	196	2.29	.84	REE	146	2.16	.80
ATI	160	2.20	.81					RED	113	2.05	.76
ASH	143	2.16	.80	HRE	153	2.18	.80				
ATE	135	2.13	.79	HIR	106	2.03	.75				
								STO	202	2.31	.84
COM	136	2.13	.79	ION	260	2.41	.88	SIX	146	2.16	.80
				ING	226	2.35	.86	SEV	131	2.12	.78
DAS	140	2.15	.79	IVE	225	2.35	.86	STA	115	2.06	.76
DER	101	2.00	.74	INE	192	2.28	.83				
DRE	100	2.00	.74	IGH	140	2.15	.79	TIO	221	2.34	.85
				IRT	105	2.02	.75	THI	211	2.32	.85
ENT	569	2.76	.99					TEE	174	2.24	.82
EEN	196	2.29	.84	MEN	131	2.12	.78	TOP	174	2.24	.82
EVE	177	2.25	.82					TWE	170	2.23	.82
EST	176	2.25	.82	NIN	207	2.32	.85	TWO	163	2.21	.81
ERE	138	2.14	.79	NTH	171	2.23	.82	THR	158	2.20	.81
EIG	135	2.13	.79	NTY	157	2.20	.81	TER	115	2.06	.76
ERS	126	2.10	.78	NET	118	2.07	.77	TED	112	2.05	.76
EQU	114	2.06	.76								
ERI	109	2.04	.76	OUR	211	2.32	.85	UND	125	2.10	.78
				ONE	210	2.32	.85				
FOR	218	2.34	.85	ORT	146	2.16	.80	VEN	190	2.28	.83
FOU	152	2.18	.80								
FIV	135	2.13	.79	PER	115	2.06	.76	WEN	153	2.18	.80

TABLE 10-C.—*The 56 trigraphs appearing 100 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their central letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
DAS	140	2.15	.79	IGH	140	2.15	.79	ENT	569	2.76	.99
								AND	228	2.36	.86
EEN	196	2.29	.84					ING	226	2.35	.86
VEN	190	2.28	.83					ONE	210	2.32	.85
TEE	174	2.24	.82	THI	211	2.32	.85	INE	192	2.28	.83
WEN	153	2.18	.80	GHT	196	2.29	.84	UND	125	2.10	.78
REE	146	2.16	.80	THR	158	2.20	.81				
MEN	131	2.12	.78								
SEV	131	2.12	.78								
NET	118	2.07	.77	TIO	221	2.34	.85				
PER	115	2.06	.76	NIN	207	2.32	.85	ION	260	2.41	.88
TER	115	2.06	.76	SIX	146	2.16	.80	FOR	218	2.34	.85
RED	113	2.05	.76	EIG	135	2.13	.79	TOP	174	2.24	.82
TED	112	2.05	.76	FIV	135	2.13	.79	FOU	152	2.18	.80
DER	101	2.00	.74	HIR	106	2.03	.75	COM	136	2.13	.79

~~RESTRICTED~~

TABLE 10-C, Concluded.—*The 56 trigraphs appearing 100 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their central letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
EQU	114	2.06	.76	EST	176	2.25	.82	OUR	211	2.32	.85
				ASH	143	2.16	.80				
HRE	153	2.18	.80								
ORT	146	2.16	.80	STO	202	2.31	.84	IVE	225	2.35	.86
ERE	138	2.14	.79	NTH	171	2.23	.82	EVE	177	2.25	.82
ERS	126	2.10	.78	ATI	160	2.20	.81				
ERI	109	2.04	.76	NTY	157	2.20	.81				
IRT	105	2.02	.75	ATE	135	2.13	.79	TWE	170	2.23	.82
DRE	100	2.00	.74	STA	115	2.06	.76	TWO	163	2.21	.81

TABLE 10-D.—*The 56 trigraphs appearing 100 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their final letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
STA	115	2.06	.76	THI	211	2.32	.85	TER	115	2.06	.76
				ATI	160	2.20	.81	HIR	106	2.03	.75
AND	228	2.36	.86	ERI	109	2.04	.76	DER	101	2.00	.74
UND	125	2.10	.78								
RED	113	2.05	.76	COM	136	2.13	.79	DAS	140	2.15	.79
TED	112	2.05	.76					ERS	126	2.10	.78
IVE	225	2.35	.86	ION	260	2.41	.88				
ONE	210	2.32	.85	NIN	207	2.32	.85	ENT	569	2.76	.99
INE	192	2.28	.83	EEN	196	2.29	.84	GHT	196	2.29	.84
EVE	177	2.25	.82	VEN	190	2.28	.83	EST	176	2.25	.82
TEE	174	2.24	.82	WEN	153	2.18	.80	ORT	146	2.16	.80
TWE	170	2.23	.82	MEN	131	2.12	.78	NET	118	2.07	.77
HRE	153	2.18	.80					IRT	105	2.02	.75
REE	146	2.16	.80	TIO	221	2.34	.85				
ERE	138	2.14	.79	STO	202	2.31	.84	FOU	152	2.18	.80
ATE	135	2.13	.79	TWO	163	2.21	.81	EQU	114	2.06	.76
DRE	100	2.00	.74								
								FIV	135	2.13	.79
ING	226	2.35	.86	TOP	174	2.24	.82	SEV	131	2.12	.78
EIG	135	2.13	.79								
				FOR	218	2.34	.85				
NTH	171	2.23	.82	OUR	211	2.32	.85	SIX	146	2.16	.80
ASH	143	2.16	.80	THR	158	2.20	.81				
IGH	140	2.15	.79	PER	115	2.06	.76	NTY	157	2.20	.81

TABLE 11–A.—*The 54 tetragraphs appearing 50 or more times in the 50,000 letters of Governmental plan-text telegrams, arranged according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{274}(2F)$		F	$L_{10}(F)$	$L_{274}(2F)$		F	$L_{10}(F)$	$L_{274}(2F)$
TION	218	2.34	.99	THIR	104	2.02	.87	ASHT	64	1.81	.79
EVEN	168	2.23	.95	EENT	102	2.01	.87	HUND	64	1.81	.79
TEEN	163	2.21	.94	REQU	98	1.99	.86	DRED	63	1.80	.79
ENTY	161	2.21	.94	HIRT	97	1.99	.86	RIOD	63	1.80	.79
STOP	154	2.19	.93	COMM	93	1.97	.85	IVED	62	1.79	.78
WENT	153	2.18	.93	QUES	87	1.94	.84	ENTS	62	1.79	.78
NINE	153	2.18	.93	UEST	87	1.94	.84	FFIC	62	1.79	.78
TWEN	152	2.18	.93	EQUE	86	1.93	.84	FROM	59	1.77	.78
THRE	149	2.17	.93	NDRE	77	1.89	.82	IRTY	59	1.77	.78
FOUR	144	2.16	.92	OMMA	71	1.85	.81	RTEE	59	1.77	.78
IGHT	140	2.15	.92	LLAR	71	1.85	.81	UNDR	59	1.77	.78
FIVE	135	2.13	.91	OLLA	70	1.85	.81	NAUG	56	1.75	.77
HREE	134	2.13	.91	VENT	70	1.85	.81	OURT	56	1.75	.77
DASH	132	2.12	.91	DOLL	68	1.83	.80	UGHT	56	1.75	.77
EIGH	132	2.12	.91	LARS	68	1.83	.80	STAT	54	1.73	.76
SEVE	121	2.08	.89	THIS	68	1.83	.80	AUGH	52	1.72	.76
ENTH	114	2.06	.89	PERI	67	1.83	.80	CENT	52	1.72	.76
MENT	111	2.05	.88	ERIO	66	1.82	.80	FICE	50	1.70	.75

TABLE 11–B.—*The 54 tetragraphs appearing 50 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their initial letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{274}(2F)$		F	$L_{10}(F)$	$L_{274}(2F)$		F	$L_{10}(F)$	$L_{274}(2F)$
ASHT	64	1.81	.79	HREE	134	2.13	.91	REQU	98	1.99	.86
AUGH	52	1.72	.76	HIRT	97	1.99	.86	RIOD	63	1.80	.79
COMM	93	1.97	.85	HUND	64	1.81	.79	RTEE	59	1.77	.78
CENT	52	1.72	.76	IGHT	140	2.15	.92	STOP	154	2.19	.93
DASH	132	2.12	.91	IVED	62	1.79	.78	SEVE	121	2.08	.89
DOLL	68	1.83	.80	IRTY	59	1.77	.78	STAT	54	1.73	.76
DRED	63	1.80	.79	LLAR	71	1.85	.81	TION	218	2.34	.99
EVEN	168	2.23	.95	LARS	68	1.83	.80	TEEN	163	2.21	.94
ENTY	161	2.21	.94	MENT	111	2.05	.88	TWEN	152	2.18	.93
EIGH	132	2.12	.91	NINE	153	2.18	.93	THRE	149	2.17	.93
ENTH	114	2.06	.89	NDRE	77	1.89	.82	THIR	104	2.02	.87
EENT	102	2.01	.87	NAUG	56	1.75	.77	THIS	68	1.83	.80
EQUE	86	1.93	.84	OMMA	71	1.85	.81	UEST	87	1.94	.84
ERIO	66	1.82	.80	OLLA	70	1.85	.81	UNDR	59	1.77	.78
ENTS	62	1.79	.78	OURT	56	1.75	.77	UGHT	56	1.75	.77
FOUR	144	2.16	.92	PERI	67	1.83	.80	VENT	70	1.85	.81
FIVE	135	2.13	.91	QUES	87	1.94	.84	WENT	153	2.18	.93
FFIC	62	1.79	.78								
FROM	59	1.77	.78								
FICE	50	1.70	.75								

TABLE 11-C.—*The 54 tetragraphs appearing 50 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their second letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{464}(2F)$
DASH	132	2.12	.91
LARS	68	1.83	.80
NAUG	56	1.75	.77
NDRE	77	1.89	.82
TEEN	163	2.21	.94
WENT	153	2.18	.93
SEVE	121	2.08	.89
MENT	111	2.05	.88
EENT	102	2.01	.87
REQU	98	1.99	.86
UEST	87	1.94	.84
VENT	70	1.85	.81
PERI	67	1.83	.80
CENT	52	1.72	.76
FFIC	62	1.79	.78
IGHT	140	2.15	.92
UGHT	56	1.75	.77
THRE	149	2.17	.93
THIR	104	2.02	.87
THIS	68	1.83	.80
TION	218	2.34	.99
NINE	153	2.18	.93
FIVE	135	2.13	.91
EIGH	132	2.12	.91
HIRT	97	1.99	.86
RIOD	63	1.80	.79
FICE	50	1.70	.75
LLAR	71	1.85	.81
OLLA	70	1.85	.81
OMMA	71	1.85	.81
ENTY	161	2.21	.94
ENTH	114	2.06	.89
ENTS	62	1.79	.78
UNDR	59	1.77	.78
FOUR	144	2.16	.92
COMM	93	1.97	.85
DOLL	68	1.83	.80
EQUE	86	1.93	.84
HREE	134	2.13	.91
ERIO	66	1.82	.80
DRED	63	1.80	.79
FROM	59	1.77	.78
IRTY	59	1.77	.78
ASHT	64	1.81	.79
STOP	154	2.19	.93
RTEE	59	1.77	.78
STAT	54	1.73	.76
QUES	87	1.94	.84
HUND	64	1.81	.79
OURT	56	1.75	.77
AUGH	52	1.72	.76
EVEN	168	2.23	.95
IVED	62	1.79	.78
TWEN	152	2.18	.93

TABLE 11-D.—*The 54 tetragraphs appearing 50 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their third letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{464}(2F)$
LLAR	71	1.85	.81
STAT	54	1.73	.76
FICE	50	1.70	.75
UNDR	59	1.77	.78
EVEN	168	2.23	.95
TEEN	163	2.21	.94
TWEN	152	2.18	.93
HREE	134	2.13	.91
QUES	87	1.94	.84
DRED	63	1.80	.79
IVED	62	1.79	.78
RTEE	59	1.77	.78
EIGH	132	2.12	.91
AUGH	52	1.72	.76
IGHT	140	2.15	.92
ASHT	64	1.81	.79
UGHT	56	1.75	.77
THIR	104	2.02	.87
THIS	68	1.83	.80
ERIO	66	1,82	.80
FFIC	62	1.79	.78
OLLA	70	1.85	.81
DOLL	68	1.83	.80
COMM	93	1.97	.85
OMMA	71	1.85	.81
WENT	153	2.18	.93
NINE	153	2.18	.93
MENT	111	2.05	.88
EENT	102	2.01	.87
VENT	70	1.85	.81
HUND	64	1.81	.79
CENT	52	1.72	.76
TION	218	2.34	.99
STOP	154	2.19	.93
RIOD	63	1.80	.79
FROM	59	1.77	.78

TABLE 11-D, Concluded.—*The 54 tetragraphs appearing 50 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their third letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
REQU	98	1.99	.86	DASH	132	2.12	.91	FOUR	144	2.16	.92
				UEST	87	1.94	.84	EQUE	86	1.93	.84
THRE	149	2.17	.93					NAUG	56	1.75	.77
HIRT	97	1.99	.86								
NDRE	77	1.89	.82	ENTY	161	2.21	.94				
LARS	68	1.83	.80	ENTH	114	2.06	.89				
PERI	67	1.83	.80	ENTS	62	1.79	.78	FIVE	135	2.13	.91
OURT	56	1.75	.77	IRTY	59	1.77	.78	SEVE	121	2.08	.89

TABLE 11-E.—*The 54 tetragraphs appearing 50 or more times in the 50,000 letters of Governmental plain-text telegrams, arranged first alphabetically according to their final letters, and then according to their absolute frequencies, accompanied by the logarithms of their assigned probabilities*

	F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$		F	$L_{10}(F)$	$L_{224}(2F)$
OMMA	71	1.85	.81	DASH	132	2.12	.91	QUES	87	1.94	.84
OLLA	70	1.85	.81	EIGH	132	2.12	.91	THIS	68	1.83	.80
				ENTH	114	2.06	.89	LARS	68	1.83	.80
				AUGH	52	1.72	.76	ENTS	62	1.79	.78
FFIC	62	1.79	.78								
				PERI	67	1.83	.80				
								WENT	153	2.18	.93
HUND	64	1.81	.79					IGHT	140	2.15	.92
DRED	63	1.80	.79	DOLL	68	1.83	.80	MENT	111	2.05	.88
RIOD	63	1.80	.79	COMM	93	1.97	.85	EENT	102	2.01	.87
IVED	62	1.79	.78	FROM	59	1.77	.78	HIRT	97	1.99	.86
								UEST	87	1.94	.84
				TION	218	2.34	.99	VENT	70	1.85	.81
NINE	153	2.18	.93	EVEN	168	2.23	.95	ASHT	64	1.81	.79
THRE	149	2.17	.93	TEEN	163	2.21	.94	OURT	56	1.75	.77
FIVE	135	2.13	.91	TWEN	152	2.18	.93	UGHT	56	1.75	.77
HREE	134	2.13	.91					STAT	54	1.73	.76
SEVE	121	2.08	.89	ERIO	66	1.82	.80	CENT	52	1.72	.76
EQUE	86	1.93	.84								
NDRE	77	1.89	.82	STOP	154	2.19	.93				
RTEE	59	1.77	.78					REQU	98	1.99	.86
FICE	50	1.70	.75	FOUR	144	2.16	.92				
				THIR	104	2.02	.87				
				LLAR	71	1.85	.81	ENTY	161	2.21	.94
NAUG	56	1.75	.77	UNDR	59	1.77	.78	IRTY	59	1.77	.78

TABLE 12.—*Average length of words and messages*

Number of letters in word x	Number of times x-letter word appears	Number of letters
1	378	378
2	973	1,946
3	1,307	3,921
4	1,635	6,540
5	1,410	7,050
6	1,143	6,858
7	1,009	7,063
8	717	5,736
9	476	4,284
10	274	2,740
11	161	1,771
12	86	1,032
13	23	299
14	23	322
15	4	60
	9,619	50,000

(1) Average length of words .... 5.2 letters.
(2) Average length of messages .... 217 letters.
(3) Modal (most frequent) length .... 105–114 letters.
(4) It is extremely unusual to find five consecutive letters without at least one vowel.
(5) The average number of letters between vowels is two.

TABLE 13.—*Checkerboard individual frequencies*[1]

[Based on a count of 5,000 digraphs]

P₁					C₁				
A	B	C	D	E	244	225	375	394	197
F	G	H	I J	K	125	98	193	271	95
L	M	N	O	P	229	199	188	350	251
Q	R	S	T	U	148	162	258	427	295
V	W	X	Y	Z	42	12	34	91	97
212	317	358	308	249	A	B	C	D	E
120	108	216	256	85	F	G	H	I J	K
216	140	152	435	269	L	M	N	O	P
206	121	306	364	284	Q	R	S	T	U
38	29	21	147	43	V	W	X	Y	Z
C₂					P₂				

[1] The numbers in the C_1 C_2 squares represent the frequency of the individual components of the cipher digraph used to replace a given P_1 P_2 digraph in accordance with a digraphic checkerboard system where P_1 and P_2 are the plain-text squares.

TABLE 14.—*Relative logarithmic values of frequencies of English digraphs*

[Based on a count of 5,000 digraphs. To obtain logarithm to base 10 (Log 10) divide by 100]

SECOND LETTER

FIRST 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	48	78	115	143	00	60	78	30	123	00	30	151	115	181	30	108	*	164	161	167	111	85	48	*	108	*
B	60	*	*	*	126	*	*	*	30	00	*	78	00	*	60	*	*	30	00	00	30	*	*	*	85	*
C	130	*	48	00	151	00	*	115	85	*	60	70	00	00	161	*	*	60	00	115	60	*	00	*	00	*
D	151	60	60	90	152	90	30	30	143	00	*	48	70	60	120	70	30	108	111	118	70	48	60	*	00	*
E	154	60	151	178	162	126	60	85	143	00	*	146	115	205	108	130	108	194	173	157	48	130	85	85	60	00
F	70	*	30	00	100	104	00	*	159	*	*	30	00	*	160	00	*	95	48	104	48	*	00	*	00	*
G	85	*	30	00	115	30	00	130	70	00	*	30	00	48	78	30	*	70	48	60	30	*	00	*	*	*
H	130	00	48	30	130	70	*	*	152	*	*	00	30	48	130	00	00	123	60	145	90	*	00	*	00	*
I	90	30	135	78	111	100	128	*	*	*	30	136	95	188	161	85	*	143	154	143	*	140	*	118	*	30
J	00	*	*	*	30	*	*	*	*	*	*	*	*	*	30	*	*	*	*	*	30	*	*	*	*	*
K	00	*	00	*	78	*	*	*	30	*	*	00	*	00	*	*	*	*	00	*	*	*	*	*	*	*
L	145	48	48	95	157	48	00	00	130	*	*	143	30	00	111	48	*	30	78	90	30	30	30	*	100	*
M	156	78	48	00	141	00	*	00	95	*	*	*	111	*	100	90	*	30	60	30	30	*	*	*	30	*
N	141	30	128	172	176	95	43	60	148	00	30	70	70	90	126	48	00	60	138	191	85	48	48	*	70	*
O	85	60	90	108	48	140	30	48	70	00	30	128	140	189	78	140	*	181	115	128	157	85	90	00	30	*
P	115	00	00	00	136	30	*	48	78	*	*	111	60	00	123	104	*	126	78	90	48	00	00	*	00	*
Q	*	*	*	*	*	*	*	*	*	*	*	*	00	*	*	*	*	00	*	*	118	*	*	*	*	*
R	159	30	95	123	199	78	85	48	148	00	00	70	95	85	145	111	*	104	149	162	70	70	60	*	95	*
S	138	48	111	70	169	108	30	142	153	*	00	30	48	60	118	100	*	70	128	180	104	00	60	*	00	*
T	145	48	78	78	185	85	00	189	165	*	*	70	78	85	170	30	00	123	128	128	70	*	156	*	161	00
U	70	48	48	48	104	00	90	*	70	*	*	78	70	132	00	30	*	149	108	108	*	00	*	*	*	*
V	78	*	*	*	176	*	*	*	108	*	*	*	*	*	00	*	*	*	*	00	*	*	*	*	*	*
W	108	*	*	*	134	*	*	60	111	*	*	00	*	30	128	*	*	00	00	*	*	*	*	*	00	*
X	30	*	30	00	00	00	*	00	30	*	*	*	*	00	00	30	*	00	00	85	*	*	*	*	*	*
Y	78	30	60	60	95	104	00	00	48	*	*	30	30	78	100	48	*	60	104	118	00	*	00	*	*	*
Z	00	*	*	*	30	*	*	*	00	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

*In computations, assign a value of −100 as the log for these digraphs. These combinations do not usually occur in 5,000 digraphs. Do not assign "0" to these combinations as that is the logarithmic value for a frequency of one, and these combinations have a frequency of less than one.

TABLE 15.—*Relative logarithmic values (Log. 222) of frequencies of English digraphs* *

[Based on a count of 5,000 digraphs]

SECOND LETTER

FIRST 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	.33	.45	.61	.73	.13	.38	.45	.25	.64	.13	.25	.76	.61	.89	.25	.58	0	.82	.80	.83	.59	.48	.33	0	.58	0
B	.38	0	0	0	.66	0	0	0	.25	.13	0	.45	.13	0	.38	0	0	.25	.13	.13	.25	0	0	0	.48	0
C	.67	0	.33	.13	.76	.13	0	.61	.48	0	.38	.42	.13	.13	.80	0	0	.38	.13	.61	.38	0	.13	0	.13	0
D	.76	.38	.38	.51	.77	.51	.25	.25	.73	.13	0	.33	.42	.88	.63	.42	0	.58	.59	.62	.42	.33	.38	0	.13	0
E	.78	.38	.76	.88	.81	.66	.38	.48	.73	.13	0	.74	.61	.99	.58	.67	.58	.94	.86	.79	.33	.67	.48	.48	.38	.13
F	.42	0	.25	.13	.55	.56	.13	0	.80	0	0	.25	.13	0	.80	.13	0	.53	.33	.56	.33	0	.13	0	.13	0
G	.48	0	.25	.13	.61	.25	.13	.67	.42	.13	0	.25	.13	.33	.45	.25	0	.42	.33	.38	.25	0	.13	0	0	0
H	.67	.13	.38	.25	.67	.42	0	0	.77	0	0	.13	.25	.33	.67	.13	.13	.64	.38	.74	.51	0	.13	0	.13	0
I	.51	.25	.69	.45	.59	.55	.67	0	0	0	.25	.70	.53	.92	.80	.48	0	.73	.78	.73	0	.72	0	.62	0	.25
J	.13	0	0	0	.25	0	0	0	0	0	0	0	0	0	.25	0	0	0	0	0	.25	0	0	0	0	0
K	.13	0	.13	0	.45	0	0	0	.25	0	0	.13	0	.13	0	0	0	0	.13	0	0	0	0	0	0	0
L	.74	.33	.33	.53	.79	.33	.13	.13	.67	0	0	.73	.25	.13	.59	.33	0	.25	.45	.51	.25	.25	.25	0	.55	0
M	.78	.45	.33	.13	.72	.13	0	.13	.53	0	0	0	.59	0	.55	.51	0	.25	.38	.25	.25	0	0	0	.25	0
N	.72	.25	.67	.85	.87	.53	.73	.38	.75	.13	.25	.42	.42	.51	.66	.33	.13	.38	.71	.93	.48	.33	.33	0	.42	0
O	.48	.38	.51	.58	.33	.72	.25	.33	.42	.13	.25	.67	.72	.92	.45	.72	0	.89	.61	.67	.79	.48	.51	.13	.25	0
P	.61	.13	.13	.13	.70	.25	0	.33	.45	0	0	.59	.38	.13	.64	.56	0	.66	.45	.51	.33	.13	.13	0	.13	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	.13	0	0	0	0	.13	0	0	.62	0	0	0	0	0
R	.80	.25	.53	.64	.96	.45	.48	.33	.75	.13	.13	.42	.53	.48	.74	.59	0	.56	.75	.81	.42	.42	.38	0	.53	0
S	.71	.33	.59	.42	.84	.58	.25	.72	.77	0	.13	.25	.33	.38	.62	.55	0	.42	.67	.88	.56	.13	.38	0	.13	0
T	.74	.33	.45	.45	.91	.48	.13	.92	.82	0	0	.42	.45	.48	.84	.25	.13	.64	.67	.67	.42	0	.78	0	.80	.13
U	.42	.33	.33	.33	.56	.13	.51	0	.42	0	0	.45	.42	.68	.13	.25	0	.75	.58	.58	0	.13	0	0	0	0
V	.45	0	0	0	.87	0	0	0	.58	0	0	0	0	0	.13	0	0	0	0	.13	0	0	0	0	0	0
W	.58	0	0	0	.69	0	0	.38	.59	0	0	.13	0	.25	.67	0	0	.13	.13	0	0	0	0	0	.13	0
X	.25	0	.25	.13	.13	.13	0	.13	.25	0	0	0	0	.13	.13	.25	0	.13	.13	.48	0	0	0	0	0	0
Y	.45	.25	.38	.38	.53	.56	.13	.13	.33	0	0	.25	.25	.45	.55	.33	0	.88	.56	.62	.13	0	.13	0	0	0
Z	.13	0	0	0	.25	0	0	0	.13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

* See pages 11–12 for details.

1197

Table 16-A.--Frequency distribution of digraphs, based on 64,365 letters of decrypted U. S. Government messages in which Z was used as a word-separator and X was used for both X_p and Z_p.

1st Ltr. \ 2d Ltr.	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	28	154	142	137	17	90	99	13	118	16	43	220	157	427	18	112	2	625	526	347	56	52	20	3	66	546
B	63	14	7	1	193	1		1	43	33		148	6	18	61	2		59	17	8	15	1	1	3	60	19
C	123	1	19	8	260	22	28	183	115		48	95	390	5	414	3	1	63	66	161	47	1	5	3	27	122
D	360	12	33	30	270	4	16		141	2	1	7	4	6	102	11	11	33	32	34	38	38	17	1	11	1026
E	180	34	226	383	620	131	35	13	275	3	6	185	144	768	75	118	91	857	329	187	40	210	28	76	29	1715
F	44	16	10	3	100	122	4	1	365	2		28	23	4	536	68		114	8	32	34	1	1	2	3	343
G	78	29	7	18	258	5	31	260	25		1	11	5	31	20	18		73	29	17	25	2			1	275
H	194	1	6	12	193	14	1	24	213	3	9	7	2	24	93	3	24	229	26	257	17	2	6	1	3	428
I	85	10	209	30	152	53	330	5	5	1	46	181	40	704	200	92	1	128	303	217	2	272	2	193	1	56
J	26		3	2	31	3		1	18	20		3	1	4	35	1		5	2	18	7	2	1		2	19
K	28		2	6	108	2			54	3	20	11	3	10	9		1	1	9	2	1	1	2	1	10	59
L	159	6	6	48	328	14		4	194	2	1	237	20	65	120	5		5	41	25	41	5		1	71	296
M	581	68	36	12	198	1	58	1	92	4	1	2	62	4	43	101		10	53	20	17	1	3	6	86	231
N	112	13	157	286	733	77	244	4	234		14	15	9	76	169	16	16	13	135	267	64	10	7	7	14	910
O	25	67	46	100	56	317	66	26	23	6	23	161	230	873	59	57	2	418	129	143	413	49	59	92	13	916
P	304	5	8	363	169		2	37	27	3		75	46	9	145	104	3	153	26	351	44	2	2	1	4	122
Q	2	1	1		7			4	1			1	5	11	1	1	9	5	7		117		1			46
R	241	5	44	86	967	26	59	5	191	5	30	61	122	45	570	310	4	72	208	179	60	19	14	13	74	733
S	143	14	66	6	389	85	52	426	334	1	16	16	34	6	99	47	13	5	143	305	138	13	12	1	43	788
T	171	1	67	22	357	32	6	572	275	2	10	27	18	49	372	9	2	119	99	156	37	1	313	10	48	1106
U	45	48	26	60	87	4	61	2	35	1	3	56	61	96	32	38		453	140	48	5	5	5	1	1	44
V	39		10	2	496	1	1		91		1	3	1	8	19	4	1	3	4	7	1	9	1	1	7	34
W	111	1	3	7	34	1	11	33	107	2	1	10		12	367	7	2	3	11	5			13	13	2	30
X	9	1	8	7	350	9		2	10	1	2		2	2	10	20	3	12	9	32	1	1		32	3	203
Y	8	3	6	3	14	6	3	2	5			4	9	10	49	27		3	18	8	4	1	1		8	432
Z	902	264	1058	613	364	844	120	171	328	98	69	135	294	349	750	823	36	700	768	1046	130	46	278	271	42	
	4081	768	2266	2245	6751	1864	1227	1790	3319	208	345	1699	1628	3606	4368	1997	222	4161	3198	3872	1354	744	792	732	629	10499

In the text which gave rise to this and the following two tables, the frequently-used punctuation signs "comma" and "period" were abbreviated as CMA and PD, respectively, and the procedure term "repeat" was abbreviated as RPT; thus, the digraphs CM, PD, PT, and RP, which usually do not occur frequently (see Table 6-A), are of relatively high frequency here.

Table 16-B.--Frequency distribution of digraphs, based on the text used for Table 16-A, from which the Z word-separator has been omitted (total: 53,866 letters).

2d Ltr.

1st Ltr.	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	78	175	190	164	40	136	111	26	189	19	52	227	166	439	58	147	3	657	619	395	65	58	40	23	67	
B	63	14	9	2	193	5		1	43	32		149	6	18	62	2		62	17	13	15	2	3	3	60	
C	133	1	31	20	263	32	29	184	119		48	98	393	11	416	8	2	78	79	180	47	1	6	4	27	
D	443	66	102	74	307	86	26	13	183	7	5	23	32	22	151	97	16	142	118	153	59	40	55	2	18	
E	299	70	384	481	690	283	48	37	326	21	12	201	190	855	181	278	93	931	476	367	53	215	87	136	34	
F	60	19	42	25	109	137	7	2	380	3	1	39	25	10	582	80	1	148	56	67	49	3	9	3	7	
G	102	39	20	59	266	19	32	262	37	5	2	12	10	41	45	38	4	91	53	38	31	2	3	7	1	
H	270	8	34	28	215	54	13	31	220	14	11	8	13	34	139	14	23	239	64	315	18	3	16	5	3	
I	86	10	213	41	156	55	330	8	5	1	46	182	40	705	202	96	1	148	303	218	2	270	3	196	1	
J	28		7	2	31	7		1	21	20		3	1	5	36	2		6	2	19	7	2	2		2	
K	35	4	7	10	108	10	2		56	3	20	11	4	13	12	7	1	6	11	5	2	1	4	1	10	
L	197	21	38	61	338	47	2	13	207	7	4	243	26	68	134	19		21	59	50	44	8	14	1	72	
M	595	72	66	18	206	22	64	4	96	6	1	6	67	17	63	123	3	26	61	40	22	2	10	15	86	
N	218	27	280	336	748	139	254	12	263	6	19	31	47	86	234	92	24	66	202	352	75	23	28	28	17	
O	63	82	191	155	93	426	72	47	37	13	27	172	252	910	99	112	2	473	204	214	417	51	68	170	17	
P	311	7	16	388	170	5	3	40	29	4		76	46	11	150	111	3	179	37	365	44	2	2	1	5	
Q	14	4	3		7	2		4	5		1	2	5	11	8	2	9	10	10	2	117		3	1		
R	298	12	131	146	1011	84	66	14	207	17	40	69	142	59	639	369	8	103	266	263	67	19	29	30	74	
S	237	37	145	31	396	149	55	453	369	5	19	25	60	36	173	129	16	62	178	385	144	14	34	2	43	
T	277	30	167	70	400	97	21	592	308	14	16	43	67	100	463	95	5	195	150	282	52	12	338	30	57	
U	48	48	33	61	88	7	61	2	36	4	4	56	61	97	35	40		454	148	50	6	5	6	6	1	
V	41		13	5	499	7	1		92		2	4	3	8	21	6	2	4	9	8	1	9	1	1	7	
W	113	6	6	9	37	2	12	35	107	3	1	10	1	14	367	10	2	3	11	6	1		13	13	4	
X	18	2	23	22	361	20		4	12	3	10	2	7	11	24	41	3	26	29	47	4	1		54	3	
Y	59	14	57	37	19	33	18	5	22	1	4	7	22	25	74	77	1	31	36	38	10	1	18		13	
Z																										
	4081	768	2206	2245	6751	1864	1227	1790	3319	208	345	1699	1688	3606	4368	1997	222	4161	3198	3872	1354	744	792	732	629	

Table 16-C.--The 53 digraphs from Table 6-A which comprise 50% of the total, arranged according to frequencies reduced to a base of 5,000 digraphs, shown with the corresponding frequencies of the same digraphs from Table 16-B (also reduced to a base of 5,000).[1]

Dig.	6-A	16-B	Dig.	6-A	16-B
EN	111	79	FO	40	54
RE	98	94	FI	39	35
ER	87	86	RA	39	28
NT	82	33	ET	37	34
TH	78	55	LE	37	31
ON	77	84	OU	37	39
IN	75	65	MA	36	55
TE	71	36	TW	36	31
AN	64	41	EA	35	28
OR	64	44	IS	35	28
ST	63	36	SI	34	34
ED	60	45	DE	33	29
NE	57	69	HI	33	20[1]
VE	57	46	AL	32	21[1]
ES	54	44	CE	32	24
ND	52	31	DA	32	41
TO	50	43	EC	32	36
SE	49	37	RS	31	25
AT	47	37	UR	31	42
TI	45	29	NI	30	24
AR	44	61	RI	30	19[1]
EE	42	64	EL	29	19[1]
RT	42	24	HT	28	29
AS	41	57	LA	28	18[1]
CO	41	39	RO	28	59
IO	41	19[1]	TA	28	26
TY	41	5[1]			

[1] With the exception of AL, EL, HI, IO, LA, RI, TY, the digraphs of this table are all from among the 65 digraphs from Table 16-B which comprise 50% of the total.

APPENDICES

APPENDIX 3

WORD AND PATTERN LISTS - ENGLISH

A. LIST OF WORDS USED IN MILITARY TEXT ARRANGED ALPHABETICALLY ACCORDING TO WORD LENGTH

TWO LETTER WORDS

AM	BY	EM	IN	MM	OK	TO
AN	CO	GO	IS	MP	ON	US
AS	CP	HE	IT	MY	OR	WD
AT	CQ	HQ	MC	NO	QM	WE
BE	DO	IF	ME	OF	SO	WO
BN						

THREE LETTER WORDS

ACT	BID	DUN	HAS	MIX	PVT	TEN
ADD	BIG	EAT	HER	NAN	QMC	THE
ADJ	BOX	END	HIM	NET	RED	TIN
AGE	BUT	EYE	HIS	NEW	RID	TON
AGO	BUY	FAR	HOW	NOT	ROB	TOO
AID	CAM	FEW	ILL	NOW	RUN	TOP
AIM	CAN	FIT	ITS	OFF	SAW	TRY
AIR	CAR	FIX	JIG	OLD	SAY	TUB
ALL	CAV	FOR	JOB	ONE	SEA	TWO
AND	COL	FOX	KEG	OUR	SEE	USE
ANY	CPL	GAL	LAW	OUT	SET	VAT
APT	CUT	GAS	LAY	OWE	SGT	WAR
ARC	CWT	GEN	LET	OWN	SHE	WAS
ARE	DAY	GET	LOT	PAR	SIX	WAY
ARM	DID	GHQ	LOW	PAY	SPY	WET
ASK	DIE	GOT	MAJ	PEN	SUM	WGT
BAD	DOG	GUN	MAN	PER	SUN	WON
BAG	DRY	HAD	MAT	PIN	TAN	YET
BAR	DUE	HAM	MEN	PUT	TAX	YOU

FOUR LETTER WORDS

ABLE	BOTH	EACH	FLEE	HIGH	LATE	MAIN
AIDE	BULB	EAST	FORM	HILL	LEAD	MANY
ALLY	BULK	EASY	FOUR	HITS	LEAK	MASK
ALSO	CALL	EDGE	FROM	HOLD	LEFT	MASS
AREA	CELL	EYES	FULL	HOOK	LESS	MEAT
ARMY	CITY	FALL	FUSE	INTO	LIEU	MEET
ASIA	CODE	FARM	FUZE	ITEM	LINE	MESS
AWAY	COOK	FAST	GUNS	JOIN	LIST	MIKE
AXIS	DARK	FEEL	HALF	JULY	LOAD	MILE
BACK	DASH	FEET	HALT	JUNE	LONG	MINE
BASE	DATE	FELL	HAND	JUST	LOOK	MORE
BEEN	DAYS	FILE	HARD	KEEP	LOSS	MOVE
BLUE	DIRT	FIRE	HAVE	KIND	LOST	MTCL
BODY	DOWN	FIRM	HEAD	KING	LOVE	MULE
BOMB	DRAW	FIVE	HERD	LAND	MADE	NAVY
BOOK	DUMP	FLAG	HERE	LAST	MAIM	NEAR

FOUR LETTER WORDS—Continued

NEXT	PARK	REAR	SHOT	TEAM	TOOK	WEST
NINE	PASS	RIOT	SIDE	TENT	TOOL	WHAT
NOON	PIPE	ROAD	SOME	TEXT	TOWN	WHEN
NOTE	PLAN	ROUT	SOON	THAN	TYPE	WILL
OBOE	POST	RULE	STOP	THAT	UNIT	WIRE
OMIT	PUMP	RUSH	SUNK	THEM	VARY	WITH
ONCE	PUSH	SAID	TAKE	THEN	VERY	XRAY
ONLY	RAID	SAME	TALK	THEY	WEAK	YOKE
OPEN	RAIL	SANK	TANK	THIS	WEEK	ZERO
ORAL	RAIN	SEEN	TARE	TIME	WELL	ZONE
OVER	RANK	SHIP	TASK	TONS	WERE	

FIVE LETTER WORDS

ABOUT	BOATS	DECKS	FIGHT	LATER	PRIOR	SHIPS	TITLE
AFTER	BOMBS	DEFER	FIRES	LEAST	PROOF	SHORE	TODAY
AGAIN	BOOTH	DELAY	FIRST	LEAVE	PROVE	SIEGE	TOTAL
AGENT	BREAK	DEPOT	FLANK	LEVEL	QUEEN	SIGHT	TRACT
ALARM	BRIBE	DEPTH	FLARE	LIGHT	QUICK	SIXTH	TRAIN
ALERT	BROKE	DOCKS	FLATS	LIMIT	QUIET	SIXTY	TROOP
ALIGN	BURST	DRAWN	FLEET	LOCAL	RADIO	SLOPE	TRUCE
ALINE	CANAL	DRESS	FOGGY	MAJOR	RAFTS	SMALL	TRUCK
ALLOW	CASES	DRILL	FORCE	MARCH	RAIDS	SMOKE	UNDER
ALONG	CAUSE	DRIVE	FORTY	METER	RALLY	SOUTH	UNION
AMONG	CEASE	EAGER	FRESH	MILES	RANGE	SPEED	UNITS
ANNEX	CHECK	EARLY	FRONT	MOTOR	RAPID	SPELL	USUAL
APPLY	CHIEF	EIGHT	GATES	NAVAL	REACH	SPLIT	VALOR
APRIL	CLEAR	ENEMY	GAUGE	NIGHT	READY	SQUAD	VISIT
AREAS	CLERK	ENTER	GIVEN	NINTH	REFER	STAFF	VITAL
ARMOR	CLOSE	EQUAL	GOING	NORTH	REPEL	STAKE	VOCAL
ASSET	COAST	EQUIP	GROUP	ORDER	RIDGE	START	VOICE
AWAIT	COLON	ERASE	GUARD	OTHER	RIGHT	STEEL	WAGON
AWARD	COMMA	ERROR	GUEST	PACKS	RIGID	SUGAR	WEIGH
BAKER	CORPS	ETHER	HEAVY	PAIRS	RIVER	TAKEN	WHEEL
BANKS	COUNT	EVERY	HONOR	PARTY	ROGER	TANKS	WHERE
BARGE	COVER	FATAL	HORSE	PETER	ROUTE	TENTH	WHICH
BEACH	CREEK	FEARS	HOURS	PLACE	SCALE	THEIR	WIDTH
BEGIN	CREST	FERRY	HOUSE	PLAIN	SEIZE	THERE	WIPED
BEING	CROSS	FIELD	ISSUE	PLANS	SEVEN	THESE	WOODS
BLACK	CURVE	FIFTH	JAPAN	POINT	SHELL	THIRD	YARDS
BLIND	DAILY	FIFTY	LARGE	PRESS	SHIFT	THREE	ZEBRA

SIX LETTER WORDS

ACCEPT	BOMBED	DEGREE	FIERCE	LESSON	OTHERS	RESUME	SUFFER
ACCESS	BOMBER	DEPART	FILING	LETTER	OUTPUT	RETIRE	SUMMER
ACROSS	BOTTOM	DEPEND	FINISH	LINING	PANAMA	RETURN	SUMMIT
ACTION	BRANCH	DEPLOY	FIRING	LIQUID	PARADE	REVIEW	SUMMON
ACTIVE	BREACH	DESERT	FLIGHT	LITTER	PARLEY	RIDING	SUNDAY
ADJUST	BREEZE	DETACH	FLYING	LITTLE	PASSED	ROCKET	SUNKEN
ADVICE	BRIDGE	DETAIL	FOLLOW	LOCATE	PASSES	ROUTED	SUNSET
ADVISE	BROKEN	DEVICE	FORCES	LOSSES	PATROL	ROUTES	SUPPLY
AFFAIR	BUREAU	DEVISE	FORMAL	MANAGE	PERIOD	RUBBER	SURVEY
ALASKA	CANADA	DIRECT	FORMED	MANNER	PICKET	RUNNER	SWITCH
ALLEGE	CANCEL	DIVERT	FOUGHT	MANUAL	PINCER	SALARY	SYSTEM
ALLIED	CANNOT	DIVIDE	FOURTH	MEAGER	PISTOL	SCHEME	TABLES
ALLIES	CANVAS	DOCTOR	FRIDAY	MEDIUM	PLACES	SCHOOL	TANKER
ALWAYS	CASUAL	DOLLAR	FUTURE	MEMBER	PLANES	SCORED	TARGET
ANIMAL	CAUSED	DOWNED	GARAGE	METHOD	POINTS	SCREEN	TATTOO
ANNUAL	CENTER	DRYRUN	GEORGE	METRIC	POISON	SEAMAN	TERROR
ANYWAY	CHANGE	DUGOUT	GREASE	MINING	POLICE	SEAMEN	THIRTY
APPEAR	CHARGE	DURING	GROUND	MINUTE	PONTON	SEARCH	THOUGH
ARABIA	CHEESE	EFFECT	GUNNER	MIRROR	POSTAL	SECOND	THREAT
ARMIES	CHURCH	EFFORT	HALTED	MOBILE	PREFER	SECTOR	TRAINS
ARMORY	CIPHER	EIGHTH	HAMMER	MONDAY	PROMPT	SECURE	TRENCH
ARREST	CIRCLE	EIGHTY	HAPPEN	MORALE	PROPER	SELECT	TROOPS
ARRIVE	COFFEE	EITHER	HARBOR	MORTAR	PURSUE	SERIAL	TURRET
ASSETS	COLORS	ELEVEN	HELPER	MOVING	RADIAL	SETTLE	TWELVE
ASSIST	COLUMN	EMBARK	HIGHER	MURDER	RAIDED	SEVERE	TWENTY
ASSURE	COMBAT	EMPLOY	HOURLY	MUZZLE	RATION	SHELLS	UNABLE
ATTACH	COMMIT	ENCODE	INDEED	NAUGHT	RAVINE	SIGCOM	UNITED
ATTACK	COMMON	ENGAGE	INFORM	NEARER	RECORD	SIGNAL	UNLESS
ATTAIN	CONVEY	ENGINE	INLAND	NINETY	REDUCE	SINGLE	VALLEY
AUGUST	CONVOY	ENROLL	INTEND	NORMAL	REFILL	SLIGHT	VERBAL
BANNER	COURSE	ENTIRE	INTENT	NOTING	REFUGE	SPHERE	VERIFY
BARBED	CREDIT	ERASER	INVENT	NOUGHT	REFUSE	SPOOLS	VESSEL
BARGES	CRISIS	ESCORT	ISLAND	NOVICE	REJECT	SPOONS	VICTIM
BATTEN	CRITIC	EUROPE	ISSUES	NOZZLE	RELIEF	STATES	VICTOR
BATTLE	DAMAGE	EXCEPT	KEEPER	NUMBER	REMAIN	STATUS	VISITS
BEETLE	DEBARK	EXCESS	KILLED	OCCUPY	REMEDY	STRAFE	VISUAL
BEFORE	DECIDE	EXCITE	LADDER	OFFEND	REPAIR	STREET	WEIGHT
BETTER	DECODE	EXPECT	LANDED	OFFICE	REPORT	STRESS	WIRING
BEYOND	DECREE	EXPELS	LAUNCH	OPPOSE	RESCUE	STRIPS	WITHIN
BILLET	DEFEAT	EXPEND	LEADER	ORDERS	RESIST	SUBMIT	WOODED
BITTER	DEFECT	EXTEND	LEAGUE	ORIENT	RESULT	SUDDEN	ZIGZAG
BODIES	DEFEND	EXTENT					

SEVEN LETTER WORDS

ABANDON	ALMANAC	APPOINT	ASIATIC	AVIATOR	BATTERY	BETWEEN
ABSENCE	AMMETER	APPROVE	ASSAULT	AWKWARD	BATTLES	BICYCLE
ADDRESS	ANALYZE	ARMORED	ATTACKS	BAGGAGE	BEARING	BINDING
ADVANCE	ANOTHER	ARRANGE	ATTEMPT	BALLOON	BECAUSE	BIVOUAC
AGAINST	ANTENNA	ARRIVAL	AVERAGE	BARRAGE	BEDDING	BOMBARD

SEVEN LETTER WORDS—Continued

BOMBERS	DEBOUCH	FITTING	LANDING	PACKAGE	REQUEST	SUPPOSE
BOMBING	DECIDED	FOGHORN	LEADING	PASSAGE	REQUIRE	SURPLUS
BOYCOTT	DECLARE	FORCING	LECTURE	PASSIVE	RESERVE	SUSPEND
BRIBERY	DECODED	FORGING	LIAISON	PATROLS	RESPECT	TACTICS
BRIGADE	DEFENSE	FORWARD	LIBRARY	PAYROLL	RESPOND	TALKING
CALIBER	DELAYED	FOXHOLE	LICENSE	PLACING	RETIRED	TARGETS
CALIBRE	DELIVER	FUELOIL	LIFTING	PLATOON	RETREAT	TERRAIN
CAPTAIN	DERRICK	FURNISH	LOADING	POUNDER	REVENUE	THATTHE
CAPTIVE	DESTROY	FURTHER	LOGICAL	PRAIRIE	REVERSE	THROUGH
CARRIER	DETRAIN	GASSING	LOOKOUT	PRECEDE	REVOLVE	TOBACCO
CAVALRY	DETRUCK	GENERAL	MACHINE	PREPARE	ROUTINE	TONIGHT
CENTRAL	DEVELOP	GETTING	MANDATE	PRESENT	RUNNING	TONNAGE
CHANGES	DIAGRAM	GLASSES	MANNING	PRESSED	SAILORS	TORPEDO
CHANNEL	DISCUSS	GRADUAL	MAPPING	PRIMARY	SATISFY	TRACTOR
CHARLIE	DISEASE	GRENADE	MARCHED	PROCEED	SECRECY	TRAFFIC
CHASSIS	DISMISS	GUARDED	MARSHAL	PROGRAM	SECTION	TRAWLER
CIRCUIT	DISTILL	HALTING	MARTIAL	PROMOTE	SECTORS	TRIGGER
CLIPPER	DROPPED	HASBEEN	MAXIMUM	PROPOSE	SERVICE	TUESDAY
COASTAL	EASTERN	HEADING	MEDICAL	PROTECT	SESSION	TWELFTH
COLLECT	ECHELON	HEAVIER	MESSAGE	PROTEST	SETBACK	UNKNOWN
COLLEGE	ELEMENT	HIGHEST	MESSING	PROVOST	SEVENTH	UNUSUAL
COLONEL	ELEVATE	HOLDING	MILITIA	PURPOSE	SEVENTY	USELESS
COMMAND	EMBASSY	HORIZON	MINIMUM	PURSUIT	SEVERAL	UTILITY
COMMEND	ENCODED	HOSTILE	MISFIRE	PUSHING	SHELLED	VACANCY
COMMENT	ENEMIES	HUNDRED	MISSING	QUARTER	SHORTLY	VARYING
COMMUTE	ENFORCE	ICEBERG	MISSION	QUICKLY	SIGNIFY	VESSELS
COMPANY	ENGAGED	ILLEGAL	MORNING	RADIATE	SIMILAR	VICTORY
COMPASS	ENTENTE	ILLNESS	NATURAL	RAIDING	SIMPLEX	VILLAGE
CONCEAL	ENTRAIN	INCLUDE	NEAREST	RAILWAY	SINKING	VISIBLE
CONDEMN	ENTRUCK	INFLICT	NIGHTLY	RAINING	SIXTEEN	VISITOR
CONDUCT	ENVELOP	INITIAL	NOTHING	RAPIDLY	SLOPING	WARFARE
CONFINE	EXCLUDE	INQUIRE	NUMBERS	REACHED	SMOKING	WARSHIP
CONTACT	EXPLAIN	INQUIRY	OBSERVE	RECEIPT	SOLDIER	WEATHER
CONTAIN	EXPRESS	INSPIRE	OCTOBER	RECEIVE	STARTER	WESTERN
CONTROL	EXTRACT	INSTALL	OFFENSE	RECOVER	STATION	WHETHER
CORRECT	EXTREME	INSTANT	OFFICER	RECRUIT	STEAMER	WILLIAM
COUNCIL	FALLING	INVADED	OMITTED	REDUCED	STOPPED	WINDAGE
COURIER	FARTHER	ISLANDS	OPERATE	REFUGEE	STORAGE	WITHOUT
COVERED	FEDERAL	ISSUING	OPINION	REGULAR	SUCCESS	WITHTHE
CROSSED	FIFTEEN	JANUARY	ORDERED	RELEASE	SUGGEST	WITNESS
CRUISER	FIGHTER	JUMPOFF	OUTPOST	RELIEVE	SUMMARY	WOUNDED
CURRENT	FILLING	KITCHEN	OUTSIDE	REPAIRS	SUNRISE	WRECKED
CYCLONE	FINDING	KILLING	PACIFIC	REPLACE	SUPPORT	WRITTEN
DAMAGED	FISHING					

EIGHT LETTER WORDS

ACTIVITY	ADVANCED	AIRBORNE	AIRPLANE	ANNOUNCE	APPROACH	ASSEMBLE
ACTUALLY	ADVANCES	AIRCRAFT	ALTITUDE	ANTITANK	APPROVAL	ASSEMBLY
ADJACENT	ADVISING	AIRDROME	AMERICAN	APPARENT	ARMAMENT	ASSIGNED
ADJUTANT	ADVISORY	AIRFIELD	ANALYSIS	APPEARED	ARRESTED	ASSOONAS

EIGHT LETTER WORDS—Continued

ATLANTIC
ATTACKED
ATTEMPTS
AVIATION
BARRACKS
BARRAGES
BATTERED
BATTLING
BESEIGED
BILLETED
BOUNDARY
BREAKING
BUILDING
BULLETIN
BUSINESS
CALAMITY
CAMPAIGN
CANISTER
CAPACITY
CAPTURED
CARELESS
CARRIAGE
CARRIERS
CARRYING
CASUALTY
CAUSEWAY
CEMETERY
CENTERED
CHAPLAIN
CHEMICAL
CIRCULAR
CITATION
CIVILIAN
CLERICAL
CODEBOOK
COMMANDS
COMMENCE
COMMERCE
COMPLETE
COMPOSED
CONCLUDE
CONCRETE
CONFLICT
CONGRESS
CONTINUE
CONTRACT
CORPORAL
CORRIDOR
COVERING
CRITICAL
CRITIQUE
CROSSING
CRUISERS
DAMAGING
DARKNESS
DAYLIGHT
DECEMBER
DECIPHER
DECISION
DECISIVE
DECLARED
DECREASE
DEDICATE
DEFEATED
DEFENDED
DEFENDER
DEFENSES
DEFERRED
DEFINITE
DELAYING
DEMANDED
DEPARTED
DEPLOYED
DEPORTED
DESCRIBE
DESERTED
DESERTER
DESPATCH
DETACHED
DETECTOR
DETONATE
DEVELOPE
DICTATED
DICTATOR
DIMINISH
DIRECTOR
DISARMED
DISASTER
DISLODGE
DISPATCH
DISPERSE
DISTANCE
DISTRESS
DISTRICT
DIVIDING
DIVISION
DOCTRINE
DOMINANT
DRESSING
DRIFTING
EASTERLY
EASTWARD
ECONOMIC
EFFECTED
EFFICACY
EIGHTEEN
ELEMENTS
ELEVENTH
ELIGIBLE
EMPLOYEE
EMPLOYER
ENCIPHER
ENCIRCLE
ENFILADE
ENGAGING
ENGINEER
ENLISTED
ENORMOUS
ENROLLED
ENTERING
ENTRENCH
ENVELOPE
EQUALIZE
EQUIPAGE
ESCORTED
ESTIMATE
EUROPEAN
EVACUATE
EXCAVATE
EXCHANGE
EXERCISE
EXPANDED
EXPEDITE
EXPELLED
EXPENDED
EXPENSES
EXTENDED
EXTERIOR
FACTIONS
FATALITY
FEBRUARY
FERRYING
FIGHTERS
FIGHTING
FINISHED
FLANKING
FLEXIBLE
FOOTHOLD
FORENOON
FORTRESS
FOURTEEN
FRONTAGE
FUSELAGE
GARRISON
GROUNDED
GROUPING
GUARDING
HAVEBEEN
HINDERED
HOSPITAL
HOWITZER
IDENTIFY
IGNITION
IMPROPER
IMPROVED
INCIDENT
INDICATE
INDIRECT
INFANTRY
INFECTED
INITIATE
INSECURE
INSIGNIA
INSTRUCT
INTEREST
INTERIOR
INTERNAL
INTRENCH
INVADING
INVASION
INVENTED
JETPLANE
JUNCTION
LANGUAGE
LATITUDE
LETTERED
LIMITING
LOCATION
LUMINOUS
MAINTAIN
MANDATED
MANEUVER
MARCHING
MARITIME
MATERIAL
MATERIEL
MECHANIC
MEDICINE
MEMORIAL
MERCIFUL
MESSAGES
MIDNIGHT
MILITARY
MISFIRES
MISSIONS
MOBILIZE
MONOPOLY
MOUNTAIN
MOVEMENT
NATIONAL
NAUTICAL
NINETEEN
NORTHERN
NOVEMBER
OBSERVED
OBSERVER
OBSOLETE
OBSTACLE
OCCUPIED
OFFENDED
OFFICERS
OFFICIAL
OPERATOR
OPPOSING
OPPOSITE
ORDINATE
ORDNANCE
OUTBOARD
OUTGUARD
OUTPOSTS
PAINTING
PARALLAX
PARALLEL
PASSPORT
PLANNING
POLITICS
PONTOONS
POSITION
POSITIVE
POSSIBLE
POSTPONE
PREPARED
PRESERVE
PRESSING
PRESSURE
PRINTING
PRIORITY
PRISONER
PROBABLE
PROBABLY
PROGRESS
PROHIBIT
PROTESTS
PROTOCOL
PURPOSES
QUARTERS
RAILHEAD
RAILROAD
RALLYING
RECEIVER
RECORDER
REDCROSS
REENLIST
REGIMENT
REGISTER
REJECTED
REJECTOR
REMEDIES
REMEMBER
REPAIRED
REPEATER
REPELLED
REPLACED
REPORTED
REPULSED
REQUIRED
RESEARCH
RESERVES
RESPECTS
RESTORED
RETIRING
RETURNED
REVIEWED
REVOLVER
RIGOROUS
SABOTAGE
SANITARY
SATURDAY
SCHEDULE
SEABORNE
SEALEVEL
SELECTED
SENTENCE
SENTINEL
SEPARATE
SERGEANT
SHELLING
SHIPPING
SIGHTING
SKIRMISH
SOLDIERS
SOUTHERN
SPECIFIC
SPOTTING
SQUADRON
STANDARD
STATIONS
STRATEGY
SUFFERED
SUITABLE
SUPERIOR
SUPPLIES
SURPRISE
SURROUND
SURVIVED
SUSPENSE
SWEEPING
SWIMMING
TACTICAL
TAXATION
TELEGRAM
TERRIBLE
TERRIFIC
THATHAVE
THIRTEEN
THOUSAND
THURSDAY
TOMORROW
TOTALING
TRAILERS
TRAINING
TRANSFER
TRAVERSE
TRAWLERS
VEHICLES
VICINITY
VIGOROUS
WARSHIPS
WESTERLY
WESTWARD
WINDWARD
WIRELESS
WITHDRAW
WITHDREW

NINE LETTER WORDS

ACCESSORY	CENTERING	DEVELOPED	FORMATION	MOVEMENTS	PROTECTOR
ACCOMPANY	CHALLENGE	DIETITIAN	FORTIFIED	MUNITIONS	PROTESTED
ACCORDING	CHARACTER	DIFFERENT	FRONTLINE	NAVALBASE	PROVISION
ADDRESSED	CHAUFFEUR	DIFFICULT	GROUPMENT	NECESSARY	PROXIMITY
ADDRESSES	CHRONICAL	DIMENSION	GYROMETER	NECESSITY	RADIATION
ADMISSION	CIGARETTE	DIRECTION	HOSTILITY	NEGLIGENT	RADIOGRAM
ADVANCING	CIRCULATE	DIRIGIBLE	HURRICANE	NEWSPAPER	READINESS
ADVANTAGE	CIVILIANS	DISAPPEAR	IDENTICAL	NORTHEAST	REARGUARD
AERODROME	CLEARANCE	DISCUSSED	IMMEDIATE	NORTHERLY	REBELLION
AEROPLANE	COALITION	DISINFECT	IMPORTANT	NORTHWARD	RECEIVING
AFTERNOON	COLLAPSED	DISMISSAL	IMPRESSED	NORTHWEST	RECOGNIZE
AGREEMENT	COLLISION	DISPERSED	INCENTIVE	NUMBERING	RECOMMEND
AIRDROMES	COMBATANT	DISTRICTS	INCIDENCE	OBJECTION	REENFORCE
AIRPLANES	COMMANDED	DIVISIONS	INCIDENTS	OBJECTIVE	REFERENCE
ALLOTMENT	COMMANDER	DOMINANCE	INCLINING	OBTAINING	REFILLING
ALLOWANCE	COMMITTEE	DOMINATED	INCLUDING	OCCUPYING	REGARDING
ALTERNATE	COMPANIES	ECHELONED	INCLUSIVE	OFFENSIVE	REINFORCE
AMBULANCE	COMPELLED	EFFECTIVE	INCREASED	OFFICIALS	REINSTATE
AMUSEMENT	COMPLETED	EFFICIENT	INDEMNITY	OPERATING	REMAINDER
ANNOUNCED	CONDEMNED	ELABORATE	INDICATED	OPERATION	REMAINING
ANONYMOUS	CONDENSED	ELEVATION	INFLATION	OSCILLATE	REPRESENT
APPARATUS	CONDITION	ELSEWHERE	INFLICTED	OUTSKIRTS	REPRISALS
APPOINTED	CONFERRED	EMBASSIES	INFLUENCE	PARACHUTE	REQUESTED
ARBITRARY	CONFIDENT	EMERGENCY	INHABITED	PARAGRAPH	REQUIRING
ARTILLERY	CONFLICTS	EMPLOYING	INSTANTLY	PARTITION	RESOURCES
ASCENSION	CONQUERED	ENDURANCE	INTEGRITY	PASSENGER	RESTRAINT
ASSAULTED	CONTINUAL	ENGINEERS	INTENSIVE	PATRIOTIC	RETENTION
ASSISTANT	CONTINUED	ENLISTING	INTENTION	PENETRATE	RETURNING
ASSOCIATE	CONTINUES	ENTRAINED	INTERCEPT	PERMANENT	REVIEWING
ASSURANCE	COOPERATE	EQUIPMENT	INTERDICT	PERSONNEL	SCREENING
ATTACKING	CORRECTED	ESTABLISH	INTERFERE	PLACEMENT	SEAPLANES
ATTEMPTED	CRITICISE	ESTIMATED	INTERMENT	POLITICAL	SECRETARY
ATTENTION	CRITICISM	ESTIMATES	INTERPOSE	POPULATED	SEMICOLON
AUTOMATIC	DEBARKING	EXCESSIVE	INTERRUPT	POSITIONS	SEMIRIGID
AVAILABLE	DECREASED	EXCLUSION	INTERVENE	PRACTICAL	SEPTEMBER
BALLISTIC	DEFECTIVE	EXCLUSIVE	INTERVIEW	PRECEDING	SERIOUSLY
BAROMETER	DEFENSIVE	EXECUTIVE	INVENTION	PREFERRED	SERVICING
BATTALION	DEFICIENT	EXERCISES	IRREGULAR	PREMATURE	SEVENTEEN
BATTERIES	DEPARTURE	EXHIBITED	KILOMETER	PREPARING	SHELLFIRE
BEACHHEAD	DEPENDENT	EXPANSION	LAUNCHING	PRESIDENT	SITUATION
BEGINNING	DESCRIBED	EXPANSIVE	LIABILITY	PRINCIPAL	SIXTEENTH
BLOCKADED	DESIGNATE	EXPENSIVE	LOGISTICS	PRINCIPLE	SOUTHEAST
BOMBARDED	DESTITUTE	EXPLOSION	LONGITUDE	PRISONERS	SOUTHWARD
BRIGADIER	DESTROYED	EXPLOSIVE	MAINTAINS	PROCEDURE	SOUTHWEST
BUILDINGS	DESTROYER	EXTENDING	MANGANESE	PROCEEDED	SPEARHEAD
CABLEGRAM	DETENTION	EXTENSION	MECHANISM	PROJECTOR	STANDARDS
CAMPAIGNS	DETERMINE	EXTENSIVE	MEMORANDA	PROMOTION	STATEMENT
CANCELLED	DETONATED	FIFTEENTH	MESSENGER	PROPOSALS	STRAGGLER
CARTRIDGE	DETRAINED	FIREALARM	MOTORIZED	PROTECTED	STRATEGIC

NINE LETTER WORDS—Continued

SUBMITTED	SUSPENDED	TELEPHONE	THEREFORE	UNTENABLE	WEDNESDAY
SUCCEEDED	SUSPICION	TENTATIVE	TRANSPORT	VARIATION	WITNESSES
SURRENDER	TECHNICAL	TERRITORY	TWENTIETH	WATERTANK	YESTERDAY
SUSPECTED	TECHNIQUE				

TEN LETTER WORDS

ACCEPTABLE	COLLISIONS	DESPATCHES	EXPENDABLE	MAINTAINED
ACCEPTANCE	COMMANDANT	DESTROYERS	EXPERIENCE	MANAGEMENT
ACCIDENTAL	COMMANDEER	DETACHMENT	EXPERIMENT	MECHANIZED
ACCORDANCE	COMMANDING	DETERMINED	EXPLOSIONS	MEMORANDUM
ACTIVITIES	COMMISSARY	DETONATION	EXTINGUISH	MILLIMETER
ADDITIONAL	COMMISSION	DETRAINING	FACILITIES	MOTORCYCLE
AIRCONTROL	COMMITMENT	DETRUCKING	FLASHLIGHT	NATURALIZE
AIRSUPPORT	COMMUNIQUE	DIFFERENCE	FORMATIONS	NAVIGATION
ALLEGIANCE	COMPENSATE	DIPLOMATIC	FOUNDATION	NEGLIGENCE
ALLOCATION	COMPLETELY	DIRECTIONS	FOURTEENTH	NEWSPAPERS
AMBASSADOR	COMPRESSED	DISCIPLINE	FRONTLINES	NINETEENTH
AMMUNITION	CONCERNING	DISCUSSION	GEOGRAPHIC	OBJECTIVES
ANTEDATING	CONCESSION	DISPATCHED	GONIOMETER	OCCUPATION
ANTICIPATE	CONCLUSION	DISPATCHER	GOVERNMENT	ONEHUNDRED
APPARENTLY	CONDITIONS	DISPATCHES	GYROSCOPIC	OPERATIONS
APPEARANCE	CONFERENCE	DISPERSION	HYDROMETER	OPPOSITION
APPROACHED	CONFESSION	DISTRESSED	HYGROMETER	OVERCOMING
ARMOREDCAR	CONFIDENCE	DISTRIBUTE	ILLITERATE	PATROLLING
ARTIFICIAL	CONNECTING	DIVEBOMBER	ILLUMINATE	PERMISSION
ASPOSSIBLE	CONNECTION	DOMINATION	ILLUSTRATE	PERSISTENT
ASSEMBLIES	CONSPIRACY	EFFICIENCY	IMPASSIBLE	PHOSPHORUS
ASSESSMENT	CONSTITUTE	EIGHTEENTH	IMPOSSIBLE	POPULATION
ASSIGNMENT	CONTINGENT	ELEMENTARY	IMPRESSION	POSSESSION
ASSISTANCE	CONTINUOUS	EMPLOYMENT	IMPRESSIVE	POSTOFFICE
ATOMICBOMB	CONTRABAND	ENCIPHERED	INCENDIARY	PRECEDENCE
ATTACHMENT	CONVENIENT	ENCIRCLING	INDICATING	PREFERENCE
ATTAINMENT	COORDINATE	ENEMYTANKS	INDICATION	PRESCRIBED
ATTEMPTING	CORRECTION	ENGAGEMENT	INDIVIDUAL	PROHIBITED
AUDIBILITY	CREDENTIAL	ENLISTMENT	INFLICTING	PROPORTION
AUTOMOBILE	CROSSROADS	ENROLLMENT	INSECURITY	PROTECTION
BALLISTICS	DEBOUCHING	ENTERPRISE	INSPECTION	PROVISIONS
BATTLESHIP	DECIPHERED	ENTRENCHED	INSTRUCTED	QUARANTINE
BEENNEEDED	DECORATION	ENTRUCKING	INSTRUCTOR	RECEPTACLE
BRIDGEHEAD	DEDICATION	EQUIVALENT	INSTRUMENT	RECREATION
CAMOUFLAGE	DEFICIENCY	ESTIMATION	INTERNMENT	RECRUITING
CAPABILITY	DEFINITION	EVACUATING	INVITATION	REENFORCED
CASUALTIES	DEMOBILIZE	EVACUATION	IRRIGATION	REENLISTED
CENSORSHIP	DEPARTMENT	EVALUATION	KILOMETERS	REGIMENTAL
CENTRALIZE	DEPENDABLE	EXCAVATION	LABORATORY	REGULATION
CIRCUITOUS	DEPLOYMENT	EXCITEMENT	LIEUTENANT	REINFORCED
COASTGUARD	DEPRESSION	EXHIBITION	LIMITATION	RESISTANCE
COLLECTING	DESIGNATED	EXPEDITING	LOCOMOTIVE	RESPECTFUL
COLLECTION	DESPATCHED	EXPEDITION	MACHINEGUN	RESTRICTED

TEN LETTER WORDS—Continued

REVOLUTION	SUBMISSION	SUSPENSION	TRANSPORTS	UNEXPENDED
SANITATION	SUBSTITUTE	SUSPICIONS	TRANSVERSE	UNSUITABLE
SEPARATION	SUCCESSFUL	SUSPICIOUS	TROOPSHIPS	VICTORIOUS
SIGNALLING	SUCCESSIVE	THIRTEENTH	TWENTYFIVE	VISIBILITY
SIMILARITY	SUFFICIENT	THREATENED	UNDERSTAND	WILLATTACK
STATISTICS	SUPPORTING	TRAJECTORY	UNDERSTOOD	WITHDRAWAL
SUBMARINES				

ELEVEN LETTER WORDS

ACCESSORIES	CONCENTRATE	EMPLACEMENT	INTERCERPTS	REAPPOINTED
AERONAUTICS	CONFINEMENT	ENCOUNTERED	INTERESTING	RECOGNITION
ALTERNATING	CONSTITUTED	ENEMYPLANES	INTERFERING	RECOMMENDED
APPLICATION	CONSUMPTION	ENFORCEMENT	INTERPRETER	RECONNOITER
APPOINTMENT	CONTINENTAL	ENGAGEMENTS	INTERRUPTED	REPLACEMENT
APPROACHING	CONTROVERSY	ENGINEERING	INTERVENING	REQUIREMENT
APPROPRIATE	COOPERATION	ESTABLISHED	INVESTIGATE	REQUISITION
APPROXIMATE	CORPORATION	ESTIMATEDAT	LEGISLATION	RESERVATION
ARBITRATION	CORRECTNESS	EXAMINATION	LIGHTBOMBER	RESIGNATION
ARMOREDCARS	CREDENTIALS	EXPLANATION	MAINTENANCE	RESPONSIBLE
ARRANGEMENT	CUSTOMHOUSE	EXTENSIVELY	MANUFACTURE	RESTRICTION
ASSESSMENTS	DEBARKATION	EXTERMINATE	MEASUREMENT	RETALIATION
ASSIGNMENTS	DEMONSTRATE	FINGERPRINT	NATIONALISM	RETROACTIVE
ASSOCIATION	DESCRIPTION	FIRECONTROL	NATIONALITY	SCHOOLHOUSE
BATTLEFIELD	DESCRIPTIVE	HEAVYBOMBER	NAVALATTACK	SEVENTEENTH
BATTLESHIPS	DESIGNATION	HEAVYLOSSES	NAVALBATTLE	SEVENTYFIVE
BELLIGERENT	DESTRUCTION	HOSTILITIES	NAVALFORCES	SIGNIFICANT
BLOCKBUSTER	DETERIORATE	IMMEDIATELY	NECESSITATE	SMOKESCREEN
BOMBARDMENT	DEVELOPMENT	IMMIGRATION	OBSERVATION	STRATEGICAL
CATASTROPHE	DISAPPEARED	IMPEDIMENTA	OVERWHELMED	SUBSISTENCE
CERTIFICATE	DISCONTINUE	IMPROVEMENT	PARENTHESIS	SUITABILITY
CIRCULATION	DISCREPANCY	INCOMPETENT	PARENTHESES	SUPERIORITY
COEFFICIENT	DISINFECTED	INDEPENDENT	PENETRATION	SURRENDERED
COINCIDENCE	DISPOSITION	INFLAMMABLE	PERFORMANCE	SYNCHRONIZE
COMMUNICATE	DISTINCTION	INFORMATION	PHILIPPINES	TEMPERATURE
COMMUNIQUES	DISTINGUISH	INSPIRATION	PHOTOGRAPHY	THERMOMETER
COMPARTMENT	DYNAMOMETER	INSTITUTION	PREARRANGED	TOPOGRAPHIC
COMPETITION	ECHELONMENT	INSTRUCTION	PREPARATION	TRADITIONAL
COMPOSITION	EFFECTIVELY	INSTRUMENTS	PRELIMINARY	TRANSFERRED
COMPUTATION	ELECTRICITY	INTELLIGENT	PROGRESSIVE	WITHDRAWING
CONCEALMENT	EMBARKATION	INTERCEPTED	RANGEFINDER	

TWELVE LETTER WORDS

ADVANTAGEOUS	CARELESSNESS	CONCENTRATED	CONSIDERABLE	COORDINATION
AGRICULTURAL	COMMENCEMENT	CONCILIATION	CONSTITUTING	DECENTRALIZE
ANNOUNCEMENT	COMMENDATION	CONFIDENTIAL	CONSTITUTION	DECIPHERMENT
ANTIAIRCRAFT	COMMISSIONED	CONFIRMATION	CONSTRUCTION	DEMONSTRATED
ANTICIPATION	COMMISSIONER	CONFISCATION	CONTINUATION	DEPARTMENTAL
BREAKTHROUGH	COMPENSATION	CONFORMATION	CONVALESCENT	DIFFICULTIES
CANCELLATION	COMPLETENESS	CONSCRIPTION	CONVERSATION	DISORGANIZED

TWELVE LETTER WORDS—Continued

DISPLACEMENT
DISSEMINATED
DISTRIBUTING
DISTRIBUTION
EMPLACEMENTS
ENCIPHERMENT
ENTANGLEMENT
ENTERPRISING
FIGHTERPLANE
GENERALALARM
GENERALSTAFF
GEOGRAPHICAL
HEADQUARTERS
HEAVYBOMBERS
HYDROGRAPHIC
ILLUMINATING
ILLUMINATION
ILLUSTRATION
INAUGURATION
INCOMPETENCE
INEFFICIENCY
INSTRUCTIONS
INTELLIGENCE
INTERDICTION
INTERFERENCE
INTERMEDIATE
INTERRUPTION
INTERVENTION
INTRODUCTION
INTRODUCTORY
IRREGULARITY
LIGHTBOMBERS
MARKSMANSHIP
MEASUREMENTS
MEDIUMBOMBER
MOBILIZATION
NONCOMBATANT
NORTHWESTERN
OBSTRUCTIONS
ORGANIZATION
PREPARATIONS
PREPAREDNESS
PRESERVATION
PRESIDENTIAL
PROCLAMATION
PSYCHROMETER
RADIOSTATION
RECREATIONAL
REENLISTMENT
REGISTRATION
REPLACEMENTS
RESPECTFULLY
ROADJUNCTION
SATISFACTORY
SEARCHLIGHTS
SHARPSHOOTER
SIGNIFICANCE
SIMULTANEOUS
SOUTHWESTERN
SUBSTITUTION
SUCCESSFULLY
TRANSFERRING
TRANSMISSION
TRANSPACIFIC
UNIDENTIFIED
UNITEDSTATES
UNSUCCESSFUL
VERIFICATION
VETERINARIAN

THIRTEEN LETTER WORDS

ACCOMMODATION
APPROXIMATELY
CHRONOLOGICAL
CIRCUMSTANCES
COMMUNICATION
CONCENTRATING
CONCENTRATION
CONGRESSIONAL
CONSIDERATION
CORRESPONDING
COUNTERATTACK
DECENTRALIZED
DEMONSTRATION
DEPENDABILITY
DETERMINATION
DISAPPEARANCE
DISCREPANCIES
DISSEMINATION
DISTINGUISHED
ENTERTAINMENT
ESTABLISHMENT
EXTERMINATION
EXTRAORDINARY
FIGHTERPLANES
IMPRACTICABLE
INDETERMINATE
INSTALLATIONS
INSTANTANEOUS
INTERNATIONAL
INVESTIGATION
MEDIUMBOMBERS
MISCELLANEOUS
PRELIMINARIES
QUALIFICATION
QUARTERMASTER
REAPPOINTMENT
REENFORCEMENT
REIMBURSEMENT
REINFORCEMENT
REINSTATEMENT
REVOLUTIONARY
SPECIFICATION
TRANSATLANTIC
WARDEPARTMENT

FOURTEEN LETTER WORDS

ADMINISTRATION
ADMINISTRATIVE
CENTRALIZATION
CHARACTERISTIC
CIRCUMSTANTIAL
CLASSIFICATION
CORRESPONDENCE
DEMOBILIZATION
DISCONTINUANCE
DISTINGUISHING
IDENTIFICATION
INTERPRETATION
INVESTIGATIONS
IRREGULARITIES
METEOROLOGICAL
NATURALIZATION
RECOMMENDATION
RECONNAISSANCE
RECONNOITERING
RECONSTRUCTION
REORGANIZATION
REPRESENTATIVE
RESPONSIBILITY
SATISFACTORILY
TRANSPORTATION

B. LIST OF WORDS USED IN MILITARY TEXT ARRANGED IN RHYMING ORDER ACCORDING TO WORD LENGTH

THREE LETTER WORDS

SEA	SEE	MAJ	TAN	TOP	EAT	APT	TAX
JOB	AGE	ADJ	GEN	GHQ	MAT	BUT	FIX
ROB	SHE	ASK	MEN	BAR	VAT	CUT	MIX
TUB	THE	GAL	PEN	CAR	ACT	OUT	SIX
QMC	DIE	ALL	TEN	FAR	GET	PUT	BOX
ARC	ONE	ILL	PIN	PAR	LET	PVT	FOX
BAD	ARE	COL	TIN	WAR	NET	CWT	DAY
HAD	USE	CPL	TON	HER	SET	YOU	LAY
ADD	DUE	CAM	WON	PER	WET	CAV	PAY
RED	OWE	HAM	DUN	AIR	YET	LAW	SAY
AID	EYE	AIM	GUN	FOR	SGT	SAW	WAY
BID	OFF	HIM	RUN	OUR	WGT	FEW	ANY
DID	BAG	ARM	SUN	GAS	FIT	NEW	SPY
RID	KEG	SUM	OWN	HAS	GOT	HOW	DRY
OLD	BIG	CAN	AGO	WAS	LOT	LOW	TRY
AND	JIG	MAN	TOO	HIS	NOT	NOW	BUY
END	DOG	NAN	TWO	ITS			

FOUR LETTER WORDS

AREA	MIKE	BASE	WEEK	FELL	JOIN	PASS	LIST
ASIA	YOKE	FUSE	TALK	WELL	NOON	LESS	LOST
BULB	ABLE	DATE	BULK	HILL	SOON	MESS	POST
BOMB	FILE	LATE	RANK	WILL	DOWN	LOSS	JUST
HEAD	MILE	NOTE	SANK	FULL	TOWN	HITS	ROUT
LEAD	MULE	BLUE	TANK	TOOL	ZERO	DAYS	NEXT
LOAD	RULE	HAVE	SUNK	TEAM	ALSO	MEAT	TEXT
ROAD	SAME	FIVE	BOOK	THEM	INTO	THAT	LIEU
RAID	TIME	LOVE	COOK	ITEM	KEEP	WHAT	DRAW
SAID	SOME	MOVE	HOOK	MAIM	SHIP	FEET	XRAY
HOLD	LINE	FUZE	LOOK	FROM	DUMP	MEET	AWAY
HAND	MINE	HALF	TOOK	FARM	PUMP	LEFT	BODY
LAND	NINE	FLAG	DARK	FIRM	STOP	OMIT	THEY
KIND	ZONE	KING	PARK	FORM	NEAR	UNIT	ALLY
HARD	JUNE	LONG	MASK	THAN	REAR	HALT	ONLY
HERD	OBOE	EACH	TASK	PLAN	OVER	TENT	JULY
ONCE	PIPE	HIGH	ORAL	BEEN	FOUR	SHOT	ARMY
MADE	TYPE	DASH	MTCL	SEEN	EYES	RIOT	MANY
AIDE	TARE	PUSH	FEEL	THEN	THIS	DIRT	VARY
SIDE	HERE	RUSH	RAIL	WHEN	AXIS	EAST	VERY
CODE	WERE	WITH	CALL	OPEN	TONS	FAST	EASY
FLEE	FIRE	BOTH	FALL	MAIN	GUNS	LAST	CITY
EDGE	WIRE	LEAK	CELL	RAIN	MASS	WEST	NAVY
TAKE	MORE	BACK					

FIVE LETTER WORDS

COMMA	SCALE	ALONG	CANAL	WAGON	PRIOR	DRESS	START
ZEBRA	TITLE	AMONG	FATAL	UNION	MAJOR	PRESS	ALERT
SQUAD	ALINE	BEACH	VITAL	COLON	VALOR	CROSS	LEAST
SPEED	SLOPE	REACH	TOTAL	DRAWN	ARMOR	FLATS	COAST
WIPED	FLARE	WHICH	EQUAL	RADIO	HONOR	BOATS	CREST
RIGID	THERE	MARCH	USUAL	EQUIP	ERROR	RAFTS	GUEST
RAPID	WHERE	WEIGH	NAVAL	TROOP	MOTOR	UNITS	FIRST
FIELD	SHORE	FRESH	WHEEL	GROUP	AREAS	TRACT	BURST
BLIND	CEASE	WIDTH	STEEL	CLEAR	BOMBS	FLEET	ABOUT
GUARD	ERASE	FIFTH	REPEL	SUGAR	RAIDS	QUIET	ALLOW
AWARD	THESE	TENTH	LEVEL	UNDER	WOODS	ASSET	ANNEX
THIRD	CLOSE	NINTH	APRIL	ORDER	YARDS	SHIFT	TODAY
BRIBE	HORSE	BOOTH	SMALL	DEFER	MILES	EIGHT	DELAY
PLACE	CAUSE	DEPTH	SHELL	REFER	FIRES	FIGHT	READY
VOICE	HOUSE	NORTH	SPELL	EAGER	CASES	LIGHT	FOGGY
FORCE	ROUTE	SOUTH	DRILL	ROGER	GATES	NIGHT	DAILY
TRUCE	ISSUE	SIXTH	ALARM	ETHER	PACKS	RIGHT	RALLY
THREE	LEAVE	BREAK	JAPAN	OTHER	DECKS	SIGHT	APPLY
RIDGE	DRIVE	BLACK	QUEEN	BAKER	DOCKS	AWAIT	EARLY
SIEGE	PROVE	CHECK	TAKEN	LATER	BANKS	SPLIT	ENEMY
RANGE	CURVE	QUICK	SEVEN	METER	TANKS	LIMIT	EVERY
BARGE	SEIZE	TRUCK	GIVEN	PETER	PLANS	VISIT	FERRY
LARGE	CHIEF	CREEK	ALIGN	AFTER	SHIPS	AGENT	FIFTY
GAUGE	STAFF	FLANK	AGAIN	ENTER	CORPS	POINT	PARTY
STAKE	PROOF	CLERK	PLAIN	RIVER	FEARS	FRONT	FORTY
SMOKE	BEING	LOCAL	TRAIN	COVER	PAIRS	COUNT	SIXTY
BROKE	GOING	VOCAL	BEGIN	THEIR	HOURS	DEPOT	HEAVY

SIX LETTER WORDS

CANADA	HALTED	DEVICE	CHARGE	SEVERE	ARRIVE	TRENCH	MANUAL
ARABIA	ROUTED	NOVICE	GEORGE	RETIRE	ACTIVE	LAUNCH	ANNUAL
ALASKA	LIQUID	FIERCE	REFUGE	ENTIRE	TWELVE	SEARCH	CASUAL
PANAMA	INLAND	REDUCE	MORALE	BEFORE	BREEZE	CHURCH	VISUAL
METRIC	ISLAND	PARADE	UNABLE	SECURE	RELIEF	SWITCH	CANCEL
CRITIC	DEFEND	DECIDE	CIRCLE	ASSURE	ZIGZAG	THOUGH	VESSEL
BOMBED	OFFEND	DIVIDE	SINGLE	FUTURE	RIDING	FINISH	DETAIL
BARBED	DEPEND	DECODE	MOBILE	GREASE	FILING	EIGHTH	REFILL
RAIDED	EXPEND	ENCODE	BEETLE	CHEESE	LINING	FOURTH	ENROLL
LANDED	INTEND	COFFEE	BATTLE	ADVISE	MINING	ATTACK	SCHOOL
WOODED	EXTEND	DECREE	SETTLE	DEVISE	FIRING	DEBARK	PATROL
INDEED	SECOND	DEGREE	LITTLE	OPPOSE	WIRING	EMBARK	PISTOL
ALLIED	BEYOND	STRAFE	NOZZLE	COURSE	DURING	VERBAL	SYSTEM
KILLED	GROUND	ENGAGE	MUZZLE	REFUSE	NOTING	RADIAL	VICTIM
FORMED	METHOD	DAMAGE	SCHEME	LOCATE	MOVING	SERIAL	SIGCOM
DOWNED	PERIOD	MANAGE	RESUME	EXCITE	FLYING	ANIMAL	BOTTOM
SCORED	RECORD	GARAGE	ENGINE	MINUTE	BREACH	FORMAL	INFORM
PASSED	OFFICE	BRIDGE	RAVINE	RESCUE	DETACH	NORMAL	MEDIUM
CAUSED	POLICE	ALLEGE	EUROPE	LEAGUE	ATTACH	SIGNAL	SUDDEN
UNITED	ADVICE	CHANGE	SPHERE	PURSUE	BRANCH	POSTAL	SCREEN

SIX LETTER WORDS—Continued

SUNKEN	MORTAR	RUNNER	FORCES	COLORS	TARGET	CANNOT	MONDAY
BROKEN	RUBBER	KEEPER	BARGES	ACCESS	PICKET	ACCEPT	SUNDAY
SEAMEN	MEMBER	HELPER	BODIES	EXCESS	ROCKET	EXCEPT	ANYWAY
HAPPEN	BOMBER	PROPER	ALLIES	UNLESS	BILLET	PROMPT	REMEDY
BATTEN	NUMBER	NEARER	ARMIES	STRESS	TURRET	DEPART	VALLEY
ELEVEN	PINCER	ERASER	TABLES	ACROSS	SUNSET	DESERT	PARLEY
REMAIN	LEADER	CENTER	PLANES	ASSETS	WEIGHT	DIVERT	CONVEY
ATTAIN	LADDER	BETTER	PASSES	VISITS	FLIGHT	ESCORT	SURVEY
WITHIN	MURDER	LETTER	LOSSES	POINTS	SLIGHT	EFFORT	VERIFY
COLUMN	PREFER	BITTER	STATES	STATUS	NAUGHT	REPORT	SUPPLY
RATION	SUFFER	LITTER	ROUTES	ALWAYS	FOUGHT	ARREST	HOURLY
ACTION	MEAGER	AFFAIR	ISSUES	COMBAT	NOUGHT	RESIST	DEPLOY
COMMON	HIGHER	REPAIR	CRISIS	DEFEAT	CREDIT	ASSIST	EMPLOY
SUMMON	CIPHER	HARBOR	SHELLS	THREAT	SUBMIT	AUGUST	CONVOY
POISON	EITHER	TERROR	SPOOLS	DEFECT	COMMIT	ADJUST	OCCUPY
LESSON	TANKER	MIRROR	TRAINS	EFFECT	SUMMIT	DUGOUT	SALARY
PONTON	HAMMER	SECTOR	SPOONS	REJECT	RESULT	OUTPUT	ARMORY
RETURN	SUMMER	VICTOR	STRIPS	SELECT	ORIENT	BUREAU	NINETY
DRYRUN	BANNER	DOCTOR	TROOPS	EXPECT	INTENT	REVIEW	EIGHTY
TATTOO	MANNER	CANVAS	ORDERS	DIRECT	EXTENT	FOLLOW	TWENTY
APPEAR	GUNNER	PLACES	OTHERS	STREET	INVENT	FRIDAY	THIRTY
DOLLAR							

SEVEN LETTER WORDS

MILITIA	COVERED	REFUGEE	WARFARE	PROMOTE	FORGING	VARYING
ANTENNA	RETIRED	WINDAGE	DECLARE	COMMUTE	FISHING	ICEBERG
ALMANAC	ARMORED	BAGGAGE	PREPARE	REVENUE	PUSHING	DEBOUCH
BIVOUAC	PRESSED	PACKAGE	CALIBRE	RELIEVE	NOTHING	THROUGH
TRAFFIC	CROSSED	VILLAGE	MISFIRE	RECEIVE	TALKING	FURNISH
PACIFIC	OMITTED	TONNAGE	INSPIRE	PASSIVE	SINKING	TWELFTH
ASIATIC	DELAYED	AVERAGE	REQUIRE	CAPTIVE	SMOKING	SEVENTH
REDUCED	COMMAND	STORAGE	INQUIRE	REVOLVE	FALLING	SETBACK
INVADED	COMMEND	BARRAGE	LECTURE	APPROVE	FILLING	DERRICK
DECIDED	SUSPEND	PASSAGE	RELEASE	OBSERVE	KILLING	DETRUCK
DECODED	RESPOND	MESSAGE	DISEASE	RESERVE	RAINING	ENTRUCK
ENCODED	BOMBARD	COLLEGE	SUNRISE	ANALYZE	MANNING	MEDICAL
WOUNDED	AWKWARD	ARRANGE	LICENSE	JUMPOFF	RUNNING	LOGICAL
GUARDED	FORWARD	WITHTHE	DEFENSE	BOMBING	MORNING	CONCEAL
PROCEED	REPLACE	THATTHE	OFFENSE	PLACING	SLOPING	ILLEGAL
ENGAGED	SERVICE	CHARLIE	PROPOSE	FORCING	MAPPING	MARSHAL
DAMAGED	ADVANCE	PRAIRIE	SUPPOSE	HEADING	BEARING	INITIAL
REACHED	ABSENCE	VISIBLE	PURPOSE	LEADING	GASSING	MARTIAL
MARCHED	ENFORCE	BICYCLE	REVERSE	LOADING	MESSING	FEDERAL
WRECKED	BRIGADE	HOSTILE	BECAUSE	BEDDING	MISSING	GENERAL
SHELLED	GRENADE	EXTREME	MANDATE	RAIDING	LIFTING	SEVERAL
DROPPED	PRECEDE	CONFINE	RADIATE	HOLDING	HALTING	CENTRAL
STOPPED	OUTSIDE	MACHINE	OPERATE	LANDING	GETTING	NATURAL
HUNDRED	INCLUDE	ROUTINE	ELEVATE	BINDING	FITTING	COASTAL
ORDERED	EXCLUDE	CYCLONE	ENTENTE	FINDING	ISSUING	GRADUAL

SEVEN LETTER WORDS—Continued

UNUSUAL	ENTRAIN	ENVELOP	STARTER	SUCCESS	ASSAULT	RAILWAY
ARRIVAL	CONTAIN	SIMILAR	QUARTER	USELESS	INSTANT	SECRECY
CHANNEL	CAPTAIN	REGULAR	DELIVER	ILLNESS	ELEMENT	VACANCY
COLONEL	CONDEMN	CALIBER	RECOVER	WITNESS	COMMENT	SIGNIFY
COUNCIL	ABANDON	OCTOBER	AVIATOR	ADDRESS	CURRENT	SATISFY
FUELOIL	OPINION	OFFICER	TRACTOR	EXPRESS	PRESENT	RAPIDLY
INSTALL	SESSION	POUNDER	VISITOR	DISMISS	APPOINT	QUICKLY
DISTILL	MISSION	TRIGGER	TACTICS	DISCUSS	RECEIPT	NIGHTLY
PAYROLL	STATION	WEATHER	ISLANDS	TARGETS	ATTEMPT	SHORTLY
CONTROL	SECTION	WHETHER	CHANGES	SURPLUS	SUPPORT	COMPANY
WILLIAM	ECHELON	ANOTHER	ENEMIES	RETREAT	SUGGEST	DESTROY
DIAGRAM	BALLOON	FARTHER	BATTLES	EXTRACT	HIGHEST	PRIMARY
PROGRAM	PLATOON	FURTHER	GLASSES	CONTACT	NEAREST	SUMMARY
MINIMUM	LIAISON	SOLDIER	CHASSIS	COLLECT	PROTEST	LIBRARY
MAXIMUM	HORIZON	CARRIER	ATTACKS	RESPECT	REQUEST	JANUARY
HASBEEN	EASTERN	COURIER	VESSELS	CORRECT	AGAINST	BRIBERY
FIFTEEN	WESTERN	HEAVIER	PATROLS	PROTECT	OUTPOST	BATTERY
SIXTEEN	FOGHORN	TRAWLER	BOMBERS	INFLICT	PROVOST	INQUIRY
BETWEEN	UNKNOWN	STEAMER	NUMBERS	CONDUCT	BOYCOTT	CAVALRY
KITCHEN	TOBACCO	CLIPPER	REPAIRS	TONIGHT	WITHOUT	VICTORY
WRITTEN	TORPEDO	CRUISER	SAILORS	CIRCUIT	LOOKOUT	EMBASSY
EXPLAIN	WARSHIP	AMMETER	SECTORS	RECRUIT	SIMPLEX	UTILITY
TERRAIN	DEVELOP	FIGHTER	COMPASS	PURSUIT	TUESDAY	SEVENTY
DETRAIN						

EIGHT LETTER WORDS

INSIGNIA	EXPELLED	DICTATED	STANDARD	LANGUAGE	ENVELOPE	OPPOSITE
SPECIFIC	ENROLLED	EFFECTED	OUTBOARD	DISLODGE	INSECURE	CONTINUE
TERRIFIC	DISARMED	INFECTED	OUTGUARD	EXCHANGE	PRESSURE	CRITIQUE
ECONOMIC	ASSIGNED	REJECTED	WINDWARD	PROBABLE	DECREASE	THATHAVE
MECHANIC	RETURNED	SELECTED	EASTWARD	SUITABLE	EXERCISE	DECISIVE
ATLANTIC	APPEARED	BILLETED	WESTWARD	ELIGIBLE	SURPRISE	POSITIVE
RAILHEAD	DECLARED	INVENTED	DESCRIBE	TERRIBLE	SUSPENSE	PRESERVE
RAILROAD	PREPARED	DEPARTED	ORDNANCE	POSSIBLE	DISPERSE	EQUALIZE
REPLACED	HINDERED	DESERTED	DISTANCE	FLEXIBLE	TRAVERSE	MOBILIZE
ADVANCED	SUFFERED	ESCORTED	COMMENCE	ASSEMBLE	DEDICATE	INVADING
DEMANDED	CENTERED	DEPORTED	SENTENCE	OBSTACLE	INDICATE	DIVIDING
EXPANDED	BATTERED	REPORTED	ANNOUNCE	ENCIRCLE	INITIATE	BUILDING
DEFENDED	LETTERED	ARRESTED	COMMERCE	SCHEDULE	ESTIMATE	GUARDING
OFFENDED	REPAIRED	ENLISTED	ENFILADE	MARITIME	ORDINATE	ENGAGING
EXPENDED	REQUIRED	SURVIVED	CONCLUDE	AIRDROME	DETONATE	DAMAGING
EXTENDED	RESTORED	IMPROVED	LATITUDE	AIRPLANE	SEPARATE	MARCHING
GROUNDED	DEFERRED	OBSERVED	ALTITUDE	JETPLANE	EVACUATE	BREAKING
BESIEGED	CAPTURED	REVIEWED	EMPLOYEE	MEDICINE	EXCAVATE	FLANKING
DETACHED	REPULSED	DEPLOYED	CARRIAGE	DOCTRINE	OBSOLETE	TOTALING
FINISHED	COMPOSED	AIRFIELD	FUSELAGE	POSTPONE	COMPLETE	SHELLING
OCCUPIED	MANDATED	FOOTHOLD	EQUIPAGE	SEABORNE	CONCRETE	BATTLING
ATTACKED	DEFEATED	THOUSAND	FRONTAGE	AIRBORNE	EXPEDITE	SWIMMING
REPELLED	REPEATED	SURROUND	SABOTAGE	DEVELOPE	DEFINITE	TRAINING

EIGHT LETTER WORDS—Continued

PLANNING	ELEVENTH	CAMPAIGN	PRISONER	VEHICLES	RESPECTS	WITHDRAW
SWEEPING	ANTITANK	CHAPLAIN	IMPROPER	MISFIRES	ELEMENTS	WITHDREW
SHIPPING	CODEBOOK	MAINTAIN	REPEATER	DEFENSES	ATTEMPTS	TOMORROW
GROUPING	CHEMICAL	MOUNTAIN	DESERTER	EXPENSES	PROTESTS	PARALLAX
ENTERING	CLERICAL	BULLETIN	DISASTER	PURPOSES	OUTPOSTS	SATURDAY
COVERING	TACTICAL	INVASION	REGISTER	RESERVES	ENORMOUS	THURSDAY
RETIRING	CRITICAL	DECISION	CANISTER	ANALYSIS	LUMINOUS	CAUSEWAY
ADVISING	NAUTICAL	DIVISION	RECEIVER	BARRACKS	RIGOROUS	EFFICACY
OPPOSING	OFFICIAL	LOCATION	REVOLVER	MISSIONS	VIGOROUS	IDENTIFY
DRESSING	MATERIAL	AVIATION	OBSERVER	STATIONS	CONTRACT	STRATEGY
PRESSING	MEMORIAL	CITATION	MANEUVER	FACTIONS	INDIRECT	PROBABLY
CROSSING	NATIONAL	TAXATION	EMPLOYER	PONTOONS	CONFLICT	ASSEMBLY
DRIFTING	INTERNAL	JUNCTION	HOWITZER	WARSHIPS	DISTRICT	ACTUALLY
FIGHTING	CORPORAL	IGNITION	CORRIDOR	OFFICERS	INSTRUCT	MONOPOLY
SIGHTING	HOSPITAL	POSITION	SUPERIOR	SOLDIERS	AIRCRAFT	EASTERLY
LIMITING	APPROVAL	FORENOON	INTERIOR	CARRIERS	DAYLIGHT	WESTERLY
PAINTING	MATERIEL	SQUADRON	EXTERIOR	TRAILERS	MIDNIGHT	BOUNDARY
PRINTING	PARALLEL	GARRISON	OPERATOR	TRAWLERS	PROHIBIT	MILITARY
SPOTTING	SENTINEL	NORTHERN	DICTATOR	CRUISERS	SERGEANT	SANITARY
DELAYING	SEALEVEL	SOUTHERN	REJECTOR	FIGHTERS	DOMINANT	FEBRUARY
RALLYING	PROTOCOL	CIRCULAR	DIRECTOR	QUARTERS	ADJUTANT	CEMETERY
CARRYING	MERCIFUL	DECEMBER	DETECTOR	CARELESS	ADJACENT	ADVISORY
FERRYING	TELEGRAM	REMEMBER	ASSOONAS	WIRELESS	INCIDENT	INFANTRY
APPROACH	AMERICAN	NOVEMBER	POLITICS	BUSINESS	ARMAMENT	CAPACITY
ENTRENCH	EUROPEAN	DEFENDER	COMMANDS	DARKNESS	MOVEMENT	FATALITY
INTRENCH	CIVILIAN	RECORDER	ADVANCES	CONGRESS	REGIMENT	CALAMITY
RESEARCH	HAVEBEEN	ENGINEER	BARRAGES	PROGRESS	APPARENT	VICINITY
DESPATCH	NINETEEN	TRANSFER	MESSAGES	FORTRESS	PASSPORT	PRIORITY
DISPATCH	EIGHTEEN	DECIPHER	REMEDIES	DISTRESS	INTEREST	ACTIVITY
SKIRMISH	THIRTEEN	ENCIPHER	SUPPLIES	REDCROSS	REENLIST	CASUALTY
DIMINISH	FOURTEEN					

NINE LETTER WORDS

MEMORANDA	CANCELLED	IMPRESSED	ATTEMPTED	ASSURANCE	AERODROME
STRATEGIC	COMPELLED	DISCUSSED	PROTESTED	ALLOWANCE	HURRICANE
AUTOMATIC	DETRAINED	INDICATED	REQUESTED	INCIDENCE	AEROPLANE
PATRIOTIC	ENTRAINED	POPULATED	SUBMITTED	REFERENCE	INTERVENE
BALLISTIC	CONDEMNED	ESTIMATED	CONTINUED	INFLUENCE	FRONTLINE
BEACHHEAD	ECHELONED	DOMINATED	DESTROYED	REENFORCE	DETERMINE
SPEARHEAD	DEVELOPED	DETONATED	MOTORIZED	REINFORCE	TELEPHONE
DESCRIBED	CONQUERED	SUSPECTED	SEMIRIGID	LONGITUDE	INTERFERE
ANNOUNCED	PREFERRED	CORRECTED	RECOMMEND	COMMITTEE	ELSEWHERE
BLOCKADED	CONFERRED	PROTECTED	REARGUARD	ADVANTAGE	SHELLFIRE
SUCCEEDED	DECREASED	INFLICTED	NORTHWARD	CARTRIDGE	THEREFORE
PROCEEDED	INCREASED	COMPLETED	SOUTHWARD	CHALLENGE	PROCEDURE
COMMANDED	CONDENSED	INHABITED	AMBULANCE	AVAILABLE	PREMATURE
SUSPENDED	COLLAPSED	EXHIBITED	DOMINANCE	UNTENABLE	DEPARTURE
BOMBARDED	DISPERSED	ASSAULTED	CLEARANCE	DIRIGIBLE	NAVALBASE
FORTIFIED	ADDRESSED	APPOINTED	ENDURANCE	PRINCIPLE	MANGANESE

NINE LETTER WORDS—Continued

CRITICISE	REGARDING	PERSONNEL	INVENTION	CONTINUES	STATEMENT
INTERPOSE	ACCORDING	CABLEGRAM	PROMOTION	BUILDINGS	EQUIPMENT
ASSOCIATE	INCLUDING	RADIOGRAM	SEMICOLON	OFFICIALS	GROUPMENT
IMMEDIATE	LAUNCHING	FIREALARM	AFTERNOON	REPRISALS	INTERMENT
OSCILLATE	ATTACKING	CRITICISM	DISAPPEAR	PROPOSALS	ALLOTMENT
CIRCULATE	DEBARKING	MECHANISM	IRREGULAR	CIVILIANS	PERMANENT
DESIGNATE	REFILLING	DIETITIAN	SEPTEMBER	CAMPAIGNS	DIFFERENT
ALTERNATE	SCREENING	SEVENTEEN	COMMANDER	MAINTAINS	REPRESENT
COOPERATE	REMAINING	SUSPICION	SURRENDER	DIVISIONS	RESTRAINT
ELABORATE	OBTAINING	BATTALION	REMAINDER	MUNITIONS	INTERCEPT
PENETRATE	INCLINING	REBELLION	PASSENGER	POSITIONS	INTERRUPT
REINSTATE	BEGINNING	COLLISION	MESSENGER	ENGINEERS	TRANSPORT
CIGARETTE	RETURNING	PROVISION	BRIGADIER	PRISONERS	NORTHEAST
PARACHUTE	PREPARING	EXPANSION	STRAGGLER	READINESS	SOUTHEAST
DESTITUTE	NUMBERING	ASCENSION	NEWSPAPER	CONFLICTS	NORTHWEST
TECHNIQUE	CENTERING	DIMENSION	CHARACTER	DISTRICTS	SOUTHWEST
EXPANSIVE	REQUIRING	EXTENSION	KILOMETER	INCIDENTS	INTERVIEW
DEFENSIVE	OPERATING	EXPLOSION	BAROMETER	MOVEMENTS	YESTERDAY
OFFENSIVE	ENLISTING	ADMISSION	GYROMETER	OUTSKIRTS	WEDNESDAY
EXPENSIVE	RECEIVING	EXCLUSION	DESTROYER	ANONYMOUS	EMERGENCY
INTENSIVE	REVIEWING	RADIATION	PROJECTOR	APPARATUS	NORTHERLY
EXTENSIVE	EMPLOYING	VARIATION	PROTECTOR	DISINFECT	SERIOUSLY
EXPLOSIVE	OCCUPYING	INFLATION	CHAUFFEUR	INTERDICT	INSTANTLY
EXCESSIVE	PARAGRAPH	FORMATION	LOGISTICS	DIFFICULT	ACCOMPANY
INCLUSIVE	ESTABLISH	OPERATION	STANDARDS	COMBATANT	ARBITRARY
EXCLUSIVE	TWENTIETH	SITUATION	RESOURCES	IMPORTANT	NECESSARY
TENTATIVE	FIFTEENTH	ELEVATION	COMPANIES	ASSISTANT	SECRETARY
DEFECTIVE	SIXTEENTH	OBJECTION	BATTERIES	CONFIDENT	ARTILLERY
EFFECTIVE	WATERTANK	DIRECTION	EMBASSIES	PRESIDENT	ACCESSORY
OBJECTIVE	TECHNICAL	CONDITION	AIRDROMES	DEPENDENT	TERRITORY
INCENTIVE	CHRONICAL	COALITION	SEAPLANES	NEGLIGENT	LIABILITY
EXECUTIVE	PRACTICAL	PARTITION	AIRPLANES	DEFICIENT	HOSTILITY
RECOGNIZE	POLITICAL	DETENTION	EXERCISES	EFFICIENT	PROXIMITY
SERVICING	IDENTICAL	RETENTION	WITNESSES	PLACEMENT	INDEMNITY
ADVANCING	PRINCIPAL	INTENTION	ADDRESSES	AGREEMENT	INTEGRITY
PRECEDING	DISMISSAL	ATTENTION	ESTIMATES	AMUSEMENT	NECESSITY
EXTENDING	CONTINUAL				

TEN LETTER WORDS

ATOMICBOMB	APPROACHED	COMPRESSED	UNDERSTOOD	CONFIDENCE
GEOGRAPHIC	ENTRENCHED	DISTRESSED	COASTGUARD	NEGLIGENCE
GYROSCOPIC	DESPATCHED	DESIGNATED	POSTOFFICE	EXPERIENCE
DIPLOMATIC	DISPATCHED	RESTRICTED	ACCORDANCE	PREFERENCE
BRIDGEHEAD	THREATENED	INSTRUCTED	ALLEGIANCE	DIFFERENCE
PRESCRIBED	MAINTAINED	PROHIBITED	APPEARANCE	CONFERENCE
REENFORCED	DETERMINED	REENLISTED	ACCEPTANCE	CAMOUFLAGE
REINFORCED	ONEHUNDRED	MECHANIZED	RESISTANCE	DEPENDABLE
BEENNEEDED	DECIPHERED	CONTRABAND	ASSISTANCE	EXPENDABLE
UNEXPENDED	ENCIPHERED	UNDERSTAND	PRECEDENCE	UNSUITABLE

~~RESTRICTED~~

TEN LETTER WORDS—*Continued*

ACCEPTABLE	EVACUATING	ALLOCATION	GONIOMETER	CONTINGENT
IMPASSIBLE	COLLECTING	FOUNDATION	HYDROMETER	SUFFICIENT
IMPOSSIBLE	CONNECTING	RECREATION	HYGROMETER	CONVENIENT
ASPOSSIBLE	INFLICTING	IRRIGATION	AMBASSADOR	EQUIVALENT
RECEPTACLE	EXPEDITING	NAVIGATION	INSTRUCTOR	ENGAGEMENT
MOTORCYCLE	RECRUITING	REGULATION	BALLISTICS	MANAGEMENT
AUTOMOBILE	ATTEMPTING	POPULATION	STATISTICS	EXCITEMENT
DISCIPLINE	SUPPORTING	ESTIMATION	CROSSROADS	DETACHMENT
QUARANTINE	EXTINGUISH	DOMINATION	DESPATCHES	ATTACHMENT
ENTERPRISE	NINETEENTH	DETONATION	DISPATCHES	EXPERIMENT
TRANSVERSE	EIGHTEENTH	OCCUPATION	ASSEMBLIES	ENROLLMENT
COORDINATE	THIRTEENTH	SEPARATION	FACILITIES	ASSIGNMENT
ILLUMINATE	FOURTEENTH	DECORATION	ACTIVITIES	ATTAINMENT
ANTICIPATE	WILLATTACK	LIMITATION	CASUALTIES	INTERNMENT
ILLITERATE	ARTIFICIAL	SANITATION	FRONTLINES	GOVERNMENT
ILLUSTRATE	CREDENTIAL	INVITATION	SUBMARINES	ASSESSMENT
COMPENSATE	ADDITIONAL	EVACUATION	OBJECTIVES	COMMITMENT
DISTRIBUTE	ACCIDENTAL	EVALUATION	ENEMYTANKS	DEPARTMENT
SUBSTITUTE	REGIMENTAL	EXCAVATION	SUSPICIONS	ENLISTMENT
CONSTITUTE	INDIVIDUAL	COLLECTION	COLLISIONS	INSTRUMENT
COMMUNIQUE	WITHDRAWAL	CONNECTION	PROVISIONS	DEPLOYMENT
TWENTYFIVE	AIRCONTROL	INSPECTION	EXPLOSIONS	EMPLOYMENT
SUCCESSIVE	SUCCESSFUL	CORRECTION	FORMATIONS	PERSISTENT
IMPRESSIVE	RESPECTFUL	PROTECTION	OPERATIONS	AIRSUPPORT
LOCOMOTIVE	MEMORANDUM	EXHIBITION	DIRECTIONS	CONSPIRACY
CENTRALIZE	SUSPENSION	EXPEDITION	CONDITIONS	DEFICIENCY
NATURALIZE	DISPERSION	DEFINITION	TROOPSHIPS	EFFICIENCY
DEMOBILIZE	CONCESSION	AMMUNITION	NEWSPAPERS	COMPLETELY
COMMANDING	CONFESSION	OPPOSITION	KILOMETERS	APPARENTLY
DEBOUCHING	DEPRESSION	PROPORTION	DESTROYERS	INCENDIARY
DETRUCKING	IMPRESSION	REVOLUTION	TRANSPORTS	COMMISSARY
ENTRUCKING	POSSESSION	MACHINEGUN	SUSPICIOUS	ELEMENTARY
ENCIRCLING	SUBMISSION	BATTLESHIP	VICTORIOUS	LABORATORY
SIGNALLING	COMMISSION	CENSORSHIP	CIRCUITOUS	TRAJECTORY
PATROLLING	PERMISSION	ARMOREDCAR	CONTINUOUS	CAPABILITY
OVERCOMING	DISCUSSION	DIVEBOMBER	PHOSPHORUS	AUDIBILITY
DETRAINING	CONCLUSION	COMMANDEER	FLASHLIGHT	VISIBILITY
CONCERNING	DEDICATION	DISPATCHER	COMMANDANT	SIMILARITY
INDICATING	INDICATION	MILLIMETER	LIEUTENANT	INSECURITY
ANTEDATING				

ELEVEN LETTER WORDS

IMPEDIMENTA	SURRENDERED	CONSTITUTED	INFLAMMABLE	CERTIFICATE
TOPOGRAPHIC	ENCOUNTERED	BATTLEFIELD	RESPONSIBLE	COMMUNICATE
RECOMMENDED	TRANSFERRED	PERFORMANCE	NAVALBATTLE	INVESTIGATE
PREARRANGED	DISINFECTED	MAINTENANCE	TEMPERATURE	APPROPRIATE
ESTABLISHED	REAPPOINTED	COINCIDENCE	MANUFACTURE	APPROXIMATE
OVERWHELMED	INTERCEPTED	SUBSISTENCE	SCHOOLHOUSE	EXTERMINATE
DISAPPEARED	INTERRUPTED	CATASTROPHE	CUSTOMHOUSE	DETERIORATE

~~RESTRICTED~~

ELEVEN LETTER WORDS—Continued

CONCENTRATE	SMOKESCREEN	DISTINCTION	PHILIPPINES	CONFINEMENT
DEMONSTRATE	APPLICATION	DESTRUCTION	PARENTHESES	REQUIREMENT
NECESSITATE	ASSOCIATION	INSTRUCTION	HEAVYLOSSES	MEASUREMENT
DISCONTINUE	RETALIATION	RECOGNITION	COMMUNIQUES	IMPROVEMENT
SEVENTYFIVE	DEBARKATION	REQUISITION	PARENTHESIS	CONCEALMENT
PROGRESSIVE	EMBARKATION	COMPOSITION	CREDENTIALS	ECHELONMENT
RETROACTIVE	LEGISLATION	DISPOSITION	BATTLESHIPS	DEVELOPMENT
DESCRIPTIVE	CIRCULATION	COMPETITION	ARMOREDCARS	APPOINTMENT
SYNCHRONIZE	INFORMATION	DESCRIPTION	CORRECTNESS	COMPARTMENT
APPROACHING	EXPLANATION	CONSUMPTION	ENGAGEMENTS	BELLIGERENT
INTERVENING	DESIGNATION	INSTITUTION	ASSIGNMENTS	INCOMPETENT
ENGINEERING	RESIGNATION	LIGHTBOMBER	ASSESSMENTS	FINGERPRINT
INTERFERING	EXAMINATION	HEAVYBOMBER	INSTRUMENTS	DISCREPANCY
ALTERNATING	PREPARATION	RANGEFINDER	INTERCERPTS	PHOTOGRAPHY
INTERESTING	COOPERATION	DYNAMOMETER	ESTIMATEDAT	IMMEDIATELY
WITHDRAWING	IMMIGRATION	THERMOMETER	SIGNIFICANT	EXTENSIVELY
DISTINGUISH	INSPIRATION	INTERPRETER	INDEPENDENT	EFFECTIVELY
SEVENTEENTH	CORPORATION	RECONNOITER	INTELLIGENT	PRELIMINARY
NAVALATTACK	PENETRATION	BLOCKBUSTER	COEFFICIENT	CONTROVERSY
STRATEGICAL	ARBITRATION	AERONAUTICS	BOMBARDMENT	ELECTRICITY
TRADITIONAL	COMPUTATION	NAVALFORCES	REPLACEMENT	NATIONALITY
CONTINENTAL	OBSERVATION	ACCESSORIES	EMPLACEMENT	SUITABILITY
FIRECONTROL	RESERVATION	HOSTILITIES	ENFORCEMENT	SUPERIORITY
NATIONALISM	RESTRICTION	ENEMYPLANES	ARRANGEMENT	

TWELVE LETTER WORDS

TRANSPACIFIC	CONSTITUTING	ILLUMINATION	CONSTITUTION	EMPLACEMENTS
HYDROGRAPHIC	BREAKTHROUGH	ANTICIPATION	NORTHWESTERN	MEASUREMENTS
UNIDENTIFIED	GEOGRAPHICAL	REGISTRATION	SOUTHWESTERN	ADVANTAGEOUS
COMMISSIONED	CONFIDENTIAL	ILLUSTRATION	MARKSMANSHIP	SIMULTANEOUS
DISSEMINATED	PRESIDENTIAL	INAUGURATION	MEDIUMBOMBER	ANTIAIRCRAFT
CONCENTRATED	RECREATIONAL	COMPENSATION	COMMISSIONER	NONCOMBATANT
DEMONSTRATED	AGRICULTURAL	CONVERSATION	PSYCHROMETER	CONVALESCENT
DISORGANIZED	DEPARTMENTAL	RADIOSTATION	SHARPSHOOTER	DISPLACEMENT
SIGNIFICANCE	UNSUCCESSFUL	CONTINUATION	DIFFICULTIES	COMMENCEMENT
INTELLIGENCE	GENERALALARM	PRESERVATION	UNITEDSTATES	ANNOUNCEMENT
INTERFERENCE	VETERINARIAN	MOBILIZATION	PREPARATIONS	ENTANGLEMENT
INCOMPETENCE	TRANSMISSION	ORGANIZATION	OBSTRUCTIONS	DECIPHERMENT
CONSIDERABLE	VERIFICATION	INTERDICTION	INSTRUCTIONS	ENCIPHERMENT
FIGHTERPLANE	CONFISCATION	ROADJUNCTION	LIGHTBOMBERS	REENLISTMENT
INTERMEDIATE	COMMENDATION	INTRODUCTION	HEAVYBOMBERS	INEFFICIENCY
DECENTRALIZE	CONCILIATION	CONSTRUCTION	HEADQUARTERS	SUCCESSFULLY
GENERALSTAFF	CANCELLATION	INTERVENTION	PREPAREDNESS	RESPECTFULLY
TRANSFERRING	PROCLAMATION	CONSCRIPTION	COMPLETENESS	SATISFACTORY
ENTERPRISING	CONFIRMATION	INTERRUPTION	CARELESSNESS	INTRODUCTORY
ILLUMINATING	CONFORMATION	DISTRIBUTION	SEARCHLIGHTS	IRREGULARITY
DISTRIBUTING	COORDINATION	SUBSTITUTION	REPLACEMENTS	

THIRTEEN LETTER WORDS

TRANSATLANTIC	CHRONOLOGICAL	DETERMINATION	FIGHTERPLANES	ESTABLISHMENT
DISTINGUISHED	CONGRESSIONAL	EXTERMINATION	INSTALLATIONS	ENTERTAINMENT
DECENTRALIZED	INTERNATIONAL	CONSIDERATION	MEDIUMBOMBERS	REAPPOINTMENT
DISAPPEARANCE	SPECIFICATION	CONCENTRATION	MISCELLANEOUS	WARDEPARTMENT
IMPRACTICABLE	QUALIFICATION	DEMONSTRATION	INSTANTANEOUS	APPROXIMATELY
INDETERMINATE	COMMUNICATION	QUARTERMASTER	REENFORCEMENT	EXTRAORDINARY
CORRESPONDING	ACCOMMODATION	CIRCUMSTANCES	REINFORCEMENT	REVOLUTIONARY
CONCENTRATING	INVESTIGATION	DISCREPANCIES	REIMBURSEMENT	DEPENDABILITY
COUNTERATTACK	DISSEMINATION	PRELIMINARIES	REINSTATEMENT	

FOURTEEN LETTER WORDS

CHARACTERISTIC	RECONNOITERING	ADMINISTRATION	REORGANIZATION
RECONNAISSANCE	METEOROLOGICAL	INTERPRETATION	RECONSTRUCTION
DISCONTINUANCE	CIRCUMSTANTIAL	TRANSPORTATION	IRREGULARITIES
CORRESPONDENCE	CLASSIFICATION	CENTRALIZATION	INVESTIGATIONS
ADMINISTRATIVE	IDENTIFICATION	NATURALIZATION	SATISFACTORILY
REPRESENTATIVE	RECOMMENDATION	DEMOBILIZATION	RESPONSIBILITY
DISTINGUISHING			

C. LIST OF WORDS USED IN MILITARY TEXT ARRANGED ALPHABETICALLY ACCORDING TO WORD PATTERN

PATTERN AA

A CC EPT
A CC ORDING
O CC UPY
A DD
SU DD EN
LA DD ER
BE DD ING
FL EE
S EE
THR EE
PROC EE D
SP EE D
CR EE K
W EE K
F EE L
ST EE L
WH EE L
B EE N
FOURT EE N
HASB EE N
QU EE N
SCR EE N
S EE N
SIXT EE N
R EE NLIST
K EE P
SW EE PING
F EE T
FL EE T
M EE T
JUMPO FF
O FF
STA FF
O FF END
SU FF ER
TRA FF IC
O FF ICE
O FF ICER
E FF ORT
FO GG Y
A LL
CA LL
CE LL
DRI LL
ENRO LL
FA LL
FE LL
FU LL
HI LL
I LL
INSTA LL
PAYRO LL
REFI LL
SHE LL
SMA LL
SPE LL
WE LL
WI LL
VI LL AGE
CO LL APSED
DO LL AR
OSCI LL ATE
KI LL ED
BI LL ET
BU LL ETIN
VA LL EY
A LL IED
A LL IES
FA LL ING
PATRO LL ING
SHE LL ING
A LL OW
A LL Y
RA LL Y
CO MM A
CO MM AND
CO MM ANDER
SU MM ARY
CO MM END
CO MM ENTY
HA MM ER
SU MM ER
CO MM IT
SU MM IT
SU MM ON
CO MM UTE
TO NN AGE
CHA NN EL
BA NN ER
GU NN ER
MA NN ER
A NN EX
CA NN OT
T OO
W OO DS
PR OO F
B OO K
C OO K
H OO K
L OO K
T OO K
SCH OO L
T OO L
PLAT OO N
S OO N
TR OO PS
C OO RDINATE
B OO TH
STO PP ED
HA PP EN
CLI PP ER
MA PP ING
A PP LY
SU PP LY
A PP OINT
A PP OINTED
SU PP ORT
SU PP ORTING
A PP ROVE
TE RR AIN
CU RR ENT
A RR EST
HU RR ICANE
DE RR ICK
GA RR ISON
A RR IVE
CA RR Y
FE RR Y
ACRO SS
COMPA SS
CONGRE SS
CRO SS
DARKNE SS
DRE SS
LE SS

PATTERN AA—Continued

LO SS	A SS IGNED	BA TT EN
MA SS	CRO SS ING	WRI TT EN
ME SS	DRE SS ING	BI TT ER
PA SS	ME SS ING	LI TT ER
PRE SS	PA SS IVE	BA TT ERY
UNLE SS	LE SS ON	SPO TT ING
WITNE SS	I SS UE	BA TT LE
PA SS ED	A SS URE	BA TT LESHIP
A SS EMBLY	EMBA SS Y	MU ZZ LE
A SS ET	OMI TT ED	NO ZZ LE
PO SS IBLE	SUBMI TT ED	

MISCELLANEOUS PATTERNS

AABA	AGR EEME NT	AABCB	SU FFICI ENT
AABA	K EEPE R	AABCB	A LLEGE
AABA	CH EESE	AABCB	CO LLEGE
AABA	BR EEZE	AABCB	BI LLETE D
AABA	MA NNIN G	AABCB	A MMETE R
AABA	PLA NNIN G	AABCB	W OODED
AABA	RU NNIN G	AABCB	TE RRIFI C
AABA	L OOKO UT	AABCB	BA TTERE D
AABA	E RROR	AABCBDEB	DI FFERENCE
AABA	MI RROR	AABCC	A CCESS
AABA	TE RROR	AABCC	A CCESS ORY
AABA	GLA SSES	AABCC	CO MMISS ARY
AABA	LO SSES	AABCCB	WI LLATTA CK
AABA	PA SSES	AABCCDD	CO MMITTEE
AABA	CHA SSIS	AABCCDEFBC	A CCESSORIES
AABA	A SSIS T	AABCDA	I LLEGAL
AABAACB	A SSESSME NT	AABCDA	A TTEMPT
AABAACBDEA	A SSESSMENTS	AABCDAB	A TTEMPTE D
AABAB	PROC EEDED	AABCDB	O FFENSE
AABB	CO FFEE	AABCDB	CHA LLENGE
AABB	BA LLOO N	AABCDB	BA LLISTI C
AABBAACAC	B EENNEEDED	AABCDB	A RRESTE D
AABBCBC	SU CCEEDED	AABCDB	PA SSENGE R
AABCA	B EETLE	AABCDB	BA TTERIE S
AABCA	A NNOUN CE	AABCDBA	SU RRENDER
AABCA	F OOTHO LD	AABCDBABD	SU RRENDERED
AABCA	CA RRIER	AABCDBC	CO MMANDAN T
AABCA	A SSETS	AABCDBD	O FFENDED
AABCA	I SSUES	AABCDBEC	BA LLISTICS
AABCADEC	CO MMITMENT	AABCDC	E FFICAC Y
AABCADEC	A TTENTION	AABCDD	A DDRESS
AABCADEFEA	A NNOUNCEMEN T	AABCDD	I LLNESS
AABCB	SCR EENIN G	AABCDDCA	A DDRESSED
AABCB	SU FFERE D	AABCDDCD	A DDRESSES
AABCB	DI FFERE NT	AABCDEB	CO MMUNIQU E
AABCB	O FFICI AL	AABCDEB	TR OOPSHIP

MISCELLANEOUS PATTERNS—Continued

AABCDEB	A SSEMBLE	ABA	INVA DED
AABCDEBC	TR OOPSHIPS	ABA	LAN DED
AABCDEC	CO MMANDIN G	ABA	RAI DED
AABCDECB	BA TTLEFIEL D	ABA	WOUN DED
AABCDED	CO MMANDED	ABA	DID
AABCDEDFC	A MMUNITION	ABA	IC EBE RG
AABCDEE	CO MMANDEE R	ABA	PR ECE DING
AABCDEFA	R EENLISTE D	ABA	R ECE IPT
AABCDEFA	I RREGULAR	ABA	CR EDE NTIAL
AABCDEFB	O FFENSIVE	ABA	F EDE RAL
AABCDEFBA	A SSEMBLIES	ABA	D EFE AT
AABCDEFC	A LLOTMENT	ABA	D EFE CT
AABCDEFC	C OOPERATE	ABA	D EFE R
AABCDEFD	I LLUSTRAT E	ABA	SI EGE
AABCDEFD	A SSIGNMEN T	ABA	R EJE CT
AABCDEFDGA	A SSIGNMENTS	ABA	S ELE CT
AABCDEFGA	C OOPERATIO N	ABA	T ELE GRAM
AABCDEFGABF	R EENLISTMENT	ABA	ELE VATION
AABCDEFGD	BA TTLESHIPS	ABA	SCH EME
AABCDEFGDAE	C OORDINATION	ABA	R EME DY
AABCDEFGDE	A PPOINTMENT	ABA	DISPLAC EME NT
ABA	AGA IN	ABA	PLAC EME NT
ABA	AGA INST	ABA	ENE MY
ABA	C ALA MITY	ABA	G ENE RAL
ABA	ALA RM	ABA	R EPE L
ABA	S ALA RY	ABA	H ERE
ABA	D AMA GE	ABA	SPH ERE
ABA	M ANA GE	ABA	TH ERE
ABA	C ANA L	ABA	W ERE
ABA	ANA LYZE	ABA	WH ERE
ABA	J APA N	ABA	CONQU ERE D
ABA	P ARA CHUTE	ABA	COV ERE D
ABA	P ARA DE	ABA	TH ESE
ABA	SEP ARA TION	ABA	PR ESE NT
ABA	F ATA L	ABA	D ESE RT
ABA	N AVA L	ABA	COMPL ETE
ABA	N AVA LFORCES	ABA	KILOM ETE R
ABA	C AVA LRY	ABA	M ETE R
ABA	EXC AVA TION	ABA	P ETE R
ABA	AWA IT	ABA	D EVE LOP
ABA	AWA RD	ABA	S EVE N
ABA	AWA Y	ABA	S EVE NTH
ABA	PRO BAB LE	ABA	S EVE NTY
ABA	PRO BAB LY	ABA	S EVE RAL
ABA	BI CYC LE	ABA	EVE RY
ABA	CYC LONE	ABA	EYE
ABA	BLOCKA DED	ABA	FIF TH
ABA	GROUN DED	ABA	FIF TY
ABA	GUAR DED	ABA	EIG HTH

MISCELLANEOUS PATTERNS—Continued

ABA	L IAI SON	ABA	CA RTR IDGE
ABA	PROH IBI T	ABA	D RYR UN
ABA	SERV ICI NG	ABA	DI SAS TER
ABA	RA IDI NG	ABA	CA SES
ABA	R IDI NG	ABA	RE SIS T
ABA	R IGI D	ABA	SUS PEND
ABA	F ILI NG	ABA	SYS TEM
ABA	M ILI TARY	ABA	S TAT ION
ABA	MOB ILI ZE	ABA	DIC TAT OR
ABA	S IMI LAR	ABA	TIT LE
ABA	L IMI T	ABA	AL TIT UDE
ABA	PROX IMI TY	ABA	LA TIT UDE
ABA	F INI SH	ABA	TOT AL
ABA	F IRI NG	ABA	TOT ALING
ABA	RET IRI NG	ABA	A UGU ST
ABA	W IRI NG	ABA	USU AL
ABA	V ISI BLE	ABA	F UTU RE
ABA	D ISI NFECT	ABA	SUR VIV ED
ABA	ADV ISI NG	ABAA	HAV EBEE N
ABA	DEC ISI ON	ABAA	SESS ION
ABA	V ISI T	ABAACC	TATTOO
ABA	V ISI TOR	ABAB	DETRA ININ G
ABA	POL ITI CS	ABAB	L ININ G
ABA	CR ITI QUE	ABAB	M ININ G
ABA	POS ITI VE	ABAB	OBTA ININ G
ABA	MEM ORIAL	ABAB	RA ININ G
ABA	NAN	ABAB	REMA ININ G
ABA	DOMI NAN CE	ABAB	TRA ININ G
ABA	ORD NAN CE	ABAB	CR ISIS
ABA	DOMI NAN T	ABAB	WI THTH E
ABA	NIN E	ABAB	PAR TITI ON
ABA	NIN ETY	ABACA	C ANADA
ABA	MOR NIN G	ABACA	P ANAMA
ABA	NIN TH	ABACA	PR ECEDE
ABA	OBO E	ABACA	ELEME NT
ABA	C OLO N	ABACA	ELEME NTARY
ABA	SEMIC OLO N	ABACA	ELEVE N
ABA	C OLO RS	ABACA	C EMETE RY
ABA	AUT OMO BILE	ABACA	S EVERE
ABA	PR OMO TE	ABACA	AUD IBILI TY
ABA	H ONO R	ABACA	EXH IBITI ON
ABA	VIG ORO US	ABACA	V ICINI TY
ABA	M OTO R	ABACA	M ILITI A
ABA	M OTO RIZED	ABACA	FAC ILITI ES
ABA	PR OVO ST	ABACA	D IMINI SH
ABA	PIP E	ABACA	L IMITI NG
ABA	POP ULATED	ABACA	INITI AL
ABA	LIB RAR Y	ABACA	DEF INITI ON
ABA	AI RDR OME	ABACA	D IRIGI BLE

MISCELLANEOUS PATTERNS—Continued

ABACA	SEM IRIGI D	ABACDA	R EVENUE
ABACA	REQU ISITI ON	ABACDA	U NKNOWN
ABACA	C IVILI AN	ABACDA	PR OMOTIO N
ABACA	D IVISI ON	ABACDAAC	S EVENTEEN
ABACA	L OCOMO TIVE	ABACDAACD	S EVENTEENT H
ABACA	M ONOPO LY	ABACDAC	D ESERTER
ABACA	PR OTOCO L	ABACDAD	D EFENSES
ABACA	CONS TITUT E	ABACDAED	AVAILABL E
ABACA	UNUSU AL	ABACDAEEC	N AVALBATTL E
ABACADA	V ISIBILI TY	ABACDB	F ATALIT Y
ABACADB	DEF INITION	ABACDB	A NONYMO US
ABACADBA	PR ECEDENCE	ABACDB	C OLONEL
ABACADC	INITIAT E	ABACDBA	TH EREFORE
ABACADD	COMPL ETENESS	ABACDC	R ECEIVI NG
ABACADDA	N AVALATTA CK	ABACDC	DYNA MOMETE R
ABACADEC	D IVISIONS	ABACDCA	L IMITATI ON
ABACB	V ACANC Y	ABACDCCA	NINETEEN
ABACB	COMB ATANT	ABACDCCAD	NINETEENT H
ABACB	C ATAST ROPHE	ABACDCEA	S TATEMENT
ABACB	D ETECT OR	ABACDCECFGHIE	M ETEOROLOGICAL
ABACB	V ISITS	ABACDD	FIFTEE N
ABACB	MEMBE R	ABACDD	FO RTRESS
ABACBDEC	D ETENTION	ABACDDEC	FIFTEENT H
ABACBDEC	R ETENTION	ABACDEA	ELEVATE
ABACBDEFGFAG	NONCOMBATANT	ABACDEA	D EVELOPE
ABACC	R EBELL ION	ABACDEA	VER IFICATI ON
ABACC	N ECESS ARY	ABACDEA	S IMILARI TY
ABACC	N ECESS ITY	ABACDEAD	SUSPENSE
ABACC	CAR ELESS	ABACDEAFGE	SUSPENSION
ABACC	WIR ELESS	ABACDEB	EXPL ANATION
ABACCA	P ARALLA X	ABACDEB	T OPOGRAP HIC
ABACCA	R EPELLE D	ABACDEBFA	R ECEPTACLE
ABACCA	T OMORRO W	ABACDEC	ABANDON
ABACCDACC	CAR ELESSNESS	ABACDEC	D AMAGING
ABACCDC	P ARALLEL	ABACDEC	QU ARANTIN E
ABACCDEFEA	N ECESSITATE	ABACDECA	P ENETRATE
ABACDA	ALASKA	ABACDECFBA	D ETERIORATE
ABACDA	ARABIA	ABACDECFGB	P ENETRATION
ABACDA	N AVALBA SE	ABACDED	C APABILI TY
ABACDA	R ECEIVE	ABACDED	M OTORCYC LE
ABACDA	D ECEMBE R	ABACDED	SUSPICI ON
ABACDA	D EFENSE	ABACDEDEDC	G ENERALALAR M
ABACDA	R EJECTE D	ABACDEDFBA	SUSPICIOUS
ABACDA	R ELEASE	ABACDEDFGA	SUSPICIONS
ABACDA	S ELECTE D	ABACDEFA	D EFECTIVE
ABACDA	R EMEDIE S	ABACDEFA	D EFENSIVE
ABACDA	EMERGE NCY	ABACDEFA	T ELEPHONE
ABACDA	ENEMIE S	ABACDEFA	D ETERMINE
ABACDA	R EPEATE D	ABACDEFA	D EVELOPME NT

MISCELLANEOUS PATTERNS—Continued

ABACDEFA	EXERCISE	ABBA	SH IPPI NG
ABACDEFAF	EXERCISES	ABBA	M ISSI NG
ABACDEFB	DEDICATE	ABBA	ADM ISSI ON
ABACDEFB	ENEMYTAN KS	ABBA	M ISSI ON
ABACDEFC	DEDICATI ON	ABBA	PERM ISSI ON
ABACDEFCDFE	V ETERINARIAN	ABBA	F ITTI NG
ABACDEFCFD	ELECTRICIT Y	ABBA	AFTER NOON
ABACDEFD	SUSPECTE D	ABBA	NOON
ABACDEFDF	SUSPENDED	ABBA	F OLLO W
ABACDEFE	ANALYSIS	ABBA	C OMMO N
ABACDEFGA	EXECUTIVE	ABBA	OPPO SE
ABACDEFGB	POPULATIO N	ABBA	OPPO SITE
ABACDEFGBA	ENEMYPLANE S	ABBA	B OTTO M
ABACDEFGBA	S EVENTYFIVE	ABBAB	B AGGAG E
ABACDEFGBEHF	D ETERMINATION	ABBAB	WITN ESSES
ABACDEFGDHH	G ENERALSTAFF	ABBACA	APPARA TUS
ABACDEFGE	MEMORANDA	ABBACA	L ETTERE D
ABACDEFGHA	MEMORANDUM	ABBACB	V ESSELS
ABACDEFGHIA	D ECENTRALIZE	ABBACDA	M ESSENGE R
ABBA	AFFA IR	ABBACDA	EFFECTE D
ABBA	APPA RENT	ABBACDB	M ISSIONS
ABBA	APPA RENTLY	ABBACDEA	IRRIGATI ON
ABBA	B ARRA CKS	ABBACDEDA	OPPOSITIO N
ABBA	B ARRA GE	ABBACDEFA	EFFECTIVE
ABBA	ARRA NGE	ABBACDEFA	D IFFICULTI ES
ABBA	P ASSA GE	ABBACDEFA	IMMIGRATI ON
ABBA	ASSA ULT	ABBACDEFCD	ILLITERATE
ABBA	ATTA CH	ABBACDEFDB	ATTAINMENT
ABBA	ATTA CK	ABBACDEFEC	ARRANGEMEN T
ABBA	ATTA IN	ABBACDEFGB	ATTACHMENT
ABBA	B ATTA LION	ABBCA	ANNUA L
ABBA	IN DEED	ABBCA	APPEA R
ABBA	EFFE CT	ABBCA	DIS APPEA R
ABBA	COMP ELLE D	ABBCA	C ARRIA GE
ABBA	SH ELLE D	ABBCA	S ETTLE
ABBA	CONF ERRE D	ABBCA	ISSUI NG
ABBA	COMPR ESSE D	ABBCA	FOUR TEENT H
ABBA	IMPR ESSE D	ABBCA	SIX TEENT H
ABBA	PR ESSE D	ABBCA	CHA UFFEU R
ABBA	V ESSE L	ABBCA	S URROU ND
ABBA	CIGAR ETTE	ABBCADAEFC	APPEARANCE
ABBA	B ETTE R	ABBCADAEFC	DIS APPEARANCE
ABBA	L ETTE R	ABBCADC	APPEARE D
ABBA	D IFFI CULT	ABBCBBDA	P OSSESSIO N
ABBA	W ILLI AM	ABBCBDA	ASSISTA NCE
ABBA	F ILLI NG	ABBCBDAED	ASSISTANT
ABBA	K ILLI NG	ABBCCDAB	ASSOONAS
ABBA	REF ILLI NG	ABBCDA	ALLOWA NCE
ABBA	SW IMMI NG	ABBCDA	APPROA CH

MISCELLANEOUS PATTERNS—Continued

ABBCDA	ARRIVA L	ABCA	ADVA NCE
ABBCDA	ASSURA NCE	ABCA	DI AGRA M
ABBCDA	M ESSAGE	ABCA	EV ALUA TION
ABBCDA	ILLUMI NATE	ABCA	ALWA YS
ABBCDAB	M ESSAGES	ABCA	C AMPA IGN
ABBCDAB	C ORRIDOR	ABCA	M ANDA TE
ABBCDAEA	B ELLIGERE NT	ABCA	M ANUA L
ABBCDAEFC	ALLOCATIO N	ABCA	J ANUA RY
ABBCDAEFC	IMMEDIATE	ABCA	C ANVA S
ABBCDAEFGAE	ILLUMINATIN G	ABCA	CH APLA IN
ABBCDAEFGAHE	ILLUMINATION	ABCA	C APTA IN
ABBCDAEFGAHE	D ISSEMINATION	ABCA	AREA
ABBCDBCEA	APPROPRIA TE	ABCA	DEB ARKA TION
ABBCDCA	EFFICIE NT	ABCA	EMB ARKA TION
ABBCDCA	C OLLISIO N	ABCA	ASIA
ABBCDCAED	EFFICIENC Y	ABCA	CO ASTA L
ABBCDCAED	C OLLISIONS	ABCA	C ASUA L
ABBCDCEFA	ADDITIONA L	ABCA	C ASUA LTY
ABBCDDCA	C OMMISSIO N	ABCA	AVIA TOR
ABBCDDCA	C OMMISSIO NER	ABCA	BARB ED
ABBCDDCEAFGC	ACCOMMODATIO N	ABCA	BOMB
ABBCDEA	ACCOMPA NY	ABCA	BOMB ARD
ABBCDEA	APPROVA L	ABCA	BOMB ER
ABBCDEA	ASSOCIA TE	ABCA	LIGHT BOMB ER
ABBCDEA	SH ELLFIRE	ABCA	BRIB E
ABBCDEA	T ERRIBLE	ABCA	BULB
ABBCDEAFB	ACCORDANC E	ABCA	CANC EL
ABBCDEAFB	REENFORCE	ABCA	CHEC K
ABBCDEAFBC	ACCEPTANCE	ABCA	CIRC LE
ABBCDEAFBGBC	REENFORCEMEN T	ABCA	CIRC ULATE
ABBCDEAFD	APPLICATI ON	ABCA	CONC EAL
ABBCDEAFEC	ASSOCIATIO N	ABCA	CONC LUDE
ABBCDEAFGC	ACCEPTABLE	ABCA	HUN DRED
ABBCDEAFGC	ALLEGIANCE	ABCA	L EADE R
ABBCDEAFGHF	C ORRESPONDIN G	ABCA	EAGE R
ABBCDEFGA	ACCIDENTA L	ABCA	M EAGE R
ABBCDEFGA	APPROXIMA TE	ABCA	S EAME N
ABBCDEFGA	OCCUPATIO N	ABCA	ST EAME R
ABBCDEFGBAHAC	IRREGULARITIE S	ABCA	N EARE ST
ABBCDEFGBA	IRREGULARI TY	ABCA	C EASE
ABBCDEFGEA	ILLUSTRATI ON	ABCA	GR EASE
ABBCDEFGHAD	C OMMENDATION	ABCA	INCR EASE D
ABCA	P ACKA GE	ABCA	L EAVE
ABCA	EV ACUA TING	ABCA	ECHE LON
ABCA	EV ACUA TION	ABCA	WR ECKE D
ABCA	R ADIA L	ABCA	INF ECTE D
ABCA	R ADIA TE	ABCA	EDGE
ABCA	ADJA CENT	ABCA	S EIZE
ABCA	GR ADUA L	ABCA	R ELIE F

MISCELLANEOUS PATTERNS—Continued

ABCA	H ELPE R	ABCA	I NFAN TRY
ABCA	TW ELVE	ABCA	CO NFIN E
ABCA	NOV EMBE R	ABCA	U NION
ABCA	ABS ENCE	ABCA	SU NKEN
ABCA	LIC ENSE	ABCA	FLA NKIN G
ABCA	C ENTE R	ABCA	I NLAN D
ABCA	ENTE R	ABCA	I NTEN D
ABCA	ENVE LOP	ABCA	CO NTIN UAL
ABCA	R EQUE ST	ABCA	CO NTIN UE
ABCA	FI ERCE	ABCA	I NVEN T
ABCA	S ERGE ANT	ABCA	OCTO BER
ABCA	MAT ERIE L	ABCA	D OCTO R
ABCA	REV ERSE	ABCA	F OGHO RN
ABCA	OBS ERVE	ABCA	P OISO N
ABCA	R ESPE CT	ABCA	C OMPO SED
ABCA	W ESTE RLY	ABCA	C ONVO Y
ABCA	W ESTE RN	ABCA	EN ORMO US
ABCA	ETHE R	ABCA	EXPL OSIO N
ABCA	MAN EUVE R	ABCA	PUMP
ABCA	R EVIE W	ABCA	PURP OSE
ABCA	EXCE PT	ABCA	HA RBOR
ABCA	EXPE CT	ABCA	AI RBOR NE
ABCA	EXPE ND	ABCA	MU RDER
ABCA	EXTE ND	ABCA	O RDER
ABCA	GAUG E	ABCA	O RDER S
ABCA	GEOG RAPHIC	ABCA	REAR
ABCA	FOR GING	ABCA	RECR UIT
ABCA	W HICH	ABCA	COU RIER
ABCA	HIGH	ABCA	P RIOR
ABCA	HIGH ER	ABCA	SUPE RIOR
ABCA	HIGH EST	ABCA	A RMOR
ABCA	V ICTI M	ABCA	A RMOR Y
ABCA	M IDNI GHT	ABCA	P ROGR AM
ABCA	DR IFTI NG	ABCA	MO RTAR
ABCA	L IFTI NG	ABCA	QUA RTER
ABCA	S IGNI FY	ABCA	QUA RTER S
ABCA	BU ILDI NG	ABCA	FEB RUAR Y
ABCA	INDI CATE	ABCA	FO RWAR D
ABCA	INDI RECT	ABCA	CEN SORS HIP
ABCA	DESCR IPTI ON	ABCA	SUNS ET
ABCA	L IQUI D	ABCA	IMPOR TANT
ABCA	A IRFI ELD	ABCA	S TART
ABCA	REPR ISAL	ABCA	PRO TECT
ABCA	M ISFI RE	ABCA	TENT
ABCA	F ISHI NG	ABCA	TENT H
ABCA	W ITHI N	ABCA	PRO TEST
ABCA	FUE LOIL	ABCA	TEXT
ABCA	MAIM	ABCA	THAT
ABCA	LA NDIN G	ABCA	S TRAT EGIC

MISCELLANEOUS PATTERNS—Continued

ABCA	S TRAT EGY
ABCA	D UGOU T
ABCA	UNSU ITABLE
ABCA	P URSU E
ABCA	P URSU IT
ABCA	O UTGU ARD
ABCAA	D ECREE
ABCAA	D EGREE
ABCAA	B ETWEE N
ABCAA	DI SCUSS
ABCAA	A SPOSS IBLE
ABCAAB	P ONTOON
ABCAAB	THATTH E
ABCAACDEB	P REARRANGE D
ABCAB	W ARFAR E
ABCAB	S ECREC Y
ABCAB	OBS ERVER
ABCAB	W HETHE R
ABCAB	B INDIN G
ABCAB	F INDIN G
ABCAB	S INKIN G
ABCAB	PA INTIN G
ABCAB	PR INTIN G
ABCAB	I NTENT
ABCAB	P ONTON
ABCAB	C ORPOR AL
ABCAB	RECRE ATION
ABCAB	P RIORI TY
ABCAB	SUPE RIORI TY
ABCAB	DI SEASE
ABCAB	PRO TECTE D
ABCAB	PRO TESTE D
ABCAB	O UTPUT
ABCABA	INT ERFERE
ABCABB	D ISMISS
ABCABB	D ISMISS AL
ABCABC	THATHA VE
ABCABCA	ENTENTE
ABCABDA	S ENTENCE
ABCABDB	REPRESE NT
ABCABDBEFGFHIB	REPRESENTATIVE
ABCABDBEFGFHIED	REPRESENTATIONS
ABCABDC	RETREAT
ABCABDED	M ANGANESE
ABCABDEFA	C ORPORATIO N
ABCABDEFGHD	RECREATIONA L
ABCAC	ARMAM ENT
ABCAC	N EARER
ABCAC	PROPO SE

ABCAC	P RAIRI E
ABCAC	PRO TESTS
ABCACA	D IETITI AN
ABCACB	O RDERED
ABCACBDEC	PROPORTIO N
ABCACDEFD	PROPOSALS
ABCADA	ALMANA C
ABCADA	R ELIEVE
ABCADA	C ENTERE D
ABCADA	B ESIEGE D
ABCADA	R EVIEWE D
ABCADAB	CO NTINENT AL
ABCADAC	S EALEVEL
ABCADAC	INDIVID UAL
ABCADAEC	IGNITION
ABCADAEFB	TENTATIVE
ABCADAEFC	S IGNIFICAN T
ABCADAEFCE	S IGNIFICANC E
ABCADAEFGHF	SUBSISTENCE
ABCADB	ATLANT IC
ABCADB	BRIBER Y
ABCADB	CIRCUI T
ABCADB	W EDNESD AY
ABCADB	LOG ISTICS
ABCADB	EXPL OSIONS
ABCADB	PREPAR ING
ABCADB	IM PROPER
ABCADB	PROPER
ABCADBA	INSIGNI A
ABCADBC	PREPARE
ABCADBCEFCGG	PREPAREDNESS
ABCADBD	PREPARA TION
ABCADBEFD	CIRCUITOU S
ABCADC	R ADIATI ON
ABCADC	ST ANDARD
ABCADC	V ARIATI ON
ABCADC	ASIATI C
ABCADC	AVIATI ON
ABCADC	R EVIEWI NG
ABCADC	EXTENT
ABCADC	I NVENTE D
ABCADC	TACTIC S
ABCADC	S TARTER
ABCADC	ZIGZAG
ABCADCA	CO NVENIEN T
ABCADCB	CO NDENSED
ABCADCB	TACTICA L
ABCADCEFBGABC	ENTERTAINMENT
ABCADCEFGED	CONCENTRATE

~~RESTRICTED~~

MISCELLANEOUS PATTERNS—Continued

ABCADCEFGEHC	CONCENTRATIN G
ABCADCEFGEHBC	CONCENTRATION
ABCADD	D EPRESS ION
ABCADD	EXCESS
ABCADD	D ISTILL
ABCADD	P OSTOFF ICE
ABCADD	B OYCOTT
ABCADDA	AMBASSA DOR
ABCADDA	EXPELLE D
ABCADDECCFA	UNSUCCESSFU L
ABCADDEFA	EXCESSIVE
ABCADEA	ADVANTA GE
ABCADEA	ADVANTA GEOUS
ABCADEA	D ECREASE
ABCADEA	S EPTEMBE R
ABCADEA	R EQUESTE D
ABCADEA	D ISCIPLI NE
ABCADEAB	CO NTINGENT
ABCADEAE	EXPENDED
ABCADEAE	EXPENSES
ABCADEAE	EXTENDED
ABCADEAFA	ELSEWHERE
ABCADEAFGA	EXPERIENCE
ABCADEB	C ENTERIN G
ABCADEB	ENTERIN G
ABCADEB	R ESPECTS
ABCADEB	INCIDEN T
ABCADEB	M ISFIRES
ABCADEBCE	INCIDENCE
ABCADEC	M ANDATED
ABCADEC	S ECRETAR Y
ABCADEC	GYR OSCOPIC
ABCADECA	REARGUAR D
ABCADECAFD	D ISTINCTION
ABCADECFC	CONCERNIN G
ABCADEDA	CO NFINEMEN T
ABCADEDAFB	INVITATION
ABCADEDBD	SUBSTITUT E
ABCADEDBDE	SUBSTITUTI ON
ABCADEDC	LI EUTENANT
ABCADEDFGA	ENTERPRISE
ABCADEDFGDBC	CONCILIATION
ABCADEDFGFB	ENTERPRISIN G
ABCADEE	P ROGRESS
ABCADEEBFGHC	CANCELLATION
ABCADEED	CANCELLE D
ABCADEEFBC	CONCESSION
ABCADEEFGD	P ROGRESSIVE
ABCADEFA	ECHELONE D
ABCADEFA	ENVELOPE
ABCADEFA	EXPEDITE
ABCADEFA	EXPERIME NT
ABCADEFAB	INDICATIN G
ABCADEFAB	D ISTINGUIS H
ABCADEFABGADE	D ISTINGUISHING
ABCADEFAGB	INDICATION
ABCADEFB	ADVANCED
ABCADEFBA	EXT RAORDINAR Y
ABCADEFC	BOMBARDM ENT
ABCADEFC	CIRCULAR
ABCADEFC	U NTENABLE
ABCADEFCGHB	RETROACTIVE
ABCADEFD	ADVANCIN G
ABCADEFD	EXTENDIN G
ABCADEFD	EXTERIOR
ABCADEFE	CONCRETE
ABCADEFE	EXPEDITI NG
ABCADEFE	EXPEDITI ON
ABCADEFE	OBSOLETE
ABCADEFE	G ONIOMETE R
ABCADEFE	PURPOSES
ABCADEFE	RECRUITI NG
ABCADEFEA	C OMPOSITIO N
ABCADEFGA	EXPENSIVE
ABCADEFGA	EXTENSIVE
ABCADEFGAF	ECHELONMEN T
ABCADEFGB	C ASUALTIES
ABCADEFGB	CIRCULATI ON
ABCADEFGBC	CONCLUSION
ABCADEFGC	INDICATED
ABCADEFGC	S TRATEGICA L
ABCADEFGD	EXTENSION
ABCADEFGDC	CONCEALMEN T
ABCADEFGE	REPRISALS
ABCADEFGF	BOMBARDED
ABCADEFGHAB	C ONFORMATION
ABCADEFGHCA	EXTERMINATE
ABCADEFGHCFIG	EXTERMINATION
ABCADEFGHEIGCF	REORGANIZATION
ABCADEFGHH	R ESPECTFULL Y
ABCADEFGHIAJF	CIRCUMSTANCES
ABCADEFGHIB	RETROACTIVE
ABCADEFGHIE	GEOGRAPHICA L
ABCADEFGHIGBH	CIRCUMSTANTIA L
ABCBA	COMP LETEL Y
ABCBA	AWKWA RD
ABCBA	CAPAC ITY
ABCBA	PA CIFIC

~~RESTRICTED~~

MISCELLANEOUS PATTERNS—Continued

ABCBA	SPE CIFIC	ABCBDEBA	RECEIVER
ABCBA	HIN DERED	ABCBDEBA	REPEATER
ABCBA	DIVID E	ABCBDEFA	REJECTOR
ABCBA	GARAG E	ABCBDEFA	STATIONS
ABCBA	C ITATI ON	ABCBDEFBA	DEVELOPED
ABCBA	LEVEL	ABCBDEFGA	R ESISTANCE
ABCBA	P REFER	ABCBDEFGBA	DETERMINED
ABCBA	REFER	ABCBDEFGHFA	DISINFECTED
ABCBA	P RESER VATION	ABCBDEFGHIJBA	DECENTRALIZED
ABCBA	RESER VATION	ABCCA	LITTL E
ABCBA	TAXAT ION	ABCCA	PASSP ORT
ABCBA	HOS TILIT Y	ABCCA	S TREET
ABCBA	U TILIT Y	ABCCABDEC	C ROSSROADS
ABCBA	AC TIVIT Y	ABCCBADED	MILLIMETE R
ABCBAA	U SELESS	ABCCBCA	BE GINNING
ABCBAAB	P REFERRE D	ABCCBDA	INF LAMMABL E
ABCBAB	DIVIDI NG	ABCCDA	COLLEC T
ABCBAB	AC TIVITI ES	ABCCDA	CORREC T
ABCBABDEB	P REFERENCE	ABCCDA	T RIGGER
ABCBABDEB	REFERENCE	ABCCDA	RUBBER
ABCBADA	MINIMUM	ABCCDA	RUNNER
ABCBADB	P RESERVE	ABCCDA	SPOOLS
ABCBADB	RESERVE	ABCCDA	SPOONS
ABCBADB	REVERSE	ABCCDA	SUGGES T
ABCBADBC	RESERVES	ABCCDA	SUPPOS E
ABCBADEB	SPE CIFICATI ON	ABCCDA	TURRET
ABCBCDBA	REMEMBER	ABCCDAA	SUCCESS
ABCBDA	DEFEND	ABCCDAAEB	SUCCESSFU L
ABCBDA	DEPEND	ABCCDAAEBFF	SUCCESSFULL Y
ABCBDA	MU NITION S	ABCCDAAEFD	SUCCESSIVE
ABCBDA	RESEAR CH	ABCCDAB	P RESSURE
ABCBDA	STATES	ABCCDAEC	TERRITOR Y
ABCBDA	STATUS	ABCCDAED	CORRECTE D
ABCBDA	IN TEREST	ABCCDAEFB	COLLECTIO N
ABCBDAB	DEFENDE R	ABCCDAEFB	CORRECTIO N
ABCBDAB	E NGAGING	ABCCDAEFBC	CONNECTION
ABCBDABA	DEFENDED	ABCCDAEFC	CONNECTIN G
ABCBDABD	DEPENDEN T	ABCCDAEFDGG	CORRECTNESS
ABCBDABDEA	STATISTICS	ABCCDEA	GASSING
ABCBDAEFGB	DEPENDABLE	ABCCDEA	GETTING
ABCBDAEFGHG	DEPENDABILI TY	ABCCDEA	ST RAGGLER
ABCBDCBA	PARAGRAP H	ABCCDEA	IN TERRUPT
ABCBDDBA	DEFERRED	ABCCDEAB	IN TERRUPTE D
ABCBDEA	E CONOMIC	ABCCDEAD	COMMENCE
ABCBDEA	DAMAGED	ABCCDEAD	COMMERCE
ABCBDEA	PO LITICAL	ABCCDEADCDE	COMMENCEMEN T
ABCBDEAEC	MANAGEMEN T	ABCCDEBFGHDA	DISSEMINATED
ABCBDEBA	DEFEATED	ABCCDEFA	COMMUNIC ATE
ABCBDEBA	DESERTED	ABCCDEFA	SUPPLIES

~~RESTRICTED~~

MISCELLANEOUS PATTERNS—Continued

ABCCDEFAGHFBE	COMMUNICATION	ABCDA	INSPI RE
ABCCDEFBGHDGAD	CORRESPONDENCE	ABCDA	LOCAL
ABCCDEFGA	R EAPPOINTE D	ABCDA	LAU NCHIN G
ABCCDEFGHAFG	R EAPPOINTMENT	ABCDA	CO NDEMN
ABCDA	S ABOTA GE	ABCDA	MACHI NEGUN
ABCDA	R AILWA Y	ABCDA	NOTIN G
ABCDA	ANIMA L	ABCDA	EXPA NSION
ABCDA	S ANITA RY	ABCDA	CO NTAIN
ABCDA	M ARSHA L	ABCDA	MOU NTAIN
ABCDA	M ARTIA L	ABCDA	I NTERN AL
ABCDA	E ASTWA RD	ABCDA	FRO NTLIN E
ABCDA	N ATURA L	ABCDA	I NTREN CH
ABCDA	N ATURA LIZE	ABCDA	C ONTRO L
ABCDA	TE CHNIC AL	ABCDA	H ORIZO N
ABCDA	COUNC IL	ABCDA	OUTBO ARD
ABCDA	R EACHE D	ABCDA	PROMP T
ABCDA	L EAGUE	ABCDA	RECOR D
ABCDA	EASTE RLY	ABCDA	REPOR T
ABCDA	EASTE RN	ABCDA	RETUR N
ABCDA	W EATHE R	ABCDA	P RIMAR Y
ABCDA	H EAVIE R	ABCDA	RIVER
ABCDA	INS ECURE	ABCDA	ROGER
ABCDA	S ECURE	ABCDA	FA RTHER
ABCDA	R EDUCE	ABCDA	FU RTHER
ABCDA	SCH EDULE	ABCDA	NO RTHER LY
ABCDA	B EFORE	ABCDA	SATIS FY
ABCDA	R EFUGE	ABCDA	SHIPS
ABCDA	R EFUSE	ABCDA	WAR SHIPS
ABCDA	R EGIME NT	ABCDA	THIRT Y
ABCDA	R EGIME NTAL	ABCDA	WI THOUT
ABCDA	EITHE R	ABCDA	EX TRACT
ABCDA	FUS ELAGE	ABCDA	TRACT
ABCDA	D ELIVE R	ABCDA	INS TRUCT
ABCDA	GR ENADE	ABCDA	DES TRUCT ION
ABCDA	ERASE	ABCDA	TWENT Y
ABCDA	OP ERATE	ABCDA	B UREAU
ABCDA	R ESCUE	ABCDA	WESTW ARD
ABCDA	PR ESIDE NT	ABCDAA	R EFUGEE
ABCDA	R ESUME	ABCDAA	C ODEBOO K
ABCDA	D EVICE	ABCDAA	BU SINESS
ABCDA	D EVISE	ABCDAA	DI STRESS
ABCDA	GOING	ABCDAA	STRESS
ABCDA	T HOUGH	ABCDAAD	F ORENOON
ABCDA	C HURCH	ABCDAB	DECIDE
ABCDA	F IGHTI NG	ABCDAB	DECODE
ABCDA	INFLI CT	ABCDAB	SP EARHEA D
ABCDA	EXT INGUI SH	ABCDAB	R EDUCED
ABCDA	INQUI RE	ABCDAB	ENTREN CH
ABCDA	INQUI RY	ABCDAB	ERASER

~~RESTRICTED~~

MISCELLANEOUS PATTERNS—Continued

ABCDAB	GEORGE	ABCDAECD	L ABORATOR Y
ABCDAB	POSTPO NE	ABCDAECE	OUTPOSTS
ABCDAB	RETIRE	ABCDAECFD	EX AMINATION
ABCDAB	ES TIMATI ON	ABCDAED	T RAVERSE
ABCDABA	DECIDED	ABCDAEE	ACTUALL Y
ABCDABAB	INCLININ G	ABCDAEE	EXPRESS
ABCDABC	M AINTAIN	ABCDAEE	THIRTEE N
ABCDABC	M AINTAIN ED	ABCDAEEFAB	THIRTEENTH
ABCDABCEFD	PHOSPHORUS	ABCDAEFA	OV ERWHELME D
ABCDABEFA	ENTRENCHE D	ABCDAEFAB	INFLICTIN G
ABCDAC	L ANGUAG E	ABCDAEFB	P RESCRIBE D
ABCDAC	ANYWAY	ABCDAEFBE	O NEHUNDRED
ABCDAC	GOV ERNMEN T	ABCDAEFC	M ANUFACTU RE
ABCDAC	I NSTANT	ABCDAEFC	PR ESIDENTI AL
ABCDAC	I NSTANT LY	ABCDAEFC	D ISTRIBUT E
ABCDAC	DI SPERSE	ABCDAEFCA	D ISTRIBUTI NG
ABCDAC	RES TRICTI ON	ABCDAEFCA	D ISTRIBUTI ON
ABCDAC	PA TRIOTI C	ABCDAEFD	F LASHLIGH T
ABCDACB	CO NDEMNED	ABCDAEFD	C ONTROVER SY
ABCDACDAEFGB	I NSTANTANEOUS	ABCDAEFD	A SCENSION
ABCDACEFDAF	COINCIDENCE	ABCDAEFD	WINDWARD
ABCDAD	MOVEME NT	ABCDAEFDB	RESTRICTE D
ABCDAD	A MUSEME NT	ABCDAEFDE	RESTRICTI ON
ABCDAD	RIGORO US	ABCDAEFE	PAR ENTHESIS
ABCDADC	S ANITATI ON	ABCDAEFE	RETURNIN G
ABCDADEDAFB	INSTITUTION	ABCDAEFEGE	RE SPONSIBILI TY
ABCDADEFEAGC	ANTIAIRCRAFT	ABCDAEFF	REDCROSS
ABCDAEA	EXTREME	ABCDAEFGAHB	INSPIRATION
ABCDAEA	MAXIMUM	ABCDAEFGC	REGARDING
ABCDAEAB	SU ITABILIT Y	ABCDAEFGD	RESTRAINT
ABCDAEABD	UNI TEDSTATES	ABCDAEFGFE	TR ANSPACIFIC
ABCDAEAE	PAR ENTHESES	ABCDAEFGHC	TWENTYFIVE
ABCDAEB	F IGHTING	ABCDAEFGHFBC	CONSCRIPTION
ABCDAEB	S IGHTING	ABCDBA	PR ACTICA L
ABCDAEB	RAILROA D	ABCDBA	W ATERTA NK
ABCDAEB	REPORTE D	ABCDBA	DIV EBOMBE R
ABCDAEB	RETURNE D	ABCDBA	ENGINE
ABCDAEB	TRACTOR	ABCDBA	S ENTINE L
ABCDAEB	INS TRUCTOR	ABCDBA	R EVOLVE
ABCDAEBA	RECORDER	ABCDBA	S ITUATI ON
ABCDAEBC	DE TONATION	ABCDBAA	ENGINEE R
ABCDAEBFBDC	U NIDENTIFIED	ABCDBAAEDBC	ENGINEERING
ABCDAEBFC	SATISFACT ORY	ABCDBAB	LIABILI TY
ABCDAEC	AVERAGE	ABCDBAD	RE TALIATI ON
ABCDAEC	D ISTRICT	ABCDBAEAD	D ISPOSITIO N
ABCDAEC	OUTPOST	ABCDBAEBE	U NEXPENDED
ABCDAECA	TWENTIET H	ABCDBBA	ANTENNA
ABCDAECAB	I NTERNMENT	ABCDBBA	D ISCUSSI ON
ABCDAECB	D ISTRICTS	ABCDBBDEA	TRA NSMISSION

MISCELLANEOUS PATTERNS—*Continued*

ABCDBCAEB		INTENTION
ABCDBCEA	A	ERODROME
ABCDBEA		INCENDI ARY
ABCDBEA	PR	OTECTIO N
ABCDBEA	IN	TERCEPT
ABCDBEAB	IN	TERCEPTE D
ABCDBEAE	C	ONTINUOU S
ABCDBEAFB		INVENTION
ABCDBEAFCDB	QU	ARTERMASTER
ABCDBEAFD		INCENTIVE
ABCDBEAFD		INTENSIVE
ABCDBECA	E	NCIRCLIN G
ABCDBEFAGABC		ENTANGLEMENT
ABCDBEFAGEB		TEMPERATURE
ABCDBEFBA		DECREASED
ABCDBEFCDAB	C	ONTINUATION
ABCDBEFGA		YESTERDAY
ABCDBEFGAB		ARMOREDCAR
ABCDBEFGBCHIA		DISTINGUISHED
ABCDBEFGHA	P	ERFORMANCE
ABCDCA		AIRCRA FT
ABCDCA		CRITIC
ABCDCA		CRITIC AL
ABCDCA	D	EFICIE NT
ABCDCA		ENGAGE
ABCDCA	P	OSITIO N
ABCDCA	PR	OVISIO N
ABCDCA	FI	REALAR M
ABCDCAAC		PHILIPPI NES
ABCDCAB		ANTITAN K
ABCDCABCA	I	NDEPENDEN T
ABCDCAC		CRITICI SE
ABCDCAC		CRITICI SM
ABCDCAD		OPINION
ABCDCAEAB		ENGAGEMEN T
ABCDCAEB	P	OSITIONS
ABCDCAED	D	EFICIENC Y
ABCDCAED	PR	OVISIONS
ABCDCAEFD		CHARACTER
ABCDCAEFDGHEGA		CHARACTERISTIC
ABCDCBABC	IN	TERPRETER
ABCDCBCEA	HO	STILITIES
ABCDCEA	BRI	DGEHEAD
ABCDCEA	M	EDICINE
ABCDCEA	D	EFINITE
ABCDCEA	S	EPARATE
ABCDCEA		SURPRIS E
ABCDCEAFC	QU	ALIFICATI ON
ABCDCEAFE	P	ERSISTENT
ABCDCEBA		ELIGIBLE
ABCDCECA	D	ESTITUTE
ABCDCECDA	CO	NSTITUTIN G
ABCDCEFGAB		PHOTOGRAPH Y
ABCDCEFGCA	DEM	OBILIZATIO N
ABCDCEFGCA	M	OBILIZATIO N
ABCDDA	R	ECOMME ND
ABCDDA	T	OBACCO
ABCDDA		SHELLS
ABCDDAB	B	EACHHEA D
ABCDDAEACBE		INEFFICIENC Y
ABCDDAEFAF	R	ECOMMENDED
ABCDDAEFGHICE	R	ECOMMENDATION
ABCDDEA		DROPPED
ABCDDEA	AI	RSUPPOR T
ABCDDEA	A	RTILLER Y
ABCDDEAEC		COEFFICIE NT
ABCDDECDFA		SCHOOLHOUS E
ABCDDEFCGHA	MI	SCELLANEOUS
ABCDDEFEACGE		CLASSIFICATI ON
ABCDDEFGGEDBA	R	ECONNAISSANCE
ABCDEA		AERONA UTICS
ABCDEA	R	AILHEA D
ABCDEA		AIRPLA NE
ABCDEA		AMBULA NCE
ABCDEA	CO	ASTGUA RD
ABCDEA	M	ATERIA L
ABCDEA	S	ATURDA Y
ABCDEA	C	AUSEWA Y
ABCDEA	N	AUTICA L
ABCDEA		BLOCKB USTER
ABCDEA	ME	CHANIC
ABCDEA		CHEMIC AL
ABCDEA		CONDUC T
ABCDEA		DISLOD GE
ABCDEA		DOWNED
ABCDEA	B	ECAUSE
ABCDEA	D	ECIPHE R
ABCDEA	D	ECLARE
ABCDEA	OBJ	ECTIVE
ABCDEA	L	ECTURE
ABCDEA	V	EHICLE S
ABCDEA		ENCODE
ABCDEA	COMP	ENSATE
ABCDEA		ENTIRE
ABCDEA	R	EPLACE
ABCDEA	R	EPULSE D
ABCDEA	CONSID	ERABLE
ABCDEA	INT	ERPOSE

MISCELLANEOUS PATTERNS—Continued

ABCDEA	S ERVICE	ABCDEABFD	NATIONALI SM
ABCDEA	EUROPE	ABCDEABFDC	NATIONALIT Y
ABCDEA	EUROPE AN	ABCDEABFE	MARKSMANS HIP
ABCDEA	EXCITE	ABCDEABFFGHD	SHARPSHOOTER
ABCDEA	T HROUGH	ABCDEABFGDHF	W ARDEPARTMENT
ABCDEA	IDENTI CAL	ABCDEAC	AUTOMAT IC
ABCDEA	IDENTI FY	ABCDEAC	AI RCONTRO L
ABCDEA	INHABI TED	ABCDEACFB	ANTEDATIN G
ABCDEA	D IRECTI ON	ABCDEAD	CONTACT
ABCDEA	MEDIUM	ABCDEAD	V ICTORIO US
ABCDEA	SY NCHRON IZE	ABCDEAD	C RUISERS
ABCDEA	JU NCTION	ABCDEADFD	THREATENE D
ABCDEA	CO NFIDEN T	ABCDEAE	ENCODED
ABCDEA	NOTHIN G	ABCDEAE	P ERMANEN T
ABCDEA	E NTRAIN	ABCDEAE	FORTIFI ED
ABCDEA	L OCATIO N	ABCDEAE	REQUIRI NG
ABCDEA	REV OLUTIO N	ABCDEAEFGC	TRADITIONA L
ABCDEA	DEC ORATIO N	ABCDEAFA	R EPLACEME NT
ABCDEA	T ORPEDO	ABCDEAFAGE	EXCITEMENT
ABCDEA	OVERCO MING	ABCDEAFAGHEAID	IDENTIFICATION
ABCDEA	T RAILER S	ABCDEAFB	CLERICAL
ABCDEA	T RAWLER	ABCDEAFB	INVASION
ABCDEA	DI RECTOR	ABCDEAFBC	RESOURCES
ABCDEA	REPAIR	ABCDEAFC	DES IGNATION
ABCDEA	NO RTHWAR D	ABCDEAFC	RES IGNATION
ABCDEA	C RUISER	ABCDEAFC	CO NFIDENTI AL
ABCDEA	I SLANDS	ABCDEAFD	D IMENSION
ABCDEA	STRIPS	ABCDEAFE	ADJUTANT
ABCDEA	SUNRIS E	ABCDEAFE	INTERIOR
ABCDEA	TARGET	ABCDEAFE	I NFLUENCE
ABCDEA	NOR THEAST	ABCDEAFF	R EADINESS
ABCDEA	THREAT	ABCDEAFGA	D ECIPHERME NT
ABCDEA	NOR THWEST	ABCDEAFGAFB	MEDIUMBOMBE R
ABCDEA	TWELFT H	ABCDEAFGD	LEGISLATI ON
ABCDEA	L UMINOU S	ABCDEAFGE	CO MPARTMENT
ABCDEAA	EIGHTEE N	ABCDEAFGEE	SMOKESCREE N
ABCDEAAE	SUBMISSI ON	ABCDEBA	DELAYED
ABCDEAAFED	EIGHTEENTH	ABCDEBA	D ETONATE
ABCDEAB	INVADIN G	ABCDEBA	INDEMNI TY
ABCDEAB	F LEXIBLE	ABCDEBA	D ISPERSI ON
ABCDEAB	NATIONA L	ABCDEBA	RECOVER
ABCDEAB	REQUIRE	ABCDEBA	SURPLUS
ABCDEAB	RESTORE D	ABCDEBAB	ARBITRAR Y
ABCDEAB	OU TSKIRTS	ABCDEBAED	ARBITRATI ON
ABCDEABA	DEMANDED	ABCDEBFA	B RIGADIER
ABCDEABD	IMPEDIME NTA	ABCDEBFAGA	ENCOUNTERE D
ABCDEABE	AT OMICBOMB	ABCDEBFCAGBF	INTERNATIONA L
ABCDEABF	REPAIRED	ABCDEBFDGA	NAVIGATION
ABCDEABFB	REQUIREME NT	ABCDEBFGAF	H EADQUARTER S

MISCELLANEOUS PATTERNS—Continued

ABCDEBFGHA	R ESPONSIBLE
ABCDEBFGHBCGIA	NATURALIZATION
ABCDECA	E NLISTIN G
ABCDECA	PRINCIP AL
ABCDECA	PRINCIP LE
ABCDECA	SKIRMIS H
ABCDECAB	I NTERMENT
ABCDECAC	I NTERVENE
ABCDECACFE	M AINTENANCE
ABCDECAFCDA	TRANSATLANT IC
ABCDECBA	NEGLIGEN T
ABCDECBA	REVOLVER
ABCDECBA	P ROTECTOR
ABCDECBAFB	NEGLIGENCE
ABCDECCFA	DISCUSSED
ABCDECDCAFC	I NTERFERENCE
ABCDECFA	ENCIRCLE
ABCDECFA	EVACUATE
ABCDECFBA	SEAPLANES
ABCDECFEA	STANDARDS
ABCDEDA	N EWSPAPE R
ABCDEDA	MARITIM E
ABCDEDA	CO NTRABAN D
ABCDEDA	C OALITIO N
ABCDEDA.	BA ROMETER
ABCDEDA	GY ROMETER
ABCDEDA	HYD ROMETER
ABCDEDA	HYG ROMETER
ABCDEDA	PSYCH ROMETER
ABCDEDAB	C ONDITION
ABCDEDAC	REC OGNITION
ABCDEDAFC	N EWSPAPERS
ABCDEDFA	DICTATED
ABCDEDFA	EXCAVATE
ABCDEDFA	EXHIBITE D
ABCDEDFAC	ANTICIPAT E
ABCDEDFAC	CLEARANCE
ABCDEDFACDGB	ANTICIPATION
ABCDEDFCAB	INTERESTIN G
ABCDEDFCGAHB	INAUGURATION
ABCDEDFDA	ARTIFICIA L
ABCDEDFDEAB	C ONSTITUTION
ABCDEDFDGHAIF	CHRONOLOGICAL
ABCDEDFGA	PR OCLAMATIO N
ABCDEDFGA	P RELIMINAR Y
ABCDEDFGABHED	INDETERMINATE
ABCDEDFGADB	P RELIMINARIE S
ABCDEDFGHAGD	ADMINISTRATI VE
ABCDEDFGHAGDIE	ADMINISTRATION

ABCDEEA	ENROLLE D
ABCDEEA	P ERSONNE L
ABCDEEA	IMPASSI BLE
ABCDEEA	IMPOSSI BLE
ABCDEEACB	S IGNALLING
ABCDEEAFDBC	INTELLIGENT
ABCDEEAFDBGD	INTELLIGENCE
ABCDEEDFGBA	RECONNOITER
ABCDEEDFGBAFE	RECONNOITERIN G
ABCDEEFAB	ENROLLMEN T
ABCDEEFAB	C ONFESSION
ABCDEEFAE	EMBASSIES
ABCDEEFDGFA	DISAPPEARED
ABCDEEFGCAHB	INTERRUPTION
ABCDEFA	C ABLEGRA M
ABCDEFA	AMERICA N
ABCDEFA	C AMOUFLA GE
ABCDEFA	CHRONIC AL
ABCDEFA	CONFLIC T
ABCDEFA	DIS CREPANC Y
ABCDEFA	S EABORNE
ABCDEFA	EMPLOYE R
ABCDEFA	ENCIPHE R
ABCDEFA	ENFORCE
ABCDEFA	ENLISTE D
ABCDEFA	D EPLOYME NT
ABCDEFA	EQUIPME NT
ABCDEFA	FIGHT ERPLANE
ABCDEFA	ESCORTE D
ABCDEFA	D ESCRIBE
ABCDEFA	J ETPLANE
ABCDEFA	EXCLUDE
ABCDEFA	INCLUSI VE
ABCDEFA	LOGICAL
ABCDEFA	F ORMATIO N
ABCDEFA	T RANSFER
ABCDEFA	REGULAR
ABCDEFA	P RISONER
ABCDEFA	SAILORS
ABCDEFA	SECTORS
ABCDEFA	SERIOUS LY
ABCDEFA	E STABLIS H
ABCDEFA	TONIGHT
ABCDEFAA	EMPLOYEE
ABCDEFAAF	T RANSFERRE D
ABCDEFAAGC	T RANSFERRIN G
ABCDEFAB	INCLUDIN G
ABCDEFAB	RADIOGRA M
ABCDEFAB	P REMATURE

MISCELLANEOUS PATTERNS—Continued

ABCDEFABA	EMPLACEME NT
ABCDEFAC	INTEGRIT Y
ABCDEFAC	P RISONERS
ABCDEFACB	IN TRODUCTOR Y
ABCDEFACD	ALTERNATE
ABCDEFACGF	ALTERNATIN G
ABCDEFAD	CONTRACT
ABCDEFAD	D ESTROYER
ABCDEFAD	INTERVIE W
ABCDEFAD	OPERATOR
ABCDEFAD	FI RECONTRO L
ABCDEFAD	P ROCEDURE
ABCDEFADB	D ESTROYERS
ABCDEFADF	T RANSVERSE
ABCDEFAE	D ISCONTIN UE
ABCDEFAEGHEC	D ISCONTINUANC E
ABCDEFAF	EXPANDED
ABCDEFAF	I MPROVEME NT
ABCDEFAFCD	R ADIOSTATIO N
ABCDEFAGA	ENCIPHERE D
ABCDEFAGAB	ENFORCEMEN T
ABCDEFAGB	AEROPLANE
ABCDEFAGB	D ETACHMENT
ABCDEFAGB	INFLATION
ABCDEFAGB	REINFORCE
ABCDEFAGB	TRAJECTOR Y
ABCDEFAGBDB	REIMBURSEME NT
ABCDEFAGBHBD	REINFORCEMEN T
ABCDEFAGC	INTERDICT
ABCDEFAGCAHB	INTERDICTION
ABCDEFAGE	D EPARTMENT
ABCDEFAGEC	D EPARTMENTA L
ABCDEFAGFD	REGISTRATI ON
ABCDEFAGHAB	ENCIPHERMEN T
ABCDEFAGHEBC	CONFISCATION
ABCDEFAGHFD	INVESTIGATE
ABCDEFAGHFAIB	INVESTIGATION
ABCDEFAGHFAIBE	INVESTIGATIONS
ABCDEFAGHIF	B REAKTHROUGH
ABCDEFBA	DECLARED
ABCDEFBA	DEPARTED
ABCDEFBA	DEPLOYED
ABCDEFBA	DEPORTED
ABCDEFBA	DETACHED
ABCDEFBA	EMPLOYME NT
ABCDEFBA	ENTRAINE D
ABCDEFBA	REGISTER
ABCDEFBA	P ROJECTOR
ABCDEFBAB	MEASUREME NT
ABCDEFBABGHD	MEASUREMENTS
ABCDEFBGA	ENDURANCE
ABCDEFBGBA	DECIPHERED
ABCDEFCA	ESTIMATE
ABCDEFCA	NORTHERN
ABCDEFCAB	ESTIMATES
ABCDEFCAD	D OMINATION
ABCDEFCAGFC	ESTIMATEDAT
ABCDEFCBA	DETONATED
ABCDEFCCFA	DISTRESSED
ABCDEFCEA	DISPERSED
ABCDEFCGA	ELABORATE
ABCDEFDA	D EPARTURE
ABCDEFDAB	C USTOMHOUS E
ABCDEFDBAB	INTERVENIN G
ABCDEFDBCAGB	INTERVENTION
ABCDEFDEAB	INTERFERIN G
ABCDEFDGAB	DEM ONSTRATION
ABCDEFDGAHCD	INTERMEDIATE
ABCDEFDGHA	HYDROGRAPH IC
ABCDEFEA	R EINSTATE
ABCDEFEAB	F INGERPRIN T
ABCDEFEAGACE	R EINSTATEMENT
ABCDEFEAGDB	CERTIFICATE
ABCDEFECACD	THERMOMETER
ABCDEFECAE	CONFERENCE
ABCDEFEDCGCAHB	INTERPRETATION
ABCDEFEFA	C OMPETITIO N
ABCDEFEGA	D EMOBILIZE
ABCDEFEGA	C OMPUTATIO N
ABCDEFFA	UN DERSTOOD
ABCDEFFA	IMPRESSI ON
ABCDEFFAGE	IMPRESSIVE
ABCDEFFEDAGBC	INSTALLATIONS
ABCDEFFGAB	C ONGRESSION AL
ABCDEFGA	DISARMED
ABCDEFGA	M ECHANIZE D
ABCDEFGA	T ECHNIQUE
ABCDEFGA	R ECOGNIZE
ABCDEFGA	ENFILADE
ABCDEFGA	EQUALIZE
ABCDEFGA	EQUIPAGE
ABCDEFGA	EQUIVALE NT
ABCDEFGA	D ESIGNATE
ABCDEFGA	EXCHANGE
ABCDEFGA	GROUPING
ABCDEFGA	GUARDING
ABCDEFGA	INSECURI TY
ABCDEFGA	D IPLOMATI C

MISCELLANEOUS PATTERNS—Continued

ABCDEFGA	E NTRUCKIN G	ABCDEFGDHFAE	ORGANIZATION
ABCDEFGA	NUMBERIN G	ABCDEFGEA	H EAVYBOMBE R
ABCDEFGA	OBJECTIO N	ABCDEFGEHA	D ESCRIPTIVE
ABCDEFGA	OPERATIO N	ABCDEFGFABF	I NCOMPETENCE
ABCDEFGA	SOLDIERS	ABCDEFGFAG	I NCOMPETENT
ABCDEFGA	DI SPATCHES	ABCDEFGGAG	H EAVYLOSSES
ABCDEFGA	WITHDRAW	ABCDEFGHA	CONSPIRAC Y
ABCDEFGA	WITHDREW	ABCDEFGHA	DOMINATED
ABCDEFGAB	D ESPATCHES	ABCDEFGHA	C ENTRALIZE
ABCDEFGAB	U NDERSTAND	ABCDEFGHA	EXCLUSIVE
ABCDEFGAB	WITHDRAWI NG	ABCDEFGHA	EXPANSIVE
ABCDEFGABF	ENLISTMENT	ABCDEFGHA	EXPLOSIVE
ABCDEFGAC	I NSTRUMENT	ABCDEFGHA	MECHANISM
ABCDEFGAC	F OUNDATION	ABCDEFGHAB	C ONSUMPTION
ABCDEFGACB	I NSTRUMENTS	ABCDEFGHADB	INFORMATION
ABCDEFGAD	SOUTHEAST	ABCDEFGHAGC	CONVALESCEN T
ABCDEFGAD	SOUTHWEST	ABCDEFGHBA	DESIGNATED
ABCDEFGADG	SOUTHWESTE RN	ABCDEFGHBA	DESPATCHED
ABCDEFGAEHBC	CONSTRUCTION	ABCDEFGHBIKA	DISORGANIZED
ABCDEFGAFE	IMPRACTICA BLE	ABCDEFGHCAEB	INTRODUCTION
ABCDEFGAG	WITHDRAWA L	ABCDEFGHCAEB	D ISCREPANCIES
ABCDEFGAHB	INSPECTION	ABCDEFGHDAB	C ONFIRMATION
ABCDEFGAHCGIDE	RECONSTRUCTION	ABCDEFGHDGCA	NORTHWESTERN
ABCDEFGBA	DESCRIBED	ABCDEFGHDIKA	REVOLUTIONAR Y
ABCDEFGBA	DESTROYED	ABCDEFGHEEHA	COUNTERATTAC K
ABCDEFGBA	DETRAINED	ABCDEFGHFA	D EMONSTRATE
ABCDEFGBA	REMAINDER	ABCDEFGHFCAG	AGRICULTURAL
ABCDEFGBA	TRANSPORT	ABCDEFGHIA	DISPATCHED
ABCDEFGBACAHGD	TRANSPORTATION	ABCDEFGHIA	OBSERVATIO N
ABCDEFGBAE	TRANSPORTS	ABCDEFGHIA	SUBMARINES
ABCDEFGBHA	ESTABLISHE D	ABCDEFGHIAB	C ONVERSATION
ABCDEFGBHIAKC	ESTABLISHMENT	ABCDEFGHIAE	C OMPENSATION
ABCDEFGCAG	CONFIDENCE	ABCDEFGHIAF	R OADJUNCTION
ABCDEFGCHEA	RANGEFINDER	ABCDEFGHIDAB	C ONSIDERATION
ABCDEFGDAHB	INSTRUCTION	ABCDEFGHIFKA	SEARCHLIGHTS
ABCDEFGDAHBC	INSTRUCTIONS	ABCDEFGHIGBA	DEMONSTRATED
ABCDEFGDBFHA	CE NTRALIZATION	ABCDEFGHIJDA	SIMULTANEOUS
ABCDEFGDHAIC	OBSTRUCTIONS		

D. DIGRAPHIC IDIOMORPHS: GENERAL

AB AB

-G EN ER	AL AL	AR M-
NE	ED ED	
-P RO CE	ED ED	
-S UC CE	ED ED	
-D ET RA	IN IN	G-
-L	IN IN	G-
-M	IN IN	G-
OB TA	IN IN	G-
QU	IN IN	E-
RA	IN IN	G-
RE MA	IN IN	G-
SH	IN IN	G-
-T RA	IN IN	G-
CR	IS IS	
PO SI TI	ON ON	
-A	RE RE	EN FO RC ED
-A	SU SU	AL
BO	TH TH	E-
WI	TH TH	E-
-P AR	TI TI	ON
RE PE	TI TI	ON
	VI VI	D-

AB — AB

-M	AI NT AI	N-
RE	AR GU AR	D-
	CH UR CH	
	DE CI DE	
	DE CO DE	
	DI VI DI	NG
SP	EA RH EA	D-
-R	ED UC ED	
-S CH	ED UL ED	
-B	EE NN EE	DE D-
	EM BL EM	
AM	EN DM EN	T-
CO NT	EN TM EN	T-
-S EV	EN TE EN	
-S EV	EN TE EN	TH
	EN TR EN	CH
	ER AS ER	

AB — AB

TH	ER EF ER	EN CE
TH	ER ES ER	VE
WH	ER EV ER	
-C AR EL	ES SN ES	S-
	GE OR GE	
SC	HO OL HO	US E-
-I LL UM	IN AT IN	G-
	IN CL IN	E-
-F IR	IN GL IN	E-
MA	IN TA IN	
-I NF AL	LI BI LI	TY
-A	ME ND ME	NT
SO	ME TI ME	
-O	NE NI NE	
	NO TK NO	WN
	NO WK NO	WN
-A PP OI	NT ME NT	
-C ON TE	NT ME NT	
-C	OM PR OM	IS E-
-P	ON TO ON	
-T HR	OU GH OU	T-
-N	OW KN OW	N-
	PH OS PH	OR US
	PO ST PO	NE
TR OO	PS HI PS	
PA	RA PH RA	SE
-P	RE FE RE	NC E-
	RE FE RE	NC E-
-T HE	RE FO RE	
-P	RE PA RE	
	RE TI RE	
	RE VE RE	NT
-C	RO SS RO	AD S-
CA RE LE	SS NE SS	
AT	TE MP TE	D-
	TH AT TH	E-
-F OR	TH WI TH	
-I NV ES	TI GA TI	ON
ES	TI MA TI	ON
-D ES	TI NA TI	ON
AC	TI VI TI	ES
-H	UM DR UM	

~~RESTRICTED~~

AB — — AB

–P	AN	AM	AC	AN	AL
	AR	BI	TR	AR	Y–
	AS	SO	ON	AS	
AC	CE	PT	AN	CE	
	EM	PL	AC	EM	EN T–
–Q UA RT	ER	MA	ST	ER	
–I NT	ER	PR	ET	ER	
–A CC	ES	SO	RI	ES	
	IN	CL	UD	IN	G–
–D	IR	EC	TF	IR	E–
TO	MO	RR	OW	MO	RN IN G–
PA	NA	MA	CA	NA	L–
–I	NT	ER	ME	NT	
–I	NT	ER	VE	NT	IO N–
CO	NT	IN	GE	NT	
–C	ON	DI	TI	ON	
–T OM	OR	RO	WM	OR	NI NG
	RA	DI	OG	RA	M–
	RE	AS	SU	RE	
–P	RE	MA	TU	RE	
–D EF EN	SI	VE	PO	SI	TI ON
IN	TE	RD	IC	TE	D–
QU AR	TE	RM	AS	TE	R–
IN	TE	RP	RE	TE	R–
IN	TE	RR	UP	TE	D–
–F OR	TI	FI	CA	TI	ON

AB — — — AB

	AR	MO	RE	DC	AR	
	EN	FO	RC	EM	EN	T–
RE	EN	FO	RC	EM	EN	TS
	IN	DE	TE	RM	IN	AT E–
	IN	TE	RE	ST	IN	G–
	IN	TE	RF	ER	IN	G–
	IN	TE	RV	EN	IN	G–
–I	NC	OM	PE	TE	NC	E–
–C	ON	GR	ES	SI	ON	AL
–D EM	ON	ST	RA	TI	ON	
–C	ON	SU	MP	TI	ON	
	PH	OT	OG	RA	PH	
	TH	IR	TE	EN	TH	

AB — — — — AB

–I	NS	TA	LL	AT	IO	NS
–C	ON	CE	NT	RA	TI	ON
–C	ON	FL	AG	RA	TI	ON
–C	ON	SI	DE	RA	TI	ON

AB — AB AB

	IN	CL	IN	IN	G–
MA	IN	TA	IN	IN	G–

~~RESTRICTED~~

E. DIGRAPHIC IDIOMORPHS: PLAYFAIR

<u>AB BA</u>

SC	AB BA	RD
	AF FA	BL E-
	AF FA	IR
-B	AG GA	GE
-H AW	AI IA	N-
	AL LA	RE AS
-B	AL LA	ST
-F	AL LA	CY
IN ST	AL LA	TI ON S-
-P AR	AL LA	X-
	AP PA	RA TU S-
	AP PA	RE L-
	AP PA	RE NT
	AP PA	RE NT LY
	AR RA	NG E-
	AR RA	Y-
-B	AR RA	CK S-
-B	AR RA	GE
-E MB	AR RA	SX SE D-
-N	AR RA	TI ON
	AS SA	IL AN T-
	AS SA	UL T-
-A MB	AS SA	DO R-
-I MP	AS SA	BL E-
-M	AS SA	CR E-
-P	AS SA	GE
	AT TA	CH
	AT TA	CK
	AT TA	IN
-B	AT TA	LI ON
-R	AT TA	N-
	BO OB	YT RA P-
IN	DE ED	
-W	EB BE	D-
	EF FE	CT
	EF FE	CT IV E-
CO MP	EL LE	D-
-E XC	EL LE	NC E-
-E XC	EL LE	NT
-E XP	EL LE	D-
-I MP	EL LE	D-
-P	EL LE	T-
PR OP	EL LE	D-
-R EP	EL LE	D-

SH	EL LE	D-
-H	EM ME	DI N-
ST	EM ME	D-
ST	EP PE	D-
AV	ER RE	D-
CO NF	ER RE	D-
-I NT	ER RE	D-
-R EF	ER RE	D-
	ES SE	NC E-
	ES SE	NT IA L-
AD DR	ES SE	S-
-C OM PR	ES SE	D-
CO NF	ES SE	D-
IM PR	ES SE	D-
-L	ES SE	N-
-M	ES SE	NG ER
PR	ES SE	D-
PR OF	ES SE	D-
-P RO GR	ES SE	D-
-S TR	ES SE	D-
-S TR	ES SE	S-
-V	ES SE	L-
WI TN	ES SE	S-
AB	ET TE	D-
-C IG AR	ET TE	S-
-B	ET TE	R-
-L	ET TE	R-
-E IG	HT TH	RE E-
-R	IB BI	NG
FO RB	ID DI	NG
-D	IF FI	CU LT
-B	IL LI	ON
-F	IL LI	NG
-K	IL LI	NG
-M	IL LI	ME TE R-
-M	IL LI	NG
-M	IL LI	ON
SH	IL LI	NG
SP	IL LI	NG
-T	IL LI	NG
-W	IL LI	AM
-W	IL LI	NG
	IM MI	GR AN T-
	IM MI	GR AT IO N-

~~RESTRICTED~~

			AB	BA				
			IM	MI	NE	NT		
		SW	IM	MI	NG			
	–B	EG	IN	NI	NG			
		SP	IN	NI	NG			
		–W	IN	NI	NG			
		CL	IP	PI	NG			
		SH	IP	PI	NG			
	–S	TR	IP	PI	NG			
			IR	RI	GA	TI	ON	
		–M	IS	SI	NG			
		–M	IS	SI	ON			
	–A	DM	IS	SI	ON			
		EM	IS	SI	ON			
		–H	IS	SI	NG			
	PE	RM	IS	SI	ON			
TR	AN	SM	IS	SI	ON			
		EM	IT	TI	NG			
		–F	IT	TI	NG			
	–S	PL	IT	TI	NG			
	PE	RM	IT	TI	NG			
–A	FT	ER	NO	ON				
	FO	RE	NO	ON				
			NO	ON	TI	ME		
		–F	OL	LO	W–			
		–H	OL	LO	W–			
		–C	OM	MO	N–			
		–C	OM	MO	TI	ON		
PO	SI	TI	ON	NO	RT	HO	F–	
	–R	EC	ON	NO	IT	ER		
			OP	PO	RT	UN	E–	
			OP	PO	RT	UN	IT	Y–
			OP	PO	SE			
			OP	PO	SI	TE		
			OP	PO	SI	TI	ON	
		–C	OR	RO	BO	RA	TE	
		–C	OR	RO	DE			
	–T	OM	OR	RO	W–			
		–B	OT	TO	M–			
		–C	OT	TO	N–			
		CA	RE	ER				
		–S	UC	CU	MB	ED		

			AB	—	BA		
		PR	AC	TI	CA	BL	E–
		PR	AC	TI	CA	L–	
		–T	AC	TI	CA	L–	
	–D	IV	EB	OM	BE	R–	
			EN	GI	NE	ER	
		–G	EN	UI	NE		
	–I	NT	ER	FE	RE		
	–I	NT	ER	FE	RE	NC	E–
	–P	EN	ET	RA	TE		
		–R	EV	OL	VE	R–	
			IN	FI	NI	TE	
		–D	IS	PO	SI	TI	ON
		–S	IT	UA	TI	ON	
		CA	NA	DI	AN		
VE	TE	RI	NA	RI	AN		
		NI	NE	TE	EN		
		NI	NE	TE	EN	TH	
			PE	RC	EP	TI	ON
		–P	RE	MI	ER		
	–S	UR	RE	ND	ER		
	–O	UR	SE	LV	ES		
	TH	EM	SE	LV	ES		
		DE	SE	RV	ES		
		RE	SE	RV	ES		
			SE	RV	ES		

~~RESTRICTED~~

	AB	—	—	BA		
	DE	BA	RK	ED		
	DE	CL	AR	ED		
	DE	FE	ND	ED		
	DE	MA	ND	ED		
	DE	PA	RT	ED		
	DE	PL	OY	ED		
	DE	PO	RT	ED		
	DE	SE	RT	ED		
	DE	TA	CH	ED		
PR	EC	ED	EN	CE		
	EM	PL	OY	ME	NT	
	EN	TR	AI	NE	D–	
	ME	AS	UR	EM	EN	T–
	NE	GL	IG	EN	CE	
	NO	TA	TI	ON		
	PA	RA	GR	AP	H–	
	RE	CE	IV	ER		
	RE	CO	RD	ER		
	RE	GI	ST	ER		
	RE	PE	AT	ER		
	RE	PO	RT	ER		
	RE	VO	LV	ER		
–P	RO	JE	CT	OR		
AS	SE	MB	LI	ES		

	AB	—	—	—	BA	
	DE	SE	CR	AT	ED	
	DE	SI	GN	AT	ED	
	DE	SP	AT	CH	ED	
	EN	EM	YP	LA	NE	S–
–D	ET	ER	IO	RA	TE	
–S	EV	EN	TY	FI	VE	
	IR	RE	GU	LA	RI	TY
	NO	MI	NA	TI	ON	
	SU	SP	IC	IO	US	

AB	—	—	—	—	BA
DE	MO	NS	TR	AT	ED
NO	TI	FI	CA	TI	ON

RESTRICTED

F. DIGRAPHIC IDIOMORPHS: FOUR-SQUARE[1]

(Grouped by number of significant letters in the idiomorphic pattern)

Two letters

A- A-

B LO CK AD ED
I NV AD ED
D AM AG E
CO MM AN DS
I SL AN DS
A IR PL AN ES
E NE MY PL AN ES
DE SI GN AT ED
E ST IM AT ED
I ND IC AT ED
C AV AL RY
N AV AL
P RO CE DU RE
ME CH AN IZ ED
IM ME DI AT EL Y
WI TH DR AW
WI TH DR EW
EM ER GE NC Y
L IE UT EN AN T
FI FT EE N
FI FT H
FI FT Y
BR ID GE HE AD
V IC IN IT Y
W IT HD RA W
A DD IT IO NA L
A MM UN IT IO N
CO ND IT IO N
RE CO GN IT IO N
E LE ME NT
MI LI TA RY
MI NI MU M
NI NT H
P OI NT
T OM OR RO W
PO NT ON
RE QU ES T
RE QU IR E
P RI SO NE R
RE SI ST AN CE
D IS PO SI TI ON
PO SI TI ON
SO UT H

A- A-

SQ UA DR ON
FI GH TE RP LA NE
MO TO RI ZE D
D EP AR TU RE
UN US UA L

A- -- A-

S AB OT AG E
D ET AC HM EN T
H AS BE EN
BA TT AL IO N
BO MB ED
CA SU AL TI ES
CA SU AL TY
CO MB AT
CO OR DI NA TE S
DI RE CT IO N
DI SP AT CH
ME DI UM BO MB ER
DI VE BO MB ER
R OA DJ UN CT IO N
R EP LA CE ME NT
R ET RE AT
S EV ER AL
JU NC TI ON
CO NF IR MA TI ON
I NF OR MA TI ON
I NT EL LI GE NC E
PA TR OL
SA BO TA GE
SE VE RE
AC TI VI TY
A TT EN TI ON
S UC CE SS FU LL Y

A- -- -- A-

AR TI LL ER Y
AT TA CK ED
R EE NF OR CE
R EE NF OR CE ME NT
ID EN TI FY
IM PA SS IB LE
IM PO SS IB LE

A- -- -- A-

MO VE ME NT
E MP LA CE ME NT
PE RS ON NE L
A RT IL LE RY

A- -- -- -- A-

CO MM UN IC AT IO NS
CO NC EN TR AT E
R EO RG AN IZ AT IO N
LI EU TE NA NT
CO NS TR UC TI ON

A- -- -- -- -- A-

CO MM IS SI ON ED

-B -B

UN AB LE
OB ST AC LE
AD VA NC E
AG AI NS T
R AI LH EA D
PR EP AR AT IO N
A SS AU LT
B OM BA RD
A IR BO RN E
S EA BO RN E
A DV AN CI NG
VI CI NI TY
DE TA CH
DE TA CH ME NT
H AV EB EE N
M OV EM EN T
EN EM Y
R ES ER VE
R ET UR N
FL AN K
FO LL OW
B AG GA GE
HA SB EE N
A PP RO AC HI NG
DE BO UC HI NG
L AU NC HI NG
I MM ED IA TE LY

[1] See subpar. ___, Section IX.

RESTRICTED

Two letters (cont.)

```
       -B -B
 IN IT IA TE
     F IF TH
 TE RR IT OR Y
     S IX TY
M IS CE LL AN EO US
     E LE VA TI ON
     E LE VE N
       LI AI SO N
    DA MA GE
       MO RN IN G
     U NU SU AL
       OB JE CT IV E
     C OL ON
     C OL ON EL
SU PE RI OR IT Y
     M OT OR IZ ED
       OU TS KI RT S
 EQ UI PM EN T
   A VE RA GE
   B AR RA GE
    AI RC RA FT
AN TI AI RC RA FT
       RE MA IN
R EQ UI RE ME NT
   M IS SI NG
```

```
          -B -B
     P ER SO NN EL
ES TI MA TE DA T
     P LA TO ON
        S UP PL Y
        S UP PO RT
       NA VA LB AS E
     F OR WA RD
    WI ND WA RD

          -B -- -B
        C AS UA LT Y
        P AT RO LS
  B AT TL ES HI PS
          GE NE RA L
     W IL LA TT AC K
     T RA NS MI SS IO N
     R EC OG NI TI ON
     T RO OP SH IP
          RE GI ME NT
       CA RR IE RS
       MI SS IO NS
          TW EN TY
     R EQ UE ST ED
```

```
  -B -- -- -B
I DE NT IF IC AT IO N
M EC HA NI ZE D
D EP LO YM EN T
M ES SE NG ER
D ES TR OY ER
A IR SU PP OR T
V IS IB IL IT Y
  ME SS EN GE R
I MP AS SI BL E
I MP OS SI BL E
A NT IA IR CR AF T
C OM MA ND IN G
  OP ER AT IO N
  PR IS ON ER
  PR OC ED UR E
  RE EN FO RC E
  TR AN SP OR TA TI ON
  YE ST ER DA Y

    -B -- -- -- -B
  R EC OM ME ND ED
    HE AV YL OS SE S
R EC OM ME ND AT IO N
  C OM MU NI CA TI ON
R EC ON NO IT ER IN G
```

Three letters

```
    A- A- A-
  N AV AL BA SE
R EQ UI SI TI ON
```

```
A- A- -- A-
RE QU ES TE D
```

```
       -B -B -B
B OM BA RD ME NT
     EL EM EN TS
  EN GA GE ME NT
```

Four letters

```
   AB A- -B
 H EA DQ UA RT ER S
   EL EV EN

   AB -B A-
   CA NC EL
RE CO NN AI SS AN CE

   AB -B -- A-
   AD VA NC ED
   EN EM YT AN KS

   AB -- A- -B
   SI GH TI NG
```

```
  A- AB -B
  AD DI TI ON AL

  A- AB -- -B
  SO UT HW ES T

  A- A- -B -B
W IT HD RA WA L

  A- A- -- A- A-
  CO MM AN DI NG

  A- A- -- -B -B
  RE QU IR EM EN T
```

```
  A- -B AB
M OR NI NG
P OS TP ON E

  A- -B -B -- A-
  RE CO NN OI TE R

  A- -B -- AB
  IN TE RD IC T

  A- -B -- A- -B
S AT IS FA CT OR Y

  A- -- A- C- C-
  DI SP AT CH ES
```

Four letters (cont.)

```
A- -- -- C- A- C-
RO AD JU NC TI ON

      -B AB A-
DI SP OS IT IO N
    P OS IT IO N
      PR ES EN T
   RE PR ES EN T

   -B A- AB
RE PE AT ED

   -B A- A- -B
   DE ST RO YE R

   -B A- -B -- A-
   UN ID EN TI FI ED

  -B A- -- AB
U NS UC CE SS FU L

-B A- -- A- -B
ME DI UM BO MB ER

-B A- -- -B A-
VI SI BI LI TY

-B A- -- -- AB
IN FO RM AT IO N

-B A- -- -- A- -B
IN ST AL LA TI ON

-B -D -B -- -D
CR OS SR OA DS

   -B -D -D -B
AI RS UP PO RT

  -B -D -- -D -B
  IN ST RU CT IO N
C ON ST RU CT IO N

        -B -- A- AB
F IG HT ER PL AN ES

  -B -- A- -- -- AB
E ST AB LI SH ME NT

  -B -- -B A- A-
  EN CO UN TE RE D

  -B -- -- -B -D -D
  RE IN FO RC EM EN T
```

Five letters

```
     A- -B AB -- -B
     NA VA LA TT AC K

     A- -B -- -B AB
R EC ON NA IS SA NC E

-B A- A- -- AB
DI ST RI BU TI ON

   -B A- -B AB
RE PL AC EM EN T

-B -D -- -D -B -D
IN ST RU CT IO NS
```

Six letters

```
  AB CB C- A-
P OS IT IO NS

  AB -D -D AB
C ON DI TI ON
  RA DI OG RA M

A- A- -B AB A-
RE QU IS IT IO N

     A- CB -- A- CB
Q UA RT ER MA ST ER

A- CB -- CB A-
SC HO OL HO US E

A- -- CB A- -- CB
ID EN TI FI CA TI ON

  -B AE AD -D
A DM IN IS TR AT IV E
```

Seven letters

```
-B AD -- -B -D AD
RE EN FO RC EM EN T
```

Eight letters

```
AB -B AD -- -B AD
QU AR TE RM AS TE R

AB -B C- AB CB
EM PL AC EM EN T

AB -D C- AD C- -B
IN TE RD IC TI ON
```

(BLANK)

APPENDIX 10

COMMUNICATION INTELLIGENCE OPERATIONS

1. Communication intelligence processes. The principal processes of communication intelligence operations are as follows:

a. Interception of communication signals or messages and forwarding raw traffic[1] to communication intelligence centers for study.

b. Radio direction finding and radio position finding operations; identification of transmitters and radio operators by means of radio fingerprinting and Morse operator analysis, respectively.

c. Traffic analysis, or the study of the external characteristics of communications, without recourse to cryptanalysis of the message texts.

d. Cryptanalysis or solution of the texts of messages.

e. Translation and emendation of the message texts.

f. Large-scale production or exploitation of communication intelligence, after the initial break-in.

g. Evaluation of information, yielding military intelligence.

h. Collation, correlation and comparison of communication intelligence with other intelligence sources.

i. Distribution of communication intelligence to consumers.

2. Interception, radio direction finding, and radio position finding.--a. Messages transmitted by radio can be manually copied or automatically recorded by suitably adjusted radio apparatus located within range of the transmitter. Some messages transmitted over wire lines can likewise be manually copied or automatically recorded by special apparatus suited for the purpose. Correspondents have no way of knowing whether or not radio transmissions are being copied by the enemy, since the interception does not interfere in the slightest degree with

[1] Raw traffic is unprocessed intercepted traffic.

(Appendix to NSA text, "Military Cryptanalytics, Part I", Dec 1952)

signals being transmitted. Interception of wire traffic is much more difficult than of radio, mainly because the equipment either must be located very near the wire line, or connected directly to it.

b. It is possible to determine, with a fair degree of accuracy, the direction of a radio transmitter from a given location and, by establishing the direction from two or more locations, it is possible to determine the geographical location of the transmitter. The science which deals with the means and methods of determining the direction in which a radio transmitter lies is called radio direction finding; the method of determining the geographical location of a radio transmitter, by the use of two or more direction-finding installations, is called radio position finding.

3. Radio fingerprinting and Morse operator analysis.--a. Radio fingerprinting is one of the valuable adjuncts of signal analysis, a communications-engineering sister of traffic analysis. Radio fingerprinting consists of the analysis of the characteristics of the emissions of an individual radio transmitter by means of oscillograms of the emitted radio waves. The oscillograms of the emissions of unidentified radio stations are compared with those of known transmitters or radio stations, and thus it is possible to equate different call signs or different frequencies which have been used by the same transmitting station. Radio fingerprinting is normally not considered conclusive in itself, but is correlated with other analyses or confirmations.

b. Another valuable adjunct of communication intelligence operations is Morse operator analysis. This analysis deals with the radio operators' characteristics when hand-sending is used; the analysis is based on the relative lengths and spacing of the dots and dashes composing the various Morse characters. It is a rarity when a radio operator will transmit a Morse character perfectly, i.e., make the dashes the correct length in respect to the dots, without any individuality (known as the "fist" or "swing") in the sending. Most operators do have certain individual characteristics or tendencies in the sending of certain Morse characters. In past decades, radio operators have identified characteristic "fists" of other operators based on the aural recognition of the rhythm of certain Morse characters. This art has been made more scientific through the use of actual physical measurement and through the assignment of a classificatory coding to the individualities present in the undulator-tape recording[2] of a Morse transmission. By matching measurements, individual radio operators may be identified, in spite of changes of call signs and other elements of the transmission.

4. Traffic Analysis.--a. A great deal of information of military value can be obtained by studying signal communications without solving encrypted messages constituting the traffic. The procedure and the methods used have yielded results of sufficient importance to warrant the

---

[2] Such recordings take the form of a wavy inked line on a paper tape, being a visual representation of the dots and dashes as transmitted.

application of a special term to this field of study; namely, traffic analysis, which is the study of signal communications and intercepted or monitored traffic for the purpose of gathering military information without recourse to cryptanalysis.

b. In general terms, traffic analysis is the careful inspection and study of signal communications for the purpose of penetrating camouflage superimposed upon the communication network for purposes of security. Specifically, traffic analysis reconstructs radio communication networks by: (1) noting volume, direction, and routing of messages; (2) correlating transmission frequencies and schedules used among and within the various networks; (3) determining directions in which transmitters lie, by means of radio direction finding; (4) locating transmitters geographically, by radio position finding; (5) developing the system of assigning and changing radio call signs; and (6) studying all items that constitute messages originated by operators and exchanged among themselves on a radio net.[3]

c. From a correlation of general and specific information derived by means of the foregoing procedures, traffic analysis is able not only to ascertain the geographic location and disposition of troops and military units (technically called "Order of Battle") and important troop movements, but also to predict with a fair degree of reliability the areas and extent of immediately pending or future activities. Traffic analysis procedures are followed to obtain information of value concerning the enemy, and to determine what information concerning our own forces is made available to the enemy through our own signal communications. Specifically, enemy military plans and operations may be revealed as follows:

(1) Unit movements and preparations for military activity may be indicated by rising and falling traffic volumes and changes in the structure of the network.

(2) The military function of a network may be revealed by the characteristic traffic pattern which results from transmissions incidental to planning, supply, or transportation.

(3) Change of grouping, disposition of forces and fleets, and probable tactical developments may be manifested in the redeployment of the radio stations which serve military elements.

d. These very important results are obtained without actually reading the texts of the intercepted messages; the solution and translation of messages are the functions of cryptanalysis and not traffic analysis. However, the cryptanalyst is frequently able to make good use of bits of information disclosed by traffic analysis such as faults noted in message routing and errors in cryptography causing messages to be duplicated or canceled. Cryptanalysis can provide important

---

[3] Such operators' communications are termed "chatter" or simply "chat."

information for traffic analysis, since the solution of messages often yields data on impending changes in signal communication plans, operating frequencies and schedules, etc. Cryptanalysis also yields data on specific channels, networks, or circuits which are most productive of intelligence, so that effective control and direction of intercept agencies for maximum results can be achieved.

5. Cryptanalysis. The most important steps of practical, operational cryptanalysis are listed below. These steps are more or less in the order in which they are followed, but in particular cases some of these steps may be interchanged, or omitted entirely.

a. The study of patent characteristics of message texts.

b. The study of any available collateral information, including that obtained from previous solutions.

c. The search for and study of indicators.

d. The determination of the type of cryptosystem used.

e. The separation of the traffic into groups of messages in the same or related keys.

f. The search for repetitions within and between messages.

g. The study of the beginnings and endings of messages.

h. The preparation of statistical counts of letters, groups, etc.

i. The reduction of the encrypted texts to simplest terms.

j. The test for probable words, stereotypes, isologs, etc.

k. The recovery of the plain texts.

6. Other Intelligence Sources. In addition to (1) traffic analysis and (2) cryptanalysis as means of obtaining information relating to communications, further data may be obtained (3) by the use of secret agents for espionage, (4) by the capture and interrogation of prisoners, (5) by the capture of headquarters or command posts with records more or less intact, and (6) by defection or carelessness on the part of personnel who handle communications. Of these six main sources, traffic analysis and cryptanalysis are the most valuable, due in great part to their reliability; they may be likened to "reading the innermost thoughts of the enemy" The amount of vital information they furnish cannot be accurately estimated as it fluctuates with time, place, circumstances, equipment, and personnel. For most effective operation, the results of both cryptanalysis and traffic analysis can be fitted together to yield a unified picture of the communications scheme. Therefore, if all transmitting

stations can be located quickly and if all communications can be intercepted and solved, extremely valuable information concerning strength, disposition of forces, and proposed moves will be continually available.

7. Time needed for cryptanalysis and its dependent factors.--

a. In military operations time is a vital element. The influence or effect that analysis of military cryptograms may have on the tactical situation depends on various time factors.

b. Of these factors, the following are the most important:

(1) The length of time necessary to transmit intercepted enemy cryptograms to solving headquarters. This factor is negligible only when signal communication agencies are properly and specifically organized to perform this function.

(2) The length of time required to organize raw materials, to make traffic analysis studies and to solve the cryptograms, and the time required to make copies, tabulate, and record data.

(3) The nature of information disclosed by traffic analysis studies and solved cryptograms; whether it is of immediate or operational importance in impending action, or whether it is of historical interest only in connection with past action.

(4) The length of time necessary to transmit information to the organization or bureau responsible for evaluating the information. Only after information has been evaluated and correlated with information from other sources does it become intelligence.[4]

(5) The length of time necessary to transmit the resulting intelligence (military, naval, air, etc.) to the agency or agencies responsible for tactical operations, and the length of time necessary for the agency to prepare orders for the action determined by the intelligence and to transmit such orders to the combat units concerned. The last sentence under (1) above applies here also.

c. Of the factors mentioned in b above, the only one of direct interest in this text is the length of time required to solve the cryptograms. This is subject to great variation, dependent upon other factors, of which the following are the most important:

(1) The degree of resistance of the system to cryptanalytic attack. This is dependent upon the technical soundness of the system itself, the technical soundness of the regulations and procedures governing the use of the system, and the extent to which cipher clerks follow these regulations and procedures.

[4] Often referred to as finished intelligence.

(2) The volume of cryptographic text available for study. As a rule, the greater the volume of text, the more easily and speedily it can be solved. A single cryptogram in a given system may present an almost hopeless task for the cryptanalyst, but if many cryptograms of the same system or in the same or closely related specific keys are available for study, the solution may be reached in a very short time.

(3) The number, skill, and efficiency of organization and cooperation of communication intelligence units assigned to the work. Cryptanalytic units range in size from a comparatively few persons in the forward echelons to many persons in the rear echelons. Such organization avoids duplication of effort and, especially in forward areas where spot intelligence is most useful, makes possible the quick interpretation of cryptograms in already solved systems. In all these units, proper organization of highly skilled workers is essential for efficient operation.

(4) The amount and character of collateral information and intelligence available to the cryptanalytic organization. Isolated cryptograms exchanged between a restricted, small group of correspondents about whom and whose business no information is available may resist the efforts of even a highly organized, skilled cryptanalytic organization indefinitely. If, however, a certain amount of such information is obtained, the situation may be entirely changed. In military operations usually a great deal of collateral information is available, from sources indicated in paragraph 6, above. As a rule, a fair amount of definite information concerning specific cryptograms is at hand, such as proper names of persons and places, and events in the immediate past or future.[5] Although the exchange of information between intelligence and cryptanalytic staffs is very important, the collection of information derived from an intensive study of already solved traffic is equally as important because it yields extremely valuable cryptanalytic intelligence which greatly facilitates the solution of new cryptograms from the same sources.

8. Cryptanalytic records and reports.--a. In practical cryptanalytic work the systematization of records and the maintenance of adequate files are of considerable importance. Likewise, the preparation of clear and concise reports, both technical and non-technical, is a major facet of practical cryptanalytic operations.

b. All messages coming into the cryptanalytic section are assigned a reference number, and a log is kept of these messages showing pertinent data such as the call signs, the date and time of interception, the group count, etc. Duplicate messages (i.e., different intercepts of the same transmission, or intercepts of retransmissions of the same message) are stapled together and garbles are corrected. Other records and files are maintained for special studies; for example, there may be card files on

---

[5] In this connection, see the remarks on cribs and probable words in subpars. 2d and 49c.

the message indicators[6] that have appeared in the traffic, card files of keys used in past and current systems, etc.

c. Cryptanalytic reports fall into two main categories: (a) technical reports intended for cryptanalytic personnel designed to give a summary of the cryptographic features of a system, with the steps that were taken to diagnose the system and effect a solution; and (b) nontechnical reports destined for intelligence consumers[7], which reports consist for the most part of message decrypts. In the latter category all decrypts might be furnished verbatim, or complete decrypts of important messages only, the rest of the messages being furnished in "gists" or in condensed form.

d. In technical reporting, clarity and detail are paramount.[8] A complete résumé of the diagnostic techniques employed in the identification of the system should be included, as well as a comprehensive outline of the steps taken to arrive at the initial solution.[9] It goes without saying that close attention should be paid to precise cryptologic terminology in all descriptions of methods and techniques, so as to lessen the chance of ambiguity or possible misunderstanding on the part of the reader. A cryptologic glossary should be freely consulted in all cases where there is an element of doubt in the mind of a writer as to the exact meaning of a term he is about to use.

e. In the next paragraph there is found an example of what may be considered as typical of a cryptanalytic technical report. Of course there is no fixed standard format for such a report, as the particular way in which a report is prepared, and the information included therein, depends upon the circumstances and situation at the time of the report. However, the hypothetical report in the next paragraph is intended as illustrative of the amount of detail that might be included in a report of this nature.

---

[6] In this connection, the location of groups of a message is designated by the terms A1, A2, A3...if reference is made to the first, second, third...positions from the beginning of the encrypted text, and by the terms ZØ, Z1, Z2...if reference is made to the last, penultimate, antepenultimate...positions from the end of the encrypted text.

[7] These reports are invariably highly classified, and their dissemination is strictly controlled on a special distribution list of those who must have a "need-to-know." This limitation is absolutely essential in order to protect the information, and prevent drying up the source and negating the work of the many weeks, months, or even years that are represented by the fruits of the communication intelligence effort. In this latter connection, when information derived through communication intelligence efforts is included in military intelligence reports, it is disguised in such a way as to protect the source of the information.

[8] For an excellent exposition on the art of technical writing, see Joseph N. Ulman, Jr., Technical Reporting, New York, 1952.

[9] See also the remarks made in subpar. 47f, on pp. 94-95.

9. Illustrative example of a technical report. The following represents a hypothetical technical report on the cryptanalysis of a newly-encountered system:

(CLASSIFICATION)
Special Distribution

Copy No. ____
of ____ copies

REPORT ON THE SOLUTION
OF THE "CALOX" SYSTEM

5 January 19__

I - BACKGROUND

1. On 12 December 19__, a new discriminant, CALOX, appeared in the enemy's traffic. The discriminant appears in the usual position, the A1 group of the message.

2. Traffic analysis indicates that CALOX traffic is being passed on air defense nets. From the characteristics of the transmission of this traffic and associated procedures, it appears that CALOX is an administrative system rather than an operational system. It also appears that CALOX does not replace an existing system, but rather is a new system introduced for some special purpose. On the enemy's air defense nets, both cipher and code systems have been encountered.

3. CALOX traffic was segregated and logged in as received, together with the worksheet reference numbers assigned to all incoming traffic by the Traffic Handling Section.

II - PRELIMINARY ANALYSIS

4. The first step in treating the CALOX system was to complete the plain-component sequence on one of the messages, on the hypothesis of direct and reversed standard alphabets, using a strip board for this purpose. (The enemy has used standard alphabets in the past in one system, changing the juxtaposition of the components after the encryption of every few letters.) This disclosed nothing of significance.

5. Uniliteral frequency distributions made for each of the six messages intercepted on 12 December were flat; the average I.C. of 1.1 indicates that it is most unlikely that the underlying cryptosystem is a monoalphabetic system involving single-letter cipher units. However, a rather odd manifestation in the distributions for each message was that C_c, D_c, H_c, L_c, and V_c were usually consistently predominating, while S_c, Y_c, and Z_c were consistently of very low frequency. No explanation for this phenomenon was forthcoming at the time.

6. A check was made on previously solved enemy systems used on his air defense and other nets, to disclose any similarity between the CALOX characteristics and those of another system; this proved fruitless, as the uniliteral frequency manifestations of CALOX were unique to that system. A check was also made to find any possible isologs between CALOX messages and those of another readable system; however, this too proved fruitless, as did the examination of chatter associated with the CALOX messages in an attempt to reveal any clues as to the system or to uncover possible cryptographic service messages, etc.

7. Digraphic distributions were made of the messages of 12 December, but no unusual phenomena were visible. The ϕ^2 approached that of ϕ_r^2, and there was no evidence to support any matching of the rows or columns of the distributions if the hypothesis of a variant system with a small matrix were assumed.

8. Triliteral frequency distributions were made of each message to disclose repetitions; these repetitions were underlined in the messages, and a comparison was made of those repetitions occurring between messages of the same day. Many short repetitions of 3, 4, and 5 letters were disclosed, the number of these repetitions being considerably above that expected for random; however, no longer repetitions were uncovered, and the intervals between the repetitions had no common factors.

9. Every day's accumulation of traffic was examined statistically with a view to revealing possible key changes, and the phenomena in par. 5, above, continued. When on 19 December the predominant peaks and troughs no longer corresponded to the norms observed in par. 5, a change of keys was assumed.

10. A typical message in the first key period is given below:

LRZ DE CKS (Intercepted 17 December, on a frequency of 5600 Kcs)

CALOX JOLDJ JLAPP DRELF QXEDZ QIHFN WMGUH DMAYM IMNDY OMZCC

OMMYE HQDAH YEMNB VUGHD IMXOG LDHUX MACJV VRNEK LCHEJ DZCDO

PRILM UGBOC DLXJL UBVVW TRAFX LKNPA HSJNE HVCAC OQTHU FJVTH

DIKQW MCGIW HRMAF LKGBE FNPOG JROGM WGUDM XJIJL BWEDK QCUMR

TJXAN BLTUR KMTOR CFIHV QCEKH HUJNQ ATBWZ VNERI LHFOQ MLUMX

LAXVY HEQBX RIKRK YACSV LPOQP NOBKU XGLED FNPAG JRRAB JLEBW

DKIQC MRADN VNURB TOPBH LIKLH EPVTR BGYMA MYQWI PVLEM GLEGH

ODMXT DHONG XNXEL DWXGA LDIGB GCILM ZQLAC LXODQ

III - THE SOLUTION

11. The following peculiar sets of similar sequences of cipher letters were noted during the examination of the 32 messages available in the first key period. The message reference numbers are given, together with the position in the message of the first letter of the sequence. (The position given is the text position, excluding the discriminant.)

	Msg No.	Pos.	
a)	60208	057	H F I J V T H O D K Q W M E C G A W H R
	61492	216	H F J V O T H D K Q A W M E C G W H A R
	60317	139	H U F J V T H D I K Q W M C G I W H R

b)	60317	123	N P A H S J N E H V C A C Q T
	62350	098	N P H I S J N O H V C C U Q T
c)	60317	184	J L B W E D K Q C U M R
	60317	291	J L E B W D K I Q C M R
d)	60295	114	T P I Q K Z H E H V P U V P B
	61007	253	T P O Q K Z H A H V P E V P B
e)	60943	147	H V G G A K W Q S O V R N
	62156	064	H V U G G K E W Q S V I R N

The behavior of the letters comprising these sequences indicates that A_c, E_c, I_c, O_c, and U_c, most likely are nulls. On this hypothesis, evidence from the lengths of the repetitions now disclosed, and the intervals between repetitions, indicates a digraphic grouping of the cipher text. On checking back to the digraphic distributions, it is noted that there are no vowel-vowel contacts in the cipher text, except for combinations with Y_c. Furthermore, in retrospect it is seen that most of the cipher groups contain 1 or 2 vowels, never more; this significance escaped notice until the near-repetitions above were observed.

12. New digraphic distributions, omitting the 5 vowels, were made for the messages in the first key period. No matching qualities were manifested in the new distributions; but this time the $\phi 2$ very closely approximated the ϕ_p^2, thus it appeared that, in spite of the limitation of only 21 ciphertext letters remaining after the null vowels were discarded, the system was basically a digraphic system. (This would not exclude, however, a matrix containing a few frequent trigraphs or tetragraphs, etc.) Work sheets were now made for several of the best messages from the first key period, the messages selected being long ones that existed in more than one intercept copy so that garbles might be corrected.

13. On 28 December the first message was solved; this was Message #60317 which was one of the longest, and which was copied by three different intercept operators. One more cryptographic idiosyncrasy of the CALOX system was now brought to light: that of the peculiarity of behavior of Y_c which had been previously overlooked. This peculiarity was that Y_c was always present in pairs, fairly close together; every Y_c was followed by another Y_c, with from 2 to 10 letters intervening. This Y_c turned out to be a number indicator, and the cipher digraphs between the indicators represented single digits.

14. From the original solution, an equivalent digraphic matrix was reconstructed with the consonant coordinates in normal alphabetical order, as shown below:

2d Letter

	B	C	D	F	G	H	J	K	L	M	N	P	Q	R	S	T	V	W	X	Y	
B					BY	CC		AY	CH								CE	CK	2		
C		TE	TI	TO											TH						
D						F6	FE	EN		ER	FW		EL	EY							ES
F											GE							FF	FR		
G	MA	LL				MB			ME	LR								MM			
H				AI			AT						1	AN	AD		AR		AC		
J									RE		QU			RA			RD				
K						VE				UL											
L		OL	OP	OU		PA		OM		OR									ON		
M		CO		DC						D4	DE			DG							DA
N			8		HI			GT				HE	GR								
P					NE						ND	NC	MO	NF			NI				
Q		RI				SA				RP						RS		SE	RM		
R	LB							Ø	LE												
S																					
T	EE					EF	EI					EA		ED							
V									IS		IL			IK				IV			
W					OB					NU											NV
X					ST		TB		TA		SS					SO	T				
Y																					
Z													YC								

Noting evidences of symmetry in the matrix, the matrix coordinates were rearranged to yield the primary matrix which is shown below, including values which were interpolated on the basis of likelihood and alphabetical sequence.

2d Letter

	Q	C	K	X	S	D	M	Z	T	F	P	N	G	R	B	H	V	L	J	W
H	A1	AA	AB	AC	AD	AE	AF	AG	AH	AI	AK	AL	AM	AN	AO	AP	AR	AS	AT	AU
B	AV	AW	AY	B2	BA	BE	BI	BL	BO	BR	BS	BU	BY	C3	CA	CC	CE	CH	CI	CK
M	CL	CO	CR	CT	CU	CY	D4	DA	DB	DC	DD	DE	DF	DG	DH	DI	DL	DM	DN	DO
T	DP	DQ	DR	DS	DT	DU	DV	DW	DY	E5	EA	EB	EC	ED	EE	EF	EG	EH	EI	EK
D	EL	EM	EN	EO	EP	EQ	ER	ES	ET	EU	EV	EW	EX	EY	EZ	F6	FA		FE	FF
F	FI	FL	FO	FR	FS	FT	FU		FY	G7	GA	GE	GF	GH	GI	GL	GM	GN	GO	GP
N	GR	GS	GT	GU	GY	H8	HA	HB	HC	HD	HE		HI							
V												IL	IM	IN	IO	IP	IR	IS	IT	IV
R	IW	IX	JØ	JA	JE	JO	JU	K	KA	KE	KI	KO	L	LA	LB	LC	LD	LE	LF	LH
G	LI	LL	LM	LN	LO	LP	LR	LS	LT	LU	LV	LW	LY	M	MA	MB		ME	MI	MM
P	MO	MP	MR	MS	MT	MU	MY	N	NA	NB	NC	ND	NE	NF	NG	NH	NI	NL	NM	NN
W	NO	NP	NQ	NR	NS	NT	NU	NV	NW	NY	O	OA	OB	OC	OD	OE	OF	OG	OH	OI
L	OK	OL	OM	ON	OO	OP	OR	OS	OT	OU	OV	OW	OX	OY	P	PA				
J											Q	QU	R	RA	RB	RC	RD	RE	RF	RG
Q	RH	RI	RL	RM	RN	RO	RP	RQ	RR	RS	RT	RU		RY	S	SA	SB	SC	SD	SE
X	SF	SG	SH	SI	SL	SM	SN	SO	SP	SQ	SR	SS	ST	SU	SW	SY	T	TA	TB	TC
C	TD	TE	TF	TG	TH	TI	TL	TM	TN	TO										
K							UL									VE				
S																				
Z	YC																			

By comparison with other messages in the same period, and with messages in subsequent periods, it was possible to recover the values inside the matrix in their entirety, as follows:

A	AA	AB	AC	AD	AE	AF	AG	AH	AI	AK	AL	AM	AN	AO	AP	AR	AS	AT	AU
AV	AW	AY	B	BA	BE	BI	BL	BO	BR	BT	BU	BY	C	CA	CC	CE	CH	CI	CK
CL	CO	CR	CT	CU	CY	D	DA	DB	DC	DD	DE	DF	DG	DH	DI	DL	DM	DN	DO
DP	DQ	DR	DS	DT	DU	DV	DW	DY	E	EA	EB	EC	ED	EE	EF	EG	EH	EI	EJ
EL	EM	EN	EO	EP	EQ	ER	ES	ET	EU	EV	EW	EX	EY	EZ	F	FA	FC	FE	FF
FI	FL	FO	FR	FS	FT	FU	FY	G	GA	GC	GE	GF	GG	GH	GI	GL	GN	GO	GP
GR	GS	GT	GU	GW	H	HA	HB	HC	HD	HE	HF	HI	HL	HM	HN	HO	HR	HS	HT
HU	HY	I	IA	IB	IC	ID	IE	IF	IG	IK	IL	IM	IN	IO	IP	IR	IS	IT	IV
IX	IZ	J	JA	JE	JO	JU	K	KA	KE	KI	KS	L	LA	LB	LC	LD	LE	LF	LG
LI	LL	LM	LN	LO	LP	LR	LS	LT	LU	LV	LW	LY	M	MA	MB	MC	ME	MI	MM
MO	MP	MR	MS	MT	MU	MY	N	NA	NB	NC	ND	NE	NF	NG	NH	NI	NK	NL	NM
NN	NO	NP	NR	NS	NT	NU	NV	NW	NY	O	OA	OB	OC	OD	OE	OF	OG	OH	OI
OK	OL	OM	ON	OO	OP	OR	OS	OT	OU	OV	OW	OX	OY	P	PA	PE	PF	PH	PI
PL	PM	PN	PO	PP	PR	PS	PT	PU	PY	Q	QU	R	RA	RB	RC	RD	RE	RF	RG
RH	RI	RL	RM	RN	RO	RP	RR	RS	RT	RU	RV	RW	RY	S	SA	SB	SC	SD	SE
SF	SG	SH	SI	SK	SL	SM	SN	SO	SP	SR	SS	ST	SU	SW	SY	T	TA	TB	TC
TD	TE	TF	TG	TH	TI	TL	TM	TN	TO	TP	TR	TS	TT	TU	TW	TY	TZ	U	UA
UB	UC	UD	UE	UG	UI	UL	UM	UN	UP	UR	US	UT	V	VA	VE	VI	VO	W	WA
WE	WH	WI	WL	WN	WO	WR	WY	X	XA	XC	XE	XF	XI	XN	XP	XT	Y	YA	YB
YC	YD	YE	YF	YG	YH	YI	YL	YM	YN	YO	YP	YR	YS	YT	YW	Z	ZA	ZE	ZI

It will be noted that the matrix contains the 26 letters, and 374 of the highest frequency digraphs. When encrypting numbers, the cipher value for 1 is the cipher equivalent of A_p, the cipher value for 2 is the $\overline{\theta\theta}_c$ for B_p, etc., to $\emptyset_p = \overline{\theta\theta}_c$ (J_p).

15. In the matrix coordinates for the first key period, the non-random phenomena in the grouping of the coordinate letters was noticed, suggesting that some systematic method for producing these sequences was used. It evolved that these sequences were derived by simple columnar transposition using the following rectangles:

For the rows:

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

For the columns:

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

Thus the key words for the first period are HYDRAULIC and QUESTIONABLY (with, of course, the vowels omitted) for the row and column coordinates, respectively.

IV - CONTINUITY OF KEY CHANGES; SUMMARY

16. Having solved the CALOX system for the first period (12-18 Dec), the second period (19-26 Dec) was easily solved by the discovery of a pair of cross-key isologs on 19 December; the third period (27-31 Dec) was speedily solved by means of a signature crib; while the fourth period (beginning on 1 Jan) had to be solved by the general method of digraphic frequencies and digraphic idiomorphs. The row and column key words for the second period were COPYRIGHTED and DOCUMENT; for the third period, CHIMPANZEE and MANDRILL; but for the fourth period the same key word, MNTVD (Montevideo?), was used for both the row and column coordinates. The coordinate sequences were derived by simple columnar transposition, as in the first period.

~~CONFIDENTIAL~~

17. If the enemy has found that two different sequences for the row and column coordinates is too inconvenient cryptographically and therefore continues to use the single key word procedure started in the fourth period, a statistical technique has been devised for establishing the identity of some (or even all) of the letters of the coordinates, based on a consideration of the relative frequencies of the ciphertext letters. This technique is founded on the fact that in a single key word procedure the combination of row 19 and column 19 of the basic matrix will yield a low frequency cipher letter, as will the combinations of row 20-column 20, and row 9-column 9; on the other hand, the combinations row 17-column 17, row 5-column 5, row 13-column 13, and row 14-column 14 will yield high frequency cipher letters. With a single key word procedure being used, the following is the expected descending frequency order of the twenty row-column combinations:

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

Even if two key words are employed for the coordinates, a modification of the statistical method is feasible, in those instances where any difficulty might be encountered in a new key period. The statistical techniques and the methods of their employment will be described in a later report.

18. No trouble is anticipated in keeping current with key changes in the CALOX system; traffic should be readable now on the first day of a key change. If the enemy used another set of 5 letters as nulls, instead of the vowels, the new nulls can be identified by searching for and examining near-repetitions, as shown in par. 11. A similar procedure would be used to identify a new number indicator, even though solution would not be impeded by this latter factor.

19. The traffic analysis report on the CALOX traffic gives complete statistics on the links on which CALOX is found, as well as a detailed summary on the number of messages intercepted, etc.

~~CONFIDENTIAL~~

(BLANK)

M. Friedman

~~CONFIDENTIAL~~

NATIONAL SECURITY AGENCY
Washington 25, D. C.

COURSE — Military Cryptanalysis, Part I

LESSON 8 — Monoalphabetic substitution with irregular-length cipher units: monome-dinome systems and others

TEXT ASSIGNMENT — Section X

1. Solve the following monome-dinome cryptogram and recover the original matrix:

78131 76784 31174 50078 76343 47807
41346 53334 01331 01799 78318 76441
31917 92478 74179 10834 76033 55723
40178 31347 46554 65323 41305 86131
34767 30345 77787 48763 77689 76072
76747 88123 11278 31788 76503 47753
17807 67921 07276 07310 17997 88878
74703 05323 15777 71034 76371 33764
47117 37607 88390 00666 33300 03985
79531 31533 78342 47800 17230 75560
34850 74547 83189

2. The following monome-dinome cryptogram is believed to contain the probable word "DIVISION". Solve the text and recover the original matrix:

17832	00066	16927	80635	28420	04596
95220	01900	21500	40563	26746	12576
80705	88123	53921	31118	13281	29159
46465	61576	52844	90033	94526	59400
25284	30032	00457	80758	80707	00526
73941	20854	56640	59352	91625	97612
46977	89125	05945	22008	41401	51129
31702	91067	53763	59062	38071	67003
84670	04267	78579	20084	17919	60266
43595	65697	00036	12004	97616	87202
60045	70787	05971	26122	81200	19003
00841	76912	09599	72673		

3. The following cryptogram was intercepted on a link which has been known to be passing traffic in two different monome-dinome systems, one involving a matrix of the type shown in Fig. 75 of the text, the other involving a matrix of the type in Fig. 77. Solve the text of the message and recover the original matrix.

47631	82870	14628	31274	12741	16263
16054	63152	84662	60736	97728	46198
46972	13808	46287	46364	83788	72846
60846	28738	27578	87073	18279	62736
97462	83107	36977	45636	26962	73168
62763	12138	08462	87316	06379	82647
28467					

4. The following messages, intercepted on a link known to be passing monome-dinome traffic, are believed to be isologs. Solve the texts and recover the original matrices.

Message "A"

94872	33935	61227	89316	23405	09079
43810	57678	93386	41999	83809	08334
94194	76279	99496	30576	79199	54343
57683	04186	07981	43349	83529	09638

Message "B"

94378	11935	62887	39326	81405	09079
41320	57673	93136	41999	81309	03114
94194	76879	99496	10576	79199	54343
57631	04136	07982	43149	31589	09613

5. The following messages are believed to be isologous monome-dinome ciphers. Solve the texts and recover the original matrices:

Message "A"

73507	09885	01652	37531	09804	39858
14983	12316	52371	12890	93312	42689
30741	59012	54398	50563	98460	77297
30415	65075	43098	13500	74379	06814
51983	12316	52371	13559	33124	39842
16361	80772	97056	29092	58145	15465
07901	10121	98617	56398	94163	84731
35039	04398				

Message "B"

36713	45807	18921	63867	55406	58179
56296	89216	37798	07485	62909	18085
43072	74292	56571	84650	14339	73640
72171	32564	58871	43063	74180	79875
62968	92163	77676	85629	06509	89612
34339	73484	97424	81798	72517	13747
74292	78017	08465	26896	80036	88716
74065					

4. The following messages, intercepted on a link known to be passing monome-dinome traffic, are believed to be isologs. Solve the texts and recover the original matrices.

Message "A"

94872	33935	61227	89316	23405	09079
43810	57678	93386	41999	83809	08334
94194	76279	99496	30576	79199	54343
57683	04186	07981	43349	83529	09638

Message "B"

94378	11935	62887	39326	81405	09079
41320	57673	93136	41999	81309	03114
94194	76879	99496	10576	79199	54343
57631	04136	07982	43149	31589	09613

5. The following messages are believed to be isologous monome-dinome ciphers. Solve the texts and recover the original matrices:

Message "A"

73507	09885	01652	37531	09804	39858
14983	12316	52371	12890	93312	42689
30741	59012	54398	50563	98460	77297
30415	65075	43098	13500	74379	06814
51983	12316	52371	13559	33124	39842
16361	80772	97056	29092	58145	15465
07901	10121	98617	56398	94163	84731
35039	04398				

Message "B"

36713	45807	18921	63867	55406	58179
56296	89216	37798	07485	62909	18085
43072	74292	56571	84650	14339	73640
72171	32564	58871	43063	74180	79875
62968	92163	77676	85629	06509	89612
34339	73484	97424	81798	72517	13747
74292	78017	08465	26896	80036	88716
74065					

CONFIDENTIAL

6. Solve the following monome-dinome-trinome cryptogram and recover the original matrix:

61745 04120 43950 43238 65332 06382
01503 20682 61661 20436 53513 17150
68412 19203 16204 38543 12043 20150
35350 12335 45039 44171 20186 50929
78509 23850 46204 84739 45049 62065
82820 43532 01561 93231 65184 71533
53842 04541 62453 32043 85421 68564

7. Solve the following uniliteral-biliteral cryptogram, and recover all keys:

PVOYA CKRTE AUOOD KNWOI BKEIA UBTAP
WOIDG OBKNT AENXB TAEBG YAEUI ENLCT
EOBZF HOOBL YIEBG UUONT BXPXR MIBKA
CWOIE PKCGP VAYEF TEINM PKSGE YAODK
UEDLR ZEYAN GCWUY AUPKP MEOIA CVPWY
RWOYC WAPWO IYAOR WSVCH EIRVC KYYPK
OICKY NWODH RKDGE AEBXU ERXDM EYABT
EUCWN GRTDW PHOAO PGUNG RKCVY ONZBG
UENTX

8. The following cryptogram, enciphered in a Playfair-type digraphic-monographic system, is suspected to begin with the probable stereotype "MORNING REPORT FOR MONDAY NOVEMBER TWENTY FIRST." Solve the text and recover all keys.

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

9. The following cryptogram was enciphered in a dinome-trinome digraphic system employing matrices similar to those in Figs. 90a and b, except that the internal numerical sequences have been changed. The message is suspected to end with the signature VINCENT ANDERSEN COL INF. Solve the text and reconstruct the matrices involved.

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

~~CONFIDENTIAL~~

10. The following cryptogram, based on a Morse code system, is suspected to begin with a spelled-out number. Solve it and recover all keys:

71430	62809	18592	35607	61572	04953
79012	87548	65983	04037	95327	30751
34904	56564	20813	01258	16408	97156
64597	60410	83159	34702	68032	95357
25173	02589	41582	60360	91754	

(BLANK)

RESTRICTED

2. The following three messages were selected from a volume of Converter M-209 traffic because they were enciphered with the same message indicators, a serious violation of cryptographic security. It is furthermore suspected that stereotypic beginnings may have been used in the drafting of the three messages. Solve the messages, and submit the plain text to Message "A".

Message "A"

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

Message "B"

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

Message "C"

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

RESTRICTED

~~CONFIDENTIAL~~

~~SECURITY INFORMATION~~

Plain	16				20					25					30					35					40	41
A	Y	D	S	Q	P	K	B	L	A	J	N	C	G	M	W	E	Z	X	G	U	R	O	I	H	F	V
B	I	Q	L	J	S	G	D	A	U	*	R	H	K	P	E	C	B	W	N	X	M	V	Z	O	S	Y
C	R	V	B	L	T	O	M	V	J	G	D	X	W	U	K	N	S	H	F	E	Z	Q	A	P	Y	C
D	Z	J	R	M	K	T	H	E	B	V	U	S	I	L	Q	F	D	C	X	O	Y	N	W	A	P	T
E	A	E	T	X	D	N	V	Q	O	X	L	I	F	Z	Y	W	M	P	U	J	H	G	B	S	C	R
F	F	W	G	V	E	I	X	B	H	R	Z	U	S	B	P	M	J	D	C	A	Q	T	Y	N	L	K
G	L	G	X	H	W	F	J	Y	C	I	S	A	V	T	C	Q	N	*	E	D	B	R	U	Z	O	M
H	S	B	F	U	Y	E	O	W	R	P	Y	M	J	G	A	Z	X	N	Q	V	K	I	H	C	T	D
I	K	I	Y	B	G	V	T	S	N	E	O	D	M	Q	F	J	P	Z	H	C	A	J	X	U	R	L
J	M	L	J	Z	C	H	W	U	T	O	F	P	E	N	R	G	K	Q	A	I	D	B	K	Y	V	S
K	X	U	O	N	L	B	E	J	Y	W	V	Q	H	R	G	P	T	I	M	S	C	K	F	D	M	A
L	G	P	D	A	X	R	Q	O	E	H	M	B	Z	Y	T	K	U	J	S	W	L	P	V	F	N	I
M	B	Y	V	P	O	M	C	F	K	Z	X	W	R	I	S	H	Q	U	J	N	T	D	L	G	E	N
N	H	X	A	F	U	S	R	M	D	N	C	L	P	E	I	O	Y	G	B	Z	I	W	T	Q	K	J
O	N	M	H	Y	I	X	G	K	Z	D	J	T	B	W	U	D	R	O	L	F	E	C	S	V	A	P
P	V	F	U	D	H	W	A	G	Q	Y	T	R	A	O	L	I	C	*	Z	P	S	X	M	K	J	E
Q	T	N	M	K	A	D	I	X	V	U	P	G	Q	F	O	S	H	L	R	B	F	E	C	L	Z	W
R	U	A	K	S	N	L	U	I	F	C	W	V	T	J	M	R	G	E	D	Y	P	Z	O	X	B	Q
S	J	H	Q	E	B	Y	S	R	P	F	I	N	C	A	Z	U	L	V	K	T	X	M	Q	W	G	O
T	W	Z	E	T	R	Q	L	C	M	B	K	O	D	H	N	X	F	A	Y	H	V	S	P	J	I	G
U	D	S	W	C	M	U	P	N	W	K	H	E	Y	X	V	L	O	T	I	G	F	A	R	B	Q	Z
V	E	T	C	G	V	Z	F	P	X	S	Q	Z	N	K	H	B	A	Y	O	R	W	L	J	I	D	U
W	Q	O	N	I	Z	J	Y	H	L	*	E	K	U	C	X	V	E	S	P	M	G	F	D	T	W	B
X	P	K	I	R	F	C	Z	T	S	Q	G	J	O	D	B	A	V	M	W	L	U	Y	N	R	X	H
Y	C	R	P	O	J	A	K	Z	I	M	B	F	L	V	D	Y	W	F	T	Q	N	H	G	E	U	X
Z	P	K	I	R	F	C	Z	T	S	Q	G	J	O	D	B	A	V	M	W	L	U	Y	N	R	X	H
-	O	C	Z	W	Q	P	N	D	G	L	A	Y	X	S	J	T	I	R	V	K	O	U	E	M	H	F

__Fig 1__. Enciphering alphabets 16-41. __Purple__ values consitute the partial alphabets obtained through depth reading; __red__ values are derived by examination of the basic cipher-text sequences (using the entire set of 100 partial alphabets).

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

__Fig 2__. The Friedman square for the fast rotor, obtained by rearranging the rows of Fig 1 (using the isomorphic patterns of the rows) so as to yield identical sequences for all the diagonals.

~~CONFIDENTIAL~~

~~CONFIDENTIAL~~

~~SECURITY INFORMATION~~

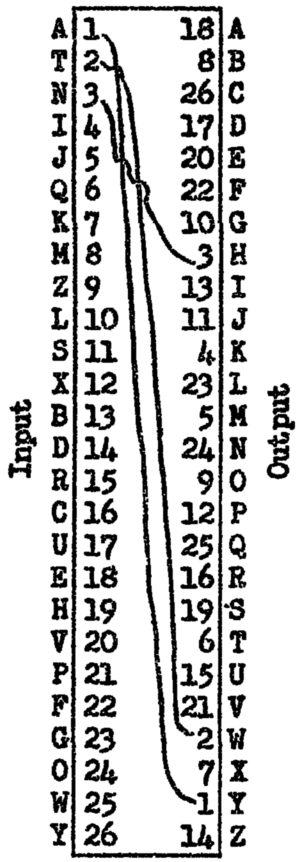


Fig 5. Wiring recovery of fast rotor.

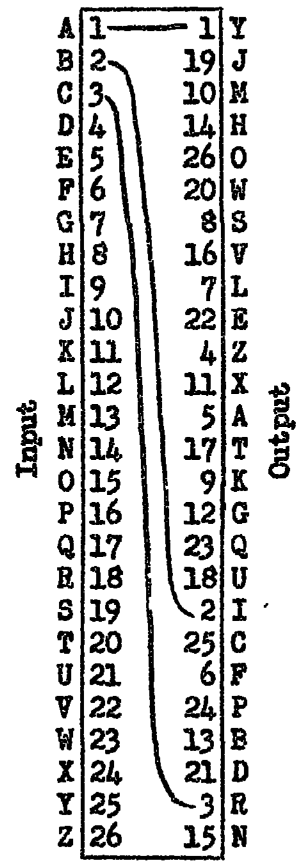


Fig 6. Wiring recovery of medium rotor.

~~CONFIDENTIAL~~

Plain	16	42	68	94
A	Y	W	P	A
B	I	J	S	I
C	R	C	M	E
D	Z	N	F	T
E	A	X	Y	L
F	F	T	A	W
G	L	P	K	S
H	S	E	B	M
I	K	V	Z	O
J	M	G	L	U
K	X	H	Q	N
L	G	A	O	C
M	B	Q	T	G
N	H	D	U	X
O	N	O	R	R
P	V	Y	W	D
Q	T	L	J	H
R	U	K	E	P
S	J	I	G	V
T	W	M	C	J
U	D	S	H	Y
V	E	R	V	K
W	Q	Z	N	F
X	P	U	X	Q
Y	C	B	I	B
Z	O	F	D	Z

Fig 3. Enciphering alphabets 16, 42, 68, and 94. These alphabets, spaced at intervals of 26 along the keying cycle, resulted from four successive positions of the medium rotor at points along the cycle where the fast rotor had returned to its initial position. Note that a diagonal chain may be completed, using alphabets 16, 42, and 68.

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

Fig 4 A fragment of the Friedman square for the medium rotor, with one of the repeating diagonals extended.

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

Fig 7. Reconstruction of machine with "mixed separator".

Direction of motion

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

Fig 8. Reconstruction of complete machine; rotors in position for alphabet 94.

~~RESTRICTED~~

~~SECURITY INFORMATION~~

ENCIPHERING ALPHABETS, "ZEBAB" PROBLEM

					5					10					15					20					25					30					35					40					45					50	
A	V	T	J	M	R	G	E	D	Y	P	Z				Q	Y	D	S	Q	P	K	B	L	A		N	C		M	W	E	Z	X	G	U	R	O		H	F	V	W	Z	E	T	R		L	C	M	A
B																		L	J								H	K	P	E					X		V		O	S								S	R		B
C	Z		K	H	B		Y	O	R	W			I	D		R			L		O	M		J	G		X	W	U		N	S	H	F	E	Z			P	Y	C	C	R	P		J			Z	I	C
D	Q	H		G	P	T		M	S	C	K	F				Z	J			K	T		E	B					L	Q	F		C	X	O	Y	N	W		P		N	M	H	Y	I	X	G	K	Z	D
E	F	L	V	D	Y	W	F	T	Q	N	H	G	E	U	X	A	E	T	X	D	N	V	Q	O	X	L		F	Z	Y	W	M	P	U	J	H	G	B	S	C	R	X	U	O	N	L	B	E	J	Y	E
F	R	A	O	L					P	S	X			J	E	F	W	G	V	E			B		R			S		P	M	J		C			T	Y	N		K				K		D	I			F
G	O		H		X	F			H			P	J					X	H		F	J	Y	C			A	V	T	C						B	R	U		O	M				R			Z			G
H	G	Q			S	H		R	B		E	C	L	Z	W			F			E	O	W				M		G			X		Q	V		I						T			V	Z	F		X	H
I	N		A	Z	U	L	V		T	X	M	Q	W	G	O	K	I	Y	B	G	V	T	S	N	E	O	D	M	Q	F	J	P	Z	H	C	A		X	U	R	L	V		U	D	H	W	A	G	Q	I
J																														R																					J
Plain K																						E															K														K
L		O	D	B		V	M							X	H							Q	O	E	H	M				T	K		J	S	W	L	P			N	I	A		T		D	N			O	L
M	P		N		G	K	Q			D				V		B	Y	V	P	O		C	F	K	Z	X	W	R	I	S	H		U			T	D	L	G		N	Q			I			Y	H		M
N		W	U	K	N		H	F	E	Z	Q	A	P	Y	C	H	X	A	F	U	S	R	M	D	N	C	L	P	E	I	O	Y	G	B	Z	I	W		Q	K	J	D	S	W	C	M	U	P	N	W	N
O		J	G	A	Z	X			V	K	I	H	C	T	D	N	M	H	Y	I		G	K	Z	D	J	T	B	W	U	D	R	O	L	F	E	C		V	A	P	O	C	Z	W	Q	P	N	D	G	O
P	I	F	Z	Y	W	M		U	J		G	B	S	C			F		D	H	W			Q	Y	T	R			L	I					S			K					S	Q		K		L		P
Q			C		V		S												K					V			G				S																				Q
R	C	G	M	W	E	Z	X	G	U	R	O		H	F	V	U	A	K	S	N	L	U	I	F	C	W	V	T	J	M	R	G	E	D	Y	P	Z	O	X	B	Q	K	I	Y	B		V	T	S	N	R
S	Y		S	J		I	R		K	O		E	M	H	F	J	H	Q	E	B	Y	S	R	P	F	I	N	C	A	Z	U	L	V		T	X	M	Q	W	G	O	I	Q	L	J	S	G		A	U	S
T	W	R		S		Q	U	J	N	T	D	L	G	E	N	W	Z	E	T	R	Q	L	C	M		K	O	D	H		X	F	A	Y	H		S	P	J	I	G	M	L	J	Z	C	H	W	U	T	T
U	A	V		C		N		E	D	B			Z		M	D		W	C	M	U	P	N	W	K	H	E	Y	X	V		O		I	G	F		R	B	Q	Z	S	B	F		Y	E			R	U
V			Y		K			S			P		F			E						F											Y						I				V						V		V
W	U					J								L																			S	P				D						R			T			B	W
X		K																																	L														I		X
Y								B			W		Q	K						J										D								G		U			Y			O			F		Y
Z																																		W																	Z
-			W	U	D		O	L	F	E	C	S	V	A	P	O	C	Z	W	Q	P	N	D	G	L	A	Y	X	S	J	T	I	R	V	K	O	U	E	M	H	F	F	W	G	V	E	I	X	B	H	-

~~RESTRICTED~~

NATIONAL SECURITY AGENCY
Washington 25, D. C.

"TAN" PROBLEM

(A problem in plaintext recovery from a shallow depth in a Baudot system)

Training Division
9 November 1953

International Teleprinter Code

UPPER CASE	WEATHER SYMBOLS	↑	⊕	○	↗	3	→	↘	↓	8	↙	←	↖	.	⊕	9	Ø	1	4	Δ	5	7	⊕	2	/	6	+	-	<	≡			
	COMMUNICATIONS	-	?	:	$	3	!	&	£	8	'	(	)	.	,	9	Ø	1	4	Δ	5	7	;	2	/	6	"	≀≀	<	≡			
LOWER CASE		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	BLANK	C.R.	L.F.	SPACE	LTR. SHIFT	FIG. SHIFT
	1	O	O		O	O	O				O	O						O		O		O		O	O	O	O					O	O
	2	●		O				O		O	O	O	O				O	O	O			O	O	O						O		O	O
	FEED HOLES	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	3			O			O		O	O		O		O	O		O	O		O		O	O		O	O					O	O	
	4		O	O	O		O	O			O	O		O	O	O			O				O		O				O			O	O
	5		O					O	O				O	O		O	O	O			O		O	O	O	O	O					O	O

Vernam Encipherment Table

θ^2

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

"TAN" PROBLEM

The following messages represent three radioprinter intercepts obtained on the same day. From analysis of previous traffic, it is known that the first four letters of each message constitute the indicator; a single space is used between all words, and a message center idiosyncrasy is the use of the doubled CR CR LF LF for the carriage return-line feed operation.

In the representation of the six special functions, the following notation has been used:

Figures	2
Space	3
Letters	4
Line feed	5
Blank	6
Carriage return	7

It is suspected that one of the messages intercepted in the morning contains the probable word REPORT. Solve the texts of the morning messages; and, after correctly juxtaposing the third message along the keying cycle, recover the plain text to the third message.

Message "A"

RDA DE IZX 0930Z 17 DEC

IVIVP LWPZX ENO2C UBHVZ EEKWF MFN4F 26SDK XEB5F VD6F3 CTUMV
-RE PORT-

YEYUV QXI2H HCW5C KSKO2 YIMHC FHBWC C423G KXVPH I4AVS SUUO6

Message "B"

RDA DE IZX 0945Z 17 DEC

IVIVU PEHKV FVCIU NIMAP E4WP7 RIVEF UMGEN UZIWN GI2TY IZVEZ
PHE R-YOU

UD6US 67E5Z 7YVLW XOOLZ GM42E DNNRC SQ6OB 4IGGM NQ6MZ S25TQ

4YXG7 IIISII

Message "C"

IZX DE RDA 1425Z 17 DEC

BMBMY XYY36 FFICJ EBU45 NZY7G EXLVZ BR7WE SARU5 STZ4G GHKSG

VSU2S 6NN4C L4EVP HUTB5

(BLANK)