

Numbering of anagrams and cryptography games

Abstract

Three versions of numbering of n -block of plain text are proposed and a method how to restore the plane text from the given serial number in each case. These ciphering methods admit randomization of serial numbers reducing communication to the exchange of random numbers. Such procedure as is one-time pad approach does not include repetitions of serial numbers and looks satisfying Shannon's criterion of an unbreakable code. The detail description is given of two mutually reciprocal algorithms computing the n -block serial number from the plain text (encoding) and restoring the plain text from the serial number with the removed random element (decoding). The recipient need only know how the random number is implemented in a block that serves as a secret key. Specific features of the algorithms within this general scheme: the "dictionary method" identifies a false cypher message in 99.95% of cases without decoding. The "anagram" method separates procedures for the letter contents of n -block and for the letter order. The simplest "division" method is the Euclidian algorithm adjusted for coding.

Keywords: anagram, positioning of integer among binomial coefficients, one-time pad, random number

Volume 9 Issue 1 - 2025

M Mestechkin

Independent Scholar, USA

Correspondence: Mikhail Mestechkin, Independent Scholar, 12773 Seabreeze Farms Dr. # 33, San Diego CA 92130, USA, Tel 1-858-847-9029

Received: April 18, 2025 | **Published:** July 14, 2025

Introduction

Preamble

This article is addressed to beginners in mathematical physics, natural sciences, and engineering. We try to demonstrate that their knowledge already at the first stages of education can be successfully applied in a field, which seems not belonging to these advanced high specialties and may be considered by young specialists as math games and instructive excursus to origins of exact sciences.

Due to developments in computer technology and increased emphasis on information security, the secret communication has attracted much attention, and many volumes have been published in this field, e.g.¹⁻⁴ The initial viewpoint of most books and articles on this topic proceeds from the supposition of identity of a secret communication and cryptography. The latter is often reduced to garbling a block of plain text to make it unreadable by changing the order or replacing characters.³ There is a direct step from this to a variety of mathematical means like matrix and number theory, theory of finite fields, theory of computational complicity, etc. that transforms cryptography into a part of pure mathematics, based on certain introduced postulates of type "the adversary may know the encryption and decryption algorithms which are being used, but does not know additional piece of information to be kept secret".⁴ As is clear, mentioned algorithms used for garbling. On the other hand, secret communication also include as a part the engineering art, construction of submarine cables, double bottom suitcases, making tattoo-letters on bold head before growing hair, using of special note pads, etc.

Here we propose another approach remaining messages untouched and using only numbering of these objects. Numbering principle resembling the construction of a dictionary, where position (number) of a word has nothing to do with its sense. So, any sentence may be considered as a set of numbers of all words of it, taken from some very large dictionary, which contains all grammatic forms of each (say English) existing word.

Using more standard terminology, we endow any n -block with a serial number, but the plain text itself formally remains untouched. The communication is carried out through the exchange of serial

numbers, as if the receiver has secret gigantic dictionary of numbered all words, all possible combinations of letters, or other symbols, etc. Thus, any imaginable message would be among them, and communication requires only to send the proper number. This article is about how to reproduce (prepare) the message from its serial number. The difference between garbling and numbering methods is like the difference between continuous and Rieman's indicator functions in mathematical analysis. Obviously, disproving the adversary of the possibility to extract plane text from the serial number makes communication secret, but here this is not a single possibility. We can supply the adversary with distorted serial number. Hence, the role of "key" (and "postulate") here is changed.

These problems resemble unraveling anagrams. Anagrams have been used for ciphering since the Middle Ages, when scientists wrote their discoveries in the form of anagrams to prove their priority and deciphered the anagrams once their discoveries had been verified. The anagram connected to Galileo Galilei's discovery of Saturn's rings is the one of the most famous (see, e.g. Perelman's "Amusing Astronomy", and a lot of other sources).⁴

Anagrams and cyphers

We consider the ciphered anagram as a set of l distinct letters (k_1, k_2, \dots, k_l) written in alphabetical order, in which any letter may be repeated more than once, namely, n_1, n_2, \dots, n_l times. The total number of characters in the anagram is $n = n_1 + n_2 + \dots + n_l$. The anagram is, in fact, the contents of the block cipher, and n and l are natural block parameters. The multiplicity of each letter (n_k) in the deciphered anagram **must be the same**, by definition, as in the ciphered anagram.

Galilei announced his observation in the anagram *Smaismrmilmepoetaleumibuvnenugtavlras*. The well-known astronomer Kepler assumed that Galilei had discovered two satellites of Mars. Kepler combined the letters into the Latin phrase "*Salve umbestinium geminatum Martia proles*," (which means, "Hello to you, twins, Mars' origination.")

Galilei decoded the anagram after he verified his discovery. "*Altissimum planetam tergeminum observavi*," which means, "I observed the highest planet as triple." Kepler's failure to decode the anagram shows the height of genius (two Mars satellites unobservable

at Kepler’s time were discovered two centuries later) and the computational complexity of combinatorics, where even with the knowledge of letter types and quantities too many possibilities exist to construct meaningful phrases.

Below, we shall prove that knowledge of **two** “magic” natural numbers is sufficient to restore the initial text of any anagram. These two numbers are an example of an ***n*-block cipher**.

The *n*-block ciphers, considered in this article, rely on a solution to a purely mathematical problem of numbering of natural ordered vectors (NOV) introduced in the first paragraph of this Sec. It is possible to build $(m|n)$ copies of the NOV from the natural numbers of the initial segment $[1, m]$ of a natural series ($m > n$). (We use the “one-floor” notation $(m|n)$ for binomial coefficients.) The $(m|n)$ copies of the NOV can be organized in the same way as words in a dictionary, *i.e.*, lexicographically. Hence, the NOV can be assigned serial numbers.

The serial number, of (k_1, k_2, \dots, k_n) is higher than the serial number of (q, k_2, \dots, k_n) if $0 < q < k_1$ because the latter precedes (k_1, k_2, \dots, k_n) and has number $(m - q | n - 1)$. Indeed, any NOV starting with q followed by $n - 1$ integers, taken from the “tail” $[m - q + 1, \dots, m]$ of the initial segment, is ahead of (k_1, k_2, \dots, k_n) according to the lexicographic rule. Since q varies from 1 to $k_1 - 1$, the total number of NOV is

$$\sum_{q=1}^{k_1} (m - q | n - 1) = (m - q | n) - (m - k_1 + 1 | n), \text{ as is clear from}$$

the identity

$$\sum_{i=l}^u (i | s) = (u + 1 | s + 1) - (l | s + 1). \tag{1}$$

However, these are not all NOV, which proceed (k_1, k_2, \dots, k_n) . If q occupies the place of k_2 , taking a value between k_1 and k_2 , such NOV precedes (k_1, k_2, \dots, k_n) , increasing the desired number by

$$\sum_{q=k_1+1}^{k_2-1} (m - q | n - 2) = (m - k_1 | n - 1) - (m - k_2 + 1 | n - 1).$$

Being in the j^{th} place, q contributes similarly

$$\sum_{q=k_j+1}^{k_{j+1}-1} (m - q | n - j - 1) = (m - k_j | n - j) - (m - k_{j+1} + 1 | n - j), j = 0, 1, \dots, n - 1 \tag{2}$$

We account for each time that the “tail” of the initial segment decreases by 1 by using Eq. (1). Adding 1 to the sum of the contributions of all intervals, we obtain the **serial number of *n*-long NOV** selected from the initial segment of length m of a natural series:

$$N_{(k_1 < k_2 < \dots < k_n)}^{(m)} = (m | n) - \sum_{j=1}^n (m - k_j | n - j + 1). \tag{3}$$

In example $m=30, n=3$ the serial number of NOV $(7, 19, 22)$ is $N^{(30)}_{(7, 19, 22)} = (30|3) - (23|3) - (11|2) - (8|1) = 5 \cdot 29 \cdot 28 - 23 \cdot 11 \cdot 7 - 11 \cdot 5 - 8 = 2226$, and of $N^{(30)}_{(28, 29, 30)} = (30|3) - (2|3) - (1|2) - (0|1) = 5 \cdot 29 \cdot 28 = 4060$. Generally, the serial number of NOV $= (1, 2, \dots, n)$ is

$$N_{(1, 2, \dots, n)}^{(m)} = (m | n) - \sum_{j=0}^{n-1} (m - n + j | m - n - 1) = (m | n) - (m | m - n) + (m - n | m - n) = 1;$$

the result follows from Eq. (1). Similarly, $k_1 = m - n + 1$ generates the last NOV since

$$N_{(m-n+1, \dots, m)}^{(m)} = (m | n) - \sum_{j=1}^n (n - j | n - j + 1) = (m | n):$$

each sum term is obviously zero.

The last result coincides with the number of all combinations of n elements taken from the set of m ones.

Numbered set of natural vectors as *n*-block cypher

Returning to Eq. (3), we restore NOV noting that the last binomial coefficient in sum $(30 - k_3 | 1) + (30 - k_2 | 2) + (30 - k_1 | 3) = (30 | 3) - 2226 = 1834$ is the largest, but $30 - k_1$ cannot be 24 since $(24 | 3) = 2024 > 1834$. This suggests to choose $30 - k_1 = 23, k_1 = 7$, reducing equation to simpler one $(30 - k_3 | 1) + (30 - k_2 | 2) = 63$. Going further, we take $30 - k_2 = 11, k_2 = 19$ because of $(11 | 2) = 55, (12 | 2) = 66$, and finally $k_3 = 30 - 63 + 55 = 8$.

Thus, the main problem at each step is finding the best approximation of the sum of binomial coefficients by a single binomial coefficient with the maximal second symbol. Steadily increasing the first symbol of the same binomial coefficient as in the inequalities from the first step we obtain: $(23 | 3) = 1771 < 1834 < 2024 = (24 | 3)$.

It is possible to apply this technique in a general case, too. Namely, we should find for two natural numbers N and k the integer p , such that the large N is between: $(p | k) \leq N < (p + 1 | k)$. It is supposed that p is much smaller than N . The result we write as $p(k, N) = p$.

Starting from some p , we increase or decrease p one by one until the inequality $(p | k) \leq N < (p + 1 | k)$ is satisfied, and N is trapped between two “adjacent” binomial coefficients. Finding N would be not difficult if a table of the necessary binomial coefficients was available. However, if the number of NOV is large (as in a telephone directory of a big city: hundreds of millions), determining the interval (4) borders requires plenty of calculations of binomial coefficients with large entries that makes, in fact, the problem unsolvable.

However, we can find an approximate value of p in the inequality $(p | k) \leq N < (p + 1 | k)$:

$$p(k, N) = \left\lfloor \sqrt[k]{k!N} + \frac{k}{2} \right\rfloor. \tag{4}$$

where $\lfloor \dots \rfloor$ denotes the integer part. Indeed, it is easy to connect N and p if N is in the middle between $(p | k)$ and $(p + 1 | k)$ *i.e.*, $N = \frac{1}{2} \{ (p | k) + (p + 1 | k) \}$.

The binomial coefficient definition gives

$$Nk! = p(p - 1) \dots (p - k + 2)(p + 1 - \frac{1}{2}k). \tag{5}$$

To estimate the right part, we consider that $p > k$ and replace each of k factors from the right by p . Then $N = p^k / k!$ becomes much greater than $(p | k)$, and N in the middle is shifted right significantly and can go behind the right interval border. To reduce this effect, we replace each factor not by the maximal one p , but by the arithmetic mean $[p + (p - 1) + \dots + (p - k + 2) + (p + 1 - \frac{1}{2}k)] / k = [(k - 1)p - \frac{1}{2}(k - 1)(k - 2) + p + 1 - \frac{1}{2}k] / k = p + 1 - \frac{1}{2}k$

that gives Eq. (5) form $Nk! = (p + 1 - \frac{1}{2}k)^k$.

Therefore, it is rational to start tests of p from the one presented in Eq. (4) in which 1 is omitted to close the arithmetic mean to the geometric mean (the power of right part Eq. (5)).

In our example, $P(3,1834) = [(3!1834)^{1/3} + 1.5] = [23.742...] = 23$. Already the first attempt places 1834 between (23|3) and (24|3). A similar result is usually obtained for bigger numbers, e.g., $N = 71452963455130$, $k = 8$, give $p = [(8!71452963455130)^{1/8} + 4] = [206.975...] = 206$ verifying $(206|8) = 70090194034625$ and $(207|8) = 72907890277275$ satisfy $(206|8) < N < (207|8)$.

Similarly: $k = 6$, $N = 8863660966$ lead to $p = [(6!8863660966)^{1/6} + 3] = 139$, which guarantees $(138|6) < N_1 < (139|6)$ since $(139|6) = 8979650478$ and $(138|6) = 8592039666$ etc., without long cumbersome process of calculation of sequence of binomial coefficients.

Thus, we built a **method of numbering of NOV and back finding the lexicographic sequence of NOV components from NOV given serial number**. To create a practical secure communication method, it only remains to extend these results on unordered natural vectors, which able to “carry” any, say, English text. Let our “building material” is the empty space (between letters) and the English alphabet: $1, A_2, B_3, C_4, D_5, E_6, F_7, G_8, H_9, I_{10}, J_{11}, K_{12}, L_{13}, M_{14}, N_{15}, O_{16}, P_{17}, Q_{18}, R_{19}, S_{20}, T_{21}, U_{22}, V_{23}, W_{24}, X_{25}, Y_{26}, Z_{27}$, equivalent to the initial segment $[1, 27]$ of natural series.

To deal still with NOV, we introduce a set of ordered auxiliary numbers $s_1 = k_1, s_2 = k_1 + k_2, s_3 = s_2 + k_3, \dots, s_n = s_{n-1} + k_n : s_1 < s_2 < \dots < s_n$ (the initial k_j may be chaotic and mutually equal). The largest possible $s_n = nm$ appears when all $k_j = m$. Therefore, the basic segment for s_j is $[1, mn]$. We can find the serial number of NOV of type (s_1, s_2, \dots, s_n) among all NOV built of s_j from $[1, mn]$ by means of Eq. (3), remembering that now the largest s is nm . We can give the unordered vector (k_1, k_2, \dots, k_n) the same serial number as that of (s_1, s_2, \dots, s_n) , where s_1, s_2, \dots, s_n are the parameters from which (k_1, k_2, \dots, k_n) are built: $k_1 = s_1, k_j = s_j - s_{j-1}, j = 2, \dots, n$.

$$N_{(k_1 k_2 \dots k_n)}^{(m)} = (mn | n) - \sum_{j=1}^n (mn - s_j | n - j + 1), s_j = \sum_{i=1}^j k_i, j = 1, \dots, n. \quad (6)$$

In particular, the “zero word”, corresponding to $k_1 = k_2 = \dots = k_n = 1$, means an empty interval of length n . It has number 1 since the sum in Eq. (6) equals

$$\sum_{j=0}^{n-1} (mn - n + j | mn - n - 1) = (mn | mn - n) - (mn - n | mn - n)$$

according to Eq. (1). If the last letter of a word is k , and the remaining positions are empty, k is the number of the whole word since $N_{(1,1,\dots,1,k)}^{(m)} = 1 + mn - n - mn + n + k - 1 = k$, where the preceding result has been used.

Now we can collect the results in a method of encrypted messaging. The serial number of the plain text of an n -block should be calculated by means of Eq. (6) and sent to the intended recipient at the decoding terminal. To convert the n -block serial number into a plain text, the recipient builds a set of n auxiliary numbers N_j and uses them to restore s_j and then k_j . The key point is finding a trial parameter p_j and comparing $(p_j | k_j)$ with the known N_j that determines the final value p_j , which restricts N_j from below $(p_j | k_{j-1}) < N_j < (p_j + 1 | k_{j-1})$. Namely,

$$N_1 = (mn | n) - N_{k_1 k_2 \dots k_j}, p_1 = \left\lceil \sqrt[n]{(n!N_1)} + \frac{n}{2} \right\rceil, (p_1 | n), p_1 = p(n, N_1), s_1 = k_1 = mn - p_1. \quad (7)$$

$$N_{j+1} = N_j - (p_j | n - j + 1), p_{j+1} = \left\lceil \sqrt[n-j+1]{((n-j+1)!N_{j+1})} + \frac{n-j+1}{2} \right\rceil, (p_{j+1} | n - j + 1), p_{j+1} = p(k_j, N_{j+1}), s_{j+1} = mn - p_{j+1}, k_{j+1} = s_{j+1} - s_j, j = 1, 2, \dots, n-1. \quad (8)$$

If all n -blocks are written in order of their serial numbers, then all decoded plain texts are ordered lexicographically as in a dictionary. It is natural to refer to this technique as the **cryptographic dictionary method**.

A specific self-defense mechanism is inherent in dictionary method. The serial numbers of the NOV given by Eq. (7) contain n arbitrary parameters s_j from the interval $[1, mn]$. Not all of them can correspond to words. The total number of words is m^n , while the total number of all possible sequences (s_1, s_2, \dots, s_n) is greater: $(mn | n)$. For instance, the ratio

$$(270 | 10) / 27^{10} = \left(\frac{479322759878148681}{205891132094649} \right) = 2328.04...$$

The system will distinguish one possible serial number from 2300 “trash” cases. Hence, only one case from a random flow of n -blocks requires decoding to test the correctness. For instance, the last NOV has serial number $(mn | n)$, while the last word $mm \dots m$ corresponds to s -set $(m, 2m, 3m, \dots, m^2)$ with the serial number $(mn | n) - \sum_{j=1}^{n-1} (mj | j + 1)$.

The trivial false example is an NOV with the first number s_1 , which exceeds m . Such s_1 cannot lead to a word of an m -long alphabet. Similarly, if Eq. (9) gives $s_{j+1} - s_j > m$, the result goes to “trash.”

Now we give a detailed step by step demonstration of these procedures on a popular example. Each acting telephone has a number, which easy to find in telephone directory. However, the reverse problem is much more difficult. Finding of the subscriber of a given telephone number is often used in cryptography as an example of a one-way function. Equation (3), allows us to endow the telephone owner by the serial number and prepare “the directory of telephone owners”, e.g., subscribers with a one-letter family name have it as serial number, but all other subscribers like “Wilson” (24,10,13,20,16,15) are identified by Eq. (6):

$$N_{(24,10,13,20,16,15)}^{(27)} = (162 | 6) - (138 | 6) - (128 | 5) - (115 | 4) - (95 | 3) - (79 | 2) - (64 | 1) = 22860316584 - 8592039666 - 264566400 - 6913340 - 138415 - 3081 - 64 = 1399 665 5618.$$

If we know this number, we restore letters of his name one-by-one, retaining each time only the biggest remaining term of the sum in Eq. (6), beginning with the first, containing only $k_1: N_1 = (162|6) - 13996655618 = 8863660966$. The described operations verify that $p_1 = [(6!8863660966)^{1/6} + 3] = 139$, $(139|6) = 8979650478$. Hence, $p(n, N_1) = 138$ since $(138|6) < N_1 < (139|6)$, and $s_1 = k_1 = 162 - 138 = 24$, i.e. the first letter is W_{24} .

Further, $N_2 = 8863660966 - (138|6) = 271621300$, $p_2 = [(5!271621300)^{1/5} + 2.5] = 129$, $(129|5) = 275234400$, then $p(k_1, N_2) = 128$ as $(128|5) < N_2 < (129|5)$, and $s_2 = k_1 + k_2 = 162 - 128 = 34$, $k_2 = 10 \rightarrow I_{10}$. The rest of letters are obtained similarly. $N_3 = 271621300 - (128|5) = 7054900$, $p_3 = [(4!7054900)^{1/4} + 2] = 116$, $(116|4) = 7160245$, $(115|4) < N_3 < (116|4)$, $p(k_2, N_3) = 115$, $s_3 = k_1 + k_2 + k_3 = 162 - 115 = 47$, $k_3 = 13 \rightarrow L_{13}$. $N_4 = 7054900 - (115|4) = 141560$, $p_4 = [(3!141560)^{1/3} + 1.5] = 96$, $(96|3) = 142880$, $(95|3) < N_4 < (96|3)$, $p(k_3, N_4) = 95$, $s_4 = 162 - 95 = 67$, $k_4 = 20 \rightarrow S_{20}$. $N_5 = 141560 - (95|3) = 3145$, $p_5 = [(2!3145)^{1/2} + 1] = 80$, $(80|2) = 3160$, $(79|2) < N_5 < (80|2)$, $p(k_4, N_5) = 79$, $s_5 = 162 - 79 = 83$, $k_5 = 16 \rightarrow O_{16}$. $N_6 = 3145 - (79|2) = 64$, $s_6 = 162 - 64 = 98$. After $k_6 = 15 \rightarrow N_{15}$ appears, we hail the owner of 1-399-665-5618: “Wilson!”. That is how one-way function works.

Numbering anagrams

Coding of anagrams gives the cue to a more convenient technique of numbering n -blocks than the dictionary method. It endows each

n -block with **two** serial numbers, which could be sent to different addressees, making decoding possible only after unification of both messages.

The anagram is specified by the number, l , of different sorts of characters in it and by l natural numbers n_1, n_2, \dots, n_l (multiplicities), which define numbers of characters of each sort. The text of an n -block is considered as an **anagram**. The net sum of multiplicities is fixed and defines $n_1+n_2+\dots+n_l=n$, the n -block size. All these parameters are identical in the ciphered and deciphered anagram. The ciphered Galilei-Kepler anagram, quoted above, can be written as $a^4 b e^4 g i^4 l^2 m^3 n^2 o p r^2 s^3 t^3 u^3 v$ meaning that $n_a=4, n_b=1, n_e=4, n_g=1, \dots, n_u=3, n_v=1, n=37, l=15$ since the order of letters in it is arbitrary. This type of notation is symbolized as

$$k_1^{n_1} k_2^{n_2} \dots k_j^{n_j} \dots k_l^{n_l}, \tag{9}$$

k_j and index j of superscript are replaced by the letter-symbol with number j in alphabet. ("Exponents" n_1, n_2, \dots, n_m do not mean powers of letters' numbers). The total number of anagrams with fixed l and n equals to the number of combinations with repetitions of n objects of l different kinds (it is assumed that n objects of each kind are available):

$$N = (n + l - 1 | l - 1) \tag{10}$$

This formula is confirmed by induction. If new n_{l+1} letters are placed in the same block, $n - n_{l+1}$ positions remain for the former letters, which can be placed in these positions by $(n - n_{l+1} + l - 1 | l - 1)$ ways according to the induction hypothesis. All admissible situations are created by n_{l+1} values from 0 to n :

$$N = \sum_{n_{l+1}=0}^n (n - n_{l+1} + l - 1 | l - 1).$$

After applying Eq. (1) with the summation index replaced by j as $n_{l+1} \rightarrow n - j + l - 1$, we obtain an expression with l increased by

$$1: \sum_{j=l-1}^{n+l-1} (j | l - 1) = (n + l | l) \text{ that proves Eq. (10) for the last}$$

possible serial number of the ciphered anagram with $l+1$ types of letters.

Let us find the serial number for a cyphered anagram among similar ones arranged in the lexicographic order. If all letters of a ciphered anagram are n copies only of the first one, k_1 , there is a single letter distribution: k_1^n having number 1. If there are n_1 copies of k_1 and n_2 copies of k_2 ($n_1+n_2=n$), the distribution $k_1^{n_1} k_2^{n_2}$ is preceded by all pairs $k_1^{n_1} k_2^{n_2}$ with $0 \leq j < n_2$, and there are n_2 such pairs in total. Hence, the anagram $k_1^{n_1} k_2^{n_2}$ will be the n_2+1 -th one, *i.e.*, its serial number is $N=n_2+1$. This means that the total number of possible ordered occupations of n places by two letters k_1, k_2 is $n+1$.

When the third letter, k_3 , is added, the total number of occupations, as Eq. (10) says, becomes $(n+2|2)$. The serial number of the distribution $k_1^{n_1} k_2^{n_2} k_3^{n_3}$ consists of the same n_2+1 distributions, which were

ahead of $k_1^{n_1} k_2^{n_2}$ (now in the presence of $k_3^{n_3}$). All $n+1$ distributions coming from k_1 and k_2 posted in n places and not including k_3 ,

are also ahead of $k_1^{n_1} k_2^{n_2} k_3^{n_3}$. The presence of one k_3 remains $n-1$ places for placing k_1 and k_2 . The presence of two k_3 remains $n-2$

places, etc. This gives in total $\sum_{j=0}^{n_3-1} (n+1-j) = (n+1)n_3 - (n_3 | 2)$,

plus n_2+1 already mentioned, *i.e.*, the serial number i

$$N_{n_1, n_2, n_3} = nn_3 + n_2 + n_3 + 1 - (n_3 | 2) = (n + 2 | 2) - n_1 - (n_1 + n_2 + 1 | 2). \tag{11}$$

The second, more convenient version of Eq. (11) is obtained after substitution of $n_3 = n - n_1 - n_2$. The result contains n_3 implicitly and makes formula (11) symmetric in the remaining multiplicities, which suggests the general expression

$$N_{n_1, n_2, \dots, n_l} = (n + l - 1 | l - 1) - \sum_{j=1}^{l-1} (s_j + j - 1 | j), \quad s_j = \sum_{i=1}^j n_i, s_l = n, \tag{12}$$

in which the last occupation number n_l is not explicitly present. This allows us to carry out the induction and complete the proof of Eq. (12).

All N_{n_1, n_2, \dots, n_l} anagrams, the last of which is $k_1^{n_1} k_2^{n_2} \dots k_l^{n_l}$, precede $k_1^{n_1} k_2^{n_2} \dots k_l^{n_l} k_{l+1}^{n_{l+1}}$. Hence, N_{n_1, n_2, \dots, n_l} is the first contribution to $N_{n_1, n_2, \dots, n_{l+1}}$. The anagram $k_1^{n_1} k_2^{n_2} \dots k_l^{n_l} k_{l+1}^{n_{l+1}}$ brings with itself $(n-1+l-1|l-1)$ anagrams more. The latter are formed by all possible distributions of k_1, k_2, \dots, k_l over $n-1$ remaining places, and all are ahead of $k_1^{n_1} k_2^{n_2} \dots k_l^{n_l} k_{l+1}^{n_{l+1}}$. Similar reasoning is related to $k_1^{n_1} k_2^{n_2} \dots k_l^{n_l} k_{l+1}^{n_{l+1}}$, but now there are $n-2$ remaining places, and so on. The last case is related to n_{l+1} and leaves $n - n_{l+1}$ places.

The total contribution is according to the rule (1)

$$\sum_{j=1}^{n_{l+1}} (n - j + l - 1 | l - 1) = (n + l - 1 | l) - (n - n_{l+1} + l - 1 | l) \tag{13}$$

Adding Eq. (13) to Eq. (12), we recognize in the sum the first term, $(n+l|l)$, of $N_{n_1, n_2, \dots, n_{l+1}}$ and the last term of sum (12) for $l+1$ in the subtrahend of (13) since $n - n_{l+1} = s_l$. This demonstrates the validity of Eq. (12) for $l+1$ and proves it for any l .

Note, the trivial case $l=1$ of Eq. (12): $N_n = 1$ ($l=1$) demonstrate an important peculiarity of parameter l . It can be arbitrary when the same situation is described by suitable multiplicities. The same case $N_{n, 0, \dots, 0} = 1$ follows from the equality of all s_j to n that turn the sum in Eq. (12) into $(n + l - 1 | n) - 1$ according to Eq. (1) and $N_{n, 0, \dots, 0}$ into 1. Another example: if letter k_l occupies l -th place,

$N_{0, \dots, 0, n} = (n + l - 1 | l - 1)$ as all s_j in the sum are 0, and the sum is 0, too.

The anagram, which consists of copies only of one (any) letter k , has a serial number $(n + k - 1 | n)$. Indeed, $k \leq l-1$ is an **admissible value** of the summation index j in Eq. (12) and appears as a subscript at n_k in the sum for s_j . Therefore, n_k is absent, as well as all other multiplicities in all s_j , with $j < k$ turning them (and, hence, the next terms) into 0. All

$$s_j = n \text{ in the remaining sum } \sum_{j=k}^{l-1} (n + j - 1 | j) \text{ that, thanks to Eq. (1), reduces the whole expression (12) to } (n + k - 1 | n).$$

The same idea allows us to extend the summation in Eq. (12) to the entire alphabet of m letters without any other changes to Eq. (12). The extended summation domain of sub-subscripts up to m simultaneously implies that multiplicities n_j of the absent letters are zero (and omitted in the multi-index at N).

$$N_{n_1, n_2, \dots, n_m} = (n + m - 1 | m - 1) - \sum_{j=1}^{m-1} (s_j + j - 1 | j), \quad s_j = \sum_{i=1}^j n_i, \sum_{i=1}^m n_i = n.$$

$$(14)$$

This does not mean that the anagram necessarily should contain all m letters, but rather that the summation index runs through the values of all letters of the alphabet. However, only if the summation index in (12) equals any of l letters, which are present in the anagram, the corresponding parameter s_j does (for $j=k_p$) jump to n_{k_p} . All terms are 0 before $j=k_1$ (since $s_j=0$), s_j remains unchanged if j is between of any neighboring k_p , and all $s_j=n$ after k_r . The last part of the sum is calculated by means of Eq. (1), which replaces m with k_r .

$$N_{n_1, k_1; n_2, k_2; \dots; n_m, k_m} = (n+m-1 | m-1) - \sum_{j=k_1}^{l-1} (s_j + j - 1 | j) - \sum_{j=k_l}^{m-1} (n+j-1 | j) = (n+k_l-1 | k_l-1) - \sum_{j=k_1}^{k_l-1} (s_j + j - 1 | j). \tag{15}$$

The obtained sum breaks up into segments with constant s_j , giving the final formula.

$$N_{n_1, k_1; n_2, k_2; \dots; n_l, k_l} = (n+k_l-1 | k_l-1) + \sum_{j=1}^{l-1} [(s_j + k_j - 1 | s_j) - (s_j + k_{j+1} - 1 | s_j)], s_j = \sum_{i=1}^j n_i. \tag{16}$$

Implementation of the letter numbers k_j into Eq. (16) allows a new interpretation of this result. It looks merely like a number of a party of objects of given kinds (individual numbers, k_j , in the list) in fixed number (n_j) of copies. The serial number of an anagram is a particular case: n_j and k_j determine the ciphered anagram content.

Equation (16) suggests the following organization of the secret communication line between a storage and consumer. Let k_j be the numbers of products in some catalog available to the customer and seller. Then customer sends to the seller **inventory number** $N_{n_1, k_1; n_2, k_2; \dots; n_l, k_l}$, which contains all information of his purchase and can be made secret after randomization procedure (described in Sec. 8). The seller solving the reverse problem for $N_{n_1, k_1; n_2, k_2; \dots; n_l, k_l}$ determines the customer purchase.

Reverse problem for list of numbered objects

A solution of the reverse problem for Eq. (16) converts the serial number $N_{n_1, k_1; n_2, k_2; \dots; n_l, k_l}$ into the set of letters k_j and multiplicities n_j . The solution uses as before the properties of the sum of binomial coefficients. The binomial coefficient with the largest absolute value, which presents in sums from Eqs. (3), (6), (12), and (16), is greater than the sum of all remaining terms. This was the base of procedures (4). The presence of negative sum complicates the situation for Eq. (16), which nevertheless is still solved by means of the same function $p(k, N)$, but now operates simultaneously with two sequences beginning from their opposite ends. The value of l is not needed beforehand to begin the process; l rather indicates the process ending when a certain parameter $M^{(l)}$ turns into 0; k_1 appears at the algorithm end. (It is convenient to supply k_1 by additional index – the superscript, indicating the step of algorithm, on which the given k_j appears: $k_j \equiv k^{(j+1)}$, $k_1 \equiv k^{(l)}$. Letters $k^{(j)}$ appear on the odd steps, and multiplicities on the even steps through their sums s_j , which are also supplied by subscripts: $s^{(j+1)} \equiv s_j$.)

This notation is in accordance with the calculation algorithm, which begins by using $s^{(1)}=s_1=n$ and finishes producing $s^{(l)}=s_1$. The latter give multiplicities at the end of the computation $s_1=n_1, s_2-s_1=n_2, \dots, s_p-s_{p-1}=n_p$, while letters k_j appear in natural order during the algorithm:

$$N^{(1)} = N_{n_1, k_1; n_2, k_2; \dots; n_l, k_l}, s^{(1)} = n; \tag{17}$$

$$k^{(j)} = p(s^{(j)}, N^{(j)}) - s^{(j)} + 2, M^{(j)} = (k^{(j)} + s^{(j)} - 1 | s^{(j)}) - N^{(j)}; j = 1, 2, \dots, s^{(j+1)} = p(k^{(j)} - 1, M^{(j)}) - k^{(j)} + 2, N^{(j+1)} = (k^{(j)} + s^{(j+1)} - 1 | k^{(j)} - 1) - M^{(j)}. \tag{18}$$

In a peculiar situation, when the left basic inequality for the function (4) turns into an exact equality, the last 2 in the equation for $k^{(j)}$ should be replaced by 1. The appearance of $N^{(l-1)}=0$ finalizes calculations and determines the number of different kinds, l , of products, beforehand unknown to the seller. The reciprocal algorithm for the list with one copy of each product is simpler (compare also with Eqs. 7, 8):

$$N^{(l)} = N_{1, k_1; 1, k_2; \dots; 1, k_l}, k^{(j)} = p((l-j+1), N^{(j)}) - l + j + 1, N^{(j+1)} = N^{(j)} - (k^{(j)} + l - j - 1 | l - j + 1). \tag{19}$$

Reverse problem for anagram

Since the serial number of the ciphered anagram is a particular case of the inventory number, the recipient has all n -block letters and multiplicities after solving the reverse problem (18). However, there is a principal difference between letters and all other products: letters can group in sensible words. Therefore, after decipher of anagram contents, recipient finds himself in Kepler’s position attempting to get the meaning of Galileo’s anagram. The ciphered anagram of type (9)

admits $\frac{n!}{n_1! n_2! \dots n_l!}$ distinguishable locations of the given letters

for fixed multiplicities. Thus, obtaining the inventory number is only the first step to make anagram numeration into a cryptographic tool. The second step is to find the unique locations of the given letters.

Let us determine letter’s locations in n -block by means of the following two-row and n -column table (called below a collocation). Each row contains n positions (in agreement with the size of n -block).

$$\begin{pmatrix} k_1 & k_2 \dots & k_l & k_1' & k_2' \dots & k_r' & k_1' & k_2' \dots & k_l' & k_1^\infty & k_2^\infty \dots & k_r^\infty \\ v_1^{(1)} & v_1^{(2)} \dots & v_1^{(l)} & v_2^{(1)} & v_2^{(2)} \dots & v_2^{(r)} \dots & v_1^{(l)} & v_1^{(2)} \dots & v_1^{(r)} & v_\infty^{(1)} & v_\infty^{(2)} \dots & v_\infty^{(r)} \end{pmatrix} \tag{20}$$

The first l positions of the first row are occupied by numbers of all **different** letters k_j of anagram (9), in alphabetic order (or by corresponding letter-symbols themselves). The numbers $v_j^{(1)}$ of the first position of each of these letters in n -block are situated under the corresponding letter in the second row. Then, the part (or all) of the same letters (primed) follows, which occupy in n -block at least two positions, with the numbers $v_j^{(2)}$ under them, etc. The letters with maximal multiplicities fall into the last group.

A letter, which is repeated in the n -block n_j times is present in the first row also n_j times. This is why the total number of all letters in the first row is

$$\sum_{j=1}^l n_j = n.$$

The numbers 1, 2, 3, ..., n of all possible letter positions of the n -block are preset in chaotic order in the second row, but their sum

$$\sum_{j=1}^n v_i^{(j)} = \sum_{j=1}^n j = (n+1 | 2) \text{ is fixed.} \tag{21}$$

Let the columns are rearranged in such a way that all numbers in the second row are ordered. Then the first row turns into the deciphered anagram. This suggests that we can apply the technique used above to number collocations (20). In particular,

the number, given in Eq. (14), being applied to (20), requires to substitute $n_j \rightarrow v_i^{(j)}$, the number m of parameters n_j by $m \rightarrow n$ and allow us to endow the collocation (20) with the serial number

$$N_{v_1, v_2, \dots, v_n} = \left(\frac{n(n+3)}{2} - 1 | n-1 \right) - \sum_{j=1}^{n-1} (\sigma_j + j - 1 | j), \quad (22)$$

$$\sigma_j = \sum_{i=1}^{n-1} v_i, \sum_{i=1}^n v_i = (n+1 | 2)$$

and to use the same reciprocal algorithms (7) – (8) as for Eq. (3).

Illustrations of anagram communication method

If we intend to spread our exaggerated estimation of the results obtained: “IT IS IT”=10,21,1,10,20,1,10,21 (see letter numbers in Sec. 3), it can be done only secretly. Starting encoding we take parameters (9), which enter Eq. (16): $l=4, k_1=1, k_2=10, k_3=20, k_4=21; n_1=2, n_2=3, n_3=1, n_4=2; s_1=2, s_2=5, s_3=6, s_4=8=n$ and find the serial number:

$$N_{2,1;3,10;1,20;1,2,21} = (28|8) + (2|2) - (11|2) + (14|5) - (24|5) + (25|6) - (26|6) = 3108105 + 1 - 55 + 2002 - 42504 + 177100 - 230230 = 3014419,$$

hiding the anagram content. Next, we conceal the letter’s order.

Parameters v_j , which enter Eq. (22), can be taken from the second row of collocation (20) for our message:

$$\begin{pmatrix} \square & I & S & T & \square & I & T & I \\ 3 & 1 & 5 & 2 & 6 & 4 & 8 & 7 \end{pmatrix}$$

written in terms of letter-symbols in lexicographic order (including the empty spaces) in the first row. Thus, $v_1=3, v_2=1, v_3=5, v_4=2, v_5=6, v_6=4, v_7=8, v_8=7; \sigma_1=3, \sigma_2=4, \sigma_3=9, \sigma_4=11, \sigma_5=17, \sigma_6=21, \sigma_7=29$ transform Eq. (22) into $N_{3,1,5,2,6,4,8,7} = (43|7) - (3|1) - (5|2) - (11|3) - (14|4) - (21|5) - (26|6) - (35|7) = 32224114 - 6724520 - 230230 - 20349 - 1001 - 165 - 10 - 3 = 25247836$. The two-number message: $G=3014419, K=25247836$ is ready.

Using these two numbers and the block size $n=8$ (which the recipient already knows), the recipient restores the full text of the message. Starting from Eq. (17): $N^{(1)}=3014419, s^{(1)}=8$, he moves to $j=1: k^{(1)}=p(8,3014419)-6=27-6=21, M^{(1)}=(28|8)-3014419=93686, s^{(2)}=p(20,93686)-19=25-19=6, N^{(2)}=(26|20)-93686=136544; then j=2, k^{(2)}=p(6,136544)-4=24-4=20, M^{(2)}=(25|6)-136544=40556, s^{(3)}=p(19,40556)-18=23-18=5, N^{(3)}=(24|19)-40556=1948; at last j=3 k^{(3)}=p(5,1948)-3=13-3=10, M^{(3)}=(14|5)-1948=54, s^{(4)}=p(9,54)-8=14-12=2, N^{(4)}=(11|9)-54=1, and according to the proviso $k^{(4)}=1$. Finally, $n_1=s_1=s^{(4)}=2, n_2=s^{(3)}-s^{(4)}=5-2=3, n_3=s^{(2)}-s^{(3)}=6-5=1, n_4=s^{(1)}-s^{(2)}=8-6=2$. (Defining p -function inequalities were checked on each step). Hence, the ciphered anagram obtained by the recipient from 3014419 is such: $\square^2 I^3 S T^2$.$

The number N_{v_1, v_2, \dots, v_n} (20) generates the following algorithm to establish the order of the letters in the n -block (consisting of $n=8$ steps in this case).

$$N_i = \left(\frac{n^2 + 3n - 2}{2} | n-1 \right) - N_{v_1, v_2, \dots, v_n}, \sigma_{n-1} = p(n-1, N_i) - n + 2, v_n = (n+1 | 2) - \sigma_{n-1}, N_{j+1} = N_j - (\sigma_{n-j} + n - j - 1 | n - j), \sigma_{n-j-1} = p(n-j-1, N_{j+1}) - n + j + 2, v_{n-j} = \sigma_{n-j} - \sigma_{n-j-1}, j = 1, 2, \dots, n-1. \quad (23)$$

To begin, the recipient transforms the ciphered anagram into the first row of the collocation $\square I S T \square I T I$, then starts to build the second row from the right, performing simple calculations:

$$N_1 = (43|7) - 25247836 = 6976278, \sigma_7 = p(7, 6976278) - 6 = 35 - 6 = 29, v_8 = 36 - 29 = 7;$$

$$N_2 = 6976278 - (35|7) = 251758, \sigma_6 = p(6, 251793) - 5 = 26 - 5 = 21, v_7 = 29 - 21 = 8;$$

$$N_3 = 251758 - (26|6) = 21528, \sigma_5 = p(5, 21528) - 4 = 21 - 4 = 17, v_6 = 21 - 17 = 4;$$

$$N_4 = 21528 - (21|5) = 1179, \sigma_4 = p(4, 1179) - 3 = 14 - 3 = 11, v_5 = 17 - 11 = 6;$$

$$N_5 = 1179 - (14|4) = 178, \sigma_3 = p(3, 178) - 2 = 11 - 2 = 9, v_4 = 11 - 9 = 2;$$

$$N_6 = 178 - (11|3) = 13, \sigma_2 = p(2, 13) - 1 = 5 - 1 = 4, v_3 = 9 - 4 = 5;$$

$$N_7 = 13 - (5|2) = 3, \sigma_1 = p(1, 3) = 3, v_2 = 4 - 3 = 1;$$

$$N_8 = 3 - (3|1) = 0, v_1 = \sigma_1 = 3.$$

Now, the recipient writes the characters $\square I S T \square I T I$ in their places 3,1,5,2,6,4,8,7: $I T \square I S \square I T$.

Another example compares the dictionary and anagram methods. The message is the family name of the owner of telephone number 1-399-665-5618 from our “symmetric telephone directory.” The anagram (9) “ILNOSW” gives $l=6, k_1=10, k_2=13, k_3=15, k_4=16, k_5=20, k_6=24$, which produces the serial number (17) $N_{1, k_1; 1, k_2; \dots; 1, k_6} = 10 + 78 + 560 + 3060 + 33649 + 376740 = 414097$, and Eq. (22) with $v_1=2, v_2=3, v_3=6, v_4=5, v_5=4, v_6=1$, leading to $\sigma_1=2, \sigma_2=5, \sigma_3=11, \sigma_4=16, \sigma_5=20, \sigma_6=21$ and the second serial number $N_{10,13,15,16,20,24} = (26|5) - (2|1) - (6|2) - (13|3) - (19|4) - (24|5) = 65780 - 2 - 15 - 286 - 3876 - 42504 = 19097$, which allows the recipient to start the reciprocal algorithm:

$$N_{1, k_1; 1, k_2; \dots; 1, k_l} = 414097 \cdot k^{(j)} = p((l-j+1), N^{(j)}) - l + j + 1, N^{(j+1)} = N^{(j)} - (k^{(j)} + l - j - 1 | l - j + 1), j = 1 \div l.$$

$$k^{(1)} = p(6, 414097) - 4 = 24 = k_6, N^{(2)} = 414097 - (28|6) = 414097 - 376740 = 37357.$$

$$k^{(2)} = p(5, 37357) - 3 = 20 = k_5, N^{(3)} = 37357 - (23|5) = 37357 - 33649 = 3708$$

$$k^{(3)} = p(4, 3708) - 2 = 16 = k_4, \text{ because } (18|4) = 3060 < 3708 < 3876 = (19|4). N^{(4)} = 3708 - (16+2|4) = 3708 - 3060 = 648.$$

$$k^{(4)} = p(3, 648) - 1 = 15 = k_3, \text{ because } (16|3) = 560 < 648 < 680 = (17|3), N^{(5)} = 648 - (15+1|3) = 648 - 560 = 88.$$

$$k^{(5)} = p(2, 88) = 13 = k_2, \text{ because } (13|2) = 78 < 88 < 91 = (14|2), N^{(6)} = 88 - (13|2) = 10.$$

$$k^{(6)} = p(1, 10) = 10 = k_1, N^{(7)} = 10 - (10|1) = 0.$$

Thus, the letter 1 is 10th in the alphabet, the letter 2 is 13th, etc., and the recipient sees “ILNOSW.”

According to (23) $N_1 = 65780 - 19097 = 46683, \sigma_5 = 24 - 6 + 2 = 20, v_6 = 21 - 20 = 1. N_2 = 46683 - (20+4|5) = 46683 - 42504 = 4179, \sigma_4 = p(4, 4179) - 3 = 16, v_5 = 20 - 16 = 4, N_3 = 4179 - (16+3|4) = 4179 - 3876 = 303, \sigma_3 = p(3, 303) - 2 = 11, v_4 = 16 - 11 = 5, N_4 = 303 - 11 + 2|3 = 303 - 286 = 17, \sigma_2 = p(2, 17) - 1 = 5, v_3 = 11 - 5 = 6, N_5 = 17 - (5+1|2) = 2, \sigma_1 = p(1, 2) = 2, v_2 = 5 - 2 = 3, N_6 = 2 - (2|1) = 0, \sigma_0 = 0, v_1 = 2$. Therefore, letter 6 from ILNOSW goes in the first place,

letter 5 goes in the fourth place, etc.,

$$\begin{pmatrix} I & L & N O & S & W \\ 2 & 3 & 6 & 5 & 4 & 1 \end{pmatrix},$$

and WILSON can be identified operating with numbers that are 33800 times smaller than his telephone number.

Program

Now we assume that numerous and monotonous computations in preceding sections could convince readers that it is rationally to program all manipulations and to deal only with results. To estimate the range of parameters, which PC can operate let us return to the classic

anagram we began from. Multiplicities of the “inventory” n -block of Kepler anagram (including interval) $a^4b^1e^4g^1P^1m^5n^2opr^2s^3t^3u^3v^3$ ($n=40$, $l=16$) are: $n_1=4, n_2=1, n_3=4, n_4=1, n_5=4, n_6=2, n_7=5, n_8=2, n_9=1, n_{10}=1, n_{11}=2, n_{12}=3, n_{13}=3, n_{14}=3, n_{15}=1, n_{16}=3$ if letters are numbered as they appear here. The corresponding sums $s_1=4, s_2=5, s_3=9, s_4=10, s_5=14, s_6=16, s_7=21, s_8=23, s_9=24, s_{10}=25, s_{11}=27, s_{12}=30, s_{13}=33, s_{14}=36, s_{15}=37, s_{16}=40$ lead to the serial number (12) $N_{n_1, n_2, \dots, n_{16}} = (55|15)-(4|1)-(6|2)-(11|3)-(13|4)-(18|5)-(21|6)-(27|7)-(30|8)-(32|9)-(34|10)-(37|11)-(41|12)-(45|13)-(49|14)-(51|15)=7953850584061$, which is calculated using the computation tutorial.⁷ However, deciphering Kepler’s anagram is out of the question for a PC because the binomial coefficient in Eq. (22) is

$\binom{n(n+3)}{2} - 1 | n - 1 = (860|39)$ for $n=40$, which cannot be stored on a 64-bit PC: $2^{64}-1=18446744073709551615$; $2^{64}-1$ is ≈ 6 times greater than this binomial coefficient $(134|14)$ for $n=15$.

The last value allows to program on PC the **anagram method for a 15-block cypher**. The expository description of such program created by Mestechkina T. M. PhD, is shown in Figure 1, where

$K = N_{v_1, v_2, \dots, v_n}, G = N_{k_1, k_2, \dots, k_m, k_m}$. The basic ideas, embodied in the Program, first were set forth in the bid.⁸



Figure 1 General view of the screen. The first row: list of possible actions, described under this Figure. The second row: the window for decoding n -block. The result is shown besides. The third row the window for encoding numbers G and K . The encoded block is shown besides. The button of action is in the middle of the rows. Some known encryptions are chosen as illustrations.

Method: «Dictionary, Anagram, Simple Division»

Edit «Copy Encoded Number, Paste Number to be decoded, Avail Encoded Number for Decoding»

Options: «Change Security Key, View word Length»

Help: «Begin with selecting encryption method and setting security key. Type word in encoding box, press blue arrow, and watch numbers appear. Enter or copy/paste adequately formatted set of numbers in decoding box; activate decoding arrow to obtain original word. Use copy/paste option when switching between multiple instances of the application».

As an example, this program can encrypt Galilei’s message with **six** numbers if the anagram is divided on three blocks: (59059963533, 2868281597381061420) → Altissimum pl; (12633704788, 2733493453175537856) → anetam tergem; (52805510514, 2814866773854152765) → inum observavi, and demonstrate that Kepler’s decipher of Galilei’s message contains **six** different numbers: (19640950549, 2545897308837069827) → Salve umbesti; (37202717155, 2884593943602529808) → nium geminatum; (44243906929, 286854222983839786) → martia prole.

One-time pads

A major challenge of modern cryptography according to, e.g., Ref. (3) is the possibility of cooperation between an adversary and a lawful user, which “makes cryptography powerless”. The following section helps clear this hurdle by excluding anyone except the sender and recipient from understanding their messages.³

Some way to reduce the risk posed by foul play is to encrypt the communication based on principles of advanced mathematics, which at least limits the circle of possible adversaries to highly educated mathematicians. The widespread encryption methods of AES (the advanced encryption standard of NIST) and RSA (Rivest, Shamir, and Adleman, Patent No. 4,405,829) apply obscure concepts in sophisticated mathematical constructions, substitution-permutation procedures, properties of primes, or Fermat’s and Euler’s theorems (RSA), and Galois’ finite field theory (AES).

While the latest methods relying on elliptic curves narrow the circle of adversaries, their approach still does not satisfy Shannon’s criterion for unbreakability, as only the one-time pad (OTP) method does. G.S. Vernam received Patent No. 1,310,719 for an early version of the OTP method with a pre-shared key text at the beginning of 20th century. The method was used to send clandestine telegraph messages combining each plain character on a ticker tape with a random character that was typed on the corresponding row of the same tape. However, a duplicate of this key tape with the random characters had to be delivered to the recipient in advance.

The basic principle of the dictionary and anagram methods of ciphering n -blocks by serial numbers eliminates the requirement for systematic preliminary delivery of random elements. The numbering principle can be compared to communicating by pronouncing only word numbers in the dictionary instead of words themselves. It is as though the recipient already has every imaginable version of the n -numbered plain text and need only get the number of the correct version to read,⁵ in contrast to garbling method. There is no difference between the smallest (equal to 1) deviation from the correct block number and any other deviation since two adjacent blocks can differ drastically, while two remote blocks can have much in common (just like words in dictionary).

Thus, the one-time pad can be used in concert with the numbering methods of encryption (dictionary and anagram). Instead of replacing a plain character in a row of ticker tape with a random character in the corresponding row, the serial number of a block can be replaced with a modified serial number and sent to the recipient. The manner of modification (for example, adding 1 to a certain digit in the serial number) should be known to the recipient in advance. The recipient could then reverse the transformation (in our example, by subtracting 1) and use the correct serial number for decoding. In Fig.1, if we use $G = 55895275495; K = 2880934958582915600$ instead of $G = 55885275495; K = 2880934958582915600$, we obtain OHID ZVBTYTYVY, but not SEND STATUTE. This opens unrestricted possibilities for agreement between the sender and receiver how to make cipher unbreakable: for example, by sending OHID ZVBTYTYVY instead of SEND STATUTE, etc.

A random number may be incorporated in the serial number, making the serial number also random. The change of random number even by 1 destroy all information of the n -block. Then the communication is carried out by exchange of random numbers, new in each cycle of exchange as in the case of one-time pads, satisfying Shannon’s criterion for an unbreakable cypher.

For example, the sender and recipient may agree in advance not only on the position of a random number in the block, but also on a rule for the preliminary alteration of the plain text and serial numbers that depends on this random number, which is present in the text. For instance, they can agree (like in Vernan’s procedure) to replace the numbers of all letters in an n -block by the sum (mod 27) of each letter number and its distance from random position. Then the n -block

number to be sent is calculated for these modified letter numbers, and the recipient restores the letters by applying modular subtraction to the decoded letter numbers. Etc.

Illustration of randomization procedure

To demonstrate the randomization procedure, we first propose a much simpler numbering technique than the dictionary and anagram methods to concentrate attention on the use of random numbers. The “**division method**,” based on the Euclidean algorithm, is available even to those who are not familiar with binomial coefficients.

Let us introduce a set of polynomials $E_n(k_1, k_2, \dots, k_n)$ of n variables:

$$\begin{aligned} E_0 &= 1, E_1(k_1) = k_1, E_2(k_1, k_2) \\ &= k_1 k_2 + 1, E_3(k_1, k_2, k_3) \\ &= k_1 k_2 k_3 + k_1 + k_3, E_4(k_1, k_2, k_3, k_4) \\ &= k_1 k_2 k_3 k_4 + k_1 k_2 + k_1 k_4 + k_3 k_4 + 1, \dots, E_n \\ &= k_n E_{n-1} + E_{n-2}, \dots \text{ and denote} \\ \alpha &= E_{n-1}(k_1, \dots, k_{n-1}), \beta = E_n(k_1, \dots, k_n) \end{aligned} \quad (24)$$

The coefficients of the Diophantine linear equations are recognizable in polynomials E_n .⁷

Integers k_n, \dots, k_2, k_1 can be used for numbering letters in n -block. If the recipient has α and β , he can restore the whole chain of k_n, \dots, k_2, k_1 by means of the well-known Euclidean algorithm used for finding the greatest common divisor of α and β . The Euclidean algorithm consists of subsequent **divisions with a remainder** (quotient q and remainder r), which we denote as: $\beta/\alpha=q|r$. All intermediate quotients are k_n, \dots, k_2, k_1 . This procedure is named the “division method”.

The already used example demonstrates this procedure. Before sending “IT IS IT”= 10,21,1,10,20,1,10,21, we calculate $\alpha=E_7(10,21,1,10,20,1,10)=559261$ and $\beta=E_8(10,21,1,10,20,1,10,21)=11795543$, which are our message: α, β . The recipient carries out the Euclidean algorithm: $11795543/559261=21|51062$, $559261/51062=10|48641$, $51062/48641=1|2421$, $48641/2421=20|221$, $2421/221=10|211$, $221/211=1|10$, $211/10=21|1$, $10/1=10|0$, and reading the quotients from right to left, he receives the message. Acting similarly, the adversary recognizes our secret. How random number can change the situation?

Assume the sender and recipient have memorized a “magic” number 3 and placed a random letter, say F_7 , in the third position. There is no need to preliminarily send this letter, F_7 , to the recipient; he must know only its position and how to use it (as agreed on with the recipient beforehand). Let, e.g., the agreement was to subtract (mod 27) the random letter number from rest of numbers on even places and to add (mod 27) on odd places. Then instead of IT IS IT, Sender should send $10+7=17$, $21-7=14$, 7 , $1+7=8$, $10-7=3$, $20+7=27$, $1-7=-6 \rightarrow 21$, $10+7=17$, $21-7=14$. Thus Recipient will see PMFGBZTPM, but he know that he must perform back transformations: $P_{17} \rightarrow 17-7=10 \rightarrow I_{10}$, $M_{14} \rightarrow 14+7=21 \rightarrow I_{10}$, etc.

Adversary will also see PMFGBZTPM and different combinations at each next interception of the same message because random letter

will be different, and the agreement Sender-Recipient is not available for him. This is something like secret key in practical cryptography. However, it depends not on special strict mathematical theories, but on flight of human fantasy with its unrestricted possibilities (probabilities).

Conclusion

Two formulas are deduced, which endows by the serial number any n -block, built of natural numbers. The first one depends on block letter numbers and their partial sums, the second one depends on block letter numbers and their multiplicities. The specific feature of the first one is that it lets easy to distinguish the false serial number. The second formula can be modified in such a way that it determines also the order of letters in the block. For an anagram, the first serial number is numbering the ciphered anagram, the second serial number decipheres it.

Creating reverse algorithms (i.e., restoration of the block plain text from its serial number) turns this math construction into a method of ciphered communication. By the way, promulgation of the above-described technique in made it public and deprived the secrecy of the restoration algorithm remaining only distortion (randomization) a single mechanism for the secret communication.⁸

Acknowledgments

The author expresses his deep thanks to A. Belinskiy for comments on historic anagrams and current trends in cryptography and fruitful cooperation in this work and also to X. Tashlitsky, Esq. for help in editing the article.

Funding

None.

Conflicts of interest

The author declares that there are no conflicts of interest.

References

- Schneier B. *Applied Cryptography*. New York, NY: Wiley. 1996.
- Koblitz N. *A course in number theory and cryptography*. New York, NY: Springer-Verlag. 1994.
- Rubin F. *Secret key cryptography ciphers: from simple to unbreakable*. New York, NY: Manning Publications Co. 2022.
- Goldwasser S, Bellare M. *Lecture notes on cryptography*. 2008.
- Perelman YI. *Entertaining astronomy*, 9th Ed. Foreign languages publishing house. 1958.
- Mestechkin M. Natural solutions of diophantine linear equation with n unknowns. *J Comp Meth Sci & Eng*. 2016;16(1):175.
- Beyer HW. *CRC Standard mathematical tables*. Boca Raton, FL: CRC Press. 1973.
- Mestechkin M, Mestechkina T. *Unbreakable communication cipher without a preliminary sharing of random elements*. US patent application US 20200092080A1.