GENETIC CODES I: BINARY SUB-ALPHABETS, BI-SYMMETRIC MATRICES AND GOLDEN SECTION

Sergei V. Petoukhov

Physicist, biologist, mathematician (b. Moscow, Russia, 1946); Chairman, Advisory Board of International Symmetry Association (ISA); Honorary chairman, ISA’s “International Society of Symmetry in Bioinformatics”; Vice-president, National Academy of Intellectual and Social Technologies (Russia); Member of the Academy of Quality Problems (Russia). Address: Department of Biomechanics, Mechanical Engineering Research Institute, Russian Academy of Sciences, M. Kharitonyevsky pereulok 4, Moscow, Russia 101990. E-mail: petoukhov@hotmail.com, personal website: http://www.petoukhov.narod.ru.

Fields of interest: biosymmetry, self-organizing systems, geometry, solitons, genetic code, biomechanics.

Awards: State Prize Award, 1986.


INTRODUCTION

This article describes beautiful properties of a special class of mathematical matrices to analyze symmetrical relations among elements of the genetic code system. In particular, these matrices reveal a new unexpected connection between numeric peculiarities of genetic codes and golden section, which is a remarkable value of the history of mathematics, physiology and culture. The next part of these original researches is published in the author’s second article in this issue.

Biological organisms differ from each other in many aspects: in their sizes, colors, types of motions, etc. But molecular genetics revealed a principal unity of all biological organisms from the viewpoint of basic structures of their genetic code. In other words, a great union of all biological organisms happened in modern science in the past century. In addition to this great union it was found that these basic structures proved to be simple enough to make people (or at least the scientific community) astonished. A universal set of basic structures of genetic codes of all biological organisms consists of 4 nitrogenous bases, 64 triplets and 20 amino acids, which are coded by these 64 triplets. A collection of 4 nitrogenous bases is interpreted as a 4-elements “alphabet” ordinarily and includes adenine (A), cytosine (C), guanine (G), uracil (U) {or thymine (T) in DNA}. 
One may assume that such a simple structure of the genetic code has an appropriate, simple and “natural” mathematics of relations between its separate components, a set of which is working as a single whole. Of course, any complex system of nature can be modeled mathematically from very different initial positions. In the author’s opinion, one of the perspective variants for modeling and analyzing the genetic code system (and of a wide class of other biological systems too) is connected with the application of mathematical matrices. This field is very developed in modern mathematics and its applications in physical sciences. Let us begin the article with a description of our special application of mathematical matrices to genetic code.

**Binary sub-alphabets of genetic code and bi-periodic matrices.** The given set of four letters A, C, G, U is usually considered as an elementary alphabet of the genetic code. The modern science does not know why the alphabet of genetic language has four letters (it could have any other number of the letters in principle) and why just these four nitrogenous bases are chosen by nature as elements of the genetic alphabet from billions possible chemical compounds.

The author paid attention to the fact that these four nitrogenous bases represent specific poly-nuclear constructions with the special biochemical properties. The set of these four structures has in itself a substantial system of binary-opposite attributes. The system of such attributes divides the four-letter alphabet into various pairs of letters, which are equivalent from a viewpoint of one of these attributes or its absence: 1) C=U and A=G (according to binary-opposite attributes: “pyrimidine” or “non-pyrimidine”, that is purine); 2) A=C and G=U (according to attributes: amino-mutating or non-amino-mutating under action of nitrous acid HNO₂ (see Wittmann 1961); 3) C=G and A=U (according to attributes: three or two hydrogen bonds are materialized in these complementary pairs).

Let us denote these binary-opposite attributes by numbers \(N = 1, 2, 3\) and let’s ascribe to each of the four letters of this alphabet the symbol “1ₙ” in that case when a letter has an attribute under number “\(N\)”, and the symbol “0ₙ” in the opposite case, when a letter hasn’t such attribute. In result we receive a representation of the four-letter alphabet of a code in system of its three “binary sub-alphabets to attributes” (see Table 1). The four-letter alphabet of a code is curtailed into the two-letter alphabet on the basis of each kind of attributes. For example, to the first kind of binary-opposite attributes we have (instead of the four-letter alphabet) the alphabet from two letters \(0₁\) and \(1₁\), which the author names as “the binary sub-alphabet to the first kind of binary attributes”.

---

**Table 1**

<table>
<thead>
<tr>
<th>Letter</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1₁</td>
</tr>
<tr>
<td>C</td>
<td>0₁</td>
</tr>
<tr>
<td>G</td>
<td>0₂</td>
</tr>
<tr>
<td>U</td>
<td>1₂</td>
</tr>
</tbody>
</table>
Bi-periodic tables of genetic codes. Let us consider a matrix $P$ of second order (2x2) with four elements C, A, G, U. The matrix $P$ is printed in bold in Table 2, left. Each of its columns consists only of the elements which are equivalent to each other from the viewpoint of the binary sub-alphabet $\#1$ (see Table 1). More precisely, both elements of the first column are pyrimidines (C and U) and have the same symbol 1 in this binary sub-alphabet. Both elements of the second column are purines (A and G) and have the symbol 0 in the binary sub-alphabet $\#1$. So both columns of the matrix $P$ are numerated naturally by symbols 1 and 0 from the viewpoint of the binary sub-alphabet $\#1$ connected with the pair of binary-opposite attributes “pyrimidines-purines”.

Concerning the numeration of rows of the matrix $P$, each its row consists of those letters only, which are equivalent to each other from the viewpoint of the binary sub-alphabet $\#2$. More precisely, both elements C and A of the first row have amino-mutating property and so they have the symbol 1 in the binary sub-alphabet $\#2$. Both elements U and G of the second row have not amino-mutating property and they are symbolized by 0 in the binary sub-alphabet $\#2$. So both rows of the matrix $P$ are numerated by
symbols 1 and 0 naturally from the viewpoint of the binary sub-alphabet № 2. (Of course, one can use symbols 1, 0, 1, 0; 0, 1, 0, 0; used by us in this tabular case).

The matrix \( P \) is the simplest representative (specimen) of a set of bi-periodic matrices (or tables) of the genetic code system. It has a vertical periodicity of the matrix elements from the viewpoint of the binary sub-alphabet № 1 and it has a horizontal periodicity of the matrix elements from the viewpoint of the binary sub-alphabet № 2.

Let us remember the well-known operation of tensor (or Kronecker) multiplication of matrices, which is applied in mathematics and physics widely (for example, see Gazale 2000). If the matrix \( P \) is raised to the power 2 in the sense of tensor multiplication, a result is the \((4x4)\) matrix of 16 genetic duplets from the table 2 (right). If \( P \) is raised to the tensor power 3, a result is the \((8x8)\) matrix of 64 triplets from the table 3 (right). A symbol of tensor multiplication is “\( \otimes \)”. A symbol \( P^{(n)} \) means, that a matrix \( P \) is raised to the power “\( n \)” in a sense of tensor multiplication. A matrix \( P^{(n)} \) consists of \( n \)-plets as its matrix elements, where \( n \) can be equal to arbitrary positive integer number. \( P^{(n)} \) has an order \((2^n x 2^n)\). All matrices \( P^{(n)} \) will be named conditionally as genetic matrices because of their connections with elements of the genetic code.

A binary numeration of columns and rows of the matrix \( P^{(n)} \) is connected with binary symbols of letters C, A, G, U in the binary sub-alphabets №№ 1 and 2 respectively. More precisely, to get a numeration number of a column of the matrix \( P^{(n)} \), one should take a sequence of letters of any \( n \)-plet from this column and write a corresponding sequence of binary symbols for these letters from the viewpoint of the binary sub-alphabet № 1. This binary sequence is a numeration number of this column automatically. For example, let’s consider the matrix \( P^{(3)} \) and its column with a triplet CAU. From the viewpoint of the binary sub-alphabet № 1 (where C=U=1 and A=0), the sequence of letters CAU is equivalent to a binary sequence 101. This binary number 101 is a numeration number of the column.

A binary number of a row of the matrix \( P^{(n)} \) is constructed in a similar algorithmic way by interpretation of any \( n \)-plet of this row from the viewpoint of the binary sub-alphabet № 2. For example, let’s consider the same triplet CAU in the matrix \( P^{(3)} \). From the viewpoint of the binary sub-alphabet № 2 (where C=A=1 and U=0), the sequence of letters CAU is equivalent to a binary sequence 110. This binary number 110 is a numeration number of the matrix row with the triplet CAU.
It can be checked easily that all matrices $P^{(n)}$ are bi-periodic matrices. Actually any column of such a matrix consists of only the n-plets which are equivalent to each other from the viewpoint of binary sub-alphabet № 1. And any row of a matrix $P^{(n)}$ consists of those n-plets only, which are equivalent to each other from the viewpoint of binary sub-alphabet № 2.

A matrix $P^{(n)}$ has a binary coordination number for each of its n-plet. All sets of its n-plets have a series of binary coordination numbers algorithmically, which are equivalent to a series of integers $0, 1, 2, ..., (4^n - 1)$ in decimal numeration system. Such a coordination number of a concrete n-plet is constructed by combination of binary numbers of a row and a column in a single whole in a form of $2n$-digit binary number. The first half of such a coordination number of any n-plet coincides with a binary numeration number of its matrix row, and its second half coincides with a binary numeration number of its column. For example, the considered triplet CAU has an individual coordination number 110101 in the matrix $P^{(3)}$. At translation of such $2n$-digit binary numbers into a decimal numeration system, all n-plets receive their individual decimal numbers from the mentioned series of integers $0, 1, 2, ..., (4^n - 1)$. Correspondingly all n-plets reveal their natural ordering connected with such numerations. For example, all triplets in a matrix $P^{(3)}$ have their natural ordering connected with their numerations by coordination numbers from a series of integers $0, 1, 2, ..., 63$. Such natural numerations and ordering of n-plets are useful for investigation of rules of symmetric relations among elements of various genetic codes (see a special rule of a disposition of stop-codons in genetic codes below). In this way we obtain an important opportunity to work with numbers in genetic codes. In other words, we digitize genetic codes by analogy with a digitizing of information in computers.

The family of matrices $P^{(n)}$ is important because all possible genetic sequences of arbitrary length belong to its matrices. A sub-family of matrices $P^{(3n)}$ is interesting especially because it contains all possible genetic sequences of triplets in a form of $3n$-plets. A set of such sequences has $3n$-plets with real coding sense (they code proteins) and without it (they do not code proteins and are a “garbage” or genetic materials with unclear meaning). For example, if a protein is coded by a sequence of 999 nitrogenous bases (333 triplets), this sequence is located in a family matrix $P^{(999)}$. Special tasks are generated from this matrix approach for future investigations in the field of molecular genetics (Petoukhov, 2003c, 2004).

A disposition of amino acids in the bi-periodic octet table of triplets. Till this moment we spoke about nitrogenous bases and triplets. But 64 triplets encode 20 amino acids and termination signals.
It seems improbable, that 20 amino acids could be arranged automatically by symmetrical order into 64 cells of bi-periodic matrix $P^{3(3)}$. One has at least two important reasons to think so. Firstly, a great set of intermediate biochemical agents (ferments, nucleic acids, etc.) and of processes exist between DNA’s sequence of triplets and a final action of assembling of different amino acids into protein chain. Cooperative synchronous work of these numerous agents is unclear for modern science in many aspects. Nobody can deduce an order of 20 amino acids in the matrix $P^{3(3)}$ from modern data about a mutual work of these biochemical agents. The second reason is that the matrix $P^{3(3)}$ for 64 triplets was constructed by the author for triplets only without any initial information about amino acids. This octet matrix of triplets was constructed in this article by absolutely formal manner in a form of a third tensor power of the matrix $P$, which was built by a formal way also from four nitrogenous bases C, A, G, U.

<table>
<thead>
<tr>
<th>111</th>
<th>110</th>
<th>101</th>
<th>100</th>
<th>011</th>
<th>010</th>
<th>001</th>
<th>000</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCC</td>
<td>CCA</td>
<td>ACC</td>
<td>ACA</td>
<td>AAC</td>
<td>AAT</td>
<td>AAU</td>
<td>AAA</td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td>Thr</td>
<td>Thr</td>
<td>Asn</td>
<td>Lys</td>
<td>Asn</td>
<td>Lys</td>
</tr>
<tr>
<td>63</td>
<td>62</td>
<td>59</td>
<td>58</td>
<td>57</td>
<td>56</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>CCG</td>
<td>GAC</td>
<td>CGC</td>
<td>GCG</td>
<td>GCC</td>
<td>GCA</td>
<td>GCA</td>
<td>GAA</td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td>Arg</td>
<td>Arg</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>54</td>
<td>55</td>
<td>44</td>
<td>45</td>
<td>27</td>
<td>26</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>CUA</td>
<td>CUU</td>
<td>AUU</td>
<td>UUC</td>
<td>GCC</td>
<td>GCA</td>
<td>GAC</td>
<td>GGA</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ile</td>
<td>Phe</td>
<td>Val</td>
<td>Ala</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td>46</td>
<td>39</td>
<td>35</td>
<td>31</td>
<td>19</td>
<td>26</td>
<td>25</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 3. An enriched representation of the bi-periodic octet matrix $P^{3(3)} = P \otimes P \otimes P$ for a case of the vertebrate mitochondria genetic code. The matrix consists of binary numeration numbers of its columns and rows, 64 triplets with their coordination numbers from 0 up to 63 in decimal system, and 20 amino acids with their traditional abbreviations Pro, His, Gln, etc. Stop-codonts are marked as “stop”. Each row has two pairs of neighbor cells with repeated amino acids (these pairs are marked by dark color).
However, if a distribution of 20 amino acids in the octet matrix $P^{(3)}$ realizes a symmetrical order suddenly, such improbable fact could be serious argument for the benefit of such matrix mathematical approach to investigate structures of the genetic system. And such improbable fact of symmetry comes true really.

Table 3 demonstrates the matrix $P^{(3)}$ with 64 triplets in enriched representation. Each matrix cell has an appropriate triplet with its coordination number in decimal system. Beside all, each matrix cell has the amino acid (or a symbol “stop” for the case of stop-codons), which is coded by this triplet in the vertebrate mitochondria genetic code. As it is described with appropriate arguments in genetic literature, this vertebrate mitochondria genetic code is the most ancient and “perfect” variant of the genetic code (for example, see Frank-Kamenetskii 1988, pp. 65-68). The author begins with this variant of genetic code all his structural genetic investigations as a rule.

Table 3 demonstrates the octet matrix $P^{(3)}$ with very symmetric disposition of 20 amino acids and stop-codons. One can mark four kinds of symmetries easily, which are reflected phenomenological properties of genetic code:

1. This matrix $P^{(3)}$ consists of four pairs of neighbor rows with even and odd numeration numbers in each pair: 1-2, 3-4, 5-6, 7-8. The rows of each pair are equivalent to each other from the viewpoint of a disposition of the same amino acids in their appropriate cells.

2. Each matrix row has two pairs of neighbor cells with repeated amino acids (these pairs are marked by dark color) and two pairs of neighbor cells with different amino acids (these pairs are marked by white color). So, we have a black-and-white mosaic for amino acids and stop-codons. Quantities of black and white cells are equal to each other and are equal to 32.

3. The left and right halves of the matrix mosaic are mirror-anti-symmetric to each other in its colors: any pair of cells, disposed by mirror-symmetrical manner in these halves, has opposite colors.

4. The black-white matrix mosaic, determined by the disposition of 20 amino acids and stop-codons, has a symmetric figure of a diagonal cross: diagonal quadrants of the matrix are equivalent to each other from the viewpoint of their mosaic.

Strongly pronounced symmetric disposition of amino acids in our bi-periodic matrix $P^{(3)}$ of triplets testifies, that purely phenomenological facts of encoding of amino acids by triplets have a regular and adequate matrix presentation, which is connected with described binary sub-alphabets. It demands a prolongation of such matrix researches of genetic structures.
About peculiarities of the genetic matrix $P^{(3)}$. The matrix $P^{(3)}$ has many interesting peculiarities. For example, each matrix quadrant with order (4x4) consists of all 16 triplets, which are begun with the same letter. In each quadrant, each of its sub-quadrant with order (2x2) consists of all 4 triplets, which have identical letters in their second position. This list of symmetrical peculiarities of the bi-periodic table of triplets can be continued (increased, enlarged) (Petoukhov 2001-2004). One of its interesting peculiarities is connected with the matrix $P_{\text{COORD}}^{(3)}$ of Table 4, which demonstrates a set of coordination numbers of triplets in the matrix $P^{(3)}$. The binary numeric matrix $P_{\text{COORD}}^{(3)}$ is realized when all 64 triplets in the matrix $P^{(3)}$ are changed by their binary coordination numbers. The index “coord” marks in matrices $P_{\text{COORD}}^{(n)}$, that we consider matrices of binary coordination numbers of n-plets instead of matrices $P^{(n)}$ itself.

It should be noted specially that the matrix $P_{\text{COORD}}^{(3)}$ coincides with the historically famous table of 64 hexagrams according with Fu-Xi’s order from the Ancient Chinese “The Book of Changes”, which was written several thousand years ago (Petoukhov 1999, 2001). A small difference is that the Ancient Chinese table uses binary symbols of broken and unbroken lines instead of modern binary symbols 0 and 1. The matrix $P_{\text{COORD}}^{(6)}$ is produced automatically when the matrix $P_{\text{COORD}}^{(1)}$ is raised to the power “n” in a sense of tensor multiplication. In particular, this tensor connection between $P_{\text{COORD}}^{(1)}$ and $P_{\text{COORD}}^{(3)}$ is useful to analyze a hidden connection between basic elements of the symbolic system of “The Book of Changes”, but results, obtained by the author in this historical field are out of this article. The table of 64 hexagrams is used in Oriental medicine (acupuncture, pulse-diagnostics of Tibetan medicine, etc.) as a basic structural system in connection with knowledge about chronocycles of biological organisms from the ancient time (see the second author’s article in this issue about a chronocyclic theory in the field of genetic codes).

<table>
<thead>
<tr>
<th></th>
<th>111</th>
<th>110</th>
<th>101</th>
<th>100</th>
<th>011</th>
<th>010</th>
<th>001</th>
<th>000</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>111111</td>
<td>111110</td>
<td>111101</td>
<td>111100</td>
<td>111011</td>
<td>111010</td>
<td>111001</td>
<td>111000</td>
</tr>
<tr>
<td>110</td>
<td>110111</td>
<td>110110</td>
<td>110101</td>
<td>110100</td>
<td>110011</td>
<td>110010</td>
<td>110001</td>
<td>110000</td>
</tr>
<tr>
<td>101</td>
<td>101111</td>
<td>101110</td>
<td>101101</td>
<td>101100</td>
<td>101011</td>
<td>101010</td>
<td>101001</td>
<td>101000</td>
</tr>
<tr>
<td>100</td>
<td>100111</td>
<td>100110</td>
<td>100101</td>
<td>100100</td>
<td>100011</td>
<td>100010</td>
<td>100001</td>
<td>100000</td>
</tr>
<tr>
<td>011</td>
<td>011111</td>
<td>011110</td>
<td>011101</td>
<td>011100</td>
<td>011011</td>
<td>011010</td>
<td>011001</td>
<td>011000</td>
</tr>
<tr>
<td>010</td>
<td>010111</td>
<td>010110</td>
<td>010101</td>
<td>010100</td>
<td>010011</td>
<td>010010</td>
<td>010001</td>
<td>010000</td>
</tr>
<tr>
<td>001</td>
<td>001111</td>
<td>001110</td>
<td>001101</td>
<td>001100</td>
<td>001011</td>
<td>001010</td>
<td>001001</td>
<td>001000</td>
</tr>
<tr>
<td>000</td>
<td>000111</td>
<td>000110</td>
<td>000101</td>
<td>000100</td>
<td>000011</td>
<td>000010</td>
<td>000001</td>
<td>000000</td>
</tr>
</tbody>
</table>

Table 4: The matrix $P_{\text{COORD}}^{(3)}$, where binary coordination numbers of 64 triplets are presented, coincides with the famous table of 64 hexagrams according with Fu-Xi’s order from the Ancient Chinese “The Book of Changes”.
Multiplicative numeric matrices of genetic systems. Till this moment, we have used two first kinds of binary-opposite attributes of Table 1 for the construction of a family of matrices \( P^{(i)} \). Now we will pay attention to the third kind of described attributes connected with hydrogen bonds. More precisely, letters C and G have 3 hydrogen bonds (C=G=3) and letters A and U have 2 hydrogen bonds (A=U=2). Let us replace each n-plet in any matrix \( P^{(i)} \) by the product of these numbers of its hydrogen bonds. In this case, we get multiplicative matrices marked as \( P_{\text{MULT}}^{(i)} \) conditionally (another family of additive matrices was considered in works (Petoukhov 1999, 2001, 2003c)). For example, the triplet CAU will be replaced by number \( 12 = (3*2*2) \) in the matrix \( P_{\text{MULT}}^{(3)} \). Table 5 demonstrates the multiplicative matrix \( P_{\text{MULT}}^{(3)} \) constructed in this way.

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>18</td>
<td>18</td>
<td>12</td>
<td>18</td>
<td>12</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>18</td>
<td>27</td>
<td>12</td>
<td>18</td>
<td>12</td>
<td>18</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>27</td>
<td>18</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>18</td>
<td>12</td>
<td>8</td>
<td>27</td>
<td>18</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>8</td>
<td>12</td>
<td>18</td>
<td>27</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>12</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>12</td>
<td>125</td>
</tr>
<tr>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>1000</td>
</tr>
</tbody>
</table>

Table 5: The multiplicative matrix \( P_{\text{MULT}}^{(3)} \) with cells, which contain products of numbers of complementary hydrogen bonds for triplets: C=G=3, A=U=2. The right column shows sums of numbers of each row. The lower row shows sums of numbers of each column. Bold frames mark diagonal cells.

All matrices \( P_{\text{MULT}}^{(i)} \) are symmetrical relative to both diagonals and they can be named “bi-symmetric matrices”. All rows and all columns of this matrix are differentiated from each other by the sequences of their numbers. But the sums of all numbers in the cells of each row and of each column in a matrix \( P_{\text{MULT}}^{(i)} \) are identical to each other. In the case of the matrix \( P_{\text{MULT}}^{(3)} \), these sums are equal to 125 = 5^3 and the total sum of tabular numbers is equal to 1000. A rank of this matrix is equal to 8. Its determinant is equal to 5^{12}. Eigenvalues of \( P_{\text{MULT}}^{(3)} \) are 1, 5, 5, 5, 5, 5, 5, 5. The matrix \( P_{\text{MULT}}^{(3)} \) has four kinds of numbers only: 8, 12, 18 and 27. The certain laws are observed in their disposition, which are connected with a few interesting properties of this matrix, including the property of invariance of its numeric mosaic under many mathematical matrix operations with this matrix (Petoukhov 2003b, 2003c, 2004).
It was much unexpected for the author to discover, that the family of matrices $P_{\text{MULT}}^{(n)}$ has a close connection with the famous golden section.

**Genetic matrices and the golden section.** In biology, a genetic system provides the self-reproduction of biological organisms in their generations. In mathematics, the so-called “golden section” (or “divine proportions”) and its properties were a mathematical symbol of self-reproduction from the Renaissance (for example, see the website “Museum of Harmony and Golden Section” by A. Stakhov, www.goldenmuseum.com). Is there any connection between these two systems? Yes, a research made by the author has found out such connection.

A golden section is a value $\phi = (1+5^{0.5})/2 = 1, 618…$ (Sometimes the inverse of this value is called a golden section in literature). If the simplest genetic matrix $P_{\text{MULT}}^{(1)}$ is raised to the power $1/2$ in ordinary sense (that is, if we take the square root), the result is a bi-symmetric matrix $\Phi = (P_{\text{MULT}}^{(1)})^{1/2}$, the matrix elements of which are equal to the golden section and to its inverse value. Table 6 demonstrates the matrices $P_{\text{MULT}}^{(1)} = \Phi^2$, $(P_{\text{MULT}}^{(1)})^{1/2} = \Phi$, $(P_{\text{MULT}}^{(2)})^{1/2} = \Phi \otimes \Phi = \Phi^{(2)}$.

<table>
<thead>
<tr>
<th>$P_{\text{MULT}}$</th>
<th>$\Phi$</th>
<th>$(P_{\text{MULT}}^{(2)})^{1/2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 2 1</td>
<td>$\phi$</td>
<td>$\phi^2 (\phi/\phi)^2$</td>
</tr>
<tr>
<td>2 3 1</td>
<td>$\phi^{-1}$</td>
<td>$\phi^{-2} (\phi/\phi)^2$</td>
</tr>
</tbody>
</table>

**Table 6:** The beginning of a family of golden genetic matrices $(P_{\text{MULT}}^{(n)})^{1/2} = \Phi^{n}$, where “$\phi$” is the golden section.

If the multiplicative genetic matrix $P_{\text{MULT}}^{(n)}$ is raised to the power $1/2$, the result is a matrix $\Phi^{(n)} = (P_{\text{MULT}}^{(n)})^{1/2}$, the matrix elements of which are constructed from a combination of $\phi$ and $\phi^{-1}$ by the following algorithm. We take a corresponding $n$-plet of the matrix $P^{(n)}$ and change its letters C and G to $\phi$. Then we take letters A and U in this n-plet and change each of them to $\phi^{-1}$. As a result, we obtain a chain with “n” links, where each link is $\phi$ or $\phi^{-1}$. The product of all such links gives the value of corresponding matrix elements in the matrix $\Phi^{(n)}$. For example, in the case of the matrix $\Phi^{(3)} = (P_{\text{MULT}}^{(3)})^{1/2}$, let us calculate a matrix element, which is disposed at the same place as the triplet CAU in the matrix $P^{(3)}$. According to the described algorithm, one should change the letter C to $\phi$ and the letters A and U to $\phi^{-1}$. In the considered example, we obtain the following product: $(\phi * \phi^{-1} * \phi^{-1}) = \phi^{-1}$. This is the desired value of the considered matrix element for the matrix $\Phi^{(3)}$. Table 7 demonstrates the matrix $\Phi^{(3)} = (P_{\text{MULT}}^{(3)})^{1/2}$ itself.
The matrix elements of the bi-symmetric matrices \( \Phi^{(n)} = (P_{\text{MULT}}^{(n)})^{1/2} \) consist of a single number in different powers only: golden section \( \varphi \), which is raised to powers taken from a symmetrical set of positive and negative integers. The golden matrix \( \Phi^{(3)} \) has only four kinds of numbers, generated from a single value of golden section: \( \varphi^1, \varphi^{-1} \) and \( \varphi^{-3} \). They have the same disposition (and the same numeric mosaic) in the octet matrix as four kinds of numbers 27, 18, 12 and 8 in the initial matrix \( P_{\text{MULT}}^{(3)} \).

Table 7: The golden octet genetic matrix \( \Phi^{(3)} = (P_{\text{MULT}}^{(3)})^{1/2} \). Here \( \varphi = 0.5(1+5^{1/2}) = 1.618… \) is a golden section.

The author has given a principal new definition to a golden section on the basis of the matrix specifics of genetic code systems: a golden section \( \varphi \) and its inverse value \( \varphi^{-1} \) are single matrix elements of the bi-symmetric matrix \( \Phi \), which is the square root from a bi-symmetric matrix \( P_{\text{MULT}} \) (2x2) with its numeric matrix elements from a set of numbers of complementary hydrogen bonds in genetic nitrogenous bases: C=G=3, A=U=2. This definition does not use conceptions of line segments, their proportions, etc., which are traditional for the definition of golden section. This new definition is based on genetic matrix peculiarities. In our opinion, many realizations of the golden section in the nature are connected with its matrix essence and with its matrix representation. It should be investigated specially and systematically, where in natural systems and phenomena we have the bi-symmetric matrix \( P_{\text{MULT}} \) with its matrix elements 3 and 2 in a direct or masked form (for example, in a form of pairs of numbers 6 and 4, or 9 and 6, or 12 and 8, etc. with the same proportion 3:2, which is so frequent for ratios of elements in genetic codes, as Petoukhov’s second article in this issue describes). One can hope to discover many new system phenomena and connections between them in the nature in this way.

The new theme of golden section in genetic matrices seems to be very important because many physiological systems and processes are connected with it. It is known that proportions of a golden section characterize many physiological processes: cardiovascular processes, respiratory processes, electric activities of brain, locomotion activity, etc. The author hopes that the algebra of bi-symmetric genetic matrices,
proposed by him in connection with a theme of golden section, will be useful for explanation and the numeric forecast of separate parameters in different physiological sub-systems of biological organisms with their cooperative essence and golden section phenomena. In our opinion, structures of all physiological systems have been developed by nature in evolutionary agreement with basic structures of genetic code for their evolutionary surviving by re-creation in next generations. The idea of golden matrices can also be useful for non-biological phenomena.

Matrices with matrix elements, all of which are equal to golden section \( \phi \) in different powers only, can be referred to as “golden matrices”. Only some of them have a direct connection with genetic code system and can be referred to as “genetic golden matrices”. Questions and peculiarities of golden matrices form a new mathematical branch with interesting applications. Many results, obtained by the author in this new field, cannot be described in this restricted article; they will be published later.

For a few examples only, let us mark that recurring transformation of a vector by golden matrices, which are modified by a simple way specially, produces interesting spirals (“spirals of golden matrices”), which can be used for the modeling of numerous spirals in biological bodies and, in particular, for the modeling of famous phyllotaxis laws of morphogenesis. Secondly, golden matrix \( \Phi \) with second order \([2x2]\) has a recurrent formula (which is an analogy of a famous recurrent formula of golden section: \( \phi^k - \phi^{k-1} = \phi^{k-2} \)):

\[
\Phi^K - \Phi^{K-1} = (\phi^2 - \phi^{-2})^{K-1} \phi^{-1} \cdot [1 1; 1 1] = (\phi + \phi^{-1})^{K-1} \cdot [\phi^{-1} \phi^{-1}; \phi^{-1} \phi^{-1}] 
\]

(1)

We have also deduced the following recurring equations for these matrices:

\[
\Phi^K - 2\phi \cdot \Phi^{K-1} = (\phi - \phi^{-1}) \cdot \Phi^{K-2} \quad \text{or} \quad \Phi^K + (\phi + \phi^{-1}) \cdot \Phi^{K-2} = 2\phi \cdot \Phi^{K-1}
\]

(2)

The considered series of golden matrices \( \Phi^0, \Phi^1, \Phi^2, \Phi^3, \ldots \Phi^K, \ldots \) has for its initial members the following numeric appearance:

<table>
<thead>
<tr>
<th>( K=0 )</th>
<th>( K=1 )</th>
<th>( K=2 )</th>
<th>( K=3 )</th>
<th>( K=4 )</th>
<th>( K=5 )</th>
<th>( K=6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0</td>
<td>1.618; 0.618</td>
<td>3; 2</td>
<td>6.090; 5.090</td>
<td>13; 12</td>
<td>28.450; 27.450</td>
<td>63; 62</td>
</tr>
<tr>
<td>0 1</td>
<td>0.618; 1.618</td>
<td>;</td>
<td>2; 3</td>
<td>;</td>
<td>5.090; 6.090</td>
<td>;</td>
</tr>
</tbody>
</table>

The classical numeric series of a golden section consists of irrationals only. The considered series of golden matrices is richer by their numeric properties. Each of its members with an even number is a matrix with integer elements and each of its members with an odd number is the matrix with irrational numeric elements (a series of matrices of Yin-Yang type). Each column and each row of a golden matrix \( \Phi^{(0)} \) has an
equal sum of its numbers, which is equal to \((5^0)^{0.5}\). Each column and each row of a golden matrix \(\Phi^{(0)}\) has an equal product of its numbers, which is equal to 1.

But let us return to bi-symmetric genetic matrices.

**Mosaic-invariant property of bi-symmetric genetic matrices.** The genetic matrices \(P_{\text{MULT}}^{(3)}\) (and \(\Phi^{(0)}\) also) have unexpected group-invariant property, which is connected with multiplications of matrices (Petoukhov 2003b, 2003c, 2004). We will explain this property through the example of the matrix \(P_{\text{MULT}}^{(3)}\) from Table 5. This matrix consists of four numbers: 8, 12, 18 and 27 only with their special disposition. The numbers 8 and 27 are disposed at matrix diagonals separately in the form of a diagonal cross. The numbers 12 are disposed in matrix cells, a set of which produces a special mosaic (such mosaic can be referred to as “symbol 69” conditionally). The numbers 18 are disposed in matrix cells, a set of which produces a mirror-symmetrical mosaic in comparison with a 69-mosaic of previous case. Table 8 demonstrates these two cases by means of a set of dark matrix cells with numbers 12 (left matrix) and with numbers 18 (right matrix).

![Table 8: A mosaic of cells with number 12 (left, the cells marked by dark) and a mosaic of cells with number 18 (right) from the multiplicative matrix \(P_{\text{MULT}}^{(3)}\) of Table 5.](image)

It is known that if an arbitrary octet matrix with four kinds of numbers as its matrix elements is raised to the power of \(n\), the resulting matrix will have usually much more kinds of numbers with very different disposition (up to 64 kinds of numbers for 64 matrix cells). But our bi-symmetrical genetic matrices have the unexpectedly wonderful property of invariance of their numeric mosaic after such operation of raising to the power of \(n\). For example, if the octet matrix \(P_{\text{MULT}}^{(3)}\) of Table 5 is raised to the power of 2, the resulting octet matrix \((P_{\text{MULT}}^{(3)})^2\) will have a new set of four numbers 2197, 2028,
It is essential that this beautiful property of invariance of the numeric mosaic of the genetic matrix is independent of values of numbers. This property is realized for such matrices with the arbitrary set of four numbers a, b, c, d, if they are located in the same manner inside a matrix. Moreover, if we have one matrix $X$ with a set of four numbers $a$, $b$, $c$, $d$ and another matrix $Y$ with another set of four numbers $k$, $m$, $p$, $q$, then the product of these matrices will be the matrix $Z=X*Y$ with a set of new four numbers $r$, $g$, $v$, $z$ and with the same mosaic of their disposition (Table 9):

$$X = \begin{pmatrix} a & b & c & d \\ e & f & g & h \end{pmatrix} \quad Y = \begin{pmatrix} i & j & k & l \\ m & n & o & p \end{pmatrix}$$

$$Z = X*Y = \begin{pmatrix} r & g & v & z \\ s & t & u & w \end{pmatrix}$$

It is obvious that the four symbols (for example, $a$, $b$, $c$, $d$) in such matrices can be not only ordinary numbers, but also arbitrary mathematical objects: complex numbers, matrices, functions of time (for example, it can be that $a=R*\cos wt$, $b=T*\sin wt$, …), etc. In particular, the possibility of the modeling of chronocyclic functions by means of such mosaic-invariance matrices is useful for the chronocyclic theory of degeneracy of genetic codes, which is described in the next article in this issue. Such a mosaic-invariant property of these genetic matrices is an expression of cooperative behavior of its elements, but not the result of the individual behavior of each kind of elements. This property reminds many aspects of the cooperative behavior of the elements of biological organisms.

An inverse matrix of the matrix form Table 3 has another set of new four numbers but with the same disposition (mosaic) also. This set of matrices and other similar sets give many opportunities to represent them in algebraic terms of groups, rings, algebras, theory of square forms, etc.

On the article by Konopel’chenko and Rumer. As the author got to know recently, the first scientists who produced an effective idea on the modeling of subsystems of genetic codes in a form of mathematical matrices on the basis of a very formal approach, were
Russian scientists B.G. Konopel’chenko and Y.B. Rumer by name. In 1975 they published a short article, in which they presented a set of nitrogenous bases C, G, A, U in a form of a mathematical matrix with order [2x2]. Moreover, they used a tensor multiplication from a field of calculus of matrices to obtain a matrix with order [4x4], which had 16 duplets of genetic code. But their pioneer work had no prolongation in other publications for a quarter of a century approximately, as we know. Only in the past few years did some authors return to the Konopel’chenko-Rumer’s article.

Our investigations in their part of matrix analysis of genetic codes can be considered as a prolongation of Konopel’chenko-Rumer’s ideas. But our article has essential novelties and differences from the Konopel’chenko-Rumer’s article. Let us indicate the main differences.

The Konopel’chenko-Rumer’s article considered not the matrix \( P \), but another genetic symbolic matrix \( R \), which is transformed into the numeric matrix \( R_{\text{MULT}} \) in the case of \( C=G=3 \) and \( A=U=2 \):

\[
R = \begin{bmatrix}
C & G \\
A & U
\end{bmatrix}; \quad R_{\text{MULT}} = \begin{bmatrix}
3 & 3 \\
2 & 2
\end{bmatrix}
\]

No numeration numbers or coordination numbers of genetic elements were introduced in their work. This matrix \( R \) has non-diagonal location of complementary pairs of nitrogenous bases. All corresponding matrices \( R_{\text{MULT}} \) and \( R_{\text{MULT}}^{(n)} \), which can be constructed from the matrix \( R \), are singular: values of their determination are equal to 0. All matrices \( R_{\text{MULT}}^{(n)} \) are not bi-symmetrical. Singular matrices \( R_{\text{MULT}}^{(n)} \) are not so interesting in comparison with non-singular matrices \( P_{\text{MULT}}^{(n)} \), which are described in our article, in many aspects. In particular, we cannot take the square root from a matrix \( R_{\text{MULT}}^{(n)} \) to obtain a golden matrix. Konopel’chenko and Rumer did not consider any numeric matrices, including the matrix \( R_{\text{MULT}} \). They considered matrices \( R \) and \( R^{(2)} \), but they did not mention matrices \( R^{(3)}, R^{(4)}, \ldots R^{(n)} \) with genetic triplets and polyplets there at all. Our ideas about binary sub-alphabets of genetic codes (which lead to idea of digitizing genetic codes and to the idea of numerous parallel genetic languages of binary type) and about the bi-periodic matrices, connected with them, have no analogies in their article.

_Circular notation of tensor products for matrices._ This section of our article is devoted to the possible use of a circular notation of system sets of elements of genetic codes in heuristic aims. Such notation can be useful to analyze hidden rules of genetic sequences.
Previous sections of the article discuss mathematical matrices, which are written in a squared shape. Under tensor multiplication of two rectangular matrices, the final result is written usually in the form of a new rectangular matrix in accordance with the following rule. All separate matrix elements of a first matrix are combined with the whole second matrix into a rectangular frame step by step (for example, see Gazale 2002, pp. 210-219). However, such rectangular forms of matrices (and of a result of a tensor multiplication of two matrices also) are not the single possible variant of their notation at all, though this variant is very comfortable and habitual for modern mathematics. It has alternative variants, one of which is a circular notation considered below.

Figure 1: A circular notation of the matrix $P^{(3)} = P \otimes P \otimes P$
The author proposes to use a circular notation of matrices in appropriate cases of investigations of genetic codes and of genetic sequences. Figure 1 demonstrates an example of such circular notation for an octet matrix of the system of 64 genetic triplets. As stated above, the octet matrix \( P^{(3)} \) of 64 triplets is generated when the basic genetic matrix \( P \) (see Table 2), which has four matrix elements C, A, U and G, is raised to the power 3 in the sense of tensor multiplication. Figure 1 demonstrates the matrix \( P^{(3)} = P \otimes P \otimes P \) so that the first of these three matrices \( P \) corresponds to a central circle with a large radius (or to a circle of a first level) and with elements C, A, U and G, which are disposed in accordance with four sides of the world.

The set of four circles of the second level with a medium radius, which are located around elements C, A, U and G of the central circle, consists of 16 elements already. They depict 16 duplets of genetic code and they correspond to the matrix \( P^{(2)} = P \otimes P \); four elements of each such circle of the second level should be read beginning from an additional element, which is shown in the center of the circle. For example, the four elements of the top circle of a second level should be read as CC, CA, CG, CU.

Finally, 16 circles with a small radius (circles of a third level), which are disposed around elements of circles of the second level and which corresponds to the matrix \( P^{(3)} \), consists of 64 elements. They correspond to 64 triplets and they should be read beginning from two additional elements: firstly from a corresponding element of a circle of the first level, and secondly from a corresponding element of a circle of the second level. For example, the topmost element of this figure is the triplet CCC.

It is obvious that in a similar way, by introducing circles of subsequent levels with corresponding numbers of their elements, one can get a circular notation of a matrix, which is a result of tensor product of many initial matrices. Such circular notation of matrices with their tensor multiplications can be useful, additionally, for the interpretation of a series of Ancient Chinese and other ancient schematic pictures from the viewpoint of modern mathematics. It is very probable that, for example, Ancient Chinese wise men worked with the idea of tensor multiplication in simple forms, but they depicted results in specific schemes, which are unusual for modern science.

It should be mentioned that the circular notation of mutual relations among parts of structured systems is one of the most ancient and widespread form of system presentations. For example, the Ancient Chinese “The Book of Changes” and its satellite literature sources used a circular notation of 8 trigrams and 64 hexagrams widely. In modern science, unknown symmetrical rules have been discovered recently.
by the use of similar circular notations for analyzing poetic compositions. These rules were described by their author in the book with a typical title “Symmetry as the empress of poetry” (Porter 2003). We use a circular notation to analyze genetic sequences. In this field much unexpected questions arise. For example, what is common between poetry and sugary diabetes? What is common is that the insulin’s gene, defects of which lead to diabetes, has such symmetric structure in certain aspects, which is similar to symmetry in poetry structures.

Moreover, let us consider the circular notation of black-white mosaic of the genetic matrix $P_{MULT}^{(3)}$ from Table 5. Figure 2 represents the octet matrix $P_{MULT}^{(3)}$ in the form of eight concentric rings with sequences of black and white sectors there. Each concentric ring corresponds to one row of the matrix. A sequence of black and white sectors in this ring corresponds to a sequence of numbers of the tabular row: black sectors are for numbers 27 and 12, white sectors are for numbers 18 and 8. Similar black-white circles with appropriate quantities of concentric rings can be constructed for other cases of bi-periodic matrices with order $[2n \times 2n]$.

![Figure 2: A circular notation of a black-white mosaic of the matrix $P_{MULT}^{(3)}$ for 64 triplets (see Table 5). Numbers shown, which are agreed with coordination numbers of appropriate triplets from the bi-periodic Table 3.](image)

These black-and-white “genetic” circles are interesting, additionally, because they can be used in attempts to find an influence (or stamp) of structures of genetic codes at peculiarities of macro biological structures. For example, the psychophysics of visual perception knows of visual illusions, which appear when an observer looks at a rotating circle (Banham’s disk) with concentric rings and black-and-white sectors there (for example, see Gregory (1970). These illusions are consisted in appearance of a color
painting of this black-and-white circle. The author makes experiments about physiological influences of an observation of rotating circles with different black-and-white mosaics, which correspond to black-and-white mosaics of the matrices of genetic codes. These experiments study cases of a rotation of such circles in both – clockwise and anti-clockwise – directions.

A continuation of these researches is presented in this issue in the author’s next paper about numeric rules of degeneracy of genetic codes and about a chronocyclic theory.

ENDNOTE

1 An interesting aspect is concerned with the fact, that each quadrant of Table 3 has a ratio between the quantities of black and white cells which is equal to 3 : 1 (or 1 : 3). The famous genetic law by Mendel demonstrates the same 3 : 1 ratio for dominant and recessive attributes. The author proposes to apply the generalized Mendel-laws for the explanation of structural peculiarities of genetic codes, presented in black-and-white mosaics of the bi-periodic genetic matrices (Petoukhov, 2001, p.106). The results of an investigation of this new problem of a possible application of Mendel-type laws to a set of elements of genetic code will be published by the author in a special article.

ACKNOWLEDGMENT

I am grateful to K. Frolov, G. Darvas, B. Szántó, J. Böhm, Y. Ne’eman, M. He, K. Kovács for their support of these investigations. This study has been prepared in the framework of theme №7 of “The thematic plan of scientific collaboration between Hungarian and Russian Academies of sciences on 2002-2004 years”.

REFERENCES


Petoukhov, S. V. (2003b) Fractal matrices, genetic matrices of Fibonacci and a positioning in biological code. Elements of a matrix mechanics of development (the article was deposited in All-Russian Institute of Scientific and Technical Information (VINITI) of the Russian Academy of Sciences 14.08.2003, № 1593-B2003), (Directory of VINITI Deposited scientific works, № 10), (in Russian).


Petoukhov, S. V. (2003e) Bi-symmetric matrices of genetic code and golden section (the article was deposited in All-Russian Institute of Scientific and Technical Information (VINITI) of the Russian Academy of Sciences 25.11.2003, № 2042-B2003), (Directory of VINITI Deposited scientific works, № 1, 2004), (in Russian).


