Symmetry: Culture and Science
Vols. 14- 15, 281-307, 2003-2004

ATTRIBUTIVE CONCEPTION OF GENETIC CODE, ITS BI-PERIODIC TABLES AND PROBLEM OF UNIFICATION BASES OF BIOLOGICAL LANGUAGES*

Sergei V. Petoukhov

Address: Department of Biomechanics, Mechanical Engineering Research Institute, Russian Academy of Sciences, M. Kharitonyevsky pereulok 4, Moscow, Russia 101830;
E-mail: petoukhov@hotmail.com, http://www.petoukhov.narod.ru

The paper is devoted to symmetrological analysis of system of genetic coding and its parallelisms with linguistic languages.

From the information viewpoint, living organisms are information essences. They live due to receiving of the genetic information from the ancestors, and they exist to transfer the genetic information to their descendants. In the biological literature it is possible quite often to meet the statement that living organisms are the texts since a molecular level of their organization. Just from the information-hereditary point of view all living organisms are unified wonderfully: all of them have identical bases of system of genetic coding.

For realization of the genetic messages, which encode sequences of amino acids in proteins, all kinds of organisms utilize in their molecules of heredity DNA (and RNA – ribonucleic acid) the "alphabet" consisting of only four "letters" or nitrogenous bases: adenine (A), cytosine (C), guanine (G), thymine (T) {or uracil (U) in RNA}. Linear sequences of these four letters on strings of molecules of heredity (DNA and RNA) contain the genetic information for protein synthesis in all living bodies - from bacteria up to a whale or from worm up to a bird and even man. The given set of four letters is usually considered as the elementary alphabet of a genetic code. The modern science

* This study has been prepared in the framework of the theme №7 of the Thematic Plan of Collaboration between the Hungarian and Russian Academies of Sciences

does not know why the alphabet of genetic language has four letters (it could have any other number of 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 constructions has on 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 & A=G (according to binary-opposite attributes: "pyrimidine" or "non-pyrimidine", that is purine); 2) A=C & G=U (according to attributes: amino-mutating or non-amino-mutating under action of nitrous acid HNO_2); 3) C=G & A=U (according to attributes: three or two hydrogen bonds are materialized in these complementary pairs).

Let's number these binary-opposite attributes by numbers N=1,2,3 and let's ascribe to each of four letters of code the symbol "1_N" in case of presence at it of an attribute under number "N" and the symbol "0_N" in case of its opposite absence. In result we receive representation of the four-alphabetic alphabet of a code in system of its three "binary sub-alphabets to attributes" (see Table 1).

	Symbols of genetic "letter" from a viewpoint of a kind of binary-opposite attributes	C	A	G	U	
№1	**1_1** – pyrimidines (one ring in a molecule); **0_1** – purines (two rings in a molecule)	1_1	0_1	0_1	1_1	
№2	**1_2** – a letter with amino-mutating property; **0_2** – a letter without it	1_2	1_2	0_2	0_2	
№3	**1_3** – a letter with three hydrogen bonds; **0_3** – a letter with two hydrogen bonds	1_3	0_3	1_3	0_3	

Table 1: Three binary sub-alphabets according to three kinds of binary-opposite attributes in a set of nitrogenous bases C, A, G, U

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 O1 and 11, which

the author names as “the binary sub-alphabet to first kind of binary attributes”.

Accordingly, a genetic message as a sequence of four elements C, A, G, U consists of three parallel and various binary texts or three different sequences of zero and unit (such binary sequences are used at a storage and transfer of the information in computers). Each from these parallel binary texts, based on objective biochemical attributes, can provide its own subsection of protein synthesis. It’s very probable, that a process of protein synthesis is divided to relative independent subsections connected with these sub-alphabets.

The “elementary” four-letters alphabet of genetic code is not elementary at all. On the basis of these binary sub-alphabets, the bi-periodic octet table of genetic code with fractal properties has been constructed (for more details see [Petoukhov 1999; 2001, Figure 2.3.1 and Table 2.3.1]). The author published two variants of this bi-periodic table earlier, which were differed from each other with replacing of two middle rows only. Table 2 represents a variant of the bi-periodic octet table with decreasing order of numbers of rows. This main variant has interesting mathematical properties, which will be described in the next page (Table 2).

A principle of construction of the octet Table 2 is very simple. All its columns are numbered by binary numbers in decreasing order 111, 110, … , 000 for a disposition in each column those eight triplets, which have appropriate binary numbers of their letters from a viewpoint of the first binary sub-alphabet of the Table 1. For example, a column with the binary number 110 accumulates eight triplets, which have pyrimidines (C, G) in two first positions and a purine (A,U) in a third position: CCA, CCG, … , UUG. All rows of the Table 2 are numbered in the same manner by binary numbers 111, 110, … , 000, but now these binary symbols “1” and “0” are interpreted from the viewpoint of a second sub-alphabet of the Table 1. For example, a row with the binary number 110 accumulates eight triplets, which have amino-mutating elements (C, A) in two first positions and non-amino-mutating elements (U, G) in a third position. In the Table each triplet has its own coordinates in a form of three-digit numbers of its row and its column. These numbers can be combined in a six-digit binary number to characterize each triplet individually. For example, in this case binary number 110101 characterizes triplet CAU. At translation of these six-digit binary numbers into decimal numeration, all triplets receive their individual decimal numbers from 0 to 63. Let’s analyze a few mathematical properties of this bi-periodic table or matrix.

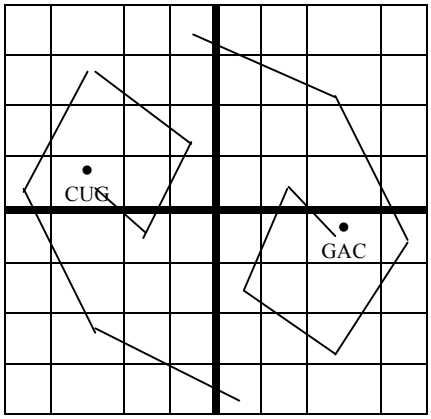
	111 CHYAN	110 TUI	101 LI	100 CHEN	011 HSUN	010 KAN	001 KEN	000 KUN
111 CHYAN	CCC 63	CCA 62	CAC 61	CAA 60	ACC 59	ACA 58	AAC 57	AAA 56
110 TUI	CCU 55	CCG 54	CAU 53	CAG 52	ACU 51	ACG 50	AAU 49	AAG 48
101 LI	CUC 47	CUA 46	CGC 45	CGA 44	AUC 43	AUA 42	AGC 41	AGA 40
100 CHEN	CUU 39	CUG 38	CGU 37	CGG 36	AUU 35	AUG 34	AGU 33	AGG 32
011 HSUN	UCC 31	UCA 30	UAC 29	UAA 28	GCC 27	GCA 26	GAC 25	GAA 24
010 KAN	UCU 23	UCG 22	UAU 21	UAG 20	GCU 19	GCG 18	GAU 17	GAG 16
001 KEN	UUC 15	UUA 14	UGC 13	UGA 12	GUC 11	GUA 10	GGC 9	GGA 8
000 KUN	UUU 7	UUG 6	UGU 5	UGG 4	GUU 3	GUG 2	GGU 1	GGG 0

Table 2: The octet bi-periodic table of genetic triplets with their numeration

1. DOUBLE HELIX DNA AND A DOUBLE TRAJECTORY OF THE BI-PERIODIC TABLE

G A C	A U G	G C U	G G C	G A A	A A U	C A A
III II III	II II III	III III II	III III III	III II II	II II II	III II II
C U G	U A C	C G A	C C G	C U U	U U A	G U U

Each pair “codon-anticodon” has an inverse-symmetrical disposition in the bi-periodic Table 2 (and invariant sum 63 of their numerations). So, for any sequence of triplets from a double helix of DNA we have in the bi-periodic table an appropriate double trajectory, both branches of which are located in inverse-symmetric manner relative to the tabular center. For example, let’s consider following complementary sequences in both strands of DNA:



Then we have for these complementary sequences a following double trajectory with two inverse-symmetrical branches in Table 2.

2. GENETIC MATRICES AND THEIR NUMERIC MOSAICS

A matrix that can be obtained from the bi-periodic table of genetic code by one or another natural way will be named "genetic matrix" (or "genomatrix"). Till this moment we have used two first kinds of binary-opposite attributes of the Table 1 for a construction of the Table 2. Now we'll 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's replace each triplet in Table 2 by the product of these numbers of its hydrogen bonds. For example, the triplet CAU will be replaced by number 12 (=3*2*2). In result the octet "multiplicative" matrix $\mathbf{M_8}$, submitted in the Table 3, will be produced.

The matrix $\mathbf{M_8}$ is symmetrical relative to both diagonals and can be named a "bi-symmetrical matrix". All rows and all columns of this matrix are differed from each other by sequences of their numbers. But the sums of all numbers in cells of each row and of each column are identical to each other and are equal to $125 = 5^3$. A determinant of this matrix is equal to 5^{12}. Eigenvalues of this matrix: 1, 5, 5, 5, 5^2, 5^2, 5^2, 5^3.

The matrix $\mathbf{M_8}$ has four kinds of numbers only: 8, 12, 18, 27. The certain laws are observed in their disposition. Only a few of them will be noted here.

27	18	18	12	18	12	12	**8**	125
18	**27**	12	18	12	18	**8**	12	125
18	12	**27**	18	12	**8**	18	12	125
12	18	18	**27**	**8**	12	12	18	125
18	12	12	**8**	**27**	18	18	12	125
12	18	**8**	12	18	**27**	12	18	125
12	**8**	18	12	18	12	**27**	18	125
8	12	12	18	12	18	18	**27**	125
125	125	125	125	125	125	125	125	1000

Table 3: Multiplicative matrix $\mathbf{M_8}$ of the bi-periodic table 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.

Each diagonal of the matrix has one kind of numbers – 8 or 27 exclusively. So, a set of cells with these numbers visualizes a figure of a diagonal cross. Non-diagonal cells contain numbers 12 and 18 only. Let's consider a disposition of cells with number 12 in Table 3. For this purpose these cells will be marked dark, and all other cells will be white. For the better presentation of arising symbols more dark color is used for painted cells in one half of the table, than in another. The result is illustrated in Table 4 (left).

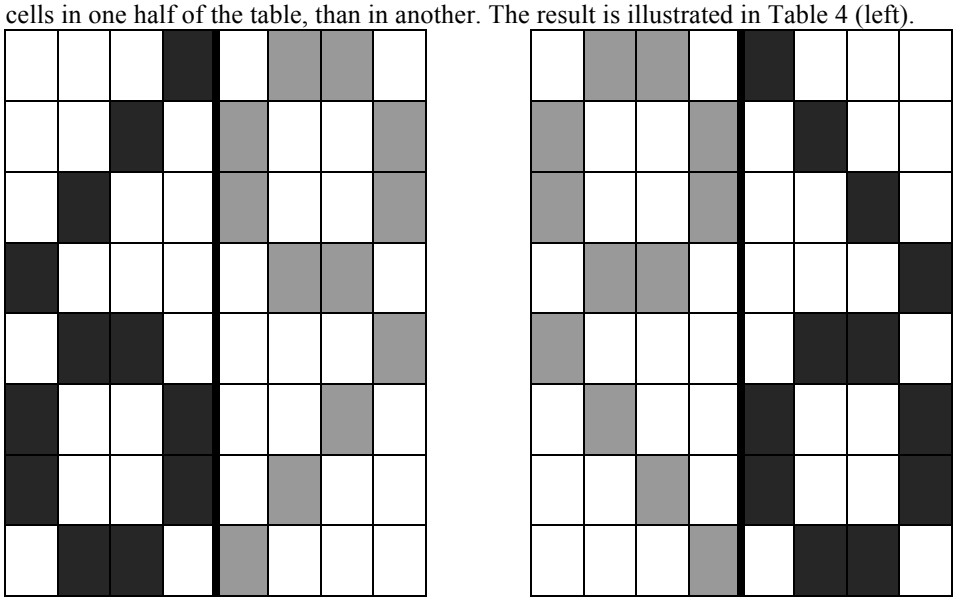


Table 4: A mosaic of cells with number 12 (left) and a mosaic of cells with number 18 (right) from the multiplicative matrix $\mathbf{M_8}$ of Table 3.

A symbol of number **69** is visualized in this mosaic unexpectedly. This number is constituted with blacker figure **6** from the left half of the table and with brighter figure **9** from the right half. Pictures **6** and **9** are inverse-symmetrical each other. Triplets from tabular cells, which are included in structure of the figure **6**, are anticodons in relation to triplets of those cells, which form the figure **9**.

A similar mosaic of tabular cells with number 18 (these cells are located in mirror-symmetrical positions concerning cells with number 12 in the Table 3) visualizes a mirror image of the figure **69** (see Table 4, right). If we rotate these matrices with 90^0 simultaneously, mosaics of symbols 69 and its mirror image are changed by their places in both matrices of Table 4.

The matrix $\mathbf{M_8}$ of Table 3 and the same matrix in a position rotated on 90^0 are complementary matrices in the following sense. If numbers from two cells with the same dispositions in both matrices are multiplied each with other, a product will be equal to 216 for each pair of appropriate cells. If numbers from two cells with the same dispositions in both matrices are added to each other, a sum will be equal to 35 for each pair of appropriate diagonal cells and to 30 for each pair of appropriate non-diagonal cells.

A trace of the matrix $\mathbf{M_8}$ is equal to 216. If the matrix $\mathbf{M_8}$ is divided by this number 216, new normalized matrix $\mathbf{M_{8N}}$ is appeared with a trace 1. This matrix $\mathbf{M_{8N}}$ has a set of new four numbers 1/27, 1/18, 1/12, 1/8, which are reciprocal numbers for 27, 18, 12, 8 from the initial matrix $\mathbf{M_8}$. The matrix $\mathbf{M_{8N}}$ has analogies with “matrices of density” from statistical quantum mechanics but peculiarities and a meaning of these analogies are out of this paper.

3. GENETIC MATRICES AND A GOLDEN SECTION

In biology a genetic system provides a self-reproduction of biological organisms in their generations. In mathematics so named “golden section” (or “divine proportions”) and its properties were a mathematical symbol a self-reproduction from Renaissance (for example, see WEB-site "*Museum of Harmony and Golden Section*" by A. Stakhov http://www.goldenmuseum.com). Is there any connection between these two systems? Yes, author’s researches found out such connection. We discovered that, if the multiplicative matrix $\mathbf{M_8}$ of Table 3 is raised to the power ½ (that is if we take the square root), a result would be a matrix $\mathbf{\Phi_8}$ presented in Table 5, where $\varphi = (1+5^{0.5})/2 =$

1, 618… is a golden section. A resulting matrix $\mathbf{\Phi_8}$ can be named as a golden octet matrix, because it has only four kinds of numbers, generated from a single value of golden section: φ^3, φ^1, φ^{-1}, φ^{-3}. They have the same disposition (and the same numeric mosaic) in the octet matrix as four kinds of numbers 27, 18, 12, 8 in the initial matrix $\mathbf{M_8}$. This new theme of golden matrices in genetic systems has many aspects, which will be presented in special publications.

φ^3	φ^1	φ^1	φ^{-1}	φ^1	φ^{-1}	φ^{-1}	φ^{-3}
φ^1	φ^3	φ^{-1}	φ^1	φ^{-1}	φ^1	φ^{-3}	φ^{-1}
φ^1	φ^{-1}	φ^3	φ^1	φ^{-1}	φ^{-3}	φ^1	φ^{-1}
φ^{-1}	φ^1	φ^1	φ^3	φ^{-3}	φ^{-1}	φ^{-1}	φ^1
φ^1	φ^{-1}	φ^{-1}	φ^{-3}	φ^3	φ^1	φ^1	φ^{-1}
φ^{-1}	φ^1	φ^{-3}	φ^{-1}	φ^1	φ^3	φ^{-1}	φ^1
φ^{-1}	φ^{-3}	φ^1	φ^{-1}	φ^1	φ^{-1}	φ^3	φ^1
φ^{-3}	φ^{-1}	φ^{-1}	φ^1	φ^{-1}	φ^1	φ^1	φ^3

Table 5: The golden octet genetic matrix $\mathbf{\Phi_8} = (\mathbf{M_8})^{0.5}$. Here $\varphi = 0.5*(1+5^{0.5}) = 1, 618…$ is a golden section.

4. MOSAIC-INVARIANT MATRICES

The genetic matrix of Table 3 has unexpected group-invariant property corresponding to a multiplication of matrices. If this matrix, which has a set of four numbers 8, 12, 18, 27, is raised to a power of n, a resulting matrix will have a set of four (another) numbers with the same specific mosaic of their disposition. For example, if this matrix is raised to a second power, a resulting matrix will have the same form of numeric mosaic or the same disposition of a set of new four numbers: $1728 = 2^6x3^3$ (instead of 8), $1872 = 2^4x3^2x13$ (instead of 12), $2028 = 2^2x3x13^2$ (instead of 18), $2197 = 13^3$ (instead of 27). In other words, a numeric mosaic in a resulting matrix will be the same as in initial matrix. It's essential, that this beautiful property of invariance of the numeric mosaic of the genetic matrix is independent on values of numbers. This property is realized for such matrices with arbitrary set of four numbers *a*, *b*, *c*, *d*, if they are located in the same manner inside a matrix. More over, 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 a product of these matrices will be a matrix **Z**=**X***Y with a set of new four numbers *r*, *g*, *v*, *z* and with the same mosaic of their disposition (Table 6):

d	*c*	*c*	*b*	*c*	*b*	*b*	*a*
c	*d*	*b*	*c*	*b*	*c*	*a*	*b*
c	*b*	*d*	*c*	*b*	*a*	*c*	*b*
b	*c*	*c*	*d*	*a*	*b*	*b*	*c*
c	*b*	*b*	*a*	*d*	*c*	*c*	*b*
b	*c*	*a*	*b*	*c*	*d*	*b*	*c*
b	*a*	*c*	*b*	*c*	*b*	*d*	*c*
a	*b*	*b*	*c*	*b*	*c*	*c*	*d*

*

q	*p*	*p*	*m*	*p*	*m*	*m*	*k*
p	*q*	*m*	*p*	*m*	*p*	*k*	*m*
p	*m*	*q*	*p*	*m*	*k*	*p*	*m*
m	*p*	*p*	*q*	*k*	*m*	*m*	*p*
p	*m*	*m*	*k*	*q*	*p*	*p*	*m*
m	*p*	*k*	*m*	*p*	*q*	*m*	*p*
m	*k*	*p*	*m*	*p*	*m*	*q*	*p*
k	*m*	*m*	*p*	*m*	*p*	*p*	*q*

=

z	*v*	*v*	*g*	*v*	*g*	*g*	*r*
v	*z*	*g*	*v*	*g*	*v*	*r*	*g*
v	*g*	*z*	*v*	*g*	*r*	*v*	*g*
g	*v*	*v*	*z*	*r*	*g*	*g*	*v*
v	*g*	*g*	*r*	*z*	*v*	*v*	*g*
g	*v*	*r*	*g*	*v*	*z*	*g*	*v*
g	*r*	*v*	*g*	*v*	*g*	*z*	*v*
r	*g*	*g*	*v*	*g*	*v*	*v*	*z*

X(*a*, *b*, *c*, *d*) **Y**(*k*, *m*, *p*, *q*) **Z**(*r*, *g*, *v*, *z*)

Table 6: Multiplication of mosaic-invariant matrices **X** and **Y** gives a new matrix **Z** with the same mosaic of the disposition of its four kinds of numbers. For illustration, cells with numbers *b*, *m*, *s* in matrices **X**, **Y**, **Z** are marked by dark color.

It's obvious, that 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. Such mosaic--invariant property of these genetic matrices is an expression of cooperative behavior of its elements, but not a result of individual behavior of each kind of elements. This property reminds many aspects of cooperative behavior of 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. Numeric mosaics in such genetic matrices are conserved under operations of matrix addition and of multiplication by number. This set of matrices and other sets that will be described in this article below, give many opportunities to describe them in algebraic terms of groups, rings, algebras, theory of square forms, etc.

5. BI-PERIODIC MATRICES OF GENETIC CODE AND A TENSOR MULTIPLICATION

Let's analyze of a disposition of triplets in the bi-periodic Table 2 more attentively with interpreting of this table as a matrix. Table 7 (right) shows it in comfortable form.

C	A
U	G

⇒

CC	CA	AC	AA
CU	CG	AU	AG
UC	UA	GC	GA
UU	UG	GU	GG

⇒

⇒

CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
CCU	CCG	CAU	CAG	ACU	ACG	AAU	AAG
CUC	CUA	CGC	CGA	AUC	AUA	AGC	AGA
CUU	CUG	CGU	CGG	AUU	AUG	AGU	AGG
UCC	UCA	UAC	UAA	GCC	GCA	GAC	GAA
UCU	UCG	UAU	UAG	GCU	GCG	GAU	GAG
UUC	UUA	UGC	UGA	GUC	GUA	GGC	GGA
UUU	UUG	UGU	UGG	GUU	GUG	GGU	GGG

Table 7: Fractal structure of bi-periodic matrix of genetic code and a possible model of its 3-steps evolution with algorithmic fragmentation of each matrix cell into four cells and with a repeated cross-order of complementary "letters" C-G and A-U in new positions (from [Petoukhov, 2001, p. 106]).

Each quadrant (4x4) of the bi-periodic matrix has all 16 triplets, which have the same letter in their first position. So, according to a name of a letter in a first position of their triplets, we have C-quadrant, G-quadrant, A-quadrant and U-quadrant with cross-dispositions of C-G-quadrants and A-U-quadrants in the table. Each sub-quadrant (2x2) of each named quadrant has triplets with the same letter in their second position and with the same cross-dispositions of complementary letters C-G and A-U there. A set of sub-sub-quadrants (1x1) has the same cross character in relative to letters C-G and A-U in third positions of their triplets.

Due to this peculiarity, a construction of the bi-periodic matrix can be deduced from a special matrix (2x2) of nitrogenous bases C, A, U, G (see Table 7, left). Such deduction is fulfilled with operation of tensor (or Kronecker) multiplication of matrices according to ordinary rules (for example, see [Gazale, 2000; Konopel'chenko, Rumer, 1975]). If such matrix (2x2) is raised to the power 2 in the sense of tensor multiplication, a result is the matrix (4x4) of 16 duplets from the table 7 (middle). If it's raised to the tensor

power 3, a result is the matrix (8x8) of 64 triplets from the table 7 (right). A symbol of tensor multiplication is "⊗". A symbol $\mathbf{P}^{(n)}$ means, that a matrix **P** is raised to the power "*n*" in a sense of tensor multiplication.

$$\mathbf{P} = \begin{vmatrix} C & A \\ U & G \end{vmatrix}, \quad \mathbf{P}^{(2)} = \mathbf{P}\otimes\mathbf{P} = \begin{vmatrix} C\begin{vmatrix} C & A \\ U & G \end{vmatrix}, & A\begin{vmatrix} C & A \\ U & G \end{vmatrix} \\ U\begin{vmatrix} C & A \\ U & G \end{vmatrix}, & G\begin{vmatrix} C & A \\ U & G \end{vmatrix} \end{vmatrix} = \begin{vmatrix} CC & CA & AC & AA \\ CU & CG & AU & AG \\ UC & UA & GC & GA \\ UU & UG & GU & GG \end{vmatrix}$$

$\mathbf{P}^{(3)} = \mathbf{P}\otimes\mathbf{P}\otimes\mathbf{P} =$

CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
CCU	CCG	CAU	CAG	ACU	ACG	AAU	AAG
CUC	CUA	CGC	CGA	AUC	AUA	AGC	AGA
CUU	CUG	CGU	CGG	AUU	AUG	AGU	AGG
UCC	UCA	UAC	UAA	GCC	GCA	GAC	GAA
UCU	UCG	UAU	UAG	GCU	GCG	GAU	GAG
UUC	UUA	UGC	UGA	GUC	GUA	GGC	GGA
UUU	UUG	UGU	UGG	GUU	GUG	GGU	GGG

Table 8: A connection among matrices of Table 7 on a base of tensor multiplications of the matrix P.

6. A FAMILY OF BI-PERIODIC MATRICES $\mathbf{P}^{(N)}$

Each matrix $\mathbf{P}^{(n)}$ is a bi-periodic matrix from a viewpoint of appropriate binary sub-alphabets of Table 1. Really, all n-plets of each column of a matrix $\mathbf{P}^{(n)}$ are equivalent to each other from a viewpoint of the first binary sub-alphabet of genetic code (Table 1). In other words, these n-plets have the same sequence of pyrimidines (C, U) and purines (A, G). All *n*-plets of each row of a matrix $\mathbf{P}^{(n)}$ are equivalent to each other from a viewpoint of the second binary sub-alphabet of genetic code (Table 1). In other words, these n-plets have the same sequence of amino-mutating (C, A) and non-amino-mutating (G, U) bases. It's obvious that we can introduce an appropriate binary numeration for columns and rows of each matrix $\mathbf{P}^{(n)}$ in accordance with first and second genetic sub-alphabets from Table 1 by analogy with the binary numeration of columns and rows of the bi-periodic matrix for triplets (Table 2). Then each *n*-plet will have its individual number in appropriate matrix $\mathbf{P}^{(n)}$, etc.

The bi-periodic matrix **P** is a basic matrix for a construction of a formal family of bi-periodic matrices $\mathbf{P}^{(n)}$, where "*n*" can be equal not only to 1, 2, 3, but also to arbitrary positive integer number. In this family, a matrix $\mathbf{P}^{(n)}$ has an order ($2^n * 2^n$) and contains appropriate number of *n*-plets, that is a sequence of n letters in each cell. For example, a matrix $\mathbf{P}^{(4)}$ has an order (16*16) and contains 256 tetraplets. This family of genetic

matrices has commutative and other valuable algebraic properties. This family is important because of all possible genetic sequences of arbitrary length are belong to its matrices. A sub-family of matrices $\mathbf{P}^{(3n)}$ is interesting specially because it contains all possible genetic sequences of triplets in a form of 3n-plet. A set of such sequences has 3*n*-plets with real coding sense (they code proteins) and without it (they don't 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 $\mathbf{P}^{(999)}$.

An important task for a nearest future is the following. We need to take all known 3n-plets from banks of genetic data and locate them in appropriate genetic matrix $\mathbf{P}^{(n)}$. Then we should analyze what matrix cells are occupied by 3*n*-pets with coding sense and what mosaics these cells form. An aim of such symmetrological investigation proposed by the author is a clarification (or modeling) of reasons of disintegration a whole set of 3*n*-plets of nitrogenous bases into sub-sets of 3*n*-pets with coding sense and without it, and a prediction of valuable genetic peculiarities, which are unknown for modern science yet. It's essential that special mathematical apparatus, which is generated here in connection with genetic binary sub-alphabets, is appropriate to study not only 64 triplets and 20 amino acids but also a primary structure of proteins simultaneously.

It can be noted additionally, that genetic elements C, A, U, G can be interpreted conditionally as young and old Ying and Yang in the frame of parallelisms between genetic code system and a famous symbolic system of the Ancient Chinese "The Book of Changes" [Petoukhov, 1999, 2001]. If we use a binary numeration C = 11, A = 10, U = 01, G = 00, then a famous table of 64 hexagrams from "The Book of Changes" will be equivalent to a matrix $\mathbf{P}^{(3)}$. It demonstrates a mathematical (tensor) connection between basic Chinese ideas about "young and old Ying and Yang" and the Chinese table of 64 hexagrams in Fu-Xi's order.

7. A SUB-FAMILY OF BI-SYMMETRICAL GENETIC MATRICES AND THEIR MOSAICS

Let's return to a case C=G=3 and A=U=2 for numbers of hydrogen bonds (see Table 3) and change all multiplets in matrices of Table 7 by products of these numbers. For this case we have the following multiplicative matrices:

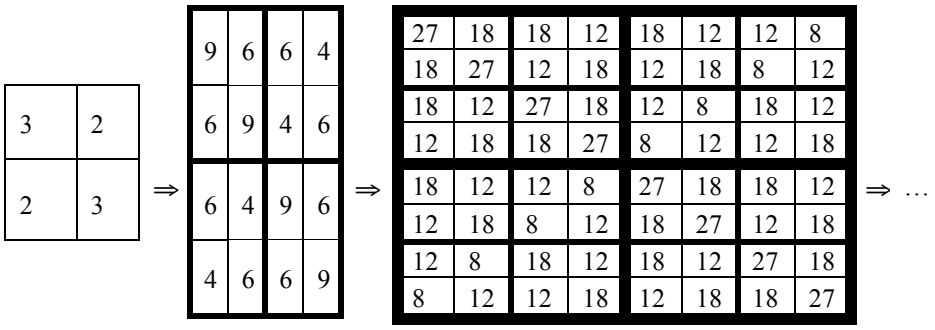


Table 9: Multiplicative bi-symmetrical matrices $\mathbf{M^{(n)}}$ from Tables 7, 8 for a case of a complementary hydrogen bonds C=G=3, A=U=2. $\mathbf{M^{(3)}}$ is identical to the matrix $\mathbf{M_8}$ from Table 3.

All numeric matrices of this family are bi-symmetrical: they are symmetrical with regard to both diagonals. Let's investigate numeric mosaics in these matrices. A first matrix $\mathbf{M^{(1)}}$ (2x2) of this family has two kinds of numbers (2 and 3) with a disposition, which forms two numeric mosaics of diagonal lines (their aggregate figure is a diagonal cross). Second matrix $\mathbf{M^{(2)}}$ with order (4x4) has three kinds of numbers (9, 6, 4), which are formed in each quadrant (2x2) of this matrix the same figure of diagonal cross. Such mosaics, which are realized in quadrants and sub-quadrants of matrices, will be named "local mosaics" in contrast to "global mosaics", which exist in entire matrix. So, the matrix $\mathbf{M^{(2)}}$ has "a memory" about the global mosaics of $\mathbf{M^{(1)}}$ in a form of local mosaics in its quadrants of appropriate order (2x2). But the matrix $\mathbf{M^{(2)}}$ has its global mosaics also in a form of bigger diagonal lines (which form a bigger diagonal cross) with their numbers 9 and 4 and in a form of a circle (a set of cells with number 6). These global mosaics of the matrix $\mathbf{M^{(2)}}$ (4x4) are repeated in the next matrix generation as local mosaics in quadrants (4x4) of the matrix $\mathbf{M^{(3)}}$. The matrix $\mathbf{M^{(3)}}$ has four kinds of numbers (8, 12, 18, 27) and its own global mosaics for each kind of these numbers also, that were considered in Tables 3 and 4 early.

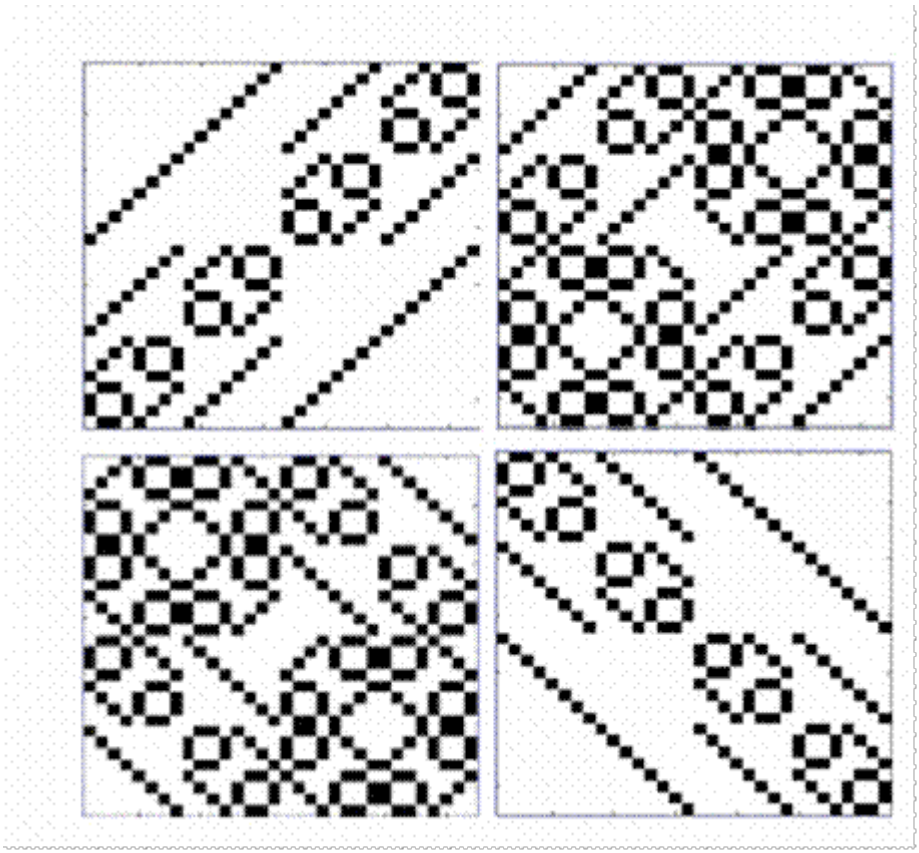


Table 10: All non-trivial mosaics for a matrix $\mathbf{M}^{(5)}$ with order (32x32), which has a set of integers 25, 24*3, 23*32, 22*33, 2*34, 35, are shown (for a case C=G=3, A=U=2). Cells with a considered number are marked by dark color. Topper: mosaics of matrix cells with number 24*3 (left) and with number 23*32 (right). Lower series: mosaics of matrix cells with number 22*33 (left) and with number 2*34 (right).

Any matrix $\mathbf{M}^{(n)}$ of this family has a set of (n+1) kinds of numbers and (n+1) global mosaics correspondingly. This set has numbers 2n, 2 (n-1) * 3, …, 2 * 3(n-1), 3n, which are forming a geometric progression with a coefficient 3/2. Also $\mathbf{M}^{(n)}$ has (in its appropriates quadrants and sub-quadrants) local mosaics which repeat all global mosaics of all previous matrix generations. In other words, this family of genetic matrices is

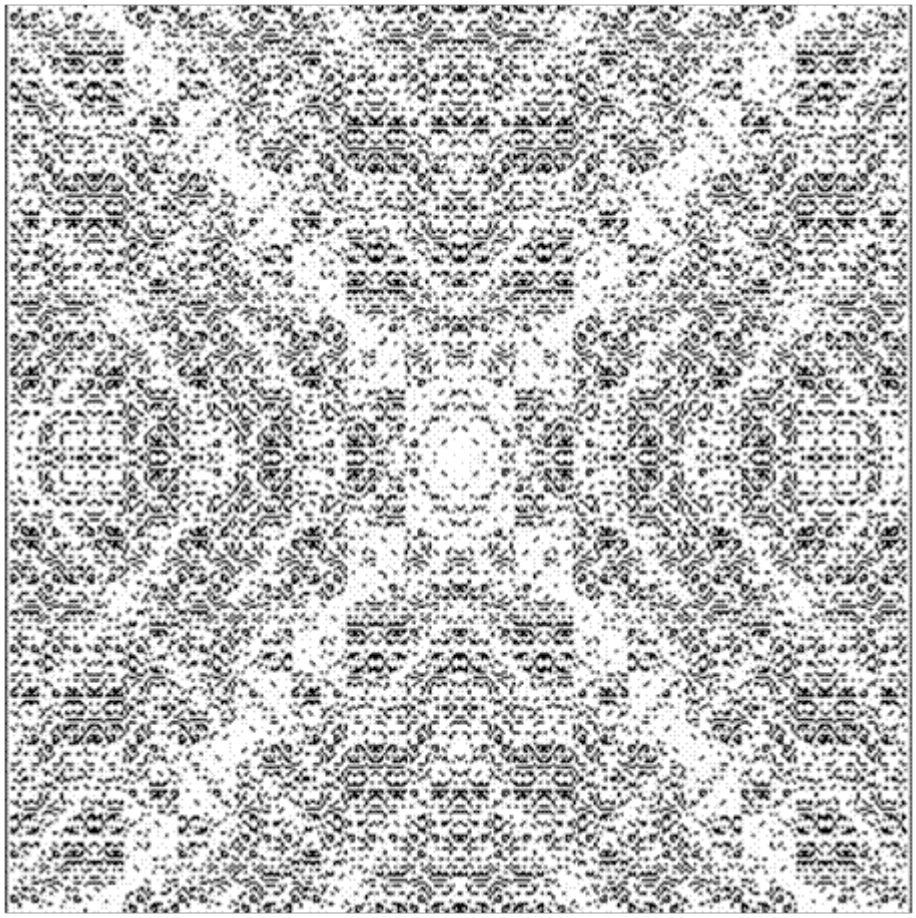


Table 11: A mosaic of cells (dark color) with number 25*35 from a matrix M(10) with order (1024x1024) is shown (for a case C=G=3, A=U=2).

characterized by a property of "a memory of generations" of matrix mosaics. This property can be used for matrix modeling many physiological and creative phenomena, which have properties of memory (for example, in morphogenesis, linguistics, logic, etc). Each matrix $\mathbf{M}^{(n)}$ has its inverse matrix $(\mathbf{M}^{(n)})^{-1}$, which has the same quantity of kinds of numbers with the same mosaics of their matrix disposition. It's important from an algebraic viewpoint.

It seems interesting to investigate, what mosaic pictures are visualized in such bi-matrices of high orders. The author made an album of all mosaics of this matrix family till to a matrix $\mathbf{M^{(10)}}$ with order (1024x1024). All $\mathbf{M^{(n)}}$ have two "primitive" mosaics in

Table 12: A beginning of a family of golden genetic matrices Φ(n), where "φ" is a golden section. The matrix Φ(3) is equal to the golden octet matrix Φ8 from Table 6.

$$\mathbf{P} = \begin{Vmatrix} 3 & 2 \\ 2 & 3 \end{Vmatrix}; \quad \mathbf{P^{1/2}} = \mathbf{\Phi} = \begin{Vmatrix} \varphi & \varphi^{-1} \\ \varphi^{-1} & \varphi \end{Vmatrix}; \quad \mathbf{\Phi \otimes \Phi} = \mathbf{\Phi^{(2)}} = \begin{Vmatrix} \varphi^2 & \varphi^0 & \varphi^0 & \varphi^{-2} \\ \varphi^0 & \varphi^2 & \varphi^{-2} & \varphi^0 \\ \varphi^0 & \varphi^{-2} & \varphi^2 & \varphi^0 \\ \varphi^{-2} & \varphi^0 & \varphi^0 & \varphi^2 \end{Vmatrix}; \quad \mathbf{\Phi^{(3)} = \Phi_8}$$

forms of their diagonal lines, where extreme numbers 3*n* and 2*n* of their numeric set are located. But other mosaics have complex, non-trivial structures. Table 10 shows all non-trivial mosaics for a matrix $\mathbf{M^{(5)}}$, which has a set of integers 25, 24*3, 23*32, 22*33, 2*34, 35. Table 11 shows a mosaic for number 25*35 in a matrix $\mathbf{M^{(10)}}$, which has a set of 11 integers.

According to our data, many matrices from this infinite family have a physiological value, which is under systematic testing now in connection with psychological and medical ideas by Carl G. Jung about archetypes or instinctive figures, etc. It seems very interesting to find similar mosaics in ancient ornamental patterns also.

8. A FAMILY OF GOLDEN GENETIC MATRICES

If a basic genetic bi-symmetrical matrix **M** (for a case of hydrogen bonds numbers C=G=3, A=U=2) is raised to the power ½ (in ordinary sense), a result will be a bi-symmetrical matrix **Φ** presented in Table 12, where $\varphi = (1+5\mathbf{0.5})/2 = 1,618\ldots$ is a golden section.

When the bi-symmetrical matrix **Φ** is raised to the power "*n*" in the sense of tensor multiplication, a result is the bi-symmetrical matrix $\mathbf{\Phi^{(n)}}$ with its order (2*n*x2*n*). Each matrix $\mathbf{\Phi^{(n)}}$ has only one kind of number – the golden section φ -, but in appropriate powers. A whole family of golden matrices $\mathbf{\Phi^{(n)}}$ is formed by such way. The family of golden matrices $\mathbf{\Phi^{(n)}}$ and the family of genetic bi-symmetrical matrices $\mathbf{M^{(n)}}$ from Table

9 are connected closely: when any matrix $\mathbf{\Phi}^{(n)}$ is raised to the second power, a result matrix is the matrix $\mathbf{M}^{(n)}$ (for a considered case C=G=3, A=U=2). In this reason all theory of the genetic matrix family $\mathbf{M}^{(n)}$ and its applications can be constructed on a base of the family of golden matrices $\mathbf{\Phi}^{(n)}$. Figuratively speaking, for the considered case the matrix family $\mathbf{M}^{(n)}$ has a masked substrate in a form of the golden matrix family $\mathbf{\Phi}^{(n)}$. Both matrix families have identical mosaics. Each golden matrix $\mathbf{\Phi}^{(n)}$ has its inverse golden **matrix** $(\mathbf{\Phi}^{(n)})^{-1}$, which contains the same quantity of kinds of numbers with the same mosaics of their matrix disposition. Families of genetic golden matrices have many interesting mathematical properties additionally, which will be described in special publication.

We can give a principal new definition of a golden section on the base of specifics of genetic code system: a golden section φ and its inverse value φ^{-1} are elements of the bi-symmetrical matrix $\mathbf{\Phi}$, which is a square root from a bi-symmetrical matrix **M** (2x2) with numeric elements from a set of numbers of complementary hydrogen bonds in genetic nitrogenous bases: C=G=3, A=U=2. This definition doesn't use habitual conceptions of line segments, their proportions, etc., but has a genetic matrix nature. In our opinion, many realizations of the golden section in the nature are connected with its matrix essence. It should be investigated specially and systematically, where in natural systems and phenomena we have the bi-symmetrical matrix **M** with elements 2 and 3 in a direct or masked form (just as in Table 13, right). We can hope to discover many new system phenomena and connections between them in the nature on this way.

The new theme of a golden section in genetic matrices is very important because many physiological systems and processes are connected with it. It's known, that proportions of a golden section characterize cardio-vascular processes, respiratory processes, electric activity of brain, locomotion activity, etc. The author hopes, that the algebra of bi-symmetric genetic matrices, proposed and developed by him in connection with a theme of golden section, will be useful for explanation and numeric forecast of separate parameters in a set of different physiological sub-systems of biological organisms with their cooperative essence and golden section phenomena. The conception of golden matrices can be useful for non-biological phenomena also.

9. ADDITIONAL KINDS OF GOLDEN AND FIBONACCI GENETIC MATRICES

Careful search of our bi-periodic matrices of genetic code discovered additional types of

golden genetic matrices.

Let's oversimplify the octet matrix of Table 2 for triplets and make three separate matrices of monoplets, which are disposed in first, second and third positions of triplets of Table 2 correspondingly. The resulting monoplets matrices for a considered case of hydrogen bonds C=G=3, A=U=2 are presented in Table 13.

In a similar way, any bi-periodic matrix $\mathbf{P}^{(n)}$ with its n-multiplets (see Table 8) can be transformed into n matrices of monoplets with a discovery of their analogical interrelation with such types of golden matrices for a case C=G=3, A=U=2.

$$\begin{vmatrix} 3&3&3&3&2&2&2&2\\ 3&3&3&3&2&2&2&2\\ 3&3&3&3&2&2&2&2\\ 3&3&3&3&2&2&2&2\\ 2&2&2&2&3&3&3&3\\ 2&2&2&2&3&3&3&3\\ 2&2&2&2&3&3&3&3\\ 2&2&2&2&3&3&3&3 \end{vmatrix} \quad \mathbf{B_1} \qquad \begin{vmatrix} \varphi&\varphi&\varphi&\varphi&\tau&\tau&\tau&\tau\\ \varphi&\varphi&\varphi&\varphi&\tau&\tau&\tau&\tau\\ \varphi&\varphi&\varphi&\varphi&\tau&\tau&\tau&\tau\\ \varphi&\varphi&\varphi&\varphi&\tau&\tau&\tau&\tau\\ \tau&\tau&\tau&\tau&\varphi&\varphi&\varphi&\varphi\\ \tau&\tau&\tau&\tau&\varphi&\varphi&\varphi&\varphi\\ \tau&\tau&\tau&\tau&\varphi&\varphi&\varphi&\varphi\\ \tau&\tau&\tau&\tau&\varphi&\varphi&\varphi&\varphi \end{vmatrix} \quad \mathbf{\Phi_{(1)} = 2*(B_1)^{1/2}}$$

$$\begin{vmatrix} 3&3&2&2&3&3&2&2\\ 3&3&2&2&3&3&2&2\\ 2&2&3&3&2&2&3&3\\ 2&2&3&3&2&2&3&3\\ 3&3&2&2&3&3&2&2\\ 3&3&2&2&3&3&2&2\\ 2&2&3&3&2&2&3&3\\ 2&2&3&3&2&2&3&3 \end{vmatrix} \quad \mathbf{B_2} \qquad \begin{vmatrix} \varphi&\varphi&T&\tau&\varphi&\varphi&\tau&\tau\\ \varphi&\varphi&T&\tau&\varphi&\varphi&\tau&\tau\\ \tau&\tau&\varphi&\varphi&\tau&\tau&\varphi&\varphi\\ \tau&\tau&\varphi&\varphi&\tau&\tau&\varphi&\varphi\\ \varphi&\varphi&T&\tau&\varphi&\varphi&\tau&\tau\\ \varphi&\varphi&T&\tau&\varphi&\varphi&\tau&\tau\\ \tau&\tau&\varphi&\varphi&\tau&\tau&\varphi&\varphi\\ \tau&\tau&\varphi&\varphi&\tau&\tau&\varphi&\varphi \end{vmatrix} \quad \mathbf{\Phi_{(2)} = 2*(B_2)^{1/2}}$$

$$\begin{vmatrix} 3&2&3&2&3&2&3&2\\ 2&3&2&3&2&3&2&3\\ 3&2&3&2&3&2&3&2\\ 2&3&2&3&2&3&2&3\\ 3&2&3&2&3&2&3&2\\ 2&3&2&3&2&3&2&3\\ 3&2&3&2&3&2&3&2\\ 2&3&2&3&2&3&2&3 \end{vmatrix} \quad \mathbf{B_3} \qquad \begin{vmatrix} \varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi&\tau\\ \tau&\varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi\\ \varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi&\tau\\ \tau&\varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi\\ \varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi&\tau\\ \tau&\varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi\\ \varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi&\tau\\ \tau&\varphi&\tau&\varphi&\tau&\varphi&\tau&\varphi \end{vmatrix} \quad \mathbf{\Phi_{(3)} = 2*(B_3)^{1/2}}$$

Table 13: Above: octet matrices $\mathbf{B_1}$, $\mathbf{B_2}$, $\mathbf{B_3}$ received from the bi-periodic Table 2 for its elements in first, second and third positions of triplets correspondingly, for a case of hydrogen bonds numbers C=G=3, A=U=2. Below: three golden matrices $\mathbf{\Phi_{(j)}}$, all elements of which are equal to a golden section $\varphi = (1 + 5^{1/2})/2 = 1,618\ldots$ or to its reciprocal value $\tau = \varphi^{-1}$. Each matrix $\mathbf{\Phi_{(j)}}$ is equal to doubled matrix $(\mathbf{B_j})^{1/2}$, where $j = 1, 2, 3$.

On examination of the bi-periodic tables and binary sub-alphabets of genetic code, one can find their connection with Fibonacci numbers F_n, which are very essential for famous phyllotaxis laws of morphogenesis and which are related with a golden section. The author has constructed an extension (generalization) of famous Fibonacci **Q**-matrix (2x2):

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \qquad Q^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}$$

This extension is represented by families of Fibonacci $\mathbf{G}_i$-matrices with their orders (2^Kx2^K), where K is an integer number. Matrices $\mathbf{G}_i^n$ of these families have neighboring Fibonacci numbers as their matrix elements only. For example, three octet matrices $\mathbf{G}_i$ (i=1, 2, 3) are demonstrated below:

$\mathbf{G_1}$=

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	0	0	0	0
1	1	1	1	0	0	0	0
1	1	1	1	0	0	0	0
1	1	1	1	0	0	0	0

$\mathbf{G_2}$=

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0

$\mathbf{G_3}$=

1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	0
1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	0
1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	0
1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	0

A connection of these matrices with Fibonacci numbers is revealed when these matrices are raised to the power of n. For example,

$(\mathbf{G_2})^7 = 4^6 *$

21	21	13	13	21	21	13	13
21	21	13	13	21	21	13	13
13	13	8	8	13	13	8	8
13	13	8	8	13	13	8	8
21	21	13	13	21	21	13	13
21	21	13	13	21	21	13	13
13	13	8	8	13	13	8	8
13	13	8	8	13	13	8	8

The following equation is fulfilled for these matrices (in analogy with the recurrent equation for Fibonacci numbers):

$$(\mathbf{G_i})^{n+2} = (\mathbf{G_i})^{n+1} + (\mathbf{G_i})^n$$

In our opinion, Fibonacci **Q**-matrix is related to phenomena of a tetra-segregation with a segregation ratio 3:1 in genetic codes additionally (these phenomena are revealed in many aspects of genetic encoding systems from a meiosis of gametes till structures of

genetic codes and Jung's archetypes). Generalized Fibonacci $\mathbf{G}_i$-matrices seems to be useful in investigations of genetic codes.

10. A DISPOSITION OF AMINO ACIDS IN THE BI-PERIODIC OCTET TABLE OF TRIPLETS

Till this moment we spoke about nitrogenous bases and triplets. But triplets encode 20 amino acids and stop-codons, which can be inserted into the bi-periodic table of genetic code. Table 15 shows such inserting for a case of genetic code of human and vertebrate mitochondria. This variant of encoding of amino acids is described in genetic literature as the most ancient and "perfect" variant of genetic code (for example, see [Frank-Kamenetskii, 1988, p.65-68). It should be mentioned, that modern science knows a few variants of genetic codes, which are realized in different biological objects. All these variants are differed each from other slightly. One can see more details on this theme at special website "National Center for Biotechnology Information – The Genetic Codes".

One can see very symmetric disposition of amino acids in Table 15 with three kinds of symmetries. Firstly, this table consists of four pairs of neighbor rows with even and odd numbers: 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.

Secondly, each 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 new black-and-white mosaic for amino acids and stop-codons. This mosaic has a symmetric disposition with a figure of a diagonal cross: diagonal quadrants are equivalent to each other according to their mosaic.

1)	CCC Pro 63	CCA Pro 62	CAC His 61	CAA Gln 60	ACC Thr 59	ACA Thr 58	AAC Asn 57	AAA Lys 56
2)	CCU Pro 55	CCG Pro 54	CAU His 53	CAG Gln 52	ACU Thr 51	ACG Thr 50	AAU Asn 49	AAG Lys 48
3)	CUC Leu 47	CUA Leu 46	CGC Arg 45	CGA Arg 44	AUC Ile 43	AUA Met 42	AGC Ser 41	AGA Stop 40
4)	CUU Leu 39	CUG Leu 38	CGU Arg 37	CGG Arg 36	AUU Ile 35	AUG Met 34	AGU Ser 33	AGG Stop 32
5)	UCC Ser 31	UCA Ser 30	UAC Tyr 29	UAA Stop 28	GCC Ala 27	GCA Ala 26	GAC Asp 25	GAA Glu 24
6)	UCU Ser 23	UCG Ser 22	UAU Tyr 21	UAG Stop 20	GCU Ala 19	GCG Ala 18	GAU Asp 17	GAG Glu 16
7)	UUC Phe 15	UUA Leu 14	UGC Cys 13	UGA Trp 12	GUC Val 11	GUA Val 10	GGC Gly 9	GGA Gly 8
8)	UUU Phe 7	UUG Leu 6	UGU Cys 5	UGG Trp 4	GUU Val 3	GUG Val 2	GGU Gly 1	GGG Gly 0

Table 15: A disposition of 20 amino acids and stop-codons in the bi-periodic octet Table 2 for genetic code of human and vertebrate mitochondria. Short names of amino acids are Pro, His, Gln,… in this Table. “Stop” is a name of stop-codons. Each row has two pairs of neighbor cells with repeated amino acids (these pairs are marked by dark color).

Thirdly, in Table 15 left and right tabular halves are mirror-anti-symmetric to each other in its colors.

Strongly pronounced symmetric disposition of amino acids in our bi-periodic table of triplets testifies into a favor of adequacy of tensor multiplications of genomatrices in the field of genetic code.

11. METRIC GENOTENSORS, GENOTENSOR FIELDS AND A THEORY OF TENSOR BIOINFORMATICS

One of the most important facts of our investigations is a revealing a great meaning of

tensor multiplications of genomatrices to discover hidden regularities and symmetries in structures of genetic codes. For instance, such tensor multiplications have permitted to create interrelated sets of genetic *n*-plets (see Table 8); to reveal a family of numeric genomatrices connected with a family of golden matrices (see Table 12); to construct octet bi-periodic table of 64 triplets with symmetrical disposition of 20 amino acids and stop-codons there (see Table 15); to discover two branches of independent biological evolution within genetic codes, etc.

One can constructed many arbitrary "multiplicative" operations to produce the third matrix from initial two matrices. A tensor multiplication has an important difference: a tensor multiplication of two tensors produces a new tensor always. Tensor analysis knows a quadratic relation between two square matrices for metric tensors, which are one of the most important geometrical objects. From this viewpoint, interrelated families of genomatrices $\mathbf{M}^{(n)}$ and $\mathbf{\Phi}^{(n)}$ (Tables 9, 12) have a valuable tensor interpretation. Any genomatrix $\mathbf{M}^{(n)}$ can be interpreted as a metric tensor in 2^n-dimensional Euclidean space with those affine frame, which has basic vectors with their coordinates from rows (or from columns) of an appropriate golden matrix $\mathbf{\Phi}^{(n)}$. For instance, the genomatrix $\mathbf{M}^{(1)} = [3 \;\; 2; 2 \;\; 3]$ (Table 9, left) is a metric tensor of Euclidean plane in affine frame with "golden" basic vectors $e1(\varphi, \varphi^{-1})$ and $e2(\varphi^{-1}, \varphi)$, coordinates of which reproduce rows of the matrix $\mathbf{\Phi}^{(1)} = [\varphi \;\; \varphi^{-1}; \;\; i^{-1} \;\; \varphi]$ (here matrices are written in a line form as in a famous software MatLab).

All set of results, received due to tensor multiplications of genomatrices, testifies into a favor of a new viewpoint, proposed by the author, about a close relation between genetic encoding systems and metric tensors, or, briefly, about tensor nature of genetic encoding. It is known, that a metric tensor is the most important characteristic of geometrical properties of space. Metric tensors are the main tool in Riemannian geometries as well. Modern physics utilizes Riemannian geometries and a conception of metric tensor fields widely, first of all, in a theory of anisotropic physical fields and anisotropic mediums concerning to their elastic, optical, thermodynamic, dielectric, piezomagnetic and other physical magnitudes. In 1913 year A. Einstein and M.Grossmann equated a gravitational field with a 10-component metric tensor and proposed an appropriate theory of gravitation. From that time, a formal identification of physical fields with tensor fields became traditional one in science.

Note that chromosomes are anisotropic ones also. They have a sophisticated configuration of a bunch of helix anisotropic structures, physical-geometrical properties of which can be described by means of an appropriate tensor fields. Chromosomes are

bearers of many tensor fields. But all these peculiarities of chromosomes and many other genetic structures don't be taken into account by modern "sequential" approach to genetic codes. This very simplified "sequential" theory of genetic codes considers mainly sequences of nitrogenous bases in molecules of heredity and studies a correlation between their lists of triplets and sequences of amino acids in proteins. As a result this theory cannot study many important phenomena of heredity, related with physical-geometrical aspects: heredity of morphological forms of biological bodies, heredity of spatial representations, etc. Molecular genetics knows that such "sequential" approach, simplified from the very beginning, does not explain, how genes and their ensembles form a phenotype. Our knowledge about functions of genes is limited ordinarily by information how gene *X* encodes a heritable trait Y; we have null information on what is happened in an interval between genes and a phenotype, and it is true for all genes and all heritable traits. A tensor approach to genetic codes has much richer abilities in its beginnings than a "sequential" approach, which is not rejected by a tensor viewpoint but which is generalized in a frame of a genotensor theory developing by the author. This new theory interprets interrelated elements of genetic codes as special tensors, which are interrelated into an entire system by tensor multiplications and other tensor operations. Families of tensors $\mathbf{M}^{(n)}$ and $\mathbf{\Phi}^{(n)}$ (Tables 9, 12) play an important role in this theory. This tensor approach permits to propose, for example, formal models of relations between genetic code structures and morphogenesis phenomena (for instance, models of phyllotaxis on the base of Fibonacci and golden genomatrices). Besides all, it permits to consider from a viewpoint of genetic codes a classical problem of physiological foundations of geometry studied attentively by H. Poincaré and other scientists. This problem is related closely with famous phenomena of heritable knowledge about spatial metric in the case of human and animal organisms.

A heritable metrization of biological manifolds is a biological fact, which should be described by means of an appropriate geometrical (tensor) language. A biological organism is a tensor construction and it is naturally that tensor tools should encode it. According to described theory, a principle "tensors encode tensors" exists in living nature. A chain of biological generations can be considered as a chain of genotensor fields, and a biological evolution can be studied as a history of development of systems of genotensor fields. An appropriate language is a language of genotensor fields, theory of which should be combined with a theory of chronobiological rhythms. A tensor approach to heritable information is a valuable step for a famous problem of geometrization of physics and biology.

This way can lead to new understanding of a question "what is life?". For instance, can life be represented as a developing tensor-hierarchical system of genotensor fields? Can genetic codes be an evolutionary product of specific genotensor fields? The future will answer. It is probably that science will discover an irreducibility of genotensor fields to known physical fields (electric, magnetic, gravitational fields, etc.) and will prefer to consider genotensor fields as a new kind of physical fields.

Studying biological phenomena of phyllotaxis (related with Fibonacci numbers) from a viewpoint of metric genotensors, we paid attention that Fibonacci matrices

$$\mathbf{Q}_{\mathbf{left}}^{n} = [0\ 1;\ 1\ 1]^{n} \qquad \text{and} \qquad \mathbf{Q}_{\mathbf{right}}^{n} = [1\ 1;\ 1\ 0]^{n}$$

can be considered as metric tensors as well (they satisfy all demands for metric tensors). Each pair of Fibonacci matrices $\mathbf{Q}^{2n}$ and $\mathbf{Q}^{n}$ is a pair of a metric tensor and a matrix of basic vectors of a respective affine frame. This studying revealed new interesting formula of a tensor relation between the golden genomatrix $\mathbf{\Phi}$ (see Table 12) and two kinds of Fibonacci matrices

$$\mathbf{Q}_{\mathbf{left}} = [0\ 1;\ 1\ 1] \quad \text{and} \quad \mathbf{Q}_{\mathbf{right}} = [1\ 1;\ 1\ 0]: \mathbf{\Phi}^{2} = \mathbf{Q}_{\mathbf{left}}^{2} + \mathbf{Q}_{\mathbf{right}}^{2} \quad \text{or}$$

$$\begin{vmatrix} \varphi & \varphi^{-1} \\ \varphi^{-1} & \varphi \end{vmatrix}^{2} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix}^{2} + \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}^{2}$$

This formula demonstrates that metric tensor $\mathbf{\Phi}^{2}$, constructed on an affine frame with "golden" basic vectors $\mathbf{e}_1(\varphi, \varphi^{-1})$ and $\mathbf{e}_2(\varphi^{-1}, \varphi)$, is equal to a sum of two metric tensors, constructed on affine frames with "Fibonacci" basic vectors $\mathbf{u}_1(0, 1)$, $\mathbf{u}_2(1, 1)$ and $\mathbf{h}_1(1, 1)$, $\mathbf{h}_2(1,0)$, coordinates of which are taken from rows of Fibonacci matrices $\mathbf{Q}_{\mathbf{left}}$ and $\mathbf{Q}_{\mathbf{right}}$. Simultaneously this square formula is similar to an ordinary formula of length in vector space. It reveals new kind of a relation between the golden section and Fibonacci numbers as well.

An additional direction of development of theory of tensor bioinformatics is related with Hermite metric genotensors for unitary space. It takes place when appropriate complex numbers are introduced in genomatrices.

A briefly presented theory of tensor bioinformatics has many possible applications in biotechnology, computer informatics, etc.

12. PARALLELS WITH A THEORY OF BLOCK CODES

The described matrix representation of a system of genetic encoding gives also a possibility of its analysis in relation with advanced mathematical theory of block codes from computer informatics. In computers, a code is defined as a presentation of a set of symbols by sequences of units and zeros. One of many ways to provide a unique decoding of such sequences is an encoding all symbols by binary sequences of an equal length. Such code is named a block code. All elements of each considered genomatrices **P** and $\mathbf{P}^{(3)}$ (Table 8), which consist of 4 monoplets and 64 triplets correspondingly, are code-genetic sequences of equal length. But they are not binary sequences because they are formed by means four-letters alphabet (A, C, G, U/T) of genetic code. It complicates an application of effective theory of block codes to study them.

However, a genetic alphabet has three binary sub-alphabets (Table 1). Therefore, one can replace each symbolic *n*-plet by its binary equivalent from the viewpoint of one of three sub-alphabets. For instance, the symbolic triplet CGA is transformed into the binary triplet 100 from a viewpoint of the first sub-alphabet; or into the binary triplet 101 from a viewpoint of the second sub-alphabet; or into 110 from a viewpoint of the third sub-alphabet. By this way, each genomatrix $\mathbf{P}^{(n)}$ is a hidden union of three different binary matrices $\mathbf{P}^{(n)}{}_1$, $\mathbf{P}^{(n)}{}_2$, $\mathbf{P}^{(n)}{}_3$, which are corresponded to these sub-alphabets. The mathematical theory of block codes (with its concepts of generating matrices, Hamming distance, Gray code, dual code, etc.) can be applied to these binary matrices now.

13. A PROBLEM OF UNIFICATION BASES OF BIOLOGICAL LANGUAGES

The parallel presence of ensembles of the binary sub-alphabets at different fields of information physiology is marked: at genetics, at color perception in vision physiology, etc. For example, an ensemble of three binary oppositions exists in a color circle of visual perception; on the basis of this parallelism, the author offers to denote genetic triplets by color and to consider color variant of bi-periodic table of triplets (Petoukhov, 2001; 2002).

The special attention should be given to linguistics, in which there is long ago a concept of binary oppositions and their ensembles as structural base of the different linguistic languages (N.S. Trubetskoi, R.O. Jacobson, Ch.J.L. Baily, F. Jacob, etc.). For a long time there is an opinion, that languages of human dialogue were formed not on an empty place, and they are continuation of genetic language or, anyway, are closely

connected to it, confirming the idea of unification of information bases of organisms. The title of the monograph by Baily [1982] "*On the Yin and Yang nature of language*" is characteristic in this view. Makovskii marks in his "*Linguistic genetics*" (1992):

> "The opinion about language as about living organism, which is submitted to the natural laws of a nature, ascends to a deep antiquity … Research of a nature, of disposition and of reasons of isomorphism between genetic and linguistic laws is one of the most important fundamental problems for linguistics of our time".

In a view of large number of such physiologic and linguistic materials, the author put forward a hypothesis that ensembles of binary sub-alphabets play a role of unification (unitized) base not only for linguistic languages but also for all or for majority of biological languages. It seems that genotensor approach, described above, should be useful in this theme as well.

ACKNOWLEDGMENT

I am grateful to Frolov K., Darvas Gy., Szántó B., Böhm J., Ne'eman Y., He M., Kovács K. 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 the Hungarian and Russian academies of sciences for the years 2002-2004.

REFERENCES

Baily, Ch.J.L. (1982) *On the Yin and Yang nature of language*. – Ann Arbor. Frank-Kamenetskii, M.D. (1988) *The most principal molecule*, Moscow, Nauka, 174 p.

Gazale, M.J. (2000) Gnomon. From *Pharaons to Fractals*. – Pergamon Press, Princeton.

Jacob, F. (1974) Le modele linquistique en biologie, *Critique, Mars, tom*e XXX, № **322**, 306-315.

He, M. (2003) Symmetry in Structure of Genetic Code. *Proceedings of the Third All-Russian Interdisciplinary Scientific Conference"Ethics and the Science of Future. Unity in Diversity", Moscow, February 12-14, 2003*.

Konopel'chenko, B.G. and Rumer, Y.B. (1975) Classification of codons in genetic code. *Dokl. Academii Nauk SSSR*, 1975, **223**, № 2, p.471-474.

Makovskii, M.M. (1992) *Linguistic Genetics*. Moscow, Nauka (in Russian).

Petoukhov, S.V. (1999) Genetic Code and the Ancient Chinese "Book of Changes", *Symmetry: Culture and Science*. **10**, № 3-4, p.211-226.

Petoukhov, S.V. (2001) *The Bi-periodic Table of Genetic Code and Number of Protons*, Foreword of K. V. Frolov. Moscow, 258 p. (in Russian).

Petoukhov, S.V. (2002) Binary sub-alphabets of genetic language and problem of unification bases of biological languages, *IX International Conference "Mathematics, computer, education"*. Russia, Dubna, January 28-31, **191** (in Russian).

Petoukhov, S.V. (2003) The Bi-periodic Table and Attributive Conception of Genetic Code. Problem of Unification Bases of Biological Languages. *Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences METMBS-2003, Las Vegas, Nevada, USA, June 23-26,* 2003. p. 495-501.

Wilhelm, H. (1960) *Change: Eight Lectures on the I Ching*, N.Y.