Symmetry: Culture and Science
Vol. 16, No. 3, 247-266, 2005

HADAMARD MATRICES AND QUINT MATRICES IN MATRIX PRESENTATIONS OF MOLECULAR GENETIC SYSTEMS

Sergey 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

Genetic information is transferred by means of discrete elements: 4 letters of genetic alphabet, 64 amino acids, etc. General theory of processing of discrete signals utilizes encoding such signals by means of special mathematical matrices and spectral representation of signals with the main aim of increase of reliability and efficiency of information transfer but not for prevention of reading this information by undesirable reader at all [Sklar, 2001; Ahmed, Rao, 1975, etc]. A typical example of such matrices with appropriate properties is the family of Hadamard matrices. Rows of such matrices form an orthogonal system of functions Hadamard-Walsh, which is used for a spectral presentation and transfer of discrete signals [Ahmed, Rao, 1975; Geramita, 1979; Yarlagadda, Hershey, 1997].

The author investigates molecular structures of genetic code from the viewpoint of matrix methods of encoding and transfer of discrete signals. His results, which have been obtained in this original way, reveal hidden interconnections, symmetries and evolutionary invariants in genetic code systems [Petoukhov, 2001-2005; He, Petoukhov, Ricci, 2004]. Simultaneously they testify in a favor of that genetic molecules are important part of specific maintenance of noise immunity and efficiency of discrete information transfer. These molecular structures are related also with the general harmony of organism.

A Hadamard matrix of dimension "n" is a square $(n \times n)$-matrix $H(n)$ with elements +1 and -1; it satisfies to a condition $H(n)*H(n)^T = n*I_n$, where $H(n)^T$ is the transposed

matrix and I_n is the *(n x n)*-identity matrix. The Hadamard matrices of dimension 2^k are given by the recursive formula $H(2^k) = H(2)^{(k)} = H(2) \otimes H(2^{k-1})$ for $2 \leq k \in N$, where $\otimes$ denotes the Kronecker product, brackets at exponent *(k)* mean the Kronecker exponentiation, *k* and *N* are integers, *H(2)* is demonstrated in Fig.1.

$H(2) =$	1	1	; $H(4) =$	1	1	1	1	; ... $H(2^k) =$	$H(2^{k-1})$	$H(2^{k-1})$
	-1	1		-1	1	-1	1		$-H(2^{k-1})$	$H(2^{k-1})$
				-1	-1	1	1			
				1	-1	-1	1			

Figure 1: The family of Hadamard matrices $H(2^k)$ based on the Kronecker product.

Rows of a Hadamard matrix are mutually orthogonal. This means that every two different rows in a Hadamard matrix represent two perpendicular vectors. The element *"*-1*"* can be disposed in any of four positions in the Hadamard matrix *H(2)* without a loss of main matrix properties. Such matrices are used in many fields due to their advantageous properties: in error-correcting codes such as the Reed-Muller code; in spectral analysis and multi-channels spectrometers with Hadamard transformations; in quantum computers with Hadamard gates (or logical operators), etc. The author revealed unexpectedly that Hadamard matrices reflect essential peculiarities of molecular genetic systems.

MATRIX PRESENTATION OF GENETIC POLYPLETS BASED ON THE KRONECKER PRODUCT

It is well-known, that polyplets are one of main peculiarities of genetic code. Really, the alphabet of genetic code is a set of four monoplets (nitrogenous bases): A (adenine), C (cytosine), G (Guanine), T/U (thymine in DNA or uracil in RNA); 64 triplets encode amino acids; each protein is encoded by more or less long polyplets (each protein with *n* amino acids is encoded by a 3*n*-plet). Due to the idea of a possible analogy between discrete signals processing in computers and in a genetic code system, the author has represented all sets of genetic polyplets as appropriate parts of a mutual family of symbolic square matrices (genomatrices) $P^{(n)}$, where brackets at exponent (*n*) mean the Kronecker exponentiation [Petoukhov, 2001-2005].

CC	*CA*	*AC*	*AA*

$P =$

C	A
T	G

; $P^{(2)} = P \otimes P =$

CT	CG	AT	AG
TC	TA	GC	GA
TT	TG	GT	GG

$P^{(3)} =$

CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
CCT	CCG	CAT	CAG	ACT	ACG	AAT	AAG
CTC	CTA	CGC	CGA	ATC	ATA	AGC	AGA
CTT	CTG	CGT	CGG	ATT	ATG	AGT	AGG
TCC	TCA	TAC	TAA	GCC	GCA	GAC	GAA
TCT	TCG	TAT	TAG	GCT	GCG	GAT	GAG
TTC	TTA	TGC	TGA	GTC	GTA	GGC	GGA
TTT	TTG	TGT	TGG	GTT	GTG	GGT	GGG

Figure 2: The beginning of the family of symbolic genomatrices $P^{(n)}$ for $n = 1, 2, 3$. Here (n) means the Kronecker exponentiation. This family is transformed into a family of Hadamard matrices from the viewpoint of essential parameters of nitrogenous bases (see below).

The basic *(*2x2*)*-genomatrix *P* of this family (Fig. 2) is constructed from four letters A, C, G, T/U. Each genomatrix $P^{(n)}$ contains a complete set of *n*-plets as its matrix elements. For example, the (8x8*)*-genomatrix $P^{(3)}$ contains all 64 triplets which encode *20* amino acids (for more details about this matrix presentation of genetic subsystems, see [Petoukhov, 2003, 2005]).

This matrix representation of genetic polyplets seems to be adequate because of the following reasons:

- It represents complete sets of *n*-plets in the universal mathematical form;
- Each column in the matrix $P^{(3)}$ represents one of eight classical octets by Wittmann [1961], which reflect real biochemical properties of triplets.
- This representation has revealed unexpectedly a very symmetrical disposition of 20 amino acids in the formal constructed genomatrix $P^{(3)}$ in the case of the vertebrate mitochondria genetic code [Petoukhov, 2001, 2003] which is considered mainly as the most ancient and "perfect" variant of genetic code.
- Moreover, this matrix representation has revealed a division of a set of 20 amino acids into two important subsets: a subset of 8 high-degenerated amino

acids, which are encoded by 4 or more triplets, and a subset of 12 low-degenerated amino acids, which are encoded by 3 or less triplets. This division into two subsets with 8 and 12 amino acids is an evolution invariant for the set of 20 amino acids according to investigation of quantities of low-degenerated and high-degenerated amino acids in all 17 variant of genetic codes which are known in modern science (see more details in [Petoukhov, 2001-2005]).

To develop matrix analyses of genetic polyplets, one can investigate properties of the basic genomatrix *P* of four genetic letters more deeply. 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. In our opinion, this choice has a deep sense, and in this reason symmetric and asymmetric peculiarities of polyatomic constructions in the alphabetic system of four nitrogenous bases (Fig. 3) can give essential helps to reveal secrets of genetic code. Let us pay attention to these peculiarities.

HADAMARD GENOMATRICES

It is known that amino-groups NH_2 play an important role in molecular genetics. For instance, the amino-group in amino acids provides a function of recognition of this amino acid by ferment [Shapeville F., Haenni A.-L., 1974]. A detachment of amino-groups in nitrogenous bases A and C in RNA under action of nitrous acid HNO_2 determines a property of amino-mutating of these bases, which was used to divide 64 triplets in eight natural families with 8 triplets in each in the work mentioned above [Wittmann, 1961].

But how the amino-groups are represented in genetic alphabets? One can note that each of three nitrogenous bases A, C, G has one amino-group (these groups are marked by circles in Fig. 3) and one base U/T has not it. This fact of existence or absence of amino-group in certain genetic letters can be represented in alphabetic genomatrix P by symbols “+1” and “-1” instead letters A, C, G and U/T correspondingly. In this case the

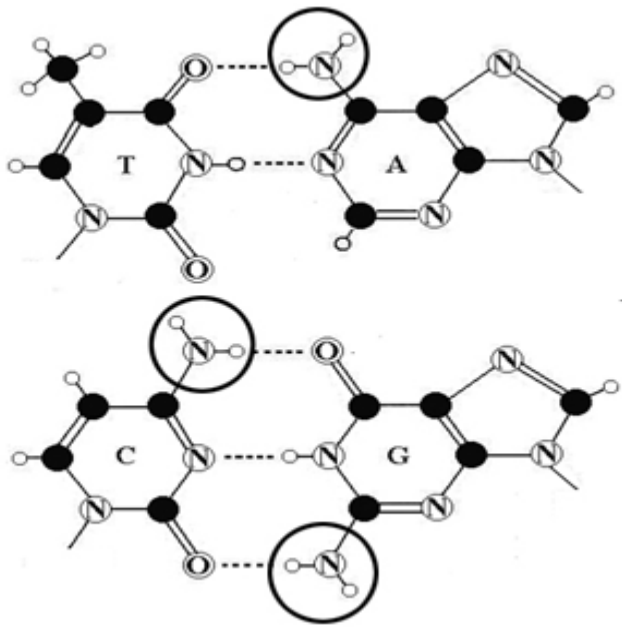

Figure3: Complementary pairs of four nitrogenous bases in DNA: A - T (adenine and thymine), C - G (cytosine and guanine). All amino-groups are marked by three big circumferences. Hydrogen bonds in complementary pairs are specified by dotted lines. Black small circles are atoms of carbon, small white circles - hydrogen, circles with the letter N - nitrogen, and circles with the letter O – oxygen.

Hadamard genomatrix $P_{H(2)} = H(2)$ is appeared (Fig.1). It is obvious that all genomatrices $P_{H(2)}^{(n)}$ will be Hadamard matrices as well. One can suppose that Hadamard genomatrices can be used in genetic systems by analogy with applications of Hadamard matrices in different fields of science and technology generally speaking: discrete information processing, quantum computers, multi-channel spectrometers, error-correcting and other codes, etc.

Simultaneously other facts can be noted about relations between genetic code system and Hadamard matrices. They are connected with a symmetrical disposition of 20 amino acids in genomatrix $P^{(3)}$ of 64 triplets in the case of the vertebrate mitochondria genetic code [Petoukhov, 2001-2003], which is considered traditionally as the most ancient and "perfect" variant of genetic code. $P^{(3)}$ was constructed above by means of the mathematical operation with the genomatrix P of the genetic alphabet without a mention about amino acids at all. But unexpectedly the disposition of these 20 amino

acids in the genomatrix $P^{(3)}$ (Fig. 4, right) has demonstrated not only very symmetrical character, but relations with Hadamard matrices also. Let us show a few of these facts.

Each sub-quadrant (2x2) of the matrix $P^{(3)}$ contains a subfamily of those four triplets, which are equivalent to each other by two first letters. Such quadruple of triplets will be named “a subfamily of NN-triplets”. A complete set of 16 subfamilies of NN-triplets is segregated by the nature into two subsets with 8 subfamilies in each. The first subset, marked by dark on Fig. 4, contains those subfamilies of NN-triplets, coded values of which are independent of their third letter. In this reason, all four triplets of such subfamily encode the same amino acids. On the contrary, the second subset, marked by white color, contains those subfamilies of NN-triplets, coded values of which are dependent of their third letter. In this reason, each such subfamily has triplets, which encode different amino acids or stop-signals.

A disposition of these “black” and “white” subfamilies of NN-triplets in the matrix $P^{(3)}$ is very symmetric. For instance, left and right halves of this matrix 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. Diagonal quadrants of the matrix are equivalent to each other from the viewpoint of their mosaic. The rows 1-2, 3-4, 5-6, 7-8 are equivalent to each other from the viewpoint of a *disposition* of the same amino acids in their appropriate cells, etc. From the set of 20 amino acids, 8 amino acids (Ala, Arg, Gly, Leu, Pro, Ser, Thr, Val) belong to the “black” sub-families of NN-triplets and other *12* amino acids (Asn, Asp, Cys, Gln, Glu, His, Ile, Lys, Met, Phe, Trp, Tyr) are presented in the “white” sub-families. These two subsets of amino acids are evolutionary invariant practically in relation to 17 variants of genetic codes, which are known in modern science (see details in [Petoukhov, 2001-2005].

Fig.4, left demonstrates the black-and-white genomatrix $P^{(2)}$ of 16 duplets, where black (white) cells correspond to those duplets, two letters of which are the beginning of appropriate black (white) families of NN-triplets of the genomatrix $P^{(3)}$ on Fig.4, right.

One can consider that each black (white) cell of matrices in Fig.4 is equal to +1 (-1). In this case many relations between these genomatrices and Hadamard matrices can be demonstrated. Firstly, each black-and-white row in these both matrices correspond to an appropriate Rademacker function, which are connected with Hadamard-Walsh functions of Hadamard matrices (see [Ahmed, Rao, 1975] about Rademacker functions and their connections with rows of Hadamard matrices). Each quadrant (2x2) of the genomatrix (4x4) in Fig.4 is identical to an appropriate Hadamard matrix (2x2).

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

;

CCC PRO	CCA PRO	CAC HIS	CAA GLN	ACC THR	ACA THR	AAC ASN	AAA LYS
CCU PRO	CCG PRO	CAU HIS	CAG GLN	ACU THR	ACG THR	AAU ASN	AAG LYS
CUC LEU	CUA LEU	CGC ARG	CGA ARG	AUC ILE	AUA MET	AGC SER	AGA STOP
CUU LEU	CUG LEU	CGU ARG	CGG ARG	AUU ILE	AUG MET	AGU SER	AGG STOP
UCC SER	UCA SER	UAC TYR	UAA STOP	GCC ALA	GCA ALA	GAC ASP	GAA GLU
UCU SER	UCG SER	UAU TYR	UAG STOP	GCU ALA	GCG ALA	GAU ASP	GAG GLU
UUC PHE	UUA LEU	UGC CYS	UGA TRP	GUC VAL	GUA VAL	GGC GLY	GGA GLY
UUU PHE	UUG LEU	UGU CYS	UGG TRP	GUU VAL	GUG VAL	GGU GLY	GGG GLY

Figure 4: Right: a representation of the genomatrix $P^{(3)}$ of 64 triplets for a case of the vertebrate mitochondria genetic code. The matrix consists of 64 triplets and 20 amino acids with their traditional abbreviations. Stop-codons are marked as "stop". Left: a representation of the appropriate genomatrix $P^{(2)}$ of 16 duplets.

Both matrices in Fig. 4 are become Hadamard matrices as a whole by means of the following simple algorithm, which is connected with the special value "*-1*" of the letter T, mentioned above (see comment to the Fig. 3): each duplet or triplet in the black-and-white genomatrices of Fig.4 should change its color into opposite color each time when the letter T stands in an odd position (in the first or in the third position) inside such polyplet. For example, the triplet TTA, which is in the white cell of the genomatrix $P^{(3)}$ in Fig. 4, is disposed in the black cell in Fig. 5 because of the letter T in its first position. Or the triplet TTT, which is in the white cell of the genomatrix $P^{(3)}$ in Fig. 4, is disposed in the white cell in Fig. 5 because of the letter T in its first and third positions (the color of this triplet is changed twice according to the described algorithm). The triplet ACG does not change its color because the letter T is absent in this triplet at all, etc.

One can check that both matrices in Fig. 5 satisfy to the general condition of Hadamard matrices $H(n)*H(n)^T = n*I_n$, mentioned above. Application of the same algorithm to the Hadamard matrices from Fig. 5 transfers them into the mosaic genomatrices of Fig. 4.

Each (2x2)-quadrant (or sub-quadrant) of these matrices in Fig.5 is a Hadamard matrix (2x2) as well. Each quadrant (4x4) in the matrix (8x8) in Fig.5 is a Hadamard matrix

also (this situation can be named “Hadamard fractals” conditionally). It means that the complete genomatrices of duplets and triplets in Fig. 5 consist of smaller Hadamard matrices inserted into them. Perhaps described algorithm of sign inversion is connected with a biological mechanism of a mutual change of the letters U and T in a course of transition of genetic sequences from RNA to DNA.

CC	*CA*	*AC*	*AA*
CT	*CG*	*AT*	*AG*
TC	*TA*	*GC*	*GA*
TT	*TG*	*GT*	*GG*

;

CCC	***CCA***	***CAC***	***CAA***	***ACC***	***ACA***	***AAC***	***AAA***
CCT	***CCG***	***CAT***	***CAG***	***ACT***	***ACG***	***AAT***	***AAG***
CTC	***CTA***	***CGC***	***CGA***	***ATC***	***ATA***	***AGC***	***AGA***
CTT	***CTG***	***CGT***	***CGG***	***ATT***	***ATG***	***AGT***	***AGG***
TCC	***TCA***	***TAC***	***TAA***	***GCC***	***GCA***	***GAC***	***GAA***
TCT	***TCG***	***TAT***	***TAG***	***GCT***	***GCG***	***GAT***	***GAG***
TTC	***TTA***	***TGC***	***TGA***	***GTC***	***GTA***	***GGC***	***GGA***
TTT	***TTG***	***TGT***	***TGG***	***GTT***	***GTG***	***GGT***	***GGG***

Figure 5: Two Hadamard matrices, which are received algorithmically from appropriate genomatrices of Fig. 4. Black (white) cells are equal to +1 (-1) in a traditional presentation of Hadamard matrices (explanation in the text).

In Fig. 5 one can note additionally that the right matrix (8x8) can be obtained from the left matrix (4x4) by a substitution of Hadamard matrix $H(2)$ (see Fig.1) into all black cells of the matrix (4x4) and a simultaneous substitution of inverted Hadamard matrix *“-$H(2)$”* into all white cells of the same matrix.

Rows of Hadamard matrices represent orthogonal systems of Hadamard-Walsh functions. Such orthogonal system can be a natural base to organize storage and transfer of discrete genetic information by means of decomposition of genetic sequences with these orthogonal systems and by means of using of orthogonal and bi-orthogonal codes. Genetic molecules are objects of quantum mechanics, where normalized Hadamard matrices play important role as unitary operators (it is known that an evolution of closed quantum system is described by unitary transformation). In particularly, quantum computers use these matrices as Hadamard gates (logical operators, Hadamard element, etc.) [Nielsen M.A., Chuang I.L., 2001]. In this connection new theoretical possibilities are appeared to transfer achievements of quantum computer conceptions into the field of molecular genetics and to consider genetic system as a quantum computer.

From the viewpoint of significance of quantum mechanics and its unitary operators, first of all, Hadamard operators, a possible answer on a fundamental question about reasons for the nature to choose four-letters genetic alphabet is the following one: the important reason is that simplest unitary matrices in two-dimensional space, first of all, Hadamard matrices (and also Pauli matrices, etc.) consist of four elements exactly. It

seems very probably that principles of quantum mechanics and quantum computers underlie structural peculiarities of genetic code.

It is essential that orthogonal systems of Hadamard-Walsh functions are also revealed by a few authors in macro-physiological systems (visual systems and others) which should be agreed structurally with genetic system for transferring along a chain of generations [Shiozaki, 1980; Carl J.V., 1974; Ginsburg A.P. et all, 1974; etc.].

This article presents only a few revealed facts of realization of Hadamard genomatrices on the base of natural parameters of molecular-genetic system. On the whole they testify that many of advantages of Hadamard matrices, which are used widely in many fields of science and technology, can be exploited generally speaking in genetic system and processes of biological self-organization, for example, for spectral analysis of genetic sequences and genetic role of emission spectrums of genetic elements. Described facts are valuable arguments that polyatomic peculiarities of genetic letters are chosen by the nature to provide reliability and effectiveness of heredity information transfer by those methods which use Hadamard matrices and spectral presentation of discrete information signal.

It should be noted that Hadamard genomatrices are appeared not only if one takes the case of an attribute of "existence or absence of amino-group" in nitrogenous bases. Hadamard genomatrices are appeared also if one takes into account an attribute of "existence or absence of oxygen atoms" in nitrogenous bases: adenine A has not such atom and other three nitrogenous bases have them (Fig.3). Hadamard matrices are appeared in DNA also in the case of an attribute "existence or absence of five atoms of carbon" because only cytosine C has not them. It seems that described facts are connected additionally with an actual theme of DNA computers [Bebenson, Shapiro, 2006].

By the way, one can note that black-and-white genomatrix *(4x4)* on Fig. 4 has an identical mosaic for each pair of quadrants *(2x2)* along matrix diagonals. This non-orthogonal genomatrix can be represented as a sum of two orthogonal matrices, each of which is named "uncontrolled gate" in the field of quantum computers due to their special meaning (http://en.wikipedia.org/wiki/Quantum_gate):

+1	-1	+1	-1

+1	-1	0	0

0	0	+1	-1

+1	+1	-1	-1	=	+1	+1	0	0	+	0	0	-1	-1
+1	-1	+1	-1		0	0	+1	-1		+1	-1	0	0
-1	-1	+1	+1		0	0	+1	+1		-1	-1	0	0

Figure 6: A representation of non-orthogonal genomatrix from Fig.4 as a sum of two orthogonal matrices.

QUINT GENOMATRICES

The described family of Hadamard genomatrices is a particular case of a family of genomatrices $P^{(n)}$ mentioned above. What types of numerical matrices are appeared if one takes other essential parameters of nitrogenous bases to install them into $P^{(n)}$?

Each pair of complementary nitrogenous bases has special number of hydrogen bonds, which were suspected in their important information meaning long ago: C and G have 3 hydrogen bonds, A and T have 2 ones. If we replace each genetic letter in alphabetic genomatrix P by the numbers of its hydrogen bonds (C=G=3, A=U=2), we receive a numeric genomatrix P_{MULT} (Fig.2) and a corresponding family of genomatrices $P_{MULT}^{(n)}$. Fig. 7 demonstrates the matrices P_{MULT} and $P_{MULT}^{(3)}$.

P_{MULT} =

3	**2**
2	**3**

; $P_{MULT}^{(3)}$ =

27	**18**	**18**	**12**	**18**	**12**	**12**	**8**
18	**27**	**12**	**18**	**12**	**18**	**8**	**12**
18	**12**	**27**	**18**	**12**	**8**	**18**	**12**
12	**18**	**18**	**27**	**8**	**12**	**12**	**18**
18	**12**	**12**	**8**	**27**	**18**	**18**	**12**
12	**18**	**8**	**12**	**18**	**27**	**12**	**18**
12	**8**	**18**	**12**	**18**	**12**	**27**	**18**
8	**12**	**12**	**18**	**12**	**18**	**18**	**27**

Figure 7: The numeric genomatrices $P_{MULT}^{(1)}$ and $P_{MULT}^{(3)}$ as the representatives of the family of quint genomatrices $P_{MULT}^{(n)}$.

The bisymmetric genomatrices $P_{MULT}^{(n)}$ have interesting mathematical properties [Petoukhov, 2001-2005]. This article takes into account only one of them. Each genomatrix of the family $P_{MULT}^{(n)}$ has the ratio 3:2 at different levels: between numerical sums in top and bottom quadrants, sub-quadrants, sub-sub-quadrants, etc. including quint ratios between neighbor numbers in them. For example, $P_{MULT}^{(3)}$ contains 4

numbers – 27, 18, 12, 8 - with the quint ratio between them: 27/18=18/12=12/8=3/2. The ratio 3:2 is named the quint (or the perfect fifth) in a theory of musical harmony. In this reason described genomatrices $P_{MULT}^{(n)}$ are named "quint genomatrices" conditionally.

Each quint genomatrix $P_{MULT}^{(n)}$ contains (n+1) kinds of numbers from a geometrical progression, factor of which is equal to the quint 3/2 [Petoukhov, 2003-04, tables 9-11]:

$P_{MULT}^{(1)} \Rightarrow$ 3, 2
$P_{MULT}^{(2)} \Rightarrow$ 9, 6, 4
$P_{MULT}^{(3)}) \Rightarrow$ 27, 18, 12, 8 (1)
…………………………………………….
$P_{MULT}^{(6)} \Rightarrow$ 729, 486, 324, 216, 144, 96, 64
…………………………………………………………….

In other words, such sequences reproduce quint scales which were known previously in theory of musical harmony (first of all, in Pythagorean musical scale and in old Chinese music scale [Needham, 1962]).

The four numbers 8=2*2*2, 12=2*2*3, 18=2*3*3, 27=3*3*3, which are presented in the genomatrix $P_{MULT}^{(3)}$ on Fig. 7, characterize those four kinds of triplets, which are differed by their numbers of hydrogen bonds of nitrogenous bases. For instance, number 18=2*3*3 belongs to those triplets, which have one nitrogenous base with *2* hydrogen bond and two bases with *3* hydrogen bonds (mathematics of genomatrices is testify in a favor of that products of numbers of hydrogen bonds should be taken into attention here but not their sums; it has precedents and the justification in information theories, in particular, in the theory of parallel channels of encoding and processing of the information). Each gene and each part of DNA have their own numerical sequences of such numbers 8, 12, 18, 27 of triplets with the quint relation between them. This fact is a part of author's materials about "music of genetic information" or "genomusic" mentioned in his previous publications [Petoukhov, 1999; 2001a, pp.43, 224, 225; 2001b]. Now the author conducts experiments about a possible physiological activity of acoustical sequences of musical sounds which are constructed artificially by means of computer on the base of appropriate numerical sequences of numbers 8, 12, 18, 27 of triplets along genes. He creates a computer bank of such "music" sequences of genes and proteins for theoretical and practical needs now.

The author notes a symmetrological fact also that the quint is the typical ratio for two other important parameters of DNA, which form their own quint sequences along DNA:

- Pyrimidines C and T have 40 protons in their rings; purines A and G have 60 protons in their rings [Petoukhov, 1999]. (Each complementary pair has 100 protons in their rings precisely). The ratio 60:40 is equal to the quint 3:2.
- Quantity of non-hydrogen atoms in molecular rings of pyrimidines C and T is equal to 6, and quantity of non-hydrogen atoms in molecular rings of purines A and G is equal to 9. The ratio 9:6 is equal to the quint 3:2.

Special families of quint genomatrices by analogy with $P^{(n)}$ can be founded for these parameters easily.

All physiological systems should be coordinated structurally with genetic code for their encoding and transfer to next generations and for a survival in a course of biological evolution. In this reason we collect examples of the quint (and other main harmonic ratios) in structures and functions of supra-molecular biological systems. For example, the quint 3:2 exists between:

- durations of phases of activity and rest in a cardio-cycle (0.6 sec and 0.4 sec correspondingly);
- plasmatic and globular volumes of blood (60% and 40%);
- albumens and globulins of blood (60% and 40%).

Let us say one additional remark. If one takes into account that the nitrogenous base T/U can be characterized by number "-1" because it has no amino-group (see the first part of this article and Fig.3) then another family of numerical genomatrices $T_{MULT}^{(n)}$ is appeared from genomatrices $P^{(n)}$ (Fig. 8), where the letter T from the matrix P is symbolized by number "-2". Rows of $T^{(n)}$ are mutually orthogonal and, from theoretical viewpoint, this matrix can be used in principle in discrete signals processing to increase reliability and effectiveness of discrete signals transfer by means of a spectral presentation of signals, etc.

THE MATHEMATICAL SCALE OF THE GOLDEN WURF

Quint genomatrices $P_{MULT}^{(n)}$ have a hidden relation with the famous golden section $\varphi = (1+5^{0.5})/2 = 1{,}618\ldots$ [Petoukhov, 2001b]. If we take the square root from any genomatrix $P_{MULT}^{(n)}$ in ordinary sense, the result is a bi-symmetric matrix $\Phi_{MULT}^{(n)} = (P_{MULT}^{(n)})^{1/2}$, the elements of which are equal to the golden section φ in different integer powers (Fig. 9). In this reason they are named the golden genomatrices.

T_{MULT} = | 3 | 2 |
| -2 | 3 |

; $T_{MULT}^{(3)}$ =

27	18	18	12	18	12	12	8
-18	27	-12	18	-12	18	-8	12
-18	-12	27	18	-12	-8	18	12
12	-18	-18	27	8	-12	-12	18
-18	-12	-12	-8	27	18	18	12
12	-18	8	-12	-18	27	-12	18
12	8	-18	-12	-18	-12	27	18
-8	12	12	-18	12	-18	-18	27

Figure 8: The numeric genomatrices T_{MULT} and $T_{MULT}^{(3)}$ as the representatives of the family of quint genomatrices $T_{MULT}^{(n)}$ with mutually orthogonal rows (explanation in the text).

$(P_{MULT})^{1/2}$ = Φ_{MULT}

φ	φ^{-1}
φ^{-1}	φ

; $(P_{MULT}^{(3)})^{1/2}$ = $\Phi_{MULT}^{(3)}$ =

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

Figure 9: The numeric genomatrices Φ_{MULT} and $\Phi_{MULT}^{(3)}$ as the representatives of the family of the golden genomatrices $\Phi_{MULT}^{(n)}$.

For instance, Fig. 9 demonstrates the matrix $\Phi_{MULT}^{(3)} = (P_{MULT}^{(n)})^{1/2}$. This matrix has only two pairs of inverse numbers: φ^1 and φ^{-1}, φ^3 and φ^{-3}. The golden section φ is a mathematical symbol of a self-reproduction for many centuries. It is well known that the golden section is shown in different physiological systems: cardio-vascular systems, respiratory systems, electric activities of brain, locomotion activity, *etc*. The discovered matrix relation between the golden section φ and basic parameters of genetic codes testifies into a favor of a molecular-genetic providing such physiological phenomena.

This fact of close formal connection between genomatrices and the golden section seems to be interesting additionally from the viewpoint of an actual problem of genetic bases of aesthetics and inborn feeling of harmony additionally. According to words by famous physicist Richard Feynman about feeling of musical harmony, "we may question whether we are any better off than Pythagoras in understanding why only certain sounds are pleasant to our ear. The general theory of aesthetics is probably no further advanced now than in the time of Pythagoras" [Feynman, 1963, Chapter 50].

Just as a quint genomatrix $P_{MULT}{}^{(n)}$ contains a sequence (1) of $(n+1)$-kinds of numbers from a geometrical progression with the quint coefficient *3/2*, a corresponding golden genomatrix $\Phi^{(n)}$ contains a sequence of $(n+1)$-kinds of numbers from a geometric progression, the coefficient of which is equal to $\varphi^2 = 2{,}618\ldots$.:

$$\begin{aligned}
&\Phi^{(1)} \Rightarrow \varphi^1, \ \varphi^{-1} \\
&\Phi^{(2)} \Rightarrow \varphi^2, \ \varphi^0, \ \varphi^{-2} \\
&\Phi^{(3)} \Rightarrow \varphi^3, \ \varphi^1, \ \varphi^{-1}, \ \varphi^{-3} \qquad (2) \\
&\ldots\ldots\ldots\ldots
\end{aligned}$$

The journal "Symmetry; Culture and Science" considers traditionally how methods and algorithms of symmetry permit to reveal structural analogies in different fields of science and culture. It is known that quint sequences (1) lead by means of a special algorithm to a construction of the Pythagorean musical scale which is a set of special tone-intervals and semitone-intervals (two kinds of frequency ratios 9/8 and 256/243 between adjacent musical notes) inside one octave (for example see [Hightower; Pont; Voloshinov, 2000]). One can show that the same Pythagorean algorithm permits to get new interesting mathematical scale from the new sequences (2). This new scale is similar to the Pythagorean musical scale with small structural differences. It is interesting from the viewpoint of the theory of musical harmony.

More precisely, a known algorithm of creation of a sequence of interval coefficients of Pythagorean musical scale consists of the following stages:

1. Construction of the first members of geometrical progression with the quint coefficient 3/2 beginning from inverse value of the quint: $(3/2)^{-1}$: 2/3, 1, 3/2, 9/4, 27/8, 81/16, 243/32, … ;
2. Transference of values of all these members into a base octave (with values from 1 to 2) by means of their division or multiplication with octave number *2* the necessary number of times (two members - 1 and 3/2 – of the sequences lay inside of an octave initially because the coefficient 3/2 less than 2). As a result of this operation, a new sequence is appeared (this sequence can be named "a geometrical progression with return into an octave "): 4/3, 1, 3/2, 9/8, 27/16, 81/64, 243/128...;
3. Rearrangement of first seven members of this sequence according to their increasing values from 1 to 2 (ends of an octave 1 and 2 are included in this sequence):

1, 9/8, 81/64, 4/3, 3/2, 27/16, 243/128, 2 (3)

In this sequence, a ratio of greater number to the neighbor smaller number refers to as interval factor. Two kinds of interval factors exist in this sequence only: 9/8, which is named tone-interval T, and 256/243, which is named semitone-interval *S*. A sequence of interval factors in (3) is *T-T-S-T-T-T-S.* These five tone-intervals and two semitone-intervals cover an octave precisely: $(9/8)^5 * (256/243)^2 = 2$.

Let us construct a new mathematical scale by means of this algorithm, using (instead of the quint coefficient 3/2) a coefficient $p = \varphi^2/2$. From the viewpoint of construction of "a geometrical progression with return into an octave", this coefficient p is equivalent to a coefficient φ^2. This is so, because the coefficient φ^2 is more than 2 and in this reason it should be divided by 2 to return into the main octave interval from 1 to 2 according to ordinary Pythagorean rule. In this case the coefficient *p* and the quint 3/2 belong both to this main octave interval. All three stages of the Pythagorean algorithm are reproduced:

1. A sequence of the first members of a geometrical progression with a coefficient $p = \varphi^2/2$, *is:* $p^{-1}, p^0, p, p^2, p^3, p^4, p^5, p^6, ,\ldots$
2. A sequence with returned members into the base octave by means of their division or multiplication with *2* the necessary number of times: $(1/p)*2$ *;* $p^0, p, p^2, p^3/2, p^4/2, p^5/2, p^6/4, \ldots$
3. Rearrangement of first members of this sequence according to their increasing values from 1 to 2 (ends of an octave 1 and 2 are included in this sequence):

1, $p^3/2$, $p^6/4$, p, $p^4/2$, $(1/p)*2$, p^2, $p^5/2$, 2 (4)

Two kinds of interval factors exist in this sequence (4) only: $p^3/2 = 1.1215...$ (new tone-interval) and $4*p^{-5} = 1.0407\ldots$ (new semitone-interval). Tone-interval is repeated *5* times and semitone interval – 3 times; they cover an octave precisely: $(p^3/2)^5 * (4*p^{-5})^3 = 2$. The condition of precise covering of octave by means of two kinds of interval factors was the main one to choice the first eight members of a considered geometrical progression. (8, 5 and 3 are Fibonacci numbers; an interesting connection between sequences, produced by this algorithm, however Fibonacci numbers are out of the scope of this article).

Fig. 10 demonstrates a minimal difference between sequences of interval factors inside one octave in the both scales. The first half of sequences coincides completely, but an additional semitone-interval is appeared in the second half of new scale. This additional semitone-interval exists due to the fact that the coefficient p=1.309… is less than the quint 3/2.

7-steps Pythagorean scale of *C* major	1	1	1/2	1		1	1	1/2
8-steps scale of the golden wurf	1	1	1/2	1	1/2	1	1	1/2

Figure 10: Sequences of interval factors in considered mathematical scales. Tone-intervals are marked by 1, semitone-intervals – 1/2.

It should be noted that the base coefficient $\varphi^2/2 = p$ is known in bio-morphology under the name "the golden wurf" (a wurf or a double ratio is the main invariant of projective geometry; it is interesting that finite projective-geometric plane is connected with Hadamard matrices considered in the first part of this article [Sachkov, 1982]). The notion "wurf" is known in the field of non-Euclidian geometries for a long time; its translation from German language means "throw". The golden wurf was introduced in works [Petoukhov, 1981, 1989] which were devoted to non-Euclidean bio-symmetries. The golden wurf has a status of ontogenetic and phylogenetic invariant of aggregated proportions of three-component kinematical blocks of human and animal bodies. The value of the golden wurf concerns acoustic perception also: the human ear cochlea consists of three patterns (three coils of a spiral), the ratios of whose lengths form a geometrical progression with the golden section as a coefficient (Fig. 11). The double ratio of these three lengths is equal to the golden wurf: $p = \varphi^2/2 = 1.309...$ [Petoukhov, 1989].

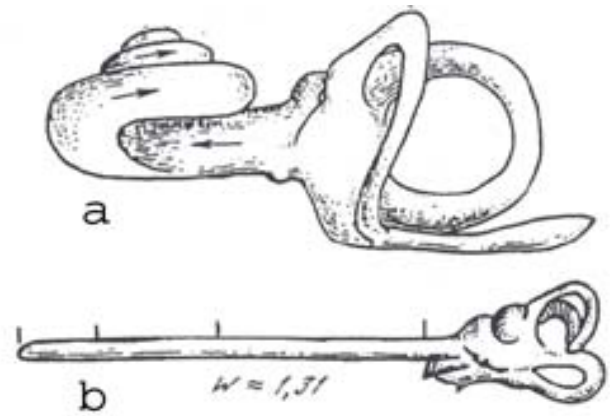

Figure 11: The helical structure of the human ear cochlea (from [Cook, 1914]) with a projective geometry proportion of the golden wurf: a – cochlea, coiled into a helix; b – cochlea uncoiled into a straight line.

Due to this name of the coefficient $p = \varphi^2/2$, the whole described mathematical scale is named "the scale of the golden wurf". Using of a sequence (4) of tone-intervals, the author constructs a sequence of tones (musical notes), which can be named "wurf-scale of *C* major" by analogy with Pythagorean scale of *C* major (Fig. 12). A choice of frequencies of these tones of the first octave is made in such way that this scale contains a frequency 440 Hz which corresponds to note "la" in Pythagorean scale and in equal temperament scale and which is used traditionally for tuning in musical instruments. Fig. 12 compares Pythagorean 7-steps scale *C* major and 8-steps scale of the golden wurf (or briefly, wurf-scale) for the first octave. Taking into account a minimal difference between both scales, the majority of notes of the wurf-scale are named by analogy with appropriate notes of Pythagorean scale but with the letter "m" in the end (for instance, "rem" instead "re"). An additional fifth note is named "pim".

Pyth	260.74 Do_1	293.33 Re	330 Mi	347.65 Fa		391.11 Sol	**440** La	495.00 Si	521.48 Do_2
wurf	256.78 Dom_1	287.98 Rem	322.98 Mim	336.13 Fam	376.98 Pim	392.32 Solm	**440** Lam	493.47 Sim	513.56 Dom_2

Figure 12: The upper row "Pyth" demonstrates frequencies of tones in 7-steps Pythagorean scale of *C* major in the first octave. The bottom row „Wurf" demonstrates frequencies of tones in 8-steps scale of the golden wurf of *C* major in the similar octave. Numbers mean frequencies in Hz. Names of notes are given.

The scale of the golden wurf has many analogies with Pythagorean scale in its symmetries and proportions. The main difference is connected with irrational values of interval factors in the wurf-scale (irrational values are used in classical equal-temperament scale also). Additional details about the mathematical scale of the golden wurf, its musical meaning and possible physiological applications, its proportions and its connection with Fibonacci numbers will be published specially.

A history of attempts of a creation of new musical scales knows many famous authors: J.Kepler, R.Descartes, G.Leibniz, L.Euler, etc. But they didn't use analogies with structures of genetic code for such aim. This part of the article describes a new attempt to improve knowledge about possible musical scales with a physiological meaning.

HARMONICAL RATIOS AND CRYSTALS

Many authors compare a living substance to crystals. For instance, E. Schrodinger named it aperiodic crystal. Do annals of modern science contain any data about a connection between music harmony and crystals? Yes, certain data exist and they have a long history. For example, they were collected by famous Russian crystallographer I.I.Shafranovskiy (for example, see [Berger, 1997, p.270-281]). Below we describe a few examples from this collection.

In 1818, C.S. Weiss, who has discovered crystallographic systems and who was one of founders of crystallography, has emphasized a musical analogy in crystallographic systems. He investigated ratios among segments, which are formed by faces of crystals of the cubic system. Weiss has shown that these ratios are identical absolutely to ratios between musical tones.

In 1829, J. Grassman, who has written a known book "Zur physischen Kristallonomie und geometrishen Combinationslehre" and has developed many mathematic methods in crystallography, has noted impressive musical analogies in the field of crystallography. The speech is about many analogies described by him between ratios of musical tones and segments, formed by faces of the same zone of crystals. According to his figurative expression, "crystal polyhedron is a fallen asleep chord - a chord of the molecular fluctuations made in time of its formation".

In the end of 1890th years the outstanding crystallographer V. Goldschmidt has returned to the same ideas. The prominent Russian mineralogist and geochemist

A.E.Fersman wrote about his thematic publications: “These works represent the historical page in crystallography, which has lead Goldschmidt to revealing by him laws of harmonic ratios. Goldschmidt has extended these laws logically from the world of crystals into the world of other correlations in the regions of paints, colors, sounds and even biological correlations. It has become one of the most favorite themes of philosophical researches by Goldschmidt” (from [Berger, 1997, p.270]). This list of historical examples can be continued including the work [Shafranovskiy, 1986], etc.

CONCLUSION

Described materials permit to study symmetrical and structural peculiarities of interrelated genetic code systems from a system viewpoint due to their presentation as parts of the mutual family of special matrices. Simultaneously this way permits to apply in this molecular-genetic field ideas and methods from a few modern sciences: digital signal processing, spectral theory, unitary operators of quantum mechanics, quantum computers, theory of musical harmony, etc.

Described researches were made by the author in the frame of a long-term cooperation between Russian and Hungarian Academies of Sciences. The author is grateful to Frolov K.V., Darvas Gy., Bohm J., Ne’eman Y., He M., Kovacs K., Smolianinov V.V., Vladimirov Y.S. for their support of these investigations.

REFERENCES

Ahmed N.U., Rao K.R. (1975) *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag New York, Inc., 253 p.

Berger L.G. (1997) *Epistomology of Culture*, Russkiy mir, Moscow (in Russian).

Carl J.V. (1974). On the use of Walsh functions in mane-made and biological pattern recognition systems., *Applications of Walsh function and sequence theory*: Proc. 1974 Symposium, Washington, p. 9-25

Cook T.A. (1914) *The Curves of Life*. L.: Constable and Co, 490 p.

Feynman R. (1963) *The Feynman Lectures*, N-Y, Pergamon Press.

Geramita, A. V. (1979) *Orthogonal Designs: Quadratic Forms and Hadamard Matrices.* New York: Dekker, 1979.

Ginsburg A.P., Carl J.V., Kabrisky M., Hall C.F., Gill R.A. (1974) Physiological aspects of a model for the classification of visual images, *Adv. Cybern. and Syst*., London e.a., v.3, p.1289-1305.

He M., Petoukhov S.V., Ricci P. (2004) Genetic Code, Hamming Distance and Stochastic Matrices, *Bulletin for Mathematical Biology*, v.66, p.1405-1421.

Hightower T.V., The creation of musical scales from a mathematic and acoustic point of view, http://home22.inet.tele.dk/hightower/scales.htm

Needham J. (1962) *Science and Civilization in China*, v.4, Cambridge University Press.

Nielsen M.A., Chuang I.L. (2001) *Quantum computation and quantum information*, Cambridge University Press.

Petoukhov S.V. (1981) *Biomechanics, bionics and symmetry*, Moscow, Nauka, 239 p.(in Russian)

Petoukhov S.V. (1989) Non-Euclidean geometries and algorithms of living bodies, *Computers & Mathematics with Applications*, 17, 4-6,.505-534.

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

Petoukhov S.V. (2001b) Genetic Codes I: Binary sub-alphabets, bi-symmetric matrices and golden section; Genetic Codes II: Numeric Rules of Degeneracy and a Chronocyclic Theory - *Symmetry in Genetic Information*, ed. Petoukhov S.V., special issue of *Symmetry: Culture and Science*, Budapest: Intern. Symmetry Foundation, 255-306.

Petoukhov S.V. (2003-2004) Attributive conception of genetic code, its bi-periodic tables and problem of unification bases of biological languages, *Symmetry: Culture and Science*, 14-15, part 1, 281-307.

Petoukhov S.V. (2005) The rules of degeneracy and segregations in genetic codes. The chronocyclic conception and parallels with Mendel's laws. - In: *Advances in Bioinformatics and its Applications*, Series in Mathematical Biology and Medicine, v.8, pp.512-532 (London, World Scientific)

Pont G. Philosophy and Science of Music in Ancient Greece: The Predecessors of Pythagoras and their Contribution, *Nexus Network Journal,* 6, 1, http://emis.elibrary.ru/journals/NNJ/Pont-v6n1.html

Sachkov V.N. (1982) Introduction in Combinatory Methods of Discrete Mathematic, Moscow: Nauka.

Shafranovskiy I.I. (1986) *Crystallography*, 31, 3, 51- 60.

Shiozaki A. (1980) A model of distributed type associated memory with quantized Hadamard transform., *Biol.Cybern*.,.38, 1, 19-22.

Schmidt-Jones C. *Octaves and the Major-Minor Tonal System*, http://cnx.org/content/m10862/latest/

Shapeville F., Haenni A.-L. (1974) *Biosynthese des proteines*, Paris: Hemann Collection.

Shults G.E., Schirmer R.H. (1979) *Principles of protein structure*. Berlin-Heidelberg-New York, Springer-Verlag, 350 p.

Sklar B. (2001) Digital Communication. Fundamentals and Applications, New Jersey: Prentice Hall PTR.

Voloshinov A.V. (2000) *Mathematics and Arts*, Moscow: Prosveschenie, 400 p. (in Russian)

Wittmann, H.G. (1961) Ansatze zur Entschlusselung des genetischen Codes, *Die Naturwissenschaften*, 48, 24, 55.

Bebenson Y, Shapiro E. (2006) Bringing DNA computers to life, *Scientific American*, May, 44-51.

Yarlagadda R., Hershey, J.(1997) Hadamard Matrix Analysis and Synthesis With Applications to Communications and Signal/Image Processing, Boston, MA: Kluwer Academic Publishers.