Matrix Genetics and Algebraic Properties of the Multi-Level System of Genetic Alphabets

Sergey V. Petoukhov

Abstract

The article is devoted to algebraic properties of the multi-level system of molecular-genetic alphabets. It leads to help solve the problem of algebraic unity of inherited information systems in living matter. These algebraic properties are revealed by means of Kronecker families of matrix forms of a presentation of molecular-genetic alphabets. A family of genetic (8x8)-matrices shows unexpected connections of the genetic system with Rademacher and Walsh functions and with special Hadamard matrices which are well-known in theory of noise-immunity coding and digital communication. Decompositions of such genetic (8x8)-matrices on the basis of the known principle of dyadic-shifts lead to sets of 8 sparse matrices. Each of these sets is closed in relation to multiplication and defines a special algebra of 8-dimensional hypercomplex numbers. Mathematical aspects of these 8-dimensional algebras are presented in connection with metric vector spaces, the sequency theory by Harmuth and some methods of spectral analysis. The diversity of known dialects of the genetic code can be analyzed from the viewpoint of these algebras. Our results are discussed taking into account the important role of dyadic shifts, hypercomplex numbers and Hadamard matrices in mathematics, informatics, theoretical physics, etc. These results testify that living matter has a profound algebraic essence which is interconnected with 8-dimensional vector spaces. In our opinion these results lead to a new way of knowledge of living matter in the field of algebraic biology and its mathematical modeling. The idea of a biological meaning of Kronecker multiplication of matrices is based on the structure of Punnett squares in the field of Mendelian genetics. The author believes that the Mendelian laws of independent inheritance of traits have revealed just the tip of an algebraic iceberg of informational structure of living matter and that matrix genetics has contributed to the next steps to disclose this important iceberg.

Key Words: genetic code, alphabet, Punnett squares, dyadic shift, hypercomplex numbers, Hadamard matrices, symmetry

NeuroQuantology 2011; 4: 799-820

Introduction

Science has led to a new understanding of life itself: “Life is a partnership between genes and mathematics” (Stewart, 1999). But what kind of mathematics is a partner with the genetic code?

Trying to find such mathematics, we have turned to study the multi-level system of interrelated molecular-genetic alphabets. On this way we were surprised to find connections of this genetic system with well-known formalisms of the engineering theory of noise-immunity coding: 1) Kronecker products of matrices; 2) orthogonal systems of functions by Rademacher and Walsh; 3) Hadamard matrices; 4) a group of dyadic shifts; 5) hypercomplex number systems, etc. This article is devoted to some of our results of such studying the system of interrelated genetic alphabets.

Corresponding author: Sergey V. Petoukhov
Addresse: Russia, 101990, Moscow, Malyi Kharitonievskiy pereulok, 4 , Mechanical Engineering Research Institute, Russian Academy of Sciences
Phone: + 7(499)-243-3567
Fax: + 7(499)-135-4065
e-mail: spetoukhov@gmail.com
Received Nov 12, 2011. Revised Dec 5, 2011.
Accepted Dec 8, 2011.

Alphabets play a basic role in communication technologies. In any communication system of "transmitter-receiver" the receiver always knows the alphabet of signals which are used by the transmitter. In linguistics, each alphabet has a complex multi-level structure because it contains sets of vowels and consonants where, in some languages, the set of vowels is divided into sub-sets of short and long sounds, and the set of consonants is divided into subsets of voiced and voiceless consonants, etc. Quantities of members in all of these parts of linguistic alphabets are not interrelated by means of known regularities of algebraic connections. We have discovered that the situation in the multi-level system of genetic alphabets is quite different: many parts of this system are closely interconnected by means of deep algebraic regularities and formalisms which are well-known in communication technologies as said before.

It is known that the molecular-genetic system of living matter includes the following alphabets each of which can be considered as a separate alphabet or as a part of a complex alphabetic system:
- 4-letter alphabet of nitrogenous bases;
- 64-letter alphabet of triplets;
- 2-letter alphabet of "weak and strong roots" of triplets;
- 20-letter alphabet of amino acids;
- 2-letter alphabet "purines vs. pyrimidines";
- 2-letter alphabet "strong vs. weak hydrogen bonds"
- 2-letter alphabet "keto vs. amino", etc. See the wide list of genetic alphabets in (Karlin, Ost, Blaisdell, 1989).

So, the molecular-genetic system is a multi-lingual system. Any sequence of nucleotides can be read from viewpoints of different genetic languages depending on the reader alphabet. It can be added that the typical expression "the genetic code" means an interrelation between elements of two of these genetic alphabets: the alphabet of triplets and the alphabet of amino acids and stop-codons.

Genetic information from the micro-world of genetic molecules dictates constructions in the macro-world of living organisms under strong noise and interference. The Mendel's law of independent inheritance of different traits (for example, colors of hair, skin and eyes are inherited independently of each other) testifies that this dictation is realized through different independent channels by means of unknown algorithms of multi-channel noise-immunity coding. It means that each living organism is an algorithmic machine of multi-channel noise-immunity coding. To understand this machine we should use the theory of noise-immunity coding. Genetic information is transferred by means of discrete elements. General theory of signal processing utilizes the encoding of discrete signals by means of special mathematical matrices and spectral representations of signals to increase reliability and efficiency of information transfer (Ahmed and Rao, 1975; Blahut, 1985; Peterson and Weldon, 1972; Sklar, 2001). A typical example of such matrices is the family of Hadamard matrices. Rows of Hadamard matrices form an orthogonal system of Walsh functions which is used for the spectral presentation and transfer of discrete signals (Ahmed and Rao, 1975; Geramita, 1979; Yarlagadda and Hershey, 1997). An investigation of structural analogies between digital informatics and genetic informatics is one of the important tasks of modern science in connection with the development of DNA-computers and bioinformatics. The author investigates molecular structures of the system of genetic alphabets by means of matrix methods of discrete signal processing (Petoukhov, 2001a, b; 2005a, b; 2008a, b; 2011; Petoukhov and He, 2010; He and Petoukhov, 2011).

The article describes author's results about relations of matrix forms of presentation of the system of genetic alphabets with special systems of 8-dimensional hypercomplex numbers (they differ from the Cayley's octonions). The discovery of these relationships is significant from some viewpoints. For example, it is interesting because systems of 8-dimensional hypercomplex numbers (first of all, Cayley's octonions and split-octonions) are one of key objects of mathematical natural sciences today. They relate to a number of exceptional structures in mathematics, among them the exceptional simple Lie algebras; they have applications in many fields such as string theory, special

relativity, the supersymmetric quantum mechanics, quantum logic, etc. (Capra, 2000; Dray, Manoque, 2009; Dixon, 1994). The term "octet" is also used frequently in phenomenologic laws of science: the Eightfold way by M. Gell-Mann and Y. Ne'eman (1964) in physics; the octet rule in chemistry, etc. In view of these facts one can think that genetic systems of 8-dimensional numbers will become one of the interesting parts of mathematical natural sciences.

In addition, hypercomplex numbers are widely used in digital signal processing (Bulow, 2001; Chernov, 2002; Felberg, 2001; Furman *et al.,* 2003; Sin'kov, 2010; Toyoshima, 1999, 2002). Formalisms of multi-dimensional vector spaces are one of basic formalisms in digital communication technologies, systems of artificial intelligence, pattern recognition, training of robots, detection of errors in the transmission of information, etc. Revealed genetic types of hypercomplex numbers can be useful to answer many questions of bioinformatics and to develop new kinds of genetic algorithms.

Hadamard matrices and orthogonal systems of Walsh functions are among the most used tools for error-correcting coding information, and for many other applications in digital signal processing (Ahmed and Rao, 1975; Geramita, 1979; Yarlagadda and Hershey, 1997). As noted in the article (Seberry *et al.,* 2005), many tens of thousands of works are devoted to diverse applications of Hadamard matrices for signal processing. Our discovery of relations of the system of genetic alphabets with special 8-dimensional hypercomplex numbers and with special Hadamard matrices helps to establish the kind of mathematics which is a partner of the molecular-genetic system.

The Appendix A describes a mathematical operation of Kronecker multiplication of matrices, and the Appendix B describes a notion of modulo-2 addition and dyadic shifts.

1. Mendel laws, Punnett squares and Kronecker multiplication of matrices

What kind of matrix approaches should be chosen to study the system of molecular-genetic alphabets? For this choice the author has used known facts and constructions from Mendelian genetics taking into account the conception of the unity of the organism. Below we will analyze an algebraic structure of Punnett squares which are one well-known tool of genetics. It was introduced by the British geneticist R. C. Punnett in 1905 (Crew, 1968) to help in prediction of inherited traits of offspring. It is a quite popular method among most text-books of Mendelian genetics. To help in revealing algebraic properties of Punnett squares, the Appendix A describes a mathematical operation of Kronecker multiplication of matrices, and the Appendix B describes a notion of dyadic shifts and matrices of dyadic shifts.

Punnett squares represent alphabets of genotypes or, more precisely, alphabets of possible combinations of male and female gametes in Mendelian crosses of organisms from a viewpoint of a certain amount of inherited traits taken into account. We will show a possibility of interpreting a set of Punnett squares for polyhybrid crosses not as tables but as square matrices (we term them "Punnett matrices") of Kronecker products of Punnett (2*2)-matrices for monohybrid crosses. Based on Kronecker multiplication of matrices this approach gives a simple algebraic method to construct Punnett squares for complex cases of multi-hybrid crosses. In addition we show that dyadic-shift decompositions of these "Punnett matrices" lead in some cases to a certain method of classification of different sub-sets of combinations of alleles from male and female gametes.

Heredity is the passing of traits from parent to offspring. Traits are controlled by genes. The different forms of a gene for a certain trait are called alleles. There are two alleles for every trait. Alleles can be dominant or recessive. Each cell in an organism's body contains two alleles for every trait. One allele is inherited from the female parent and one allele is inherited from the male parent. Punnett square is a simple method for predicting the ways in which alleles can be combined. The Punnett square is a summary of every possible combination of one maternal allele with one paternal allele for each gene being studied in the cross. In a Punnett square, dominant and recessive alleles are usually represented by letters. An uppercase letter represents a dominant allele (we will use for dominant

alleles symbols H, B, C,...), and a lowercase letter represents a recessive allele (we will use for recessive alleles symbols h, b, c,...). An organism is homozygous if it has identical alleles for a particular trait, for example HH or hh. An organism is heterozygous if it has non-identical alleles for a particular trait, for example Hh. There are three possible combinations of alleles of an organism for a particular trait: homozygous dominant (HH), heterozygous (Hh), and homozygous recessive (hh). In the classic way of constructing Punnett square, alleles of a maternal gamete are put on one side of the square (usually on the top) and alleles of a paternal gamete are put on the left side; the possible progeny are produced by filling the squares with one allele from the top and one allele from the left to produce the progeny genotypes. By tradition an uppercase letter of dominant allele stands before a lowercase letter of a recessive allele (for example Hh, but not hH).

If only one trait is being considered in a genetic cross, the cross is called monohybrid. If two or three traits are being considered in a genetic cross, the cross is called dihybrid or trihybrid correspondingly. Figure 1 shows an example of Punnet squares for one of cases of trihybrid cross of parents with the maternal genotype HhbbCc and the paternal genotype HhBbCc (this particular example of initial genotypes is taken from a site about Punnett squares http://www.changbioscience.com/genetics/punnett.html). In this example each set of gametes include 8 gametes; the set of female gametes is different from the set of male gametes. One can note that this is not the easiest thing to construct Punnett squares for different cases of multi-hydrid crosses by means of the classical way.

		maternal gametes							
		HbC	Hbc	HbC	Hbc	hbC	hbc	hbC	hbc
paternal gametes	HBC	HHBbCC	HHBbCc	HHBbCC	HHBbCc	HhBbCC	HhBbCc	HhBbCC	HhBbCc
	HBc	HHBbCc	HHBbcc	HHBbCc	HHBbcc	HhBbCc	HhBbcc	HhBbCc	HhBbcc
	HbC	HHbbCC	HHbbCc	HHbbCC	HHbbCc	HhbbCC	HhbbCc	HhbbCC	HhbbCc
	Hbc	HHbbCc	HHbbcc	HHbbCc	HHbbcc	HhbbCc	Hhbbcc	HhbbCc	Hhbbcc
	hBC	HhBbCC	HhBbCc	HhBbCC	HhBbCc	hhBbCC	hhBbCc	hhBbCC	hhBbCc
	hBc	HhBbCc	HhBbcc	HhBbCc	HhBbcc	hhBbCc	hhBbcc	hhBbCc	hhBbcc
	hbC	HhbbCC	HhbbCc	HhbbCC	HhbbCc	hhbbCC	hhbbCc	hhbbCC	hhbbCc
	hbc	HhbbCc	Hhbbcc	HhbbCc	Hhbbcc	hhbbCc	hhbbcc	hhbbCc	hhbbcc

Figure 1. A typical Punnett square for a trihybrid cross; the maternal genotype is HhbbCc and the paternal genotype is HhBbCc.[2]

	maternal gametes	
	H	h
H	HH	Hh
h	Hh	hh

;

	maternal gametes	
	b	b
B	Bb	Bb
b	bb	bb

;

	maternal gametes	
	C	c
C	CC	Cc
c	Cc	cc

Figure 2. Three Punnett squares of monohybrid cross for the case of the Punnett square of the trihybrid cross on Figure 1.

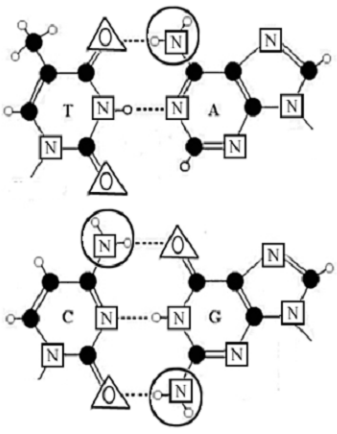

Figure 3. The complementary pairs of the 4 nitrogenous bases in DNA: A-T (adenine and thymine), C-G (cytosine and guanine). Hydrogen bonds in these pairs are shown by dotted lines. Black circles are atoms of carbon; small white circles are atoms of hydrogen; squares with the letter N are atoms of nitrogen; triangles with the letter O are atoms of oxygen. Amides (or amino-groups) NH_2 are marked by big circles (Petoukhov and He, 2010).

[2] from http://www.changbioscience.com/genetics/punnett.html

But the author pays attention that this Punnett (8*8)-matrix (in the bold frame on Figure 1) is identical to a result of Kronecker product of three Punnett (2*2)-matrices of monohybrid cross relative to each of the traits (see Appendix A about Kronecker product). Figure 2 shows three Punnett squares of monohybrid cross for the considered case of the maternal genotype HhbbCc and the paternal genotype HhBbCc.

Cases of multihybrid crosses with identical maternal and paternal genotypes demonstrate Punnett matrices which are connected with their dyadic-shift decomposition (DS-decomposition) in some aspects (see Appendix B about dyadic shifts). For example in the case of the trihybrid cross of the maternal genotype HhBbCc and the paternal genotype HhBbCc, the Punnett (8*8)-matrix can be DS-decomposed into 8 sparse matrices p_0, p_1, p_2, ..., p_7 by analogy with the matrix of dyadic shifts on Figure B.1 (Appendix B). Then the first of the sparse matrices p_0 (where only all cells on the main diagonal contain non-zero entries) will contain all the homozygous combinations of alleles HHBBCC, HHBBcc, HHbbCC, HHbbcc, hhBBCC, hhBBcc, hhbbCC, hhbbcc. The second of the sparse matrices p_1 will contain (in its 8 non-zero cells) combinations of alleles where only the third trait is heterozygous: HHBBCc, HHbbCc, hhBBCc, hhbbCc, etc. In other words by means of DS-decompositions we get a method of classification of different sub-sets of homozygous or heterozygous organisms in these cases of multi-hybrid crosses. Below we show similar algebraic structures at a molecular-genetic level which represents a completely different level of biological organization in a comparison with the level of holistic organisms.

2. Matrix presentations of molecular-genetic alphabets and Rademacher functions

Now let us proceed to the system of molecular-genetic alphabets. All living organisms have the same main set of molecular-genetic alphabets in which genetic multiplets play a significant role. Four monoplets forms the alphabet of nitrogenous bases: A (adenine), C (cytosine), G (guanine), U/T (uracil in RNA or thymine in DNA); 64 triplets encode amino acids and termination signals; each protein is encoded by more or less long multiplets (n-plets). Each complete set of n-plets contains 4^n different n-plets and can be considered formally as a separate alphabet; the sets of 16 duplets and 64 triplets are the most interesting for us in this article. On the basis of the idea about analogies between computer informatics and genetic informatics, each of the formal alphabets of genetic n-plets can be presented in a general form of the relevant square matrix (genomatrix) $[C\ A;\ U\ G]^{(n)}$ of the Kronecker family (Figure 5) (Petoukhov, 2001a, 2005, 2008a,b). Here A, C, G, U are the letters of the alphabet of 1-plets of nitrogenous bases (Figure 5), (n) means the Kronecker exponentiation. This operation of Kronecker multiplication of matrices was met in the previous section about Punnett squares. One can remind additionally that Kronecker product of matrices is an important tool in the theory of linear spaces and operators; it is associated with the tasks of a special union of two linear spaces with dimensions “n” and “m” into a linear space with the higher dimension “n*m”, and also with the tasks of constructing the matrix operators in this (n*m)-space on the basis of operators of the initial spaces of smaller dimensions “n” and “m” (Halmos, 1974). Each genomatrix $[C\ A;\ U\ G]^{(n)}$ contains a complete set of n-plets as its matrix elements (Figure 5). For example, the (8x8)-genomatrix $[C\ A;\ U\ G]^{(3)}$ contains all 64 triplets which encode 20 amino acids and stop-signals. We will show that these genetic matrices have phenomenological relations with the logic of dyadic shifts described in Appendix B.

The 4-letter alphabet of nitrogenous bases contains the 4 specific poly-atomic constructions with the special biochemical properties (Figure 3). The set of these 4 constructions is not absolutely heterogeneous, but it bears the substantial symmetric system of distinctive-uniting attributes (or, more precisely, pairs of "attribute-antiattribute"). This system of pairs of opposite attributes divides this 4-letter alphabet into various three pairs of letters by all three possible ways; letters of each such pair are equivalent to each other in accordance with one of these attributes or with its absence.

This system of pairs of opposite attributes divides this 4-letter alphabet into the following three pairs of letters, which are equivalent from a viewpoint of one of these attributes or its absence: 1) C = U & A = G (according to the binary-opposite attributes: "pyrimidine" or "non-pyrimidine", which is purine); 2) A = C & G = U (according to the attributes: amino-mutating or non-amino-mutating under action of nitrous acid HNO_2 (Wittmann, 1961; Ycas, 1969); the same division is given by the attributes "keto" or "amino" (Karlin *et al.,* 1989); 3) C = G & A = U (according to the attributes: three or two hydrogen bonds are materialized in these complementary pairs). The possibility of such division of the genetic alphabet into three binary sub-alphabets is known (Karlin *et al.,* 1989). We utilize these known sub-alphabets in the field of matrix genetics which studies matrix forms of presentation of the genetic alphabets. Each of the monoplets A, C, G, U/T can be symbolized by appropriate binary symbols "0" or "1" on the basis of these sub-alphabets. Then these binary symbols can be used for a binary numeration of columns and rows of the matrices in the Kronecker family of genomatrices of n-plets.

Let us mark these three kinds of binary-opposite attributes by numbers N = 1, 2, 3 and ascribe to each of the four genetic letters the symbol "0_N" (the symbol "1_N") in a case of presence (or of absence correspondingly) of the attribute under number "N" in this letter. As a result we obtain the representation of the genetic 4-letter alphabet in the system of its three "binary sub-alphabets corresponding to attributes" (Figure 4).

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

Figure 4. Three binary sub-alphabets according to three kinds of binary-opposite attributes in the 4-letter alphabet of nitrogenous bases C, A, G, U. The scheme on the right side explains graphically the symmetric relations of equivalence between the pairs of letters from the viewpoint of the separate attributes 1, 2, 3 (Petoukhov, 2001a).

The table on Figure 4 shows that, on the basis of each kind of the attributes, each of the letters A, C, G, U/T possesses three "faces" or meanings in the three binary sub-alphabets. On the basis of each kind of the attributes, the 4-letter alphabet is curtailed into the 2-letter alphabet. For example, on the basis of the first kind of binary-opposite attributes we have (instead of the 4-letter alphabet) the alphabet from two letters 0_1 and 1_1, which one can name "the binary sub-alphabet of the first kind of the binary attributes".

On the basis of the idea about a possible analogy between discrete signals processing in computers and in the genetic code system, one can present the genetic 4-letter alphabet in the following matrix form [C A; U G] (Figure 5). Then the Kronecker family of matrices with such alphabetical kernel can be considered: $[C\ A;\ U\ G]^{(n)}$, where (n) means the integer Kronecker (or tensor) power. This form possesses the analogy with the Kronecker family of Hadamard matrices $[1\ 1;\ -1\ 1]^{(n)}$ from discrete signals processing (Ahmed and Rao, 1975). Figure 5 shows the first genetic matrices of such a family. One can see on this Figure that each of the matrices contains all the genetic multiplets of equal length in a strict order: [C A; U G] contains all the 4 monoplets; $[C\ A;\ U\ G]^{(2)}$ contains all the 16 duplets; $[C\ A;\ U\ G]^{(3)}$ contains all the 64 triplets.

All the columns and rows of the matrices on Figure 4 are binary numerated and disposed in a monotonic order by the following algorithm which uses biochemical features of the genetic nitrogenous bases and which can be used in bio-computers of any organism. Numerations of columns and rows are formed automatically if one interprets multiplets of each column from the viewpoint of the first binary sub-alphabet (Figure 4) and if one interprets multiplets of each row from the viewpoint of the second binary sub-alphabet. For example, the column 010 contains all the triplets of the form "pyrimidine-purine-pyrimidine"; the row 010 contains all the triplets of the form "amino-keto-amino". Each of the triplets in the matrix $[C\ A;\ U\ G]^{(3)}$ receives its dyadic-shift numeration by means of modulo-2 addition of numerations of its column and row (see Appendix B). For example, the triplet CAG receives its dyadic-shift

numeration 010 (or 2 in decimal notation) because it belongs to the column 011 and the row 001. Any codon and its anti-codon are disposed in inversion-symmetrical manner relative to the centre of the genomatrix [C A; U G]$^{(3)}$ (Figure 5). In this genomatrix any codon and its anti-codon possess the same dyadic-shift numeration; each of the dyadic-shift numerations denotes a subset of 8 triplets with 4 pairs "codon-anticodon".

It should be noted additionally that each of 64 triplets in the genomatrix [C A; U G]$^{(3)}$ receives the same dyadic-shift numeration if each of its three nitrogenous bases is numerated from the viewpoint of the binary sub-alphabet № 3 from Figure 4 in which C=G=0, A=U/T=1. For example in this way the triplet CAG receives again the dyadic-shift numeration 010 if each of its letters is replaced by its binary symbol (C=G=0, A=U/T=1). This alternative way of dyadic-shift numeration of multiplets is simpler for using and it is termed "the direct dyadic-shift numeration".

[C A; U G] =

	0	1
0	C 0	A 1
1	U 1	G 0

; [C A; U G]$^{(2)}$ =

	00(0)	01(1)	10(2)	11(3)
00 (0)	CC 00 (0)	CA 01(1)	AC 10 (2)	AA 11 (3)
01 (1)	CU 01 (1)	CG 00 (0)	AU 11 (3)	AG 10 (2)
10 (2)	UC 10 (2)	UA 11(3)	GC 00 (0)	GA 01 (1)
11 (3)	UU 11 (3)	UG 10 (2)	GU 01 (1)	GG 00 (0)

	000 (0)	001 (1)	010 (2)	011 (3)	100 (4)	101 (5)	110 (6)	111 (7)
000 (0)	CCC 000 (0)	CCA 001 (1)	CAC 010 (2)	CAA 011 (3)	ACC 100 (4)	ACA 101 (5)	AAC 110 (6)	AAA 111 (7)
001 (1)	CCU 001 (1)	CCG 000 (0)	CAU 011 (3)	CAG 010 (2)	ACU 101 (5)	ACG 100 (4)	AAU 111 (7)	AAG 110 (6)
010 (2)	CUC 010 (2)	CUA 011 (3)	CGC 000 (0)	CGA 001 (1)	AUC 110 (6)	AUA 111 (7)	AGC 100 (4)	AGA 101 (5)
011 (3)	CUU 011 (3)	CUG 010 (2)	CGU 001 (1)	CGG 000 (0)	AUU 111 (7)	AUG 110 (6)	AGU 101 (5)	AGG 100 (4)
100 (4)	UCC 100 (4)	UCA 101 (5))	UAC 110 (6)	UAA 111 (7)	GCC 000 (0)	GCA 001 (1)	GAC 010 (2)	GAA 011 (3)
101 (5)	UCU 101 (5)	UCG 100 (4)	UAU 111 (7)	UAG 110 (6)	GCU 001 (1)	GCG 000 (0)	GAU 011 (3)	GAG 010 (2)
110 (6)	UUC 110 (6)	UUA 111 (7)	UGC 100 (4)	UGA 101 (5)	GUC 010 (2)	GUA 011 (3)	GGC 000 (0)	GGA 001 (1)
111 (7)	UUU 111 (7)	UUG 110 (6)	UGU 101 (5)	UGG 100 (4)	GUU 011 (3)	GUG 010 (2)	GGU 001 (1)	GGG 000 (0)

Figure 5. The first genetic matrices of the Kronecker family [C A; U G]$^{(n)}$ with binary numerations of their columns and rows on the basis of the binary sub-alphabets № 1 and № 2 from Figure 4. The lower matrix is the genomatrix [C A; U G]$^{(3)}$. Each of matrix cells contains a symbol of a multiplet, a dyadic-shift numeration of this multiplet and its expression in decimal notation. Decimal numerations of columns, rows and multiplets are written in brackets. Black and white cells contain triplets and duplets with strong and weak roots correspondingly for cases of the Standard Code, the Vertebrate Mitochondrial Code and the most other dialects of the genetic code (see more details in Petoukhov, 2001a, 2005, 2010, 2011; Petoukhov and He, 2010).

The genetic code is termed the degeneracy code because its 64 triplets encode 20 amino acids, and different amino acids are encoded by means of different quantities of triplets. Modern science knows many variants (or dialects) of the genetic code, which are variants of the correspondence between the alphabet of triplets and the alphabet of amino acids and stop-codons. Data about these variants are presented on NCBI's Web site, http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi. Seventeen variants (or dialects) of the genetic code exist, which differ one from another by some details of correspondences between triplets and objects encoded by them. Most of these dialects (including the Standard Code and

the Vertebrate Mitochondrial Code which are presented on Figure 6) have a general scheme of their degeneracy, where 32 triplets with “strong roots” and 32 triplets with “weak roots” exist (Petoukhov, 2001a, 2005, 2008, 2011; Petoukhov and He, 2010). Cells with these triplets are marked by black and white colors correspondingly inside the matrix [C A; U G]$^{(3)}$ on Figure 5. Here it should be mentioned that a combination of letters on the two first positions of each triplet is termed a “root” of this triplet; a letter on its third position is termed a “suffix”.

The set of 64 triplets contains 16 possible variants of such roots. Taking into account properties of triplets, the set of 16 possible roots is divided into two subsets with 8 roots in each. The first of such octets contains roots CC, CU, CG, AC, UC, GC, GU and GG. These roots are termed "strong roots" (Konopel’chenko and Rumer, 1975) because each of them defines four triplets with this root, coding values of which are independent on their suffix. For example, four triplets CGC, CGA, CGU, CGG, which have the identical strong root CG, encode the same amino acid Arg, although they have different suffixes (Figure 6). The second octet contains roots CA, AA, AU, AG, UA, UU, UG and GA. These roots are termed “weak roots” because each of them defines four triplets with this root, coding values of which depend on their suffix. An example of such a subfamily in Figure 6 is represented by four triplets CAC, CAA, CAU and CAC, two of which (CAC, CAU) encode the amino acid His and the other two (CAA, CAG) encode the amino acid Gln.

THE STANDARD CODE	
8 subfamilies of triplets with strong roots (“black triplets”) and the amino acids, which are encoded by them	8 subfamilies of triplets with weal roots (“white triplets”) and the amino acids, which are encoded by them
CCC, CCU, CCA, CCG → Pro	CAC, CAU, CAA, CAG → His, His, Gln, Gln
CUC, CUU, CUA, CUG → Leu	AAC, AAU, AAA, AAG → Asn, Asn, Lys, Lys
CGC, CGU, CGA, CGG → Arg	AUC, AUU, AUA, AUG → Ile, Ile, Ile, Met
ACC, ACU, ACA, ACG → Thr	AGC, AGU, AGA, AGG → Ser, Ser, Arg, Arg
UCC, UCU, UCA, UCG → Ser	UAC, UAU, UAA, UAG → Tyr, Tyr, Stop, Stop
GCC, GCU, GCA, GCG → Ala	UUC, UUU, UUA, UUG → Phe, Phe, Leu, Leu
GUC, GUU, GUA, GUG → Val	UGC, UGU, UGA, UGG → Cys, Cys, Stop, Trp
GGC, GGU, GGA, GGG → Gly	GAC, GAU, GAA, GAG → Asp, Asp, Glu, Glu
THE VERTEBRATE MITOCHONDRIAL CODE	
CCC, CCU, CCA, CCG → Pro	CAC, CAU, CAA, CAG → His, His, Gln, Gln
CUC, CUU, CUA, CUG → Leu	AAC, AAU, AAA, AAG → Asn, Asn, Lys, Lys
CGC, CGU, CGA, CGG → Arg	AUC, AUU, AUA, AUG → Ile, Ile, Met, Met
ACC, ACU, ACA, ACG → Thr	AGC, AGU, AGA, AGG → Ser, Ser, Stop, Stop
UCC, UCU, UCA, UCG → Ser	UAC, UAU, UAA, UAG → Tyr, Tyr, Stop, Stop
GCC, GCU, GCA, GCG → Ala	UUC, UUU, UUA, UUG → Phe, Phe, Leu, Leu
GUC, GUU, GUA, GUG → Val	UGC, UGU, UGA, UGG → Cys, Cys, Trp, Trp
GGC, GGU, GGA, GGG → Gly	GAC, GAU, GAA, GAG → Asp, Asp, Glu, Glu

Figure 6. The Standard Code and the Vertebrate Mitochondrial Code possess the basic scheme of the genetic code degeneracy with 32 triplets of strong roots (“black triplets”) and 32 triplets of weak roots.[3]

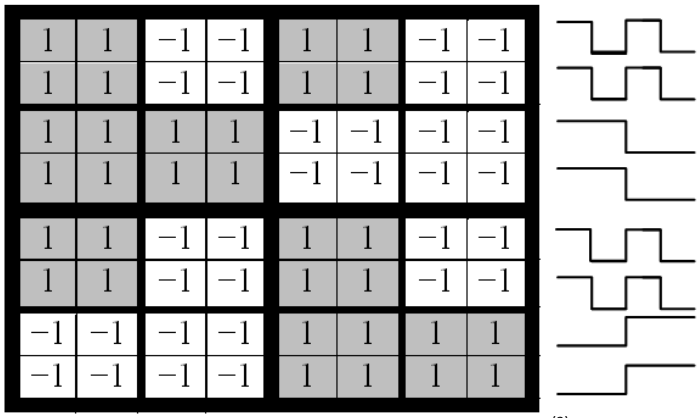

Figure 7. Rademacher form R of presentation of the genomatrix [C A; U G]$^{(3)}$ from Figure 5. A relevant system of Rademacher functions is shown at the right side (Petoukhov, 2011).

[3] Initial data from http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi

How these two subsets of triplets with strong and weak roots are disposed in the genomatrix [C A; U G]$^{(3)}$ (Figure 5) which was constructed formally on the base of the 4-letter alphabet of nitrogenous bases and Kronecher multiplications without any mention about the degeneracy of the genetic code and about amino acids? Can one anticipate any symmetry in their disposition? It should be noted that the huge quantity 64! ≈ 10^{89} of variants exists for dispositions of 64 triplets in the (8x8)-matrix. One can note for comparison, that the modern physics estimates the time of existence of the Universe in 10^{17} seconds. It is obvious that in such a situation an accidental disposition of the 20 amino acids and the corresponding triplets in a (8x8)-matrix will give almost never any symmetry in their disposition in matrix halves, quadrants and rows.

But it is a phenomenological fact that the disposition of the 32 triplets with strong roots ("black triplets" in Figure 5) and the 32 triplets with weak roots ("white triplets") has a unique and exceptional symmetric character (see Figure 5). For example the left and right halves of the matrix mosaic are mirror-anti-symmetric to each other in its colors: any pair of cells, disposed by mirror-symmetrical manner in these halves, possesses the opposite colors. One can say that each row of this mosaic matrix corresponds to an odd function. In addition each row of the mosaic matrix [C A; U G]$^{(3)}$ has a meander-line character (the term "meander-line" means here that lengths of black and white fragments are equal to each other along each row). But the theory of discrete signal processing uses such odd meander functions for a long time under the known name "Rademacher functions". Rademacher functions contain elements "+1" and "-1" only. Each of the matrix rows presents one of the Rademacher functions if each black (white) cell is interpreted such that it contains the number +1 (−1). Figure 7 shows a transformation of the mosaic matrix [C A; U G]$^{(3)}$ of (Figure 5) into a numeric matrix in the result of such replacements of black and white triplets by numbers "+1" and "-1" correspondingly.

3. The dyadic-shift decomposition of the genomatrix [C A; U G]$^{(3)}$ and the first type of 8-dimensional hypercomplex numbers

The Rademacher form R of the genomatrix [C A; U G]$^{(3)}$ (Figure 7) can be decomposed into sum of 8 sparse matrices r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7 (Figure 8) in accordance with the structural principle of the dyadic-shift matrix (Appendix B). More precisely any sparse matrix r_k (k=0, 1, ..., 7) contains entries "+1" or "-1" from the matrix R on Figure 7 in those cells which correspond to cells of the dyadic-shift matrix with the same dyadic-shift numeration "k"; all the other cells of the matrix r_k contain zero. For example, the sparse matrix r_2 contains entries "+1" and "-1" from the matrix R (Figure 7) only in those cells which correspond to cells with the dyadic numeration "2" in the dyadic-shift matrix on Figure B.1. Determinants of all the sparse matrices r_k are equal to 1.

$R = r_0+r_1+r_2+r_3+r_4+r_5+r_6+r_7 =$

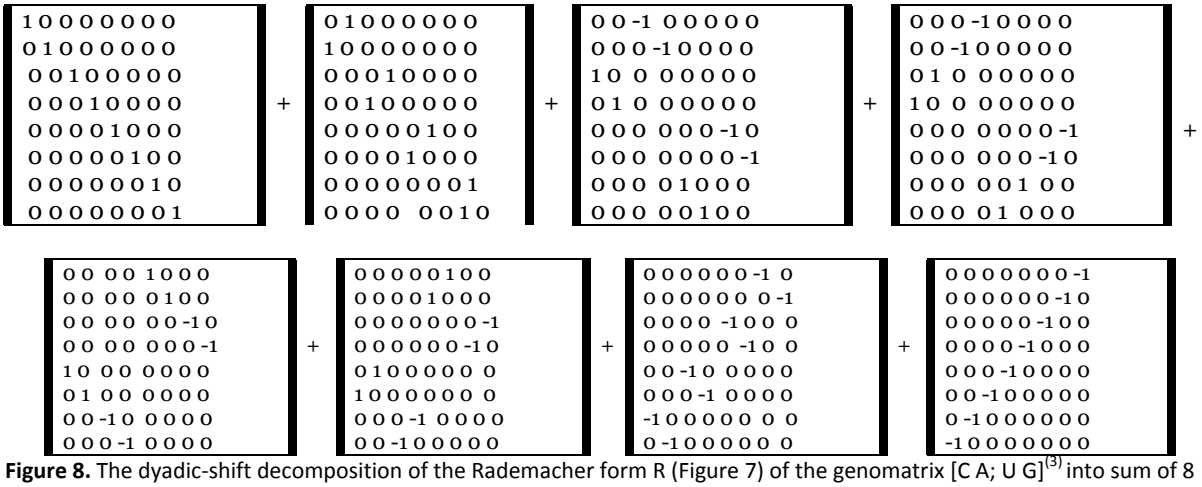

Figure 8. The dyadic-shift decomposition of the Rademacher form R (Figure 7) of the genomatrix [C A; U G]$^{(3)}$ into sum of 8 sparse matrices r_0, r_1,..., r_7 (Petoukhov, 2011).

The author has discovered that this set of 8 matrices r_0, r_1,... , r_7 (where r_0 is identity matrix) is closed in relation to multiplication and it satisfies the table of multiplication on Figure 9. The multiplication table on Figure 9 is asymmetrical relative to the main diagonal and it corresponds to the non-commutative associative algebra of 8-dimensional hypercomplex numbers, which are an extension of double numbers (or numbers by Lorentz). This matrix algebra is non-division algebra because it has zero divisors. It means that such non-zero hypercomplex numbers exist whose product is equal to zero. For example (r_3+r_4) and (r_2+r_5) (see Figure 8) are non-zero matrices, but their product is equal to zero matrix. These genetic 8-dimensional hypercomplex numbers are different from Cayley's octonions. The algebra of Cayley's octonions is non-associative algebra and correspondingly it does not possess a matrix form of its presentation (each of matrix algebras is an associative algebra). The known term "octonions" is not appropriate for the case of the multiplication table on Figure 9 because this mathematical term is usually used for members of normed division non-associative algebra.

	1	$\mathbf{r_1}$	$\mathbf{r_2}$	$\mathbf{r_3}$	$\mathbf{r_4}$	$\mathbf{r_5}$	$\mathbf{r_6}$	$\mathbf{r_7}$
1	**1**	$\mathbf{r_1}$	$\mathbf{r_2}$	$\mathbf{r_3}$	$\mathbf{r_4}$	$\mathbf{r_5}$	$\mathbf{r_6}$	$\mathbf{r_7}$
$\mathbf{r_1}$	$\mathbf{r_1}$	**1**	$\mathbf{r_3}$	$\mathbf{r_2}$	$\mathbf{r_5}$	$\mathbf{r_4}$	$\mathbf{r_7}$	$\mathbf{r_6}$
$\mathbf{r_2}$	$\mathbf{r_2}$	$\mathbf{r_3}$	**-1**	$\mathbf{-r_1}$	$\mathbf{-r_6}$	$\mathbf{-r_7}$	$\mathbf{r_4}$	$\mathbf{r_5}$
$\mathbf{r_3}$	$\mathbf{r_3}$	$\mathbf{r_2}$	$\mathbf{-r_1}$	**-1**	$\mathbf{-r_7}$	$\mathbf{-r_6}$	$\mathbf{r_5}$	$\mathbf{r_4}$
$\mathbf{r_4}$	$\mathbf{r_4}$	$\mathbf{r_5}$	$\mathbf{r_6}$	$\mathbf{r_7}$	**1**	$\mathbf{r_1}$	$\mathbf{r_2}$	$\mathbf{r_3}$
$\mathbf{r_5}$	$\mathbf{r_5}$	$\mathbf{r_4}$	$\mathbf{r_7}$	$\mathbf{r_6}$	$\mathbf{r_1}$	**1**	$\mathbf{r_3}$	$\mathbf{r_2}$
$\mathbf{r_6}$	$\mathbf{r_6}$	$\mathbf{r_7}$	$\mathbf{-r_4}$	$\mathbf{-r_5}$	$\mathbf{-r_2}$	$\mathbf{-r_3}$	**1**	$\mathbf{r_1}$
$\mathbf{r_7}$	$\mathbf{r_7}$	$\mathbf{r_6}$	$\mathbf{-r_5}$	$\mathbf{-r_4}$	$\mathbf{-r_3}$	$\mathbf{-r_2}$	$\mathbf{r_1}$	**1**

Figure 9. The multiplication table of basic matrices r_0, r_1,... , r_7 (where r_0 is identity matrix) which corresponds to the 8-dimensional algebra over the field of real numbers. It defines the 8-dimensional numeric system of genetic R_{123}-octetons (Petoukhov, 2011).

For this reason we term these hypercomplex numbers, which are revealed in matrix genetics, as "dyadic-shift genetic octetons" (or briefly "genooctetons" or simply "octetons"). In addition we term such kinds of matrix algebras, which are connected with dyadic-shift decompositions, as dyadic-shift algebras (or briefly DS-algebras). The author supposes that DS-algebras are very important for genetic systems. It is interesting that all the basic matrices r_0, r_1,..., r_7 are disposed in the multiplication table (Figure 9) in accordance with the structure of the dyadic-shift matrix (Appendix B). The numeric system of dyadic-shift genooctetons differs cardinally from the system of genetic 8-dimensional numbers which have been described by the author in matrix genetics previously (Kappraff and Petoukhov, 2009; Petoukhov, 2008a,c; Petoukhov and He, 2010, Section 3) and which are termed "8-dimensional bipolars" or "8-dimensional Yin-Yang genonumbers".

Below we describe another variant of genooctetons which is connected with Hadamard genomatrices (H-octetons). For this reason we term the first type of genooctetons (Figures 7-9) as R_{123}-octetons (here R is the first letter of the name Rademacher; the index 123 means the order 1-2-3 of positions in triplets).

A general form of R_{123}-octetons (Figure 8) is the following:

$$R_{123} = x_0*\mathbf{1} + x_1*\mathbf{r_1} + x_2*\mathbf{r_2} + x_3*\mathbf{r_3} + x_4*\mathbf{r_4} + x_5*\mathbf{r_5} + x_6*\mathbf{r_6} + x_7*\mathbf{r_7} \quad (1)$$

where coefficients x_0, x_1,.. , x_7 are real numbers. Here the first component x_0 is a scalar. Other 7 components $x_1*\mathbf{r_1}$, $x_2*\mathbf{r_2}$, $x_3*\mathbf{r_3}$, $x_4*\mathbf{r_4}$, $x_5*\mathbf{r_5}$, $x_6*\mathbf{r_6}$, $x_7*\mathbf{r_7}$ are imaginary units.

The multiplication table (Figure 9) shows that R_{123}-octetons contain 2 sub-algebras of two-dimensional complex numbers z in their (8*8)-matrix forms of presentation: $z=x_0*\mathbf{r_0}+x_2*\mathbf{r_2}$ and $z=x_0*\mathbf{r_0}+x_3*\mathbf{r_3}$. It is interesting because, as science knows, two-dimensional complex numbers, which are a sum of a real item and imaginary item, have appeared as outstanding instruments for the development of theories and calculations in the field of physical problems of heat, light, sounds, vibrations, elasticity, gravitation, magnetism, electricity, liquid streams, and phenomena of the micro-world. Our results in the field of matrix genetics reveal a bunch of (8*8)-matrix forms of presentation of two-dimensional complex numbers in 8-dimensional vector spaces of R_{123}-octetons (below we describe

cases of positional permutations in triplets which lead to additional (8*8)-matrix forms of presentations of complex numbers and of double numbers). From a formal viewpoint, the mentioned branches of theoretical physics can be interpreted in a frame of the 8-dimensional vector space of octetons. Simultaneously the multiplication table (Figure 9) shows that R_{123}-octetons contain 5 sub-algebras of two-dimensional double numbers "d" in their (8*8)-matrix forms of presentation: $d=x_0*\mathbf{r_0}+x_1*\mathbf{r_1}$, $d=x_0*\mathbf{r_0}+x_4*\mathbf{r_4}$, $d=x_0*\mathbf{r_0}+x_5*\mathbf{r_5}$, $d=x_0*\mathbf{r_0}+x_6*\mathbf{r_6}$, $d=x_0*\mathbf{r_0}+x_7*\mathbf{r_7}$. It is known that each complex number can be interpreted in three ways: as a point, or a vector or an operator in a two-dimensional vector space. The same is true for the case of each complex numbers in 8-dimensional vector spaces of octetons. The work (Petoukhov, 2011) describes some interesting mathematical properties and possible biological applications of genetic octetons. Moreover an application of this 8-dimensional genetic algebra allows discovering new phenomenological rules of diversity of dialects of the genetic code (Petoukhov, 2011).

Let us mention here one interesting aspect of R_{123}-octetons. One can replace the symbolic genomatrix [C A; U G](3) (see Figure 5) by a special numeric presentation where each of the triplets is replaced by twice the number of hydrogen bonds in its suffix (it is a letter on the third position of the triplet). For example, all the triplets with suffixes C and G (which possess 3 hydrogen bonds in their complementary pairs in DNA) are replaced by number 2*3=6, and all the triplets with suffixes A and U (which possess 2 hydrogen bonds in their complementary pairs) are replaced by number 2*2=4. These numbers should be taken with the sign "+" for triplets in black cells of the genomatrix on Figure 4 and with the sign "-" for triplets in white cells. One can check that in this case the final numeric matrix is the matrix form of presentation of the following R_{123}-octetons (2) with a separation of coordinate axes with odd and even indexes:

$$R_{123} = 6*(r_0+r_2+r_4+r_6) + 4*(r_1+r_3+r_5+r_7) \quad (2)$$

This special R_{123}-octeton (2) with integer values of all the coordinates is interesting because its square root gives the "golden" R_{123}-octeton (3) whose coordinates are equal to irrational numbers of the golden section $f=(1+5^{0.5})/2$ and of its inverse value f^{-1}:

$$R_{123} = f*(r_0 + r_2 + r_4 + r_6) + f^{-1}*(r_1 + r_3 + r_5 + r_7) \quad (3)$$

The theme of the golden section and Fibonacci numbers in structural properties of genetic alphabets appears unexpectedly in many cases (Petoukhov, 2008a; Petoukhov and He, 2010). This theme has attracted an increased attention of the scientific community because it is related with quasi-crystals by D.Shechtman (1984) and his Nobel Prize-2011 (Shechtman, 2011).

4. Hadamard genomatrices and 8-dimensional hypercomplex numbers

By definition a Hadamard matrix of dimension "n" is the (n*n)-matrix $H_{(n)}$ with elements "+1" and "-1". It satisfies the condition $H_{(n)}*H_{(n)}^T = n*I_n$, where $H_{(n)}^T$ is the transposed matrix and I_n is the identity (n*n)-matrix. Rows of Hadamard matrices are termed Walsh functions. Hadamard matrices are widely used in error-correcting codes such as the Reed-Muller code and Hadamard codes; in the theory of compression of signals and images; in spectral analysis and multi-channel spectrometers with Hadamard transformations; in quantum computers with Hadamard gates (Nielsen and Chuang, 20010); in a realization of Boolean functions by means of spectral methods; in the theory of planning of multiple-factor experiments and in many other branches of science and technology. The works (Petoukhov, 2005b, 2008a, b) have revealed that Kronecker families of genetic matrices are related to some kinds of Hadamard matrices ("Hadamard genomatrices") by means of so termed U-algorithm.

This section describes that the dyadic-shift decompositions of Hadamard genomatrices lead to special 8-dimensional hypercomplex numbers. More precisely Hadamard genomatrices are the sum of basic matrices of the DS-algebra of 8-dimensional hypercomplex numbers; or, in other words, Hadamard (8x8)-genomatrices are the 8-dimensional hypercomplex number whose coordinates are equal to 1.

Let us begin the description of relations of Hadamard matrices with the system of genetic alphabets. For the U-algorithm, phenomenological facts are essential that the letter U in RNA (and correspondingly the letter T in DNA) is a very special letter in the 4-letter alphabet of nitrogenous bases in the following two aspects:

- Each of three nitrogenous bases A, C, G has one amino-group NH_2 (Figure 3), but the fourth basis U/T does not have it. From the viewpoint of the existence of the amino-group (which is very important for genetic functions) the letters A, C, G are identical to each other and the letter U is opposite to them;
- The letter U is a single letter in RNA, which is replaced in DNA by another letter, the T.

This uniqueness of the letter U can be utilized in genetic computers of organisms. Taking into account the unique status of the letter U in this genetic alphabet, the author has shown the existence of the following formal "U-algorithm", which demonstrates the close connection between Hadamard matrices and the matrix mosaic of the degeneracy of the genetic code (Petoukhov, 2005b, 2008a, b).

By definition the U-algorithm contains two steps: 1) on the first step, each of the triplets in the black-and-white genomatrix (for example, in the genomatrix [C A; U G](3) on Figure 5) should change its own color into opposite color each time when the letter U stands in an odd position (in the first or in the third position) inside the triplet; 2) on the second step, black triplets and white triples are interpreted as entries "+1" and "-1" correspondingly. For example, the white triplet UUA (see Figure 5) should become the black triplet (and its matrix cell should be marked by black color) because of the letter U in its first position; for this reason the triplet UUA is interpreted finally as "+1". Or the white triplet UUU should not change its color because of the letter U in its first and third positions (the color of this triplet is changed twice according to the described algorithm); for this reason the triplet UUU is interpreted finally as "-1". The triplet ACG does not change its color because the letter U is absent in this triplet.

By means of the U-algorithm, the black-and-white mosaic of the genomatrix [C A; U G](3) (Figure 5) is transformed into relevant numeric genomatrix H_{123} (Figure 10).

$$H_{123} = h_0+h_1+h_2+h_3+h_4+h_5+h_6+h_7 =$$

$$\begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1 \end{bmatrix} + \begin{bmatrix} 0&1&0&0&0&0&0&0\\ -1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&-1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&-1&0&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&-1&0 \end{bmatrix} + \begin{bmatrix} 0&0&-1&0&0&0&0&0\\ 0&0&0&-1&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&0&0&0&0&-1&0\\ 0&0&0&0&0&0&0&-1\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0 \end{bmatrix} + \begin{bmatrix} 0&0&0&-1&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ -1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&-1\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&-1&0&0&0 \end{bmatrix} +$$

$$\begin{bmatrix} 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&-1&0\\ 0&0&0&0&0&0&0&-1\\ -1&0&0&0&0&0&0&0\\ 0&-1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0 \end{bmatrix} + \begin{bmatrix} 0&0&0&0&0&1&0&0\\ 0&0&0&0&-1&0&0&0\\ 0&0&0&0&0&0&0&-1\\ 0&0&0&0&0&0&1&0\\ 0&-1&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&-1&0&0&0&0&0 \end{bmatrix} + \begin{bmatrix} 0&0&0&0&0&0&-1&0\\ 0&0&0&0&0&0&0&-1\\ 0&0&0&0&-1&0&0&0\\ 0&0&0&0&0&-1&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0 \end{bmatrix} + \begin{bmatrix} 0&0&0&0&0&0&0&-1\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&-1&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&-1&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ -1&0&0&0&0&0&0&0 \end{bmatrix}$$

Figure 11. The dyadic-shift decomposition of the Hadamard genomatrix H_{123} (Figure 10) into sum of 8 sparse matrices h_0, h_1,..., h_7.

H_{123} =

(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	(0)	(3)	(2)	(5)	(4)	(7)	(6)
(2)	(3)	(0)	(1)	(6)	(7)	(4)	(5)
(3)	(2)	(1)	(0)	(7)	(6)	(5)	(4)
(4)	(5)	(6)	(7)	(0)	(1)	(2)	(3)
(5)	(4)	(7)	(6)	(1)	(0)	(3)	(2)
(6)	(7)	(4)	(5)	(2)	(3)	(0)	(1)
(7)	(6)	(5)	(4)	(3)	(2)	(1)	(0)

Figure 10. The Hadamard genomatrices H_{123} which are received from the genomatrix [C A; U G][3] (Figure 5) by means of the U-algorithm. Brackets contain dyadic numerations of cells by analogy with the matrix on Figure 5. Black color and white color of cells mean entries "+1" and "-1" in these cells correspondingly.

One can make the dyadic-shift decomposition of this Hadamard genomatrix H_{123} (Figure 10) by analogy with the described decomposition of the relevant matrix R on Figures 7 and 8. In the result a new set of sparse matrices h_0, h_1, h_2, h_3, h_4, h_5, h_6, h_7 arise where h_0 is an identity matrix (Figure 11). It is unexpected but this set of the Hadamard genomatrix is closed in relation to multiplication and it corresponds to a new multiplication table on Figure 12.
The existence of the multiplication n table (Figure 12) means that a new 8-dimensional DS-algebra or a new system of 8-dimensional hypercomplex numbers (4) exists on the base of this Hadamard genomatrix H_{123}.

$$H = x_0*\mathbf{1} + x_1*\mathbf{h_1} + x_2*\mathbf{h_2} + x_3*\mathbf{h_3} + x_4*\mathbf{h_4} + x_5*\mathbf{h_5} + x_6*\mathbf{h_6} + x_7*\mathbf{h_7} \quad (4)$$

	1	$\mathbf{h_1}$	$\mathbf{h_2}$	$\mathbf{h_3}$	$\mathbf{h_4}$	$\mathbf{h_5}$	$\mathbf{h_6}$	$\mathbf{h_7}$
1	**1**	$\mathbf{h_1}$	$\mathbf{h_2}$	$\mathbf{h_3}$	$\mathbf{h_4}$	$\mathbf{h_5}$	$\mathbf{h_6}$	$\mathbf{h_7}$
$\mathbf{h_1}$	$\mathbf{h_1}$	**-1**	$\mathbf{h_3}$	$\mathbf{-h_2}$	$\mathbf{h_5}$	$\mathbf{-h_4}$	$\mathbf{h_7}$	$\mathbf{-h_6}$
$\mathbf{h_2}$	$\mathbf{h_2}$	$\mathbf{h_3}$	**-1**	$\mathbf{-h_1}$	$\mathbf{-h_6}$	$\mathbf{-h_7}$	$\mathbf{h_4}$	$\mathbf{h_5}$
$\mathbf{h_3}$	$\mathbf{h_3}$	$\mathbf{-h_2}$	$\mathbf{-h_1}$	**1**	$\mathbf{-h_7}$	$\mathbf{h_6}$	$\mathbf{h_5}$	$\mathbf{-h_4}$
$\mathbf{h_4}$	$\mathbf{h_4}$	$\mathbf{h_5}$	$\mathbf{h_6}$	$\mathbf{h_7}$	**-1**	$\mathbf{-h_1}$	$\mathbf{-h_2}$	$\mathbf{-h_3}$
$\mathbf{h_5}$	$\mathbf{h_5}$	$\mathbf{-h_4}$	$\mathbf{h_7}$	$\mathbf{-h_6}$	$\mathbf{-h_1}$	**1**	$\mathbf{-h_3}$	$\mathbf{h_2}$
$\mathbf{h_6}$	$\mathbf{h_6}$	$\mathbf{h_7}$	$\mathbf{-h_4}$	$\mathbf{-h_5}$	$\mathbf{h_2}$	$\mathbf{h_3}$	**-1**	$\mathbf{-h_1}$
$\mathbf{h_7}$	$\mathbf{h_7}$	$\mathbf{-h_6}$	$\mathbf{-h_5}$	$\mathbf{h_4}$	$\mathbf{h_3}$	$\mathbf{-h_2}$	$\mathbf{-h_1}$	**1**

Figure 12. The multiplication table for the dyadic-shift decomposition of Hadamard genomatrix H_{123} (Figures 10 and 11). Here the identity matrix h_0 is replaced by the symbol **1** (Petoukhov, 2011).

We term these new 8-dimensional hypercomplex numbers as H-octetons (here "H" is the first letter in the name Hadamard) because they differ from R-octetons (Figure 9) and from Cayley's octonions. The DS-algebra of H-octetons (Figure 12) is non-commutative associative non-division algebra. It has zero divisors: for example (h_3+h_4) and (h_2-h_5) are non-zero H-octetons, but their product is equal to zero.
It should be noted that Hadamard matrices play important roles in many tasks of discrete signal processing; tens of thousands of publications devoted to them (see a review in (Seberry et *al.,* 2005)). For example, codes based on Hadamard matrices have been used on spacecrafts "Mariner" and "Voyager", which allowed obtaining high-quality photos of Mars, Jupiter, Saturn, Uranus and Neptune in spite of the distortion and weakening of the incoming signals.

Only a few symmetrical Hadamard matrices are usually used in the field of discrete signal processing (Ahmed and Rao, 1975; Trahtman and Trahtman, 1975). But, as we have checked, dyadic-shift decompositions of these "engineering" Hadamard matrices do not lead to any 8-dimensional hypercomplex numbers in contrast to the asymmetrical Hadamard genomatrices described in our article. Moreover the author knows no publications about the facts that Hadamard matrices can be the base for matrix forms of presentation of 8-dimensional hypercomplex numbers. It seems that the genetic code has led the author to discovering the new interesting fact in the field of theory of Hadamard matrices about the unexpected relation of some Hadamard matrices with multidimensional DS-algebras of hypercomplex numbers. This fact can be useful for many applications of Hadamard genomatrices for simulating of bioinformation phenomena, for technology of discrete signal processing, etc. A great number of Hadamard (8x8)-matrices exists (according to some experts, their quantity is equal to approximately 5 billion). Perhaps, only the genetic Hadamard matrices, which represent a small subset of a great set of all the Hadamard matrices, are related with multidimensional DS-algebras but it is an open question now.

Why living nature uses just such the genetic code that is associated with Hadamard genomatrices? We suppose that

its reason is related with solving in biological organisms the same information tasks which lead to a wide use of Hadamard matrices in digital signal processing and in physics. The orthogonal systems of Walsh functions in Hadamard genomatrices can be a natural basis to organize storage and transfer of genetic information with noise-immunity properties, etc. The described Hadamard genomatrices can be considered as a patent by the living nature which was discovered through matrix genetics. It should be emphasized that few authors have applied orthogonal systems of Walsh functions from "engineering" Hadamard matrices to study macro-physiological systems (Shiozaki, 1980; Carl, 1974; Ginsburg *et al.*, 1974). It seems that applications of "genetic" Walsh functions from Hadamard genomatrices are right now more relevant for such macro-physiological researches.

5. Matrix algebras in cases of positional permutations in triplets

The theory of discrete signal processing pays special attention to permutations of information elements. This section shows that all the possible permutations of positions inside all the triplets in the genetic matrix [C A; U G]$^{(3)}$ (Figures 5 and 7) lead to new mosaic genomatrices whose Rademacher forms of presentation are connected with the same DS-algebra (Figure 9). Let us describe this situation in more details.

It is obvious that a simultaneous permutation of positions in triplets transforms the most of the triplets in cells of the initial genomatrix [C A; U G]$^{(3)}$ and also transforms the black-and-white mosaic of this matrix. For example, in the case of the cyclic transformation of the order 1-2-3 of positions into the order 2-3-1, the black triplet CGA is transformed into the white triplet GAC, etc. Because each of the triplets is connected with the binary numeration of its column and row, these binary numerations are also transformed correspondingly; for example, the binary numeration 011 is transformed into 110.

Six variants of the order of positions inside triplets are possible: 1-2-3, 2-3-1, 3-1-2, 3-2-1, 2-1-3, 1-3-2. This set of variants includes two sub-sets of cyclic permutations of positions: 1) 1-2-3, 2-3-1, 3-1-2; 2) 3-2-1, 2-1-3, 1-3-2. The initial genomatrix [C A; U G]$_{123}^{(3)}$ is related with the first of these orders (Figure 5). Other five genomatrices [C A; U G]$_{231}^{(3)}$, [C A; U G]$_{312}^{(3)}$, [C A; U G]$_{321}^{(3)}$, [C A; U G]$_{213}^{(3)}$, [C A; U G]$_{132}^{(3)}$, which correspond to other five orders, have their black-and-white mosaics where each row corresponds again to one of Rademacher functions (subscripts indicate the order of positions in triplets). The replacement of all the triplets with strong and weak roots by entries "+1" and "-1" correspondingly transforms these genomatrices into their Rademacher forms R_{231}, R_{312}, R_{321}, R_{213}, R_{132} (see Figure 13, upper) (Petoukhov, 2011).

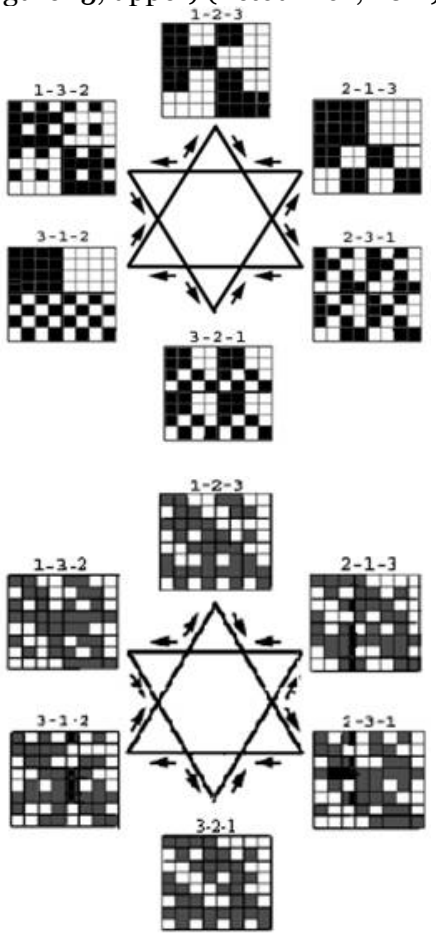

Figure 13. Upper: six Rademacher forms of genetic matrices [C A; U G]$_{123}^{(3)}$, [C A; U G]$_{231}^{(3)}$, [C A; U G]$_{312}^{(3)}$, [C A; U G]$_{321}^{(3)}$, [C A; U G]$_{213}^{(3)}$, [C A; U G]$_{132}^{(3)}$. **Below:** six Hadamard forms of the same genetic matrices. Black color and white color of cells mean entries "+1" and "-1" in these cells correspondingly. Arrows and triangles show interrelations in the two sub-sets of the genetic matrices with cyclic permutations of positions in indexes: 1) 1-2-3, 2-3-1, 3-1-2; 2) 3-2-1, 2-1-3, 1-3-2.

Each of the Rademacher forms R_{231}, R_{312}, R_{321}, R_{213}, R_{132} can be again decomposed

into the sum of 8 sparse matrices r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7 in accordance with dyadic numerations of its cells. In the result 6 different sets of eight sparse matrices r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7 arise (r_0 is identity matrix in all the sets). Unexpected facts are that, firstly, each of these sets is closed in relation to multiplication and, secondly, each of these sets corresponds to the same multiplication table from Figure 9. It means that this genetic DS-algebra of 8-dimensional hypercomplex numbers possesses at least 6 different matrix forms of its presentation (see more details in (Petoukhov, 2011)). In particular the Rademacher forms R_{123}, R_{231}, R_{312}, R_{321}, R_{213}, R_{132} are different matrix forms of presentation of the same R-octeton whose coordinates are equal to 1 ($x_0=x_1=...=x_7=1$). Our results demonstrate that this DS-algebra of genetic octetons possesses a wonderful stability (invariance) relative to many of the variants of positional permutations in triplets and relative to the dyadic shifts on the base of modulo-2 addition.

The similar situation holds true for the Hadamard presentation of the six genetic matrices $[C\ A;\ U\ G]_{123}^{(3)}$, $[C\ A;\ U\ G]_{231}^{(3)}$, $[C\ A;\ U\ G]_{312}^{(3)}$, $[C\ A;\ U\ G]_{321}^{(3)}$, $[C\ A;\ U\ G]_{213}^{(3)}$, $[C\ A;\ U\ G]_{132}^{(3)}$ (Figure 13, below). One can make the dyadic-shift decomposition of each of these six Hadamard genomatrices H_{123}, H_{231}, H_{312}, H_{321}, H_{213}, H_{132} by analogy with the described decompositions of the relevant genomatrices R_{123}, R_{231}, R_{312}, R_{321}, R_{213}, R_{132}. In the result, six new different sets of basic sparse matrices h_0, h_1, h_2, h_3, h_4, h_5, h_6, h_7 arise (where h_0 is the identity matrix). It is unexpectedly but each of these six sets for Hadamard genomatrices is closed in relation to multiplication. Moreover each of these sets h_0, h_1, h_2, h_3, h_4, h_5, h_6, h_7 corresponds to the same multiplication table on Figure 12. The six Hadamard genomatrices H_{123}, H_{231}, H_{312}, H_{321}, H_{213}, H_{132} are different matrix forms of presentation of the same H-octeton whose coordinates are equal to 1 ($x_0=x_1=...=x_7=1$) (Petoukhov, 2011).

Discussion

The author has revealed a close relation of the genetic code with 8-dimensional hypercomplex numbers (first of all, R-octetons and H-octetons) and with dyadic shifts and Hadamard matrices. This relation is interesting in many aspects. Some of them are the following.

Numeric presentations of genetic sequences are useful to study hidden genetic regularities (Cristea, 2002, 2010; Petoukhov and He, 2010). On the base of the described results, new approaches of numeric presentations of genetic sequences can be proposed for such aims. It seems appropriate to interpret genetic sequences as sequences of 8-dimensional vectors where genetic elements are replaced by their special numeric presentations which are connected with the described DS-algebras. Then Hadamard spectrums, dyadic distances and some other characteristics of these vector sequences can be studied.

Molecular genetics possesses examples of a special phenomenological role of number 8. For instance, in eukaryote cells, filaments of DNA are coiled around nucleosomes, each of which is an octamer shank consisting of the histones of the four types: H2A, H2B, H3 and H4. More precisely, a single nucleosome contains the ensemble of eight histones, where two histones of each of the four types H2A, H2B, H3 and H4 are included. The DNA molecule is reeled up on this octamer shank. The octamer structure of nucleosome plays the main role in the packing of DNA on all levels. The wide-known concept of the histones code exists in molecular genetics. One can hope that mathematical formalisms of genetic octetons can be used to simulate some phenomenological facts in this fundamental concept.

Let us go further. The following question arises additionally: why the molecular-genetic system is connected with associative algebras of genetic octetons but not with non-associative algebras of Cayley's octonions and split-octonions which are very popular in theoretical physics of non-living matter? The possible reason is that associativity is very important in the field of noise-immunity coding, where many types of codes (for example cyclic codes) are widely used which are based on the concept of algebraic groups and of Galois fields. But the notions of a group and Galois fields contain conditions of an associativity of their elements. Consequently, associative algebras of octetons have the fundamental advantage

in the field of noise-immunity coding in comparison with the non-associative algebras of Cayley's octonions and split-octonions.

The wide using of Hadamard matrices and their Walsh functions in digital technologies is based on their properties. Firstly, stair-step functions by Walsh are implemented technologically much simpler than trigonometric functions and many others. Secondly, the using of Walsh functions allows providing the processing of digital signals by means of operations of addition and subtraction only without using multiplication and division (Ahmed and Rao, 1975). Since division operations are not required here, digital informatics can use algebras without dividing, for example the described DS-algebras of octetons. This distinguishes computer informatics from theoretical physics, where multiplication and division are essential, and therefore attention of theorists is drawn to the Cayley algebra of octonions which includes a division operation. It is essential that for molecular-genetic systems operations of additions and subtractions can be organized simply by means of interconnections or disconnections of molecular elements which is easier than molecular organization of multiplication operations. It should be mentioned that some works about applications of Walsh functions for spectral analysis of genetic sequences and genetic algorithms are known (Geadah and Corinthios, 1977; Goldberg, 1989; Shiozaki, 1980; Vose and Wright, 1998; Waterman, 1999). These functions can be useful in researches in algebraic biology (Pellionisz *et al.,* 2011; Petoukhov and He, 2010, etc).

It is known that multi-dimensional numeric systems are used in the field of multi-dimensional digital signal processing in cases of multi-channel communication to provide some functional advantages. For example 2-dimensional complex numbers are used in digital processing of 2-dimensional signals (Lyons, 2004, Chapter 8); in this case the real part of complex numbers corresponds to a signal of the first channel and the imaginary part of complex numbers corresponds to a signal of the second channel. But biological organisms are systems of multi-channel informatics. For example a retina of eyes possesses an inherited set of receptors that define a multi-channel transmission of information about a projected image from a great number of separate receptors into the nervous system. But a cooperative principle is needed for systems with a great number of independent channels to operate efficiently with multi-channel flows of signals; such cooperative principle can be realized in a form of an algebraic system of multi-dimensional hypercomplex numbers.

We suppose that the described genetic octetons are used in inherited biological multi-channel informatics for this purpose. One can add that in the field of molecular genetics whole families of proteins should work to provide physiological processes in an active and living context (a separate protein is not a living substance). But these families should contain individual quantities of proteins of different types. The genetic system successfully solves this numeric task of a genetic determination of amounts of proteins of each type inside separate families of proteins for various physiological processes. The described genetic octetons can be additionally used to construct mathematical models of this genetic phenomenon. Generally speaking, we suppose that in the phenomenological field of molecular genetics and inherited physiological systems, the 8-dimensional genetic algebras (which have been revealed in our works) can be a natural genetic basis to simulate numeric regularities in inherited families of information elements. In our opinion, these genetic algebras can be used also in development of "science of consciousness" (Penrose, 1996) and in solving many other questions of the genetic system and inherited physiological functions (see for example (Castro-Chavez, 2011)).

Many practical tasks are connected with hypercomplex numbers in the field of informatics. For example, different hypercomplex numbers are used to encode information in digital communication (Bulow, 2001; Chernov, 2002; Felberg, 2001; Furman *et al.,* 2003; Sin'kov, 2010; Toyoshima, 1999, 2002). Genetic 8-dimensional hypercomplex numbers and their metric 4-dimensional subspaces can be used to construct new genetic algorithms and new decisions in the field of artificial intelligence, robotics, etc. Some works try to

analyze genetic sequences on the base of complex and hypercomplex numbers but in these attempts a difficult problem exists: what kind of hypercomplex numbers should be chosen for the analysis from a great set of all the types of hypercomplex numbers (Cristea, 2002, 2010; Shu and Li, 2010; Shu and Ouw, 2004)? It seems obviously that hypercomplex numbers, which are described in our article, should be actively used in analyzing genetic sequences.

Orthogonal systems of Walsh functions play the main role in the fruitful sequency theory by Harmuth for signal processing (Harmuth, 1970, 1977, 1981, 1989). Rows of Hadamard genomatrices correspond to very special kinds of Walsh functions ("genosystems" of Walsh functions) which define special ("genetic") variants of sequency analysis in signal processing. The author believes that this "genetic" sequency analysis, whose bases have been revealed in the field of matrix genetics, can be a key to understand important features not only of genetic informatics but also of many other inherited physiological systems (morphogenetic, sensori-motor, etc.).

In comparison with spectral analysis by means of sine waves, which is applicable to linear time-invariant systems, the sequency analysis is based on non-sinusoidal waves and it is used to study systems which are changed in time (biological systems belong to such systems) (Harmuth, 1977, 1989). The author believes that mechanisms of biological morphogenesis are closely associated with spatial and temporal filters from the field of sequency analysis for genetic systems (the general theory of spatial and temporary filters of sequency theory has been described by Harmuth). From this viewpoint, many inherited morphogenetic abnormalities are consequences of disruptions in physiological spatial and temporal filters of sequency types. Taking into account the sequency theory by Harmuth together with our data about Hadamard genomatrices and genetic H-octetons, one can assume that biological evolution can be interpreted largely like the evolution of physiological spatial and temporal filters of sequency types. In this direction of thoughts, the author develops in his Laboratory new approaches in bioinformatics, bioengineering and medicine.

A special attention should be paid to algebraic analogies at different levels of biological organization from molecular-genetic alphabets up to independent inheritance of traits in holistic organisms (algebraic properties of Punnett squares, etc.). In our opinion these analogies are included as a particular part in our conception of "alphabetic-molecular Mendelism" (Petoukhov, 2004). This conception argues that supramolecular phenomena of inheritance of traits in holistic biological organisms, which are described by the Mendel's laws, did not arise at an empty place, but they are a continuation of molecular-structural phenomena which are defined by deeper laws of the alphabetic-molecular level. These deeper laws possess important analogies with Mendel's laws. A set of inherited traits in biological organisms can be considered as a special alphabetic system of a high level. From such viewpoint, biological evolution can be interpreted in some extend as an evolution of multilevel systems of inherited and interconnected biological alphabets beginning at least from the molecular-genetic level. We suppose an existence of universal bio-algorithms of evolutionary producing of interrelated biological alphabets at higher and higher levels of organization. Kronecker multiplications of genetic matrices, algorithms of dyadic shifts and some other algorithmic operations described in our works can be used to simulate this evolutionary process of a complication of a multilevel system of biological alphabets.

Appendix A
The Kronecker product of matrices

The Kronecker product is an operation on two matrices of arbitrary size resulting in a block matrix (http://en.wikipedia.org/wiki/Kronecker_product).

The Kronecker product should not be confused with usual matrix multiplication, which is an entirely different operation. It is named after German mathematician Leopold Kronecker. If one has two square matrices A = || a_{ij} || and B = || b_{kp} ||, where i, j = 1,..., m and k, p = 1,..., n, then a square block matrix

$$C = A \otimes B = \begin{Vmatrix} a_{ij}*b_{kp} \end{Vmatrix}$$

is called the Kronecker product of the matrices A and B. The Kronecker product of matrices arises in a natural way in a problem of searching a matrix. The eigenvalues of matrix A⊗B are equal to a product of c_i*d_j, where c_i and d_j are eigenvalues of the matrices A and B (Bellman, 1960). The Kronecker product is connected with fractal structures (Gazale, 1999) and is used widely in many fields of mathematics, theoretical physics, informatics, etc.

Appendix B.
Modulo-2 addition, dyadic shifts and matrices of dyadic shifts

For special decompositions of genetic matrices we use structures of matrices of dyadic shifts long known in theory of discrete signal processing, sequency theory by Harmuth, etc. (Ahmed and Rao, 1975; Harmuth, 1977, §1.2.6). These matrices are constructed on the basis of mathematical operation of modulo-2 addition for binary numbers.

Modulo-2 addition is utilized broadly in the theory of discrete signal processing as a fundamental operation for binary variables. By definition, the modulo-2 addition of two numbers written in binary notation is made in a bitwise manner in accordance with the following rules:

$$0 + 0 = 0,\ 0 + 1 = 1,\ 1 + 0 = 1,\ 1 + 1 = 0 \quad \text{(B.1)}$$

For example, modulo-2 addition of two binary numbers 110 and 101, which are equal to 6 and 5 respectively in decimal notation, gives the result 110 ⊕ 101 = 011, which is equal to 3 in decimal notation (⊕ is the symbol for modulo-2 addition). The set of binary numbers

$$000,\ 001,\ 010,\ 011,\ 100,\ 101,\ 110,\ 111 \quad \text{(B.2)}$$

forms a diadic group, in which modulo-2 addition serves as the group operation (Harmuth, 1989). The distance in this symmetry group is known as the Hamming distance. Since the Hamming distance satisfies the conditions of a metric group, the diadic group is a metric group. The modulo-2 addition of any two binary numbers from (B.2) always results in a new number from the same series. The number 000 serves as the unit element of this group: for example, 010 ⊕ 000 = 010. The reverse element for any number in this group is the number itself: for example, 010 ⊕ 010 = 000. The series (B.2) is transformed by modulo-2 addition with the binary number 001 into a new series of the same numbers:

$$001,\ 000,\ 011,\ 010,\ 101,\ 100,\ 111,\ 110 \quad \text{(B.3)}$$

Such changes in the initial binary sequence, produced by modulo-2 addition of its members with any of binary numbers (B.2), are termed diadic shifts (Ahmed and Rao, 1975; Harmuth, 1989). If any system of elements demonstrates its connection with diadic shifts, it indicates that the structural organization of its system is related to the logic of modulo-2 addition. The article shows that the structural organization of genetic systems is related to logic of modulo-2 addition.

	000 (0)	001 (1)	010 (2)	011 (3)	100 (4)	101 (5)	110 (6)	111 (7)
000 (0)	000 (0)	001 (1)	010 (2)	011 (3)	100 (4)	101 (5)	110 (6)	111 (7)
001 (1)	001 (1)	000 (0)	011 (3)	010 (2)	101 (5)	100 (4)	111 (7)	110 (6)
010 (2)	010 (2)	011 (3)	000 (0)	001 (1)	110 (6)	111 (7)	100 (4)	101 (5)
011 (3)	011 (3)	010 (2)	001 (1)	000 (0)	111 (7)	110 (6)	101 (5)	100 (4)
100 (4)	100 (4)	101 (5)	110 (6)	111 (7)	000 (0)	001 (1)	010 (2)	011 (3)
101 (5)	101 (5)	100 (4)	111 (7)	110 (6)	001 (1)	000 (0)	011 (3)	010 (2)
110 (6)	110 (6)	111 (7)	100 (4)	101 (5)	010 (2)	011 (3)	000 (0)	001 (1)
111 (7)	111 (7)	110 (6)	101 (5)	100 (4)	011 (3)	010 (2)	001 (1)	000 (0)

Figure B.1. An example of a (8*8)-matrix of dyadic shifts. Parentheses contain expressions of numbers in decimal notation.

Figure B.1 shows an example of matrices of dyadic shifts. Each row and each column are numerated by means of binary numbers from the dyadic group (B.2). Each of matrix cells has its binary numeration from the same dyadic group (B.2). This numeration of any cell is the result of modulo-2 addition of binary numerations of the column and the row of this cell. For example, the cell from the column 110 and the row 101 obtains the binary numeration 011 by means of such addition. Such numerations of matrix cells are termed “dyadic-shift numerations” (or simply “dyadic numeration”). One can see that the disposition of even and odd numbers 0, 1, 2,..., 7 of the dyadic group in this matrix of dyadic shifts is identical to the disposition of white and black cells in a chessboard.

Let us show a connection of the Punnett matrices with matrices of dyadic shifts on the example of the trihybrid cross with identical maternal and paternal genotypes HhBbCc (Figure B.2). This Punnett matrix can be presented (with a preciseness up to the order of the elements) in a form of Kronecker multiplication of three Punnett matrices for monohybrid crosses (Figure B.2): [HH Hh; Hh hh]⊗ [BB Bb; Bb bb]⊗[CC Cc; Cc cc].

	maternal gametes	
	H	h
H	HH	Hh
h	Hh	hh

;

	maternal gametes	
	B	b
B	BB	Bb
b	Bb	bb

;

	maternal gametes	
	C	c
C	CC	Cc
c	Cc	cc

Figure B.2. Three Punnett matrices of monohybrid crosses. Their Kronecker multiplication gives the Punnet matrix of the trihybrid cross on Figure B.3.

	HBC(000)	HBc(001)	HbC(010)	Hbc(011)	hBC(100)	hBc(101)	hbC(110)	hbc(111)
HBC (000)	HHBBCC (000)	HHBBCc (001)	HHBbCC (010)	HHBbCc (011)	HhBBCC (100)	HhBBCc (101)	HhBbCC (110)	HhBbCc (111)
HBc (001)	HHBBCc (001)	HHBBcc (000)	HHBbCc (011)	HHBbcc (010)	HhBBCc (101)	HhBBcc (100)	HhBbCc (111)	HhBbcc (110)
HbC (010)	HHBbCC (010)	HHBbCc (011)	HHbbCC (000)	HHbbCc (001)	HhBbCC (110)	HhBbCc (111)	HhbbCC (100)	HhbbCc (101)
Hbc (011)	HHBbCc (011)	HHBbcc (010)	HHbbCc (001)	HHbbcc (000)	HhBbCc (111)	HhBbcc (110)	HhbbCc (101)	Hhbbcc (100)
hBC (100)	HhBBCC (100)	HhBBCc (101)	HhBbCC (110)	HhBbCc (111)	hhBBCC (000)	hhBBCc (001)	hhBbCC (010)	hhBbCc (011)
hBc (101)	HhBBCc (101)	HhBBcc (100)	HhBbCc (111)	HhBbcc (110)	hhBBCc (001)	hhBBcc (000)	hhBbCc (011)	hhBbcc (010)
hbC (110)	HhBbCC (110)	HhBbCc (111)	HhbbCC (100)	HhbbCc (101)	hhBbCC (010)	hhBbCc (011)	hhbbCC (000)	hhbbCc (001)
hbc (111)	HhBbCc (111)	HhBbcc (110)	HhbbCc (101)	Hhbbcc (100)	hhBbCc (011)	hhBbcc (010)	hhbbCc (001)	hhbbcc (000)

Figure B.3. The Punnett matrix of the trihybrid cross in the case of identical maternal and paternal genotypes HhBbCc. Numeration of columns and rows is presented not only in the letter form but also in numeric bitwise manner by means of the denotation of each of uppercase letters H, B, C as number 0 and each of the lowercase letters h, b, c as number 1. Binary numeration of each cell (or each of combinations of alleles) is received by means of modulo-2 addition of binary numerations of its column and row by analogy with the matrix of dyadic shifts on Figure B.1.

One of simple ways to show the mentioned connection with dyadic shifts is the following. In the three-letter numerations of columns and rows of the Punnett matrix (Figure B.3), each of the uppercase letters H, B, C can be denoted by means of the number 0, and each of the lowercase letters h, b, c can be denoted by means of the number 1. In this case for example the letter numerations of columns and rows HBC, HBc, HbC, Hbc, hBC, hBc, hbC, hbc are becoming numerical numerations 000, 001, 010, 011, 100, 101, 110, 111 correspondingly (Figure B.3). Using modulo-2 addition of these binary numeration of columns and rows of the Punnett matrix by analogy with the matrix of dyadic shifts on Figure B.1 gives the following binary numeration of cells of the Punnett matrix (or combinations of alleles inside the matrix): 000, 001, …, 111 (in decimal notation 0, 1,…, 7) (Figure B.3).

One can see on Figure B.3 that such dyadic-shift numeration devides the set of combination of alleles in the Punnett matrix into 8 different sub-sets of homozygous or heterozygous combinations:

- the binary numeration 000 denotes all the homozygous combinations of alleles HHBBCC, HHBBcc, HHbbCC, HHbbcc, hhBBCC, hhBBcc, hhbbCC, hhbbcc (they stand on the main diagonal of the matrix);
- the binary numeration 001 denotes all the combinations of alleles where only the third trait is heterozygous: HHBBCc, HHbbCc, hhBBCc, hhbbCc;
- the binary numeration 010 denotes all the combinations of alleles where only the second trait is heterozygous: HHBbCC, HHBbcc, hhBbCC, hhBbcc;
- the binary numeration 011 denotes all the combinations of alleles where only the first trait is homozygous: HHBbCc, hhBbCc;
- the binary numeration 100 denotes all the combinations of alleles where only the first trait is heterozygous: HhBBCC, HhBBcc, HhbbCC, Hhbbcc;
- the binary numeration 101 denotes all the combinations of alleles where only the second trait is homozygous: HhBBCc, HhbbCc;
- the binary numeration 110 denotes all the combinations of alleles where only the third trait is homozygous: HhBbCC, HhBbcc;
- the binary numeration 111 denotes all the combinations of alleles where only heterozygous traits exist: HhBbCc.

It is obvious that by analogy with the DS-decompositions of the genetic (8*8)-matrices (Figures 7, 8, 10, 11), this Punnett matrix can be decomposed on the base of DS-decomposition into sum of 8 sparse matrices each of which contains only one of these 8 sub-sets of combinations of alleles. This example illustrates that by means of DS-decompositions of Punnett matrices we get the algorithmic method of classification of different sub-sets of homozygous or heterozygous organisms in cases of multi-hybrid crosses.

Acknowledgments

Described researches were made by the author in the frame of a long-term cooperation between Russian and Hungarian Academies of Sciences and in the frames of programs of “International Society of Symmetry in Bioinformatics” (USA) and of “International Symmetry Association” (Hungary, http://symmetry.hu/). The author is grateful to Ganiev R.F., Darvas G., He M., Kappraff J., Adamson G., Cristea P., Kulakov Y.I., Negadi T., Pavlov D.G., Pellionisz A.J, Smolianinov V.V., Stepanian I.V., Vladimirov Y.S., Katanov D., Svirin V.I. for their support.

References

Ahmed NU, Rao KR. Orthogonal transforms for digital signal processing, Springer-Verlag New York, Inc., 1975.

Bellman R. Introduction in Matrix Analysis. N.Y.: McGraw Hill Book Company, 1960.

Blahut RE. Fast algorithms for digital signal processing. California: Addison-Wesley Publishing Company, 1985.

Bulow T. Non-commutative Hypercomplex Fourier Transforms. Geometric Computing with Clifford Algebras, Springer-Verlag, Berlin, 2001: 187-207.

Capra F. The Tao of physics: an exploration of the parallels between modern physics and eastern mysticism. New Jersey: Shambhala Publications, Inc., 2000.

Carl JV. On the use of Walsh functions in man-made and biological pattern recognition systems. Applications of Walsh function and sequence theory: Proc. Symposium, Washington, 1974, pp9-25.

Castro-Chavez F. The quantum workings of the rotating 64-grid genetic code. NeuroQuantology 2011; 9(4): 728-746.

Chernov VM. Two-dimensional FFT-like Algorithms with Overlapping in Some Hypercomplex Algebras. Optical Memory & Neural Networks 2002; 11 (1): 29-38.

Crew FA. R. C. Punnett. Genetics 1968; 58 (1): 1–7.

Cristea PD. Conversion of nucleotides sequences into genomic signals. J Cell Mol Med 2002, 6 (2): 279-303.

Cristea PD. Symmetries in genomics. Symmetry: Culture and Science 2010, 21 (1-3): 71-86.

Dixon G. Division Algebras: Octonions, Quaternions, Complex Numbers and Algebraic Design of Physics. Springer, 1994.

Dray T and Manoque C. Octonionic Cayley Spinors and E6. http://arxiv.org/abs/0911.2255 Accessed date: February 25, 2010.

Felberg M. Commutative Hypercomplex Fourier Transforms of Multidimensional Signals. Geometric Computing with Clifford Algebras. Springer-Verlag, Berlin, 2001, pp209-229.

Furman YA, Krevetskiy AV, Rozhentsov AA. Complex and hypercomplex system in the processing of multidimensional signals. Publishing house Moscow Phys Mat Lit, 2003.

Gazale MJ. Gnomon. From Pharaons to Fractals. NJ: Princeton Univ. Press., 1999.

Geadah YA, Corinthios MJ. Natural, dyadic and sequency order algorithms and processors for the Walsh-Hadamard transform. IEEE Trans. Comput 1977; 26: 435-442.

Gell-Mann M., Ne'eman Y. The Eightfold Way. N.Y.: W. A. Benjamin, 1964.

Geramita, A. V. Orthogonal designs: quadratic forms and Hadamard matrices. Dekker, 1979.

Ginsburg AP., Carl JV, Kabrisky M, Hall CF., Gill RA. Physiological aspects of a model for the classification of visual images. Adv Cybern Syst 1974; 3: 1289-1305.

Goldberg DE. Genetic algorithms and Walsh functions. Complex systems 1989; 3(2): 129-171.

Halmos PR. Finite–Dimensional Vector spaces. Berlin: Springer, 1974.

Harmuth HF. Transmission of information by orthogonal functions. Berlin: Springer, 1970.

Harmuth HF. Sequency theory. Foundations and applications. N.-Y.: Academic Press, 1977.

Harmuth HF. Nonsinusoidal waves for radar and radio communication. N.-Y.: Academic Press, 1981.

Harmuth HF. Information theory applied to space-time physics. Washington: The Catholic University of America, DC., 1989.

He M, Petoukhov SV. Mathematics of bioinformatics: theory, practice, and applications. USA: John Wiley & Sons, Inc., 2011. http://symmetry.hu/isabm/petoukhov.html

Kappraff J, Petoukhov SV. Symmetries, generalized numbers and harmonic laws in matrix genetics. Symmetry: Culture and Science 2009; 20 (1-4): 23-50.

Karlin S, Ost F, Blaisdell BE. Patterns in DNA and amino acid sequences and their statistical significance", in Mathematical methods for DNA sequences, M. S. Waterman (Ed.), Florida: CRC Press, 1989, pp167-197.

Konopel'chenko BG., Rumer UB. The classification of codons in the genetic code. Doklady Akademii Nauk of the USSR 1975; 223 (2): 471-474.

Lyons RG. Understanding Digital Signal Processing. New Jersey: Prentice Hall, 2004.

Nielsen MA., Chuang IL. Quantum computation and quantum information, Cambridge University Press, 2001.

Pellionisz AJ, Graham R, Pellionisz PA, Perez JC. Recursive genome function of the cerebellum: geometric unification of neuroscience and genomics. - Chapter in: Handbook: "The Cerebellum and Cerebellar Disorders". Ed. Manto M., Springer Verlag. Submitted Oct. 20, Accepted Nov. 1, 2011 (In press).

Penrose, R. Shadows of the Mind: A Search for the Missing Science of Consciousness. Oxford University Press, 1996.

Petoukhov SV. Genetic codes I: binary sub-alphabets, bi-symmetric matrices and golden section; Genetic codes II: numeric rules of degeneracy and the chronocyclic theory. Symmetry: Culture and Science 2001a; 12 (3-4): 255-306.

Petoukhov SV. The bi-periodic table of the genetic code and numbers of protons. Moscow: publishing house MKTS, 2001b (in Russian).

Petoukhov SV. Symmetries in biology. Supplement of the book: Shubnikov AV, Koptsik VA,Symmetry in Art, Moscow, Regular and Chaotic Dinamics, 3-d edition, 2004; 489-546.

Petoukhov SV. 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, World Sci 2005a; 8: 512-532.

Petoukhov SV. Hadamard matrices and quint matrices in matrix presentations of molecular genetic systems. Symmetry: Culture and Science 2005b; 16 (3): 247-266.

Petoukhov SV. Matrix genetics, algebras of the genetic code, noise-immunity. Moscow: Regular and Chaotic Dynamics, 2008a (in Russian; summary in English is on the http://www.igi-global.com/book/symmetrical-analysis-techniques-genetic-systems/962)

Petoukhov SV. The degeneracy of the genetic code and Hadamard matrices. http://arxiv.org/abs/0802.3366 Accessed date: February 22, 2008b.

Petoukhov SV. Matrix genetics, part 2: the degeneracy of the genetic code and the octave algebra with two quasi-real units (the genetic octave Yin-Yang-algebra). http://arxiv.org/abs/0803.3330 Accessed date: March 23, 2008c.

Petoukhov SV. Matrix genetics, part 1: permutations of positions in triplets and symmetries of genetic matrices. http://arxiv.org/abs/0803.0888 Accessed date: March 29, 2010.

Petoukhov SV. The genetic code, 8-dimensional hypercomplex numbers and dyadic shifts. 2011,

http://arxiv.org/abs/1102.3596v4 Accessed date: October 14, 2011.
Petoukhov SV, He M. Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications. Hershey, USA: IGI Global, 2010
Schrodinger E. What is life? The physical aspect of the living cell. Cambridge, University Press, 1955.
Seberry J, Wysocki BJ., Wysocki TA. On some applications of Hadamard matrices. Metrica 2005, 62: 221-239.
Shechtman D., Blech I., Gratias D., Cahn J W. Metallic phase with long-range orientational order and no translational symmetry. Physical Review Letters 1984; 53: 1951—1953.
Shiozaki A. A model of distributed type associated memory with quantized Hadamard transform. Biol Cybern 1980; 38 (1): 19-22.
Shu JJ, Li Y. Hypercomplex cross-correlation of DNA sequences. Journal of Biological Systems 2010; 18 (4): 711-725.
Shu JJ., Ouw LS. Pairwise alignment of the DNA sequence using hypercomplex number representation. Bull Math Biol 2004; 66 (5): 1423-1438.
Sin'kov MV. Finite Hypercomplex Numerical System; Foundations of the Theory; Applications. Kiev, IPPRI NAN Ukrainy, 2010.
Sklar B. Digital Communication. Fundamentals and Applications. Prentice Hall, 2001.
Stent GS. The Coming of the Golden Age. N-Y, The Natural History Press, 1969.
Stewart I. Life's other secret: The new mathematics of the living world. New-York: Penguin, 1999.
Toyoshima H. Design of Hypercomplex All-Pass Filters to Realize Complex Transfer Functions. Toyoshima H., Higuchi S. ed. Proc. Second Int Conf Information, Communications and Signal Processing 1999;B3.4:1-5.
Toyoshima H. Computationally Efficient Implementation of Hypercomplex Digital Filters IEICE Trans Fundamentals 2002; B85-A: 1870-1876.
Trahtman AM., Trahtman VA. The foundations of the theory of discrete signals on finite intervals. Moscow: Sovetskoie Radio, 1975.
Vose M, Wright A. The simple genetic algorithm and the Walsh transform. - Journal of Evolutionary Computation 1998; 6(3): 253-274.
Waterman MS. Editor. Mathematical methods for DNA sequences. Florida: CRC Press, 1999.
Wittmann HG. Ansatze zur Entschlusselung des genetischen Codes. Die Natur wissenschaften 1961; 48 (24): 55-83.
Yarlagadda R, Hershey J. Hadamard matrix analysis and synthesis with applications to communications and signal/image processing. Kluwer Academic Publ., 1997.
Ycas M. The biological code. North-Holland Pub. Co, 1969.