# Algebraic rules for the percentage composition of oligomers in genomes

Sergey V. Petoukhov[1]

[1]-Mechanical Engineering Research Institute of Russian Academy of Sciences.
Russia, 101990, Moscow, M. Kharitonievskiy pereulok, 4,
http://eng.imash.ru/, info@imash.ru

**Comment:** Some data of this article were presented by the author at the following conferences: the Fourth International Conference of Artificial Intelligence, Medical Engineering, Education (03-04 October 2020, Moscow, Russia, http://www.icics.net/conf/2020/AIMEE2020/index.html) and the Second International Symposium on Computer Science, Digital Economy and Intelligent Systems (18-20 December 2020, Moscow, Russia, http://www.icics.net/conf/2020/CSDEIS2020/index.html).

**Abstract**. The article presents the author's results of studying hidden rules of structural organizations of long DNA sequences in eukaryotic and prokaryotic genomes. The results concern some rules of percentages (or probabilities) of $n$-plets in genomes. To reveal such rules, the author considers genomic DNA nucleotide sequences as multilayers sequences of $n$-plets and studies the percentage contents of $n$-plets in different layers. Unexpected rules of invariance of total sums of percentages in certain tetra-groupings of $n$-plets in different layers of genomic DNA sequences are revealed. These discovered rules are candidates for the role of universal genomic rules. A tensor family of matrix representations of interrelated DNA-alphabets of 4 nucleotides, 16 doublets, 64 triplets, and 256 tetraplets is used in the study. This matrix approach allows revealing algebraic properties of the mentioned genetic rules of probabilities, which are useful for developing algebraic and quantum biology. Some analogies of the discovered genetic phenomena with phenomena of Gestalt psychology are noted and discussed. The author connects the received results about the genomic percentages rules with a supposition of P. Jordan, who is one of the creators of quantum mechanics and quantum biology, that life's missing laws are the rules of chance and probability of the quantum world. Additional attention is paid to the algebraic features of the system of structured DNA alphabets and their relationship with the methods of algebraic holography, known in the technique of processing discrete signals. The concept of algebraic-holographic genetics is being developed for the understanding of inherited holographic properties of organisms.

**Keywords**: DNA alphabets, genomes, percentages of $n$-plets, binary-oppositions, tensor family of matrices, tetra-groupings, quantum biology, algebraic holography.

## Contents

## 1. Introduction.

The article continues publications [Petoukhov, 2020a-c] of the author's results of studying hidden rules of structural organizations of long DNA nucleotide sequences (that is, DNA-texts) in eukaryotic and prokaryotic genomes.

One of the founders of quantum mechanics, who introduced also the term "quantum biology," P. Jordan noted the main difference between living and inanimate objects: inanimate objects are controlled by the average random movement of their millions of particles, whose individual influence is negligible, while in a living organism selected – genetic - molecules have a dictatorial influence on the whole living organism. Besides this, claimed that life's missing laws were the rules of chance and probability of the quantum world [Jordan, 1932; McFadden, Al-Khalili, 2018]. From the standpoint of Jordan's statement, the study of probabilities or percentages of *n*-plets (monoplets, doublets, triplets, etc., that is, oligomers with lengths n) in long DNA sequences is important for discovering hidden biological laws and for developing quantum biology. In his previous articles [Petoukhov, 2020a-c], the author described the universal hyperbolic rules of the oligomer cooperative organization of DNA nucleotide sequences in eukaryotic and prokaryotic genomes.

As it is known, there are DNA alphabets of 4 nucleotides, 16 doublets, 64 triplets, etc. (each such alphabet consists of $4^n$ elements of length *n*). In the previous works and in the presented article, the study of genomes is carried out based on the author's method of representing long DNA texts in the form of composite, multilayered texts, in which each $n^{th}$ layer is a sequence of *n*-plets (or of oligomers of fixed length *n*); in other words, each $n^{th}$ layer is a separate DNA-text written in its own alphabet of $4^n$ *n*-plets. For example, in the text ACCTGTAACG... the first layer is a sequence of nucleotides (A-C-C-C-T-G- ...), the second layer is a sequence of doublets (AC-CT-GT-AA-CG- ...), the third is a sequence of triplets (ACC- TGT- AAC- ..), etc. In each $n^{th}$ layer, we calculate the percentages of each of the $4^n$ types of

*n*-plets and study the relationship between all of them in different layers. This approach to the analysis of long DNA-text as a set of many parallel texts, each of which is written in its own language (but alphabets of all such languages are interrelated), reveals important algebraic patterns in the genetic informatics of higher and lower organisms.

The previously formulated rules concerned total amounts of certain classes of *n*-plets. In this new preprint, the author focuses on searching possible rules of probabilities (or percentages) of *n*-plets in genomes in line with the mentioned supposition of Jordan about the existence of such rules.

This research uses a well-known fact of binary-oppositional features of DNA nucleotides (adenine A, thymine T, cytosine C, and guanine G), which allows the constructing a family of square tables for DNA alphabets of 4 nucleotides, 16 doublets, 64 triplets, …, $4^n$ *n*-plets. Each of *n*-plets occupies its strong individual place in this family of tables, which form a tensor family of square matrices. By placing the percentage of each *n*-plet (calculated inside the $n^{th}$ layer of given DNA-text) into the cell occupied by this *n*-plet in the appropriate square matrix, we obtain the numerical matrices of probabilities of all *n*-plets inside the $n^{th}$ layer of given DNA-text. Analysis of this family of probability matrices for *n*-plets reveals hidden regularities in the structural organization of the studied genomic DNA-texts. Below these regularities for cases *n* = 1, 2, 3, 4 are described and discussed.

## 2. The matrix representation of the DNA alphabets based on binary-oppositional traits of nucleotides.

As it is known, the DNA alphabet of 4 nucleotides A, T, C, and G is endowed with a system of binary-opposition traits or indicators [Fimmel, Petoukhov, 2020; Petoukhov, 2008; Petoukhov, He, 2010; Stambuk, 1999]:

   1) two of these molecules are purines with two rings (A and G), and the other two are pyrimidines with one ring (C and T). In terms of these oppositional indicators, C = T = 1,  A = G = 0 ;

   2) the two letters are keto molecules (T and G), and the other two - amino molecules (C and A). In terms of these oppositional indicators, C = A = 1, T = G = 0.

   In the DNA alphabet of 4 nucleotides, each of the letters C, A, T, and G is uniquely determined by its mentioned binary indicators. With this in mind, it is convenient to present sets of 4 DNA nucleotides, their 16 doublets and 64 triplets in the form of square tables, the columns of which are numbered with binary indicators "pyrimidine or purine" (C = T = 1, A = G = 0), and the rows are numbered with binary indicators "amino or keto "(C = A = 1, T = G = 0). In such tables, all 4 nucleotides, 16 doublets, and 64 triplets of DNA automatically occupy their individual places in strict order (Fig. 2.1).

|   | 1 | 0 |
|---|---|---|
| 1 | C | A |
| 0 | T | G |

|    | 11 | 10 | 01 | 00 |
|----|----|----|----|----|
| 11 | CC | CA | AC | AA |
| 10 | CT | CG | AT | AG |
| 01 | TC | TA | GC | GA |
| 00 | TT | TG | GT | GG |

|     | 111 | 110 | 101 | 100 | 011 | 010 | 001 | 000 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 111 | CCC | CCA | CAC | CAA | ACC | ACA | AAC | AAA |
| 110 | CCT | CCG | CAT | CAG | ACT | ACG | AAT | AAG |
| 101 | CTC | CTA | CGC | CGA | ATC | ATA | AGC | AGA |
| 100 | CTT | CTG | CGT | CGG | ATT | ATG | AGT | AGG |
| 011 | TCC | TCA | TAC | TAA | GCC | GCA | GAC | GAA |
| 010 | TCT | TCG | TAT | TAG | GCT | GCG | GAT | GAG |
| 001 | TTC | TTA | TGC | TGA | GTC | GTA | GGC | GGA |
| 000 | TTT | TTG | TGT | TGG | GTT | GTG | GGT | GGG |

|      | 1111 | 1110 | 1101 | 1100 | 1011 | 1010 | 1001 | 1000 | 0111 | 0110 | 0101 | 0100 | 0011 | 0010 | 0001 | 0000 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1111 | CCCC | CCCA | CCAC | CCAA | CACC | CACA | CAAC | CAAA | ACCC | ACCA | ACAC | ACAA | AACC | AACA | AAAC | AAAA |
| 1110 | CCCT | CCCG | CCAT | CCAG | CACT | CACG | CAAT | CAAG | ACCT | ACCG | ACAT | ACAG | AACT | AACG | AAAT | AAAG |
| 1101 | CCTC | CCTA | CCGC | CCGA | CATC | CATA | CAGC | CAGA | ACTC | ACTA | ACGC | ACGA | AATC | AATA | AAGC | AAGA |
| 1100 | CCTT | CCTG | CCGT | CCGG | CATT | CATG | CAGT | CAGG | ACTT | ACTG | ACGT | ACGG | AATT | AATG | AAGT | AAGG |
| 1011 | CTCC | CTCA | CTAC | CTAA | CGCC | CGCA | CGAC | CGAA | ATCC | ATCA | ATAC | ATAA | AGCC | AGCA | AGAC | AGAA |
| 1010 | CTCT | CTCG | CTAT | CTAG | CGCT | CGCG | CGAT | CGAG | ATCT | ATCG | ATAT | ATAG | AGCT | AGCG | AGAT | AGAG |
| 1001 | CTTC | CTTA | CTGC | CTGA | CGTC | CGTA | CGGC | CGGA | ATTC | ATTA | ATGC | ATGA | AGTC | AGTA | AGGC | AGGA |
| 1000 | CTTT | CTTG | CTGT | CTGG | CGTT | CGTG | CGGT | CGGG | ATTT | ATTG | ATGT | ATGG | AGTT | AGTG | AGGT | AGGG |
| 0111 | TCCC | TCCA | TCAC | TCAA | TACC | TACA | TAAC | TAAA | GCCC | GCCA | GCAC | GCAA | GACC | GACA | GAAC | GAAA |
| 0110 | TCCT | TCCG | TCAT | TCAG | TACT | TACG | TAAT | TAAG | GCCT | GCCG | GCAT | GCAG | GACT | GACG | GAAT | GAAG |
| 0101 | TCTC | TCTA | TCGC | TCGA | TATC | TATA | TAGC | TAGA | GCTC | GCTA | GCGC | GCGA | GATC | GATA | GAGC | GAGA |
| 0100 | TCTT | TCTG | TCGT | TCGG | TATT | TATG | TAGT | TAGG | GCTT | GCTG | GCGT | GCGG | GATT | GATG | GAGT | GAGG |
| 0011 | TTCC | TTCA | TTAC | TTAA | TGCC | TGCA | TGAC | TGAA | GTCC | GTCA | GTAC | GTAA | GGCC | GGCA | GGAC | GGAA |
| 0010 | TTCT | TTCG | TTAT | TTAG | TGCT | TGCG | TGAT | TGAG | GTCT | GTCG | GTAT | GTAG | GGCT | GGCG | GGAT | GGAG |
| 0001 | TTTC | TTTA | TTGC | TTGA | TGTC | TGTA | TGGC | TGGA | GTTC | GTTA | GTGC | GTGA | GGTC | GGTA | GGGC | GGGA |
| 0000 | TTTT | TTTG | TTGT | TTGG | TGTT | TGTG | TGGT | TGGG | GTTT | GTTG | GTGT | GTGG | GGTT | GGTG | GGGT | GGGG |

**Fig. 2.1**. The square tables of the DNA-alphabets of 4 nucleotides, 16 doublets, 64 triplets, and 256 tetraplets, which are constructed by the method of the binary numbering of their rows and columns and which are members of a tensor family of matrices $[C, A; T, G]^{(n)}$ under n = 1, 2, 3, 4 (see explanations in the text).

These four tables are not simple tables but they form a single tensor family of matrices: the second, the third, and the fourth tensor powers of the (2*2)-matrix $[C, A; T, G]$ automatically give this (4*4)-matrix of 16 doublets, this (8*8)-matrix of 64 triplets, and this (16*16)-matrix of 256 tetraplets (Fig. 1). Using the same method of the binary numbering of rows and columns of square matrices of DNA alphabets of *n*-plets, one can similarly construct square tables of 1024 pentaplets, and so on. These new tables will also be members of the unified tensor family of symbolic matrices $[C, A; T, G]^{(n)}$ for values n = 5, 6, ...

The tensor family of matrices $[C, A; T, G]^{(n)}$ was first used by the author for a comparative analysis of the percentage of different *n*-plets in the DNA-texts of various genomes. Let us explain our analytical approach using a specific example of the DNA of the first human chromosome, which contains a sequence of about 250 million nucleotides C, A, T, and G (initial data on this chromosome were taken in the GenBank: https://www.ncbi.nlm.nih.gov/nuccore/NC_000001.11). One can remind here that genomic sequences in the GenBank sites usually contain some letters N, indicating that there can be any nucleotide in this place (https://www.ncbi.nlm.nih.gov/books/NBK21136/). For this reason, the total amount of all nucleotides A, T, C, G, which are calculated for the sequence from the GenBank, is slightly less than the complete length of the DNA sequence, which is indicated in the GenBank. But practically this is not essential for the resulting values of percentages of separate nucleotides in the analyzed genomic sequences.

At the first step of the author's approach, percents of each of the nucleotides C, A, T, and G in this chromosome are calculated: %C ≈ 0.2085 , %G ≈ 0.2089, %A ≈ 0.2910, %T ≈ 0.2917 (here percents are shown in fractions of one, and their values are rounded to the fourth decimal place). These percent values are used to be indicated in appropriate cells of the matrix of nucleotides [C, A; T, G] instead of nucleotide symbols for receiving a numeric matrix of nucleotides percents [Fig. 2]. Here and below, percentages are rounded to the fourth decimal place.

| C | A | | %C | %A | | 0.2085 | 0.2910 |
|---|---|---|----|----|---|--------|--------|
| T | G | ➔ | %T | %G | = | 0.2918 | 0.2087 |

**Fig. 2.2**. The transformation of the symbolic matrix of 4 nucleotides into numeric matrix of nucleotides percents in the case of human chromosome №1.

One can note that %C ≈ %G and %A ≈ %T by the second Chargaff's rule [Albrecht-Buehler, 2006; Chargaff, 1971; Prahbu, 1993].

At the second step of the described approach, the DNA-text of the analyzed chromosome is represented as a text of doublets (for example, the text TAACCCTA… is represented as TA-AC-CC-TA-…) and percents of each of 16 doublets are calculated. Then these percents are indicated in appropriate cells of the (4*4)-matrix [C, A; T, G]$^{(2)}$ shown in Fig. 2.1. Fig. 2.3 presents the resulting matrix of percent of 16 doublets.

| %CC | %CA | %AC | %AA | | 0.05409 | 0.07274 | 0.05033 | 0.09504 |
|-----|-----|-----|-----|---|---------|---------|---------|---------|
| %CT | %CG | %AT | %AG | | 0.07134 | 0.01031 | 0.07429 | 0.07137 |
| %TC | %TA | %GC | %GA | = | 0.06008 | 0.06312 | 0.04402 | 0.06008 |
| %TT | %TG | %GT | %GG | | 0.09568 | 0.07286 | 0.05046 | 0.05419 |

**Fig. 2.3**. The matrix of percents of the 16 doublets in the DNA-sequence of doublets in the human chromosome №1.

At the third step of the described approach, the DNA-text of the analyzed chromosome is represented as a text of triplets (for example, the text TAACCCTAG… is represented as TAA-CCC-TAG-…) and percents of each of 64 triplets are calculated. Then these percents are indicated in appropriate cells of the (8*8)-matrix [C, A; T, G]$^{(3)}$ shown in Fig. 2.1. The resulting matrix of percents of 64 triplets is presented in Fig. 2.4.

| %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 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.0138 | 0.0188 | 0.0152 | 0.0186 | 0.0118 | 0.0198 | 0.0145 | 0.0369 |
| 0.0185 | 0.0029 | 0.0179 | 0.0210 | 0.0162 | 0.0025 | 0.0238 | 0.0199 |
| 0.0176 | 0.0127 | 0.0025 | 0.0023 | 0.0132 | 0.0194 | 0.0144 | 0.0224 |
| 0.0201 | 0.0209 | 0.0026 | 0.0029 | 0.0239 | 0.0178 | 0.0161 | 0.0185 |
| 0.0159 | 0.0196 | 0.0110 | 0.0199 | 0.0125 | 0.0146 | 0.0096 | 0.0196 |
| 0.0223 | 0.0023 | 0.0194 | 0.0128 | 0.0144 | 0.0025 | 0.0133 | 0.0176 |
| 0.0197 | 0.0198 | 0.0146 | 0.0195 | 0.0096 | 0.0112 | 0.0126 | 0.0160 |
| 0.0372 | 0.0188 | 0.0199 | 0.0190 | 0.0145 | 0.0153 | 0.0119 | 0.0138 |

**Fig. 2.4**. The matrix of percents of the 64 triplets in the DNA-sequence of triplets in the human chromosome №1.

At the fourth step of the described approach, the DNA-text of the analyzed chromosome is represented as a text of tetraplets (such as TAAC-CCTA-…) and percents of each of 256 tetraplets are calculated. Then these percents are indicated in appropriate cells of the (16*16)-matrix $[C, A; T, G]^{(4)}$ shown in Fig. 2.1. The resulting matrix of percent of 256 tetraplets is presented in Fig. 2.5.

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| .0033 | .0055 | .0042 | .0044 | .0040 | .0056 | .0032 | .0070 | .0030 | .0042 | .0040 | .0053 | .0032 | .0059 | .0055 | .0149 |
| .0041 | .0010 | .0044 | .0058 | .0047 | .0010 | .0040 | .0044 | .0041 | .0005 | .0051 | .0054 | .0047 | .0006 | .0095 | .0071 |
| .0050 | .0029 | .0008 | .0006 | .0036 | .0039 | .0049 | .0059 | .0037 | .0032 | .0006 | .0006 | .0040 | .0070 | .0037 | .0066 |
| .0048 | .0058 | .0006 | .0009 | .0057 | .0047 | .0045 | .0058 | .0049 | .0044 | .0007 | .0007 | .0071 | .0057 | .0049 | .0047 |
| .0052 | .0057 | .0027 | .0038 | .0010 | .0006 | .0003 | .0006 | .0033 | .0045 | .0032 | .0063 | .0042 | .0049 | .0039 | .0078 |
| .0058 | .0009 | .0035 | .0028 | .0007 | .0003 | .0005 | .0008 | .0049 | .0005 | .0064 | .0035 | .0046 | .0007 | .0049 | .0059 |
| .0048 | .0036 | .0046 | .0051 | .0005 | .0004 | .0008 | .0006 | .0047 | .0056 | .0033 | .0050 | .0031 | .0037 | .0047 | .0057 |
| .0073 | .0044 | .0053 | .0058 | .0006 | .0010 | .0005 | .0010 | .0096 | .0040 | .0051 | .0044 | .0046 | .0047 | .0040 | .0041 |
| .0046 | .0049 | .0042 | .0051 | .0024 | .0045 | .0028 | .0078 | .0030 | .0041 | .0029 | .0038 | .0023 | .0037 | .0029 | .0073 |
| .0057 | .0006 | .0051 | .0052 | .0037 | .0004 | .0056 | .0036 | .0047 | .0008 | .0033 | .0046 | .0030 | .0005 | .0046 | .0048 |
| .0057 | .0040 | .0005 | .0006 | .0030 | .0055 | .0026 | .0041 | .0032 | .0027 | .0006 | .0005 | .0025 | .0031 | .0032 | .0057 |
| .0068 | .0058 | .0006 | .0006 | .0070 | .0039 | .0032 | .0029 | .0037 | .0048 | .0006 | .0008 | .0041 | .0036 | .0036 | .0051 |
| .0051 | .0063 | .0034 | .0063 | .0041 | .0049 | .0031 | .0061 | .0023 | .0031 | .0017 | .0035 | .0031 | .0042 | .0023 | .0052 |
| .0077 | .0006 | .0063 | .0038 | .0049 | .0006 | .0045 | .0057 | .0038 | .0004 | .0032 | .0027 | .0042 | .0010 | .0034 | .0052 |
| .0073 | .0078 | .0038 | .0051 | .0037 | .0046 | .0042 | .0051 | .0030 | .0028 | .0029 | .0042 | .0023 | .0024 | .0030 | .0046 |
| .0150 | .0072 | .0055 | .0045 | .0059 | .0057 | .0043 | .0054 | .0055 | .0032 | .0040 | .0042 | .0032 | .0039 | .0030 | .0033 |

**Fig. 2.5**. The matrix of percents of the 256 tetraplets in the DNA-sequence of tetraplets in the human chromosome №1.

### 3. Positional tetra-groupings rule of percentage composition of *n*-plets.

At first glance, the set of percent in the resulting matrices (Figs. 2.3-2.5) is quite chaotic. It has the following features regarding the percent of separate *n*-plets:

- Percent of presented *n*-plets significantly depends on the order of letters in them. For example, the percent of doublets CG and GC, having the same letter composition, differ several times: %CG = 0.0103, and %GC = 0.0440. Similarly, the percent of triplets of the same letter composition CAT, CTA, ACT, ATC, TCA, TAC are significantly different: %CAT=0.0179, %CTA=0.0127, %ACT=0.0162, %ATC=0.0132, %TCA=0.0196, %TAC=0.0110, and so on;
- Accordingly, the numerical percent matrices for doublets, triplets, and tetraplets (Figs. 2.3-2.5) are not tensor powers of the nucleotide percent (2*2)-matrix (Fig. 2.2). The percentages of doublets, triplets and tetraplets in the

chromosome are not equal at all to the product of the percentages of separate nucleotides in their composition: %C = 0.2085, %G = 0.2087, %A = 0.2910, %T = 0.2917.

But unexpectedly these values %C=0.2085, %G=0.2087, %A=0.2910, %T=0.2917 showed themselves in the block organization of percentages of different *n*-plets in various layers of the genomic DNA-text as one can calculate from data of percent matrices in Figs. 2.3-2.5:

- The total sum Σ%CN of percentages of all 4 doublets CN (hereinafter, the symbol N denotes any of the nucleotides A, T, C, and G), which start with the nucleotide C, is equal to %C, that is, Σ%CN ≈ %CC + %CA +%CT + %CG ≈ 0.0541+0.0727+0.0713+0.0103 ≈ 0.2085 ≈ %C;
- The total sum Σ%NC of percentages of all 4 doublets NC, which have the nucleotide C at their second positions, is practically equal to %C, that is, Σ%NC ≈ %CC +%AC + %TC +%GC = 0.0541+0.0503+0.0601+0.0440 ≈ 0.2085 ≈ %C as well;
- The total sum Σ%CNN of percentages of all 16 triplets CNN, which have the nucleotide C at their first position, is practically equal to %C, that is, Σ%CNN ≈ 0.0284 ≈ %C as well;
- The total sum Σ%NCN of percentages of all 16 triplets NCN, which have the nucleotide C at their second position, is practically equal to %C, that is, Σ%NCN ≈ 0.0285 ≈ %C as well;
- The total sum Σ%NNC of percentages of all 16 triplets NNC, which have the nucleotide C at their third position, is practically equal to %C, that is, Σ%NNC ≈ 0.0285 ≈ %C as well;
- The total sum Σ%CNNN of percentages of all 64 tetraplets CNNN, which have the nucleotide C at their first position, is practically equal to %C, that is, Σ%CNNN ≈ 0.0285 ≈ %C as well;
- The total sum Σ%NCNN of percentages of all 64 tetraplets NCNN, which have the nucleotide C at their second position, is practically equal to %C, that is, Σ%NCNN ≈ 0.0285 ≈ %C as well;
- The total sum Σ%NNCN of percentages of all 64 tetraplets NNCN, which have the nucleotide C at their third position, is practically equal to %C, that is, Σ%NNCN ≈ 0.0285 ≈ %C as well;
- The total sum Σ%NNNC of percentages of all 64 tetraplets NNNC, which have the nucleotide C at their fourth position, is practically equal to %C, that is, Σ%NNNC ≈ 0.0285 ≈ %C as well.

Similar equalities turn out to be valid also for the total sums Σ of the considered *n*-plets with nucleotides A, T, G at the analogical positions, as shown in Fig. 3.1.

| %C ≈ 0.2085 | %G ≈ 0.2087 | %A ≈ 0.2910 | %T ≈ 0.2918 |
|---|---|---|---|
| Σ%CN ≈ 0.2085 | Σ%GN ≈ 0.2088 | Σ%AN ≈ 0.2910 | Σ%TN ≈ 0.2917 |
| Σ%NC ≈ 0.2085 | Σ%NG ≈ 0.2087 | Σ%NA ≈ 0.2910 | Σ%NT ≈ 0.2918 |
| Σ%CNN ≈ 0.2084 | Σ%GNN ≈ 0.2088 | Σ%ANN ≈ 0.2910 | Σ%TNN ≈ 0.2917 |
| Σ%NCN ≈ 0.2085 | Σ%NGN ≈ 0.2088 | Σ%NAN ≈ 0.2910 | Σ%NTN ≈ 0.2917 |
| Σ%NNC ≈ 0.2085 | Σ%NNG ≈ 0.2087 | Σ%NNA ≈ 0.2910 | Σ%NNT ≈ 0.2918 |
| Σ%CNNN ≈ 0.2085 | Σ%GNNN ≈ 0.2088 | Σ%ANNN ≈ 0.2910 | Σ%TNNN ≈ 0.2917 |
| Σ%NCNN ≈ 0.2085 | Σ%NGNN ≈ 0.2087 | Σ%NANN ≈ 0.2910 | Σ%NTNN ≈ 0.2918 |
| Σ%NNCN ≈ 0.2085 | Σ%NNGN ≈ 0.2088 | Σ%NNAN ≈ 0.2910 | Σ%NNTN ≈ 0.2918 |
| Σ%NNNC ≈ 0.2085 | Σ%NNNG ≈ 0.2087 | Σ%NNNA ≈ 0.2910 | Σ%NNNT ≈ 0.2918 |

**Fig. 3.1**. Percentages of nucleotides C, G, A, T, and the sum Σ of percent of *n*-plets with these nucleotides at certain positions for the case of DNA of the human chromosome №1. The symbol N denotes any of the nucleotides.

Briefly speaking, the following relations (3.1) hold true - with a high level of accuracy - regarding percentages of the nucleotides C, G, A, and T, and the considered *n*-plets in the human chromosome №1:

%C ≈ Σ%CN ≈ Σ%NC ≈ Σ%CNN ≈ Σ%NCN ≈ Σ%NNC ≈
Σ%CNNN ≈ Σ%NCNN ≈ Σ%NNCN ≈ Σ%NNNC;
%G ≈ Σ%GN ≈ Σ%NG ≈ Σ%GNN ≈ Σ%NGN ≈ Σ%NNG ≈
Σ%GNNN ≈ Σ%NGNN ≈ Σ%NNGN ≈ Σ%NNNG;
%A ≈ Σ%AN ≈ Σ%NA ≈ Σ%ANN ≈ Σ%NAN ≈ Σ%NNA ≈
Σ%ANNN ≈ Σ%NANN ≈ Σ%NNAN ≈ Σ%NNNA;
%T ≈ Σ%TN ≈ Σ%NT ≈ Σ%TNN ≈ Σ%NTN ≈ Σ%NNT ≈
Σ%TNNN ≈ Σ%NTNN ≈ Σ%NNTN ≈ Σ%NNNT          (3.1)

These equalities (3.1) can also be written in the form (3.2) of the equality of 4-dimensional vectors of genomic percentages, the coordinates of which are the percentages of individual nucleotides or the percent sums of *n*-plets with nucleotides at certain positions:

[%A, %T, %C, %G] ≈ [Σ%AN, Σ%TN, Σ%CN, Σ%GN] ≈
[Σ%NA, Σ%NT, Σ%NC, Σ%NG] ≈
[Σ%ANN, Σ%TNN, Σ%CNN, Σ%GNN] ≈
[Σ%NAN, Σ%NTN, Σ%NCN, Σ%NGN] ≈
[Σ%NNA, Σ%NNT, Σ%NNC, Σ%NNG] ≈
[Σ%ANNN, Σ%TNNN, Σ%CNNN, Σ%GNNN] ≈
[Σ%NANN, Σ%NTNN, Σ%NCNN, Σ%NGNN] ≈
[Σ%NNAN, Σ%NNTN, Σ%NNCN, Σ%NNGN] ≈
[Σ%NNNA, Σ%NNNT, Σ%NNNC, Σ%NNNG].          (3.2)

Knowing the percentages of nucleotides %A, %T, %C, and %G, it is possible to predict with high accuracy the sums of percentages of *n*-plets of the noted groupings. The ability of such predictions based on equalities (3.1) or (3.2) holds not only for the considered human chromosome №1 but also for many eukaryotic and prokaryotic genomes, which were analyzed by the author till now. One should note that percentages of nucleotides %A, %T, %C, and %G can be essentially different in various genomes. (Appendix I contains one of many possible examples of percent matrices related to the genome of bacteria *Bradyrhizobium japonicum* where %A ≈ 0.1819, %T ≈ 0.1815, %C ≈ 0.3184, and %G ≈ 0.3182 in contrast to the considered case of the human chromosome #1). This indicates a universal cooperative organization of percentages compositions of *n*-plets in genomic DNA-texts, which is reflected in very special block-mosaic structures of the percent matrices of *n*-plets (Fig. 2.3-2.5).

The four columns in Fig. 3.1 show that in each of the presented layers of the genomic DNA-text, there exist corresponding tetra-groupings of *n*-plets with the same

percentage sums. The following **rule of probabilities** can be formulated based on such results:

- All $n$ layers of a long DNA-text, each of which consists of a sequence of $4^n$ $n$-plets, has approximately the same sum of percentages of all those $n$-plets, which contain the considered nucleotide (C, G, A, or T) at a fixed $m$th position ($m \leq n$); here $n = 1, 2, 3,..$ but is not too large compared to the length of DNA.

By this rule, the percentage sums of $n$-plets in such groupings in different layers of a considered long DNA-text are equal to the percentage of a corresponding nucleotide in the first layer of the DNA-text, although the percent values of individual $n$-plets, that are summands in these sums, can differ significantly.

For example, in the second layer of DNA-text of human chromosome № 1, the sum of the percentages of all 4 doublets with nucleotide C in the first position ($m = 1$) is equal to $\Sigma\%CN \approx \%CC+\%CA+\%CT+\%CG \approx 0.0541 + 0.0727 + 0.0713 + 0.0103 \approx 0.2085$. In the same second layer, the sum of the percentages of all 4 doublets with a nucleotide C in the second position ($m = 2$) is equal to the same number, although the summands in this sum are significantly different: $\Sigma\%NC \approx \%CC+\%AC+\%TC+\%GC = 0.0541 + 0.0503 + 0.0601 + 0.0440 \approx 0.2085$. These two equal total values are equal to the percentage of nucleotide C in the first layer of the given genomic DNA-text: $\%C \approx 0.0285$.

This rule of probabilities can be also formulated in a form of an exclusion principle as a supposed universal principle of exclusion in genetics:

- Sums of percentages of all those $n$-plets, which contain the considered nucleotide (C, G, A, or T) at a fixed $m$th position ($m \leq n$), cannot be essentially different from each other in various n layers of any genomic DNA-text (here $n = 1, 2, 3, ...$ but is not too large compared to the length of the DNA-text).

This rule of the constancy of the sums of percentages in the different layers of long DNA-texts (in the indicated tetra-groupings of $n$-plets), which is relatively independent of the percent values of separate $n$-plets, the author calls as the **Gestalt genetic rule** of percentages of $n$-plets for their C-, G-, A-, and T-groupings in multilayer DNA-texts of genomes. The names "Gestalt genetic phenomena" and "Gestalt genetic rules", introduced by the author, are due to a similarity of these genetic phenomena with the phenomena of Gestalt psychology (this similarity is described below in more detail). In Gestalt psychology literature, the term "Gestalt" is traditionally capitalized; due to this tradition, we also capitalize this term in the name "Gestalt genetics".

Each of the tetra-groupings of $n$-plets, which is defined by a disposition of attributive nucleotides C, G, A, and T at a certain position $m$ ($m \leq n$) in these $n$-plets, we call as an *m-positional tetra-grouping*. For example, a tetra-grouping corresponding to sets ANN, TNN, CNN, and GNN is called as a 1-positional tetra-grouping; a tetra-grouping corresponding to NAN, NTN, NCN, NGN is called as a 2-positional tetra-grouping, and so on. The described phenomena, reflected in the Gestalt genetic rule, have analogies with the phenomena of holography, in which it is possible to reconstruct the image of a whole object from the image of its piece. Indeed, the knowledge of the sums of the percentages of $n$-plets in $m$-positional tetra-groupings of one layer of a long DNA-text gives knowledge about the sums of the percentages of $n$-plets in $m$-positional tetra-groupings of other layers.

Returning for a moment to the tensor family of matrices $[C, A; T, G]^{(n)}$ (Fig. 2.1), let us consider conditional (or model) percentages of $n$-plets in a tensor family of

percentage matrices $[\%C, \%A; \%T, \%G]^{(n)} \approx [0.2085, 0.2910; 0.2917, 0.2089]^{(n)}$. Figs. 3.2 and 3.3 show the received matrices for $(n) = 2, 3$, whose percent entries are significantly differ from real percentages of *n*-plets shown above in Figs. 2.3-2.4.

   But unexpectedly the Gestalt rule holds for these model percentages of *n*-plets by the expression (2) as well:

$[\%A, \%T, \%C, \%G] \approx [\Sigma\%AN, \Sigma\%TN, \Sigma\%CN, \Sigma\%GN] \approx$
$[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG] \approx$
$[\Sigma\%ANN, \Sigma\%TNN, \Sigma\%CNN, \Sigma\%GNN] \approx$
$[\Sigma\%NAN, \Sigma\%NTN, \Sigma\%NCN, \Sigma\%NGN] \approx$
$[\Sigma\%NNA, \Sigma\%NNT, \Sigma\%NNC, \Sigma\%NNG] \approx [0.2910, 0.2918, 0.2085, 0.2087]$.

| %CC | %CA | %AC | %AA | | 0.0435 | 0.0607 | 0.0607 | 0.0847 |
|-----|-----|-----|-----|---|--------|--------|--------|--------|
| %CT | %CG | %AT | %AG | | 0.0608 | 0.0435 | 0.0849 | 0.0607 |
| %TC | %TA | %GC | %GA | = | 0.0608 | 0.0849 | 0.0435 | 0.0607 |
| %TT | %TG | %GT | %GG | | 0.0851 | 0.0609 | 0.0609 | 0.0436 |

**Fig. 3.2**. The matrix of conditional or model percents for 16 doublets, which is the second tensor power of the matrix [0.2085, 0.2910; 0.2917, 0.2089] (all values are rounded to the fourth decimal place).

| %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 |

=

| .00906382 | .01265035 | .01265035 | .01765605 | .01265035 | .01765605 | .01765605 | .02464250 |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| .01268315 | .00907448 | .01770183 | .01266522 | .01770183 | .01266522 | .02470639 | .01767681 |
| .01268315 | .01770183 | .00907448 | .01266522 | .01770183 | .02470639 | .01266522 | .01767681 |
| .01774773 | .01269806 | .01269806 | .00908514 | .02477045 | .01772264 | .01772264 | .01268011 |
| .01268315 | .01770183 | .01770183 | .02470639 | .00907448 | .01266522 | .01266522 | .01767681 |
| .01774773 | .01269806 | .02477045 | .01772264 | .01269806 | .00908514 | .01772264 | .01268011 |
| .01774773 | .02477045 | .01269806 | .01772264 | .01269806 | .01772264 | .00908514 | .01268011 |
| .02483467 | .01776859 | .01776859 | .01271298 | .01776859 | .01271298 | .01271298 | .00909582 |

**Fig. 3.3**. The matrix of conditional or model percents for 64 triplets, which is the third tensor power of the matrix [0.2085, 0.2910; 0.2917, 0.2089].

   It should be especially noted that in the case of a tensor family of percentage matrices $[\%C, \%A; \%T, \%G]^{(n)}$ we have the absolute accuracy of fulfillment of the Gestalt rule. For example, if $\%A = 0.291001313$, $\%T = 0.291755765$, $\%C = 0.208498924$, $\%G = 0.208743998$, then the expressions (2) is fulfilled precisely:

$[\%A, \%T, \%C, \%G] = [\Sigma\%AN, \Sigma\%TN, \Sigma\%CN, \Sigma\%GN] =$
$[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG] =$
$[\Sigma\%ANN, \Sigma\%TNN, \Sigma\%CNN, \Sigma\%GNN] =$
$[\Sigma\%NAN, \Sigma\%NTN, \Sigma\%NCN, \Sigma\%NGN] =$
$[\Sigma\%NNA, \Sigma\%NNT, \Sigma\%NNC, \Sigma\%NNG] =$
$[\Sigma\%ANNN, \Sigma\%TNNN, \Sigma\%CNNN, \Sigma\%GNNN] =$
$[\Sigma\%NANN, \Sigma\%NTNN, \Sigma\%NCNN, \Sigma\%NGNN] =$
$[\Sigma\%NNAN, \Sigma\%NNTN, \Sigma\%NNCN, \Sigma\%NNGN] =$
$[\Sigma\%NNNA, \Sigma\%NNNT, \Sigma\%NNNC, \Sigma\%NNNG] =$
$[0.291001313, 0.291755765, 0.208498924, 0.208743998].$          (3.3)

It gives pieces of evidence that fundamental genetic phenomena, reflected in the formulated Gestalt rule, are connected with the algebraic operation of the tensor product, which is so important in quantum mechanics and quantum informatics. Accordingly, the tensor family of percentage matrices $[\%C, \%A; \%T, \%G]^{(n)}$ and their percentage entries can be considered in each specific case as a certain standard of comparison in the analysis of long DNA-texts. Below we will show that the difference between the tetra-groupings of real percentages and these reference percentages is related to unitary operators. The percentage values of *n*-plets in the tensor family of matrices $[\%C, \%A; \%T, \%G]^{(n)}$ we will conditionally call as reference percentages of *n*-plets in the analyzed long DNA-text.

## 4. DNA epi-chains and the Gestalt rule of percentages of *n*-plets.

This Section presents some results of the analogical study of percentages of *n*-plets in special subsequences of long nucleotide sequences in single-stranded DNA. These subsequences are termed «DNA epi-chains» [Petoukhov, 2019, 2020a-c]. The author's initial results testify that the above-described equalities (3.1) and (3.2) of total sums of percentages of *n*-plets hold for these epi-chains as well.

By definition, in a nucleotide sequence $N_1$ of any DNA strand $N_1$ (Fig. 4.1a) with sequentially numbered nucleotides 1, 2, 3, 4, ... (Fig. 4.1a), epi-chains of different orders $k$ are such subsequences that contain only those nucleotides, whose numeration differ from each other by natural number $k = 1, 2, 3, 4, …$ For example, in any single-stranded DNA, one can consider its epi-chain of the second-order $N_2$, in which its nucleotide sequence numbers differ by $k = 2$: an epi-chain $N_2$ contains nucleotides with numerations 1, 3, 5, … (Fig. 4.1b). By analogy, an epi-chain of the third-order $N_3$ is connected with $k = 3$ and contains a subsequence of nucleotides with numerations 1, 4, 7, 10, ... (Fig. 4.1d).
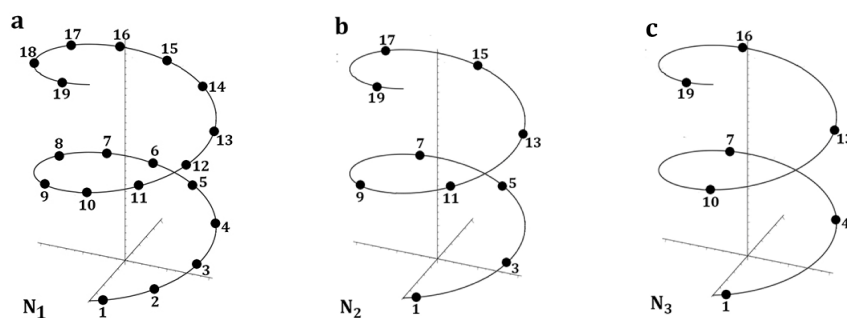
**Fig. 4.1**. Schematic representations of a single-stranded DNA and its initial epi-chains of numerated nucleotides, denoted by black circles. **a**, a sequence $N_1$ of numerated nucleotides of the DNA strand; **b**, an epi-chain of the second-order $N_2$ having nucleotides with numbers 1-3-5-7-…; **c**, an epi-chain of the third-order $N_3$ nucleotides numbers 1-4-7-10-… .

Each genomic DNA epi-chain of $k$th order (if $k$ = 2, 3, 4, ….) contains $k$ times fewer nucleotides than the DNA strand and has its own arrangements of nucleobases A, T, C, and G. Each of these epi-chains contain different percentages of corresponding $n$-plets. But unexpectedly the total sums of percentages of $n$-plets in C-, G-, A-, and T-groupings practically coincide for each of the epi-chains and for the complete genomic DNA-text (at this stage of the research, the author studied the percentages of $n$-plets in epi-chains only in cases of epi-chains with relatively small orders $k$).

To illustrate this result, Fig. 4.2 shows the percent matrices for epi-chains of the second, third and fourth orders ($k$ = 2, 3, 4) in the DNA of the human chromosome №1.

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

≈

| 0.05409 | 0.07274 | 0.05033 | 0.09504 |
|---|---|---|---|
| 0.07134 | 0.01031 | 0.07429 | 0.07137 |
| 0.06008 | 0.06312 | 0.04402 | 0.06008 |
| 0.09568 | 0.07286 | 0.05046 | 0.05419 |

| Epi-chain of the 2nd order | | | |
|---|---|---|---|
| 0.0492 | 0.0524 | 0.0538 | 0.0985 |
| 0.0591 | 0.0478 | 0.0800 | 0.0588 |
| 0.0612 | 0.0788 | 0.0443 | 0.0613 |
| 0.0988 | 0.0530 | 0.0539 | 0.0492 |

| Epi-chain of the 3rd order | | | |
|---|---|---|---|
| 0.0487 | 0.0562 | 0.0579 | 0.0912 |
| 0.0571 | 0.0464 | 0.0849 | 0.0570 |
| 0.0608 | 0.0826 | 0.0411 | 0.061 |
| 0.0918 | 0.0565 | 0.0578 | 0.0489 |

| Epi-chain of the 4th order | | | |
|---|---|---|---|
| 0.0487 | 0.0558 | 0.0574 | 0.0940 |
| 0.0566 | 0.0475 | 0.0831 | 0.0565 |
| 0.0579 | 0.0835 | 0.0445 | 0.0579 |
| 0.0943 | 0.0561 | 0.0577 | 0.0487 |

**Fig. 4.2.** Matrices of percentages of 16 doublets in the single-stranded DNA of the human chromosome №1. Upper row: the percentages matrices of the genomic DNA ($k$ = 1). Bottom row: percent matrices for the DNA epi-chains of the second, third, and fourth orders ($k$ = 2, 3, 4).

One can see from data in Fig. 4.2 that for considered epi-chains, different percentages matrices contain essential different percentages in corresponding separate cells. For example, %CG = 0.0103 in the complete single-stranded DNA; %CG = 0.0478 in the epi-chain of the 2nd order; %CG = 0.0464 in the epi-chain of the 3rd order; %CG = 0.0475 in the epi-chain of the 4th order. But in each of these four percentages matrices the equalities (3.2) are realized with a high level of accuracy:

$$[\%A, \%T, \%C, \%G] \approx [\Sigma\%AN, \Sigma\%TN, \Sigma\%CN, \Sigma\%GN] \approx$$
$$[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG] \approx [0.2910, 0.2918, 0.2085, 0.2087] \qquad (4.1)$$

It illustrates that in the considered DNA epi-chains, the sums of the percentages in the tetra-groupings of $n$-plets in each of the epi-chains practically do not depend on the percent of its separate $n$-plets and coincide with corresponding

sums in the genomic DNA-text. This can be considered as the separate Gestalt rule for genomic epi-chains (in addition to the above-presented gestalt rule of percentages of *n*-plets in long multilayer DNA-texts).

## 5. Regarding the development of Gestalt genetics based on analogies with Gestalt psychology.

The discovered and described genomic phenomena of the relative independence of the sums of the percentages of *n*-plets (in the indicated tetra-groupings) from the values of individual percentages in these sums resemble the phenomenon of perceiving a musical melody: a musical melody can be reproduced by different musical instruments and in different frequency ranges, that is, under significantly changing the frequency of the sound of each of its note elements, but despite these changes, the melody remains generally recognizable.

Many such phenomena of perception, in which there is relative independence of the integral form from its constituent individual components, are studied in Gestalt psychology. The described universal regularities in the preservation of total percentages in tetra-groupings of *n*-plets relatively regardless of the percentage of individual *n*-plets in genomic multilayer DNA-texts allow the author to develop Gestalt genetics. It seems natural to think that Gestalt genetics is interrelated with Gestalt psychology, which studies some genetically inherited properties of our brain regarding the perception of the environment.

The phenomena of perception of visual, auditory, and other images, studied by Gestalt psychology, reflect the fundamental inherited property of the psyche - to seek in a disparate whole. Thanks to the ability to think in Gestalts, you can understand the sentence, even if you change the order of the letters in each word and leave only the beginning and end in place. For example, you can easily understand the phrase, strongly "mutated" by local permutations: "Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are» (this "mutated" phrase example is taken from https://www.dictionary.com/e/typoglycemia/ ).

The analogies between Gestalt phenomena of perception and the genetic phenomena, which are reflected, in particular, in the above-formulated gestalt rule of genomes for the tetra-groupings, allow the author developing Gestalt genetics. This scientific direction studies holistic genetic patterns that are relatively independent of particular components. Gestalt genetics comes in contact with the teachings of the creator of analytical psychology C. Jung and his associate Nobel laureate in physics W. Pauli about the archetypes of the unconscious; in particular, they linked these archetypes with the Yin-Yang schemes of the ancient Chinese book "I-Ching" and its table of 64 hexagrams, which have deep structural analogies with the DNA alphabets (these analogies were noted by some reputable geneticists - G.S. Stent in 1969, and Nobel laureate F. Jacob in 1977; extended information on these analogies is in publications [Petoukhov, 1999, 2008; Petoukhov, He, 2010]).

In our opinion, the origins of the genetically inherited ability of the brain to work with Gestalt images should be sought in Gestalt genetics. In particular, Gestalt genetics is capable of providing new approaches to understanding the noise immunity of genetic information under mutations of DNA-texts.

In addition to the phenomena of Gestalt psychology, in living organisms, there are many genetically inherited physiological phenomena, in which the same whole pattern is realized in conditions of a wide variety of constituent elements and which can be attributed to Gestalt biology (this new name is proposed by the author as

uniting genetically inherited Gestalt-like phenomena of different types). For example, Gestalt biology includes some genetically inherited phenomena of morphogenesis (laws of phyllotaxis, spiralization of biological structures at various levels and branches of biological evolution), as well as some functional phenomena (homeostasis at different stages of ontogenesis; the processing of sensory information from different sense organs according to main psychophysical law of Weber-Fechner; the biomechanical phenomenon, known as Bernstein's problem, that the general target task of body movement is performed exactly regardless of the inaccuracies of its constituent motor subtasks). In particular, Gestalt genetics includes an observation of the embryology classic K. Baer that chick embryos are vastly different, while the resulting adult organisms are remarkably similar. The author thinks that various genetically inherited phenomena of Gestalt biology are based on Gestalt genetics.

One should also recall that the molecular composition of a living body is constantly changing while maintaining the shape of the body. Our body's proteins are involved in continuous life-death cycles of their assembling and disassembling into amino acids. For example, the half-life of the hormone insulin is 6-9 minutes, etc. In other words, genetically inherited parts of our body are constantly dying and reborn. Taking into account such phenomena, the renowned physiologist A.G. Gurvich claimes: "*The main problem in biology is maintaining shape while constantly renewing the substrate*" [Gurvich, 1977]. In our opinion, the described phenomena of Gestalt genetics with its Gestalt rules of percentages in genomic multilayer DNA-texts are directly related to this fundamental problem of biology.

Gestalt genetics provides new approaches to understanding ontogeny. Using the terminology of Gestalt psychology and Gestalt therapy, the author suggests interpreting ontogenetic processes as the stepwise processes of "closing" certain genetic gestalts. One should note that all physiological systems are forced to bear the structural stamp of the genetic code since they should be genetically encoded for transmission to descendants and survival.

Gestalt genetics also concerns the phenomenon of the biomechanics of movements, described by the classic of biomechanics N.A. Bernstein: the general target task of the movement is performed exactly regardless of the inaccuracies of its constituent motor subtasks [Bernstein, 1967]. For example, when repeating an exact hit with a hammer on a nail, a person each time uses different trajectories, speeds, and accelerations of body parts with changes in both flexions in the joints and the activity of many muscles of each joint with many motoneurons of each muscle. This question of how the central nervous system is capable of adequately controlling the many degrees of freedom of the musculoskeletal system was first addressed by Bernstein and is now known usually as the «Bernstein problem». The degrees of freedom problem in motor control states that there are multiple ways for humans or animals to perform a movement to achieve the same goal using redundant neurophysiological degrees of freedom. How the nervous system "chooses" a subset of these near-infinite degrees of freedom is an overarching difficulty in understanding motor control and motor learning. In other words, under normal circumstances, no simple one-to-one correspondence exists between a motor problem (or task) and a motor solution to the problem.

Gestalt genetics provides new opportunities for understanding the known physiological phenomena of innate knowledge about the scheme of our body. For example, mature people with congenital absence of limbs feel a phantom of limb's existence with a feeling of pain in them, etc., although they have no lifetime experience of using them [Vetter, Weinstein, 1967; Weinstein, Sersen, 1961].

Gestalt psychology comes into contact with the well-known phenomena of perception constancy, reflecting the ability of the brain to stably recognize the shapes of objects and other structures of the external world under conditions of a significant change in the conditions of their presentation to the sense organs. For example, a change in the color illumination of a room does not prevent a person from recognizing the shapes and colors of objects in it, although other light frequencies from objects come to the retina. Similar properties one can be waiting for information phenomena in Gestalt genetics, which preseeds phenomena of Gestalt psychology.

According to Mendel's law of independent inheritance of traits, information from the level of DNA molecules dictates the macrostructures of living bodies through many independent channels, despite strong noises and interferences. For example, hair, eye, and skin colors are inherited independently of each other. Accordingly, each organism is a machine of multichannel noise-immune coding. Gestalt genetics, with its genomic Gestalt rules, can help in understanding this multi-channel noise immunity.

## 6. Gestalt genetics and quantum informatics.

Many authors have long suggested the quantum-informational nature of living bodies. For example, R. Penrose, treating the body as a quantum computer, appeals to tubulin proteins that can switch from one state to another by analogy with triggers [Penrose, 1996]. One can show that the analysis of living bodies at a deeper - genomic - level leads to more convincing and fruitful pieces of evidence about organisms as quantum information entities (or as quantum-like entities whose modeling can use the formalisms of quantum informatics).

In quantum mechanics and quantum informatics, when analyzing the probabilities of events, the amplitudes of these probabilities, equal to the square root of their values, are traditionally considered. The genomic Gestalt rule for percentages of $n$-plets (in C-, G-, A-, and T-groupings) was revealed under analyzing the sums of the probabilities of $n$-plets from certain groupings, for example, all doublets starting with the letter C: $\Sigma\%CN = \%CC + \%CA + \%CT + \%CG$. Each of these probabilities $\%CC$, $\%CA$, $\%CT$, and $\%CG$ has its own "amplitude" in the form of its square root.

It can be noted that the sum of the percentages (that is, the probabilities) of $n$-plets in each of these four groupings can be interpreted as the square of the length of a vector whose components are equal to the square roots of the probabilities of the corresponding $n$-plets. For example, the sum $\%CC+\%CG+\%CA+\%CT$ is the square of the length of the 4-dimensional vector $V_{CN} = [\sqrt{\%CC}, \sqrt{\%CG}, \sqrt{\%CA}, \sqrt{\%CT}]$. Accordingly, the sum $\Sigma\%CNN$ is interpreted as the square of the length of an 8-dimensional vector $V_{CNN} = [\sqrt{\%CCC}, \sqrt{\%CCG}, \sqrt{\%CCA}, \sqrt{\%CCT}, \sqrt{\%CGC}, \sqrt{\%CGG}, \sqrt{\%CGA}, \sqrt{\%CGT}, \sqrt{\%CAC}, \sqrt{\%CAG}, \sqrt{\%CAA}, \sqrt{\%CAT}, \sqrt{\%CTC}, \sqrt{\%CTG}, \sqrt{\%CTA}, \sqrt{\%CTT}]$, and so on.

From this point of view, equalities (3.1) mean the constancy of the length of the corresponding $2^n$-dimensional vectors, whose coordinates are the amplitudes of the probabilities of the corresponding $n$-plets. This metric approach allows for developing new methods of comparative vector analysis in genetics, which are now being studied in our laboratory.

Let us compare all $m$-positional tetra-groupings - under different values $m$ - in the case of real percentages of $n$-plets (similar to those shown in Figs. 2.3 – 2.5) and in the case of reference percentages (similar to those shown in Figs. 3.2 and 3.3). In

all compared tetra-groupings, the corresponding $2^n$-dimensional vectors of amplitudes of probabilities have different coordinate values. But the lengths of these vectors in all compared cases are equal to each other due to the Gestalt rule, which says about the equality of the sums of the percentages of $n$-plets in each grouping to the same value, that is, to the percentage of the corresponding nucleotide in the analyzed genomic DNA-text. Accordingly, these vectors of equal length can be transformed into each other by unitary transformations that do not change the length of the vectors and are either rotations or mirror reflections.

Thus, algebraic Gestalt genetics turns out to be connected with unitary operators, which are key for quantum informatics: all calculations in quantum computers and quantum search algorithms are based on unitary operators as quantum gates. Moreover, any unitary matrix can serve as a quantum gate. In quantum mechanics, the evolution of a closed quantum system is described by unitary transformations. Since this article deals with many layers of genomic DNA texts, it can be assumed that, in particular, a whole set of quantum search algorithms work in multilayer DNA texts, each of which is individually oriented to a particular text layer. The articles [Patel, 2001a-c] suppose that the genetic code is related to the quantum Grover's algorithm.

Along the way, one can note that the entire genetically inherited kinematic scheme of movements of our body with its parts is built on rotations in the joints and mirror reflections, that is, on unitary transformations that have physiological significance. Turtles and crocodiles, when hatched from an egg, immediately crawl to the water with coordinated movements using innate algorithms with the same unitary transformations of rotations and mirror reflections.

The materials of this new article supplement the author's previously published works on the topic of quantum biology and formalisms of quantum informatics in biology [Petoukhov, 2008, 2020a,c, 2021; Petoukhov, He, 2010; Petoukhov, Petukhova, Svirin, 2019]. The revealed connection of genetics with quantum informatics opens up the possibility of introducing rich formalisms of quantum mechanics and quantum informatics into algebraic biology for the mutual enrichment of these sciences and the inclusion of biology in the field of developed mathematical natural science. There are about 100 trillion cells in the human body, forming a single system. The formalisms of quantum informatics and Gestalt genetics can help in understanding such coherence phenomena.

One should mention suppositions of many authors that formalisms of quantum informatics can be effectively used for deep understanding and modeling biological bodies (see, for example: [Igamberdiev, 1993; Matsuno, 1999; Matsuno, Paton, 2000; Abbott, Davies, Pati, 2008]).


## 7. Percentages matrices of *n*-plets and algebras of hyperbolic numbers.

This section shows that each of the vectors of percentages sums of $n$-plets in the equalities (2) corresponds to such mosaic matrix, which is connected with well-known algebras of $2^n$-dimensional hyperbolic numbers (hyperbolic numbers are termed also as double numbers, Lorentz numbers, split-complex numbers, and perplex numbers). Consequently, the arrangements of the described tetra-groupings of $n$-plets inside these $(2^n * 2^n)$-matrices of percentages obey special algebraic rules. It is additionally interesting since, as it is known, structures of some genetically inherited biological phenomena are related to $2^n$-dimensional hyperbolic numbers [Petoukhov,

2021]. 2-dimensional hyperbolic numbers form an algebra over the field of real numbers [Harkin, Harkin, 2004; Kantor, Solodovnikov, 1989].

To demonstrate these connections, we will take initially from equalities (2) the vector of percentage sums of 16 doublets [$\Sigma\%AN$, $\Sigma\%TN$, $\Sigma\%CN$, $\Sigma\%GN$]; its coordinates correspond to summary percentages of tetra-groupings of doublets, which start with one of the nucleotides C, G, A, or T. Let us denote matrix cells, which are occupied by 8 doublets CN and GN, by number +1, and matrix cells, which are occupied by 8 doublets AN and AT, by number -1. (These two groupings of doublets occupy appropriate quadrants along diagonals of the matrix as it it shown in Figs. 2.1 and 2.3). Fig. 7.1 gives a result of these denotations, which transform a symbolic matrix of doublets into numeric matrix $M_2$ with entries +1 and -1.

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

➔  $M_{21} =$

| +1 | +1 | -1 | -1 |
|----|----|----|----|
| +1 | +1 | -1 | -1 |
| -1 | -1 | +1 | +1 |
| -1 | -1 | +1 | +1 |

**Fig. 7.1**. Transformation of the symbolic (4*4)-matrix of 16 doublets into the numeric matrix $M_{21}$ which shows the arrangements of cells, occupied by doublets CN and GN (black cells) and also by AN and TN (white cells).

The received matrix can be decomposed into sum of four sparse matrices $e_0$, $e_1$, $e_2$, $e_3$: $M_{21} = e_0 + e_1 + e_2 + e_3$ (Fig. 7.2). The set of sparse matrices turns out to be closed under the multiplication operation: the product of any two matrices of this set by each other gives a matrix of the same set. This defines the multiplication table of these matrices (Fig. 7.2), which is known as the multiplication table of the bases units of the algebra of 4-dimensional hyperbolic numbers (or algebra of 4-dimensional hyperbolic matrions [Petoukhov, 2008; Petoukhov, He, 2010]). It should be noted that the decomposition of the matrix $M_{21}$ was done based on the known method of the dyadic-shift decompositions [Petoukhov, 2008; Petoukhov, He, 2010], which is used in this section below for other matrices as well.

$$M_{21} = e_0 + e_1 + e_2 + e_3 = \begin{vmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{vmatrix} + \begin{vmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{vmatrix} + \begin{vmatrix} 0&0&-1&0 \\ 0&0&0&-1 \\ -1&0&0&0 \\ 0&-1&0&0 \end{vmatrix} + \begin{vmatrix} 0&0&0&-1 \\ 0&0&-1&0 \\ 0&-1&0&0 \\ -1&0&0&0 \end{vmatrix}$$

| * | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
|-----|-----|-----|-----|-----|
| $e_0$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_1$ | $e_1$ | $e_0$ | $e_3$ | $e_2$ |
| $e_2$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ |
| $e_3$ | $e_3$ | $e_2$ | $e_1$ | $e_0$ |

**Fig. 7.2**. The decomposition of the matrix $M_{21} = e_0 + e_1 + e_2 + e_3$ from Fig. 7.1 and the multiplication table (at right) of these sparse matrices $e_0$, $e_1$, $e_2$, and $e_3$. The bold box in the multiplication table refers the subalgebra of 2-dimensional hyperbolic numbers.

By analogy, one can take from the equalities (2) another vector $[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG]$; its coordinates correspond to summary percentages of those tetra-groupings of 16 doublets, which have nucleotides C, G, A, and T at their second positions and occupy a new set of cells in the percentages matrix of doublets. Let us denote matrix cells, which are occupied by 8 doublets NC and NG, by number +1, and cells, which are occupied by 8 doublets NA and NT, by number -1. Fig. 7.3 shows a numeric matrix $M_{22}$, which appears in the result and which has a mosaic resembling a chessboard.

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

➔ $M_{22} =$

| +1 | -1 | +1 | -1 |
|----|----|----|----|
| -1 | +1 | -1 | +1 |
| +1 | -1 | +1 | -1 |
| -1 | +1 | -1 | +1 |

**Fig. 7.3**. Transformation of the symbolic matrix of doublets into the numeric matrix $M_{22}$, which shows the arrangements of cells, occupied by doublets NC and NG (black cells) and also by NA and NT (white cells).

The received matrix can be decomposed into sum of four new sparse matrices $e_0, e_1, e_2, e_3$: $M_{22} = e_0 + e_1 + e_2 + e_3$ (Fig. 7.4). The set of these sparse matrices turns out to be closed under the multiplication as well: the product of any two matrices of this set by each other gives a matrix of the same set. This defines the multiplication table of these matrices (Fig. 7.4), which coincides with the multiplication table shown in Fig. 7.2. It means that both matrices $M_{12}$ and $M_{22}$ are different matrix representations of 4-dimensional hyperbolic numbers with unit coordinates.

$$M_{22} = e_0 + e_1 + e_2 + e_3 = \begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{bmatrix} + \begin{bmatrix} 0&-1&0&0 \\ -1&0&0&0 \\ 0&0&0&-1 \\ 0&0&-1&0 \end{bmatrix} + \begin{bmatrix} 0&0&1&0 \\ 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \end{bmatrix} + \begin{bmatrix} 0&0&0&-1 \\ 0&0&-1&0 \\ 0&-1&0&0 \\ -1&0&0&0 \end{bmatrix}$$

|       | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
|-------|-------|-------|-------|-------|
| $e_0$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_1$ | $e_1$ | $e_0$ | $e_3$ | $e_2$ |
| $e_2$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ |
| $e_3$ | $e_3$ | $e_2$ | $e_1$ | $e_0$ |

**Fig. 7.4**. The decomposition of the matrix $M_{22} = e_0 + e_1 + e_2 + e_3$ from Fig. 7.3 and the multiplication table (at right) of these sparse matrices $e_0, e_1, e_2,$ and $e_3$. The bold box in the multiplication table refers the subalgebra of 2-dimensional hyperbolic numbers.

By analogy, one can take from the equalities (2) another vector $[\Sigma\%ANN, \Sigma\%TNN, \Sigma\%CNN, \Sigma\%GNN]$; its coordinates correspond to summary percentages of those tetra-groupings of 64 triplets, which start with nucleotides C, G, A, and T. Let us denote matrix cells, which are occupied by 32 triplets CNN and GNN, by number +1, and matrix cells, which are occupied by 32 triplets ANN and TNN, by number -1. These two groupings of triplets occupy appropriate quadrants of the matrix as it is

shown in Figs. 2.1 and 2.4. Fig. 7.5 shows a result of these denotations, which lead to a numeric matrix $M_{31}$ with entries +1 and -1.

| %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 |

➔

$$M_{31} = \begin{bmatrix} +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 \\ -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 \\ -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 \\ -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 \end{bmatrix}$$

**Fig. 7.5**. Transformation of the symbolic (8*8)-matrix of 64 triplets into the numeric matrix $M_{31}$, which shows the arrangements of cells, occupied by triplets CNN and GNN (black cells) and also by triplets ANN and TNN (white cells).

The received matrix $M_{31}$ can be decomposed by the method of the dyadic-shift decompositions into sum of 8 sparse matrices: $M_{31} = e_0 + e_1 + e_2 + e_3 + e_4 + e_5 + e_6 + e_7$ (Fig. 7.6). The set of these 8 sparse matrices turns out to be closed under the multiplication operation and defines the multiplication table of these matrices (Fig. 7.6), which is known as the multiplication table of the basis units of the algebra of 8-dimensional hyperbolic numbers (or 8-dimensional hyperbolic matrions [Petoukhov, 2008; Petoukhov, He, 2010]).

$M_{31} = e_0+e_1+$
$e_2+e_3+e_4+e_5+$
$e_6+e_7 =$

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

$$
\begin{matrix}
0\ 0\ 0\ 1\ 0\ 0\ 0\ 0 \\
0\ 0\ 1\ 0\ 0\ 0\ 0\ 0 \\
0\ 1\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
0\ 0\ 0\ 0\ 0\ 0\ 0\ 1 \\
0\ 0\ 0\ 0\ 0\ 0\ 1\ 0 \\
0\ 0\ 0\ 0\ 0\ 1\ 0\ 0 \\
0\ 0\ 0\ 0\ 1\ 0\ 0\ 0
\end{matrix}
\ +\
\begin{matrix}
0\ 0\ 0\ 0\ \text{-}1\ 0\ 0\ 0 \\
0\ 0\ 0\ 0\ 0\ \text{-}1\ 0\ 0 \\
0\ 0\ 0\ 0\ 0\ 0\ \text{-}1\ 0 \\
0\ 0\ 0\ 0\ 0\ 0\ 0\ \text{-}1 \\
\text{-}1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
0\ \text{-}1\ 0\ 0\ 0\ 0\ 0\ 0 \\
0\ 0\ \text{-}1\ 0\ 0\ 0\ 0\ 0 \\
0\ 0\ 0\ \text{-}1\ 0\ 0\ 0\ 0
\end{matrix}
\ +\
\begin{matrix}
0\ 0\ 0\ 0\ 0\ \text{-}1\ 0\ 0 \\
0\ 0\ 0\ 0\ \text{-}1\ 0\ 0\ 0 \\
0\ 0\ 0\ 0\ 0\ 0\ 0\ \text{-}1 \\
0\ 0\ 0\ 0\ 0\ 0\ \text{-}1\ 0 \\
0\ \text{-}1\ 0\ 0\ 0\ 0\ 0\ 0 \\
\text{-}1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
0\ 0\ 0\ \text{-}1\ 0\ 0\ 0\ 0 \\
0\ 0\ \text{-}1\ 0\ 0\ 0\ 0\ 0
\end{matrix}
\ +
$$

$$
\begin{matrix}
0\ 0\ 0\ 0\ 0\ 0\ \text{-}1\ 0 \\
0\ 0\ 0\ 0\ 0\ 0\ 0\ \text{-}1 \\
0\ 0\ 0\ 0\ \text{-}1\ 0\ 0\ 0 \\
0\ 0\ 0\ 0\ 0\ \text{-}1\ 0\ 0 \\
0\ 0\ \text{-}1\ 0\ 0\ 0\ 0\ 0 \\
0\ 0\ 0\ \text{-}1\ 0\ 0\ 0\ 0 \\
\text{-}1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
0\ \text{-}1\ 0\ 0\ 0\ 0\ 0\ 0
\end{matrix}
\ +\
\begin{matrix}
0\ 0\ 0\ 0\ 0\ 0\ 0\ \text{-}1 \\
0\ 0\ 0\ 0\ 0\ 0\ \text{-}1\ 0 \\
0\ 0\ 0\ 0\ 0\ \text{-}1\ 0\ 0 \\
0\ 0\ 0\ 0\ \text{-}1\ 0\ 0\ 0 \\
0\ 0\ 0\ \text{-}1\ 0\ 0\ 0\ 0 \\
0\ 0\ \text{-}1\ 0\ 0\ 0\ 0\ 0 \\
0\ \text{-}1\ 0\ 0\ 0\ 0\ 0\ 0 \\
\text{-}1\ 0\ 0\ 0\ 0\ 0\ 0\ 0
\end{matrix}
$$

| * | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|---|---|---|---|---|---|---|---|---|
| $e_0$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
| $e_1$ | $e_1$ | $e_0$ | $e_3$ | $e_2$ | $e_5$ | $e_4$ | $e_7$ | $e_6$ |
| $e_2$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ | $e_6$ | $e_7$ | $e_4$ | $e_5$ |
| $e_3$ | $e_3$ | $e_2$ | $e_1$ | $e_0$ | $e_7$ | $e_6$ | $e_5$ | $e_4$ |
| $e_4$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_5$ | $e_5$ | $e_4$ | $e_7$ | $e_6$ | $e_1$ | $e_0$ | $e_3$ | $e_2$ |
| $e_6$ | $e_6$ | $e_7$ | $e_4$ | $e_5$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ |
| $e_7$ | $e_7$ | $e_6$ | $e_5$ | $e_4$ | $e_3$ | $e_2$ | $e_1$ | $e_0$ |

**Fig. 7.6**. The decomposition of the (8*8)-matrix $M_{31}$ of 64 triplets from Fig. 7.5 and the multiplication table (at the bottom) of the sparse matrices $e_0$, $e_1$, …, $e_7$, which are shown above. The bold boxes in the multiplication table refer to the subalgebras of 2-dimensional and 4-dimensional hyperbolic numbers.

By analogy, one can take from the equalities (2) another vector [Σ%NAN, Σ%NTN, Σ%NCN, Σ%NGN]; its coordinates correspond to summary percentages of those tetra-groupings of 64 triplets, which have nucleotides C, G, A, and T at their second positions and occupy a new set of cells in the percentages matrix of triplets. Let us denote matrix cells, which are occupied by 32 triplets NCN and NGN, by number +1, and matrix cells, which are occupied by 32 triplets NAN and NTN, by number -1. Fig. 7.7 shows a result of such denotations, which lead to a numeric matrix $M_{32}$ with entries +1 and -1.

$$
M_{32} = \begin{array}{|cccc|cccc|}
\hline
+1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\
+1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\
-1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 \\
-1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 \\
\hline
+1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\
+1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\
-1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 \\
-1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 \\
\hline
\end{array}
$$

**Fig. 7.7**. Transformation of the symbolic (8*8)-matrix of 64 triplets (Figs. 2.1 and 2.4) into the numeric matrix $M_{32}$, which shows the arrangements of cells, occupied by 32 triplets NCN and NGN (black cells) and also by 32 triplets NAN and NTN (white cells).

The received numeric matrix $M_{32}$ can be also decomposed by the same method of the dyadic-shift decomposition into a sum of new 8 sparse matrices (as it was made above for matrices $M_{21}$, $M_{22}$, and $M_{31}$); the set of these 8 sparse matrices turns out to be closed about the multiplication again and defines the same multiplication table of these matrices as in the case of the matrix $M_{31}$ (see Fig. 7.6, at the bottom).

One can take from the equalities (2) also another vector [Σ%NNA, Σ%NNT, Σ%NNC, Σ%NNG]; its coordinates correspond to summary percentages of those tetra-groupings of 64 triplets, which have nucleotides C, G, A, and T at their third positions and occupy a new set of cells in the percentages matrix of triplets. Let us denote matrix cells, which are occupied by 32 triplets NNC and NNG, by number +1, and matrix cells, which are occupied by 32 triplets NNA and NNT, by number -1. Fig. 7.8 shows a result of such denotations, which lead to a numeric matrix $M_{33}$ with entries +1 and -1.

$$
M_{33} = \begin{array}{|cccc|cccc|}
\hline
+1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\
-1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 \\
+1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\
-1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 \\
\hline
+1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\
-1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 \\
+1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\
-1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 \\
\hline
\end{array}
$$

**Fig. 7.8**. Transformation of the symbolic (8*8)-matrix of 64 triplets (from Figs. 2.1 and 2.4) into the numeric matrix $M_{33}$, which shows the arrangements of cells, occupied by 32 triplets NNC and NNG (black cells) and also by 32 triplets NNA and NNT (white cells).

The received numeric matrix $M_{33}$ can be also decomposed by the same method of the dyadic-shift decomposition into a sum of new 8 sparse matrices (as it was made above for matrices $M_{21}$, $M_{22}$, $M_{31}$, and $M_{32}$); the set of these 8 sparse matrices turns out to be closed about the multiplication again and defines the same

multiplication table of these matrices as in the case of the matrix $M_{31}$ (see Fig. 7.6, at the bottom).

The same analysis with similar results can be done for tetra-groupings of matrix cells regarding 4-dimensional vectors of percentages sums of tetraplets from the equalities (2). One can note else that the sparse matrices arising in the dyadic-shift decomposition of matrix representations of $2^n$-dimensional hyperbolic numbers (Figs. 7.2, 7.4, 7.6) are symmetric self-adjoint matrices. This shows another connection between genetic structures and quantum mechanics, in which the concept of self-adjoint matrices plays an important role.

Let us return to the table (Fig. 3.1) of tetra-groupings of the sums of genomic percentages of *n*-plets in the layers of the DNA-text and see how the *n*-plets of these groupings are located in the cells of the initial tensor family of genetic matrices $[C, A; T, G]^{(n)}$. Taking into account the belonging of each *n*-plet to one of the 4 columns of the table, that is, to one of the C-, G-, A-, and T-groupings, we introduce the following symbolic notation for each of the *n*-plets:
- *n*-plets of the first column (C-column) are denoted by the symbol ♠;
- *n*-plets of the second column (G-column) are denoted by the symbol ♣;
- *n*-plets of the third column (A-column) are denoted by the symbol ♥;
- *n*-plets of the fourth column (T-column) are denoted by the symbol ♦.

Figs. 7.9-7.12 show the arrangement of *n*-plets belonging to C-, G-, A-, T-groupings of *n*-plets in the matrices of the tensor family $[C, A; T, G]^{\wedge}(n)$ under n = 1, 2, 3, 4.
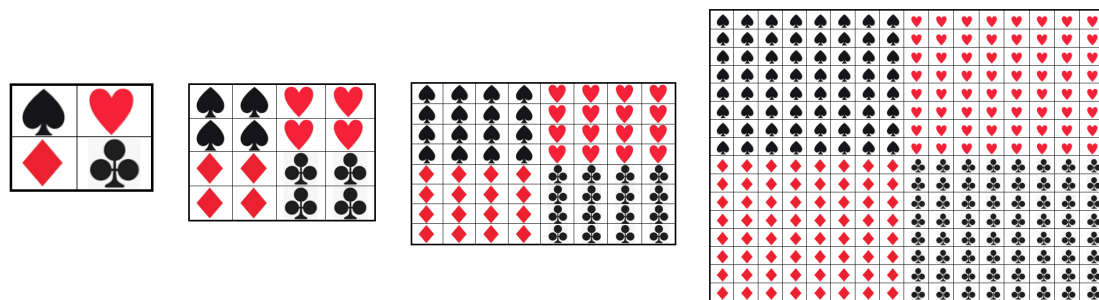


**Fig. 7.9**. The arrangements of *n*-plets, which start with the nucleotides C, or G, or A, or T, in the matrices of the tensor family $[C, A; T, G]^{(n)}$ under *n* = 1, 2, 3, 4 (see explanations in the text).
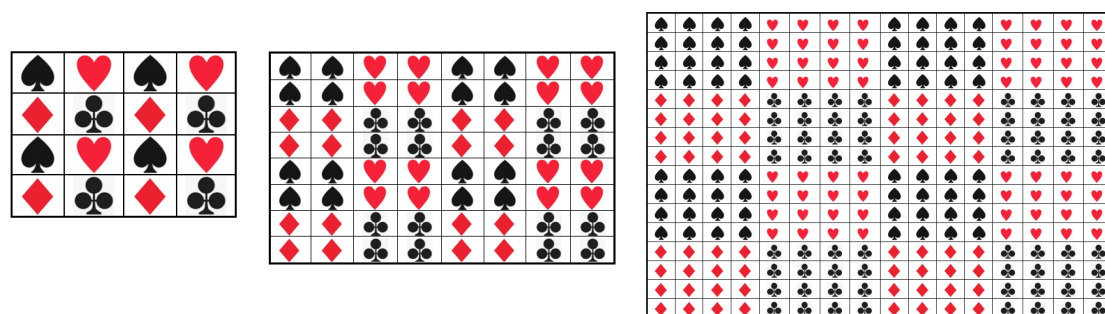


**Fig. 7.10**. The arrangements of *n*-plets, which have the nucleotides C, or G, or A, or T at their second position, in the matrices of the tensor family $[C, A; T, G]^{(n)}$ under *n* = 2, 3, 4 (see explanations in the text).
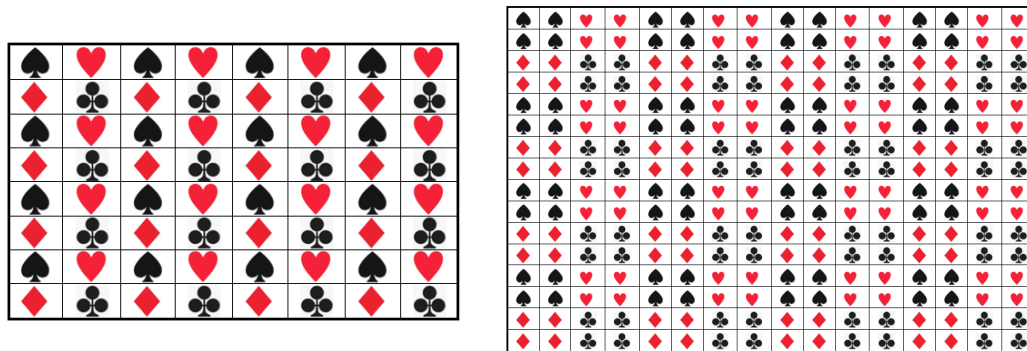
**Fig. 7.11**. The arrangements of 64 triplets and 256 tetraplets, which have the nucleotides C, or G, or A, or T at their third position, in the matrices of the tensor family $[C, A; T, G]^{(n)}$ under $n = 3, 4$ (see explanations in the text).
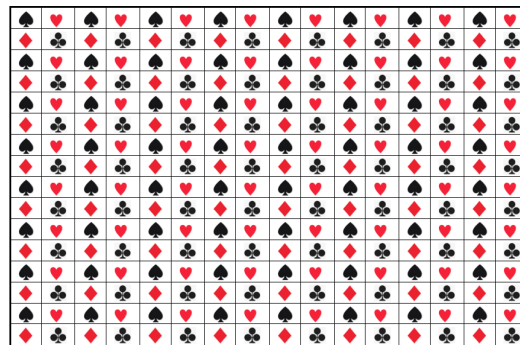


**Fig. 7.12**. The arrangements of 256 tetraplets, which have the nucleotides C, or G, or A, or T at their fourth position, in the matrix $[C, A; T, G]^{(4)}$ (see explanations in the text).

One can see that all mosaic matrices in Figs. 7.9-7.12 are bi-symmetrical about arrangements of black and red symbols. They also possess meander-like patterns: each column and each row consists of sequences of fragments of equal lengths with black and red symbols. These meander-like patterns resemble meander functions of Rademacher and Walsh, which well-known in the theory of noise-immune coding of information and which were used previously in matrix genetics [Petoukhov, 2008; Petoukhov, He, 2010].

## 8. Projectors and tetra-groupings of the percentage composition of *n*-plets.

Let us continue studying properties of the percentage content of *n*-plets in the multilayer presentation of long DNA-texts. Till now we considered the Gestalt rule related to *m*-positional tetra-groupings of *n*-plets, which are connected with the algebra of hyperbolic numbers (see the previous section). For this reason, we additionally call such *m*-positional tetra-groupings hyperbolic tetra-groupings and the corresponding Gestalt rule hyperbolic Gestalt rule. This section is devoted to another type of tetra-groupings called projector tetra-groupings of *n*-plets since they are

related to oblique projectors. The projector tetra-groupings are connected with the corresponding projector Gestalt rule described below.

Our visual perception and a "projection method" in the drawing are based on projections of external objects on the retina and a drawing plane (Fig. 8.1). In mathematics, such operations of projections are expressed using square matrices, which are called "projection operators" (or briefly, "projectors").
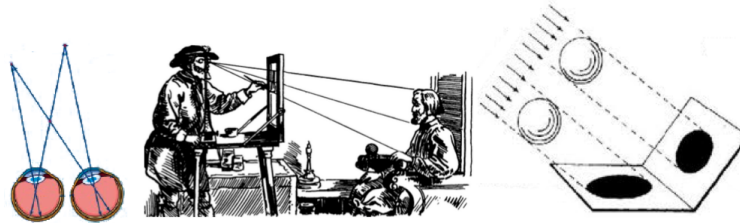


**Fig. 8.1**. Examples of projection. Left: the path of light rays entering the eyes. Middle: Ray projection in the old manner of painting. Right: oblique and orthogonal projection.

The notion of projection operators is one of the important notions in many non-biological sciences: physics, including quantum mechanics; mathematics; computer science, and informatics, including the theory of digital codes; chemistry; mathematical logic, etc. The matrix P (Fig. 8.2) is an example of a projector, making a projection of vector [x, y, z] on the plane [x, y, 0]:
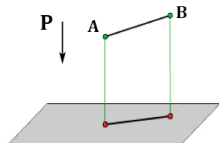
$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad P\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$



**Fig. 8.2.** Example of a matrix operator P, which provides a projection operation.

Projectors are expressed using square matrices (http://mathworld.wolfram.com/ProjectionMatrix.html). A necessary and sufficient condition that a matrix P is a projection operator is the fulfillment of the following condition: $P^2 = P$. A set of projectors is separated into two sub-sets:
- Orthogonal projectors, which are expressed by symmetric matrices. Their theory is well developed and has a lot of applications;
- Oblique projectors, which are expressed by non-symmetric matrices. Their theory and its applications are developed much weaker. Namely oblique projectors will be the main objects of attention in this Section.

Concerning the theme of projectors in inherited biological phenomena, one can note that our genetically inherited visual system works on the principle of projection of external objects at the retina. The described author's results give pieces of evidence that the meaning of projectors for bioinformatics is not limited to this single fact of the biological significance of projection operators, but that the whole system of genetic informatics uses them actively.

Now another type of tetra-groupings of *n*-plets in different layers of genomic DNA-texts will be presented, whose groupings percentage sums conserve constant values both in different layers and in different projector tetra-groupings of the $n^{th}$

layer ($n$ = 2, 3, 4, …, but much less than the length of the genomic DNA text considered). We conventionally call tetra-groupings of this new type the projector tetra-groupings to distinguish them from the above-considered $m$-positional tetra-groupings presented in expressions (3.1) and (3.2) and additionally called hyperbolic tetra-groupings.

　　　　Let us return to the percentage matrices of 16 doublets and 64 triplets in the human chromosome №1 (Figs. 2.3 and 2.4). To facilitate a visual presentation of these projector tetra-groupings of $n$-plets, the percentage matrices from Figs. 2.3 and 2.4 are transformed in corresponding symbolic matrices in Fig. 8.3 with 4 types of playing card symbols in their cells: spades (♠), hearts (♥), diamonds (♦), and clubs (♣).
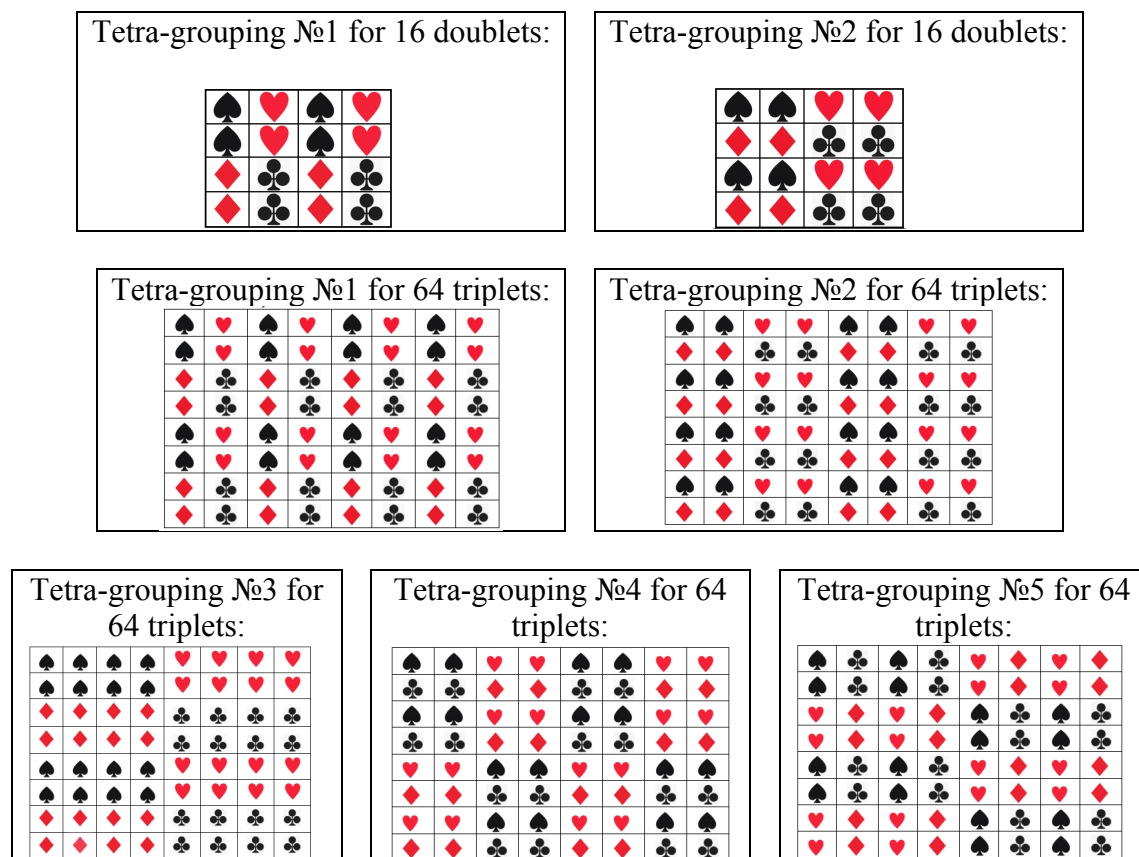


**Fig. 8.3**. Arrangements of $n$-plets from different projector tetra-groupings of in matrices for 16 doublets and 64 triplets from Fig. 2.1. Corresponding cells of each of the four groupings is denoted by one of the symbols of playing cards.

　　　　The amazing result is constancy of percentage sums Σ%♥, Σ%♦, Σ%♠, and Σ%♣ of $n$-plets, which are denoted by the same symbol, in all matrices with different mosaics of tetra-groupings shown in Fig. 8.3:

　　　Σ%♥ ≈ constant, Σ%♦ ≈ constant, Σ%♠ ≈ constant, Σ%♣ ≈ constant.　　　(8.1)

　　　　For example, for the human chromosome №1, whose doublets percentage (4*4)-matrix is shown in Fig. 2.3, the following sums exist for the projector tetra-grouping №1 of 16 doublets shown in Fig. 8.3:

$\Sigma\%\heartsuit = \%CA + \%CG + \%AA + \%AG \approx$
$\qquad \approx 0.07274 + 0.01031 + 0.09504 + 0.07137 \approx \mathbf{0.250},$
$\Sigma\%\diamondsuit = \%TC + \%TT + \%GC + \%GT \approx$
$\qquad \approx 0.06008 + 0.09568 + 0.04402 + 0.05046 \approx \mathbf{0.250},$
$\Sigma\%\spadesuit = \%CC + \%CT + \%AC + \%AT \approx$
$\qquad \approx 0.05409 + 0.07134 + 0.05033 + 0.07429 \approx \mathbf{0.250},$
$\Sigma\%\clubsuit = \%TA + \%TG + \%GA + \%GG \approx$
$\qquad \approx 0.06312 + 0.07286 + 0.06008 + 0.05419 \approx \mathbf{0.250}.$ (8.2)

The tetra-grouping №2 for 16 doublets, shown in Fig. 8.3, has another combination of 4 doublets in each grouping but the same percentage sums by analogy with (8.2):

$\Sigma\%\heartsuit = \%AC + \%AA + \%GC + \%GA \approx$
$\qquad \approx 0.05033 + 0.09504 + 0.04402 + 0.06008 \approx \mathbf{0.250},$
$\Sigma\%\diamondsuit = \%CT + \%CG + \%TT + \%TG \approx$
$\qquad \approx 0.07134 + 0.01031 + 0.09568 + 0.07286 \approx \mathbf{0.250},$
$\Sigma\%\spadesuit = \%CC + \%CA + \%TC + \%TA \approx$
$\qquad \approx 0.05409 + 0.07274 + 0.06008 + 0.06312 \approx \mathbf{0.250},$
$\Sigma\%\clubsuit = \%AT + \%AG + \%GT + \%GG \approx$
$\qquad \approx 0.07429 + 0.07137 + 0.05046 + 0.05419 \approx \mathbf{0.250}$ (8.3)

A similar rule of total percent constancy holds for all five (8\*8)-matrices in Fig. 8.3, whose mosaics present different projector tetra-groupings for the alphabet of 64 triplets. Each reader can check this by direct summations of corresponding percents. We illustrate this constancy here only for a total percentage $\Sigma\%\spadesuit$ of 16 triplets marked by the symbol $\spadesuit$, using mosaic matrices in Fig. 8.3 and percentage matrices in Fig. 2.4.

The tetra-grouping №1 for 64 triplets in Fig. 8.3 gives the following:
$\Sigma\%\spadesuit = \%CCC+\%CCT+\%CAC+\%CAT+\%ACC+\%ACT+\%AAC+\%AAT+$
$\qquad +\%TTC+TCT+\%TAC+\%TAT+\%GCC+\%GCT+\%GAC+\%GAT \approx$
$\qquad \approx 0.0138+0.0185+0.0152+0.0179+0.0118+0.0162+0.0145+0.0238+0.0159$
$\qquad +0.0223+0.0110+0.0194+0.0125+0.0144+0.0096+0.0133 \approx \mathbf{0.250}.$ (8.4)

The projector tetra-grouping №2 for 64 triplets in Fig. 8.3 gives the following:

$\Sigma\%\spadesuit = \%CCC+\%CCA+\%ACC+\%ACA+\%CTC+\%CTA+\%ATC+\%ATA+$
$\qquad +\%TCC+\%TCA+\%GCC+\%GCA+\%TTC+\%TTA+\%GTC+\%GTA \approx$
$\qquad \approx 0.0138+0.0188+0.0118+0.0198+0.0176+0.0127+0.0132+0.0194+0.0159$
$\qquad +0.0196+0.0125+0.0146+0.0197+0.0198+0.0096+0.0112 \approx \mathbf{0.250}.$ (8.5)

The tetra-grouping №3 for 64 triplets in Fig. 8.3 gives the following:

$\Sigma\%\spadesuit = \%CCC+\%CCA+\%CAC+\%CAA+\%CCT+\%CCG+\%CAT+\%CAG+\%TCC+$
$\qquad +\%TCA+\%TAC+\%TAA+\%TCT+\%TCG+\%TAT+\%TAG \approx$
$\qquad \approx 0.0138+0.0188+0.0152+0.0186+0.0185+0.0029+0.0179+0.0210+0.0159+$
$\qquad +0.0196+0.0110+0.0199+0.0223+0.0023+0.0194+0.0128 \approx \mathbf{0.250}.$ (8.6)

The tetra-grouping №4 for 64 triplets in Fig. 8.3 gives the following:

$\Sigma\%\spadesuit$ = %CCC+%CCA+%ACC+%ACA+%CTC+%CTA+%ATC+%ATA+%TAC
     +%TAA+%GAC+%GAA+%TTA+%TGA+%GGC+%GGA ≈
  ≈ 0.0138+0.0188+0.0118+0.0198+0.0176+0.0127+0.0132+0.0194+
   +0.0110+0.0096+0.0196+0.0146+0.0195+0.0126+0.0160 ≈ **0.250**.          (8.7)

The projector tetra-grouping №5 for 64 triplets in Fig. 8.3 gives the following:

$\Sigma\%\spadesuit$ = %CCC+%CCT+%CAC+%CAT+%ATC+%ATT+%AGC+%AGT+%TCC
     +%TCT+%TAC+%TAT+%GTC+%GTT+%GGC+%GGT ≈
  ≈ 0.0138+0.0185+0.0152+0.0179+0.0132+0.0239+0.0144+0.0161+0.0159+
   +0.0223+0.0110+0.0194+0.0096+0.0145+0.0126+0.0119 ≈ **0.250**.          (8.8)

In the particular numeric example of the human chromosome №1, the expressions (8.2-8.8) show that $\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit \approx \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \approx 0.250$. But in the general case, for genomic DNA-texts, the following two approximate equalities are fulfilled:

$$\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit, \qquad \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \qquad\qquad (8.9)$$

For example, in the case of the bacterial genome of *Bradyrhizobium japonicum* (see Appendix 1) the following equalities are fulfilled for projector tetra-groupings of *n*-plets in all 7 mosaic matrices of doublets and triplets in Fig. 8.3:

$$\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit \approx 0.277, \qquad \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \approx 0.223 \qquad\qquad (8.10)$$

The author considers this genomic phenomenon of the constancy of summary percents (8.10) in projector tetra-groupings, having a different percentage of separate *n*-plets in various layers of long DNA-texts, as a basis of a Gestalt genetic rule of probabilities for projector tetra-groupings and as a basis of a corresponding supposed exclusion principle in genetics.

It should be noted a significant difference of the Gestalt phenomena in the case of *m*-positional tetra-groupings, which are related to the expressions (3.1) or (3.2), and in the case of projector tetra-groupings, which are related to the expression (8.9). In *m*-positional tetra-groupings, values of percentage sums of *n*-plets in each of the 4 groupings are equal to a percentage of appropriate nucleotides A, T, C, or G. In projector tetra-groupings, values of percentage sums of *n*-plets in each of the 4 groupings have no obvious connection with the percentage of separate nucleotides.

One should mention that only some examples of projector tetra-groupings of *n*-plets are presented in this article. More number of projector tetra-groupings of *n*-plets exist in the set of 256 tetraplets. It seems natural to suppose that projector tetra-groupings exist also in sets of $4^n$ kinds of *n*-plets under *n* = 5, 6, 7, … but the author didn't check this supposition.

The author considers this genomic phenomenon of the constancy of summary percent (8.10) in different projector tetra-groupings, having different percentages of separate *n*-plets in the described layers of long DNA-texts, as a basis of a corresponding **Gestalt genetic rule for projector tetra-groupings**.

Similar percentage sums for these 7 tetra-groupings (Fig. 8.3) we have in the case of the model percentages of 16 doublets and 64 triplets in the matrices (Figs. 3.2 and 3.3), which are generated as the second and the third tensor powers of the

percentage matrix of nucleotides and which have their own percentages of separate doublets and triplets:

$$\mathbf{\Sigma\%}\spadesuit \approx 0.250; \ \ \mathbf{\Sigma\%}\clubsuit \approx 0.250; \ \ \mathbf{\Sigma\%}\heartsuit \approx 0.250; \ \ \mathbf{\Sigma\%}\diamondsuit \approx 0.250. \qquad (8.11)$$

Fig. 8.4 shows that some examples of (4*4)-matrices with Gestalt projector tetra-groupings for 16 doublets and of (8*8)-matrices with Gestalt projector tetra-groupings for 64 triplets (presented in Fig. 8.3) have a connection based on an algorithm of the tensor product by the (2*2)-matrix [1, 1; 1, 1].



**Fig. 8.4**. Tensor interrelations between some projector (4*4)- and (8*8)-matrices from Fig. 8.3. The symbol $\otimes$ refers to the tensor product.

One can mention that a list of connections of oblique projection operators with the percentage content of *n*-plets in the multilayer presentation of genomic DNA-texts is far from being limited to the tetra-groupings noted above in this article.

**9. Projector decompositions of the matrices bearing projector tetra-groupings.**

This section shows a special decomposition of each of the symbolic matrices in Fig. 8.3 and explains why the tetra-groupings of *n*-plets in these matrices are conditionally called the projector tetra-groupings.

Fig. 9.1 shows a decomposition of the symbolic matrix of tetra-grouping №1 for 16 doublets from Fig. 8.3, which presents this matrix as the sum of four symbolic sparse matrices. Arrangements of symbols in these sparse matrices are not accidental but correspond to arrangements of non-zero entries in a set of four numeric matrices $e_0$, $e_1$, $e_2$, and $e_3$, each of which is an oblique projector since $e_0^2 = e_0$, $e_1^2 = e_1$, $e_2^2 = e_2$, and $e_3^2 = e_3$. Besides the set of these projectors $e_0$, $e_1$, $e_2$, and $e_3$ is closed relative to multiplication: multiplication of any two of these projectors gives a projector from this set. The multiplication table, which is defined by these 4 oblique projectors, coincides with the multiplication table of a non-commutative algebra of 4-dimensional numbers (Fig. 9.1, bottom row). Correspondingly each of these projectors $e_0$, $e_1$, $e_2$, and $e_3$ is a bases unit of this algebra.

**Fig. 9.1**. Upper row: the decomposition of the symbolic matrix of the projector tetra-grouping №1 for 16 doublets (from Fig. 8.3) into the sum of four symbolic sparse matrices. Bottom row: corresponding four numeric sparse matrices, each of which is an oblique projector, define the multiplication table of a non-commutative algebra of 4-dimensional numbers (shown at right).

Fig. 9.2 shows a decomposition of the symbolic matrix of the projector tetra-grouping №2 for 16 doublets from Fig. 8.3, which presents this matrix as the sum of four symbolic sparse matrices. Arrangements of symbols in these sparse matrices correspond to arrangements of non-zero entries in a set of four numeric matrices $e_0$, $e_1$, $e_2$, and $e_3$, each of which is an oblique projector since $e_0^2 = e_0$, $e_1^2 = e_1$, $e_2^2 = e_2$, and $e_3^2 = e_3$. Besides the set of these projectors $e_0$, $e_1$, $e_2$, and $e_3$ is closed relative to multiplication as well. The multiplication table, which is defined by these 4 oblique projectors, is the same multiplication table of the non-commutative algebra of 4-dimensional numbers, which was shown in Fig. 9.1.



**Fig. 9.2**. Upper row: the decomposition of the symbolic matrix of the projector tetra-grouping №2 for 16 doublets (from Fig. 8.3) into the sum of four symbolic sparse matrices. Bottom row: corresponding four numeric sparse matrices, each of which is an oblique projector; they define the same multiplication table of the non-commutative algebra of 4-dimensional numbers as in Fig. 9.1.

Fig. 9.3 shows a decomposition of the symbolic matrix of the projector tetra-grouping №1 for 64 triplets from Fig. 8.3, which presents this matrix as the sum of 8 symbolic sparse matrices. Arrangements of symbols in these sparse matrices correspond to arrangements of non-zero entries in a shown set of 8 numeric matrices $e_0$, $e_1$, $e_2$, $e_3$, $e_4$, $e_5$, $e_6$, and $e_7$, each of which is an oblique projector satisfying the criterion $e_i^2 = e_i$ $(i = 0, 1, 2,\ldots,7)$. Besides the set of these projectors $e_0$, $e_1$, $\ldots$, $e_7$ is closed relative to multiplication as well. A multiplication table, which is defined by these 8 oblique projectors, is the multiplication table of a non-commutative algebra of 8-dimensional numbers.



| * | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|---|---|---|---|---|---|---|---|---|
| $e_0$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ |
| $e_1$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ |
| $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ |
| $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ |
| $e_4$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ |
| $e_5$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ |
| $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ |
| $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ |

**Fig. 9.3**. Upper part: the decomposition of the symbolic matrix of the projector tetra-grouping №1 for 64 triplets (from Fig. 8.3) into the sum of 8 symbolic sparse matrices. Middle part: corresponding 8 numeric sparse matrices, each of which is an oblique projector. Bottom part: the multiplication table of a non-commutative algebra of 8-dimensional numbers, which is defined by these 8 sparse matrices and which has 4 sub-algebras marked by bold borders.

Fig. 9.4 shows a decomposition of the symbolic matrix of the projector tetra-grouping №2 for 64 triplets from Fig. 8.3, which presents this matrix as the sum of 8 symbolic sparse matrices. Arrangements of symbols in these sparse matrices correspond to arrangements of non-zero entries in a shown set of 8 numeric matrices $e_0$, $e_1$, $e_2$, $e_3$, $e_4$, $e_5$, $e_6$, and $e_7$, each of which is an oblique projector satisfying the criterion $e_i^2 = e_i$ ($i = 0, 1, 2,…,7$). Besides the set of these projectors $e_0$, $e_1$, …, $e_7$ is closed relative to multiplication as well. A multiplication table, which is defined by these 8 oblique projectors, is the multiplication table of a non-commutative algebra of 8-dimensional numbers.

$$e_0+e_1+e_2+e_3+e_4+e_5+e_6+e_7 =$$

$$
\begin{vmatrix}
11000000 \\
00000000 \\
11000000 \\
00000000 \\
11000000 \\
00000000 \\
11000000 \\
00000000
\end{vmatrix}
+
\begin{vmatrix}
00001100 \\
00000000 \\
00001100 \\
00000000 \\
00001100 \\
00000000 \\
00001100 \\
00000000
\end{vmatrix}
+
\begin{vmatrix}
00110000 \\
00000000 \\
00110000 \\
00000000 \\
00110000 \\
00000000 \\
00110000 \\
00000000
\end{vmatrix}
+
\begin{vmatrix}
00000011 \\
00000000 \\
00000011 \\
00000000 \\
00000011 \\
00000000 \\
00000011 \\
00000000
\end{vmatrix}
+
$$

$$
\begin{vmatrix}
00000000 \\
11000000 \\
00000000 \\
11000000 \\
00000000 \\
11000000 \\
00000000 \\
11000000
\end{vmatrix}
+
\begin{vmatrix}
00000000 \\
00001100 \\
00000000 \\
00001100 \\
00000000 \\
00001100 \\
00000000 \\
00001100
\end{vmatrix}
+
\begin{vmatrix}
00000000 \\
00110000 \\
00000000 \\
00110000 \\
00000000 \\
00110000 \\
00000000 \\
00110000
\end{vmatrix}
+
\begin{vmatrix}
00000000 \\
00000011 \\
00000000 \\
00000011 \\
00000000 \\
00000011 \\
00000000 \\
00000011
\end{vmatrix}
+
$$

| * | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|---|---|---|---|---|---|---|---|---|
| $e_0$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_1$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_2$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_3$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_4$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
| $e_5$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
| $e_6$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
| $e_7$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |

**Fig. 9.4**. Upper part: the decomposition of the symbolic matrix of the projector tetra-grouping №2 for 64 triplets (from Fig. 8.3) into the sum of 8 symbolic sparse matrices. Middle part: corresponding 8 numeric sparse matrices, each of which is an oblique projector. Bottom part: the multiplication table of a non-commutative algebra of 8-dimensional numbers, which is defined by these 8 sparse matrices and which has 4 sub-algebras marked by bold borders.

Fig. 9.5 shows a decomposition of the symbolic matrix of the projector tetra-grouping №3 for 64 triplets from Fig. 8.3, which presents this matrix as the sum of 8 symbolic sparse matrices. Arrangements of symbols in these sparse matrices correspond to arrangements of non-zero entries in a shown set of 8 numeric matrices $e_0$, $e_1$, $e_2$, $e_3$, $e_4$, $e_5$, $e_6$, and $e_7$, each of which is an oblique projector satisfying the criterion $e_i^2 = e_i$ (i = 0, 1, 2,…,7). Besides the set of these projectors $e_0$, $e_1$, …, $e_7$ is closed relative to multiplication as well. A multiplication table, which is defined by these 8 oblique projectors, is the multiplication table of a non-commutative algebra of 8-dimensional numbers.

$$
\begin{aligned}
&e_0+e_1+e_2+e_3+e_4+e_5\\
&+e_6+e_7 \qquad =
\end{aligned}
\quad
\begin{array}{l}
11110000\\
00000000\\
00000000\\
00000000\\
11110000\\
00000000\\
00000000\\
00000000
\end{array}
+
\begin{array}{l}
00000000\\
11110000\\
00000000\\
00000000\\
00000000\\
11110000\\
00000000\\
00000000
\end{array}
+
\begin{array}{l}
00001111\\
00000000\\
00000000\\
00000000\\
00001111\\
00000000\\
00000000\\
00000000
\end{array}
+
\begin{array}{l}
00000000\\
00001111\\
00000000\\
00000000\\
00000000\\
00001111\\
00000000\\
00000000
\end{array}
+
$$

$$
\begin{array}{l}
00000000\\
00000000\\
11110000\\
00000000\\
00000000\\
00000000\\
11110000\\
00000000
\end{array}
+
\begin{array}{l}
00000000\\
00000000\\
00000000\\
11110000\\
00000000\\
00000000\\
00000000\\
11110000
\end{array}
+
\begin{array}{l}
00000000\\
00000000\\
00001111\\
00000000\\
00000000\\
00000000\\
00001111\\
00000000
\end{array}
+
\begin{array}{l}
00000000\\
00000000\\
00000000\\
00001111\\
00000000\\
00000000\\
00000000\\
00001111
\end{array}
$$

| *   | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|
| $e_0$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ |
| $e_1$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ |
| $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ |
| $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ |
| $e_4$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ |
| $e_5$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ |
| $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ |
| $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ |

**Fig. 9.5**. Upper part: the decomposition of the symbolic matrix of the projector tetra-grouping №3 for 64 triplets (from Fig. 8.3) into the sum of 8 symbolic sparse matrices. Middle part: corresponding 8 numeric sparse matrices, each of which is an oblique projector. Bottom part: the multiplication table of a non-commutative algebra of 8-dimensional numbers, which is defined by these 8 sparse matrices and which has 4 sub-algebras marked by bold borders.

Fig. 9.6 shows a decomposition of the symbolic matrix of the projector tetra-grouping №4 for 64 triplets from Fig. 8.3, which presents this matrix as the sum of 8 symbolic sparse matrices. Arrangements of symbols in these sparse matrices correspond to arrangements of non-zero entries in a shown set of 8 numeric matrices $e_0$, $e_1$, $e_2$, $e_3$, $e_4$, $e_5$, $e_6$, and $e_7$, each of which is an oblique projector satisfying the criterion $e_i^2 = e_i$ ($i = 0, 1, 2,…,7$). Besides the set of these projectors $e_0$, $e_1$, …, $e_7$ is closed relative to multiplication as well. A multiplication table, which is defined by these 8 oblique projectors, is the multiplication table of a non-commutative algebra of 8-dimensional numbers.

$e_0+e_1+e_2+e_3+e_4+e_5+e_6+e_7 =$

| * | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|---|---|---|---|---|---|---|---|---|
| $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ |
| $e_1$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ |
| $e_2$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ | $e_0$ | $e_2$ | $e_2$ | $e_0$ |
| $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ | $e_3$ | $e_1$ | $e_1$ | $e_3$ |
| $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ |
| $e_5$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ |
| $e_6$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ | $e_4$ | $e_6$ | $e_6$ | $e_4$ |
| $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ | $e_7$ | $e_5$ | $e_5$ | $e_7$ |

**Fig. 9.6**. Upper part: the decomposition of the symbolic matrix of the projector tetra-grouping №4 for 64 triplets (from Fig. 8.3) into the sum of 8 symbolic sparse matrices. Middle part: corresponding 8 numeric sparse matrices, each of which is an oblique projector. Bottom part: the multiplication table of a non-commutative algebra of 8-dimensional numbers, which is defined by these 8 sparse matrices.

Fig. 9.7 shows a decomposition of the symbolic matrix of the projector tetra-grouping №5 for 64 triplets from Fig. 8.3, which presents this matrix as the sum of

8 symbolic sparse matrices. Arrangements of symbols in these sparse matrices correspond to arrangements of non-zero entries in a shown set of 8 numeric matrices $e_0$, $e_1$, $e_2$, $e_3$, $e_4$, $e_5$, $e_6$, and $e_7$, each of which is an oblique projector satisfying the criterion $e_i^2 = e_i$ $(i = 0, 1, 2,\ldots,7)$. Besides the set of these projectors $e_0$, $e_1$, $\ldots$, $e_7$ is closed relative to multiplication as well. A multiplication table, which is defined by these 8 oblique projectors, is the multiplication table of a non-commutative algebra of 8-dimensional numbers.



| * | $e_0$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|---|---|---|---|---|---|---|---|---|
| $e_0$ | $e_0$ | $e_3$ | $e_0$ | $e_3$ | $e_4$ | $e_7$ | $e_4$ | $e_7$ |
| $e_1$ | $e_2$ | $e_1$ | $e_2$ | $e_1$ | $e_6$ | $e_5$ | $e_6$ | $e_5$ |
| $e_2$ | $e_2$ | $e_1$ | $e_2$ | $e_1$ | $e_6$ | $e_5$ | $e_6$ | $e_5$ |
| $e_3$ | $e_0$ | $e_3$ | $e_0$ | $e_3$ | $e_4$ | $e_7$ | $e_4$ | $e_7$ |
| $e_4$ | $e_0$ | $e_3$ | $e_0$ | $e_3$ | $e_4$ | $e_7$ | $e_4$ | $e_7$ |
| $e_5$ | $e_2$ | $e_1$ | $e_2$ | $e_1$ | $e_6$ | $e_5$ | $e_6$ | $e_5$ |
| $e_6$ | $e_2$ | $e_1$ | $e_2$ | $e_1$ | $e_6$ | $e_5$ | $e_6$ | $e_5$ |
| $e_7$ | $e_0$ | $e_3$ | $e_0$ | $e_3$ | $e_4$ | $e_7$ | $e_4$ | $e_7$ |

**Fig. 9.7**. Upper part: the decomposition of the symbolic matrix of the projector tetra-grouping №5 for 64 triplets (from Fig. 8.3) into the sum of 8 symbolic sparse

matrices. Middle part: corresponding 8 numeric sparse matrices, each of which is an oblique projector. Bottom part: the multiplication table of a non-commutative algebra of 8-dimensional numbers, which is defined by these 8 sparse matrices.

As far as the author can judge, the oblique projectors, which are shown in Figs. 9.1-9.7, have not been previously used in science and are representatives of a new class of projectors (the article describes only a few examples of this wide class of oblique projectors). The author conventionally calls representatives of this class as genetic Gestalt projectors due to the following:
- 1) they were discovered in research on Gestalt genetics;
- 2) they have interesting algebraic properties that allow them to be used as operators in modeling the generation of stably structured states of multiparameter systems from their unstructured states; in particular, such modeling is very interesting for algebraic biology studying structural states of biological systems.

Their algebraic properties are briefly described in the next section.

## 10. Some algebraic properties of genetic Gestalt projectors.

To begin with, consider a 4-parameter system, whose state vector $\overline{X} = [x_0(t), x_1(t), x_2(t), x_3(t)]$ has coordinates, for example, in form of 4 functions $x_0(t)$, $x_1(t)$, $x_2(t)$, and $x_3(t)$, which are independent each other (that is, we consider a "chaotic" set of initial 4 coordinates). Let us study, for example, the action on this state vector $\overline{X}$ of the genetic Gestalt projectors $e_0$, $e_1$, $e_2$, and $e_3$, presented in Figure 9.1

$$[x_0(t), x_1(t), x_2(t), x_3(t)]*e_0 = [x_0(t) + x_1(t), x_0(t) + x_1(t), 0, 0], \qquad (10.1)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t)]*e_1 = [0, x_2(t) + x_3(t), 0, x_2(t) + x_3(t)], \qquad (10.2)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t)]*e_2 = [0, x_0(t) + x_1(t), 0, x_0(t) + x_1(t)], \qquad (10.3)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t)]*e_3 = [x_2(t) + x_3(t), 0, x_2(t) + x_3(t), 0]. \qquad (10.4)$$

Expressions (10.1-10.4) show that under the action of these oblique projectors, the state vector $\overline{X} = [x_0(t), x_1(t), x_2(t), x_3(t)]$ with unbound coordinates is transformed into a state vector with strongly bound and ordered coordinates. For example, the expression (10.1) shows that the new state vector has two coordinates with an identical function $x_0(t)+x_1(t)$, and the other two coordinates are zero. If $x_0(t)$, $x_1(t)$, $x_2(t)$, $x_3(t)$ are cyclic or wave functions, then two non-zero coordinates of the state vector are synchronously and cyclically changed in time, and the new state vector of the 4-parameter system is also cyclically changed in the 2-dimensional subspace defined by these two coordinates. One should separately note arising symmetry in the set of coordinates of the new state vector.

Below we will show an example of similar results of the ordering and synchronizing action of genetic Gestalt projectors on 8-dimensional state vectors, having random sets of coordinates, but first, we note the importance of modeling the consistency of genetically inherited cyclic processes in living bodies.

The coordinated cyclic behavior of genetically inherited parts is one of the characteristic features of living bodies. Many mysterious biological phenomena are associated with this. For example, as it is known, the first thing that comes into

motion in a three-week-old human embryo, which has a size of 4 millimeters only, is the heart; it begins to pulsate due to inner reasons since at this stage it has no nerve connections and there is no blood to pump. The living organism as a whole is a huge chorus of coordinated cyclic processes. Tibetan pulse diagnosis is based on the relationship of cyclic processes in the body, and so on.

One can add that if the state vector $\overline{X} = [x_0, x_1, x_2, x_3]$ corresponds to a quantum mechanical or quantum informational system with random amplitudes $x_0$, $x_1$, $x_2$, $x_3$ of probabilities to find the system in one of its 4 quantum states (or in one of its 4 computational basis states), then the results of the action of the genetic Gestalt projectors $e_0$, $e_1$, $e_2$, and $e_3$, presented by the expressions (10.1-10.4), draw special attention to a particular state of this system with identical non-zero probabilities of two bases states and two zero probabilities of two other bases states.

Let us consider one more example of a 8-parameter system, whose state vector $\overline{X} = [x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]$ has coordinates, for example, in form of 8 functions $x_0(t)$, $x_1(t)$, $x_2(t)$, $x_3(t)$, $x_4(t)$, $x_5(t)$, $x_6(t)$, and $x_7(t)$, which are independent each other. Expressions (10.5-10.12) show results of the action on this state vector $\overline{X}$ of the genetic Gestalt projectors $e_0$, $e_1$, $e_2$, $e_3$, $e_4$, $e_5$, $e_6$, and $e_7$, presented in Fig. 9.3.

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_0 =$$
$$= [x_0(t)+x_1(t), 0, x_0(t)+x_1(t), 0, x_0(t)+x_1(t), 0, x_0(t)+x_1(t), 0], \qquad (10.5)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_1 =$$
$$= [x_4(t)+x_5(t), 0, x_4(t)+x_5(t), 0, x_4(t)+x_5(t), 0, x_4(t)+x_5(t), 0], \qquad (10.6)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_2 =$$
$$= [0, x_0(t)+x_1(t), 0, x_0(t)+x_1(t), 0, x_0(t)+x_1(t), 0, x_0(t)+x_1(t)], \qquad (10.7)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_3 =$$
$$= [0, x_4(t)+x_5(t), 0, x_4(t)+x_5(t), 0, x_4(t)+x_5(t), 0, x_4(t)+x_5(t)], \qquad (10.8)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_4 =$$
$$= [x_2(t)+x_3(t), 0, x_2(t)+x_3(t), 0, x_2(t)+x_3(t), 0, x_2(t)+x_3(t), 0], \qquad (10.9)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_5 =$$
$$= [x_6(t)+x_7(t), 0, x_6(t)+x_7(t), 0, x_6(t)+x_7(t), 0, x_6(t)+x_7(t), 0], \qquad (10.10)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_6 =$$
$$= [x_2(t)+x_3(t), 0, x_2(t)+x_3(t), 0, x_2(t)+x_3(t), 0, x_2(t)+x_3(t), 0], \qquad (10.11)$$

$$[x_0(t), x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)]*e_7 =$$
$$= [0, x_6(t)+x_7(t), 0, x_6(t)+x_7(t), 0, x_6(t)+x_7(t), 0, x_6(t)+x_7(t)]. \qquad (10.12)$$

The expressions (10.5-10.12) confirm the ordering, synchronizing, and symmetrizing action of genetic Gestalt projectors on 8-dimensional state vectors, having random sets of coordinates.

The tensor product of these genetic Gestalt projectors, when acting on a state vector with a random set of coordinates $x_i(t)$, also generates a vector with an ordered, synchronized and symmetrized set of coordinates. This makes it possible to extend this algorithm for generating ordered, synchronized and symmetrized state vectors to

the general case of $2^n$-dimensional vectors. For example, the action of the tensor product of two projectors e0 and e3 from Fig. 9.1 on 16-dimensional vectors, having random sets of coordinates, gives the following result:

$$[x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}] * (e_0 \otimes e_3) =$$
$$= [x_2+x_3+x_6+x_7, 0, x_2+x_3+x_6+x_7, 0, 0, 0, 0, 0,$$
$$\quad x_2+x_3+x_6+x_7, 0, x_2+x_3+x_6+x_7, 0, 0, 0, 0, 0] \tag{10.13}$$

The author proposes such genetic Gestalt projectors as algebraic tools for modeling the fundamental ability of biological bodies to create a genetically inherited biological order from disorder. Regardless of the chaotic set of coordinates of the initial state vector, the result of the action of these Gestalt projectors is always the same: ordering, synchronization, and symmetrization of coordinates in the new state vector. This interconnected ordering, cyclicity, and symmetry of the coordinates of the emerging state vector of a multiparameter system can be considered as a kind of genetic Gestalt for further developing Gestalt genetics and algebraic biology. It should be additionally noted that here an interesting theme of information waves of collective probability arises in the layers of genomic DNA-texts and genomic DNA epi-chains; this topic is currently being investigated in the author's laboratory.

## 11. Gestalt percentage rules for sets of *n*-plets starting with certain k-plets (*k<n*).

Above, in section 3, we described *m*-positional tetra-groupings of *n*-plets, which are defined by a disposition of attributive nucleotides C, G, A, and T at a certain position *m* ($m \leq n$) inside these *n*-plets. Now in this section, we describe other positional tetra-groupings of *n*-plets where a role of attributive elements is played by not separate nucleotides but separate doublets, or separate triplets, and so on. As it is revealed, these new tetra-groupings of *n*-plets obey Gestalt rules but have their own values for percentage sums in each of the groupings. Let us explain this by some examples using percentages of *n*-plets in different layers of the DNA-text of the human chromosome №1 (Figs. 2.3 - 2.5).

For the beginning, consider the percentage contents of positional tetra-groupings, each of which is defined by one of 16 doublets, which is its attributive positional element. In a general case, we call such tetra-groupings as *mm*-positional tetra-groupings or, in individual cases, as a CA-tetra-grouping, or a TA-tetra-grouping, etc. Below special denotations are used for percentage sums of *n*-plets, which are explained by the following examples (here N refers to any of nucleotides A, T, C, and G):

- $\Sigma\%$TAN means a percentage sum of all 4 triplets, which start with the doublet TA in an analyzed DNA-text presented as a sequence of triplets;
- $\Sigma\%$NTA means a percentage sum of all 4 triplets, which end with the doublet TA in an analyzed DNA-text presented as a sequence of triplets;
- $\Sigma\%$TANN means a percentage sum of all 16 tetraplets, which start with the doublet TA in an analyzed DNA-text presented as a sequence of tetraplets;
- $\Sigma\%$NTAN means a percentage sum of all 16 tetraplets, which have the doublet TA in their middle in a DNA-text presented as a sequence of tetraplets;
- $\Sigma\%$NNTA means a percentage sum of all 16 tetraplets, which end with the doublet TA in an analyzed DNA-text presented as a sequence of tetraplets.

Let us calculate these percentage sums of *n*-plets for corresponding layers of the DNA-text of the human chromosome #1 using the data presented in Figs 2.3-2.5

about percents of separate doublets, triplets, and tetraplets. This calculation gives equalities shown in Fig. 11.1.

| |
|---|
| %AA ≈ Σ%AAN ≈ Σ%NAA ≈ Σ%AANN ≈ Σ%NAAN ≈ Σ%NNAA ≈ 0.095 |
| %AT ≈ Σ%ATN ≈ Σ%NAT ≈ Σ%ATNN ≈ Σ%NATN ≈ Σ%NNAT ≈ 0.074 |
| %AC ≈ Σ%ACN ≈ Σ%NAC ≈ Σ%ACNN ≈ Σ%NACN ≈ Σ%NNAC ≈ 0.050 |
| %AG ≈ Σ%AGN ≈ Σ%NAG ≈ Σ%AGNN ≈ Σ%NAGN ≈ Σ%NNAG ≈ 0.071 |
| %TA ≈ Σ%TAN ≈ Σ%NTA ≈ Σ%TANN ≈ Σ%NTAN ≈ Σ%NNTA ≈ 0.063 |
| %TT ≈ Σ%TTN ≈ Σ%NTT ≈ Σ%TTNN ≈ Σ%NTTN ≈ Σ%NNTT ≈ 0.096 |
| %TC ≈ Σ%TCN ≈ Σ%NTC ≈ Σ%TCNN ≈ Σ%NTCN ≈ Σ%NNTC ≈ 0.060 |
| %TG ≈ Σ%TGN ≈ Σ%NTG ≈ Σ%TGNN ≈ Σ%NTGN ≈ Σ%NNTG ≈ 0.073 |
| %CA ≈ Σ%CAN ≈ Σ%NCA ≈ Σ%CANN ≈ Σ%NCAN ≈ Σ%NNCA ≈ 0.073 |
| %CT ≈ Σ%CTN ≈ Σ%NCT ≈ Σ%CTNN ≈ Σ%NCTN ≈ Σ%NNCT ≈ 0.071 |
| %CC ≈ Σ%CCN ≈ Σ%NCC ≈ Σ%CCNN ≈ Σ%NCCN ≈ Σ%NNCC ≈ 0.054 |
| %CG ≈ Σ%CGN ≈ Σ%NCG ≈ Σ%CGNN ≈ Σ%NCGN ≈ Σ%NNCG ≈ 0.010 |
| %GA ≈ Σ%GAN ≈ Σ%NGA ≈ Σ%GANN ≈ Σ%NGAN ≈ Σ%NNGA ≈ 0.060 |
| %GT ≈ Σ%GTN ≈ Σ%NGT ≈ Σ%GTNN ≈ Σ%NGTN ≈ Σ%NNGT ≈ 0.050 |
| %GC ≈ Σ%GCN ≈ Σ%NGC ≈ Σ%GCNN ≈ Σ%NGCN ≈ Σ%NNGC ≈ 0.044 |
| %GG ≈ Σ%GGN ≈ Σ%NGG ≈ Σ%GGNN ≈ Σ%NGGN ≈ Σ%NNGG ≈ 0.054 |

**Fig. 11.1**. Percentage sums are presented for *mm*-positional tetra-groupings related to 64 triplets and 256 tetraplets in the appropriate triplet- and tetraplet-layers of the DNA-text of the human chromosome №1. Each of these tetra-groupings is defined by one of 16 doublets as its attributive positional element disposed of in the beginning, or in the middle, or at the end of the *n*-plets. Numerical values of percentage sums are calculated based on data about the percents of separate *n*-plets in Figs. 2.3-2.5.

Data in Fig. 11.1 show that percentage sums in each of these mm-tetra-groupings are equal to a percentage of that doublet, which plays the role of the attributive positional element for these tetra-groupings. The presented equality of these percentage sums in all tetra-groupings of each row of Fig. 11.1 occurs even though summands in them in each case are very different. Other genomes, analyzed by the author, have similar Gestalt properties for *mm*-positional tetra-groupings in the layers of their DNA-texts. This new genomic Gestalt phenomenon and the corresponding Gestalt rule for *mm*-positional tetra-groupings (Fig. 11.1) are analogs of the Gestalt phenomenon and the Gestalt rule described above for *m*-positional tetra-groupings in section 3 and its expressions (3.1) and (3.2).

Now consider percentage contents of 64 *mmm*-positional tetra-groupings, each of which is defined by one of 64 triplets, which is its attributive positional element. The analogical denotations are used for these tetra-groupings:
- Σ%TAAN means a percentage sum of all 4 tetraplets, which start with the triplet TAA in an analyzed DNA-text presented as a sequence of triplets;
- Σ%NTAA means a percentage sum of all 4 tetraplets, which end with the triplet TAA.

Let us calculate these percentage sums of n-plets for corresponding layers of the DNA-text of the human chromosome №1 using the data presented in Figs 2.4-2.5 about the percents of separate triplets and tetraplets. This calculation gives equalities shown in Fig. 11.2.

| |
|---|
| %CCC ≈ Σ%CCCN ≈ Σ%NCCC ≈ 0.014 |
| %CCG ≈ Σ%CCGN ≈ Σ%NCCG ≈ 0.003 |
| %CGC ≈ Σ%CGCN ≈ Σ%NCGC ≈ 0.003 |
| %CGG ≈ Σ%CGGN ≈ Σ%NCGG ≈ 0.003 |
| %GCC ≈ Σ%GCCN ≈ Σ%NGCC ≈ 0.013 |
| %GCG ≈ Σ%GCGN ≈ Σ%NGCG ≈ 0.003 |
| %GGC ≈ Σ%GGCN ≈ Σ%NGGC ≈ 0.013 |
| %GGG ≈ Σ%GGGN ≈ Σ%NGGG ≈ 0.014 |
| %CCA ≈ Σ%CCAN ≈ Σ%NCCA ≈ 0.019 |
| %CCT ≈ Σ%CCTN ≈ Σ%NCCT ≈ 0.019 |
| %CGA ≈ Σ%CGAN ≈ Σ%NCGA ≈ 0.002 |
| %CGT ≈ Σ%CGTN ≈ Σ%NCGT ≈ 0.003 |
| %GCA ≈ Σ%GCAN ≈ Σ%NGCA ≈ 0.015 |
| %GCT ≈ Σ%GCTN ≈ Σ%NGCT ≈ 0.014 |
| %GGA ≈ Σ%GGAN ≈ Σ%NGGA ≈ 0.016 |
| %GGT ≈ Σ%GGTN ≈ Σ%NGGT ≈ 0.012 |
| %CAC ≈ Σ%CACN ≈ Σ%NCAC ≈ 0.015 |
| %CAG ≈ Σ%CAGN ≈ Σ%NCAG ≈ 0.021 |
| %CTC ≈ Σ%CTCN ≈ Σ%NCTC ≈ 0.018 |
| %CTG ≈ Σ%CTGN ≈ Σ%NCTG ≈ 0.021 |
| %GAC ≈ Σ%GACN ≈ Σ%NGAC ≈ 0.010 |
| %GAG ≈ Σ%GAGN ≈ Σ%NGAG ≈ 0.018 |
| %GTC ≈ Σ%GTCN ≈ Σ%NGTC ≈ 0.010 |
| %GTG ≈ Σ%GTGN ≈ Σ%NGTG ≈ 0.015 |
| %CAA ≈ Σ%CAAN ≈ Σ%NCAA ≈ 0.019 |
| %CAT ≈ Σ%CATN ≈ Σ%NCAT ≈ 0.018 |
| %CTA ≈ Σ%CTAN ≈ Σ%NCTA ≈ 0.013 |
| %CTT ≈ Σ%CTTN ≈ Σ%NCTT ≈ 0.020 |
| %GAA ≈ Σ%GAAN ≈ Σ%NGAA ≈ 0.020 |
| %GAT ≈ Σ%GATN ≈ Σ%NGAT ≈ 0.013 |
| %GTA ≈ Σ%GTAN ≈ Σ%NGTA ≈ 0.011 |
| %GTT ≈ Σ%GTTN ≈ Σ%NGTT ≈ 0.014 |
| %ACC ≈ Σ%ACCN ≈ Σ%NACC ≈ 0.012 |
| %ACG ≈ Σ%ACGN ≈ Σ%NACG ≈ 0.003 |
| %AGC ≈ Σ%AGCN ≈ Σ%NAGC ≈ 0.014 |
| %AGG ≈ Σ%AGGN ≈ Σ%NAGG ≈ 0.018 |
| %TCC ≈ Σ%TCCN ≈ Σ%NTCC ≈ 0.016 |
| %TCG ≈ Σ%TCGN ≈ Σ%NTCG ≈ 0.002 |
| %TGC ≈ Σ%TGCN ≈ Σ%NTGC ≈ 0.015 |
| %TGG ≈ Σ%TGGN ≈ Σ%NTGG ≈ 0.019 |
| %ACA ≈ Σ%ACAN ≈ Σ%NACA ≈ 0.020 |
| %ACT ≈ Σ%ACTN ≈ Σ%NACT ≈ 0.016 |
| %AGA ≈ Σ%AGAN ≈ Σ%NAGA ≈ 0.022 |
| %AGT ≈ Σ%AGTN ≈ Σ%NAGT ≈ 0.016 |
| %TCA ≈ Σ%TCAN ≈ Σ%NTCA ≈ 0.020 |
| %TCT ≈ Σ%TCTN ≈ Σ%NTCT ≈ 0.022 |
| %TGA ≈ Σ%TGAN ≈ Σ%NTGA ≈ 0.019 |
| %TGT ≈ Σ%TGTN ≈ Σ%NTGT ≈ 0.020 |

| |
|---|
| %AAC ≈ Σ%AACN ≈ Σ%NAAC ≈ 0.014 |
| %AAG ≈ Σ%AAGN ≈ Σ%NAAG ≈ 0.020 |
| %ATC ≈ Σ%ATCN ≈ Σ%NATC ≈ 0.013 |
| %ATG ≈ Σ%ATGN ≈ Σ%NATG ≈ 0.018 |
| %TAC ≈ Σ%TACN ≈ Σ%NTAC ≈ 0.011 |
| %TAG ≈ Σ%TAGN ≈ Σ%NTAG ≈ 0.013 |
| %TTC ≈ Σ%TTCN ≈ Σ%NTTC ≈ 0.020 |
| %TTG ≈ Σ%TTGN ≈ Σ%NTTG ≈ 0.019 |
| %AAA ≈ Σ%AAAN ≈ Σ%NAAA ≈ 0.037 |
| %AAT ≈ Σ%AATN ≈ Σ%NAAT ≈ 0.024 |
| %ATA ≈ Σ%ATAN ≈ Σ%NATA ≈ 0.019 |
| %ATT ≈ Σ%ATTN ≈ Σ%NATT ≈ 0.024 |
| %TAA ≈ Σ%TAAN ≈ Σ%NTAA ≈ 0.020 |
| %TAT ≈ Σ%TATN ≈ Σ%NTAT ≈ 0.019 |
| %TTA ≈ Σ%TTAN ≈ Σ%NTTA ≈ 0.020 |
| %TTT ≈ Σ%TTTN ≈ Σ%NTTT ≈ 0.037 |

**Fig. 11.2**. Percentage sums are presented for *mmm*-positional tetra-groupings related to 256 tetraplets in the tetraplet-layer of the DNA-text of the human chromosome №1. Each of these tetra-groupings is defined by one of 64 triplets as its attributive positional element disposed of in the beginning or in the end of the tetraplets. Numerical values of percentage sums are calculated based on data about percents of separate triplets and tetraplets in Figs. 2.4-2.5.

Data in Fig. 11.2 show that percentage sums in each of these *mmm*-tetra-groupings are equal to a percent of the triplet, which plays the role of the attributive positional element for these tetra-groupings. The presented equality of these percentage sums in all tetra-groupings of each row of Figure 11.2 occurs even though summands in them in each case are very different. Other genomes, analyzed by the author, have similar Gestalt properties for such *mmm*-positional tetra-groupings in the appropriate layers of their DNA-texts.  This new genomic Gestalt phenomenon and the corresponding Gestalt rule for *mmm*-positional tetra-groupings (Fig. 11.2) are also analogs of the Gestalt phenomenon and the Gestalt rule described above for *m*-positional tetra-groupings in section 3 and its expressions (3.1) and (3.2).

The author suggests that similar Gestalt phenomena exist also for layers of *n*-plets with *n* = 5, 6, 7,... of genomic DNA-texts of various types of organisms but he did not make corresponding calculations at this stage of research.

At the end of section 3, the concept of reference percentages of *n*-plets in the corresponding layers of genomic DNA texts was introduced based on the tensor family of matrices $[\%C, \%A; \%T, \%G]^{(n)}$. Figs. 3.2 and 3.3 showed these reference percentages of *n*-plets for the first layers of the DNA-text of human chromosome №1. It should be noted that for these reference percentages $[\%C, \%A; \%T, \%G]^{(n)}$ similar Gestalt rules hold for *mm*-positional and *mmm*-positional tetra-groupings in the appropriate layers of genomic DNA-texts.  For example, for the human chromosome #1 in line with its data in Figs. 3.2 and 3.3, we have the following equalities for a case when 16 doublets play the role of positional attributive elements:

%CCC + %CCA + %CCT + %CCG ≈ %CC, because
0.00906382 + 0.01265035 + 0.01268315 + 0.00907448 ≈ 0.0435;

%CAC+%CAA+%CAT+%CAG ≈ %CA, because
0.01265035 + 0.01765605 + 0.01770183 + 0.01266522 ≈ 0.0607,                    (11.1)

and so on for all doublets as attributive elements in *mm*-positional tetra-groupings.

## 12. On relationships of the tetra-groupings of n-plets and Hadamard matrices.

In the previous sections of the article, individual examples of mosaic matrices with phenomenological tetra-groupings of n-plets were shown (Figs. 7.9-7.12, 8.3, 9.1-9.7), which reflect the patterns of relationships in the percentage composition of n-plets in different layers of genomic DNA-texts. Each of these matrices, having mosaics by 4 card symbols, can be decomposed into 4 sparse matrices by these card symbols. Fig. 12.1 shows examples of such decompositions for two mosaic matrices, presenting two different tetra-groupings, which are taken from Figs. 7.10 and 9.2.



**Fig. 12.1**. Two mosaic matrices with denoted arrangements of their projector tetra-groupings of 16 doublets in the second layer of genomic DNA-texts. These arrangements coincide with the arrangements of elements «+1» in 4 sparse matrices shown at right for each of the symbolic matrices. Upper row: the matrix of the projector tetra-grouping from Fig. 9.2 is presented. Bottom row: the matrix of the *m*-positional tetra-grouping from Fig. 7.10 is shown.

Let us take one of the unitary Hadamard (2*2)-matrices $H_2 = [1, -1; 1, 1]/2^{0.5}$, which presents a rotation transformation in a 2-dimensional space. Its second tensor power $H_4 = H_2^{(2)} = 0.5*[1, -1, -1, 1; 1, 1, -1, -1, 1, -1, 1, -1; 1, 1, 1, 1]$ is a unitary matrix of a rotation transformation in a 4-dimensional space. Shown in Fig. 12.1, the set of sparse matrices $e_0, e_1, e_2, e_3$ is transformed into the set of sparse matrices $s_0, s_1, s_2, s_3$ using an action of this Hadamard matrix $H_4$ by the following expressions:

$$e_0*H_4^3 = s_1, \quad e_1*H_4 = s_0; \quad e_3*H_4 = s_2; \quad e_2*H_4^3 = s_3 \qquad (12.1)$$

One can show also other examples of the role of unitary Hadamard matrices as connectors between different sets of sparse matrices received by decompositions of mosaic matrices, which present different tetra-groupings of *n*-plets in different layers of genomic DNA-texts. This noted system of interrelations of different mosaic $(2^n*2^n)$-matrices, which show appropriate tetra-groupings of *n*-plets, with Hadamard matrices should be studied systematically in the future. Hadamard matrices have wide

applications in many scientific fields and technologies [Ahmed, Rao, 1975; Seberry, Wysocki, Wysocki, 2005]. In these applications, they are usually used for actions on vectors. In contrast to this, in the described cases of matrix Gestalt genetics, they are used for actions upon the square matrix operators to receive new matrix operators by an analogy with the expressions (12.1).

## 13. Unitary matrices and the Gestalt phenomenon of biological spiralization

The author draws special attention to unitary matrices in developing Gestalt genetics since he considers unitary rotation transformation as a basis for modeling a well-known morphogenetic Gestalt phenomenon of spiralization in biology that is, the existence of helical and spiral morphological configurations at different levels and branches of biological organization independently on their genetically inherited bio-material content. Goethe even called spirals "symbols of life" because of multiple implementations of inherited spiral structures and processes in living bodies at various lines and levels of biological evolution. In the human body spiral and helical structures genetically inherited from one generation to another are presented in the muscles, heart, blood vessels, bones, nerves, an organ of hearing (the cochlea), cellular organization of the embryo (zygote), etc. The structure of tendons and ligaments consists of spirals and helices, which in turn are composed of collagen that has a triple helical structure. Spiral motions (nutation) are observed during the growth of roots and shoots, tendrils of plants are spirally wrapped, a tissue in the trunks of trees grows spirally, etc. Because of spiral bioconfigurations, all the fluids in the body (blood, lymph, and urine) are spiral. The title of a book about bio-spirals – "Lines of Life" [Cook, 1914] - reflects their importance for living matter. Previous author's publications describe structural connections of molecular-genetic systems with unitary matrices as well [Petoukhov, 2008, 2018; Petoukhov, He, 2010].

In this article, the following new structural connection of unitary matrices with genomic DNA-texts is noted. When considering a DNA double helix containing complementary nucleotide pairs C-G and A-T, the percentages of nucleotides C and G are exactly equal to each other (%C = %G), as are the percentages of nucleotides A and T (%A = %T). Taking into account genetic matrices from Figs. 2.1 and 2.2, the matrix of percentages of nucleotides in the DNA double helix is a bisymmetric doubly stochastic matrix, in which the sum of the percentages in each row and each column is equal to 1 (Figure 12.2 on the left). Such matrices are connected with unitary matrices as the following theorem claims [Prasolov, 1995]:

• if a square ($n*n$)-matrix M= $|m_{ij}|$ is unitary then a ($n*n$)-matrix B=|bij|, where

   $b_{ij} = |m_{ij}|^2$, is doubly stochastic.

In line with this theorem, Fig. 13.1 shows an interrelation between four unitary (2*2)-matrix and a doubly stochastic (2*2)-matrix of percentages of nucleotides in DNA double helices where %C = %G and %A = %T. Two of the four unitary matrices present transformations of mirror reflections. The other two unitary matrices are transformations of rotations of vectors clockwise and counterclockwise and simultaneously they are matrix presentations of complex number z = $\sqrt{2\%C}$ + i*$\sqrt{2\%A}$, where i² = -1 is the imaginary unit of complex numbers.
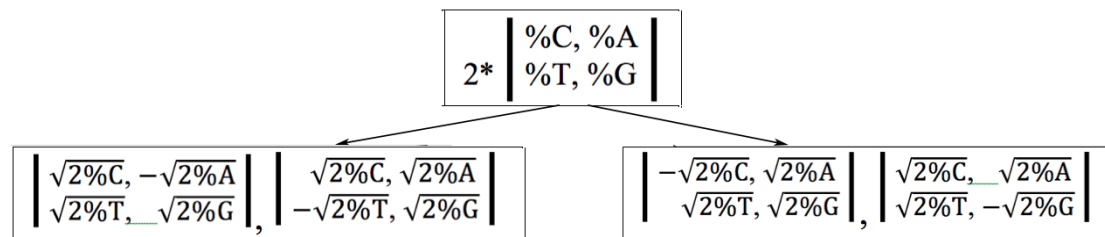
$$2* \begin{vmatrix} \%C, \%A \\ \%T, \%G \end{vmatrix}$$

$$\begin{vmatrix} \sqrt{2\%C}, -\sqrt{2\%A} \\ \sqrt{2\%T}, \ \sqrt{2\%G} \end{vmatrix}, \begin{vmatrix} \sqrt{2\%C}, \sqrt{2\%A} \\ -\sqrt{2\%T}, \sqrt{2\%G} \end{vmatrix} \qquad \begin{vmatrix} -\sqrt{2\%C}, \sqrt{2\%A} \\ \sqrt{2\%T}, \sqrt{2\%G} \end{vmatrix}, \begin{vmatrix} \sqrt{2\%C}, \ \sqrt{2\%A} \\ \sqrt{2\%T}, -\sqrt{2\%G} \end{vmatrix}$$

**Fig. 13.1**. The interrelation between the doubly stochastic matrix (upper row) of the percentage of nucleotides in the DNA double helix and four unitary matrices (bottom row). Two unitary matrices (at left) are transformations of rotation clockwise and counterclockwise and present complex number $z = \sqrt{2\%C} + i*\sqrt{2\%A}$. The other two unitary matrices (at right) present transformations of mirror reflection.

This interrelation of the structural features of DNA double helices, presented in the percentage matrix (upper row in Fig. 13.1), with unitary matrices of rotations clockwise and counterclockwise can be used for algebraic modeling of well-known Gestalt morphogenetic phenomena of left and right spiralizations realized in a great number of genetically inherited biological structures.

The described connection of matrices of nucleotide percentage in DNA double helices with complex numbers, which are presented by noted unitary matrices, is important for algebraic and quantum mechanical modeling of genetic phenomena since quantum mechanics is closely related to complex numbers.

Such interrelations of (2*2)-matrices, shown in Figs. 13.1, can be generalized for $(2^n*2^n)$-matrices produced by tensor powers of the initial unitary (2*2)-matrix and the doubly stochastic (2*2)-matrix (as it is known, tensor powers of unitary matrices generate new unitary matrices of increased orders). Tensor powers of (2*2)-matrices, which present complex numbers, produce $(2^n*2^n)$-matrices, which present algebraic extensions of complex numbers including many systems of multidimensional numbers. These algebraic extensions of the described complex numbers belong, first of all, to the field of algebraic biology and Gestalt genetics; the received results are additionally studied concerning general problems of theoretical physics.

## 14. Regarding the unified two-level scheme of dichotomy in genetics

In its beginning, the article described the binary-oppositional features of the DNA-alphabet of 4 nucleotides A, T, C, and G. They allowed constructing $(2^n*2^n)$-matrices (Fig. 2.1) used many times in the article for algebraic analysis of structural properties of the genomic DNA-texts. Reflecting these binary-oppositional features, Fig. 14.1 represents the DNA-alphabet of 4 nucleotides in a form of a two-level scheme of dichotomy.
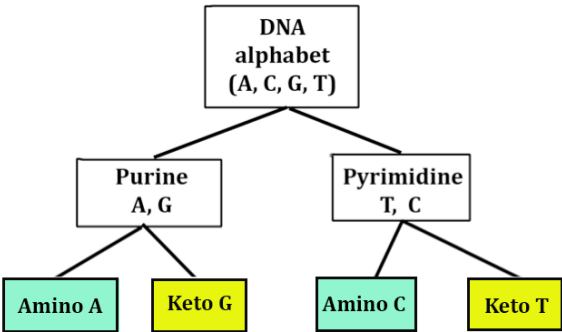
**Fig. 14.1**. The two-level scheme of dichotomy reflecting the two-level binary-oppositional structure of the DNA alphabet of 4 nucleotides A, C, G, and T.

The author notes that similar two-level schemes of dichotomy hold for Gestalt rules of percentage contents of the described different tetra-groupings of *n*-plets in different layers of various genomic DNA-texts. Figs. 14.2-14.4 show only some of many examples of such dichotomy for the described tetra-groupings of *n*-plets in the first layers of the DNA-text of the human chromosome №1.
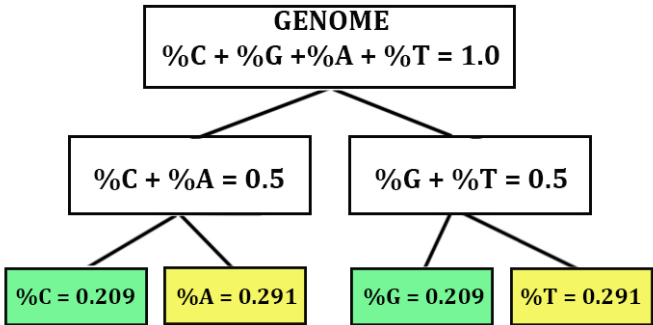


**Fig. 14.2**. The two-level scheme of dichotomy for the percentage content of 4 nucleotides in the first layer of the DNA-text of the human chromosome №1.
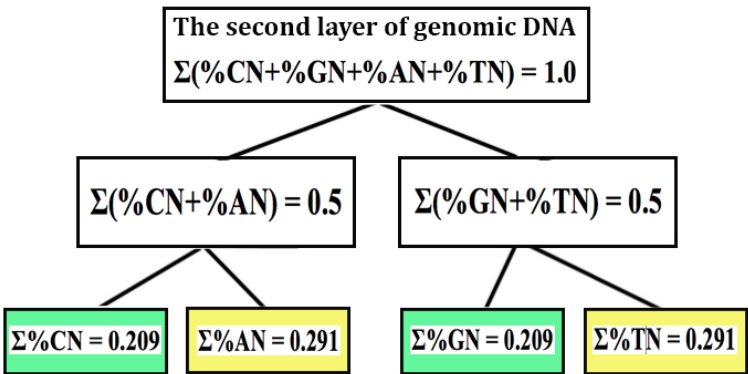


**Fig. 14.3**. The two-level scheme of dichotomy for the percentage content of 16 doublets of the first-positional tetra-grouping in the second layer of the DNA-text of the human chromosome №1.
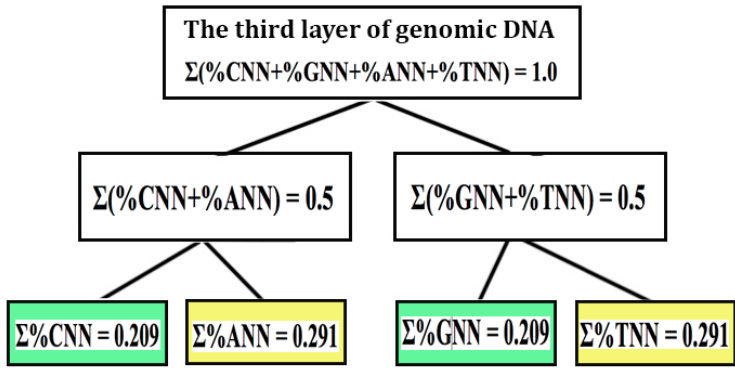
**Fig. 14.4**. The two-level scheme of dichotomy for the percentage content of
64 triplets of the first-positional tetra-grouping in the third layer of the
DNA-text of the human chromosome №1.

The author notes that when studying the structures of the genetic system with its binary oppositions at different levels, associations often arise with the concept of Yu.I. Kulakov on binary eidos (black and white, male and female eidos) as the basis of the laws of physics and geometry [Kulakov, 2004, 2018; Simonov, Kulakov, Vityaev, 2014]. This concept is presented in his theory of physical structures. The mathematical part of this theory is actually a universal algebraic theory of systems of real relations on one and two sets of elements. In the works [Vladimirov, 2017, 2019, 2020], the results of Kulakov and his colleagues were generalized to the case of complex relations, and the theory of binary systems of complex relations and the related binary geometrophysics were developed. Within the framework of this generalized theory, the problem of deriving classical space-time representations from an independent system of concepts and regularities inherent in the physics of the microworld is solved. Wherein the concept of binary pregeometry arises, concerning which the model concepts of space and time, traditionally used in physical theories, are secondary [Vladimirov, 2017, 2019,  2020]. It can be noted that molecular informational genetics is based on a certain system of binary relations, which can be considered as an analog of the named binary pregeometry. Developing an analogy of a binary-structured genetic system with Vladimirov's binary pregeometry can lead to formal models of biological time and space as secondary concepts arising from such a pregeometry of binary relations.

The above-described phenomenon of a systematic separation of *n*-plets in different layers of genomic DNA-texts into 4 groupings (by the Gestalt rules of the percentage content of *n*-plets) allows denoting all members of these 4 groupings correspondingly by digits 0, 1, 2, and 3. In this case, each of the long DNA sequences of *n*-plets in different DNA layers is transformed into a long sequence of digits 0, 1, 2, and 3, which can be represented by a special fractal-like pattern by the method of the Chaos Game Representation described in [Jeffrey, 1990] for long sequences of nucleotides. This approach produces a great number of corresponding fractal patterns for different layers of long DNA-texts of various biological species. A comparative analysis of sets of such fractal patterns gives interesting information for many problems of biological evolution. A corresponding construction of such fractal-like pictures for different genomes and long genes and their comparison analysis is conducted now in the author's laboratory for publication some later.

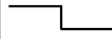## 15. Matrix presentations of alphabets of *n*-plets having strong or weak roots
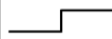
This section is devoted to one more kind of rule of the percentage of *n*-plets tetra-groupings in different layers of genomic DNA sequences.

One of the important types of binary-oppositional features in the genetic coding system was noted by Yu.Rumer [Rumer, 1968; Fimmel, Strüngmann, 2016; Petoukhov, 2008]. He described that the set of 64 triplets, according to their features of coding 20 amino acids, is divided into two subsets of 32 triplets in each:
- 32 triplets with "strong roots" (that is, with 8 doublets CC, CT, CG, AC, TC, GC, GT, GG at the beginning of triplets), whose code meanings do not depend on the letter in their third position;
- 32 triplets with "weak roots" (that is, with the remaining 8 doublets CA, AA, AT, AG, TA, TT, TG, GA at the beginning of triplets), whose code meanings depend on the letter in their third position.

Fig. 15.1 shows the beginning of the tensor family of matrices [C, A; T, G]$^{(n)}$ under *n* = 2, 3, which were used for representations of the alphabets of 16 doublets and 64 triplets in many of our previous works [Petoukhov, 2008; Petoukhov, He, 2010, etc]. The mentioned strong doublets and also triplets with strong roots are marked by black color in symbolic matrices in Fig. 15.1 by contrast to other doublets and triplets. In addition, the author also considers the genetic matrices [C, A; T, G]$^{(n)}$ under *n* = 4, 5, … for alphabets of 256 tetraplets, 1024 pentaplets,... with corresponding black-and-white mosaics, where *n*-plets with strong roots (that is, with 8 doublets CC, CT, CG, AC, TC, GC, GT, GG at their beginnings) are analogically marked by black color by contrast to other *n*-plets having weak roots. These mosaic matrices are conditionally termed «rooted mosaic matrices». An example of such rooted mosaic matrix in the case of the DNA alphabet of 256 tetraplets is shown in Fig. 15.1, bottom row.

|    | 11 | 10 | 01 | 00 |
|----|----|----|----|----|
| 11 | CC | CA | AC | AA |
| 10 | CT | CG | AT | AG |
| 01 | TC | TA | GC | GA |
| 00 | TT | TG | GT | GG |

|     | 111 | 110 | 101 | 100 | 011 | 010 | 001 | 000 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 111 | CCC | CCA | CAC | CAA | ACC | ACA | AAC | AAA |
| 110 | CCT | CCG | CAT | CAG | ACT | ACG | AAT | AAG |
| 101 | CTC | CTA | CGC | CGA | ATC | ATA | AGC | AGA |
| 100 | CTT | CTG | CGT | CGG | ATT | ATG | AGT | AGG |
| 011 | TCC | TCA | TAC | TAA | GCC | GCA | GAC | GAA |
| 010 | TCT | TCG | TAT | TAG | GCT | GCG | GAT | GAG |
| 001 | TTC | TTA | TGC | TGA | GTC | GTA | GGC | GGA |
| 000 | TTT | TTG | TGT | TGG | GTT | GTG | GGT | GGG |

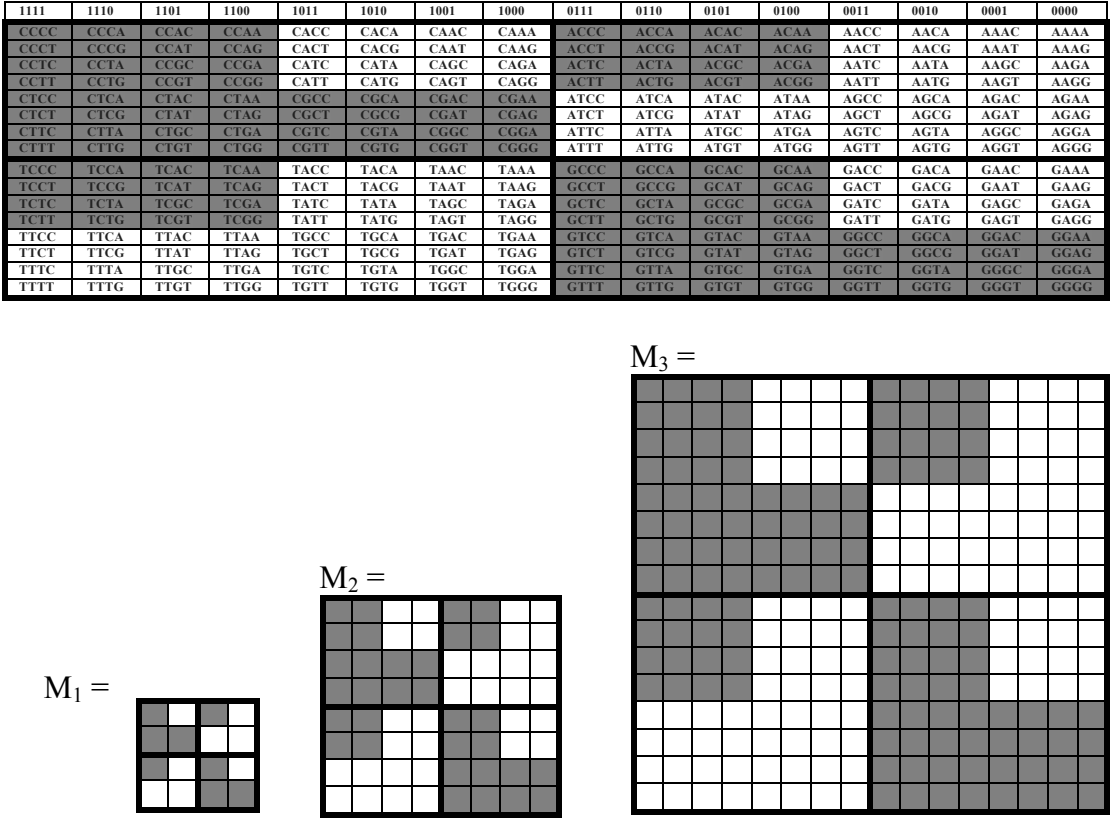| 1111 | 1110 | 1101 | 1100 | 1011 | 1010 | 1001 | 1000 | 0111 | 0110 | 0101 | 0100 | 0011 | 0010 | 0001 | 0000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CCCC | CCCA | CCAC | CCAA | CACC | CACA | CAAC | CAAA | ACCC | ACCA | ACAC | ACAA | AACC | AACA | AAAC | AAAA |
| CCCT | CCCG | CCAT | CCAG | CACT | CACG | CAAT | CAAG | ACCT | ACCG | ACAT | ACAG | AACT | AACG | AAAT | AAAG |
| CCTC | CCTA | CCGC | CCGA | CATC | CATA | CAGC | CAGA | ACTC | ACTA | ACGC | ACGA | AATC | AATA | AAGC | AAGA |
| CCTT | CCTG | CCGT | CCGG | CATT | CATG | CAGT | CAGG | ACTT | ACTG | ACGT | ACGG | AATT | AATG | AAGT | AAGG |
| CTCC | CTCA | CTAC | CTAA | CGCC | CGCA | CGAC | CGAA | ATCC | ATCA | ATAC | ATAA | AGCC | AGCA | AGAC | AGAA |
| CTCT | CTCG | CTAT | CTAG | CGCT | CGCG | CGAT | CGAG | ATCT | ATCG | ATAT | ATAG | AGCT | AGCG | AGAT | AGAG |
| CTTC | CTTA | CTGC | CTGA | CGTC | CGTA | CGGC | CGGA | ATTC | ATTA | ATGC | ATGA | AGTC | AGTA | AGGC | AGGA |
| CTTT | CTTG | CTGT | CTGG | CGTT | CGTG | CGGT | CGGG | ATTT | ATTG | ATGT | ATGG | AGTT | AGTG | AGGT | AGGG |
| TCCC | TCCA | TCAC | TCAA | TACC | TACA | TAAC | TAAA | GCCC | GCCA | GCAC | GCAA | GACC | GACA | GAAC | GAAA |
| TCCT | TCCG | TCAT | TCAG | TACT | TACG | TAAT | TAAG | GCCT | GCCG | GCAT | GCAG | GACT | GACG | GAAT | GAAG |
| TCTC | TCTA | TCGC | TCGA | TATC | TATA | TAGC | TAGA | GCTC | GCTA | GCGC | GCGA | GATC | GATA | GAGC | GAGA |
| TCTT | TCTG | TCGT | TCGG | TATT | TATG | TAGT | TAGG | GCTT | GCTG | GCGT | GCGG | GATT | GATG | GAGT | GAGG |
| TTCC | TTCA | TTAC | TTAA | TGCC | TGCA | TGAC | TGAA | GTCC | GTCA | GTAC | GTAA | GGCC | GGCA | GGAC | GGAA |
| TTCT | TTCG | TTAT | TTAG | TGCT | TGCG | TGAT | TGAG | GTCT | GTCG | GTAT | GTAG | GGCT | GGCG | GGAT | GGAG |
| TTTC | TTTA | TTGC | TTGA | TGTC | TGTA | TGGC | TGGA | GTTC | GTTA | GTGC | GTGA | GGTC | GGTA | GGGC | GGGA |
| TTTT | TTTG | TTGT | TTGG | TGTT | TGTG | TGGT | TGGG | GTTT | GTTG | GTGT | GTGG | GGTT | GGTG | GGGT | GGGG |



**Fig. 15.1**. **Three upper rows**: rooted mosaic matrices $[C, A; T, G]^{(2)}$, $[C, A; T, G]^{(3)}$, and $[C, A; T, G]^{(4)}$ for DNA alphabets of 16 doublets, 64 triplets, and 256 tetraplets. Black color denotes 8 strong doublets and also 32 triplets and 128 tetraplets with strong roots. For (8*8)-matrix of 64 triplets, graphics of Rademacher functions are shown additionally at the right to emphasize their analogical meander patterns. **Bottom row**: mosaic matrices $M_1$, $M_2$, $M_3$ where black (white) cells denote elements +1 (-1); each of the meander-like rows corresponds to an appropriate meander Rademacher function.

Dispositions of "black" and "white" doublets, triplets, and tetraplets in the matrices in Fig. 15.1 form the very symmetric black-and-white mosaics. For instance, left and right halves of each matrix are mirror-anti-symmetric to each other in their colors: any pair of cells, disposed of by the mirror-symmetrical manner in these halves, has opposite colors (correspondingly a mosaic of each row represents a complementary or inverse palindrome). Both quadrants along each of the diagonals are identical in their mosaics and – for this reason – these mosaic matrices remind a figure of the cross.

A sequence of blocks of black and white cells in each of the rows has a meander character (contains blocks of equal lengths) and corresponds to one of Rademacher functions [Petoukhov, 2008; Petoukhov, He, 2010]. This orthonormal system of functions is well known in the theory of discrete signals and expressed by the expression $r_n(t) = \text{sign}(\sin 2^n \pi t)$, $n = 1, 2, 3, \ldots$ ; the functions take only two values

"+1" and "−1" [https://encyclopediaofmath.org/wiki/Rademacher_system] (see the bottom row in Fig. 15.1). The system $r_n(x)$ is a typical example of a system of stochastically independent functions and has applications both in probability theory and in the theory of orthogonal series. Several theorems in probability theory can be formulated in terms of Rademacher functions. Rademacher functions are particular cases of Walsh functions, which are widely used in the algebraic encoding of information in communication technology and also in logical holography (see below about algebraic holography in modeling genetic informatics). The complete orthogonal system of Walsh functions, which is connected with Hadamard matrices in the theory of noise-immune coding and quantum computing, can be constructed as a product of Rademacher functions. Besides this, numeric matrices $M_1$, $M_2$, and $M_3$ (Fig. 15.1), whose mosaics coincide with phenomenological mosaics in non-symmetric genetic matrices $[C, A; T, G]^{(n)}$, are connected with oblique projectors mentioned above, since $(0.5*M_1)^2 = 0.5*M_1$, $(0.25*M_2)^2 = 0.25*M_2$, $(0.125*M_3)^2 = 0.125*M_3$. Below it will be shown that these mosaic matrices $M_1$, $M_2$, and $M_3$ are ensembles of unitary sparse matrices as well. On the whole, it means that the described phenomenology of DNA-alphabets is based on significant algebraic principles.

The author discovered that cooperative percentage organization of a set of different layers of genomic DNA nucleotide sequences is connected with this structural division of the $n$-plets alphabets into equal sub-alphabets of $n$-plets with strong roots and weak roots and also with meander patterns of Rademacher functions in rows of the genetic matrices in Fig. 15.1. To demonstrate this, let us initially calculate separately for left and right halves of each of rooted mosaic matrices in Fig. 15.1, a total percentage of $n$-plets in their black cells and their white cells for the case of the human chromosome №1, whose $n$-plets percentages in different DNA-texts are shown above in Figs. 2.3-2.5. Results of such calculation are shown in Figs. 15.2 and 15.3.

| LEFT HALVES OF GENETIC MATRICES $[C, A; T, G]^{(n)}$ | | |
|---|---|---|
| | Sums of percents for $n$-plets with strong roots **(black cells)** | Sums of percents for $n$-plets with weak roots **(white cells)** |
| Mosaic matrix of 16 doublets | %CC+%CT+%TC+%CG ≈ 0.05409+0.07134+0.06008+ 0.01031 ≈ **0.1958**. | %CA+%TT+%TA+%TG ≈ 0.07274+0.09568+0.06312+ 0.0728 ≈ **0.3044**. |
| Mosaic matrix of 64 tripets | %CCC+%CCT+%CTC+ %CTT+%TCC+%TCT+ %CCA+%CCG+%CTA+ %CTG+%TCA+%TCG+ %CGC+%CGT+%CGA+ %CGG ≈ 0.0138+0.0185+ 0.0176+0.0201+0.0159+ 0.0223+0.0188+0.0029+ 0.0127+0.0209+0.0196+ 0.0023+0.0025+0.0026+ 0.0023+0.0029 ≈ **0.1957**. | %CAC+%CAT+%CAA+ %CAG+%TAC+%TAT+ %TGC+%TGT+%TAA+ %TAG+%TGA+%TGG+ %TTC+%TTT+%TTA+ %TTG≈ 0.0152+0.0179+0.0186+0.0210+ 0.0110+0.0194+0.0146+0.0199+ 0.0199+0.0128+0.0195+0.0190+ 0.0197+0.0372+0.0198+  0.0188 ≈ **0.3043**. |
| Mosaic matrix of 256 tetraplets | The total sum of percentages of 64 tetraplets in black cells ≈ **0.1958**. | The total sum of percentages of 64 tetraplets in white cells ≈ **0.3044**. |

**Fig. 15.2**. Comparison of sums of the chromosomal percents of doublets, triplets, and tetraplets with strong and weak roots located in the left halves of the mosaic matrices in Fig. 15.1 for the case of the human chromosome # 1.

In line with phenomenological numeric data in Fig. 15.2, percentage sums of 4 doublets, 16 triplets, and 64 tetraplets with strong roots, which are located in black cells of the left halves of the mosaic matrices $[C, A; T, G]^{(n)}$ (Fig. 15.1), are approximately identical to each other with high accuracy ( $\approx$ **0.1958**). Simultaneously, percentage sums of 4 doublets, 16 triplets, and 64 tetraplets with weak roots, which are located in the left halves of these matrices, are also approximately identical to each other with high accuracy ($\approx$ **0.3044**).

| RIGHT HALVES OF GENETIC MATRICES $[C, A; T, G]^{(n)}$ | | |
|---|---|---|
| | Sums of percents for *n*-plets with strong roots **(black cells in matrices)** | Sums of percents for *n*-plets with weak roots **(white cells in matrices)** |
| Mosaic matrix of 16 doublets | %AC+%%GC+%GT+%GG≈ 0.05033+0.04402+0.05046+ 0.05419 ≈ **0.1990** | %AA+%AT+%AG+%GA = 0.09504+0.07429+0.07137+ 0.06008 ≈ **0.3008** |
| Mosaic matrix of 64 tripets | %ACC+%ACT+%ACA+ %ACG+%GCC+%GCT+ %GTC+%GTT+%GCA+ %GCG+%GTA+%GTG+ %GGC+%GGT+%GGA+ %GGG ≈ 0.0118+0.0162+ 0.0198+0.0025+0.0125+ 0.0144+0.0096+0.0145+ 0.0146+0.0025+0.0112+ 0.0153+0.0126+0.0119+ 0.0160+0.0138 ≈ **0.1992** | %ATC+%ATT+%ATA+ %ATG+AAC+%AAT+ %AGC+%AGT+%AAA+ %AAG+%AGA+%AGG+ %GAC+%GAT+%GAA+ %GAG ≈ 0.0132+0.0239+ 0.0194+0.0178+0.0145+ 0.0238+0.0144+0.0161+ 0.0369+0.0199+0.0224+ 0.0185+0.0096+0.0133+ 0.0196+0.0176 ≈ **0.3009** |
| Mosaic matrix of 256 tetraplets | The total sum of percentages of 64 tetraplets in black cells ≈ **0.1990** | The total sum of percentages of 64 tetraplets in white cells ≈ **0.3008** |

**Fig. 15.3**. Comparison of sums of the chromosomal percents of doublets, triplets, and tetraplets with strong and weak roots located in the right halves of the matrices in Fig. 15.1 for the case of the human chromosome # 1.

In line with phenomenological numeric data in Fig. 15.3, percentage sums of 4 doublets, 16 triplets, and 64 tetraplets with strong roots, which are located in black cells of the right halves of the mosaic matrices $[C, A; T, G]^{(n)}$ (Fig. 15.1), are approximately identical to each other with high accuracy ( $\approx$ **0.1990**). Simultaneously, percentage sums of 4 doublets, 16 triplets, and 64 tetraplets with weak roots, which are located in the right halves of these matrices, are also approximately identical to each other with high accuracy ($\approx$ **0.3008**).

Thus, the mysterious fact of dividing the DNA alphabets of *n*-plets into two equal sub-alphabets - with strong and weak roots - is associated with Gestalt genetics: four types of sub-alphabets of *n*-plets presented in Figs. 15.2 and 15.3 (two in the left half of the mosaic matrices and two in their right half) form a new kind of tetra-

groupings of *n*-plets, where the percentage sums in each of the four groupings are practically the same in different *n*-plets layers representing genomic DNA sequences.

Additional kinds of cooperative tetragroupings of *n*-plets in Gestalt genetics are revealed under analyzing genomic percentage sums of *n*-plets in quadrants along the main diagonal of the matrices (Fig. 15.1) and also in quadrants along the second diagonal of these matrices (in other words, the black-and-white mosaic along each of two diagonals has its own tetragrouping with constant percentage sums in its 4 parts). For example, in considering human chromosome № 1, the both quadrants along the main diagonal give the following practically invariant percentage sums of their n-plets in all the matrices:

- In the upper left quadrant, a percentage sum of *n*-plets in black cells is approximately equal to 0.1357, and a percentage sum of *n*-plets in white cells is approximately equal to 0.0727;
- In the lower right quadrant, a percentage sum of *n*-plets in black cells is approximately equal to 0.1486, and a percentage sum of *n*-plets in white cells is approximately equal to 0.0600.

Analogically, both quadrants along the second diagonal give the following practically invariant percentage sums of their *n*-plets in all the matrices for the considering human chromosome № 1:

- In the upper right quadrant, a percentage sum of *n*-plets in black cells is approximately equal to 0.0503, and a percentage sum of *n*-plets in white cells is approximately equal to 0,2407;
- In the lower-left quadrant, a percentage sum of *n*-plets in black cells is approximately equal to 0.0601, and a percentage sum of *n*-plets in white cells is approximately equal to 0.2317.

Now let us turn to one more phenomenological interrelation of different rooted mosaic matrices $[C, A; T, G]^{(n)}$ in Fig. 15.1, which is based on percentage sums of *n*-plets with strong and weak roots in their different rows, whose mosaics correspond to Rademacher-Walsh functions. The analysis carried out by the author revealed genomic percentage interrelations of different rows from these matrices. For example, in the case of the considered human chromosome №1 using percentage data in Figs. 2.3-2.5, for rows of matrices in Fig. 15.1, we have the following interrelations between percentage sums of *n*-plets with strong roots and also for *n*-plets with weak roots:

- The percentage sum of each of two oppositional subsets of doublets in row 1 (upper row in the matrix of 16 doublets in Fig. 15.1) is approximately equal to the percentage sum of corresponding subset of triplets in rows 1 and 2 (two upper rows in the matrix of 64 triplets in Fig. 15.1), and also is approximately equal to the percentage sum of the corresponding subset of tetraplets in rows 1-4 (four upper rows in the matrix of 256 tetraplets in Fig. 15.1). This numeric result is shown in Fig. 15.4;
- The percentage sum of each of two oppositional subsets of doublets in row 2 of the matrix of 16 doublets in Fig. 15.1 is approximately equal to the percentage sum of the corresponding subset of triplets in rows 3 and 4 in the matrix of 64 triplets in Fig. 15.1, and also is approximately equal to the percentage sum of the corresponding subset of tetraplets in rows 5-8 of the matrix of 256 tetraplets in Fig. 15.1. This numeric result is shown in Fig. 15.5;
- The percentage sum of each of two oppositional subsets of doublets in row 3 of the matrix of 16 doublets in Fig. 15.1 is approximately equal to the

percentage sum of the corresponding subset of triplets in rows 5 and 6 in the matrix of 64 triplets in Fig. 15.1, and also is approximately equal to the percentage sum of the corresponding subset of tetraplets in rows 9-12 of the matrix of 256 tetraplets in Fig. 15.1. This numeric result is shown in Fig. 15.6;

- The percentage sum of each of two oppositional subsets of doublets in row 4 of the matrix of 16 doublets in Fig. 15.1 is approximately equal to the percentage sum of the corresponding subset of triplets in rows 7 and 8 in the matrix of 64 triplets in Fig. 15.1, and also is approximately equal to the percentage sum of the corresponding subset of tetraplets in rows 13-16 of the matrix of 256 tetraplets in Fig. 15.1. This numeric result is shown in Fig. 15.7.

| | Σ% of *n*-plets with strong roots | Σ% of *n*-plets with weak roots |
|---|---|---|
| **Row 1 in matrix of 16 doublets** | %CC+%AC ≈ 0.05409+0.05033 ≈ **0.1044**. | %CA+%AA ≈ 0.07274+0.09504 ≈ **0.1678**. |
| **Rows 1-2 in matrix of 64 triplets** | %CCC+%CCA+%ACC+%AC+ %CCT+%CCG+%ACT+%ACG ≈ 0.0138+0.0188+0.0118 +0.0198+0.0185+0.0029+0.0162 +0.0025 ≈ **0.1043**. | %CAC+%CAA+%AAC+%AAA +%CAT+%CAG+%AAT+%AAG ≈ 0.0152+0.0186+0.0145+0.0369 +0.0179+0.0210+0.0238+0.0199 ≈ **0.1678**. |
| **Rows 1-4 in matrix of 256 tetraplets** | Percentage sum of 32 tetraplets with strong roots ≈ **0.1044**. | Percentage sum of 32 tetraplets with weak roots ≈ **0.1678**. |

**Fig. 15.4**. The percentage sums of those *n*-plets with strong and weak roots, which are located correspondingly in black and white matrix cells (Fig. 15.1) in named rows for the DNA of the human chromosome № 1. Corresponding percent values for separate *n*-plets are shown in Figs. 2.3-2.5.

| | Σ% of *n*-plets with strong roots | Σ% of *n*-plets with weak roots |
|---|---|---|
| **Row 2 in matrix of 16 doublets** | %CT+%CG ≈ 0.07134+0.01031 ≈ **0.0817**. | %AT+%AG ≈ 0.07429+0.07137 ≈ **0.1457**. |
| **Rows 3-4 in matrix of 64 triplets** | %CTC+%CTA+%CGC+%CGA +%CTT+%CTG+%CGT+%CGG ≈  0.0176+0.0127+0.0025+0.0023 +0.0201+0.0209+0.0026+0.0029  ≈ **0.0816**. | %ATC+%ATA+%AGC+%AGA +%ATT+%ATG+%AGT+%AGG ≈ 0.0132+0.0194+0.0144+0.0224 +0.0239+0.0178+0.0161+0.0185 ≈ **0.1457**. |
| **Rows 5-8 in matrix of 256 tetraplets** | Percentage sum of 32 tetraplets with strong roots ≈ **0.0816**. | Percentage sum of 32 tetraplets with weak roots ≈ **0.1456**. |

**Fig. 15.5**. The percentage sums of those *n*-plets with strong and weak roots, which are located correspondingly in black and white matrix cells (Fig. 15.1) in named rows for the DNA of the human chromosome № 1. Corresponding percent values for separate *n*-plets are shown in Figs. 2.3-2.5.

| | Σ% of *n*-plets with strong roots | Σ% of *n*-plets with weak roots |
|---|---|---|
| **Row 3 in matrix of 16 doublets** | %TC+%GC ≈ 0.06008+0.04402 ≈ **0.1041**. | %TA+%GA ≈ 0.06312+0.06008 ≈ **0.1232**. |
| **Rows 5-6 in matrix of 64 triplets** | %TCC+%TCA+%GCC+%GCA +%TCT+%TCG+%GCT+%GCG ≈0.0159+0.0196+0.0125+0.0146 +0.0223+0.0023+0.0144+0.0025 ≈ **0.1041**. | %TAC+%TAA+%GAC+%GAA+ %TAT+%TAG+%%GAT+%GAG ≈0.0110+0.0199+0.0096+0.0196 +0.0194+0.0128+0.0133+0.0176 ≈ **0.1232**. |
| **Rows 9-12 in matrix of 256 tetraplets** | Percentage sum of 32 tetraplets with strong roots ≈ **0.1041**. | Percentage sum of 32 tetraplets with strong roots ≈ **0.1232**. |

**Fig. 15.6**. The percentage sums of those *n*-plets with strong and weak roots, which are located correspondingly in black and white matrix cells (Fig. 15.1) in named rows for the DNA of the human chromosome № 1. Corresponding percent values for separate *n*-plets are shown in Figs. 2.3-2.5.

| | Σ% of *n*-plets with strong roots | Σ% of *n*-plets with weak roots |
|---|---|---|
| **Row №4 in matrix of 16 doublets** | %GT+%GG ≈ 0.05046+0.05419 ≈ **0.1047**. | %%TT+%TG ≈ 0.09568+0.07286 ≈ **0.1685**. |
| **Rows 7-8 in matrix of 64 triplets** | %GTC+%GTA+%GGC+%GGA +%GTT+%GTG+%GGT+%GGG ≈ 0.0096+0.0112+0.0126+0.0160 0.0145+0.0153+0.0119+0.0138 ≈ **0.1049**. | %TTC+%TTA+%TGC+%TGA +%TTT+%TTG+%TGT+%TGG ≈ 0.0197+0.0198+0.0146+0.0195 +0.0372+0.0188+0.0199+0.0190 ≈ **0.1685**. |
| **Rows 13-16 in matrix of 256 tetraplets** | Percentage sum of 32 tetraplets with strong roots ≈ **0.1047**. | Percentage sum of 32 tetraplets with weak roots ≈ **0.1685**. |

**Fig. 15.7**. The percentage sums of those *n*-plets with strong and weak roots, which are located correspondingly in black and white matrix cells (Fig. 15.1) in named rows for the DNA of the human chromosome № 1. Corresponding percent values for separate *n*-plets are shown in Figs. 2.3-2.5.

The above described Gestalt-genetic phenomena, which are connected with the tensor family of alphabetical mosaic matrices $[C, A; T, G]^{(n)}$ (Fig. 15.1) seem to be additionally significant taking into account the following. As previously described in the author's publications (for example, in [Petoukhov, 2008; Petoukhov, He, 2010] and many articles presented at the personal website http://petoukhov.com/), this tensor family of mosaic matrices has algebraic properties that are important for the development of algebraic biology. In particular, this tensor family is associated with Walsh functions, Hadamard matrices, projection operators, Cockle's split-quaternions, cyclic groups based on sums of oblique projectors, gene-logical coding, logical holography, etc. All these properties are now associated with the author's concepts of Gestalt genetics, algebraic-holographic genetics, and statistical coding of genetic information like arithmetic coding known in the theory of information compression.

One should additionally remind about one more connection of DNA alphabets of *n*-plets with complete orthogonal sets of Walsh functions, which was noted in the author's publications [Petoukhov, 2008, Chapter 3; Petoukhov, He, 2010, Chapter 6]. The molecular composition of the letters A, C, G, T of the genetic alphabet is characterized by a binary-oppositional separation: the three nitrogenous bases A, C, G have one amides (amino-group) $NH_2$, but the fourth basis T has not it (Fig. 15.8); this amino-group of amino acids bears an important function to provide recognition of an amino acid by enzyme [Shapeville, Haenni, 1974]. Taking this binary-opposition into account, the following numeric representation of nucleotides arises: A = C = G = +1, T = -1.

This binary-oppositional numeric representation of the alphabet of 4 nucleotides is also supported by another known fact of the very special status of nucleotide T: only nucleotide T is replaced by another nucleotide U (uracil) at the transition from DNA to RNA for an unknown reason.

This numeric representation A = C = G = +1, T = -1 transforms the tensor family of symbolic matrices $[C, A; T, G]^{(n)}$ in Fig. 2.1 into the tensor family of numeric matrices $[1, 1; -1, 1]^{(n)}$, whose sets of rows represent appropriate complete orthogonal systems of Walsh functions for $2^n$-dimensional spaces. The matrices of the type $[1, 1; -1, 1]^{(n)}$ or $[1, 1; 1, -1]^{(n)}$ are frequently termed Walsh-Hadamard matrices, and transformations of appropriate vectors by them are termed Walsh-Hadamard transformations. For example, Walsh-Hadamard matrices are actively used in noise-immunity coding of information, including spacecraft "Mariner" and "Voyager". These spacecraft transmit to Earth pictures of different planets by electromagnetic signals, which are strongly deformed under pass through millions of kilometers of noises. These matrices and transformations are used below in the next section.
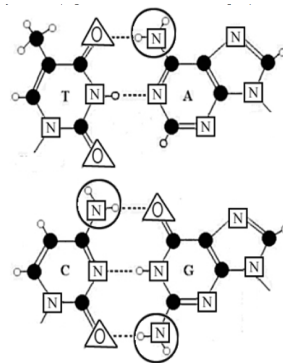


**Fig. 15.8**. The complementary pairs of the four nitrogenous bases in DNA. A-T (adenine and thymine), C-G (cytosine and guanine). Hydrogen bonds in these pairs are shown by dotted lines. Black circles are atoms of carbon; small white circles are atoms of hydrogen; squares with the letter N are atoms of nitrogen; triangles with the letter O are atoms of oxygen. Amides (or amino-groups) $NH_2$ are marked by big circles.

The detection of natural realization of Hadamard matrices (and orthogonal systems of Walsh functions, which are connected with them) based on structures of the molecular-genetic system gives pieces of evidence for the following: all known advantages of Walsh functions and Hadamard matrices can be utilized in bioinformatics and its model approaches including the algebraic-logical holography approach noted below.

**16. Algebra-logical holography and genetic informatics.**

Living organisms possess inherited properties, which seem to be analogical to properties of holography with its non-local record of information. For example, German embryologist Hans Driesch [Driesch, 1921] separated from each other two or four blastomeres of sea urchin eggs or on changing their mutual positions. The main result of Driesch's experiments was that fairly normal (although proportionally diminished) larvae with all of their organs properly arranged could be obtained from a single embryonic cell (blastomere) containing no more than 1∕2 (if two first blastomeres were separated) or even 1∕4 (in the case of four blastomeres separation) of the entire egg's material. Rather soon these effects (defined by Driesch as "embryonic regulations") were numerously confirmed and extended to the species belonging to almost all taxonomic groups of metazoans, from sponges to mammalians [Belousov, 2015; Driesch, 1921]. In 1901 Hans Spemann has conducted an experiment on the separation of the amphibian embryo into individual cells, from which quite normal tadpoles have grown (in 1935 he won the Nobel Prize for the discovery of organizing effects in embryonic development). These experimental results testify that complete sets of "causes" required for further development are contained not only within whole eggs/embryos but also in their halves, quarters, etc. Similar properties exist in holograms, where one can restore a whole holographic image of a material object from a part of the hologram. A hologram has such property since each element of the hologram possesses information about all elements of the object (by contrast to ordinary photos).

Physical holography discovered by D. Gabor provided science and technology with surprisingly effective means of storing and processing information based on its following properties:
- Holographic memory is distributed over the entire volume (area) of the carrier;
- Reducing the volume of the medium only worsens the detail of the image, but does not destroy its parts;
- The possibility of many high-quality images on one medium by changing the angle of the laser beam during their recording and recreation (in contrast to ordinary photography, in which shooting different images at the same frame of film spoils each of them);
- all information is recorded on a flat 2-dimensional plate, however, a three-dimensional volumetric image is recreated on its basis.

These amazing properties of physical holography have given rise to many new approaches in biology and physics, many of which are associated with them. For example, in biology, the concept of the connection of brain functions with the properties of holography is widely known, presented in Pribram's book "Languages of the Brain" [Pribram, 1971]. This book emphasizes that holographic description is unmatched for explaining problems of perception, especially problems of image formation and fantastic ability to recognize. There is a lot of physiological data to support this concept. For example, Lashley [Lashley, 1929] defined the following: no matter which part of the rat's brain we remove, we cannot destroy the memories of how to perform the complex actions that the rat learned before surgery. Many physiological phenomena, the properties of which resemble holographic ones, are described, for example, in the book "The Holographic Universe" [Talbot, 1996].

But the brain and the nervous system have appeared at a relatively late stage of biological evolution. A great number of species of organisms lived perfectly up to

this and are now living without neuronal networks. It is also evident that mentioned holographic-like phenomena of embryology also cannot be reduced to the work of the nervous system. The origins of the similarity between holography and nonlocal informatics of living organisms need to look at the level of the genetic system.

Physical holography, which possesses, in particular, the highest properties of noise-immunity, is based on a record of standing waves from two coherent physical waves of the object beam and of the reference beam. Its impressive properties provoked wishes in the scientific community to create such algebraic and logical methods of coding and processing of digital information sequences and images, which would have some similar properties without using photo materials, laser beams, etc. As a result, the following three algebraic methods of holography-like processing of digital sequences were proposed by different authors:
- 1) Bit-reversal holography (or bit-reversal permutation);
- 2) Holographic coding by Walsh–Hadamard transformation of randomized and permuted data;
- 3) Logical holography by Walsh functions.

The author revealed that structures of genetic DNA informatics correspond to key mathematical aspects of these three algebra-holographic methods, which can be used jointly for the development of algebraic biology and a profound understanding of genetic phenomena. This allows creating new algebra-holographic approaches in genetics and also such genetic-like versions of artificial intelligence, which imitate genetic structural properties; it gives pieces of evidence in favor of the author's conception as well that inherited holographic-like properties of brain functions are based on appropriate algebra-holographic properties of genetic informatics. In this section, the author's data are presented in favor of the algebraic-holographic nature of genetic informatics. Descriptions of each of the noted three methods are briefly given below jointly with some of their modeling applications to biological phenomena.

## 16.1. Relations of genetic DNA alphabets to bit-reversal holography

In digital signal processing, the method of bit-reversal permutations of the binary numbering of columns and rows in square matrices, which sometimes is termed bit-reversal holography, plays an important role; the method is connected, in particularly, noise-immunity coding and algorithms of fast Fourier transform [Gold and Rader, 1969; Karp, 1996; Lyons, 2010; Shiman, Patsey, 2013; Shishmintsev, 2012; Yang at al., 2013]. In the case of two-dimensional fast Fourier transform - for example, in image processing - algorithms with interleaving in frequency are considered optimal [Lyons, 2010].

The method of bit-reversal permutations is based on a backward order of reading of binary numbering of columns and rows (001 becomes 100, etc.) in the square $(2^n*2^n)$-matrices with appropriate permutations of columns and rows. For example, under bit-reversal permutations, the binary-numeric sequence 111, 110, 101, 100, 011, 010, 001, 000 (its decimal analog is the sequence 7, 6, 5, 4, 3, 2, 1) is transformed to the binary sequence 111, 011, 101, 001, 110, 010, 100, 000 (its decimal analog is 7, 3, 5, 1, 6, 2, 4, 0, where all odd numbers are collected inside the left half of the sequence and all even numbers – in the right half). Such renumbering of the columns and rows of the matrix leads to a corresponding permutation of its cells and the appearance of a new matrix. Applying a bit-reversal permutation to the binary numbering of the columns and rows of this new matrix returns the original matrix. One should explain why the method of bit-reversal permutations is termed

algebra-holographic.

   For the explanation, we use an example from the work [Shishmintsev, 2012] showing a square matrix of 512*512 pixels with a large letter A drawn in it (Fig. 16.1.1a-d). Then a backward reading of the binary numbering of columns and rows of pixels in this matrix shuffles columns and rows, transforming the whole matrix image. If a part of the image in this new matrix is removed by coloring (Fig. 16.1.1c), then the second backward reading restores the original image of the image "A" largely where all defects occur to be dispersed and do not interfere with the recognition of the original image (a quasi-holographic property).



a)      b)      c)      d)
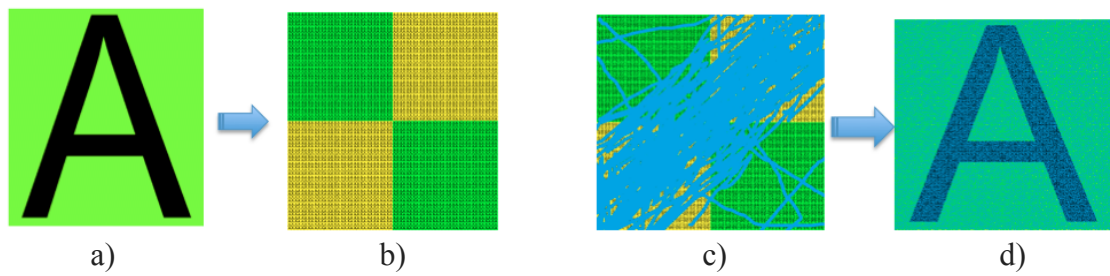
**Fig. 16.1.1**. An example of holographic-like properties related with the method of bit-reversal permutations (from [Shishmintsev, 2012]. a) an initial image of the letter A inside a matrix M having 512*512 pixels with appropriate 9-bit numerations of its columns and rows; b) the result of the transformation of the matrix M to a new matrix $M_{reverse}$ by bit-reversal permutations inside numerations of its columns and rows; c) distortion of a significant part of the pixels in the matrix $M_{reverse}$; d) the result of applying a bit-reverse permutation to the numbering of columns and rows of a distorted matrix $M_{reverse}$: a recognizable initial image of the letter A is restored.

   Fig. 16.1.2a-d shows another example of holographic-like properties of the method of bit-reversal permutations (or bit-reversal interleaving) with image restoring even at 50% loss (from [Shiman, Patsey, 2013]).



a)      b)      c)      d)
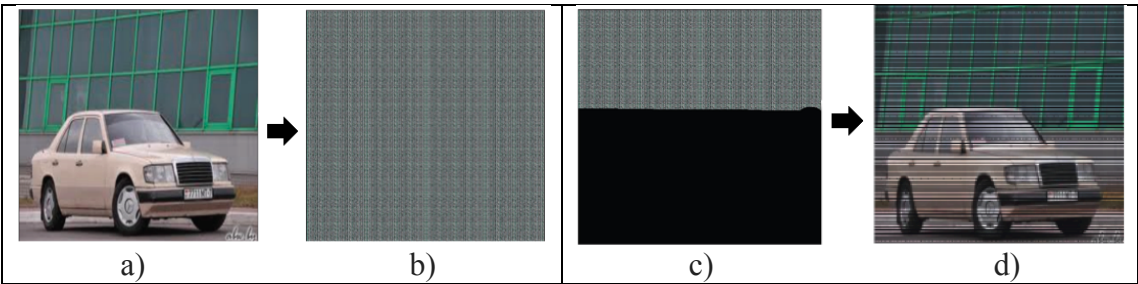
**Fig. 16.1.2**. An example of holographic-like properties related to the method of bit-reversal permutations (from [Shiman, Patsey, 2013]). a) an initial image inside a matrix M having 512*512 pixels with appropriate 9-bit numerations of its columns and rows; b) the result of the transformation of the matrix M to a new matrix $M_{reverse}$ by bit-reversal permutations inside numerations of its columns and rows; c) distortion of 50% of the matrix $M_{reverse}$; d) the result of restoring the initial image by a bit-reverse

permutation inside the numbering of columns and rows of a distorted matrix $M_{reverse}$.

The author's application of the holographic method of bit-reversal permutations to alphabetic mosaic matrices from their tensor family $[C, A; T, G]^{(n)}$, whose first members are shown above in Fig. 15.1, has given interesting results illustrated by Fig. 16.1.3.
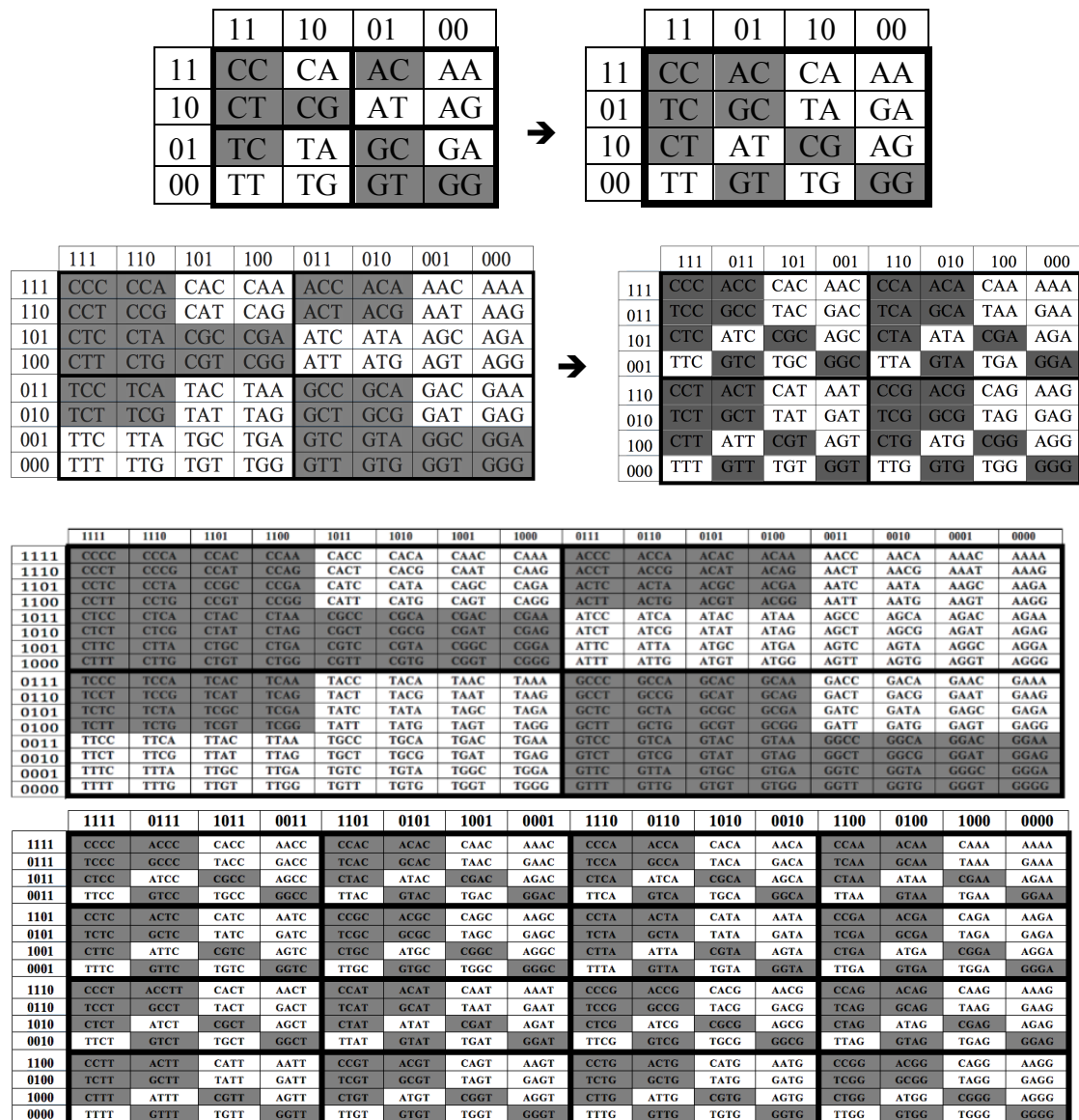
|    | 11 | 10 | 01 | 00 |
|----|----|----|----|----|
| 11 | CC | CA | AC | AA |
| 10 | CT | CG | AT | AG |
| 01 | TC | TA | GC | GA |
| 00 | TT | TG | GT | GG |

→

|    | 11 | 01 | 10 | 00 |
|----|----|----|----|----|
| 11 | CC | AC | CA | AA |
| 01 | TC | GC | TA | GA |
| 10 | CT | AT | CG | AG |
| 00 | TT | GT | TG | GG |

|     | 111 | 110 | 101 | 100 | 011 | 010 | 001 | 000 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 111 | CCC | CCA | CAC | CAA | ACC | ACA | AAC | AAA |
| 110 | CCT | CCG | CAT | CAG | ACT | ACG | AAT | AAG |
| 101 | CTC | CTA | CGC | CGA | ATC | ATA | AGC | AGA |
| 100 | CTT | CTG | CGT | CGG | ATT | ATG | AGT | AGG |
| 011 | TCC | TCA | TAC | TAA | GCC | GCA | GAC | GAA |
| 010 | TCT | TCG | TAT | TAG | GCT | GCG | GAT | GAG |
| 001 | TTC | TTA | TGC | TGA | GTC | GTA | GGC | GGA |
| 000 | TTT | TTG | TGT | TGG | GTT | GTG | GGT | GGG |

→

|     | 111 | 011 | 101 | 001 | 110 | 010 | 100 | 000 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 111 | CCC | ACC | CAC | AAC | CCA | ACA | CAA | AAA |
| 011 | TCC | GCC | TAC | GAC | TCA | GCA | TAA | GAA |
| 101 | CTC | ATC | CGC | AGC | CTA | ATA | CGA | AGA |
| 001 | TTC | GTC | TGC | GGC | TTA | GTA | TGA | GGA |
| 110 | CCT | ACT | CAT | AAT | CCG | ACG | CAG | AAG |
| 010 | TCT | GCT | TAT | GAT | TCG | GCG | TAG | GAG |
| 100 | CTT | ATT | CGT | AGT | CTG | ATG | CGG | AGG |
| 000 | TTT | GTT | TGT | GGT | TTG | GTG | TGG | GGG |

|      | 1111 | 1110 | 1101 | 1100 | 1011 | 1010 | 1001 | 1000 | 0111 | 0110 | 0101 | 0100 | 0011 | 0010 | 0001 | 0000 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1111 | CCCC | CCCA | CCAC | CCAA | CACC | CACA | CAAC | CAAA | ACCC | ACCA | ACAC | ACAA | AACC | AACA | AAAC | AAAA |
| 1110 | CCCT | CCCG | CCAT | CCAG | CACT | CACG | CAAT | CAAG | ACCT | ACCG | ACAT | ACAG | AACT | AACG | AAAT | AAAG |
| 1101 | CCTC | CCTA | CCGC | CCGA | CATC | CATA | CAGC | CAGA | ACTC | ACTA | ACGC | ACGA | AATC | AATA | AAGC | AAGA |
| 1100 | CCTT | CCTG | CCGT | CCGG | CATT | CATG | CAGT | CAGG | ACTT | ACTG | ACGT | ACGG | AATT | AATG | AAGT | AAGG |
| 1011 | CTCC | CTCA | CTAC | CTAA | CGCC | CGCA | CGAC | CGAA | ATCC | ATCA | ATAC | ATAA | AGCC | AGCA | AGAC | AGAA |
| 1010 | CTCT | CTCG | CTAT | CTAG | CGCT | CGCG | CGAT | CGAG | ATCT | ATCG | ATAT | ATAG | AGCT | AGCG | AGAT | AGAG |
| 1001 | CTTC | CTTA | CTGC | CTGA | CGTC | CGTA | CGGC | CGGA | ATTC | ATTA | ATGC | ATGA | AGTC | AGTA | AGGC | AGGA |
| 1000 | CTTT | CTTG | CTGT | CTGG | CGTT | CGTG | CGGT | CGGG | ATTT | ATTG | ATGT | ATGG | AGTT | AGTG | AGGT | AGGG |
| 0111 | TCCC | TCCA | TCAC | TCAA | TACC | TACA | TAAC | TAAA | GCCC | GCCA | GCAC | GCAA | GACC | GACA | GAAC | GAAA |
| 0110 | TCCT | TCCG | TCAT | TCAG | TACT | TACG | TAAT | TAAG | GCCT | GCCG | GCAT | GCAG | GACT | GACG | GAAT | GAAG |
| 0101 | TCTC | TCTA | TCGC | TCGA | TATC | TATA | TAGC | TAGA | GCTC | GCTA | GCGC | GCGA | GATC | GATA | GAGC | GAGA |
| 0100 | TCTT | TCTG | TCGT | TCGG | TATT | TATG | TAGT | TAGG | GCTT | GCTG | GCGT | GCGG | GATT | GATG | GAGT | GAGG |
| 0011 | TTCC | TTCA | TTAC | TTAA | TGCC | TGCA | TGAC | TGAA | GTCC | GTCA | GTAC | GTAA | GGCC | GGCA | GGAC | GGAA |
| 0010 | TTCT | TTCG | TTAT | TTAG | TGCT | TGCG | TGAT | TGAG | GTCT | GTCG | GTAT | GTAG | GGCT | GGCG | GGAT | GGAG |
| 0001 | TTTC | TTTA | TTGC | TTGA | TGTC | TGTA | TGGC | TGGA | GTTC | GTTA | GTGC | GTGA | GGTC | GGTA | GGGC | GGGA |
| 0000 | TTTT | TTTG | TTGT | TTGG | TGTT | TGTG | TGGT | TGGG | GTTT | GTTG | GTGT | GTGG | GGTT | GGTG | GGGT | GGGG |

→

|      | 1111 | 0111 | 1011 | 0011 | 1101 | 0101 | 1001 | 0001 | 1110 | 0110 | 1010 | 0010 | 1100 | 0100 | 1000 | 0000 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1111 | CCCC | ACCC | CACC | AACC | CCAC | ACAC | CAAC | AAAC | CCCA | ACCA | CACA | AACA | CCAA | ACAA | CAAA | AAAA |
| 0111 | TCCC | GCCC | TACC | GACC | TCAC | GCAC | TAAC | GAAC | TCCA | GCCA | TACA | GACA | TCAA | GCAA | TAAA | GAAA |
| 1011 | CTCC | ATCC | CGCC | AGCC | CTAC | ATAC | CGAC | AGAC | CTCA | ATCA | CGCA | AGCA | CTAA | ATAA | CGAA | AGAA |
| 0011 | TTCC | GTCC | TGCC | GGCC | TTAC | GTAC | TGAC | GGAC | TTCA | GTCA | TGCA | GGCA | TTAA | GTAA | TGAA | GGAA |
| 1101 | CCTC | ACTC | CATC | AATC | CCGC | ACGC | CAGC | AAGC | CCTA | ACTA | CATA | AATA | CCGA | ACGA | CAGA | AAGA |
| 0101 | TCTC | GCTC | TATC | GATC | TCGC | GCGC | TAGC | GAGC | TCTA | GCTA | TATA | GATA | TCGA | GCGA | TAGA | GAGA |
| 1001 | CTTC | ATTC | CGTC | AGTC | CTGC | ATGC | CGGC | AGGC | CTTA | ATTA | CGTA | AGTA | CTGA | ATGA | CGGA | AGGA |
| 0001 | TTTC | GTTC | TGTC | GGTC | TTGC | GTGC | TGGC | GGGC | TTTA | GTTA | TGTA | GGTA | TTGA | GTGA | TGGA | GGGA |
| 1110 | CCCT | ACCT | CACT | AACT | CCAT | ACAT | CAAT | AAAT | CCCG | ACCG | CACG | AACG | CCAG | ACAG | CAAG | AAAG |
| 0110 | TCCT | GCCT | TACT | GACT | TCAT | GCAT | TAAT | GAAT | TCCG | GCCG | TACG | GACG | TCAG | GCAG | TAAG | GAAG |
| 1010 | CTCT | ATCT | CGCT | AGCT | CTAT | ATAT | CGAT | AGAT | CTCG | ATCG | CGCG | AGCG | CTAG | ATAG | CGAG | AGAG |
| 0010 | TTCT | GTCT | TGCT | GGCT | TTAT | GTAT | TGAT | GGAT | TTCG | GTCG | TGCG | GGCG | TTAG | GTAG | TGAG | GGAG |
| 1100 | CCTT | ACTT | CATT | AATT | CCGT | ACGT | CAGT | AAGT | CCTG | ACTG | CATG | AATG | CCGG | ACGG | CAGG | AAGG |
| 0100 | TCTT | GCTT | TATT | GATT | TCGT | GCGT | TAGT | GAGT | TCTG | GCTG | TATG | GATG | TCGG | GCGG | TAGG | GAGG |
| 1000 | CTTT | ATTT | CGTT | AGTT | CTGT | ATGT | CGGT | AGGT | CTTG | ATTG | CGTG | AGTG | CTGG | ATGG | CGGG | AGGG |
| 0000 | TTTT | GTTT | TGTT | GGTT | TTGT | GTGT | TGGT | GGGT | TTTG | GTTG | TGTG | GGTG | TTGG | GTGG | TGGG | GGGG |

**Fig. 16.1.3**. Transformations of the alphabetic mosaic matrices $[C, A; T, G]^{(n)}$ from Fig. 15.1 into new alphabetic mosaic matrices $[C, A; T, G]^{(n)}_{bit\text{-}reversal}$ by the holographic method of bit-reversal permutations inside binary numeration of their columns and rows.

The first interesting result of such bit-reversal transformations of genetic matrices from the tensor family $[C, A; T, G]^{(n)}$ is that all received mosaic matrices $[C, A; T, G]^{(n)}_{bit\text{-}reversal}$ under $n = 3, 4, \ldots$ (Fig. 16.1.3) are block matrices: the family of their mosaics is tensor colonies or crystal-like unions of the same mosaic (4*4)-block

representing the bit-reversal mosaic (4*4)-matrix of the DNA alphabet of doublets (Fig. 16.1.3, on the right in the top row). This fact additionally emphasizes an algebraic nature of the system of structured DNA and RNA alphabets and its close connection with the bit-reversal holography. The author conditionally termes these matrix crystal-like colonies of the same mosaic (4*4)-block as "matrix crystals" of DNA- and RNA- alphabets of $n$-plets. The alphabetical matrix crystals can be briefly termed α-crystals (using that α is the first letter of the Greek alphabet and that the Latin letter A and the Cyrillic A, which are the first letters in corresponding alphabets, were originated from the letter α). Fig. 16.1.4 illustrates this mosaic block interrelation of bit-reversed matrices of many structured DNA-alphabets in the numeric form. The numeric (4*4)-matrix [1, 1, -1, -1; 1, 1, -1, -1; 1, -1, 1, -1; -1, 1, -1, 1], which plays a role of general cell in the numeric matrix crystals in Fig. 16.1.4, we term α-matrix.
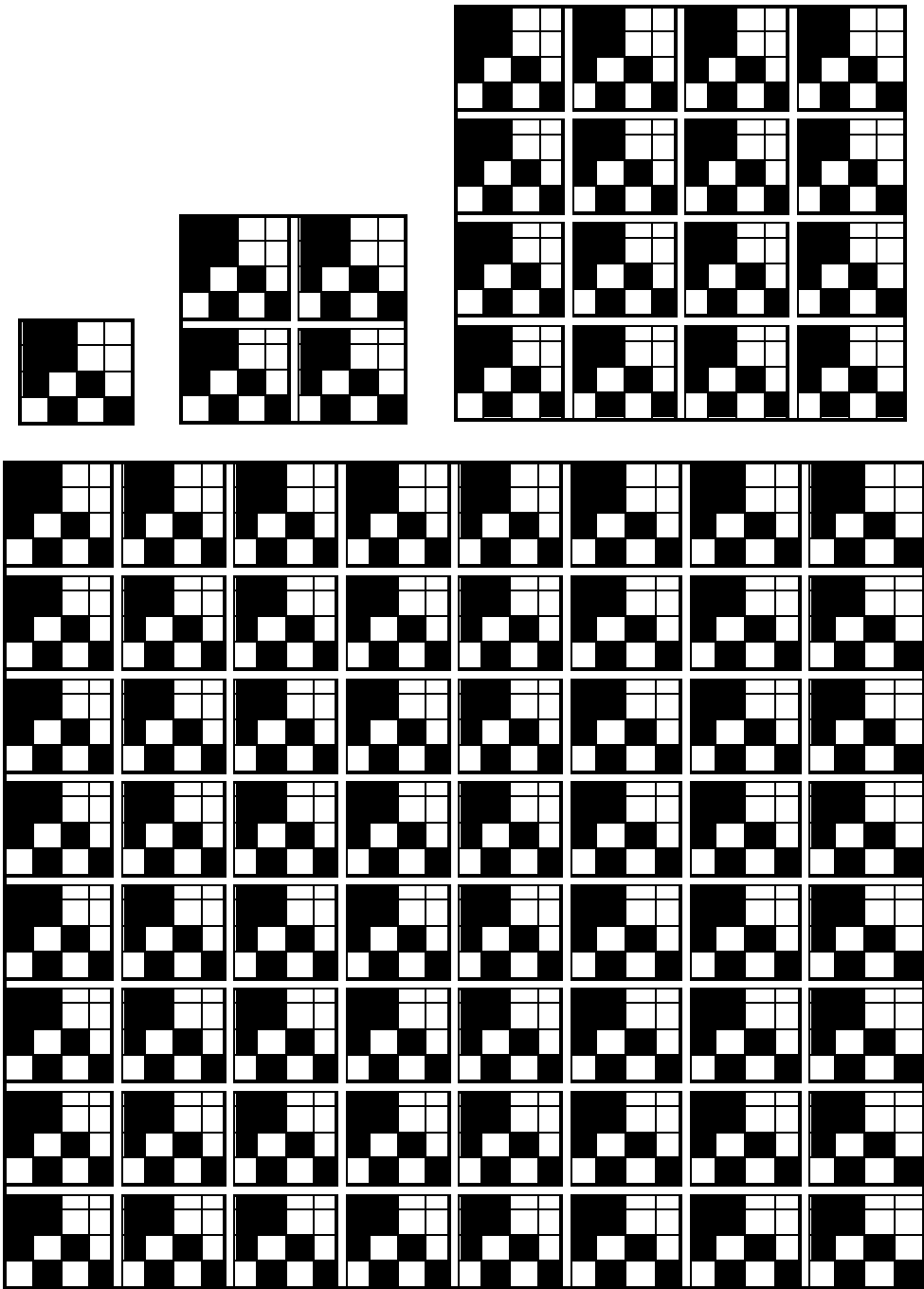
**Fig. 16.1.4**. "Matrix crystals" representing in the numeric form the interrelated mosaics of symbolic matrices $[C, A; T, G]^{(n)}_{bit-reversal}$ of DNA-alphabets from Fig. 16.1.3. In the numeric matrices, each of the black cells contains "+1" and each of the white cells contains "-1". The shown multi-block matrices are colonies or unions of the same numeric (4*4)-matrix ($\alpha$-matrix) representing the symbolic matrix $[C, A; T, G]^{(2)}_{bit-reversal}$ of the alphabet of 16 doublets.

Each of the numeric matrices in Fig. 16.1.4 can be constructed from the $\alpha$-matrix by the algorithm (16.1) based on its tensor product with the matrix [1, 1; 1, 1]:

$$[1, 1; 1, 1]^{(n)} \otimes [1, 1, -1, -1; 1, 1, -1, -1; 1, -1, 1, -1; -1, 1, -1, 1] \qquad (16.1)$$

The matrix [1, 1; 1, 1] used in (16.1) is the matrix representation of a 2-dimensional hyperbolic number with unit coordinates. Deep structural connections of genetic informatics with $2^n$-dimensional hyperbolic numbers are described in [Petoukhov, 2021].

The second interesting result of the described bit-reversal transformations of genetic matrices from the tensor family $[C, A; T, G]^{(n)}$ was received from analysis of the question: does this basic $\alpha$-matrix [1, 1, -1, -1; 1, 1, -1, -1; 1, -1, 1, -1; -1, 1, -1, 1], which is algorithmically repeated in the matrix crystals in Fig. 16.1.4, have any essential algebraic content, or is it just a random collection of numbers +1 and -1?

The answer received by the author reveals an essential algebraic content of this genetic $\alpha$-matrix. Fig. 16.1.5 shows that the $\alpha$-matrix is a sum of 4 sparse matrices $j_0+j_1+j_2+j_3$, the set of which is closed relative to multiplication (multiplying any two matrices with each other generates a matrix from the same set). The corresponding multiplication table of these 4 sparse genetic matrices unexpectedly coincides with the multiplication table of base elements of the known algebra of 4-dimensional split-quaternions or coquaternions introduced by J. Cockle in 1849. In particular, split-quaternions are used in hyperbolic geometry for describing hyperbolic motions in the Poincare disk model [Karzel, Kist, 1985; https://en.wikipedia.org/wiki/Split-quaternion ].

$$\alpha = \begin{vmatrix} 1, & 1, & -1, & -1 \\ 1, & 1, & -1, & -1 \\ 1, & -1, & 1, & -1 \\ -1, & 1, & -1, & 1 \end{vmatrix} = \begin{vmatrix} 1,0,0,0 \\ 0,1,0,0 \\ 0,0,1,0 \\ 0,0,0,1 \end{vmatrix} + \begin{vmatrix} 0,0,-1,0 \\ 0,0,0,-1 \\ 1,0,0,0 \\ 0,1,0,0 \end{vmatrix} + \begin{vmatrix} 0,0,0,-1 \\ 0,0,-1,0 \\ 0,-1,0,0 \\ -1,0,0,0 \end{vmatrix} + \begin{vmatrix} 0,1,0,0 \\ 1,0,0,0 \\ 0,0,0,-1 \\ 0,0,-1,0 \end{vmatrix} = j_0+j_1+j_2+j_3$$

| * | $j_0$ | $j_1$ | $j_2$ | $j_3$ |
|---|---|---|---|---|
| $j_0$ | $j_0$ | $j_1$ | $j_2$ | $j_3$ |
| $j_1$ | $j_1$ | $-j_0$ | $j_3$ | $-j_2$ |
| $j_2$ | $j_2$ | $-j_3$ | $j_0$ | $-j_1$ |
| $j_3$ | $j_3$ | $j_2$ | $j_1$ | $j_0$ |
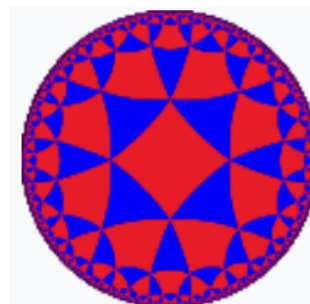
**Fig. 16.1.5**. The decomposition of the α-matrix gives 4 sparse matrices (α = $j_0+j_1+j_2+j_3$) whose set is closed relative to multiplication (top row) and corresponds to the multiplication table of split-quaternions by J.Cockle. One of the possible graphical symbols of the Poincare disk model of Lobachevsky hyperbolic geometry is also shown here to remind about using split-quaternions for describing hyperbolic motions in this model (the symbol is taken from https://commons.wikimedia.org/wiki/Category:Poincar%C3%A9_disk_models ).

If each of the repeated α-matrices in the multi-block matrices in Fig. 16.1.4 is replaced by the shown symbol of the Poincare disk model, then artistic representations of these matrices appear whose examples are given in Fig. 16.1.6. Such artistic representations are useful for developing geometric intuition.
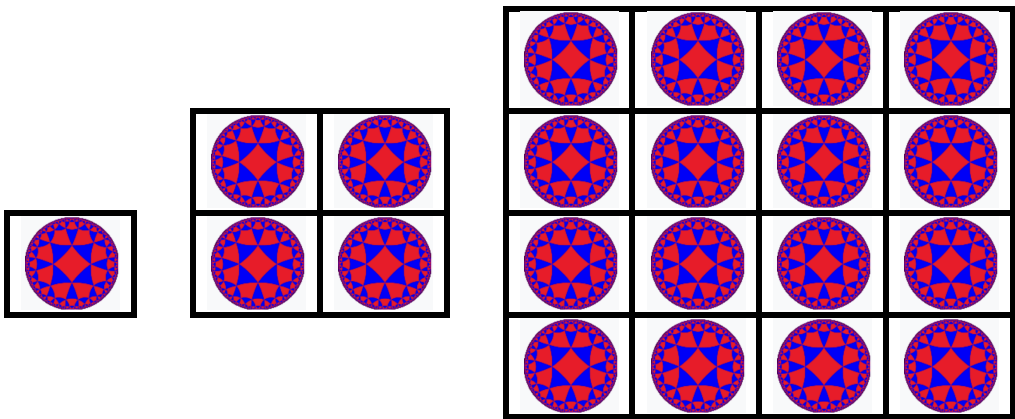


**Fig. 16.1.6**. Examples of the artistic representations of genetic multi-block matrices from Fig. 16.1.4 are received by replacements of their blocks (α-matrices) with the symbol of the Poincare disk model.

The presented study of the algebraic-holographic structures of multi-alphabetic genetic informatics unexpectedly reveals the structural parallels of genetic informatics with holographic quantum codes and algebraic-holographic principles, actively developed in modern theoretical physics by many famous authors. To illustrate these parallels, let us use excerpts from works on holographic quantum error-correcting codes [Pastawski et al., 2015; Preskill, 2016]. The Poincaré disk model and the corresponding hyperbolic space tilings are used in these works. Fig. 16.1.7, which is taken from these works, briefly illustrates these parallels by using the same symbol of the Poincaré disk model, which is presented above in connection with the algebraic-holographic tensor properties of genetic informatics in Fig. 16.1.5 and 16.1.6.



Holographic correspondence
Quantum error correction
Are they closely related?
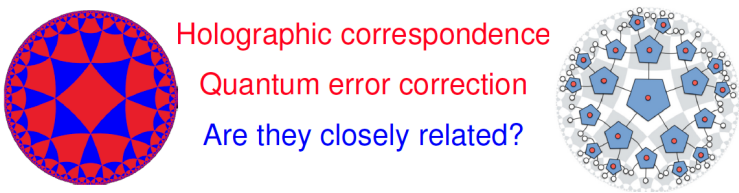
**Fig. 16.1.7**. The illustration to the theme of holographic quantum error-correcting codes (from [Preskill, 2016]).

The following quotes from these works, which describe quantum error-correcting codes with a tensor network structure, clarify these parallels in more detail. "*The tensor network is supported on a uniform tiling of a hyperbolic space, known as a hyperbolic tessellation. … We shall focus on examples based on tilings of two-dimensional hyperbolic space, which are specific realizations of uniform hyperbolic tilings known as hyperbolic tessellations*" [Pastawski et al., 2015]. «*Holographic quantum error-correcting codes are constructed by contracting perfect tensors according to a tilling of hyperbolic space by polygons*» [Preskill, 2016].

Let us return to the family of genetic mosaic matrices $[C, A; T, G]^{(n)}_{bit-reversal}$ (Fig. 16.1.3), which were received by the algebra-holographic method, to study the following question about the genetic code of amino acids: how 20 amino acids and stop-codons, which are encoded by 64 triplets, are located in the bit-reversal (8*8)-matrix $[C, A; T, G]^{(3)}_{bit-reversal}$? Does their arrangement have a pronounced law character, which can additionally confirm the deep consistency of the amino acid genetic coding system with the holographic method of bit-reverse permutations, which led to this matrix?

Fig. 16.1.8 shows interrelated locations of triplets, amino acids, and stop-codons in the bit-reversal (8*8)-matrix $[C, A; T, G]^{(3)}_{bit-reversal}$ (from Fig. 16.1.3) for the case of the Vertebrate Mitochondrial Code, which is the most symmetrical among known dialects on the genetic code and which is called in genetics the most ancient and "ideal" [Frank-Kamenetskii, 1988] (other dialects of the genetic code have small differences from this basic one). It can be seen from Fig. 16.1./8 that the upper and lower halves of the matrix $[C, A; T, G]^{(3)}_{bit-reversal}$ are unexpectedly identical to each other in the composition and arrangement of amino acids and stop-codons. All so-termed high-degenerate amino acids (Pro, Thr, Ser, Ala, Leu, Arg, Val, Gly), which are encoded by triplets with strong roots (black cells), are located in 4 quadrants identically. These strict regularities give additional pieces of evidence that genetic informatics is closely connected with principles of algebraic holography.

| CCC | ACC | CAC | AAC | CCA | ACA | CAA | AAA |
|-----|-----|-----|-----|-----|-----|-----|-----|
| **Pro** | **Thr** | **His** | **Asn** | **Pro** | **Thr** | **Gln** | **Lys** |
| UCC | GCC | UAC | GAC | UCA | GCA | UAA | GAA |
| **Ser** | **Ala** | **Tyr** | **Asp** | **Ser** | **Ala** | **Stop** | **Glu** |
| CUC | AUC | CGC | AGC | CUA | AUA | CGA | AGA |
| **Leu** | **Ile** | **Arg** | **Ser** | **Leu** | **Met** | **Arg** | **Stop** |
| UUC | GUC | UGC | GGC | UUA | GUA | UGA | GGA |
| **Phe** | **Val** | **Cys** | **Gly** | **Leu** | **Val** | **Trp** | **Gly** |
| CCU | ACU | CAU | AAU | CCG | ACG | CAG | AAG |
| **Pro** | **Thr** | **His** | **Asn** | **Pro** | **Thr** | **Gln** | **Lys** |
| UCU | GCU | UAU | GAU | UCG | GCG | UAG | GAG |
| **Ser** | **Ala** | **Tyr** | **Asp** | **Ser** | **Ala** | **Stop** | **Glu** |
| CUU | AUU | CGU | AGU | CUG | AUG | CGG | AGG |
| **Leu** | **Ile** | **Arg** | **Ser** | **Leu** | **Met** | **Arg** | **Stop** |
| UUU | GUU | UGU | GGU | UUG | GUG | UGG | GGG |
| **Phe** | **Val** | **Cys** | **Gly** | **Leu** | **Val** | **Trp** | **Gly** |

**Fig. 16.1.8.** Arrangement of amino acids and stop-codons in the alphabetical matrix $[C, A; T, G]^{(3)}_{bit-reversal}$ (from Fig. 16.1.3) of 64 triplets, which encode them (the case of the Vertebrate Mitochondria Code is presented).

Besides this, each of the quadrants of $(2^n*2^n)$-matrices contains all the *n*-plets with the same letter at the last position. Therefore, for the corresponding quadrants (for example, for quadrants with the nucleotide C at the last position of *n*-plets), the Gestalt rule is fulfilled about approximate equality of the sums of the percentages of *n*-plets in genomic DNA sequences (see above Section 3 and Fig. 3.1).

In this Section, we met the fact that the holographic method of bit-reversal permutations, when applied to genetic alphabetical matrices $[C, A; T, G]^{(n)}$, reveals a connection of genetic structures with split-quaternions, which are connected with the Poincare disk model of hyperbolic space and used for describing hyperbolic motions there. Let us else remind some additional information about split-quaternions, Poincare disk model, and hyperbolic motions.

The Poincaré disk model, which is also called the conformal disk model, is related to the hyperboloid model of hyperbolic geometry [Reid, Szendroi, 2005; Reynolds, 1993; https://en.wikipedia.org/wiki/Hyperbolic_motion#Disk_model_motions; https://en.wikipedia.org/wiki/Poincar%C3%A9_disk_model]. Inversive geometry and Mobius transformations are used for describing hyperbolic motions there. The hyperboloid model is also known as the Minkowski model (in a general case it is a model of *n*-dimensional hyperbolic geometry using a two-sheeted hyperboloid in $(n+1)$-dimensional Minkowski space).

In connection with Möbius transformations and conformal geometry in this algebra-genetic study, one can recall the phenomenon of conformal-geometric biosymmetries in genetically inherited morphological configurations [Petoukhov, 1989]. Let us also remind that by the pioneering work [Luneburg, 1950], the space of binocular visual perception is described by hyperbolic geometry. These findings were followed by many papers in various countries, where the idea of a non-Euclidean space of visual perception was extended and refined. The Luneburg approach was thoroughly tested by G. Kienle [Kienle, 1964]. In the main series of his experiments, where about 200 observers were involved, Kienle obtained about 1300 visual patterns of various kinds. The experiments confirmed not only that the space of visual perception is described by hyperbolic geometry but also that the Poincare disk (or conformal) model was an adequate model of that geometry. He concluded his paper by writing: "*Poincare's model of hyperbolic space, applied for the first time for a mapping of the visual space, shows a reasonably good agreement with experimental results*" [Kienle, 1964, p. 399].

An interesting study about «Non-Euclidean geometries for grid cells» in mammalian brains should be also reminded here [Urdapilleta et al., 2015; Sissa Medialab, 2015]. According this study, published in Interface, the journal of the Royal Society, «*grid cells, space-mapping neurons of the entorhinal cortex of rodents, could also work for hyperbolic surfaces*». This study is coordinated by neuroscientist A.Treves, who notes: "*It took human culture millennia to arrive at a mathematical formulation of non-Euclidean spaces, but it's very likely that our brains could get there long before. In fact, it's likely that the brain of rodents gets there very naturally every day*". The speech is about grid cells, neurons of the entorhinal cortex of rodents that fire in a characteristic way when the animal moves in an arena. The discovery of grid cells by Edvard and May-Britt Moser has been awarded the Nobel Prize in 2014 (for the discovery of the system of cells in the brain, which allows you to navigate in space).

The article [Smolyaninov, 2000] presents the results of 20 years of author's research on the locomotion of a wide variety of animals and humans. According to these results, spatio-temporal organization of locomotion control is related – in a special manner - with hyperbolic rotations and with Minkovsky geometry. Based on this analysis, Smolyaninov put forward his "Locomotor theory of relativity" and wrote about a relativistic brain and relativistic biomechanics.

The works [Bodnar, 1992, 1994] analyzed growth transformations of geometric lattices in famous morphogenetic phenomena of phyllotaxis, which exist in plant and animal bodies at various levels and branches of biological evolution. On the basis of this analysis, the author declared that living matter is structurally related to Minkowski geometry.

A special analysis should be conducted in the future about possible relations of the mentioned genetic phenomena with hyperbolic structures on 3-manifolds described in [Thurston, 1982, 1986].

Appendix III contains additional data concerning extensions of split-quaternion algebra in connection with the «matrix crystals» in Figs. 16.1.4 and 16.1.6.

The author notes that usually the matrix representation of split-quaternions is given based on two pairs of conjugate complex numbers. But the genetic $\alpha$-matrix [1, 1, -1, -1; 1, 1, -1, -1; 1, -1, 1, -1; -1, 1, -1, 1], which represents split-quaternions with unit coordinates, is connected with the following non-traditional matrix representation of split-quaternions on the basis of two pairs of conjugate hyperbolic numbers x+iy, x-iy, and -(z+iv), z-iv, where $i^2 = +1$ is the imaginary unit of hyperbolic numbers, and x, y, z, v are real numbers (Fig. 16.1.9). This additionally shows the important structural role of hyperbolic numbers in genetic informatics described in [Petoukhov, 2021].

$$\begin{bmatrix} x+iy, & -(z+iv) \\ z-iv, & x-iy \end{bmatrix} = \begin{bmatrix} x, & y, & -z, & -v \\ y, & x, & -v, & -z \\ z, & -v, & x, & -y \\ -v, & z, & -y, & x \end{bmatrix}$$

**Fig. 16.1.9**. The matrix representation of split-quaternions $x + zj_1 + vj_2 + yj_3$ on the based of 2-dimensional hyperbolic numbers (where $i^2 = +1$); $j_1, j_2, j_3$ are shown above in Fig. 16.1.5.

One should else note the actions of the $(2^n*2^n)$-matrices (Figs. 16.1.4 and 16.1.6) to vectors. The expression (16.1) shows the algorithm for constructing such matrices based on the $\alpha$-matrix. Let us consider their actions to voluntary $2^n$-dimensional vectors having arbitrary coordinates $a_i(t)$, which are not connected to each other and can be, for example, cyclic functions of time.

The action of the (4*4)-matrix $\alpha$ = [1, 1, -1, -1; 1, 1, -1, -1; 1, -1, 1, -1; -1, 1, -1, 1] to the vector $\bar{x}_4$ = [$a_0(t), a_1(t), a_2(t), a_3(t)$] with coordinates unrelated to each other gives a new block-structured vector with interrelated coordinates combined in pairs by a binary-oppositional principle:

$$\bar{x}_4*\alpha = [v(t), w(t), -w(t), -v(t)], \qquad (16.2)$$

where $v(t) = a_0(t) + a_1(t) + a_2(t) - a_3(t)$, $w(t) = a_0(t) + a_1(t) - a_2(t) + a_3(t)$. In this generated vector, its coordinates $v(t)$ and $-v(t)$ (and also $w(t)$ and $-w(t)$) are equal to each other up to the sign and their changes are synchronized in time.

The action of the $(8*8)$-martix $\alpha_4 = [1, 1; 1, 1] \otimes \alpha$ to the vector $\bar{x}_8 = [a_0(t), a_1(t), a_2(t), a_3(t), a_4(t), a_5(t), a_6(t), a_7(t)]$ with coordinates unrelated to each other gives a new block-structured vector with interrelated coordinates combined in pairs by the similar binary-oppositional principle:

$$\bar{x}_8 * \alpha_4 = [v(t), w(t), -w(t), -v(t), v(t), w(t), -w(t), -v(t)], \qquad (16.3)$$

where $v(t) = a_0 + a_1 + a_2 - a_3 + a_4 + a_5 + a_6 - a_7$, $w(t) = a_0 + a_1 - a_2 + a_3 + a_4 + a_5 - a_6 + a_7$. This generated vector (16.3) has a block structure and paired coordinates $v(t)$ and $-v(t)$ (the same for $w(t)$ and $-w(t)$), which are equal to each other up to the sign and their changes are coordinated and synchronized in time.

The action of the $(16*16)$-matrix $\alpha_8 = [1, 1; 1, 1]^{(2)} \otimes \alpha$ to a 16-dimensional vector $\bar{x}_{16} = [a_0(t), a_1(t), a_2(t), a_3(t), a_4(t), a_5(t), a_6(t), a_7(t), a_8(t), a_9(t), a_{10}(t), a_{11}(t), a_{12}(t), a_{13}(t), a_{14}(t), a_{15}(t)]$ with coordinates unrelated to each other gives a new block-structured vector with interrelated coordinates combined in pairs by the similar binary-oppositional principle:

$$\bar{x}_{16} * \alpha_8 = [v(t), w(t), -w(t), -v(t), v(t), w(t), -w(t), -v(t),$$
$$v(t), w(t), -w(t), -v(t), v(t), w(t), -w(t), -v(t)] \qquad (16.4)$$

where $v(t) = a_0 + a_1 + a_2 - a_3 + a_4 + a_5 + a_6 - a_7 + a_8 + a_9 + a_{10} - a_{11} + a_{12} + a_{13} + a_{14} - a_{15}$, $w(t) = a_0 + a_1 - a_2 + a_3 + a_4 + a_5 - a_6 + a_7 + a_8 + a_9 - a_{10} + a_{11} + a_{12} + a_{13} - a_{14} + a_{15}$. This generated vector (16.4) has a block structure and paired coordinates $v(t)$ and $-v(t)$ (the same for $w(t)$ and $-w(t)$), which are equal to each other up to the sign and their changes are coordinated and synchronized in time.

In a general case, actions of $(2^n*2^n)$-matrices $[1, 1; 1, 1]^{(n-2)} \otimes \alpha$ to a voluntary $2^n$-dimensional vector with coordinates unrelated to each other gives a new block-structured vector with interrelated coordinates combined in pairs by the similar binary-oppositional principle. If initial vectors $\bar{x}_4, \bar{x}_8, \bar{x}_{16}, \ldots$ are state vectors of systems with unrelated coordinates in configuration spaces (or unrelated behavior in different configuration subspaces), then actions of matrix operators $[1, 1; 1, 1]^{(n)} \otimes \alpha$ transform the system states into new ordered forms with interrelated behavior in different configuration subspaces. For example, if coordinates of an initial state vector of a system are cyclic or wave functions, then the received new state vector have block-structured sets of coordinates whose blocks are equal to each other up to the sign, and their cyclic changes are synchronized in time (that is, cyclic behavior of the considered system in its different configuration subspaces becomes strongly interrelated). Figuratively speaking, these matrix operators (Figs. 16.1.4 and 16.1.6), obtained by analyzing the system of structured DNA alphabets, create order out of chaos.

This algebra-genetic result is essential for the modeling of biological systems since the coordinated cyclic behavior of genetically inherited parts is one of the characteristic features of living bodies. The living organism as a whole is a huge chorus of coordinated cyclic processes. For example, the inherited normal work of the heart is based on the coordinated cyclic work of the parts of the heart muscle; violation of this coordination leads to disease - cardiac arrhythmias. Tibetan pulse

diagnosis is based on the relationship of inherited cyclic processes in the body, and so on.

It should be noted holographic-like features of the received multi-block vectors: knowing only one of the blocks of such vector allows restoring all the vector. For example, knowing two adjacent coordinates v(t) and w(t) in any part of the 16-dimensional vector in (16.4), it is possible to recover the entire vector.

## 16.2. Holographic coding by Walsh–Hadamard transformation of randomized and permuted data

Above, in Section 15, we showed that the set of algebra-structured alphabets of DNA and RNA has essential relations with Walsh functions and Hadamard matrices. Besides this, binary-oppositional features of these genetic alphabets are connected with the logical operation of modulo-2 addition using in computer informatics [Petoukhov, 2016a,b]. But Walsh functions and the operation of modulo-2 addition are bases of a known algebra-holographic method of processing binary data in digital technologies, which is termed as "holographic coding by Walsh-Hadamard transformation of randomized and permuted data" [Berend et al. 2016; Dolev, Frenkel, 2010; Dolev, Frenkel, Cohen, 2012]. It is holographic in the sense that any portion of the coded information (which is any subset of corresponding codewords) represents a blurred image of the entire data. The Walsh–Hadamard transform (multiplication by the Walsh–Hadamard matrix, which is a specific Hadamard matrix having sizes ($2^n*2^n$) is a holographic in the sense that each coefficient (entry in the output vector) is a result of a simple computation involving all the entries of the original input vector. The Hamming distance between the original and the reconstructed binary files is considered as a blurriness measure.

In particular, this algebra-holographic approach is used to simulate holographic "brain" memory [Berend et al., 2016]. The encoding of data is performed on randomized information, which is then represented by a set of Walsh–Hadamard coefficients. These coefficients turn out to have holographic properties. Namely, any portion of the set of coefficients defines a ''blurry image'' of the original data. The authors describe a built-in error correction technique - enlarging the width of the matrix used in the Walsh-Hadamard transform to produce a rectangular Hadamard matrix. They randomize a binary input vector by xoring it with a random input vector (that is, making modulo-2 addition of them). In this case, the data can bear more errors, resulting in a system that is not affected by missing coefficients up to a certain threshold. Above this threshold, the loss of data is reflected by getting a ''blurry image'' rather than concentrated damage. Besides a heuristic analysis of the ability of the technique to correct errors, as well as showing an example of an image saved using the system, the authors give an example of a simple implementation of this algebra-holographic approach using neural networks as a proof of concept.

But the noted articles, considering their algebra-holographic method as a possible for simulation of holographic properties of the human brain, pay neither attention to the fact that the properties of the human brain are genetically inherited and by this reason should be connected in any way with structural properties of the genetic code system. Our above-presented results show that structural properties of the genetic code system correspond to formalisms of the method of holographic coding by Walsh–Hadamard transformation.

Additional acknowledgment of the relation of the system of bit-reverse alphabetical matrices [C, A; T, G]$^{(n)}$_bit-reversal with Walsh-Rademacher functions gives

the phenomenological Gestalt-rule of the percentage composition of DNA sequences in genomes of higher and lower organisms. The meander-like mosaics in rows of alphabetic matrices $[C, A; T, G]^{(n)}$ and of their bit-reverse analogs $[C, A; T, G]^{(n)}_{\text{bit-reversal}}$ corresponds to the meander-like Rademacher-Walsh functions. At the same time, the corresponding rows in these two types of matrices are the same in the composition of $n$-plets (they differ only in the order of $n$-plets in them with a change in the order of black and white cells). Therefore, the above-noted genomic Gestalt-rule (Figs. 15.4-15.7) of approximate constancy of the sums of percentages in the sets of $n$-plets with strong roots and $n$-plets with weak roots is fulfilled for corresponding rows of the tensor families of matrices of both types.

### 16.3. Logical holography by Walsh functions.

The pioneering work about logical «*holography by Walsh waves*» was [Morita, Sakurai, 1973]. The work was devoted to Walsh waves (or Walsh functions), which propagate through electronic circuits - composed of logical and analog elements - by the analogy with the optical Fourier transform holography. In this algebra-logical Walsh-holography, objects, whose digital holograms should be made, are represented in forms of $2^n$-dimensional vectors. Each component of these vectors corresponds to one of $2^n$ input channels of appropriate electric circuits; the same is true for $2^n$ output channels, which are related with components of resulting vectors. Examples of electrical circuits for logic holography are shown in the works [Morita, Sakurai,1973; Soroko, 1974]. Due to the application of Walsh transformation, information about such vector is written in each component of the appropriate hologram, which is also a $2^n$-dimensional vector, to provide a nonlocal character of storing information.

One should especially note that Walsh functions are closely related to dyadic groups since Walsh functions are characters of the dyadic groups of binary numbers [Fine, 1949]. Therefore, the Fourier analysis on dyadic groups is defined in terms of Walsh functions. In the same way, the discrete Walsh functions are characters of the finite dyadic groups, on which the switching functions are defined. Therefore, the Fourier analysis for switching functions considered as a subset of complex-valued functions is formulated in terms of Walsh functions [Karpovsky, Stankovic, Astola, 2008].

This algebra-logical Walsh-holography under the title «logical holography» was also considered later in [Derzhypolskyy, Melenevskyy, Gnatovskyy, 2007; Golubov, Efimov, Skvortsov, 1987]. All these and other works about logical Walsh-holography considered possibilities of its applications in engineering technologies of digital signals processing without any supposition of its application in biology, in particular, in genetics. In the author's publications [Petoukhov, 2008a,b,c, 2016a,b; Petoukhov, Petukhova, 2017], deep connections of the genetic code system with Walsh functions, Hadamard matrices, dyadic groups, and logical modulo-2 addition are described. Due to these results and some other materials, the author has put forward in 2016 the hypothesis that the genetic system is related to principles of logical holography and appropriate logic operations.

This hypothesis leads to a new class of mathematical models of genetic structures and phenomena on the basis of logical holography and appropriate logical operations. Correspondingly the author develops the theory of «genetic logical holography», where the mathematics of logical holography and logical operations is used for modeling genetic phenomena. The mathematical basis of this modeling

approach is lattice functions, logical operations with them, dyadic spaces, dyadic groups of binary numbers, logic modulo-2 addition, dyadic convolution, dyadic derivatives as well as isomorphic and homomorphic transformations of Walsh in close relation with peculiarities of molecular-genetic systems. In our opinion, algebra-logical formalisms of logical holography are effective for modeling holographic-like phenomena in genetics and algebraic biology. Based on such approach, the author proposed the concept of the geno-logic code and also algebra-logical models of the following genetic phenomena (see [Petoukhov, 2016b; Petoukhov, Petukhova, 2017]):

- the phenomenon of zipper-like reproductions of DNA molecules (Fig. 16.3.1);
- phenomena of different repetitions in nucleotide sequences of DNA and RNA where a great number of complementary palindromes and simple repetitions exists;
- Fibonacci numbers in inherited biological structures.
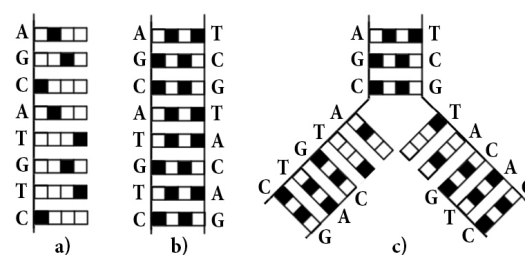


**Fig. 16.3.1**. The illustration to the model of the zipper-like reproduction of DNA (from [Petoukhov, 2016b]). Nitrogenous bases of DNA are denoted as informational vectors: $\overline{C} = [0.5, 0, 0, 0]$, $\overline{G} = [0, 0, 0.5, 0]$, $\overline{A} = [0, 0.5, 0, 0]$, $\overline{T} = [0, 0, 0, 0.5]$. These vectors are shown schematically as structures with 4 cells in each, where each black cell means 0.5 and each white cell means 0. a) A separate DNA-chain; b) two complementary chains of DNA; c) unzipping two chains of DNA.

**Some concluding remarks**.

The presented results of the study of the regularities in the distribution of the percentages of *n*-plets in long DNA-texts of various organisms are consistent with Jordan's claiming that life's missing laws are the rules of chance and probability of the quantum world [Jordan, 1932; McFadden, Al-Khalili, 2018]. The described author's results show the existence of previously unknown genetic regularities. These results were obtained baseed on new methods of analysis and modeling of DNA-texts, which are connected with mathematical formalisms of quantum mechanics and quantum informatics, algebras of multi-dimensional hypercomplex numbers, metric analysis, algebraic holography, and the theory on noise-immune coding of informatics.

Considering the views of Jordan and Schrödinger about the dictatorial role of the structured informatics of genetic molecules for the whole organism [McFadden, Al-Khalili, 2018], it is natural to think that the structural features of DNA informatics leave their marks on all genetically inherited biological systems and phenomena. This is consistent with the fact that all physiological systems must be structurally aligned with genetic coding to be transmitted in genetically encoded form to offspring. This is also consistent with the point of view that the main task of living organisms is to transfer genetic information along the chain of generations. The described algebraic-

genetic results give pieces of evidence that the system of genetic coding is based on methods of probability coding, which are studied now in the author's laboratory.

E.Schrodinger noted: "*from all, we have learned about the structure of living matter, we must be prepared to find it working in a manner that cannot be reduced to the ordinary laws of physics… because the construction is different from anything we have yet tested in the physical laboratory*» [Schrodinger, 1944]. For comparison, the enzymes in biological organisms work a million times more effectively than catalysts in the laboratory. What makes the enzyme in the body for 1 second, a catalyst in the laboratory can make only for 100 thousand years. We believe that such ultra-efficiency of enzymes in biological bodies is defined not only by laws of physics, but also by quantum-logical relations in the genetic system, and therefore - in line with Schrodinger - this ultra-efficiency cannot be reduced to the ordinary laws of physics. As far as we understand, the found Gestalt rules of "dictatorial" DNA-texts are not derived from the known laws of physics, and therefore refer to the special laws of the structuring of living things that Schrödinger spoke about.

The received results can be used, in particular, for developing our knowledge about principles of brain activities and about the relation between 'living' and non-living matter. These themes are actively discussed in the scientific community. For example, concerning a relation between 'living' and non-living matter, W. Pauli stated that the mental and the material domain are governed by common ordering principles, and can be understood as "*complementary aspects of the same reality*" [Pauli, 1994; Geesink, Meijer, 2016].

Regarding the metric features of biological phenomena, it can be noted that thoughts about metric spaces are spread by some authors even to mathematics as a high form of intellectual activity. For example, the book [Hofstadter, 1980, page 612] notes that a mathematician feels that in mathematics there is a certain metric that unites ideas - that all mathematics is a network of results that are interconnected by a huge number of connections; had we been able to introduce this highly developed sense of mathematical closeness - the mental metric of a mathematician - into the program, we could create a primitive artificial mathematician.

By these statements of Hofstadter, the book [Nalimov, 2015, p. 115] emphasizes: "*In other words, artificial intelligence could be brought closer to mathematical thinking, if it were possible to realize the metric properties of the human thinking space ... We are ready to go further and say that consciousness itself is geometrically structured: existentially, a person is geometric ...  In our minds, when constructing texts through which we perceive the World, something very similar to what happens in morphogenesis happens. We are ready to see in the depths of consciousness the same geometric images that are revealed in morphogenesis*".

This article is devoted to algebraic Gestalt phenomena in genetics and also to applications of algebraic holography methods in genetics for understanding of inherited holographic-like properties in organisms. The presented author's results show previously unknown abilities and directions for developing of genetical informatics on holographic principles in parallels with holographic approaches actively developed in modern theoretical physics and science on brain functions.

For example, let us recal about D.Bohm, who has been described as one of the most significant theoretical physicists of the 20th century and who contributed unorthodox ideas to quantum theory, neuropsychology and the philosophy of mind [https://en.wikipedia.org/wiki/David_Bohm]. Bohm believed that a separate human consciousness interacts with the consciousness of other people in a holographic way. Human consciousness is part of the universal hologram of the entire human race, so

that all the data of this great whole can be extracted from a separate consciousness. David Bohm was convinced that the entire universe has a holographic structure and that with its help many events in the universe and in human life can be explained. Bohm worked with Pribram on the theory that the brain operates in a manner that is similar to a hologram, in accordance with quantum mathematical principles and the characteristics of wave patterns [Pribram, 1998]. Nobel laureate in physics Gerard 't Hooft has put forward the holographic principle, which is currently being developed by many physicists in string theory, quantum gravity, etc. [https://en.wikipedia.org/wiki/Holographic_principle]. According to this principle, which connects physics with informatics, the description of a volume of space can be thought of as encoded on a lower-dimensional boundary to the region.

One should also note about perspectives of using of DNA as an ultrahigh-density storage medium that could meet exponentially growing worldwide demand for archival data storage [Banal et al., 2021; Chandak et al., 2019]. Our results about algebraic features of DNA informatics can help in solving some engineering problems in this technological challenge.

In general, the presented results of the author's studies of the structural rules of genetic informatics give pieces of evidence in favor of the effectivity of a model approach to living organisms as quantum-informational algebraic-harmonic essences having algebraic-holographic properties.

**Appendix I. An example of *n*-plets percentages in one of the bacterial genomes.**

The Appendix presents data about percentages of *n*-plets in the rhizobacteria *Bradyrhizobium japonicum* strain E109, complete genome, 9224208 bp (initial data were taken from https://www.ncbi.nlm.nih.gov/nuccore/CP010313.1?report=genbank; the length of this sequence is more than 20 times shorter than the length of the sequence of the human chromosome №1). Figs. AI.1-AI.5 show matrices of phenomenological percentages of 4 nucleotides, 16 doublets, 64 triplets, and 256 tetraplets in corresponding layers of this genomic DNA. These percentages matrices for the bacteria and the percentages matrices (Figs. 2.2-2.5) for human chromosome №1 have similar properties for *m*-positional tetra-groupings of *n*-plets, which are expressed by the equalities (3.1) and (3.2).

| C | A | | %C | %A | | 0.3184 | 0.1819 |
|---|---|---|----|----|---|--------|--------|
| T | G | ➔ | %T | %G | = | 0.1815 | 0.3182 |

**Fig. AI.1**. The transformation of the symbolic matrix of 4 nucleotides into the numeric matrix of nucleotides percents in the case of the genomic DNA of the *Bradyrhizobium japonicum* strain E109.

| %CC | %CA | %AC | %AA | | 0.0781 | 0.0587 | 0.0467 | 0.0367 |
|-----|-----|-----|-----|---|--------|--------|--------|--------|
| %CT | %CG | %AT | %AG | | 0.0515 | 0.1302 | 0.0469 | 0.0517 |
| %TC | %TA | %GC | %GA | = | 0.0720 | 0.0145 | 0.1216 | 0.0719 |
| %TT | %TG | %GT | %GG | | 0.0365 | 0.0584 | 0.0468 | 0.0779 |

**Fig. AI.2**. The matrix of phenomenolical percents of the 16 doublets in the genomic DNA-sequence of doublets in the *Bradyrhizobium japonicum* strain E109. (Compare with Fig. 2.3).

| %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 |

=

| 0.0131 | 0.0154 | 0.0136 | 0.0129 | 0.0145 | 0.0083 | 0.0097 | 0.0068 |
|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0148 | 0.0344 | 0.0145 | 0.0185 | 0.0053 | 0.0182 | 0.0075 | 0.0126 |
| 0.0174 | 0.0040 | 0.0446 | 0.0328 | 0.0205 | 0.0043 | 0.0205 | 0.0106 |
| 0.0126 | 0.0183 | 0.0185 | 0.0343 | 0.0075 | 0.0144 | 0.0052 | 0.0146 |
| 0.0149 | 0.0133 | 0.0044 | 0.0021 | 0.0353 | 0.0210 | 0.0195 | 0.0153 |
| 0.0105 | 0.0323 | 0.0043 | 0.0039 | 0.0210 | 0.0443 | 0.0209 | 0.0171 |
| 0.0151 | 0.0020 | 0.0206 | 0.0133 | 0.0195 | 0.0044 | 0.0356 | 0.0150 |
| 0.0069 | 0.0125 | 0.0082 | 0.0153 | 0.0098 | 0.0133 | 0.0148 | 0.0133 |

**Fig. AI.3**. The matrix of phenomenological percents of the 64 triplets in the genomic DNA-sequence of triplets in the *Bradyrhizobium japonicum* strain E109 (compare with Fig. 2.4).

| .0023 | .0025 | .0033 | .0030 | .0052 | .0017 | .0035 | .0021 | .0022 | .0037 | .0018 | .0019 | .0023 | .0023 | .0016 | .0015 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| .0026 | .0057 | .0038 | .0055 | .0014 | .0051 | .0027 | .0045 | .0029 | .0058 | .0024 | .0022 | .0015 | .0036 | .0019 | .0018 |
| .0048 | .0011 | .0119 | .0082 | .0066 | .0011 | .0081 | .0030 | .0014 | .0006 | .0057 | .0050 | .0025 | .0010 | .0045 | .0028 |
| .0039 | .0052 | .0048 | .0096 | .0021 | .0044 | .0018 | .0052 | .0014 | .0019 | .0029 | .0047 | .0018 | .0022 | .0015 | .0039 |
| .0036 | .0024 | .0016 | .0004 | .0141 | .0072 | .0094 | .0057 | .0042 | .0044 | .0010 | .0006 | .0053 | .0043 | .0026 | .0026 |
| .0020 | .0090 | .0015 | .0004 | .0073 | .0162 | .0085 | .0091 | .0035 | .0085 | .0012 | .0016 | .0039 | .0072 | .0035 | .0019 |
| .0056 | .0005 | .0068 | .0037 | .0082 | .0016 | .0173 | .0055 | .0024 | .0005 | .0048 | .0034 | .0018 | .0006 | .0058 | .0036 |
| .0019 | .0043 | .0022 | .0055 | .0035 | .0051 | .0059 | .0056 | .0019 | .0027 | .0024 | .0039 | .0015 | .0013 | .0030 | .0027 |
| .0027 | .0031 | .0035 | .0031 | .0012 | .0009 | .0006 | .0005 | .0059 | .0062 | .0048 | .0047 | .0058 | .0036 | .0042 | .0029 |
| .0035 | .0057 | .0034 | .0036 | .0006 | .0016 | .0005 | .0005 | .0059 | .0173 | .0049 | .0067 | .0018 | .0081 | .0025 | .0057 |
| .0040 | .0010 | .0103 | .0092 | .0017 | .0004 | .0013 | .0010 | .0069 | .0013 | .0167 | .0102 | .0098 | .0017 | .0070 | .0040 |
| .0027 | .0030 | .0050 | .0082 | .0011 | .0011 | .0006 | .0010 | .0045 | .0080 | .0058 | .0118 | .0024 | .0066 | .0014 | .0047 |
| .0036 | .0031 | .0007 | .0004 | .0065 | .0033 | .0035 | .0031 | .0037 | .0036 | .0010 | .0007 | .0094 | .0065 | .0037 | .0035 |
| .0026 | .0058 | .0006 | .0004 | .0043 | .0069 | .0044 | .0024 | .0026 | .0095 | .0010 | .0016 | .0053 | .0141 | .0041 | .0036 |
| .0028 | .0005 | .0047 | .0030 | .0036 | .0009 | .0061 | .0032 | .0040 | .0005 | .0048 | .0034 | .0059 | .0011 | .0059 | .0027 |
| .0015 | .0021 | .0020 | .0029 | .0024 | .0016 | .0036 | .0026 | .0016 | .0034 | .0018 | .0033 | .0023 | .0052 | .0021 | .0024 |

**Fig. AI.4**. The matrix of phenomenological percents of the 256 tetraplets in the genomic DNA-sequence of tetraplets in the *Bradyrhizobium japonicum* strain E109 (compare with Fig. 2.5).

| %C ≈ 0.3184 | %G ≈ 0.3182 | %A ≈ 0.1819 | %T ≈ 0.1815 |
|-------------|-------------|-------------|-------------|
| Σ%CN ≈ 0.3185 | Σ%GN ≈ 0.3182 | Σ%AN ≈ 0.1820 | Σ%TN ≈ 0.1814 |
| Σ%NC ≈ 0.3183 | Σ%NG ≈ 0.3182 | Σ%NA ≈ 0.1818 | Σ%NT ≈ 0.1816 |
| Σ%CNN ≈ 0.3197 | Σ%GNN ≈ 0.3201 | Σ%ANN ≈ 0.1806 | Σ%TNN ≈ 0.1797 |
| Σ%NCN ≈ 0.3166 | Σ%NGN ≈ 0.3173 | Σ%NAN ≈ 0.1836 | Σ%NTN ≈ 0.1825 |
| Σ%NNC ≈ 0.3189 | Σ%NNG ≈ 0.3173 | Σ%NNA ≈ 0.1815 | Σ%NNT ≈ 0.1824 |

| $\Sigma\%CNNN \approx 0.3184$ | $\Sigma\%GNNN \approx 0.3183$ | $\Sigma\%ANNN \approx 0.1816$ | $\Sigma\%TNNN \approx 0.1817$ |
|---|---|---|---|
| $\Sigma\%NCNN \approx 0.3184$ | $\Sigma\%NGNN \approx 0.3185$ | $\Sigma\%NANN \approx 0.1816$ | $\Sigma\%NTNN \approx 0.1815$ |
| $\Sigma\%NNCN \approx 0.3185$ | $\Sigma\%NNGN \approx 0.3181$ | $\Sigma\%NNAN \approx 0.1823$ | $\Sigma\%NNTN \approx 0.1811$ |
| $\Sigma\%NNNC \approx 0.3182$ | $\Sigma\%NNNG \approx 0.3180$ | $\Sigma\%NNNA \approx 0.1820$ | $\Sigma\%NNNT \approx 0.1818$ |

**Fig. AI.5**. Percentages of nucleotides C, G, A, T, and the sums of percent of *n*-plets with these nucleotides at their attributive positions in corresponding tetra-groupings of *n*-plets in the case of the genomic DNA-text of the *Bradyrhizobium japonicum* strain E109 (compare with Fig. 3.1).

Regarding projector tetra-groupings of *n*-plets in different layers of the genomic DNA of *Bradyrhizobium japonicum* strain E109, it was noted above in the expressions (8.9) and (8.10) that the following equalities are fulfilled in all 7 mosaic matrices of doublets and triplets in Fig. 8.3:

$$\Sigma\%\heartsuit \approx \Sigma\%\blacklozenge \approx 0.277, \qquad \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \approx 0.223$$

For example, in the case of the mosaic (4\*4)-matrix with the tetra-grouping №1 for 16 doublets (see Fig. 8.3), the following percentage sums appear for this genomic DNA based on percentages of doublets in Fig. AI.2:

$\Sigma\%\heartsuit \approx \%CA + \%CG + \%AA + \%AG \approx 0.0587 + 0.1302 + 0.0367 + 0.0517 \approx 0.277;$
$\Sigma\%\blacklozenge \approx \%TC + \%TT + \%GC + \%GT \approx 0.0720 + 0.0365 + 0.1216 + 0.0468 \approx 0.277;$
$\Sigma\%\spadesuit \approx \%CC + \%CT + \%AC + \%AT \approx 0.0781 + 0.0515 + 0.0467 + 0.0469 \approx 0.223;$
$\Sigma\%\clubsuit \approx \%TA + \%TG + \%GA + \%GG \approx 0.0145 + 0.0584 + 0.0719 + 0.0779 \approx 0.223.$

## Appendix II. Regarding Yin-Yang schemes of ancient Chinese book "I-Ching".

This appendix describes analogies between mosaic genetic matrices, which are related to the described genetic Gestal phenomena and which represent the arrangements of *m*-positional tetra-groupings of 64 triplets in Figs. 7.9-7.11, and a table of 64 hexagrams from the famous ancient Chinese book I-Ching (or "Book of Cyclic Changes").

The genetic coding system has binary-oppositional structures at different levels of its organization. As it is known, the ancient Chinese book I-Ching, which was written a few thousand years ago, introduced the system of symbols Yin and Yang (equivalents of 0 and 1). This book had a powerful impact on the culture, medicine, and science of ancient China and several other countries. The system of I-Ching is represented by the schemes with 4 bigrams, 8 trigrams, and 64 hexagrams. Similar to this, the genetic code is constructed on DNA molecules using 4 nitrogenous bases, 16 doublets, and 64 triplets. Structural analogies of Yin-Yang schemes of I-Ching with the alphabets of DNA have long been noted by various authors, including prominent geneticists: Stent G.S. in 1969 and Nobel laureate F. Jacob in 1977. More details about such analogies one can read in publications [Hu, Petoukhov, Petoukhova, 2017; Petoukhov, 1999, 2008; Petoukhov, He, 2010].

A great number of literature sources are devoted to I-Ching. Many of these sources label I-Ching as one of the greatest and most mysterious human creations. From the viewpoint of the Chinese culture (the most ancient among all cultures, which continue their existence on the Earth), I-Ching represents something even more

considerable: the creation made by the Superperson, who embodied a secret of the Universe in special symbols and signs. This book has a fundamental paradigmatic influence on the whole culture of traditional China and the adjacent countries. The ideas expressed in it have created an original world-view and methodology. They have influenced to a huge extent the development of philosophy, religion, natural sciences, literature, and arts in Ancient China. The symbols and principles of I-Ching penetrated into all spheres of life of traditional China from theoretical conceptions and high art to household subjects and decorations. Confucius said: "*If it would be possible to extend my years, I would have given fifty of them for the study of I-Ching*" [Shchutskii, 1997]. I-Ching declares a universality of the cyclic principle of organization in nature. Traditional Oriental medicine is based on the viewpoints of this book.

Many western scientists studied and used I-Ching. For example, the creator of analytical psychology C. Jung developed his doctrine about the collective unconscious in connection with this book. According to Jung and his fellow campaigner Nobel laureate in physics W. Pauli, the trigrams and the hexagrams of I-Ching "*fix a universal set of archetypes (innate psychic structures)*" [Shchutskii, 1997, p. 12]. Niels Bohr chose the Yin-Yang symbol as his personal emblem. Many modern physicists, who feel the unity of the world, connect their theories with the ideas of traditional Oriental culture, which unite all nature. For example, this has been reflected in the title "the eightfold way" of the famous book of [Gell-Mann, Ne'eman, 1964]. The intensive development of the self-organizing and nonlinear dynamics of complex systems (synergetics) promotes the strengthening attention of western scientists to the traditional eastern world-view (e.g., see [Capra, 2000]). Special groups study I-Ching in many eastern and western universities. Several conferences "I-Ching and modern sciences" were conducted in Moscow, Russia. A great number of sites on the Internet are devoted to similar studies. The influence of I-Ching is widely presented in the modern life of Eastern countries. For example, the national flag of South Korea bears the symbols of trigrams. In a great number of specialized schools, the instructors teach the students about methodological aspects of the practical application of relevant ancient knowledge in medicine and in other fields.

According to Ancient Chinese, "*trigrams, hexagrams and their components in all possible combinatory combinations form a universal hierarchy of classification schemes. These schemes in visual patterns embrace any aspects of reality - spatial parts, time intervals, the elements, numbers, colors, body organs, social and family conditions, etc.*" [Shchutskii, 1997, p. 10]. The Ancient Chinese claimed that this table of 64 hexagrams (Fig. 1) is the universal archetype of nature. The Ancient Chinese knew nothing about the genetic code, but the genetic code is arranged following the I-Ching in many aspects.

One should note that I-Ching was written several thousand years before the occurrence of modern Academies of sciences. It represents a set of statements of unclear origin. From the point of view of modern science, the book justifies a universal conformity of these statements to the structure of nature without their appropriate substantiation. Historically, the western academic science and its scientific laws (for example, Newton's laws, etc.) were developed without any connection with I-Ching using another methodology, although inspirational conditions, in which the person suddenly receives the complete picture of the answer to his questions, are well-known in western science as well.

In I-Ching, broken and unbroken (solid) lines symbolize Yin and Yang correspondingly. The main table in the I-Ching book is the famous square table of 64

hexagrams in Fu-Xi's order (Fig. 1), which Chinese tradition considers as a universal natural archetype. Each hexagram is a pile of six broken and unbroken (solid) lines. According to the western tradition, these broken and unbroken lines are shown in the form of the binary symbols "0" and "1" and each hexagram represents a sequence of such six binary symbols. Fig. AII.1 demonstrates hexagrams in both of these forms. Each position in all hexagrams has its own individual number: in the western numeric presentation of a hexagram, positions of its binary symbols are numbered left-to-right by the numbers from 1 to 6; in the Chinese graphical presentation, numbering of the lines of each hexagram is read in the sequence bottom-up.



**Fig. AII.1**. The table of 64 hexagrams in Fu-Xi's order where each of hexagrams is shown in the symbolic Chinese tradition and also in a binary representation: each solid or broken line corresponds to 1 or 0 correspondingly. Chinese names of trigrams are also shown.

In the table of 64 hexagrams (Fig. AII.1), 8 trigrams (three-digit binary numbers or piles of three lines) indicate its rows and columns. Each of the hexagrams is a concatenation of 2 trigrams: the trigram numerating its row and the trigram numerating its column. It should be emphasized that exactly according to a similar scheme of column and row numbering by ternary digital symbols, the (8*8)-matrix of 64 triplets of the genetic code in Fig. 2.1 in this article was built. By this analogy, each of the 64 genetic triplets in Fig. 2.1 can be denoted by a 6-bit binary number; for example, the triplet CAC is denoted by 111101 [Petoukhov, 2008; Petoukhov, He, 2010]. The traditional Chinese point of view is the following: "*Hexagrams are not*

*trigrams, which are alloyed together, but they are two trigrams, which are located on a vertical one over another*” [Shchutskii, 1997, p. 101].

Let us study some analogies between the table of 64 hexagrams (Fig. AII.1) and genetic mosaic (8*8)-matrices (Figs. 7.9-7.11), which present the arrangement of *m*-positional tetra-groupings of 64 triplets in them.  A key element for identifying these analogies is the following ancient Chinese doctrine about the important role of pairs of the correlative positions 1-4, 2-5, and 2-6 in Chinese hexagrams: “*The theory of I-Ching considers that a bottom trigram concerns an internal life… and a top trigram concerns to an external world…. Similar positions in a top trigram and in a bottom trigram have the nearest relation to each other. Because of this, the first position relates by analogy to the fourth position, the second position – to the fifth position, and the third position relates by analogy to the sixth position…. If these correlative positions (1-4, 2-5, 3-6) are occupied by various lines, it is considered that “conformity exists” between them, and in the case when these correlative positions are occupied by identical lines, it is considered that “conformity is absent” between them*" [Shchutskii, 1997, p. 86].

Figs. AII.2-AII.4 show the table of 64 hexagrams (Fig. AII.1) with European denotations of the hexagrams by binary numbers. Pairs of binary digits that are in hexagrams at correlative positions 1-4, 2-5, and 3-6 are highlighted in 4 colors: blue color corresponds to a pair of digits 1 and 1; red color corresponds to a pair of numbers 1 and 0; yellow corresponds to a pair of digits 0 and 1; green corresponds to a pair of numbers 0 and 0. These 4 pairs of digits, standing in the mentioned positions, define appropriate tetra-groupings of 64 hexagrams, which are called the 1-positional tetra-grouping of hexagrams (it relates to the correlative positions 1-4), the 2-positional tetra-grouping (it relates to the correlative positions 2-5), and the 3-positional tetra-grouping (it relates to the correlative positions 3-6).

In each of Figs. AII.2-AII.4, the arrangement of these 4 colors at correlative positions (1-4, 2-5, and 3-6) in the 64 hexagrams completely coincide with the arrangement of the 4 symbols ♠, ♥, ♦, and ♣ in the shown genetic matrices, which are connected with the described genomic Gestalt phenomena of probabilities and which present the arrangement of 64 triplets in the cases of the 1-positional tetra-groupings (Fig. AII.2), 2-positional tetra-groupings (Fig. AII.3), and 3-positional tetra-groupings (Fig. AII.4).

More precisely, in all these mutually corresponding *m*-positional tetra-groupings of   64 triplets and 64 hexagrams, the symbol ♠ corresponds to the hexagram pair of digits 1 and 1; the symbol ♥ corresponds to the hexagram pair 1 and 0; the symbol ♦ corresponds to the pair 0 and 1; the symbol ♣ corresponds to the pair 0 and 0.



**Fig. AII.2.** Identity between the arrangements of the 1-positional tetra-grouping of 64 genetic triplets in the genetic matrix from Fig. 7.9, and the arrangements of a 1-positional tetra-grouping of 64 hexagrams in the ancient Chinese table of I-Ching; this tetra-grouping is defined by four binary pairs (1 and 1, 1 and 0, 0

and 1, 0 and 0) standing inside hexagrams in the traditional correlative positions 1-4. Each of the 4 groupings is marked by its own color.



**Fig. AII.3**. Identity between the arrangements of the 2-positional tetra-grouping of 64 genetic triplets in the genetic matrix from Fig. 7.10, and the arrangements of a 2-positional tetra-grouping of 64 hexagrams in the ancient Chinese table of I-Ching; this tetra-grouping is defined by four binary pairs (1 and 1, 1 and 0, 0 and 1, 0 and 0) standing inside hexagrams in the traditional correlative positions 2-5. Each of the 4 groupings is marked by its own color similar to Fig. AII.2.
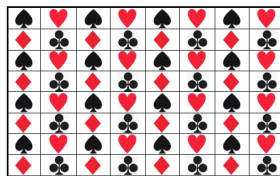


**Fig. AII.4**. Identity between the arrangements of the 3-positional tetra-grouping of 64 genetic triplets in the genetic matrix from Fig. 7.11, and the arrangements of a 3-positional tetra-grouping of 64 hexagrams in the ancient Chinese table of I-Ching; this tetra-grouping is defined by four binary pairs (1 and 1, 1 and 0, 0 and 1, 0 and 0) standing inside hexagrams in the traditional correlative positions 3-6. Each of the 4 groupings is marked by its own color similar to Figs. AII.2 and AII.3.

The parallels, which are shown in Figs. AII.2-AII4, give pieces of evidence that the discovered genomic Gestalt phenomena of probabilities are structurally connected with the Yin-Yang schemes of I-Ching including the table of 64 hexagrams. Here it can be again recalled that a creator of analytical psychology C.Jung and his fellow campaigner Nobel laureate in physics W. Pauli, who were experts regarding I-Ching, believed that the trigrams and the hexagrams of I-Ching *"fix a universal set of archetypes (innate psychic structures)"* [Shchutskii, 1997, p. 12]. Correspondingly all these data confirm in particular that the discovered genetic Gestalt phenomena and the described genetic Gestalt rules have deep relations to properties of genetically inherited psychological phenomena and to a rich theme of nature's archetypes.

One can additionally mention that matrices of 4 bigrams, 16 tetragrams, and 64 hexagrams of I-Ching can be represented as interrelated members of a tensor

family of matrices by analogy with the tensor family of genetic matrices of *n*-plets in Fig. 2.1 (see detail in [Petoukhov, 2008; Petoukhov, He, 2010]). Such tensor families of different ($2^n*2^n$)-matrices allow searching new "*innate psychic structures*" and nature's archetypes.

I-Ching is also called the "Book of Cyclical Changes". There is a connection between various ensembles of binary numbers in the table of 64 hexagrams based on the logical operation of modulo-2 addition. For example, ensembles of binary numbers from different parts of the tetra-grouping indicated in Fig. AII.2 can be transformed into each other based on this logical operation. In particular, those two ensembles of binary numbers of the table 64 hexagrams, whose locations correspond to the location of the symbols ♠ and ♦ in the mosaic matrix, which represents the 1-positional tetra-grouping of 64 genetic triplets in Fig. AII.2, are transformed into each other by the logical operation of modulo-2 addition of the binary number 100000 with members of the ensembles. The study of cyclical binary relationships between various tetra-groupings in matrices representing genetic Gestalt phenomena is a separate interesting topic.

The table of 64 hexagrams is also connected with the described mosaic matrices of projector tetra-groupings of triplets but this connection is more difficult to see and describe. A system of cyclical binary relationships, based on the logical operation of modulo-2 addition, also exists between various projector tetra-groupings and should be studied in the future.

### Appendix III. Extensions of split-quaternion algebra in connection with bit-reversal permutations in mosaic genetic matrices

Above in Section 16.1, we considered alphabetical mosaic matrices [C, A; T, G]$^{(n)}_{\text{bit-reversal}}$ (Fig. 16.13), which were constructed from the alphabetic mosaic matrices [C, A; T, G]$^{(n)}$ by the holographic method of bit-reversal permutations inside binary numeration of their columns and rows. Appropriate numeric multi-block matrices arise, which are colonies or unions of the same numeric (4*4)-matrix (α-matrix) as their general block (Figs. 16.1.4, 16.1.5). The α-matrix reflects structural features of the DNA-alphabet of 16 doublets and is a matrix representation of split-quaternions, which was demonstrated by the special decomposition of α-matrix into 4 sparse matrices (Fig. 16.1.5).

This Appendix is devoted to a similar analysis of the (8*8)-matrix, which reflects structural features of the DNA-alphabet of 64 triplets and is a colony or union of 4 α-matrices (Figs. 16.1.4 and 16.1.6). Fig. AIII.1 shows this matrix and its decomposition into 8 sparse matrices $k_0+k_1+k_2+k_3+k_4+k_5+k_6+k_7$, the set of which is closed relative to multiplication and defines a corresponding multiplication table.

Similarly, it is possible to analyze dyadic-shift decompositions of mosaic numerical ($2^n*2^n$)-matrices representing the structural features of the DNA alphabets of tetraplets, pentaplets, etc., each time obtaining the corresponding block multiplication table for a set of sparse matrices of this decomposition. These data additionally draw attention to the algebraic nature of the structures of genetic informatics.

The action of such ($2^n*2^n$)-matrices obtained using the algorithm (16.1) on a $2^n$-dimensional vector with arbitrary coordinates generates a vector with a batch (cluster) binary-oppositional organization of coordinates. For example, if unrelated coordinates of an initial state vector of a system are cyclic or wave functions, then the

generated new state vector has block-structured sets of coordinates whose blocks are equal to each other up to the sign, and their cyclic changes are synchronized in time (that is, cyclic behavior of the considered system in its different configuration subspaces becomes strongly interrelated).

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

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

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

$$=k_0+k_1+k_2+k_3+k_4+k_5+k_6+k_7$$

| * | $k_0$ | $k_1$ | $k_2$ | $k_3$ | $k_4$ | $k_5$ | $k_6$ | $k_7$ |
|---|---|---|---|---|---|---|---|---|
| $k_0$ | $k_0$ | $k_1$ | $k_2$ | $k_3$ | $k_4$ | $k_5$ | $k_6$ | $k_7$ |
| $k_1$ | $k_1$ | $k_0$ | $k_3$ | $k_2$ | $k_5$ | $k_4$ | $k_7$ | $k_6$ |
| $k_2$ | $k_2$ | $-k_3$ | $-k_0$ | $k_1$ | $k_6$ | $-k_7$ | $-k_4$ | $k_5$ |
| $k_3$ | $k_3$ | $-k_2$ | $-k_1$ | $k_0$ | $k_7$ | $-k_6$ | $-k_5$ | $k_4$ |
| $k_4$ | $k_4$ | $k_5$ | $k_6$ | $k_7$ | $k_0$ | $k_1$ | $k_2$ | $k_3$ |
| $k_5$ | $k_5$ | $k_4$ | $k_7$ | $k_6$ | $k_1$ | $k_0$ | $k_3$ | $k_2$ |
| $k_6$ | $k_6$ | $-k_7$ | $-k_4$ | $k_5$ | $k_2$ | $-k_3$ | $-k_0$ | $k_1$ |
| $k_7$ | $k_7$ | $-k_6$ | $-k_5$ | $k_4$ | $k_3$ | $-k_2$ | $-k_1$ | $k_0$ |

**Fig. AIII.1**. The decomposition of the numeric multi-block (8*8)-matrix, which reflects structural properties of the DNA-alphabet of 64 triplets and contains 4 α-matrix as its blocks (see Figs. 16.1.4 and 16.1.6). Each of the black (white) cells in the presented matrix corresponds to number +1 (-1). This (8*8)-matrix is a sum of shown 8 sparse matrices: $k_0+k_1+k_2+k_3+k_4+k_5+k_6+k_7$. The multiplication table of these 8 matrices, which is shown in the bottom row, has a block structure: its quadrants along each of the diagonals, are identical.

**Acknowledgments**

**References.**

**Abbott D., Davies P.C.W., Pati A.K**. (eds.), Foreword by Sir Roger Penrose. *Quantum Aspects of Life* (2008). ISBN-13: 978-1-84816-253-2

**Ahmed N.U., Rao K.R**. *Orthogonal transforms for digital signal processing*. Springer-Verlag New York, Inc. (1975).

**Albrecht-Buehler G**. Asymptotically increasing compliance of genomes with Chargaff's second parity rules through inversions and inverted transpositions. *Proc Natl Acad Sci U S A*. November 21; 103(47): 17828–17833 (2006).

**Banal J., Shepherd T., Berleant J., Huang H., Reyes M., Ackerman Ch.M.,**

**Blainey P.C., Bathe M.** Random access DNA memory using Boolean search in an archival file storage system. *Nature Materials*, June 2021, DOI: 10.1038/s41563-021-01021-3.

**Belousov L**. *Morphomechanics of Development*. Springer International Publishing Switzerland, NEW YORK, USA, 195 p.) (Driesch H (1921) Philosophie des Organischen. Engelmann, Leipzig (2015).

**Berend D., Dolev Shl., Frenkel S., Hanemann A**. Towards holographic ''brain'' memory based on randomization and Walsh–Hadamard transformation. *Neural Networks*, 77, p.87–94 (2016).

**Bernstein N.A**. *The co-ordination and regulation of movements*. Oxford: Pergamon Press (1967).

**Bodnar O.Ya**. Geometry of phyllotaxis. *Reports of the Academy of Sciences of Ukraine*, №9,  pp. 9-15 (1992).

**Bodnar O.Ya**. *Golden Ratio and Non-Euclidean Geometry in Nature and Art*. Lviv: Publishing House "Sweet" (1994).

**Capra F**. *The Tao of Physics: an Exploration of the Parallels between Modern Physics and Easterm Mysticism*. Shambala, Boston USA.

**Chandak Sh., Neu J., Tatwawadi K., Mardia J., Lau B., Kubit M.,**

**Hulett R., Griffin P., Wootters M., Weissman Ts., Ji H**. Overcoming high nanopore basecaller error rates for DNA storage via basecaller-decoder integration and convolutional codes. - bioRxiv preprint first posted online Dec. 20, 2019; doi: http://dx.doi.org/10.1101/2019.12.20.871939.

**Chargaff E**. Preface to a Grammar of Biology: A hundred years of nucleic acid research. - *Science*, 172, p. 637-642 (1971).

**Cook T.A**. *The Curves of Life*. London: Constable and Co, 1914.

**Dolev S**., **Frenkel S**. Multiplication free holographic coding. In: 2010 IEEE 26th convention of electrical and electronics engineers in Israel, IEEEI, p.146-150

(2010).

**Derzhypolskyy A., Melenevskyy D., Gnatovskyy A**. A Comparative Analysis of Associative Properties of Fourier versus Walsh Digital Holograms. - *Acta Physica Polonica A*, vol. 112, No. 5, pp. 1101-1106. Proceedings of the International School and Conference on Optics and Optical Materials, ISCOM07, Belgrade, Serbia, September 3-7, 2007 (2007).

**Dolev S., Frenkel S.L., Cohen A**. Holographic coding by Walsh–Hadamard transformation of randomized and permuted data. *Inform. Primen*., 2012, Vol. 6, Issue 4, p. 76–83, http://www.mathnet.ru/links/556fb0a28ccf40fbc4fbc89c4c02f750/ia236.pdf

**Driesch H.** *Philosophie des Organischen*. Engelmann, Leipzig (1921).

**Fimmel E., Petoukhov S.V**. Development of Models of Quantum Biology Based on the Tensor Product of Matrices. In: Hu Z., Petoukhov S., He M. (eds). Advances in Intelligent Systems and Computing, v. 1126, p.126-135. Springer, Cham (2020), DOI https://doi.org/10.1007/978-3-030-39162-1_12.

**Fimmel E., Strüngmann L**.Yury Borisovich Rumer and his 'biological papers' on the genetic code. Phil. Trans. R. Soc. A, 374: 20150228 (2016). http://dx.doi.org/10.1098/rsta.2015.0228 .

**Fine, N.J**. On the Walsh functions. *Transactions of the American Society*, 3, pp. 372–414 (1949).

**Frank-Kamenetskiy M.D**. The most principal molecule. Moscow, Nauka (1988) (in Russian).

**Geesink H.J.H., Meijer D.K.F**. Quantum wave information of life revealed. *NeuroQuantology*, vol. 14, issue 1 (March 2016), doi:10.14704/nq.2016.14.1.911.

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

**Gold B., Rader C**. M. *Digital Processing of Signals*. McGraw–Hill: New York, USA (1969).

**Golubov B.I., Efimov A.V., Skvortsov V.A**. *Series and Fourier transforms*. Moscow: Nauka (1987) (in Russian).

**Gurvich A.G**. *Selected Works*. Moscow: Medicine (1977).

**Harkin A.A., Harkin J.B**. Geometry of Generalized Complex Numbers. *Mathematics Magazine*, v. 77(2), p. 118-29 (2004).

**Hofstadter D.R**. Gödel, Escher, Bach. An Eternal Golden Braid. N.Y.: Vintage Books, 777 p. (1980).

**Hu Z.B., Petoukhov S.V., Petukhova E.S**. I-Ching, dyadic groups of binary numbers and the geno-logic coding in living bodies. *Progress in Biophysics and Molecular Biology*, vol. 131, pp. 354-368 (December 2017).

**Igamberdiev A.U**. Quantum mechanical properties of biosystems: a framework for complexity, structural stability, and transformations. *Biosystems*, 31(1), p. 65–73 (1993).

**Jeffrey H.J**. Chaos game representation of gene structure. *Nucleic Acids Research*, Vol. 18, No. 8, p. 2163-2170 (1990).

**Jordan P.** Die Quantenmechanik und die Grundprobleme der Biologie und Psychologie. *Naturwissenschaften* 20, 815–821 (1932), (doi:10.1007/BF01494844)

**Kantor I.L., Solodovnikov A.S**. *Hypercomplex numbers*. Berlin, New York: Springer-Verlag (1989). ISBN 978-0-387-96980-0.

**Karp A.H**. Bit reversal on uniprocessors. *SIAM Review*, v. 38 (1), p.1–26 (1996),

doi:10.1137/1038001, MR 1379039.

**Karpovsky M.G., Stankovic R.S., Astola J.T**. Spectral Logic and its Applications for the Design of Digital Devices. - New Jersey: John Wiley & Sons, 598 p. (2008).

**Karzel H., Kist G**. Kinematic Algebras and their Geometries, in *Rings and Geometry*, R. Kaya, P. Plaumann, and K. Strambach editors, pp. 437–509, esp 449,50, D. Reidel, (1985). ISBN 90-277-2112-2.

**Kienle G**. Experiments concerning the non-Euclidean structure of visual space. In Bioastronautics; Pergamon Pree: New York, NY, USA, 1964; pp. 386–400.

**Kulakov Yu.I**. *The theory of physical structures*. Novosibirsk: Alta Vista (2004). 847 p. (in Russian).

**Kulakov Yu.I**. The theory of physical structures as the basis of mathematics and physics. *Metaphysics*, No. 1 (27), p. 49-53 (2018, in Russian).

**Lashley R.S**. *Brain mechanisms and intelligence*. University of Chicago Press (January 1, 1929), ISBN-10 : 1135563772, ISBN-13 : 978-1135563776.

**Luneburg R**. The metric of binocular visual space. *J. Opt. Soc. Am*. 1950, 40, 627–642.

**Lyons R**. Understanding Digital Signal Processing. Pearson; 3rd edition, 954 pages, (2010), ISBN-10: 0137027419, ISBN-13: 978-0137027415.

**Matsuno K**. Cell motility as entangled quantum coherence. *BioSystems*, 51, p. 15–19 (1999).

**Matsuno K., Paton R.C**. Is there a biology of quantum information? *BioSystems,* 55, p. 39–46 (2000).

**McFadden J., Al-Khalili J**. The origins of quantum biology. *Proceedings of the Royal Society A*, Vol. 474, Issue 2220, p. 1-13, 12 December 2018, https://doi.org/10.1098/rspa.2018.0674.

**Morita Y**., **Sakurai Y**. Holography by Walsh Waves. *Proceedings of the IV Symposium «Application of Walsh functions» Held at the Catholic University of America*, Washington, D. C. on 16-18 April 1973, pp. 122-126 (1973).

**Nalimov V.V**. *I am scattering thoughts* (in Russian: *Razbrasyvaiu mysli*). Moscow, Center for Humanitarian Initiatives (2015). ISBN 978-5-98712-521-2.

**Pastawski F., Yoshida B., Harlow D., Preskill J.** Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *J. High Energ. Phys.*, 149 (2015). https://doi.org/10.1007/JHEP06(2015)149, https://link.springer.com/article/10.1007%2FJHEP06%282015%29149#citeas.

**Patel A**. Quantum algorithms and the genetic code. *Pramana – J. Phys*. 56(2–3), p. 367–381, arXiv:quant-ph/0002037 (2001a).

**Patel A**. Testing quantum dynamics in genetic information processing. *J. Genet*. 80(1), p. 39–43 (2001b).

**Patel A**. Why genetic information processing could have a quantum basis. *J. Biosci*. 26(2), p. 145–151 (2001c).

**Pauli W**. *Writings on Physics and Philosophy*. Ed. by C.P. Enz and K. von Meyenn, Springer, Berlin (1994).

**Penrose R**. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press, USA, 480 p. (1996).

**Petoukhov S.V**. Non-Euckidean geometries and algorithms of living bodies. *Computers and Mathematics with Applications*, v. 17, № 4-6, p. 505-534 (1989).

**Petoukhov S.V**. Genetic code and the Ancient Chinese «Book of Changes». *Symmetry:Culture and Science*, vol. 10, №3-4, pp. 211-226 (1999).

**Petoukhov S.V.** *Matrix genetics, algebrases of genetic code, noise immunity*. Moscow, RCD, 316 p. (2008a, in Russian). ISBN 978-5-93972-643-6.

**Petoukhov S.V**. The degeneracy of the genetic code and Hadamard matrices. - HTTP://ARXIV:0802.3366, p. 1-26  (The first version is from February 22, 2008b).

**Petoukhov S.V.** Matrix genetics, part 2: the degeneracy of the genetic code and the octave algebra with two quasi-real units (the "Yin-Yang octave algebra").  - http://arXiv:0803.3330, p. 1-23. Retrieved  March 23 (2008c).

**Petoukhov S.V**. Symmetries of the genetic code,Walsh functions and the theory of genetic logical holography. Symmetry Cult. Sci., 27, p. 95–98 (2016a).

**Petoukhov S.V**. The Genetic Code, 8-Dimensional Hypercomplex Numbers and Dyadic Shifts. 11th Version of the Article. Available online: http://arxiv.org/abs/1102.3596 (accessed on 15 July 2016b).

**Petoukhov S.V.** The Genetic Coding System and Unitary Matrices. *Preprints* 2018, 2018040131 (2018), doi: 10.20944/preprints201804.0131.v2.

**Petoukhov S.V**. Nucleotide Epi-Chains and New Nucleotide Probability Rules in Long DNA Sequences. Preprints 2019, 2019040011, 17 pages, (2019), doi: 10.20944/preprints201904.0011.v1, https://www.preprints.org/manuscript/201904.0011/v1

**Petoukhov S.V**. Hyperbolic Rules of the Cooperative Organization of Eukaryotic and Prokaryotic Genomes. *Biosystems*, 198, 104273 (2020a).

**Petoukhov S.V**.  Hyperbolic Rules of the Oligomer Cooperative Organization of Eukaryotic and Prokaryotic Genomes. *Preprints* 2020, 2020050471 (2020b), doi:10.20944/preprints202005.0471.v2, https://www.preprints.org/manuscript/202005 .0471/v2.

**Petoukhov S.V**. The rules of long DNA-sequences and tetra-groups of oligonucleotides. arXiv:1709.04943v6 , 6th version from  22.05.2020 (2020c).

**Petoukhov S.V**. Modeling inherited physiological structures based on hyperbolic numbers, *BioSystems*, Vol. 199, 104285 (2021), ISSN 0303-2647, https://doi.org/10.1016/j.biosystems.2020.104285.

**Petoukhov S.V., He M**. *Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications*. Hershey, USA, IGI Global (2010).

 **Petoukhov S.V., Petukhova E.S**. Symmetries in genetic systems and the concept of geno-logical coding. - *Information*,  8(1), 2 (2017); doi:10.3390/info8010002, http://www.mdpi.com/2078-2489/8/1/2/pdf

**Petoukhov S.V., Petukhova E.S., Svirin V.I**. Symmetries of DNA alphabets and quantum informational formalisms. *Symmetry: Culture and Science,* Vol. 30, No. 2, p.161-179 (2019), https://doi.org/10.26830/symmetry_2019_2_161.

**Prabhu V. V**. Symmetry observation in long nucleotide sequences.  *Nucleic Acids Res*., 21, 2797-2800 (1993).

**Prasolov V.V**. *Problems and Theorems in Linear Algebra* (Translations of Mathematical Monographs, Vol. 134). American Mathematical Society, 225 pages (1994).

**Preskill J**. Stability, topology, holography: the many facets of quantum error correction. - *Presentation at American Physical Society*, 16 March 2016, 57 slides, http://theory.caltech.edu/~preskill/talks/APS-March-2016-preskill.pdf.

**Pribram K**. *Languages of the Brain; Experimental Paradoxes and Principles in Neuropsychology*. Englewood Cliffs, N. J.: Prentice-Hall (1971).

**Pribram K**. The holographic brain, (1998),
https://web.archive.org/web/20060518075852/http://homepages.ihug.co.nz/~sai/pribr
am.htm

**Reid M., Szendroi B**. *Geometry and topology*.§3.11. Hyperbolic motions. Cambridge
University Press, Cambridge (2005). ISBN: 978-0-521-84889-3.

**Reynolds W.F**. Hyperbolic geometry on a hyperboloid. *American Mathematical
Monthly*, 100: 442-455 (1993).

**Rumer Yu.B**. Codon systematization in the genetic code. *Doklady Akademii Nauk
SSSR, 183*(1), 225-226 (1968).

**Schrödinger E**. *What is life?* Cambridge Press (1944).

**Seberry J., Wysocki D.J., Wysocki T.A**. Some applications of Hadamard matrices.
*Metrika*, 62, p. 221-239 (2005).
https://www.uow.edu.au/~jennie/WEBPDF/2005_12.pdf

**Shapeville F., Haenni A. L**. *Biosynthese des proteins*. Paris: Hermann (1974).

**Shchutskii Y.K**. *The Chinese Classical "I Ching"*. Moscow, Vostochnaya literatura,
(1997, in Russian).

**Shiman D.V., Patsey N.V**. Application of methods of interleaving when transmitting
graphic information. - Proceedings of the Belarusian State Technological
University, No. 6. Physics and mathematics and informatics, pp. 108-110
(2013).

**Shishmintsev S**. Holographic properties of a bit-reversal permutation.
https://habrahabr.ru/post/155471/ ( accessed 22 October 2012)  (in Russian).

**Simonov A.A., Kulakov Y.I., Vityaev E.E.** On an algebraic definition of
laws.  *Journal of Mathematical Psychology*, vol. 58, p. 13-20 (January 2014),
ISSN 0022 2496.

**Sissa Medialab**. Non-Euclidean geometries for grid cells." ScienceDaily, 6 May
2015, www.sciencedaily.com/releases/2015/05/150506084636.htm

**Smolyaninov V.V**. Spatio-temporal problems of locomotion control. *Uspekhi
Fizicheskikh Nauk*, v.170,  N 10, pp. 1063–1128 (2000),  DOI:
https://doi.org/10.3367/UFNr.0170.200010b.1063.

**Soroko L.M**. "Experience" of Young-Walsh and Logical Holography. - *Materials of
6th national conference on holography* (11-17.02.1974, Yerevan), Leningrad
Institute of Nuclear Physics (1974),
http://bsfp.media- security.ru/school6/5.htm.

**Stambuk N**. Circular coding properties of gene and protein sequences. *Croat. Chem.
Acta*, 72, pp. 999-1008 (1999).

**Talbot M**. *The Holographic Universe*. HarperCollins Publishers Ltd., 352 p. (1996).

**Thurston W**. Three-dimensional manifolds, Kleinian groups and hyperbolic
geometry, Bull. Amer. Math. Soc. (N.S.) 6, 357–381 (1982).

**Thurston W**. Hyperbolic structures on 3-manifolds. I. Deformation of acylindrical
manifolds. *Ann. of Math.* (2) 124, no. 2, 203–246 (1986).

**Urdapilleta E., Troiani Fr., Stella F., Treves A.** Can rodents conceive hyperbolic
spaces? *Interfaces*, May 2015 DOI: 10.1101/015057.

**Vladimirov Yu.S**. *Leibniz-Mach relational concept*. Moscow: LENAND (2017, in
Russian).

**Vladimirov Yu.S**. *From geometrophysics to metaphysics: the development of
relational, geometric and field-theoretical paradigms in Russia at the end of the
20th - beginning of the 21st century. State and prospects*. Moscow: LENAND
(2019, in Russian).

**Vladimirov Yu.S**. Principles and consequences of relative world understanding.

*Foundations of the fundamental physics and* mathematics. Materials of the IV Russian Conference, Moscow, RUDN University, December 11-12, 2020 (edited by Yu.S. Vladimirov and V.A. Panchelyuga). - Moscow: RUDN, 244 pages (2020, in Russian). ISBN 978-5-209-10381-3.

**Vetter R.J., Weinstein S.**, The history of the phantom in congenitally absebt limbs. – Neuropsychologia, № 5, p. 335-338 (1967).

**Weinstein S., Sersen E.A**. Phantoms in cases of congenital absence of limbs. Neurology,  № 11, p. 905–911 (1961).

**Yang Q., Ellis J., Mamakani K., Ruskey F**. In-place permuting and perfect shuffling using involutions. *Information Processing Letters,* v. 113 (10-11), pp. 386-391 (2013)*,*  doi*:*10.1016/j.ipl.2013.02.017*,*  MR 3037467.