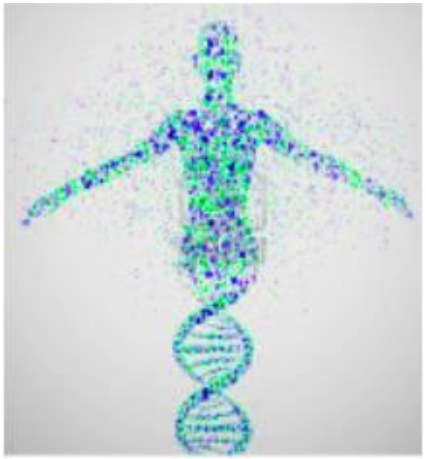
THE WINTER SESSION OF DAYALBAGH SCIENCE OF CONSCIOUSNESS.

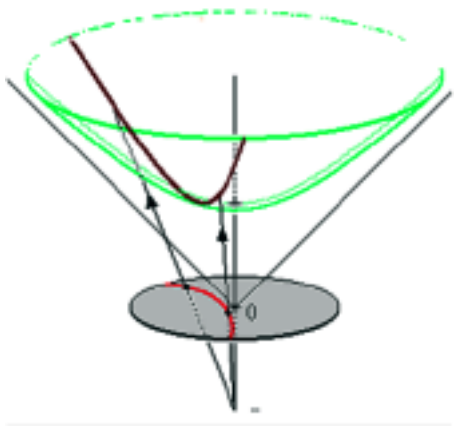
INDIA, 01.01.2022

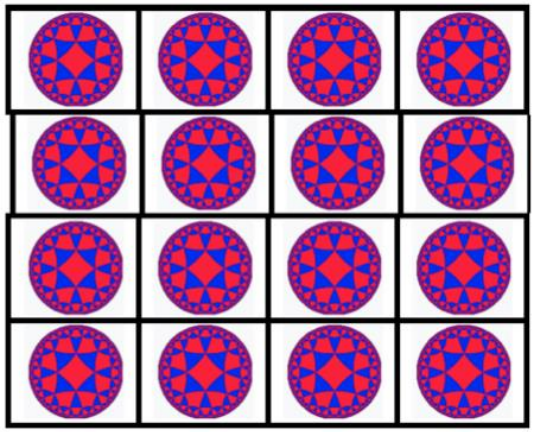
(conference website - www.dsc-dei.in).

Genetics and consciousness: algebraic holography, gestalt biology, and dualism “probability-vs-determinism”.

S.V. Petoukhov
Head of Lab of Biomechanical Systems, Institute of Mechanical Engineering Research Institute, Russian Academy of Sciences
http://petoukhov.com/ , spetoukhov@gmail.com







THE MAIN DIFFERENCE OF LIVING BODIES

The creators of quantum mechanics P. Jordan and (later) E. Schrödinger pointed out **the main difference between living bodies and inanimate ones**: in living bodies, **genetic molecules play a dictatorial role** (inanimate objects have no such dictatorial molecules since they are governed by the average random motion of millions of particles). The study of properties of DNA informatics is some of the key tasks of modern science (see the history of "quantum biology", 2018, https://royalsocietypublishing.org/doi/full/10.1098/rspa.2018.0674)

Pascual Jordan

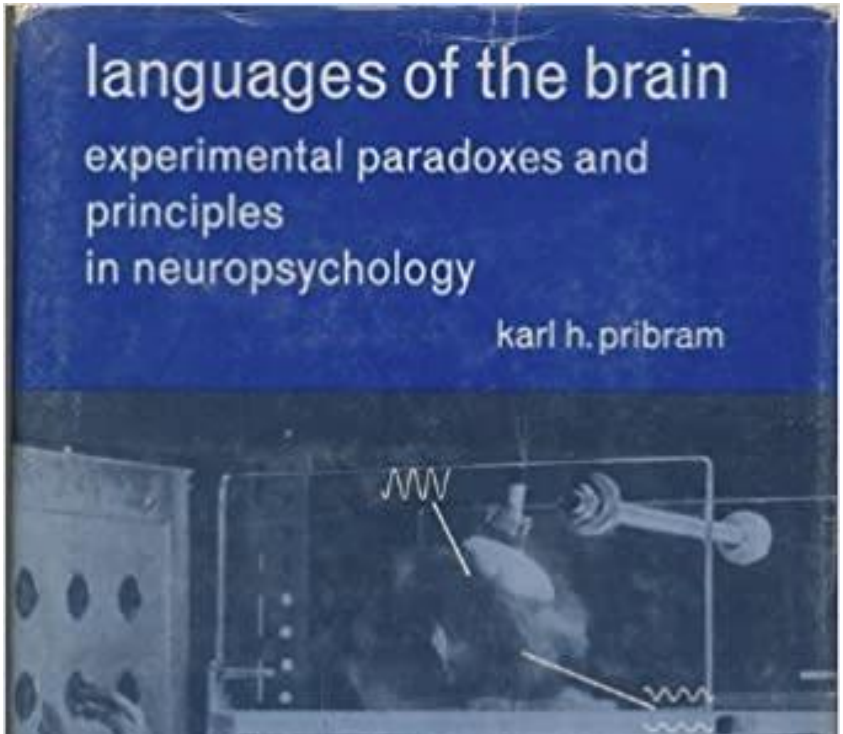
Erwin Schrödinger

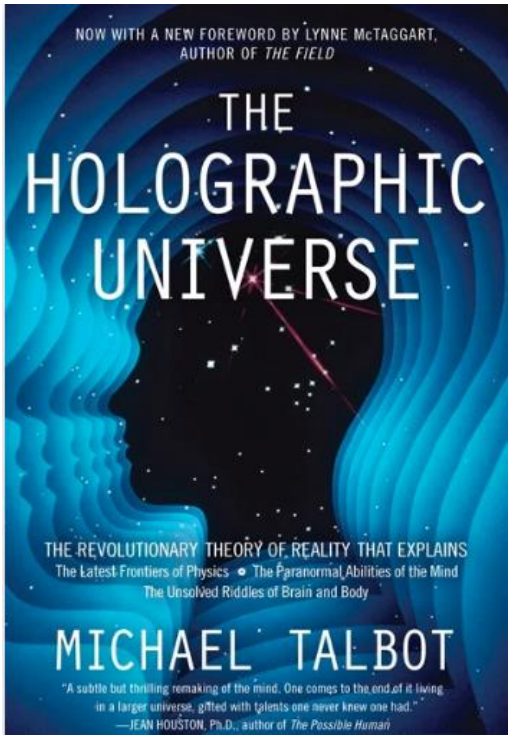
All living organisms are endowed in some degree with consciousness, the physiological foundations of which are genetically inherited. This consciousness provides for organisms the ability to perform actions of an intellectual nature to get food, save themselves from predators, etc.

When creating artificial intelligence systems, one should **look at the principles and patterns of genetic inheritance of such abilities from living nature. The talk is devoted to some of the results of such a peeping through the prism of algebra**.

It is known that living organisms have properties reminiscent of the properties of holography with its non-local informatics.

For example, K. Pribram's book "Languages of the Brain" about the holographic principles of the brain work emphasizes that the holographic description has no equal for explaining the problems of perception.

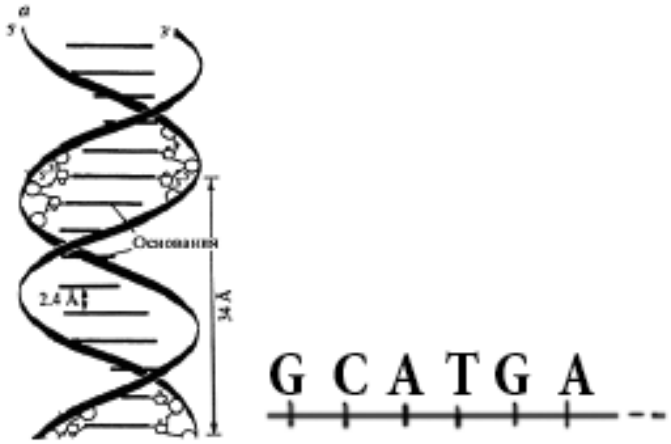


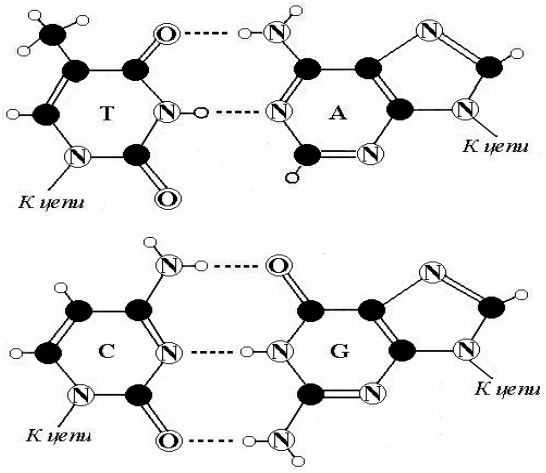


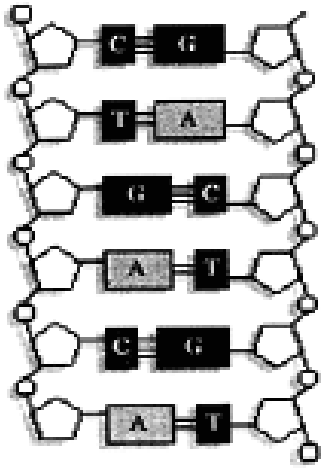
But all physiological structures are genetically inherited, and therefore it is natural to look for the original holographic properties in a structured genetic coding system, relying on the methods of algebraic holography from theoretical physics and digital informatics.

Let us recall some data on molecular genetic informatics of living organisms.

Genetic information is recorded in DNA molecules by means of long sequences of four nucleotides (molecular "letters") - adenine A, thymine T, cytosine C, and guanine G - and their combinations forming alphabets of 16 duplets, 64 triplets, 256 tetraplets, etc.

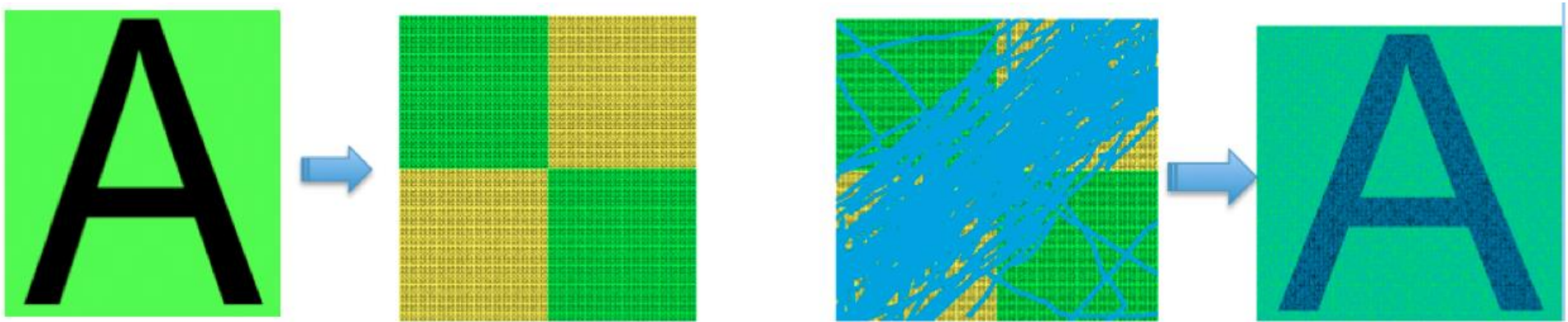




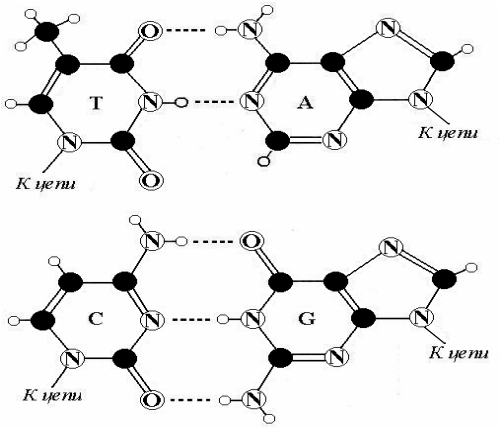


Let us show the consistency of the system of structured DNA alphabets (discovered by the speaker) with the well-known algebraic bit-reversible holography. Bit-reversion means reading binary numbers in the opposite direction: for example, 001 becomes 100 (which corresponds to decimal numbers 1 and 5).

The Figure shows an image A inside a 512*512 pixel matrix, in which all columns and rows are sequentially numbered with binary numbers. Reading each of these binary numbers in reverse order leads to renumbering of columns and rows and their new placement with a corresponding transformation of the entire image (on the second frame). If now part of this frame is corrupted, then reapplying the bit-reverse reading with a permutation will restore the original image in an identifiable form (the example is taken from.https://habrahabr.ru/post/155471/).



As it is known, the four DNA-nucleotides A, T, C, and G are endowed with binary-oppositional traits:
- 1) two of these nucleotides are purines (A and G) and the other two (C and T) are pyrimidines. From the standpoint of these oppositional indicators: **C = T = 1, A = G = 0** ;
- 2) two of these nucleotides are keto-molecules (T and G) and the other two (C and A) are amino molecules, which give a representation **C = A = 1, T = G = 0**.



This symmetric property allows representing DNA alphabets of 4 letters, 16 duplets, 64 triplets, 256 tetraplets, etc. in the form of square tables, whose columns are numbered with binary indicators "pyrimidine or purine" (**C = T = 1**, **A = G = 0**) , and whose rows - with binary indicators "amino or keto" (**C = A = 1**, **T = G = 0**). In such tables, all nucleotides, duplets, triplets, etc. automatically take their strictly individual place. These DNA alphabet tables turn out to be members of the tensor family of alphabetical matrices [C, A; T, G]$^{(n)}$ (the tensor product of matrices plays important role in quantum mechanics and quantum informatics).

	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

The phenomenological fact of the binary oppositional division of the DNA alphabet of 64 triplets - according to their code properties - into two equal sub-alphabets is known: 32 triplets with strong roots (i.e. triplets beginning with 8 duplets CC, CT, CG, AC, TC, GC, GT, GG) and 32 triplets with weak roots (triplets beginning with the other 8 duplets) [Yu.B.Rumer, 1975]. In the constructed matrices of DNA alphabets of 16 duplets, 64 triplets, and 256 tetraplets, we mark *n*-plets with strong roots in black:

	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

The bit-reversible permutation in these mosaic matrices generates a family of multi-block matrices whose mosaic consists of repeating a mosaic of a (4*4)-matrix of 16 duplets:

	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

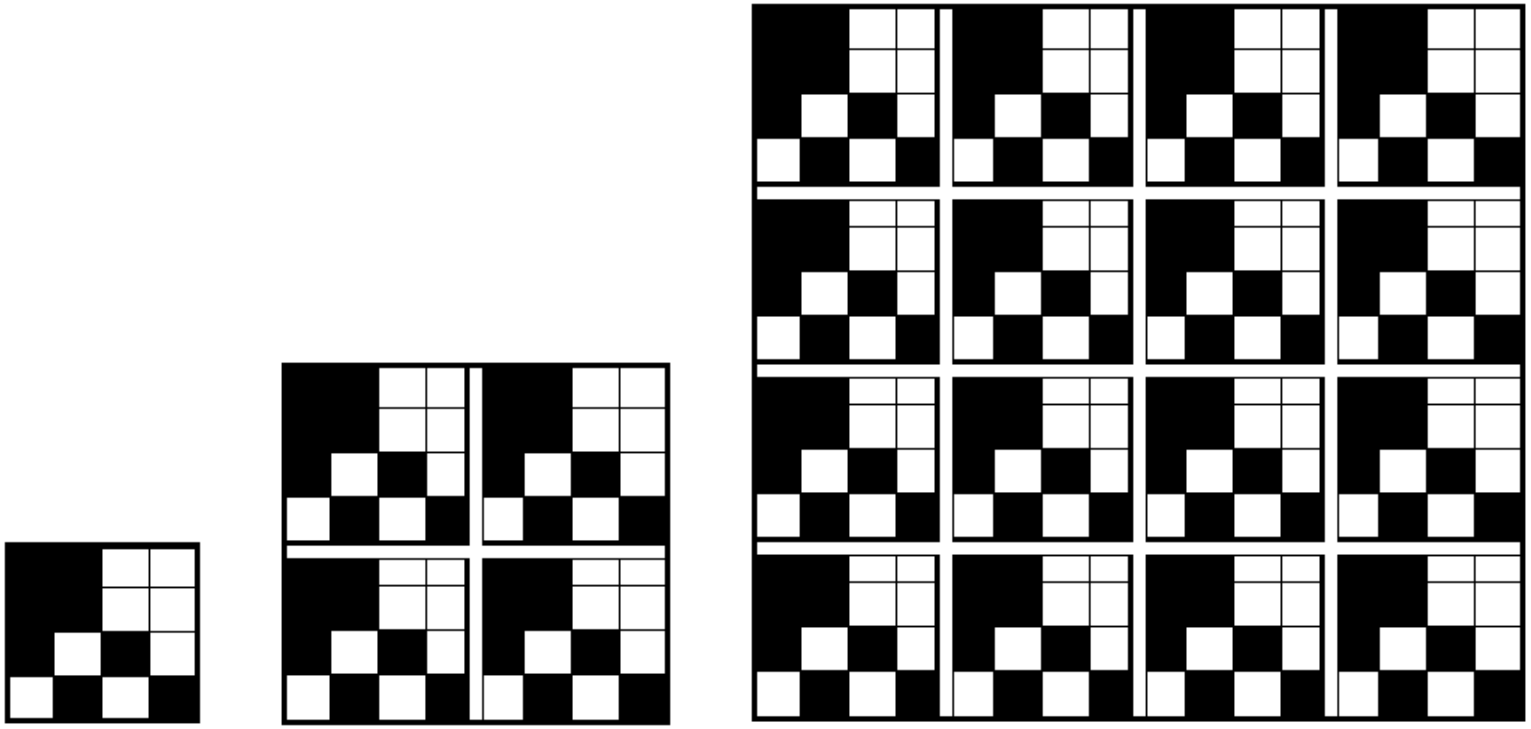
→

	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	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	ACCTT	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

Black and white cells of the matrices can be represented by elements +1 and -1 in them. The mosaics of all emerging DNA-alphabetical matrices are a repetition of the mosaic of the alphabet matrix of 16 duplets, forming "matrix crystals".

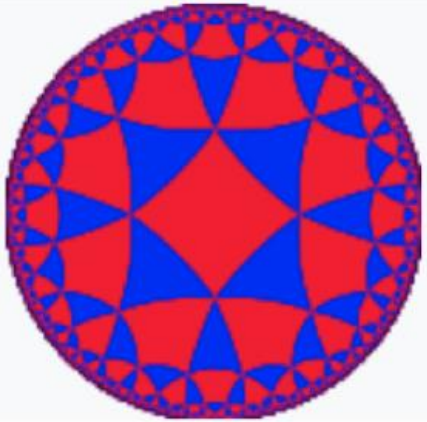
Does this "universal" repeating block (showed in the bottom at left) have a considerable algebraic meaning? Yes, it has.

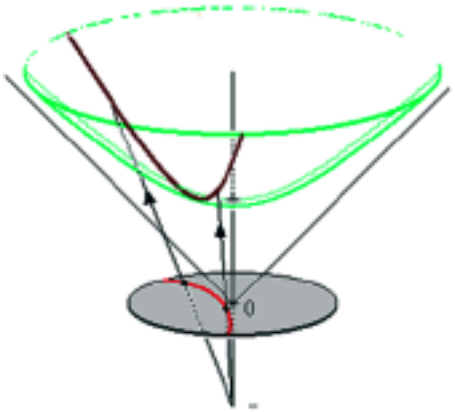


1	1	-1	-1
1	1	-1	-1
1	-1	1	-1
-1	1	-1	1

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

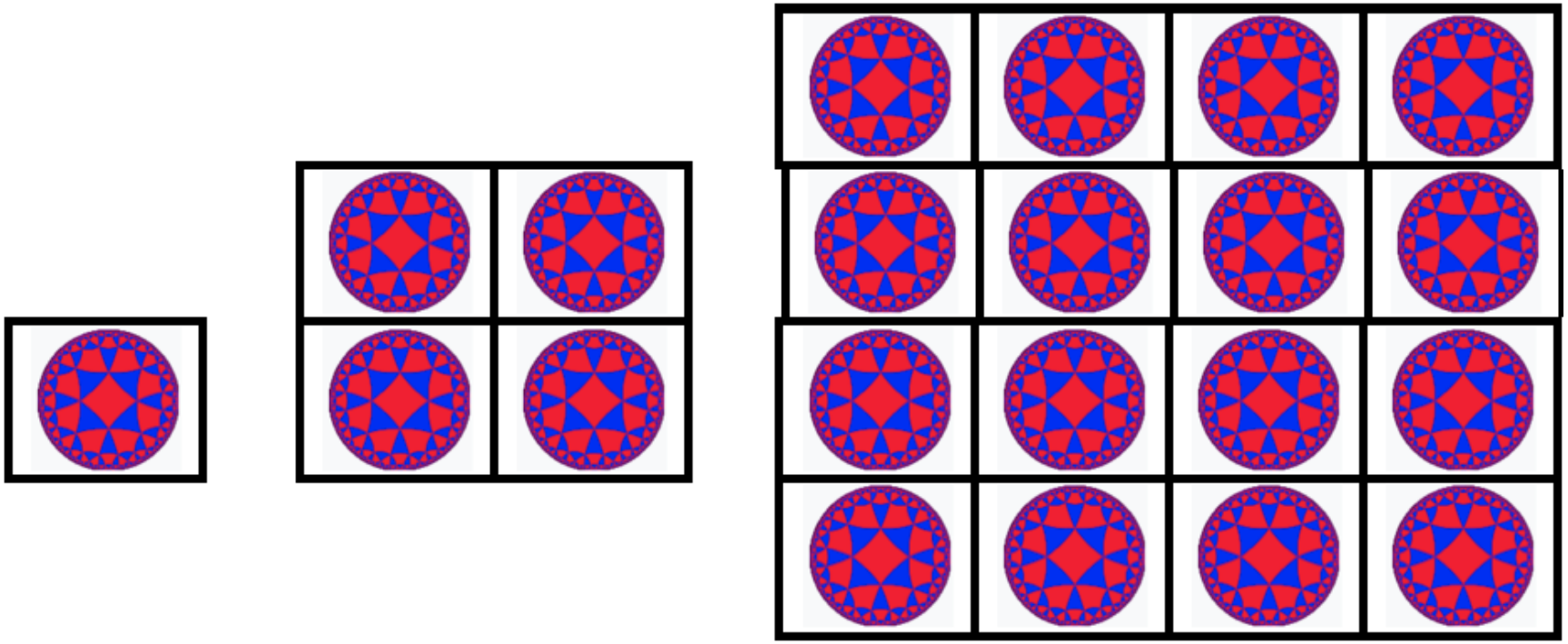
*	j_0	j_1	j_2	j_3
j_0	j_0	j_1	j_2	j_3
j_1	j_1	$-j_0$	j_3	$-j_2$
j_2	j_2	$-j_3$	j_0	$-j_1$
j_3	j_3	j_2	j_1	j_0



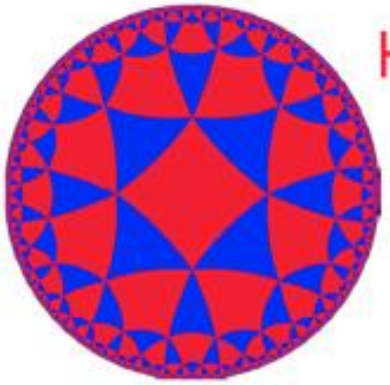


It turns out that this block is the sum of shown 4 sparse matrices, whose set is closed under multiplication and determines the corresponding multiplication table, known as the multiplication table of the algebra of 4-dimensional split-quaternions of Cockle (1849). Split-quaternions are used in the Poincaré disk model to describe hyperbolic motions in Lobachevsky's hyperbolic geometry (the symbol of the Poincaré model is shown).

Using this symbol of the Poincaré disk model of hyperbolic geometry, the resulting block mosaic matrices of structured DNA alphabets are represented in an artistic form to facilitate heuristic associations:



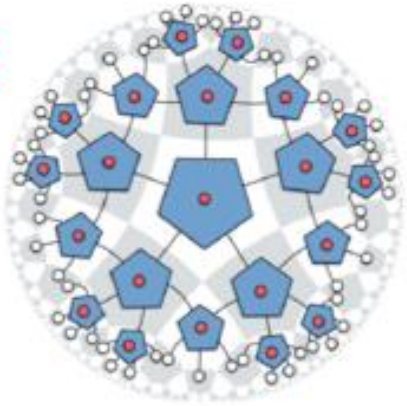
The family of block-unified matrices of DNA-alphabets, whose blocks are coupled with the disk model of hyperbolic geometry, unexpectedly echoes the theme of "**holographic quantum codes that correct errors**". This topic is being developed at the California Institute of Technology USA in connection with the same Poincaré disk model and its tiling:



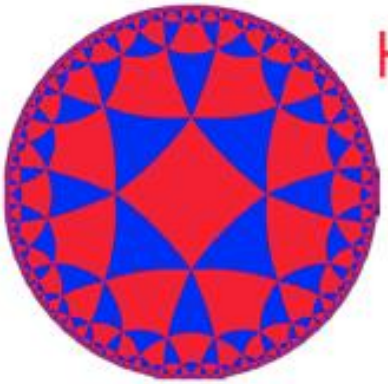
Holographic correspondence

Quantum error correction

Are they closely related?



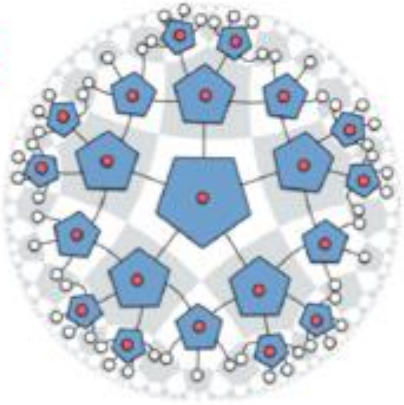
(From a presentation at the American Physical Society 2016http://theory.caltech.edu/~preskill/talks/APS-March-2016-preskill.pdf). The combination of these directions suggests itself.



Holographic correspondence

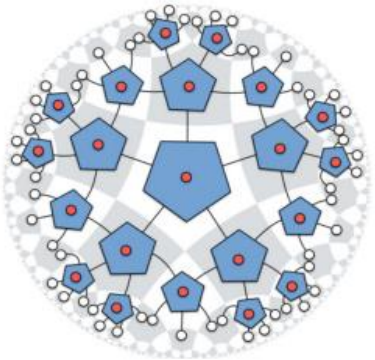
Quantum error correction

Are they closely related?

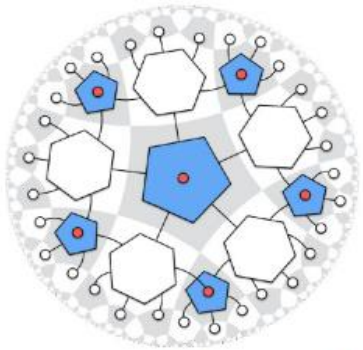


This holographic topic is headed by J.Preskill, director of the Caltech Institute for Quantum Information and Matter. The topic also includes the consideration of space-time as a quantum error-correcting code: «Is spacetime a quantum error-correcting code?» (arXiv:1503.06237).

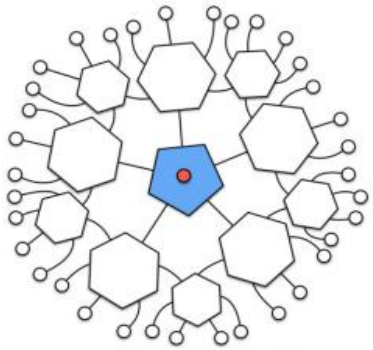
Holographic quantum codes



pentagon code



pentagon/hexagon code

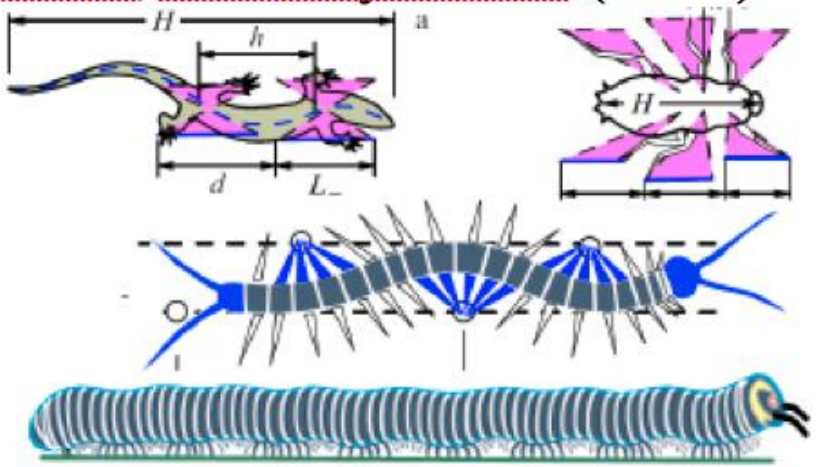


one encoded qubit

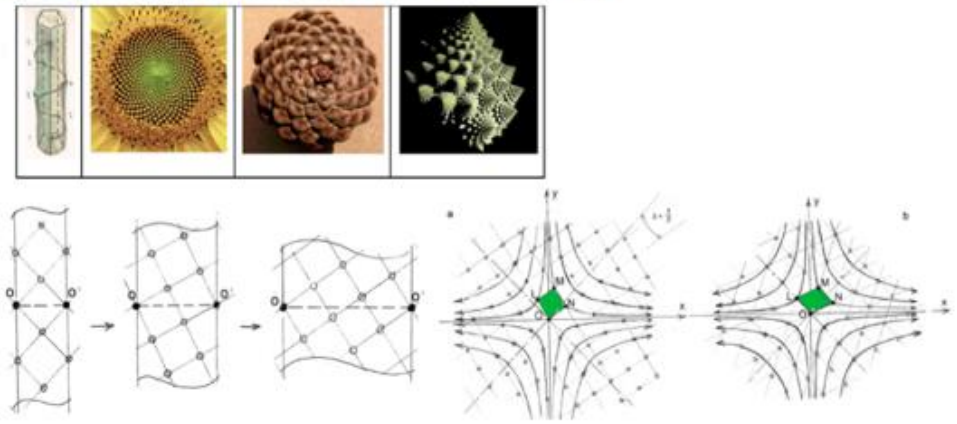
Revealing the connections of the genetic coding system with the Poincaré disk model of hyperbolic geometry testifies to the genetic basis of the known facts of the relationship between physiological phenomena and hyperbolic geometry, for example, described by:

- Luneburg (1950), Kienle (1964) about the space of visual perception;
- Smolyaninov (2000) on locomotion of animals and humans;
- Bodnar (1992) on rearrangements of phyllotaxis lattices during the growth of organisms.

From V.Smolyaninov (2000):



From O.Bodnar (1992):



One should note that arrangement of 20 amino acids and stop-codons in the matrix of 64 triplets and its bit-reversible analogue are very symmetrical (the case of the Vertebrate Mitochondria genetic code is shown):

	111	110	101	100	011	010	001	000
111	CCC Pro	CCA Pro	CAC His	CAA Gln	ACC Thr	ACA Thr	AAC Asn	AAA Lys
110	CCT Pro	CCG Pro	CAT His	CAG Gln	ACT Thr	ACG Thr	AAT Asn	AAG Lys
101	CTC Leu	CTA Leu	CGC Arg	CGA Arg	ATC Ile	ATA Met	AGC Ser	AGA Stop
100	CTT Leu	CTG Leu	CGT Arg	CGG Arg	ATT Ile	ATG Met	AGT Ser	AGG Stop
011	TCC Ser	TCA Ser	TAC Tyr	TAA Stop	GCC Ala	GCA Ala	GAC Asp	GAA Glu
010	TCT Ser	TCG Ser	TAT Tyr	TAG Stop	GCT Ala	GCG Ala	GAT Asp	GAG Glu
001	TTC Phe	TTA Leu	TGC Cys	TGA Trp	GTC Val	GTA Val	GGC Gly	GGA Gly
000	TTT Phe	TTG Leu	TGT Cys	TGG Trp	GTT Val	GTG Val	GGT Gly	GGG Gly

	111	011	101	001	110	010	100	000
111	CCC Pro	ACC Thr	CAC His	AAC Asn	CCA Pro	ACA Thr	CAA Gln	AAA Lys
011	TCC Ser	GCC Ala	TAC Tyr	GAC Asp	TCA Ser	GCA Ala	TAA Stop	GAA Glu
101	CTC Leu	ATC Ile	CGC Arg	AGC Ser	CTA Leu	ATA Met	CGA Arg	AGA Stop
001	TTC Phe	GTC Val	TGC Cys	GGC Gly	TTA Leu	GTA Val	TGA Trp	GGA Gly
110	CCT Pro	ACT Thr	CAT His	AAT Asn	CCG Pro	ACG Thr	CAG Gln	AAG Lys
010	TCT Ser	GCT Ala	TAT Tyr	GAT Asp	TCG Ser	GCG Ala	TAG Stop	GAG Glu
100	CTT Leu	ATT Ile	CGT Arg	AGT Ser	CTG Leu	ATG Met	CGG Arg	AGG Stop
000	TTT Phe	GTT Val	TGT Cys	GGT Gly	TTG Leu	GTG Val	TGG Trp	GGG Gly

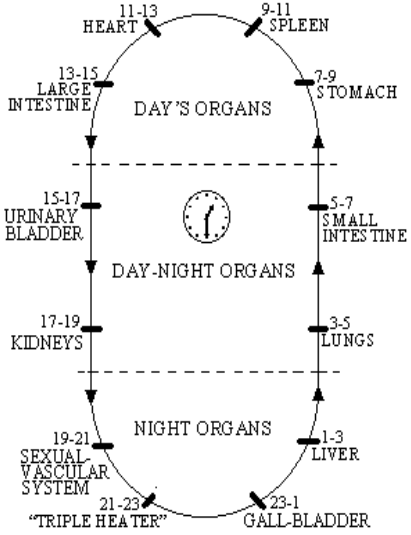
The matrix (at left) of the encoded amino acids and stop-codons consists of pairs of adjacent rows, identical in composition of amino acids and stop-codons (indicated by color). Its bit-reversible analogue (at right) has identical upper and lower halves in composition and arrangement of amino acids and stop-codons. **So, genetic informatics is closely related to the principles of algebraic holography.**

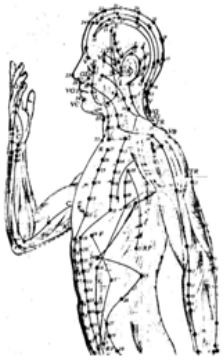
It is interesting that the matrix of 64 triplets in its binary oppositional structure is similar to the famous table of **64 Yin-Yang hexagrams of the ancient Chinese "Book of Changes" (I-Ching)**:

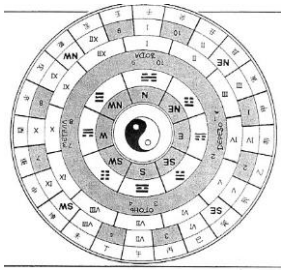
	111	110	101	100	011	010	001	000
111	CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
110	CCT	CCG	CAT	CAG	ACT	ACG	AAT	AAG
101	CTC	CTA	CGC	CGA	ATC	ATA	AGC	AGA
100	CTT	CTG	CGT	CGG	ATT	ATG	AGT	AGG
011	TCC	TCA	TAC	TAA	GCC	GCA	GAC	GAA
010	TCT	TCG	TAT	TAG	GCT	GCG	GAT	GAG
001	TTC	TTA	TGC	TGA	GTC	GTA	GGC	GGA
000	TTT	TTG	TGT	TGG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

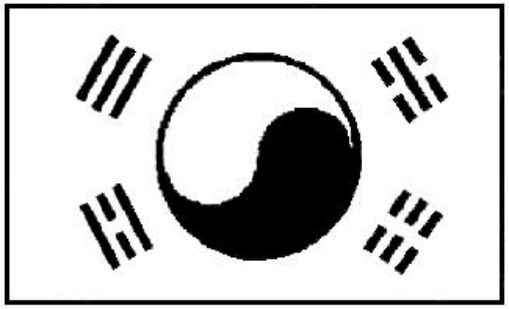
	111 CHYAN	110 TUI	101 LI	100 CHEN	011 HSUN	010 KAN	001 KEN	000 KUN
111 CHYAN	111111	111110	111101	111100	111011	111010	111001	111000
110 TUI	110111	110110	110101	110100	110011	110010	110001	110000
101 LI	101111	101110	101101	101100	101011	101010	101001	101000
100 CHEN	100111	100110	100101	100100	100011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

I-Ching was written several thousand years ago and has had a powerful influence on oriental medicine and culture. In connection with it, acupuncture and Tibetan pulse diagnostics were developed, indicating the holographic principles of the body. Carl Jung believed that its trigrams and hexagrams fix a universal set of archetypes (innate mental structures).









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

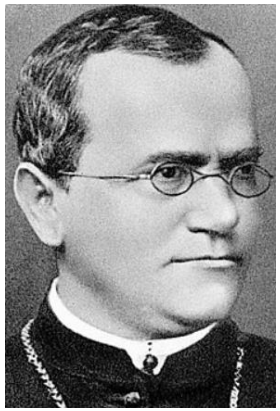
The ancient Chinese claimed that this table of 64 yin-yang hexagrams is a universal natural archetype. They did not know anything about the genetic code, but it with its 64 genetic triplets and binary properties turns out to be surprisingly similar to this table and schemes of the I-Ching.

Holographic principles have long been used for the physical understanding of the world. Thus, the famous physicist David Bohm believed that human consciousness is part of the universal hologram of the entire human race and that the entire Universe has a holographic structure. Bohm worked with Pribram on the theory that the brain works like a hologram in accordance with quantum mathematical principles and characteristics of wave patterns.

Nobel laureate in physics Gerard t'Hooft put forward the principle of holography in the structure of the world, which is actively developed in modern physics by many authors.

Dualism "probability-vs-determinism" in genetics

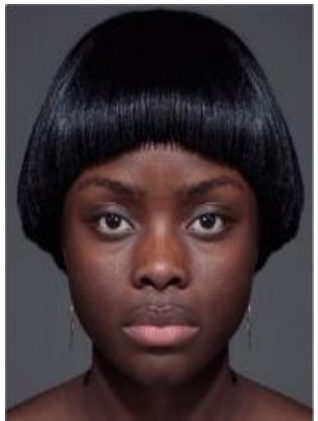
Genetics as a science began with Mendel's discovery of the stochastic rules of inheritance of traits in experiments on the crossing of organisms. Many processes in living bodies are of a stochastic nature. The well-known expressions "gene noise" or "cell noise" reflect the fact that even genetically identical cells within the same tissue exhibit different levels of protein expression, different sizes and structures due to the stochastic nature of interactions of individual molecules in cells.



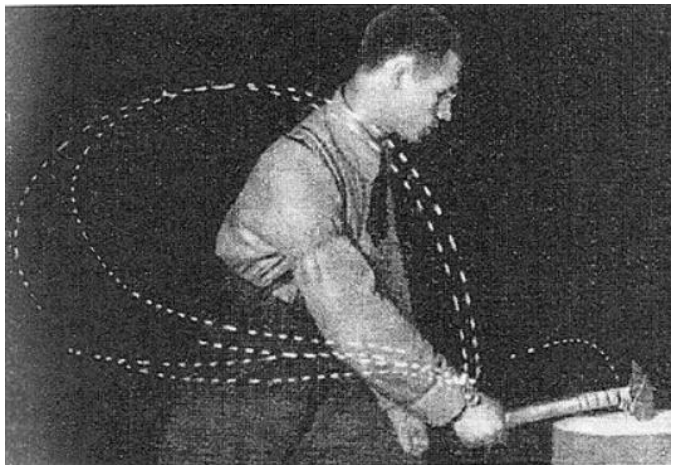
Gr. Mendel (1822-1884)

This stochastic nature of genetic processes in the "small" does not interfere with the inheritance of the traits determined in the "big". For example, fingerprints are different for all people, although fingers in general have a typical shape and composition.

According to Mendel's law of independent inheritance of traits, information from the level of DNA molecules dictates the macrostructure of living bodies through many independent channels, despite strong noises. Thus, hair, eye and skin colors are inherited independently of each other. Accordingly, **each organism is a machine of multichannel noise-immunity encoding**.



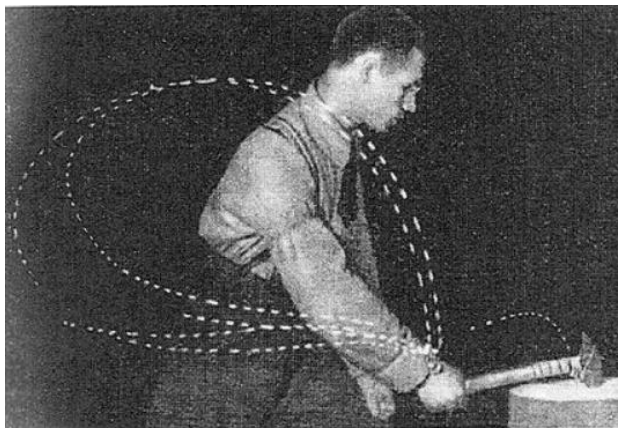
One of the examples of the connection between stochastic and deterministic features in biology is given by the phenomenon of biomechanics of movements by the classic of biomechanics N.A. Bernstein: the general target task of movement is performed exactly regardless of the inaccuracies of its constituent motor subtasks. For example, when repeating an exact hit with a hammer on a nail, each time a person uses different trajectories, speeds and accelerations of body parts with a change in both the flexion angles in the joints and the activity of many muscles of each joint with many motor neurons of each muscle.



This example shows that the intelligence of an organism works with the **probabilistic** characteristics of local movements of the body's links, giving out a **deterministic** general result of movement - hitting the target.

The speaker is convinced that the knowledge of the laws of the dualism "probability-vs-determinism" in biology is important for the creation of artificial intelligence systems of the next generation.

The described phenomenon of biomechanics of movements belongs to the broad topic of **Gestalt-biology**, in which stable holistic patterns are formed regardless of the variability of their constituent parts.



For example, Gestalt psychology studies the innate properties of the brain to form holistic images that are relatively independent of their particular components. Thus, a musical melody is recognized by us, even when it is played on different instruments and in different frequency ranges.

This is an inherited fundamental property of the psyche: to seek in a disparate whole.

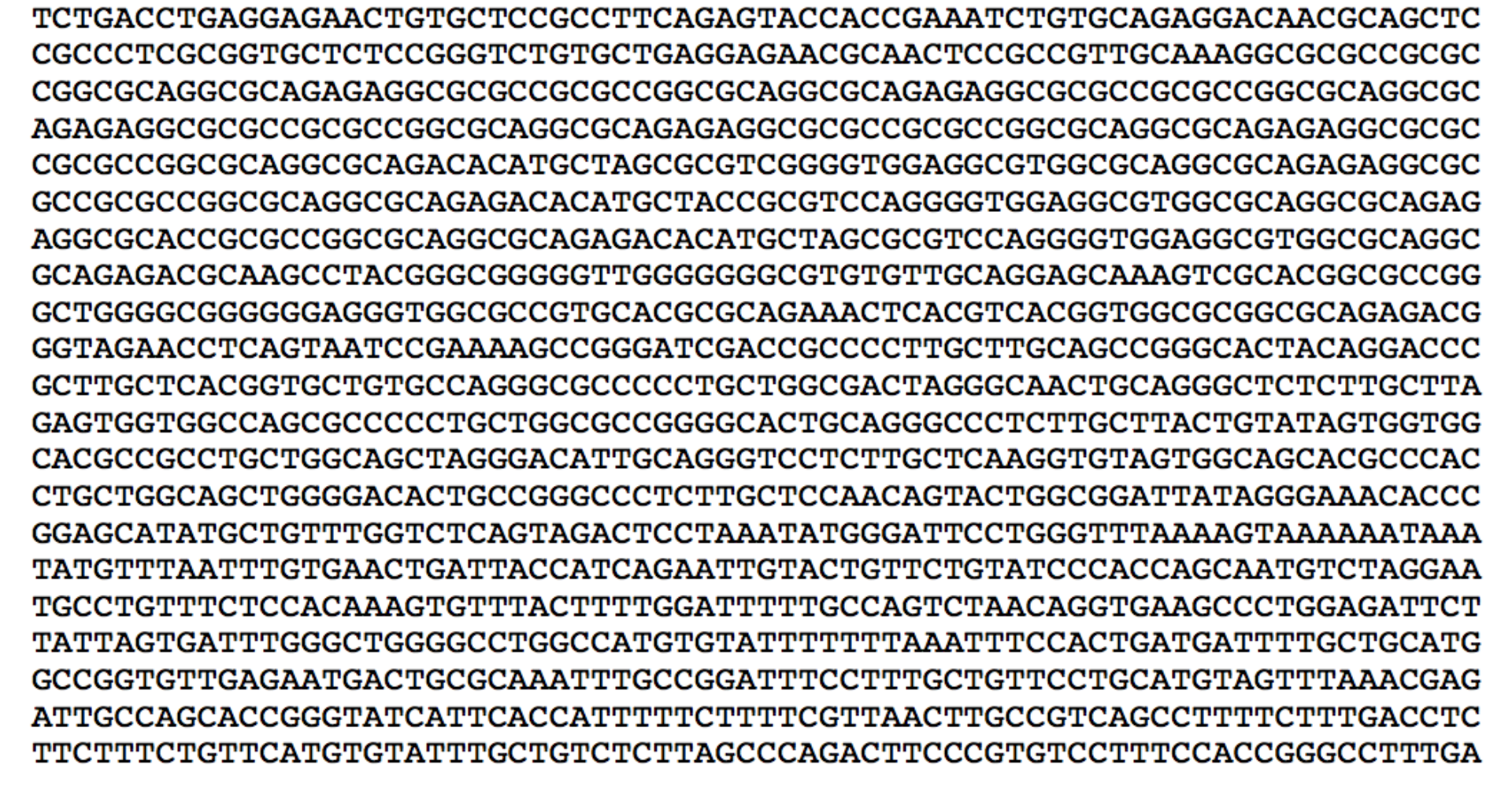
In biology, there are many other genetically inherited gestalt phenomena - morphogenetic, homeostatic, sensory, etc. - in which a stable holistic pattern is realized in conditions of a wide variability of its constituent components.

For example, the molecular composition of a living body is constantly changing while maintaining the shape of the body. Its proteins are involved in continuous cycles "life-death" of stochastic assembly and disassembly 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. According to the physiologist A.G. Gurvich, "***the main problem of biology is maintaining the shape with constant renewal of the substrate***".

All physiological systems are genetically inherited through their genetic coding. Therefore, one should look for the origins of these inherited gestalt phenomena of physiology in the genetic system of DNA informatics.

In the course of such searches, the speaker discovered that already in the information system of DNA molecules of the genomes of higher and lower organisms, universal gestalt phenomena of their stochastic organization are embedded. Let us explain this.

Long DNA sequences for each of many genomes are presented in the public GenBank in the following standard form occupying thousands of pages:



For clarification of the discovered universal rules of genomic DNAs, let's turn to the DNA of the first human chromosome, which contains a sequence of about 250 million nucleotides. Calculating in this DNA the percentages of each member of the DNA **alphabet of four nucleotides** (%C,%A,%T,%G), we obtain a table of their probabilities:

%C	%A	=	0.2085	0.2910
%T	%G		0.2918	0.2087

Then we represent the same DNA as a text of two-letter words (such as CA-TT-GA-) based on the **alphabet of 16 duplets** and, calculating the percentage of each of the types of duplets, we get a table of percentages of 16 types of duplets:

%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

Similarly, presenting the same DNA as a text of three-letter words, and then as a text of four-letter words, etc., we obtain the corresponding tables of percentages of kinds of n-plets in these texts, written respectively using the alphabets of 64 triplets, 256 tetraplets, etc. ... For example, for the DNA-text of triplets we have a percentage of 64 triplets:

%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

=

0.01385	0.01878	0.01524	0.01861	0.01183	0.01977	0.01447	0.03693
0.01853	0.00291	0.01789	0.02104	0.01622	0.00254	0.02375	0.01988
0.01758	0.01275	0.00251	0.00227	0.01317	0.01942	0.01441	0.02237
0.02009	0.02088	0.00259	0.00291	0.02388	0.01781	0.01614	0.01848
0.01588	0.01964	0.01103	0.01986	0.01255	0.01456	0.00962	0.01960
0.02226	0.00233	0.01939	0.01284	0.01437	0.00253	0.01327	0.01756
0.01972	0.01981	0.01457	0.01947	0.00956	0.01115	0.01256	0.01600
0.03725	0.01884	0.01988	0.01895	0.01445	0.01534	0.01185	0.01382

Thus, according to this method, the nucleotide sequences in each single-stranded genomic DNA turn out to be a whole bunch of parallel texts, each of which is written in its own alphabet of n-plets, and the DNA as a whole appears as a multilingual object. This method proved to be effective, leading to the identification of universal rules and symmetries of the stochastic organization of genomes.

Taking this into account, figuratively speaking, any genomic DNA speaks many languages, is a polyglot!

At first glance, the resulting sets of percentages of n-plets in these n-plets DNA texts are quite chaotic. Moreover, the percentage of types of n-plets in these texts depends on the order of the letters. For example, in the duplet text of the DNA under consideration, the percentage of CG and GC duplets of the same letter composition differs several times: %CG = 0.0103, a %GC = 0.0440.

%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

But in this seeming chaos, there are many universal rules for n-plet groupings that are valid for all studied genomes.

I will give some of the examples of universal equalities for the sums of percentages of n-plets, using percentages in the DNA of the first human chromosome. Let us turn to the tensor family of matrices of DNA alphabets $[C, A; T, G]^{(n)}$ with binary numbered rows and columns:

	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

It turns out that in these matrices any pair of rows and any pair of columns numbered with bit-reversed numbers (mutual reversal 0↔1) has, with high precision, the same sums of percentages of their n-plets within each pair, although separate percentages in these rows and columns differ significantly.

	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

For example, in the case of the duplet text of the DNA under consideration, we have practically equal sums of percentages of types of duplets in pairs of columns and rows with reversed numbers, despite the fact that the terms in these sums are different:

Столбцы	Суммы процентов в столбцах 16 дуплетов
11	0.05409+0.07134+0.06008+0.09568 = **0.28119**
00	0.09504+0.07137+0.06008+0.05419 = **0.28068**
10	0.07274+0.01031+0.06312+0.07286 = **0.21903**
01	0.05033+0.07429+0.04402+0.05046 = **0.21910**

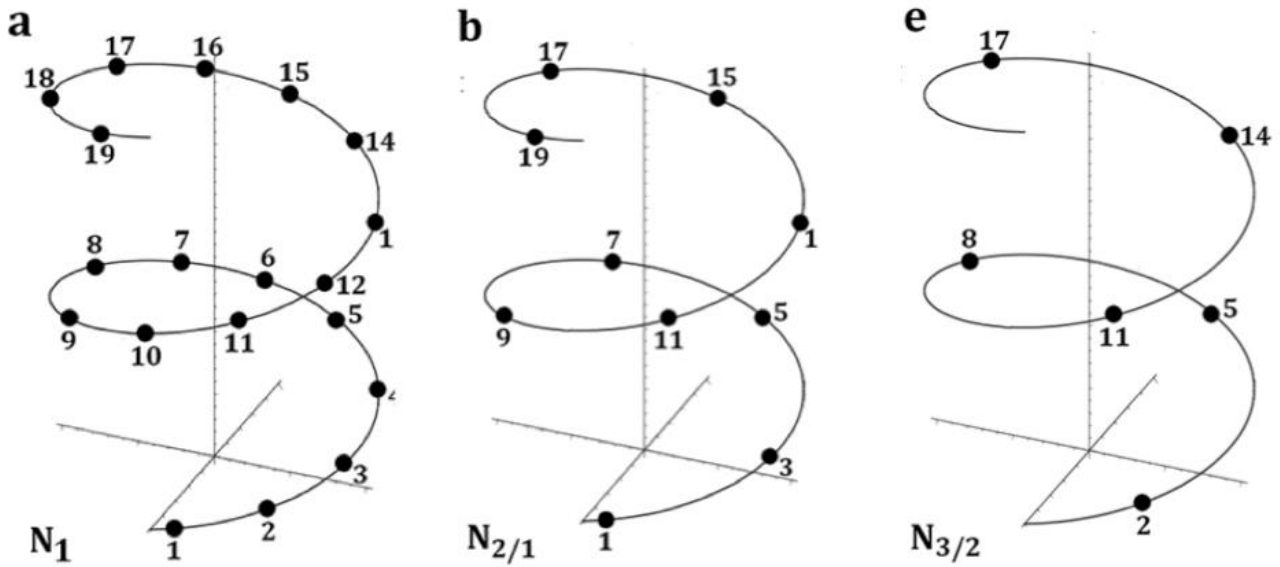
Строки	Суммы процентов в строках 16 дуплетов
11	0.05409+0.07274+0.05033+0.09504 = **0.27219**
00	0.09568+0.07286+0.05046+0.05419 = **0.27319**
10	0.07134+0.01031+0.07429+0.07137 = **0.22730**
01	0.06008+0.06312+0.04402+0.06008 = **0.22731**

In the case of a triplet text of the same DNA, we have equalities of the sums of the percentages of the types of triplets in pairs of columns and rows with binary-reversed enumerations, though the values of the terms in these sums are different:

Столбцы	Суммы процентов в столбцах 64 триплетов
111	0.01385+0.01853+0.01758+0.02009+0.01588+0.02226+0.01972+0.03725 = **0.16516**
000	0.03693+0.01988+0.02237+0.01848+0.01960+0.01756+0.01600+0.01382 = **0.16464**
110	0.01878+0.00291+0.01275+0.02088+0.01964+0.00233+0.01981+0.01884 = **0.11594**
001	0.01447+0.02375+0.01441+0.01614+0.00962+0.01327+0.01256+0.01185 = **0.11607**
101	0.01524+0.01789+0.00251+0.00259+0.01103+0.01939+0.01457+0.01988 = **0.10310**
010	0.01977+0.00254+0.01942+0.01781+0.01456+0.00253+0.01115+ 0.01534 = **0.10312**
100	0.01861+0.02104+0.00227+0.00291+0.01986+0.01284+0.01947+0.01895 = **0.11595**
011	0.01183+0.01622+0.01317+0.02388+0.01255+0.01437+0.00956+0.01445 = **0.11603**

Строки	Суммы процентов в строках 64 триплетов
111	0.01385+0.01878+0.01524+0.01861+0.01183+0.01977+0.01447+0.03693=**0.14946**
000	0.03725+0.01884+0.01988+0.01895+0.01445+0.01534+0.01185+0.01382=**0.15039**
110	0.01853+0.00291+0.01789+0.02104+0.01622+0.00254+0.02375+0.01988=**0.12275**
001	0.01972+0.01981+0.01457+0.01947+0.00956+0.01115+0.01256+0.01600=**0.12285**
101	0.01758+0.01275+0.00251+0.00227+0.01317+0.01942+0.01441+0.02237=**0.10448**
010	0.02226+0.00233+0.01939+0.01284+0.01437+0.00253+0.01327+0.01756=**0.10456**
100	0.02009+0.02088+0.00259+0.00291+0.02388+0.01781+0.01614+0.01848=**0.12277**
011	0.01588+0.01964+0.01103+0.01986+0.01255+0.01456+0.00962+0.01960=**0.12273**

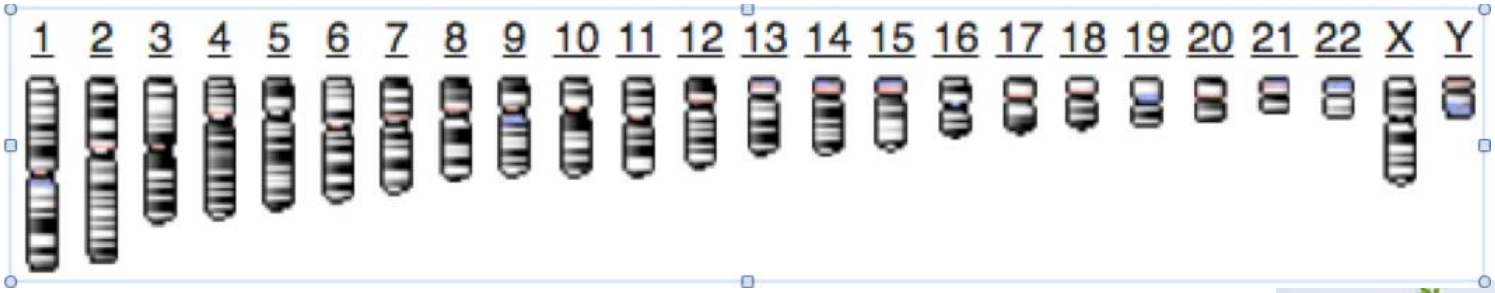
Interesting results on the universal percentage regularities of the fractal-like structure of genomic DNA are obtained from the analysis of such sparse sequences of these DNAs, which consist only of nucleotides spaced by a fixed number of m = 2, 3, 4, ... from each other in this DNA. These sparse truncated sequences are called DNA epichains, and data on their percentage characteristics are presented in the article: Petoukhov S.V., Nucleotide epi-chains and new nucleotide probability rules in long DNA sequences (2019, https://www.preprints.org/manuscript/201904.0011/v2).

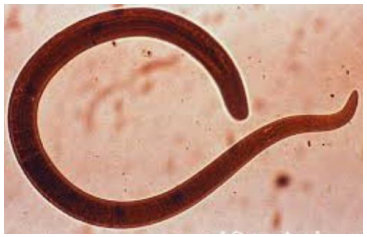


In general, in the genetic foundations of living matter, we are faced with a very special and universal type of "**regularly limited stochasticity**": the randomness of alphabetic elements in the n-texts of genomic DNA of higher and lower organisms turns out to be limited by sets of deterministic numerical relationships between groupings of n-plets (this multi-layer vaguely reminiscent of polyphony in music).

The described method analyzed the percentage of n-texts of the following genomic DNA: 1) all 24 human chromosomes; 2) all the chromosomes of Drosophila, mouse, worm, many plants; 3) 19 genomes of bacteria and archaea; 4) many extremophiles living in extreme conditions, including, for example, radiation with a level exceeding 1000 times fatal to humans.

Due to this, the rules and symmetries of the stochastic organization of genomes, which are common to all these genomic DNAs, and therefore are candidates for the role of universal genomic rules, have been identified.





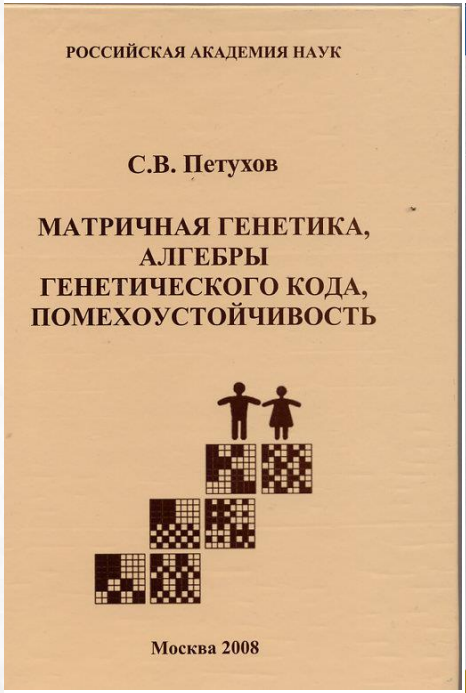
The universal Gestalt rules and symmetries identified by the author in the stochastic organization of DNA of genomes of higher and lower organisms indicate the existence of nontrivial algebraic invariants of a **globally** genomic nature, which remain unchanged over millions of years of biological evolution, during which millions of species of organisms die off and new ones appear (although **locally** genomic sequences are modified by mutations, natural selection mechanisms, etc.). This seems to be connected with the question of the origin of life, since these patterns are presented also in archaea and bacteria.

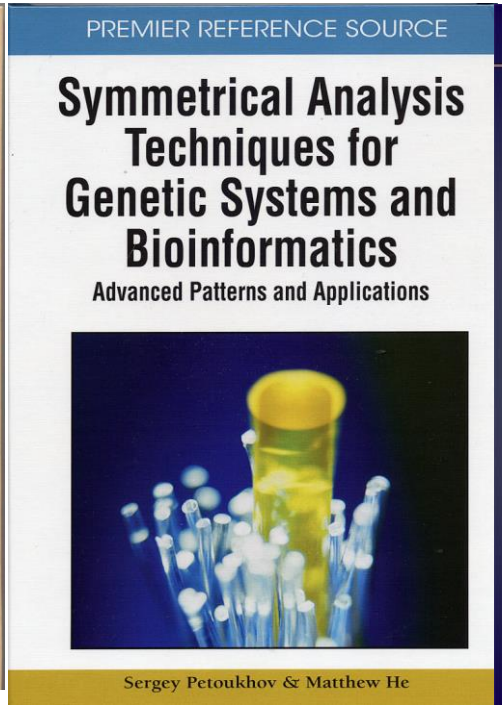
The discovered universal Gestalt rules of the block-stochastic organization of genomic DNA indicate the existence of quantum-mechanical long-range links in the nucleotide DNA sequences.

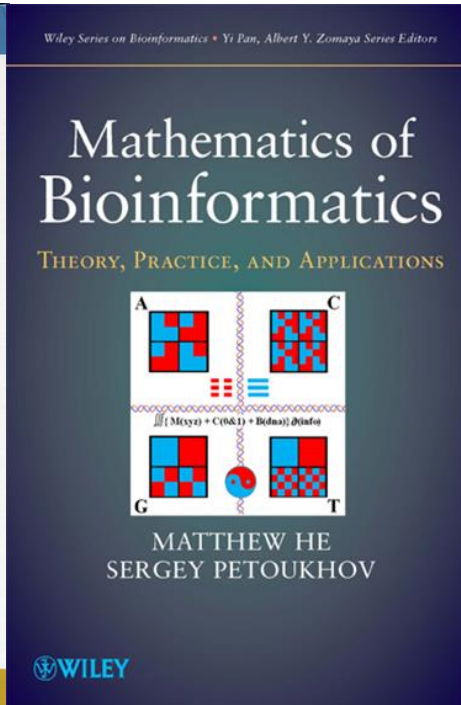
They are used to create new approaches to artificial intelligence and a new view of organisms as algebraic entities associated with the formalisms of quantum informatics.

Let us turn to the question of the relationship between innate knowledge and knowledge acquired in the course of life. The ancient Greek philosopher Plato stated that knowledge is a matter of recollection of the state before one is born (briefly speaking, to know means to recollect). Close to this is the opinion that our body with its nervous system already carries in itself the fullness of knowledge, which partially come into our consciousness under their persistent request. But our body grows out of a single sex cell that carries genomic DNA. It follows that the entire amount of our knowledge for life and science is either contained in this genome (which is doubtful), or, as the author believes, genomic DNA is a complex fractal antenna for receiving some space knowledge and programs. These questions are being studied in the laboratory of the author.

This 30-minute talk presents only fragments of our works in holographic genetics and algebraic biology. Many other materials are available on the speaker's website: http://petoukhov.com/ .







Some conclusions

1) Molecular genetic structures are associated with the formalisms of algebraic holography, quantum informatics and Poincaré's disk model of hyperbolic geometry.

2) Genomic DNA are packets of parallel texts, each of which is written in one of the n-plets DNA alphabets.

3) Revealing universal numerical rules in the stochastic organization of genomes of higher and lower organisms is important for understanding the dualism "probability-vs-determinism", the development of algebraic biology and the development of new approaches to artificial intelligence systems.