Internet Electronic Journal of Molecular Design **2004**, *3*, 247–270 ISSN 1538–6414
BioChem Press http://www.biochempress.com

Inter*net* Electronic Journal of Molecular Design

May 2004, Volume 3, Number 5, Pages 247–270

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Nenad Trinajstić on the occasion of the 65th birthday
Part 11

Guest Editor: Douglas J. Klein

Symmetry Groups for the Rumer–Konopel'chenko–Shcherbak "Bisections" of the Genetic Code and Applications

Tidjani Négadi[1]

[1] Physics Department, Faculty of Science, Oran University, Es–Sénia, 31100, Oran, Algeria

Received: September 30, 2003; Revised: January 30, 2004; Accepted: March 1, 2004; Published: May 31, 2004

Citation of the article:
T. Négadi, Symmetry Groups for the Rumer–Konopel'chenko–Shcherbak "Bisections" of the Genetic Code and Applications, *Internet Electron. J. Mol. Des.* **2004**, *3*, 247–270, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

Symmetry Groups for the Rumer–Konopel'chenko–Shcherbak "Bisections" of the Genetic Code and Applications[#]

Tidjani Négadi[1,*]

[1] Physics Department, Faculty of Science, Oran University, Es–Sénia, 31100, Oran, Algeria

Received: September 30, 2003; Revised: January 30, 2004; Accepted: March 1, 2004; Published: May 31, 2004

Internet Electron. J. Mol. Des. **2004**, *3* (5), 247–270

Abstract

We derive, in a new way, the discrete symmetry groups for (*i*) the 4–base set {U, C, A, G}, (*ii*) the 16–doublet set and (*iii*) the 64–codon set, as collections of adjacency matrices of selected graphs on the Wittmann sub–sets of the above respective sets. In the case of the genetic code 64 codons system, we re–derive the chain of groups $D_8 \supset V \supset C_2$ and show that the last member of the chain, C_2, leaves 16 codons of type GNN invariant and this invariance is maintained across all species with respect to their "non–standard" use of the genetic code, including nuclear genomes as well as mitochondrial genomes. Moreover, we show that this symmetry is suited, in fact it fits, the "bisections" of the set of 64 codons, used by Shcherbak to derive many striking arithmetical regularities and balances, involving the nucleon numbers in the amino acids. Besides the symmetry aspects, our next new result concerns the derivation, using only the concept of matrix–norms in traditional linear algebra, of some (striking) numbers which appear to be characteristic of the genetic code. Finally, by using only the RNA–components, *i.e*., the four nitrogenous bases mentioned above, we introduce matrices encoding the hydrogen–bond attribute and other matrices encoding a certain "molecular size index" for the bases and derive the ratio of their trace, and of their norms, which appear to be equal in both cases to Shcherbak's "*Prime Quantum*" 037.

Keywords. Genetic code; symmetry; Rumer transformation; invariant codons; Wittmann multiplets; matrix–norms; autopoietic numbers.

1 INTRODUCTION

When the writing of the results, corresponding to part of this work, was made the first cut in, we were focusing on symmetry aspects of the genetic code (Rumer symmetry), and were far from imagining that the final aspect of this invited paper, dedicated to Professor Nenad Trinajstić on the occasion of his 65th birthday, was going to become richer by the addition of a new section. The reason is the recent publishing in *Biosystems* of an important paper by Shcherbak, summing up some fifteen years of published research leading to the mentioned paper entitled "*Arithmetic inside the universal genetic code*" [1], in which the author describes striking arithmetical regularities in the genetic code, revealed when certain partitions of the 64–codon set are made. We shall, in section 4,

[#] Dedicated to Professor Nenad Trinajstić on the occasion of the 65th birthday.

* Correspondence author; E–mail: tidjani_negadi@yahoo.com.

return to the work of Shcherbak but let us say only that Rumer's symmetry plays an important role in revealing these regularities. Moreover, we take the opportunity given by the added section to (*i*) present the derivation of some (no less striking) numbers we have found, appearing to be characteristic of the genetic code, which results *would* have been presented elsewhere[1] and (*ii*) to propose a simple derivation of Shcherbak's "*Prime Quantum*" 037, a distinguished (decimal) number which plays an important role in the work of Shcherbak where it is associated to a *Prime Quantum Divisibility* feature.

In this paper, which is a continuation of recent works [2,3,4], we shall present, in a new way, the symmetries associated to the four fundamental nitrogenous bases, uracil, cytosine, adenine and guanine, the building blocks of RNA, and to the *k*–plet sets constructed out of them: the set of 16 doublets ($k = 2$) and the set of 64 triplets or codons ($k = 3$). (A *k*–plet set is also called, in the computing science community, the k^{th} extension of the quaternary alphabet {U/T, C, A, G} (Yockey [20]). Our construction, which was introduced in [2], starts with the constituents of the bases themselves: the atoms, written as 2×2 matrices and, as will be shown in section 2, these sets (for $k = 1, 2, 3$) are obtained as matrices of dimension $2^k \times 2^k$. They constitute our *classification* matrices. There exist also an "experimental" manner to construct all the members of these *k*–plet sets which goes back to Wittmann who, in the heroic days of the deciphering of the genetic code in the early sixties, used mutagens like nitrous acid to investigate the genetic code. In particular, in [5], he studied the genetic code in Tobacco Mosaic Virus and deduced the *Octet Model* in which the 64 possible codons are partitioned into eight Octets of eight triplet–codons each (see section 3).

As this is important for what follows, let us say some words on this construction. Wittmann, considering the phenomenon of deamination induced by nitrous acid, formulated the conversions of the bases as follows: U, unchanged, C→U, G→(X)→G and A→(H)→G. X (xantine) and H (hypoxantine), are two minor bases both with very similar hydrogen bonding properties to guanine. Using these rules, one could start with sub–sets containing *only* A and/or C, *i.e.*, bases whose deamination can produce a mutation and, for each member of these sub–sets, construct the other co–members by following the rules.

As a simple example, in the case of the four–base set ($k = 1$), the sub–set is constituted by C and A. C will give U and A will give G so that the two Wittmann Doublets are C→U and A→G. For $k = 2$, the sub–set has $2^2 = 4$ base–doublets containing C and/or A: CC, CA, AC and AA. Take CA, for example. It produces UA and CG. UA could give only UG and CG could give only UG and, the process stops. The four Wittmann Quartets Q_i ($i = 1, 2, 3, 4$) are shown below (Figure 1).

[1] In the Congress "Festival Symmetry" 2003 (Budapest, August, 16–22, 2003) or in the "Third International Symposium on Quantum Theory and Symmetries", Cincinnati, Sept. 10 – 14, 2003). These two invitations, and another one didn't go, unfortunately, to an end because of the blindness of local rulers.

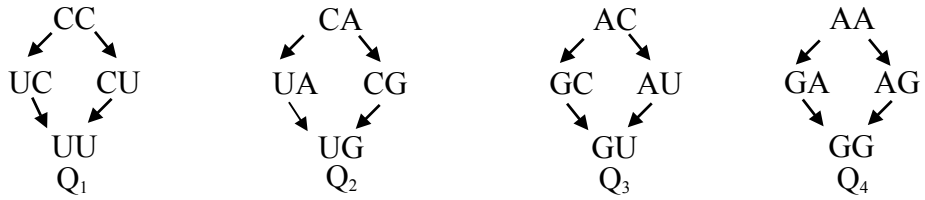


Figure 1. The four Wittmann Quartets Q_i.

Finally, in the case $k = 3$, the starting sub–set contains eight (2^3) combinations of A and C: AAA, AAC, ACA, CAA, ACC, CAC, CCA and CCC. As above, each one of these triplets will be the first top member of a descending cascade governed by the Wittmann rules. One such Octet (Ω_2, top member AAC), as an example, could be visualized as presented in Figure 2.

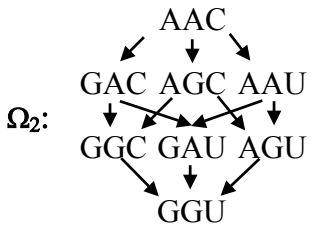


Figure 2. One of the eight Wittmann Octets (Ω_2), see text.

In this way, a total of eight Wittmann Octets (Ω_i, $i = 1, \ldots, 8$) are constructed comprising all 64 codons; we have shown only one of them, Ω_2. Now, our own way to construct these objects, which relies on a matrix representation of the base, doublet and codon sets, is linked to the Wittmann Doublets, Quartets and Octets in a very simple manner: the latter are the *rows* of our classification matrices, respectively for $k = 1, 2, 3$. This very fact will let us realize two objectives in this work: (*i*) a guided search of the symmetries, through the use of permutation graphs between the Wittmann multiplets, and the use of the adjacency matrices of these graphs as transformation operators on our classification matrices (sections 2 and 3) and (*ii*) a mathematically correct establishment of certain striking numbers (mentioned briefly above), all relevant for (and characteristic of) the genetic code, which appear to be attached to certain matrices coding either the *hydrogen bond attribute* or a *molecular–size–index* relying on the total number of atoms in the bases, either as "norms", sizing the matrices and, at the same time, sizing the Wittmann Doublets and Octets, or as invariants (traces). This will be done in section 4. Shcherbak [1], besides using the Rumer symmetry, has chosen another attribute, the number of nucleons of the amino acids and, in fact in his work, everything seems to be associated with a special number, he calls the "*Quantum Prime*" 037, which divides almost everything. We propose, briefly in the fourth section, a possible *raison d'être* for this number, which is not transparent in the above work. In particular, we show that it is independent of the amino acids.

In the second section, we sum up the construction of our classification matrices. In the third, we

construct the symmetry groups which will act on the classification matrices. These matrix groups are obtained as collections of adjacency matrices for selected graphs on the Wittmann sets. In the fourth section, we use some tools supplied by linear algebra (matrix–norms and traces) to establish some interesting numbers.

2 BASES, DINUCLEOTIDES AND CODONS

2.1 The base–matrix

In [2], we have designed a 2×2 matrix representation of atoms and molecules, among them, the four nitrogenous bases U (uracil: $C_4N_2H_4O_2$), C (cytosine: $C_4N_3H_5O$), A (adenine: $C_5N_5H_5$) and G (guanine: $C_5N_5H_5O$), the building blocks of RNA. Their particular (matrix) form, which is the consequence of their detailed atomic composition, is such that they could be easily united in the following base–matrix [2]

$$\mathbf{B} = \begin{pmatrix} U & C \\ A & G \end{pmatrix}, \tag{1}$$

where the numerical values are U = $6^4 7^2 8^2$ = 4064256, C = $6^4 7^3 8$ = 3556224, A = $6^5 7^5$ = 130691232, G = $6^5 7^5 8$ = 1045529856 and are the products of the atomic numbers of the constituent atoms: carbon (Z = 6), nitrogen (Z = 7), oxygen (Z = 8) and hydrogen (Z = 1). Eq. (1) is the basis of the next constructions.

2.2 The doublet–matrix

The matrix of the 16 possible base–doublets is constructed using the (recently introduced) Kronecker product with *concatenation*[2] [3], which incorporates the non–commutativity of the bases, unlike the ordinary Kronecker product. From Eq. (1), we have:

$$\mathbf{D} = \mathbf{B} \,||\, \mathbf{B} = \begin{pmatrix} UU & UC & CU & CC \\ UA & UG & CA & CG \\ AU & AC & GU & GC \\ AA & AG & GA & GG \end{pmatrix}, \tag{2}$$

The numerical value of each one of the 16 base–doublets, the matrix elements in the form XY, is obtained by concatenating the corresponding values X and Y from the values given above in Eq. (1). For example UG = 40642561045529856 is different from GU = 10455298564064256. In this

[2] For two matrices M and N, with integer entries, this product is given by $(M||N)_{js,kt}=M_{jk}||N_{st}$, and the computation is the same as for the ordinary Kronecker product except that the product of matrix elements is replaced by the *concatenation* of these (see ref. [3]).The concatenation of any two integers a and b, a||b, is just their juxtaposition ab.

way the 16 objects are all numerically distinct. Recall that the four Wittmann Doublets, Q_i, are respectively the four rows in Eq. (2). Also and importantly (see section 4), the doublets belonging to the same column have the same total number of hydrogen bonds (recall that U and A have both two hydrogen bonds while C and G have three each.

2.3 The Codon–Matrix

To obtain the matrix representing the 64 (triplet) codons of the genetic code, it is sufficient to repeat the operation on the base–doublet matrix in Eq. (2). One has:

$$\mathbf{C} = \mathbf{D} \parallel \mathbf{B} = \begin{pmatrix} UUU & UUC & UCU & UCC & CUU & CUC & CCU & CCC \\ UUA & UUG & UCA & UCG & CUA & CUG & CCA & CAG \\ UAU & UAC & UGU & UGC & CAU & CAC & CGU & CGC \\ UAA & UAG & UGA & UGG & CAA & CAG & CGA & CGG \\ AUU & AUC & ACU & ACC & GUU & GUC & GCU & GCC \\ AUA & AUG & ACA & ACG & GUA & GUG & GCA & GCG \\ AAU & AAC & AGU & AGC & GAU & GAC & GGU & GGC \\ AAA & AAG & AGA & AGG & GAA & GAG & GGA & GGG \end{pmatrix}. \quad (3)$$

Here also, each codon, as a matrix element, has an associated value obtained by concatenating three numbers. For example, UUC = 406425640642563556224, UCU = 406425635562244064256, CUU = 355622440642564064256, so that the non–commutativity of the bases, in the codons, is also implemented. Here also, the rows in Eq. (3) are the eight Wittmann Octets and the triplets belonging to the same column have a constant total number of hydrogen bonds (see Eq. (21)). As said above, these two properties will gain all their importance in section 4.

3 THE RUMER–KONOPEL'CHENKO SYMMETRY

Now, we turn to symmetry considerations. In 1966 [7], Rumer introduced an interesting partition of the 64–codon set into two equal sub–sets with 32 codons each, by "breaking" the degeneracy number six (see below). The two sets, we call them M_1 and M_2, are exchanged under the Rumer transformation UCAG<–>GACU. We shall, in the following three sub–sections, build a group theoretical framework for the Rumer symmetry and its extensions, following Konopel'chenko and Rumer [8] (Cf. [1]).

3.1 C_2 Rumer–symmetry for B

This is the simplest case. In order to find a symmetry (matrix) transformation for **B**, we consider the two Wittmann Doublets, WD_1 and WD_2 which, as we have said, are the rows in the matrix in Eq. (1). Alternatively, we can consider the columns of this matrix which contain complementary bases, with fixed number of hydrogen bonds, respectively two and three, and forming the sets {U, A} and {C, G}. Now with two objects, here the Wittmann Doublets, WD_1 and WD_2, we can

consider the two graphs shown below (self–loops allowed), representing links, or transformations, between these two objects (see figure 3, below)

Figure 3. The two possible graphs representing transformations between the two Wittmann Doublets WD_1 and WD_2.

In the first, we have the identity, while in the second, WD_1 and WD_2 are exchanged. The adjacency matrices for these two graphs are given respectively by:

$$\mathbf{e} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\eta} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4}$$

Recall the adjacency matrix for a graph (V, E) with $n = |V|$ vertices is an $n \times n$ matrix **M** such that $\mathbf{M}_{ij}$ is 1 if and only if there is an edge from vertex i to vertex j; otherwise $\mathbf{M}_{ij}$ is zero. The set {**e**, **η**} close under matrix multiplication ($\boldsymbol{\eta}^2 = \mathbf{e}$) and constitutes the simplest group with two elements, the cyclic group C_2. From the point of view of graph theory, this set is also what is called a *cellular algebra* or Bose–Mesner algebra (in fact the simplest) verifying $\mathbf{e}+\boldsymbol{\eta}=\mathbf{J}$, with **J** the all–1 matrix. Now, as we have explained in the introduction, these two matrices, obtained as adjacency matrices of graphs on the Wittmann multiplets (or, alternatively, graphs on the columns), will be used as symmetry transformations for the matrix **B**. For **e**, one has the identity and, for **η** ($\boldsymbol{\eta}^{-1} = \boldsymbol{\eta}$), one gets

$$\boldsymbol{\eta}^{-1}\mathbf{B}\,\boldsymbol{\eta} = \begin{pmatrix} G & A \\ C & U \end{pmatrix} \tag{5}$$

We see therefore that the transformation **η**, which acts as a *similarity transformation*, exchanges pyrimidines and purines: U<–>G and C<–>A. It exchanges at the same time the rows (the two Wittmann Doublets WD_1 and WD_2) and the columns. This is the Rumer transformation, mentioned above and noted UCAG <–> GACU. There exist also a *combined* transformation UCAG<–>CUGA *and* UCAG<–>AGUC, equivalent to the Rumer transformation UCAG <–> GACU, [9, 10]. The latter two transformations do not really exist individually but, jointly, they reconstitute the Rumer transformation. (They are sometimes called secondary Rumer transformations.) This could be seen with the aid of the second graph, above. This graph means, as we have said, that the two rows *and* the two columns are exchanged (or permuted) which is just the result of applying one time the matrix **η**, see Eq. (5). Below, we shall meet this kind of transformations and shall call them simply *transitions*, for UCAG<–>CUGA, and *transversions*, for UCAG<–>AGUC. Note that the Rumer transformation is a transversion.

3.2 Rumer and Rumer–Konopel'chenko Symmetry for D

Since the seminal work by Rumer, many people have studied (and still do today) the structure of the genetic code in terms of the 16–doublet set (see, only but a few examples [11,12,13]). This corresponds to what Shcherbak calls the "compressed" representation of the genetic code at scale 1 with known symmetries. In this sub–section, we shall describe these symmetries in terms of a small discrete group, known as the Klein group V. Danckwert and Neubert [11], already in 1975, used this group but in a formal way, *i.e.*, without a concrete representation. Here, we shall define *concretely* two Klein groups with different actions on the set of doublets.

We take, as a starting set, the four Wittmann Quartets which are the four rows in matrix (2); four is the required number of objects to get 4×4 adjacency matrices able to act on our 4×4 matrix **D**. Now, we have found two different graph–sets of four graphs each. The first one is given by the following four graphs (Figure 4), symbolizing the transformations (the nodes represent the Wittmann Quartets numbered from 1 to 4 from left to right and from top to bottom)

Figure 4. The first graph–set for the four Wittmann's Quartets (see text).

The adjacency matrices for these graphs are given respectively by:

$$\mathbf{V}_1: \quad \mathbf{T}_1=\begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}, \quad \mathbf{T}_2=\begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&0&1\\ 0&0&1&0 \end{pmatrix}, \mathbf{T}_3=\begin{pmatrix} 0&0&1&0\\ 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0 \end{pmatrix}, \quad \mathbf{T}_4=\begin{pmatrix} 0&0&0&1\\ 0&0&1&0\\ 0&1&0&0\\ 1&0&0&0 \end{pmatrix}. \tag{6}$$

Note that theses matrices are also permutation matrices. It has been shown in [3] that the set {$\mathbf{T_1}$, $\mathbf{T_2}$, $\mathbf{T_3}$, $\mathbf{T_4}$} close under matrix multiplication and constitute a commutative group, the four group V, also known as Klein's 4–group[3]; call it V_1. The second set of graphs, comprising the same identity transformation (the first graph above) which we do not duplicate below, is the following (Figure 5):

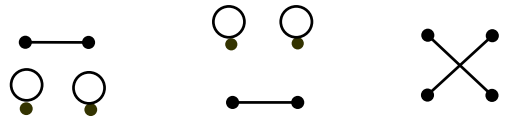


Figure 5. The second graph–set for the four Wittmann's Quartets.

Here, the adjacency matrices are the following:

[3] The Klein group V could be defined, for example, via the following permutations on four objects, {1, (12)(34), (13)(24), (14)(23)} or simply by its Cayley table (elements e, a, b, c): $a^2= b^2= c^2=e$, ab=ba=c, ac=ca=b, bc=cb=a.

$$\mathbf{V}_2 : \mathbf{S}_1=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\quad \mathbf{S}_2=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\ \mathbf{S}_3=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\ \mathbf{S}_4=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}. \tag{7}$$

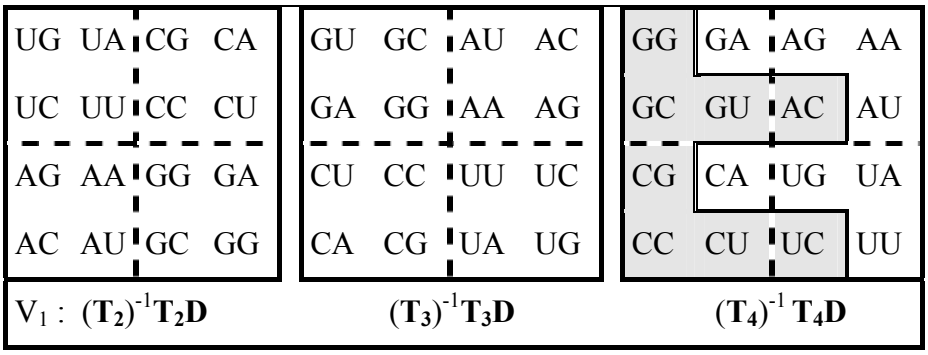
In this case, the set {$\mathbf{S_1}$, $\mathbf{S_2}$, $\mathbf{S_3}$, $\mathbf{S_4}$} also close under matrix multiplication and constitute a second Klein's 4–group, V_2, as it could be easily verified from (7). Note that the first set, {$\mathbf{T_1}$, $\mathbf{T_2}$, $\mathbf{T_3}$, $\mathbf{T_4}$} is a cellular algebra while it is not the case for the set {$\mathbf{S_1}$, $\mathbf{S_2}$, $\mathbf{S_3}$, $\mathbf{S_4}$}. We reproduce for ease of comparison Eq. (2) as the following representation:

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

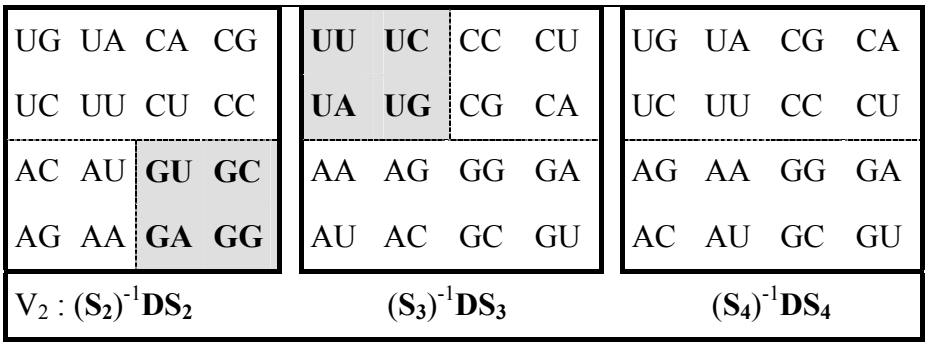
D, Eq. (2)

We also add, before continuing, some comments on **D** which prepare us for the next sub–section. First, the set of 16 elements is partitioned in two sub–sets M_1 (UC, CU, CC, CG, AC, GU, GC, GG) and M_2 (UU, UA, UG, CA, AU, AA, AG, GA) the elements of which are exchanged by the Rumer transformation UCAG <–> GACU and separated by the double lines. When going to the codon level (see below), all the codons from M_1 will correspond to amino acids for which the third base is irrelevant (the quartets and the quartet part of the sextets), while those of M_2, except for the singlets, will necessitate a third base (doublets, triplet). The idea to "break" the 64 codons set into two equal parts M_1 and M_2, by invoking the partition of the greatest degeneracy 6 as 4+2, goes back to Yuriĭ Borisovich Rumer, in 1966 [7]. (As a short historical note, Ginzburg, Mikhaĭlov (Rumer) and Pokrovskiĭ have written an entire paper, in 2001, on his life as a physicist and an untiring teacher and also on his contribution to biology, in *Physics Uspekhi* **2001**, *44*, 1075–1081. They say, for example, "He wrote a paper on the classification of codons in the genetic code using the symmetry principles and some linguistic arguments".) Second, viewing otherwise, **D** could be partitioned into four sub–sets, of four doublets each, harboring the same *first* base.We shall call these sub–sets $U^{(1)}$, $C^{(1)}$, $A^{(1)}$ and $G^{(1)}$. Each of these sub–sets (one of the four quadrants) have doublets belonging to codons sharing the same *first* base and, therefore, are related by *biosynthetic pathways* [14], (see also [4,6]). There exist another partitioning where these sub–sets correspond to doublets (codons) sharing the same *second* base and being related by similar physico–chemical properties [14] (see below, Eq. (8), and in the sequel for the codon set). We shall return later to these various "views".

Let us now consider the action of our groups. For the first set, the transformations act as follows:



In this case, the Rumer transformation is implemented in all cases by V_1. $\mathbf{T_2}$ leaves $U^{(1)}$, $C^{(1)}$, $A^{(1)}$ and $G^{(1)}$ globally invariant by implementing the Rumer transformation at the *second* position. $\mathbf{T_3}$ at the same time exchanges $U^{(1)}$, $C^{(1)}$, $A^{(1)}$ and $G^{(1)}$, according to the Rumer transformation, and implements this latter only at the *first* position. $\mathbf{T_4}$ acts as $\mathbf{T_3}$ but, here, the two positions are concerned. Importantly, $\mathbf{T_4}$ exchanges M_1 and M_2 and the Rumer transformation acts on both positions. Now, let us consider the action of the second group V_2. We have in this case:



Here, the situation is different from the preceding one. In all three cases, $U^{(1)}$, $C^{(1)}$, $A^{(1)}$ and $G^{(1)}$ are globally conserved (*first* base invariant) with a specific action on each one of them at the *second* base position:

$\mathbf{S_2}$: conserves *strictly* $G^{(1)}$, implements UCAG<–>CUGA in $A^{(1)}$ *and* UCAG<–>AGUC in $C^{(1)}$ *and* the Rumer transformation in $U^{(1)}$.

$\mathbf{S_3}$: conserve *strictly* $U^{(1)}$, implements UCAG<–>CUGA in $C^{(1)}$ *and* UCAG<–>AGUC in $A^{(1)}$ *and* the Rumer transformation in $G^{(1)}$.

$\mathbf{S_4}$: implements the Rumer transformation in all cases.

Finally, and for later use, we consider the action of V_2 on the following transformed form of **D**, **D'**, which classifies together the doublets sharing the same second base:

$$\mathbf{D}'=\mathbf{X}_1^{-1}\mathbf{D}\ \mathbf{X}_1=\begin{pmatrix} UU & CU & UC & CC \\ AU & GU & AC & GC \\ UA & CA & UG & CG \\ AA & GA & AG & GG \end{pmatrix},\ \mathbf{X}_1=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (8)$$

The group V_2 act as follows on **D'**:

GU AU AC GC	**UU CU** CC UC	GU AU GC AC
CU UU UC CC	**AU GU** GC AC	CU UU CC UC
CA UA **UG CG**	AA GA GG AG	GA AA GG AG
GA AA **AG GG**	UA CA CG UG	CA UA CG UG
V_2 : $(\mathbf{S_2})^{-1}\mathbf{D'U_2}$	$(\mathbf{S_3})^{-1}\mathbf{D'S_3}$	$(\mathbf{S_4})^{-1}\mathbf{D'S_4}$

Here, the situation is analogous to the preceding one, but the roles of the base–position are exchanged, *i.e.*, $U^{(2)}$, $C^{(2)}$, $A^{(2)}$ and $G^{(2)}$ are globally conserved (*second* base invariant) with a specific action on each one of them at the *first* base position. We have in the detail

$\mathbf{S_2}$: conserves *strictly* $G^{(2)}$, implements UCAG<–>CUGA in $A^{(2)}$ *and* UCAG<–>AGUC in $C^{(2)}$ *and* the Rumer transformation in $U^{(2)}$.

$\mathbf{S_3}$: conserve *strictly* $U^{(2)}$, implements UCAG<–>CUGA in $C^{(2)}$ *and* UCAG<–>AGUC in $A^{(2)}$ *and* the Rumer transformation in $G^{(2)}$.

$\mathbf{S_4}$: implements the Rumer transformation in all cases.

3.3 Rumer and Rumer–Konopel'chenko symmetry for C

Now, we turn to the 64 codons matrix in Eq. (3), and study its symmetries. As for the set of doublets, we consider the eight Wittmann Octets (WO), which are the eight rows in Eq. (3) numbered from 1 to 8, from top to bottom. From the many ways to define graphs between these WOs, we have retained the ones below because first, the eight graphs contain one, the adjacency matrix of which corresponds to the Rumer transformation, a basic transformation. Second, the adjacency matrices of these graphs constitute a symmetry group which is a dihedral group D_8 with eight elements containing, as symmetry sub–groups, a Klein group V and a cyclic group with two elements C_2. We find again, in a new way, a result [6] that this end–of–the–chain group C_2 could describe an invariant part of the 64–codon set, more exactly 16 codons, which appear to be invariant across all species with respect to their use, in their respective genetic codes (see the end of this section). Third, these symmetry groups seem to fit nicely into the Rumer–division and the "5'–bisections" of the 64–codon set used by Shcherbak, (see [1] and the references therein), to establish the many arithmetical regularities inside the genetic code. We shall return to this latter point in section 4. Consider now the following eight graphs, G_i (i = 1, 2, …, 8):

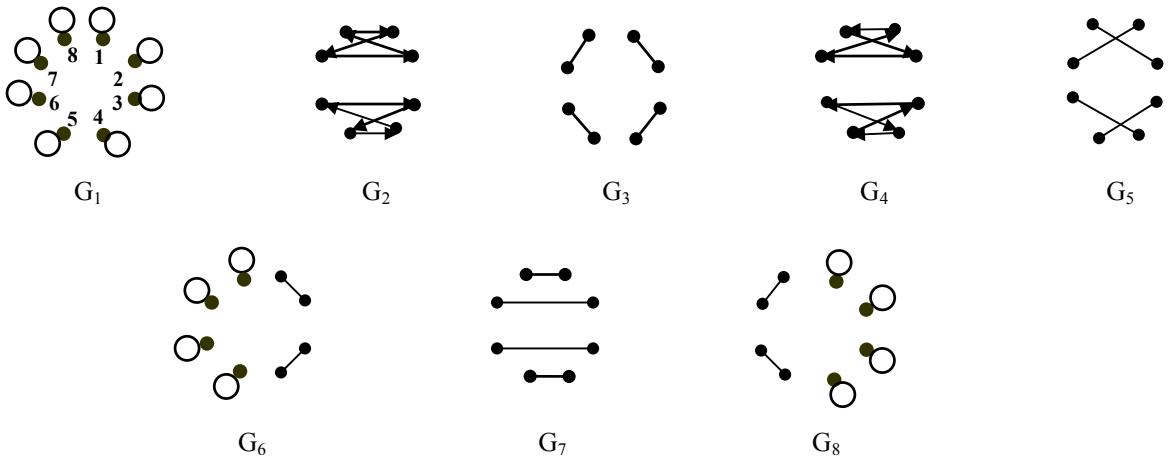


Figure 6. Graph–set for the eight Wittmann's Octets (see text).

These are permutation graphs on the eight Wittmann Octets. For example G_2 corresponds to the permutation (7 8 2 1)(5 6 4 3) and G_6 to the permutation (2 1)(4 3)(5 6 7 8). It is immediate to write down the adjacency matrices for these graphs. They are:

$$\mathbf{R_1}=\begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{pmatrix},\quad \mathbf{R_2}=\begin{pmatrix}0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&1&0&0&0&0\\0&0&1&0&0&0&0&0\\0&1&0&0&0&0&0&0\\1&0&0&0&0&0&0&0\end{pmatrix},\quad \mathbf{R_3}=\begin{pmatrix}0&1&0&0&0&0&0&0\\1&0&0&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\end{pmatrix},\tag{9.1}$$

$$\mathbf{R_4}=\begin{pmatrix}0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&0&0\\0&0&0&0&1&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\end{pmatrix},\quad \mathbf{R_5}=\begin{pmatrix}0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\end{pmatrix},\quad \mathbf{R_6}=\begin{pmatrix}0&1&0&0&0&0&0&0\\1&0&0&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{pmatrix},\tag{9.2}$$

$$\mathbf{R_7}=\begin{pmatrix}0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&0&0\\0&0&0&0&1&0&0&0\\0&0&0&1&0&0&0&0\\0&0&1&0&0&0&0&0\\0&1&0&0&0&0&0&0\\1&0&0&0&0&0&0&0\end{pmatrix},\quad \mathbf{R_8}=\begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\end{pmatrix}.$$

With respect to the ordinary matrix product, we obtain the following multiplication table (Table 1) of these matrices, or the Cayley table, as it is called in group theory.

Table 1. Multiplication table for the eight matrices $\mathbf{R_i}$

	$\mathbf{R_1}$	$\mathbf{R_2}$	$\mathbf{R_3}$	$\mathbf{R_4}$	$\mathbf{R_5}$	$\mathbf{R_6}$	$\mathbf{R_7}$	$\mathbf{R_8}$
$\mathbf{R_1}$	$\mathbf{R_1}$	$\mathbf{R_2}$	$\mathbf{R_3}$	$\mathbf{R_4}$	$\mathbf{R_5}$	$\mathbf{R_6}$	$\mathbf{R_7}$	$\mathbf{R_8}$
$\mathbf{R_2}$	$\mathbf{R_2}$	$\mathbf{R_3}$	$\mathbf{R_4}$	$\mathbf{R_1}$	$\mathbf{R_8}$	$\mathbf{R_5}$	$\mathbf{R_6}$	$\mathbf{R_7}$
$\mathbf{R_3}$	$\mathbf{R_3}$	$\mathbf{R_4}$	$\mathbf{R_1}$	$\mathbf{R_2}$	$\mathbf{R_7}$	$\mathbf{R_8}$	$\mathbf{R_5}$	$\mathbf{R_6}$
$\mathbf{R_4}$	$\mathbf{R_4}$	$\mathbf{R_1}$	$\mathbf{R_2}$	$\mathbf{R_3}$	$\mathbf{R_6}$	$\mathbf{R_7}$	$\mathbf{R_8}$	$\mathbf{R_5}$
$\mathbf{R_5}$	$\mathbf{R_5}$	$\mathbf{R_6}$	$\mathbf{R_7}$	$\mathbf{R_8}$	$\mathbf{R_1}$	$\mathbf{R_2}$	$\mathbf{R_3}$	$\mathbf{R_4}$
$\mathbf{R_6}$	$\mathbf{R_6}$	$\mathbf{R_7}$	$\mathbf{R_8}$	$\mathbf{R_5}$	$\mathbf{R_4}$	$\mathbf{R_1}$	$\mathbf{R_2}$	$\mathbf{R_3}$
$\mathbf{R_7}$	$\mathbf{R_7}$	$\mathbf{R_8}$	$\mathbf{R_5}$	$\mathbf{R_6}$	$\mathbf{R_3}$	$\mathbf{R_4}$	$\mathbf{R_1}$	$\mathbf{R_2}$
$\mathbf{R_8}$	$\mathbf{R_8}$	$\mathbf{R_5}$	$\mathbf{R_6}$	$\mathbf{R_7}$	$\mathbf{R_2}$	$\mathbf{R_3}$	$\mathbf{R_4}$	$\mathbf{R_1}$

This is the Cayley table of the dihedral group with eight elements D_8. Consider now the codon–matrix **C**, in Eq. (3), which we reproduce below (for ease of comparison) as table **C** ($\mathbf{R_1}$ is the identity so it corresponds to **C**). In the other seven tableaux, we show the action of the remaining seven transformation matrices. They are the following:

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

C

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

$(\mathbf{R_6})^{-1}\mathbf{C}\,\mathbf{R_6}$

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

$(\mathbf{R_8})^{-1}\mathbf{C}\mathbf{R_8}$

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

$(\mathbf{R_3})^{-1}\mathbf{C}\,\mathbf{R_3}$

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

$(\mathbf{R_7})^{-1}\mathbf{CR_7}$

GGG	GGA	GAG	GAA	AGA	AGG	AAA	AAG
GGC	GGU	GAC	GAU	AGU	AGC	AAU	AAC
GCG	GCA	GUG	GUA	ACA	ACG	AUA	AUG
GCC	GCU	GUC	GUU	ACU	ACC	AUU	AUC
CGC	CGU	CAC	CAU	UGU	UGC	UA**U**	UA**C**
CGG	CGA	CAG	CAA	UGA	UG**G**	UA**A**	UA**G**
CCC	CCU	CUC	CUU	UC**U**	UC**C**	UU**U**	UU**C**
CCG	CCA	CUG	CUA	UC**A**	UC**G**	UU**A**	UU**G**

$(\mathbf{R_2})^{-1}\mathbf{CR_2}$

GG**U**	GG**C**	GA**U**	GA**C**	AGC	AGU	AAC	AAU
GG**A**	GG**G**	GA**A**	GA**G**	AGG	AGA	AAG	AAA
GC**U**	GC**C**	GU**U**	GU**C**	ACC	ACU	AUC	AUU
GC**A**	GC**G**	GU**A**	GU**G**	ACG	ACA	AUG	AUA
CGA	CGG	CAA	CAG	UGG	UGA	UAG	UAA
CGU	CGC	CAU	CAC	UGC	UGU	UAC	UAU
CCA	CCG	CUA	CUG	UCG	UCA	UUG	UUA
CCU	CCC	CUU	CUC	UCC	UCU	UUC	UUU

$(\mathbf{R_4})^{-1}\mathbf{CR_4}$

GG**U**	GG**C**	GA**U**	GA**C**	AGU	AGC	AAU	AAC
GG**A**	GG**G**	GA**A**	GA**G**	AGA	AG**G**	AA**A**	AA**G**
GC**U**	GC**C**	GU**U**	GU**C**	ACU	ACC	AUU	AUC
GC**A**	GC**G**	GU**A**	GU**G**	ACA	AC**G**	AU**A**	AU**G**
CG**U**	CG**C**	CA**U**	CA**C**	UGU	UGC	UAU	UAC
CG**A**	CG**G**	CA**A**	CA**G**	UGA	UG**G**	UA**A**	UA**G**
CC**C**	CC**U**	CU**C**	CU**U**	UCU	UCC	UUU	UUC
CC**G**	CC**A**	CU**G**	CU**A**	UC**A**	UC**G**	UU**A**	UU**G**

$(\mathbf{R_5})^{-1}\mathbf{CR_5}$

In this sub–section, we shall continue to use M_1 and M_2 to name the two members of the Rumer bisection, see section 3, (M_1 is in gray). Now, from the above tables, we see that the general action is concentrated mainly at the third base–position, and the Rumer transformation, acting exclusively, concerns only the first and second base–positions. The group D_8 has, as an interesting sub–group, the Klein group V with elements {$\mathbf{R_1}$, $\mathbf{R_3}$, $\mathbf{R_6}$, $\mathbf{R_8}$}. This latter group leaves M_1 and M_2 invariant and, consequently, leaves also invariant all 16 "family boxes", *i.e.*, all quartets of (four) codons sharing the same first two bases (separated from each other by small dashed lines in the tableaux). Thus, this means that, possibly, only the third base will be altered. Next, the transformation $\mathbf{R_7}$ implements the Rumer transformation UCAG<–>GACU at all *three* base–positions. As for the other three transformations, $\mathbf{R_2}$, $\mathbf{R_4}$ and $\mathbf{R_5}$ of D_8, they alter the first two base–positions according to the Rumer transformation with, possibly, full invariance or alteration by the (secondary Rumer

transformations mentioned in section 3.1) at the third base–position, see the tableaux. As in section 3.2, we could consider either the Bio–Synthetic Classes or Physico–Chemical Classes, using the same notation as for the doublets in the last section: $U^{(1)}$, $C^{(1)}$, $A^{(1)}$ and $G^{(1)}$ and $U^{(2)}$, $C^{(2)}$, $A^{(2)}$ and $G^{(2)}$, respectively. For the former, we shall unit the four Bio–Synthetic Classes into two big ones, the Pyrimidine Bio–Synthetic Class (Pyr–BC), made of $U^{(1)}$ and $C^{(1)}$, and the Purine Bio–Synthetic Class (Pur–BC), made of $A^{(1)}$ and $G^{(1)}$. The first class occupies the four top rows, in matrix (3), and the second class occupy the last four. In this way, the transformations induced by the matrices $\mathbf{R_i}$ act as follows: $\mathbf{R_1}$, $\mathbf{R_3}$, $\mathbf{R_6}$, $\mathbf{R_8}$, the sub–group V, conserve (globally) Pyr–BC and Pur–BC while $\mathbf{R_2}$, $\mathbf{R_4}$ and $\mathbf{R_5}$ exchange them. Moreover, $\mathbf{R_6}$ conserves *strictly* the 16 codons of $G^{(1)}$ and $\mathbf{R_8}$ conserves *strictly* the 16 codons of $U^{(1)}$. For the latter, it is interesting to transform the codon–matrix **C** as follows:

$$\mathbf{C}' = \mathbf{X}^{-1}\,\mathbf{C}\,\mathbf{X} = \begin{pmatrix} UUU & UUC & CUU & CUC & UCU & UCC & CCU & CCC \\ UUA & UUG & CUA & CUG & UCA & UCG & CCA & CCG \\ AUU & AUC & GUU & GUC & ACU & ACC & GCU & GCC \\ AUA & AUG & GUA & GUG & ACA & ACG & GCA & GCG \\ UAU & UAC & CAU & CAC & UGU & UGC & CGU & CGC \\ UAA & UAG & CAA & CAG & UGA & UGG & CGA & CGG \\ AAU & AAC & GAU & GAC & AGU & AGC & GGU & GGC \\ AAA & AAG & GAA & GAG & AGA & AGG & GGA & GGG \end{pmatrix}, \tag{10}$$

where **X** is given by (see Eq. (8)):

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{11}$$

Here, the symbol $\otimes$ is for the usual Kronecker product of matrices. In (10), $U^{(2)}$, $C^{(2)}$, $A^{(2)}$ and $G^{(2)}$ occupy the four quadrants and we could apply the transformations $\mathbf{R_i}$, as for **C**. The transformation patterns are the same. Proceeding as before, we have that the Physico–Chemical Classes Pyr–PC and Pur–PC, which are invariant under the action of $\mathbf{R_3}$, $\mathbf{R_6}$, $\mathbf{R_8}$, are exchanged by $\mathbf{R_7}$, $\mathbf{R_2}$, $\mathbf{R_4}$ and $\mathbf{R_5}$. (Alternatively, we could let the transformations defined recently in reference [6] act directly on the matrix **C** and we obtain the same results.) Concerning the physico–chemical properties, Pelc, in 1965 ([22], see also [23]), analyzed the correlation between the triplets codons and the structure of the amino acids and found that codons with U or C, as a *second* base, correspond mainly to hydrophobic and weakly polar amino acids, whereas those containing A or G, as a *second* base, involve strongly polar amino acids. To close this section, let us give a last example of a symmetry group for **C** in which the family boxes of a given Bio–Synthetic Class are mixed. Consider the transformations corresponding to the permutations (1)(23)(45678), (123458)(67) and (their product) (1458)(23)(67) of the Wittmann Octets. As adjacency matrices of the corresponding graphs (not drawn here) they read, respectively (without the identity $\mathbf{W_1}$):

$$\mathbf{W_2}=\begin{pmatrix}1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{pmatrix},\ \mathbf{W_3}=\begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\end{pmatrix},\ \mathbf{W_4}=\begin{pmatrix}1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\end{pmatrix}. \qquad (12)$$

The set $\{\mathbf{W_1}, \mathbf{W_2}, \mathbf{W_3}, \mathbf{W_4}\}$ constitutes also a Klein group V. This group acts analogously to $\mathbf{R_3}$, $\mathbf{R_6}$, $\mathbf{R_8}$ but, here, $\mathbf{W_2}$, $\mathbf{W_3}$ leave invariant 36 codons and $\mathbf{W_4}$ only 16. The Rumer sets become mixed but the Bio–Synthetic Classes $U^{(1)}$, $C^{(1)}$, $A^{(1)}$ and $G^{(1)}$ are globally conserved with possible mixing between family boxes. The group above acts as follows (compare to the action of the first group V, in Eq. (9.2)):

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

C

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

$(\mathbf{W_2})^{-1}\mathbf{C}\,\mathbf{W_2}$

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

$(\mathbf{W_3})^{-1}\mathbf{C}\mathbf{W_3}$

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

$(\mathbf{W_4})^{-1}\mathbf{C}\,\mathbf{W_4}$

We see that the Bio–Synthetic Classes keep their individuality (globally invariant) and the codons of the amino acids of a given class get exchanged in various manners. Take $\mathbf{W_2}$ for example. It acts in $U^{(1)}$ in such a way that it reproduce the action of $\mathbf{X}_1$ in Eq. (8): the exchanges are vertical, horizontal and diagonal. The exchanges in $C^{(1)}$ are only vertical and in $A^{(1)}$ only horizontal. As for $G^{(1)}$, its 16 codons are strictly conserved, as with the action of $\mathbf{R_6}$ (see above). A similar analysis could be made with $\mathbf{W_3}$ and $\mathbf{W_4}$. Before closing this section, let us return to our first Klein group V: {$\mathbf{R_1}$, $\mathbf{R_3}$, $\mathbf{R_6}$, $\mathbf{R_8}$}. As we have mentioned in the introduction, the manner we have built these transformations is guided. In reference [6], we were motivated by the search of transformation groups that lead to invariance of groups of codons across all species. It is well known that some living species use differently the genetic code. They use a variant form of the "standard" genetic code where, in general and in each case, only few codons are concerned and there are several known variants, see [15,16]. These variant codons code evidently for a different amino acid, other than the one in the above table which is the standard form The table below (Table 2), represents Eq. (3) of section 2.3 with the additional information about the amino acids and their molecular weights or total number of nucleons. The codons which are subject to variations, considering all (known) genomes, nuclear and mitochondrial, and taken from [16], are underlined. Examining the above table reveals that the Bio–Synthetic Class $G^{(1)}$, and only this class among the four, has 16 codons which are never subject to variations, [6]. The authors of reference [15] explain that there are several reasons why the codons in the form GNN (N: any base, U, C, A and G) are thought to be the most primitive. Loomis, [17], in the beautiful article entitled "Origin of life" mentions also the five amino acids glycine, valine, alanine, aspartic acid and glutamic acid (those of class $G^{(1)}$) when discussing the appearance of the hypercycles in the pre–biotic soup.

Table 2. The standard genetic code table

UUU Phe	**UUC** 165	**UCU**	**UCC**	**CUU**	**CUC**	**CCU**	CCC
Leu **UUA**	131 **UUG**	Ser **UCA**	105 UCG	Leu CUA	131 CUG	Pro **CCA**	115 **CCG**
UAU Tyr	UAC 181	**UGU** Cys	UGC 121	CAU His	**CAC** 155	CGU	**CGC**
Stop1 **UAA**	Stop2 UAG	Stop3 UGA	Trp204 **UGG**	Gln **CAA**	146 CAG	Arg CGA	174 **CGG**
AUU Ile	AUC 131	ACU Thr	**ACC** 119	**GUU** Val	GUC 117	GCU Ala	**GCC** 89
AUA	Met149 AUG	**ACA**	ACG	GUA	**GUG**	GCA	**GCG**
AAU Asn	**AAC** 132	AGU Ser	AGC 105	GAU Asp	GAC 133	**GGU**	**GGC**
Lys AAA	146 **AAG**	Arg **AGA**	174 **AGG**	Glu **GAA**	147 **GAG**	Gly **GGA**	75 GGG

As a result, we have [6], that the cyclic group C_2, a sub–group of a certain Klein group, itself sub–group of a dihedral group D_8 according to the chain $D_8 \subset V \subset C_2$, as an end–of–chain group implements the invariance of the above mentioned 16 codons of $G^{(1)}$. In the context of the present work, the group C_2 corresponds to the set {$\mathbf{R_1}$, $\mathbf{R_6}$ }, sub–group of V ($\mathbf{R_1}$, $\mathbf{R_3}$ $\mathbf{R_6}$, $\mathbf{R_8}$), itself sub–group of D_8 ($\mathbf{R}_i$, i = 1, 2, …, 8).

The guiding idea, in considering this problem, was to define graphs with a maximum of four invariant rows, and their variations. For example, in the graph $\mathbf{G_6}$, of section 3.3, the nodes 5, 6, 7 and 8 are connected only to themselves (the self–loops) while the nodes 1, 2, 3 and 4 are subject to permutations with an edge between the nodes 1 and 2, in the one hand and 3 and 4, on the other. Furthermore, as we explained in section 2, these transformations between nodes concern at the same time the rows and the columns. In this way, we constructed the corresponding adjacency matrix, $\mathbf{R_6}$, which gives us, in this case, strict invariance in $G^{(1)}$ and transitions in $A^{(1)}$, both belonging to the four last rows and only transversions in $U^{(1)}$ and $C^{(1)}$, both belonging to the first four rows. For all other transformations in this work, we proceeded analogously.

4 MATRIX–NORMS, TRACES, AND SHCHERBAK'S *QUANTUM PRIME* 037

As mentioned in the introduction, we end this article by presenting some (striking) numbers which arise by "sizing" our classification matrices for the 16 doublets and the 64 codons. Before giving these results let us, first, recall some numbers pertaining to the genetic code functioning. Besides the (ubiquitous) existence of 64 codons and 20 (canonical) amino acids, there is a total of 44 *degenerate* codons (64–20) of which 24 are in M_1 and 20 in M_2 (counting the three stop codons). Note that not including the latter reduces this number to 41, see below.

Another (complementary) manner to reckon, which goes back to the physicist Georges Gamow (1954) see [1], is based on the division of the 64 triplet codons set into three sub–sets, depending on their composition (identical and unique base) regardless of their types and positions in the codons. This partitioning is shown in Tables 3–5 (in the first three rows in each table)

Table 3. The "**36**" Family (two identical bases)

CUU	UUA	UUG	CUC	CCA	CCG	AAU	AAC	AGA	GGU	GGC	AGG
UCU	AUU	GUU	CCU	ACC	GCC	UAA	CAA	AAG	UGG	CGG	GAG
UUC	UAU	UGU	UCC	CAC	CGC	AUA	ACA	GAA	GUG	GCG	GGA
444	555	666	555	777	888	777	888	1110	999	1110	1221

Table 4. The "**24**" Family one unique base)

CAU	CUA	CUG	GUC	UGA	AGU	AGC	ACG
UCA	ACU	GCU	CGU	AUG	UAG	CAG	GAC
AUC	UAC	UGC	UCG	GAU	GUA	GCA	CGA
666	666	777	777	888	888	999	999

Table 5. The "**4**" Family (three identical bases)

UUU	CCC	AAA	GGG
111	222	333	444

There are therefore 36 codons in the "**36**" Family (Table 3), 24 codons in the "**24**" Family (Table 4) and 4 codons in the "**4**" Family (Table 5). Note that 60 (36 + 24) is also an important number which arise in the combinatorial necklace model of the genetic code [18], where the structure consists of 64 beads (the codons) of four colors (the bases) and the colored beads form hanging vertical decorations or chains of three ($x = 3$) beads. In this model, there are $y = 4^3$ distinct vertical chains. The total number of possible vertical decorations containing at least two colors each is $y^x - y = 60$ and $y = 4$ decorations are of the same color.

Now, we turn to our results. We have found interesting to assign *attributes* to the bases and *norms* to the (classification) matrices. Let us begin by the norms, more exactly the matrix–norms. (The latter, are widely used in many applications of linear algebra, and have been termed the "yardsticks". Also, they are often used in matrix–based algorithms.) These norms capture the essential notions of size and distance in vector spaces. The matrix–norms of a matrix A, we shall use in this paper, are the 1–norm, $||\mathbf{A}||_1$, which is equal to the *maximum column sum* and the ∞–norm, $||\mathbf{A}||_\infty$, which is equal to *the maximum row sum*. They are defined as follows

$$\|\mathbf{N}\|_1 = \max_{1 \le \mathbf{j} \le \mathbf{n}} \sum_{\mathbf{i}=1}^{\mathbf{n}} |\mathbf{a_{ij}}| \qquad \|\mathbf{N}\|_\infty = \max_{1 \le \mathbf{j} \le \mathbf{n}} \sum_{\mathbf{j}=1}^{\mathbf{n}} |\mathbf{a_{ij}}| \tag{13}$$

where the modules in the sums are vector norms. We have therefore a mean to compare the rows between themselves and the columns between themselves. Now, concerning the attributes of the bases, we shall concentrate on their "molecular size" and also on their hydrogen bonding properties. We shall take as the "molecular size index" of a given base the following expression

$$\mu = 2\,(n_N - n_C) + n_O + 2 \tag{14}$$

Where the numbers refer to the number of nitrogen, carbon and oxygen atoms, respectively, in a given base (U, C, A and G). This formula has been introduced in 1991 by Rosen [19], (see also [2]). The size index, μ, captures the monotonic size gradation of the four bases: U = 0, C = 1, A = 2, G = 3 (see section 2.1). This ordering, among others, is often considered (see [13] for a recent example). As for the hydrogen bond attribute, it is unambiguous: U = A = 2 and C = G = 3. Now, let us gather these numbers into the two following simple 2×2 matrices, respectively $\boldsymbol{\mu}$ and **h**.

$$\boldsymbol{\mu} = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} + \mathbf{J} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{h} = \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}. \tag{15}$$

We have slightly translated the (would–be) matrix $\boldsymbol{\mu}$ by one unit by adding the matrix **J**, the all–1 2×2 matrix, mentioned in section 3.1, just to avoid the zero. This is justified because this "attribute" is certainly rough but, as we shall see, it has interesting consequences. Let us begin by considering the set of 16 doublets, in Eq. (2). In this case, the rows which are sized by $||\mathbf{A}||_\infty$, are the Wittmann Doublets and the columns which are compared by $||\mathbf{A}||_1$ are the sub–sets of doublets sharing the same total number of hydrogen bonds. The matrix encoding the sizes of the doublets is computed as follows:

$$\boldsymbol{\Pi}^{(\mathbf{ms})} = \boldsymbol{\mu} \otimes \mathbf{J} + \mathbf{J} \otimes \boldsymbol{\mu} = \begin{pmatrix} 2 & 3 & 3 & 4 \\ 4 & 5 & 5 & 6 \\ 4 & 5 & 5 & 6 \\ 6 & 7 & 7 & 8 \end{pmatrix}. \tag{16}$$

Were **J** is the 2×2 all–1 matrix mentioned above. For the second attribute, the following matrix gives the total number of hydrogen bonds in each doublet.

$$\boldsymbol{\Pi}^{(\mathbf{hb})} = \mathbf{h} \otimes \mathbf{J} + \mathbf{J} \otimes \mathbf{h} = \begin{pmatrix} 4 & 5 & 5 & 6 \\ 4 & 5 & 5 & 6 \\ 4 & 5 & 5 & 6 \\ 4 & 5 & 5 & 6 \end{pmatrix}. \tag{17}$$

For these matrices we have:

$$\left\|\boldsymbol{\Pi}^{(\mathrm{ms})}\right\|_1 = 24, \quad \left\|\boldsymbol{\Pi}^{(\mathrm{ms})}\right\|_\infty = 28, \quad \left\|\boldsymbol{\Pi}^{(\mathrm{hb})}\right\|_1 = 24, \quad \left\|\boldsymbol{\Pi}^{(\mathrm{hb})}\right\|_\infty = 20, \quad \left\|\boldsymbol{\Pi}^{(\mathrm{hb})}\right\|_1 + \left\|\boldsymbol{\Pi}^{(\mathrm{hb})}\right\|_\infty = 44. \tag{18}$$

Note that the traces of the matrices (16) and (17) are both equal to 20. It is also interesting to compute the sum of the (hydrogen bond) attribute over the sets M_1 and M_2 (see the above sections). We have:

$$\Sigma_{M_1}^{(\mathrm{hb})} = 44, \quad \Sigma_{M_2}^{(\mathrm{hb})} = 36. \tag{19}$$

Now, for the codons, we have:

$$\mathbf{\Theta^{(ms)}} = \boldsymbol{\mu} \otimes \mathbf{J} \otimes \mathbf{J} + \mathbf{J} \otimes \boldsymbol{\mu} \otimes \mathbf{J} + \mathbf{J} \otimes \mathbf{J} \otimes \boldsymbol{\mu} = \begin{pmatrix} 3 & 4 & 4 & 5 & 4 & 5 & 5 & 6 \\ 5 & 6 & 6 & 7 & 6 & 7 & 7 & 8 \\ 5 & 6 & 6 & 7 & 6 & 7 & 7 & 8 \\ 7 & 8 & 8 & 9 & 8 & 9 & 9 & 10 \\ 5 & 6 & 6 & 7 & 6 & 7 & 7 & 8 \\ 7 & 8 & 8 & 9 & 8 & 9 & 9 & 10 \\ 7 & 8 & 8 & 9 & 8 & 9 & 9 & 10 \\ 9 & 10 & 10 & 11 & 10 & 11 & 11 & 12 \end{pmatrix} \quad (20)$$

and

$$\mathbf{\Theta^{(hb)}} = \mathbf{h} \otimes \mathbf{J} \otimes \mathbf{J} + \mathbf{J} \otimes \mathbf{h} \otimes \mathbf{J} + \mathbf{J} \otimes \mathbf{J} \otimes \mathbf{h} = \begin{pmatrix} 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \\ 6 & 7 & 7 & 8 & 7 & 8 & 8 & 9 \end{pmatrix} \quad (21)$$

It is important to carry out the products in the order displayed, to reproduce the correct number of the sum of the molecular indices or the total number of hydrogen bonds in all the 64 codons. In both definitions (20) and (21), the first term gives the contribution of the first base, the second that of the second base and finally the third term adds the contribution of the third base. This is also true for the doublet matrices in Eqs. (16) and (17) with only two bases. For the matrix in (20), we have $|| \mathbf{\Theta^{(ms)}} ||_1 = 72$, $|| \mathbf{\Theta^{(ms)}} ||_\infty = 84$, and $\mathrm{Tr}(\mathbf{\Theta^{(ms)}}) = 60$. For the one in Eq. (21), we obtain $|| \mathbf{\Theta^{(hb)}} ||_1 = 72$, $|| \mathbf{\Theta^{(hb)}} ||_\infty = 60$, and $\mathrm{Tr}(\mathbf{\Theta^{(hb)}}) = 60$. It is interesting to note that the last number, 60, plays also an important role in the work on the genetic code by Petoukhov [25], who constructed a similar matrix to matrix (21). We could also use, in place of the Kronecker product, the Kronecker product with *concatenation* (see section 2.2 and footnote 2) to obtain the interesting matrix (the double vertical bars are for concatenation):

$$\mathbf{\Theta_c^{(ms)}} = \boldsymbol{\mu} \parallel \boldsymbol{\mu} \parallel \boldsymbol{\mu} = \begin{pmatrix} 111 & 112 & 121 & 122 & 211 & 212 & 221 & 222 \\ 113 & 114 & 123 & 124 & 213 & 214 & 223 & 224 \\ 131 & 132 & 141 & 142 & 231 & 232 & 241 & 242 \\ 133 & 134 & 143 & 144 & 233 & 234 & 243 & 244 \\ 311 & 312 & 321 & 322 & 411 & 412 & 421 & 422 \\ 313 & 314 & 323 & 324 & 413 & 414 & 423 & 424 \\ 331 & 332 & 341 & 342 & 431 & 432 & 441 & 442 \\ 333 & 334 & 343 & 344 & 433 & 434 & 443 & 444 \end{pmatrix} \quad (22)$$

This operation, using the concatenation of matrix elements, has been introduced in [3] to make

sure that the non–commutativity of the bases in the triplets is respected. For example, in the above matrix the triplets UUC, UCU and CUU (see Table 2) are assigned the numbers 112, 121 and 211, respectively, which are all different, as we want. In fact, all the 64 triplets are numerically distinct. Had we taken only the ordinary Kronecker product, these assignments would have been 2 in all three cases, because the latter product involves ordinary multiplication of numbers. Moreover, the codon–numbers, generated by concatenation are interesting because when they are grouped into 20 triplets (of triplets codons) that is 12 triplets in the "**36**" Family and 8 triplets in the "**24**" Family, they show regular numerical patterns, see Tables 3–5 and below. Now, we have for the norms and the trace, in this case:

$|| \mathbf{\Theta_c}^{(ms)} ||_1 = 2664, || \mathbf{\Theta_c}^{(ms)} || \infty = 3108, \mathrm{Tr}(\mathbf{\Theta_c}^{(ms)}) = 2220.$

The comparison of the numbers in Eqs. (18) and (19) and those described at the beginning of this section is, at least, striking. Even the number 28, which was not yet mentioned, could possibly finds its place in the quasi 28–gon (or *Icosikaioctagon*) recent theory of the genetic code by Yang, which is based solely on the set of the 16 doublets [13]. What is really amazing is that if one permutes the two numbers, 3 and 4, in the second row of the first matrix (μ) in Eq. (15), which corresponds to another ordering of the bases (U = 0, C = 1, G = 2, A = 3, see above) chosen by Yang, and relying on another attribute, the sp^2 N–atom number, where N is for nitrogen, then the obtained doublet–matrix as in Eq. (16) has for its trace and norms $\mathrm{Tr}(\mathbf{\Pi}^{(ms)}) = 16, || \mathbf{\Pi}^{(ms)} ||_1 = 20$, by $|| \mathbf{\Pi}^{(ms)} || \infty = 28$, *i.e.*, just the right three numbers in Yang's theory, which we recall, is based solely on the set of 16 doublets. Also the numbers 60 and 72 are interesting. From Eq. (21), we see that all the Wittmann Octets, the rows, have the same total number of hydrogen bonds which is 60. We have already shown the relevance of this number to the combinatorial necklace model [18]. As for the second, 72, interestingly, we have found in the recent literature connected with the use of graph theory to study the structural properties of the metabolic pathways (Fell and Wagner, [21]), that 72 happens to be the total number of reactions in the nucleotide and nucleoside biosynthesis. At least, this could happen to be just a happy coincidence.

Now, before ending this article, let us say some words on the work by Shcherbak, as promised in the introduction. Shcherbak, by using the Rumer–Konopel'chenko partitioning of the 64 codons and the number of nucleons in the amino acids, separating the side chain, which is variable, from the standard blocks (with 74 nucleons), which are the same for all amino acids, established several number of amazing *exact balances* and *regularities*. Let us mention only some representative ones, but there are many other interesting ones in [1]. First, by retaining the Rumer division, one has (see Table 2):

- Total number of nucleons (whole molecule) in M_1: 925
- Total number of nucleons in the blocks, in M_1: 592
- Total number of nucleons in the side chains, in M_1: 333

- Total number of nucleons in the side chains, in M_2: 1110
- Total number of nucleons in the blocks, in M_2: 1110

For M_2, there is *exact balance*. For M_1, by invoquing the Prime Quantum Divisibility by 37, one is left with the first Pythagorean triple (3, 4, 5): $333/37 = 3^2$, $592/37 = 4^2$ and $925/37 = 5^2$. As a second example, consider the Pyrimidine Bio–Synthetic Class (Pyr–BC), and the Purine Bio–Synthetic Class (Pur–BC), considered in the last section. They correspond in [1] to the series with the 5'–pyrimidine bases (5' bisection) and to the series with the 5' purine bases, respectively. In the first cases there is again an *exact balance*: the sum of nucleons in the standard blocks is 814 and the sum in the side chains is also 814, with a grand total 1628. In the other case there is no balance but the total number of nucleons is 1517. Let us note that the above numbers, 1628 and 1517, like all others are both divisible by 37. We have that 1628/37 = 44 and 1517/37 = 41. We have already met these two numbers, 44 and 41 (see the beginning of this section). Note, importantly, that the difference between 44 and 41 (that is with or without the 3 stop–codons), is correctly "managed": the stop–codons belong to the Pyrimidine Bio–Synthetic Class (Pyr–BC) or the 5' bisection, see Table 2. Shcherbak has shown also that the digital representation of the balances, as consequence of the criterion of divisibility by 37, acquires the unique form aaa with a = 1, 2,…, 9. It is interesting to note that similar forms arise also for the codons independently of the amino acids and the numbers of their nucleons, when they are grouped *à la* Gamow (see the "**36**", "**24**" and "**4**" Families above), into triplets (of codons). The codons in the "**36**" Family are in bold characters, in Table 2. In the fourth rows of Tables 1, 2 and 3, we have computed the sum of the numbers of the codons from the matrix (22). There is nothing magical with these particular forms: they are simply a consequence of the mathematical fact that when one sums three three–digit decimal numbers by carrying out the (three) permutations, one obtains abc + cab + bac= 111 × (a + b + c) which is, first, of the required form and, second, divisible by 37, as 111 is. Finally, let us show briefly that the number 37, which plays a basic role in the work of Shcherbak, could be derived with the help of the above size and hydrogen bond matrices. As a matter of fact, this number is obtained in two ways and independently of the amino acid nucleon numbers. In the first, it is obtained as the ratio of two traces (invariants):

$$\frac{\mathrm{Tr}(\boldsymbol{\Theta}_c^{(ms)})}{\mathrm{Tr}(\boldsymbol{\Theta}^{(hb)})}=\frac{\mathrm{Tr}(\boldsymbol{\Theta}_c^{(ms)})}{\mathrm{Tr}(\boldsymbol{\Theta}^{(ms)})}=\frac{2220}{60}=37 \tag{23}$$

In the second, it is obtained as the ratio of homologous norms, one from the hydrogen bond–matrix Eq. (22), and the other from the molecular size–matrix Eq. (20):

$$\frac{\left\|\boldsymbol{\Theta}_c^{(ms)}\right\|_1}{\left\|\boldsymbol{\Theta}^{(ms)}\right\|_1}=\frac{2664}{72}=37,\quad \frac{\left\|\boldsymbol{\Theta}_c^{(ms)}\right\|_\infty}{\left\|\boldsymbol{\Theta}^{(ms)}\right\|_\infty}=\frac{3108}{84}=37, \tag{24}$$

Note that one could also obtain 37 as the ratio of the total number of nucleons in the *amino acids* in M_2 (2220, see above or compute it from Table 2) and the trace of $\Theta^{(ms)}$ or $\Theta^{(hb)}$ of *base–made–codons* matrices so that, while the (equational) symmetry is lost, the link between the nucleic acids and the amino acids is made. Note that fifteen amino acids are represented in M_2, which is 75% of the total of 20. The present author would be almost prone to re–name the title of this last section as "autopoietic numbers in the 16 doublets and the 64 codons". We stop here but we shall return to these questions in a following publication [24].

Note added. One may wonder, as suggested by the reviewers, if there could be any applications of the above numerical relations (in section 4). This is a delicate but nevertheless an interesting question. Independently of (possible) applications, one tempting occurrence would be to look at these numbers *as* some (new) kind of *numeric genetic information* appearing "written" on the physical basic units of life (RNA–DNA, amino acids) and revealed when these are arranged into definite sets with patterns exhibiting symmetry, see reference [26], for a recent work. Now, since the submission of this paper and, perhaps, as a beginning answer in the direction of applications, we have derived, for example, *explicitly*, the number 37 which has a tantamount importance in Shcherbak's work, by considering a *real* system (RNA–and–DNA components), in place of just deriving it from abstract matrices and norms, as we did in the present work. In fact, we did more than this: we have established a connection between Shcherbak's work, with its ubiquitous (and omnipresent) number 37, and the recent 28–gon polyhedral theory of the genetic code by Yang, [24].

Acknowledgment

The author acknowledges Professor Sergei Pethoukhov (Moscow University), Professor Vladimir Karasev (Saint Petersbourg University) and Professor Chi Ming Yang (Nankai University) for internet exchanges and positive feedback. Professor Vladimir Shcherbak (al–Faraby Kazakh National University, Kazakhstan) is warmly acknowledged for sending me a copy of his very recent and important article in BioSystems. The anonymous referees are acknowledged for rising an important question.

5 REFERENCES

[1] V. I. Shcherbak, Arithmetic inside the universal genetic code, *BioSystems* **2003**, *70*, 187–209.
[2] T. Négadi, Cracking the genetic code(s) with a Modular Determinative degree: An algebraic approach, *Int. J. Quantum Chem.* **2003**, *91*, 651–662.
[3] T. Négadi, Rumer's transformation, in biology, as the negation, in classical logic, *Int. J. Quantum Chem.* **2003**, *94*, 65–74.
[4] T. Négadi, On the symmetries of the 16 genetic code–doublets; in: *Proceedings of the XXIV International Colloquium on Group Theoretical Methods in Physics* (Paris, July 15–20, 2002), IOP Conference Series, GROUP 24 Physical and Mathematical Aspects of Symmetries, Eds. J.P. Gazeau, J.P. Antoine, S. Métens, J.Y. Thibon, 2003, pp 687–690.
[5] H.G.Wittmann, Studies on the Genetic Code in Tobacco Mosaic Virus; in: Informational macromolecules, A symposium Eds.H.J.Vogel,V.Bryson and J.O.Lampen *Academic Press N.Y.* and London, 1963, p.177.
[6] T. Négadi, Unpublished work on the symmetries in Mendelian Genetics (in matrix form) where, also, we introduce the group C_2, as invariance group for the 64 codons set (2002).

[7] Yu. B. Rumer, About systematization of the codons of the genetic code, *Dokl. Acad. Nauk. SSSR* **1966**, *167*, 1393–1394.
[8] B.G. Konopel'chenko and Yu. B. Rumer, The codon classification inside the genetic code, *Dokl. Acad. Nauk. SSSR* **1975**, *223*, 471–474.
[9] V. I. ShCherbak, Rumer's rule and transformation in the context of co–operative symmetry of the genetic code, *J. Theor. Biol.* **1989**, *139*, 271–276.
[10] V. A. Karasev and S. G. Sorokin, Topological structure of the genetic code, *Russian J. Genetics* **1997**, *33*, 622–628.
[11] H. –J. Danckwerts and D. J. Neubert, Symmetries of the genetic code doublets, *J. Mol. Evol.* **1975**, *5*, 327–332.
[12] M. A. Jiménez Montaño, C. R. de la Mora Basáñez and T. Poshel , The hypercube structure of the genetic code explains conservative and non–conservative amino acid substitutions *in vivo* and *in vitro*, *BioSystems* **1996**, *39*, 379–384.
[13]] C. M. Yang, The bi–pyramidal nature, the Lucas series in the genetic code and their relation to aminoacyl–tRNA Synthetases; *arxiv.org*/ftp/q– bio/papaers/0309008, july **2003**; *arxiv.org*/ftp/q– bio/papaers/0309008, july **2003**; The naturally designed spherical symmetry in the genetic code, *arxiv.org*/ftp/q– bio/papaers/0309014, june **2003**.
[14] F. J. R. Taylor and D. Coates, The code within the codons, *BioSystems* **1989**, *22*, 177–187.
[15]] R. D. Knight, S. J. Freeland and L. F. Landweber, Selection, history and chemistry: the three faces of the genetic code, *TIBS*, **1999**, *24*, 241–246.
[16] R. D. Knight, S. J. Freeland and L. F. Landweber, Rewriting the keyboard: evolvability of the genetic code, *Nature Reviews*, *Genetics*, **2001**, *2*, 49–57.
[17] W. F. Loomis, Origin of life, *Encyclopedia of Microbiology*, **2000**, *3*, 490–501.
[18]] N. Štambuk, On circular coding properties of gene and protein sequences, *Croat. Chem. Acta*, **1999**, *72*(*4*), 999–1008.
[19] G. Rosen, Rook's tour representation of the genetic code, *Bull, Math. Biol.* **1991**, *53*, 845–851.
[20] H. P. Yockey, Origin of life on earth and Shannon's theory of communication, *Computers and chemistry*, **2000**, *24*, 105–123.
[21] D. Fell and A. Wagner, The small world of metabolism, *Nature Biotechnology*, **2000**, *18*, 1221–1222.
[22] S. R. Pelc, Correlation between codong–triplets and amino–acids, *Nature*, **1965**, *207*, 597–599.
[23] V. A. Karasev and V. E. Stefanov, *J. Theor. Biol.*, **2001**, *209*, 303–317.
[24] T. Négadi, A connection between Shcherbak's arithmetical and Yang's 28–gon polyhedral "views" of the genetic code, work presented at the Internet Electronic Conference of Molecular Design, **2003**, *IECMD*, Nov. 23– Dec. 6. http://www.biochempress.com/iecmd_2003_04.html
[25] S.V. Petoukhov, Genetic code and the Ancien Chinese Book of Change, *Symmetry: Culture and Science*, **2001**, Vol.*10*, Nos 3–4, 211–226.
[26] T. Négadi, Symmetry and proportion in the genetic code, and genetic information from the basic units of life, invited paper submited to *Symmetry: Culture and Science* (Dec. **2003**).

Biographies

Tidjani Négadi is professor of Theoretical Physics at the Faculty of Science, Oran University, Oran, Algeria. (Some bio–data could be found at: http://www.geocities.com/tidjani_negadi/vitae.pdf)