

Article

Revealing the Genetic Code Symmetries through Computations Involving Fibonacci-like Sequences and Their Properties

Tidjani Négadi

Physics Department, Faculty of Exact and Applied Science, University Oran1 Ahmed Ben Bella, Oran 31100, Algeria; negadi.tidjani@univ-oran1.dz

Abstract: In this work, we present a new way of studying the mathematical structure of the genetic code. This study relies on the use of mathematical computations involving five Fibonacci-like sequences; a few of their “seeds” or “initial conditions” are chosen according to the chemical and physical data of the three amino acids serine, arginine and leucine, playing a prominent role in a recent symmetry classification scheme of the genetic code. It appears that these mathematical sequences, of the same kind as the famous Fibonacci series, apart from their usual recurrence relations, are highly intertwined by many useful linear relationships. Using these sequences and also various sums or linear combinations of them, we derive several physical and chemical quantities of interest, such as the number of total coding codons, 61, obeying various degeneracy patterns, the detailed number of H/CNOS atoms and the integer molecular mass (or nucleon number), in the side chains of the coded amino acids and also in various degeneracy patterns, in agreement with those described in the literature. We also discover, as a by-product, an accurate description of the very chemical structure of the four ribonucleotides uridine monophosphate (UMP), cytidine monophosphate (CMP), adenosine monophosphate (AMP) and guanosine monophosphate (GMP), the building blocks of RNA whose groupings, in three units, constitute the triplet codons. In summary, we find a full mathematical and chemical connection with the “ideal sextet’s classification scheme”, which we alluded to above, as well as with others—notably, the Findley–Findley–McGlynn and Rumer’s symmetrical classifications.

Keywords: genetic code symmetries; Fibonacci-like sequences; amino acids; ribonucleotides; patterns; hydrogen atoms; atoms; molecular mass



Citation: Négadi, T. Revealing the Genetic Code Symmetries through Computations Involving Fibonacci-like Sequences and Their Properties. *Computation* **2023**, *11*, 154. <https://doi.org/10.3390/computation11080154>

Academic Editor: Sergei Abramovich

Received: 8 June 2023

Revised: 29 July 2023

Accepted: 1 August 2023

Published: 7 August 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A novel approach to studying the genetic code’s mathematical and chemical structure is presented in this paper. More precisely, using a small set of Fibonacci-like sequences and, occasionally, some (useful) well-known elementary functions from number theory, the whole and detailed chemical content of the set of amino acids, as structured by several well-known symmetry patterns, including their degeneracy, is revealed. Also, several other original applications, using the above sequences, are carried out.

This paper, in addition to presenting new research results, also has an educational dimension, that of introducing the interested reader to an aspect of the mathematical study of the genetic code. It could therefore also be read (the computations easily worked out) by non-experts with mathematical backgrounds.

1.1. The Genetic Code

The genetic code is the basis of life on Earth and was masterfully deciphered in the 1960s [1]. It is the great biological “dictionary” that translates the language of DNA/RNA, which transmits the inherited information located in the genes, to the language of proteins that carry out the biological constructions and functions. It is well known that the “alphabet” of the former language consists of four fundamental units, the nitrogenous bases T (thymine), C (cytosine), A (adenine) and G (guanine) for DNA and U (uracil), C, A and

G for RNA. As for the “alphabet” of the second language, it comprises a set of 20 amino acids. In the process of translation between these two languages, in the ribosome for short, there are $64 = 4^3$ “words”, the codons. Each group of three bases in mRNA constitutes a codon, and each (sense) codon specifies a particular amino acid. Multiple codons can encode the same amino acid; they are known as “synonymous” codons. This phenomenon is also called degeneracy. In the standard genetic code, 61 sense codons are translated into 20 amino acids, which are organized into five “multiplets”, and three other (nonsense) codons serve as termination or stop signals. These “multiplets” are the following:

- Three sextets: each coded by six codons serine (Ser), arginine (Arg) and leucine (Leu);
- Five quartets: each coded by four codons proline (Pro), alanine (Ala), threonine (Thr), valine (Val) and glycine Gly);
- One triplet: coded by three codons isoleucine (Ile);
- Nine doublets: each coded by two codons phenylalanine (Phe), tyrosine (Tyr), cysteine (Cys), histidine (His), glutamine (Gln), glutamic acid (Glu), aspartic acid (Asp), asparagine (Asn) and lysine (Lys);
- Two singlets: each coded by one codon methionine (Met) and tryptophane (Trp).

Table 1 shows the relationship between the amino acids, represented in their three-letter code (see above), and the codons that encode them. For example, the codon UUU codes for the amino acid phenylalanine (UUU-Phe). The three stop codons are indicated in black.

Table 1. The genetic code table.

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

In this work, the “anomalous” three amino acids serine, arginine and leucine, each coded by six codons, will play a prominent role. Contrary to the 17 other amino acids, the codons of which share the same first base, the three mentioned amino acids have, each, their six codons distributed over two separate *family boxes*. There are 16 such family boxes in the genetic code table, and each one of them is a set of four codons sharing the same *first and second* base (see Table 1). The structure of the three sextets is the following serine: {UCN, AGY}, arginine {CGN, AGR}, leucine {CUN, UUR} (N for any base, Y for pyrimidine U or C and R for purine A or G).

There are more and more voices rising to underline or put emphasis on the singular nature of the three sextets and also bring experimental data which tend to show it [2,3].

A few years ago, a published work [4] claimed that the number of “codon families” has to be increased to 23 by considering the quartet part and the doublet part of each one of the three sextets as *distinct*. A “codon family”, a term used by the authors of the above reference, not to be confused with the “family box” alluded above, is a group of synonymous codons. In the case of the standard genetic code, each member in the five multiplets mentioned above, taken individually, constitutes such a “codon family” because its codons are synonymous and encode the same amino acid. For example, the triplet of codons AUU, AUC and AUA, in Table 1, encode isoleucine. Also, in the special case of the five quartets and the three quartet parts of the three sextets, the “codon family” and “family box” represent the same thing. This identification is no longer valid in the other cases where each one of the eight remaining “family boxes” contains groups of non-synonymous codons. For example, in the “family box” AAN, the two synonymous codons AAU and AAC encode asparagine, and the two codons synonymous AAA and AAG encode lysine.

In their work, the above authors present a new “effective number of codon families”, called N_c , to characterize codon usage bias in the analysis of protein-coding genes, which improves existing ones. (An “effective number of codons” is a widely used index in bioinformatics, see the above mentioned reference [4].) Specifically, they show that N_c is a better predictor when its value is increased from 20 to 23; in particular, each sixfold codon set (each sextet, as it is called in this work) is considered to be composed of *separate* fourfold and twofold parts. These six entities are $\text{Ser}^{\text{II,IV}}$, $\text{Arg}^{\text{II,IV}}$ and $\text{Leu}^{\text{II,IV}}$ which, added to the 17 remaining amino acids with no “degeneracy” at the *first base position*, as mentioned above, give a total of 23. This number (of codons), together with the remaining degenerate codons, 38, constitutes what we call the pattern “23 + 38” (see Section 4.1 and elsewhere in the paper). Of the kind of approaches mentioned above (i.e., Refs. [2–4]), there is one that is particularly relevant to the present work: the “Ideal” symmetry classification scheme, introduced a few years ago. It will be summarized in Section 2.3, and we present its numerous connections with the present work in Section 4.2.2.

1.2. Previous Works

At this point, before continuing this introduction, let us linger a bit to emphasize the novelty of this work compared to all that has been conducted by us and published so far. In all our previous works on the genetic code, the results obtained, concerning either its degeneracy structure or the derivation of the chemical content of the coded amino acids, were scattered over several publications, and, in these, the mathematical methods used were, in each case, different. Let us mention here only a few of them.

In [5], we considered, as a starting point, the unique number $23!$, the order of the permutation group of 23 objects, in its two representations, the decimal representation and the prime factorization representation, to derive the multiplet structure of the 64 codons.

In [6], we started by considering an empirical inventory of the degeneracies in the 64 codons table, put as a sequence of numbers, and then applied a Gödel’s encoding procedure to this sequence to derive, as an output, the number $23!$, which we started with in the previous reference.

In [7], moving away from the previous methods, we considered the number of atoms in the four ribonucleotides UMP, CMP, AMP and GMP, 144, the twelfth Fibonacci number, as a *unique* starting point determinant and also Euler’s phi function to find, again, the previous results mentioned above

1.3. The Novelty in This Work

The present work, on the other hand, is entirely new in its methods and unified in its structure. It is based on a completely different and new mathematical formalism, namely, that using Fibonacci-like sequences, with carefully chosen initial conditions, i.e., their first two terms (called “seeds” in this paper), some of them chemically “dressed”, that is, having values from the chemical data of three special amino acids having great importance. Using these sequences and their mathematical properties allows us, as we will see in the sequel, to find, again, a few results of the previous works, such as the number of amino acids, the degeneracy and the chemical composition according to degeneracy. However, the overwhelming majority of the results presented in this paper, which is concerned with the symmetries of the genetic code, are new and reported here for the first time in Sections 4–7, all from the unified and integrated mathematical formalism described in Section 3.

In Section 2, we summarize three important symmetries of the genetic code, Rumer’s symmetry, [8], the Findley-Findley-McGlynn third base symmetry, [9], and the Rosandić-Paar “ideal” symmetry, [10,11]. In Section 3, we present our new Fibonacci-like sequences and their properties, which are the main mathematical tools used in this paper. In Section 4, we apply these sequences to derive the degeneracy structure of the 61 sense codons (in Section 4.1), as well as the hydrogen atom content (in Sections 4.2.1–4.2.4), the atom content (in Section 4.3) and the nucleon number content (in Section 4.4), in the side chains of all the encoded amino acids, as structured by the symmetries described in Section 2, as

well as various other remarkable patterns. We have also included, at the end of Section 4 (in Section 4.2.5), a discussion concerning the choice, and its justification, of the initial conditions of our Fibonacci-like sequences, defined in Section 3. In Section 5, still using some elements of our sequences, we make contact with the work by shCherbak, [12], concerning the singular structure of proline and derive a mathematical form of the shCherbak–Makukov “activation” key, [13], which, as is well known, led to many remarkable and beautiful nucleon number patterns comprising, in particular, those related to Rumer’s symmetry. In Section 6, using the “seeds” of our Fibonacci-like sequences, that is, their initial conditions, and *only these*, we find that they are capable, on their own, to provide the very hydrogen atom content of the amino acids, derived in the various patterns considered in Section 4. Finally, in Section 7, we present some (new) results concerning the *vertebrate mitochondrial genetic code*, a case that arose while finishing this paper. We strongly recommend that the reader, at this point, before going to the next sections and getting a comfortable reading of them, take a careful look at Appendix A, which gives the chemical data of all 20 amino acids, in Table A1, and also includes some hints for the evaluation of several quantities when the degeneracy is involved. (Several of these quantities, evaluated from the table, are to be compared with their equivalents, derived mathematically in this paper, from our Fibonacci-like sequences and their properties.) In Appendix B, a few other mathematical tools used in this paper are defined with the presentation of some computation examples. We have also included a third Appendix C, where we explain how the use of mathematical software, containing a built-in “Fibonacci” function, could help the reader to carry out the various computations presented in this paper. We also give several examples.

2. The Symmetries of the Genetic Code

2.1. Rumer’s Symmetry

The oldest known symmetry of the genetic code was discovered by Rumer in 1966, see [8]. This symmetry, which is defined by the transformation $U \leftrightarrow G, A \leftrightarrow C$, divides the genetic code 8×8 table into two equal halves of 32 codons each; we call them M_1 and M_2 . In Table 2 below, which is a duplicate of Table 1, we show, in addition, such a division. The set M_1 , shown in a grey background, comprises eight quartets of codons, each having the same two first bases and coding for the same amino acid, the third base being irrelevant. In this set, among the eight quartets, three correspond to the quartet part of the three sextets serine, arginine and leucine. The set M_2 comprises group-I amino acids (two singlets), group-II amino acids (nine doublets), group-III amino acid (one triplet) and also three stops or termination codons. The point here, concerning symmetry, is that under Rumer’s transformation, performed on all three bases, the sets M_1 and M_2 are exchanged: $M_1 \leftrightarrow M_2$.

Table 2. Rumer’s division of the genetic code table.

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

2.2. The Third Base Symmetry Classification

In 1982, Findley et al. (see [9]), by viewing the genetic code as an f -mapping, extracted a fundamental symmetry for the doubly degenerate codons (group-II). Below, to ease the reading, we reproduce a few elements from the above reference to help the reader understand what the f -mapping is. The authors consider the 64-codons set, C , and define $C_k = \{C_{ijk} \in C \mid i, j \in B\}$, $k \in B$, where i, j and k designate the first, second and third base in the codon C_{ijk} (B is for base, U, C, A, G). $C_k, k \in B$, partitions C into four disjoint

subsets, where each subset contains only codons having the same third base. Each of these subsets may be mapped by f into members of the amino acids set A , with the image being denoted $f(C_k)$; this is shown in Table 3 below. One has, therefore, $f(C_U) = f(C_C)$ and $f(C_A) \neq f(C_G)$. With this f -mapping, the authors also establish relations that define a one-to-one correspondence between one member of a doubly degenerate codon pair and the other member (see the reference above for details). These relations could be stated, in words, as follows: (i) if a codon for an amino acid has the third base U, then there is a codon for the same amino acid having the third base C and vice versa, OR (ii) if a codon for an amino acid has the third base A, then there is a codon for the same amino acid having the third base G and vice versa. For a doubly degenerate codon pair, (i) and (ii) are mutually exclusive. For order four, or quartets, (i) and (ii) hold simultaneously. For order six, the sextets, the quartet part obeys (i) AND (ii), and for the doublet part, one has (i) OR (ii). For the odd-order degenerate codons (Ile, Met and Trp), however, there is a slight deviation from symmetry. In Table 3, we show this classification.

Table 3. The third base classification of the 64 codons [9].

UCU		UCC		UCA		UCG	
AGU	Ser (6)	AGC	Ser (6)	AGA	Ser (3)	AGG	Ser (3)
CGU	Arg (10)	CGC	Arg (10)	CGA	Arg (20)	CGG	Arg (20)
CUU	Leu (9)	CUC	Leu (9)	CUA		CUG	
GCU	Ala (4)	GCC	Ala (4)	UUA	Leu (18)	UUG	Leu (18)
GUU	Val (7)	GUC	Val (7)	GCA	Ala (3)	GCC	Ala (3)
CCU	Pro (5)	CCC	Pro (5)	GUA	Val (7)	GUG	Val (7)
GGU	Gly (1)	GGC	Gly (1)	CCA	Pro (5)	CCG	Pro (5)
ACU	Thr (5)	ACC	Thr (5)	GGA	Gly (1)	GGG	Gly (1)
UUU	Phe (7)	UUC	Phe (7)	ACA	Thr (5)	ACG	Thr (5)
UAU	Tyr (7)	UAC	Tyr (7)	CAA	Gln (6)	CAG	Gln (6)
UGU	Cys (3)	UGC	Cys (3)	AAA	Lys (10)	AAG	Lys (10)
CAU	His (5)	CAC	His (5)	GAA	Glu (5)	GAG	Glu (5)
GAU	Asp (3)	GAC	Asp (3)	UAA		UAG	Stop
AAU	Asn (4)	AAC	Asn (4)	UGA	Stop	UGG	Trp (8)
AUU	Ile (9)	AUC	Ile (9)	AUA	Ile (9)	AUG	Met (7)
Hydrogen	84		84		92		98
Nucleons		1728				1676	

2.3. The Weak/Strong, Purine/Pyrimidine and Keto/Amino Symmetries

The main idea behind the “ideal” symmetry classification scheme by Rosandić and Paar mentioned earlier ([10]; see also [11]) is to consider the three sextets serine, arginine and leucine, each encoded by six codons, as “initial generators”, with serine playing the central role. This scheme divides the 64 codons table into two groups of 32 codons each, the “leading” group and the “nonleading” group, and each one of them consists of A+U rich and G+C rich (equal) parts. The “ideal” classification scheme is generated by combining the six codons of serine, arginine and leucine in the following manner: serine, the “initial” generator with its six codons, arginine, also with its six codons, and leucine, with *only* the quartet part of its six codons part, define the “leading” group (with 32 codons). The remaining doublet part of leucine, on the other hand, constitutes a “seed” for the construction of the “nonleading” group (with 32 codons). The whole set $\{ \text{Ser}^{\text{IV-II}}, \text{Arg}^{\text{IV-II}}, \text{Leu}^{\text{IV-II}} \}$ is called, by the above authors, the “core”; its members are underlined in Table 4 below.

Table 4. The “ideal” symmetry classification scheme [10].

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

In the above table, which is also a duplicate of Table 1, the “leading” group is indicated in a light green background. As explained, at length, by the authors in [10], the genetic code table in this new scheme is created by codons sextets based on exact *purine/pyrimidine* symmetries (YR: (U, C, A, G) → (C, U, G, A)), A+U-rich/C+G-rich symmetries, *strong/weak*, or complementary, symmetries (SW: (U, C, A, G) → (A, G, U, C)) and *keto/amino* symmetries (KM: (U, C, A, G) → (G, A, C, U)). By starting with serine, the *initial generator* with its six codons, the whole “leading” group (32 codons) is created using transformations among those mentioned above and some mapping rules. Analogously, starting from the two codons of leucine (Leu^{II}) as “seeds”, the whole “nonleading” group is constructed. There is also a simple relation between the “leading” group and the “nonleading” group. We show, in Table 4, for visualization, these two groups by using our own format of the genetic code table. We also find it noteworthy to mention that, under Rumer’s transformation $U \leftrightarrow G, A \leftrightarrow C$, the “leading” group remains *globally invariant* whether the transformation is applied to the first base only, to the first two bases only or to all three bases, and the same is true for the “nonleading” group.

Below, in Section 4.2, we will show that the three amino acids serine, arginine and leucine will also play a prominent role as mathematical (and chemically inspired) “seeds” in computing the chemical content of the twenty amino acids, including degeneracy.

3. A Rich Set of Fibonacci-like Sequences and Their Properties

Let us introduce, as stated in the introduction, four Fibonacci-like sequences that will prove resource-rich and prolific in their applications throughout this work. (Another fifth sequence, just as interesting, will be introduced later, in Equation (26).) They are also called (p, q)-Fibonacci sequences and are well known in mathematics. What characterizes them, in this paper, is the specific choice of the initial conditions (see below). They are defined by the following common defining relation:

$$pF_{n-1} + qF_{n-2}, \tag{1}$$

where F_n is an ordinary Fibonacci number. These four sequences differ only by the data of the numbers p and q, which play the role of initial conditions or “seeds”, as we will call them throughout this paper. Below, we shall explain and *justify* the choice of these “seeds”, but for the moment, we introduce the four sequences by giving a name to each one of them while assigning their “seeds”: (i) $a_n : p = 1, q = 6$, (ii) $a'_n : p = 6, q = 1$, (iii) $b_n : p = 9, q = 13$, (iv) $c_n : p = 5, q = 30$. In Table 5 below, we give the first few terms.

Table 5. The first few terms of the Fibonacci-like sequences a_n, a'_n, b_n and c_n .

	n	1	2	3	4	5	6	7	8	9	10	11	12	13
$p = 1,$ $q = 6$	a_n	6	1	7	8	15	23	38	61	99	160	259	419	678
$p = 6,$ $q = 1$	a'_n	1	6	7	13	20	33	53	86	139	225	364	589	953
$p = 9,$ $q = 13$	b_n	13	9	22	31	53	84	137	221	358	579	937	1516	2453
$p = 5,$ $q = 30$	c_n	30	5	35	40	75	115	190	305	495	800	1295	2095	3390

These sequences obey several linear relations (or identities), some of which will prove very useful in view of their applications in this work. They are presented below, in Equation (2), and could be checked (see Appendix C, where concrete examples are also presented)

$$\begin{aligned}
 & \text{(i) } a_n + b_{n+1} = a_{n+4}, \\
 & \text{(ii) } a_n + a_{n+6} = 2b_{n+2}, \\
 & \text{(iii) } b_n + b_{n+2} = c_{n+2}, \\
 & \text{(iv) } b_n + c_{n+1} = 2b_{n+1}, \\
 & \text{(v) } c_n + 2b_{n-1} = b_{n+2}, \\
 & \text{(vi) } b_n + c_{n+3} = b_{n+4}, \\
 & \text{(vii) } a_n + c_{n+3} = 2a_{n+5}, \\
 & \text{(viii) } a_n + a_{n+2} = b_n, \\
 & \text{(ix) } c_n + b_{n-1} = 2b_n, \\
 & \text{(x) } a_n + b_{n+2} = 4a_{n+2}.
 \end{aligned} \tag{2}$$

It is worth noting here that the difference

$$a_n - a'_{n-1}, \tag{3}$$

gives the (slightly modified) Fibonacci sequence noted as F'_n

$$F'_n : 1, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots, \tag{4}$$

in an unusual but interesting form: its “seeds” here are inverted with respect to the usual Fibonacci sequence. Also, the sum of any of its first members until a certain index gives an exact Fibonacci number, contrary to the usual Fibonacci sequence with the seeds 0 and 1, which always gives one unit less than a Fibonacci number. For example, in our case, for $n = 9$, we obtain $\sum_1^9 F'_n = 34$. (Note that the indexing is shifted here, but the recurrence relation is still valid.) There is also another relation linking the sequences a'_n and b_n . It writes

$$a'_n - b_{n-2} = 2F'_{n-5}. \tag{5}$$

For $n = 7$, the sequences a' and b take the same value: $a'_7 - b_5 = 0$. Also, for $n = 8$, $a'_8 = 86$ and $b_6 = 84$, and their difference is 2. These relations will have applications in the following sections. Importantly, the sequences in Table 5 together with the one defined in Equations (26) and (27) below either display several numbers highly relevant in this work, directly as members in Table 5 (shown in a dark red color), or lead to significant sums to be evaluated in the following sections. We have also discovered that the above sequences, including the one defined in Equation (26), can all be shown to exhibit a *bilateral* symmetry and other symmetry properties, in the line of thought of those established for the ordinary Fibonacci sequence by Edge, see [14]. These findings will be reported elsewhere.

4. The Symmetries of the Genetic Code Revealed

4.1. The Multiplet Structure

Let us consider, in this section, the first sequence a_n . It is full of meaningful numbers and underlying sums. First, we have $a_4 = 8$, $a_5 = 15$ and their sum $a_4 + a_5 = 8 + 15 = 23$. These are, respectively, the number of amino acids in Rumer's sets M_1 and M_2 , regardless of degeneracy, and the sextets are counted *twice*, once in M_1 and once in M_2 . Second, we have $a_6 + a_7 = 23 + 38 = a_8 = 61$. This is the pattern, "23 + 38", for 23 amino acids (see above) and 38 degenerate codons. This latter pattern will be mentioned frequently in this paper. The above relationships will also let us derive the detailed multiplet structure of the genetic code. Consider the following sum, which will be used occasionally in this paper

$$\sum_1^k a_n = a_{k+2} - 1. \quad (6)$$

It is the analog of the one for the ordinary Fibonacci sequence and could be checked either with a pocket calculator directly from Table 5, for low values of k , or using the same computations as those performed for the examples in Appendix C. For $k = 5$, we have $6 + 1 + 7 + 8 + 15 = 37 = 38 - 1$. By grouping the first three terms on the one hand and the remaining two on the other, we have

$$(6 + 1 + 7) + (8 + 1 + 15) = 14 + 24 = 38. \quad (7)$$

The unit is transferred to the left. Using the sum mentioned above ($a_4 + a_5 = 8 + 15 = a_6 = 23$) and adding it to the preceding relation gives (by appropriately arranging the terms)

$$(15 + 14) + (8 + 24) = 29 + 32 = 61. \quad (8)$$

It appears that there are 15 amino acids and 14 degenerate codons in Rumer's set M_2 , while there are 8 amino acids and 24 degenerate codons in Rumer's set M_1 (see above).

Let us now go into the details by examining, first, the set M_2 . The number 15 could be partitioned in two ways. The first consists in using the above sum for $k = 3$ to obtain $6 + (1 + 7 + 1) = 6 + 9 = 15$. Using the second way, we can apply the useful A_0 function and its properties (see below and Appendix B) to the number 15 ($= 3 \times 5$): $A_0(15) = A_0(3) + A_0(5) = 6 + 9 = 15$, which gives the same result as above, where we have used the *additivity* property. Finally, the number 6, a *perfect* number, could be written as the sum of its proper divisors $\{1, 2, 3\}$ so that $15 = 1 + 2 + 3 + 9$. We interpret this relation as one triplet, two singlets, three doublet parts of the three sextets and nine doublets. On the other hand, for the degeneracy part, 14, which writes $6 + 1 + 7$ (see above), we can, again, write 6 as the sum of its divisors, arrange the terms and obtain $14 = 3 + (1 + 1) + (2 + 7) = 3 + 2 + 9$. Here, we have three degenerate codons for the three doublet parts of the three sextets, two degenerate codons for the triplet and nine degenerate codons for the nine doublets. For the set M_1 , things are simpler. The degeneracy part from Equation (8) above writes $24 = (8 + 1) + 15 = 9 + 15$. As for the number of amino acids, eight, as a Fibonacci number, it could simply be written as $5 + 3$. This is the structure of the set M_1 . Table 6, below, summarizes all of these results for the two Rumer's sets, which are thus completely described using the Fibonacci-like sequence a_n .

4.2. Hydrogen Atom Content and the Symmetries

In this section, we examine the hydrogen atom content in each one of the symmetry cases summarized in Section 2: Rumer's symmetry (Section 2.1), the third base symmetry (Section 2.2) and the weak/strong, purine/pyrimidine and keto/amino symmetries or "ideal" symmetry (Section 2.3). Before developing these topics, let us consider, first, the hydrogen atom content in the side chains of all the amino acids coded by the 61 sense codons.

Table 6. The derived multiplet structure of the amino acids in Rumer’s division.

	multiplets	# amino acids	# degenerate codons	total
M ₁	quartets	5	15	20
	quartet parts of the sextets	3	9	12
	total	8	24	32
	multiplets	# amino acids	# degenerate codons	total
M ₂	doublets	9	9	18
	doublet parts of the sextets	3	3	6
	triplet	1	2	3
	singlets	2	0	2
	total	15	14	29

4.2.1. The Hydrogen Atom Content

From Table A1 in Appendix A, the total number of hydrogen atoms in the side chains of all the amino acids coded by the 61 sense codons is equal to 358. Let us note from the start that, in this count, we take for the (singular) *imino acid proline*, as a special case, five hydrogen atoms in its side chain. We will return to this important point later, in Section 5, with brand new results. A quick look at Table 5 of our Fibonacci-like sequences reveals that the number of hydrogen atoms, mentioned above, is showing itself in multiple instances: first, ostensibly, as the ninth member of the sequence b_n ($b_9 = 358$); second, from the relation (viii) in Equation (2) which, we recall, is valid, particularly for $n = 9$: $a_9 + a_{11} = 99 + 259 = 358 = b_9$; third, from the recurrence relation of the sequence $b_n : b_7 + b_8 = 137 + 221 = b_9 = 358$; fourth, from the sum

$$\sum_1^9 a'_n = 358. \tag{9}$$

This last equation will be considered in detail below, as it has great importance concerning the computation of the degeneracy of the genetic code in various formats. By isolating the last term a'_9 , we have

$$\sum_1^8 a'_n + a'_9 = 219 + 139 = 358. \tag{10}$$

This relation is important and will play a prominent role in this section and later (in Section 6). Equation (10) gives the number of hydrogen atoms in the amino acids’ side chains, distributed into two parts: 139 hydrogen atoms in 23 amino acids (17 amino acids with no “degeneracy” at the *first base* position and the six entities $\text{Ser}^{\text{IV-II}}$, $\text{Arg}^{\text{IV-II}}$ and $\text{Leu}^{\text{IV-II}}$), on the one hand, and 219 hydrogen atoms in the remaining side chains of the amino acids encoded by the 38 degenerate codons, on the other (see Appendix A for the calculations from the table). This is the equivalent “23 + 38” pattern for the hydrogen content. Next, as we have $139 = 53 + 86 = 22 + 31 + 86$ from the recurrence relation of the sequence b_n , we can cast the relation above as follows:

$$(219 + 22) + (31 + 86) = 241 + 117 = 358. \tag{11}$$

This is the hydrogen atom content in the usual pattern “20 + 41” (117 hydrogen atoms in the side chains of 20 amino acids and 241 hydrogen atoms in the side chains of the amino acids coded by the 41 degenerate codons; see Table A1 in Appendix A). Note that 22 is the number of hydrogen atoms in the side chains of serine, arginine and leucine, corresponding to one codon for each one of them (see Table A1 in Appendix A). It is also just the right factor that connects the two patterns “23 + 38” and “20 + 41”.

By restricting the sum in Equation (10), as shown below, we have

$$\sum_1^7 a'_n + (a'_8 + a'_9) = 133 + 225 = 358. \tag{12}$$

This hydrogen atom partition corresponds to Rakočević’s *Cyclic Invariant Periodic System* (CIPS) classification of the amino acids, where there are 133 (225) hydrogen atoms in the amino acids side chains in the secondary superclass (primary superclass), [15]. The above hydrogen atom partition is only one unit from another one, which is twice relevant. By transferring the first member of the sequence, $a'_1 = 1$, from the sum to the other factor, we obtain

$$\sum_2^7 a'_n + (a'_1 + a'_8 + a'_9) = 132 + 226 = 358. \tag{13}$$

First, this hydrogen atom pattern corresponds to 132 hydrogen atoms in all the side chains of the 3 sextets coded by 18 codons, on the one hand, and 226 hydrogen atoms in all the side chains of the remaining 17 amino acids coded by 43 codons, on the other (see below). Here, we see that the three sextets are set apart, and this has, we think, a link with the subject of Section 4.2.2 below. Second, this pattern also describes the distribution of hydrogen atoms in the side chains of the amino acids in the two classes of the aminoacyl t-RNA synthetases: 226 hydrogen atoms in the side chains of all the amino acids coded by 29 codons in Class-I and 132 hydrogen atoms in the side chains of all the amino acids coded by 32 codons in Class-II; see [7]. Note the codon pattern “29 + 32”, the same as in Equation (8) above.

4.2.2. The Hydrogen Atom Content in the “Ideal” Symmetry Classification Scheme

In this section, we consider the hydrogen atom content for the “ideal” symmetry classification scheme, [10], which occupies an important place in this work, as it has a tight relation with the choice of the “seeds” of our Fibonacci-like series. As promised at the beginning of Section 3, this is the right place to explain and justify the choice of the initial conditions of the sequences b_n and c_n , as defined in Section 3, having importance in this section (more will be said about the “seeds” of the other sequences in Section 4.2.5, which is devoted to their choice). Concerning b_n , the “seeds” are 13 and 9 (see Table 5). These are chosen, respectively, to be the number of *hydrogen atoms* in arginine’s and serine’s side chains (10 + 3) and in leucine’s side chain (9). Their sum, which is the recurrence relation, $b_1 + b_2 = 13 + 9 = b_3 = 22$, is the number of hydrogen atoms in the side chains of these three amino acids (see Equation (12)). The “seeds” of c_n , 30 and 5, are chosen to be, respectively, the number of *atoms* in the side chains of arginine and leucine (17 + 13) and in the side chain of serine (5). Here, as for hydrogen, we have the recurrence relation $c_1 + c_2 = 17 + 13 = c_3 = 30$, which is the number of atoms in the side chains of these three amino acids (see Table A1 in Appendix A).

We show, in this section and also in the next ones, using all the resources offered by our Fibonacci-like series and their properties, that these three sextets (more precisely, their hydrogen and atoms numbers), as “seeds”, will create the entire hydrogen atom, atom and even nucleon content of the whole set of amino acids, including the degeneracy, much like the creation of the 64 codons from the three sextets in the “ideal” symmetry scheme, [10], mentioned above.

Now, we return to the subject of this section. First, using the relation (v) $c_n + 2b_{n-1} = b_{n+2}$ in Equation (2), we can derive the hydrogen atom content in the two sets: the “leading” group and the “nonleading” group. We have, for $n = 7$ (see Table 5 and also Appendix C)

$$190 + 2 \times 84 = 358. \tag{14}$$

It can be seen, from Table 4 and also, in parallel, from an evaluation using the data in Table A1 in Appendix A, that there are 190 and 168 hydrogen atoms in the side chains of the amino acids in the “leading” group and in the “nonleading” group, respectively. Moreover, concerning the latter, there are 84 hydrogen atoms in the side chains of the amino acids, the codons of which have the *same* first two bases, UU, CC, AA and GG (in the four corners of Table 4), and 84 hydrogen atoms in the side chains of the amino acids located in the four boxes in the center of the table, the codons of which have *different* first two bases, UG, GU, AC and CA. Equation (14) above faithfully describes, therefore, this pattern. Now, we

move further to accurately describe the hydrogen atom content involving the amino acids of the “core” comprising serine, arginine and leucine. To see this, we invoke the following two relations:

$$5a_n + 2b_{n-1} = b_{n+2}, \quad (15)$$

$$3a_n + 4a_{n+1} = b_{n+2}. \quad (16)$$

It could be verified that they give the same result and both hold (see Appendix C). They can also be transformed into each other, using the relation (viii) in Equation (2), $a_n + a_{n+2} = b_n$. For $n = 7$, they give $190 + 168$ and $114 + 244$, respectively, with a common value of 358, the total number of hydrogen atoms in the side chains of all the amino acids encoded by the 61 sense codons. These relations are of interest for what follows. In the first relation, as we have seen above, 190 is the number of hydrogen atoms in the side chains of the amino acids in the “leading” group, and 168 is the number of hydrogen atoms in the side chains of the amino acids in the “nonleading” group. In the second relation, 114 is the number of hydrogen atoms in the side chains of the amino acids in the part of the “core” belonging to the “leading” group ($\text{Ser}^{\text{IV}/\text{II}}$, $\text{Arg}^{\text{IV}/\text{II}}$, Leu^{IV}), and 244 is the number of hydrogen atoms in the side chains of all the remaining amino acids in the other part of Table 4, comprising, in particular, the part of the “core” belonging to the “nonleading” group, that is, Leu^{II} . The authors write in their paper [10], “The sextets as initial building blocks for the creation of their new scheme of the genetic code generate by themselves the patterns of A+U rich/C+G rich, purine/pyrimidine, weak-strong and amino-keto symmetries”. They also add that, in their approach, “the symmetries are a consequence of sextet’s dynamics”. To go further and show agreement with what has just been said, we can use our Fibonacci-like sequences to reveal the exact hydrogen atom content of the “core”, constituted by the three sextets. As mentioned above, the “core” has two parts: one that belongs to the “leading” group and the other that belongs to the “nonleading” group. Let us consider the former with 114 hydrogen atoms. Using Euler’s totient function φ and also the so-called “reduced” totient function or Carmichael’s function $\lambda(n)$ (see Appendix B), we have for the number 114 $\varphi(114) = 36$ and $\lambda(114) = 18$. Subtracting these from the number 114, we obtain $114 - 36 - 18 = 60$, and by rearranging, we obtain

$$114 = 60 + 36 + 18. \quad (17)$$

This is the correct content of the part of the “core” in the “leading” group: 60 hydrogen atoms (6×10) in arginine’s side chain ($\text{Arg}^{\text{IV}/\text{II}}$), 36 hydrogen atoms (4×9) in leucine’s side chain (Leu^{IV}) and 18 hydrogen atoms (6×3) in serine’s side chain ($\text{Ser}^{\text{IV}/\text{II}}$). Let us, alternatively, add the above-mentioned two functions to the number 114. We have

$$114 + 36 + 18 = (114 + 36) + 18 = 150 + 18 = 168. \quad (18)$$

This is the number of hydrogen atoms in the side chains of the amino acids of the “nonleading” group, where the isolated number 18 is now *re-interpreted* as the number of hydrogen atoms in the side chain of leucine (2×9), the “seed” of the “nonleading” group, that is, Leu^{II} (see above). We have thus established the exact hydrogen atom content in the “ideal” symmetry scheme of the genetic code where the sextets play a prominent role. Note, finally, that, as $\lambda(114) = 18$ has been used two times, once as the number of hydrogen atoms in $\text{Ser}^{\text{IV}/\text{II}}$ and once as the number of hydrogen atoms in Leu^{II} , we can summarize all of what has been said above by adding $\lambda(114) = 18$ to Equation (17) and write the exact hydrogen atom content of the entire “core” $60 + (36 + 18) + 18 = 132$ constituted by $\text{Arg}^{\text{IV}/\text{II}}$, $(\text{Leu}^{\text{IV}} + \text{Leu}^{\text{II}})$ and $\text{Ser}^{\text{IV}/\text{II}}$, respectively. (The 18 codons of the “core” are underlined in Table 4.) Of course, after subtracting the number 132 from the total sum 358 in Equation (14) above, we are left with 226, the number of hydrogen atoms in the side chains of the 17 amino acids outside the “core”. We have thus seen that the “seeds” of the

sequences b_n and c_n are capable of creating the hydrogen atom structure in good agreement with the “ideal” symmetry classification scheme (see also Section 4.2.4 below).

As a by-product of the results obtained in this section, we have found, unexpectedly, a way to derive from the number of hydrogen atoms in the part of the “core” in the “leading” group, 114, and in the rest, 244, comprising the part of the “core” in the “nonleading” group (see above), and only from these, the very chemical structure of the building blocks of RNA: the four ribonucleotides *uridine monophosphate* (UMP), *cytidine monophosphate* (CMP), *adenosine monophosphate* (AMP) and *guanosine monophosphate* (GMP). Using the functions A_0 and λ (see Appendix B), we have $A_0(114) = 38$, $A_0(244) = 88 = 61 + 1 + 18 + 4 + 4$ and $\lambda(114) = 18$ (see Appendix B, where the details of the computations are given as examples). First, we have, from these three quantities, $[A_0(114) + \lambda(114)] + A_0(244) = 56 + 88 = 144$. This is the total number of atoms in the four ribonucleotides: 56 in the four nucleotides U (12 atoms), C (13 atoms), A (15 atoms) and G (16 atoms) and 88 in the four identical “backbones”, each with 22 atoms (see [7] for the details of the calculation, which also includes a mathematical derivation of the number 22 above, which is part of the “condensation” equation for the assembly of a ribonucleotide from the three units: a nucleotide, a ribose and a phosphate group with the release of two water molecules, also derived). Now, as there are 30 codons in the “leading” group (two stop codons not counted) and 31 codons in the “nonleading” group (one stop codon also not counted) (see Table 4), we can use this decomposition for the number 61 above and finally write the relations above in the form $(30 + 4) + (31 + 4) + (2 \times 18 + 1) + 38 = 34 + 35 + 37 + 38$. Note that the above decomposition of the number 61 could also be obtained in another way, by directly using the properties of the sequence a_n ; see Table 5. We have, in this case, $a_8 = 61 = 23 + 38$, $a_7 = 38 = 23 + 15$ and $a_5 = 15 = 7 + 8$, so by combining them, we obtain $61 = (23 + 7) + (23 + 8) = 30 + 31$. The above-computed quantities 34, 35, 37 and 38 are, respectively, the number of atoms in the four ribonucleotides UMP ($C_9H_{13}N_2O_9P$), CMP ($C_9H_{14}N_3O_8P$), AMP ($C_{10}H_{14}N_5O_7P$) and GMP ($C_{10}H_{14}N_5O_8P$), where we have indicated their elemental composition.

4.2.3. The Hydrogen Atom Content in Rumer’s Symmetry

Now, we return to the symmetries and examine the second case, Rumer’s symmetry (Section 2.1). Let us reconsider Equation (10) and write it in the following form:

$$\sum_1^7 a'_n + a'_8 + a'_9 = (133 + 53) + 2 \times 86 = 186 + 2 \times 86 = 358, \tag{19}$$

where we have used the recurrence relation of the sequence a'_n to write the number 139 as $86 + 53$ (see Table 5). We have already mentioned in the examples following Equation (5) that, for $n = 8$, one has $86 - 84 = 2$ or $86 = 84 + 2$. Inserting this quantity in the above equation results in

$$186 + (84 + 88) = 358. \tag{20}$$

This is the hydrogen atom content in Rumer’s division: 186 hydrogen atoms in the side chains of the amino acids in M_2 and 172 hydrogen atoms in the side chains of the amino acids in M_1 , where, in this latter, we have the correct partition into 84 hydrogen atoms (4×21) in the side chains of the amino acids constituting the 5 quartets and 88 hydrogen atoms (4×22) in the side chains of the amino acids constituting the 3 sextets. To obtain the details concerning the number of hydrogen atoms in M_2 , 186, we first isolate the sum of the first four numbers in the sum in Equation (19), that is, $1 + 6 + 7 + 13 = 27 = 3^3 = 3 \times 9$. This is equal to the number of hydrogen atoms in the triplet isoleucine (see below). We are left, in the sum, with the three terms 3×53 . By writing the number 53 once as $15 + 38$ from the relation (viii) in Equation (2), with $n = 5$, and twice as $22 + 31$ from the recurrence relation of the sequence b_n , we obtain

$$2 \times 50 + 2 \times 22 + 27 + 7 + 8 = 186. \tag{21}$$

Here, $2 \times 50 = 2 \times 31 + 38 = 2 \times (31 + 19)$ and $15 = 7 + 8$ from the recurrence relation of the sequence a_n . We have, therefore, in detail, the correct number of hydrogen atoms in M_2 : $100 = 2 \times 50$ in the 9 doublets, $44 = 2 \times 22$ in the doublets of the 3 sextets, $27 = 3 \times 9$ in the triplet, 7 in the singlet methionine and 8 in the singlet tryptophane.

4.2.4. The Hydrogen Atom Content in the Third Base Symmetry

In Section 2.2, we explained that the authors extracted an inherent basic symmetry linked to the third base by partitioning the 64-codons set into four pair-wise subsets, where each one of them contains only codons having the same third base. In this way, a one-to-one correspondence exists between one member of a doubly degenerate codon pair and the other member. Here, also, for this symmetry, we could describe the hydrogen atom content, using our Fibonacci-like series. Take the relation (v) in Equation (2), the one we already considered above in Equation (14)

$$2 \times 84 + 190 = 358. \tag{22}$$

This relation, *as it is*, is the pattern shown in Table 3 for the gross third-base division UC/AG; more exactly, we have from the Table 3. $2 \times 84 + (92 + 98) = 2 \times 84 + 190 = 358$. Here, we note that this relation already describes, nicely, the equality of the number of hydrogen atoms in the columns third base U and third base C, where the amino acids are the same (see the penultimate row in the Table 3). We can do better by invoking two more relations. First, we have the relation (x) in Equation (2): $a_n + b_{n+2} = 4a_{n+2}$ which, for $n = 4$, gives $8 + 84 = 92$ (see Appendix C). Second, we have the relation $2b_n + b_{n+1} = c_{n+2}$, which also holds and gives, for $n = 5$, $2 \times 53 + 84 = 190$. Inserting the number $84 = 92 - 8$, from the relation just above, in the second one results in $190 = 92 + 98$. Collecting these results in Equation (22) above gives, finally,

$$2 \times 84 + 92 + 98 = 358. \tag{23}$$

This last relation completely describes, therefore, the hydrogen atom content pattern of Table 3. The third base classification mentioned above can also be supported by the following calculation. We know, from Section 2.2, that the doubly degenerate codons (group-II) obey a fundamental symmetry, so they must play a basic role, including, we will show, in the hydrogen atom content. We have, using the sequence a_n ,

$$\sum_1^9 a_n = 258. \tag{24}$$

By subtracting this sum from the right side of Equation (22) above, which gives the total number of hydrogen atoms in the side chains of all the amino acids coded by the 61 sense codons, we obtain, by arranging,

$$100 + 258 = 358. \tag{25}$$

These two numbers can be interpreted as follows: 100 hydrogen atoms in the side chains of the amino acids constituting the 9 doublets and 258 hydrogen atoms in the side chains of the amino acids constituting the remaining multiplets (5 quartets, 3 sextets, 2 singlets and 1 triplet); see Equation (21) and below it. This same relation, Equation (25), could also be obtained, in another way, from the relation mentioned in Section 4.1, $a_9 + a_{11} = 99 + 259 = b_9 = 358$, noting that the sum in Equation (24) above is also equal to $259 - 1$ (recall $\sum_1^k a_n = a_{k+2} - 1$, with $k = 9$). We then get back to our result as follows: $(99 + 1) + 258 = 100 + 258$. Note also that $2 \times \varphi(258) = 2 \times 84$ and $358 - 2 \times \varphi(258) = 190$ or $2 \times 84 + 190$, which is nothing but the hydrogen atoms pattern of the present classification (see Equation (22) and Table 3). (The function φ is defined in Appendix B, and the factor two, which has been introduced above, is for “doubly” degenerate codons.)

4.2.5. On the Choice of the “Seeds” of the Fibonacci-like Sequences

We have explained and justified, in Section 4.2.2, our choice of the “seeds” of the Fibonacci-like sequences b_n and c_n ; they are related, respectively, to the hydrogen and atom numbers of the three sextets serine, arginine and leucine, which play a prominent role in the “ideal” symmetry classification scheme. The choice of the “seeds” of the remaining sequences, a_n , a'_n and g_n , is of another nature. These “seeds” have been found (by a trial-and-error thought process) to be fruitful. These “seeds” may, perhaps, also have some deep connection with the nature of the codons; let us outline below how.

Consider, first, the sequence a_n . First, we have, using Equation (6), $6 + 1 + 7 + 8 + 1 = 23$, with the “seeds” being the first two numbers 6 and 1, and a unit was transferred from the right side of the equation to the left side. From the Fibonacci relation $F_{2n} = F_{n+1}^2 - F_{n-1}^2$, with $n = 3$, we have $8 = 9 - 1$ or $8 + 1 = 9$. Next, it could be easily shown that the sequence F'_n , in Equation (4), is related to the Lucas sequence, $L_n = F'_n + F'_{n+2}$, so that, for $n = 5$, we have $7 = 2 + 5$. Finally, we call, exceptionally, the term $a_0 = -5$, which also obeys the recurrence relation $a_0 + a_1 = a_2$, that is, $-5 + 6 = 1$, or, equivalently, $6 = 5 + 1$. Putting together all these pieces, we end up with $(5 + 1) + 1 + 2 + 5 + 9 = 23$. The last four terms on the left side could be interpreted as 1 triplet, 2 singlets, 5 quartets and 9 doublets, which are the 17 amino acids outside the “core” of the “ideal” symmetry classification scheme, discussed in Section 4.2.2. As for the first two terms, in the parenthesis, they are just enough to describe the five entities $Ser^{IV/II}$, $Arg^{IV/II}$ and Leu^{IV} , forming the part of the “core” belonging to the “leading” group, on the one hand, and one for Leu^{IV} , the part of the “core” belonging to the “nonleading” group, on the other. (The “seeds” of the sequence a_n , leading to the sequence of numbers 8, 15, 23, 38 and 61, also allowed us to establish the multiplet structure of the amino acids and the Rumer’s division of the genetic code table in Section 4.1).

Consider the “seeds” of sequence a'_n . They also lead to meaningful results. From Equation (49), defined below in Section 5, we have, for $n = 3$, $1 + 6 + 7 = 20 - 6$ or $1 + 6 + 7 + 6 = 20$. Analogously to what we accomplished above, we call the index $n = 0$ and the recurrence relation $a'_0 + a'_1 = a'_2$, that is, $5 + 1 = 6$; this is the first number six in the equation above. The second number six, which is also a perfect number, could be written as the sum of its proper divisors: $6 = 1 + 2 + 3$ (this trick was also useful in Section 4.1). By bringing together these terms and arranging, we obtain, finally, $1 + (1 + 1) + 5 + 3 + (7 + 2) = 1 + 2 + 5 + 3 + 9 = 20$. This last relation could be interpreted as the sum of the number of multiplets of the standard genetic code: 1 triplet, 2 singlets, 5 quartets, 3 sextets and 9 doublets, that is, 20 amino acids (see the introduction). (The “seeds” of the sequence a'_n also lead to meaningful results, like the distribution of hydrogen atoms in Equation (10), which, in turn, is in agreement with Equation (34); see just below).

Finally, the sequence g_n , defined below in Equation (26), together with its “seeds”, 23 and -3 , will lead us to establish Equations (34) and (35), below in the next section, and these latter are also shown to agree with the “ideal” symmetry classification scheme of Section 4.2.2.

4.3. The Atom Content and Degeneracy

Over the course of writing this paper, we have discovered one more Fibonacci-like sequence, tailor-made for the description of the atom number content in Equation (29) below. It is defined as follows:

$$g_n = -3F_{n-1} + 23F_{n-2}. \tag{26}$$

where the numbers 23 and -3 are the “seeds”. The first few terms are shown below:

$$g_n : 23, -3, 20, 17, 37, 54, 91, 145, 236, 381, \dots \tag{27}$$

This sequence is related to the sequences a_n and b_n , as follows:

$$b_n + g_n = 6a_n, \tag{28}$$

which can be shown to hold (see Appendix C). The case $n = 9$ is particularly relevant. We have, from Table 5 and the series in Equation (27) above,

$$358 + 236 = 594, \tag{29}$$

and we see that it gives the total number of atoms in the side chains of all the amino acids coded by the 61 sense codons, distributed into 358 hydrogen atoms (see Section 4.2.1) and 236 atoms (C/N/O/S); see Table A1 in Appendix A (180 carbon atoms and 56 N/O/S atoms). Now, we have the relation

$$\sum_1^k g_n = g_{k+2} - g_2 = g_{n+2} - (-3) = g_{n+2} + 3, \tag{30}$$

which can also be shown to hold for any k , which is the analog of the sum of the first k Fibonacci numbers. For $k = 7$, it gives $236 + 3 = 239$ or $236 = 239 - 3$. By inserting this latter in the above equation, we obtain

$$239 + (358 - 3) = 239 + 355 = 594. \tag{31}$$

Here, we have the number of atoms, also in the “23 + 38” pattern: 239 atoms in all the side chains of the amino acids encoded by 23 codons (the sextets with 35 atoms are counted two times) and 355 atoms in the side chains of the amino acids encoded by the remaining 38 degenerate codons (see Table A1 in Appendix A). Let us, at this stage, remember the sequence c_n , especially its “seeds” $a_1 = 30$ and $a_2 = 5$ with the sum $a_1 + a_2 = 35$. They were chosen, intentionally, as the sum of the number of atoms in arginine and leucine, equal to 30 ($= 17 + 13$), on the one hand, and the number of atoms in serine, equal to 5, on the other (see Section 4.2.2). Their sum is therefore just the right thing to add and subtract from Equation (31) above to obtain

$$(239 - 35) + (355 + 35) = 204 + 390 = 594, \tag{32}$$

which is the correct partition of the number of atoms—this time, in the pattern “20 + 41” (see the comments between Equations (11) and (12) in Section 4.2.1 for hydrogen). We have 204 atoms in the side chains of 20 amino acids, on the one hand, and 390 atoms in the side chains of the amino acids encoded by 41 degenerate codons (see Table A1 in Appendix A). Now, the use of the above sum in Equation (30), for $k = 8$, gives $\sum_1^8 g_n = 384$, which appears also doubly significant; see below. By subtracting this latter number from the total sum, 594, and arranging, we have

$$210 + 384 = 594. \tag{33}$$

This partition of the number of atoms also has an interpretation: there are 210 atoms in the side chains of the six entities (the sextets) Ser^{IV-II} , Arg^{IV-II} and Leu^{IV-II} (35×6) encoded by 18 codons and 384 atoms in the side chains of the remaining 17 amino acids encoded by 43 codons (taking into account the degeneracy). It is worth noting that the first two recurrence relations of the sequence g_n $23 - 3 = 20$ and $20 - 3 = 17$, together, lead to the relation

$$23 = 17 + (3 + 3), \tag{34}$$

which is in line with the above result for the atom numbers and also with the “ideal” symmetry scheme (as depicted below):

$$(3 + 3) \leftrightarrow (Ser^{IV}, Arg^{IV}, Leu^{IV}) + (Ser^{II}, Arg^{II}, Leu^{II}). \tag{35}$$

Finally, we could also derive the partition of the number of atoms for Rumer’s sets M_1 and M_2 . Consider, again, the equation above, $210 + 384 = 594$ —more precisely, the number 384, which was calculated from Equation (30), with $k = 8$. By partitioning this sum in two parts: the first, for $k = 4$, gives $54 - (-3) = 54 + 3$, and the second, which is equal to $g_5 + g_6 + g_7 + g_8$, gives 327. By inserting these two parts in Equation (33) and arranging, we obtain

$$(210 + 54) + (327 + 3) = 264 + 330 = 594. \tag{36}$$

This is the content in atoms in $M_1(264)$ and in $M_2(330)$; see Table A1 in Appendix A. We can also reveal the details for the multiplets. Considering, first, M_1 , let us present the following (new) relation connecting the sequences b_n and c_n :

$$c_n + b_{n+2} = 4b_n, \tag{37}$$

It could also be checked following the hints in Appendix C. For $n = 3$, it gives $35 + 53 = 4 \times 22 = 88$. Using a recurrence relation for b_n , we have $53 = 31 + 22$, and by combining the above two relations, we obtain $35 + 31 + 22 = 4 \times 22$, or, by subtracting 22 from both sides, we obtain $31 + 35 = 3 \times 22 = 66$. Multiplying this latter equation by any number does not change it, particularly by 4, keeping in mind that the eight quartets composing the set M_1 each have four codons, and we have $4 \times 31 + 4 \times 35 = 264$. This is the detailed number of atoms in M_1 : 4×31 in the five quartets and 4×35 in the three quartet parts of the three sextets (see Table A1 in Appendix A). The above equation, $31 + 35 = 66$, which was used as an intermediate of the calculation above, could also be exploited for the set M_2 . Consider Equation (5), $a'_n - b_{n-2} = 2F'_{n-5}$, for $n = 6$: $33 - 31 = 2$. The insertion of this difference in the above equation gives $33 + 35 = 68$. Now, the following relation linking the Fibonacci and Lucas numbers $L_n + 3F_n = 2F_{n+2}$, for $n = 7$, gives $29 + 3 \times 13 = 2 \times 34 = 68$. If, moreover, we use the recurrence relation for the Lucas number $29 = 11 + 18$, we obtain $3 \times 13 + 11 + 18 = 68$. This perfectly matches the number of atoms in the triplet isoleucine (3×13) and in the two singlets methionine (11) and tryptophane (18); see Table A1 in Appendix A. We showed above that there are 330 atoms in the set M_2 . Subtracting the above number of atoms, 68, in the triplet and in the two singlets, we are left with 262 atoms. To obtain the right partition of these, it suffices to take the sum of the first three members of the sequence c_n : $30 + 5 + 35 = 2 \times 35 = 70$, which appears to be the right number of atoms in the doublet parts of the three sextets. Adding and subtracting this latter from 262 gives 192, which is the number of atoms in the nine doublets, $2 \times 96 = 192$ (see Table A1 in Appendix A). In summary, we have

$$\begin{aligned} M_1 : 4 \times 31 + 4 \times 35 &= 264, \\ M_2 : 192 + 2 \times 35 + 3 \times 13 + 11 + 18 &= 330, \end{aligned} \tag{38}$$

which is the precise and detailed partition. Finally, let us note that the number 384, mentioned below Equation (32), also has another relevant interpretation. It is equal to the number of atoms in the 20 amino acids, this time adding to the side chains their 20 identical backbones with 9 atoms each: $204 + 9 \times 20 = 384$.

4.4. Derivation of Several Nucleon Number Patterns

In this section, we use our Fibonacci-like series to derive several patterns for the nucleon number (or *integer molecular mass*) content. Before starting, let us make an important remark about the sequence c_n (see Table 5). There is a simple relation between the sequences a_n and c_n ; the latter is simply five times the former: $c_n = 5a_n$. One may wonder how the use of c_n would bring something significant, as it is simply related to a_n . In fact, it does, and we will show that below. First, let us consider the following sum:

$$\sum_1^9 a_n + 2 \sum_1^9 b_n + \sum_1^9 c_n = 3404. \tag{39}$$

It appears that this number, 3404, is the number of nucleons in the side chains of all the amino acids coded by the 61 sense codons (see Table A1 in Appendix A). This is nice, but we could do more. Consider again the “seeds” of the sequence c_n , 30 and 5 with the sum 35, the number of atoms in the side chains of the three sextets serine (5), arginine (17) and leucine (13). Here, we call Zeckendorf’s theorem which states that every positive integer can be represented uniquely as the sum of one or more non-consecutive Fibonacci numbers. It is not difficult, by applying this theorem to the number 30 ($= 21 + 8 + 1$) and the fact that $21 = 13 + 8$, to show that the sum of the “seeds” takes the form $13 + 17 + 5 = 35$, i.e., the correct atom numbers in the three sextets, mentioned above. Now, by isolating the sum of the above “seeds” of c_n from the third sum in Equation (39) and including it in the two other sums, we obtain

$$2149 + 1255 = 3404. \tag{40}$$

Here, we have a significant result: there are 1255 nucleons in the side chains of the 20 amino acids (see Table A1 in Appendix A) and 2149 nucleons in the side chains of the amino acids encoded by the 41 degenerate codons, following, again, the pattern “20 + 41” (see Equations (11) and (32)). Let us now exploit the relation between the two sequences a_n and c_n ($c_n = 5a_n$), mentioned above, and write the sum in Equation (39) in the form

$$\left[4 \sum_1^9 a_n + \sum_1^9 b_n \right] + \left[2 \sum_1^9 a_n + \sum_1^9 b_n \right] = 1960 + 1444 = 3404. \tag{41}$$

Recall the sum $\sum_1^k a_n = a_{k+2} - 1$, mentioned in Equation (6) of Section 4.1. In the present case, for its use in Equation (41), we have $\sum_1^9 a_n = 259 - 1$ for $k = 9$. By considering this latter relation in *only one* such sum in the *first* bracket of the above equation and including the unit “-1” in the *second* bracket, we obtain

$$1961 + 1443 = 3404. \tag{42}$$

One recognizes here the nucleon number, in the pattern “38 + 23” (see above and Appendix A): 1443 nucleons in the side chains of the amino acids coded by 23 codons (the sextets counted two times) and 1961 nucleons in the side chains of the remaining amino acids encoded by 38 degenerate codons. We can also, from the above relations, make contact with the “ideal” symmetry scheme of Section 4.2, at the level of the nucleon numbers. To do this, let us first remark that the number 114 appears *twice*, once as the number of hydrogen atoms in the part of the “core” belonging to the “leading” group of the “ideal” symmetry scheme (see Section 4.2.2) and once as the number of nucleons in Leu^{II} (2×57), the part of the “core” belonging to the “nonleading” group (see Table A1 in Appendix A). This will prove significant in the following. Consider the sum

$$\sum_1^9 a_n + 2 \sum_1^9 b_n = 2114. \tag{43}$$

The number 2114 by itself is not very interesting, but its φ -function is. We have $\varphi(2114) = 900$ (see Equation (A3) in Appendix B) and, adding to this two times the number 114 gives $900 + 2 \times 114 = 1128$. This is the number of nucleons in the “core”: $31 \times 6 + 100 \times 6 + 57 \times 6 = 1128$. Arranging the sum as $(900 + 114) + 114 = 1014 + 114$ gives the partition of the nucleon numbers between the two parts of the “core”, $31 \times 6 + 100 \times 6 + 57 \times 4 = 1014$ in the “leading” group, on the one hand, and $57 \times 2 = 114$ in the “nonleading” group, on the other:

$$1128 = 1014 + 114. \tag{44}$$

In the following, we can also derive three more results by “watering three plants with one hose”, so to speak. Consider again the sum in Equation (39), and split it as follows:

$$\left[\sum_1^9 a_n + \sum_1^9 b_n + \sum_1^7 c_n \right] + \left[\sum_1^9 b_n + \sum_8^9 c_n \right] = 1676 + 1728 = 3404. \quad (45)$$

We have here the nucleon number pattern of the third base classification of Section 2.2: 1728 nucleons in the U/C third-base division and 1676 nucleons in the A/G third-base division (see Table 3, last row). By borrowing, from the first bracket above, the sum of the first three members of the sequence c_n : $30 + 5 + 35 = 2 \times 35 = 70$, the one we used earlier (see above Equation (38)), to the benefit of the second bracket, we obtain (as an example of evaluation from the table in Appendix A, one obtain for the «leading» group: $31 \times 6 + 57 \times 4 + 100 \times 6 + 15 \times 4 + 59 \times 2 + 73 \times 2 + 107 \times 2 + 57 \times 3 + 75 = 1798$). Here, we recognize the number of nucleons in the “leading” group, 1798, and that in the “nonleading” group, 1606.

$$1606 + 1798 = 3404. \quad (46)$$

Finally, we could also establish the nucleon number pattern corresponding to Rumer’s division. Consider again Equation (39). We partition it as follows:

$$\left[\sum_1^8 a_n + 2\sum_1^8 b_n + \sum_1^8 c_n \right] + (a_9 + 2b_9 + c_9) = 2094 + 1310 = 3404. \quad (47)$$

It suffices now, analogously to what we did in Equation (40) above, to subtract, *once*, the sum of the “seeds” of the sequence b_n in the bracket, that is, $13 + 9 = 22$, and add it to the three terms in the parenthesis to obtain

$$2072 + 1332 = 3404. \quad (48)$$

We have, as promised above, 1332 nucleons in M_1 and 2072 nucleons in M_2 (see Table A1 in Appendix A).

5. On Proline’s Singularity and a Derivation of the shCherbak–Makukov “Activation” Key

In this section, we use our Fibonacci-like sequences to shed light, by giving concrete results, on a question relative to the special amino (more exactly, imino) acid, proline, which is an exception among the set of 20 amino acids. It is the only amino acid whose side chain is connected to its backbone twice. shCherbak, [12], to “standardize” the common backbone of the amino acids, with 74 nucleons, proposed an *imaginary* “borrowing” of one nucleon (one hydrogen atom) from the side chain of proline, which has only 73 nucleons in its backbone, to the benefit of this latter, to reach 74, as is the case for 19 other amino acids. In his next work with Makukov, [13], the above “borrowing” process, or the transfer of one nucleon, has been termed the “activation key”. Activating the key, i.e., standardizing, leads to an innumerable number of remarkable and beautiful arithmetical patterns. These authors write in their paper: “*Applied systematically without exceptions, the artificial transfer in proline enables holistic and arithmetically precise order in the code*”. Here, in this section, we establish not only a mathematical version of the “activation key” itself but also its effect on the total hydrogen atom content, with simple possible extensions to the atom and nucleon content. Let us begin by examining the action of the “activation key”. Consider, again, the sequence a'_n and the following sum:

$$\sum_1^k a'_n = a'_{k+2} - 6. \quad (49)$$

It could be shown and verified that the above relation holds for any k (see Appendix C). For $k = 9$, it gives $358 = 364 - 6$. (This low k case could simply be evaluated from Table 5 using a pocket calculator.) As established and mentioned many times previously, 358

is the number of hydrogen atoms in the side chains of all the amino acids coded by the 61 sense codons, where the special amino acid proline has 5 hydrogen atoms in its side chain. If, instead, one considers that proline’s side chain now has six hydrogen atoms, at the cost of its block, i.e., no standardization made, or the “activation key” *off* (see below), and taking into account the number of its coding codons, which is four, then we now have $362 = 358 + 4$ hydrogen atoms in the side chains of all the amino acids coded by the 61 sense codons. Let us reconsider Equation (10), the partition of the number of hydrogen atoms between the amino acids encoded by 38 degenerate codons, 219, and the amino acids encoded by 23 codons, 139, (the sextets counted twice), but now using the above relation ($358 = 364 - 6$):

$$\sum_1^8 a'_n + a'_9 = 219 + 139 = 364 - 6. \tag{50}$$

To obtain a correct partition, let us consider the *perfect number* 6 which is, as such, equal to the sum of its proper divisors: $6 = 1 + 2 + 3$ (also used in Section 4.2.5). These are just the right numbers we need. By inserting them in the above equation by selecting the odd divisors 1 and 3 and shifting them to the left while leaving the even one 2 to the right, and finally arranging them properly, we obtain

$$\sum_1^8 a'_n + a'_9 = (219 + 3) + (139 + 1) = 364 - 2 = 362, \tag{51}$$

We have here something noteworthy: one more hydrogen atom in the amino acids in the part encoded by 23 codons and 3 more hydrogen atoms for its 3 degenerate codons, still in its side chain and located in the degeneracy part.

Taking a look at the sixth term in the sequence c_n , $115 = 40 + 75$, it appears to be equal to the number of nucleons in proline’s *side chain and backbone*; see below about this latter sum. This number, 115, is “invariant” whether we make shCherbak’s “borrowing” of one nucleon or not. To obtain more insight, we consider another invariant number, the total number of hydrogen atoms in all the amino acids coded by the 61 sense codons, *including* the backbones (with 4 hydrogen atoms in each), that is, $358 + 244 = 362 + 240 = 602$. Without borrowing one nucleon from the side chain of proline in favor of its block, there are 362 hydrogen atoms in the side chains and 240 hydrogen atoms, $57 \times 4 + 4 \times 3 = 240$, in the backbones of all the amino acids coded by the 61 sense codons. Applying the “borrowing”, there are 358 hydrogen atoms in all the side chains and $244 (= 61 \times 4)$ hydrogen atoms in all the backbones. Note, in passing, the following nice relations seemingly linking the two views: $\varphi(240) + \varphi(362) = 244$ and $(240 + 362) - [\varphi(240) + \varphi(362)] = 358$.

Now, let us examine the former point, the derivation of the “activation key”. Considering the above-mentioned invariant numbers, $115 (= 5 \times 23)$ and $602 (= 2 \times 7 \times 43)$, we have, using their A_0 function (defined in Appendix B):

$$115 - A_0(115) = 115 - 42 = 73, \tag{52}$$

$$115 - A_0(602) = 115 - 74 = 41. \tag{53}$$

From these relations, we deduce that $115 = 42 + 73 = 41 + 74$, which is seen to describe, fully and precisely, the two views: $42 + 73$ (“activation key” *off*) and $41 + 74$ (“activation key” *on*). From $(\sigma(41)) = 42 = 41 + 1$, where σ is the sum of the divisors, we can also write $115 = (41 + 1) + 73 = 41 + (73 + 1)$. Also, from $\varphi(41) = 40 = 41 - 1$, we can make contact with the sequence c_n through the relation $c_6 = 115 = 40 + 75$, mentioned above: $41 + (75 - 1) = 41 + 74$. Moreover, we can alternatively exploit the number 75 itself. Calling Legendre’s three squares theorem. This theorem states that a natural number n can be represented as a sum of three squares if and only if it is not of the form $4^a (8b + 7)$ for a and b two positive integers. It could be easily verified that the number 75 cannot be written in this form so it can be represented as the sum of the following three squares.): $1^2 + 5^2 + 7^2$ or $1 + (25 + 49) = 1 + 74$. This latter form gives us, again, $(40 + 1) + 74 = 41 + 74$.

Finally, using $\varphi(41) = 40 = 41 - 1$ and the decomposition of the number 75 as the sum of three squares, mentioned above, we can write, by allocating the two units in two ways: $41 - 1 + 1 + 74 = 41 + 74 = 42 + 73$. This is, again, what we found above from Equations (52) and (53).

6. A Remarkable Imprint in the “Seeds”

Before starting this section, let us remember what has been said about the three sextets in Section 4.2. In the “ideal” symmetry classification scheme, briefly described in Section 2.3, the authors explain that, in their approach, the symmetries are a consequence of the sextet’s dynamics, and the whole set of amino acids is created starting from these three sextets, where serine plays a prominent role. In our present approach, relying on the use of Fibonacci-like series, on the other hand, we have chosen, as already mentioned, for two of them, b_n and c_n , the hydrogen atom and atom numbers of the three sextets (see Section 4.2) as “seeds”. We have also explained, in Section 4.2.5, that the ‘seeds’ of the other sequences a_n , a'_n and g_n were, as mentioned above, found by a thought process but have been shown to also lead to meaningful results as the degeneracy structure of the codons or a connection with the ‘ideal’ classification scheme. Below, we show that the “seeds” of all the Fibonacci-like sequences used in this paper, and only these, by themselves, can remarkably “create” the main hydrogen number patterns derived in this paper. The *sum and product* of the “seeds” of the sequence b_n , alone, gives

$$b_1 \times b_2 + (b_1 + b_2) = 117 + 22 = 139. \tag{54}$$

One recognizes here the number of hydrogen atoms in the side chains of the 20 amino acids, 117, augmented by the number of hydrogen atoms in the three sextets, 22. The total, 139, corresponds to 23 codons (the sextets counted two times). Let us now compute the following expression, using the *sum and product* of the “seeds” of the sequence c_n and *only* the sum of the “seeds” of the other three remaining sequences a_n , a'_n and g_n (the latter defined in Equation (26)). We have

$$\begin{aligned} c_1 \times c_2 + (c_1 + c_2) + (a_1 + a_2 + a'_1 + a'_2) + (g_1 + g_2) = \\ = 150 + 35 + 14 + 20 = 219. \end{aligned} \tag{55}$$

Here, we have the number of hydrogen atoms in the side chains of the amino acids coded by the 38 degenerate codons. Equations (54) and (55), together, constitute the “23 + 38” hydrogen atom pattern established in Section 4.1. Furthermore, borrowing the number 22 from Equation (54) to the benefit of Equation (55) gives $117 + 241 = 358$, which corresponds to the other pattern “20 + 41” (see Equations (10) and (11)). Next, we arrange Equations (54) and (55) as follows:

$$(150 + 22) + (117 + 35 + 14 + 20) = 172 + 186 = 358. \tag{56}$$

Here, we have, again, the hydrogen atom content in Rumer’s division: 172 hydrogen atoms in M_1 and 186 hydrogen atoms in M_2 ; see Section 4.2 and Equations (19) and (20). To obtain the other patterns, we call the Fibonacci (0, 1, 1, 2, 3, 5, ...) series and the Lucas (2, 1, 3, 4, 7, 11, ...) series, which, as is well known, are linked by the relation $F_n + L_{n+2} = F_{n+4}$. For $n = 5$, we have $5 + 29 = 34$, so we can replace the term $34 = 14 + 20$ in Equation (56) with the latter. By arranging, we obtain

$$(150 + 35 + 5) + (22 + 117 + 29) = 190 + 168 = 358. \tag{57}$$

This is the hydrogen atom pattern for (i) the third base classification of Section 4.2 (Equation (14)) and (ii) the “ideal” symmetry classification scheme in the same section (Equation (22)).

Finally, we reconsider Zekendorf’s theorem (see above) and apply it to the number 117, giving $89 + 21 + 5 + 2$. Writing 89, a Fibonacci number, as $55 + 34$, we can rearrange the content of the second parenthesis in Equation (57) above as $55 + 29 = 84$ and $34 + 21 + 22 + 5 + 2 = 84$, so that $168 = 2 \times 84$, which, again, describes the pattern $190 + 2 \times 84 = 358$. The fact of having used the Fibonacci and Lucas sequences here is all the more interesting in that it can also give us another remarkable result. By adding the two “seeds” of the Fibonacci and Lucas sequences, 0 and 1 and 2 and 1, respectively, to the above sum of Equations (54) and (55) and arranging, we obtain

$$(139 + 1) + (219 + 2 + 1) = 140 + 222 = 362. \tag{58}$$

which is the hydrogen atom pattern found in Section 5, devoted to the special imino acid proline and the shCherbak–Makukov “activation” key, when this latter is “off”; see Equation (51) in Section 5.

7. The Case of the Vertebrate Mitochondrial Genetic Code

One can wonder whether these findings (i) could find biological applications and/or (ii) are specific to the current standard genetic code table, especially concerning symmetry. The answer to these questions is certainly difficult, but, as a shy beginning, we have found, while ending this paper, that something could be said about the point (ii), at least for the hydrogen atom content. It is about the *vertebrate mitochondrial genetic code, the only perfect symmetry genetic code* [16–18]. In this code, there is no triplet and there are no singlets; there are only sextets, quartets and doublets (see [19]). Briefly, arginine loses its two codons (AGA and AGG) of its doublet part, which are now assigned to two new stop codons, and joins the quartet set as a sixth member. Tryptophane picks the stop codon, UGA, and becomes a doublet. Methionine absorbs the codon AUA of isoleucine to also become a doublet, leaving only a doublet isoleucine. In summary, we have 2 sextets, 6 quartets and 12 doublets; see [19]. Looking at Table A1 in Appendix A and the data below it, we have $(9 + 3) \times 6 + (21 + 10) \times 4 + (50 + 9 + 7 + 8) \times 2 = 344$ hydrogen atoms in the amino acids coded by the 60 sense codons (there are, in the present case, four stop codons). From the above relation, we can see that the count for the two sextets and the six quartets is 196, while the one for the 12 doublets is 148. Now, we apply our Fibonacci-like formalism to this case. From Table 5 of Section 3, we have, by a quick pocket calculator computation, $2\sum_1^7 a_n = 2 \times 98 = 196$ and $\sum_1^6 g_n = 148$, so that these sums correctly describe the above two counts. From Equation (6) in Section 4.1, we have $98 = 99 - 1$, and from Equation (3) in Section 3 for $n = 9$, we obtain $99 = 86 + 13$, so the above sum now writes $2\sum_1^7 a_n = 2 \times (86 + 13 - 1) = 2 \times (86 + 12) = 2 \times 86 + 2 \times 12 = 172 + 24$. By summarizing and arranging, we are left with

$$2\sum_1^7 a_n + \sum_1^6 g_n = 172 + (24 + 148) = 172 + 172 \tag{59}$$

This is a mathematical *balance*, established here by computation, and has a precise equivalent for the actual hydrogen atom count in the two Rumer sets M_1 and M_2 (see the data in Table A1 in Appendix A):

$$[(9 + 3) + (21 + 10)] \times 4 + [(50 + 9 + 7 + 8) + (9 + 3)] \times 2 = 172 + 172 \tag{60}$$

where we put the quartet part of the two sextets with the quartets and their doublet part with the other doublets. It is even possible to separate, in M_1 , the hydrogen atom count of the quartets from that for the quartet part of the two sextets by writing the above term, 2×86 , as $2 \times (84 + 2) = 2 \times (2 \times 31 + 22 + 2) = 4 \times 31 + 4 \times 12$, where we have used the identity in Equation (5) for $n = 8$ ($86 = 84 + 2$) and also the recurrence relation of the sequence b_n to write 84 as $53 + 31$ and next as $31 + 22 + 31 = 2 \times 31 + 22$. We

have, therefore, a perfect description, via computation, of the highly symmetric *vertebrate mitochondrial genetic code* (VMC). The summary is depicted in Table 7 below, where the hydrogen atom numbers of the two parts of the sextets, the quartet part, 48, in M_1 , and the doublet part, 24, in M_2 , are set apart. (Observe that the “symmetry” of the numbers is also gracefully put on show).

Table 7. The hydrogen atom content in the VMC (Rumer’s division).

M_1	M_2
{48, 124}	{24, 148}
172	172

8. Conclusions

In this work, we have strayed a little off the beaten paths in genetic code mathematical research. Starting with a handful of Fibonacci-like sequences, in Section 3, we have derived not only the degeneracy structure of the genetic code, in Section 4.1, but also the hydrogen atom content, in Sections 4.2.1–4.2.4. We have also included, in Section 4.2.5, a discussion devoted to the choice of the initial conditions of our Fibonacci-like sequences. Next, we derived the atom number content, in Section 4.3, and also the integer molecular mass (nucleon) content of the set of 20 amino acids, as structured in the 64-codon table, in Section 4.4. As a by-product of our mathematical formalism, we derived the atomic (elemental) content of the building blocks of RNA, the four ribonucleotides UMP, CMP, AMP and GMP, in Section 4.2.2.

Still using the above mathematics, we bring, for the first time, in Section 5, an additional brick to shCherbak’s theory, concerning the role of the special imino acid proline whose virtual “double” structure renders possible, via the use of the “activation key”, a large number of remarkable and beautiful arithmetical patterns.

In Section 6, we show that the “seeds” of our Fibonacci-like sequences and only these, by themselves, are capable of reproducing the main hydrogen number patterns derived in this paper.

Finally, in Section 7, we have applied, successfully, our Fibonacci-like formalism to the highly symmetrical vertebrate mitochondrial genetic code as well as a numerical hydrogen atom balance inherent to Rumer’s division of the genetic code table.

Our main findings, such as the total hydrogen atom content, the total atom content, the total molecular mass content of the 20 amino acids, including the degeneracy, as well as other relevant quantities related to the symmetries of the genetic code, are found directly, either as ostensible members of the Fibonacci-like sequences or from the summation properties of the latter.

Let us note that the hydrogen atom, atom and nucleon contents of the amino acids considered in this work are the ones corresponding to their neutral state. This choice has also been considered in [12]. Now, it is well known that few amino acids are charged in their normal (physiological) state. This case can also lead to the existence of remarkable (nucleon or integer mass) balances; see [13] and also [20]. We have found that this latter case could also be handled using the mathematical formalism used in the present work. The corresponding results, which are in progress, will be submitted soon for publication.

Below, we give a brief summary of the paper, in a “one-liner” format, showing only the main “parent” relations whose numerous “offsprings”, which are derived in the different sections, disclose the symmetries of the genetic code.

1. Hydrogen atoms in all the amino acid side chains coded by 61 sense codons (Section 4.2)

$$\sum_{i=1}^9 a'_i + a'_9 = 219 + 139 = 358$$
2. Atoms (H/CNOS) in all the amino acid side chains coded by 61 sense codons (Section 4.3)

$$b_9 + g_9 = 6a_9 = 358 + 236 = 594$$
3. Integer molecular mass (nucleon number) in all the amino acid side chains coded by 61 sense codons (Section 4.4)

$$\sum_1^9 a_n + 2 \sum_1^9 b_n + \sum_1^9 c_n = 3404$$
4. Hydrogen atoms in all the amino acid side chains coded by 60 sense codons in the vertebrate mitochondrial genetic code (Rumer's division, Section 7)

$$2 \sum_1^7 a_n + \sum_1^6 g_n = 344 = 172 + 172$$

Funding: This research received no external funding.

Data Availability Statement: No data availability Statement.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

In the table of this appendix, we give the detailed elemental composition of the side chains of the 20 amino acids. H stands for hydrogen, C for carbon, N for nitrogen, O for oxygen and S for sulfur. The calculated values of some important quantities, taking into account the degeneracies, are indicated in the last five rows; they are useful to know when reading the main text (*those shown in red color are all mathematically derived in this paper using the present new approach*). In the table, the first column, M, gives the number of codons which code for an amino acid (four for a quartet, six for a sextet, two for a doublet, three for a triplet and one for a singlet). In column six, we provide the number of atoms in the side chains, and the number of nucleons (protons and neutrons), which is also the integer molecular mass of an amino acid, is displayed in column 7. Below the table, we offer hints for computing some of them. The table is in the "standardized" form, that is, proline has 5 hydrogen atoms in its side chain, and all 20 amino acids, including proline, have 74 nucleons in each of their backbones; see Section 5. The general chemical (linear) formula of an amino acid is



where R is the radical, also called the side chain, and the rest of the molecule constitutes the backbone. Also, the side chain is bound to the α -carbon. In the special case of proline, its side chain from the α -carbon connects to the nitrogen N, forming a pyrrolidine loop. (It is the side chain that gives an amino acid its specific functional properties.) To calculate, for example, the nucleon numbers or the integer molecular mass of an amino acid, the molecular masses of the chemical elements are those of the most abundant isotopes: hydrogen (1), carbon (12), nitrogen (14), oxygen (16) and sulfur (32). From the formula above, one easily computes the integer molecular mass of the backbone: $2 \times 12 + 1 \times 14 + 2 \times 16 + 4 \times 1 = 74$. In the (*unique*) case of proline, as mentioned above, there is one less hydrogen atom in the backbone, and the nucleon number is $73 = 74 - 1$; this is the non-standardized form ("activation key" off) (see Section 5).

Table A1. The elemental composition of the 20 amino acids.

M	Amino Acid	# H	# C	# N/O/S	# Atoms	# Nucleons
4	Proline (Pro)	5	3	0	8	41
	Alanine (Ala)	3	1	0	4	15
	Threonine (Thr)	5	2	0/1/0	8	45
	Valine (Val)	7	3	0	10	43
	Glycine (Gly)	1	0	0	1	1
6	Serine (Ser)	3	1	0/1/0	5	31
	Leucine (Leu)	9	4	0	13	57
	Arginine (Arg)	10	4	3/0/0	17	100
	Phenylalanine (Phe)	7	7	0	14	91
	Tyrosine (Tyr)	7	7	0/1/0	15	107
	Cysteine (Cys)	3	1	0/0/1	5	47
2	Histidine (His)	5	4	2/0/0	11	81
	Glutamine (Gln)	6	3	1/1/0	11	72
	Asparagine (Asn)	4	2	1/1/0	8	58
	Lysine (Lys)	10	4	1/0/0	15	72
	Aspartic Acid (Asp)	3	2	0/2/0	7	59
3	Glutamic Acid (Glu)	5	3	0/2/0	10	73
	Isoleucine (Ile)	9	4	0	13	57
	Methionine (Met)	7	3	0/0/1	11	75
1	Tryptophane (Trp)	8	9	1/0/0	18	130
	Total (20)	117	67	20	204	1255
	Total (23)	139	76	24	239	1443
	Total (38)	219	104	32	355	1961
	Total (61)	358	180	56	594	3404
	M ₁ /M ₂	172/186			264/330	1332/2072

Obtaining the results in the second of the last five rows from the first one, it suffices to count the values of the sextets two times. For the rest, to ease the calculations, one can use the following pre-calculated sums for the hydrogen atom content: 5 quartets 21, 3 sextets 22, 9 doublets 50, 1 triplet 9 and 2 singlets 15 = 7 + 8. For the atom number, it is: 5 quartets 31, 3 sextets 35, 9 doublets 96, 1 triplet 13 and 2 singlets 29 = 11 + 18. For the nucleon numbers, it is: 5 quartets 145, 3 sextets 188, 9 doublets 660, 1 triplet 57 and 2 singlets 205 = 75 + 130.

In the calculations, the reader also needs to know what we mean by degeneracy. This latter is defined as the number of codons coding for an amino acid *minus one*. Therefore, for a quartet, the degeneracy is 3 = 4 - 1; for a doublet, it is 1 = 2 - 1; for a triplet, it is 2 = 3 - 1 and for a singlet, it is 0 = 1 - 1. For the special case of the sextets, there are two possibilities related to the two patterns mentioned several times in this paper: "20 + 41 = 61" and "23 + 38 = 61". In the first case, the degeneracy is 3 + 2 = 5 (three for the quartet part and two for the doublet part whose two codons are both considered degenerate). In the second case, the quartet part and the doublet part of each sextet are considered as *separate* entities (e.g., Ser^{IV} and Ser^{II}), so the degeneracy is equal to 3 + 1 = 4, three for the quartet part and one for the doublet part, which, here, is considered as a doublet. In this way, for the number of amino acids and the total number of coding codons, we have 20 = 5 + 3 + 9 + 1 + 2 and 41 = 5 × 3 + 3 × 5 + 9 × 1 + 1 × 2 in the first case and 23 = 5 + (3 + 3) + 9 + 1 + 2 and 38 = 5 × 3 + 3(3 + 1) + 9 × 1 + 1 × 2 in the second one. With these definitions, it is not difficult

to carry out the rest of the computations. Let us give a few examples from the table above for the number of hydrogen atoms for the pattern “23 + 38”: $139 = 21 + 22 \times 2 + 50 + 9 + 7 + 8$, $219 = 21 \times 3 + 22 \times 4 + 50 \times 1 + 9 \times 2$, $358 = 21 \times 4 + 22 \times 6 + 50 \times 2 + 9 \times 3 + 7 + 8$.

Appendix B

In this appendix, we mention a few other additional mathematical elements used in this paper: (i) Euler’s phi totient function, (ii) the Carmichael lambda function and (iii) our function A_0 . All these functions rely on the *Fundamental Theorem of Arithmetic*, which states that every integer n (except the number one) can be represented, uniquely, as a product of prime numbers, irrespective of their order:

$$n = p_1^{n_1} \times p_2^{n_2} \dots \times p_k^{n_k} \tag{A1}$$

First, there is Euler’s totient function for an integer n , $\varphi(n)$, which is extensively used in many scientific areas such as in cryptography and graph theory. It counts the number of positive integers less than or equal to n which are relatively prime to n (also called coprimes). For example, 24 has 8 coprimes (1, 5, 7, 11, 13, 17, 19, 23): $\varphi(24) = 8$. A simple formula for computing this function is the following (see [21])

$$\varphi(n) = n \prod_{i=1}^m \left(1 - \frac{1}{p_i}\right) \tag{A2}$$

where m is the distinct prime factors in the factorization (A1). Let us take two examples from the text: $\varphi(2114) = 900$ (see below Equation (43)) in Section 4.4 and $\varphi(114) = 36$ (mentioned above Equation (17)) in Section 4.2.2. The prime factorizations of these two numbers are given by $2114 = 2^1 \times 7^1 \times 151^1$ and $114 = 2^1 \times 3^1 \times 19^1$. From Equation (A2), we have, respectively,

$$\varphi(2114) = 2114 \times \left(1 - \frac{1}{2}\right) \times \left(1 - \frac{1}{7}\right) \times \left(1 - \frac{1}{151}\right) = 1 \times 6 \times 150 = 900 \tag{A3}$$

$$\begin{aligned} \varphi(114) &= 114 \times \left(1 - \frac{1}{2}\right) \times \left(1 - \frac{1}{3}\right) \times \left(1 - \frac{1}{19}\right) \\ &= 1 \times 2 \times 18 = 36 \end{aligned} \tag{A4}$$

Second, there is the Carmichael λ -function, also called the *reduced* totient function, which is, in fact, used only once in Section 4.2, where it appears to be useful. It is defined as the smallest positive divisor of Euler’s totient function that satisfies Euler’s Theorem, [22], which states that if n is a positive integer and a and n are coprime, then $a^{\varphi(n)} \equiv 1 \pmod{n}$, where $\varphi(n)$ is Euler’s totient function. For example, $\lambda(24) = 2$. (The reader could easily find good online calculators for these functions for checking.) Here, there also exists a simple formula for computing this function, using Equation (A1):

$$\lambda(n) = lcm \left[(p_i - 1) p_i^{n_i - 1} \right]_i \tag{A5}$$

where $p_i^{n_i}$ is the prime factors of n from Equation (A1) and *lcm* is the *least common multiple*. Let us give, as an example, the computation of $\lambda(114)$, mentioned above in Equation (17) in Section 4.2.2. From its prime factorization above and Equation (A5), we have

$$\lambda(114) = lcm(1, 2, 18) = 18 \tag{A6}$$

Finally, there is the A_0 function, which is defined by

$$A_0(n) := a_0(n) + SPI(n) + \Omega(n),$$

where $a_0(n)$ is the sum of the prime factors of the integer n , including the multiplicities, $p_1 \times n_1 + p_2 \times n_2 + \dots + p_k \times n_k$, $SPI(n)$ is the Sum of the Prime Indices $PI(p_1) \times n_1 + PI(p_2) \times n_2 + \dots + PI(p_k) \times n_k$, where $PI(2) = 1, PI(3) = 2, PI(5) = 3$ and so on, also including the multiplicities and, finally, $\Omega(n)$, the so-called Big Omega function, is the number of the number of the prime factors $n_1 + n_2 + \dots + n_k$. Consider, as an example, the number 192, whose prime factorization is $2^6 \times 3^1$. We have

$$\begin{aligned} A_0(192) &= a_0(2^6 \times 3^1) + SPI(2^6 \times 3^1) + \Omega(2^6 \times 3^1) \\ &= (6 \times 2 + 1 \times 3) + (6 \times 1 + 1 \times 2) + (6 + 1) = 30. \end{aligned}$$

The function A_0 also enjoys the useful additivity (“logarithmic”) property $A_0(n \times m \times p \times \dots) = A_0(n) + A_0(m) + A_0(p) + \dots$. Let us give a few other illustration examples, taken from Section 4.2, concerning the computation of $A_0(114)$ and $A_0(244)$. For the first, we have $114 = 2^1 \times 3^1 \times 19^1$ such that $A_0(114) = (2 + 3 + 19) + (1 + 2 + 8) + 3 = 38$. For the second, we have $244 = 2^2 \times 61^1$. To obtain the result established in the end of Section 4.2, it makes sense to use the additivity property mentioned above: $A_0(244) = A_0(2^1) + A_0(2^1) + A_0(61^1) = 4 + 4 + (61 + 18 + 1) = 88$. This form, which sets apart the two factors four proved useful in revealing the structure of the four ribonucleotides (in Section 4.2).

Appendix C

In this appendix, we give some hints to the interested reader who wants either to verify the identities in Equation (2) of Section 3 or to carry out the various computations presented in the different sections by himself/herself. In the latter case, where only low values of n are involved, it suffices to use a pocket calculator, along with the data in Table 5 of Section 3. For more complicated cases, like the verification of the identities in Equation (2), especially for large or even very large values of n , a computer is necessary. In this vein, a mathematical software, to the extent that it contains a *built-in* “fibonacci” function, generally written as “fibonacci(i)”, as it exists in Maple, Matlab, Mathematica, etc., could be used. Those familiar with programming languages, like, for example, Python or C++, could use the source codes for the Fibonacci sequence, available in the following links: [23,24], respectively. Given this function, the reader only needs, for performing the verifications or the calculations, to write the five functions a_n, a'_n, b_n, c_n and g_n together with their “seeds” in terms of the fibonacci function, from their definition in Equation (1) of Section 3, as follows:

$$\begin{aligned} a[n] &:= fibonacci(n - 1) + 6 * fibonacci(n - 2) \\ a'[n] &:= 6 * fibonacci(n - 1) + fibonacci(n - 2) \\ b[n] &:= 9 * fibonacci(n - 1) + 13 * fibonacci(n - 2) \\ c[n] &:= 5 * fibonacci(n - 1) + 30 * fibonacci(n - 2) \\ g[n] &:= -3 * fibonacci(n - 1) + 23 * fibonacci(n - 2) \end{aligned} \tag{A7}$$

Let us give some examples.

Example A1. *The verification of the identity (x) $a_n + b_{n+2} = 4a_{n+2}$ in Equation (2) of Section 4.2.3. For $n = 4$, we have $a[4] = 8, b[6] = 84, 4 * a[6] = 92$ and $a[4] + b[6] = 4 * a[4] = 92$. (This can be checked simply by hand from Table 5.) For larger values of n , a computer must be used. For $n = 100$ (taking a value for n that is not too large to save the place), one obtains*

$$\begin{aligned} a[100] + b[102] &= 10793987732357554298204 \\ 4 * a[102] &= 10793987732357554298204 \end{aligned} \tag{A8}$$

Example A2. *The verification of the identity $b_n + g_n = 6a_n$ in Equation (28). For $n = 9$, we have, from Table 5: $b[9] = 358, g[9] = 236, 6a[9] = 594, b[9] + g[9] = 358 + 236 = 6a[9] = 594$.*

Example A3. The verification of the identity $(v) c_n + 2b_{n-1} = b_{n+2}$ in Equation (2). The case $n = 7$, which was involved in Equation (14), gives immediately from Table 5: $c[7] = 190$, $2b[6] = 2 \times 84 = 168$, $b[9] = 358$ and $c[7] + 2b[6] = 190 + 168 = b[9] = 358$.

For $n = 150$, one obtains

$$\begin{aligned} c[150] + 2b[149] &= 274774599627602176762968441359741 \\ b[152] &= 274774599627602176762968441359741 \end{aligned} \quad (\text{A9})$$

Once the functions a_n , a'_n , b_n , c_n and g_n are written, one can use a simple built-in summation function for them to evaluate the various sums in the text, which all involve only low values of the index n . As an example, let us compute the two parts of Equation (10) of Section 4.2.1 and their sum. We have

$$\sum_{i=1}^8 a'[i] = 219, \quad a'[9] = 139, \quad \sum_{i=1}^8 a'[i] + a'[9] = 358 \quad (\text{A10})$$

References

- Nirenberg, M.; Leder, P.; Bernfield, M.; Brimacombe, R.; Trupin, J.; Rottman, F.; O'Neal, C.N.A. Codewords and Protein Synthesis, VII. On the General Nature of the RNA Code. *Proc. Natl. Acad. Sci. USA* **1965**, *53*, 1161–1168. [[CrossRef](#)] [[PubMed](#)]
- Inouye, M.; Takino, R.; Ishida, Y.; Inouye, K. Evolution of the genetic code; Evidence from codon use disparity in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 28572–28575. [[CrossRef](#)] [[PubMed](#)]
- Zwick, A.; Regier, J.C.; Zwickl, D. Resolving Discrepancy between Nucleotides and Amino Acids in Deep-Level Arthropod Phylogenomics: Differentiating Serine Codons in 21-Amino-Acid Models. *PLoS ONE* **2012**, *7*, e47450. [[CrossRef](#)] [[PubMed](#)]
- Sun, X.; Yang, Q.; Xia, X. An improved implementation of effective number of codons (N_c). *Mol. Biol. Evol.* **2013**, *30*, 191–196. [[CrossRef](#)] [[PubMed](#)]
- Négadi, T. The genetic code multiplet structure, in one number. *Symmetry Cult. Sci.* **2007**, *18*, 149–160. [[CrossRef](#)]
- Négadi, T. The Genetic Code via Gödel Encoding. *Open Phys. Chem. J.* **2008**, *2*, 1–5. [[CrossRef](#)]
- Négadi, T. The genetic code invariance: When Euler and Fibonacci meet 2014. *Symmetry Cult. Sci.* **2014**, *25*, 261–278.
- Rumer, Y. About systematization of the genetic code. *Dok. Akad. Nauk SSSR* **1966**, *167*, 1393–1394.
- Findley, G.I.; Findley, A.M.; McGlynn, S.P. Symmetry characteristics of the genetic code. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 7061–7065. [[CrossRef](#)] [[PubMed](#)]
- Rosandić, M.; Paar, V. Codons sextets with leading role of serine create “ideal” symmetry classification scheme of the genetic code. *Gene* **2014**, *543*, 45–52. [[CrossRef](#)] [[PubMed](#)]
- Rosandić, M.; Paar, V. The novel Ideal Symmetry Genetic Code table-Common purine-pyrimidine symmetry net for all RNA and DNA species. *J. Theor. Biol.* **2021**, *524*, 110748. [[CrossRef](#)] [[PubMed](#)]
- shCherbak, V. The Arithmetical origin of the genetic code. In *The Codes of Life: The Rules of Macroevolution*; Barbieri, M., Ed.; Springer Publishers: New York, NY, USA, 2008; pp. 153–185.
- shCherbak, V.; Makukov, M. The “wow! Signal” of the terrestrial genetic code. *Icarus* **2013**, *224*, 228–242. [[CrossRef](#)]
- Edge, M. Symmetry in Fibonacci numbers. *Symmetry Cult. Sci.* **2009**, *20*, 393–408.
- Rakočević, M.M. Genetic Code: The unity of the stereochemical determinism and pure chance. *arXiv* **2009**, arXiv:0904.1161v1.
- Shu, J.J. A new integrated symmetrical table for genetic codes. *Biosystems* **2017**, *151*, 21–26. [[CrossRef](#)] [[PubMed](#)]
- Lehmann, J. Physico-chemical constraints connected with the coding properties of the genetic system. *J. Theor. Biol.* **2000**, *202*, 129–144. [[CrossRef](#)] [[PubMed](#)]
- Gonzalez, D.L.; Giannerini, S.; Rosa, R. On the origin of the mitochondrial genetic code. Towards a unified mathematical framework for the management of genetic information. *Nat. Prec.* **2012**, *2012*, 1–20. [[CrossRef](#)]
- Available online: <https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?chapter=tgencodes#SG2> (accessed on 27 July 2023).
- Downes, A.M.; Richardson, B.J. Relationships between genomic base content and distribution of mass in coded proteins. *J. Mol. Evol.* **2002**, *55*, 476–490. [[CrossRef](#)] [[PubMed](#)]
- Available online: <https://www.dcode.fr/euler-totient> (accessed on 27 July 2023).

22. Available online: <https://t5k.org/glossary/page.php?sort=EulersTheorem> (accessed on 27 July 2023).
23. Available online: <https://www.programiz.com/python-programming/examples/fibonacci-sequence> (accessed on 27 July 2023).
24. Available online: <https://www.programiz.com/cpp-programming/examples/fibonacci-series> (accessed on 27 July 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.