Unification of Four Approaches to the Genetic Code

M. Pitkänen,

June 19, 2019
Email: matpitka6@gmail.com.
http://tgdtheory.com/public_html/.
Recent postal address: Rinnekatu 2-4 A 8, 03620, Karkkila, Finland.

Contents

Abstract

A proposal unifying four approaches to genetic code is discussed.

The first approach is introduced by myself and is geometric: genetic code is interpreted as an imbedding of the aminoacid space to DNA space possessing a fiber bundle like structure with DNAs coding for a given aminoacid forming a discrete fiber with a varying number of points. Also Khrennikov has proposed an analogous approach based on the identification of DNAs coding for a given aminoacid as an orbit a discrete flow defined by iteration of a map of DNA space to itself.

Much later (2014) I have introduced a variant of this scenario in which the fiber space structure is by assigning aminoacids to the 20 vertices of icosahedron. This model allows to understand the degeneracies of genetic code group theoretically.

Second approach starts from the 5-adic approach of Dragovich and Dragovich. Codons are labelled by 5-adic integers n which have no non-vanishing 5-digits so that the n is in the range $[31, 124]$. The number of primes in the range $[31, 124]$ is 20. This suggests the labelling of aminoacids by these primes. This inspires an additional condition on the geometric code: if possible, one of the integers n projected to p equals to $p(n)$. This condition fails only for the primes 53,79,101,103 for which some of 5-digits vanishing in 5-ary expansion.

The third approach relies on the generalization of the basic idea of the so called divisor code proposed by Khrennikov and Nilsson. The requirement is that the number of factors for integer n labelling one of DNAs, call it n_d coding for a given aminoacid is the total number of codons coding for the aminoacid, its degeneracy. Therefore a given aminoacid labelled by prime p with no non-vanishing 5-digits is coded by DNAs labelled by p itself and by n_d. A group theoretic and physical interpretation for the origin of the divisor code is proposed.

The fourth approach is a modification of the earlier 4-adic number theoretic thermodynamics approach of Pitkänen.

1. 5-adic thermodynamics involving a maximization of number theoretic negentropy $N_p(n) = -S_p(n) > 0(!)$ as a function of p-adic prime p labelling aminoacids assigns a unique prime to the codon. If no prime in the range divides S_p, the codon is identified as a stopping codon.
2. The number theoretic thermodynamics is assigned with the partitions P of the integer $n_{2)}$ determined by the first two letters of the codon (16 integers belonging to the range $[6, 24]$). The integer valued number theoretic Hamiltonian $h(P) \in Z_{25}$ appearing in the Boltzmann weight $5^{h(P)/T_5}$ is assumed to depend on the number r of summands for the partition only. $h(r)$ is assumed to be tailored by evolution so that it reproduces the code.
3. The effect of the third nucleotide is described in terms of 5-adic temperature $T_5 = 1/n$, $n \in [0, 24]$: the variation of T_5 explains the existence of variants of genetic code and its temporal variation the observed context sensitivity of the codon-aminoacid correspondence for some variants of the code.

A numerical calculation scanning over $N \sim 10^{30}$ candidates for $h(r)$ allows only 11 Hamiltonians and with single additional symmetry inspired condition there are 2 solutions which differ only for 5 largest values of r. Due to the limited computational resources available only 24 percent of the available candidates have been scanned and the naive expectation is that the total number of Hamiltonians is about about 45 unless one poses additional conditions.

The problem of the number theoretic models is that they do not predict but only reproduce. This is in sharp contrst to the model based on dark proton sequences, which leads to a radically new vision about the evolution of prebiotic life and to the vision about how immune system and genetic code evolved and what is the meaning of the genetic code.

1 Introduction

A proposal unifying four approaches to genetic code is discussed.

The first approach is introduced by myself and is geometric: genetic code is interpreted as an imbedding of the aminoacid space to DNA space possessing a fiber bundle like structure with DNAs coding for a given aminoacid forming a discrete fiber with a varying number of points. Also Khrennikov has proposed an analogous approach based on the identification of DNAs coding for a given aminoacid as an orbit a discrete flow defined by iteration of a map of DNA space to itself.

Much later (2014) I have introduced a variant of this scenario in which the fiber space structure is by assigning aminoacids to the 20 vertices of icosahedron. This model allows to understand the degeneracies of genetic code group theoretically.

Second approach starts from the 5-adic approach of Dragovich and Dragovich. Codons are labelled by 5-adic integers n which have no non-vanishing 5-digits so that the n is in the range $[31, 124]$. The number of primes in the range $[31, 124]$ is 20. This suggests the labelling of aminoacids by these primes. This inspires an additional condition on the geometric code: if possible, one of the integers n projected to p equals to $p(n)$. This condition fails only for the primes 53,79,101,103 for which some of 5-digits vanishing in 5-ary expansion.

The third approach relies on the generalization of the basic idea of the so called divisor code proposed by Khrennikov and Nilsson. The requirement is that the number of factors for integer n labelling one of DNAs, call it n_d coding for a given aminoacid is the total number of codons coding for the aminoacid, its degeneracy. Therefore a given aminoacid labelled by prime p with no non-vanishing 5-digits is coded by DNAs labelled by p itself and by n_d. A group theoretic and physical interpretation for the origin of the divisor code is proposed.

The fourth approach is a modification of the earlier 4-adic number theoretic thermodynamics approach of Pitkänen.

1. 5-adic thermodynamics involving a maximization of number theoretic negentropy $N_p(n) = -S_p(n) > 0(!)$ as a function of p-adic prime p labelling aminoacids assigns a unique prime to the codon. If no prime in the range divides S_p, the codon is identified as a stopping codon.

2. The number theoretic thermodynamics is assigned with the partitions P of the integer $n_{2)}$ determined by the first two letters of the codon (16 integers belonging to the range $[6, 24]$). The integer valued number theoretic Hamiltonian $h(P) \in Z_{25}$ appearing in the Boltzmann weight $5^{h(P)/T_5}$ is assumed to depend on the number r of summands for the partition only. $h(r)$ is assumed to be tailored by evolution so that it reproduces the code.

3. The effect of the third nucleotide is described in terms of 5-adic temperature $T_5 = 1/n$, $n \in [0, 24]$: the variation of T_5 explains the existence of variants of genetic code and its temporal variation the observed context sensitivity of the codon-aminoacid correspondence for some variants of the code.

A numerical calculation scanning over $N \sim 10^{30}$ candidates for $h(r)$ allows only 11 Hamiltonians and with single additional symmetry inspired condition there are 2 solutions which differ only for 5 largest values of r. Due to the limited computational resources available only 24 percent of the available candidates have been scanned and the naive expectation is that the total number of Hamiltonians is about about 45 unless one poses additional conditions.

The problem of the number theoretic models is that they do not predict but only reproduce. This is in sharp contrst to the model based on dark proton sequences, which leads to a radically new vision about the evolution of prebiotic life and to the vision about how immune system and genetic code evolved and what is the meaning of the genetic code.

The appendix of the book gives a summary about basic concepts of TGD with illustrations. There are concept maps about topics related to the contents of the chapter prepared using CMAP realized as html files. Links to all CMAP files can be found at `http://tgdtheory.fi/cmaphtml.html` [L5]. Pdf representation of same files serving as a kind of glossary can be found at `http://tgdtheory.fi/tgdglossary.pdf` [L6].

2 Unifying Various Approaches To The Genetic Code

The understanding of genetic at deeper level has gained increasing attention: mention only the proposals of Khrennikov [A5, A8], Pitkänen [K9, K8, K4], and Dragovich and Dragovich [A7]. Quite recently Khrennikov and Nilsson introduced the idea of divisor code [A9]. The idea is inspired by the observations that the numbers of divisors of integers in the range $[1, 20]$ are rather near to degeneracies of amino-acids for the genetic code. The attempts to realize this idea as such had however only a limited success and this led to a generalization of the basic idea of the divisor code and stimulated the attempt to combining four different approaches to the genetic code to single unified approach.

2.1 Geometric Approach To The Genetic Code

The geometric approach of Pitkänen [K9, K8, K4] was inspired by the basic hypothesis of TGD [L4, L3, L2] that space-times can be regarded as 4-surfaces $X^4 \subset H = M^4 \times CP_2$ of 8-dimensional imbedding space H. The idea was to replace H by the discrete space of integers labeling the 64 DNA triplets and X^4 by the discrete space of 20 amino-acids [K8]. Thus genetic code imbeds amino-acid space with points labeled by integers n_A to the DNA space labeled by some subset of integers (not necessarily $0 \leq n \leq 63$) such that the DNAs coding for a given amino-acid A form a discrete fiber like structure. One could also assume that one of the integers $n(DNA)$ labeling one of DNAs coding for A satisfies $n(DNA) = n(A)$ if possible.

As a matter fact, there exists the algebraic-geometric theory for codes based on the identification of code as a subset of subspace of G_p^k where G_p is finite field [A10]. If the points of this subset are labeled by some subset of integers m, the inclusion induces the code as a map $m \rightarrow n(m)$ where $n(m)$ consists of k G_p valued numbers. This concept of code does not apply to genetic code but the generalization is obvious: assign to the imbedding a bundle structure assigning to each point $n(m)$ a fiber consisting of points of G_p^k.

[A6] [A5] has proposed identification of codons coding for given amino-acid as an orbit of a discrete flow in the space of codons. It is possible to interpret DNA space as a bundle with fibers identified as orbits of the flow acting as a discrete group Z_n of symmetries in the fiber. The imbedding of amino-acid space to DNA space in the case of 5-adic code is however not quite equivalent with this view since four primes labeling amino-acids do not label codons.

2.2 4-Adicity And 5-Adicity As Possible Realizations Of The Symmetries Of The Genetic Code

An important physical constraint on any model is the fact that for the mitochondrial code codons have exact A-C and G-U symmetries with respect to the last codon. For eukaryote code this symmetry is broken only by two codons (Stop-Trp and Ile-Met pairs). A natural origin for this symmetry would be the formation of the 3-codons via fusion of 2-codons and 1-codons as suggested in the model of prebiotic evolution proposed in [K8].

One can consider two mathematical models for this symmetry.

1. 4-adic model of Pitkänen [K4] assumes the labeling of the codons using 4-adic numbers $n = n_0 + n_1 4 + n_3 16$, $n_i \in Z_4$ such that codons with $i = 0, 2$ and $1, 3$, which are 4-adically close to each other, correspond to symmetry related pairs. Also the model of Khrennikov and Kozyrev based on the identification of DNA space as 8×8 diadic plane (chess board!) starts from 4-adicity [A8] and interprets genetic code as a locally constant map from DNA space to amino-acid space. The number of primes $p < 64$ is 18 which leads to the idea that integers $n = 0, 1$ and the primes $p < 64$ code for amino-acids. Note however that 4-adicity as a strict symmetry needs to be assumed only for the third nucleotide.

2. For the 5-adic labeling of the codons suggested Dragovich and Dragovich [A7] codons are labeled by integers $n_0 + n_1 5 + n_2 5^2$ with $n_i \neq 0$ and vary in the range $[31, 124]$. The observation that the number of primes in this range is 20 inspires the hypothesis that that the primes in question label amino-acids. 5-adicity in the weakest sense means 5-adicity with respect to the third nucleotide so that either the codons $(n, n + 50)$ or codon pairs $(n, n + 25)$ and $(n, n + 75)$ code for the same amino-acid in the case of vertebrate mitochondrial code. There are three primes pairs $(p, p_1 = p + 50)$ [(47, 97), (53, 103), (59, 109)] so that $n \rightarrow n + 50$ symmetry is not consistent with the labeling of amino-acids by primes. Hence only $(n, n+25)$ and $(n, n + 75)$ option meaning that A-C and G-U pairs correspond to pairs of even and odd integers is acceptable and that the conjugation $n_3 \rightarrow 5 - n_3$ cannot correspond to DNA conjugation, which was the original motivation for the 5-adicity, but to the A↔C and G↔U symmetries.

2.3 Number Theoretical Thermodynamics And Genetic Code

The original thermodynamical model for the genetic code developed by Pitkänen [K4] is based on 4-adic labeling of codons. The model assumes that the number theoretical thermodynamics associated with the partitions of integers n labeling codons assigns to a given codon a unique prime labeling the amino-acid coded by DNA as the prime p for which the number theoretic negentropy $S_p = -\sum_k p_k log_p(|p_k|_p)log(p)$ is maximum: here $|x|_p$ denotes p-adic norm. S_p satisfies basic axioms of Shannon entropy but can be also negative so that its negative becomes a genuine measure of information [K11, K4]. Stopping codons would correspond to DNAs for which no prime in the allowed range of primes exists. A possible physical justification could be a breaking of conformal symmetry so that the states of given conformal weight $n = \sum n_i$ associated with the states $\prod L_{n_i}|n = 0\rangle$ are have different number theoretic "energies" depending only on the number r of integers n_i in the partition.

One can consider two variants of the number theoretical thermodynamics.

1. In the 4-adic case $n = 0$ and $n = 1$ amino-acids and codons correspond to DNAs labeled by same integers and are thus in a special role. The number theoretical thermodynamics [K4] is able to reproduce the genetic code and its variants by assuming that the integer valued Boltzmann weights of the thermodynamics are integers in a suitable range tailored by evolution in order to maximize the number theoretical negentropy. Boltzmann weights are assumed to be arbitrary integers in some range rather than powers of some prime so that genuine p-adic thermodynamics for some prime is not in question.

2. The 5-adic thermodynamics is favored by the fact that there are no special amino-acids now ($n = 0$ and $n = 1$). Preliminary calculations suggests that the 5-adic thermodynamics can be reduced to that for the 2-codons defined by the first two nucleotides labeled by integers $n_{2)} = n_0 + n_1 5$, $n_i \neq 0$ belonging to the range $[6, ..., 24]$. The integer valued Hamiltonian $h(P)$ for the thermodynamics of partitions P of $n_{2)}$ and defining Boltzmann weights $5^{h(P)}$ would depend only on the number r of summands in the partition P of n as $n = \sum_{k=1}^{r} n_k$. The dependence of the coded amino-acid on the third letter of the codon would be coded by the integer valued inverse of the 5-adic temperature $T_5 = 1/n$. A-C and G-U symmetries would correspond to the symmetry $T_5(r, k) = T_5(r, 5 - k)$ and the breaking of these symmetries would be due to the variation of temperature. The temporal variation of T_5 would explain the fact that for some variants of code same codon can code for either an amino-acid or stopping sign [K4], [I1].

2.4 Group Theoretic Interpretation Of The Divisor Code

The basic question is why the product decompositions of integer n characterizing one of the DNAs coding for a given amino-acid labeled by prime would determine the number of DNAs coding for the amino-acid. The original suggestion was that explanation is group theoretical. The fundamental role of discrete subgroups of rotation group in quantum TGD [K17, K7] suggests that finite subgroups $H \subset G$ of $G \subset SU(2)$ are involved with the code. Finite symmetry groups are indeed naturally associated with codes and the first observation is that product decompositions of integer n correspond naturally to the decompositions of an Abelian group G order n to products of subgroups with orders r and s, $n = r \times s$.

The hypothesis is that integer n characterizing the amino-acid corresponds to the order of G and that the factor pairs (r, s) of $n = rs$ correspond to its subgroups $H_r \times H_s \subset G$. The codons coding for amino-acid characterized by n would correspond to a normal sub-groups of G in general case and to any subgroup in the Abelian case. The simplest identification of G is as the cyclic group Z_n. That the product decompositions (r, s) and (s, r), $r \times s = n$ must be counted as separate can be understood if a wave function invariant under $Z_r = Z_n/Z_s$ characterizes the codon labeled by (r, s). Z_n would naturally act as a symmetry group in the discrete fiber of the fiber bundle defined by the DNA space and defining a discrete flow in the fiber. The p-adic prime p assigned to the amino-acid could in turn characterize the p-adicity of corresponding space-time sheet [K18].

The physical interpretation suggested by TGD and to be discussed later is that the wave functions of (say) free electron pairs (possibly Cooper pairs) defined in the set of points defined by the orbit of $Z_n \subset G_a$ are invariant under the subgroup of $Z_r = Z_n/Z_s \subset Z_n$ for DNA labeled by (r, s), $r \times s = n$. Thus the codons coding for an amino-acid having Z_n as a symmetry group would be characterized by wave functions for free electron pairs transforming under representations of Z_n and remaining invariant under $Z_r \subset Z_n$ and thus reducing to representations of Z_s. Note that $r = 1$ corresponds to all irreps of Z_n and $r = n$ to singlets under Z_n.

2.5 Divisor Code

The idea of divisor code discussed in [A9] is inspired by the following observations.

1. Consider the number $N(n)$ of integer divisors for integers n in the range $[1, 21]$ corresponding to amino-acids with stopping sign counted as amino-acid.

2. Denote the number of integers $n \leq 21$ for which the number of divisors is k by $B(k)$. Also stopping sign is counted as an amino-acid and $n = 0$ corresponds to amino-acid also. This number $N(k)$ varies in the range $[1, 6]$. $B(k)$ has the values $(1, 8, 2, 5, 1, 3)$ where k runs from 1 to 6.

3. Denote by $A(k)$ the number of amino-acids coded by k DNA codons. $A(k)$ has the values $2, 9, 2, 5, 0, 3$.

The spectrum of $A(k)$ is very similar to that of $B(k)$ and this raises the question whether one could understand genetic code as a divisor code in the sense that the degeneracy of amino-acid would be dictated by the number of the integers $1 \leq n \leq 21$ coding it. One might also ask whether the amino-acids which are abundant and thus important are coded by integers with a large number of divisors. Also one can ask whether the divisor structure possibly correlates with the structure of the amino-acid.

Divisor code in this form would be only approximate and one can wonder could try to imagine some simple symmetry breaking mechanism. In this respect the crucial observations might be following.

1. The number of DNAs needed to realize divisor code would be 70 instead of 64. One must drop 6 codons and by choosing them suitably one might hope of getting correct degeneracies.

2. The most natural manner to break the symmetry is to drop the 4 codons from the codons coding for 5-plet which would thus become 1-plet. 5-plet corresponds to integer $n = 16$ and its product compositions $(16, 1), (1, 16), (2, 8), (8, 2), (4, 4)$ correspond to the DNAs coding for it. $(4, 4)$ would naturally correspond to singlet.

3. By dropping 2 codons from some 4-plet one obtains 2-plet and correct degeneracies. One candidate for 4-plet corresponds to $n = 8$ and its product decompositions $(1, 8), (8, 1), (2, 4), (4, 2)$. By dropping two of these one obtains correct degeneracies. It might that power of 2 property of $n = 8$ and $n = 16$ somehow relates to 2-adicity and to the special role of these amino-acids.

4. A possible interpretation is in terms of symmetry based on cyclic group $Z(n)$ serving as a symmetry of DNA codons coding for amino-acid labeled n. Z_n allows decompositions $Z_n = Z_{n_1} \times Z_{n_2}$, $n = n_1 \times n_2$ and if the representations are invariant under Z_{n_2} and thus reduce to those of Z_{n_1} codons coding for a given amino-acid correspond to the product decompositions. Symmetry breaking would be due to the lacking 6 codons and would mean that only Z_4 invariant states would be realized for Z_{16} and Z_1 and Z_8 of Z_2 and Z_4 invariant states are realized for $n = 8$. $n = 4$ could correspond to triplet of stopping codons so that powers of 2 would be in special role for vertebrate code suggesting 4-adicity. 4-adicity is also suggests by the almost exact A-G and T-C symmetries of the last nucleotide.

2.6 Topological Interpretation Of The Divisor Code In TGD Framework

The most concrete physical interpretation of the divisor code found in TGD framework is topological and based on TGD inspired vision about the role of dark matter in biology.

1. The generalized 8-D imbedding space has a book like structure with pages glued together along back which is 4-D surface of $H = M^4 \times CP_2$ [K7, K14]. Particles at different pages are dark relative to each other since they cannot have local interactions (appear in the same vertex of Feynman diagram). The pages are partially characterized by the value of Planck constant which can be arbitrary large. This explains the macroscopic quantum coherence of living matter. Matter can leak between different pages meaning a phase transition changing Planck constant.

2. The notion of magnetic body with flux tubes carrying dark matter and connecting different bio-molecules central for the TGD inspired model of living matter [K6]. Magnetic bodies of bio-molecules can be also connected by magnetic flux tubes, even those in different pages of the book. For instance, the phase transition reducing $\hbar$ reduces the distance between two bio-molecules connected in this manner and forces them near to each other. This explains the extreme selectivity of bio-catalysis and the miraculous ability of two bio-molecules to find each other in the dense soup of bio-molecules. In particular, DNA and its conjugate codons, mRNA codons, and tRNA would be connected by this kind of flux tubes. Also amino-acids would be connected to tRNA codons in this manner since tRNA molecules catch the amino-acids and bring them to the mRNA-amino-acid translation site. Genetic code could reduce to the selection rules for the flux tube connections connecting in general situation magnetic bodies belonging to different pages of the book.

3. The pages of book are almost copies of $M^4 \times CP_2$. This means that M^4 is replaced with n_a-fold singular covering and CP_2 with n_b-fold singular covering. The coverings have cyclic groups Z_{n_a} and Z_{n_b} act as discrete symmetries for the wave functions of particles in the covering. A given page is thus labeled by two pager numbers (n_a, n_b). Two pages contain common points and thus a direct tunnelling of 3-surfaces between these pages is possible only if the number n_{a_1} of the sheets of covering divides n_{a_2} or vice versa. Same holds true for n_{b_1} and n_{b_2}. This rule is just the basic rule about how symmetries of system can change in phase transition. This number theoretic rule could be behind genetic code and the extreme selectivity of bio-catalysis.

4. Suppose that both bio-molecules correspond to ordinary matter with $n_a = n_b = 1$ but that the magnetic body of a given amino-acid corresponds to $(n_a(A), n_b(A))$ and DNA, RNA, and tRNA codon to $(r_a(DNA), r_b(DNA))$. Since the flux tube from tRNA codon to the amino-acid page is essential for the process in which amino-acid is attached to tRNA, only tRNA for with $r_a(tRNA)$ divides $n_a(A)$ can catch an amino-acid labeled by n_a. Same applies to r_b and n_b.

5. Without the presence of the integer n_b the code would fail since DNA codon labeled by r_a would code for all amino-acids for which n_a has r_a as a factor. n_b can indeed save the situation. Suppose that one has $r_b(tRNA) = n_b(A)$ if DNA codes for an amino-acid. Assume also that $n_b(a)$ is prime: $n_b(A) = p_b(A)$, and different for each amino-acid. This prime does not correspond to p-adic prime, which is expected to be very large in the length scales of atomic physics (electron corresponds to $M_{127} = 2^{127} - 1$). Note that the assumption that amino-acids are labeled by small primes was made in both TGD inspired number theoretical models of the genetic code.

6. The assumptions mean that tRNA and amino-acid can be connected by a magnetic flux tube only if one has

$$p_b(tRNA) = p_b(A)$$

and $r_a(tRNA)$ divides $n_a(A)$. If the pages numbers n_a vary in the range $[1, 21]$ the divisor code follows from the argument of the previous section. Taking the previous argument seriously, one should also understand why there is no amino-acid labeled by $n_a = 4$ and why corresponding DNAs correspond to prime characterizing $n_a = 4$, why the number of DNA codons labeled by the factors of $n_a = 8$ is two, and why the number of codons associated with $n_a = 16$ only one.

Some further comments are in order.

1. The realization of the genetic code is not unique since the integers r_a and n_a could be replaced with Nn_a, where N is a product of primes larger than $p = 19$. It is also enough that the integers characterizing amino-acids are relative primes (have not common factors). The simplest assumption would be that the primes $p(A)$ satisfy $p(A) > 19$ so that $p(A)$ does not divide $n(A)$ for any A. If $p(A)$ is as small as possible the value spectrum of $p(A)$ is

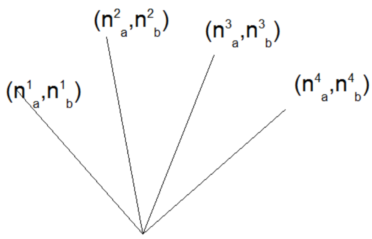

Figure 1: Illustration of the book-like structure of the generalized imbedding space.

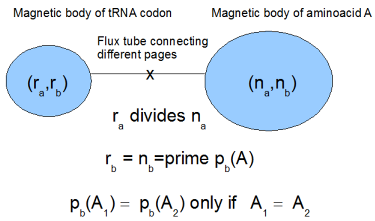

Figure 2: Illustration of the selection rules for magnetic flux tubes connecting magnetic bodies of tRNA and amino-acid.

$$\{23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109\} \ .$$

If one assumes that the two additional amino-acids coded in some cases by non-vertebrate genetic code correspond to primes also the primes $113, 127$ are included.

What is interesting is that Mersenne prime $M_7 = 2^7 - 1 = 127$ appears in the model of genetic code based on the notion of Combinatorial Hierarchy [K9]. This model assumes that DNA codons correspond to 64 integers in the range $1, ..., 127$. This realization of the genetic code cannot however be consistent with the divisor code realized in the proposed manner since it would require that the integers $n(A)p(A)$ belong to the range $1, .., 127$. The prime factors of these integers can however belong to this range.

2. The model in principle allows an infinite number of analogous codes and an interesting question is whether the bio-catalysis involves this kind of codes. The quantum antenna model for remote replication discussed in [K10] allows a dynamical interpretation for the flux tube realization of the genetic code as a divisor code in terms of quantum antenna hypothesis [K13], and predicts that sequences of DNA codons serve as names for polar molecules quite generally so that genetic code would define a universal language in living matter. This leads to an identification of the basic mechanism responsible for the functioning and evolution of the immune system.

3. The quantum states of dark baryons realize vertebrate genetic code with very general assumptions group theoretically [L1, K10, K15], [L1]. Since dark matter is involved in both cases, one might wonder whether these codes could be related somehow. A one-one correspondence between the quantum states of dark nucleons representing codon and the integers r_a, p_b is required in order to have this connection. The simplest possibility is that that energy minimization implies that given dark nucleon resides with high probability at flux tube labeled by unique value of r_a. Same applies to amino-acids.

The number theoretical model discussed in this chapter emerged before the topological explanation of the divisor code. Hence the model becomes somewhat obsolete. In particular, it involves un-necessarily strong assumptions. p-Adic thermodynamics might be used to understand the possible equivalence of the divisor code and the dark baryon code discussed in [K15] but this problem will not be discussed here.

2.7 Is The Fusion Of Geometric, Thermodynamical, And Divisor Code Approaches Possible In The 5-Adic Case?

A very attractive general idea is that genetic code could be understood in two dual manners: as an assignment $n \rightarrow p(n)$ and as an assignment $p \rightarrow n(p)$.

1. Genetic code could be understood in terms of a 5-adic thermodynamics for the partitions of integers characterizing codons. Here $6 \leq n_{2)} = n_0 + n_1 5 \leq 24$, $n_k \neq 0$, labels the 2-codons formed by the first two letters of the codon. This approach would predict the assignment $n \rightarrow p(n)$ once the number theoretic thermodynamics is specified.

2. Genetic code could be understood as a geometric imbedding $p \rightarrow n(p)$ of amino-acid space labeled by 20 primes $31 \leq p < 124$ to DNA space such that one has $n(p) = p$ if possible. This cannot the case for 4 primes ($p = 53, 79, 101, 103$). Also the interpretation as an induction of number theoretical bundle structure over amino-acid base space from DNA space is possible. $n(p) = p$ constraint obviously poses strong constraints on the model but it turns out that it is possible to satisfy these constraints for other than exceptional primes.

3. Also the basic idea of the divisor code could be included to the model via the condition that the number of divisors of the integer $n_{2)}$ for one of the DNAs coding for a given amino-acid equals to the number of DNAs coding for the amino-acid. There would be thus *two* labelings of amino-acids so that the model would become highly predictive.

The natural starting point is the vertebrate mitochondrial code with full $A \leftrightarrow C$ and $G \leftrightarrow U$ symmetries and one could interpret the breaking of these symmetries in the case of eukaryote code in terms of the context sensitivity characterized by the number theoretic temperature T_5. The large number of constraints raises the hope that a rather unique code could result. It will be found that for the number theoretic Hamiltonian depending only on the number partitions r of the integer $n_{2)}$ characterizing the first two letters of the 5-adic codon, only 4 solutions to the conditions can be found in the set of $N \sim 10^{30}$ candidates for $h(r)$.

3 5-Adicity Or 4-Adicity?

It seems that 5-adic representation of $A - C$ and $T - G$ symmetries allows the unification of the geometric view about genetic code with the number theoretic thermodynamics view and the idea of the divisor code.

3.1 The Problems Of The 4-Adic Model Of The Divisor Code

The 4-adic model for the divisor code has some problems.

1. 4-adic model is not consistent with the assumption that the set of DNAs coding for given amino-acid contains both the integer characterizing the degeneracy of the amino-acid as a number of its divisors and the codon labeled by the prime labeling the amino-acid. Hence

the geometric realization must be given up unless one assumes that the primes associated with amino-acids associated with columns not containing primes are mapped to the integers in the columns by imbedding map. Even this option fails.

2. It is not easy to understand the emergence of singlets without assuming breaking of the number theoretical symmetries.

3. The proposed TGD inspired topological interpretation of the divisor code is not consistent with the presence of $n = 0$ codons. Also $n = 1$ codons are problematic.

4. There is no obvious connection with the maximization of the number theoretic negentropy assigning primes to amino-acids. 5-adic thermodynamics can do this and one could have dual descriptions. Geometric description in terms of imbedding of amino-acid space to DNA space (assigning DNAs to amino-acids) and thermodynamics description in terms of 5-adic thermodynamics assigning amino-acids to DNAs.

3.2 5-Adic Model Works For Thermodynamics Based On Partitions

5-adic variant of the model can overcome the problems of the 4-adic model.

3.2.1 Basic assumptions

1. Stopping codons do not correspond to formal amino-acids. The natural hypothesis is that the stopping codons do not possess negentropy maximizing prime in the range considered.

2. The question is whether conjugation $k \to 5-k$ for the last nucleotide corresponds 1) to DNA conjugation as in [A7] or 2) to a symmetry of the last codon. The naive guess would be 1). The guess turns out to be wrong since it implies that 3 4-plets contain symmetry related primes so that the number of amino-acids would be reduced by 3 due to the $n \to n + 50$ symmetry of the last nucleotide. On the other hand, $k \to 5-k$ as a representation of $A \leftrightarrow C$ and $G \leftrightarrow U$ symmetries takes odd integers to even integers so that there are no problems.

3. DNA codons correspond to 5-adic integers in the range [31, 124] having no vanishing 5-digits. Amino-acids are labeled by the 20 primes in the same range. They are mapped to DNA triplets. For 16 primes this imbedding is just the identification $n(p) = p$. The 4 "outsider" primes $53, 79, 101, 103$, which have a vanishing 5-digit, have necessarily $n(p) \neq p$. The first guess is that the outsider primes $53, 79, 101, 103$ correspond to amino-acids that are somehow special. It turns out that a possible identification for the amino-acids is as Trp, Lys, Met, Gln but that Lys, Gln pair can be replaced by any pair in the set $\{Gln, Lys, Glu\}$. One could also argue that the amino-acids corresponding to 53 and $103 = 53 + 50$ should be related by some kind of symmetry. Trp and Met indeed have the comment feature that a codon coding for them can also act as stopping codon. On the other hand, also Lys, Gln, and Glu share the property of being polar amino-acids.

3.2.2 Further constraints

The observation that there are two 4-columns containing no primes when combined with some facts about the genetic code and its variants give strong constraints on the code.

1. One of the prime-free columns must correspond to shared Ser-Arg column which transforms to Ser-Stop column for mitochondrial code. Otherwise one prime coding for an amino-acid would be lost.

2. In the case of the yeast mitochondria Thr is coded 8 times and Leu only twice. This forces the conclusion that second prime-free 4-column corresponds to Leu.

3. Since Leu must be coded by prime, Leu-Phe 4-column must correspond to the second 4-plet containing two primes. Hence the two 4-columns containing 2 primes give rise to three doublets. 6 additional doublets for eukaryote code and 9 additional doublets for mitochondrial code must be identified.

4. Thr 4-plet should contain n possessing 8 divisors. Only 3 4-columns contain $n = 8$ and correspond to 321, 131, and 231 columns.

3.2.3 Detailed identification of the code

Consider now a more detailed identification of the code.

1. Mitochondrial code is obtained as follows. 4 outsider primes which do not label DNAs directly are imbedded into 4-columns containing single prime. This gives 8 doublets altogether. Stopping codons in the 4-column containing Tyr and corresponding prime give one additional doublet so that a correct number of doublets result.

2. The breaking of the mitochondrial code to eukaryote code is easy to understand in the proposed framework. Trp and Met become singlets and Ile becomes triplet so that 9 doublets result.

3. Outsider primes would in this model correspond to Gln, Lys, Trp, Met. Gln and Lys could be replaced with any pair in the set {Gln, Lys, Glu} for the simple reason that corresponding amino-acid doublets cannot be distinguished from each other number theoretically. The identifications of the integers associated with amino-acids coded by 4 entire 4-column (Val, Ala, Pro, Gly) are unique apart from $4! = 24$ permutations of these amino-acids. It should be noticed that Lys, Gln, Glu belong to the group of 11 polar amino-acids and Met and Trp belong to the group of 8 hydrophobic amino-acids.

4. The multiplet containing Met is unique since there is only single codon ($n = 11^2 = 121$) for which the number of divisors is 3.

5. One can say that Ile and Met compete: either Ile^3-Met results when Ile wins. Ile^2-Met^2 results when Met wins. One can argue that Trp as outsider prime can also correspond to singlet or that Stop can "eat" any any prime and reduce the degeneracy. 5-adicity is broken for the first two nucleotides, which is not surprising.

These number theoretic constraints do not allow a unique identification of the code but pose considerable restrictions. **Table 1** represents one example consistent with these conditions. Note that the table does not fix how the primes 53, 79, 101, and 103 are assigned to Trp, Lys, Met, and Gln. Trp and Met are indeed special since they can be replaced by stopping codon some variants of the code.

It will be found that under rather general conditions (roughly 10^{30} candidates for the Hamiltonian $h(r)$ characterizing the thermodynamics of partitions) there are only 4 choices of $h(r)$ reproducing the eukaryote code, vertebrate mitochondrial code as well as other variations of the code. If one requires that the polar amino-acids Lys and Gln (or any pair in the set {Gln, Lys, Glu}) correspond to the conjugation related primes 53 and 103 only single solution for $h(r)$ is found. The 5-adic thermodynamics based on spin-spin interaction fails as do also other simple models.

(114,106,4,UG)	(214,107,2,GU)	(314,108,GC)	(414,109,2,GA)
(113,81) Trp	(21,82,4) Val	(313,83,2, Ala)	(413,84, Glu)
(112,56,4,Cys)	(212,57,4)	(312,58)	(412,59,2,Asp)
(111,31,2)	(211,32)	(311,33,4)	(411,34,4)
(124,111,4,GG)	(224,112,UA)	(324,113,2,AC)	(424,114,6,CG)
(123,86,4,Gly)	(223,87,4,Stop)	(323,88,8,Thr)	(423,89,2,Arg)
(122,61,2)	(222,62,4,Tyr)	(322,63,4)	(422,64)
(121,36)	(221,37,2)	(321,38,4)	(421,39,4)
(134,116,6,CC)	(234,117,6,UC)	(334,118,4,AA)	(434,119,4,AG)
(133,91,4,Pro)	(233,92,6,Ser)	(333,93,4,Lys)	(433,94,4,Arg)
(132,66,8)	(232,67,2)	(332,68,6,Asn)	(432,69,4,Ser)
(131,41,2)	(231,42,8)	(331,43,2)	(431,44,6)
(144,121,3,AU)	(244,122,4,UU)	(344,123,4,CA)	(444,124,6,CU)
(143,96,Met)	(243,97,2,Leu)	(343,98,6,Gln)	(443,99,6,Leu)
(142,71,2,Ile)	(242,72,Phe)	(342,73,2,His)	(442,74,4)
(141,46,4)	(241,47,2)	(341,48)	(441,49,3)

Table 1: An example of a code obeying approximate 5-adic symmetry $k \leftrightarrow 5-k$ with respect to the last codon. Given are the integers associated with the codons of given 4-column in 5-adic and decimal notion, the number of divisors appearing if it belongs to the range of allowed values, and the 2-codon associated with the 4-column. Note that 5-adic symmetry for the first two nucleotides is broken.

4 5-Adic Thermodynamical Model For The Genetic Code

The challenge is to guess the number theoretic Hamiltonian characterizing the thermodynamical model and the dependence of the 5-adic temperature T_5 on third nucleotide describing the splitting of 4-plets to doublets and further splitting of the doublets in the case of eukaryote code. There are two options concerning the choice of the Hamiltonian.

1. The Hamiltonian depends only on the number r of integers in the partition $n_{2)} = \sum n_k$ of $6 \leq n \leq 24$ of integer $n_{2)} = n_0 + n_1 5$ characterizing the first two nucleotides of the codon. Hamiltonian is tailored by evolution to reproduce the genetic code and its variants.

2. Hamiltonian is a direct analog of spin spin interaction $J \sum n_k n_l$ with n_k interpreted as spin associated with n_k Cooper pairs.

4.1 The Simplest Model For The 5-Adic Temperature

The simplest model for 5-adic temperature applies irrespective of the number theoretic Hamiltonian h and relies on the assumption inspired by the comparison of the mitochondrial and eukaryote code tables.

1. $T_5(n_3) = T_5$ hold true for common 4-plets, 4-plet parts of 6-plets, and 6-plets of the mitochondrial and eukaryote codes.

2. $T_5(n_3) = T_5(5-n_3)$ holds true for common 2-plets (A-C and T-G symmetries with respect to the third nucleotide) of eukaryote and mitochondrial code and for all 2-plets of mitochondrial code.

3. For eukaryote code this symmetry of 5-adic temperature would fail for Ile^3-Met, Cys^2-Stop-Trp and only for the second pair of values of n_3 corresponding to Met-Met$\rightarrow$ Ile-Met and Trp-Trp $\rightarrow$ Ttop-Trp $[n_3, 5-n_3) = (2,3)]$. Ser-Stop-Ser-Stop to Ser-Arg-Ser-Arg transition

would in turn be induced by the change of 5-adic temperature. Stop would correspond to a 5-adic temperature for which no prime coding amino-acid divides the partition function.

The condition that the model reproduces correctly the $n \rightarrow p(n)$ correspondence to be discussed later in principle allows to fix number theoretic Hamilton and $T_5(n_3)$ to a high degree.

4.2 The Simplest Possible Model For Thermodynamics

Before dwelling into complex calculations it is useful to ask what could be the simplest model for the 5-adic thermodynamics.

1. Calculational simplicity would suggest that the partition function must be as small as possible and thus satisfy $Z(n) < 125$. This restriction also maximizes the probability that the prime divisors are in the range $31 \leq p \leq 113$ with stopping codons involving only divisors $p < 31$. This together with the 5-adicity at the level of partition function would suggest that the definition of $Z(n)$ should involve 5-adic cutoff in the form $Z(n) \rightarrow Z(n)\ mod\ 5^3$. The natural constraint on the values h of the number theoretical Hamiltonian would thus be $h \in \{0,1,2\} \in Z_3$. Modulo three arithmetics fits also nicely with the triplet structure of codons.

2. In this model the effect of changing 5-adic temperature form $T_5 = 1$ to $T_5 = 1/n$, $n = 1,2$ would be expressed as $h(r) \rightarrow n \times h(r)$. Only two possible 5-adic temperatures would be possible and the symmetries of the vertebrate mitochondrial code would be predicted automatically. The symmetry breaking down to eukaryote code could be described in terms of 5-adic temperature if one allows formally infinite temperature for which one would have effectively $h(r) \rightarrow h(r) = 0$ so that partition function equivalent with $Z = 1$ would result and the codon in question would code for stopping sign. This is indeed the case for the codon coding originally Trp. For the breaking of Ile-Met doublet the splitting to triplet and singlet can be also understood as the dependence of T_5 on codon in symmetry breaking manner.

3. The simplest possible model would correspond to $Z(n) = p(n) = \sum p_k 5^k$ so that p_k would have interpretation as degeneracies of states modulo 5: this would imply that the doublets would correspond to primes related by exchange of p_1 and p_2, which does not make sense. Hence the integers p_k cannot directly correspond to the degeneracies of states with different energies and the partition function must be obtained via $Z \rightarrow Z\ \ mod\ 125$ prescription from a more complex partition function having values $Z > 125$. The three digits p_k for 5-adic code and Z_3 valuedness of $h(r)$ might relate naturally to 3-letter structure of codons.For $n = p(n)$ one would simply have $Z(n) = n = p(n)$. For the four exceptional amino-acid primes $p = 53, 79, 101, 103$ this would not hold true. The most general model would allow small integer $k \leq 4$ as an additional factor of $Z(n) \leq 124$.

Unfortunately, this simple model does not allow any obvious number theoretical realization. In particular, the models based thermodynamics of partitions and on spin-spin interaction fail with Z_3 valued $h(r)$ and Z_{125} valued $Z(n)$. The simplicity and explanatory power of the model encourage however to keep mind open for the existence of this kind of model.

4.3 Number Theoretic Hamilton Depending On The Number OfPartitions Of Integer Characterizing DNA

The number theoretic model for the genetic code discussed in [K4] was based on the assumption that the number theoretic Hamiltonian depends only on the number of summands in the partition $n = \sum_k n_k$.

Generalizing to the recent context, the Hamiltonian $h(r)$ for the 5-adic thermodynamics should depend only on the number r of summands in the partition $n_{2)} = \sum_{k=1}^{r} n_k$. The deviations from the standard code would be explained in terms of the variation 5-adic temperature which has values $T = 1/n$, n positive integer, implying Boltzmann weights $5^{h(r)/T_5}$. The fact that same codon does not always code same amino-acid [K8] , [I1], could be understood in terms of temporal variation of 5-adic temperature. A possible interpretation is in terms of a breaking of conformal invariance characterized completely the number r of subsets in the partition.

A further assumption motivated by 5-adicity is the replacement $X \equiv h(r)/T_5$ in Boltzmann weight with $X \ mod \ N$, where N characterizes the highest power of 5 appearing in partition function. $N = 3$ would be the minimal option but it turns that only $N = 25$ works. It will be assumed that evolution has gradually tailored $h(r)$ so that the observed genetic code maximizes for a given DNA the p-adic information measure defined by the prime $p(DNA)$ coding the corresponding amino-acid in practice this means that partition function is divisible by a power of $p(DNA)$.

The interpretation in terms of the number of sub-condensates of Cooper pairs containing n_k spin 1 Cooper pairs is an alternative interpretation and would look attractive physically but in this case the Hamilton depending on the number r of partitions only does no look natural. The number theoretic Hamiltonian would depend on the number r of bound states only if the interaction energy $E(n_k, n_l)$ between two sub-condensates with n_k and n_l Cooper pairs is a constant integer $E(n_k, n_l) = E$, so that the interaction energy between sub-condensates would behave as $r(r-1)E \ mod \ N$. This could give rise to a rather random looking behavior of $h(r)$ as a function of r. The modulo arithmetic constraint would restrict considerably the number of choices of $h(r)$. This model does not reproduce realistic genetic code.

4.3.1 Formula for the partition function

The formula for the partition function is given as

$$\begin{aligned} Z &= \sum_r d(n,r)5^{H(r)} \ , \\ H(r) &= \frac{h(r)}{T_5} \ mod \ 25 \ . \end{aligned} \tag{4.1}$$

$T_5 = 1/n$ varies in the range $n \in [1, 24]$.

The partition numbers appearing in are conveniently calculated by using the recurrence relation [A2]

$$d(n,r) = P(n,r) = P(n-1,r-1) + P(n-r,r) \ , \ P(n,1) = 1 \ . \tag{4.2}$$

4.3.2 The structure of the calculation

The flow of calculation proceeds along the rows of the code table as given in **Table ??** coding for the constraints coming from the assumption that the number of divisors for of the integers labeling DNAs is same as the degeneracy of corresponding amino-acid and from the consistency with the geometric model of the code.

1. It is assumed $0 \leq h(r) \leq h_{max} = 2$ for $r > 1$. $h(1) = 0$ can be assumed without a loss of generality if one assumes that $r = 1$ (trivial partition) corresponds to the most probable minimum energy partition in the sense of 5-adic thermodynamics. This implies that 3^{23} candidates for $h(r)$ must be scanned. All possible $4! = 24$ assignments of Trp, Lys, Met, Gln with the primes $p = 53, 101, 79, 103$ which do not label codons are considered.

2. At the first step those guesses for $h(r)$, $r \leq 6$, for which the DNA-Cys correspondence with $p(Cys) = 31$ is reproduced and stored.

3. At the next step calculation branches to four separate calculations corresponding to the four possible values of $p(Trp) \in \{53, 101, 79, 103\}$. 5-adic temperature T_5 is varied and it is found whether the $p(Trp)$ can be reproduced for some value of $T_5 \in \{1, 2, ..., 24\}$. If this is not possible, the candidate for $h(r)$, $r \leq 6$ is rejected. After this the calculation proceeds for given $p(Trp)$ assignment through next values of $h(r)$ to $r = 18$ where one checks whether $p(Asn) = 43$ can be reproduced. In the transitions to new row corresponding to $r = 10, 11$ and $r = 15, 16$ two values of of $0 \leq h(r) \leq 2$ appear and bring in additional degrees of freedom. In $Glu - Asp$ column at the end of the first row T_5 is varied to see whether also $p(Asp) = 59$ can be reproduced.

(1, 1) tlmg	(1, 2) tglm	(1, 3) tmgl	(1, 4) tlgm	(1, 5) tgml	(1, 6) tmlg
(2, 1) ltmg	(2, 2) gtlm	(2, 3) mtgl	(2, 4) ltgm	(2, 5) gtml	(2, 6) mtlg
(3, 1) lmtg	(3, 2) mgtl	(3, 3) gltm	(3, 4) lgtm	(3, 5) gmtl	(3, 6) mltg
(4, 1) lmgt	(4, 2) glmt	(4, 3) mglt	(4, 4) lgmt	(4, 5) gmlt	(4, 6) mlgt

Table 2: There are 24 different solution types depending on which permutation $xyzu$ of (Trp, Lys, Met, Gln) corresponds to the exceptional primes $(53, 79, 101, 103)$. For instance, lmtg means $(Lys, Met, Trp, Gln) \rightarrow (53, 79, 101, 103)$, and tglm means $(Trp, Gln, Lys, Met) \rightarrow (53, 79, 101, 103)$. It is convenient to label the 24 possibilities by pairs of integers (m, n). $m = 1, 2, 3, 4$ according to whether Trp, Lys, Met or Gln corresponds to $p = 53$. The second integer $n = 1, ..., 6$ specifies which of the six permutations of remaining three amino-acids corresponds to $(79, 101, 103)$ in a manner expressed by the table. For instance, for $(m, n) = (1, 1) \leftrightarrow (tlmg)$ codes for $(Trp, Lys, Gln, Met) \rightarrow (53, 79, 101, 103)$.

4. After this the calculation for given value of $p(Trp)$ branches to 6 alternatives corresponding to different assignments of remaining exceptional primes to Lys, Met, Gln. Since Arg-Ser four-column does not give any conditions the values of $h(r)$ for $r = 19, 20, 21$ appear as free parameters. This part of calculation is especially critical since the first 4-columns of the last row of the table contain only doublets. The last 4-column (Leu) corresponding to $r = 24$ does not pose any conditions on $h(24)$ unless one requires that also $n = 49$ gives partition function for $p(Leu) = 97$ is the maximizing prime.

4.3.3 Results

The difficulties involved with the numerical computation were considerable since only University MATLAB was available and for the extensive computations involved its functioning turned out to be somewhat unreliable and reasons for this could not be identified. 22 solutions to the conditions expressed in Table 2 has been found from the set of about 10^{30} candidates, and have been checked separately to satisfy all the conditions.

The 11 number theoretic Hamiltonians $h(r)$ for $r = 1, 2, ..., 23$ are given in**Table 3** with conventions expressed in **Table ??**

One can consider additional symmetry assumptions reducing the number of solutions.

1. One might argue that the "unstable" amino-acids Trp and Met naturally correspond to the conjugation related primes 53 and 103. The are only 2 solutions (h_1 and h_2 in Table 3) corresponding to the assignment $(Trp, Met) \rightarrow (53, 103)$ or vice versa (the integer pairs (m, n) corresponding to txym and mxyt in Table 2 are (1, 2), (1, 4), (4, 3), (4, 6)). These two solutions differ only for last 5 values of r.

2. One might also argue that the polar amino-acids Lys and Gln (or any pair in the set $\{Lys, Gln, Glu\}$) correspond to the conjugation related primes 53 and 103 (the integer pairs (m, n) corresponding to lxyg and gxyl in Table 2). There are 3 solutions (h_6, h_7 and h_8 in Table 3) corresponding to the assignment $(Lys, Gln) \rightarrow (53, 103) or vice versa$ (the integer pairs (m, n) corresponding to txym and mxyt in Table 2 are (2, 1), (2, 4), (3, 1), (3, 4)).

That not too many solutions exist to the conditions together with the fact that the model is consistent with the basic ideas of geometric code and of divisor code and results from 5-adic thermodynamics, raises the hope that something more than a mere complex parameterization of the genetic code might be in question. For $r = 2$ $h(r)$ only the values $h(r) \leq 5$ have been scanned (the reasons were the strange problems that made the continuation of calculations very difficult) so that a portion $6/25 = 24$ per cent of all possible candidates for $h(r)$ are scanned. The number

m	1	1	1	1	2	3	3	3	3	3	4
n	2	2	5	5	2	1	1	1	2	6	2
r	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	h_{11}
1	0	0	0	0	0	0	0	0	0	0	0
2	1	1	4	5	3	2	0	0	0	0	0
3	3	3	24	23	11	0	10	10	13	13	13
4	19	19	12	24	2	14	16	16	4	4	4
5	3	3	13	15	9	18	21	21	12	12	12
6	0	0	19	6	5	2	9	9	12	12	12
7	1	1	12	4	14	5	16	16	9	9	9
8	15	15	16	0	10	18	20	20	7	7	7
9	17	17	7	15	9	2	14	14	12	12	12
10	3	3	17	10	15	12	14	14	16	16	16
11	17	17	9	22	3	1	24	24	5	5	5
12	8	8	14	12	18	3	4	4	11	11	11
13	4	4	24	3	17	12	5	5	19	19	19
14	16	16	5	11	19	6	4	4	18	18	18
15	13	13	9	19	3	16	1	1	7	7	7
16	11	11	20	11	20	7	2	2	7	7	7
17	23	23	14	5	17	22	14	14	21	21	21
18	7	7	13	3	4	1	5	5	6	6	6
19	14	16	1	11	8	6	11	14	9	4	4
20	16	14	1	22	22	1	6	12	7	17	23
21	6	19	17	11	19	12	13	15	13	23	22
22	14	0	6	22	2	7	19	5	15	21	16
23	13	12	6	17	7	2	7	12	12	4	15

Table 3: Table represents the 11 solutions found for the Hamiltonian of partition thermodynamics consistent with the code table represented in Table 1. The integer pair (m, n) given in the first two rows codes for the correspondence between amino-acids (Trp, Lys, Met, Gln) and exceptional primes $(53, 79, 101, 103)$ according via the correspondence given in **Table 2**

m	n	$\beta(1)$	$\beta(4)$	$\beta(11)$	$\beta(13)$	$\beta(14)$	$\beta(15)$
1	2	19	11	6	5	24	21
1	2	19	11	6	5	23	7
1	5	21	5	15	6	4	7
1	5	15	13	10	23	21	13
2	2	10	16	23	15	16	21
3	1	6	17	16	17	3	19
3	1	10	2	23	17	20	11
3	1	10	2	23	4	4	12
3	2	5	6	5	18	18	7
3	6	5	6	5	8	23	16
4	2	11	6	5	24	23	18

Table 4: Inverse 5-adic temperatures $\beta = 1/t_5$ for doublets of the vertebrate mitochondrial code. The notational conventions and the ordering of solutions are same as in the previous table.

of solutions found is 11. If the solutions are distributed evenly, the estimate for the total number solutions is about 45.

The 5-adic temperature is $T_5 = 1$ for all lower doublets in the code table (the two smallest values of $n(DNA)$ in a given 4-column). The values of 5-adic temperature for the upper vertebrate mitochondrial doublets are given by **Table 4** for some cases. For eukaryote code symmetry breaking means only a change of 5-adic temperature for the symmetry breaking codon so that it codes for either Stop as in case of Trp-Cys doublet or for Ile instead of Met. Also the context dependence observed for some variants of the genetic code [I1] can be understood in terms of a temporary change of the 5-adic temperature. Note however that the amino-acid coded temporarily does not belong to the group of standard amino-acids.

For the stopping codon $1/T_5 = 2$ is the minimum temperature implying that no prime $31 \leq p \leq 113$ divides the partition function.

4.4 Number Theoretical Hamiltonian Identified As Spin-Spin Interaction

The hypothesis that Hamiltonian depends on the number r of summands in the partition is of course only a very simple working hypothesis allowing a relatively easy numerical search of the Hamiltonian (in the original model one had $n \leq 63$ so that rather large numbers of partitions had to be considered). If one takes seriously the idea about sub-condensates of spin 1 Cooper pairs, one could argue that the interaction energy between blocks of Cooper pairs is spin-spin interaction proportional to the product of net spins of electrons and is therefore of form $E(n_k, n_l) = Jn_kn_l$, $k \neq l$. A number theoretical analog of rather spin glass variant of Ising model would be in question.

In this case one would have $h = J\sum_{k,l} n_kn_l = \sum_k n_k(n - n_k) = n^2 - \sum_k n_k^2$ and thermodynamically equivalent with $h = J\sum_k n_k^2$. This Hamiltonian or rather, its modulo N variant ($N = 3$ in the minimal case), would distinguish between partitions with the same value of r. In the recent model one has $6 \leq n_{2)} \leq 24$ so that the numbers of partitions are quite reasonable.

What makes this Hamiltonian so attractive would be its clear physical interpretation and involve a minimal amount of ad hoc elements.

The simplest working option is that third nucleotide affects only the 5-adic temperature so that one would have

$$h(n_1, ..., n_r) = \frac{J}{T_5} \times \sum_{pairs} n_kn_l \ ,$$

where one has $T_5 = 1, 2$. This interpretation conforms with the idea about living matter as spin glass like structure for which interaction strengths for spin-spin interactions are variable parameters. This would also conform with the general vision about TGD Universe as a four-dimensional spin glass like structure [L2].

4.4.1 Calculation of the partition function for a model based on spin-spin interaction

The task is to calculate the partition function $Z(T(n_3)) = \sum_P 5^{h(n_2,P)/T_5}$. To achieve this one can generalize the recursion formulas for the numbers $d(n,r)$ of partitions of n to sum of r terms.

1. One can arrange the integers in the partition so that one has always $n_k \leq n_{k+1}$ and start the recursive calculation from $h_r(1,....1,n-r+1) = (n-r+1)(r-1)$.

2. This gives rise to general recursion formula given by

$$\begin{aligned} h_r(n_1,...,n_{r-1},n-r+1-k_1) &= J(n-r+1-k_1)(r-1+k_1) \\ &+ h_{r-1}(n_1,...,n_{r-1}) \ . \end{aligned} \tag{4.3}$$

   Using this recursion formula one can express the formula for Hamiltonian as

$$\begin{aligned} &\frac{1}{J}h_r(k_r+1,k_{r-1}+1-k_r,...,k_2+1-k_3,k_1+1-k_2,n-r+1-k_1) \\ &= (n-r+1-k_1)(r-1+k_1)+(k_1-r+2-k_2)1)(r-2+k_2) \\ &+...+(k_{s-1}-r+s-k_s)1)(r-s+k_s)+...+(k_{r-1}-k_r)k_r \end{aligned} \tag{4.4}$$

   In this formula $h \rightarrow h \ mod \ 25$ operation is not written explicitly.

The expression for the partition function can be written as

$$\begin{aligned} Z(n) &= \sum_r Z(n,r) \\ Z(n,r) &= \sum_{k_1,....,k_r} 5^{h_r(k_r+1,k_{r-1}+1-k_r,...,k_2+1-k_3,k_1+1-k_2,n-r+1-k_1)} \ . \end{aligned} \tag{4.5}$$

The lower and up upper bounds for k_s in the summation can be deduced as follows. An upper bound for k_1 obtained from the condition $rk_1 = n$ and gives $k_1 \leq k_{max} = [n/r]$ where $[x]$ denotes the integer $n \leq x$ nearest to x. The corresponding upper bound for k_s reads as $k_s \leq [k_{s-1}/r-s+1]$. A lower bound for k_s comes from the requirement $n_s \geq 1$ and gives $k_s \leq k_{s-1}$.

To avoid problems caused by the fact that the numbers for various loops are dynamical, one can use recursion to calculate $Z(n,r)$ such that the module in question calculates $h(...)$ by calling itself repeatedly. What simplifies the calculation dramatically is that it is not necessary to store the data about the values of Hamiltonian since partition function is all that is needed.

1. At s^{th} level the module first adds to the Hamiltonian of a given branch the contribution from that level and after that adds the contributions from from $(s+1)^{th}$ level.

2. The calculation branches which means a a loop over the values of k_{s+1}. This means that module calls itself at each step of the loop to calculate the contributions of the next level to the Hamiltonian at a given branch.

3. The module adds also to Z the contribution from $(s+1)^{th}$ level is added. The addition is trivial until the r^{th} level is reached and all contributions to the Hamilton are known.

4. At the last level of tree the situation looks like follows. At given branch of the tree at $(r-1)^{th}$ level the module adds in loop-wise manner to Z the contributions from r^{th} level for that branch. After the return to $(r-2)^{th}$ branch next branch at $(r-1)^{th}$ level is selected and same process is repeated. Etc...

5. In order to avoid overflow problems it is safest to express the terms of the partition function in pinary series with respect to the p-adic prime $31 \leq p \leq 113$ considered and perform the addition of contributions to Z in terms of the pinary series.

4.4.2 Structure of the calculation

The general structure of the calculation is following.

1. Perform a loop over n labeling the 2-codons and find for each of them the prime p for which negentropy $S_p(n)$ is minimum and look whether for a suitable choice of T_5 the resulting assignment $n \to p(n)$ is consistent with the geometric model of the code and with the basic idea of the divisor code.

2. For a given n perform a loop over allowed values of p to see whether anyone of them appears as a divisor of the partition function and which of them maximizes the number theoretic negentropy. Unless this occurs the codon in question is identified as a stopping codon. The proposed geometric model of course fixes the integers n associated with the stopping codon.

3. For given n and p perform a loop over the values of r and sum their contributions to the partition function $Z(n,r)$ by applying the recursive procedure described in the previous subsection. In order to avoid overflow problems (possibly appearing in the case of MATLAB), the calculation must be performed for each value of p separately using pinary expansions for $Z(n,r)$. If Hamiltonian belongs to Z_3, overflow problems are of course avoided automatically.

4. An alternative manner to view the calculation is to take the proposal for the $n \to p(n)$ correspondence represented as a table at the end of previous section as an input and by a suitable selection of $0 \leq J(n_{2)}) \leq 2$ try to reproduce it. Note that the correspondence between primes 53, 79, 101, 103 and amino-acids Trp, Met, Gln, Lys if not fixed by the model represented in the table.

5. The most practical manner to perform the calculation is to take $J = 1$ and allow T_5 to run from 1 to 2 for every value of n and look whether the resulting spectrum of primes is consistent with the proposed $n \to n(p)$ correspondence or possible modification of it. At the roughest level the calculation serves as a test for 5-adicity that is whether the integer $n = n_0 + n_1 5$ corresponds to prime of form $n + 25$ or $n + 75$.

4.4.3 Results

The proposed spin-spin interaction model allowing varying value of T_5 cannot reproduce the model summarized by **Table 3**. The roughest test for the model is whether 5-adic description of A-C and T-G symmetries works. For mod 25 thermodynamics with $n = n_0 + n_1 5$ determining the thermodynamics the fails to be consistent with the predictions of the simplest model.

5 A Possible Physical Interpretation Of Various Codes In TGD Framework

The inspiration for attempts to interpret physically the origin of various codes in TGD framework (summaries of quantum TGD, TGD inspired theory of consciousness, and TGD inspired view about quantum biology are given in articles [L4, L3, L2]) springs from the following ideas.

1. At fundamental level quantum TGD reduces to almost topological quantum field theory at light-like 3-surfaces of $H = M^4 \times CP_2$ having also interpretation as random light-like orbits of 2-dimensional partons, which can have arbitrarily large sizes. Quantum TGD involves fusion of real physics and its p-adic variants relying crucially to the assumption that S-matrix involves only data at intersections of real 2-surfaces and their p-adic counterparts obeying same algebraic equations consisting of rational pointsandalgebraic points in the algebraic extension of p-adic numbers characterization physical states in question. These intersections consist of discrete points giving rise to cognitive representations which should naturally relate to the genetic code.

2. TGD based view about dark matter as a hierarchy of quantum coherent phases labeled by symmetry groups $G_a \times G_b \subset SU(2) \times SU(2) \subset SL(2,C) \times SU(3)$, where $SL(2,C)$ is Lorentz group and $SU(3)$ corresponds to the gauge group of color interactions. These phases

are characterized by arbitrarily large values of Planck constants and are assumed to be responsible for the quantum control in living matter.

3. The generalization of the notion of imbedding space $H = M^4 \times CP_2$ based on the geometric realization of the dark matter hierarchy and involving a hierarchy of discrete sub-groups $G_a \times G_b$.

The basic idea is that the maximal cyclic subgroup Z_n of G_a could correspond to the group Z_n assigned with amino-acid and corresponding codons in the proposed group theoretic interpretation of the divisor code. n would give the order of the maximal cyclic subgroup $Z_n \subset G_a$ acting as symmetry group of wave functions of free electron pairs and (r, s), $rs = n$ could define a decomposition of $Z_n = Z_r \times Z_s$ with Z_r leaving invariant the electronic wave function.

5.1 Generalization Of Imbedding Space And Interpretation Of Discrete Bundle Like Structures

One should understand how the discrete number theoretical structures associated with various realizations of the genetic code emerge from TGD based physics. TGD suggests a very general geometric realization of the geometric codes in terms of points in the intersection of p-adic and real space-time sheets (actually a 2-D "partonic" surfaces having arbitrarily large size) consisting of algebraic points and of the TGD based generalization of imbedding space obtained by gluing together infinite number of copies of the imbedding space having singular bundle structure $H = M^4 \times CP_2 \rightarrow H/G_a \times G_b$, where one has $G_a \times G_b \subset SU(2) \times SU(2) \subset SL(2, C) \times SU(3)$.

G_a would manifest itself directly as discrete rotational symmetries of biomolecules basically due the presence of dark matter having G_a as exact group of rotational symmetries. Hence only G_a would be interesting in the recent case. In fact, the maximal cyclic subgroup Z_n for arbitrary G_a is in a special physical role and it might be possible to identify the group characterizing amino-acid and DNA as this group.

The bundle structure $H \rightarrow H/G_a \times G_b$ has singular points corresponding to the points of H for which $G_a \times G_b$ or its subgroup acts as an isotropy group leaving the point invariant. Quite generally, the singular points, in particular those for which G_a acts as isotropies, are involved with the phase transitions changing Planck constant and interpreted as a leakage of 3-surfaces between sectors of H labeled by different groups $G_a \times G_b$.

The interpretation of G_r characterizing DNA as an isotropy of singular point of bundle structure does not seem however natural. Rather, the wave functions of (say) free electron pairs (possibly Cooper pairs) defined in the set of points defined by the orbit of $Z_n \subset G_a$ could be invariant under some subgroup of $Z_r \subset Z_n$ for DNA labeled by (r, s), $r \times s = n$. Thus codons coding for an amino-acid having Z_n as a symmetry group would be characterized by wave functions for free electron pairs transforming under representations of Z_n and remaining invariant under $Z_r \subset Z_n$ and thus reducing to representations of $Z_s = Z_n/Z_r$. Note that $r = 1$ corresponds to all irreps of Z_n and $r = n$ to singlets under Z_n.

5.2 A Possible Interpretation For The Divisor Code

Consider now a model for what might happen in the coding of amino-acid by DNA.

1. Suppose that the maximal cyclic subgroup $Z_n \subset G_a$ acts as symmetries of "dark" space-time sheets and wave functions of "dark" free electron pairs for the amino-acid and corresponding DNAs so that the 2-surfaces in question are n-fold coverings of CP_2 points by M^4 points (corresponding to positions of say 5 molecules in a cyclic molecule) and corresponding codons. Free electron pairs could correspond to the dark matter in question.

2. Suppose that DNA characterized by n and its particular divisor r has electronic wave functions invariant under Z_r and thus forming irreducible representations of $Z_s = Z_n/Z_r$, $n = r \times s$. The electronic wave functions assignable to the amino-acid would in general transform according to some irreducible representations of $Z_n = \prod_i Z_{p_i}$, $n = \prod_i p_i$, where same prime p_i can appear several times. This assumption would explain why the product decompositions (r, s) and (s, r) are not equivalent.

5.3 About The Geometric Interpretation For The Thermodynamics Of Partitions Of $N_{2)}$

Suppose that the maximization of the information content for the thermodynamics for the partitions of the integer $n_2 = n \ mod \ 5^2$ belonging to the range $[6, 24]$ and labeling 2-codons provides a dual manner to understand the genetic code. $n \to n \ mod \ 25$ would have an interpretation in terms of reduction to a subset of the finite finite field $G(5, 2)$ and would be natural in 5-adic context.

One could try to interpret the modulo arithmetics in terms of the generalized notion of imbedding space.

1. One could label the points of M^4 covering of CP_2 by integers $0 \leq m \leq n$. The sheets points m and $m + k25$ should be equivalent from the point of view of mitochondrial genetic code so that Z_{25} equivalence classes would give rise to $n_{2)}$ points.
2. A more concrete interpretation would be that first nucleotide along gives rise to n_0-fold covering, second nucleotide adds $5n_1$ sheets so that $n_{2)} = n_0 + 5n_1$-fold covering results, and third nucleotide adds $n_3 5^2$ sheets so that to $n = n_{2)} + n_3 \times 5^2$-fold covering results. The sheets contributed by the third nucleotide would not participate in the partition thermodynamics and the third nucleotide would only determine the 5-adic temperature $T_5 = 1/n$.

5.4 About The Physical Interpretation For The Thermodynamics Of Partitions Of $N_{2)}$

The 5-adic thermodynamics relies on the partitions of $n_2 = n \ mod \ 5^2$. n_2 could have interpretation both as a net conformal weight or spin associated with spin one electronic Cooper pairs.

1. Modulo 5^2 property could be due to the invariance of electronic wave functions under Z_{25} acting as rotations. There would be 25-periodicity of physics in the covering, the analog of a lattice structure in angle degree of freedom with sub-lattices forming dynamical units. Also quantum group with quantum phase $q = exp(i\pi/25)$ implies the analog of lattice structure in angle degrees of freedom.
2. Each equivalence class analogous to a sub-lattice with points having distance of 25 units would effectively carry one unit conformal weight or one unit of spin (L_0 and iL_0 act as infinitesimal scaling and rotation respectively). At the concrete physical level the following alternative interpretations suggest themselves.

5.4.1 The interpretation in terms of conformal symmetry

The partitions of the integer $n_{2)} = n_0 + n_1 5$, $n_i \neq 0$ could have interpretation as partitions of the set of equivalence classes to a union of subsets with the number n_k of elements in the subset giving the total conformal weight created by L_{n_k} rather than L_1^k. These partitions could be interpreted as partitions of a molecular Z_{25} equivalence classes of building blocks of the molecular structure with Z_n rotational symmetry to subsets of basic building blocks and Virasoro generators L_{n_k} would act on various building blocks. A formation of bound states each binding single particle states associated with n_k sheets and created by L_1 suggests itself. The reduction of Virasoro algebra defined in Z to a Virasoro algebra defined in the finite field $G(5, 2)$ or in the ring Z_{25} is natural in this framework.

5.4.2 Interpretation in terms of irreducible representations of symmetric group and braids

Partitions label the conjugacy classes of symmetric group S_n consisting of the permutations of n objects. The summand n_k corresponds to a cyclic permutation of n_k objects. Partitions label also the irreducible representations of S_n. S_n can be defined by generators e_m representing permutation of m^{th} and $(m + 1)^{th}$ object satisfying the conditions

$$\begin{aligned} e_m e_m &= e_n e_m \ \text{for} \ |m-n|>1 , \\ e_n e_{n+1} e_n &= e_n e_{n+1} e_n e_{n+1} \ \text{for} \ n=1,...,n-2 \ , \\ e_n^2 &= 1 \ . \end{aligned} \tag{5.1}$$

By dropping the condition $e_n^2 = 1$ one obtains the defining relations of the braid group B_n of braid consisting of n strands. The irreducible representations of B_n are projective representations of S_n and give as a special case the representations of S_n.

1. Could the dynamics for partitions of n correspond to the dynamics for irreducible representations of S_n?

S_n brings in mind braids and topological quantum computation and the suggestion of [K16] that DNA and/or RNA might act as a topological quantum computer. The so called number theoretical braids, which provide representations for Galois groups permuting roots of an n^{th} order irreducible polynomial are subgroups of S_n (and equal to S_n in the generic case), are in a central role in the formulation of quantum TGD [K3], [A1].

This interpretation would assign to a given codon a braid with n strands, whose states would correspond to irreducible representations of S_n [A3]. The thermodynamics would be for the irreducible representations of S_n with the number n of braids varying in the range $[6, 24]$. Braid would be a 5-adic thermodynamical system such that all $d(n, r)$ irreducible representations with a given value of r would have the same value of the 5-adic Hamiltonian $h(r)$ (definitely not the most general dynamics now). The reason for the absence of n-braids for which n has zeros in its 5-adic expansion could relate to the fact that the quantum phase $q = exp(i\pi/m)$ defines a universal topological quantum computer for $m \geq 5$. $m = 5$ is suggested strongly in case of DNA since it manifests itself in the geometry of DNA (twisting angle for single nucleotide and the presence of 5-cycles).

2. More general dynamics?

The alternative interpretation forces to reconsider the definition of 5-adic thermodynamics. Let us denote by (n, r, i) the irrep of S_n corresponding to a particular partition of n with r summands. It would seem natural to interpret the dimension $D_{n,r,i}$ of the irrep as the additional degeneracy factor replacing $d(n, r)$ so that the number $d(n, r)$ of partitions with r summands (subsets) would be replaced by the degeneracy factor

$$D(n,r) = \sum_i D_{n,r,i} \ ,$$

and $D_{n,r,i}$ is the dimension of the irrep in question. The irreps $d(n, r, i)$ are in one-one-correspondence with Yang tableaus consisting of n boxes in r rows and $D(n, r, i)$ can be calculated using standard formulas [A4].

One might hope that this modification could allow to simplify the dynamics. The best one might dream of would be that $h(r)$ could be taken to be Z_3 valued: $0 \leq h(r) \leq 2$. One could also check whether the definition of the partition sum using modulo 125 arithmetics as $Z = \sum_r D(n, r)5^{h(r)} \ mod \ 125$ gives sensible results. Only two possible temperatures $1/T_5 = 1, 2$ besides $1/T_5 = 0$ corresponding to stopping codon are possible so that doublets pose very strong conditions on the model. The transformation $Z = Z_0 + Z_1 5 + Z_2 5^2 \rightarrow Z_0 + Z_2 5 + Z_1 5^2$ corresponds to the temperature scaling by 2. Hence it is not surprising that the simplest model does not work. In any case, the modification of earlier computational model to this case involves only the replacement of $d(n, r)$ with $D(n, r)$.

The dynamics could be however much more flexible. The 5-adic thermodynamics for irreducible representations of S_n instead of partitions allows the replacement of $h(r)$ with $h(n, r, i)$, say $h(n, r, d(n, r, i))$, where $d(n, r, i)$ is the dimension of the representation in question. The dynamics for a given n would be independent of the dynamics for other values of n unless one assumes that $h(n, r, d(n, r, i))$ is some simple function, say $h = d(n, r, i) \ mod \ 3$. In the most general case the number of parameters $h(n, r, i)$ would be the number of irreps given by the number $d(n) = \sum_r d(n, r)$ of partitions. For $n = 6$ one has $d(6) = 11$ partitions and $n = 24$ would give $d(24) = 3^2 \times 5^2 \times 7 = 1575$ partitions. Even for $h(n, r, i) \in Z_3$ this would increase the number of parameters dramatically and might allow to reproduce the genetic code in consistency with the constraints from the divisor code.

3. What could be the physical interpretation?

One can ask how this picture could relate to the picture provided by the divisor code in which representations of cyclic group Z_n reduced to some of its subgroup with integer $31 \leq n \leq 124$ being one of the integers associated with a given amino-acid. Is there place in TGD Universe for these two discrete symmetries? This might be the case if one takes seriously both the hierarchy of Planck constants involving the generalization of the imbedding space concept and the notion of number theoretic braid.

1. The permutation group $S_{n_{2)}}$, $n_{2)} \in [6, 24]$ for braid strands associated with the first two letters of the codon $n > n_{2)}$ would act on the number theoretical braids with $n_{2)}$ strands. The increase of n_2 could have interpretation as an increase of complexity in the sense that the number of braid strands increases.

2. The cyclic group Z_n, $n \in [31, ..., 124]$, possibly associated with electron pairs, could correspond to the G_a covering of $M^4_\pm$ defined by the hierarchy of Planck constants associated with the hierarchy of fiber bundle structures $H_\pm = M^4_\pm \times CP_2 \rightarrow H_\pm/G_a \times G_b$, $G_a \times G_b \subset SU(2) \times SU(2) \subset SL(2, C) \times SU(3)$. Cyclic group Z_n would be identified as the maximal cyclic group of G_a. Note however that topological quantum computer considerations would suggest that G_a has Z_5 as maximal cyclic subgroup so that Z_n cannot correspond to the number of sheets in the cyclic covering essential for topological quantum computation. A more natural interpretation would be as a cyclic group of symmetries for the magnetic flux quanta action as rotations permuting the flux tubes of the topologically quantized dipole type magnetic field. What remains a mystery is why $n_1 = n \ mod \ 5$, $n_2 = n - n_1 \ mod \ 5^2$, ... cannot vanish. Could the irreducible representation of $S_{n_{2)}}$ corresponding to the partition $n_{2)} = \sum_k n_k 5^k$ defined by 5-adic expansion and having $r = 2$ summands have a special role? IR could the sub-group $\prod_k S_{n_k 5^k}$ of S_n have a special role?

5.4.3 The interpretation in terms of decomposition to many-particle states consisting of free electron pairs or Cooper pairs

The fact that iL_0 corresponds to rotations allows to consider also the interpretation of the partitions in terms of decompositions of the state to a product of angular momentum eigen states with values of $J_z = n_k$. Basic building blocks could have spin $S_z = 1$ so that codon would be characterized by its total spin $S_z = n_2 = n \ mod \ 5^2$ possible associated with dark Cooper pairs with spin quantum number $S_z = 1$. The blocks of the partition would be coherent sub-Bose-Einstein condensates of dark Cooper pairs and the number theoretic Hamiltonian would characterize the change of energy like quantity as this kind of state is formed.

This interpretation conforms with the general TGD based view about living matter. High T_c superconductivity indeed plays a key role in TGD based model of living matter [K1, K2] and there is experimental evidence that DNA can have anomalously high conductivity [I2]. TGD based model [K2] relies on the hypothesis that free electron pairs associated with the 5- and/or 6-rings of sugars in the backbone of DNA correspond to dark matter with Planck constant $\hbar = n\hbar_0$, $n = 5$ and/or $n = 6$. Also the observation that the twist angle of single nucleotide in double helix is $\pi/5$ is suggestive of 5-adicity. Note that $n = 5$ defines the minimum value of n making possible universal topological quantum computation and in [K16] it is proposed that DNA and/or RNA could act as topological quantum computer.

5.5 A Possible Interpretation For The P-Adic Prime Labeling Amino-Acid And DNAs Coding It

The notion of field body or magnetic body is central for the TGD inspired model of living matter [K5] , [L2]. This notion is justified by so called topological quantization of classical fields making it possible to assign to a given physical system a field body which is typically much larger than the physical body. For instance, in case of brain the magnetic body is of astrophysical size (EEG wavelengths are of order Earth size). Dark magnetic body containing Bose Einstein condensates of ions with large value of Planck constant would be the fundamental bio-controller utilizing biological body as a sensory receptor and motor instrument [K5].

A possible interpretation for the p-adic prime labeling amino-acid and DNAs coding for it could be as a characterizer of the effective p-adic topology associated with their magnetic bodies and the genuine p-adic topology for their p-adic counterparts obeying same algebraic equations. This is possible since for large values of Planck constant possibly associated with the magnetic body the small p-adic primes could correspond to size scales of order EEG wave lengths. Notice however that the p-adic primes characterizing elementary particles are much larger. For instance, electron is characterized by Mersenne prime $M_{127} = 2^{127} - 1$.

The preferred values of n_a and n_b are given by $n_i = 2^k \prod F_i$, where F_i are distinct Fermat primes (only four of them corresponding to $F = 3, 5, 17, 257, 2^{16} + 1$ are known). The 2-adic hierarchy $n_a = 2^k$ could provide a deeper justification for the p-adic length scales hypothesis.

The 2-adic sub-hierarchy $n_a = 2^{k11}$, $k = 0, 1, 2...$ is especially interesting. For $n_b = 1$ $k = 11$ would correspond to the time scale $T_{121} = T(127)/64$, $T_{127}(2) = .1$ s, which defines the fundamental 10 Hz biorhythm. $T_{121} \simeq 1.6$ ms corresponds to a typical time scale for nerve pulse activity. For this option primary *resp.* secondary p-adic length scales associated with an amino-acid labeled by prime p would be $T_p = \sqrt{p}T_{121}$ *resp.* $T_p = pT_{121}$ and could define a small-p p-adic hierarchy of time scales of neuronal activity.

Obviously, the maximal cyclic subgroup of G_a containing 2^{121} elements and acting naturally as symmetries of magnetic and electric flux tube structures accompanying DNA and amino-acids cannot correspond to the group Z_n, $n \leq 124$ associated with DNA and amino-acid molecules.

Acknowledgements.

I want to thank Andrei Kozyrev and Marcus Nilsson for interesting discussions about models of genetic code, in particular the idea of divisor code.

6 Appendix: 4-Adic Realization Of $N \to N + 32$ Symmetry,Divisor Code, And Labeling Of Amino-Acids By Primes Are Not Mutually Consistent

For the four-adic realization of the divisor code geometrically 18 amino-acids would correspond to primes $p < 63$ whereas the integers $n = 0$ and $n = 1$ would correspond to special amino-acids. $n \to n + 32$ symmetry means that 4-columns of the code table contain either even or odd integers depending on whether the row is odd or even. Hence the 4-columns containing even integers cannot contain the prime coding for the amino-acid so that the geometric realization in which DNAs coding amino-acid contain both prime labeling for the amino-acid and the integer characterizing the degeneracy of the amino-acid as the number of its divisors is not possible.

One could weaken the condition by requiring that $n(p) = p$ holds true only when one of the coding codons is labeled by a prime. This however leads to a further difficulty since the primes $(5, 5 + 32 = 27)$ and $(11, 11 + 32 = 43)$ belong to same 4-column and should code for same amino-acid. Hence the assumption that amino-acids correspond to $n = 0, 1$ and 18 primes $p < 63$ does not look natural. One could however consider a less ambitious realization of the divisor code by giving up this requirement altogether and requiring only that one of the DNAs is labeled by an integer for which the number of divisors equals to the degeneracy of the corresponding codon.

For eukaryote code Met would naturally correspond to $n = 1$. For mitochondrial code the multiplets containg $n = 0$ and $n = 1$ DNA would contain also second DNA. The problem is that the number of its divisors should be $n = 2$ for the mitochondrial code for both Met and Ile and one end ups with a contradiction unless one somehow loosens the rules. One could say that the prime $n = 17$ determines the degeneracy of Ile for mitochondrial code so that Met takes the rest.

The multiplet coding for a particular amino-acid would contain DNA labeled by the prime coding for amino-acid and an integer with a number of divisors equal to the degeneracy of the codon. For odd rows of the code table 4-columns contain only even primes so that primes are contained in 4-columns in even rows of the table.

Table 5 represents the best variant found hitherto. One of the integers in 4-column is consistent with the degeneracy of amino-acid according to divisor code and for each amino-acid one of DNAs

UCC Ser	AGC Ser	CCC Pro	CUC Leu
UCA Ser	AGA Stop	CCA Pro	CUA Leu (16)
UCU Ser 20	AGU Ser	CCU Pro	CUU Leu 0
UCG Ser (4)	AGG Stop 8	CCG Pro 12	CUG Leu
(49) AUC Ile 53	CAC His 57	GUC Val 61	UUC Leu (33)
AUA Ile (37)	CAA Gln (41)	GUA Val (45)	UUA Phe 17
AUU Ile	CAU His	GUU Val 29	UUU Leu 1
AUG Met 5	CAG Gln (9)	GUG Val 13	UUG Phe
CGC Arg	GCC Ala	ACC Thr	GGC Gly 34
GGA Arg	GCA Ala	ACA Thr	GGA Gly 18
GGU Arg	GCU Ala	ACU Thr	GGU Gly 2
GGG Arg 6	GCG Ala 10	ACG Thr 14	GGG Gly
GAC Asp	UGC Cys 59	AACAsn 63	UAC Tyr
GAA Glu 39	UGA Trp (43)	AAA Lys (47)	UAA Stop 19
GAU Asp 23	UGU Cys	AAU Asn (31)	UAU Tyr 3
GAG Glu 7	UGG Trp 11	AAG Lys (15)	UAG Stop

Table 5: Best variant of the code table

corresponds to the integers consistent with the degeneracy. For Trp in case of eukaryote code stop breaks the symmetry. 7 codes only for a singlet (Trp).

REFERENCES

Mathematics

[A1] Langlands program. Available at: `http://en.wikipedia.org/wiki/Langlands_program`.

[A2] Partition function P. Available at: `http://mathworld.wolfram.com/PartitionFunctionP.html`.

[A3] Representation theory of the symmetric group. Available at: `http://en.wikipedia.org/wiki/Representation_theory_of_the_symmetric_group`.

[A4] Yang tableau. Available at: `http://en.wikipedia.org/wiki/Young_tableau`.

[A5] Khrennikov A. Gene expression from polynomial dynamics in the 4-adic information space, 2006.

[A6] Khrennikov AYu. p-Adic Probability and Statistics. *Dokl Akad Nauk*, (6), 1992.

[A7] Dragovich A Dragovich B. Available at: `http://arxiv.org/abs/q-bio/0607018`, 2006.

[A8] Kozyrev S Khrennikov A. Genetic code on the diadic plane, 2006.

[A9] Nilsson M Khrennikov A. A number theoretical observation about the degeneracy of the genetic code. Available at: `http://arxiv.org/abs/q-bio/0612022`, 2006.

[A10] Goppa VD. *Geometry and codes*. Kluwer, 1988.

Biology

[I1] The Genetic Code. Available at: `http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/C/Codons.html`.

[I2] Barton K et al. *Sci*, 275, March 1997.

Books related to TGD

[K1] Pitkänen M. Bio-Systems as Super-Conductors: part I. In *Quantum Hardware of Living Matter*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/bioware.html#superc1, 2006.

[K2] Pitkänen M. Bio-Systems as Super-Conductors: part II. In *Quantum Hardware of Living Matter*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/bioware.html#superc2, 2006.

[K3] Pitkänen M. Construction of Quantum Theory: Symmetries. In *Towards M-Matrix*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/tgdquantum.html#quthe, 2006.

[K4] Pitkänen M. Could Genetic Code Be Understood Number Theoretically? In *Genes and Memes*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/genememe.html#genenumber, 2006.

[K5] Pitkänen M. Dark Matter Hierarchy and Hierarchy of EEGs. In *TGD and EEG*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/tgdeeg.html#eegdark, 2006.

[K6] Pitkänen M. DNA as Topological Quantum Computer. In *Genes and Memes*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/genememe.html#dnatqc, 2006.

[K7] Pitkänen M. Does TGD Predict the Spectrum of Planck Constants? In *Hyper-finite Factors and Dark Matter Hierarchy*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/neuplanck.html#Planck, 2006.

[K8] Pitkänen M. Evolution in Many-Sheeted Space-Time. In *Genes and Memes*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/genememe.html#prebio, 2006.

[K9] Pitkänen M. Genes and Memes. In *Genes and Memes*. Online book. Available at: http://www.tgdtheory.fi/tgdhtm/genememel.html#genememec, 2006.

[K10] Pitkänen M. Homeopathy in Many-Sheeted Space-Time. In *Bio-Systems as Conscious Holograms*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/holography.html#homeoc, 2006.

[K11] Pitkänen M. Negentropy Maximization Principle. In *TGD Inspired Theory of Consciousness*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/tgdconsc.html#nmpc, 2006.

[K12] Pitkänen M. Nuclear String Hypothesis. In *Hyper-finite Factors and Dark Matter Hierarchy*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/neuplanck.html#nuclstring, 2006.

[K13] Pitkänen M. Quantum Antenna Hypothesis. In *Quantum Hardware of Living Matter*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/bioware.html#tubuc, 2006.

[K14] Pitkänen M. Quantum Hall effect and Hierarchy of Planck Constants. In *Hyper-finite Factors and Dark Matter Hierarchy*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/neuplanck.html#anyontgd, 2006.

[K15] Pitkänen M. The Notion of Wave-Genome and DNA as Topological Quantum Computer. In *Genes and Memes*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/genememe.html#gari, 2006.

[K16] Pitkänen M. Topological Quantum Computation in TGD Universe. In *Genes and Memes*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/genememe.html#tqc, 2006.

[K17] Pitkänen M. Was von Neumann Right After All? In *Hyper-finite Factors and Dark Matter Hierarchy*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/padphys.html#vNeumann, 2006.

[K18] Pitkänen M. *p-Adic length Scale Hypothesis*. Online book. Available at: http://www.tgdtheory.fi/tgdhtml/padphys.html, 2013.

Articles about TGD

[L1] Pitkänen M. Further Progress in Nuclear String Hypothesis. Available at: http://tgdtheory.fi/public_html/articles/nuclstring.pdf, 2007.

[L2] Pitkänen M. TGD Inspired Quantum Model of Living Matter. Available at: http://tgdtheory.fi/quantumbio.pdf, 2008.

[L3] Pitkänen M. TGD Inspired Theory of Consciousness. Available at: http://tgdtheory.fi/tgdconsc.pdf, 2008.

[L4] Pitkänen M. Topological Geometrodynamics: What Might Be the Basic Principles. Available at: http://tgdtheory.fi/public_html/articles/tgd2008.pdf, 2008.

[L5] Pitkänen M. CMAP representations about TGD. Available at: http://www.tgdtheory.fi/cmaphtml.html, 2014.

[L6] Pitkänen M. CMAP representations about TGD, and TGD inspired theory of consciousness and quantum biology. Available at: http://www.tgdtheory.fi/tgdglossary.pdf, 2014.