Progress in the understanding of the icosa tetrahedral realization of the genetic code

September 29, 2024

Matti Pitkänen

orcid:0000-0002-8051-4364.
email: matpitka6@gmail.com,
url: http://tgdtheory.com/public_html/,
address: Rinnekatu 2-4 A 8, 03620, Karkkila, Finland.

Abstract

TGD leads to two models of the genetic code. The first model emerges from a model of music harmony based on the combination of icosahedral and tetrahedral geometries. The second model relies on the representation of the genetic codons as entangled triplets of dark protons at the monopoles flux tubes defining the dark variant of DNA accompanying the ordinary DNA.

It took quite a long time to understand why both icosahedra and tetrahedra are needed and how the two models are related. The solution of the puzzle came from a universal model of the genetic code based on a completely unique tessellation of 3-D hyperbolic space H^3 realized as the light-cone proper time constant hyperboloid of the Minkowski space. This icosa tetrahedral tessellation (ITT) (known also as tetrahedral-icosahedral tessellation) makes sense in all scales and I have proposed its realization at the level of DNA. The model involves several intuitive elements and the best way to proceed is to try to improve the existing understanding and to identify the possible weaknesses of the model.

This article provides an answer to the question how many icosahedrons, octahedrons and tetrahedrons meet at the vertex of ITT: the numbers would be 3, 3, and 1. The study of the vertex figure of ITT in turn suggests that the ITT can be constructed as a "blow" up of the icosahedral tessellation by replacing vertices with tetrahedra and gluing to the faces of tetrahedra octahedra. It remains to be seen whether this information is consistent with the earlier heuristic model of the fundamental region of the ITT.

Contents

1 Introduction

TGD leads to two models of the genetic code. The first model emerges from a model of music harmony based on the combination icosahedral and tetrahedral geometries [K1] [L2]. The second

model relies on the representation of the genetic codons as entangled triplets of dark protons at the monopoles flux tubes defining the dark variant of DNA accompanying the ordinary DNA [L1].

It took quite a long time to understand why both icosahedra and tetrahedra are needed and how the two models are related. The solution of the puzzle came from a universal model of the genetic code based on a completely unique tessellation of 3-D hyperbolic space H^3 realized as the light-cone proper time constant hyperboloid of the Minkowski space. This icosa tetrahedral tessellation (ITT) (known also as tetrahedral-icosahedral tessellation) makes sense in all scales and I have proposed its realization at the level of DNA in [L4]. The model involves several intuitive elements and the best way to proceed is to try to improve the existing understanding and to identify the possible weaknesses of the model.

I am not a professional hyperbolic crystallographer and my view of ITT (see this) relies on guesswork guided by physical and biological intuition based on what I call icosa tetrahedral model for the genetic code.

I am not a professional hyperbolic crystallographer and my view of ITT (see this) relies on guesswork guided by physical and biological intuition based on what I call icosa tetrahedral model for the genetic code. In this article I represent some results based on using standard results from Platonic solids to deduce the numbers of tetrahedrons, octahedrons and icosahedrons emanating from a given vertex of the tessellation. The study of the vertex figure of ITT leads to a rather plausible guess for a manner to obtain ITT as a "blow-up" of icosahedral tessellation.

The improved understanding of the icosa tetrahedral tessellation allows to answer some long standing questions related to the detailed realization of the genetic code. It however turns out that the notion of the super-icosahedron discussed in [L4] is not consistent with the improved view. However the 3-D generalization of the vertex figure of the ITT as its inverse image under projection permutes the numbers for the Platonic solids appearing in the super-icosahedron.

2 More precise views about some aspects of the icosa tetrahedral realization of the genetic code

The improved understanding of the icosa tetrahedral tessellation allows to answer some long standing questions related to the detailed realization of the genetic code.

2.1 How many tetrahedra, octahedra and icosahedra meet at a given vertex of icosa tetrahedral tessellation?

One basic question is how many tetrahedra, icosahedra and octahedra emerge from a given vertex. I realized that one can answer this question just from the knowledge of the solid angles associated with vertices of these Platonic solids. The solid angles are naturally defined as ratios of spherical areas to the radial distance squared and at the limit of very small hyperbolic radial distance approaching Euclidean distance, the total solid angle is 4π as in the Euclidean case. Hyperbolic 3-space H^3 has a 3-D rotation group as isometries so that the notion of Platonic solid applies also in H^3.

The lines emanating from the vertex are shared by neighboring T, O, and I emanating from the vertex. Two neighboring lines are associated with a triangular face shared by two Platonic solids involved.

The basic condition for the numbers $n(i)$ of the Platonic solids involved is $\sum_{i\in\{T,O,I\}} n_i\Omega_i = 4\pi$. One can find the general formulas for the solid angles from Wikipedia (see this).

1. Platonic solids are classified by 2 integers $\{q,p\}$ stating that q $p-$polygons meet at a given vertex. In the recent case one has only 3-polygons, that is triangles, for all Platonic solids involved. One has

$$[q(T),q(O),q(I)] = [3,4,5] \quad , \quad [p(T),p(O),p(I)] = [3,3,3] \quad .$$

2. Dihedral angle angle is the interior angle between the faces of the Platonic solid and satisfies the general formula

$$\theta(q,p) = 2asin(\frac{cos(\pi/q)}{sin(\pi/p)}) \ .$$

The solid angle at the vertex is given as

$$\Omega(q,p) = q\theta(q,p) - (q-2)\pi \ .$$

3. Suppose that all vertices are identical as the fact that there is only single vertex figure, rhombicosadodecahedron (RID) (see this). The vertex of ITT, call it V, involves $n(T) \equiv n(3)$ tetrahedrons, $n(O) \equiv n(4)$ octahedrons and $n(I) \equiv n(5)$ icosahedrons. The sum of the solid angles equals to 4π, which gives

$$\sum_{q\in\{3,4,5\}} n(q)[q\theta(q,3) - (q-2)\pi] = 4\pi \ .$$

This gives

$$\sum_{q\in\{3,4,5\}} n(q)qasin(\frac{cos(\pi/q)}{sin(\pi/p)}) - (n(3) + 2n(4) + 3n(5))\pi = 4\pi \ .$$

4. One can try to guess the solution to the condition by starting from the icosa tetrahedral model [L1, L2] for the genetic code, which is a fusion of 3 icosahedral codes associated with Hamiltonian cycles (HCs) with symmetry groups Z_6, Z_4, Z_2 and of a single tetrahedral code defined by the unique tetrahedral HC. In the proposed model based on ITTs [L3, L4], the octahedron is passive and does not contribute to the code. A reasonable guess based on this model is $n(I) = 3$ and $n(T) = 1$.

   The normalized vertex solid angles are

$$\frac{[\Omega(3), \Omega(4), \Omega(5)]}{4\pi} = [0.043871, 0.1082, 0.2097] \ .$$

   The consistency condition is

$$\frac{n(T)\Omega(T) + n(O)\Omega(O) + n(I)\Omega(I)}{4\pi} = 1 \ .$$

   This leaves only the guess $[n(T), n(O), n(I)] = [1, 3, 3]$ under consideration giving for the sum the value 0.9974 in the accuracy used partially determined by the approximation $\pi \simeq 3.14159$. It would seem that the heuristic guess makes sense.

From this piece of data one can try to get some idea what ITT looks like.

1. Consider the vertex. There is a single T associated with the vertex and 3 O:s or 3 I:s have a common face with T. The remaining 3 T:s or 3O:s have common faces. The 4 vertices associated with T share the same T so that there is one T per 4 vertices. The set of T:s form a disjoint lattice-like structure.

2. The vertex figure of the ITT, obtained by cutting a 3-sphere around a vertex is a 2-D object, is rhombicosidodecahedron (RID) (see this). This is an Archimedean solid, one of thirteen convex isogonal nonprismatic solids constructed of two or more types of regular polygon faces. RID has 20 regular disjoint triangular faces (as also I has), 30 square faces which share their corners with other square faces, 12 regular disjoint pentagonal faces (as D has), 60 vertices, and 120 edges. Note that 3 vertices of RID are associated with T. The remaining $57 = 3 \times 19$ vertices could each have 19 T:s each shared by 3 vertices. The T:s in IIT are disjoint

3. The illustrations of RID gives a 2-D analog for what it means that the tessellation has different 3-D Platonic solids as building bricks. One can say that RD is obtained from I by replacing each vertex of I with pentagon and each edge of the pentagon of D or form D by replacing each vertex with triangle and each edge of the triangle by a square. Interestingly, the faces of RID are faces of the duals of the Platonic solids T, O and and I. In RID the triangles are disjoint and share sides with squares.

   Since RID and ITT are combinatorially slosely related, this suggests that the disjoint triangles of RID correspond disjoint T:s of ITT and the squares of RID having sharing only corners correspond to O:s of ITT sharing only edges whereas D:s would correspond to I:s.

Could an analog for the construction of RID allow to deform the hyperbolic icosahedral tessellation (IT) to ITT?

1. The construction would rely on the correspondence *triangle - self dual T, pentagon - I as dual of D, square - O as dual of cube*. One could generalize the correspondence to *triangle*$\rightarrow$ *T , pentagon* $\rightarrow$ *I, and square* $\rightarrow$ *O.*

   The recipe would be as follows. Start from the hyperbolic icosahedral tessellation (IT) $\{(3,5,3\}$ with 3 I:s $\{5,3\}$ meeting at each edge. One could blow-up the icosahedral vertices to T:s and glue to the faces of a given T 4 O:s. O:s would also share faces with other T:s and I:s but not with O:s if there is a combinatorial analogy with RID.

2. One can also ask how the 3 I:s and 3 O:s emanating from V share their faces. One expects by the analogy with the RID (squares have only shared corners) that O:s do not share faces but meet along common edges. Each O shares 1 face with T and this leaves 3+3+3 faces for O:s. 3 I:s have 5+5+5 = 15 faces and each O shares 3+3+3 faces with 5+5+5 faces of Is. This suggests that I:s share 2+2+2 faces with each other so that each I shares 2 faces with its nearest neighbor I and 3 faces with O:s.

The remaining challenge is to deduce the fundamental domain. It is not yet clear whether the above picture is consistent with the rather heuristic model of the ITT discussed in [L4] leading to a concrete model of DNA double strand..

2.2 Some comments related to the detailed realization of the genetic code

ITT emerged as a mathematization of the icosa tetrahedral realization of the genetic code and it is interesting to see whether the new results allow us to gain some understanding about the issues related to the detailed realizations.

In the original vision [K1], it was unclear whether there are 3 different I:s or only a single one realizing one of the 3 HCs at time. Also the relationship between T and I was unclear. The proposal was that there is a single I and T shares a common face with it. The idea about a common face was however somewhat fuzzy and I have discussed several ways to understand the details of the genetic code, in particular those assignable to stop codons.

The recent view of ITT however fixes the picture.

1. 3 icosahedral HCs, having symmetries Z_6, Z_4 and Z_2, are involved. Could one assign them to the 3 I:s associated with the vertex of ITT or with the 3 I:s meeting along the edge of the IT, whose "blow-up" ITT would be? There is only a single T associated with 3 I:s at V and this forces us to conclude that each vertex, and therefore each T, corresponds to a single letter of the genetic codon. One of the 3 HCs associated with the 3 I:s would be selected at each vertex. A single genetic codon would correspond to 3 T:s and associated 9 I:s and 9 O:s.

2. There would be 3 I:s and T associated with the letter. The T:s would share some faces and could not allow several I:s to be active simultaneously ("active" means that there is codon realized by adding dark protons to the vertices of the corresponding triangle). The HCs associated with I:s and T would have only a single point in common. In the original heuristic vision this assumption looked strange.

Can one understand why a genetic codon has 3 letters in this view?

1. If each vertex of IT is replaced with T in the blow-up identified as ITT, a single icosahedral triangle of IT would be replaced with 3 T:s. A natural identification would be in terms of a genetic codon with 3 letters, one T per letter.

2. The problem of the icosahedral realization has been following. The icosahedral HC with Z_6 symmetry corresponds to 3 Z_6 orbits with 6 triangles and one orbit with 2 triangles (6+6+6+2=20). This corresponds to 5 amino acids (AAs) identified as orbits of Z_6. The HC with Z_4 symmetry contains 5 orbits with 4 triangles ($5 \times 4 = 20$) and gives 5 AAs. The HC with Z_2 symmetry gives 10 orbits and therefore 10 AAs. One has 19 AAs altogether. One AA is missing.

3. Tetrahedral cycle involves 4 triangles and Z_3 symmetry is natural. This would give 2 AAs corresponding to the triangle opposite to V coding for an AA and the remaining triangles related by Z_3 and coding for a single AA (vertex of T and the vertices opposite to it). The triplet could correspond to stop codons coding for no physical AA. Could the singlet code for the missing AA?

4. There are also problems related to the chemical realization of the dark code. There are several slightly different chemical realizations of the code, which are not complete and violate the symmetries, which are exact for the dark realization. For instance the dark doublet pair (ile,ile)-(met,met) is replaced with (ile,ile)-(ile,met). A possible interpretation is as symmetry breaking.

There are also questions related to the HC and its realization and also the realization of the codons as cyclotron frequency triplets.

1. Icosahedral HC corresponds to a sequence of 12 vertices to which one can assign T:s. The basic idea of bioharmony is that one assigns to each vertex a note of 12-note scale and the notes associated with the triangular faces define the 20 chords of the harmony for a given T as dark counterparts of DNA codons.

   The 12 edges connecting the faces in turn represent an interval, which in the simplest model corresponds to a scaling of frequency by 3/2 or by $2^{7/12}$ corresponding to Pythagorean and well-tempered scales (HC as quint cycle module octave equivalence).

   Various HCs correspond to different realizations of the genetic code in terms of 3-chords realized as cyclotron frequency triplets assignable to the triangular faces of I and interpreted as different harmonies as representations of moods: this aspect is absent in the standard view.

2. How the cyclotron frequencies are assigned with the vertices of I. One could consider the situation also from the point of view of IIT as a blow-up of IT. Each vertex of I has T as a blow-up. One should assign a cyclotron frequency with this particular vertex of T.

3. Could one assign the frequency triplets with the 3 T:s associated with the blow-ups of the triangular face of I? A given cyclotron frequency is most naturally associated with the vertex which it shares with an active I. This seems necessary since the letters of the codon defined by the 3 T:s should be independent.

   Therefore the frequencies assignable to the vertices of T are not free unless the I:s associated with some of the 4 vertices of T are passive so that the codonic sub-tessellation does not continue to these directions. If it continues only in 1 direction as for the *linear* realization of the genetic code, 2 frequencies remain free and 2 are fixed by incoming and I codons. The possible frequency triplets would have 1 fixed frequency. Note that TGD predicts that genetic code is universal and has also higher-dimensional realization as sub-tesseslations of ITT [L4]. For instance, cell membrane could realize genetic code and brain might provide a 3-D realization.

4. The frequency triplet as a 3-chord is fixed if the tetrahedral frequency scaling in the step along the tetrahedral HC is fixed. The full T triangle must correspond to a scaling by octave if one accepts the octave equivalence. This requires that a single step represents a scaling by $2^{1/3}$ representing CE-interval, rather than the icosahedral scaling $2^{7/12}$ representing the quint interval CG. The T triangle would correspond to the 3-chord $CEG\sharp$. It would seem that one can insert these codons to the sequence of I codons if the sub-tessellation is linear.

5. There is however a potential problem. The frequency triplet opposite to the preferred vertex corresponds to $EG\sharp G\sharp$. If one can allow a triplet, which effectively reduces to a doublet, it must be very special from the point of view of the genetic code. The obvious candidates for this special codon are met appearing as start codon and the codon coding for trp. Trp is the largest of all twenty amino acids in the translational toolbox. Its side chain is indole, which is aromatic with a binuclear ring structure, whereas those of Phe, Tyr, and His are single-ring aromatics. Trp occurs at strategic places in proteins and trp often appears in lipid-water interfaces and plays an anchoring role.

6. In the case that the linear codon sequence ends, one has a situation in which a single O fixes only one T frequency. The other 3 frequencies are free unless one poses octave equivalence. This would allow only one frequency triplet unless the order of frequencies $C, E, G\sharp$ matters. At the level of DNA it seems to matter but at the level of AAs it cannot matter. The T triplets could correspond to different DNA codons coding for the same AA.

   The situation is very similar for the start codon and stop codons. The start codon of the gene coding for the met is very special. Could one assign the start codon to the triangle opposite to the active vertex of T, so that it would effectively reduce to a doublet and the three codons coding for ile to the other faces of T?

What DNA codons do the T codons correspond to? One can consider two options for which T codons would code for DNA triplet and singlet and there would be no symmetry at the level of T.

1. At the level of the ordinary DNA, T codons could correspond to 3 stop codons and a codon coding trp or to 3 ile codons and met. symmetry breaking. There would be symmetry breaking for the I quartet giving rise to (ile,ile,ile, met) but this is not possible since there are 5 AAs coded by 4 DNAs. Therefore this option fails.

2. For the second option dark T codons correspond to DNA codons coding for (ile,ile,ile,met). One doublet would code for (stop,stp) and a second doublet would break Z_2 symmetry and code for (stop,trp). This option might also allow us to understand the small deviations from the standard genetic code for which stop codons occasionally code for a real AA.

2.3 Can the proposed model of dark DNA sequences be consistent with the earlier model?

In [L4] proposed a heuristic model for the 10-codon piece of DNA sequence a candidate for the fundamental region of IIT. The idea was that it corresponds to what I called super-icosahedron (SI) having icosahedrons as 12 super-edges, tetrahedrons as 20 super-faces, and 30 octahedrons as super-vertices. What is worrying is that 2 DNAs would be missing so that there would be 10 Is. What goes wrong is that in the recent view there should be 30 T:s for a 10-codon piece.

The recent view suggests that DNA sequence corresponds to a linear sub-tessellation of IIT allowing to have also T codons. Each of the 30 letters would correspond to T and this is in conflict with the earlier proposal. It is not quite clear what the number of O:s could be. Each T of ITT has 3 O:s as neighbors. The fact that the O:s are passive with respect to the genetic code allows us to consider the possibility that the linear sub-tessellation contains. The number of O:s could vary in the range 0, 10, 20, 30. 10:s and 30 T:s would be the minimal option. The maximal option would be 10 I:s, 30 T:s and 30 O:s co be compared with a super icosahedron. 12 DNAs would not correspond to a fundamental region ITT but something else.

Intriguingly, the RID has 20 regular disjoint triangular faces (as also I has), 30 square faces, which share their corners with other square faces, 12 regular disjoint pentagonal faces (as D has) plus 60 vertices, and 120 edges. The triplet (20 triangles, 30 squares, 12 pentagons) contains the

same numbers as appear in SI. Could the correct identification of SI be, not as a fundamental domain of IIT, but as a 3-D combinatorial analog of RID obtained by the replacement (triangle → O, square → T, pentagon → I). Could the sequence of 12 DNAs correspond to (30 T:s, 20 O:s, 12 I:s) as the 3-D inverse image of the RID?

REFERENCES

Books related to TGD

[K1] Pitkänen M. Geometric Theory of Bio-Harmony. In *Genes and Memes: Part II. https://tgdtheory.fi/tgdhtml/Bgenememe2.html*. Available at: https://tgdtheory.fi/pdfpool/harmonytheory.pdf, 2023.

Articles about TGD

[L1] Pitkänen M. An overall view about models of genetic code and bio-harmony. Available at: https://tgdtheory.fi/public_html/articles/gcharm.pdf., 2019.

[L2] Pitkänen M. How to compose beautiful music of light in bio-harmony? https://tgdtheory.fi/public_html/articles/bioharmony2020.pdf., 2020.

[L3] Pitkänen M. Is genetic code part of fundamental physics in TGD framework? Available at: https://tgdtheory.fi/public_html/articles/TIH.pdf., 2021.

[L4] Pitkänen M. About tessellations in hyperbolic 3-space and their relation to the genetic code . https://tgdtheory.fi/public_html/articles/tessellationH3.pdf., 2023.