

Article

Algebraic and Geometric Representations of the Genetic Code

Richard L Amoroso^{*,1}, Peter Rowlands²

¹Noetic Advanced Studies Institute, USA

²University of Liverpool, UK

Abstract

Algebraic and geometric representations of the genetic code are used to show their functions for coding amino acids. The algebra is a 64-part vector quaternion combination, and the related geometry is based on the structure of the regular icosidodecahedron. An almost perfect pattern suggesting that this is a biologically significant way of representing the genetic code that may lead to a deeper understanding of a relationship between geometry and teleological life principles of complex self-organization.

Keywords: Genetic code, triplet codons, amino acids, vector-quaternion algebra, teleology, Icosidodecahedron.

1. Introduction

We explore an interesting way to represent the genetic code using a correspondence between algebra and geometry. The algebraic component is based on an Icosian calculus with a non-commutative algebraic structure discovered by William Rowan Hamilton in 1856, which he called quaternions. In modern terms, Hamilton produced a group presentation of the icosahedral rotation group by generators and relations.

Hamilton's discovery was derived from his attempts to find an algebra of 'triplets' that he believed would reflect the three Cartesian axes in a manner extending the complex numbers, which took the form, $i^2 = j^2 = k^2 = ijk = -1$. The symbols of the icosian calculus can be equated to moves between vertices on a dodecahedron.

2. The Algebraic Representation

In previous work [1-3] we have used various mathematical structures to represent the genetic code, including a 64-part vector quaternion algebra, which is isomorphic to the algebra of the quantum mechanical Dirac equation, and a combination of the faces and vertices of a regular icosidodecahedron. Here, we aim to show that it is possible to represent the codon structures both algebraically and geometrically in a way that relates to their function in coding for amino acids.

*Correspondence: Prof. Richard L. Amoroso, Director of Physics Lab., Noetic Advanced Studies Institute, Utah, USA.
<https://orcid.org/0000-0003-2405-9034>; <http://www.noeticadvancedstudies.us> E-mail: amoroso@noeticadvancedstudies.us

It is based on a vector-quaternion algebra whose units can be represented as follows [4-6]:

i	j	k	vector
i			pseudoscalar
i	j	k	quaternion
1			scalar

They can be considered as the units of two spaces: ('real', constructed from i, j, k) and ('vacuum', constructed from i, i, j, k, 1). In principle, any self-organizing system, whether physical, chemical or biological, forms a space, which has a kind of distorted mirror image in another 'space' representing the rest of the universe. The double space creates the entire combination of system and 'vacuum' as a zero totality.

In principle, any self-organizing system, whether physical, chemical or biological, forms a space, which has a kind of distorted mirror image in another 'space' representing the rest of the universe. The double space creates the entire combination of system and 'vacuum' as a zero totality. The algebraic structure has an exact parallel with a geometric one which can be represented using Platonic or Archimedean solids in which each structure has a dual which could be imagined as constructed in another space.

The units of the vector-quaternion algebra constructing the double space can be represented as follows:

i	j	k	ii	ij	ik	i	1	-i	-j	-k	-ii	-ij	-ik	-i	-1
i	j	k	ii	ii	ik			-i	-j	-k	-ii	-ii	-ik		
ii	ij	ik	iii	iii	ik			-ii	-ij	-ik	-iii	-iii	-ik		
ji	ij	jk	iji	iji	ijk			-ji	-ij	-jk	-iji	-iji	-ijk		
ki	kj	kk	iki	iki	ikk			-ki	-kj	-kk	-iki	-iki	-ikk		

An alternative ordering would separate the four complex numbers from 12 nilpotent structures, each formed from 5 units. Here, we create a subset of 60 units, which has significance in the dodecahedral and icosahedral representations and in Hamilton's Icosian calculus:

1						-1
ii	ij	ik	ik	j		-ii -ij -ik -ik -j
ji	jj	jk	ii	k		-ji -jj -jk -ii -k
ki	kj	kk	ij	i		-ki -kj -kk -ij -i
i						-i
iii	iiij	iiik	ik	j		-iii -iiij -iiik -ik -j
iji	ijjj	ijjk	ii	k		-iji -ijjj -ijjk -ii -k
iki	ikjj	ikk	ij	i		-iki -ikjj -ikk -ij -i

One way of generating the 64 units is by taking the product of 4 options × 4 options × 4 options, as is done in the case of the genetic code, where each of three bases may be U (or T), G, A or C. To represent this algebraically, we may use the vector units i, j, k and 1 for the options U, G, A

and C for the first base. Then, for the second base, we may represent U, G, A and C by the quaternion units i , j , k and 1 . Then U, G, A and C on the third base may be represented by the units of complex algebra 1 , i , -1 , $-i$. Using three different algebras (vectors, quaternions and complex numbers) allows us to track the three bases individually according to their positions in the codon.

We have previously grouped the amino acids produced by the genetic code mechanism according to the second base in the codon which produced it. The second base seems, in this respect, the most important, and the third base the least, becoming in some sense almost redundant. Using this division, the 64 codons fall naturally into 4 groups of 16:

amino acid codon first base second base third base

Group I					Group II				
Phe	UUU	i	i	1	Cys	UGU	i	j	1
	UUC	i	i			UGC	j	i	
Leu	UUA	i	i	-1	Trp	UGG	i	j	$-i$
	UUG*	i	i	$-i$	STOP	UGA	i	j	-1
	CUU	1	i	1	Gly	GGU	j	j	1
	CUC	1	i	i		GGC	j	j	i
	CUA	1	i	-1		GGA	j	j	-1
	CUG*	1	i	$-i$		GGG	j	j	$-i$
Val	GUU	j	i	1	Ser	AGU	k	j	1
	GUC	j	i	i		AGC	j	i	
	GUA	j	i	-1	Arg	CGU	1	1	
	GUG	j	i	$-i$		CGC	1	j	i
Ile	AUU	k	i	1		CGA	1	j	-1
	AUC	k	i	i		CGG	1	j	$-i$
	AUA*	k	i	-1		AGA	k	j	-1
Met	AUG*	k	i	$-i$		AGG	k	i	$-i$
Group III					Group IV				
STOP	UAA	i	k	-1	Ser	UCU	i	1	1
	UAG	i	k	$-i$		UCC	i	1	i
Tyr	UAU	i	k	1		UCA	i	1	-1
	UAC	k	i			UCG	1	$-i$	
Asp	GAU	j	k	1	Ala	GCU	j	1	1
	GAC	k	i			GCC	j	1	i
Glu	GAA	j	k	-1		GCA	j	1	-1
	GAG	j	k	$-i$		GCG	1	$-i$	
Lys	AAA	k	k	-1	Thr	ACU	k	1	1
	AAG	k	k	$-i$		ACC	k	1	i
Asn	AAU	k	k	1		ACA	k	1	-1
	AAC	k	$-i$			ACG	1	$-i$	
His	CAU	1	k	1	Pro	CCU	1	1	1
	CAC	k	i			CCC	1	1	i
Gln	CAA	1	k	-1		CCA	1	1	-1
	CAG	k	$-i$			CCG	1	$-i$	

The

asterisks represent codons that can act as a START. Notably they are all in the same group. Conveniently also (though this is mainly an artefact of our representation) all the START and STOP codons are represented by negative units. The way that the various structures are relevant to the formation of amino acids will become clearer in the following table:

Group I

Phenylalanine

UUU UUC
ii ii

Leucine

UUA UUG* CUU CUC CUA CUG*
-ii -i ii i -i -ii

Valine

GUU GUC GUA GUG
ij ij -ij -ij

Isoleucine

AUU AUC AUA*
ik ik -ik

Methionine

AUG*
-ik

Group II

Cysteine

UGU UGC
ji ij

Tryptophan

UGG
-ij

STOP

UGA
-ji

Glycine

GGU GGC GGA GGG
-jj -ij j ij

Serine

AGU AGC
jk ik

Arginine

CGU CGC CGA CGG AGA AGG
j ij -j -ij -jk -ijk

Group III

STOP

UAA UAG
-ki -iki

Tyrosine

UAU UAC
ki iki

Aspartate

GAU GAC
kj ikj

Glutamate

GAA GAU
-kj -ikj

Lysine

AAA AAG
-kk -ikk

Asparagine

AAU AAC
kk ikk

Histidine

CAU CAC
k ik

Glutamine

CAA CAG
-k -ik

Group IV

Serine

UCU UCC UCA UCG
i ii -i ii

Alanine

GCU GCC GCA GCG
j ij -j -ij

Threonine

ACU ACC ACA ACG
k ik -k -ik

Proline

CCU CCC CCA CCG
l i -l -i

We can see the pattern is nearly perfect; illustrated, particularly by the complete regularity of Groups II and IV. Only the two serine codons in Group II are anomalous (not in their biological group), and they will be in any arrangement. Arginine and serine in this group seem to each have two codons that could, originally, have coded a different amino acid. Almost certainly, the codons for arginine and serine have become mixed at some stage in biological evolution. (We may note also that arginine seems to be an exception to the general tendency for the more

complicated amino acid molecules to be coded using fewer codon pathways.) The table we have given is only for one species, and it may be that evolutionary drift may be determined to some extent by the variations in the patterns from the assumed perfect norm. The codon for tryptophan, notably, can become the STOP codon in some species, and vice versa.

3. The Geometrical Representation

Algebraic and geometrical structures are fundamentally dual. Where there is an algebra, there is also a geometry, and vice versa. It is easy to show that this is the case here. The four groups of codons can now be represented on the faces of a regular icosidodecahedron, divided into four equal sections. (We could use the combined faces plus vertices of a dodecahedron or icosahedron.) The negative units are not shown in the figures, but can be assumed either to be represented on the corresponding vertices of the dual rhombic triacontahedron, or on the inner, rather than outer, surface of the icosidodecahedron.

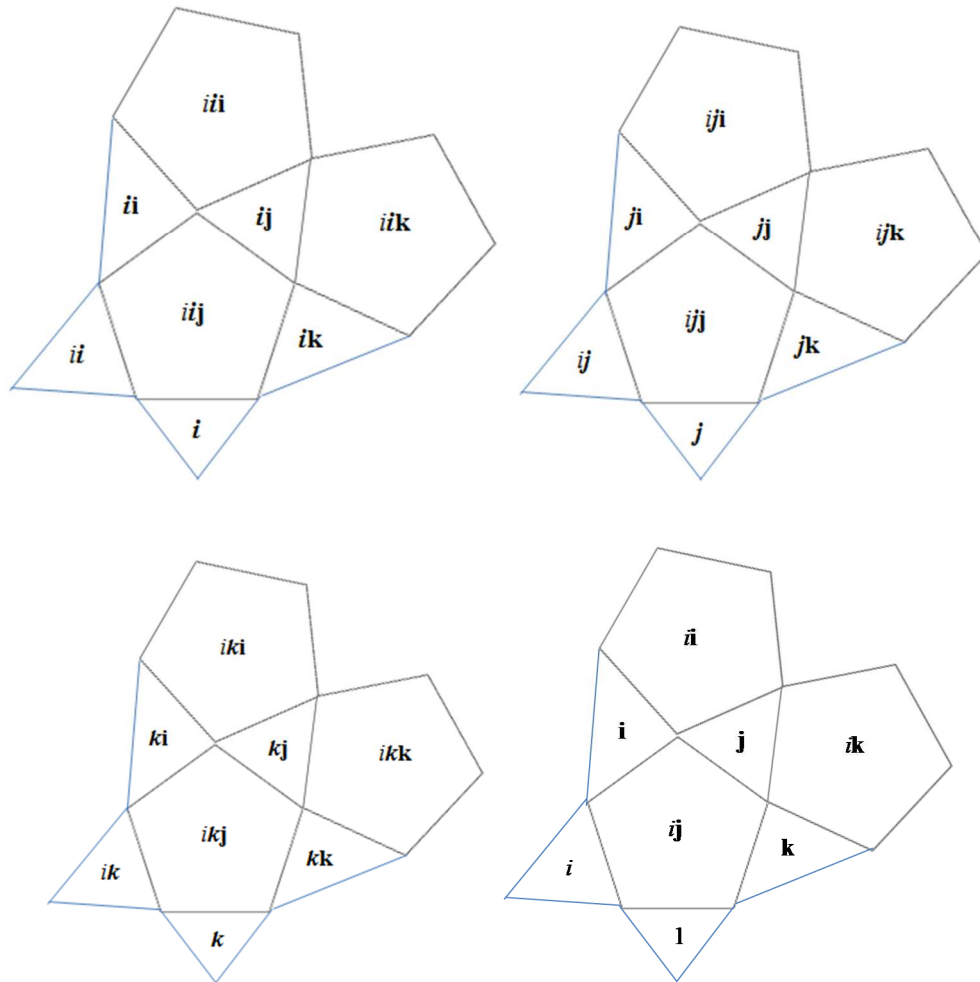


Figure 1. Algebraic geometry for representing codons on the icosidodecahedron.

The codons can be represented on these diagrams in the form:



Figure 2. codons representations

The amino acids coded can be represented as follows:

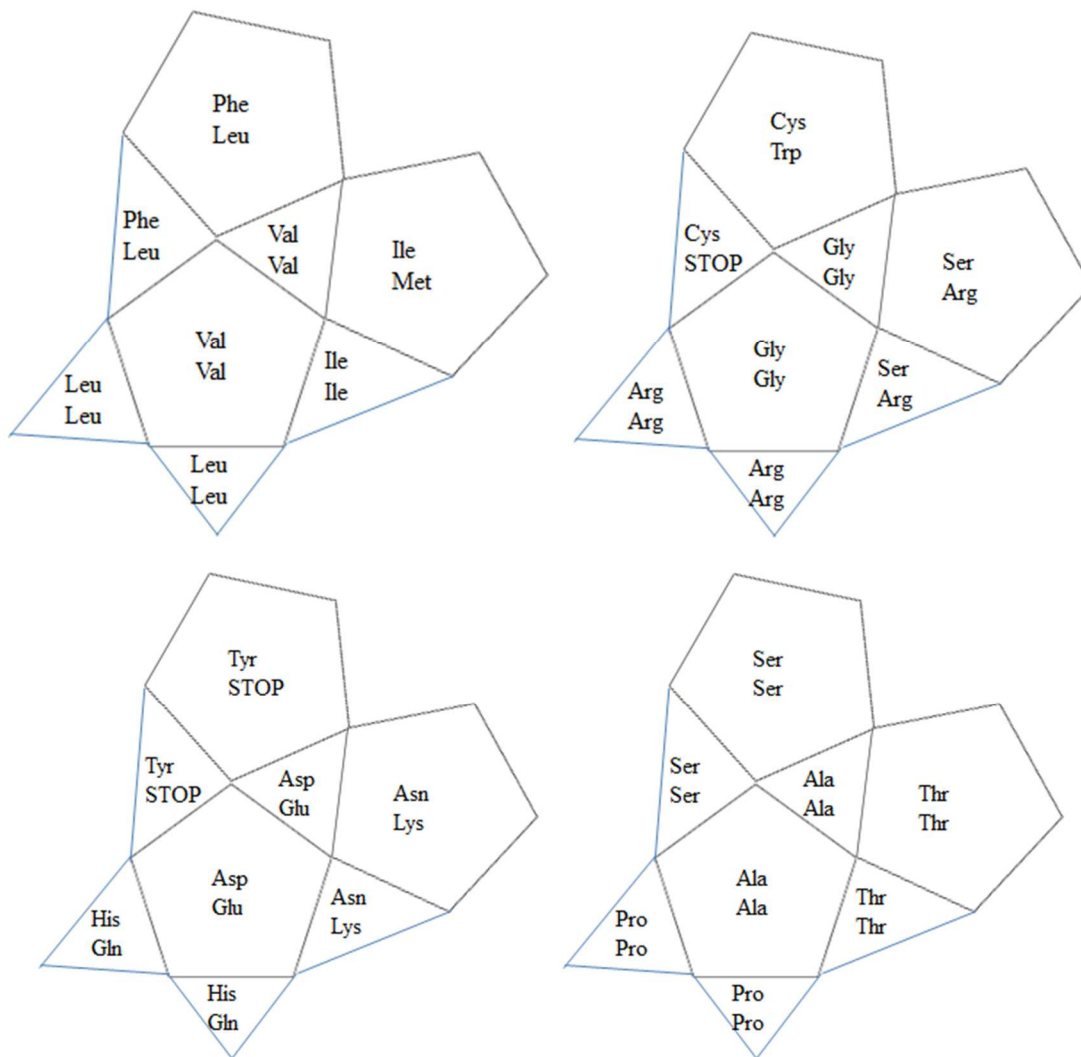


Figure 3. Amino acids on the icosidodecahedron.

Here, the position of serine in two separate groups gives an idea of how the four sections might be connected. One of the significant aspects of the first three sections in alternative arrangement (below) is that the algebraic units of the three pentagons in each, combined with those of the two outer triangles on the lowest pentagon form the basis of a nilpotent structure, such as we find in the amplitude term in the nilpotent Dirac equation ($ikE + iipx + ijpy + ikpz + jm$). At the same time, the five triangles taken together form the basis of another nilpotent structure. So, the three inner triangles and the three pentagons display a duality in that either group can be used with the two outer triangles to generate a set of nilpotent units (though with their roles switched in the two cases). This provides another way of generating 12 nilpotent structures from the algebra. Even using the first version of the icosidodecahedral sections, we can connect these triangles making up the nilpotent units with the upper pentagons, and so maintain the nilpotent structure. Nilpotency is one of the assumed bases of the overall pattern that we have described as Nature's code, and it appears to be the means by which a self-organizing system connects with its external

environment. Its presence in these geometric structures indicates the real importance of geometry in the genetic code as a route towards self-organization.

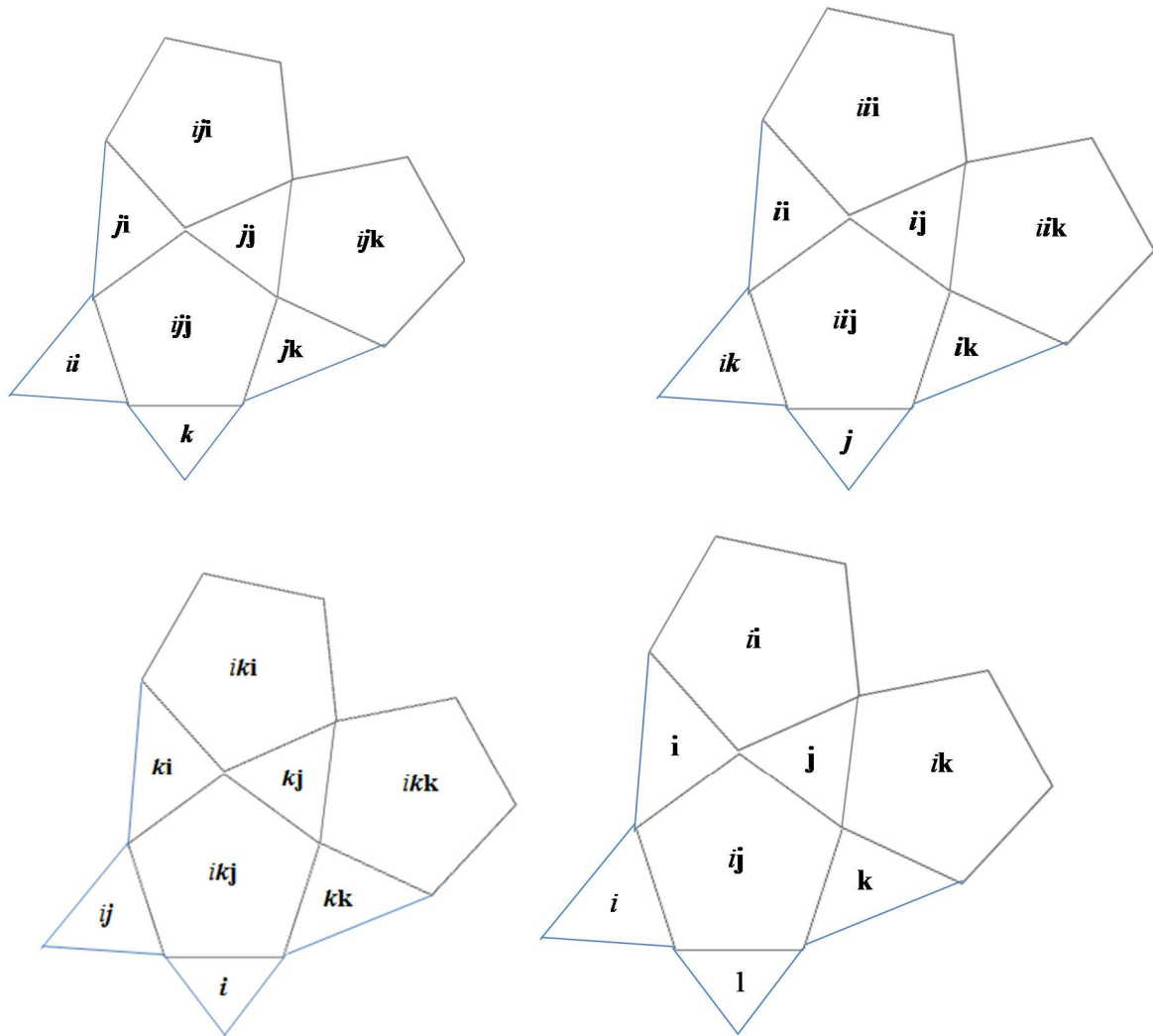


Figure 4. Quaternion algebraic representation.

Appendix: Note on the Regular Icosidodecahedron

Geometrically, an icosidodecahedron is a polyhedron with twenty (icosa) triangular faces and twelve (dodeca) pentagonal faces. An icosidodecahedron has 30 identical vertices, with two triangles and two pentagons meeting at each vertex. It also has 60 identical edges, each separating a triangle from a pentagon. Because of this, it is one of the Archimedean solids

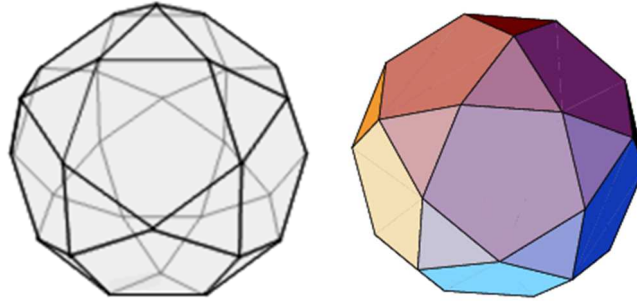


Figure 5. Two views of the icosidodecahedron.

All Archimedean solids can be produced from Platonic solids, by ‘cutting the edges’ of the platonic solid. Likewise, Platonic solids can be turned into Archimedean solids by following a series of rules for their construction.

Interestingly, in Cartesian coordinates, the vertices of an icosidodecahedron with unit edges are given by the even permutations of

$$(0, 0 \pm \varphi) \\ \left(\pm \frac{1}{2}, \pm \frac{\varphi}{2}, \frac{\varphi^2}{2} \right),$$

where φ is the golden ratio, $\frac{1+\sqrt{5}}{2}$ [7].

Received June 8, 2021; Accepted June 26, 2021

References

- [1] Hill, V.J. and Rowlands, P. (2008) Nature’s code, AIP Conference Proceedings, 1051, 117-126.
- [2] Hill, V.J. and Rowlands, P. (2010) Nature’s fundamental symmetry breaking, International Journal of Computing Anticipatory Systems, 25, 144-159.
- [3] Hill, V.J. and Rowlands, P. (2010) The numbers of nature’s code, International Journal of Computing Anticipatory Systems, 25, 160-175.
- [4] Rowlands, P. (2007) Zero to Infinity: The Foundations of Physics, Singapore and Hackensack, N.J., World Scientific.
- [5] Rowlands, P. (2010) Dual vector spaces and physical singularities, AIP Conference Proceedings, 1316, 102-111.
- [6] Amoroso, R.L., Rowlands P., Kauffman, L.H. (2013) Exploring novel cyclic extensions of Hamilton’s dual-quaternion algebra, in R.L. Amoroso, L.H. Kauffman, P. Rowlands (eds.) The Physics of Reality Space, Time, Matter, Cosmos, Proceedings of the 8th Symposium Honoring Mathematical Physicist Jean-Pierre Vigièr, pp. 81-92, Singapore: World Scientific; <https://vixra.org/pdf/1711.0468v1.pdf>.
- [7] Weisstein, E. W. (2010) Icosahedral Group, MathWorld, A Wolfram Web Resource; <https://mathworld.wolfram.com/IcosahedralGroup.html>.