Highlights

The H-bond Acceptor Scaffold for Watson-Crick Codon:Anticodon Pairing and the 4_{21} Polytope Scaffold for E_8

David Halitsky, Dr. Richard Klitzing, Greg Moxness

- Compelling E_8 group-theoretic properties are found in the set of 240 O and N H-bond acceptors which support Watson-Crick codon:anticodon pairing
- Empirically correct consequences are deduced from the group-theoretic model

The H-bond Acceptor Scaffold for Watson-Crick Codon:Anticodon Pairing and the 4_{21} Polytope Scaffold for E_8

David **Halitsky**[a,**], Dr. Richard **Klitzing**[b] and Greg **Moxness**[c,*]

[a] *Independent Researcher, South Orange, NJ, USA*
[b] *Independent Researcher, Heidenheim, Germany*
[c] *Independent Researcher, Tucson, AZ, USA*

ARTICLE INFO

Keywords:
polytopes
group theory
codon-anticodon interaction

ABSTRACT

All prior attempts to find meaningful mathematical structure in the genetic code have taken the $64 \times 3 = 192$ RNA codon bases as primitives. Instead, the primitives of this attempt are the 240 O and N H-bond acceptors which support Watson-Crick codon:anticodon pairing. These 240 primitives have a natural decomposition into (128,112) and then into ((72,56),(56,56)), parallel to well-known decompositions of the 240 roots of the group E_8 represented as vertices of the 4_{21} polytope.

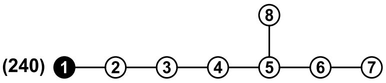
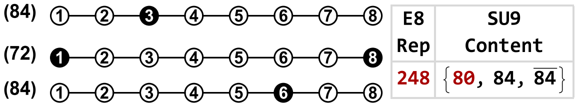


Figure 1: *Coxeter-Dynkin diagram of the E_8 4_{21} with 240 vertices.*



E8 Rep	SU9 Content
248	$\{80, 84, \overline{84}\}$

Figure 2: *E_8 Lie group A_8 maximal embeddings*

1. Introduction

This communication reports a heretofore unnoticed relationship between the symmetries of a familiar algebraic space and a familiar bio-molecular space. The algebraic space is defined by the 240 roots of E_8 represented as the 240 vertices of the semi-regular Gosset polytope termed 4_{21}. The bio-molecular space is defined by the 240 H-bond acceptors which support Watson-Crick codon:anticodon pairing. The relationship between the symmetries of these two spaces is self-evident from the fact that in both spaces, there is a natural decomposition of 240 into (128,112) and a further decomposition of (128,112) into ((72,56),(56,56)).

2. The geometry of the 8D polytope 4_{21}

The maximal exceptional symmetry of E_8 gives rise to the famous 8D polytope 4_{21} represented in Fig. 1, which Thorold Gosset found in 1900 as the dimensionally last finite semi-regular figure[4]. His attribute "semi-regular" then meant that it has only regular facets. In fact, it has a facet total of 2160 7D orthoplexes, 17280 7D simplices, and relevant in this article, exactly 240 symmetry-equivalent vertices. Its description by means of the symbol $k_{ij} = 4_{21}$ was later given by H. S. MacDonald Coxeter because of its Coxeter-Dynkin diagram which shows up with 4 links from the marked beginning (parent) node 1, and with 2 and 1 links each from the unmarked end nodes (7,8) toward the bifurcation node (5)[1]:

Because this polytope is still a convex shape, it can be considered as the hull of its vertex set. This latter property makes it easy to find subdivisions of its vertex set into subsets, the hulls of which provide a compound of the same vertex set in total. This is even more interesting when all of its compound components remain Wythoffian polytopes that belong to the common total symmetry, which provides for various solutions. Within the irreducible symmetries, we have in A_8 the compound of the polytopes shown in Fig. 2 and show a vertex subdivision according to $240 = 84 + 72 + 84$.

Please note that all Coxeter-Dynkin diagrams in Figures 2-10 have representation content boxes created from the *Mathematica*™ GroupMath package[3]. These indicate the Lie group dimension with the total number of E_8 roots (240) + E_8 rank (8). The red subgroup content text indicates which representation contains the rank, which is subtracted when counting the number of Weyl subgroup vertices in the polytope of that Coxeter-Dynkin diagram. These not only confirm the consistency of the polytope vertex counts, but also indicate deeper Lie group theoretic relationships to the codon:anticodon discussion in Sections 3-4.

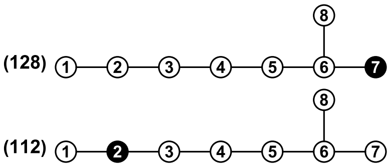
A different compound can be obtained from D_8 symmetry shown in Fig. 3 as that of the polytopes according to a vertex subdivision of $240 = 128+112$. This very subdivision will get reconsidered in further detail in this article later.

Other such subdivisions are known for reducible symmetries as well. For instance, $A_4 \times A_4$ shown in Fig. 4 would allow for the compound which corresponds to the subdivision into $240 = 20 + 20 + 50 + 50 + 50 + 50$. For $D_4 \times D_4$ shown in Fig. 5, we would get the compound which corresponds to the subdivision into $240 = 24 + 24 + 64 + 64 + 64$. For $A_2 \times E_6$ shown in Fig. 6, we have the

---

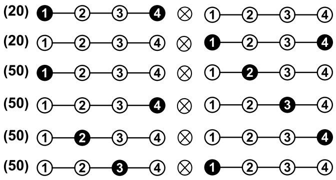
[*] Corresponding author
[**] Principal corresponding author

✉ `halitsky.d@att.net` (D. Halitsky); `richard.klitzing@t-online.de` (Dr.R. Klitzing); `jgmoxness@TheoryOfEverything.org` (G. Moxness)
ORCID(s): `0009-0007-1465-0011` (G. Moxness)



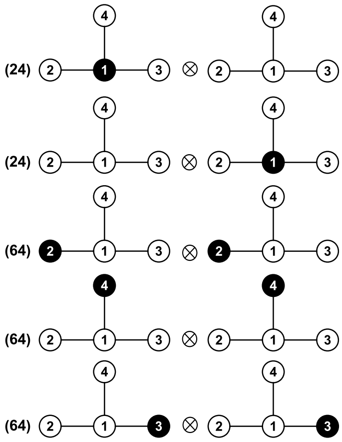
E8 Rep	SO16 Content
248	$\{120, 128'\}$

Figure 3: *E_8 Lie group D_8 maximal embeddings*



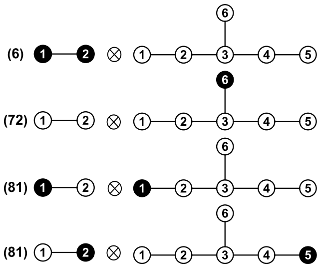
E8 Rep	SU5⊗SU5 Content
248	$\{24\otimes 1, 1\otimes 24, \overline{10}\otimes 5,$ $5\otimes 10, \overline{5}\otimes\overline{10}, 10\otimes\overline{5}\}$

Figure 4: *E_8 Lie group $A_4 \times A_4$ maximal embeddings*



SO16 Rep	SO8⊗SO8 Content
120	$\{28\otimes 1, 1\otimes 28, {8_v}^{\otimes 2}\}$
128'	$\{8_c\otimes 8_s, 8_s\otimes 8_c\}$

Figure 5: *E_8 Lie group $D_4 \times D_4$ maximal embeddings*

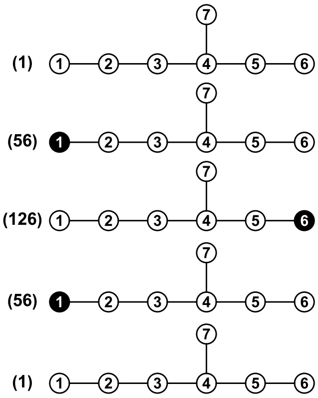


E8 Rep	SU3⊗E6 Content
248	$\{8\otimes 1, 1\otimes 78, 3\otimes\overline{27}, \overline{3}\otimes 27\}$

Figure 6: *E_8 Lie group $A_2 \times E_6$ maximal embeddings*

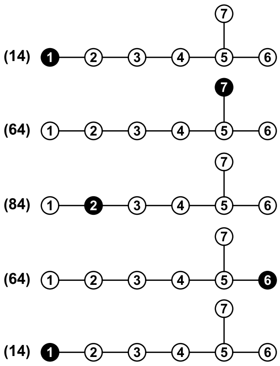
compound with diagrams that correspond to the subdivision into $240 = 6 + 72 + 81 + 81$. Clearly there are subdivisions according to reducible symmetries with even more factors too, but this would lead way beyond the scope of this article.

A quite different subdivision of vertex sets would be as stacked layers, i.e. choosing an (n-1)-dimensional subsymmetry and its orthogonal axis. Then the vertex set not only would fall into parallel hyperplanes perpendicular to that axis, but each individual vertex subset would also respect the common subsymmetry as well. In terms of polytopes, these vertex layers correspond to its parallel cross-sections at the various vertex heights. Or conversely, when considering the hulls of each of those layered subsets, the total polytope can be obtained as the lace tower of the polytope sections.



E8 Rep	E7⊗SU2 Content
248	$\{133\otimes 1, 1\otimes 3, 56\otimes 2\}$

Figure 7: *E_8 Lie group E_7 maximal embeddings (vertex first orientation)*



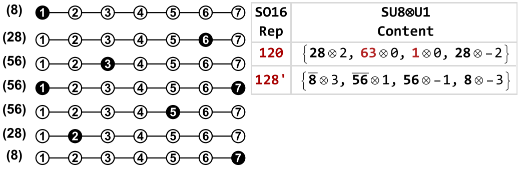
SO16 Rep	SO14⊗U1 Content
120	$\{14\otimes -2, 91\otimes 0, 1\otimes 0, 14\otimes 2\}$
128'	$\{64\otimes -1, \overline{64}\otimes 1\}$

Figure 8: *E_8 Lie group D_7 maximal embeddings (orthoplex-first orientation)*

For instance, 4_{21} can be considered as the vertex layer sequence with respect to its E_7 subsymmetry which corresponds to the subdivision into $240 = 1 + 56 + 126 + 56 + 1$ as shown in Fig. 7. This mimics its vertex-first orientation. With respect to its D_7 subsymmetry, we would get subdivisions into $240 = 14+64+84+64+14$ according to its orthoplex-first orientation shown in Fig. 8.

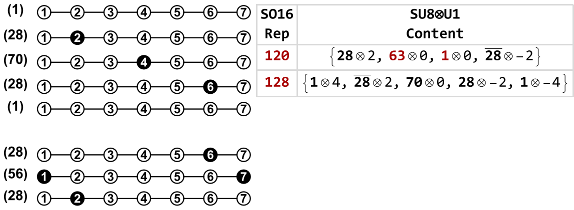
For its A_7 subsymmetry, we would get its axial subdivisions into $240 = 8 + 28 + 56 + 56 + 56 + 28 + 8$ shown in Fig. 9. If we now apply this latter axial subdivision onto the former compounding, thereby using a common orientation of all components, then for example, that D_8 symmetry representation could be further split using its A_7 subsymmetry into stacked layers which correspond to the subdivision into $240 = (1+28+70+28+1)+(28+56+28)$ shown in Fig. 10.

Note that both of those stackings with respect to the A_7 subsymmetry, either 4_{21} directly or its previously applied D_8 compound decomposition when accordingly re-grouped, provide the dissection of $240 = 56 + 56 + 56 + 72$ which indeed gets observed by the codon-anticodon pairings, as described in the subsequent sections.



SO16 Rep	SU8⊗U1 Content
120	$\{28\otimes 2,\ 63\otimes 0,\ 1\otimes 0,\ 28\otimes -2\}$
128'	$\{\overline{8}\otimes 3,\ \overline{56}\otimes 1,\ 56\otimes -1,\ 8\otimes -3\}$

Figure 9: *E_8 Lie group A_7 maximal embeddings (simplex first orientation)*



SO16 Rep	SU8⊗U1 Content
120	$\{28\otimes 2,\ 63\otimes 0,\ 1\otimes 0,\ \overline{28}\otimes -2\}$
128	$\{1\otimes 4,\ \overline{28}\otimes 2,\ 70\otimes 0,\ 28\otimes -2,\ 1\otimes -4\}$

Figure 10: *E_8 Lie group A_7 maximal embeddings (axial subdivision according to its D_8 compound components)*

3. The H-bond Acceptor Scaffold for Watson-Crick Codon:Anticodon Pairing

The structure of the Hydrogen (H) bond acceptor scaffold for Watson-Crick codon:anticodon pairing is determined by the way in which Oxygen (O) and Nitrogen (N) acceptors occur within the four RNA bases uracil (u), cytosine (c), adenine (a), and guanine (g)[2]. As shown in Fig. 11, the N acceptor occurs once within a and once within c, while the O acceptor occurs not only once within u and g, but also once within c. So this distribution of N and O acceptors among the four RNA bases results in a scaffold of 240 H-bond acceptors for Watson-Crick pairing of the $64 \times 3 = 192$ codon bases with anticodon bases.

In-as-much as these 192 bases contain 48 for each of $\{u, c, a, g\}$, there are:

- 48 O acceptors in the 48 u's
- 48 O acceptors in the 48 g's
- 48 N acceptors in the 48 a's
- 48 N acceptors in the 48 c's

with an additional:

- 48 O acceptors in the 48 c's.

Moreover, this scaffold of 240 acceptors contains certain empirically meaningful subscaffolds which can readily be discerned by considering the three standard (IUPAC) bifurcations of the 4 RNA bases:

- pyrimidine (y) vs purine (r) bifurcation, where $y = \{u, c\}$ and $r = \{a, g\}$
- keto (k) vs amino (m) bifurcation, where $k = \{u, g\}$ and $m = \{a, c\}$
- weak (w) vs strong (s) bifurcation, where $w = \{u, a\}$ and $s = \{c, g\}$

In particular, suppose that for any of $x \in \{r, y, k, m, w, s, u, c, a, g\}$, the term "$x$-centered codon" is defined to mean any codon whose center base is an x. Then there are 16 for each of $\{u, c, a, g\}$-centered codons and 32 for each of $\{r, y, k, m, w, s\}$-centered codons. Corresponding to these natural classes of x-centered codons, there are:

- u-centered subscaffold of 56 acceptors for bases of u-centered codons
- a-centered subscaffold of 56 acceptors for bases of a-centered codons
- g-centered subscaffold of 56 acceptors for bases of g-centered codons
- c-centered subscaffold of 72 acceptors for bases of c-centered codons

which in turn define:

- r-centered subscaffold of 112 acceptors
- y-centered subscaffold of 128 acceptors
- k-centered subscaffold of 112 acceptors
- m-centered subscaffold of 128 acceptors
- w-centered subscaffold of 112 acceptors
- s-centered subscaffold of 128 acceptors

4. Parallel Decompositions of the 240 4_{21} Vertices and the 240 H-bond Acceptors

From Section 2, it should be clear that the set of 240 4_{21} vertices naturally decompose into a pair of sets containing 128 and 112 vertices, which in turn decompose into two pairs of sets - one pair containing (72,56) vertices and the other containing (56,56) vertices.

And from Section 3, it should be equally clear that there are many various ways in which the scaffold of 240 H-bond acceptors can be naturally decomposed into two subscaffolds of (128,112) vertices and then into two pairs of subscaffolds with (72,56) and (56,56) vertices. For example:

- 240 acceptors = 128 m-based acceptors and 112 k-based acceptors
- 128 m-based acceptors = 72 c-based acceptors and 56 a-based acceptors
- 112 k-based acceptors = 56 u-based acceptors and 56 g-based acceptors

It is left to the reader to tabulate the other possible (128,112) and ((72,56),(56,56)) decompositions of the 240-acceptor scaffold, and also the equivalent decompositions which can be obtained simply by replacing the term "x-centered codon" with "x-initial codon" or "x-final codon", where "x-initial codon" means any codon whose first base is an x and "x-final codon" means any codon whose last base is an x.

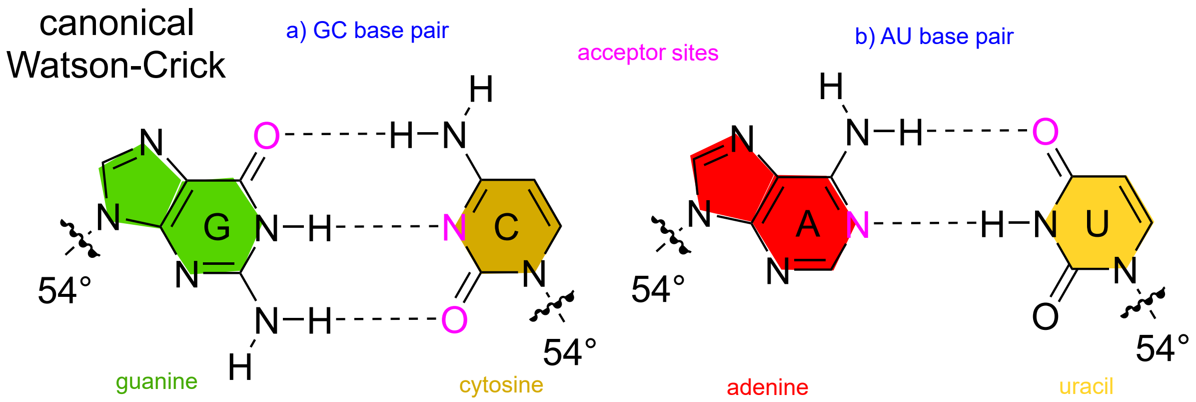


Figure 11: *Canonical Watson-Crick GC and AU base pairs*

CRediT authorship contribution statement

David Halitsky: Conceptualization and coordination of the research, developed the theory, performed computations. **Dr. Richard Klitzing:** Developed the theory, performed computations, and validation of geometric claims. **Greg Moxness:** Developed the theory, performed computations, validation of geometric claims, and wrote the manuscript.

References

[1] Coxeter, H.S.M., 1963. Regular Polytopes. 2nd ed., Macmillan, New York.

[2] Crick, F.H., 1966. Codon–anticodon pairing: the wobble hypothesis. Journal of Molecular Biology 19, 548–555. URL: `https://doi.org/10.1016/S0022-2836(66)80022-0`, doi:`10.1016/s0022-2836(66)80022-0`.

[3] Fonseca, R.M., 2021. GroupMath: A Mathematica package for group theory calculations. Computer Physics Communications 267, 108085. URL: `https://doi.org/10.1016%2Fj.cpc.2021.108085`, doi:`10.1016/j.cpc.2021.108085`.

[4] Gosset, T., 1900. On the regular and semi-regular figures in space of *n* dimensions. Messenger of Mathematics 29, 43–48.

David Halitsky (MA) is a retired consulting developer who has published in leading peer-reviewed journals (Language, Linguistic Analysis, and Mathematical Biosciences) and given presentations at MAA and AMA annual meetings.

Dr. Richard Klitzing is a trained mathematician with a PhD in theoretical physics from Eberhard Karls University of Tübingen on the theory of quasicrystals. He is an external part-time lecturer at the cooperative State University of Heidenheim. He is responsible for the discovery of several scaliform and CRF polytopes, as well as inventing segmentotopes as a pedagogical means of easy to visualize monostratic polytopes, and researched on edge-facetings.

Greg Moxness is a retired Chief Technologist of a major Aerospace and Defense company with a BSEE from University of Minnesota. He is an independent researcher who owns and operates www.TheoryOfEverything.org and is a frequent contributor to the group theory and polytope articles on Wikipedia.