1 Introduction

Viruses come in a variety of forms:

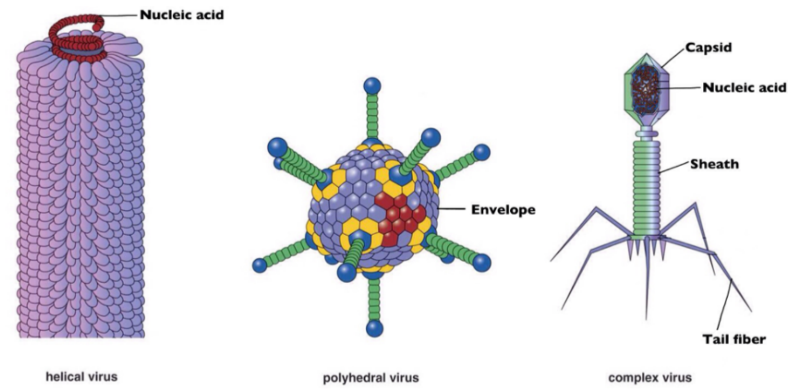


Figure 1: The different shapes of viruses[1]. Viruses can be helical in structure, they can have polyhedral symmetry (often called spherical viruses), or they can also be found in the form of a bacteriophage.

In this paper we will focus specifically on the structure of spherical viruses, which have icosahedral symmetry. Icosahedral symmetry is a group of rotational symmetries of order 60, denoted by I. The most important features that we explore are the symmetry axes that are present in objects that exhibit icosahedral symmetry. Each isocahedral polyhedron has 5-fold, 3-fold, and 2-fold symmetry axes.

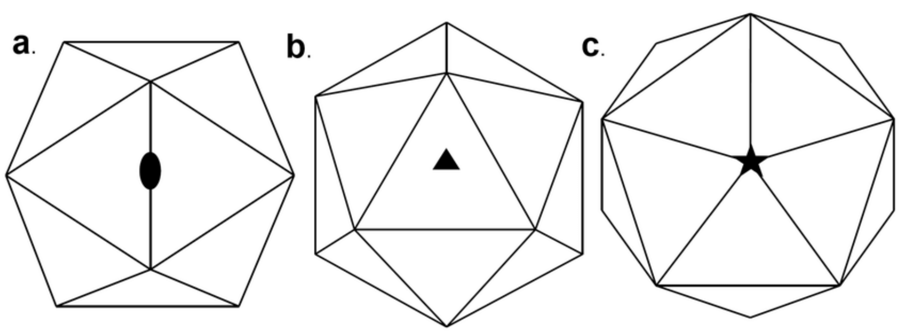
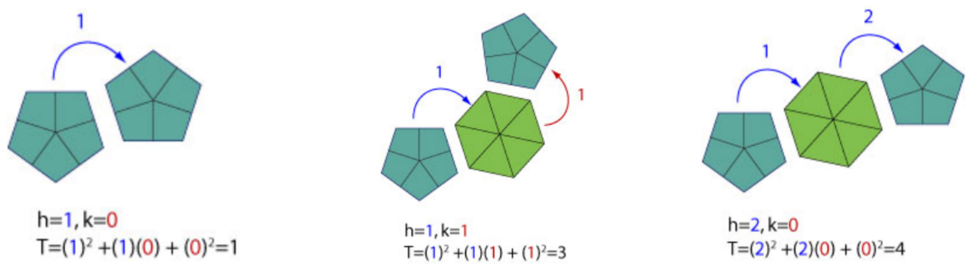


Figure 2: Symmetry axes of a regular icosahedron. (a) is a 2-fold axis, (b) is a 3-fold axis, (c) is a 5-fold axis[4]. For a two fold axis, there are two rotations that maintain the shapes orientation. If we rotate a 2 fold axis 180 degrees, then its orientation is preserved. A three fold axis needs to be rotated 120 degrees in order for it to look the same, and a 5-fold axis has 5 rotations of 72 degrees.

Up until this point, the unique orientation of proteins on the viral capsid - a protective shell that surrounds a virus's genome - has been classified using a Triangulation number (T number). The T number describes the number of steps that must be traversed in each direction, starting from a pentamer and colliding with another pentamer. The formula to determine T number is as follows:

$$T = h^2 + hk + k^2 \tag{1}$$

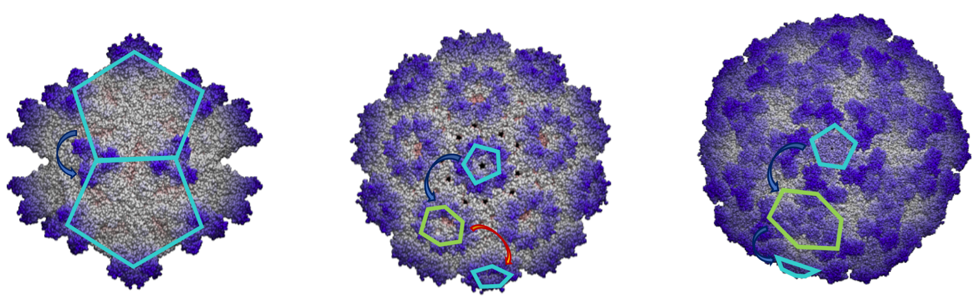
Where h and k each represent a step in each direction. Consider the following diagram.



(a) T = 1 (b) T = 3 (c) T = 4

Figure 3: A demonstration of T-number[2], where a step in each direction is represented by h and k. For example, in Figure 3.a, it only takes one step in the same direction to reach another pentamer. Therefore, h = 1, and k = 0, so the virus would be a T = 1 virus. In Figure 3.b, it takes 1 step in each direction, so both h and k are equal to 1. $T = 1^2 + (1)(1) + 1^2 = 3$. In the final case, we have an example where it requires two steps in the same direction in order to link two pentamers.

Let's overlay some of these pentamers and hexamers over a real virus to get a better feel for how the T number is used as a classification for the arrangement of proteins.



(a) Bursal Disease Virus, T=1 (b) Bacteriophage MS2, T=3 (c) Nudaurelia Capensis ω Virus, T=4

Figure 4: A demonstration of T-number, where a step in each direction is represented by h and k. Figure 4 provides a real life analog to the T-number orientations shown in Figure 3.[3]

While T-number is an excellent organization system that allows us to quickly gain information on on the number of proteins in the viral capsid (just multiply T by 60), the triangulation number unfortunately provides absolutely no information on the radial distribution of the capsid. Consider the case of the following viruses, who all have a T-number of 1, and thus all have 60 proteins on the viral capsid.

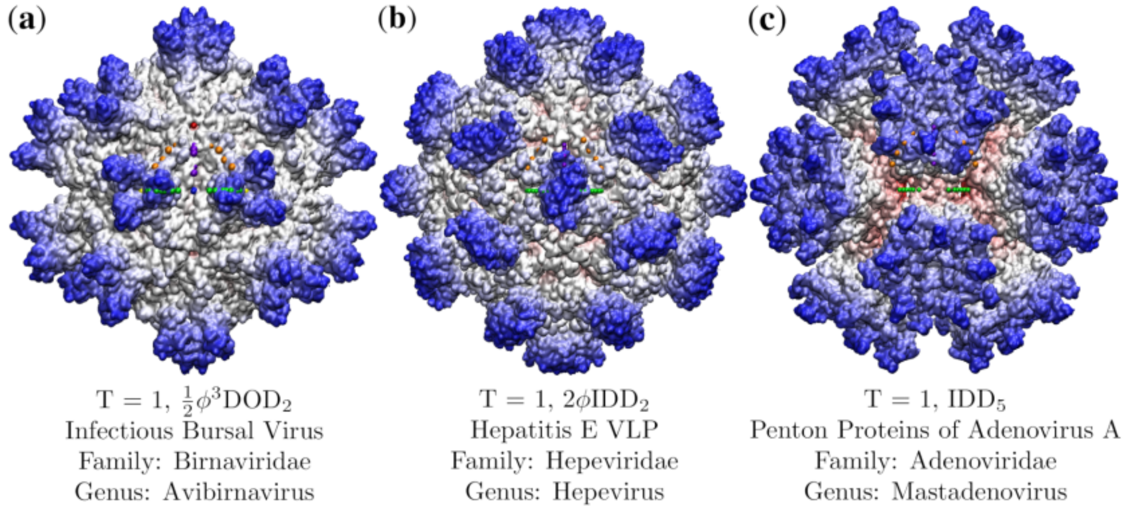


Figure 5: A demontration of the T-number failing to relate information about the radial distribution of the viral capsid[5]. The notation of 2ϕIDD$_2$ will be explained in the next section

2 Point Arrays

In order to construct more information about the 3-d assortment and symmetry of a virus at the radial level, we turn to point arrays. Point arrays are a cloud of points that are generated from two key features: a base icosahdral shape, and an extension in the direction of one of the symmetry axes. So for example, we can start with the base shape as an icosahedron, and then extend in the direction of the 5-fold axis. After we've performed the extentsion, we re-apply the I symmetry operations to generate a point cloud that is also icoshedrally symmetric.

Point Arrays are a valuable tool because they generate structures that convey information about fluctuations in radial symmetry. When applied to viruses, we can better understand certain geometric stable points, and be able to better classify structural relationships between other viruses.

3 Decomposition

The first thing I sought out to do was determine the symmetries of different radial levels of a point array. I performed my decompositions and classifications using MATLAB, by analyzing the 55 point arrays that are generated from the base shapes: icosahedron, dodecahedron, isodeodecahedron, and whose extensions are along the direction of the 2-fold, 3-fold, or 5-fold symmetry axis. After being provided the coordinates of the first 55 point arrays by my SIP research mentor Dr. Dave (Note: this Point array analysis is completely different to what my SIP work consisted of, this is a geometric exploration of a tool I was introduced to while studying viruses for my SIP), I began plotting what they looked like in 3-d space to build intuition to their structure by simply plotting points at every coordinate. Here is what the very first point array, ϕICO$_2$.

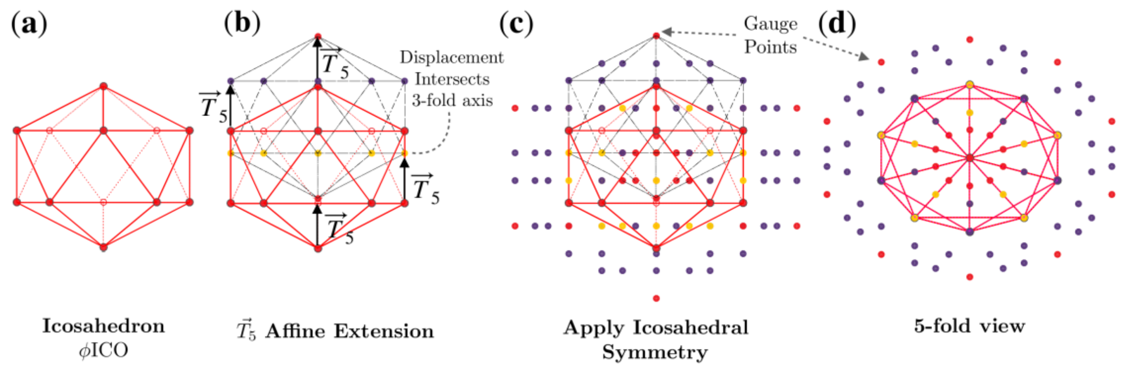


Figure 6: How point arrays are generated. The notation to describe these point arrays starts by telling you the base shape and its scale, and then a subscript indicating the direction of the extension. For example, $\phi \mathrm{IDD}_2$ tells us that we are starting with the base shape of and idodecahedron with scaling of ϕ, the golden ration, and then extending in the direction of the 2-fold symmetry axis.[5]

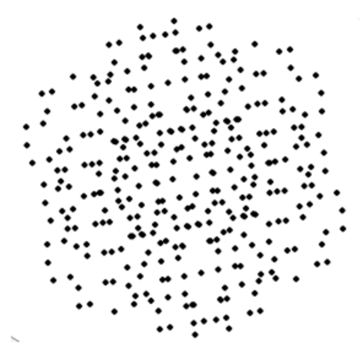


Figure 7: 3D view of the point array cloud generated by a base shape of and icosahedron extended in the 2-fold direction.

We can elucidate the structure by separating the different radial levels. Because the point arrays have icsahedral symmetry, we will see icosahedrally symmetric geometris arise from each individual radial level of the point arrays. The next thing I did was write a script that calculated the radius of every point and then kept track of what the radial values were present in the structure in an array, and finally I separated the point array into its various radial levels. Plotting these different levels provides a refined illustration of the point arrays. Take for example the decomposition of point array 1 into its various radial levels:

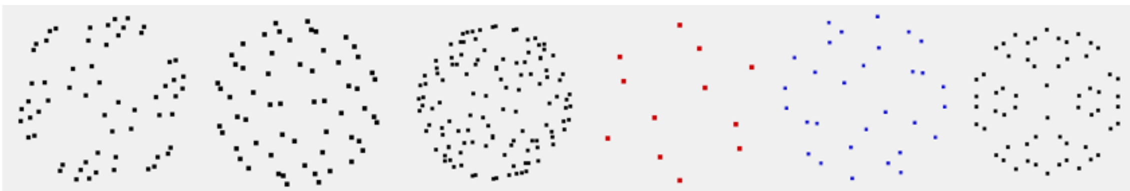


Figure 8: The first point array's various radial levels, descending from left to right.

While this 3-d view from some arbitrary angle is not enough to precisely classify the point arrays at the radial levels, we take a more objective approach by viewing every radial level from the positive z-axis, consequently looking down the 5-fold symmetry axis. Here is a representation that shows the relative radial level along with how many vertices there are at that radial level.

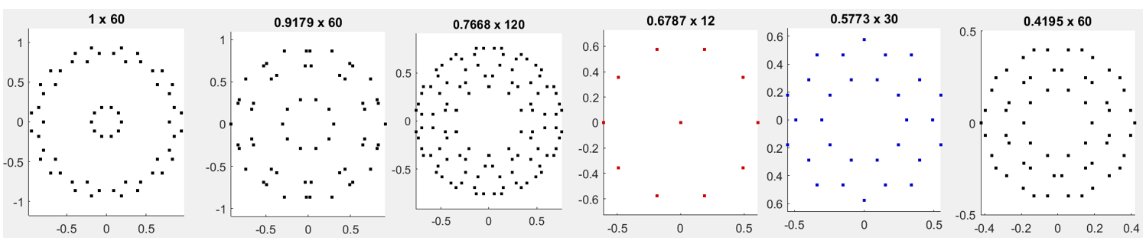


Figure 9: The first point array's various radial levels, descending from left to right. Red dots tell us that the point array takes the form of an icosahedron, blue dots indicate an icosidodecahedron, yellow dots indicate a dodecahedron, and black dots indicate a solid with 60 or 120 vertices.

For the purpose of this investigation, only the point array levels consisting of 60 vertices are going to be classified. Whereas the 5-fold orientation of points is always the same for the icosahedron, dodecahedron, and icosidodecahedron point array clouds, we already see in the previous figure that clouds consisting of 60 vertices clearly take on different spatial orientations, if not completely different shapes themselves. Here are the next two point arrays' decompositions as an exposition:

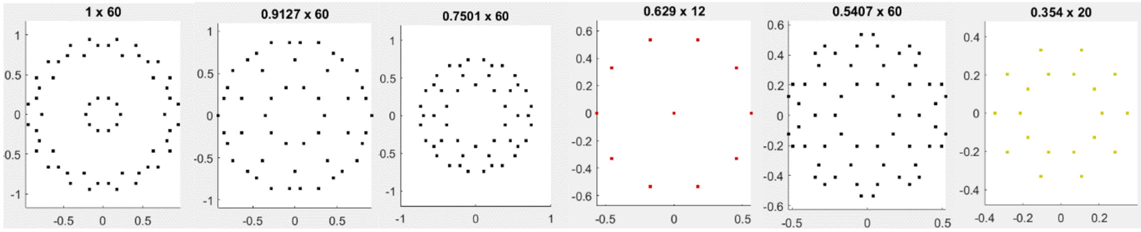


Figure 10: Point array 2, denoted as $\frac{\phi^2}{2}\mathrm{ICO}_2$

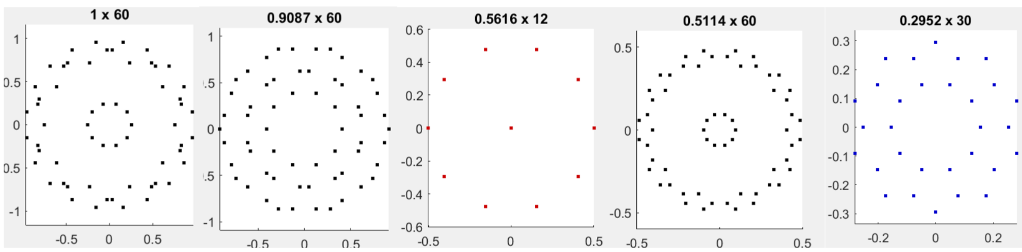


Figure 11: Point array 2, denoted as ICO_2

4 Classification

We already see a few orientations that are similar, and quite a few that are different. To classify these objectively, let's separate each orientation by examining how many points there are at each horizontal level, from top to bottom. For example, in Figure 11, if we look at the 3rd radial level (a red icosahedron), we could denote this as 2-2-3, because there are 2-points at the highest horizon, 2 points at the second horizon, and 3-points in the equator, afterwhich the symmetry is repeated so 2-2-3 encapsulates all the information that we need. If we examine all 55 point array clouds whose radial level has 60 vertices, we the following similar geomotries arise:

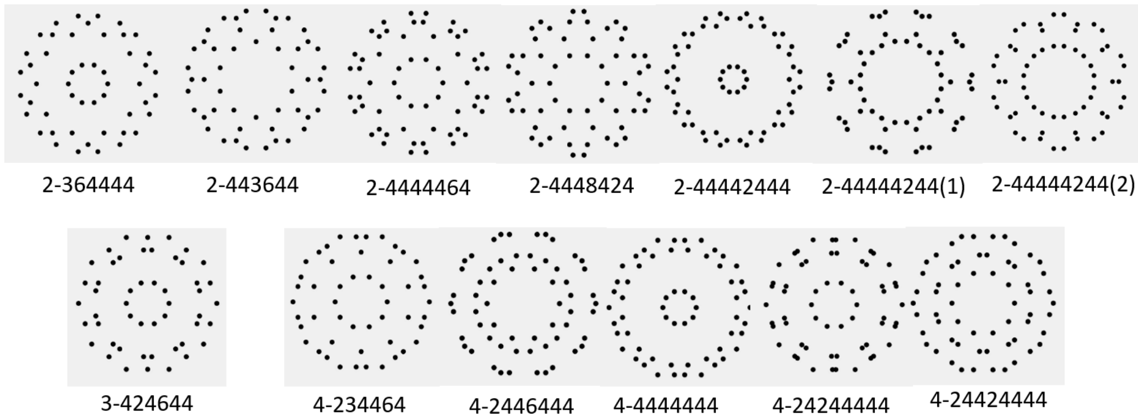


Figure 12: All the different orientations that are present for different radial levels, all clouds having 60 vertices.

See if you can't go back to the three point array radial decompositions and classify the orientations you see. The next step is to find out information about the connectivity of these point clouds. Using Matlab to manipulate these point clouds in three dimensions, I was able to extract information about the connectivity and ultimately about the shapes they formed. Consider the following drawings:

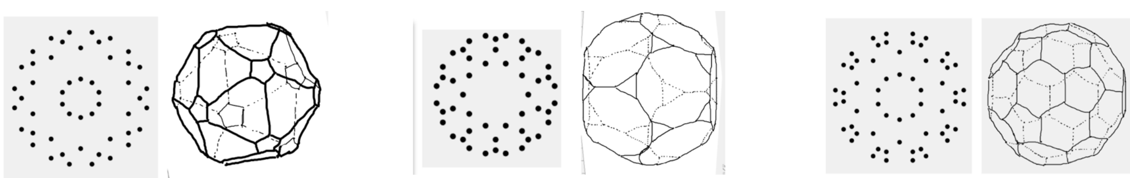


Figure 13: 3D connectivity of sample point arrays

As it turns out, the individual radial levels are represented by the icosahedral Archimedean solids.

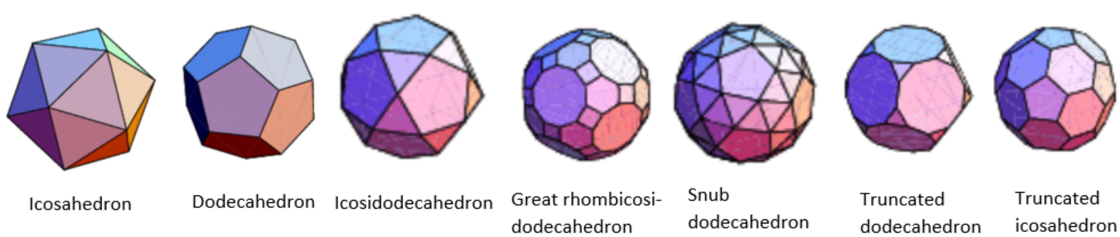


Figure 14: 3D connectivity of sample point arrays. Polyhedron taken from WolframWorld.[6]

Drawing out the rest of the point array clouds by hand (well, technically by mouse) allowed me to group the following point arrays as the same archimedean solid, just in a different spacial orientation:

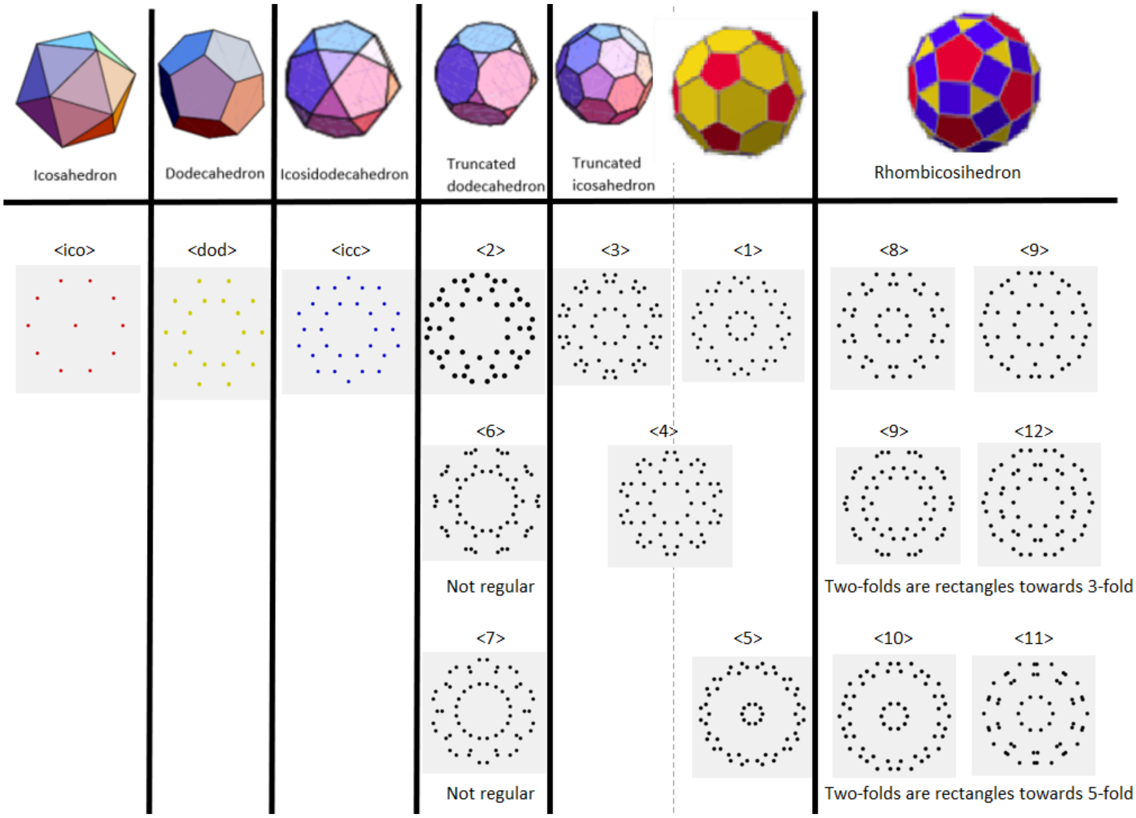


Figure 15: Classification of point arrays viewed from the positive z-axis and the archimedean solids they build, where the bracket notation indicates the order that the pattern appeared in Figure 12

We notice that for the base shapes from which the point arrays are built, there is only one orientation and form in which they appear, and that they are all regular polyhedrons. However, for the radial point arrays with 60 vertices, we see that the principal archimedean solids that they span are: truncated dodecahedrons, truncated icosahedron, and rhombicosihedrons. We also make note of the fact that given there are multiple different unique orderings of points when having a bird's eye view from the positive z-axis used to represent the same solid; this fact indicates that the shapes formed by radial point arrays sometimes have different ratios for their side lenghts, that is to say they are not made up of regular polygons. There are two archimedean solids that are not covered by the radial point arrays of the first 55 point arrays, they are the great rhombicosidodecahedron and the snub dodecahedron.

5 Impact

Why is it important that these point arrays be thouroughly classified? As mentioned before, they can serve as a practical classification system for different viruses, encoding information about the radial distribution of the viral capsid. Grouping viruses into these geometric paradigms allows us to further our understanding of critical points that are essential for a the viral capsid's stability. This exploration was inspired by the geometric structures that I learned about when researching viruses for my physics SIP.

References

[1] G. Firsove *http://microproject.jjht.org/wp-content/uploads/2012/12/Virus-Shapes.jpg*

[2] Viral Zone *https://viralzone.expasy.org/1057?outline=all_by_protein*

[3] E. Voyles *Decomposition of Virus Normal Modes into Spherical Harmonics*. Department of Physics, Kalamazoo College, Kalamazoo, Michigan, United States of America

[4] L. Poudel *Deciphering of packing signal hypothesis in bacterophage RNA recognition by the MS2 capsid protein in virus assembly*

[5] D. Wilson *Unveiling the Hidden Rules of Spherical Viruses Using Point Arrays* Department of Phhysics, Kalamazoo College, Kalamazoo, MI 49006, USA

[6] Wolfram World *https://mathworld.wolfram.com/ArchimedeanSolid.html*

[7] R. Twarock, A. Luque *Structural puzzles in virology solved with an overarching icosahedral design principle*