

Capsids and Stuff

Charles Zahara

June 9, 2015

Contents

List of Figures	3
1 Introduction	4
2 Background	5
2.1 The History of Icosahedral Symmetry	8
2.1.1 Initial Theories of Crick and Watson	8
2.1.2 The Theory of Quasi-Equivalence	11
2.1.3 Recent Work	17
3 Mathematical Problems and Solutions	19
3.1 Triangulation Numbers	19
3.1.1 Solution 1: Geometry	20
3.1.2 Solution 2: Linear Algebra	26
3.2 Protein Shape	31
3.2.1 Solution 1: Euler's Polyhedral Formula	32
3.2.2 Solution 2: Necessary Symmetry	35
4 Conclusion	37
5 References	38

List of Figures

1	Viral Life Cycle	6
2	Examples of Viruses	7
3	Rotational vs Reflexive Symmetry	10
4	Table of cubic point groups	10
5	Capsids Folded From Hexagonal Nets	14
6	Capsid Represented as Hexamer, Pentamer, and Triangular Units	15
7	Calculating T-Numbers	16
8	Classes Based Upon T-Numbers	16
9	Capsid Class Examples	17
10	Equilateral Triangle on h , k -axes	20
11	Area of Equilateral Triangle Constructions	21
12	Calculating the Base	22
13	Base Found	23
14	Locating Point P	24
15	Location of Point P	25
16	Basis Vectors	26
17	Converting b_2 to \mathcal{E}	26
18	Rotation of Basis Vectors e_1 and e_2	29
19	Rotation via Linear Transformation	30
20	$T = 13$ Capsid Model	32
21	Individual Protein Shape	35
22	Protomer Options	36

1 Introduction

Mathematics is a beautiful field full of excitement and discovery. From purely theoretical considerations to strictly applied problems the mathematician is never without mental stimulus. Her inspirations are everywhere and she needs only to be willing to ask why and how to be able to find a fantastic problem to consider. Although the sciences sometimes lack a willingness to work together, Biology is a great realm within which to discover new questions. To find challenging mathematical problems within Biology one does not need to be an applied mathematician with decades of experience in modeling. Nor does one need to have an extensive background in Biology and science. All that is required is a willingness to examine physical systems and ask questions.

The aim of this paper is to take the reader through the discovery of various mathematical puzzles inspired by the physical world. Specifically, these mathematical questions will come from the theory of spherical virus capsids. Problems left unanswered by many of the works on the spherical virus capsids will be laid out and then multiple techniques for solving these problems will be detailed. Although this paper is centered around the mathematics of why icosahedral shells work, it aims to be readable by anyone without external materials, extensive background reading, or a 6' by 20' wall of white board to work out gaps in equations. This paper will ultimately demonstrate that unique problems in mathematics are everywhere and can often be solved with very basic tools if one is willing to apply some hard work and problem solving.

A secondary objective of this paper is to be an introduction to the theory of icosahedral symmetry of virus capsids. It is intended to be a type of tour guide through the development and arguments of this theory. As such, the paper begins with a review of both past and present work on virus capsids. First, the pivotal works on the theory of symmetry in spherical virus capsids will be summarized. Next, a few selections of the current work both biologists

and mathematicians are performing in the field of virus capsids will be examined. Hopefully the reader will come away understanding not only the form and function of icosahedral capsids, but also the amazing, challenging, and ultimately fun mathematical problems that arise within Biology.

2 Background

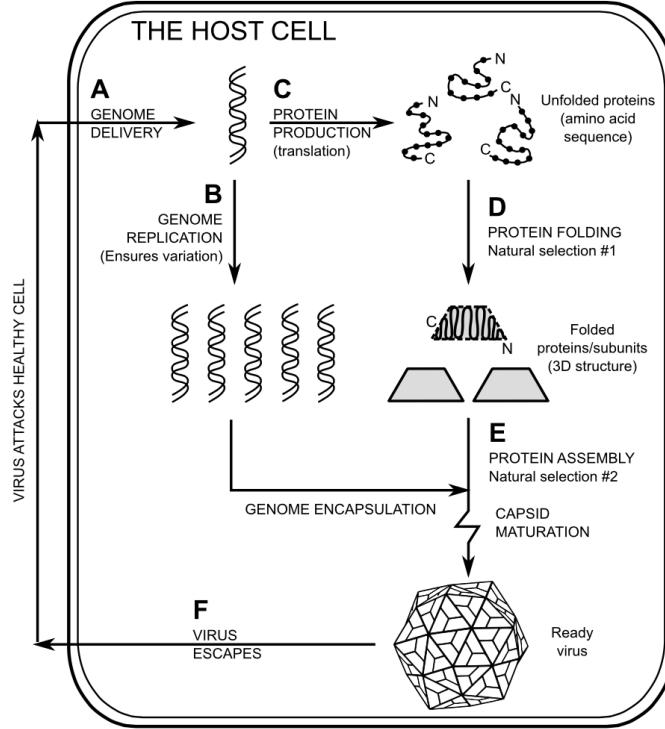
Viruses have been shaping human history for millennia. From smallpox and polio to the yearly strains of flu and common cold, every generation has battled against these tiny creatures. They have wiped out civilizations and continue to kill millions every year. As such the study of their form and function remains important.

Definition (Virus). *Viruses are infectious agents composed of nucleic acid surrounded by a protein shell, called a capsid, and in some cases a membranous envelope.* [1]

While the above definition may make viruses seem like simple little organisms, they have a multitude of unique variations. Their nucleic acid genetic material may come in the form of DNA or RNA, depending on the virus, and be either single or double stranded. They can be rod shaped, roughly spherical, or in some cases rather complex in shape. Additionally, viruses are very effective organisms. They cannot replicate or reproduce themselves, but instead must inject their genetic material into a host cell. The host cell is then taken over and forced to create copies of the virus genetic material and capsid proteins. Eventually, once it is filled to capacity, the host cell bursts open and the newly created capsids are expelled and proceed to a new host to repeat the process (Figure 1). [1]

The purpose of the capsid shell is to protect the virus's genetic material as it moves from host cell to host cell. This makes the capsid crucial in the infection process as it has been shown for some to protect the nucleic acid from ribonuclease and temperature change [3, p 5]. It also means that understanding the structure and mechanics of the capsids may allow

Figure 1: Viral Life Cycle

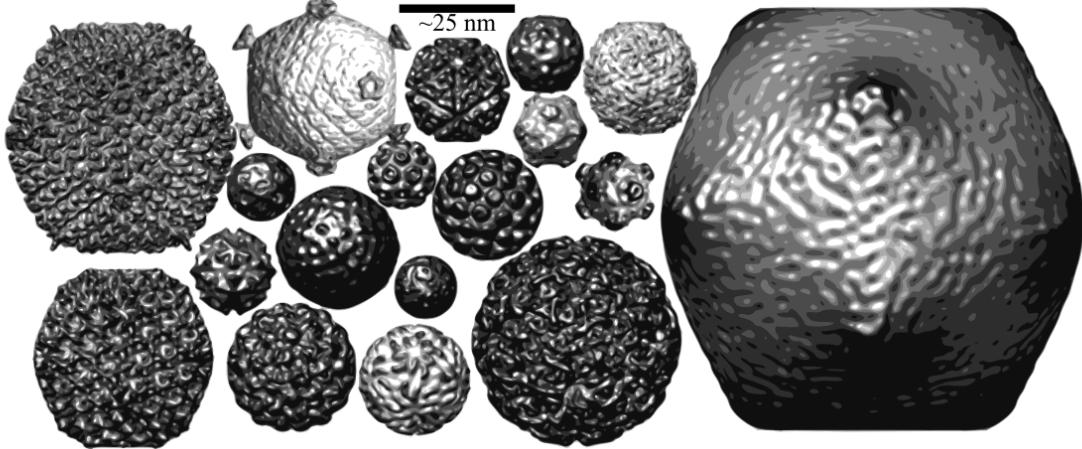


This figure demonstrates the basic life cycle of a virus. First, in **A**, the virus injects its genetic material into the host cell. The genetic material is then replicated by the host, **B**, and used to synthesize proteins, **C**. The proteins then fold into their proper shape, **D**, and irregularities and mistakes are thrown out. Finally, in **E**, the newly created proteins form into a capsid around the genetic material and burst forth from the host cell, in **F**. This figure is taken from *figure 1.1 ref. [2]*.

scientists to develop ways to disrupt their construction, thus eliminating the virus's ability to travel and infect new host cells.

As stated previously, viruses come in a variety of shapes and sizes defined by their capsid shells. Some are long tubes similar to a straw, classified as helical viruses, others are roughly spherical, classified as icosahedral viruses, and others do not fall into either of the two previously mentioned categories and are classified as complex (Figure 2). Spherical viruses are the focus of this paper. In the late 1950's and early 1960's it was suggested that spherical viruses capsids must conform to icosahedral symmetry [4] [3]. Through decades of study and advances in microscope technology the theory that most spherical capsids conform

Figure 2: Examples of Viruses



Above is an assortment of various sized virus capsids. This figure was originally presented as *figure 2.1* in [2]

to icosahedral symmetry has been confirmed both visually and computationally, although there are still exceptions which continue to be examined. What is often left unexplained in papers discussing spherical viruses is the mathematical reasoning behind the theory of their shape.

Before the history and theory of icosahedral symmetry in virus capsids are reviewed a few key definitions are stated below and the relationships between them are discussed.

Definition (Capsid). *A viral capsid is a protein shell made up multiple copies of one, but in some cases a few, small proteins, which encloses and protects the viral genome.*

Definition (Capsomere). *A capsomere is a repeated symmetric building unit for the viral capsid. They are made up of multiple proteins.*

Definition (Protomer). *A protomer is a single protein unit of a capsomere.*

The important note is that the protomeres are the smallest unit. They are single asymmetric proteins. Capsomeres are built up of multiple protomeres in a symmetric fashion. In icosahedral capsids the capsomeres come as either hexamers, which are made up of 6 pro-

motes, and pentamers, which are made up of 5 protomers. The Capsid is the final shell, made up of the capsomere building blocks.

2.1 The History of Icosahedral Symmetry

2.1.1 Initial Theories of Crick and Watson

In 1956 F. H. Crick and J. D. Watson wrote a paper entitled “Structure of Small Viruses.” In this paper they hypothesized that “a small virus contains identical sub-units, packed together in a regular manner” [4, p 473]. While this was not an entirely new idea, Crick and Watson were the first to describe the specifics of the idea and to proclaim it as a feature of all small viruses. This hypothesis was based upon the fact that all plant viruses having been studied at the time were extremely ”regular,” meaning that for a particular virus each particle was similar, if not identical, in appearance. This belief was strengthened by the fact that viruses can form crystals, a feature only possible if each individual viral molecule is identical.

At the time of their writing, viruses had only been observed to follow the rod shape of the tobacco-mosaic virus (TMV), as studied by Tubingen and Berkley, or the spherical shape of the turnip yellow mosaic virus (TYSMV), as studied by Markham. Indeed, more research had been done on the tobacco mosaic virus than any spherical virus at the time and so many conjectures were made from these observations. For instance, the tobacco mosaic virus consists of 94% protein and 6% RNA [4, p 473]. This suggests that proteins must be repeated as there is not enough RNA to code unique proteins in that ratio. Very early X-Ray work showed that the each TMV particle was made up of specific sub-units. More importantly, Crick and Watson were able to observe that the TMV particles had a screw like axis, which acted as a symmetry axis, thus placing each sub-unit in an identical environment as the rest.

While Crick and Watson never used the words capsomere and protomer in their paper, it is beneficial to differentiate them here. In the case of the TMV, the sub-units that Crick and Watson were speaking of were the individual protein units, the protomers. This is evidenced by their description of the sub-unit as the “small globular proteins” [4, p 473]. In the case of helical viruses there are no capsomeres; the repeated sub-unit packed around a helical axis are simply the individual proteins themselves [4, p 474]. It is helpful to recognize these observations of the TMV particles in order to understand the conjectures Crick and Watson made about spherical virus.

Crick and Watson extended their observations of the TMV particles to the less studied spherical viruses. It had been shown that TYMV and bushy stunt virus (BSV) would crystallize into a unit cell that had the shape of a cube. While a mathematical justification is beyond the scope of this paper, identical particles crystalizing into a cube forces the individual virus particles to have cubic symmetry. Crick and Watson were even in communication with D. L. D. Caspar who had been able to clearly demonstrate BSV particles had cubic symmetry [4, p 474].

Additionally, they observed that while the virus particle can be rotated, the individual sub-units cannot be “flipped”. This is due to the presence of asymmetric carbon atoms within the proteins [4, p 474]. This means there can be no mirror symmetry axes on the particle. A mirror axis is one over which a figure can be mirrored and result in the same figure. For a visual representation of rotational versus mirror symmetry axes see Figure 3. Using the previous mentioned observations, Crick and Watson then systematically examined all possible solids which have 3 fold rotationally axes. These are listed in Figure 4. Note that a point group is the group of rotational symmetry operations of a solid. For the purposes of this paper, it can simply be thought of as the ways in which platonic solids can be rotated and remain in the same position.

Through their exploration of cubic point groups they noticed that the number of possible

Figure 3: Rotational vs Reflexive Symmetry

	C_1	C_2	C_3	C_4	C_5	C_6
P	.	d	d d P	d P P	d d P	d d P
A	-	-	-	-	-	-

The above figure demonstrates both rotational and mirror symmetries. The columns show rotational symmetries of order 1 to 6 from left to right. The top row represents a subunit which has only rotational symmetry. Notice there is no way to mirror a **P** and result in the same figure. The second row shows how mirror symmetries are available if the subunit has a mirror axes, represented by the dashed line. Notice how the **A** can be flipped horizontally and the letter is unchanged. This figure originally appeared as *figure 2.6* in [2].

Figure 4: Table of cubic point groups

Solid:	Tetrahedron	Hexahedron	Octahedron	Dodecahedron	Icosahedron
3D shape:					
Symmetry: (elements): Order:	Tetrahedral, T (3,2) 12	Octahedral, O (4,3,2) 24			Icosahedral, I (5,3,2) 60

There are 3 types of cubic point groups with no mirror symmetries. They are classified by the types of rotational axes available. The above figure gives the platonic solids which possesses each type of symmetry, the types of axes, and the order of the group. They are considered *cubic* point groups as the 3 element is the largest common to them all. The order of the point group is the number of unique ways to arrange the points of the solid. The order corresponds to the number of identical subunits which the solid may be broken in to. Note that the dark oblong marks are 2-fold rotational axes, the triangles are 3-folds, squares are 4-folds and pentagons are 5-folds. This figure originally appeared as *figure 2.4* in [2].

asymmetric sub-units was a multiple of 12 for all of the groups. But, with 20 sides the icosa-hedron provided the largest number of sub-units. Each side can be split into 3 asymmetric

sub-units, thus creating a total of 60 which is of course a multiple of 12. Recall the reason they were concerned with the possible number of asymmetric units, not the number of faces, was because they wanted the smallest possible unit, and because the individual proteins were bound to be asymmetric simply due to their nature. Additionally, this reasoning was influenced by their observation of the TMV particles which had only protomers. Thus, they were not considering the forms the proteins may aggregate into before the final shell was constructed, but instead concerned only with the individual units being in identical environments. This was their final, and ultimately erroneous conclusion; spherical capsids are icosahedrons composed of 60 identical asymmetric units. [4, p 475].

It is important to note that their process was one of trying to apply a set of rules to the limited observations they had of both spherical and helical viruses. They were not trying to ask mathematical questions as to whether or not their results should be expected, but were instead simply trying to explain observed patterns. This technique, of describing observations, was also the route taken by Caspar and Klug a few years later, only with a more accurate outcome.

2.1.2 The Theory of Quasi-Equivalence

D. L. D. Caspar and A. Klug wrote many papers on the structure of viruses. Their most prominent work was done in 1962 in a paper entitled “Physical Principles in the Construction of Regular Viruses” [3]. In this work they built upon the ideas and groundwork set forth by Crick and Watson, only with more scientific tools at their disposal. The results validated the general idea set forth by Crick and Watson, but ultimately corrected the overall theory into what is still the basis of the scientific theory that stands today.

Much of their paper repeats in greater detail the same concepts set forth by Crick and Watson. Both groups explained why there may only be a limited number of types of proteins in any one capsid. They also explained why the number of ways to efficiently design a

container using a large number of identical proteins is very restricted. But, Caspar and Klug had greater access to scientific evidence, largely due to the greater volume of research that had been performed and published by 1962. For instance, by studying the weight of the RNA within small viruses and the average number of nucleotides required to make a protein, they were able to prove that there was only enough genetic material in small viruses to produce 2 or 3 types of proteins [3, p 1]. This proved that with small viruses the types of proteins in the capsids was limited to 1 or 2. Additionally, advances in X-ray diffraction and electron microscopy definitively demonstrated that for multiple small viruses were indeed made up of repeated identical sub-units, just as in Crick and Watson's hypothesis [3, p 2]. What differed in theory from Crick and Watson was the number of individual protomers possible in spherical viruses.

While Caspar and Klug ultimately disproved a flaw in Crick and Watson's theory for spherical viruses, it is worth noting that they also confirmed Crick and Watson's theory for helical ones. Caspar and Klug also relied heavily on the study of TMV for their understanding of helical viruses. Through advances in X-ray diffraction and other physical and chemical studies, they proved that each individual protein in a TMV capsid is structurally and chemically identical to the others. They also found that the proteins would form into the helical shell without the presence of the TMV RNA [3, p 4-5]. These pieces of evidence added to the observations that the TMV capsid consisted of repeated units around a helical symmetry axis.

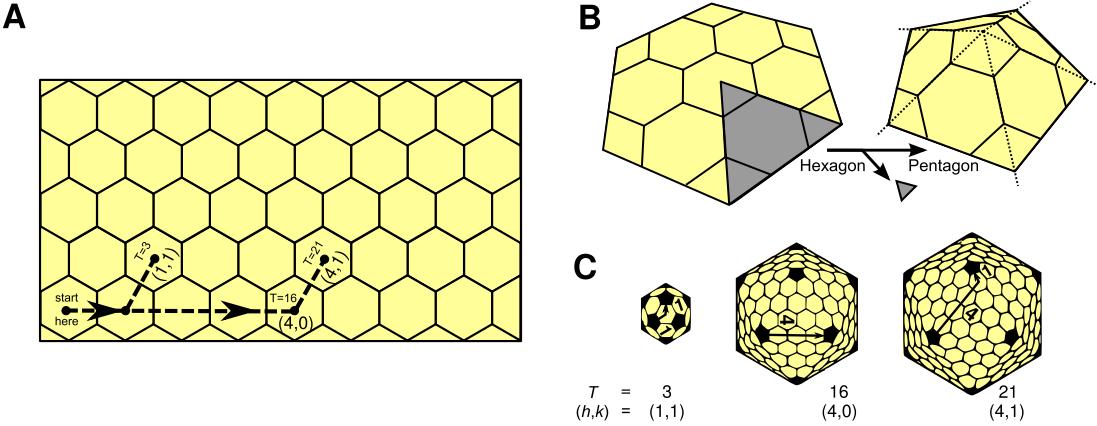
The final important observation made of the helical viruses was that they were flexible. The rod like capsids could bend and flex without breaking. Recalling that the helical capsids are made up only of protomers, the individual proteins, this meant that the inter-protein bonds did not have to be perfectly identical. The bonds were allowed to vary slightly, or be quasi-equivalent [3, p 7]. This discovery helped lead to the theory of quasi-equivalent bonds which allows for scalable spherical capsids.

Caspar and Klug agreed with the rationale of Crick and Watson that the spherical capsid must be of cubic symmetry. Similarly, they agreed that the icosahedron was the likeliest structure as it had the most sides, and the largest possible number of individual protein subunits [3, p 7-8]. While more recent X-ray diffraction photographs had confirmed that some virus particles had icosahedral symmetry, developments in electron microscopes allowed the observation that many capsids had more than 60 protein sub-units [p 9]. They even noted that it is mathematically impossible to place more than 60 identical unit on the surface of a sphere in such a way that each is identically situated [p 10]. This directly contradicted the claims of Crick and Watson, and was the inspiration for quasi-equivalence.

Inspired by Fuller geodesic domes, Caspar and Klug recognized that it is possible to construct icosahedra from hexagons and pentagons. Specifically, they sought to find a way to fold a planar net, that is a plane which is constructed of one repeated shape, into the shape of an icosahedron. Relying on the work of G. S. Pawley's "Plane groups on polyhedra," [5] they determined that it was only possible to fold planar nets made of either squares or equilateral triangles onto the surface of convex polyhedra. Because the square has 4-fold rotational symmetry it could not be used as the goal was to create a polyhedra with 2,3,5 rotational symmetry. Furthermore, it is possible to fold the equilateral triangle net into a convex polyhedra if either 3, 4, or 5 triangles are joined at the vertices, instead of 6 as in the net. Recall that the helical capsids had taught Caspar and Klug that the inter protein bonds could vary slightly and not break. Thus the goal was to deviate as little as possible, and keep everything quasi-equivalent. For this reason they chose to use all 5-vertices, which results in the icosahedron [3, p 11-13]. This process can also be thought of as taking a hexagonal net and removing $\frac{1}{5}$ of specific hexagons to make pentagons. The folding of such a net will also result in an icosahedron (Figure 5).

To get from an icosahedron made up of equilateral triangles, to a capsid constructed of pentamers and hexamers, it must be remembered that the equilateral triangle can be split

Figure 5: Capsids Folded From Hexagonal Nets

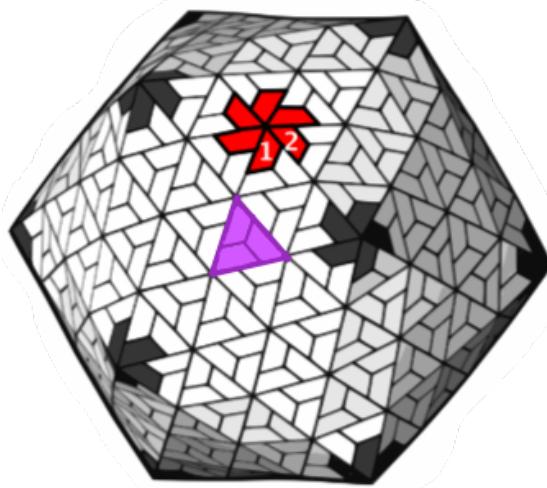


The above figure represents the process of removing triangular facets from a hexagonal grid to construct icosahedra. **A** gives an example of the hexagonal grid, and the process of counting the T number from pentamer to pentamer. **B** shows how removing a single triangular faces turns the hexagon into a pentagon. **C** shows the resulting capsids from the various T number selections represented in **A**. This figure comes from *figure 2.2* of [2].

into 3 identical sub-units, refer to Figure 6. The equilateral triangle unit that Caspar and Klug spoke of is highlighted in purple. This unit can be broken up into 3 trapezoidal sub-units. These trapezoidal sub-units are the individual protomers. Notice that when 5 of the triangles come together, there are really 5 protomers coming together to form a pentamer, highlighted in grey, and when 6 triangles come together there are really 6 protomers coming together to form a hexamer, highlighted in red. So while the capsid was constructed by folding an equilateral triangle net, it can be viewed as being entirely constructed from the red and grey capsomeres. This is the essence of quasi-equivalence, that the biological difference between the 5-fold bonds and the 6-fold bonds are insignificant.

The final important product of Caspar and Klug's theory of quasi-equivalence was the T-number, or *triangulation number*. An icosahedron has 20 faces. When the icosahedron is broken into smaller equilateral triangles, it will have $20T$ smaller triangular facets. To determine the T-number for a specific capsid, begin at one of the protomers, the 5-fold vertices of the capsid, and then count, in 2 straight lines, the number of 6-fold vertices,

Figure 6: Capsid Represented as Hexamer, Pentamer, and Triangular Units



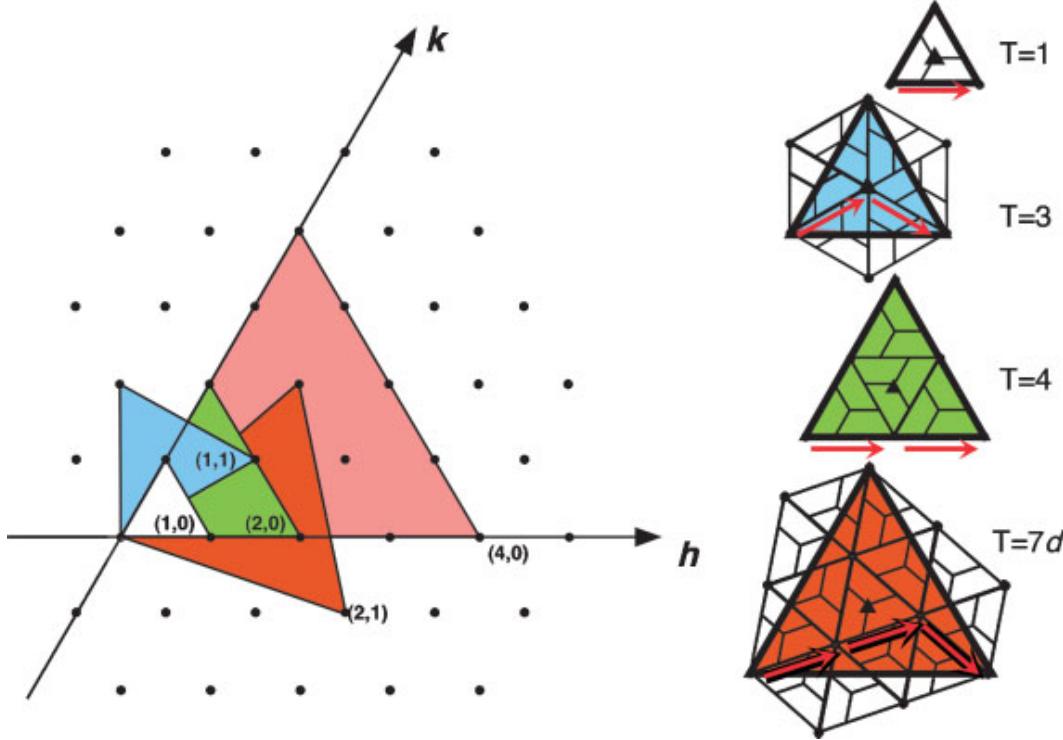
In the above image pentamers are represented in grey, a single hexamer is highlighted in red, and a single triangular facet is highlighted in purple. The individual trapezoidal protomers are left white when not being highlighted as part of a larger unit. This figure is adapted from *figure 1* in [6].

or hexameres, that it takes to get to the next 5-fold vertex. This process is demonstrated in Figure 5. Now consider the movement taken in one direction to be the number h , and the movement in the other direction to be k . Then the T-number for the specific capsid examined will be $T = h^2 + hk + k^2$. Notice that this was not explicitly proven by Caspar and Klug, but only stated to be true.

T-numbers are important because they classify the size of the capsid by determining the number of individual proteins. Since the capsid will have $20T$ equilateral triangle facets (the purple triangle in Figure 6) and each facet consists of 3 protomers, the resulting capsid will have $P = 3 * 20T = 60T$ individual protein units. Knowing the overall number of proteins also allows for the calculation of the number of hexameres. Each icosahedral capsid must have 12 5-fold vertices, since it is an icosahedron. Thus there are $12 * 5 = 60$ proteins used up in pentamers, leaving $60T - 60 = 60(T - 1)$ proteins for the hexamers. Since each hexamer is 6 proteins there are $\frac{1}{6}60(T - 1) = 10(T - 1)$ hexamers per capsid.

It is very powerful to state that the T-number uniquely defines the number of proteins per

Figure 7: Calculating T-Numbers

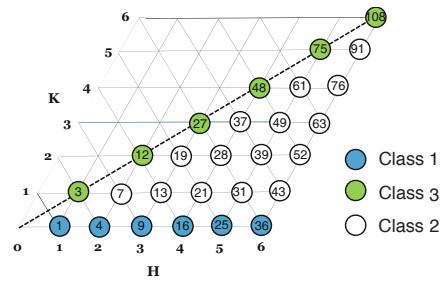


Various T-numbers are shown with their calculations. The right hand side of the figure shows how to count from one pentamer to the next on the actual capsid, whereas the left hand side shows how that counting relates to a face of the icosahedron placed on the h, k -axes. The figure is courtesy of *figure 2* [7].

capsid.

This says that given an equilateral triangle grid with 2 axes, label them the *h-axis* and *k-axis*, oriented at 60° to one another, if an equilateral triangle is drawn with one vertex at the origin, $(0, 0)$, and another at point (h, k) , then the number of smaller triangular units which fit inside the larger one is given by T (Figure 7). While this is easy to confirm in the case where either h or k is zero, it is more difficult when $h, k \neq 0$.

Figure 8: Classes Based Upon T-Numbers



T-numbers are represented by their place on the h, k -axes and color coded by Class. This figure is taken from *figure 5* in [8]

The T-number classification holds more than just the overall capsid size, it also describes the “twist” in the capsid. There are 3 general classifications of icosahedral capsids. Class 1 is where $h > k = 0$. These capsids form perfect icosahedrons. Class 2 are the skew classes, where $h > k > 0$. And Class 3 is where $h = k$ [9] [8, p 22-24]. Figure 8 provides a visual representation of the classes based upon their T-number. This classification has provided a direction in which to explore why some T-numbers, specifically those of Class 3, are less prevalent in natures [9]. Figure 9 provides a visual representation of the differences in the classes.

Figure 9: Capsid Class Examples

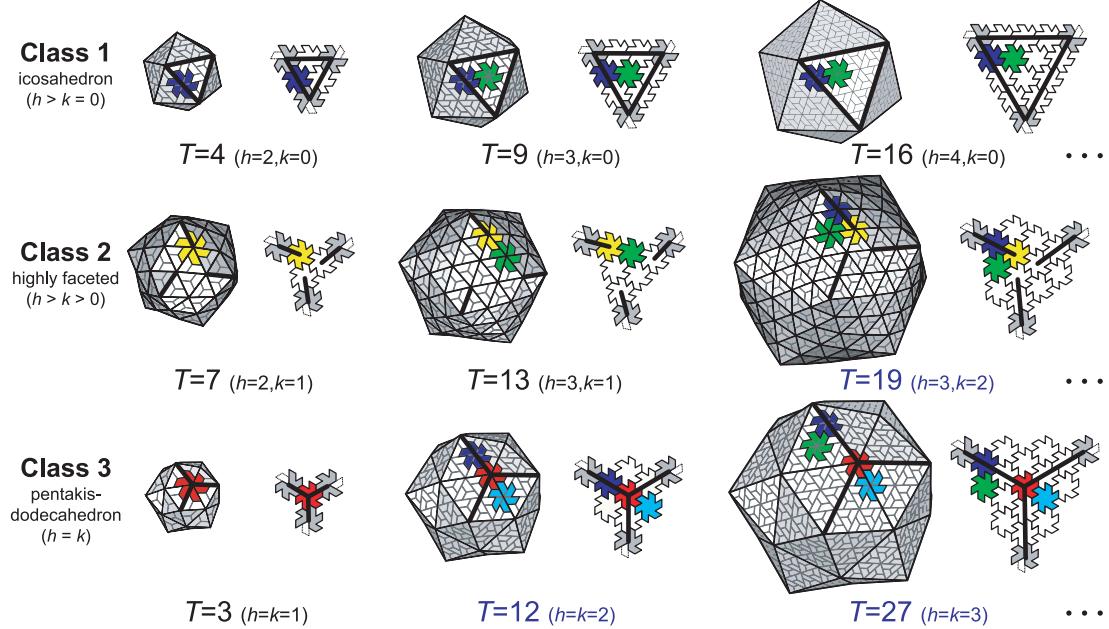


Figure originally printed as *figure 2* in [9]

2.1.3 Recent Work

While the focus of this paper is on a cursory understanding of icosahedral virus capsids and the mathematical questions which this theory inspires, it is beneficial to note some of the current work which is being performed in the field virus capsids. Reidun Twarock is one of

the foremost current mathematical biologists working to understand abnormal virus capsids. She is applying methods from the field of quasicrystals to provide new tiling methods for the abnormal capsids. As previously mentioned, there are capsids which appear spherical, but do not conform to icosahedral symmetry. One particular addition is allowing for translations, instead of only rotations as in the Caspar and Klug theory, in the group movements of the capsid tiles [10] [11].

An additional question left unanswered by the Caspar and Klug theory is why larger T-number capsids and Class 3 capsids are less frequent in nature. Ranjin Mannige and Charles L. Brooks III have worked to answer these questions. They have mathematically modeled the stress placed on the inter-protein bonds caused by a capsid being of a skew, or large T-number. As the capsid size goes up, or skewness increases, the angle between the proteins in the pentamers versus the hexamers begins to vary more and more. Notice how in Figure 9 the Class 3 capsids have more rigid peaks. They argue that this strain caused by the necessity of the altered angles explains for natures selections away from heavily skewed capsids [6].

Finally, while more and more is understood about the limitations of the capsid shape, much is still to be learned about the physical and chemical properties of the capsid construction process. Adam Zlotnick has used statistical mechanical models to examine the details of the capsid assembly chemical reactions. He has gone further to attempt to model how populations of capsids interact with one another [7]. In a simmilar fashion, Roya Zandi et all have also provided a model for capsid assembly. Their model has also gone one step further to provide mechanisms for viral genome release [12].

While most of these works are beyond the scope of this paper, it is worth noting that although the theory being examined here is about 50 years old, it is still a prominent field of science today. Understanding how viruses transmit their genetic material is extremely important, and scientists will continue to study capsid structure and properties so long as

there are viruses to be cured.

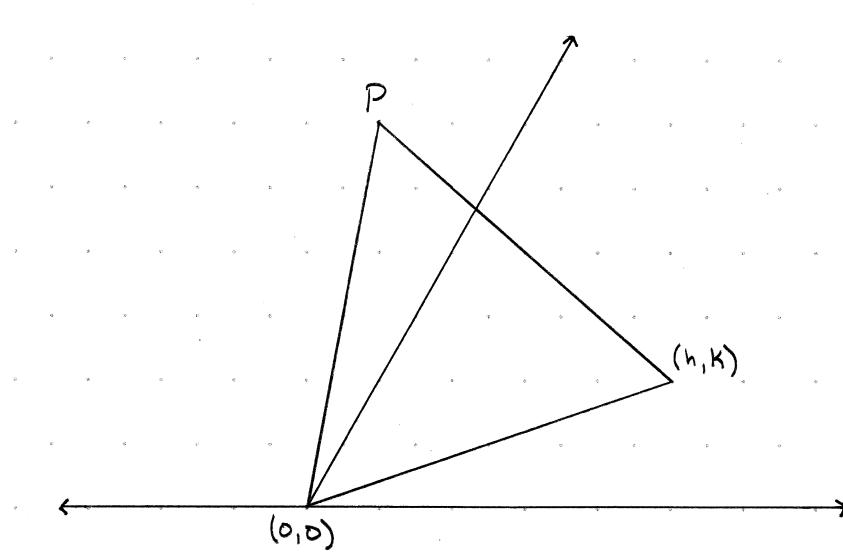
3 Mathematical Problems and Solutions

Recall that one of the goals of this paper is to explore the fun mathematical problems that can be inspired by biology. Thus, in the following sections mathematical problems from two different areas of icosahedral capsid theory will be explored. Multiple solution techniques will be given for the problems with varying degrees of mathematical difficulty.

3.1 Triangulation Numbers

There are two problems which arise from Caspar and Klug's definition and claims about the T-number. First, Caspar and Klug claimed that the T-number, $T = h^2 + hk + k^2$, defines the exact number of identical asymmetric protein sub-units (P) found in a spherical capsid through the formula $P = 60T$. Recall the 60 is because the T number measure the number of triangular facets per side of the icosahedron and each facet is composed of 3 protomeres, the protein sub-unit. Since an icosahedron has 12 sides, we would have $3 * 12 * T = 60T$ proteins per capsid. While their T-number has been confirmed as accurate through better electron microscopes, the mathematical question remains, was this expected? Will the T-number always work for integers h and k ? The first problem will be to establish that using points $(0, 0)$ and (h, k) as the base of an equilateral triangle will result in a triangle with an area equal to T times the area of an equilateral triangle with unit length, that is base equal to one. The second question that must be asked is, if two corners are placed at $(0, 0)$ and (h, k) will the third corner, call it point P, always fall on an integer valued point (Figure 10)? Two different approaches will be used to answer each question. One requires no mathematics beyond high-school, and the second only requires some knowledge of linear algebra.

Figure 10: Equilateral Triangle on h, k-axes



One face of a $T = h^2 + hk + k^2$ icosahedral capsid is represented on the h, k coordinate system. Recall when working with T numbers that the h-axis is in the standard horizontal orientation, and the k-axis is at 60° to the h-axis in a counter clockwise rotation.

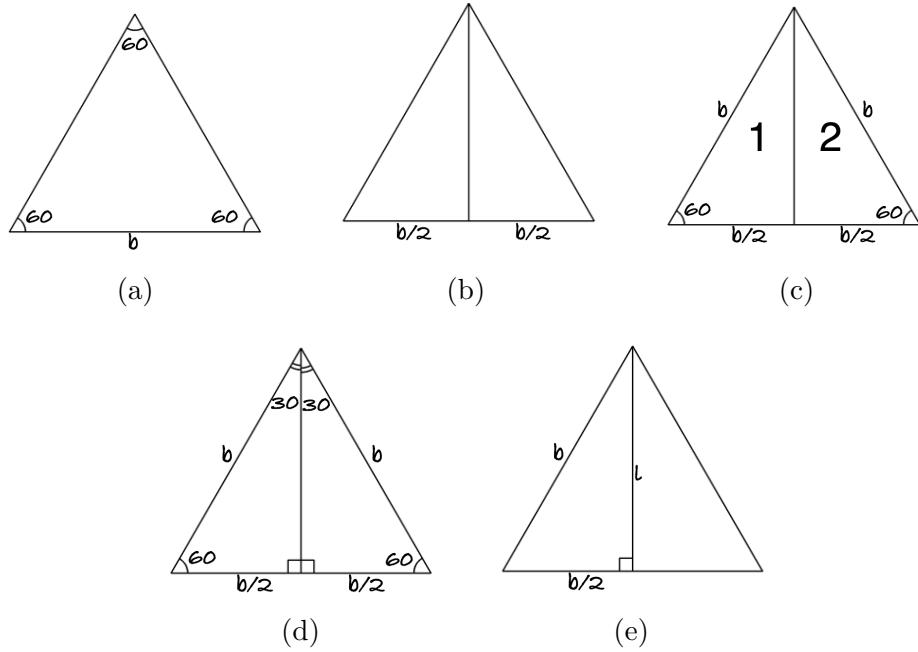
3.1.1 Solution 1: Geometry

The goal of this approach is to establish that if given a point (h, k) as defined by Caspar and Klug (that is that the h-axis is the standard x-axis and the k-axis is a 60° rotation counter-clockwise of the h-axis) that the area of an equilateral triangle with a base from $(0, 0)$ to (h, k) will equal the area of an equilateral triangle with base length 1 times T where $T = h^2 + hk + k^2$. This first solution provided utilizes no more skills than someone who has passed high school geometry.

First, recall that the area of an equilateral triangle is $A = \sqrt{3}/4 b^2$ where b is the length of any side of the triangle. If this formula is unknown to the reader, it is a simple and beneficial exercise to re-establish it. To begin, take an equilateral triangle with side length b (Figure 11a). Draw a line from any corner to the midpoint of the opposing side (Figure 11b). Now there are 2 triangles. Recalling the triangle congruency theorems, by the side-angle-side theorem these triangles are congruent (Figure 11c). This forces the angle that was cut

with the drawn line to be bisected. Since the starting triangle was an equilateral triangle which has all angles equal to 60° , the top angles are both 30° making the constructed line perpendicular to the base (Figure 11d).

Figure 11: Area of Equilateral Triangle Constructions



To find the length of the constructed line l (Figure 11e) apply the pythagorean theorem:

$$b^2 = \left(\frac{b}{2}\right)^2 + l^2$$

$$l^2 = b^2 - \frac{b^2}{4}$$

$$l^2 = \frac{3}{4} b^2$$

$$l = \sqrt{\frac{3}{4} b^2}$$

$$l = \frac{\sqrt{3}}{2} b$$

Thus, using the standard triangle area formula of $A = \frac{1}{2} \text{base} * \text{height}$ we get

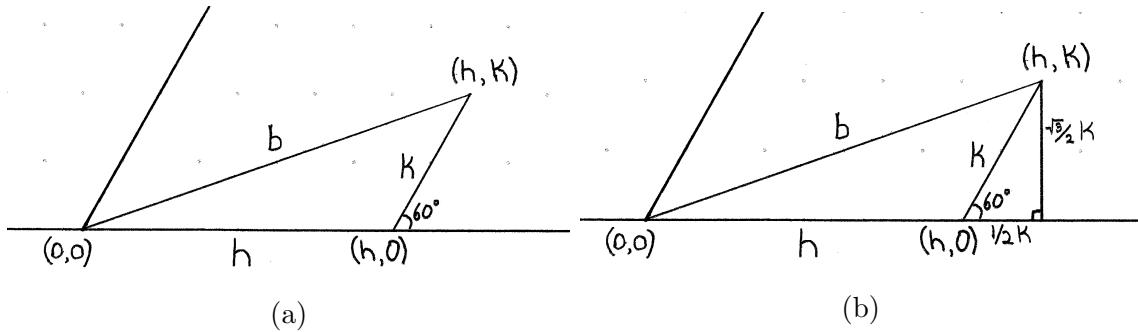
$$A = \frac{1}{2} b * \frac{\sqrt{3}}{2} b$$

$$A = \frac{\sqrt{3}}{4} b^2$$

Recall the goal is to establish that the area of an equilateral triangle with a base from $(0, 0)$ to (h, k) (Figure 10) will equal the area of an equilateral triangle with base length 1 times $T = h^2 + hk + k^2$. So from the above formula with $b = 1$, the resulting area should be $\frac{\sqrt{3}}{4} (h^2 + hk + k^2)$.

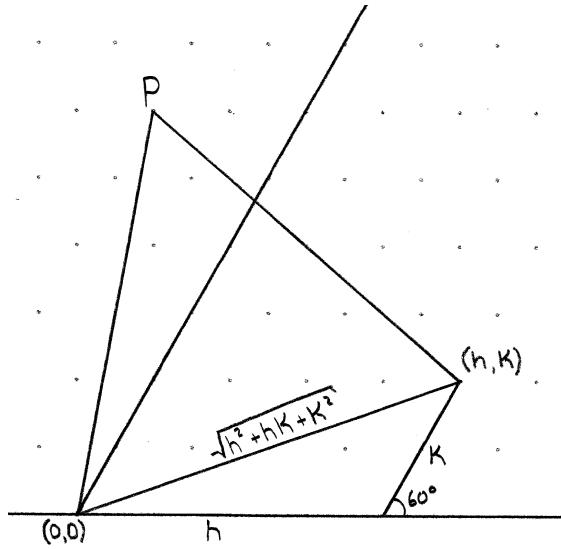
Now consider the triangle created by drawing lines between $(0, 0)$, (h, k) and $(h, 0)$. Two of the side lengths are known and side length b remains to be found (Figure 12a). From the setup of the grid, the exterior angle at point $(h, 0)$ is 60° , so if a line is dropped down from point (h, k) perpendicular to the t-axis the result is a triangle that was just worked with and has known dimensions. Since it is k units from $(h, 0)$ to (h, k) , the perpendicular just drawn has length $\sqrt{3}/2 k$ and the final leg of the triangle is $1/2 k$ (Figure 12b).

Figure 12: Calculating the Base



Now the length in question, from point $(0, 0)$ to (h, k) , is the hypotenuse of a right triangle with side lengths $\sqrt{3}/2 k$ and $h + 1/2 k$. Once again the length lies at the end of the pythagorean theorem (Figure 13).

Figure 13: Base Found



$$\begin{aligned}
 b^2 &= \left(\frac{\sqrt{3}}{2} k\right)^2 + \left(h + \frac{1}{2} k\right)^2 \\
 &= \frac{3}{4} k^2 + h^2 + hk + \frac{1}{4} k^2 \\
 &= h^2 + hk + k^2
 \end{aligned}$$

$$b = \sqrt{h^2 + hk + k^2}$$

Plugging this into the Area formula for b , the result is $A = \frac{\sqrt{3}}{4} (h^2 + hk + k^2)$, which is exactly the desired area from before! So this definitively shows that the area of an equilateral triangle with base from point $(h, 0)$ to (h, k) is equal to the area of $T = h^2 + hk + k^2$ equilateral triangles with base length 1. Recall:

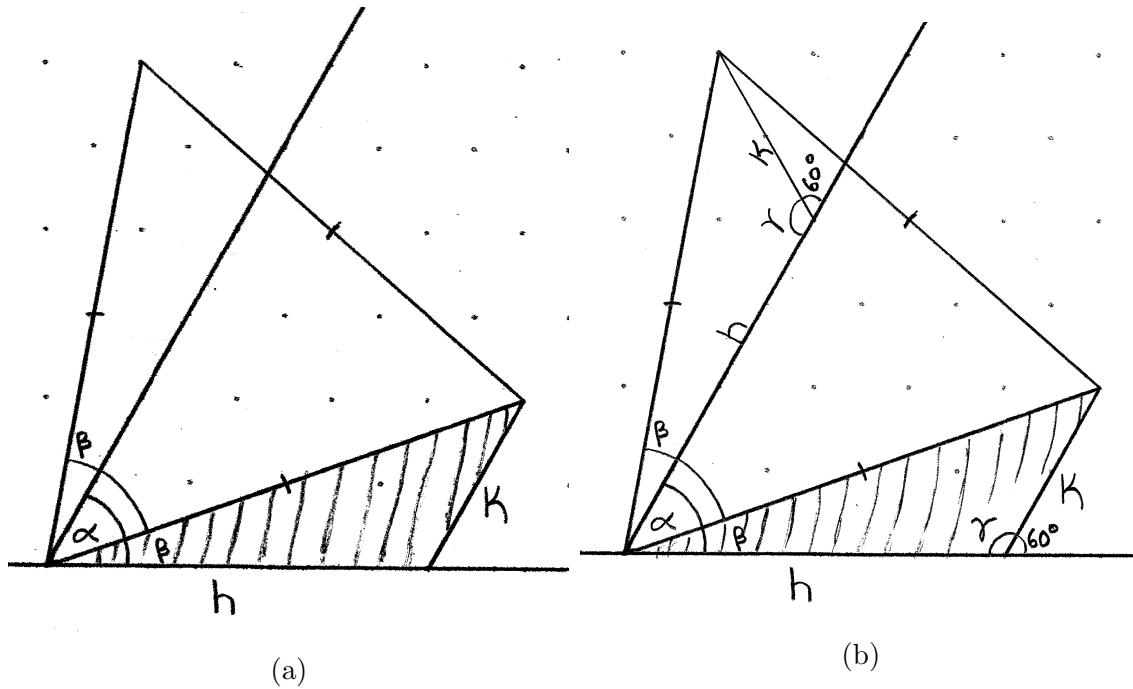
1. The triangle with base from point $(h, 0)$ to (h, k) represents one side of an icosahedron
2. Icosahedrons have 20 sides
3. The base 1 triangle represents one of the triangular facets of the capsid which is made up of 3 proteins

Put it all together and the result is that a total of $3 * 20 * (h^2 + hk + k^2) = 60T$ proteins will fit on the capsid described.

While the previous work answered the first question, it is still unknown where point P is located, or if it is even an integer valued point (recall Figure 13). A quick examination of the angles at the origin provide a good starting place from which to answer this question. The angle created by lines $\overline{P(0,0)}$ and $\overline{(0,0)(h,k)}$ is 60° since the triangle is an equilateral. Similarly, the angle created by the axes is 60° by definition. Now both of these angles share a common angle from the k -axis to line $\overline{(0,0)(h,k)}$, denoted α . This means that the remaining portion of each angle must be the same, denoted β , where $\alpha + \beta = 60^\circ$ (Figure 14a). Since all the sides of triangle $P(0,0)(h,k)$ are the same, one path would be to try and recreate the lower, shaded triangle above the k -axis.

To recreate the lower, shaded triangle, create a line from point $(0, h)$ to point P . Notice even without preset marks on the axis, this could be done simply by copying the distance h

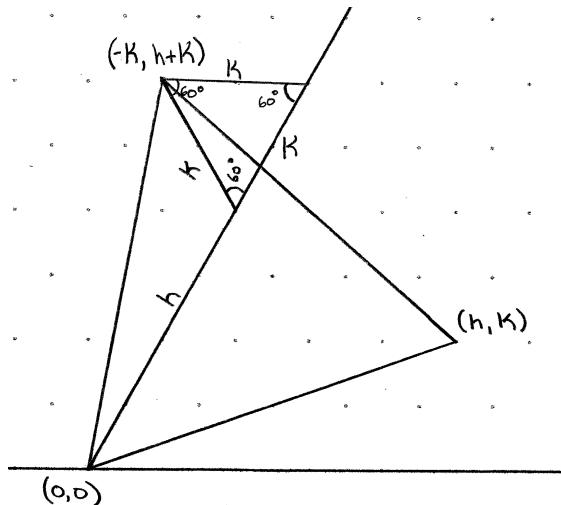
Figure 14: Locating Point P



using a compass and marking that distance from the origin on the k-axis. Now the resulting triangle has sides with length $\sqrt{h^2 + hk + k^2}$ and length h which come together at an angle of β . This matches the dimensions of the shaded triangle. Thus by the side-angle-side triangle congruency theorem the two smaller triangles are congruent (Figure 14b). This forces the unknown side length to be k , and the angle on the k axis to be 120° , marked γ .

Notice that since $\gamma = 120^\circ$, the exterior angle of the constructed triangle at point $(0, h)$ is 60° . This makes it possible to construct an equilateral triangle using the k-axis and the line $\overline{P(0, h)}$. Connect point $(0, h + k)$ to point P . Since two sides have length k and they come together at a 60° angle, the constructed triangle must be equilateral with side length k (Figure 15). Now everything needed to fully define point P is available. Begin at the origin $(0, 0)$ and go up the k-axis $h + k$ units, then go left k units, parallel to the h-axis. This means that $P = (-k, h + k)$. Look at that amazing result! The third point of the original triangle (from Figure 10) falls at an integer valued point uniquely determined by the values h and k .

Figure 15: Location of Point P



Both questions raised about the T number have now been answered. It has been fully established that the equilateral triangle with points at $(0, 0)$ and (h, k) will have its third point at $(-k, h + k)$ and have the same area as T equilateral triangles with unit sides. What

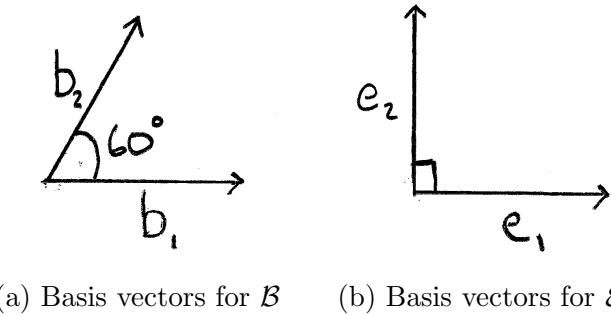
is more, this result was again found using nothing more than basic geometry construction skills.

3.1.2 Solution 2: Linear Algebra

While the geometric solution provided above does work, it can be a bit difficult to discover. It is the result of trial and error using geometric constructions and thus is the product of multiple dead ends. There is a slightly more direct method for solving the same questions by employing some slightly higher mathematical techniques, namely linear algebra. The following will explore that technique.

Because both problems occur in a space defined by the h and k axes which exist at 60° to one another, the first step in attacking these problems will be to establish a matrix for the transformation from the standard basis for \mathbb{R}^2 , call it $\mathcal{E} = \{e_1, e_2\}$, and a basis derived from the h and k axes, call it $\mathcal{B} = \{b_1, b_2\}$ (see Figure 16). This is done by finding a representation for both b_1 and b_2 in \mathcal{E} and then using those representations as the columns of a matrix.

Figure 16: Basis Vectors

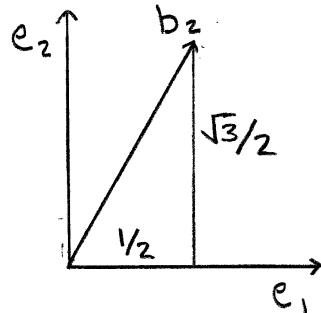


(a) Basis vectors for \mathcal{B} (b) Basis vectors for \mathcal{E}

Let T represent the transformation from basis \mathcal{B} to \mathcal{E} .

For simplicity let $T(b_1) = e_1$ so $T(1, 0) = (1, 0)$. Now using the results previously obtained about equilateral triangles and the fact that we want the basis vectors to be of unit length, it is easy to see that $T(b_2) = \frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$, or

Figure 17: Converting b_2 to \mathcal{E}



$T(0, 1) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ (Figure 17). Recall that the transformation matrix, T , from basis \mathcal{B} to \mathcal{E} is represented by using the two vectors found above as columns in a matrix. So as a matrix:

$$T = \begin{bmatrix} [b_1]_{\mathcal{E}}, [b_2]_{\mathcal{E}} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \quad (1)$$

Similarly, to go from \mathcal{E} to \mathcal{B} simply use T^{-1} . To do so, recall that for any 2x2 matrix A with non-zero determinant:

$$\text{if } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{then } A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\text{thus } T^{-1} = \frac{2}{\sqrt{3}} \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} \quad (2)$$

Now it is possible to freely transition between the bases \mathcal{B} and \mathcal{E} . Since \mathcal{E} is the standard basis for \mathbb{R}^2 , the standard distance formula defined by the Euclidean norm applies. That is, for any vector (x, y) the length of the vector is given by:

$$\|(x, y)\|_{\mathcal{E}} = \sqrt{x^2 + y^2}$$

Recall from Figure 10 that the first goal was to establish the length of vector (h, k) in \mathcal{B} . To do so, simply change bases from \mathcal{B} to \mathcal{E} and then find the distance using the standard formula above, thus $\|(h, k)\|_{\mathcal{B}} = \|T(h, k)\|_{\mathcal{E}}$.

$$T(h, k) = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} h + \frac{1}{2}k \\ \frac{\sqrt{3}}{2}k \end{bmatrix} = \left(h + \frac{1}{2}k, \frac{\sqrt{3}}{2}k \right)$$

Thus

$$\begin{aligned}
\|(h, k)\|_{\mathcal{B}} &= \left\| \left(h + \frac{1}{2}k, \frac{\sqrt{3}}{2}k \right) \right\|_{\mathcal{E}} \\
&= \sqrt{\left(h + \frac{1}{2}k \right)^2 + \left(\frac{\sqrt{3}}{2}k \right)^2} \\
&= \sqrt{h^2 + hk + \frac{1}{4}k^2 + \frac{3}{4}k^2} \\
&= \sqrt{h^2 + hk + k^2}
\end{aligned}$$

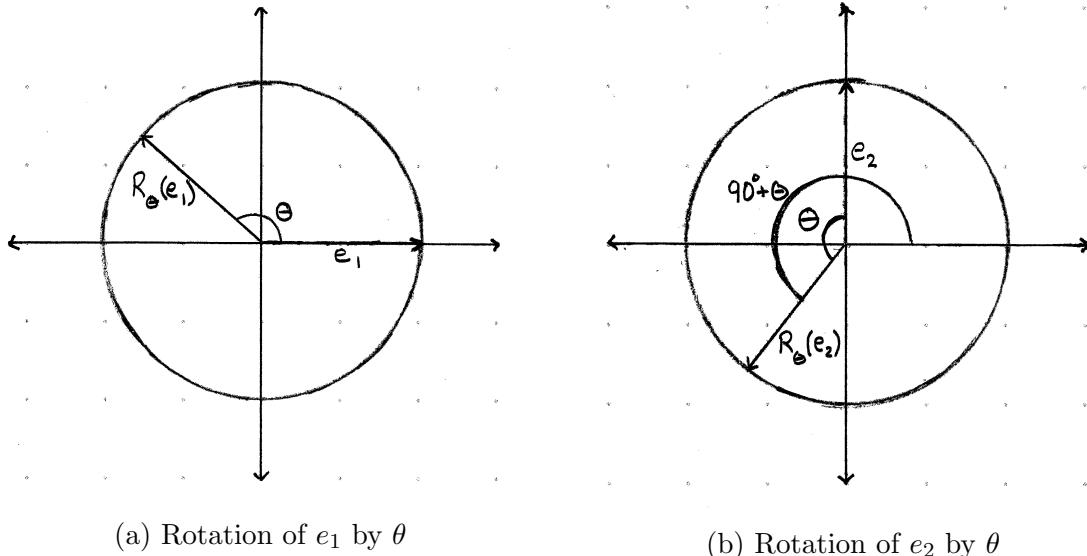
Notice that this is the exact same result obtained from the geometric solution! From here it would be a simple matter to again apply the area formula for an equilateral triangle, and arrive at the conclusion that the triangle originally shown in Figure 10 has $\text{Area} = \sqrt{3}/4 (\sqrt{h^2 + hk + k^2})^2 = \sqrt{3}/4(h^2 + hk + k^2)$. This confirms once again that $h^2 + hk + k^2$ equilateral triangles with unit sides fit in a an equilateral triangle with one side from $(0, 0)$ to (h, k) .

Approaching the second problem, the location of point P , is once again rather straightforward in linear algebra, although a beneficial side track will be taken during the work. To find the location of point P simply rotate vector (h, k) 60° counter-clockwise about the origin. Doing so will require the use of the standard rotation matrix in \mathbb{R}^2 , so the first, and only, side trip in this problem will be to establish the rotation matrix.

The idea is similar to finding the matrix which allows passage between the bases \mathcal{B} and \mathcal{E} . First, rotate each of the standard basis vectors, e_1 and e_2 , by θ° . Then represent the result as columns in a matrix. This resulting matrix is the rotation matrix for \mathbb{R}^2 , call it R_θ .

Since both basis vectors are of unit length, any rotation about the origin will leave them on the unit circle (Figure 18). This means that when e_1 is rotated by θ , the result is the point on the unit sphere given by θ , which is $(\cos \theta, \sin \theta)$ (Figure 18a). Thus $R_\theta(e_1) = (\cos \theta, \sin \theta)$.

Figure 18: Rotation of Basis Vectors e_1 and e_2



Notice that if e_2 is rotated by θ the resulting point is equivalent to the point on the unit circle defined by $\theta + 90^\circ$. Thus $R_\theta(e_2) = (\cos(\theta + 90), \sin(\theta + 90))$. At this point R_θ could be represented by the matrix:

$$\begin{bmatrix} \cos \theta & \cos(\theta + 90) \\ \sin \theta & \sin(\theta + 90) \end{bmatrix}$$

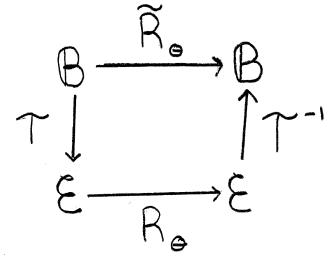
But this matrix is a bit more cumbersome than desired and can be cleaned up using some simple trigonometric identities. While they shall not be proven here, the identities $\cos(\theta + 90) = -\sin \theta$ and $\sin(\theta + 90) = \cos \theta$ are rather standard and all that are needed to clean up the above matrix. Using these identities, the rotation matrix for \mathbb{R}^2 becomes:

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (3)$$

Now since this matrix was determined using e_1 and e_2 it will only work in \mathcal{E} . Luckily, rotation in \mathcal{B} , represented by \widetilde{R}_θ is equivalent to performing a change of basis from \mathcal{B} to \mathcal{E} , using the rotation matrix R_θ , and then changing bases back to \mathcal{B} (Figure 19). Now let $\theta = 60$ as the goal was to rotate by 60° and use equations 1 2 and 3 as directed by Figure 19:

$$\begin{aligned}
 \widetilde{R}_{60} &= T^{-1} R_{60} T \\
 &= \frac{2}{\sqrt{3}} \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos 60 & -\sin 60 \\ \sin 60 & \cos 60 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix} \\
 &= \frac{2}{\sqrt{3}} \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix} \\
 &= \frac{1}{4\sqrt{3}} \begin{bmatrix} \sqrt{3} & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & \sqrt{3} \end{bmatrix} \\
 &= \frac{1}{4\sqrt{3}} \begin{bmatrix} \sqrt{3} & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 2\sqrt{3} & 2\sqrt{3} \end{bmatrix} \\
 &= \frac{1}{4\sqrt{3}} \begin{bmatrix} 0 & -4\sqrt{3} \\ 4\sqrt{3} & 4\sqrt{3} \end{bmatrix} \\
 &= \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}
 \end{aligned}$$

Figure 19: Rotation via Linear Transformation



Applying \widetilde{R}_{60} to the vector (h, k) thus rotates the vector by 60° in \mathcal{B} .

$$\widetilde{R}_{60}(h, k) = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} -k \\ h+k \end{bmatrix} = (-k, h+k)$$

So as before, since $\widetilde{R}_{60}(h, k) = (-k, h+k)$, the point P of the initial triangle in Figure 10 is an integer valued point uniquely defined by h and k , given by $P = (-k, h+k)$. Thus two complete solutions have been provided to the questions arising from Capsar and Klug's T-number.

3.2 Protein Shape

While this problem plagued me during my studies of capsids, it was not one that I was able to solve on my own. The following solutions are adaptations of the explanations provided by Ranjan Mannige in [6, p 5-7].

The second problem is not one that comes directly from the theory of quasi-equivalence or icosahedral symmetry. Instead, the second problem is a result of the standard images used in most papers on spherical viruses. When displaying capsids, most papers represent the protomeres as trapezoids (see for instance Figure 1, 6, 7, 9) with little to no explanation. The simplest explanation is that it is now possible to see the proteins in such clarity that they are known to resemble trapezoids. But this is bothersome. Should such a mathematically regular result have been expected and can it be explained? To state it another way, is there a mathematical justification for the protomer being represented by a trapezoid? Again two methods for answering this question will be provided. The first will require no mathematical skills beyond that of a high-schooler, only some problem solving and algebraic manipulation of formulas. The second, while it can be viewed as a product of group theory, will boil down to nothing more than understanding symmetry.

3.2.1 Solution 1: Euler's Polyhedral Formula

Euler's polyhedral formula is a standard, well known formula which relates the number of vertices (V), edges (E), and faces (F) of a simply connected, convex polyhedra. For the purpose of this work it is acceptable to understand simply connected as meaning, "without holes." This formula is given below:

$$V - E + F = 2 \quad (4)$$

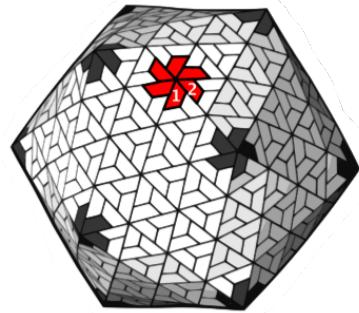
Recall from Section 2.1.2, that Caspar and Klug's theory states that all icosahedral virus capsids are made up of 12 pentagonal capsomeres (grey in Figure 20) and $10(T - 1)$ hexagonal capsomeres (red in Figure 20). Additionally, the protomers come together in groups of 3 to form equilateral triangles since the overall capsid has 3-point rotational symmetry axes. This information, along with Equation (4) should be enough to shed some light on the number of possible sides per protomer, call this number σ for "sides".

Since the number of faces of the overall polyhedra corresponds to the number of individual proteins the following identity holds:

$$F = P = 60T \quad (5)$$

Additionally, because the capsid has no holes, each edge must be shared by two faces. So the overall number of edges would be the number of sides per protein times the number of

Figure 20: $T = 13$ Capsid Model



This figure is adapted from *figure 1* in [6].

proteins divided by two, since there are two faces per edge. Thus:

$$E = \sigma \frac{F}{2} = \sigma \frac{60T}{2} = 30T\sigma \quad (6)$$

Now when it comes to the vertices, a little ingenuity is required. The theory of quasi-equivalence states that there will be 12 5-point vertices and $10(T - 1)$ 6-point vertices at the centers of the pentamers and hexamers respectively. Since these vertices occur at the center of the capsomeres let them be labeled as V_C . This means that

$$V_C = 12 + 10(T - 1) = 10T + 2 \quad (7)$$

The remaining vertices (V_R) must all be 3-point vertices, that is where 3 edges come together, as the final figure has 3-point rotational symmetry. So for each protomer, one of its edges is used up in connecting to the capsomer's center vertex, the V_C , and of the remaining $(\sigma - 1)$ sides, each one goes to a 3-point vertex which is shared with 2 other proteins. Putting this all together, remembering there are $60T$ protomers, gives:

$$V_R = \frac{60T}{3}(\sigma - 1) = 20T(\sigma - 1) = 20T\sigma - 20T \quad (8)$$

The vertices must either be the center of a capsomere (V_C), or not (V_R), and thus the total number of vertices is given my combining Equations (7) and (8).

$$V = V_C + V_R = 10T + 2 + 20T\sigma - 20T = 2 + 20T\sigma - 10T \quad (9)$$

Finally, E , F , and V are defined in terms of σ and T , and thus substituting Equations

(5), (6), and (9), into Equation (4) should provide a definition for σ in terms of T .

$$2 = V - E + F$$

$$2 = 2 + 20T\sigma - 10T - 30T\sigma + 60T$$

$$0 = 50T - 10T\sigma$$

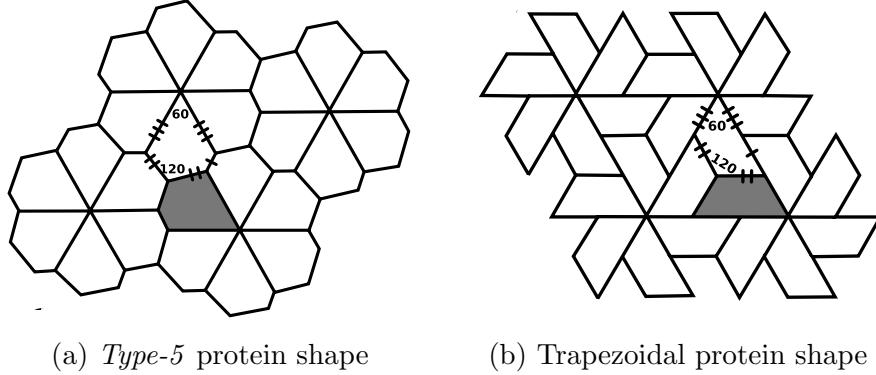
$$10T\sigma = 50T$$

$$\sigma = 5$$

This means that the individual protein units will have 5 sides, or edges, regardless of the T-number! While this is an amazing result, it is not quite what was expected. The goal was to explain why the standard protomer is represented as a trapezoid, but trapezoids have 4 sides, not 5, so what is going on? In the paper “On the Tileable Nature of Virus Capsids and the Role of Topological Constraints in Natural Capsid Design,” Ranjan V. Mannige explains that there are only a limited number of ways to cover convex pentagons with non-overlapping tilings. Furthermore, there is only one type of tile with 5 sides which can cover the icosahedron. It is classified as a *Type-5* and represented in Figure 21a. From here, all that is needed is to note that this tiling is equivalent to the trapezoidal tiling seen in nature, represented in Figure 21b.

The two different protein shapes represented in Figure 21 are equivalent because the base of the trapezoidal unit is split into two edges. When considering vertices, any place where edges diverge is a vertex. This means that the base of the trapezoidal unit must be split by a vertex because the edge of two other adjoining trapezoids join into the base. So, the natural occurrence of trapezoidal protein units within icosahedral capsids is not a random occurrence, but is to be expected due to Euler’s Polyhedral Formula (4).

Figure 21: Individual Protein Shape



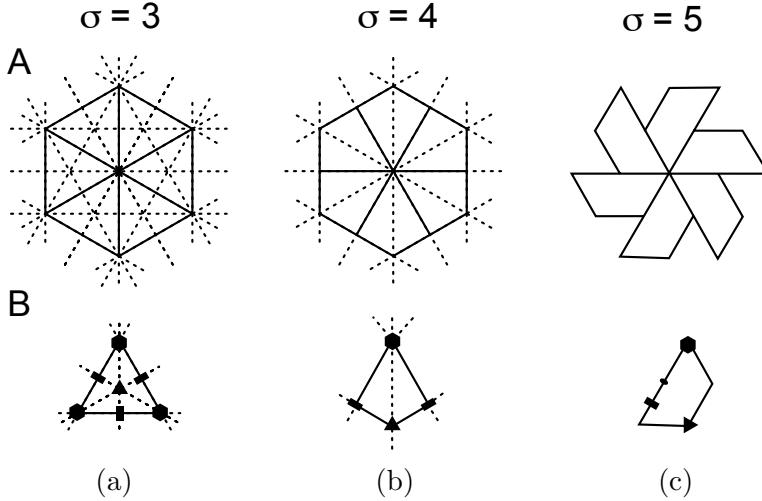
Note that the two shapes are topologically equivalent. Figure is adapted from figure 5 in [13].

3.2.2 Solution 2: Necessary Symmetry

A second technique for justifying the need for 5 sided, and therefore trapezoidal, protein units is by examining the symmetry needs of the overall capsid. It was established by Caspar and Klug, and explained previously, that the spherical capsids conform to icosahedral symmetry. Additionally, it was noted that due to the physical nature of the protomers, mirror symmetry axes are not possible because the individual proteins have a specific left and right hand orientation. Recall from Section 2.1.1, the asymmetric nature of the carbon atoms present in the capsid proteins does not allow for mirror axes. This meant that out of the two types of icosahedral symmetry, full or rotational, only the rotation symmetry was possible. The importance of this fact is that the capsid may not have any mirror axes. Knowing that six protomers come together to form the hexamer, it is possible to systematically examine the available numbers of sides for the proteins, and eliminate those that will not work.

First consider a simple 3 sided protomer, shown in Figure 22a. An equilateral triangle is the only 3 sided polygon which provides the hexamer with 2, 3, and 5-fold rotational symmetry axes. But, an equilateral triangle has a total of 6 mirror symmetry axes, as shown. If the triangle is “flipped over” across any dotted line, the resulting shape is the

Figure 22: Protomer Options



Row **A** shows 3, 4, and 5 sided protomer shapes in their hexamer form. Row **B** shows them as single units. Note that the darkened hexagons denote locations of 6-fold symmetry, the triangles denote 3-fold symmetry, and the rectangles denote 2-fold symmetry. The dashed lines denote the mirror, or reflexive, symmetry axes. This figure is an adaptation of *figure 6* in [13].

same as the original. That is to say that the resulting figure could be placed directly over the original. Thus, the 3 sided protein cannot work.

Next, try a 4 sided protomer as in Figure 22b. At first glance this may appear to be the best shape for the protein subunit. Three kites, such as that in the figure mentioned, can come together to form an equilateral triangle. It is known that the protein subunits must be able to come together to form an equilateral triangle since this is face of the icosahedron. Using the simplest shape to do so would make the most sense. But, once again, the kite has 5 mirror axes. For instance, flip the shape over any one of its edges and it can still be set back upon the original figure. So, as with the 3 sided protomer, the 4 sided will not work.

Finally, consider a 5 sided protomer, shown in Figure 22c. Recall that in Section 3.2.1 it was mentioned that the trapezoid is topologically equivalent to the *Type-5* protein shape shown in Figure 21, due to the base being split into 2 edges by the capsomer construction. Now notice that while the trapezoid has a reflexive axis through the midpoint of the top

and base, there are no reflexive axes which extend throughout the entire hexamer. Thus, the five sided trapezoid is the protein shape with the smallest number of possible sides. This is an amazing, although expected, result to have been achieved using no more than an understanding of reflexive and rotational symmetry.

4 Conclusion

This paper was not intended to be an in-depth scientific exploration, but simply and introduction to the theory and beauty of icosahedral capsids. While the scientific theory behind icosahedral virus capsids can take a while to wrap your mind around, at least some of the mathematical explorations performed should have been accessible to all readers. Hopefully this paper has demonstrated some of the wonderful mathematical exercises which arise from the physical sciences. It is to be a reminder that math does not need to be dull rote memorization, nor does it need to be the product of decades of work and theorems. Instead, this paper serves as a reminder to let math be what it is, the fun and exciting exploration of problems both inspired by the real world, and created in our minds.

5 References

- [1] N. Campbell and J. B. Reece, *Biology*. San Francisco: Benjamin Cummings, 2002.
- [2] R. V. Mannige, “Principles of virus capsid design,” Dissertation, The Scripps Research Institute, 2010. [Online]. Available: <http://gradworks.umi.com/34/01/3401319.html>
- [3] D. L. Caspar and A. Klug, “Physical principles in the construction of regular viruses.” *Cold Spring Harbor symposia on quantitative biology*, vol. 27, pp. 1–24, Jan. 1962. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/14019094>
- [4] F. H. C. Crick and J. D. Watson, “Structure of Small Viruses,” *Nature*, vol. 177, no. 4506, pp. 473–475, Mar. 1956. [Online]. Available: <http://dx.doi.org/10.1038/177473a0>
- [5] G. S. Pawley, “Plane groups on polyhedra,” *Acta Crystallographica*, vol. 15, no. 1, pp. 49–53, Jan. 1962. [Online]. Available: <http://dx.doi.org/10.1107/S0365110X62000110>
- [6] R. V. Mannige and C. L. Brooks, “Geometric considerations in virus capsid size specificity, auxiliary requirements, and buckling,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 21, pp. 8531–8536, 2009. [Online]. Available: <http://www.pnas.org/content/106/21/8531.abstract>
- [7] A. Zlotnick, “Theoretical aspects of virus capsid assembly.” *Journal of Molecular Biology*, vol. 18, no. 6, pp. 479–490, 2005.
- [8] B. V. V. Prasad and M. F. Schmid, “Principles of virus structural organization.” *Advances in experimental medicine and biology*, vol. 726, pp. 17–47, Jan. 2012. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3767311&tool=pmcentrez&rendertype=pdf>

- [9] R. V. Mannige and C. L. Brooks, “Periodic table of virus capsids: implications for natural selection and design.” *PloS one*, vol. 5, no. 3, p. e9423, Jan. 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2831995/>
- [10] R. Twarock, “A tiling approach to virus capsid assembly explaining a structural puzzle in virology,” *Journal of Theoretical Biology*, vol. 226, no. 4, pp. 477–482, 2004.
- [11] ——, “Mathematical virology: a novel approach to the structure and assembly of viruses.” *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 364, no. 1849, pp. 3357–3373, 2006.
- [12] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, J. Rudnick, and H. Reiss, “Origin of Icosahedral Symmetry in Viruses,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 44, pp. 15 556–15 560, 2004. [Online]. Available: <http://www.jstor.org/stable/3373675>
- [13] R. V. Mannige and C. L. Brooks, “Tilable nature of virus capsids and the role of topological constraints in natural capsid design,” *Phys. Rev. E*, vol. 77, no. 5, p. 51902, May 2008. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevE.77.051902>