Mathematical Musings on the External Anatomy of the Novel Corona Virus*

Part 4: Models of n-Cov

Jyotirmoy Sarkar and Mamunur Rashid

What is the shape of the novel coronavirus (n-CoV) which has turned our world upside down? Even though under a microscope, it looks dull, unattractive, and even disgusting, creative artists have attributed to it bright colors, made it look pretty, and depicted it as a thing of beauty. What can a mathematician contribute to this effort? We take a purist's point of view by imposing on it a quasi-symmetry and then deriving some consequences. In an idealistic world, far removed from reality but still constrained by the rules of mathematics, anyone can enjoy this ethereal beauty of the mind's creation, beckoning others to join in the pleasure.

Our musings end with this Part 4. We fondly hope readers have benefited from our suggestion that they indulge in their own musings, tell others about them, and propagate the good virus of mathematical thinking.

Jyotirmoy Sarkar is a Professor at IUPUI, Indianapolis, Indiana. His research includes mathematics, probability and statistics. He enjoys reading, 'riting, 'rithmetic and R-coding. He devotes his spare times to puzzle making.

Mamunur Rashid is an Associate Professor at DePauw University, Greencastle, Indiana. His research interests include statistical computing, data visualization, data analysis, and applications of statistics to public health/population studies.

Gist of Parts 1, 2 and 3[1]

In Part 1, we described the external shape of the n-Cov as a sphere with three kinds of proteins protruding out of it. In Part 2, we modeled the locations of the S-Proteins as the vertices of icosahedron and dodecahedron inscribed within a sphere, and at the midpoints of edges of the icosahedron (or the dodecahedron). In Part 3, we studied the properties of spherical triangles, which will aid us in locating M- and E-proteins.

Here in Part 4, we answer the challenge posed in Part 1 and offer

*Vol.27, No.10, DOI: https://doi.org/10.1007/s12045-022-1466-3

Keywords
Great circle, Platonic solids, duality, spherical geometry, Girard's theorem, Hamiltonian cycle, intermediate value theorem.

[1] See Vol.27, Nos.4, 5, 9, pp.609–622; pp.855–866; pp.1597–1606, 2022.

additional variations to our model.

7. How Many S-Proteins in a 2D Pictures?

Earlier, we explained that a 2D picture of the n-CoV exhibits the protrusions that will be visible when a slab (with some thickness) is cut out with planar cuts above and below any great circle, and the 3D slab is flattened onto a 2D space, making the protrusions extend out of the 2D circle. According to our model, if the slab is very thin, it is possible to find 10 or 12 S-Proteins in the 2D picture, as shown in *Figure* 1, by the thick, dashed curves—one along the equator and the other along a meridian.

Suppose the slab is made a bit thicker and parallel to one of these great circles. We may assume (1) the S-Proteins on the equator and in the portion of the slab belonging to one hemisphere are visible, but the ones in the other hemisphere are not visible, or (2) all S-Proteins on the slab are visible irrespective of which hemisphere they belong to. Then using our model, the 2D diagrams can show 10, 12, 15, 18, 20, and 24 S-Proteins.

More answers are possible if you change your perspective.

To elaborate, if you focus on the inscribed dodecahedron and imagine rolling it so that it comes to rest with one of its pentagonal faces resting on the tabletop, then the equator passes through 10 (height-wise middlemost) edge-centers. This is shown in *Figure* 1 by the dashed equatorial line. There are five near-neighbors on vertices that belong to the upper half and five more equally far neighbors on vertices that belong to the bottom half. This explains 10, 15, or 20 S-Proteins in a 2D picture. Alternatively, as in *Figure* 3 of Part 2, we may hold the inscribed icosahedron so that its two opposite vertices occupy the north and south poles, respectively. Then also, the equator passes through 10 S-Proteins on edge-centers; there are five near-neighbor face-centers belonging to the upper half and five more near-neighbor face-centers belonging to the bottom half. Regularity in angular deviations (from the spherical center) is clear for 10 or 20 S-Proteins, but not for 15.

Why then nobody depicts 15 S-Proteins in a 2D diagram?

Here is a possible explanation as to why 2D pictures do not depict 15 S-Proteins.

Perhaps angular irregularity, being not aesthetically pleasing to an

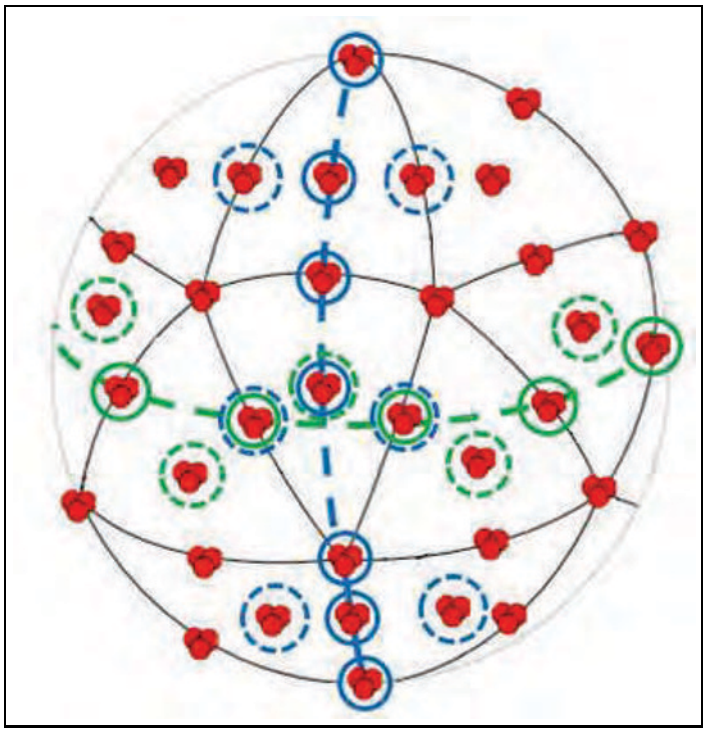

Figure 1. The number of S-Proteins on 2D pictures varies according to which great circular cut is considered, how thin or thick the slab is, and whether S-Proteins on the slab that belong to only one hemisphere or either hemisphere are assumed visible.

artist, is the reason why (and because 15 is an odd number) none of the 2D diagrams show 15 S-Proteins. But if someone did omit every fourth S-Protein after placing 20 at regular angular intervals of 18°, they can justify their omission based on our model.

On the other hand, if we roll the inscribed icosahedron until it comes to rest on the tabletop or hold an inscribed dodecahedron so that its opposite vertices occupy the north and south poles, respectively, then we can count 12 S-Proteins on the equator and 6 near-neighbors on the top half and 6 in the bottom half. This is shown in *Figure* 1 by the dashed line along a meridian. This will explain 12, 18, or 24 S-Proteins in a 2D picture. Although, according to our model, these sets of S-Proteins lack regular angular deviations, who can blame the protagonists for invoking an artistic license to make the angles equal?

The 3D regularity of our model may vanish when projected to 2D.

There is yet a third way to hold the inscribed Platonic solids—with one edge touching the flat surface of a table and the opposite edge at the very top vertically above the first edge. In such a case, for either polyhedron, the middle-most planar (equatorial) cut contains 12 S-Proteins, and they exhibit the same pattern (after a 30° rotation about the axis joining the midpoints of the top and bottom edges) as in the second way mentioned in the previous paragraph, having 6 near-neighbors in the upper half and 6 in the lower half. Therefore, the third way also explains 12, 18, or 24 S-Proteins in a 2D picture.

A second explanation for 12, 18 or 24 S-Proteins

Thus, the 2D count of S-Proteins depends on where exactly the cutting planes are placed and what assumptions are made about visibility. As such, all 2D pictures shown in *Figure* 1 of Part 1, except the one showing 16 S-Proteins, are indeed derivable using our model! To provide a mathematical basis for explaining how 16 may also be a viable answer, one can consider a slab that misses the center. Alternatively, the reader may construct a model starting from a superposition of a hexahedron (cube) and its dual, the octahedron (consisting of 26 vertices and edge-centers), and then adding more points maintaining some form of symmetry (such as by adding quartile points on each edge).

In a 2D depiction, 16 S-Proteins are justifiable.

In conclusion, either our proposed model or one of its variations can vindicate all 2D diagrams in *Figure* 1 of Part 1! Hats off to all creative artists and scientists who produced these diagrams.

8. M-Proteins and E-Proteins: How Many? And Where?

Evidence from the diagrams available on the internet is insufficient to count the number and propose the locations of M-Proteins and E-Proteins. This is where we proceed with our own creative thinking, allowing two possibilities based on the variety of pictures in *Figure* 1 of Part 1: The number of M-Proteins and E-Proteins combined is either (1) about the same as, or (2) about twice as many as the number of S-Proteins. Likewise, guided by the pictures in *Figure* 1 of Part 1, we may allow M-Proteins to be about four to twelve times more frequent than E-Proteins.

We are about to take some creative license.

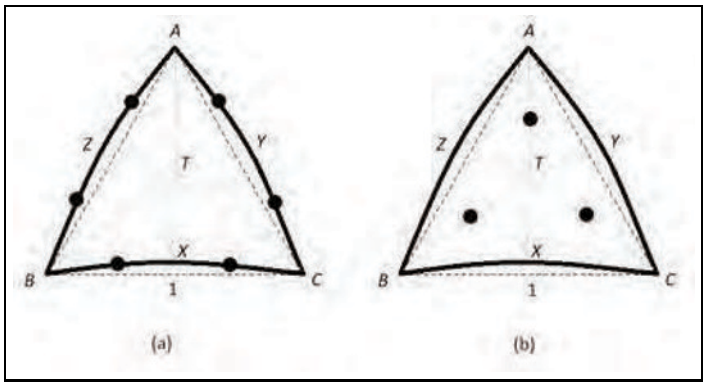

Figure 2. Two ways to add 60 locations for M-Proteins and E-Proteins.

8.1 M+E ≈ S

In the first case, in which the total number of M-Proteins and E-Proteins approximately equal the number of S-Proteins, we propose to identify a total of 60 locations (henceforth called M-E locations) for these two types of proteins, placing them either (a) at the first and the third quartile points of each side of SET/20 (the mid-point is already taken by another S-Protein), or (b) exactly midway between each vertex and the center of each SET/20. See *Figure* 2. In subcase (a), there are $30 \times 2 = 60$ M-E locations; in subcase (b), again $20 \times 3 = 60$ locations.

We offer two sets of 60 points as M-E locations.

8.2 M+E ≈ 2S

In the second case, in which the total number of M-Proteins and E-Proteins approximately equals twice the number of S-Proteins, we propose 120 M-E locations, placing them either (a) both at the quartile points of sides and also midway between vertex and face-center, or (b) at the centroids of the six similar right triangles (of side lengths $g(b) \approx 0.52, g(u) \approx 0.34, g(v) \approx 0.62$ respectively) into which each face of SET/20 is partitioned by the three medians. See *Figure* 3.

Again, two sets of 120 points for M-E locations.

We leave it to the reader to determine how far each centroid is from the nearest vertex, edge-center, and face-center, and also

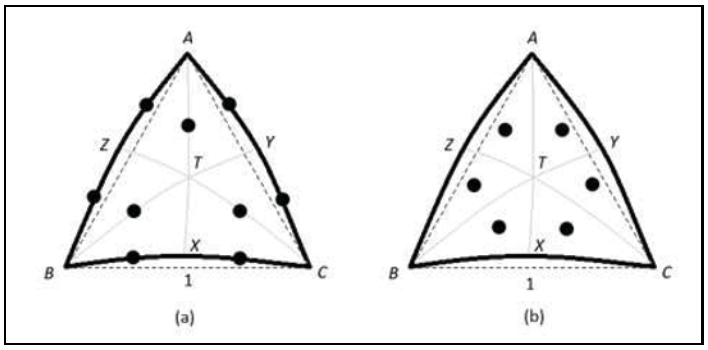

Figure 3. Two ways to add 120 locations for M-Proteins and E-Proteins.

to determine the side length of the regular (which we know by invoking symmetry) hexagon formed by all six centroids.

The last subcase (b) reminds us of another application in computer and information science and statistics. After determining the 120 M-E locations (of the M-Proteins and E-Proteins), if we remove all the S-Proteins and thereafter assign each point on the sphere to the nearest M-E location, we will reconstruct the edges and the medians of all SET/20, partitioning the entire sphere into 120 identical right (spherical) triangles, and thereby recover the S-Protein locations as vertices of these right triangles. This nearest-neighbor partitioning of space is routinely used for the classification and discrimination of massive data into subgroups.

Can you reassemble these 120 "triangles" to make the sphere?

8.3 Separating M-Proteins and E-Proteins

Having chosen the combined locations of the M-Proteins and the E-Proteins, the next task is to determine which protein will occupy which location. Suffices it to locate the E-Proteins, the less frequent of the two types; the remaining M-E locations will be allocated to M-Proteins. For the case $M + E \approx S$, we describe how to locate i E-Proteins, where i divides 24; and, for the case $M + E \approx 2S$, how to locate j E-Proteins, where j divides 40.

Which protein will occupy which location?

For the $M + E \approx S$ case, abiding by our sense of regularity, we determine E-Protein locations by imposing a Hamiltonian cycle

We impose a Hamiltonian cycle.

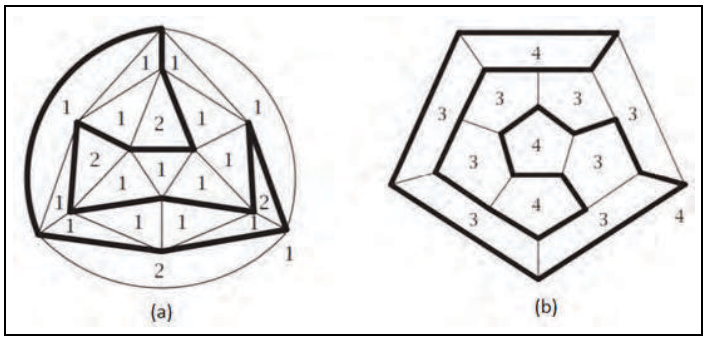

Figure 4. Hamiltonian cycle using at least one edge from each face of **(a)** an octahedron, **(b)** A dodecahedron. Printed on each face is the number of its edges used in the cycle.

connecting the vertices of an icosahedron. Hopkins (2004) [2] proves that there are multiple, topologically distinct Hamiltonian cycles on an icosahedron. For our purpose, we prefer using the one that includes at least one edge of each face. See *Figure* 4(a). As a (smart) bug walks on this Hamiltonian cycle, visiting every vertex exactly once until it returns to the starting vertex, it traverses exactly 12 edges, passing through 24 M-E locations. If we pick all, alternate, every third, fourth, sixth, eighth, or twelfth such M-E locations and assign them to E-Proteins, we can accommodate 1, 2, 3, 4, 6, 8, 12, 24 E-Proteins.

You can make an appropriate choice based on your estimate of the number of E-Proteins.

Likewise, for the $M + E \approx 2S$ case, we determine E-Protein locations by imposing a Hamiltonian cycle connecting the vertices of a dodecahedron (or, using duality, the face-centers of an icosahedron). Again, Hopkins (2004) [2] tells us there is only one topologically distinct Hamiltonian cycle. See *Figure* 4(b). This time, as the bug walks on the Hamiltonian cycle, it traverses exactly 20 edges, passing through 40 M-E locations. We pick all, alternate, every fourth, fifth, eighth, tenth, or twentieth such M-E locations and assign them to E-Proteins. Thus, we can accommodate 1, 2, 4, 5, 8, 10, 20, 40 E-Proteins.

There is only one topologically distinct Hamiltonian cycle.

Of course, we can locate other numbers of E-Proteins obtained by adding the available numbers in the above two sets. For example, we can locate 11 E-Proteins by placing 6 on the Hamiltonian cycle of the icosahedron and 5 on that of the dodecahedron. Also,

in each case, we could inscribe a tetrahedron, octahedron, or hexahedron (cube) inside the sphere and, by rotating it appropriately, match its vertices to the nearest M-E locations to be assigned to 4, 6, or 8 E-Proteins. (We say a match is good if it minimizes the sum of squared (or absolute) distances between the desired and the available points.)

9. Illustrating the Locations of the Three Types of Proteins

Now we reach the climax of our musings. We illustrate the placement, according to our model, of 62 S-Proteins, 6 E-Proteins, and 54 M-Proteins in *Figure* 5 following the strategy shown in *Figures* 2(a) and 4(a). Diametrically opposite (antipodal) points (that are not visible from this point of view) have the exact same type of protein. We only focus on the locations of the visible proteins and restrain from usurping the specialties of creative artists. We leave it to the reader to draw the corresponding diagrams for other choices of numbers of these three types of proteins that our model, or one of its cousins, declares viable.

We only focus on the locations, not the lengths of protrusions.

Here, we have proposed a model for the external anatomy of the n-CoV focusing only on the numbers and locations of three major kinds of protrusions that emerge from its spherical core. We say nothing about the internal components, and absolutely nothing about the functioning of any of the components, leaving those important tasks to experts more qualified than us. Mathematical symmetry principles guide our model's construction to the extent possible. We explained that insistence on perfect symmetry fails to account for enough protrusions. Thereafter, settling for quasi-symmetry, we have proposed 62 S-Proteins and their locations on the sphere. Finally, we gave multiple options to locate 60 (or 120) M- and E-Proteins. We have classified these latter locations into two types depending on various proposals for their number.

Because insistence on perfect symmetry fails, we have settled for quasi-symmetry.

9.1 Alternative Proposals

If someone proposes to locate a total of 90 M- and E-Proteins, we can convince them such a task is impossible to achieve while

To show impossibility, we need a logical proof.

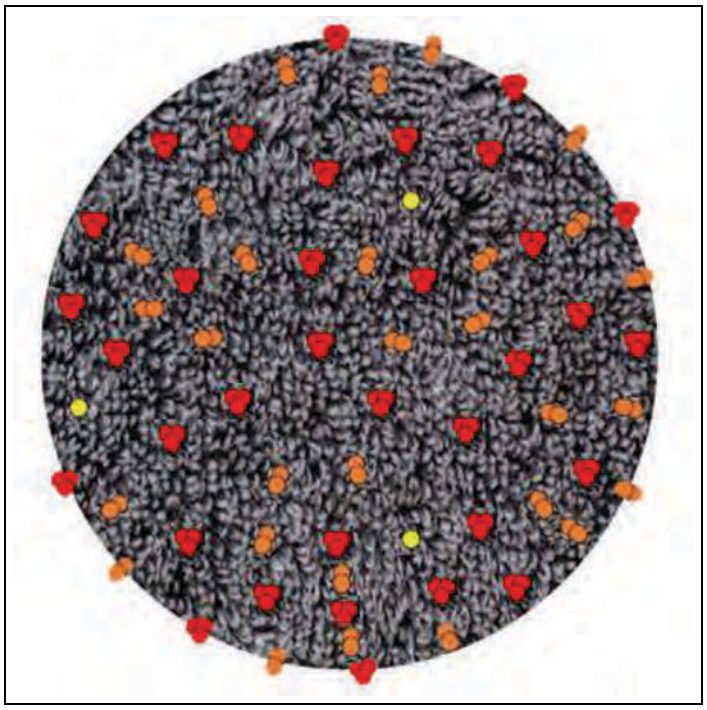

Figure 5. Locating 62 S-Proteins, 6 E-Proteins and 54 M-Proteins according to the recipe given in *Figure* 2(a) and *Figure* 4(a).

maintaining 62 S-Proteins and our adopted form of quasi-symmetry. The reason is simple: The number 90 cannot be written as $m20 + n60$, where m, n are integers, and 20 is the number of faces of an icosahedron (and vertices of a dodecahedron) while 30 is the number of edges of an icosahedron (or a dodecahedron) and edge is already bisected by an S-Protein. Neither can you write 90 as $m12 + n60$.

If another person proposes to locate a total of 100 M- and E-Proteins, we can show the impossibility of that goal as well. Even though $100 = 2(20)+1(60)$, it is not possible to write 100 as $k12+l60$, with integers k, l, for there are 12 vertices of an icosahedron (and faces of a dodecahedron). The astute reader may propose an alternative approach to accommodate 90 or 100 M- and E-Proteins.

What our model cannot do, perhaps another model can and vice versa.

In short, the proposed total number of M- and E-Proteins must be

a multiple of 60. We have already shown how to pick 60 and 120 M-E locations. If someone else insists that there ought to be 180 M-Proteins and 60 E-Proteins, we have a strategy to achieve that target. We borrow from Darwin (1859) [1] an explanation of how bees make honeycombs with hexagonal cross-sections. For the moment, we leave aside the vertices of the icosahedron and focus on its face-centers and edge-centers. Around each of the 20 face-centers, we allow spherical hexagons whose all three opposite pairs of sides are orthogonal to the three medians. Around each of the 30 edge-centers, we allow spherical hexagons whose one pair of opposite sides are orthogonal to the two medians incident at that edge-center (accounting for four vertices), and the remaining pair of opposite vertices are on the edges. These hexagons keep growing bigger and bigger at the same rate until their sides overlap. Then we stop. (The intermediate value theorem guarantees a unique stop.) Clearly, all spherical hexagons are equal in size with a side length of approximately $g(u)/\sqrt{3} = 0.200344$ (if you think the hexagon is planar), but more accurately (because the hexagon is spherical) the sides are of length

How do bees make honeycombs?

The intermediate value theorem guarantees a unique solution.

$$w = g\left(g^{-1}\left(\frac{g(u)}{2}\right)\frac{2}{\sqrt{3}}\right) = 0.2004225.$$

In other words, if we think the hexagon is planar, it is not such a bad approximation.

The vertices of the hexagons are the M-Protein locations. See the larger (orange) bullets in *Figure* 6. On each SET/20, there are six M-Proteins in the interior and six on the boundary shared with another adjacent SET/20. Hence, altogether there are $(6 + 6/2) \times 20 = 180$ M-Protein locations.

The M-Proteins on the boundaries are counted on two faces.

Next, in the interior of each SET/20, we propose to locate three E-Proteins on the medians closer towards the vertices, exactly w units away (along geodesics) from the nearest two interior M-Proteins on that same SET/20. See the smaller (yellow) bullets in *Figure* 6. Each such E-Protein is closest to a vertex of the icosahedron. Thus, looking atop each vertex of the icosahedron, we see at a fixed distance about $g(v) - \sqrt{3}w = 0.2734233$ from the vertex five E-Proteins at a regular angular spread of 72°; then, a

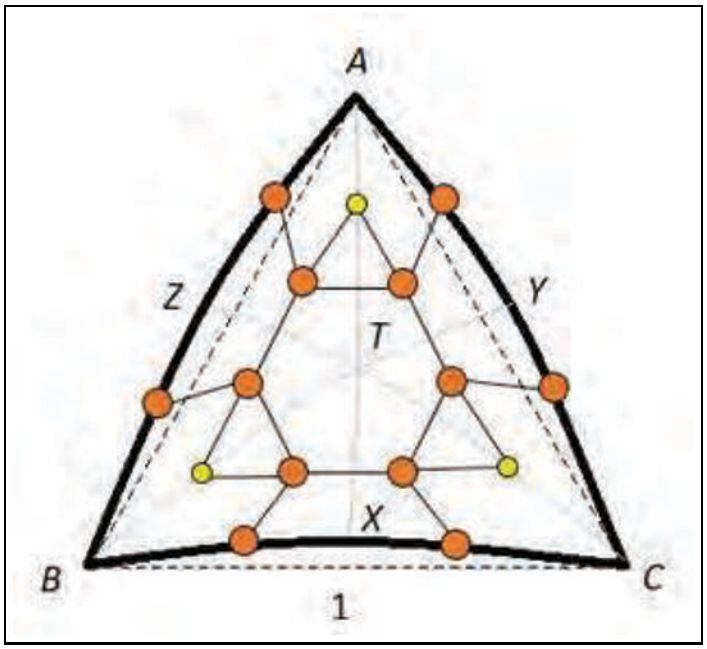

Figure 6. A strategy to locate 180 M-Proteins and 60 E-Proteins.

little further away (at a distance about $g(b) - w = 0.3261364$ from the vertices) we see 5 M-Proteins angularly midway between the E-Proteins; and finally still further away at a distance (the exact value of which we leave to the attentive reader to figure out) slightly more than $g(v) - \sqrt{3}w/2 = 0.4469264$ from the vertices) ten more M-Proteins, not forming a regular decagon, but a cyclic decagon with sides alternating in the ratio 1: $\sqrt{3}$.

Looking atop each vertex of the icosahedron, we see beautiful patterns.

In subsection 6.3 (see Part 3), we justified why the mid-points of any two adjacent edges of an icosahedron subtend an angle of 36° at the center by invoking the golden ratio between the radius of the circumscribing sphere and the length of the geodesic joining these two points, Here is a simpler reason: In *Figure* 1 of Part 3, where two vertices occupy the north and south poles, the equator passes through exactly 10 midpoints of edges at regular angular distances spanning 360° together. Hence, each must subtend 36° at the center!

A global view offers a simpler explanation.

Epilogue

Dear readers, begin/continue your own musings.

Having presented some beautiful gems gleaned from the ocean of mathematical ideas as a bouquet of offering to our dear readers during this stay-at-home pandemic period, we must now say farewell. We sincerely hope that attentive readers are sufficiently equipped to begin/continue their own musings on topics of relevance. May we suggest that they try to construct other models, perhaps starting with a radial projection of a hexahedron and its dual (an octahedron), and then embellish it with additional points chosen quasi-symmetrically to ensure the sufficient number of points on the sphere to model the novel coronavirus or any other object of this nature. Godspeed.

Address for Correspondence
Jyotirmoy Sarkar
Mathematical Sciences
Indiana University-Purdue
University Indianapolis
402 North Blackford Street
Indianapolis, IN 46202, USA.
Email: jsarkar@iupui.edu

Mamunur Rashid
Mathematical Sciences
DePauw University
2 East Hanna Street
Greencastle, IN 46135, USA.
Email: mrashid@depauw.edu

Acknowledgement

The authors thank Professor Roland Roeder for some discussions, and for drawing their attention to Smale (1998).

Suggested Reading

[1] C Darwin, *On the Origin of Species by Means of Natural Selection*, or *The Preservation of Favoured Races in the Struggle for Life*, London: John Murray, Albemerle Street, 1859.

[2] B Hopkins, Hamiltonian paths on platonic graphs, *International Journal of Mathematics and Mathematical Sciences*, 30, pp.1613–1616, 2004.