Morphological Group Theory of Soft Matter Structure - A New Mathematics of Natural Image Numbers

Ziqiang Zhou*
Institute of Materials, Shanghai University, Shanghai, China

*Corresponding author email: zqzhou408@21cn.com

Abstract. The accurate understanding and interpretation of the SARS-CoV-2 electron microscope photos (EMP) is of great significance for the scientific treatment of coronary pneumonia and precise strategy. Here I study the image mathematics and physics description of micro-images of soft matter system by the morphology group theory of material structure, and show that there is a natural image numbers equation, which can be used to describe not only the common feature of soft matter structure, but also to correct classify the states into different classes, so that to study the morphology physics for soft matter science and biology, some important basic concepts of biophysics, such as the central dogma of molecular biology, Gene sequence analysis, heuristic recognition of SARS-CoV-2 can also be newly explained and understood by the equation.

Keywords: Image-Mathematics models (IMM); Deduction Theorem for Image–Math; COVID-19; the Central Dogma of Molecular Biology; Heuristic Search and Recognition.

1. Introduction
The COVID-19 disease is threatening the lives of people all over the world. The accurate understanding and interpretation of the SARS-CoV-2 electron microscope photos (EMP) is of great significance for the scientific treatment of coronary pneumonia and precise strategy. Morgan confined genes to the chromosomes of cells and studied them on the chromosomes.[1] and the application of chemical analysis led to the development of genetics from chromosomes to molecular biology, emphasizing that chromosomes are composed of proteins and DNA or RNA.[2]. The introduction of the electron microscope (EM) expanded the scope of research from micron to nanometer scale, this kind of method provides a huge amount of valuable images for us to understand the internal structure of the soft matter system. But there exist still some difficult to be solved about how to classify these micro-images, how to study the relation among the different shape kinds of complex micro-images, and how to explain and understand these images, which is need to introduce soft matter physics[3,4] and new thinking science[5][6]. Here I want to discuss the modeling about the image mathematics model (IMM) for descriptions of the complex structures for soft matter system and biology. I begin with a morphological equation with one particle, then through introducing a method of set theory and recursion theory to construct the morphology equation of multi-particle systems and the final to interpret these complex and changing micro-images of morphology physics.

2. The Images–Math Symbol Systems for Morphological Physics
Up to now mathematics has been based on logical thinking. But the EMP is image. In order to model these complicated images one must at first have to thinking in image. It is of the key question that how to transform the complicated image question into mathematical one or how to discover the fundamentally new concepts and way of thinking in images, produce a qualitative structure for the
relations and functions about micro-images, i.e., a qualitative structure's law on micro-images for soft matter system (including biology), which are the basic questions that are important throughout IMM.

2.1. Modeling—The Images-Mathematics Symbol Systems for 2D and 3D Structure

Jordan's closed curve theorem states that a simple closed curve $C$ divides the points on the plane not on $C$ into two distinct domains, the interior and the exterior of the curve(with no points in common) of which $C$ is the common boundary. The theorem deals with geometric(topology) propositions, but does not relate to the structure and components of matter. In order to apply this theorem to describe the structure of materials, one must study the relation between the morphology of micro-structure and the meaning of topological propositions. I put forward the morphology group theory of material structure\textsuperscript{[6]} which include the morphology equation (consisting of four images-mathematics symbols $|z|$, $[J]$, $[H], M$.) For a given one observation region, $x \in M$, if there is a simple closed curve $C$ (Fig.1), then there is a \textbf{morphological deduction theorem}\textsuperscript{[6]} which consisting of four statements: (a) morphological transformation $z$ (or function) defined by a simple closed curve $C$ must divide the plane region,$x \in M$ into two kind classes, the interior and the exterior image; (b) The interior image Int M: $z|\alpha| = [J(\alpha)|z|$ called as the Jordan morphological element (ME); (c) The exterior image Ext M: $z|\alpha| = [H(\alpha)|z|$ called as Hidden ME,(d) of which the simple closed curve $C$ is the common boundary.

$$M = [H(x)|z|]$$ \hspace{1cm} (1)

In the morphological equation (1),$[\ ]$ denote a symbol of image operator, and $|z|$ stands for a morphological transformation defined by Jordan simple closed curve.\textsuperscript{[7]}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{morphological_deduction_theorem.png}
\caption{(a) The morphological deduction theorem \hspace{1cm} (b) The difference between mathematics and image mathematics}
\end{figure}

In topology the Jordan curve theorem can be extended the 3D separation theorems\textsuperscript{[7]}, so (1) can also be used to describe the 3D micro-images. Such two kinds of ME and morphological transformation may also be used to describe micro-structure of other materials, which can be used as the image-mathematical description language\textsubscript{2}, improve the description of the micro-images (Fig.1(b)).

2.2. IMM and Physics of Molecular Biology

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{imm_and_physics_of_molecular_biology.png}
\caption{(a) Non-folded state \hspace{1cm} (b) Folding state}
\end{figure}

\textbf{Figure 2.} Application of IMM to express central dogma of molecular biology.
The third step of the modeling: To apply theoretically derived morphological equations to explain practical problems. Here we need to use the equations to interpret biological problems such as cells and SARS-CoV-2. We need to translate the mathematical model language of morphological equations into the biological morphological language of cells or viruses.

2.2.1. IMM and the central dogma of molecular biology

The central dogma of molecular biology has deciphered the genetic code, that is, the way DNA carries genetic information is from DNA to RNA to protein, and can change the genetics of organisms by changing their genetic material so as to change the orientation, which is genetic engineering. The deficiency is that there is no connection between molecular composition and morphological structure and molecular biophysics, so that many problems in molecular biophysics have not been solved. In order to answer these question, we need to translate the mathematical model language of morphological equations into the biological morphological language of cells or viruses. Specifically, we translate and convert the two types of morphological elements of matter structure into two types of macromolecule (protein and RNA or DNA) of cell or virus, and then use morphological equations to describe the morphological structure of cells and viruses, identify and interpret the qualitative structure of image morphology of virus, and study the scientific issues of virus morphology.

\[ M = [H \mid z \mid J] \Rightarrow M = [H(Protein)\mid z(q)\mid J(DNA)] \] (2)

In order to solve these problems, the concept of ordered pairs must be introduced. the ordered pairs \( \sigma: [j, h] \) is defined to be \([j], [j, h]\) , this definition succeed in capturing the desired unique decomposition property; the order pair \([j, h]\) uniquely determine both what j and h are, and the order upon them. In other words, if an ordered pair can be represented in two ways

\[ \sigma \in M: [j, h] = [\alpha, \gamma] \text{or } [j, h] = [j(DNA), h(Protein)] \] (3)

Then the representation are identical in the sense that

\[ j=\alpha \text{ and } h=\gamma \text{ or } j=j(DNA) \text{ and } h=h(Protein) \] (4)

The relationship R above is a set of ordered pairs ~\([j(DNA), h(Protein)]\), using these relationships, a new model graph of the central dogma of molecular biology is constructed. Fig.3.a and b.

This is a new illustration of the Central Dogma of Molecular Biology by the language of IMM, which not only expresses the basic connotation of the way DNA carries genetic information from DNA to RNA to protein, but also expresses the following new content: (1) In living cells, DNA, RNA, and Protein can cooperate closely to participate in the life process, and even form macromolecules essential for life activities, such as nucleoprotein. (2)The structure of RNA is complex, and it can be divided into three situations according to its role. (a) As a template for forming protein, the shape of RNA produces a protein of the same shape, which is mRNA (messenger RNA); (b)In the process of protein synthesis, which is responsible for transferring the appropriate amino acid molecule to the appropriate position in the graph, namely RNA (transfer RNA); (c). Combine with protein synthesis to form a ribosome, namely rRNA (ribosome RNA), which is protein synthesis "Factory". Therefore, the new model is an important improvement and development of the central law of molecular biology.

3. Gene Sequence Combination and Sequence Analysis

According to the idea of model theory, the problem of biological sequence combination and sequence analysis should be at first abstracted as an image mathematical problem, and then a multi-particle image mathematical model should be constructed through the deduction of mathematical theory such as the induction set of image mathematics, and finally it can be used to explain the problem of gene sequence combination and sequence analysis. The derivation of natural morphological equation of matter systems with multi-particles is of a complicated question. [5]
3.1. Modeling: Mathematics of Natural Image-number (IMM)—Mathematical Induction of Natural Image-number

Natural numbers define several functions $\emptyset$, $\{\emptyset\}$ etc., in a direct way, as primitive functions for constructing a set of natural numbers, and then apply mathematical induction to deduce the set $\omega$ of natural numbers. But the symbols $\emptyset$, $\{\emptyset\}$ etc. are no clear definition of $\emptyset$, $\{\emptyset\}$, the universality and conditionality of successor operations $n^+=n+1$ are not explained, the complexity of the inductive set is not described and the completeness of number theory is not discussed. [8-10]

3.2. Mathematical Induction of Natural Image-number

To study gene sequence combination and sequence analysis, it is to need to develop a new morphological mathematics by formal languages $L_z$. The four image-math symbols $\mathcal{J}$, $\mathcal{H}$, $|z|$, $M$ are directly defined by the Jordan closed curve separation theorem, and the concept of morphological relationship $R_z$ and morphological function, and derives the morphological equation and morphological deduction theorem [6], which provides a basis for us to construct the mathematical model of the image. The morphological transformation $|z|$ is defined by a simple closed curve, which has the duality of image and number, which is the basics how we construct an image and number set.

3.2.1. Definition. For the natural image sequence, one can directly define the following function and define its range:[8]

(a) Successor function S: The value of $Sx$ is the direct successor number of $x$, $Sx = x+1$.

(b) Precursor function D: The value of $Dx$ is the direct predecessor number of $x$, $Dx = x-1$.

Thus the inductive set of natural image numbers can be obtained according to set theory. See Fig.3(a).

3.2.2. The set algebra of natural image numbers. According to the set algebra rules, the first few natural image number (in Fig.3(a)) can be characterized as follows,

For $z=1$, $M = [H(1)||z(1)||J(1)]$; it is given by Jordan curve theorem

For $z=2$, $M = [H(1,2)||z(2)||J(1)||J(2)]$; it is “closed under successor operation

$z=3$, $M = [H(1,2,3)||z(3)||J(1)||J(2)||J(3)]$.

$z=n$, $M = [H(1,2,\ldots,n)||z(n)||J(1)||J(2)\ldots||J(n)]$

Formally, let $\{I_i : i \in I\}$ be a family of sets indexed by $I$. The disjoint union of this family is the set

$$\bigcup_{i \in I} I_i = \bigcup_{i \in I}(x, i) : x \in I_i$$

So, $M = [H(n)||z(n)||\bigcup_{i \in I} I_i]$ and $z = 0 \Rightarrow M = [H(0)]$. (6)

The analysis above illustrates the relationship between natural numbers and natural image numbers. It shows that (1) under $|z(n)|$ conditions defined above, there is a morphological transformation group,
whose \( I(1) \) is the identity element, \( H(1) \) is the inverse element, \( M=H(0) \) is a zero element. 

(2) The two kinds of numbers show that \( Z = 1 \) corresponds to \( I(1) \), \( z=0 \) corresponds to \( M=H(0) \), That is, the natural image number is the extension of natural numbers, in turn, natural image number provides us with the morphological mathematical definition of natural numbers 1 and 0 (\( \emptyset, \{\emptyset\} \)), natural numbers can be regarded as special cases of natural image numbers. Thus the inductive set of natural image numbers can be obtained. See Fig.3(a)

3.3. The Formal Deduction System of IMM and Gene Sequence Combination and Sequence Analysis

Combining Equation (2) with Equation (6), we can obtain the equation

\[
M(n) = \left[ H(\text{Protein}) (n) | z(n) | \bigcup_{i \in J} (\text{RNA}) (i) \right] \tag{7}
\]

The equation is also called as the natural morphology equation, Such images mathematics equations can be used to help analyze Gene Sequence Analysis. Given a micro-image we can determine (without calculation) what objects are named by letter \( |z| I(\alpha) H(\alpha) \) and can judge the truth or falsity of such “atomic” formulas. Given any formulas \( I(\alpha) \) and \( H(\alpha) \) we can combine them in various ways to obtain new ones, the intended meaning of the new formulas can be captured by giving a truth table. One might judge if these formulas is true or false.

4. Heuristic Search and Recognition of COVID-19 Morphology (SARS-CoV-2)

Pattern recognition is one of the important parts in intelligent simulation. Pattern recognition is also required for virus morphological type recognition. Due to the complex form of the virus, it is necessary to apply the mathematical model language \( L_Z \), and use the subset (four sentences) contained in the sentence set like the mathematical deduction theorem to perform syntactic analysis. Form the framework of the morphological recognition system; then study the pattern grammar and study the semantic issues of the pattern.

Take the EMP of SARS-CoV-2 as an example, apply the morphological description language of the image mathematics model, and conduct morphological syntactic analysis through the four sentences of the morphological deduction theorem of image mathematics to determine the biology of the specific research object and the meaning of the virus, The specific steps are as follows:

1. The first step of \( S \in V_N \) starts with the downloaded electron microscope picture, and specifically includes the following sub-steps
   a. Download the coronavirus EMP from the literature, that is, our observation area \( M = \) the height and width of the picture = \( a \times b \).
   b. Analyze the observed picture \( M \), search for the discrete countable large-size red particle structure \( J \)---SARS-CoV-2 particles in the picture; look for the same type of virus particles, a total of six, the area occupied by these virus particles is equal to their disjoint union, \( J=\bigcup_{i \in J} [i] \).
   c. Find the complement between viruses, that is, the medium \( [H]=\bigcup_{i \in J} [i]=a \times b - \bigcup_{i \in J} [i] \) = Black matrix + green particles.
   d. Verification: The superposition state of virus and medium is: \( M = [H(\pi_0)] Q | \bigcup_{i \in J} [i] | = a \times b \).

2. \( V^* \) Enter the next sub-model to search, and further heuristic search for the virus medium \( [H] \), the content steps are the same as above. It is found that the medium in picture \( M \) is also complex and
uneven. It contains a black and green area, which contains much smaller particles than the coronavirus. This two-color area is analyzed using the sentence set of the morphological deduction theorem. The steps are the same as above. The result of the sub-pattern search is a morphological region different from the SARS-CoV-2, which may belong to a protein virus (prion).

3. \(V^++\) Further heuristic search for the sub-patterns of the media in the protein virus, the more refined level, enters the level of amino acids, which belongs to the level of small biological molecules and atoms. It shows that there may be changes in the composition and structure of 20 different amino acids within the protein. This poses a challenge for us to understand the internal nanophase transitions of proteins. Due to the limitation of the resolution of the electron microscope, we can only find such a level of analysis, that is \(V_T\), which is the set of terminator.

This is a special heuristic search program, Its characteristic is to use the mathematical deduction theorem to search cells or viruses layer by layer, from SARS-CoV-2 to macromolecules (such as proteins) to small molecules (such as amino acids) to atoms, each level has a corresponding particle and hidden structure and the morphological equation.

5. Conclusion
1. The principle of image mathematics based on morphological equations is a new type of natural IMM, which can be used to describe the morphological structure of cells and viruses, explain the microscopic image problem of virus structure and the morphological mathematics of biophysics.
2. Using the morphological deduction theorem to analyze the physics of soft matter, it can provide new explanations for the central dogma of molecular biology, protein folding, gene sequence combination, gene sequence analysis, and complex coronavirus structure.
3. The heuristic search and patterns recognition of COVID-19 is worthy of further study

Acknowledgements
During the course of writing the paper, I had the benefit of many conversations with Academician ZHOU Bangxin, Prof. QI Feipeng, Prof. YIN Jingbo, and Prof. YAO Meiyi, which often led to sudden clarification of an obscure point. To all of them, I wish to extend my sincere thanks.

References
[1] T.H. Morgan 1917The Theory of the Gene,(the University of Chicago press)
[2] Zhang Weijie 2011 Introduction to Life Sciences (2nd Edition) (Higher Education Press) (in Chinese)
[3] Mohamed Daoud, Claudine E. Williams ( Editors ) 1999 Soft Matter Physics (Springer)
[4] Lu Kunquan & Liu Jixing 2007 Introduction to Soft Matter Physics (Peking University Press) p598 (in Chinese)
[5] QIAN Xuesen 1989 System Science, Thinking Science and Human Body Science(Chengdu Sichuan Education Press)p 15-33(in Chinese)
[6] ZHOU Ziqiang 2018 Morphological Group Theory of Material Structure (Journal of Applied Mathematics and Physics) January 5, p.69
[7] Munkres J.R.2000 Topology. (Pearson Education) p.392
[8] Mo Shaokui 1999 Recursion Theory (Science Press) p178 (in Chinese)
[9] QIAN Weichang. 1980 Modern Mechanics and the Four Modernizations, p.86; in 2012 (Selected Works of Qian Weichang Vol.2)(Shanghai University Press) (in Chinese),
[10] Klaus Mainzer 2005 Symmetry and complexity The spirit and beauty of nonlinear science p199-237
[11] Herbert A .Simon. 1976 Computer Science as Empirical Inquiry: Symbols and Search, (Communications of the ACM), Vol.19, No.3
[12] Herbert B. Enderton. 2006 Elements of Set Theory,( Elsevier Science)
[13] XIE Mingpei 1981 Intelligent Simulation, (Defense Industry Press) (in Chinese)
[14] Gene-Wei Li & X.Sunnet Xie 2011 Central dogma at the single-molecule level in living cells,(Nature),Vol.475,21 July,p308