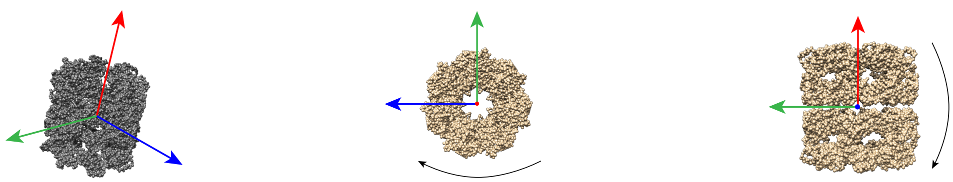
Goal: In this homework you will implement the symmetry detection algorithm and complete some exercises that provide the mathematical justification for the algorithm.

Algorithm

We begin by implementing the algorithm, without yet justifying why it works. Please hand in a Python program, e.g., a `.py` file, that completes the following tasks. Use Python comments to answer questions that aren't explicit computations. Make sure your program runs!

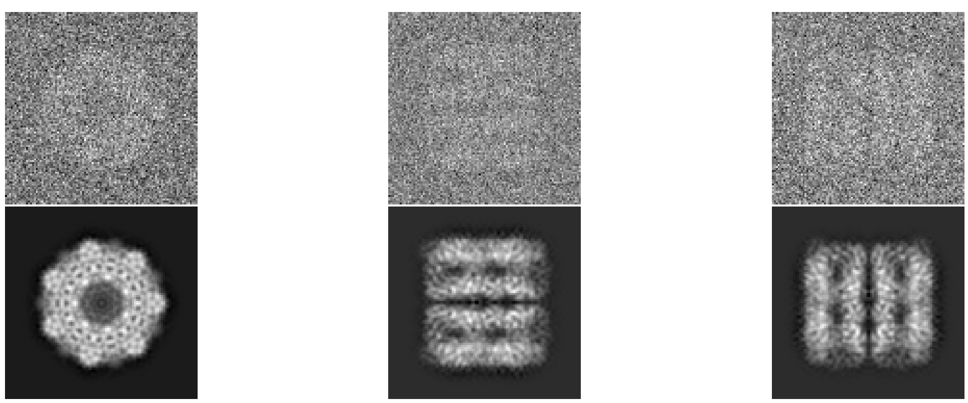
In this assignment we will implement a program that detects the symmetry group of the molecule GroEL from simulated cryo-EM images. The following is an outline of the symmetry detection algorithm. The input to the algorithm is a dataset of N cryo-EM images $I_1, \ldots, I_N$ of a molecule, see [Figure 2]. The output should be one of the possible 3D symmetry groups listed in [Figure 4].

The input images for your algorithm are available at `http://www.math.wisc.edu/~dynerman/groel.mat`.



(a) Molecule at origin of $\mathbb{R}^3$ (b) 7-fold symmetry (c) 3-fold symmetry

Figure 1: GroEL



(a) 7-fold rotational symmetry (b) 2-fold rotational symmetry (c) 2-fold rotational symmetry

Figure 2: Noisy and clean cryo-EM Images of GroEL

A cryo-EM image projected along a symmetry axis [Figure 3] will exhibit rotational symmetry [Figure 2a]. We say the <u>rotational symmetry order</u> of an image is the order of this symmetry, i.e., the number of times the image matches up with itself as we perform one complete rotation about its center. For example, the rotational symmetry order of [Figure 2a] is 7, because the image will line up with itself 7 times as we rotate it about its center.

Our algorithm will distinguish between the different possible 3D symmetry groups by finding which rotational symmetry orders appear in the set of images. We first discuss how to compute the rotational symmetry order of an image.

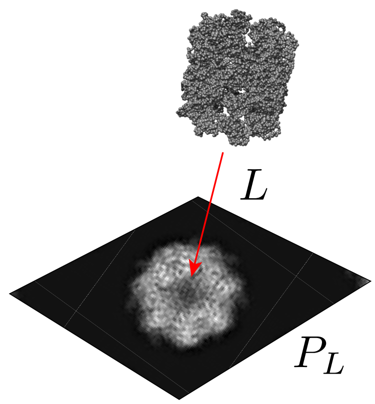


Figure 3: A cryo-EM image projected along L onto P_L.

Step 1: Compute rotational symmetry orders

In this section you will implement code to detect the order of rotational symmetry in images by computing the image's rotational auto-correlation. The rotational auto-correlation of an image I is the correlation (i.e., dot product) between I and the rotation of I by some angle θ, i.e.,

$$\mathrm{RAC}_I(\theta) = \langle I, \mathrm{rot}_\theta(I) \rangle,$$

where $\mathrm{rot}_\theta(I)$ is image I rotated about its center by θ degrees.

1. Write a function to compute the rotational auto-correlation of an image I. To rotate an image, see `rotate` in the package `scipy.ndimage.interpolation`. How do different interpolation methods effect your algorithm?

2. For a sample image I of your choosing, plot RAC_I as a function of θ.

Remark: You might find the image `http://www.math.wisc.edu/~dynerman/gauss_3_top.mat` useful to test your code.

Observe that the rotational symmetry order of I is the same as the number of peaks of RAC_I. Thus we can compute the period of this function to obtain the rotational symmetry order of I.

Claim: If a signal f of length N has period k, where k divides N, then non-zero Fourier coefficients may only appear at positions $0, 1, N/k, 2N/k, \ldots, (k-1)N/k$.

Based on the claim above, we can compute the period of RAC_I by examining its Fourier coefficients. Due to numerical imprecision, we may see "disallowed" Fourier coefficients even in a perfectly periodic signal, so rather than looking at which Fourier coefficients are non-zero, you may want to look at which coefficients are the largest.

For example, if the 3ℓ, $\ell \in \mathbb{Z}$, Fourier coefficients of RAC_I are the largest, then we can reasonably conclude that RAC_I has period $360/3$, i.e., I has 3-fold symmetry.

3. Write a function that determines the period of RAC_I. You can compute the Fourier coefficients of RAC_I by using the discrete Fourier transform (DFT). See the `numpy.fft` package.

   Remark: Your algorithm must also detect the absence of symmetry in an image. How will you detect no symmetry, e.g., the period of the rotational autocorrelation is 1? Here is one possibility: if you detect periodicity by finding large Fourier coefficients, then perhaps you can detect an aperiodic signal when no set of Fourier coefficients stands out from the rest.

Group		Rotational symmetry orders of images
Cyclic	C_n	Order n
Dihedral	D_n	Order n and order 2.
Tetrahedral	T	Order 3 and order 2.
Octahedral	O	Order 4, order 3 and order 2.
Icosahedral	I	Order 5, order 3 and order 2.

Table 1: Rotational symmetry orders appearing in images of a molecule with given 3D symmetry group.

4. Combine your code from this section into a single function `find_rotational_symmetry_order` that takes an image as input and returns the order of its rotational symmetry.

Step 2: Lookup group in flowchart

We are now ready to glue the pieces and finish the flowchart algorithm for symmetry detection. In Step 1 we computed the rotational symmetry orders of all the images in our dataset. The rotational symmetry orders we observe are enough to distinguish the group, see [Table 1].

To determine the group, we compute the rotational symmetry orders in our images, and then lookup which group they correspond to in [Table 1]. For example, if the only rotational symmetry orders we see are 7 [Figure 2a] and 2 [Figure 2b]

5. Using your `find_rotational_symmetry`, and the table on page 1, implement the `flowchart` function that takes N cryo-EM images and returns the symmetry group of the molecule. Test your algorithm on the GroEL images and verify that it outputs the correct symmetry group.

Remark: Due to noise and numerical imprecision, your algorithm may compute rotational autocorrelation orders that do not appear in the table. For instance, when processing the noisy images [Figure 2], we may see the following rotational symmetry orders in our images:

$$\{2, 2, 3, 3, 2, 2, 2, 2, 7, 7, 3, 5, 7, 7, 7, 7, 7, 2, 2, 2, 2, 7, 7, 7\}.$$

One way to deal with this is by <u>voting</u>. In the example above, we see 2 and 7 as rotational symmetry orders of many images, while 3 and 5 only appear a few times. Thus, we might conclude that the correct rotational symmetry orders are 2 and 7, corresponding to the group D_7, and decide the orders 3 and 5 were misdetected.

Theoretical justification of algorithm

In this section you will solve some exercises that provide a mathematical justification for why the symmetry detection algorithm you developed above works. Our algorithm processed cryo-EM images, so we begin by presenting a mathematical model for these images.

Cryo-EM images

We will model our molecule as a function $\phi : \mathbb{R}^3 \to \mathbb{R}$, see [Figure 1a]. This function represents the electronic density of the molecule at various spatial locations.

Recall that we can think of a cryo-EM image as a projection of ϕ along some axis $L \subset \mathbb{R}^3$, which we think of as a microscope viewing direction, see [Figure 3]. The image obtained, denoted $I_L(\phi)$, is a function on $P_L = L^\perp$, the plane in $\mathbb{R}^3$ perpendicular to L. The value of the image at a point $p \in P_L$, is modeled by

$$I_L(\phi)(p) = \int_L \phi(p + l)\, dl.$$

In other words, the intensity of the image $I_L(\phi)$ at a point p is the integral of ϕ along the line through p in the direction of the projection axis L.

Rotations of images

In our algorithm we computed the rotational symmetry order of an image. In order to explain the meaning of this number, we first discuss rotations of cryo-EM images.

Fix a microscope viewing direction L and let $\mathrm{SO}(P_L)$ be the collection of all 3D rotations that fix the plane P_L.

6. Show that $\mathrm{SO}(P_L)$ is a subgroup of $\mathrm{SO}(3)$.

If $R \in \mathrm{SO}(3)$ is a rotation, then the rotation of our molecule ϕ by R is the function given by

$$(R \cdot \phi)(v) = \phi(R^{-1}v).$$

Similarly, if $I_L(\phi)$ is an image, and $S \in \mathrm{SO}(P_L)$ is a 3D rotation fixing P_L, then the rotation of I_L by S is given by

$$(S \cdot I_L(\phi))(p) = I_L(\phi)(S^{-1}p).$$

7. Fix a projection axis L and a rotation $S \in \mathrm{SO}(P_L)$. Show that $I_L(S \cdot \phi) = S \cdot I_L(\phi)$.

This last exercise shows that a 2D rotation of an image gives the same result as first performing the corresponding 3D rotation, and then projecting.

Maximal Cyclic Subgroups

We will show that rotational symmetry order is in fact the size of a maximal cyclic subgroup of the molecule's symmetry group. We first describe the notion of a maximal cyclic subgroup.

A <u>cyclic subgroup</u> H of a group G is a subgroup that is generated by a single element $g \in G$, i.e.,

$$H = \{1_G,\ g,\ g *_G g,\ g *_G g *_G g,\ \ldots\}.$$

In this case we say H is generated by g and write $H = \langle g \rangle$. For example, if $G = C_6$, presented as the integers modulo 6, then

$$H = \{0, 2, 4\} \subset \{0, 1, 2, 3, 4, 5\} = C_6$$

is a cyclic subgroup, since H is generated by 2, which we write $H = \langle 2 \rangle$. Note that the entire group C_6 is also cyclic (it is generated by 1).

We say that $H \subset G$ is <u>maximal cyclic</u> if H is not contained inside a larger cyclic subgroup of G. For example, inside the symmetric group S_3 of permutations on $\{1, 2, 3\}$, the subgroup

$$H = \{1_{S_4},\ (1, 2)\}$$

generated by the transposition $(1, 2)$ is maximal cyclic.

8. For each element g of S_3, write down the cyclic subgroup generated by g. Indicate which of these subgroups are maximal.

A non-example is the subgroup $H = \langle 2 \rangle$ of C_6. This subgroup is <u>not</u> maximal cyclic, because it is contained in the larger cyclic group $C_6 = \langle 1 \rangle$.

Rotational symmetry order

We now describe the meaning of the rotational symmetry orders we computed in our algorithm. Suppose that the molecule ϕ has symmetry group Γ. Recall that this means that $R \cdot \phi = \phi$ for all $R \in \Gamma$.

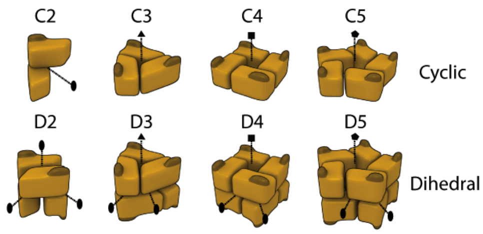
9. Let $R \in \Gamma$ be a symmetry of ϕ that rotates vectors about an axis L. Note that this implies $R \in \mathrm{SO}(P_L)$. Show that

   $$R \cdot I_L(\phi) = I_L(\phi),$$

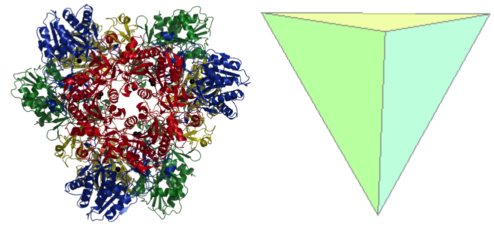
   i.e., the cryo-EM image $I_L(\phi)$ is unchanged when we rotate it by R.

10. Conclude that if Γ has a maximal cyclic subgroup C_n, generated by rotation about an axis L, then the image $I_L(\phi)$ has rotational symmetry order n.

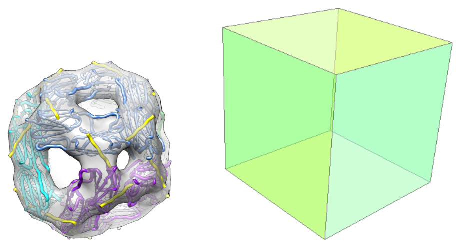
11. Now, suppose that we have an image $I_S(\phi)$ with rotational symmetry order m. Does it follow that C_m is a maximal cyclic subgroup of Γ? If yes, give an argument; If no, give a counter-example.



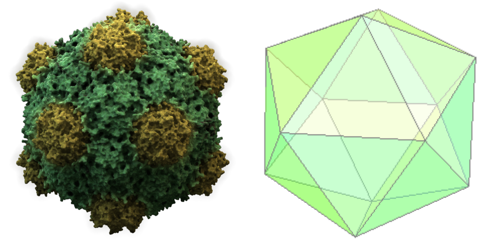
(a) C_n, D_n: Cyclic & Dihedral symmetries.



(b) T: Symmetries of a tetrahedron.



(c) O: Symmetries of a cube.



(d) I: Symmetries of an icosahedron.

Figure 4: Symmetry groups of 3D objects

Symmetry groups and their maximal cyclic subgroups

We now see that rotational symmetry orders of images tell us about maximal cyclic subgroups of a molecule's symmetry group. In this section we conclude the mathematical justification of our algorithm by seeing how finite symmetry groups of 3D objects can be distinguished by their maximal cyclic subgroups. Recall that any symmetry group is the symmetry group of one of the objects in [Figure 4].

Let us examine the maximal cyclic subgroups appearing in these groups above. First, if $G = C_n$ [Figure 4a] then the entire group G is the only maximal cyclic subgroup. If $G = D_n$, then G contains two maximal cyclic subgroups: one C_n generated by a rotation of $360/n$ degrees about the vertical axis in [Figure 4a], and one C_2 generated by a rotation of 180 degrees about the second axis coming out of the page in [Figure 4a].

If $G = T$, then G contains 7 maximal cyclic subgroups. They are (i) four C_3's generated by rotations of 360/3 degrees about an axis through the center of each triangular face, and (ii) three C_2's generated by rotations of 180 degrees about an axis through the midpoints of opposite edges, see [Figure 4b].

For $G = I$, G contains (i) six C_5's generated by rotations about axes joining opposite vertices, (ii) ten C_3's generated by rotations about axes through opposite triangular faces, and (iii) fifteen C_2's generated by rotations about axes through midpoints of opposing edges.

12. List all the maximal cyclic subgroups of $\Gamma = O$, see [Figure 4c]. For each maximal cyclic subgroup H, indicate the axis of rotation of the elements in H.

Thus, the "rotational symmetry orders" in [Table 1] are actually the sizes of the maximal cyclic subgroups appearing in these groups. By our analysis in this section, we see that the sizes of maximal cyclic subgroups are enough to distinguish between the groups in [Figure 4].

Remark: Klein's theorem tells us that the groups above, see [Figure 4], are the only possible 3D symmetry groups, so our flowchart will detect any 3D symmetry group.