SIAM J. Appl. Algebra Geometry
Vol. 1, pp. 507–535

Cohomology of Cryo-Electron Microscopy*

Ke Ye† and Lek-Heng Lim‡

Abstract. The goal of cryo-electron microscopy (EM) is to reconstruct the 3-dimensional structure of a molecule from a collection of its 2-dimensional projected images. In this paper, we show that the basic premise of cryo-EM—patching together 2-dimensional projections to reconstruct a 3-dimensional object—is naturally one of Čech cohomology with SO(2)-coefficients. We deduce that every cryo-EM reconstruction problem corresponds to an oriented circle bundle on a simplicial complex, allowing us to classify cryo-EM problems via principal bundles. In practice, the 2-dimensional images are noisy and a main task in cryo-EM is to denoise them. We will see how the aforementioned insights can be used towards this end.

Key words. oriented circle bundle, flat oriented circle bundle, co-cycle condition, impossible figure, classifying space, principal bundle, circular Radon transform

AMS subject classifications. 92E10, 46M20, 94A08, 68U10, 44A12, 55R35

DOI. 10.1137/16M1070220

1. Introduction.

The problem of cryo-electron microscopy (cryo-EM) asks for the following: Given a collection of noisy 2-dimensional (2D) projected images, reconstruct the 3-dimensional (3D) structure of the molecule that gave rise to these images. Viewed from a high level, it takes the form of an inverse problem similar to those in medical imaging [3, 4, 43], remote sensing [14, 5], or underwater acoustics [10, 36], except that for cryo-EM the data comes from an electron microscope instead of a CT scanner, radar, or sonar. However, when examined at a finer level of detail, one realizes that the cryo-EM problem possesses mathematical structures that are quite different from those of other classical inverse problems. It has inspired studies from the perspectives of representation theory [20, 21], differential geometry [47, 46], and is related to profound problems in computational complexity [6] and operator theory [7]. This paper examines the problem from an algebraic topological angle—we will show that the problem of cryo-EM is a problem of cohomology, or, more specifically, the Čech cohomology of a simplicial complex with coefficients in the Lie group SO(2) and the discrete group $\mathrm{SO}(2)_d$, i.e., SO(2) endowed with the discrete topology.

Despite its abstract appearance, the aforementioned cohomology framework is actually concrete and natural. The fact that cohomology has an important role to play in understand-

*Received by the editors April 12, 2016; accepted for publication (in revised form) April 17, 2017; published electronically August 22, 2017.

http://www.siam.org/journals/siaga/1/M107022.html

Funding: The authors were generously supported by AFOSR FA9550-13-1-0133, DARPA D15AP00109, NSF IIS 1546413, DMS 1209136, and DMS 1057064.

†Computational and Applied Mathematics Initiative, Department of Statistics, University of Chicago, Chicago, IL 60637 (kye@galton.uchicago.edu).

‡Corresponding author. Computational and Applied Mathematics Initiative, Department of Statistics, University of Chicago, Chicago, IL 60637 (lekheng@galton.uchicago.edu).

ing 2D projections of 3D objects is already evident in simple examples like the Penrose tribar or Escher brick, as we will see in section 2. Our analysis of discrete and continuous cryo-EM cocycles requires a more sophisticated type of cohomology but is essentially along the same lines. In fact, the same ideas that we use to study the cryo-EM problem also underlies the classical field theory of electromagnetism [11]. The cohomology framework allows us to classify cryo-EM cocycles: Given two different collections of 2D projected images, are they equivalent in the sense that they will give us the same 3D reconstruction? The insights gained also shed light on the denoising techniques: What are we really trying to achieve when we minimize a certain loss function to denoise cryo-EM images?

The technique of cryo-electron microscopy has been described in great detail in [17, 18] and more than adequately summarized in [20, 21, 38, 45, 47, 46, 48, 51, 52]. It suffices to provide a very brief review here. A more precise mathematical model for the following high-level description will be given in section 4. The basic idea is that one first immobilizes many identical copies of a molecule in ice and employs an electron microscope to produce 2D images of the molecule. As each copy of the molecule is frozen in some unknown orientation, each of the 2D images may be regarded as a projection of the molecule from an unknown viewing direction. The cryo-EM dataset is then the set of these 2D projected images. Such a 2D image shows not only the shape of the molecule in the plane of the viewing direction but also contains information about the density of the molecule, captured in the intensity of each pixel of the 2D image [35]. The ultimate goal of cryo-EM is to construct the 3D structure of the molecule from a cryo-EM dataset. In practice, these 2D images are very noisy due to various issues ranging from the electron dosage of the microscope to the structure of the ice in which the molecule are frozen. Hence the main difficulty in cryo-EM reconstruction is to denoise these 2D images by determining the true viewing directions of these noisy 2D images so that one may take averages of nearby images. There has been much significant progress toward this goal in recent years [38, 45, 48, 51, 52].

Our paper attempts to understand cryo-EM datasets of 2D images via Čech and singular cohomology groups. We will see that for a given molecule, the information extracted from its 2D cryo-EM images determines a cohomology class of a two-dimensional simplicial complex. Furthermore, each of these cohomology classes corresponds to an oriented circle bundle on this simplicial complex. We note that there are essentially two interpretations of cohomology: obstruction and moduli. On the one hand, a cohomology group quantifies the *obstruction* from local to global. For example, this is the sense in which cohomology is used when demonstrating the nonexistence of an impossible figure [40] or in the solution of the Mittag-Leffler problem [19, p. 34]. On the other hand, a cohomology group may also be used to describe a collection of mathematical objects, i.e., it serves as a *moduli* space for these objects. For example, when we use a cohomology group to parameterize all divisors or all line bundles on an algebraic variety [22, p. 143], it is used in this latter sense.

The line bundles example is a special case of a more general statement: A cohomology group serves as the moduli space of principal bundles over a topological space. This forms the basis for our use of cohomology in the cryo-EM reconstruction problem—as a moduli space for all possible cryo-EM datasets. Obviously, such a classification of cryo-EM datasets comes under the implicit assumption that the 2D images in a dataset are noise-free. Our classification depends on a standard mathematical *model* for molecules in the context of cryo-EM under a noise-free assumption. Here the reader is reminded that a molecule is a physical notion and not a mathematical one. A mathematical answer to the question "What is a molecule?" depends on the context. In one theory, a molecule may be a solution to a Schrödinger PDE (e.g., quantum chemistry) whereas in another, it may be a path in a $6N$-dimensional phase space (e.g., molecular dynamics). In our model, a molecule is a real-valued function on $\mathbb{R}^3$ representing potential. When our images are noisy, this model gives us a natural way, namely, the cocycle condition, to denoise them by fitting them to the model. Various methods for denoising cryo-EM images [45, 48] may be viewed as nonlinear regression for fitting the cocycle condition under additional assumptions.

2. Cohomology and 2D projections of 3D objects.

The idea that cohomology arises whenever one attempts to analyze 2D projections of 3D objects was first pointed out by Penrose, who proposed in [40] a cohomological argument to analyze Escher-type optical illusions. In the following, we present Penrose's elegantly simple example since it illustrates some of the same principles that underly our more complicated use of cohomology in cryo-EM.

We follow the spirit of Penrose's arguments in [40] but we will deviate slightly to be more in-line with our discussions of cryo-EM and to obtain a proof for the nonexistence of Penrose tribar. The few unavoidable topological jargons are defined in section 3 but they are used in such a way that one could grasp the intuitive ideas involved even without knowledge of the jargons. To be clear, a 3D object is one that can be embedded in $\mathbb{R}^3$ by an injective map J such that $J(ax + by) = aJ(x) + bJ(y)$ whenever x, y, $ax + by$ are points in this object, and $a, b \in \mathbb{R}$.

The Penrose *tribar* is defined to be a fictitious 3D object—fictitious as it does not exist in $\mathbb{R}^3$—obtained by gluing three rectangular solid cuboids (i.e., bars) L_1, L_2, L_3 in $\mathbb{R}^3$ as follows: L_i is glued to L_j by identifying a cubical portion L_{ij} at one end of L_i with a cubical portion L_{ji} at one end of L_j as depicted in Figure 1(b), $i, j = 1, 2, 3$.

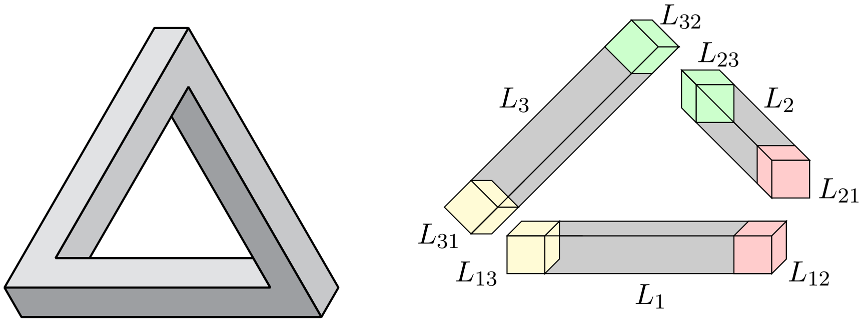


Figure 1. (a) *Projection of tribar into* $\mathbb{R}^2$. (b) *Decomposition into three overlapping pieces in* $\mathbb{R}^3$.

The tribar is more commonly shown in its 2D projected form as in Figure 1(a). Let Δ be the triangular 2D object in Figure 1(a), which appears to be the projection of the Penrose tribar, should it exist, onto a plane $H \cong \mathbb{R}^2$. Indeed, there are (infinitely) many 3D objects that, when projected onto a plane $H \cong \mathbb{R}^2$, gives Δ as an image. An example is the object in Figure 2, as we explain below.

Note that the object in Figure 2 is an abstraction of the sculpture in Figure 3, which

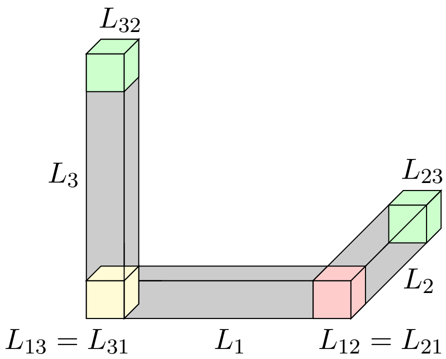


Figure 2. *A 3D object whose projection onto* $\mathbb{R}^2$ *is* Δ.

Figure 3. *The* Impossible Triangle *sculpture by Brian MacKay and Ahmad Abas, located in the Claisebrook Roundabout, Perth, Australia. Photograph by Bjørn Christian Tørrissen shared under a Creative Commons license.*

depicts how it projects to give Δ when viewed from an appropriate angle. The plane H in this case is either the viewer's retina or the camera's photographic film.

Let $H \subseteq \mathbb{R}^3$ be a hyperplane which partitions $\mathbb{R}^3$ into two half-spaces. Let $O \in \mathbb{R}^3$ be an arbitrary point in one half-space and the three bars L_1, L_2, L_3 be in the other. The reader should think of O as the position of the viewer and the viewing direction as a normal to H. Now we are going to arrange L_1, L_2, L_3 in such a way that their projections onto H give us Δ. This is clearly possible; for example, the 3D object in Figure 2, upon an appropriate rotation dependent on H and O, would give Δ as a projection.

Define $d_{ij} \in \mathbb{R}_+$ to be the distance from O to the center of L_{ij} and $d_{ii} = 1$, $i, j = 1, 2, 3$. Let $g = (g_{ij})_{i,j=1}^3$ be the 3×3 matrix of cross ratios

$$g_{ij} = \frac{d_{ij}}{d_{ji}}, \qquad i, j = 1, 2, 3.$$

Then g is a matrix with $g_{ij}^{-1} = g_{ji}$ and $g_{ii} = 1$ for all $i, j = 1, 2, 3$.

The matrix g is a function of the positions of the bars L_1, L_2, L_3, or, to be precise, a function of the centroids of these rigid bodies. These bars have a certain degree of freedom: We may *move* each of them independently along the viewing direction and this would keep their projections in $\mathbb{R}^2$ invariant, always forming Δ. This movement is a similarity transform that preserves the direction of the bar, with no rotation. Moving L_i in the viewing direction results in a rescaling of the distance d_{ij} by a factor $g_i \in \mathbb{R}_+$ for all $j \neq i$, i.e., if d'_{ij} denotes the new distance upon moving L_i's along viewing directions, then $d'_{ij} = d_{ij}/g_i$ for all $i \neq j$. Let $g' = (g'_{ij})_{i,j=1}^3$ be the new matrix of cross ratios upon moving L_i's along viewing direction. Then we have

$$g'_{ij} = \frac{d'_{ij}}{d'_{ji}} = \frac{d_{ij}/g_i}{d_{ji}/g_j} = g_{ij}\frac{g_j}{g_i}, \qquad i,j = 1,2,3. \tag{1}$$

Suppose that we could eventually move L_1, L_2, L_3 to form the tribar in $\mathbb{R}^3$. Then, in this final position, the centers of L_{ij} and L_{ji} coincide and so $d'_{ij} = d'_{ji}$ for all $i \neq j$, and thus $g'_{ij} = 1$ for all $i, j = 1,2,3$. In other words, the matrix g must be a *coboundary*, i.e.,

$$g_{ij} = \frac{g_i}{g_j}, \tag{2}$$

for some $g_i, g_j \in \mathbb{R}_+$, $i,j = 1,2,3$.

In summary, what we have shown is that if L_1, L_2, L_3 could be moved into place to form a tribar, then for L_1, L_2, L_3 in *any* positions that form Δ upon projection onto $\mathbb{R}^2$, the corresponding matrix g must be a coboundary, i.e., it satisfies (2), or equivalently, g is the identity element in the cohomology group $H^1(\mathbb{R}^2, \mathbb{R}_+)$. With this observation, we will next derive a contradiction showing that the tribar does not exist. Let L_1, L_2, L_3 be arranged as in Figure 2 and recall that their projections onto $\mathbb{R}^2$ give Δ. In this case, the matrix g is

$$g = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & g_{23} \\ 1 & g_{32} & 1 \end{bmatrix}.$$

If the tribar exists, then g is a coboundary, i.e., (2) has a solution for some $g_i, g_j \in \mathbb{R}_+$, $i,j = 1,2,3$, and so

$$g_1 = g_2 = g_3,$$

implying $g_{23} = 1$. However, as is evident from Figure 2, L_{23} does not even intersect L_{32} and so $g_{23} \neq 1$, a contradiction.

Although the tribar does not exist as a 3D object, i.e., it cannot be embedded in $\mathbb{R}^3$, it clearly exists as an abstract geometrical object (a cubical complex) defined by the gluing procedure described earlier—we will call this the *intrinsic tribar* to distinguish it from the nonexistent 3D object. In fact, the intrinsic tribar can be embedded in a 3-manifold $\mathbb{R}^3/\mathbb{Z}$, a quotient space of $\mathbb{R}^3$ under a certain action of the discrete group $\mathbb{Z}$ related to Figure 2 (see [16] for details).

We emphasize that a tribar is a geometrical object, not a topological one. It may be tempting to draw a parallel between the nonembeddability of the intrinsic tribar in $\mathbb{R}^3$ with the nonembeddability of the Möbius strip in $\mathbb{R}^2$ or the Klein bottle in $\mathbb{R}^3$. But these are different phenomena. As a topological object, a Möbius strip is only defined up to homotopy, i.e., we may freely deform a Möbius strip continuously. However, the definition of the tribar

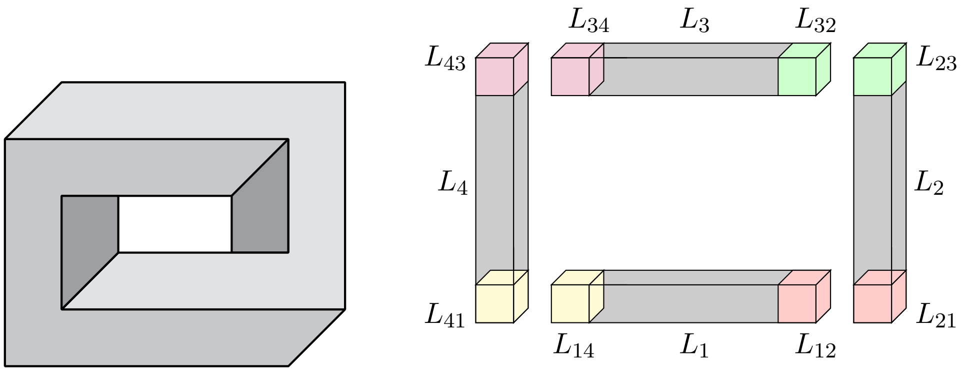


Figure 4. (a) *Projection of Escher brick into* $\mathbb{R}^2$. (b) *Decomposition into overlapping pieces in* $\mathbb{R}^3$.

does not afford this flexibility, i.e., a tribar is not homotopy invariant. For instance, we are not allowed to twist or bend the bars. In fact, had we allowed such continuous deformation, the intrinsic tribar is homotopy equivalent to a torus and, therefore, trivially embeddable in $\mathbb{R}^3$. This is much like our study of cryo-EM, where the goal is to reconstruct the 3D structure of a molecule precisely, and not just up to homotopy.

The discussions above also apply to other impossible objects in $\mathbb{R}^3$. For example, the *Escher brick*, defined as the (nonexistent) 3D object obtained by gluing four bars L_1, L_2, L_3, L_4 as in Figure 4. If the Escher brick exists in $\mathbb{R}^3$, then whenever L_1, L_2, L_3, L_4 projects onto $\mathbb{R}^2$ to form Figure 4(a), the matrix $g \in \mathbb{R}^{4\times 4}$ is necessarily a coboundary, i.e., satisfies $g_{ij} = g_i/g_j$ for some $g_i \in \mathbb{R}_+$, $i, j = 1, 2, 3, 4$. We may construct an analogue of Figure 2 whereby we glue three of the four ends in Figure 4(b). This 3D object projects onto $\mathbb{R}^2$ to form Figure 4(a) but its corresponding matrix $g \in \mathbb{R}^{4\times 4}$ is not a coboundary. Hence the Escher brick does not exist in $\mathbb{R}^3$.

3. Singular cohomology and Čech cohomology.

This paper is primarily intended for an applied and computational mathematics readership. For readers unfamiliar with algebraic topology, this section provides in one place all the required definitions and background material, kept to a bare minimum of just what we need for this paper.

We will define two types of cohomology groups associated to a topological space X and a topological group G that will be useful for our study of the cryo-EM problem: $H^n(X, G)$, the singular cohomology group with coefficients in G; and $\check{H}^n(X, G)$, the Čech cohomology group with coefficients in G. For a given X, these cohomology groups are in general different; but they would always be isomorphic for the space X that we construct from a given collection of cryo-EM images (see section 4). The reason we need both of them is that they are good for different purposes: the cohomology of cryo-EM is most naturally formulated in terms of Čech cohomology; but singular cohomology is more readily computable and facilitates our explicit calculations.

Our descriptions in the next few subsections are highly condensed, but in principle complete and self-contained. While this material is standard, our goal here is to make them accessible to practitioners by limiting the prerequisite to a few rudimentary definitions in point set topology and group theory. We provide pointers to standard sources at the beginning of each subsection.

We use $X \simeq Y$ to denote isomorphism if X, Y are groups, homotopy equivalence if X, Y are topological spaces, and bundle isomorphism if X, Y are bundles. We use $X \cong Y$ to denote homeomorphism of topological spaces.

3.1. Singular cohomology. Standard references for this section are [23, 32, 49].

The *standard n-simplex* for $n = 0, 1, 2, 3$, is the set

$$\Delta_n \coloneqq \left\{(t_0, \dots, t_n) \in \mathbb{R}^{n+1} : \sum\nolimits_{i=0}^{n} t_i = 1,\ t_i \geq 0\right\}.$$

Δ_n is the convex hull of its $n+1$ *vertices*,

$$e_0 = (0, 0, \dots, 0), \quad e_1 = (1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 1).$$

The standard 0-simplex is a point, the standard 1-simplex is a line, the standard 2-simplex is a triangle, and the standard 3-simplex is a tetrahedron.

For $n = 0, 1, 2$, the convex hull of any n vertices $e_{i_1}, \dots, e_{i_n}$ of Δ_n, where $0 \leq i_1 < \dots < i_n \leq n$, is called a *face* of Δ_n and denoted by $[i_1, \dots, i_n]$.

Let X be a topological space and let $n = 0, 1, 2, 3$. A continuous map $\sigma : \Delta_n \to X$ is called a *singular simplicial simplex* on X. We denote by $C_n(X)$ the free abelian group generated by all singular simplicial simplices on X. The *boundary maps* are homomorphisms of abelian groups

$$\partial_1 : C_1(X) \to C_0(X), \qquad \partial_2 : C_2(X) \to C_1(X), \qquad \partial_3 : C_3(X) \to C_2(X),$$

defined respectively by the linear extensions of

$$\begin{aligned}
\partial_1(\sigma) &= \sigma|_{[1]} - \sigma|_{[0]}, \\
\partial_2(\sigma) &= \sigma|_{[1,2]} - \sigma|_{[0,2]} + \sigma|_{[0,1]}, \\
\partial_3(\sigma) &= \sigma|_{[1,2,3]} - \sigma|_{[0,2,3]} + \sigma|_{[0,1,3]} - \sigma|_{[0,1,2]}.
\end{aligned}$$

Here $\sigma|_{[i]}$ denotes the restriction of σ to the face $[i]$ of Δ_1, $\sigma|_{[i,j]}$ denotes the restriction of σ to the face $[i, j]$ of Δ_2, and $\sigma|_{[i,j,k]}$ denotes the restriction of σ to the face $[i, j, k]$ of Δ_3. We set $\partial_0 : C_0(X) \to \{0\}$ to be the zero map.

The sequence of homomorphisms of abelian groups

$$C_3(X) \xrightarrow{\partial_3} C_2(X) \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\partial_0} 0 \tag{3}$$

forms a *chain complex*, i.e., it has the property that

$$\partial_0 \circ \partial_1 = 0, \qquad \partial_1 \circ \partial_2 = 0, \qquad \partial_2 \circ \partial_3 = 0, \tag{4}$$

which are easy to verify. For $n = 0, 1, 2$, let $Z_n(X) \coloneqq \operatorname{Ker} \partial_n \subseteq C_n(X)$ be the subgroup of *n-cycles* and let $B_n(X) \coloneqq \operatorname{Im} \partial_{n+1} \subseteq C_n(X)$ be the subgroup of *n-boundaries*. It follows from (4) that $B_n(X) \subseteq C_n(X)$. The quotient group

$$H_n(X) \coloneqq Z_n(X)/B_n(X)$$

is called the nth *singular homology group* of X, $n = 0, 1, 2$.

For $n = 0, 1, 2, 3$, define $C^n(X) = \operatorname{Hom}_{\mathbb{Z}}(C_n(X), \mathbb{Z})$, the set of all group homomorphisms from $C_n(X)$ to $\mathbb{Z}$. $C^n(X)$ is clearly an abelian group itself under addition of homomorphisms. The map *induced* by the boundary map $\partial_n : C_n(X) \to C_{n-1}(X)$ is defined as

$$\partial_n^* : C^{n-1}(X) \to C^n(X), \qquad \partial_n^*(f)(\sigma) = f(\partial_n(\sigma)),$$

for any $f \in C^{n-1}(X)$ and $\sigma \in C_n(X)$. The sequence of homomorphisms of abelian groups

$$0 \xrightarrow{\partial_0^*} C^0(X) \xrightarrow{\partial_1^*} C^1(X) \xrightarrow{\partial_2^*} C^2(X) \xrightarrow{\partial_3^*} C^3(X) \tag{5}$$

forms a *cochain complex*, i.e., it has the property that

$$\partial_1^* \circ \partial_0^* = 0, \qquad \partial_2^* \circ \partial_1^* = 0, \qquad \partial_3^* \circ \partial_2^* = 0, \tag{6}$$

which follows from (4). For $n = 0, 1, 2$, let $Z^n(X) := \operatorname{Ker} \partial_{n+1}^* \subseteq C^n(X)$ be the subgroup of n-*cocycles* and let $B^n(X) := \operatorname{Im} \partial_n^* \subseteq C^n(X)$ be the subgroup of n-*coboundaries*. The quotient group

$$H^n(X) := Z^n(X)/B^n(X)$$

is called the nth *singular cohomology group* of X, $n = 0, 1, 2$. More generally, let G be a group then one can define the nth *singular cohomology group* $H^n(X, G)$ *with coefficient* G of X to be the cohomology groups $Z^n(X, G)/B^n(X, G)$ of the cochain complex

$$0 \xrightarrow{\partial_0^*} C^0(X, G) \xrightarrow{\partial_1^*} C^1(X, G) \xrightarrow{\partial_2^*} C^2(X, G) \xrightarrow{\partial_3^*} C^3(X, G),$$

where $C^n(X, G) = \operatorname{Hom}_{\mathbb{Z}}(C_n(X), G)$, ∂_n^* is the map induced by $\partial_n : C_n(X) \to C_{n-1}(X)$, $n = 0, 1, 2$, and

$$Z^n(X, G) := \operatorname{Ker} \partial_{n+1}^* \subseteq C^n(X, G),$$
$$B^n(X, G) := \operatorname{Im} \partial_n^* \subseteq C^n(X, G).$$

Note that when $G = \mathbb{Z}$, $C^n(X, \mathbb{Z}) = C^n(X)$, $Z^n(X, \mathbb{Z}) = Z^n(X)$, $B^n(X, \mathbb{Z}) = B^n(X)$, $H^n(X, \mathbb{Z}) = H^n(X)$.

For the purpose of this paper, X would take the form of a *finite simplicial complex*, a collection K of finitely many simplices such that

(i) every face of a simplex in K is also contained in K;

(ii) the intersection of two simplices Δ_1, Δ_2 in K is a face of both Δ_1 and Δ_2.

We denote the union of simplices in K by $|K|$. We also say that a topological space X is a *finite simplicial complex* if X can be realized as $|K|$ for some finite simplicial complex K. For example, spheres $\mathbb{S}^n$ and tori $\mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ are finite simplicial complexes in this more general sense.

For the purpose of this paper, readers need only know that

$$H_0(\mathbb{S}^2) \simeq H_2(\mathbb{S}^2) \simeq \mathbb{Z}, \qquad H_1(\mathbb{S}^2) = 0, \qquad H^0(\mathbb{S}^2) \simeq H^2(\mathbb{S}^2) \simeq \mathbb{Z}, \qquad H^1(\mathbb{S}^2) = 0,$$

and that if X is a simplicial complex of dimension p, then $H_n(X) = 0$ for all $n > p$.

A topological space X is *contractible* if there is a point $x_0 \in X$ and a continuous map $H : X \times [0,1] \to X$ such that

$$H(x,0) = x_0 \qquad \text{and} \qquad H(x,1) = x.$$

Roughly speaking, this means that X can be continuously shrunk to a point x_0. For example, an open/closed/half-open-half-closed line segment is contractible, as is an open/closed disk or a disk with an arc on the boundary. The following is the only fact about contractible spaces that we need for this paper.

Proposition 3.1. *If X is contractible and G is an abelian group, then $H^n(X, G) = 0$ for all $n > 0$ and $H^0(X, G) = G$.*

3.2. Principal bundles and classifying spaces. Standard references for this section are [23, 25, 32, 33, 49].

Let G be a group with multiplication map $\mu : G \times G \to G$, $(x, y) \mapsto xy$ and inversion map $\iota : G \to G$, $x \mapsto x^{-1}$. If G is also a topological space such that μ and ι are continuous, then G together with this topology is called a *topological group*. Every group G is a topological group if we put the discrete topology on G; we will denote such a topological group by G_d (unless the natural topology is the discrete topology, in which case we will just write G). For example, $\mathbb{Z}$ with its natural discrete topology is a topological group. In this paper, we are primarily interested in the case where G is the group of 2×2 real orthogonal matrices. When endowed with the manifold topology, this is $\mathrm{SO}(2)$, the special orthogonal group in dimension two and is homeomorphic to the unit circle $\mathbb{S}^1$ as a topological space. On the other hand, $\mathrm{SO}(2)_d$ is just a discrete uncountable collection of 2×2 orthogonal matrices. Both $\mathrm{SO}(2)$ and $\mathrm{SO}(2)_d$ will be of interest to us.

Let X, P, F be topological spaces. We say that $\pi : P \to X$ is a *fiber bundle* with fiber F and base space X if π is a continuous surjection and every point of X has a neighborhood U such that $\pi^{-1}(U)$ is homoeomorphic to $U \times F$.

In particular, $\pi^{-1}(x) \cong F$ for all $x \in X$.

A *principal G-bundle* is a tuple (P, π, φ) where $\pi : P \to X$ is a fiber bundle with fiber G and $\varphi : G \times P \to P$ is a group action such that

(i) φ is a continuous map;
(ii) $\varphi(g, f) \in \pi^{-1}(x)$ for any $f \in \pi^{-1}(x)$;
(iii) if $\varphi(g, f) = f$ for some $f \in P$, then g is the identity element in G;
(iv) For any x and $f, f' \in \pi^{-1}(x)$, there is a $g \in G$ such that $\varphi(g, f) = f'$.

We will often say "P is a principal G-bundle on X" to mean the above, without specifying π and φ. A principal $\mathrm{SO}(2)$-bundle is called an *oriented circle bundle* and a principal $\mathrm{SO}(2)_d$-bundle is called a *flat oriented circle bundle*. We will have more to say about these in sections 4 and 5.

Let (P, π, φ) and (P', π', φ') be two principal G-bundles on X. We say that (P, π, φ) is *isomorphic* to (P', π', φ'), denoted $P \simeq P'$, if there is a homeomorphism $\vartheta : P \to P'$ compatible with the group actions φ, φ' and the projection maps π, π' in the following sense:

$$\vartheta \circ \varphi = \varphi' \circ (\mathrm{id}_G \times \vartheta) \quad \text{and} \quad \pi' \circ \vartheta = \pi.$$

Here $\mathrm{id}_G : G \to G$ is the identity map. Let $\mathcal{U} = \{U_i : i \in I\}$ be an open covering of X such that $\pi^{-1}(U_i) \cong U_i \times G$ via some isomorphism τ_i for all $i \in I$. A *transition function* corresponding to $\mathcal{U}$ is a map $\tau_{ij} := \tau_i \tau_j^{-1}$, defined for all $i, j \in I$ such that $U_i \cap U_j \neq \varnothing$. It may be regarded as a G-valued function $\tau_{ij} : U_i \cap U_j \to G$. Transition functions are important because one may construct a principal G-bundle entirely from its transition functions [25].

For $G = \mathrm{SO}(2)$, transition functions τ_{ij} of an oriented circle bundle are continuous $\mathrm{SO}(2)$-valued functions on open sets $U_i \cap U_j$. For $G = \mathrm{SO}(2)_d$, transition functions τ'_{ij} of a flat oriented circle bundle are continuous $\mathrm{SO}(2)_d$-valued functions on open sets $U_i \cap U_j$ but since $\mathrm{SO}(2)_d$ has the discrete topology, this means that τ'_{ij} are locally constant $\mathrm{SO}(2)$-valued functions on $U_i \cap U_j$. In particular, if $U_i \cap U_j$ is connected, then τ'_{ij} are constant $\mathrm{SO}(2)$-valued functions on $U_i \cap U_j$. In our case, the covering that we choose (see (13)) will have connected $U_i \cap U_j$'s and so we may regard

$$\big\{\text{isomorphism classes of flat oriented circle bundles}\big\} \\ \subseteq \big\{\text{isomorphism classes of oriented circle bundles}\big\}.$$

In other words, flat oriented circle bundles are just oriented circle bundles whose transition functions are constant-valued.

Let X, Y be topological spaces. Two maps $h_0, h_1 : X \to Y$ are *homotopic* if there is a continuous function $H : X \times I \to Y$ such that

$$H(x, 0) = h_0(x) \qquad \text{and} \qquad H(x, 1) = h_1(x).$$

Homotopy is an equivalence relation and the set of homotopy equivalent classes of maps from X to Y is denoted by $[X, Y]$. Let $\mathbb{S}^n$ be the n-sphere. We say that a topological space X is *weakly contractible* if $[\mathbb{S}^n, X]$ contains only the equivalence class of the *trivial map*, i.e., the map that sends all points in $\mathbb{S}^n$ into a fixed point of X. The *classifying space* of a topological group G is a topological space BG together with a principal G-bundle EG on BG such that EG is weakly contractible.

Proposition 3.2. *For any topological space X and topological group G, there is a one-to-one correspondence between the following two sets:*

$$[X, BG] \longleftrightarrow \{\textit{isomorphism classes of principal } G\textit{-bundles on } X\},$$

given by $h \mapsto h^*(EG)$, the principal G-bundle on X whose fiber over $x \in X$ is the fiber of EG over $h(x) \in BG$.

For the purpose of this paper, readers need only know that the classifying space $B\,\mathrm{U}(n)$ of the unitary group $\mathrm{U}(n)$ is $\mathrm{Gr}(n, \infty)$, the Grassmannian of n-planes in $\mathbb{C}^\infty$. In particular, if $n = 1$, since $\mathrm{U}(1) = \mathrm{SO}(2)$, we have

$$B\,\mathrm{SO}(2) = \mathbb{C}P^\infty. \tag{7}$$

Let G be an abelian group with identity 1. We write $\mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{Z})$ for the set of all homomorphisms from G to $\mathbb{Z}$. An element $g \in G$ is a *torsion* element if it has finite order, i.e., $g^n = 1$ for some $n \in \mathbb{N}$. The subgroup of all torsion elements in G is called its *torsion* *subgroup* and denoted G_T. For example, every element in $\mathbb{Z}/m\mathbb{Z}$ is a torsion element whereas 0 is the only torsion element in $\mathbb{Z}$. For an abelian group G, we also denote its torsion subgroup as

$$G_T = \mathrm{Ext}^1_{\mathbb{Z}}(G, \mathbb{Z}).$$

The reason for including this alternative notation is that it is very standard—a special case of Ext groups for G defined more generally [23, 24]. We now state some routine relations [24] that we will need for our calculations. Let G and G' be abelian groups. Then

$$\mathrm{Hom}_{\mathbb{Z}}(G_T, \mathbb{Z}) = 0, \qquad \mathrm{Hom}_{\mathbb{Z}}(G/G_T, \mathbb{Z}) \simeq G/G_T,$$
$$\mathrm{Hom}_{\mathbb{Z}}(G \oplus G', \mathbb{Z}) \simeq \mathrm{Hom}_{\mathbb{Z}}(G, \mathbb{Z}) \oplus \mathrm{Hom}_{\mathbb{Z}}(G', \mathbb{Z})$$

and

$$\mathrm{Ext}^1_{\mathbb{Z}}(G_T, \mathbb{Z}) = G_T, \qquad \mathrm{Ext}^1_{\mathbb{Z}}(G/G_T, \mathbb{Z}) = 0,$$
$$\mathrm{Ext}^1_{\mathbb{Z}}(G \oplus G', \mathbb{Z}) \simeq \mathrm{Ext}^1_{\mathbb{Z}}(G, \mathbb{Z}) \oplus \mathrm{Ext}^1_{\mathbb{Z}}(G', \mathbb{Z}).$$

Singular homology and singular cohomology are related via $\mathrm{Ext}^1_{\mathbb{Z}}$ and $\mathrm{Hom}_{\mathbb{Z}}$ in the following well-known theorem.

Theorem 3.3 (universal coefficient theorem). *Let X be a topological space. Then we have a natural short exact sequence*

$$0 \to \mathrm{Ext}^1_{\mathbb{Z}}(H_1(X), \mathbb{Z}) \to H^2(X) \to \mathrm{Hom}_{\mathbb{Z}}(H_2(X), \mathbb{Z}) \to 0.$$

In particular we have an isomorphism,

$$H^2(X) \simeq \mathbb{Z}^b \oplus T_1,$$

where $b := \mathrm{rank}(H_2(X)) = b_2(X)$ is the second Betti number of X and T_1 is the torsion subgroup of $H_1(X)$.

The second Betti number of X counts the number of 2D "voids" in X. In the case of interest to us, where X is a finite two-dimensional simplicial complex, the second Betti number counts the number of 2-spheres (by which we meant the boundary of a 3-simplex, which is homeomorphic to $\mathbb{S}^2$) contained in X.

We will also need the following alternative characterization [32, Chapter 22] of $H^2(X)$.

Theorem 3.4. *Let X be a topological space. Then we have*

$$[X, \mathbb{C}P^\infty] \simeq H^2(X).$$

3.3. Čech cohomology. Standard sources for this are [19, Chapter 0], [22, Chapter 3], and [27, Chapter 2].

Let G be a topological abelian group and let X be a topological space. For any open subset U of X we define an assignment

$$U \mapsto \underline{G}(U) := \text{group of } G\text{-valued continuous functions on } U$$

for all open subset $U \subseteq X$. By definition, if G is a discrete group and U is any connected open subset of X, then $\underline{G}(U) = G$. If $U \subseteq V$, then we have a *restriction map*

$$\rho_{V,U} : \underline{G}(V) \to \underline{G}(U)$$

defined by the restriction of G-valued continuous functions on V to U.

Let X be a topological space and let G be a topological abelian group on X. Let $\mathcal{U} = \{U_i : i \in I\}$ be an open covering of X. We may associate a cochain complex to X, G, and $\mathcal{U}$ as follows:

$$C^0(\mathcal{U}, G) \xrightarrow{\delta_0} C^1(\mathcal{U}, G) \xrightarrow{\delta_1} C^2(\mathcal{U}, G), \tag{8}$$

where

$$\begin{aligned} C^0(\mathcal{U}, G) &= \prod\nolimits_{i\in I} \underline{G}(U_i), \\ C^1(\mathcal{U}, G) &= \Big\{(g_{ij})_{i,j\in I} \in \prod\nolimits_{i,j\in I} \underline{G}(U_i \cap U_j) : g_{ij}g_{ji} = 1 \text{ for all } i,j \in I\Big\}, \\ C^2(\mathcal{U}, G) &= \Big\{(g_{ijk})_{i,j,k\in I} \in \prod\nolimits_{i,j,k\in I} \underline{G}(U_i \cap U_j \cap U_k) : \\ &\qquad\qquad g_{ijk}g_{ikj} = g_{ijk}g_{kji} = g_{ijk}g_{jik} = 1 \text{ for all } i,j,k \in I\Big\}, \end{aligned}$$

and

$$\begin{aligned} \big(\delta_0(g_i)_{i\in I}\big)_{j,k} &= g_k g_j^{-1} && \text{for all } j,k \in I, \\ \big(\delta_1(g_{ij})_{i,j\in I}\big)_{k,l,m} &= g_{lm}g_{mk}g_{kl} && \text{for all } k,l,m \in I. \end{aligned}$$

To be precise, we have

$$\begin{aligned} g_k g_j^{-1} &= \rho_{U_k, U_k\cap U_j}(g_k) \cdot \rho_{U_j, U_k\cap U_j}(g_j^{-1}), \\ g_{lm}g_{mk}g_{kl} &= \rho_{U_l\cap U_m, U_k\cap U_l\cap U_m}(g_{lm}) \cdot \rho_{U_k\cap U_m, U_k\cap U_l\cap U_m}(g_{mk}) \cdot \rho_{U_k\cap U_l, U_l\cap U_m\cap U_k}(g_{kl}). \end{aligned}$$

It is easy to check that $\delta_1 \circ \delta_0 = 0$ and so (8) indeed forms a cochain complex.

As in the case of singular cohomology, $\check{B}^1(\mathcal{U}, G) \coloneqq \operatorname{Im}\delta_0$ and $\check{Z}^1(\mathcal{U}, G) \coloneqq \operatorname{Ker}\delta_1$ are the groups of *Čech* 1*-coboundaries* and *Čech* 1*-cocycles*, respectively. Again we have $\check{B}^1(\mathcal{U}, G) \subseteq \check{Z}^1(\mathcal{U}, G)$. The *first Čech cohomology group associated to* $\mathcal{U}$ with coefficients in G is then defined to be the quotient group

$$\check{H}^1(\mathcal{U}, G) \coloneqq \check{Z}^1(\mathcal{U}, G)/\check{B}^1(\mathcal{U}, G).$$

Explicitly, we have

$$\check{H}^1(\mathcal{U}, G) = \frac{\{(g_{ij}) : g_{ij}g_{jk}g_{ki} = 1 \text{ for all } i,j,k\}}{\{(g_{ij}) : g_{ij} = g_j g_i^{-1} \text{ for all } i,j\}}.$$

We have in fact already encountered this notion in section 2, $\check{H}^1(\mathbb{R}^2, \mathbb{R}_+)$, the Čech cohomology group of the plane $\mathbb{R}^2$ with coefficients in the group $\mathbb{R}_+$ has appeared implicitly in our discussion.

By its definition, $\check{H}^1(\mathcal{U}, G)$ depends on the choice of open covering $\mathcal{U}$ of X. To obtain a Čech cohomology group of X independent of open covering, we take the *direct limit* over all possible open coverings of X. The *first Čech cohomology group* of X with coefficients in G is defined to be the direct limit

$$\check{H}^1(X, G) := \varinjlim \check{H}^1(\mathcal{U}, G)$$

with $\mathcal{U}$ running through all open coverings of X.

For those unfamiliar with the notion of direct limit, $\check{H}^1(X, G)$ may be defined explicitly using an equivalence relation:

$$\check{H}^1(X, G) := \left[\coprod_{\mathcal{U}} \check{H}^1(\mathcal{U}, G)\right] \Big/ \sim,$$

where $\coprod_{\mathcal{U}}$ denotes the disjoint union of $\check{H}^1(\mathcal{U}, G)$ for all possible open coverings of X. The equivalence relation $\sim$ is given as follows: For $\varphi_{\mathcal{U}} \in \check{H}^1(\mathcal{U}, G)$ and $\varphi_{\mathcal{V}} \in \check{H}^1(\mathcal{V}, G)$, $\varphi_{\mathcal{U}} \sim \varphi_{\mathcal{V}}$ if and only if

(i) there is an open covering $\mathcal{W}$ such that every open set $W \in \mathcal{W}$ is contained in $U \cap V$ for some $U \in \mathcal{U}$ and $V \in \mathcal{V}$;

(ii) there is an element $\varphi_{\mathcal{W}} \in \check{H}^1(\mathcal{W}, G)$ such that the *restriction* of $\varphi_{\mathcal{U}}$ and the restriction of $\varphi_{\mathcal{V}}$ are both equal to $\varphi_{\mathcal{W}}$.

The term "restriction" needs elaboration. Let $\mathcal{U} = \{U_i : i \in I\}$, $\mathcal{V} = \{V_\alpha : \alpha \in \Lambda\}$ be open covers of X such that for any $U_i \in \mathcal{U}$, there is some $V_{\alpha_i} \in \mathcal{V}$ with $U_i \subseteq V_{\alpha_i}$. Fix a map $\tau : I \to \Lambda$ such that $U_i \subseteq V_{\tau(i)}$. There is a natural restriction map $\rho_{\mathcal{V},\mathcal{U}} : \check{H}^1(\mathcal{V}, G) \to \check{H}^1(\mathcal{U}, G)$ induced by $\widetilde{\rho}_{\mathcal{V},\mathcal{U}} : C^1(\mathcal{V}, G) \to C^1(\mathcal{U}, G)$, where

$$(\widetilde{\rho}_{\mathcal{V},\mathcal{U}}(g_{\alpha,\beta}))_{i,j} = \rho_{V_{\tau(i)} \cap V_{\tau(j)}, U_i \cap U_j}(g_{\tau(i),\tau(j)}).$$

The image $\rho_{\mathcal{V},\mathcal{U}}(\varphi)$ of $\varphi \in \check{H}^1(\mathcal{V}, G)$ is called the *restriction* of φ to $\check{H}^1(\mathcal{U}, G)$. It does not depend on the choice of τ.

As the reader can guess, calculating the Čech cohomology group using such a definition would in general be difficult. Fortunately, the following theorem (really a special case of Leray's theorem [15]) allows us to simplify the calculation in all cases of interest to us in this paper.

Theorem 3.5 (Leray's theorem). *Let X be a topological space and let G be a topological abelian group. Let $\mathcal{U} = \{U_i : i \in I\}$ be an open cover of X such that $\check{H}^1(U_i, G) = 0$ for all $i \in I$. Then we have*

$$\check{H}^1(\mathcal{U}, G) \simeq \check{H}^1(X, G).$$

Furthermore, we will often be able to reduce calculation of Čech cohomology to calculation of singular cohomology since they are equal in the case when X is a finite simplicial complex [41].

Theorem 3.6. *If K is a finite simplicial complex and G is an abelian group, then*

$$\check{H}^1(K, G_d) \simeq H^1(K, G),$$

where G_d is the group G equipped with the discrete topology.

For a contractible space, we have $H^1(K, G) = 0$ by Proposition 3.1. So we may deduce the following from Theorem 3.6.

Corollary 3.7. *If K is a finite contractible simplicial complex and G is an abelian group, then*

$$\check{H}^1(K, G_d) = 0.$$

To check whether an oriented circle bundle on a finite simplicial complex K is flat, we have the following useful result [29, 34, 37].

Proposition 3.8. *An oriented circle bundle on K is flat if and only if its Euler class is a torsion element in $H^2(K)$.*

Defining the Euler class of an oriented circle bundle would take us too far afield and so this will be the only term left undefined in our paper. Fortunately, all we need is the following corollary of Proposition 3.8.

Corollary 3.9. *If $H^2(K)$ is torsion free, then any oriented circle bundle on K must be flat.*

A particularly important result [9, 26] for us is the following theorem that relates the Čech cohomology group with G-coefficients and principal G-bundles.

Theorem 3.10. *If G is a topological abelian group, then $\check{H}^1(X, G)$ is in canonical one-to-one correspondence with the set of isomorphism classes of principal G-bundles on X.*

4. Cohomological classification of discrete cryo-EM cocycles.

We will follow the mathematical setup for the cryo-EM problem as laid out in [20, 21]. First, recall the high-level description of the problem: Given cocycles comprising a collection of noisy 2D projected images, reconstruct the 3D structure of the molecule that gave rise to these images. The *standard mathematical model* for cryo-EM casts the problem in mathematical terms and may be described as follows:

(i) The *molecule* is described by a function $\varphi : \mathbb{R}^3 \to \mathbb{R}$, the potential function of the molecule.

(ii) A *viewing direction* is described by a point on the 2-sphere $\mathbb{S}^2$.

(iii) The *position* of an image is described by a 3×3 matrix $A = [a, b, c] \in \mathrm{SO}(3)$ where the orthonormal column vectors a, b, c are such that $\operatorname{span}\{a, b\}$ is the projection plane and c is the viewing direction.

(iv) A *projected image* ψ of the molecule φ by A is described by a function $\psi : \mathbb{R}^2 \to \mathbb{R}$ where

$$\psi(x, y) = \int_{z\in\mathbb{R}} \varphi(xa + yb + zc)\, dz.$$

The function ψ describes the *density* of the molecule along the chosen viewing direction.

Let $\Psi = \{\psi_1, \dots, \psi_n\}$ be a set of n projected images of the molecule and $c_1, \dots, c_n$ be the corresponding viewing directions. It is common to impose two mild assumptions:

(a) The function φ is generic. In particular, each image $\psi_i \in \Psi$ has a uniquely determined viewing direction. In practice, this means that the molecule has no extra symmetry. This assumption does not exclude the possibility where two images ψ_i, ψ_j may share the same viewing direction. However, it excludes the case where an image ψ_i can be obtained from projections of the molecule from two different directions.

(b) The viewing directions $c_1, \dots, c_n \in \mathbb{S}^2$ are distributed uniformly on $\mathbb{S}^2$. This is a standard assumption in cryo-EM literature although in practice, viewing directions are rarely uniformly distributed.

In addition, since each image ψ_i is associated with a viewing direction c_i, we should regard ψ_i to be a real-valued function on the tangent plane to $\mathbb{S}^2$ with unit normal in the direction of c_i. This is the point of view adopted in [48] and we will assume it throughout this paper. An important distinction between cryo-EM and other reconstruction problems in medical imaging, remote sensing, underwater acoustics, etc, mentioned in section 1 is that for the former, the viewing directions $c_1, \dots, c_n$ are *unknown* and have to be determined from the data set Ψ, whereas for the latter, we usually know in which directions the imaging instruments (CT scanner, camera, radar, sonar, etc) are pointed. In fact, determining $c_1, \dots, c_n$ from Ψ is the most crucial step in cryo-EM—our goal is to show that there is some interesting algebraic topology behind this problem.

Henceforth, by a "molecule," we will mean one in the standard mathematical model, i.e., a function φ. These include φ's that do not correspond to any actual molecules. We assume that $\varphi \in L^2(\mathbb{R}^3)$ and $\psi_1, \dots, \psi_n \in L^2(\mathbb{R}^2)$. There is a natural notion of distance [38] between projected images $\Psi = \{\psi_1, \dots, \psi_n\}$ given by

$$d(\psi_i, \psi_j) = \min_{g \in \mathrm{SO}(2)} \|g \cdot \psi_i - \psi_j\|,$$

where $\|\cdot\|$ is the norm in $L^2(\mathbb{R}^2)$ and the action of $g \in \mathrm{SO}(2)$ on a projected image ψ is

$$(g \cdot \psi)(x, y) = \psi(g^{-1}(x, y)).$$

Geometrically, the action of g on ψ is the rotation of ψ by the angle represented by $g \in \mathrm{SO}(2)$. Let g_{ij} be the element in $\mathrm{SO}(2)$ which realizes the minimum of the distance $d(\psi_i, \psi_j)$, i.e.,

$$g_{ij} \coloneqq \operatorname*{argmin}_{g \in \mathrm{SO}(2)} \|g \cdot \psi_i - \psi_j\| \tag{9}$$

for $i, j = 1, \dots, n$. Clearly, we have

$$g_{ii} = 1_n \qquad \text{and} \qquad g_{ij} g_{ji} = 1_n \tag{10}$$

for all $i, j = 1, \dots, n$, where 1_n is the $n \times n$ identity matrix, which we will henceforth denote simply as 1 when there is no cause for confusion. In general, g_{ij} is not unique since it could happen that two different rotations both minimize the distance but our assumption that the function φ is generic ensures that g_{ij} is uniquely determined by ψ_i and ψ_j. We will call

$$D \coloneqq \{g_{ij} \in \mathrm{SO}(2) : i, j = 1, \dots, n\} \tag{11}$$

the set of *pairwise angular comparisons*. This is of course derived from the raw image data set Ψ and the process of extracting D from Ψ is itself an active research topic [6, 7], particularly when the images ψ_i's are noisy. We will not concern ourselves with this auxiliary problem here.

We will use notations consistent with those introduced in section 3.1 for simplices. For any $\varepsilon > 0$, we may construct an undirected graph $G_\varepsilon = (V, E)$ where $V = \{[1], \ldots, [n]\}$ is the set of vertices corresponding to the projected images $\Psi = \{\psi_1, \ldots, \psi_n\}$, and E is the set of edges defined by

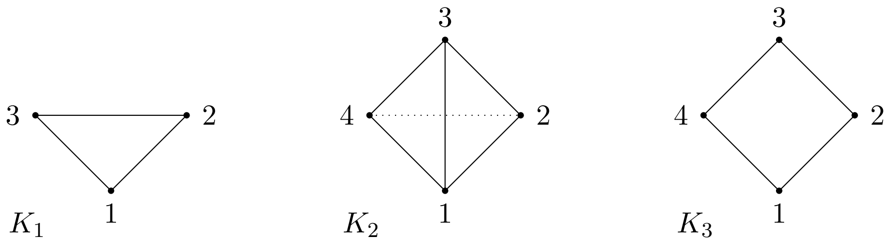
$$[i, j] \in E \quad \text{if and only if} \quad d(\psi_i, \psi_j) \le \varepsilon. \tag{12}$$

Let us first consider an ideal situation where the projected images ψ_i's are noiseless. Also we fix $\varepsilon > 0$ and the number of images n. Let G_ε be the associated undirected graph. We define the *cryo-EM complex* K_ε as follows:

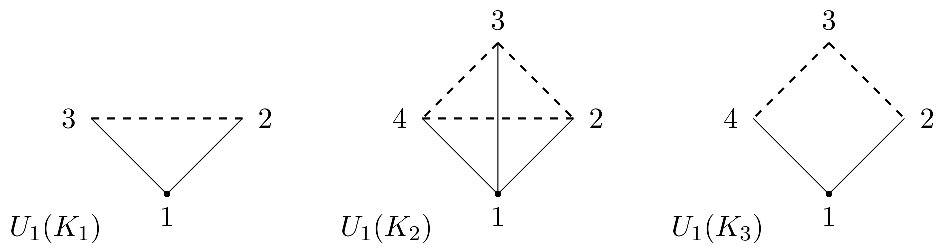
(i) the 0-simplices of K_ε are the vertices of G_ε,
(ii) the 1-simplices of K_ε are the edges of G_ε,
(iii) the 2-simplices of K_ε are the triangles $[i, j, k]$ such that $[i, j], [i, k], [j, k]$ are all edges of G_ε.

K_ε is a two-dimensional finite simplicial complex. It is the 3-*clique complex* [8, 31] of the graph G_ε. In addition, K_ε is also the *Vietoris–Rips complex* [12, 53] defined by (12) with respect to the metric d.

Some simple examples: The graph $G_1 = (V_1, E_1)$ with $V_1 = \{[1], [2], [3]\}$ and $E_1 = \{[1, 2], [1, 3], [2, 3]\}$ defines a simplicial complex K_1 that is a triangle. The graph $G_2 = (V_2, E_2)$ with $V_2 = \{[1], [2], [3], [4]\}$ and $E_2 = \{[1, 2], [1, 3], [2, 3], [1, 4], [2, 4], [3, 4]\}$ defines a simplicial complex K_2 that is the boundary of a tetrahedron or 3-simplex. The graph $G_3 = (V_3, E_3)$ with $V_3 = \{[1], [2], [3], [4]\}$ and $E_3 = \{[1, 2], [2, 3], [1, 4], [3, 4]\}$ defines a simplicial complex K_3 that is the boundary of a square.



We will regard our simplicial complex K_ε as being embedded in $\mathbb{R}^4$ and inherits the Euclidean topology from $\mathbb{R}^4$, i.e., K_ε is a geometric simplicial complex and not just an abstract simplicial complex. For each vertex $[i]$ of K_ε we define an open set $U_i(K_\varepsilon)$ to be the union of the interior of all simplices of K_ε containing the vertex $[i]$. Those familiar with simplicial complex might like to note that $U_i(K_\varepsilon)$ is just the complement of the *link* of $[i]$ in the *star* of $[i]$. For example, $U_1(K_i)$ for $i = 1, 2, 3$ are shown below. Here dashed lines are excluded from the neighborhood.



It follows from our definition of $U_i(K_\varepsilon)$ that

$$\mathcal{U} = \{U_i : [i] \text{ is a vertex of } K_\varepsilon\} \tag{13}$$

is an open covering of K_ε.

Let φ be a fixed molecule and let $\Psi = \{\psi_1, \dots, \psi_n\}$ be a set of projected images of φ. The set of pairwise angular comparisons $D = \{g_{ij} \in \mathrm{SO}(2) : i, j = 1, \dots, n\}$ contains all g_{ij}'s corresponding to every pair of images ψ_i, ψ_j. For the purpose of cryo-EM reconstruction, one does not usually need all elements in the D [48], only a much smaller subset comprising the g_{ij}'s corresponding to images ψ_i, ψ_j that are near each other, i.e., $d(\psi_i, \psi_j) \le \varepsilon$ for some small $\varepsilon > 0$. This is expected since most reconstruction methods proceed by aggregating *local* information. With this in mind, we define the following.

Definition 4.1. *Let $D = \{g_{ij} \in \mathrm{SO}(2) : i, j = 1, \dots, n\}$ be the set of pairwise angular comparisons. Let $\varepsilon > 0$ and K_ε be the cryo-EM complex. The discrete cryo-EM cocycle on K_ε is the subset of D corresponding to edges in K_ε given by*

$$z_\varepsilon^d := \{g_{ij} \in \mathrm{SO}(2) : [i, j] \in K_\varepsilon\}.$$

We may view z_ε^d as the "useful" part of the set of pairwise angular comparisons D for cryo-EM reconstruction. In fact we are unaware of any reconstruction method that makes use of g_{ij} where $[i, j] \notin K_\varepsilon$.

As we mentioned earlier in this section, we take the point of view in [48] that the projected images ψ_i's lie in tangent planes of a 2-sphere determined by their viewing directions. We also assume, as in [48], that if the images ψ_i, ψ_j, and ψ_k have viewing directions close enough, then they lie in the *same* tangent plane. This assumption is reasonable since if ψ_i and ψ_j share the same viewing direction, then they will only differ by a plane rotation. Moreover, if ψ_i, ψ_j, and ψ_k share the same viewing direction, then the angle needed to rotate ψ_i to ψ_k is the sum of the angle needed to rotate ψ_i to ψ_j and the angle needed to rotate ψ_j to ψ_k—implying that the g_{ij}'s corresponding to $\Psi = \{\psi_1, \dots, \psi_n\}$ satisfy the following 1-*cocycle condition*:

$$g_{ij} g_{jk} g_{ki} = 1. \tag{14}$$

Here 1 is the identity matrix in $\mathrm{SO}(2)$. Note that the matrices g_{ij}'s in the discrete cryo-EM cocycle always satisfy (10), irrespective of whether viewing directions are close enough.

By the preceding discussion, we will *assume* that for $\varepsilon > 0$ small enough, the g_{ij}'s will satisfy the 1-cocycle condition (14) for all edges $[i, j], [j, k], [k, i]$ of the cryo-EM complex K_ε. One motivation for this assumption is that when $\varepsilon \to 0$, images that lie in an ε-neighborhood will share the same viewing direction and thus g_{ij}'s will satisfy the cocycle condition (14). Therefore, "small enough ε" should be taken mathematically to mean the value of ε such that (14) holds, bearing in mind that (14), like any mathematical model, is ultimately only an approximation of reality. Our assumption that the 1-cocycle condition is satisfied for small enough $\varepsilon > 0$ is a basic tenet for our subsequent discussions. As far as we know, this assumption is not in existing cryo-EM literature although it is closely related to the "same tangent plane" assumption in [48]. While never explicitly stated, (14) is the implicit principle underlying many, if not most, denoising techniques for cryo-EM images [47, 48, 44], as we will see in section 6.

Given an open subset U of K_ε, any element $g \in \mathrm{SO}(2)$ can be regarded as the constant $\mathrm{SO}(2)$-valued function sending every point $x \in U$ to g, and thus we may regard z_ε^d as a cocycle in $\check{Z}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)$. We highlight this observation as follows:

Every discrete cryo-EM cocycle on K_ε is an $\mathrm{SO}(2)_d$-valued Čech 1-cocycle on K_ε.

Henceforth , we will regard

$$\check{Z}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big) = \big\{\text{all discrete cryo-EM cocycles on } K_\varepsilon\big\}.$$

The set on the right includes all possible discrete cryo-EM cocycles on K_ε corresponding to all molecules φ. A cocycle z_ε^d only tells us how to glue together local information. It is possible for two different 3D molecules to give the same discrete cryo-EM cocycle z_ε^d as long as the relations between their projected images are the same.

Given a discrete cryo-EM cocycle $z_\varepsilon^d \in \check{Z}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)$, i.e., elements in z_ε^d satisfy (14), and any arbitrary image $\psi \in L^2(\mathbb{R}^2)$, we may apply each $g \in z_\varepsilon^d$ to ψ to obtain a set of images

$$z_\varepsilon^d(\psi) := \{g \cdot \psi : g \in z_\varepsilon^d\} = \{g_{ij} \cdot \psi : [i,j] \in K_\varepsilon\}.$$

The cocycle condition (14) ensures that for any image $g \cdot \psi$ in this set, we obtain the same set of images by applying each $g \in z_\varepsilon^d$, i.e.,

$$z_\varepsilon^d(g \cdot \psi) = z_\varepsilon^d(\psi) \qquad \text{for any } g \in z_\varepsilon^d.$$

Moreover, the discrete cryo-EM cocycle obtained would be exactly z_ε^d. A set of projected images $z_\varepsilon^d(\psi)$ allows one to reconstruct the 3D molecule φ whose projected images are precisely the ones in $z_\varepsilon^d(\psi)$ [17, 18, 42, 39]. Put in another way, given a discrete cryo-EM cocycle $z_\varepsilon^d \in \check{Z}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)$ and an image $\psi \in L^2(\mathbb{R}^2)$, we may construct a 3D molecule $\varphi \in L^2(\mathbb{R}^3)$ whose discrete cryo-EM cocycle is exactly z_ε^d and one of whose projected image is ψ.

The context for the following theorem is that we are given two collections of n projected images $\Psi = \{\psi_1, \dots, \psi_n\}$ and $\Psi' = \{\psi'_1, \dots, \psi'_n\}$ of the same molecule φ. These give two discrete cryo-EM cocycles $D = \{g_{ij} \in \mathrm{SO}(2) : i,j = 1, \dots, n\}$ and $D' = \{g'_{ij} \in \mathrm{SO}(2) : i,j = 1, \dots, n\}$. Let $\varepsilon > 0$ be sufficiently small and let $z_\varepsilon^d = \{g_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$, $z'^d_\varepsilon = \{g'_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ be the corresponding discrete cryo-EM cocycles on K_ε.

Theorem 4.2 (bundle classification of discrete cryo-EM cocycles). *Let $\varepsilon > 0$ be small enough so that (14) holds and let K_ε be the corresponding cryo-EM complex. Then*

(i) *the 1-cocycle z_ε^d determines a flat oriented circle bundle on K_ε;*

(ii) *two 1-cocycles z_ε^d and z'^d_ε for the same molecule determine isomorphic flat oriented circle bundles if and only if*

$$g'_{ij} = g_{ij} g_i g_j^{-1} \tag{15}$$

for some $g_i, g_j \in \mathrm{SO}(2)$, $[i,j] \in K_\varepsilon$.

Proof. Let $\mathcal{U} = \{U_i(K_\varepsilon) : i = 1, \dots, n\}$ be the open cover defined in (13). It is easy to see that $U_i(K_\varepsilon)$ is contractible and so by Corollary 3.7,

$$\check{H}^1\big(U_i(K_\varepsilon), \mathrm{SO}(2)_d\big) = \{1\}$$

for all $i = 1, \dots, n$. We may then apply Theorem 3.5 to get

$$\check{H}^1\big(\mathcal{U}, \mathrm{SO}(2)_d\big) \simeq \check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big).$$

Therefore, it follows from Theorem 3.10 that $\check{H}^1\big(\mathcal{U}, \mathrm{SO}(2)_d\big)$ is canonically in one-to-one correspondence with the set of isomorphism classes of $\mathrm{SO}(2)_d$-principal bundles, i.e., flat oriented circle bundles. Since the subset $z_\varepsilon^d = \{g_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ is a 1-cocycle in $\check{H}^1(\mathcal{U}, \mathrm{SO}(2)_d)$, it determines an oriented circle bundle over K_ε. Part (ii) follows from the fact that the 1-cocycle $b_\varepsilon = \{g_i g_j^{-1} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ is a 1-coboundary and thus represents the trivial cohomology class. ∎

If the reader finds (15) familiar, that is because we have seen a similar version (1) in our discussion of the Penrose tribar. The difference here is that the quantities in (1) are from the group $\mathbb{R}_+$ whereas the quantities in (15) are from the group SO(2). Two cocycles $z_\varepsilon^d = \{g_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ and $z_\varepsilon'^d = \{g'_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ are said to be *cohomologically equivalent* if and only if they differ by a coboundary $b_\varepsilon = \{g_i g_j^{-1} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ in the sense of (15). Cohomologically equivalent z_ε^d and z'^d_ε define the same *cohomology class* in the quotient group and we have

$$\begin{aligned}\check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big) &:= \check{Z}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)/\check{B}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)\\ &= \big\{\text{cohomologically equivalent discrete cryo-EM cocycles on } K_\varepsilon\big\}.\end{aligned}$$

By Proposition 3.2, the cohomology group $\check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)$ can be identified as sets with the classifying space $[K_\varepsilon, B\,\mathrm{SO}(2)_d]$, which classifies the isomorphism classes of flat oriented circle bundles on K_ε. We obtain a canonical one-to-one correspondence

$$\begin{aligned}&\big\{\text{cohomologically equivalent discrete cryo-EM cocycles on } K_\varepsilon\big\}\\ &\qquad\longleftrightarrow \big\{\text{isomorphism classes of flat oriented circle bundles on } K_\varepsilon\big\}.\end{aligned} \tag{16}$$

Finally, we arrive at the following result.

Theorem 4.3. *Let $\varepsilon > 0$ be small enough so that (14) holds and let K_ε be the corresponding cryo-EM complex. Then*

(i) *every flat oriented circle bundle on K_ε is the trivial circle bundle;*
(ii) *all discrete cryo-EM cocycles on K_ε are coboundaries $b_\varepsilon = \{g_i g_j^{-1} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$.*

Proof. By Proposition 3.8, it suffices to show that $H^2(K_\varepsilon)$ is torsion free. But this follows from Theorem 3.3: By our construction of K_ϵ, the simplicial complex is actually homotopic to a one-point union of several spheres or a one-point union of several circles. This implies that either $H_1(K_\epsilon) = 0$ or $H_1(K_\epsilon) \simeq \mathbb{Z}^r$ for some integer $r \geq 1$. In particular, $H_1(K_\epsilon)$ is torsion free, i.e., $T_1 = 0$. ∎

In other words, the set on the right of (16) is a singleton comprising only the trivial bundle. Consequently, discrete cryo-EM cocycles on K_ε are all cohomologically equivalent and all correspond to the trivial circle bundle. So Theorem 4.2 does not provide an interesting classification. The reason is that a discrete cryo-EM cocycle as defined by (9), i.e., an element of $\check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)_d\big)$, is too coarse. In the next section, we will see how the classification becomes more interesting mathematically when we look at *continuous* cryo-EM cocycles.

5. Cohomological classification of continuous cryo-EM cocycles. In the standard mathematical model for cryo-EM, a projected image is a function $\psi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\psi(x, y) = \int_{z \in \mathbb{R}} \varphi(xa + yb + zc)\, dz,$$

where $A = [a, b, c] \in \mathrm{SO}(3)$ describes the orientation of the molecule in $\mathbb{R}^3$ and φ is the potential function of the molecule. For every pair of images ψ_i, ψ_j we define an SO(2)-valued function

$$h_{ij}(r) := \operatorname{argmin}_{g \in \mathrm{SO}(2)} \int_0^{2\pi} |(g \cdot \psi_i)(r \cos\theta, r \sin\theta) - \psi_j(r\cos\theta, r\sin\theta)|^2\, d\theta, \tag{17}$$

where $r = \sqrt{x^2 + y^2}$. These h_{ij}'s should be interpreted as follows: We regard a 2D image ψ_i as comprising circular "slices" of different radii as in Figure 5, i.e., each slice is the intersection of the image ψ_i with a circle of radius r. For each pair i, j, $h_{ij}(r) \in \mathrm{SO}(2)$ is the rotation that minimizes the difference between the slice of ψ_i of radius r and the slice of ψ_j of radius r.

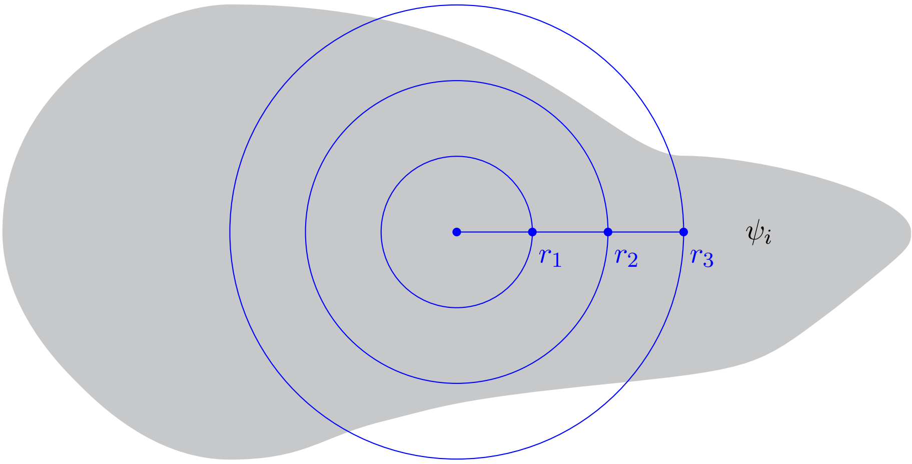


Figure 5. *Circular slices of an image* ψ_i.

The integral in (17) is in fact a restriction of the *circular Radon transform* [28], defined for a compactly supported $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$Sf : \mathbb{R}^2 \times (0, \infty) \to \mathbb{R}, \qquad Sf(\eta, \xi, r) = \int_{(x-\eta)^2 + (y-\xi)^2 = r} f(x, y)\, d\sigma(x, y),$$

where σ denotes the surface measure on the circle of radius r centered at (η, ξ). Although it has not been used, as far as we know, in cryo-EM applications, the circular Radon transform is common in a variety of other applications, e.g., thermoacoustic tomography and optoacoustic tomography [2, 13, 30, 54, 1]. If we set $(\eta, \xi) = (0, 0)$, $x = r\cos\theta$, $y = r\sin\theta$, then

$$Sf(0, 0, r) = \int_0^{2\pi} f(r\cos\theta, r\sin\theta)\, d\theta,$$

and so (17) is the circular Radon transform of $|g \cdot \psi_i - \psi_j|^2$ at $(0,0,r)$.

Let $\varepsilon > 0$ and the potential function $\varphi : \mathbb{R}^3 \to \mathbb{R}$ be chosen so that $h_{ij}(r)$ satisfies the 1-cocycle condition

$$h_{ij}(r)h_{jk}(r)h_{ki}(r) = 1 \tag{18}$$

for all $r > 0$ whenever the images ψ_i, ψ_j, and ψ_k are such that

$$d(\psi_i, \psi_j) \leq \varepsilon, \qquad d(\psi_j, \psi_k) \leq \varepsilon, \qquad d(\psi_k, \psi_i) \leq \varepsilon.$$

We remind readers that the existence of such an ε that guarantees (18) is an underlying basic tenet of our model. Since φ is compactly supported, so must its projections ψ_i's, implying that h_{ij} is eventually constant, i.e., there exists some $R > 0$ and some $g \in \mathrm{SO}(2)$ such that $h_{ij}(r) = g$ whenever $r \geq R$. In fact there is no loss of generality in assuming that $g = 1$: Since $\mathrm{SO}(2)$ is connected, we may pick a continuous curve $\gamma : [R, R'] \to \mathrm{SO}(2)$ such that $\gamma(R) = g$ and $\gamma(R') = 1$; replacing $h|_{[R,R']}$ by γ then gives an h where $h_{ij}(r) = 1$ for sufficiently large r. In particular, $\lim_{r\to\infty} h_{ij}(r) = 1$, the identity element in $\mathrm{SO}(2)$.

Recall that we write $\underline{G}(U)$ for the set of G-valued functions on an open set U. So $h_{ij} \in \underline{\mathrm{SO}(2)}(\mathbb{R}^2)$. Let $\mathcal{U}$ be the open covering of K_ε in (13). We will now define a *continuous cryo-EM cocycle*, a Čech 1-cocycle

$$z_\varepsilon^c \coloneqq \{\tau_{ij} \in \underline{\mathrm{SO}(2)}(U_i \cap U_j) : [i,j] \in K_\varepsilon\}$$

on K_ε determined by the h_{ij}'s. The process is analogous to how we obtained z_ε^d, the discrete cryo-EM cocycle on K_ε, from the set of pairwise angular comparisons D in section 4 but is a little more involved.

We first define the restriction of τ_{ij} to $U_i \cap U_j \cap U_k$ for all $k = 1, \dots, n$ and show that we can glue them together to obtain a globally defined SO(2)-valued function on $U_i \cap U_j$. By construction, the open covering $\mathcal{U}$ has the property that for any U_i, U_j, U_k, either

$$U_i \cap U_j \cap U_k = \varnothing \qquad \text{or} \qquad U_i \cap U_j \cap U_k \cong \mathbb{R}^2.$$

In the first case there is nothing to define. If $U_i \cap U_j \cap U_k \cong \mathbb{R}^2$, we fix a homeomorphism and regard $U_i \cap U_j \cap U_k$ as $\mathbb{R}^2$, then define the restriction of τ_{ij} to be

$$\tau_{ij}(x,y) = h_{ij}(r),$$

for $(x,y) \in U_i \cap U_j \cap U_k$ and $h_{ij} \in \underline{\mathrm{SO}(2)}(U_i \cap U_j)$. Although the definition of $\tau_{ij}|_{U_i \cap U_j \cap U_k}$ depends on a homeomorphism $U_i \cap U_j \cap U_k \cong \mathbb{R}^2$, two such homeomorphisms induce a homeomorphism from $U_i \cap U_j \cap U_k$ to itself. So we obtain a one-to-one correspondence between the set of τ_{ij}'s constructed from one homeomorphism and the set of τ_{ij}'s constructed from the other. This in turn induces a one-to-one correspondence between cohomology classes represented by the two sets of τ_{ij}'s. So while different homeomorphisms $U_i \cap U_j \cap U_k \cong \mathbb{R}^2$ give different τ_{ij}'s, their cohomology classes are in one-to-one correspondence.

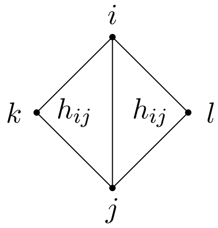
Since $U_i \cap U_j \cap U_k$ is disjoint from $U_i \cap U_j \cap U_l$ whenever k and l are distinct, to define τ_{ij} on $U_i \cap U_j$ we need only define it on the set

$$V_{ij} \coloneqq U_i \cap U_j - \bigcup\nolimits_{k \neq i,j} U_i \cap U_j \cap U_k.$$

If $V_{ij} \neq \varnothing$, then it must be the interior of the 1-simplex connecting $[i]$ and $[j]$. In this case we define τ_{ij} to be the constant $\lim_{r\to\infty} \tau_{ij}(x,y) = 1 \in \mathrm{SO}(2)$ where $(x,y) \in U_i \cap U_j \cap U_k$ and $r = \sqrt{x^2+y^2}$. Note that the limit exists as $\tau_{ij}(x,y)$ depends only on $r = \sqrt{x^2+y^2}$ and φ and ψ_i's are compactly supported. Lastly, it is obvious from its definition that τ_{ij} satisfies the 1-cocyle condition

$$\tau_{ij}(x,y)\tau_{jk}(x,y)\tau_{ki}(x,y) = 1. \tag{19}$$

To illustrate our construction of τ_{ij}, we consider an example where the two-dimensional simplicial complex K is obtained by gluing two triangles as follows:



Here $U_i \cap U_j \cap U_k$ is the interior of the triangle with vertices i, j, k. We define the values of τ_{ij}, τ_{ki}, and τ_{jk} on $U_i \cap U_j \cap U_k$ to be h_{ij}, h_{ki}, and h_{jk} respectively. One should think of $U_i \cap U_j \cap U_k$ as a copy of $\mathbb{R}^2$ and the boundary of the triangle with vertices i, j, k as the "points at infinity" of $\mathbb{R}^2$.

Since z^c_ε satisfies (19), we see that $z^c_\varepsilon \in \check{Z}^1\big(K_\varepsilon, \mathrm{SO}(2)\big)$. By an argument similar to the proof of Theorem 4.2, we obtain the following classification result.

Theorem 5.1 (bundle classification of continuous cryo-EM cocycles I). *Let $\varepsilon > 0$ be small enough so that (19) holds and let K_ε be the corresponding cryo-EM complex. Then*

(i) *the 1-cocycle z^c_ε determines an oriented circle bundle on K_ε;*

(ii) *two 1-cocycles z^c_ε and z'^c_ε for the same molecule determine isomorphic oriented circle bundles if and only if*

$$\tau'_{ij} = \tau_{ij}\tau_i\tau_j^{-1} \tag{20}$$

for some $\tau_i \in \underline{\mathrm{SO}(2)}(U_i), \tau_j \in \underline{\mathrm{SO}(2)}(U_j)$, $[i,j] \in K_\varepsilon$.

For small enough $\varepsilon > 0$, Theorem 5.1 gives us a classification of all possible continuous cryo-EM cocycles on K_ε, a canonical correspondence

$$\begin{aligned}\big\{\text{cohomologically equivalent cryo-EM cocycles on } K_\varepsilon\big\}& \\ \longrightarrow \big\{\text{isomorphism classes of oriented circle bundles on } K_\varepsilon\big\}&.\end{aligned} \tag{21}$$

By Proposition 3.2, the isomorphism classes of principal G-bundles may be identified with $[K_\varepsilon, BG]$, the homotopy classes of continuous maps from K_ε to the classifying space of G. In our case, $G = \mathrm{SO}(2) \simeq \mathbb{S}^1$, the circle group. By (7), $BG = B\,\mathrm{SO}(2) \simeq \mathbb{C}P^\infty$ and so

$$\check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)\big) \simeq [K_\varepsilon, B\,\mathrm{SO}(2)] \simeq [K_\varepsilon, \mathbb{C}P^\infty] \simeq H^2(K_\varepsilon), \tag{22}$$

where the last isomorphism is by Theorem 3.4. We will discuss the two main implications of (22) separately: $H^2(K_\varepsilon)$ gives us a homological classification of continuous cryo-EM cocycles; whereas $[K_\varepsilon, B\,\mathrm{SO}(2)]$ tells us about the *moduli* space of continuous cryo-EM cocycles.

5.1. Cohomology as obstruction. The cohomology group $H^2(K_\varepsilon)$ may be viewed as the obstruction to K_ε degenerating into a one-dimensional simplicial complex. If $H^2(K_\varepsilon) = 0$, then K_ε contains no 2-sphere—by which we mean the boundary of a 3-simplex, which is homeomorphic to $\mathbb{S}^2$. Thus K_ε is a two-dimensional simplicial complex whose 2-simplices are all contractible, and thus it is homotopic to a one-dimensional simplicial complex. Let $H^2(K_\varepsilon) = 0$. If ψ_j, ψ_k, ψ_l are three images that lie in the ε-neighborhood of an image ψ_i, then at least one of ψ_j, ψ_k, ψ_l cannot lie in the intersection of ε-neighborhoods of the other two. In terms of the graph G_ε, $H^2(K_\varepsilon) = 0$ implies that G_ε does not contain a 4-clique, i.e., a complete subgraph with four vertices.

The isomorphism wtih $H^2(K_\varepsilon)$ also allows us to calculate $\check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)\big)$ explicitly.

Theorem 5.2. *$\check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)\big) \simeq H^2(K_\varepsilon) = \mathbb{Z}^b$ where $b = b_2(K_\varepsilon)$, the second Betti number of K_ε.*

Proof. The isomorphism is (22). The equality follows from Theorem 3.3, observing that $H_1(K_\varepsilon) = 0$ by our construction of K_ε and so $T_1 = 0$. We may also derive the isomorphism directly without going through the chain of isomorphisms in (22). Snake lemma [23, 32, 49] applied to the exact sequence of groups

$$1 \to \mathbb{Z} \xrightarrow{2\pi} \mathbb{R} \xrightarrow{\mathrm{expi}} \mathbb{S}^1 \to 1,$$

where the first map is multiplication by 2π and $\mathrm{expi}(x) := \exp(ix)$, yields a long exact sequence of cohomology groups

$$\cdots \to \check{H}^1(K_\varepsilon, \mathbb{R}) \to \check{H}^1(K_\varepsilon, \mathbb{S}^1) \to \check{H}^2(K_\varepsilon, \mathbb{Z}) \to \check{H}^2(K_\varepsilon, \mathbb{R}) \to \cdots .$$

Both $\check{H}^1(K_\varepsilon, \mathbb{R})$ and $\check{H}^2(K_\varepsilon, \mathbb{R})$ are zero by the existence of partition of unity on K_ε. So $\check{H}^1(K_\varepsilon, \mathbb{S}^1) = \check{H}^2(K_\varepsilon, \mathbb{Z})$. Since $\mathbb{S}^1 = \mathrm{SO}(2)$, $\check{H}^1(K_\varepsilon, \mathbb{S}^1) = \check{H}^1\big(K_\varepsilon, \mathrm{SO}(2)\big)$. Finally, by Theorem 3.6, we get $\check{H}^2(K_\varepsilon, \mathbb{Z}) \simeq H^2(K_\varepsilon, \mathbb{Z}) = H^2(K_\varepsilon)$. ■

5.2. Cohomology as moduli. A benefit of classifying continuous cryo-EM cocycles in terms of oriented circle bundles is that these are very well understood classical objects [11, 50]. In what follows, we will refine Theorem 4.2 with explicit descriptions of the oriented circle bundles that arise in the classification of continuous cryo-EM cocycles.

Let $b_2(K_\varepsilon) = b$. Since K_ε is a finite two-dimensional simplicial complex, $b_2(K_\varepsilon) = b$ implies that K_ε contains b copies of 2-spheres. By (21) and Theorem 5.2, we expect to obtain an oriented circle bundle over K_ε for each $(m_1, \ldots, m_b) \in \mathbb{Z}^b$. An oriented circle bundle over any one-dimensional simplicial complex K must be trivial since $H^2(K) = 0$. Hence any oriented circle bundle over K_ε is uniquely determined by its restriction to the 2-spheres contained in K_ε and understanding oriented circle bundles on K_ε reduces to understanding oriented circle bundles on $\mathbb{S}^2$, which we will describe explicitly in the following.

We start by identifying the 3-sphere with the group of unit quaternions, i.e.,

$$\mathbb{S}^3 = \{a + bi + cj + dk \in \mathbb{H} : a, b, c, d \in \mathbb{R},\ a^2 + b^2 + c^2 + d^2 = 1\},$$

and identify the circle with the group of unit complex numbers, i.e.,

$$\mathbb{S}^1 = \{a + bi \in \mathbb{C} : a, b \in \mathbb{R},\ a^2 + b^2 = 1\}.$$

Elements of $\mathbb{S}^1$ may be regarded as unit quaternions with $c = d = 0$ and so $\mathbb{S}^1$ a subgroup of $\mathbb{S}^3$. In particular, $\mathbb{S}^1$ acts on $\mathbb{S}^3$ by quaternion multiplication and we have a group action

$$(23)\quad \varphi : \mathbb{S}^1 \times \mathbb{S}^3 \to \mathbb{S}^3, \quad (x+yi, a+bi+cj+dk) \mapsto xa-yb+(xb+ya)i+(xc-yd)j+(xd+yc)k.$$

As topological spaces we have

$$\mathbb{S}^3/\mathbb{S}^1 \simeq \mathbb{S}^2$$

but note that $\mathbb{S}^1$ is not a normal subgroup of $\mathbb{S}^3$ and so $\mathbb{S}^2$ does not inherit a group structure. Let

$$\pi : \mathbb{S}^3 \to \mathbb{S}^3/\mathbb{S}^1 \simeq \mathbb{S}^2$$

be the natural quotient map.

For $m \in \mathbb{N}$, let C_m be the subgroup of $\mathbb{S}^1$ generated by $\exp(2\pi i/m)$, a cyclic group of order m. Each C_m is also a subgroup of $\mathbb{S}^3$ and acts on $\mathbb{S}^3$ by quaternion multiplication. Since C_m is a subgroup of $\mathbb{S}^1$, we obtain an induced projection map

$$(24)\qquad \pi_m : \mathbb{S}^3/C_m \to \mathbb{S}^3/\mathbb{S}^1 \simeq \mathbb{S}^2$$

for each $m \in \mathbb{N}$. The following classic result [50] describes all *circle bundles* on $\mathbb{S}^2$—there are infinitely many of them, one for each nonnegative integer.

Proposition 5.3. *For each $m = 0, 1, 2, \dots,$ there is a circle bundle (A_m, π_m, φ_m) with base space $\mathbb{S}^2$ where*

$$A_0 = \mathbb{S}^1 \times \mathbb{S}^2, \qquad A_m = \mathbb{S}^3/C_m \quad \textit{for } m \in \mathbb{N}.$$

The projection to $\mathbb{S}^2$,

$$\pi_0 : A_0 \to \mathbb{S}^2, \qquad \pi_m : A_m \to \mathbb{S}^3/\mathbb{S}^1 \simeq \mathbb{S}^2,$$

is the projection onto the second factor for $m = 0$ and the quotient map (24) for $m \in \mathbb{N}$. The group action $\varphi_m : \mathbb{S}^1 \times A_m \to A_m$ is the trivial action (any element in $\mathbb{S}^1$ acts as identity on A_0) for $m = 0$ and the action induced by quaternion multiplication φ in (23) for $m \in \mathbb{N}$. Every circle bundle on $\mathbb{S}^2$ is isomorphic to an A_m for some $m = 0, 1, 2, \dots$.

Note that these are SO(2)-bundles since we regard $\mathrm{SO}(2) = \mathbb{S}^1$. A_0 is the trivial circle bundle on $\mathbb{S}^2$ and A_1 is the well-known *Hopf fibration*. As a manifold, $A_m = \mathbb{S}^3/C_m$ is orientable for all $m \in \mathbb{N}$ and so each A_m comes in two different orientations, which we denote by A_m^+ and A_m^-. For $m = 0, 1, 2, \dots,$ we write

$$B_0 \coloneqq A_0, \qquad B_m \coloneqq A_m^+, \qquad B_{-m} \coloneqq A_m^-.$$

These are the *oriented circle bundles* on $\mathbb{S}^2$.

We will next construct a *cryo-EM bundle* by gluing oriented circle bundles along the cryo-EM complex K_ε, attaching a copy of B_m for some $m \in \mathbb{Z}$ to each 2-sphere in K_ε. We then show that these bundles are in one-to-one correspondence with continuous cryo-EM cocycles on K_ε.

Let K_ε be a cryo-EM complex with $b_2(K_\varepsilon) = b$, i.e., K_ε contains b copies of 2-spheres; in fact, by its definition, K_ε is homotopic to the one-point union of b copies of $\mathbb{S}^2$, as we discussed in the proof of Theorem 4.3. Label these arbitrarily from $i = 1, \ldots, b$ and denote them $\mathbb{S}^2_1, \ldots, \mathbb{S}^2_b$. For any $(m_1, \ldots, m_b) \in \mathbb{Z}^b$, we may define a principal SO(2)-bundle $B_{m_1,\ldots,m_b}$ on K_ε as one whose restriction on the ith 2-sphere in K_ε is B_{m_i}, $i = 1, \ldots, b$, and is trivial elsewhere. We remove all the 2-spheres contained in K_ε and let the remaining simplicial complex be

$$L_\varepsilon \coloneqq \left(\overline{K_\varepsilon - \bigcup_{i=1}^{b} \mathbb{S}^2_i}\right).$$

As a topological space, $B_{m_1,\ldots,m_b}$ is the union of B_{m_i}'s corresponding to each of the 2-spheres and the trivial circle bundle on L_ε,

$$B_{m_1,\ldots,m_b} \coloneqq \left[\bigcup_{i=1}^{b} B_{m_i}\right] \cup \left[L_\varepsilon \times \mathbb{S}^1\right].$$

To see that $B_{m_1,\ldots,m_b}$ is a fiber bundle on K_ε, take the open covering

$$\mathcal{U} = \{U_1(K_\varepsilon), \ldots, U_n(K_\varepsilon)\}$$

of K_ε in section 4. By the construction of $B_{m_1,\ldots,m_b}$, its restriction to $U_i(K_\varepsilon)$ is a trivial fiber bundle since $U_i(K_\varepsilon)$ is contractible. So $B_{m_1,\ldots,m_b}$ is locally trivial and thus a fiber bundle on K_ε. Moreover, the bundle $(B_{m_1,\ldots,m_b}, \pi, \varphi)$ is an oriented circle bundle on K_ε with π and φ defined as follows. The projection map $\pi : B_{m_1,\ldots,m_b} \to K_\varepsilon$ is defined by

$$\pi(f) = \begin{cases} \pi_{m_i}(f) & \text{if } f \in B_{m_i}, \quad i = 1, \ldots, b, \\ \mathrm{pr}_1(f) & \text{if } f \in L_\varepsilon \times \mathbb{S}^1. \end{cases}$$

Here $\mathrm{pr}_1 : L_\varepsilon \times \mathbb{S}^1 \to L_\varepsilon$ is the projection onto the first factor. The group action $\varphi : \mathrm{SO}(2) \times B_{m_1,\ldots,m_b} \to B_{m_1,\ldots,m_b}$ is defined by

$$\varphi(g, f) = \begin{cases} \varphi_{m_i}(g, f) & \text{if } f \in B_{m_i}, \quad i = 1, \ldots, b, \\ f & \text{if } f \in L_\varepsilon \times \mathbb{S}^1, \end{cases}$$

for any $g \in G$ and $f \in B_{m_1,\ldots,m_b}$. Furthermore, the intersection of any two simplices in K_ε is by our construction either empty or a contractible space and so any bundle is trivial on the intersection.

Since every oriented circle bundle on K_ε is isomorphic to $B_{m_1,\ldots,m_b}$ for some $(m_1, \ldots, m_b) \in \mathbb{Z}^b$, we have the following classification theorem for continuous cryo-EM cocycles in terms of $B_{m_1,\ldots,m_b}$.

Theorem 5.4 (bundle classification of continuous cryo-EM cocycles II). *Let $\varepsilon > 0$ be small enough so that (19) holds and let K_ε be the corresponding cryo-EM complex. Let $b = b_2(K_\varepsilon)$. Then each cohomologically equivalent continuous cryo-EM cocycles z^c_ε on K_ε corresponds to an isomorphism class of an oriented circle bundle $B_{m_1,\ldots,m_b}$ on K_ε for $(m_1, \ldots, m_b) \in \mathbb{Z}^b$.*

Proof. Let $z^c_\varepsilon = \{g_{ij} \in \underline{\mathrm{SO}(2)}(U_i \cap U_j) : [i, j] \in K_\varepsilon\}$ and let $z'^c_\varepsilon = \{g'_{ij} \in \underline{\mathrm{SO}(2)}(U_i \cap U_j) : [i, j] \in K_\varepsilon\}$ be cohomologically equivalent continuous cryo-EM cocycles on K_ε, i.e., they are related by (20) for some $g_i, g_j \in \mathrm{SO}(2)$, $[i, j] \in K_\varepsilon$. By Theorem 5.1, z^c_ε and z'^c_ε must correspond to the same oriented circle bundle on K_ε. ■

6. Denoising cryo-EM images and cohomology.

Aside from providing theoretical classification results (e.g. Theorems 4.2 and 5.4) whose practical value is as yet unclear, we show here that the more elementary aspects of our cohomology framework can shed light on one aspect of cryo-EM imaging—denoising cryo-EM images. Our goal is not to propose any new method but to provide some perspectives on existing methods, which work well in practice [47, 48, 44]. We saw in section 4 that a *noiseless* discrete cryo-EM cocycle $z^d_\varepsilon = \{g_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ on K_ε satisfies the cocycle condition

$$g_{ij}g_{jk}g_{ki} = 1, \tag{25}$$

when ε is sufficiently small. In reality, a collection of projected images obtained from cryo-EM measurements, $\widehat{\Psi} = \{\widehat{\psi}_1, \dots, \widehat{\psi}_n\}$, will invariably be corrupted by noise; here a carat over a quantity signifies that it is possibly corrupted by noise. As a result, the discrete cryo-EM cocycle $\widehat{z}_\varepsilon = \{\widehat{g}_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ obtained from $\widehat{\Psi}$ will not satisfy (25) for sufficiently small $\varepsilon > 0$. To see this, let $\Psi = \{\psi_1, \dots, \psi_n\}$ be a set of noise-free projected images whose discrete cryo-EM cocycle is z^d_ε. If $\widehat{z}_\varepsilon$ also satisfies (25), then we have

$$\widehat{g}_{ij} = g_{ij}g_i g_j^{-1} \in \mathrm{SO}(2)$$

for some $g_i \in \mathrm{SO}(2)$. But this implies that $\widehat{\psi}_i$ can be obtained by rotating the noise-free image ψ_i by $g_i \in \mathrm{SO}(2)$ and so $\widehat{\psi}_i$ is also noise-free, a contradiction.

In general cryo-EM images are denoised by *class averaging* [18]. Noisy images are grouped into classes of similar viewing directions. The within-class average is then taken as an approximation of the noise-free image in that direction. The methods for grouping images into classes [47, 48, 44] are essentially all based on the observation that in the noiseless scenario, the cocycle condition (25) must hold. We will look at a few measures of deviation of discrete cryo-EM cocycles from being a cocycle.

Let $\widehat{z}^d_\varepsilon = \{\widehat{g}_{ij} \in \mathrm{SO}(2) : [i,j] \in K_\varepsilon\}$ be a discrete cryo-EM cocycle on K_ε computed from a noisy set of projected images $\widehat{\Psi}$. Since $\mathrm{SO}(2)$ can be identified with the circle $\mathbb{S}^1$, every $g \in \mathrm{SO}(2)$ corresponds to an angle $\theta \in \mathbb{S}^1$, represented by $\theta \in [0, 2\pi)$. A straightforward measure of deviation of z_ε from being a cocycle is given by

$$\delta(\widehat{z}^d_\varepsilon) = \sum\nolimits_{i,j,k\,:\,[i,j],[i,k],[j,k]\in K_\varepsilon} (\theta_{ij} + \theta_{jk} + \theta_{ki})^2,$$

where the addition in the parentheses is computed in $\mathbb{S}^1$, i.e., given by the unique number $\theta_{ijk} \in [0, 2\pi)$ such that

$$\theta_{ij} + \theta_{jk} + \theta_{ki} = \theta_{ijk} \pmod{2\pi}.$$

Lemma 6.1. *$\widehat{z}^d_\varepsilon$ is a cocycle if and only if $\delta(\widehat{z}^d_\varepsilon) = 0$.*

Let ψ be an arbitrary projected image. Then $\delta(\widehat{z}_\varepsilon)$ quantifies the obstruction of gluing images in

$$\widehat{z}^d_\varepsilon(\psi) = \{g \cdot \psi : g \in \widehat{z}^d_\varepsilon\} = \{\widehat{g}_{ij} \cdot \psi : [i,j] \in K_\varepsilon\}$$

together to get the 3D structure of the molecule. If $\delta(\widehat{z}_\varepsilon)$ is small, then $\widehat{z}^d_\varepsilon$ is already close enough to a cocycle and hence every image is good.

On the other hand, if $\delta(\widehat{z}_\varepsilon^d)$ is big, then the following measure allows us to identify subsets of good images, if any. Given an image ψ, whenever $[i,j]$ is an edge of K_ε for some j, we want the viewing direction of $g_{ij} \cdot \psi$ to be close to that of ψ. This is captured by the quantity $\rho_i(\widehat{z}_\varepsilon^d) := \delta_i(\widehat{z}_\varepsilon^d)/3\delta(\widehat{z}_\varepsilon^d)$ where

$$\delta_i(\widehat{z}_\varepsilon^d) = \sum\nolimits_{j,k\,:\,[i,j],[i,k],[j,k]\in K_\varepsilon} (\theta_{ij} + \theta_{jk} + \theta_{ki})^2, \qquad i = 1,\dots,n.$$

Clearly, $\sum_{i=1}^n \rho_i(\widehat{z}_\varepsilon^d) = 1$. For $g_{ij} \cdot \psi$ to be a good image, we want $\rho_i(\widehat{z}_\varepsilon^d) \ll 1$.

Acknowledgments. We thank Amit Singer for introducing us to this fascinating topic. We are extremely grateful to the two anonymous referees for their unusually thorough reading and for making numerous useful comments that led to a vast improvement of our manuscript.

REFERENCES

[1] G. Ambartsoumian and P. Kuchment, *A range description for the planar circular Radon transform*, SIAM J. Math. Anal., 38 (2006), pp. 681–692, https://doi.org/10.1137/050637492.

[2] G. Ambartsoumian and P. Kuchment, *On the injectivity of the circular Radon transform arising in thermoacoustic tomography*, Inverse Problems, 21 (2005), pp. 473–485, https://doi.org/10.1088/0266-5611/21/2/004.

[3] S. R. Arridge, *Optical tomography in medical imaging*, Inverse Problems, 15 (1999), pp. R41–R93, https://doi.org/10.1088/0266-5611/15/2/022.

[4] S. R. Arridge and J. C. Hebbden, *Optical imaging in medicine:* II. *Modelling and reconstruction*, Phys. Med. Biol., 42 (1997), pp. 841–853, https://doi.org/10.1088/0031-9155/42/5/008.

[5] C. Baillard and H. Maitre, 3-*D reconstruction of urban scenes from aerial stereo imagery: A focusing strategy*, Comput. Vis. Image. Und., 76 (1999), pp. 244–258, https://doi.org/10.1006/cviu.1999.0793.

[6] A. S. Bandeira, Y. Chen, and A. Singer, *Non-unique games over compact groups and orientation estimation in cryo-EM*, preprint, 2015, http://arxiv.org/abs/1505.03840.

[7] A. S. Bandeira, C. Kennedy, and A. Singer, *Approximating the little Grothendieck problem over the orthogonal and unitary groups*, Math. Program., 160 (2016), pp. 433–475, https://doi.org/10.1007/s10107-016-0993-7.

[8] C. Berge, *Hypergraphs*, North-Holland Math. Library 45, North-Holland, Amsterdam, 1989, https://doi.org/10.1016/S0924-6509(08)70093-X.

[9] J.-L. Brylinski, *Loop Spaces, Characteristic Classes and Geometric Quantization*, Birkhäuser Boston, Boston, 2008, htttps://doi.org/10.1007/978-0-8176-4731-5.

[10] U. Castellani, A. Fusiello, and V. Murino, *Registration of multiple acoustic range views for underwater scene reconstruction*, Comput. Vis. Image. Und., 87 (2002), pp. 78–89, https://doi.org/10.1006/cviu.2002.0984.

[11] S. S. Chern, *Circle bundles*, Geometry and Topology, Lecture Notes in Math. 597, Springer, Berlin, 1977, pp. 114–131, https://doi.org/10.1007/BFb0085351.

[12] V. De Silva and R. Ghrist, *Homological sensor networks*, Notices Amer. Math. Soc., 54 (2007), pp. 10–17, http://www.ams.org/notices/200701/fea-ghrist.pdf.

[13] D. Finch, S. K. Patch, and Rakesh, *Determining a function from its mean values over a family of spheres*, SIAM J. Math. Anal., 35 (2004), pp. 1213–1240, https://doi.org/10.1137/S0036141002417814.

[14] A. Fischer, T. H. Kolbe, F. Lang, A. B. Cremers, W. Förstner, L. Plümer, and V. Steinhage, *Extracting buildings from aerial images using hierarchical aggregation in* 2*D and* 3*D*, Comput. Vis. Image. Und., 72 (1998), pp. 185–203, https://doi.org/10.1006/cviu.1998.0721.

[15] O. Forster, *Lectures on Riemann surfaces*, Grad. Texts Math. 81, Springer-Verlag, New York, 1991, https://doi.org/10.1007/978-1-4612-5961-9.

[16] G. K. Francis, *A Topological Picturebook*, Springer, New York, 2007, https://doi.org/10.1007/978-0-387-68120-7.

[17] J. Frank, *Single-particle imaging of macromolecules by cryo-electron microscopy*, Annu. Rev. Biophys. Biomol. Struct., 31 (2002), pp. 303–319, https://doi.org/10.1146/annurev.biophys.31.082901.134202.
[18] J. Frank, *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in their Native State*, Oxford University Press, New York, 2006, https://doi.org/10.1093/acprof:oso/9780195182187.001.0001.
[19] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley, New York, 1994, https://doi.org/10.1002/9781118032527.
[20] R. Hadani and A. Singer, *Representation theoretic patterns in three-dimensional cryo-electron microscopy* I: *The intrinsic reconstitution algorithm*, Ann. of Math. (2), 174 (2011), pp. 1219–1241, https://doi.org/10.4007/annals.2011.174.2.11.
[21] R. Hadani and A. Singer, *Representation theoretic patterns in three-dimensional cryo-electron microscopy* II: *The class averaging problem*, Found. Comput. Math., 11 (2011), pp. 589–616, https://doi.org/10.1007/s10208-011-9095-3.
[22] R. Hartshorne, *Algebraic Geometry*, Grad. Texts Math. 52, Springer, New York, 1977, https://doi.org/10.1007/978-1-4757-3849-0.
[23] A. Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, UK, 2002, https://www.math.cornell.edu/~hatcher/AT/AT.pdf.
[24] T. W. Hungerford, *Algebra*, Grad. Texts Math. 73, Springer, New York, 1980, https://doi.org/10.1007/978-1-4612-6101-8.
[25] D. Husemöller, *Fibre Bundles*, 3rd ed., Grad. Texts Math. 20, Springer, New York, 1994, https://doi.org/10.1007/978-1-4757-2261-1.
[26] D. Husemöller, M. Joachim, B. Jurčo, and M. Schottenloher, *Basic Bundle Theory and K-Cohomology Invariants*, Lecture Notes in Phys. 726, Springer, Berlin, 2008, https://doi.org/10.1007/978-3-540-74956-1.
[27] B. Iversen, *Cohomology of Sheaves*, Universitext, Springer-Verlag, Berlin, 1986, https://doi.org/10.1007/978-3-642-82783-9.
[28] F. John, *Plane Waves and Spherical Means*, Springer-Verlag, New York, 1981, https://doi.org/10.1007/978-1-4613-9453-2.
[29] F. Kamber and Ph. Tondeur, *Flat bundles and characteristic classes of group-representations*, Amer. J. Math., 89 (1967), pp. 857–886, https://doi.org/10.2307/2373408.
[30] R. A. Kruger, P. Liu, Y. R. Fang, and C. R. Appledorn, *Photoacoustic ultrasound (*paus*) reconstruction tomography*, Med. Phys., 22 (1995), pp. 1605–1609, https://doi.org/10.1118/1.597429.
[31] L.-H. Lim, *Hodge Laplacians on Graphs*, *Geometry and Topology in Statistical Inference*, S. Mukherjee, ed., Proceedings of Symposia in Applied Math. 74, AMS, Providence, RI, 2017.
[32] J. P. May, *A Concise Course in Algebraic Topology*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, 1999, https://www.math.uchicago.edu/~may/CONCISE/ConciseRevised.pdf.
[33] J. W. Milnor and J. D. Stasheff, *Characteristic Classes*, Princeton University Press, Princeton, NJ, 1974.
[34] S. Morita, *Geometry of Characteristic Classes*, Transl. Math. Monogr. 199, AMS, Providence, RI, 2001.
[35] F. Natterer, *The Mathematics of Computerized Tomography*, Classics Appl. Math. 32, SIAM, Philadelphia, 2001, https://doi.org/10.1137/1.9780898719284.
[36] S. Negahdaripour, H. Sekkati, and H. Pirsiavash, *Opti-acoustic stereo imaging: On system calibration and* 3*-D target reconstruction*, IEEE Trans. Image Process., 18 (2009), pp. 1203–1214, https://doi.org/10.1109/TIP.2009.2013081.
[37] J. Oprea and D. Tanré, *Flat circle bundles, pullbacks, and the circle made discrete*, Int. J. Math. Math. Sci., 2005 (2005), pp. 3487–3495, https://doi.org/10.1155/IJMMS.2005.3487.
[38] P. A. Penczek, J. Zhu, and J. Frank, *A common-lines based method for determining orientations for* $N > 3$ *particle projections simultaneously*, Ultramicroscopy, 63 (1996), pp. 205–218, https://doi.org/10.1016/0304-3991(96)00037-X.
[39] P. Penczek, M. Radermacher, and J. Frank, *Three-dimensional reconstruction of single particles embedded in ice*, Ultramicroscopy, 40 (1992), pp. 33–53, https://doi.org/10.1016/0304-3991(92)90233-A.
[40] R. Penrose, *On the cohomology of impossible figures*, Structural Topology, 17 (1991), pp. 11–16, https://doi.org/10.2307/1575844.

[41] V. V. Prasolov, *Elements of Homology Theory*, Grad. Stud. Math., 81, AMS, Providence, RI, 2007.
[42] M. Radermacher, *Three-dimensional reconstruction from random projections: Orientation alignment via Radon transforms*, Ultramicroscopy, 53 (1994), pp. 121–136, https://doi.org/10.1002/jemt.1060090405.
[43] M. Schweiger and S. R. Arridge, *Optical tomographic reconstruction in a complex head model using a priori region boundary information*, Phys. Med. Biol., 44 (1999), pp. 2703–2721, https://doi.org/10.1088/0031-9155/44/11/302.
[44] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 20–36, https://doi.org/10.1016/j.acha.2010.02.001.
[45] A. Singer and Y. Shkolnisky, *Three-dimensional structure determination from common lines in cryo-EM by eigenvectors and semidefinite programming*, SIAM J. Imaging Sci., 4 (2011), pp. 543–572, https://doi.org/10.1137/090767777.
[46] A. Singer and H.-T. Wu, *Two-dimensional tomography from noisy projections taken at unknown random directions*, SIAM J. Imaging Sci., 6 (2013), pp. 136–175, https://doi.org/10.1137/090764657.
[47] A. Singer and H.-T. Wu, *Vector diffusion maps and the connection Laplacian*, Comm. Pure Appl. Math., 65 (2012), pp. 1067–1144, https://doi.org/10.1002/cpa.21395.
[48] A. Singer, Z. Zhao, Y. Shkolnisky, and R. Hadani, *Viewing angle classification of cryo-electron microscopy images using eigenvector*, SIAM J. Imaging Sci., 4 (2011), pp. 723–759, https://doi.org/10.1137/090778390.
[49] E. H. Spanier, *Algebraic Topology*, Springer, New York, 1981, https://doi.org/10.1007/978-1-4684-9322-1.
[50] N. E. Steenrod, *The classification of sphere bundles*, Ann. of Math. (2), 45 (1944), pp. 294–311, https://doi.org/10.2307/1969267.
[51] B. Vainshtein and A. Goncharov, *Determination of the spatial orientation of arbitrarily arranged identical particles of an unknown structure from their projections*, Sov. Phys. Dok., 31 (1986), pp. 278–280.
[52] M. Van Heel, *Angular reconstitution: A posteriori assignment of projection directions for 3D reconstruction*, Ultramicroscopy, 21 (1987), pp. 111–123, https://doi.org/10.1016/0304-3991(87)90078-7.
[53] L. Vietoris, *Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen*, Math. Ann., 97 (1927), pp. 454–472, https://eudml.org/doc/182647.
[54] M. Xu and L.-H. V. Wang, *Time-domain reconstruction for thermoacoustic tomography in a spherical geometry*, IEEE Trans. Med. Imag., 21 (2002), pp. 814–822, https://doi.org/10.1109/TMI.2002.801176.