REPRESENTATION THEORETIC PATTERNS IN THREE DIMENSIONAL CRYO-ELECTRON MICROSCOPY II - THE CLASS AVERAGING PROBLEM

RONNY HADANI AND AMIT SINGER

Abstract. In this paper, we continue to develop the representation theoretic setup for 3D cryo-electron microscopy (cryo-EM) that was initiated in [7]. In particular, we provide a complete spectral analysis of the local parallel transport operator on the two dimensional sphere. This is then used to prove the admissibility (correctness) and the numerical stability of the intrinsic classification algorithm for identifying raw projection images of similar viewing directions in cryo-EM, that was recently introduced in [9]. This preliminary classification is of fundamental importance in determining the three dimensional structure of macromolecules from cryo-EM images.

0. Introduction

The goal in cryo-EM is to determine the 3D structure of a molecule from noisy projection images taken at unknown random orientations by an electron microscope, i.e., a random Computational Tomography (CT). Determining 3D structures of large biological molecules remains vitally important, as witnessed, for example, by the 2003 Chemistry Nobel Prize, co-awarded to R. MacKinnon for resolving the 3D structure of the Shaker K+ channel protein [1, 4], and by the 2009 Chemistry Nobel Prize, awarded to V. Ramakrishnan, T. Steitz and A. Yonath for studies of the structure and function of the ribosome. The standard procedure for structure determination of large molecules is X-ray crystallography. The challenge in this method is often more in the crystallization itself than in the interpretation of the X-ray results, since many large molecules, including various types of proteins have so far withstood all attempts to crystallize them.

Cryo-EM is an alternative approach to X-ray crystallography. In this approach, sample of identical molecules are rapidly immobilized in a thin layer of vitreous ice (this is an ice without crystals). The cryo-EM imaging process produces a large collection of tomographic projections, corresponding to many copies of the same molecule, each immobilized in a different and unknown orientation. The intensity of the pixels in a given projection image is proportional to the line integrals of the electric potential induced by the molecule along the path of the imaging electrons (see Figure 0.1). The goal is to reconstruct the three-dimensional structure of the molecule from such a collection of projection images. The main problem is that the highly intense electron beam damages the molecule and, therefore, it is problematic to take projection images of the same molecule at known different directions as in the case of classical CT[1]. In other words, a single molecule is imaged only once,

Date: October 01, 2009.

[1]We remark that there are other methods like single-or multi-axis tilt EM tomogrophy, where several lower dose/higher noise images of a single molecule are taken from known directions. These

rendering an extremely low signal-to-noise ratio (SNR), mostly due to shot noise induced by the maximal allowed electron dose.

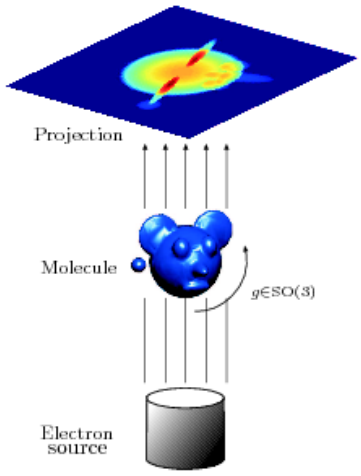

Figure 0.1. Schematic drawing of the imaging process: Every projection image corresponds to some unknown spatial orientation of the molecule.

0.1. **Mathematical model.** When modeling the experimental set-up of cryo-EM, it is often more convenient to think of the molecule as fixed and the electron microscope as changing orientations, thus we think of the orientation as associated with the microscope instead of with the molecule.

In more details: the molecule is modeled by a real valued function ϕ on

$$V = \mathbb{R}^3,$$

which describes the electric potential due to the charge density in the molecule.

A spatial orientation of the microscope is modeled by an oriented orthonormal basis $F = (f_1, f_2, f_3)$ in V, also called a *frame*. Equivalently, a frame F can be thought of as an element of the special rotation group $SO(3)$ whose first,second and third columns are f_1, f_2 and f_3 respectively. The vector f_3 is called the *viewing direction* and is denoted by $\pi(F)$.

Finally, the corresponding projection image is a real valued function I on the coordinate plane of the camera which is a copy of $\mathbb{R}^2$. The function I is given by the X-ray transform of ϕ along the viewing direction $\pi(F)$, that is

$$I_F(x,y) = \text{X-ray}_{\pi(F)}(\phi) = \int_{-\infty}^{\infty} \phi(xf_1 + yf_2 + tf_3)\, dt,$$

for every $(x, y) \in \mathbb{R}^2$.

The data collected from the experiment is a set consisting of N projection images $\mathcal{P} = \{I_1, .., I_N\}$. If we assume that the molecule is generic, that is, ϕ admits no non-trivial symmetries, then each image $I_i \in \mathcal{P}$ is associated with a unique frame $F_i \in SO(3)$. Another empirical assumption that we make is that the frames F_i,

methods are used for example when one has an organic object in vitro or a collection of different objects in the sample. There is a rich litrature for this field starting in the work of Crowther, DeRosier and Klug in the early 1960s.

$i = 1, .., N$ are uniformly distributed in $SO(3)$, according to the normalized Haar measure on the group. The main problem of cryo-EM is to reconstruct the unknown frame $F_i \in SO(3)$ associated with each projection image $I_i \in \mathcal{P}$.

0.2. **Class averaging.** As projection images in cryo-EM have extremely low SNR[2] (see Figure 0.2), a crucial initial step in all reconstruction methods is "class averaging" [2]. Class averaging is the grouping of a large data set of noisy raw projection images into clusters, such that images within a single cluster have similar viewing directions. Averaging rotationally aligned noisy images within each cluster results in "class averages"; these are images that enjoy a higher SNR and are used in later cryo-EM procedures such as the angular reconstitution procedure, [12], that requires better quality images. Finding consistent class averages is challenging due to the high level of noise in the raw images.

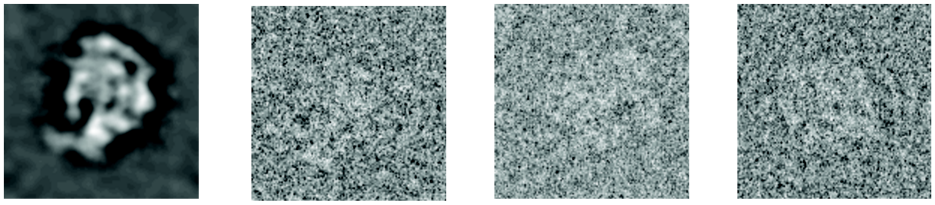

Figure 0.2. The left most image is a clean simulated projection image of the E.coli 50S ribosomal subunit. The other three images are real electron microscope images of the same subunit.

The starting point for the classification is the idea that visual similarity between projection images suggests vicinity between viewing directions of the corresponding (unknown) frames. The similarity between images I_i and I_j is measured by their invariant distance (introduced in [6]) which is the Euclidean distance between the images when they are optimally aligned with respect to in-plane rotations, namely

$$d(I_i, I_j) = \min_{g \in SO(2)} \|R(g) I_i - I_j\|, \tag{0.1}$$

where

$$R(g) I(x, y) = I\left(g^{-1}(x, y)\right),$$

for any function $I : \mathbb{R}^2 \to \mathbb{R}$.

One can choose some threshold value ϵ, such that $d(I_i, I_j) \leq \epsilon$ is indicative that perhaps the corresponding frames F_i and F_j have nearby viewing directions. The threshold ϵ defines an undirected graph $G = (\text{Vertices}, \text{Edges})$ with vertices labeled by numbers $1, .., N$ and an edge connecting vertex i with vertex j iff the invariant distance between the projection images I_i and I_j is smaller then ϵ, namely

$$\{i, j\} \in \text{Edges} \Longleftrightarrow d(I_i, I_j) \leq \epsilon.$$

In an ideal noiseless world, the graph G acquires the geometry of the unit sphere $S(V)$, namely, two images are connected by an edge iff their corresponding viewing directions are close on the sphere, that is, they belong to some small spherical cap of opening angle $a = a(\epsilon)$.

[2]SNR stands for Signal to Noise Ratio, which is the ratio between the squared L^2 norm of the signal and the squared L^2 norm of the noise.

However, the real world is far from ideal as it is governed by noise, hence, it often happens that two images of completely different viewing directions have small invariant distance. This can happen when the realizations of the noise in the two images match well for some random in-plane rotation, leading to spurious neighbor identification. Therefore, the naïve approach of averaging the rotationally aligned nearest neighbor images can sometimes yield a poor estimate of the true signal in the reference image.

To summarize: From this point of view, the main problem is to distinguish the good edges from the bad ones in the graph G, or, in other words, to distinguish the true neighbors from the false ones which are also called *outliers.* The existence of outliers is why the classification problem is non-trivial. We emphasize that without excluding the outliers, averaging rotationally aligned images of small invariant distance (0.1) yields poor estimate of the true signal, rendering the problem of 3D reconstruction from cryo-EM images non-feasible. In this respect, the class averaging problem is of fundamental importance.

0.3. **Main results.** In [9], we introduced a novel algorithm, referred to in this paper as the *intrinsic classification algorithm*, for classifying noisy projection images of similar viewing directions. The main appealing property of this new algorithm is its extreme robustness to noise and to presence of outliers; in addition, it also enjoys efficient time and space complexity. These properties are explained thoroughly in [9], which includes also large number of numerical experiments.

In this paper we prove the admissibility (correctness) and the numerical stability of the intrinsic classification algorithm, thus putting it on firm mathematical grounds. The proof relies on the study of a certain operator T_h, of geometric origin, called the *localized parallel transport operator*. Specifically:

- Admissibility depends on the fact that the maximal eigenspace of T_h is a three dimensional complex vector space and that there is a canonical identification of Hermitian vector spaces between this eigenspace and the complexified vector space $W = \mathbb{C}V$.
- Numerical stability depends on the existence of a spectral gap which separates the maximal eigenvalue of T_h from the rest of the spectrum, that enables to obtain a stable numerical approximation of the corresponding maximal eigenspace and of other related geometric structures.

In this regard, the main technical result of this paper is a complete description of the spectral properties of the localized parallel transport operator. In the course, we continue to develop the mathematical set-up for cryo-EM that was initiated in [7], thus further elucidating the central role played by representation theoretic principles in this scientific discipline.

The remainder of the introduction is devoted to a detailed description of the intrinsic classification algorithm and to explaining of the main ideas and results of this paper.

0.4. **Transport data.** The first step is to extract certain geometric data from the set of projection images. When computing the invariant distance between images I_i and I_j let us also record the rotation $g \in SO(2)$ that realizes the minimum in

(0.1) and denote this special rotation by $\widetilde{T}(i,j)$, that is

$$\widetilde{T}(i,j) = \underset{g \in SO(2)}{\operatorname{argmin}} \|R(g) I_i - I_j\|. \tag{0.2}$$

We note that

$$\widetilde{T}(j,i) = \widetilde{T}(i,j)^{-1}. \tag{0.3}$$

The main observation is, that in an ideal noiseless world the rotation $\widetilde{T}(i,j)$ admits geometric interpretation in terms of parallel transport on the sphere, provided the invariant distance between the corresponding images is small: define the rotation

$$T(F_i, F_j) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

as the unique solution of the equation

$$F_i \cdot T(F_i, F_j) = t_{\pi(F_i),\pi(F_j)} F_j, \tag{0.4}$$

where $t_{\pi(F_i),\pi(F_j)}$ is the parallel transport along the unique geodesic on the sphere connecting the points $\pi(F_j)$ with $\pi(F_i)$ or, in other words, it is the rotation in $SO(V)$ that takes the vector $\pi(F_j)$ to $\pi(F_i)$ along the shortest path on the sphere and

$$F \cdot \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = (\cos(\theta) f_1 + \sin(\theta) f_2, -\sin(\theta) f_1 + \cos(\theta) f_2, f_3),$$

for $F = (f_1, f_2, f_3)$.

The precise statement is that the rotation $\widetilde{T}(i,j)$ approximates the rotation $T(F_i, F_j)$ when $\{i,j\} \in$ Edges. This statement is partly suggested from mathematical considerations and partly from empirical considerations.

- On the mathematical side, we note that $T(F_i, F_j)$ can be characterized as the unique rotation of the frame F_i around its viewing direction $\pi(F_i)$, minimizing the distance to the frame F_j. This is a standard fact from differential geometry (a direct proof of this statement appears in [9]).
- On the empirical side, it is reasonable to assume that if the function ϕ is "nice", then the optimal alignment $\widetilde{T}(i,j)$ of the projection images is correlated with the optimal alignment $T(F_i, F_j)$ of the corresponding frames. This correlation of-course improves as the distance between $\pi(F_i)$ and $\pi(F_j)$ becomes smaller. A quantitative study of the relation between $\widetilde{T}(i,j)$ and $T(F_i, F_j)$ involves considerations from image processing thus is beyond the scope of this paper.

To conclude, the "empirical" rotation $\widetilde{T}(i,j)$ approximates the "theoretical" rotation $T(F_i, F_j)$ only when the viewing directions $\pi(F_i)$ and $\pi(F_j)$ are close, that is, they belong to some small spherical cap of opening angle a. In the other case, $\widetilde{T}(i,j)$ is not related any longer to parallel transportation on the sphere. For this reason, we consider only rotations $\widetilde{T}(i,j)$ for which $\{i,j\} \in$ Edges and call this collection (local) *empirical transport data*.

0.5. **The intrinsic classification algorithm.** The intrinsic classification algorithm reconstructs the Euclidian inner products $\{(\pi(F_i),\pi(F_j)) : i,j = 1,..N\}$ from the empirical transport data $\{\widetilde{T}(i,j) : \{i,j\} \in \text{Edges}\}$. We note that given these inner products, one can identify the true neighbors of each vertex in the graph G.

The main idea of the algorithm is to construct an intrinsic model, denoted by $\mathbb{W}_N$, of the Hermitian vector space $W = \mathbb{C}^3$ which is expressed solely in terms of the empirical transport data.

The algorithm proceeds in three steps:

Step1 (Ambient Hilbert space): consider the standard N-dimensional Hilbert space

$$\mathcal{H}_N = \mathbb{C}^N.$$

Step 2 (Self adjoint operator): identify $\mathbb{R}^2$ with $\mathbb{C}$ and consider each rotation $\widetilde{T}(i,j)$ as a complex number of unit norm. Define the $N \times N$ complex matrix

$$\widetilde{T}_N : \mathcal{H}_N \to \mathcal{H}_N,$$

by putting the rotation $\widetilde{T}(i,j)$ in the (i,j) entry. Notice that the matrix $\widetilde{T}_N$ is self-adjoint by (0.3).

Step 3 (Intrinsic model): the matrix $\widetilde{T}_N$ induces a spectral decomposition

$$\mathcal{H}_N = \bigoplus_{\lambda} \mathcal{H}_N(\lambda).$$

Theorem 1. *There exists a threshold λ_0 such that*

$$\dim \bigoplus_{\lambda > \lambda_0} \mathcal{H}_N(\lambda) = 3.$$

Define the Hermitian vector space

$$\mathbb{W}_N = \bigoplus_{\lambda > \lambda_0} \mathcal{H}_N(\lambda).$$

Let us denote by $\varphi_i \in \mathbb{W}_N$ the vector

$$\varphi_i = \sqrt{2/3} \cdot \mathrm{pr}_i^*(1),$$

where $\mathrm{pr}_i : \mathbb{W}_N \to \mathbb{C}$ denotes the projection on the ith component and $\mathrm{pr}_i^* : \mathbb{C} \to \mathbb{W}_N$ is the adjoint map.

For every $F \in X$, let us denote by $\delta_F \in W = \mathbb{C}^3$ the vector $f_1 - if_2$.

The upshot is, that the intrinsic vector space $\mathbb{W}_N$ consisting the collection of vectors $\varphi_i \in \mathbb{W}_N$, $i = 1,..,N$ is (approximately[3]) isomorphic to the extrinsic vector space W consisting the collection of vectors $\delta_{F_i} \in W$, $i = 1,..,N$. This is summarized in the following theorem:

Theorem 2. *There exists a unique (approximated) isomorphism $\tau_N : W \xrightarrow{\simeq} \mathbb{W}_N$ of Hermitian vector spaces such that*

$$\tau_N(\delta_{F_i}) = \varphi_i,$$

for every $i = 1,..,N$.

[3]This approximation improves as N grows.

The main application of the above theorem is that it enables to express, in intrinsic terms, the Euclidian products between the viewing directions. More precisely, we have following identity from linear algebra:

$$(\pi(F), \pi(E)) = |\langle \delta_F, \delta_E \rangle| - 1, \tag{0.5}$$

for every $F, E \in SO(3)$, where $(\cdot,\cdot)$ is the standard Euclidian product on $V = \mathbb{R}^3$ and $\langle \cdot,\cdot \rangle$ is the standard Hermitian product on W.

Granting the validity of Theorem 2 and based on the above identity, we obtain the following basic relation

$$(\pi(F_i), \pi(F_j)) = \left|\left\langle \varphi_i, \varphi_j \right\rangle\right| - 1, \tag{0.6}$$

for every $i, j = 1, .., N$.

0.6. **Structure of the paper.** The paper consists of three sections except of the introduction.

- In Section 1, we begin by introducing the basic analytic setup which is relevant for the class averaging problem in cryo-EM. Then, we proceed to formulate the main results of this paper, which are: A complete description of the spectral properties of the localized parallel transport operator (Theorem 3), the spectral gap property (Theorem 4) and the admissibility of the intrinsic classification algorithm (Theorem 5 and Theorem 6).
- In Section 2, we prove Theorem 3, in particular, we develop all the representation theoretic machinery that is needed for the proof.
- Finally, in Appendix A, we give the proofs of all technical statements which appeared in the previous sections.

Acknowledgement: *The first author would like to thank Joseph Bernstein for many helpful discussions concerning the mathematical aspects of this work. He also thanks Richard Askey for his valuable advises about Legendre polynomials. The second author is partially supported by Award Number R01GM090200 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. This work is part of a project conducted jointly with Shamgar Gurevich, Yoel Shkolnisky and Fred Sigworth.*

1. Preliminaries and main results

1.1. **Set up.** Let $(V, (\cdot,\cdot))$ be a three dimensional, oriented, Euclidian vector space over $\mathbb{R}$. The reader can take $V = \mathbb{R}^3$ equipped with the standard orientation and standard inner product. In addition, we will consider the complexified vector space $W = \mathbb{C}V$, equipped with the Hermitian product $\langle \cdot,\cdot \rangle : W \times W \to \mathbb{C}$, given by

$$\langle u + iv, u' + iv' \rangle = (u, v) - (v, v') - i(u, v') + i(v, u').$$

Let $SO(V)$ denote the group of orthogonal transformations with respect to the inner product $(\cdot,\cdot)$ which preserve the orientation. Let $S(V)$ denote the unit sphere in V, that is, $S(V) = \{v \in V : (v, v) = 1\}$. Let $X = Fr(V)^+$ denote the manifold of oriented orthonormal frames in V, that is, a point $F \in X$ is an orthonormal basis $F = (f_1, f_2, f_3)$ of V compatible with the orientation.

The frame manifold admits two commuting actions: A left action of the group $SO(V)$, given by

$$g \triangleright (f_1, f_2, f_3) = (gf_1, gf_2, gf_3)$$

and a right action of the special orthogonal group $SO(3)$, given by

$$\begin{aligned}(f_1, f_2, f_3) \triangleleft g \quad = \quad & (a_{11}f_1 + a_{21}f_2 + a_{31}f_3, \\ & a_{12}f_1 + a_{22}f_2 + a_{32}f_3, \\ & a_{13}f_1 + a_{23}f_2 + a_{33}f_3),\end{aligned}$$

for

$$g = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

We distinguish the copy of $SO(2)$ inside $SO(3)$ consisting of matrices of the form

$$g = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and consider X as a principal $SO(2)$ bundle over $S(V)$ where the fibration map $\pi : X \to S(V)$ is given by $\pi(f_1, f_2, f_3) = f_3$. We call the vector f_3 the *viewing direction.*

1.2. **The Transport data.** For every pair of frames $F, E \in X$ such that $\pi(F) \neq \pm\pi(E)$, let $T(F, E) : \mathbb{R}^2 \to \mathbb{R}^2$ be the element in $SO(2)$, characterized by the property

$$F \cdot T(F, E) = t_{\pi(F),\pi(E)}(E),$$

where $t_{\pi(F),\pi(E)} : X_{\pi(E)} \to X_{\pi(F)}$ is the morphism of fibers given by the parallel transport mapping along the unique geodesic in the sphere $S(V)$, connecting the points $\pi(E)$ with $\pi(F)$. We identify $\mathbb{R}^2$ with $\mathbb{C}$ and consider $T(F, E)$ as a complex number of unit norm.

We have the following properties:

- **Symmetry:** For every $F, E \in X$, we have $T(E, F)$ is equal $T(F, E)^{-1}$ which coincides with the complex conjugate $\overline{T(F, E)}$. This property follows from the fact that $t_{\pi(E),\pi(F)} = t^{-1}_{\pi(F),\pi(E)}$.
- **Invariance:** For every $F, E \in X$ and element $g \in SO(V)$, we have that $T(g \triangleright F, g \triangleright E) = T(F, E)$ .This property follows from the fact that $t_{\pi(g\triangleright F),\pi(g\triangleright E)} = g \circ t_{\pi(F),\pi(E)} \circ g^{-1}$, for every $g \in SO(V)$.
- **Equivariance:** For every $F, E \in X$ and elements $g_1, g_2 \in SO(2)$, we have that $T(F \triangleleft g_1, E \triangleleft g_2) = g_1^{-1} T(F, E)\, g_2$. This property follows from the fact that $t_{\pi(F\triangleleft g_1),\pi(E\triangleleft g_2)} = t_{\pi(F),\pi(E)}$, for every $g_1, g_2 \in SO(2)$.

The collection $\{T(F, E)\}$ is called *transport data.*

1.3. **The parallel transport operator .** Let $\mathcal{H} = \mathbb{C}(X)$ denote the Hilbertian space of complex valued functions on X (here, the word Hilbertian means that $\mathcal{H}$ is not complete)[4]. In addition, $\mathcal{H}$ supports a unitary representation of the group

[4]In general, in this paper, we will not distinguish between an Hilbertian vector space and its completion and the correct choice between the two will be clear from the context.

$SO(V) \times SO(2)$, where the action of an element $g = (g_1, g_2)$ sends a function $s \in \mathcal{H}$ to a function $g \cdot s$, given by

$$(g \cdot s)(F) = s\left(g_1^{-1} \rhd F \lhd g_2\right),$$

for every $F \in X$.

Using the transport data, we define an integral operator $T : \mathcal{H} \to \mathcal{H}$ by

$$T(s)(F) = \int_{E \in X} T(F, E)\, s(E)\, dE,$$

for every $s \in \mathcal{H}$, where dE denotes the normalized Haar measure on X.

- The symmetry property implies that T is self adjoint.
- The invariance property implies that T commutes with the $SO(V)$ action, namely $T(g \cdot s) = g \cdot T(s)$ for every $s \in \mathcal{H}$ and $g \in SO(V)$.
- The implication of the equivariance property will be discussed later when we study the kernel of T.

The operator T is called *parallel transport operator*.

1.3.1. *Localized parallel transport operator.* Let us fix an angle $a \in [0, \pi]$, designating an opening angle of a spherical cap on the sphere. Consider the parameter $h = 1 - \cos(a)$, taking values in the interval $[0, 2]$.

Define an integral operator $T_h : \mathcal{H} \to \mathcal{H}$ by

$$T_h(s)(F) = \int_{E \in B(F,a)} T(F, E)\, s(E)\, dE. \tag{1.1}$$

where $B(F, a) = \{E \in X : (\pi(F), \pi(E)) > \cos(a)\}$. Similar considerations as before show that T_h is self-adjoint and, in addition, commutes with the $SO(V)$ action.

The operator T_h is a localization of the operator of parallel transport, in the sense, that only frames with close viewing directions interacts through the integral (1.1). The operator T_h is called *localized parallel transport operator*.

1.4. **Spectral properties of the localized parallel transport operator.** We focus our attention on the spectral properties of the operator T_h, in the regime $h \ll 1$, since this is the relevant regime for the class averaging application.

Theorem 3. *The operator T_h has a discrete spectrum $\lambda_n(h)$, $n \in \mathbb{N}$, such that $\dim \mathcal{H}(\lambda_n(h)) = 2n + 1$, for every $h \in (0, 2]$, Moreover, in the regime $h \ll 1$, the eigenvalue $\lambda_n(h)$ has the asymptotic expansion*

$$\lambda_n(h) = \frac{1}{2}h - \frac{1 + (n+2)(n-1)}{8}h^2 + O\left(h^3\right).$$

For a proof, see Section 2.

In fact, each eigenvalue $\lambda_n(h)$, as a function of h, is a polynomial of degree $n+1$. In Section 2, we give a complete description of these polynomials by means of a generating function. To get some feeling for the formulas that arise, we list below

the first four eigenvalues

$$\begin{aligned}\lambda_1(h) &= \frac{1}{2}h - \frac{1}{8}h^2,\\ \lambda_2(h) &= \frac{1}{2}h - \frac{5}{8}h^2 + \frac{1}{6}h^3,\\ \lambda_3(h) &= \frac{1}{2}h - \frac{11}{8}h^2 + \frac{25}{24}h^3 - \frac{15}{64}h^4,\\ \lambda_4(h) &= \frac{1}{2}h - \frac{19}{8}h^2 + \frac{27}{8}h^3 - \frac{119}{64}h^4 + \frac{7}{20}h^5.\end{aligned}$$

The graphs of $\lambda_i(h)$, $i = 1, 2, 3, 4$ are given in the figure below.

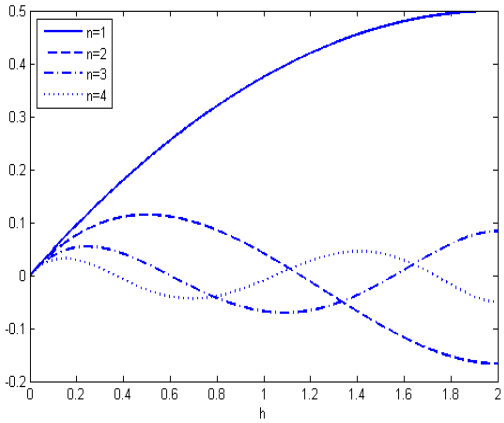

Figure 1.1. The first four eigen values of T_h

1.4.1. *Spectral gap.* Noting that $\lambda_2(h)$ attains its maximum at $h = 1/2$, we have

Theorem 4. *For $h \in [0, 2]$, the maximal eigenvalue of T_h is $\lambda_1(h)$. Moreover, for $h \in [0, 1/2]$, there is a spectral gap $G(h)$ of the form*

$$G(h) = \lambda_1(h) - \lambda_2(h) = \frac{1}{2}h^2 - \frac{1}{6}h^3.$$

For a proof, see Appendix A.

Consequently, in the regime $h \ll 1$, the spectral gap behaves like

$$G(h) \sim \frac{1}{2}h^2.$$

1.5. **Extrinsic and intrinsic models.** Our goal is to describe an **intrinsic** model of the Hermitian vector space W, which is expressed in the terms of the parallel transport operator and the Hermitian structure of $\mathcal{H}$ alone. We assume that $h < 1/2$ (see Theorem 4)).

Intrinsic model: Denote by $\mathbb{W} = \mathcal{H}(\lambda_{\max}(h))$ the maximal eigenspace of T_h.

Let $ev_F : \mathcal{H} \to \mathbb{C}$ be the evaluation morphism at the point $F \in X$. Define the morphism

$$\varphi_F = \sqrt{2/3} \cdot (ev_{F|\mathbb{W}})^* : \mathbb{C} \to \mathbb{W}.$$

The intrinsic model is the pair $(\mathbb{W}, \{\varphi_F : F \in X\})$.

Extrinsic model: Let $\delta_F : \mathbb{C} \to W$ be the unique complex morphism sending $1 \in \mathbb{C}$ to $f_1 - if_2 \in W$. The extrinsic model is the pair $(W, \{\delta_F : F \in X\})$.

The main result of this paper asserts that there exists an isomorphism between the intrinsic and the extrinsic models. Namely, that W equipped with the collection of natural maps $\{\delta_F : \mathbb{C} \to W\}$ is canonically isomorphic, as an Hermitian vector space, to $\mathbb{W}$ equipped with the collection of maps $\{\varphi_F : \mathbb{C} \to \mathbb{W}\}$.

The isomorphism can be realized explicitly as follows. Let $\tau : W \to \mathcal{H}$ denote the canonical morphism which sends a vector $v \in W$ to the function $\tau(v) \in \mathcal{H}$, defined by

$$\tau(v)(F) = \sqrt{3/2} \cdot \delta_F^*(v).$$

Theorem 5. *The morphism τ maps W isomorphically, as an Hermitian vector space, onto the subspace $\mathbb{W} \subset \mathcal{H}$. Moreover,*

$$\tau \circ \delta_F = \varphi_F,$$

for every $F \in X$.

For a proof, see Appendix A (the proof uses the results and terminology of Section 2).

The main application of Theorem 5 is, that it enables to express in intrinsic terms the inner product between viewing directions. The precise statement is

Theorem 6. *For every $F, E \in X$, we have*

$$(\pi(F), \pi(E)) = |\langle \varphi_F(v), \varphi_E(u) \rangle| - 1, \tag{1.2}$$

for any choice of unit numbers $v, u \in \mathbb{C}$.

For a proof, see Appendix A. Note that substituting $v = u = 1$ in (**??**) we obtain (0.6).

1.6. **Explanation of Theorems 1 and 2.** We end this section with an explanation of the two main statements that appeared in the introduction. The explanation is based on an approximation argument of the discrete from the continuous.

The main observation is that in the limit of infinite number of frames which are independently drawn from the normalized Haar measure on X, the transport matrix $\widetilde{T}_N$ approaches the localized parallel transport operator $T_h : \mathcal{H} \to \mathcal{H}$. This implies that the spectral properties of $\widetilde{T}_N$ for large values of N are governed by the spectral properties of the operator T_h when h lies in the regime $h \ll 1$ which is the relevant regime for the class averaging problem.

The statement of Theorem 1 is explained by the fact that the maximal eigenvalue of T_h has multiplicity three (see Theorem 3) and that there exists a spectral gap $G(h) \sim h/2$, separating it from the rest of the spectrum (see Theorem 4). The later property ensures that the numerical computation of this eigenspace makes sense.

The statement of Theorem 2 is explained by the fact that the vector space $\mathbb{W}_N$ is a numerical approximation of the theoretical vector space $\mathbb{W}$ and Theorem 5.

2. Spectral analysis of the localized parallel transport operator

In this section we study the spectral properties of the localized parallel transport operator T_h, mainly concentrating on the regime $h \ll 1$. But, first we need to introduce some preliminaries from representation theory.

2.1. **Isotypic decompositions.** The Hilbert space $\mathcal{H}$, as a unitary representation of the group $SO(2)$, admits an isotypic decomposition

$$\mathcal{H} = \bigoplus_{k\in\mathbb{Z}} \mathcal{H}_k, \tag{2.1}$$

where a function $s \in \mathcal{H}_k$ if and only if $s(Fg) = g^k s(F)$, for every $F \in X$ and $g \in SO(2)$. In turns, each Hilbert space $\mathcal{H}_k$, as a representation of the group $SO(V)$, admits an isotypic decomposition

$$\mathcal{H}_k = \bigoplus_{n\in\mathbb{N}^{\geq 0}} \mathcal{H}_{n,k}, \tag{2.2}$$

where $\mathcal{H}_{n,k}$ denotes the component which is a direct sum of copies of the unique irreducible representation of $SO(V)$ which is of dimension $2n+1$. A particularly important property is that each irreducible representation which appears in (2.2) comes up with multiplicity one. This is summarized in the following theorem:

Theorem 7 (Multiplicity one)**.** *If $n < |k|$ then $\mathcal{H}_{n,k} = 0$. Otherwise, $\mathcal{H}_{n,k}$ is isomorphic to the unique irreducible representation of $SO(V)$ of dimension $2n+1$.*

for a proof, see Appendix A.

The following proposition is a direct implication of the equivariance property of the operator T_h:

Proposition 1. *We have*

$$\bigoplus_{k\neq -1} \mathcal{H}_k \subset \ker T_h.$$

Consequently, from now on, we will consider T_h as an operator from $\mathcal{H}_{-1}$ to $\mathcal{H}_{-1}$. Moreover, since for every $n \geq 1$, $\mathcal{H}_{n,-1}$ is an irreducible representation of $SO(V)$ and since T_h commutes with the group action, by Schur's Lemma T_h acts on $\mathcal{H}_{n,-1}$ as a scalar operator, namely

$$T_{h|\mathcal{H}_{n,-1}} = \lambda_n(h)\, Id.$$

The reminder of this section is devoted to the computation of the eigenvalues $\lambda_n(h)$. The strategy of the computation is to choose a point $F_0 \in X$ and a "good" vector $u_n \in \mathcal{H}_{n,-1}$ such that $u_n(F_0) \neq 0$ and then to use the relation

$$T_h(u_n)(F_0) = \lambda_n(h)\, u_n(F_0),$$

which implies that

$$\lambda_n(h) = \frac{T_h(u_n)(F_0)}{u_n(F_0)}. \tag{2.3}$$

2.2. **Set-up.** Fix a frame $F_0 \in X$, $F_0 = (f_1, f_2, f_3)$. Under this choice, we can safely identify the group $SO(V)$ with the group $SO(3)$ by sending an element $g \in SO(V)$ to the unique element $h \in SO(3)$ such that $g \rhd F_0 = F_0 \lhd h$. Hence, from now on, we will consider the frame manifold equipped with a commuting left and right actions of $SO(3)$.

Consider the following elements in the Lie algebra $so(3)$:

$$\begin{aligned} A_1 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \\ A_2 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \\ A_3 &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

The elements A_i, $i = 1, 2, 3$ satisfy the relations

$$\begin{aligned} [A_3, A_1] &= A_2, \\ [A_3, A_2] &= -A_1, \\ [A_1, A_2] &= A_3. \end{aligned}$$

Let (H, E, F) be the following sl_2 triple in the complexified Lie algebra $\mathbb{C}so(3)$:

$$\begin{aligned} H &= -2iA_3, \\ E &= iA_2 - A_1, \\ F &= A_1 + iA_2. \end{aligned}$$

Finally, let $\left(H^L, E^L, F^L\right)$ and $\left(H^R, E^R, F^R\right)$ be the associated (complexified) vector fields on X induced from the left and right action of $SO(3)$ respectively.

2.2.1. *Spherical coordinates.* We consider the spherical coordinates of the frame manifold $\omega : (0, 2\pi) \times (0, \pi) \times (0, 2\pi) \to X$, given by

$$\omega(\varphi, \theta, \alpha) = F_0 \lhd e^{\varphi A_3} e^{\theta A_2} e^{\alpha A_3}.$$

We have the following formulas

- The normalized Haar measure on X is given by the density
$$\frac{\sin(\theta)}{2(2\pi)^2} d\varphi d\theta d\alpha.$$
- The vector fields $\left(H^L, E^L, F^L\right)$ are given by
$$\begin{aligned} H^L &= 2i\partial_\varphi, \\ E^L &= -e^{-i\varphi}\left(i\partial_\theta + \cot(\theta)\,\partial_\varphi - 1/\sin(\theta)\,\partial_\alpha\right), \\ F^L &= -e^{i\varphi}\left(i\partial_\theta - \cot(\theta)\,\partial_\varphi + 1/\sin(\theta)\,\partial_\alpha\right). \end{aligned}$$
- The vector fields $\left(H^R, E^R, F^R\right)$ are given by
$$\begin{aligned} H^R &= -2i\partial_\alpha, \\ E^R &= e^{i\alpha}\left(i\partial_\theta + \cot(\theta)\,\partial_\alpha - 1/\sin(\theta)\,\partial_\varphi\right), \\ F^R &= e^{-i\alpha}\left(i\partial_\theta + \cot(\theta)\,\partial_\alpha - 1/\sin(\theta)\,\partial_\varphi\right). \end{aligned}$$

2.3. **Choosing a good vector.**

2.3.1. *Spherical functions.* Consider the subgroup $T \subset SO(3)$ generated by the infinitesimal element A_3. For every $k \in \mathbb{Z}$ and $n \geq k$, the Hilbert space $\mathcal{H}_{n,k}$ admits an isotypic decomposition with respect to the left action of T:

$$\mathcal{H}_{n,k} = \bigoplus_{m=-n}^{n} \mathcal{H}_{n,k}^{m},$$

where a function $s \in \mathcal{H}_{n,k}^{m}$ if and only if $s\left(e^{-tA_3} \rhd F\right) = e^{imt} s(F)$, for every $F \in X$. Functions in $\mathcal{H}_{n,k}^{m}$ are usually referred to in literature as (generalized) *spherical functions*. Our plan is to choose for every $n \geq 1$, a spherical function $u_n \in \mathcal{H}_{n,-1}^{1}$ and exhibit a closed formula for the generating function

$$\sum_{n \geq 1} u_n t^n.$$

Then, we will use this explicit generating function to compute $u_n(F_0)$ and $T_h(u_n)(F_0)$ and use (2.3) to compute $\lambda_n(h)$.

2.3.2. *Generating function.* For every $n \geq 0$, let $\psi_n \in \mathcal{H}_{n,0}^{0}$ be the unique spherical function such that $\psi_n(F_0) = 1$. These functions are the well known spherical harmonics on the sphere. Define the generating function

$$G_{0,0}(\varphi, \theta, \alpha, t) = \sum_{n \geq 0} \psi_n(\varphi, \theta, \alpha) t^n.$$

The following theorem is taken from [11]:

Theorem 8. *The function $G_{0,0}$ admits the following formula:*

$$G_{0,0}(\varphi, \theta, \alpha, t) = \left(1 - 2t\cos(\theta) + t^2\right)^{-1/2}.$$

Take $u_n = E^L F^R \psi_n$. Note that indeed $u_n \in \mathcal{H}_{n,-1}^{1}$ and define the generating function

$$G_{1,-1}(\varphi, \theta, \alpha, t) = \sum_{n \geq 1} u_n(\varphi, \theta, \alpha) t^n.$$

It follows that, $G_{1,-1} = E^L F^R G_{0,0}$. Direct calculation, using the formula in Theorem 8, reveals that

$$\begin{aligned} G_{1,-1}(\varphi, \theta, \alpha, t) &= e^{-i(\alpha+\varphi)}[3\sin(\theta)^2 t^2 \left(1 - 2t\cos(\theta) + t^2\right)^{-5/2} \\ &\quad -t\cos(\theta)\left(1 - 2t\cos(\theta) + t^2\right)^{-3/2} \\ &\quad -t\left(1 - 2t\cos(\theta) + t^2\right)^{-3/2}]. \end{aligned} \tag{2.4}$$

It is enough to consider $G_{1,-1}$ when $\varphi = \alpha = 0$. We use the notation $G_{1,-1}(\theta, t) = G_{1,-1}(0, \theta, 0, t)$. By (2.4)

$$\begin{aligned} G_{1,-1}(\theta, t) &= 3\sin(\theta)^2 t^2 \left(1 - 2t\cos(\theta) + t^2\right)^{-5/2} \\ &\quad -t\cos(\theta)\left(1 - 2t\cos(\theta) + t^2\right)^{-3/2} \\ &\quad -t\left(1 - 2t\cos(\theta) + t^2\right)^{-3/2}. \end{aligned} \tag{2.5}$$

2.4. **Computation of $u_n(F_0)$.** Observe that

$$G_{1,-1}(0,t) = \sum_{n\geq 1} u_n(F_0)\, t^n.$$

Direct calculation reveals that

$$\begin{aligned} G_{1,-1}(0,t) &= -2t(1-t)^{-3} \\ &= -2t\sum_{n\geq 0}\binom{-3}{n}(-1)^n t^n \\ &= -2\sum_{n\geq 1}\binom{-3}{n-1}(-1)^{n-1} t^n. \end{aligned}$$

Since $\binom{-3}{n-1} = \frac{(-1)^{n-1}}{2} n(n+1)$, we obtain

$$u_n(F_0) = -n(n+1). \tag{2.6}$$

2.5. **Computation of $T_h(u_n)(F_0)$.** Recall that $h = 1-\cos(a)$.

Using the definition of T_h, we obtain

$$T_h(u_n)(F_0) = \int\limits_{E\in B(F_0,a)} T(F_0,E)\, u_n(E)\, dE.$$

Using the spherical coordinates, the integral on the right hand side can be written as

$$\frac{1}{(2\pi)^2}\int_0^{2\pi} d\varphi \int_0^a \frac{\sin(\theta)}{2} d\theta \int_0^{2\pi} T(F_0, \omega(\varphi,\theta,\alpha))\, u_n(\omega(\varphi,\theta,\alpha)).$$

First

$$\begin{aligned} T(F_0,\omega(\varphi,\theta,\alpha)) &= T\left(F_0, F_0 \lhd e^{\varphi A_3} e^{\theta A_2} e^{\alpha A_3}\right) \\ &= T\left(F_0, e^{\varphi A_3} \rhd F_0 \lhd e^{\theta A_2} e^{\alpha A_3}\right) \\ &= T(e^{-\varphi A_3} \rhd F_0, F_0 \lhd e^{\theta A_2} e^{\alpha A_3}) \\ &= e^{i\varphi} T(F_0, F_0 \lhd e^{\theta A_2}) e^{i\alpha}, \end{aligned} \tag{2.7}$$

where the third equality uses the invariance property of the transport data and the second equality uses the equivariance property of the transport data.

Second, since $u_n \in \mathcal{H}^1_{n,-1}$ we have

$$u_n(\omega(\varphi,\theta,\alpha)) = e^{-i\varphi} u_n\left(F_0 \lhd e^{\theta A_2}\right) e^{-i\alpha}. \tag{2.8}$$

Combining (2.7) and (2.8), we conclude

$$\begin{aligned} T_h(u_n)(F_0) &= \int_0^a \frac{\sin(\theta)}{2} T(F_0, F_0 \lhd e^{\theta A_2}) u_n\left(F_0 \lhd e^{\theta A_2}\right) d\theta \\ &= \int_0^a \frac{\sin(\theta)}{2} u_n\left(F_0 \lhd e^{\theta A_2}\right) d\theta. \end{aligned} \tag{2.9}$$

where the second equality uses the fact that $F_0 \lhd e^{\theta A_2}$ is the parallel transport of F_0 along the unique geodesic connecting $\pi(F_0)$ with $\pi\left(F_0 \lhd e^{\theta A_2}\right)$.

Denote

$$I_n(h) = \int_0^a \frac{\sin(\theta)}{2} u_n\left(F_0 \lhd e^{\theta A_2}\right) d\theta.$$

Define the generating function $I(h,t) = \sum_{n\geq 0} I_n(h)\, t^n$ and observe that

$$I(h,t) = \int_0^a \frac{\sin(\theta)}{2} G_{1,-1}(\theta,t)\, d\theta.$$

Direct calculation reveals that

$$\begin{aligned} I(h,t) &= 1/2[h\left(1+2t(h-1)+t^2\right)^{-1/2} \\ &\quad -th(2-h)\left(1+2t(h-1)+t^2\right)^{-3/2} \\ &\quad -t^{-1}(\left(1+2t(h-1)+t^2\right)^{1/2}-(1-t))]. \end{aligned} \tag{2.10}$$

2.6. **Proof of Theorem 3.** Expanding $I(h,t)$ with respect to the parameter t reveals that the function $I_n(h)$ is a polynomial in h of degree $n+1$. Then, using Equation (2.3), we get

$$\lambda_n(h) = -\frac{I_n(h)}{n(n+1)}.$$

In principle, it is possible to obtain a closed formula for $\lambda_n(h)$ for every $n \geq 1$.

2.6.1. *Quadratic approximation.* We want to compute the first three terms in the Taylor expansion of $\lambda_n(h)$:

$$\lambda_n(h) = \lambda_n(0) + \partial_h\lambda_n(0) + \frac{\partial_h^2\lambda_n(0)}{2} + O\left(h^3\right).$$

We have

$$\begin{aligned} \lambda_n(0) &= -\frac{I_n(0)}{n(n+1)}, \\ \partial_h\lambda_n(0) &= -\frac{\partial_h I_n(0)}{n(n+1)}, \\ \partial_h^2\lambda_n(0) &= -\frac{\partial_h^2 I_n(0)}{n(n+1)}. \end{aligned}$$

Observe that

$$\partial_h^k I(0,t) = \sum_{n\geq 1} \partial_h^k I_n(0).$$

Direct computation, using Formula (2.10), reveals that

$$\begin{aligned} I(0,t) &= 0, \\ \partial_h I(0,t) &= -\sum_{n\geq 1} n(n+1)\, t^n, \\ \partial_h^2 I(0,t) &= \frac{1}{4}\sum_{n\geq 1} n(n+1)\left(1+(n+2)(n-1)\right) t^n. \end{aligned}$$

Combing all the above yields the desired formula

$$\lambda_n(h) = \frac{1}{2}h - \frac{1+(n+2)(n-1)}{8}h^2 + O\left(h^3\right).$$

This concludes the proof of the theorem.

Appendix A. Proofs

A.1. **Proof of Theorem 4.** The proof is based on two technical lemmas.

Lemma 1. *The following estimates hold:*

1. *There exists* $h_1 \in (0,2]$ *such that* $\lambda_n(h) \leq \lambda_1(h)$*, for every* $n \geq 1$ *and* $h \in [0,h_1]$.
2. *There exists* $h_2 \in (0,2]$ *such that* $\lambda_n(h) \leq \lambda_2(h)$*, for every* $n \geq 2$ *and* $h \in [0,h_2]$.

The proof appears below.

Lemma 2. *The following estimates hold:*

1. *There exists* N_1 *such that* $\lambda_n(h) \leq \lambda_1(h)$*, for every* $n \geq N_1$ *and* $h \in [h_1,2]$.
2. *There exists* N_2 *such that* $\lambda_n(h) \leq \lambda_1(h)$*, for every* $n \geq N_2$ *and* $h \in [h_2,1/2]$.

The proof appears below.

Granting the validity of these two lemmas we can finish the proof of the theorem.

First we prove that $\lambda_n(h) \leq \lambda_1(h)$, for every $n \geq 1$ and $h \in [0,2]$: By Lemmas 1,2 we get that $\lambda_n(h) \leq \lambda_1(h)$ for every $h \in [0,2]$ when $n \geq N_1$. Then, we verify directly that $\lambda_n(h) \leq \lambda_1(h)$ for every $h \in [0,2]$ in the finitely many cases when $n < N_1$.

Similarly, we prove that $\lambda_n(h) \leq \lambda_2(h)$ for every $n \geq 2$ and $h \in [0,1/2]$: By Lemmas 1,2 we get that $\lambda_n(h) \leq \lambda_2(h)$ for every $h \in [0,1/2]$ when $n \geq N_2$. Then, we verify directly that $\lambda_n(h) \leq \lambda_1(h)$ for every $h \in [0,1/2]$ in the finitely many cases when $n < N_2$.

This concludes the proof of the theorem.

A.2. **Proof of Lemma 1.** The strategy of the proof is to reduce the statement to known facts about Legendre polynomials.

Recall $h = 1 - \cos(a)$. Here, it will be convenient to consider the parameter $z = \cos(a)$, taking values in the interval $[-1,1]$.

We recall that Legendre polynomials $P_n(z)$, $n \in \mathbb{N}$ appear as the coefficients of the generating function

$$P(z,t) = \left(t^2 - 2tz + 1\right)^{-(1/2)}.$$

Let

$$J_n(z) = \begin{cases} \frac{1}{2(1-z)} & n = 0 \\ \frac{1}{2(1-z)} & n = 1 \\ \partial_z \lambda_{n-1}(z) & n \geq 2 \end{cases}.$$

Consider the generating function

$$J(z,t) = \sum_{n=0}^{\infty} J_n(z)\, t^n.$$

The function $J(z,t)$ admits the following closed formula

$$J(z,t) = \frac{t + tz + t^2 + 1}{2(1-z)} \left(t^2 + 2tz + 1\right)^{-1/2}. \tag{A.1}$$

Using (A.1), we get that for $n \geq 2$

$$J_n(z) = \frac{1}{2(1-z)} \left(Q_n(z) + (1+z) Q_{n-1}(z) + Q_{n-2}(z) \right),$$

where $Q_n(z) = (-1)^n P_n(z)$. In order to prove the lemma, it is enough to show that there exists $z_0 \in (-1, 1]$ such that for every $z \in [-1, z_0]$ the following inequalities hold

- $Q_n(z) \leq Q_3(z)$ for every $n \geq 3$.
- $Q_n(z) \leq Q_2(z)$ for every $n \geq 2$.
- $Q_n(z) \leq Q_1(z)$ for every $n \geq 1$.
- $Q_n(z) \leq Q_0(z)$ for every $n \geq 0$.

These inequalities follow from the following technical proposition.

Proposition 2. *Let $n_0 \in \mathbb{N}$. There exists $z_0 \in (-1, 1]$ such that $Q_n(z) < Q_{n_0}(z)$, for every $z \in [-1, z_0]$ and $n \geq n_0$.*

The proof appears below.

Take $h_0 = h_1 = 1 + z_0$. Granting Proposition A.1, verify that $J_n(z) \leq J_2(z)$, for $n \geq 2$, $z \in [-1, z_0]$ which implies that $\lambda_n(h) \leq \lambda_1(h)$, for $n \geq 2$, $h \in [0, h_0]$ and $J_n(z) \leq J_3(z)$, for $n \geq 3$, $z \in [-1, z_0]$ which implies that $\lambda_n(h) \leq \lambda_2(h)$, for $n \geq 3$, $h \in [0, h_0]$.

This concludes the proof of the Lemma.

A.2.1. *Proof of Proposition 2.* Denote by $a_1 < a_2 < ... < a_n$ the zeroes of $Q_n(\cos(a))$ and by $\mu_1 < \mu_2 < ... < \mu_{n-1}$ the local extrema of $Q_n(\cos(a))$.

The following properties of the polynomials Q_n are implied from known facts about Legendre polynomials (Properties 1 and 2 can be verified directly), that can be found for example in the book [10]:

Property 1: $a_i < \mu_i < a_{i+1}$, for $i = 1, .., n-1$.
Property 2: $Q_n(-1) = 1$ and $\partial_z Q_{n+1}(-1) < \partial_z Q_n(-1) < 0$, for $n \in \mathbb{N}$.
Property 3: $|Q_n(\cos(\mu_i))| \geq \left| Q_n\left(\cos\left(\mu_{i+1}\right)\right) \right|$, for $i = 1, .., [n/2]$.
Property 4: $(i - 1/2)\pi/n \leq a_i \leq i\pi/(n+1)$, for $i = 1, .., [n/2]$.
Property 5: $\sin(a)^{1/2} \cdot |Q_n(\cos(a))| < \sqrt{2/\pi n}$, for $a \in [0, \pi]$.

Granting these facts, we can finish the proof.

By Properties 1,4

$$\frac{\pi}{2n} < \mu_1 < \frac{2\pi}{n+1}.$$

We assume that n is large enough so that, for some small $\epsilon > 0$

$$\sin(a_1) \geq (1 - \epsilon) a_1,$$

In particular, this is the situation when $n_0 \geq N$, for some fixed $N = N_\epsilon$.

By Property 5

$$\begin{aligned} |Q_n(\cos(\mu_1))| &< \sqrt{2/\pi n} \cdot \sin(\mu_1)^{-1/2} \\ &< \sqrt{2/\pi n} \cdot \sin(a_1)^{-1/2} \\ &< \sqrt{2/\pi n} \cdot ((1-\epsilon) a_1)^{-1/2} = \frac{2}{\pi\sqrt{1-\epsilon}}. \end{aligned}$$

Let $a_0 \in (0, \pi)$ be such that $Q_{n_0}(\cos(a)) > 2/\pi\sqrt{1-\epsilon}$, for every $a < a_0$. Take $z_0 = \cos(a_0)$.

Finally, in the finitely many cases where $n_0 \leq n \leq N$, the inequality $Q_n(z) < Q_{n_0}(z)$ can be verified directly.

This concludes the proof of the proposition.

A.3. **Proof of Lemma 2.** We have the following identity:

$$tr\left(T_h^2\right) = \frac{h}{2}, \tag{A.2}$$

for every $h \in [0,2]$. The proof of (A.2) is by direct calculation:

$$\begin{aligned} tr\left(T_h^2\right) &= \int\limits_{F\in X} T_h^2(F,F)\,dF = \\ &= \int\limits_{F\in X}\int\limits_{E\in B(F,a)} T_h(F,E)\circ T_h(E,F)\,dEdF. \end{aligned}$$

Since $T_h(F,E) = T_h(E,F)^{-1}$ (symmetry property), we get

$$tr\left(T_h^2\right) = \int\limits_{F\in X}\int\limits_{E\in B(F,a)} dEdF = \int\limits_0^a \frac{\sin(\theta)}{2}d\theta = \frac{1-\cos(a)}{2}.$$

Substituting, $a = \cos^{-1}(1-h)$, we get the desired formula $tr\left(T_h^2\right) = h/2$.

On the other hand,

$$tr\left(T_h^2\right) = \sum_{n=1}^{\infty} tr\left(T^2_{h|\mathcal{H}_{n,-1}}\right) = \sum_{n=1}^{\infty}(2n+1)\,\lambda_n(h)^2\,. \tag{A.3}$$

From (A.2) and (A.3) we obtain the following upper bound

$$\lambda_n(h) \leq \frac{\sqrt{h}}{\sqrt{4n+2}}. \tag{A.4}$$

Now we can finish the proof.

First estimate: We know that $\lambda_1(h) = h/2 - h^2/8$, hence, one can verify directly that there exists N_1 such that $\sqrt{h}/\sqrt{4n+2} \leq \lambda_1(h)$ for every $n \geq N_1$ and $h \in [h_1, 2]$, which implies by (A.4) that $\lambda_n(h) \leq \lambda_1(h)$ for every $n \geq N_1$ and $h \in [h_1, 2]$.

Second estimate: We know that $\lambda_2(h) = h/2 - 5h^2/8 + h^3/6$, therefore, one can verify directly that there exists N_2 such that $\sqrt{h}/\sqrt{4n+2} \leq \lambda_2(h)$ for every $n \geq N_2$ and $h \in [h_2, 1/2]$, which implies by (A.4) that $\lambda_n(h) \leq \lambda_2(h)$ for every $n \geq N_2$ and $h \in [h_2, 1/2]$.

This concludes the proof of the Lemma.

A.4. **Proof of Theorem 5.** We begin by proving that τ maps $W = \mathbb{C}V$ isomorphically, as an Hermitian space, onto $\mathbb{W} = \mathcal{H}(\lambda_{\max}(h))$.

The crucial observation is, that $\mathcal{H}(\lambda_{\max}(h))$ coincide with the isotypic subspace $\mathcal{H}_{1,-1}$ (see Section 2). Consider the morphism $\alpha = \sqrt{2/3}\cdot\tau : W \to \mathcal{H}$, given by

$$\alpha(v)(F) = \delta_F^*(v).$$

First claim is, that $\operatorname{Im}\alpha \subset \mathcal{H}_{-1}$, namely, that $\delta^*_{Fg}(v) = g^{-1}\delta^*_F(v)$, for every $v \in W$, $F \in X$ and $g \in SO(2)$. Denote by $\langle\cdot,\cdot\rangle_{std}$ the standard Hermitian product

on $\mathbb{C}$. Now write

$$\begin{aligned}\langle \delta_{Fg}^{*}(v), z\rangle_{std} &= \langle v, \delta_{Fg}(z)\rangle = \langle v, \delta_F(gz)\rangle \\ &= \langle \delta_F^{*}(v), gz\rangle_{std} = \langle g^{-1}\delta_F^{*}(v), z\rangle_{std}.\end{aligned}$$

Second claim is, that α is a morphism of $SO(V)$ representations, namely, that $\delta_F^{*}(gv) = \delta_{g^{-1}F}(v)$, for every $v \in W$, $F \in X$ and $g \in SO(V)$. This statement follows from

$$\begin{aligned}\langle \delta_F^{*}(gv), z\rangle_{std} &= \langle gv, \delta_F(z)\rangle = \left\langle v, g^{-1}\delta_F(z)\right\rangle \\ &= \langle v, \delta_{g^{-1}F}(z)\rangle = \langle \delta_{g^{-1}F}^{*}(v), z\rangle_{std}.\end{aligned}$$

Consequently, the morphism α maps W isomorphically, as a unitary representation of $SO(V)$, onto $\mathcal{H}_{1,-1}$, which is the unique copy of the three dimensional representation of $SO(V)$ in $\mathcal{H}_{-1}$. In turns, this implies that, up to a scalar, α and, hence τ, are isomorphisms of Hermitian spaces. In order to complete the proof it is enough to show that

$$tr(\tau^{*}\circ\tau) = 3.$$

This follows from

$$\begin{aligned}tr(\tau\circ\tau^{*}) &= \frac{3}{2}tr(\alpha^{*}\circ\alpha) \\ &= \frac{3}{2}\int\limits_{v\in S(W)}\langle \alpha^{*}\circ\alpha(v), v\rangle_{\mathcal{H}}\,dv \\ &= \frac{3}{2}\int\limits_{v\in S(W)}\langle \alpha(v), \alpha(v)\rangle_{\mathcal{H}}\,dv \\ &= \frac{3}{2}\int\limits_{v\in S(W)}\int\limits_{F\in X}\langle \delta_F^{*}(v), \delta_F^{*}(v)\rangle_{std}\,dvdF \\ &= \frac{3}{2}\int\limits_{v\in S(W)}\int\limits_{F\in X}2dvdF = 3.\end{aligned}$$

where dv denotes the normalized Haar measure on the five dimensional sphere $S(W)$.

Next, we prove that $\tau\circ\delta_F = \varphi_F$, for every $F \in X$. The starting point is the equation $ev_{F|W}\circ\alpha = \delta_F^{*}$, which follows from the definition of the morphism α and the fact that $\operatorname{Im}\alpha = \mathbb{W}$. This implies that $\varphi_F^{*}\circ\tau = \delta_F^{*}$. The statement now follows from

$$\begin{aligned}\varphi_F^{*}\circ\tau &= \delta_F^{*} \Rightarrow \varphi_F^{*}\circ(\tau\circ\tau^{*}) = \delta_F^{*}\circ\tau^{*} \\ &\Rightarrow \varphi_F^{*} = \delta_F^{*}\circ\tau^{*} \Rightarrow \varphi_F = \tau\circ\delta_F.\end{aligned}$$

This concludes the proof of the theorem.

A.5. **Proof of Theorem 6.** We use the following terminology: for every $F \in X$, let $\widetilde{\delta}_F : \mathbb{C} \to V$ be the map given by $\widetilde{\delta}_F(x+iy) = xf_1 + yf_2$ and observe that $\delta_F(v) = \widetilde{\delta}_F(v) - i\widetilde{\delta}_F(iv)$, for every $v \in \mathbb{C}$.

We proceed with the proof. Let $F, E \in X$. Denote $x = \pi(F)$, $y = \pi(E)$. Choose unit vectors $v_F, v_E \in \mathbb{C}$ such that $\widetilde{\delta}_F(v_F) = \widetilde{\delta}_E(v_E) = v$.

Write

$$\begin{aligned}\langle \delta_F(v_F), \delta_E(v_E)\rangle &= \left\langle \widetilde{\delta}_F(v_F) - i\widetilde{\delta}_F(iv_F), \widetilde{\delta}_E(v_E) - i\widetilde{\delta}_E(iv_E)\right\rangle \\ &= \left(\widetilde{\delta}_F(v_F), \widetilde{\delta}_E(v_E)\right) + \left(\widetilde{\delta}_F(iv_F), \widetilde{\delta}_E(iv_E)\right) \\ &\quad -i\left(\widetilde{\delta}_F(iv_F), \widetilde{\delta}_E(v_E)\right) + i\left(\widetilde{\delta}_F(v_F), \widetilde{\delta}_E(iv_E)\right). \end{aligned} \tag{A.5}$$

For every frame G and vector $v_G \in \mathbb{C}$, the following identity can be easily verified:

$$\widetilde{\delta}_G(iv_G) = \pi(G) \times \widetilde{\delta}_G(v_G).$$

This implies that

$$\begin{aligned}\widetilde{\delta}_F(iv_F) &= x \times \widetilde{\delta}_F(v_F) = x \times v, \\ \widetilde{\delta}_E(iv_E) &= y \times \widetilde{\delta}_E(v_E) = y \times v.\end{aligned}$$

Combining these identities with Equation (A.5), we obtain

$$\langle \delta_F(v_F), \delta_E(v_E)\rangle = (v,v) + (x\times v, y\times v) - i(x\times v, v) + i(v, y\times v).$$

Since $v \in \widetilde{\delta}_F(P_F) \cap \widetilde{\delta}_E(v_E)$, it follows that $(x\times v, v) = (v, y\times v) = 0$. In addition,

$$(x\times v, y\times v) = \det\begin{pmatrix}(x,y) & (x,v)\\ (y,v) & (v,v)\end{pmatrix} = (x,y).$$

Thus, we obtain that $\langle \delta_F(v_F), \delta_E(v_E)\rangle = 1 + (x,y)$. Since the right hand side is always ≥ 0 it follows that

$$|\langle \delta_F(v_F), \delta_E(v_E)\rangle| = 1 + (x,y). \tag{A.6}$$

Now, notice that the left hand side of A.6 does not depend on the choice of the unit vectors v_F and v_E.

To finish the proof, we use the isomorphism τ which satisfies $\tau \circ \delta_E = \varphi_E$ for every $E \in X$, and get

$$|\langle \varphi_F(v_F), \varphi_E(v_E)\rangle| = 1 + (x,y).$$

This concludes the proof of the theorem.

A.6. **Proof of Proposition 7.** The basic observation is, that $\mathcal{H}$, as a representation of $SO(V) \times SO(3)$, admits the following isotypic decomposition

$$\mathcal{H} = \bigoplus_{n=0}^{\infty} V_n \otimes U_n,$$

where V_n is the unique irreducible representation of $SO(V)$ of dimension $2n+1$, and, similarly, U_n is the unique irreducible representation of $SO(3)$ of dimension $2n+1$. This assertion, principally, follows from the Peter Weyl Theorem for the regular representation of $SO(3)$.

This implies that the isotypic decomposition of $\mathcal{H}_k$ takes the following form

$$\mathcal{H}_k = \bigoplus_{n=0}^{\infty} V_n \otimes U_n^k,$$

where U_n^k is the weight k space with respect to the action $SO(2) \subset SO(3)$. The statement now follows from the following standard fact about the weight decomposition:

$$\dim U_n^k = \begin{cases} 0 & n < k \\ 1 & n \geq k \end{cases} .$$

This concludes the proof of the theorem.

References

[1] Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., MacKinnon, R. (1998) The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity, *Science* 3 April 1998 280 (5360): pp. 69–77.

[2] Frank, J., Three-Dimensional Electron Microscopy of Macromolecular Assemblies. *Visualization of Biological Molecules in Their Native State*, Oxford (2006).

[3] Vainshtein, B., and Goncharov, A., Determination of the spatial orientation of arbitrarily arranged identical particles of an unknown structure from their projections. *Proc. llth Intern. Congr. on Elec. Mirco.*, 459-460 (1986)

[4] MacKinnon, R. (2004) Potassium Channels and the Atomic Basis of Selective Ion Conduction, 8 December 2003, *Nobel Lecture, in Bioscience Reports*, Springer Netherlands, 24 (2) pp.75–100.

[5] Natterer, F., The Mathematics of Computerized Tomography, *SIAM: Society for Industrial and Applied Mathematics, Classics in Applied Mathematics* (2001).

[6] Penczek, P.A., Zhu, J., and Frank, J. (1996) A common-lines based method for determining orientations for N > 3 particle projections simultaneously. Ultramicroscopy 63, pp. 205-218.

[7] Hadani, R. and Singer, A., Representation theoretic patterns in three dimensional cryo-electron macroscopy I - The Intrinsic reconstitution algorithm. *Submitted (2009), arXiv*:0910.3009v1.

[8] Singer, A. and Shkolnisky, Y., Three-diemensional Structure Determination From Common Lines in Cryo-EM by Eigenvectors and Semidefinite Programming, *Submitted* (2009).

[9] A. Singer., Z. Zhao., Y. Shkolnisky., R. Hadani., Viewing Angle Classification Of Cryo-Electron Microscopy Images Using Eigenvectors, *Submitted (2009)*. A PDF version can be doenloaded from http://www.math.princeton.edu/~amits/publications.html.

[10] Szegö, G., Orthogonal polynomials, *Colloquium Publications,* Volume XXIII, American Mathematical Scociety (1939).

[11] Michael, E. Taylor., Noncommutative Harmonic Analysis. *Mathematical Surveys and Monographs*, Volume **22**, American Mathematical Society.

[12] Van Heel, M. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy* **21** (2):111-23 (1987). PMID: 12425301 [PubMed - indexed for MEDLINE]

[13] Wang, L. and Sigworth, F. J., Cryo-EM and single particles. *Physiology (Bethesda),* **21**:13-8. Review. PMID: 16443818 [PubMed – indexed for MEDLINE] (2006).

Current address: Department of Mathematics, University of Texas at Austin, Austin C1200, USA.

Department of Mathematics and PACM, Princeton University, Fine Hall, Washington Road, Princeton NJ 08544-1000, USA

E-mail address: hadani@math.utexas.edu

E-mail address: amits@math.princeton.edu