DISSERTATION

KARHUNEN-LOÈVE DECOMPOSITION IN THE PRESENCE OF SYMMETRY

Submitted by
Brigitte Lahme
Department of Mathematics

In partial fulfillment of the requirements
for the Degree of Doctor of Philosophy
Colorado State University
Fort Collins, Colorado
Summer 1999

UMI Number: 9947929

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

April 21, 1999

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY BRIGITTE LAHME ENTITLED "Karhunen-Loève Decomposition in the Presence of Symmetry" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

____

____

____

____

____

Adviser

____

Department Head

ABSTRACT

THE KARHUNEN-LOÈVE DECOMPOSITION IN THE PRESENCE OF SYMMETRY

The Karhunen-Loève decomposition is widely used with data which very often exhibits some symmetry, afforded by a group action. For a finite group we derive an algorithm using group representation theory to reduce the cost of determining the KL basis. We demonstrate the technique on a Lorenz-type ODE system. For a compact group such as tori or $SO(3, \mathbb{R})$ the method also applies, and we extend the results to these cases. We then investigate the group $SO(3)$ in more detail and follow the algorithm with a specific example of data lying in the space of vector fields that are tangent to the sphere.

Brigitte Lahme
Department of Mathematics
Colorado State University
Fort Collins, Colorado 80523
Summer 1999

Contents

1 Introduction 5

I Finite Groups 8

2 The Karhunen-Loève Decomposition 9

2.1 The optimal basis 10

2.2 Properties of the KL Decomposition 13

3 Basics from Group Representation Theory 16

3.1 Definitions and Schur's Lemma 16

3.2 Projectors, and Transfer Operators 18

4 KL Decomposition with Symmetry 20

4.1 KL Decomposition and Symmetry 20

4.2 The Block Matrices B^ρ 24

4.3 Algorithm for Exploiting Symmetry in the KL Basis Computation . . 28

4.4 The Computational Efficiency 29

5 Example 31

II Compact Lie Groups 35

6 Theory 37

6.1 KL Decomposition 37

6.2 Representation Theory 39

6.3 KL Decomposition and Symmetry 40

6.4 Example: The circle group S^1 42

7 The orthogonal group SO(3) 43

7.1 SO(3) 43

7.2 Representation of SO(3) - Scalar Functions on the Sphere 44

7.2.1 The Irreducible Representations 44

7.2.2 Symmetry Subspaces for SO(3) 48

8 Example: Polynomial Vector Fields 54

8.1 Theory 54

8.1.1 Irreducible Representations 55

8.1.2 Examples 62

8.2 Divergence-Free Vector Fields 65

8.2.1 Curl and Divergence 70

8.3 Example: Ginzburg Laundau Equation 71

8.3.1 The Ginzburg Landau Equation 71

8.3.2 The Data Set 74

8.3.3 Following the Algorithm 78

8.3.4 Discussion of Results 82

9 Conclusion 91

List of Figures

8.1 $\dot{x} = y + z,\ \dot{y} = -x + z,\ \dot{z} = -x - y$ 64

8.2 $\dot{x} = y^2 + z^2,\ \dot{y} = -xy,\ \dot{z} = -xz$ 64

8.3 $\dot{x} = 2yz,\ \dot{y} = -xz,\ \dot{z} = -xy$ 65

8.4 Grid on the sphere, simplified 74

8.5 Change of coordinates at the poles 75

8.6 Absolute value of the first coordinate at a grid point on the 3rd latitude line 76

8.7 Absolute value of the first coordinate at a grid point on the 8thth latitude line 77

8.8 Absolute value of the first coordinate at the north pole latitude line . 77

8.9 Weights for the discretization of the integral. The darker the color, the bigger the value. 78

8.10 Graphical interpretation of the weights 79

8.11 Mean square error of the 5000 element data set in the different subspaces. 84

8.12 Linear basis vector, 3-dimensional irreducible representation, real part 85

8.13 Linear basis vector, 3-dimensional irreducible representation, imaginary part 85

8.14 Quadratic basis vector, 3-dimensional irreducible representation, real part 86

8.15 Quadratic basis vector. 3-dimensional irreducible representation. imaginary part 86
8.16 Quadratic basis vector. 5-dimensional irreducible representation. real part 87
8.17 Quadratic basis vector. 5-dimensional irreducible representation. imaginary part 88
8.18 Cubic basis vector. 5-dimensional irreducible representation. real part 88
8.19 Cubic basis vector. 5-dimensional irreducible representation. imaginary part 89
8.20 Cubic basis vector. 7-dimensional irreducible representation. real part 89
8.21 Cubic basis vector. 7-dimensional irreducible representation. imaginary part 90

Chapter 1

Introduction

The Karhunen-Loève (KL) Decomposition (or Principal Component Analysis) is a powerful tool for linear data compression [AG]. This technique for analyzing a finite data set in a vector space V finds the optimal basis for the ambient space in the sense that the truncation errors are minimized for every desired compression dimension.

It is quite typical in data obtained from physical modeling that the underlying set of patterns enjoys some symmetry. The partial differential equations which model many physical processes exhibit inherent symmetry, which is generally well reflected in both experimental data and simulation data. The concept of symmetry in data sets is always related to the action of a group G on V, which carries the data set onto itself.

Whenever a group acts linearly on a vector space, the space splits into irreducible subspaces for the action: for each irreducible representation ρ of the group G there are $\dim(\rho)$ corresponding projectors P^{ρ}_{ii} whose images are the "symmetry subspaces" V^{ρ}_{i}. There are natural isomorphisms P^{ρ}_{ij} between V^{ρ}_{i} and V^{ρ}_{j} when $i \neq j$. Therefore for data enjoying symmetry, we have two apparently competing decompositions of the ambient vector space: one coming from the KL basis and one coming from the group representations. In this thesis we analyze the relationship between these two decompositions, and prove the following:

Theorem 1.1 *The KL basis for the full vector space V is obtained by projecting the data set (via the orthogonal projectors P_{11}^{ρ}) into the first symmetry subspaces V_1^{ρ} for each ρ, and computing a KL basis for the projected data set in V_1^{ρ}. These separate bases are then transferred (via the isomorphisms P_{j1}^{ρ}) to the other symmetry subspaces V_j^{ρ}, and concatenated to form the full KL basis for V.*

The KL basis for a data set in $\mathbb{R}^n$ is obtained by forming the n-by-n covariance matrix of the data and taking the eigenvectors. We show that the above process of passing to the symmetry subspaces reduces the full KL problem to a series of subproblems, each of smaller dimension. The total computational complexity is correspondingly reduced.

This thesis is divided into two main parts. In Part I we cover the case of a finite group G acting on a finite dimensional vector space V. In Part II we extend the results we found earlier to the case of compact Lie groups acting on manifolds of data. The structure of both parts is similar. We first introduce basic facts about the KL decomposition, then we present the relevant facts concerning group representation theory which we require. Chapter 4 is the main part of the paper. There we combine the KL decomposition with the group theory to divide the problem into the sub-problems. An exact algorithm using group representation theory to reduce the cost of determinng the KL basis is also given. We then present the analysis of the computational efficiency of the method for a finite group, concluding with a simple example illustrating the technique.

In Part II we first discuss briefly how to adjust the derived theory to the case of compact Lie groups acting on manifolds of data. The changes are fairly mild and the results are practically the same. In this part we concentrate on investigating the group of rotations in $\mathbb{R}^3$, the special orthogonal group SO(3). In Chapter 7 we present the decomposition of the vector space of scalar valued function on the sphere under the action of SO(3). Finally, in Chapter 8 we follow the *algorithm for exploiting symmetry in the KL basis computation* step by step with a specific data set lying in the manifold of vector fields that are tangent to the sphere.

The exploitation of symmetry in KL decomposition was first recognized by Sirovich in [Si]; a recent article of Smaoui and Armbruster applies these methods to eigenfunction computations [SA].

Part I

Finite Groups

Chapter 2

The Karhunen-Loève Decomposition

The Karhunen-Loève Decomposition is a very effective tool in data compression. It mainly known under this name in the field of pattern analysis. The fact that it is known under a variety of different names in different fields shows how widely used and important it is. In statistics it is known as Principal Component Analysis. in linear algebra as Singular Value Decomposition. and other well-known terms are Proper Orthogonal Decomposition and Empirical Orthogonal Eigenfunctions: see for example [Ka] [Lo]. [Fu]. and [AG]. It seems that the Karhunen-Loève Decomposition was first suggested by Pearson 1901 in [Pe] and then again by Hotelling in 1933 [Ho]. In 1965 Watanabe introduced the method into the field of pattern analysis [Wa].

The goal of the KL decomposition is the representation of a given set of patterns or data in an optimal coordinate system. This system has to capture the main characteristics of the data while avoiding unnecessarily high dimensionality. Of course. this description simplifies the processes considerably. In this chapter we will introduce the finite dimensional KL decomposition. We will begin by deriving the KL decomposition mathematically and list the main properties of the KL basis.

2.1 The optimal basis

Let V be a complex vector space of dimension d containing a finite set U of N *pattern* vectors, say $U = \{u_i\}_{i=1}^{N}$. We assume that V is equipped with the standard Hermitian inner product defined by $(x, y) = \overline{x}^T \cdot y$. The Karhunen-Loève decomposition of this set U of patterns provides an orthonormal basis for V which is optimal for arbitrary truncation errors, which we now explain.

Let $\mathcal{B} = \{v_1, \ldots, v_d\}$ be an ordered orthonormal basis for V. Then for any vector $u_i \in U$, we may write u_i uniquely as $u_i = \sum_j a_j v_j$, without error. The D-truncation of u_i with respect to this basis $\mathcal{B}$ is the vector $u_i^{(D)} = \sum_{j \le D} a_j v_j$, and the error vector of this truncation is $e_i^{(D)} = u_i - u_i^{(D)}$.

Given the set of pattern vectors U, and the basis $\mathcal{B}$, we define the *mean square error* of the D-truncation to be

$$\epsilon_{mse}^{(D)} = \frac{1}{|U|} \sum_i \|e_i^{(D)}\|^2 .$$

The KL basis for V is that orthonormal basis for which the mean square errors $\epsilon_{mse}^{(D)}$ are minimized for every D between 1 and d. If we would like to approximate our pattern set in a D_1 dimensional vector space, we use the first D_1 vectors in the KL basis. If, on the other hand, we prefer a $D_2 > D_1$ dimensional approximation, we just include the next $D_2 - D_1$ basis vectors. In both cases, the *mean square error* for the corresponding dimension is minimal. The KL decomposition is a method to approximate a pattern set that lies in a relatively high dimensional vector space in a lower dimensional subspace with the smallest possible error.

The construction of the KL basis for V is well-known and can be derived in different ways, see for example [Fu] or [Ki]. We would just like to sketch one derivation which leads to the following theorem; recall that $u^* = \overline{u}^T$.

Theorem 2.1 *The KL basis is given by the eigenvectors of the ensemble average covariance matrix* $C = (1/|U|)\sum_i u_i u_i^*$.

It is very common to subtract the average of the data set from every pattern. Geometrically this means that we move the center of the data set to the origin of the coordinate system. From now on we will assume that the data is concentrated at the origin. Such a set of data is also called *fluctuating vectors*.

Derivation of the result

A similiar way of describing this problem can be found in [Ki]. It is our goal to find an orthonormal basis $\{\phi_i\}_{i=1}^d$ which minimizes the error

$$\begin{aligned}\epsilon_{mse}^{(D)} &= \frac{1}{N}\sum_{i=1}^{N} ||e_i^{(D)}||^2 \\ &= \frac{1}{N}\sum_{i=1}^{N} ||u_i - u_i^{(D)}||^2.\end{aligned}$$

Since $u_i = \sum_{i=1}^d a_j^i \phi_j$ and $u_i^{(D)} = \sum_{i=1}^D a_j^i \phi_j$ we have

$$\epsilon_{mse}^{(D)} = \frac{1}{N}\sum_{i=1}^{N} ||\sum_{j=D+1}^{d} a_j^i \phi_j||^2.$$

where the norm is induced by the standard Hermitian product. Furthermore we can replace a_j^i by (u_i, ϕ_j) since we require the basis vectors to be orthonormal. This leads to

$$\begin{aligned}\epsilon_{mse}^{(D)} &= \frac{1}{N}\sum_{i=1}^{N} \left|\left| \sum_{j=D+1}^{d} (u_i, \phi_j)\phi_j \right|\right|^2 \\ &= \frac{1}{N}\sum_{i=1}^{N} (\sum_{j=D+1}^{d} (u_i, \phi_j)\phi_j, \sum_{k=D+1}^{d} (u_i, \phi_k)\phi_k) \\ &= \frac{1}{N}\sum_{i=1}^{N} \sum_{j,k=D+1}^{d} (u_i, \phi_j)(u_i, \phi_k)(\phi_j, \phi_k) \\ &= \frac{1}{N}\sum_{i=1}^{N} \sum_{j=D+1}^{d} (u_i, \phi_j)^2.\end{aligned}$$

Notice that $(u, \phi)^2 = (\phi, uu^*\phi)$, and recall that $C = \frac{1}{N}\sum_{i=1}^{N} u_i u_i^*$, to obtain:

$$\begin{aligned}
\epsilon_{mse}^{(D)} &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=D+1}^{d}(\phi_j, u_i u_i^* \phi_j) \\
&= \sum_{j=D+1}^{d}\left(\phi_j, \left[\frac{1}{N}\sum_{i=1}^{N} u_i u_i^*\right]\phi_j\right) \\
&= \sum_{j=D+1}^{d}(\phi_j, C\phi_j)
\end{aligned}$$

Now we will use the method of Lagrange multipliers with constraints $(\phi_j, \phi_j) = 1$ to find the vectors ϕ_j. These constraints will guarantee a non-trivial solution for the extreme values.

We define the function

$$g(\phi) = \sum_{j=D+1}^{d}(\phi_j, C\phi_j) - \sum_{j=D+1}^{d}\lambda_j((\phi_j, \phi_j) - 1), \quad \phi = (\phi_j),$$

with the Lagrange multipliers λ_j for $j = D+1, \ldots, d$. By a theorem of restraint optimization there are λ_j, $j = D+1, \ldots, d$, such that for the optimal solution $\phi_{D+1}, \ldots, \phi_d$ we have

$$\frac{\partial g}{\partial \phi_j} = (C - \lambda_j I)\phi_j = 0,$$

or in another formulation we get

$$C\phi_j = \lambda_j \phi_j.$$

To reach a minimum, the basis vectors ϕ_j have to be eigenvectors of C with eigenvalues being the Lagrange multipliers λ_j

There is a natural ordering of the eigenvalues, namely $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. Recall that since the matrix

$$C = \frac{1}{|U|}\sum_{i=1}^{N} u_i u_i^*$$

is positive semi-definite Hermitian, it has only non-negative eigenvalues. C is called the Ensemble Average Covariance Matrix, we will occasionally refer to it just as the Covariance matrix. With this ordering of the basis vectors each pattern is approximated optimally. This can be shown rather easily. As we already mentioned we have

$$\epsilon_{mse} = \sum_{j=D+1}^{d} (o_j, C o_j).$$

Since o_j is an eigenvector of C with eigenvalue λ_j, it follows that

$$\begin{aligned}\epsilon_{mse} &= \sum_{j=D+1}^{d} (o_j, \lambda_j o_j) \\ &= \sum_{j=D+1}^{d} \lambda_j.\end{aligned}$$

Thus the error is minimal as long as the eigenvalues are arranged in decreasing order.

A different proof that shows that we have found a basis that minimizes the error is based on [Wa] and can be read in [Ki]. It is shown that for an arbitrary orthonormal basis $\{v_j\}_{j=1}^{d}$ for V the corresponding mean square error is less than or equal to the one given above: that is with $\rho_j = (v_j, C v_j)$ and $\epsilon = \sum_{j=D+1}^{d} \rho_j$, we have $\epsilon_{mse} \leq \epsilon$.

2.2 Properties of the KL Decomposition

In this section we have collected a number of properties of the KL basis vectors and their eigenvalues [Ki].

1. In an ensemble of fluctuating vectors the coefficients are zero on average.

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N} a_j^i &= \frac{1}{N}\sum_{i=1}^{N}(u_i, \phi_j) \\ &= (\frac{1}{N}\sum_{i=1}^{N} u_i, \phi_j) \\ &= (0, \phi_j) \\ &= 0\end{aligned}$$

2. The coefficients of the KL expansion are uncorrelated on average.

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N} a_j^i a_k^i &= \frac{1}{N}\sum_{i=1}^{N}(u_i, \phi_j)(u_i, \phi_k) \\ &= \frac{1}{N}\sum_{i=1}^{N}(\phi_j, u_i u_i^* \phi_k) \\ &= (\phi_j, \frac{1}{N}\sum_{i=1}^{N} u_i u_i^* \phi_k) \\ &= (\phi_j, C\phi_k) \\ &= (\phi_j, \lambda_k \phi_k) \\ &= \lambda_k \delta_{jk}\end{aligned}$$

3. As we mentioned before, the eigenvalues of C are non-negative:

$$\lambda_j = \frac{1}{N}\sum_{i=1}^{N}(a_j^i)^2 \geq 0$$

for $j = 1, \ldots, d$.

4. The sum of the eigenvalues is called the *energy* of the system:

$$E(U) = \sum_{j=1}^{d} \lambda_j.$$

5. The statistical variance of the jth coordinate direction is proportional to the jth eigenvalue. We will denote the statistical variance of the jth coordinate direction by $var(a_j^i)$. where

$$\begin{aligned} var(a_j^i) &= \frac{1}{N-1}\sum_{i=1}^{N}(a_j^i - \frac{1}{N}\sum_{k=1}^{N} a_j^k)^2 \quad (\frac{1}{N}\sum_{k=1}^{N} a_j^k = 0) \\ &= \frac{N}{N-1}\cdot\frac{1}{N}\sum_{i=1}^{d}(a_j^i)^2 \\ &= \frac{N}{N-1}\lambda_j \\ &\approx \lambda_j \end{aligned}$$

Writing the KL transformation in matrix notation and denoting the matrix whose columns are the eigenvectors of C by Φ. we get

$$a^i = (a_1^i. a_2^i. \ldots. a_d^i) = u_i^* \Phi.$$

6. The KL basis diagonalizes the covariance matrix:

$$\begin{aligned} \frac{1}{N}\sum_{i=1}^{N}(a^i)^* a^i &= \frac{1}{N}\sum_{i=1}^{N}(\Phi^* u_i)(u_i^* \Phi) \\ &= \Phi^* C \Phi \\ &= \Lambda. \end{aligned}$$

where $\Lambda_{ii} = \lambda_i$ and $\Lambda_{ij} = 0$. when $i \neq j$.

Chapter 3

Basics from Group Representation Theory

In this section we collect the basic results of the representation theory of finite groups which we require. All of the results are completely standard and may be found in many texts, e.g. [FH], [Se], or [Vi].

3.1 Definitions and Schur's Lemma

Let G be a finite group with identity I and $|G|$ elements. Let V be a finite dimensional complex vector space.

Definition 3.1 *A* representation *of G on V is a homomorphism $\rho : G \to GL(V)$ of G to the group of linear automorphisms of V.*

Whenever $V = \mathbb{C}^d$ or $V = \mathbb{R}^d$ we may think of ρ as a map sending each group element to a d-by-d matrix where d is the dimension of V. Then we have the equivalent concepts of a representation of G on V and the action of the matrices corresponding to the group elements on the vectors in V. Sometimes V is also called the representation of G. We will often write $g \cdot v$ or gv for $\rho(g)(v)$. The dimension of V is sometimes called the dimension, or degree, of ρ, and we will denote it by $\dim(\rho)$. For each V and each G there usually exist many different representations ρ, and we can decompose V into the direct sum of special subspaces. These subspaces play a special role, and because they can't be split up any further, they are called *irreducible* representations. We have to define this term a little more precisely.

Definition 3.2 *A subspace W of V is* invariant *under G if $gw \in W$ for all $g \in G$ and $w \in W$.*

Definition 3.3 *A representation V is called* irreducible *if there is no proper nonzero invariant subspace of V.*

A finite group G admits only finitely many irreducible representations up to isomorphism: the number of irreducible representations of G is equal to the number c of conjugacy classes.

A vector space can be split up into a direct sum of subspaces in several different ways. In the following section we will derive a special decomposition of the vector space V which will allow us later to reduce the cost of a KL decomposition of data in V. To derive this decomposition we will need a few more definitions and intermediate results. Very important is Schur's lemma.

Theorem 3.4 *Let V and W be irreducible representations of G, and $f : V \to W$ a G-equivariant, linear transformation. Then:*

(a) If $f \neq 0$, then f is an isomorphism.

(b) If $V = W$ and f has an eigenvalue λ, then it follows that $f = \lambda \cdot id_V$.

Suppose that W is a representation of G and $L : W \to W$ is a linear transformation which is G-equivariant, that is $L(gw) = gL(w)$ for every $g \in G$ and $w \in W$. If W is an irreducible representation, Schur's Lemma implies that every G-equivariant map is a multiple of the identity.

3.2 Projectors, and Transfer Operators

For each irreducible representation ρ we write $A^{\rho}(g)$ for the matrix $\rho(g)$: its entries are then $A^{\rho}_{ij}(g)$, for $1 \leq i, j \leq \dim(\rho)$. We have the following basic orthogonality relations for the entries of these A^{ρ} matrices:

$$\frac{1}{|G|}\sum_{g\in G} A^{\rho}_{ij}(hg)A^{\tau}_{k\ell}(g^{-1}) = \frac{1}{\dim(\rho)}A^{\rho}_{i\ell}(h)\delta_{jk}\delta_{\rho\tau} \tag{3.5}$$

(For a proof, see [Se] or [CR].)

Definition 3.6 *Define, for any representation V of G, the operator*

$$P^{\rho}_{ij} = \frac{\dim(\rho)}{|G|}\sum_{g\in G} A^{\rho}_{ji}(g^{-1})g.$$

We define $V^{\rho}_{i} = \mathrm{Image}(P^{\rho}_{ii})$ for $1 \leq i \leq \dim(\rho)$ and $V^{\rho} = \sum_{i=1}^{\dim(\rho)} V^{\rho}_{i}$.

As we will see in the following definition, for a fixed ρ all symmetry subspaces have the same dimension. This dimension is also called the *multiplicity* of ρ in V: $\dim(V^{\rho}_{i}) = d_{\rho}$.

The next Proposition now follows completely formally from the orthogonality relations:

Proposition 3.7 *The mappings $P^{\rho}_{ij} : V \to V$ have the following properties:*

(a) For ρ and τ irreducible representations of G, $P^{\rho}_{ij} \circ P^{\tau}_{kl} = P^{\tau}_{il}\delta_{\rho\tau}\delta_{jk}$.

(b) The operators $\{P^{\rho}_{ii}\}$ are orthogonal projectors.

(c) The sum of these projectors is the identity on V: $\sum_{\rho}\sum_{i=1}^{\dim(\rho)} P^{\rho}_{ii} = I$.

(d) The space V splits as the direct sum $V = \bigoplus_{\rho}\bigoplus_{i=1}^{\dim(\rho)} V^{\rho}_{i}$.

(e) *For $i \neq j$ the operator P^{ρ}_{ij} maps V^{ρ}_{j} isomorphically onto V^{ρ}_{i}, and is zero on V^{τ}_{k} if $\tau \neq \rho$ or $k \neq j$.*

(f) *For each ρ the sum $V^{\rho} = \sum_{i=1}^{\dim(\rho)} V^{\rho}_{i}$ is a direct sum, and is a G-invariant subspace of V, isomorphic to the irreducible representation ρ taken exactly $\dim(V^{\rho}_{1})$ times.*

We will say that a basis for V is *G-adapted* if it is formed by taking, for each ρ, a basis $\{v^{\rho}_{1,\alpha}\}$ for V^{ρ}_{1}, and applying the transfer operators P^{ρ}_{j1} for each j with $2 \leq j \leq \dim(\rho)$ to obtain bases $\{v^{\rho}_{j,\alpha} = P^{\rho}_{j1}(v^{\rho}_{1,\alpha})\}$ for V^{ρ}_{j}, and then concatenating these bases to obtain a basis $\{v^{\rho}_{j,\alpha}\}_{\rho,j,\alpha}$ for V.

Definition 3.8 *A basis is called G-adapted if it is constructed by the following process:*

1. *For each ρ find an arbitrary basis $\{v^{\rho}_{1,\alpha}\}$ for V^{ρ}_{1}.*
2. *Apply the transfer operators P^{ρ}_{j1} for all j with $2 \leq j \leq \dim(\rho)$ to each basis vector and thereby obtain bases $\{v^{\rho}_{j,\alpha} = P^{\rho}_{j1}(v^{\rho}_{1,\alpha})\}$ for V^{ρ}_{j}.*
3. *Collect these bases in one set to obtain a basis for V: $\{v^{\rho}_{j,\alpha}\}_{\rho,j,\alpha}$.*

We may take each irreducible representation to be unitary, so that $\rho(g)^{*}\rho(g) = I$ for every g; we will assume this in what follows. If we choose, in each subspace V^{ρ}_{1}, the basis $\{v^{\rho}_{1,\alpha}\}$ to be orthonormal, and the irreducible representations of G are all taken to be unitary, then an easy computation shows that the entire basis for V constructed in this way is orthonormal, so that $(v^{\rho}_{i,\alpha})^{*}v^{\tau}_{j,\beta} = \delta_{\rho\tau}\delta_{ij}\delta_{\alpha\beta}$.

In this case we also have that $(P^{\rho}_{ij}(v))^{*} = v^{*}P^{\rho}_{ji}$, considering the operators as matrices.

Chapter 4

KL Decomposition with Symmetry

4.1 KL Decomposition and Symmetry

Let V be a complex vector space and G a finite group of order $|G|$ acting on V. Moreover let us assume that the action is *unitary*, so that $g^*g = I$ for every $g \in G$. (We will usually suppress the notation for the mapping of G into $GL(V)$ and simply consider each $g \in G$ as a unitary matrix operator on V.)

Let V be the ambient space for a set of patterns $\mathcal{U}$: we think of $\mathcal{U}$ as the set of all possible patterns coming from the data being collected. The set $\mathcal{U}$ is typically a submanifold lying inside the vector space V.

Our symmetry hypothesis is that $\mathcal{U}$ is invariant under the action of G on V: for every $g \in G$ and $u \in \mathcal{U}$, the translate $g \cdot u$ is also in $\mathcal{U}$. Note that this symmetry assumption is *not* that each pattern in $\mathcal{U}$ is G-invariant; only that a translate of a pattern is again a possible pattern.

Suppose we have a finite set of N patterns $U = \{u_i\}_{i=1}^{N}$ which lie in $\mathcal{U}$. We want to exploit the group action and the symmetry assumption to extract information about the KL basis for U. This basis is given by the eigenvectors of the ensemble average covariance matrix $C = \frac{1}{N}\sum_{i=1}^{N} u_i u_i^*$.

Every pattern in U gives rise to $|G|$ other patterns by translating by the elements of the group; all of these patterns also will lie in the pattern space $\mathcal{U}$ by our symmetry assumption. We "symmetrize" the data by collecting these $N|G|$ patterns in the set $\tilde{U}$:

$$\tilde{U} = \{g \cdot u \mid g \in G, u \in U\}.$$

A basis for the KL subspace based on this new set of patterns is given by the eigenvectors of the ensemble average covariance matrix $\tilde{C}$ which is defined using the enlarged set of patterns $\tilde{U}$:

$$\tilde{C} = \frac{1}{N|G|} \sum_{g \in G} \sum_{u \in U} (g \cdot u)(g \cdot u)^* = \frac{1}{N|G|} \sum_{g \in G} \sum_{u \in U} (guu^* g^*) = \frac{1}{|G|} \sum_{g \in G} gCg^*$$

The matrix $\tilde{C}$ is a $d \times d$ matrix, and as such can be viewed as an operator on the vector space V. We have the following results for $\tilde{C}$.

Lemma 4.1 *The matrix $\tilde{C}$ is G-equivariant.*

Proof: For $a \in G$, we simply unravel the composition $a\tilde{C}$:

$$\begin{aligned} a\tilde{C} &= \frac{1}{|G|} \sum_{g \in G} agCg^* = \frac{1}{|G|} \sum_{h \in G} hC(a^{-1}h)^* \quad \text{(setting } h = ag) \\ &= \frac{1}{|G|} \sum_{h \in G} hCh^*(a^{-1})^* = \tilde{C}a \quad \text{(since } a \text{ is unitary).} \end{aligned}$$

qed

Theorem 4.2

(a) The operator $\tilde{C}$ commutes with the projectors and transfer operators P^{ρ}_{ij}, and therefore each subspace V^{ρ}_i is $\tilde{C}$-invariant. Hence if we change to a G-adapted basis for V, the matrix $\tilde{C}$ will be in block form, with block B^{ρ}_i for the subspace V^{ρ}_i. In particular, the eigenvectors of $\tilde{C}$ can be taken to lie in the subspaces V^{ρ}_i.

(b) *The blocks B_i^ρ are in fact independent of i: $B_i^\rho = B^\rho$ for each i. A vector $v \in V_j^\rho$ is an eigenvector for $\tilde{C}$ with eigenvalue λ if and only if the transferred vector $P_{ij}^\rho(v) \in V_i^\rho$ is an eigenvector for $\tilde{C}$ with eigenvalue λ.*

(c) *The eigenvalues of $\tilde{C}$ on $V^\rho = \sum_{i=1}^{\dim(\rho)} V_i^\rho$ each occur with multiplicity equal to a multiple of* $\dim(\rho)$.

Proof:

(a) The projectors and transfer operators $P_{ij}^\rho = \frac{dim(\rho)}{|G|}\sum_{g\in G} A_{ji}^\rho(g^{-1})g$ are linear combinations of the elements in G. By lemma 4.1, $\tilde{C}$ commutes with every group element, and thus with a linear combination of group elements. Since projectors and transfer operators are linear combinations of group elements, $\tilde{C}$ commutes with them.

Since $V_j^\rho = P_{jj}^\rho(V)$ it follows: for each $v \in V_j^\rho$ we have $v = P_{jj}^\rho(v)$. This implies the following chain of equalities:

$$\tilde{C}v = \tilde{C}P_{jj}^\rho(v) = P_{jj}^\rho\tilde{C}v = P_{jj}^\rho(\tilde{C}v).$$

Thus V_j^ρ is $\tilde{C}$-invariant.

(b) This claim can be proved by a simple computation. Let $v \in V_j^\rho$ and $\tilde{C}v = \lambda v$, i.e., v is an eigenvector of $\tilde{C}$ with eigenvalue λ. Applying the transfer operator P_{ij}^ρ to both sides of the equation and using the commutativity of P_{ij}^ρ with $\tilde{C}$, we have

$$\tilde{C}P_{ij}^\rho(v) = P_{ij}^\rho\tilde{C}(v) = P_{ij}^\rho(\lambda v) = \lambda P_{ij}^\rho(v).$$

To see that the blocks B_i^ρ are independent of i, we notice that the $\alpha\beta$-entry of

B_i^ρ is equal to

$$\begin{aligned}(v_{i\alpha}^\rho)^* \tilde{C} v_{i\beta}^\rho &= P_{i1}^\rho (v_{1\alpha}^\rho)^* \tilde{C} P_{i1}^\rho (v_{1\beta}^\rho) \\ &= (v_{1\alpha}^\rho)^* P_{1i}^\rho \tilde{C} P_{i1}^\rho v_{1\beta}^\rho \\ &= (v_{1\alpha}^\rho)^* \tilde{C} P_{1i}^\rho P_{i1}^\rho v_{1\beta}^\rho \\ &= (v_{1\alpha}^\rho)^* \tilde{C} P_{11}^\rho v_{1\beta}^\rho \\ &= (v_{1\alpha}^\rho)^* \tilde{C} v_{1\beta}^\rho .\end{aligned}$$

But this is exactly the $\alpha\beta$-entry of B_1^ρ.

(c) The last claim follows directly from (b). If λ is an eigenvalue of $\tilde{C}$ with eigenvectors in V_i^ρ and multiplicity m, then λ appears in each V_j^ρ as an eigenvalue with multiplicity m. Altogether, λ has multiplicity $m \cdot \dim(\rho)$ in V^ρ.

qed

We note the following relationship between the eigenvalues of the two relevant matrices.

Lemma 4.3 *The sum of eigenvalues of the ensemble average covariance matrix C is the same as the sum of eigenvalues of the enlarged ensemble average covariance matrix $\tilde{C}$; i.e.* $\mathrm{Tr}(C) = \mathrm{Tr}(\tilde{C})$.

Proof: Since $\tilde{C} = \frac{1}{|G|} \sum_{g \in G} gCg^*$, where we again view each $g \in G$ as a unitary matrix operator on V, we have $g^* = g^{-1}$ and therefore gCg^* and C are similar matrices for each $g \in G$. Hence they have the same trace:

$$\begin{aligned}
\mathrm{Tr}(\tilde{C}) &= \mathrm{Tr}(\frac{1}{|G|}\sum_{g\in G} gCg^*) \\
&= \frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}(gCg^*) \\
&= \frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}(C) \\
&= \mathrm{Tr}(C).
\end{aligned}$$

qed

4.2 The Block Matrices B^ρ

Let us now return to the block matrices B^ρ introduced in Theorem 4.2, because it is our goal to find the eigenvalues and eigenvectors of these blocks. To do this we want to find the most simple way of writing the entries of these matrices.

We have seen that the $\alpha\beta$ entry of B^ρ is

$$(B^\rho)_{\alpha\beta} = (v^\rho_{1\alpha})^*\tilde{C}v^\rho_{1\beta} = \frac{1}{|G||U|}\sum_{g,u}((v^\rho_{1\alpha})^*gu)((v^\rho_{1\beta})^*gu)^*.$$

The numbers $(v^\rho_{1\alpha})^*gu$ are exactly the coordinates, in the G-adapted orthonormal basis, of the vectors gu, projected into V^ρ_1. Therefore, viewing P^ρ_{11} as a map from V to V^ρ_1, we have that

$$B^\rho = \frac{1}{|G||U|}\sum_{g,u} P^\rho_{11}(gu)P^\rho_{11}(gu)^* \tag{4.4}$$

and so B^ρ can be viewed as the covariance matrix of the projected symmetrized data. The following Lemma will be useful.

Lemma 4.5 *Let ρ be an irreducible unitary representation of G, with matrices $A^\rho_{ij}(g)$ for $g \in G$.*

(a) For $v \in V$ and $g \in G$ we have $P^\rho_{ij}(gv) = \sum_k A^\rho_{jk}(g) P^\rho_{ik}(v)$.

(b) For $v \in V$ we have $\frac{1}{|G|} \sum_{g \in G} P^\rho_{11}(gv) P^\rho_{11}(gv)^* = \frac{1}{\dim(\rho)} \sum_k P^\rho_{1k}(v) P^\rho_{1k}(v)^*$.

Proof: To prove (a), we see that

$$\begin{aligned} P^\rho_{ij}(gv) &= \frac{1}{|G|} \sum_{a \in G} A^\rho_{ji}(a^{-1}) agv = \frac{1}{|G|} \sum_{b \in G} A^\rho_{ji}(gb^{-1}) bv \quad (\text{setting } b = ag) \\ &= \frac{1}{|G|} \sum_{b \in G} \sum_k A^\rho_{jk}(g) A^\rho_{ki}(b^{-1}) bv \\ &= \sum_k A^\rho_{jk}(g) \frac{1}{|G|} \sum_{b \in G} A^\rho_{ki}(b^{-1}) bv = \sum_k A^\rho_{jk}(g) P^\rho_{ik}(v). \end{aligned}$$

Now (b) follows since by (a) we have $P^\rho_{11}(gv) = \sum_k A^\rho_{1k}(g) P^\rho_{1k}(v)$ and so

$$\begin{aligned} \frac{1}{|G|} \sum_{g \in G} P^\rho_{11}(gv) P^\rho_{11}(gv)^* &= \frac{1}{|G|} \sum_{g \in G} \sum_{k,\ell} A^\rho_{1k}(g) P^\rho_{1k}(v) P^\rho_{1\ell}(v)^* \overline{A^\rho_{1\ell}(g)} \\ &= \sum_{k,\ell} \left[\frac{1}{|G|} \sum_{g \in G} A^\rho_{1k}(g) A^\rho_{\ell 1}(g^{-1}) \right] P^\rho_{1k}(v) P^\rho_{1\ell}(v)^* \\ &= \frac{1}{\dim(\rho)} \sum_k P^\rho_{1k}(v) P^\rho_{1k}(v)^*, \end{aligned}$$

since in the second line the bracketed expression equals $\frac{1}{\dim(\rho)} \delta_{k\ell}$ using (3.5). qed

Combining (4.4) with the previous Lemma implies that

$$B^\rho = \frac{1}{\dim(\rho)|U|} \sum_{u \in U} \sum_{k=1}^{\dim(\rho)} P^\rho_{1k}(u) P^\rho_{1k}(u)^*. \tag{4.6}$$

We note that this formula for B^ρ is more efficient than the formula (4.4): the sum is over the dimension of the representation ρ instead of over all the group elements. In order to use this to compute the entries of the matrix for B^ρ, we have the following lemma.

Lemma 4.7 *Let $\{v^\rho_{k\beta}\}$ be a G-adapted basis for V^ρ (so that $\{v^\rho_{1\beta}\}$ is an orthonormal basis for V^ρ_1). Then the entries $b^\rho_{\alpha\beta}$ of B^ρ in this basis are*

$$b^\rho_{\alpha\beta} = \frac{1}{\dim(\rho)|U|}\sum_u \sum_{k=1}^{\dim(\rho)} (v^\rho_{k\alpha})^* u u^* v^\rho_{k\beta}.$$

Proof: Since $\{v^\rho_{1\beta}\}$ is an orthonormal basis for V^ρ_1, the α-β entry of B^ρ is

$$b^\rho_{\alpha\beta} = (v^\rho_{1\alpha})^* B^\rho v^\rho_{1\beta} = \frac{1}{\dim(\rho)|U|}\sum_u \sum_{k=1}^{\dim(\rho)} (v^\rho_{1\alpha})^* P^\rho_{1k}(u) P^\rho_{1k}(u)^* v^\rho_{1\beta} \tag{4.8}$$

using (4.6). For any γ, we have

$$\begin{aligned}
(v^\rho_{1\gamma})^* P^\rho_{1k}(u) &= \frac{\dim(\rho)}{|G|}\sum_{g\in G} A^\rho_{k1}(g^{-1})(v^\rho_{1\gamma})^* g u \\
&= \frac{\dim(\rho)}{|G|}\sum_{g\in G} \overline{A^\rho_{1k}(g)}(g^* v^\rho_{1\gamma})^* u \\
&= \frac{\dim(\rho)}{|G|}\sum_{g\in G} \overline{A^\rho_{1k}(g) u^* (g^* v^\rho_{1\gamma})} \\
&= \overline{u^* \frac{\dim(\rho)}{|G|}\sum_{a\in G} A^\rho_{1k}(a^{-1}) a v^\rho_{1\gamma}} \quad \text{setting } a = g^* \\
&= \overline{u^* P^\rho_{k1}(v^\rho_{1\gamma})} \\
&= \overline{u^* v^\rho_{k\gamma}} \qquad \text{since the basis is } G\text{-adapted} \\
&= (v^\rho_{k\gamma})^* u
\end{aligned}$$

so that $(v^\rho_{1\alpha})^* P^\rho_{1k}(u) = (v^\rho_{k\alpha})^* u$ and $P^\rho_{1k}(u)^* v^\rho_{1\beta} = u^* v^\rho_{k\beta}$. Plugging these into the equation (4.8) gives the result. **qed**

The previous Lemma motivates the following notation. For each pattern u, define the column vector u_k^ρ (whose dimension is equal to $\dim V_1^\rho$) by setting the α coordinate equal to

$$(u_k^\rho)_\alpha = (v_{k\alpha}^\rho)^* u$$

where $\{v_{k\alpha}^\rho\}$ is the given G-adapted basis for V. Then Lemma 4.7 exactly says that

$$B^\rho = \frac{1}{\dim(\rho)|U|} \sum_u \sum_{k=1}^{\dim(\rho)} (u_k^\rho)(u_k^\rho)^*. \tag{4.9}$$

4.3 Algorithm for Exploiting Symmetry in the KL Basis Computation

With the formulation from the previous sections we can now give a precise algorithm for exploiting the symmetry assumption in determining a KL basis for V.

Algorithm for Exploiting Symmetry
in the KL basis computation:

Given: A complex vector space V of dimension d, a set of patterns U drawn from a pattern subset $\mathcal{U}$, and a group G acting unitarily on V satisfying $g \cdot u \in \mathcal{U}$ for all $g \in G$ and $u \in U$.

1. Determine the irreducible unitary representations ρ of G, and a G-adapted basis $\{v^\rho_{k\alpha}\}$ for V. Let d_ρ be the dimension of V^ρ_1.

2. For each ρ, each $k = 1, \ldots, \dim(\rho)$, and each pattern $u \in U$, form the vector u^ρ_k of dimension d_ρ by $(u^\rho_k)_\alpha = (v^\rho_{k\alpha})^* u$.

3. For each ρ, form the $d_\rho \times d_\rho$ matrix B^ρ of $\tilde{C}|_{V^\rho_1}$ by $B^\rho = \frac{1}{\dim(\rho)|U|} \sum_u \sum_{k=1}^{\dim(\rho)} (u^\rho_k)(u^\rho_k)^*$.

4. For each ρ, find the d_ρ eigenvalues λ^ρ_α and corresponding eigenvectors w^ρ_α (for $1 \leq \alpha \leq d_\rho$) of the matrix B^ρ. Write each eigenvector as $w^\rho_\alpha = (w^\rho_{\alpha 1} \cdots w^\rho_{\alpha d_\rho})^T$.

5. Define

$$z^\rho_{k\alpha} = \sum_\beta w^\rho_{\alpha\beta} v^\rho_{k\beta}$$

for each ρ, each $k = 1, \ldots, \dim(\rho)$, and each $\alpha = 1, \ldots, d_\rho$. The KL basis for V consists of these vectors $z^\rho_{k\alpha}$, each with eigenvalue λ^ρ_α (independent of k).

We note that Step 1 is to be considered as the overhead of the method. which can be done once. and subsequently applied to many different KL problems in the space V. The power of the method is that it replaces the eigenvector computation for the large $d \times d$ matrix $\tilde{C}$ with several eigenvector computations of the smaller matrices B^{ρ}.

4.4 The Computational Efficiency

In order to calculate the efficiency of the proposed method. we assume that a general method such as QR is used to compute the eigenvectors and eigenvalues of symmetric (or Hermitian) matrices: for a $k \times k$ matrix. these methods are typically of order k^3. that is. the number of flops required to find the eigenvectors and eigenvalues is approximately ak^3 for some constant a (see [Ci] for example). Therefore in our original problem. the cost of directly computing the KL basis is ad^3. since d is the dimension of the ambient space V.

The dimension of the symmetry subspaces V_k^{ρ} for each irreducible representation depends on the precise nature of the representation of the group G on the space V. However it is quite typical that this representation is a direct sum of several copies of the regular representation. or varies from this in a minor way often due to fixed points in a permutation representation. For the regular representation. which is of dimension $|G|$. the dimension of the first symmetry subspace V_1^{ρ} is $\dim(\rho)$. Therefore if V is approximately a direct sum of regular representations, then the dimension of V_1^{ρ} will be approximately $d_{\rho} = d \cdot \dim(\rho)/|G|$. Hence the cost of finding the eigenvectors in this symmetry subspace will be approximately $ad^3 \dim(\rho)^3/|G|^3$.

Solving this sub-problem for each ρ and summing, we see that the approximate ratio of the cost of solving the original problem to the cost of using the method outlined above is

$$\alpha(G) = \frac{ad^3}{ad^3 \sum_\rho \dim(\rho)^3/|G|^3} = \frac{|G|^3}{\sum_\rho \dim(\rho)^3}.$$

For a cyclic group C_m of order m, all irreducible representations ρ are one–dimensional, and there are m of them. Hence $\alpha(C_m) = m^2$.

For a dihedral group $D_{m/2}$ of order m (where m is even), all irreducible representations are either one– or two–dimensional; there are at most 4 representations of degree one and approximately $m/4$ representations of degree two, the precise numbers depending on the value of m modulo 4. Therefore $\alpha(D_{m/2}) \approx m^2/2$. For the group D_4 of symmetries of the square, the precise factor is $128/3 \approx 42$.

Finally we look at the group O of symmetries of the cube; it has order 48 and there are 4 irreducible representations of degree one, 2 of degree 2, and 4 of degree 3. Therefore

$$\alpha(O) = \frac{48^3}{4 \cdot 1^3 + 2 \cdot 2^3 + 4 \cdot 3^3} = 110592/128 = 864.$$

The other extreme situation is when each $d_\rho = 1$. In this case each submatrix B^ρ is simply a number, equal to the eigenvalue λ_1^ρ, and no eigenvector computations are required: the eigenvectors are exactly vectors $\{v_{k\alpha}^\rho\}$ of the G-adapted basis. This very special situation actually occurs more often than one might expect, especially for compact Lie groups.

Finally we remark that this method is useful when the number of patterns N is much larger than the dimension of the ambient space V.

Chapter 5

Example

In this section we present an example using data generated by a 7-dimensional ODE system which exhibits S_3 symmetry. The ODE system we consider has variables x_i, y_i for $i = 1, 2, 3$ and p; the system is given by

$$\begin{aligned}
\dot{x}_i &= \sigma(y_i - x_i) \\
\dot{y}_i &= \tau x_i - y_i - p x_i \\
\dot{p} &= -bp + \frac{a}{3}\sum_{i=1}^{3} x_i y_i.
\end{aligned}$$

This system was kindly suggested to us by Gerhard Dangelmayr. We use as parameters $\sigma = 10$, $\tau = 28$, $b = \frac{8}{3}$, and $a = 1$, which gives a symmetry-adapted Lorenz-type system exhibiting a strange attractor with a chaotic trajectory. We hope that the chaotic nature of the trajectory produces data which "fills up" the attractor, allowing us to use one trajectory to produce a good approximation to the attractor.

The group S_3 acts simply by permuting the indices of the x_i's and y_i's; the variable p is fixed. It is clear that the ODE is preserved by this action, and therefore solutions to the ODE are carried into other solutions.

After choosing an initial condition randomly, a trajectory is generated for each time step t; this gives a 7-dimensional vector $\underline{x}(t)$ for each time step. A fourth-order Runge-Kutta scheme was used to generate 5000 time steps; the first 1000 were discarded, producing a data set U of 4000 vectors in the 7-dimensional space V.

S_3 has three irreducible representations: the one-dimensional trivial representation W_1, the one-dimensional alternating representation W_2, and the two-dimensional standard representation W_3. The alternating representation does not occur in our 7-dimensional representation V: the trivial representation occurs 3 times, and the standard occurs twice. Hence there are three symmetry subspaces: V_1^1 is 3-dimensional, V_1^3 and V_2^3 are two-dimensional. G-adapted bases for these subspaces are:

$$V_1^1 : \; v_{11}^1 = \frac{\sqrt{3}}{3}(1,1,1,0,0,0,0)^\top, v_{12}^1 = \frac{\sqrt{3}}{3}(0,0,0,1,1,1,0)^\top, v_{13}^1 = (0,0,0,0,0,0,1)^\top$$
$$V_1^3 : \; v_{11}^3 = \frac{\sqrt{6}}{6}(2,-1,-1,0,0,0,0)^\top, v_{12}^3 = \frac{\sqrt{6}}{6}(0,0,0,2,-1,-1,0)^\top$$
$$V_2^3 : \; v_{21}^3 = \frac{\sqrt{2}}{2}(0,1,-1,0,0,0,0)^\top, v_{22}^3 = \frac{\sqrt{2}}{2}(0,0,0,0,1,-1,0)^\top$$

We find the 3×3 matrix $B^{(1)}$ and the 2×2 matrix $B^{(3)}$ for the two symmetry subspaces to be:

$$B^{(1)} = \begin{pmatrix} 160.596 & 160.799 & 2.393 \\ 160.799 & 203.127 & 1.063 \\ 2.393 & 1.063 & 62.959 \end{pmatrix} \quad B^{(3)} = \begin{pmatrix} 4.868 & 4.874 \\ 4.874 & 6.15 \end{pmatrix}$$

The eigenvalues and eigenvectors of $B^{(1)}$ are:

$$344.08 : w_1^1 = (0.659, 0.752, 0.008)^\top$$

$$62.97 : w_2^1 = (-0.0135, 0.023, -1)^\top$$

$$19.63 : w_3^1 = (0.7519, -0.6589, -0.02537)^\top$$

The eigenvalues and eigenvectors of $B^{(3)}$ are:

$$10.42 : w_1^3 = (0.6594, 0.7518)^\top$$

$$0.59 : w_2^3 = (0.7518, -0.6594)^\top$$

We can now form the resulting set of eigenvectors of $\tilde{C}$ by using the eigenvectors in the last step together with the G-adapted basis for V. We list the eigenvalues with their eigenvectors.

$$
\begin{aligned}
344.08 : z_{11}^1 &= w_{11}^1 v_{11}^1 + w_{12}^1 v_{12}^1 + w_{13}^1 v_{13}^1 \\
&= (0.38, 0.38, 0.38, 0.434, 0.434, 0.434, 0.008)^T \\
62.96 : z_{12}^1 &= w_{21}^1 v_{11}^1 + w_{22}^1 v_{12}^1 + w_{23}^1 v_{13}^1 \\
&= (-0.0078, -0.0078, -0.0078, 0.013, 0.013, 0.013, -1)^T \\
19.63 : z_{13}^1 &= w_{31}^1 v_{11}^1 + w_{32}^1 v_{12}^1 + w_{33}^1 v_{13}^1 \\
&= (0.434, 0.434, 0.434, -0.38, -0.38, -0.38, -0.025)^T \\
10.42 : z_{11}^3 &= w_{11}^3 v_{11}^3 + w_{12}^3 v_{12}^3 = (0.538, -0.269, -0.269, 0.614, -0.307, -0.307, 0)^T \\
z_{21}^3 &= w_{11}^3 v_{21}^3 + w_{12}^3 v_{22}^3 = (0, 0.466, -0.466, 0, 0.532, -0.532, 0)^T \\
0.59 : z_{12}^3 &= w_{21}^3 v_{11}^3 + w_{22}^3 v_{12}^3 = (0.614, -0.307, -0.307, -0.538, 0.269, 0.269, 0)^T \\
z_{22}^3 &= w_{21}^3 v_{21}^3 + w_{22}^3 v_{22}^3 = (0, 0.532, -0.532, 0, -0.466, 0.466, 0)^T
\end{aligned}
$$

Finding this set of vectors was our goal because they provide the full KL-basis for the enlarged pattern set $\tilde{U}$ (consisting of 24000 vectors) and the 7×7 enlarged ensemble average covariance matrix $\tilde{C}$. For completeness, this matrix is presented below; all computations have been done with six digit precision, the results were then rounded to three decimal places of accuracy.

$$
\tilde{C} = \begin{pmatrix}
56.778 & 51.909 & 51.909 & 56.849 & 51.975 & 51.975 & 1.382 \\
51.909 & 56.778 & 51.909 & 51.975 & 56.849 & 51.975 & 1.382 \\
51.909 & 51.909 & 56.777 & 51.975 & 51.975 & 56.849 & 1.382 \\
56.849 & 51.975 & 51.975 & 71.809 & 65.659 & 65.659 & .614 \\
51.975 & 56.849 & 51.975 & 65.659 & 71.809 & 65.659 & .614 \\
51.975 & 51.975 & 56.849 & 65.658 & 65.658 & 71.809 & .614 \\
1.382 & 1.382 & 1.382 & .614 & .614 & .614 & 62.96
\end{pmatrix}
$$

We note that 95% of the energy of the data is concentrated in the eigenspaces of the three largest eigenvalues.

Let us now compare the eigenvalues and eigenvectors of C (which is the ensemble averaged covariance matrix based on the small set of patterns U), and $\tilde{C}$. We have:

$$C = \begin{pmatrix} 26.172 & 37.429 & 48.691 & 26.205 & 37.476 & 48.752 & .966 \\ 37.429 & 53.528 & 69.633 & 37.476 & 53.595 & 69.721 & 1.382 \\ 48.691 & 69.633 & 90.584 & 48.752 & 69.721 & 90.699 & 1.797 \\ 26.205 & 37.476 & 48.752 & 33.103 & 47.341 & 61.585 & .429 \\ 37.476 & 53.595 & 69.721 & 47.341 & 67.703 & 88.074 & .614 \\ 48.752 & 69.721 & 90.699 & 61.585 & 88.074 & 114.573 & .799 \\ .966 & 1.382 & 1.797 & .429 & .614 & .799 & 62.96 \end{pmatrix}$$

The non zero eigenvalues and their eigenvectors of C are:

$$364.836 : y_1 = (0.258, 0.37, 0.481, 0.295, 0.422, 0.548, 0.008)^T$$

$$62.97 : y_2 = (-0.006, -0.008, -0.011, 0.009, 0.013, 0.017, -1)^T$$

$$20.817 : y_3 = (0.295, 0.422, 0.548, -0.258, -0.369, -0.48, -0.027)^T$$

This information tells us that all the generated patterns lie in the three dimensional subspace generated by the eigenvectors above. When we look at the enlarged pattern set, we would expect that the same thing is true, and we have indeed seen that the energy *is* concentrated in a 3-dimensional subspace. The final step is to compare these two 3-dimensional subspaces.

A comparison of the eigenvectors of these two matrices show that they span essentially the same three dimensional subspace of $\mathbb{R}^7$. Taking the dot-product of the eigenvectors makes this clear.

$$(y_i \cdot z^j_{k\ell}) = \begin{pmatrix} 0.971 & 0.0001 & 0 & -0.207 & -0.119 & 0 & 0 \\ -0.001 & 1 & -0.001 & -0.002 & -0.001 & 0.006 & 0.003 \\ 0 & 0.002 & 0.971 & 0 & 0 & -0.206 & -0.119 \end{pmatrix}$$

In a perfect system, the above 7×3 matrix would have a 3×3 identity block at the left, and zeroes elsewhere. That this is not the case may be due to several factors: the particular trajectory used may not fill up the attractor as well as desired, and the Runga-Kutta scheme used to compute the trajectory will introduce some errors also. However, the agreement between the two 3-spaces is convincingly close.

Part II

Compact Lie Groups

In the second part of this thesis we are going to investigate continuous data. such as continuous functions which are defined on manifolds. This kind of data is very important in natural science. since most of the processes one encounters in nature. or one wishes to model with differential equations. are continuous and not discrete. Thus. data such as temperature. barometric pressure. physical fields. and solutions to differential equations are continuous functions and not finite dimensional vectors.

Symmetry groups which are of interest when working with such data are often compact Lie groups. They differ from groups we have encountered in Part I of this paper because they not only possess a group structure but also a topological structure.

The theory we developed in Part I can be extended to the case of compact Lie groups with relatively little change. In Part II we will start with making the appropriate changes in the theory and then continue to demonstrate the method using examples of data on the sphere. The compact Lie group acting on the data is the group of rigid rotations in $\mathbb{R}^3$. the special orthogonal group SO(3).

Chapter 6

Theory

6.1 KL Decomposition

In many cases we are interested in data that are not just vectors in finite dimensional vector spaces but continuous functions. To investigate this kind of data with the KL analysis we can either discretize the data or use the continuous version of the KL decomposition.

Our functions are typically defined on a compact manifold M. We will denote the space of continuous $\mathbb{C}$-valued functions by $\mathcal{F}^n(M)$. The standard Hermitian inner product in this space is defined as

$$(f_1, f_2) = \frac{1}{\text{vol}(M)} \int_M f_1(m)^* f_2(m) dm.$$

Let $U \in \mathcal{F}^n(M)$ be a submanifold of data. Under the assumption that all symmetry subspaces of $\mathcal{F}^n(M)$ are finite dimensional, we can apply the same theory as in Part I to this situation to find the KL basis corresponding to the data submanifold U. As we said, the data set is typically a submanifold of data. Thus we have to replace averages over U with integrals where necessary.

The KL decomposition of the submanifold U will produce an orthonormal basis for $\mathcal{F}^n(M)$, which is optimal for arbitrary truncation dimensions. We have to go into some more detail to explain this idea of optimality.

Let $\mathcal{B} = \{v_1, v_2 \ldots\}$ be an ordered orthonormal basis for V. Then we can write each vector $u \in U$ uniquely and without error as a linear combination of the basis vectors:

$$u = \sum_j a_j v_j.$$

If we truncate the sum after $D > 0$ terms we get a D-term approximation of u:

$$u \approx u^{(D)} = \sum_{j \leq D} a_j v_j.$$

with approximation error $e(u)^{(D)} = u - u^{(D)}$.

For the given data set U and the basis $\mathcal{B}$ we define the mean square error of the D-dimensional truncation as

$$\epsilon_{mse}^{(D)} = \frac{1}{\text{vol}(U)} \int_U ||e(u)^{(D)}||^2.$$

where the norm is induced by the standard Hermitian inner product. i.e. $||f||^2 = (f, f)$.

The KL basis is precisely the ordered orthonormal basis which minimizes the mean square error $\epsilon_{mse}^{(D)}$ for each truncation dimension $D \geq 1$. The KL basis can be constructed analogously to the method we described in Part I and can be found in [Fu], [AG], [Ki]. The main result is summarized in the following theorem which is just a reformulation of the analogous theorem from Part I.

Theorem 6.1 *The KL basis is given by the eigenvectors of the ensemble average covariance matrix*

$$C = \frac{1}{\text{vol}(U)} \int_U uu^* du.$$

6.2 Representation Theory

The representation theory of compact Lie groups is very similar to the theory of finite groups, and both are often discussed together in the literature of this field (see for example [He]).

Even though the theory of finite and compact Lie groups is similar, it is not identical and we will point out some of the differences in this section. While in the case of finite groups we compute the average of a function $f : G \to \mathbb{C}$ over a group G by

$$\mu(f) = \frac{1}{|G|} \sum_{g \in G} f(g).$$

we have to define the averaging process over a compact Lie group differently, and we have to justify that it exists in the first place.

Theorem 6.2 *There exists a unique measure dg on G with*

1. $\int_G f(g)dg = \int f(gg_0)dg$ *for all functions* f *and* $g_0 \in G$.

2. $\int_G dg = 1$

This measure is called "Haar Measure" or "Haar Integral" due to Alfred Haar who published an existence theorem in 1933. For a proof refer to [FV].

Compact groups usually have infinitely many finite dimensional irreducible representations. But Schur's lemma still holds (up to isomorphism).

Theorem 6.3 *Let ρ be a unitary representation of G in the Hilbert space $\mathcal{H}$. Then ρ is irreducible if and only if: $T \in L(\mathcal{H})$ with $T\rho(g) = \rho(g)T$ implies that $T = \lambda \cdot id_{\mathcal{H}}$.*

That is, each G equivariant map on an irreducible representation is a multiple of the identity.

For each finite dimensional continuous irreducible representation ρ we denote by $A^\rho(g)$ the matrix $\rho(g)$ with entries $A^\rho_{ij}(g)$. The basic orthogonality relations (3.5) still hold in this context:

$$\int_G A^\rho_{ij}(hg)A^\tau_{k\ell}(g^{-1})dg = \frac{1}{\dim(\rho)}A^\rho_{i\ell}(h)\delta_{jk}\delta_{\rho\tau}. \tag{6.4}$$

For each finite dimensional continuous representation V of G we define the operator

$$P^\rho_{ij}(v) = \dim(\rho)\int_G A^\rho_{ji}(g^{-1})g\cdot v \;\; dg$$

as before. We let $V^\rho_i = \text{Image}(P^\rho_{ii})$ for $1 \le i \le \dim(\rho)$ and $V^\rho = \sum_{i=1}^{\dim(\rho)} V^\rho_i$.

The analog of Proposition 3.7 holds unchanged in this situation. Even if G is not finite dimensional, the operators P^ρ_{ij} are still defined and parts (a), (b) and (e) hold. And we still have that (f) holds in the case that all symmetry subspaces of V^ρ are finite dimensional. We also keep the identical definition of a G-adapted basis for V as before.

6.3 KL Decomposition and Symmetry

A typical application of this theory comes up in the case when a compact Lie group G acts on a compact manifold M and thus on the space of continuous $\mathbb{C}^n$-valued functions $\mathcal{F}^n(M)$. This action is defined by $(g\cdot f)(m) = f(g^{-1}m)$. Recall that the standard Hermitian metric on $\mathcal{F}^n(M)$ is defined by

$$(f_1, f_2) = \frac{1}{\text{vol}(M)}\int_M f_1(m)^* f_2(m)\; dm.$$

Recall that $U \subset \mathcal{F}^n(M)$ is a submanifold of data. Under the assumption that the symmetry subspaces of $\mathcal{F}^n(M)$ are all finite-dimensional, the identical theory applies to compute the KL basis for $\mathcal{F}^n(M)$ with respect to the data set U. Since U may be a submanifold of data, we have that sums over U are replaced by integrals: and with these mild alterations the algorithm holds almost without change:

Algorithm for Exploiting Symmetry in the KL Basis Computation in the Compact Lie Group Case

Given: A manifold M, a submanifold of patterns U drawn from the space of functions $\mathcal{F}^n(M)$, and a compact Lie group G acting on M preserving the set U. We assume that each symmetry subspace is finite-dimensional.

1. Determine the irreducible unitary representations ρ of G, and a G-adapted basis $\{v^\rho_{k\alpha}\}$ for $\mathcal{F}^n(M)$. Let d_ρ be the dimension of V^ρ_1.

2. For each ρ and each $k = 1, \dots, \dim(\rho)$, consider the function $\gamma^\rho_{k\alpha} : U \to \mathbb{C}$ defined by

$$\gamma^\rho_{k\alpha}(u) = (v^\rho_{k\alpha}, u) = \frac{1}{\mathrm{vol}(M)} \int_M v^\rho_{k\alpha}(m)^* u(m)\, dm.$$

3. For each ρ, form the $d_\rho \times d_\rho$ matrix B^ρ by

$$(B^\rho)_{ij} = \frac{1}{\dim(\rho)\,\mathrm{vol}(U)} \sum_{k=1}^{\dim(\rho)} \int_U \gamma^\rho_{ki}(u)\overline{\gamma^\rho_{kj}(u)}\, du.$$

4. For each ρ, find the d_ρ eigenvalues λ^ρ_α and corresponding eigenvectors w^ρ_α (for $1 \le \alpha \le d_\rho$) of the matrix B^ρ. Write each eigenvector as $w^\rho_\alpha = (w^\rho_{\alpha 1} \cdots w^\rho_{\alpha d_\rho})^\top$.

5. Define

$$z^\rho_{k\alpha} = \sum_\beta w^\rho_{\alpha\beta} v^\rho_{k\beta}$$

for each ρ, each $k = 1, \dots, \dim(\rho)$, and each $\alpha = 1, \dots, d_\rho$. The KL basis for $\mathcal{F}^n(M)$ consists of these vectors $z^\rho_{k\alpha}$, each with eigenvalue λ^ρ_α (independent of k).

Note that since G has in general infinitely many irreducible representations, the method actually is useful for finding that part of the KL basis which resides in any particular symmetry subspace, i.e., for fixed ρ.

6.4 Example: The circle group S^1

Consider the compact Lie group $G = S^1$. G acts on $M = G$ by translation, and we consider the space $\mathcal{F}$ of continuous functions on M, which of course corresponds to continuous periodic functions on the line.

The irreducible representations of G are all one-dimensional, indexed by integers n: $\rho_n(\theta) = \exp(2\pi i n\theta)$. For each n, the symmetry space is also one-dimensional, and $v_{11}^{(n)}(x) = \exp(-2\pi i n x)$; $d_\rho = d_n = 1$ for every n. Since all dimensions are one, we will drop the subscripts; so the G-adapted basis is $v^{(n)}(x) = \exp(-2\pi i n x)$. The functionals $\gamma^{(n)}$ on $\mathcal{F}$ are

$$\gamma^{(n)}(u) = (v^{(n)}, u) = \frac{1}{2\pi}\int_0^{2\pi} \exp(2\pi i n\theta)u(\theta)\, d\theta.$$

For any pattern set U, the KL basis consists of these $v^{(n)}$'s, and the only unknowns are the eigenvalues. Each matrix $B^\rho = B^n$ is 1×1, with entry the eigenvalue λ_n; by the above, it is

$$\lambda_n = \frac{1}{\mathrm{vol}(U)}\int_U \gamma^{(n)}(u)\overline{\gamma^{(n)}(u)}\, du.$$

This theory then simply reproduces the Fourier Analysis for the set U: the KL basis consists of exactly the Fourier modes $v^{(n)}(x) = \exp(-2\pi i n x)$, with the eigenvalues given above.

Chapter 7

The orthogonal group SO(3)

In this chapter we will consider examples of data which are acted on by the special orthogonal group SO(3). We will investigate the representation theory of SO(3) acting on the space of scalar valued functions on the sphere and also acting on the space of vector fields of polynomials that are tangent to the sphere. We will start by presenting the theory in both cases and finish by demonstrating the use of the algorithm on numerical data gained from solving a vector valued partial differential equation on the sphere.

7.1 SO(3)

We denote by SO(3) the group of 3×3 orthogonal matrices with real entries and determinant one. Such a matrix uniquely represents a rotation in $\mathbb{R}^3$ and can be viewed as a composition of three consecutive rotations about different axes. We can first rotate about the z-axis through angle ϕ_1, then rotate through angle θ about the x-axis, and finally rotate through angle ϕ_2 about the z-axis. The angles ϕ_1, θ, and ϕ_2, are the Euler angles as defined in [GMS]. The ranges for the angles are: $0 \leq \phi_1, \phi_2 \leq 2\pi$, $0 \leq \theta \leq \pi$.

The three rotations are given by the three 3×3 matrices:

$$g_{\phi_1} = \begin{pmatrix} \cos\phi_1 & -\sin\phi_1 & 0 \\ \sin\phi_1 & \cos\phi_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; g_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

$$g_{\phi_2} = \begin{pmatrix} \cos\phi_2 & -\sin\phi_2 & 0 \\ \sin\phi_2 & \cos\phi_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The rotation $g(\phi_1, \theta, \phi_2) = g_{\phi_2} g_\theta g_{\phi_1}$ is obtained by multiplying these matrices in reverse order:

$$\begin{pmatrix} \cos\phi_1\cos\phi_2 - \cos\theta\sin\phi_1\sin\phi_2 & -\cos\phi_1\sin\phi_2 - \cos\theta\sin\phi_1\cos\phi_2 & \sin\phi_1\sin\theta \\ \sin\phi_1\cos\phi_2 + \cos\theta\sin\phi_2\cos\phi_1 & -\sin\phi_1\sin\phi_2 + \cos\theta\cos\phi_1\cos\phi_2 & -\cos\phi_1\sin\theta \\ \sin\phi_2\sin\theta & \cos\phi_2\sin\theta & \cos\theta \end{pmatrix}$$

Let $g \in \mathrm{SO}(3)$ be given in terms of the three Euler angles ϕ_1, θ, ϕ_2. Now fix α such that $2\alpha = \phi_1 + \phi_2$, and let $\beta = \alpha - \phi_2$ (so that $2\beta = \phi_1 - \phi_2$). Note that α is not unique but that there are two choices which differ by π. Thus the choice of α changes both α and β by a difference of π.

The invariant integral or Haar integral for SO(3) is defined the following way:

$$\int f(g)dg = \frac{1}{8\pi^2} \int_0^{2\pi} \int_0^{\pi} \int_0^{2\pi} f(\phi_1, \theta, \phi_2) \sin\theta d\phi_1 d\theta d\phi_2,$$

where g is given in terms of the three Euler angles. Or in terms of α, θ and β we have:

$$\int f(g)dg = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{\pi} \int_{-\pi}^{\pi} f(\alpha, \theta, \beta) \sin\theta d\alpha d\theta d\beta.$$

7.2 Representation of SO(3) - Scalar Functions on the Sphere

7.2.1 The Irreducible Representations

There are different ways to derive the irreducible representations of SO(3). We will have a closer look at two particular ones.

The Spherical Harmonics

The most common way to describe the irreducible representations of SO(3) is in terms of spherical harmonics. These special functions come up as solutions of the Laplace equation in $\mathbb{R}^3$:

$$\Delta u = u_{xx} + u_{yy} + u_{zz} = 0, \qquad u = u(x, y, z).$$

It is a well known result that the irreducible representations of SO(3) in the space of scalar valued functions is realized in the space of spherical harmonics with odd dimension $\dim(V_k) = 2k + 1$ $(k \in \mathbb{N}_0)$, where

$$V_k = \{ \sum_{m=-k}^{k} r^k z_m Y_{km}(\theta, \phi) | z_m \in \mathbb{C}, \bar{z}_m = (-1)^m z_m \}.$$

Here we write $\vec{r} = (r, \theta, \phi)$ in spherical coordinates with $Y_{km}(\theta, \phi) = P_m^k(\cos(\theta))e^{im\phi}$ the spherical harmonics and where P_m^k are the corresponding Legendre polynomials. Since in each dimension $2k+1$ there are exactly $2k+1$ linearly independelnt spherical harmonics, they form a basis for the irreducible representation V_k (see [GSS]).

We can summarize these ideas as the following result:

Theorem 7.1 SO(3) *has exactly one irreducible representation (up to isomorphism) in each odd dimension* $2k + 1$. *The irreducible representations* V_k *are realized in the space of spherical harmonics.*

Thus we have found the KL basis for a data set of scalar-valued SO(3) symmetric functions, namely the spherical harmonics. We still have to compute the corresponding eigenvalues to complete the analysis. We see that the KL basis is independent of the data set we are investigating but the eigenvalues are not.

Since each irreducible representation has multiplicity one, we know that each symmetry subspace is one dimensional. Thus each of the block matrices B^ρ in our algorithm is one-by-one, a number, which has to an eigenvalue of the covariance matrix.

To compute these numbers we need a matrix representation of SO(3). There are many equivalent matrix representations: for each basis there is a corresponding matrix representation. Theoretically we could use the spherical harmonics to compute the matrices, but it becomes apparent quickly, that this approach is not very practical and entails incredibly tedious computations. Some books actually give formulas for the entries of these matrices, but again, they are not very explicit and require considerable effort to use. [GMS].

Because it is so complicated to compute the matrix entries using the basis of spherical harmonics, we will now describe an alternative method using the universal covering space of SO(3), which is the special unitary group SU(2).

The Universal Covering Space SU(2)

The special unitary group SU(2) is the universal covering space of SO(3), i.e. SU(3) is locally isomorphic to SO(3), [He]. An element g in SO(3) can be lifted to a corresponding element $\tilde{g}$ in the covering space SU(2). Recall that SU(2) is the group of 2×2 special unitary matrices with complex entries, i.e., the group of matrices of the form $\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$, with $|u|^2 + |v|^2 = 1$. Suppose that g is given in terms of the three Euler angles ϕ_1, θ, ϕ_2. Fix α such that $2\alpha = \phi_1 + \phi_2$ and set $\beta = \alpha - \phi_2$ (so that $2\beta = \phi_1 - \phi_2$). Note that there are two such α's, differing by π, and that the alternate choice changes both α and β by π. Then the lifting of g to an element in SU(2) is given by:

$$\tilde{g} = \begin{pmatrix} \cos(\frac{\theta}{2})\exp(i\alpha) & i\sin(\frac{\theta}{2})\exp(-i\beta) \\ i\sin(\frac{\theta}{2})\exp(i\beta) & \cos(\frac{\theta}{2})\exp(-i\alpha) \end{pmatrix}.$$

Note that the different choice of α changes the sign of this matrix, reflecting the fact that SO(3) = SU(2)/($\pm I$).

Let V_n be the space of homogeneous polynomials of degree n in X and Y with coefficients in $\mathbb{C}$. For each positive integer n, V_n is an irreducible representation of SU(2) of dimension $n+1$ (see [BD]). For each even n, $\pm I$ acts trivially on V_n, and so V_n is also an irreducible representation of SO(3) of dimension $n+1$.

The Hermitian form on V_n is induced by the standard Hermitian form on V_1. V_1 has the basis $\{X, Y\}$, and the Hermitian form is defined by

$$\langle v_1, w_1\rangle = \langle \alpha_1 X + \alpha_2 Y, \beta_1 X + \beta_2 Y\rangle = \overline{\alpha_1}\beta_1 + \overline{\alpha_2}\beta_2,$$

where $\langle X, X\rangle = \langle Y, Y\rangle = 1$ and $\langle X, Y\rangle = \langle Y, X\rangle = 0$. The induced Hermitian form on V_n is given by:

$$\langle v_1 v_2 \cdots v_n, w_1 w_2 \cdots w_n\rangle = \frac{1}{n!}\sum_{\sigma}\left(\prod_{m=1}^{n}\langle v_m, w_{\sigma(m)}\rangle\right)$$

where v_m and w_m are in V_1 for $m = 1, \ldots, n$, and σ runs through all permutations in S_n. An orthonormal basis for V_n is given by:

$$B = \left\{\sqrt{\binom{n}{m}}X^{n-m}Y^m, 0 \le m \le n\right\}$$

Let $g = \begin{pmatrix} u & v \\ -\overline{v} & \overline{u}\end{pmatrix}$ be an element in SU(2). Then the n^{th} irreducible representation, written in the orthonormal basis B, sends g to a unitary, $(n+1)\times(n+1)$ matrix M_g. The $(m+1),(j+1)$ entry of M_g is given by:

$$(M_g)_{(m+1),(j+1)} = \frac{\sqrt{\binom{n}{m}}}{\sqrt{\binom{n}{j}}}\sum_{l+k=j}\binom{n-m}{k}\binom{m}{l}(-\overline{v})^{m-l}v^k u^{n-m-k}(\overline{u})^l$$

where $0 \le m, j \le n$.

Now let g be the lifting of an element in SO(3) as shown above. i.e. $u = \cos(\frac{\theta}{2})\exp(i\alpha)$ and $v = i\sin(\frac{\theta}{2})\exp(i\beta)$. The corresponding matrix M_g has its $(m+1), (j+1)$ entry equal to

$$(M_g)_{(m+1),(j+1)} = \frac{\sqrt{\binom{n}{m}}}{\sqrt{\binom{n}{j}}}\exp(-i\beta(-m+j))\exp(i\alpha(n-m-j))T_{(m+1),(j+1)}(\theta)$$

where

$$T_{(m+1),(j+1)}(\theta) = \sum_{l+k=j}\binom{n-m}{k}\binom{m}{l}(i\sin\frac{\theta}{2})^{m-l+k}(\cos\frac{\theta}{2})^{n-m-k+l}.$$

with $0 \le m, j \le n$, $0 \le \alpha \le 2\pi$, $0 \le \theta \le \pi$ and $-\pi \le \beta \le \pi$.

7.2.2 Symmetry Subspaces for SO(3)

We now want to find basis vectors for the first symmetry subspace for each irreducible representation V_{2k+1} of SO(3), where $n = 2k$. These will be one-dimensional spaces.

We require the following standard theorem from complex analysis (see [La]).

Theorem 7.2 *Let $Q(x, y)$ be a rational function which is continuous when $x^2 + y^2 = 1$. Let $f(z) = Q(\frac{1}{2}(z+\frac{1}{z}), \frac{1}{2i}(z-\frac{1}{z}))/iz$. Then*

$$\int_0^{2\pi} Q(\cos\theta, \sin\theta)d\theta = 2\pi i\sum \textit{(residues of } f \textit{ inside the unit circle).}$$

This is used to prove the following.

Lemma 7.3 $\int_0^{2\pi}\exp(-kix)\sin^n x\cos^m x dx = 0$ *for $n, m \in \mathbb{Z}$ with $0 \le n+m < k$.*

Proof: In our case $Q(\cos\theta, \sin\theta) = (\cos\theta - i\sin\theta)^k\sin^n\theta\cos^m\theta$ is continuous everywhere. Thus

$$\begin{aligned} f(z) &= [\frac{1}{2}(z+\frac{1}{z}) - \frac{1}{2i}(z-\frac{1}{z})]^k[\frac{1}{2i}(z-\frac{1}{z})]^n[\frac{1}{2}(z+\frac{1}{z})]^m\frac{1}{zi} \\ &= \frac{1}{2}(z+\frac{1}{z}-z+\frac{1}{z})^k\frac{1}{zi}(\frac{1}{i})^n(\frac{1}{2})^{n+m}(z-\frac{1}{z})^n(z+\frac{1}{z})^m \\ &= (\frac{1}{z})^{k+1}(z-\frac{1}{z})^n(z+\frac{1}{z})^m(\frac{1}{i})^{n+1}(\frac{1}{2})^{n+m}. \end{aligned}$$

The term $\frac{1}{z}$ can only show up in the expansion of $f(z)$ if z^k appears in the binomial expansion of $(z - \frac{1}{z})^n(z + \frac{1}{z})^m$, but since $n + m < k$ this cannot happen. Therefore the residue of f inside the unit circle is zero, and thus the integral is zero by applying Theorem 7.2. qed

With the help of this Lemma we have the following.

Proposition 7.4 *The basis vector of the first symmetry subspace will be a polynomial in x and y only, for all irreducible representations V_n.*

Proof: To see this we may project any polynomial $f(x, y, z)$ in x, y and z onto the first symmetry subspace; it is convenient to choose x^k, where $k = n/2$. Recall that the projector has the following form:

$$P_k(x^k) = \frac{2k+1}{8\pi^2} \int_0^{2\pi} \int_0^{\pi} \int_0^{2\pi} A(\phi_1, \theta, \phi_2)(g \cdot x^k) d\phi_1 d\theta d\phi_2$$
$$\text{where } A(\phi_1, \theta, \phi_2) = \cos^{2k} \frac{\theta}{2} \exp(-ik(\phi_1 + \phi_2)) \sin\theta.$$

The order of integration does not matter; in fact we can write the triple integral as a sum of products of three single integrals since in the expansion the variables separate. Hence it suffices to show that the integral in ϕ_2 vanishes.

Furthermore note that for an arbitrary group element g, as shown earlier, we have

$$\begin{aligned} g \cdot x^k = ([&-\sin\phi_1 \cos\theta \sin\phi_2 + \cos\phi_2 \cos\phi_1]x \\ &+ [\sin\phi_1 \cos\theta \cos\phi_2 + \sin\phi_2 \cos\phi_1]y + \sin\theta \sin\phi_1 z)^k. \end{aligned}$$

Note that the coefficient of z does not involve $\sin\phi_2$ or $\cos\phi_2$; thus in the integrand any term in z will have the form $h(\phi_1, \theta)\exp(-ik\phi_2)\sin^l\phi_2 \cos^j\phi_2$, where $j + l < n$. Thus by Lemma 7.3 this integral is zero. qed

We can use a similar method to find the values of the triple integrals. The entire integral can be split up as a sum of integrals of the form

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} CA(\phi_1,\theta,\phi_2)\cos^j\theta\cos^m\phi_1\sin^n\phi_1\cos^M\phi_2\sin^N\phi_2 x^{k-l}y^l d\phi_1 d\theta d\phi_2$$
$$= C\int_0^{2\pi}\exp(-ki\phi_1)\cos^m\phi_1\sin^n\phi_1 d\phi_1\int_0^{\pi}\cos^{2k}\frac{\theta}{2}\cos^j\theta\sin\theta d\theta \qquad (7.5)$$
$$\cdot\int_0^{2\pi}\exp(-ki\phi_2)\cos^M\phi_2\sin^N\phi_2 d\phi_2$$

The integrals in ϕ_1 and ϕ_2 have the form of the integral in Theorem 7.2. Thus we can use it to compute the residue at zero and thus the integrals, especially by observing that $m+n=k$ and $M+N=k$ by looking at $g\cdot x^n$. An easy computation shows that the values of the two integrals is $2\pi i(-i)^{n+1}(\frac{1}{2})^k(-1)^k$ and $2\pi i(-i)^{N-1}(\frac{1}{2})^k(-1)^k$, respectively, where the dimension of the irreducible representation is $2k+1$. Also the integral with respect to θ can be computed with elementary tools, namely trigonometric identities and substitution. The value is

$$(\frac{1}{2})^k\sum_{l=0}^{k}\binom{k}{l}\frac{1-(-1)^{l+j+1}}{l+j+1}$$

We note that $j=n$, and so the triple integral (7.5) has the value

$$-C4\pi^2(\frac{1}{2})^{3k}(-i)^{n+N+2}\sum_{l=0}^{k}\binom{k}{l}\frac{1-(-1)^{l-n+1}}{l+n+1}. \qquad (7.6)$$

We can use this formula to find the basis vector of the first symmetry subspace for each irreducible representation V_n with dimension $n+1$.

Theorem 7.7 *The basis vector of the first symmetry subspace for V_{2k} of dimension $2k+1$ is $-(\frac{1}{2})^k(x-iy)^k$; i.e.*

$$P_k(x^n) = \frac{2k+1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi}A(\phi_1,\theta,\phi_2)(g\cdot x^k)d\phi_1 d\theta d\phi_2$$
$$= -(\frac{1}{2})^k(x-iy)^k.$$

Proof: Considering $P_k(x^k)$, we notice that it is a homogeneous polynomial in x and y of degree k. To prove the claim we will compute the coefficients of $x^{k-l}y^l$ for $0 \leq l \leq k$ explicitly. As we have seen

$$g \cdot x^k = ([-\sin\phi_1\cos\theta\sin\phi_2 + \cos\phi_2\cos\phi_1]x$$
$$+ [\sin\phi_1\cos\theta\cos\phi_2 + \sin\phi_2\cos\phi_1]y + \sin\theta\sin\phi_1 z)^k.$$

We can disregard the terms involving z because of Proposition 7.4. Thus we just have to consider

$$((-\sin\phi_1\cos\theta\sin\phi_2 + \cos\phi_2\cos\phi_1)x + (\sin\phi_1\cos\theta + \sin\phi_2\cos\phi_1\cos\phi_2)y)^k$$
$$= \sum_{l=0}^{k}\binom{k}{l}(-\sin\phi_1\cos\theta\sin\phi_2 + \cos\phi_2\cos\phi_1)^{k-l}x^{k-l}$$
$$(\sin\phi_1\cos\theta + \sin\phi_2\cos\phi_1\cos\phi_2)^l y^l$$

Now denote the coefficient of $x^{k-l}y^l$ in this expansion as $(g \cdot x^k)_l$, so that

$$(g \cdot x^k)_l = \binom{k}{l}\sum_{j=0}^{k-l}\sum_{p=0}^{l}\binom{k-l}{j}\binom{l}{p}(-1)^{k-l-j}(\sin\phi_1)^{k-j-p}(\sin\phi_2)^{k-l-j+p}$$
$$\cdot (\cos\theta)^{k-j-p}(\cos\phi_1)^{j+p}(\cos\phi_2)^{j+l-p}.$$

To compute the coefficients of $x^{k-l}y^l$ for $0 \leq l \leq k$ we have to evaluate the following integrals:

$$\text{coeff. of} x^{k-l}y^l = \frac{2k+1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} A(\phi_1,\theta,\phi_2)(g \cdot x^k)_l d\phi_1 d\theta d\phi_2$$
$$= \frac{2k+1}{8\pi^2}(-4\pi^2)(\frac{1}{2})^(3k)\binom{k}{l}\sum_{j=0}^{k-l}\sum_{p=0}^{l}(-i)^{2(k-j+1)-l}\binom{k-l}{j}\binom{l}{p}$$
$$(-1)^{k-l-j}\sum_{m=0}^{k}\binom{k}{m}\frac{1-(-1)^{m+k-j-p+1}}{m+k-j-p+1} \quad \text{by (7.6)}$$
$$= -\binom{k}{l}\frac{2k+1}{2}(\frac{1}{2})^{3k}(-i)^l\sum_{J=0}^{k-l}\binom{k-l}{J}\sum_{P=0}^{l}\binom{l}{P}\sum_{m=0}^{k}\binom{k}{m}\frac{1-(-1)^{m+J+P+1}}{m+J+P+1}$$

where $J = k - l - j$ and $P = l - p$

Now we just have to use the binomial formula and its integral to compute

$$
\begin{aligned}
&\sum_{J=0}^{k-l}\binom{k-l}{J}\sum_{P=0}^{l}\binom{l}{P}\sum_{m=0}^{k}\binom{k}{m}\frac{1-(-1)^{m+J+P+1}}{m+J+P+1} \\
&=\sum_{J=0}^{k-l}\binom{k-l}{J}\sum_{P=0}^{l}\binom{l}{P}\int_0^1((1+x)^k x^{J+P}+(1-x)^k(-x)^{J+P})dx \\
&=\sum_{J=0}^{k-l}\binom{k-l}{J}\int_0^1((1+x)^{k-l}x^J+(1-x)^{k-L}(-x)^J)dx \\
&=\int_0^1((1+x)^{k+l}(1+x)^{k-l}+(1-x)^{k+l}(1-x)^{k-l})dx \\
&=\int_0^1((1+x)^{2k}+(1-x)^{2k})dx \\
&=\frac{2^{2k+1}}{2k+1}
\end{aligned}
$$

Thus the coefficient of $x^{k-l}y^l$ is

$$-\binom{k}{l}\frac{2k+1}{2}\left(\frac{1}{2}\right)^{3k}(-i)^l\frac{2^{2k+1}}{2k+1}=-\binom{k}{l}(-i)^l\left(\frac{1}{2}\right)^k.$$

But that is exactly the coefficient of $x^{k-l}y^l$ in the expansion of $-(\frac{1}{2})^k(x-iy)^k$.

qed

Having found a basis vector for each first symmetry subspace we can use the transfer operators to collect a G-adapted basis for our vector space.

Comparing the basis vectors with the spherical harmonic functions, we also see that our basis vectors are exactly equal to the first spherical harmonic in each dimension.

With $x=\cos\theta\sin\phi$, $y=\sin\theta\sin\phi$, $z=\cos\phi$, we have

$$(x-iy)^k=\sin^k\phi\exp(-ik\theta)=Y_{k,-k}(\phi,\theta)$$

It is a reasonable conjecture that the basis of spherical harmonics is a G-adapted basis. Unfortunately, so far we have not succeeded in proving this conjecture.

We have found the KL basis for the space of scalar valued functions on the sphere. without even looking at data set. The KL basis seems to be independent of the data set but the eigenvalues are not. To complete the KL analysis of a particular data set in this vector space we would have to follow the algorithm from step two on (step one was completed in this section). A suitable data set. for example. would come from solving the complex Ginzburg Landau equation on the sphere. This is a partial differential equation which is equivariant under the action of SO(3). and thus a data set collected from numerically solving the equation would satisfy our symmetry assumption that every acted-on data vector is again a possible data vector. Sibylle Geiger describes the normal KL decomposition on such a data set, and symmetrized data sets in her Diplom thesis [Ge]. We will have a closer look at the Ginzburg Landau equation in the next chapter. although we are using a vector-valued version to produce vector fields that are tangent to the sphere.

Chapter 8

Example: Polynomial Vector Fields

8.1 Theory

In this section we will investigate the theoretical aspects of the representations of SO(3) on the vector space of polynomial vector fields that are tangent to the sphere. Vector fields are of interest whenever one studies flow problems such as fluid flow. magnetic flow, and athmospheric flow. Studying tangent vector fields seems like a good way to start looking at this problem.

This example promises to be more interesting than using data of scalar valued functions. since the structure of the decomposition into the direct sum of irreducible representations is more complicated. As we showed in the previous chapter the decomposition of the vector space of scalar valued functions consists of one copy of each irreducible representation. Thus all the symmetry subspaces were one dimensional and the KL basis was the same as the G-adapted basis. Considering tangent vector fields on the sphere we would expect this to change, since the data now lies in tangent planes. It seems plausible to expect a decomposition where each irreducible representation appears with multiplicity two.

We will first discuss the decomposition of the space into the direct sum of irreducible representations and find the symmetry subspaces of this representation. Later we will also make connections between tangent and divergence free vector fields.

8.1.1 Irreducible Representations

We want to investigate the vector space of polynomial vector fields that are tangent to the sphere. First we will compute the dimension of this vector space and then we will find its decomposition into irreducible representations. The main tool we will use to attain those goals is to look at exact sequences of related vector spaces.

Consider the following diagram with the indicated maps and spaces as given below.

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & & & \\
0 \arrow[r] & K_{d-2} \arrow[r] \arrow[d] & V^3_{d-2} \arrow[r, "\tau"', "\cong"] \arrow[d, "\rho"] & V_{d-1} \arrow[r] \arrow[d, "\rho"] & C_{d-1} \arrow[r, phantom, "\cong"] \arrow[d, "\cong"] & \mathbb{R} \arrow[r] & 0 \\
0 \arrow[r] & K_d \arrow[r] \arrow[d] & V^3_d \arrow[r, "\tau"] \arrow[d, "\pi"] & V_{d+1} \arrow[r] \arrow[d, "\pi"] & C_{d+1} \arrow[r, phantom, "\cong"] \arrow[d] & \mathbb{R} \arrow[r] & 0 \\
0 \arrow[r] & \overline{K_d} \arrow[r] \arrow[d] & S^3_d \arrow[r, "\tau"] \arrow[d] & S_{d+1} \arrow[r] \arrow[d] & \overline{C_{d+1}} \arrow[r, phantom, "="] & 0 & \\
 & 0 & 0 & 0 & & &
\end{tikzcd}$$

The Spaces:

$V_d =$ polynomials in x, y, z of degree less than or equal to d.

$S_d =$ polynomials of degree less than or equal to d on the sphere of radius one.

$V^3_d =$ polynomial vector fields of degree less than or equal to d.

$S_d^3 =$ polynomial vector fields of degree less than or equal to d on the sphere.

$\overline{K_d} =$ tangent vector fields of polynomials of degree less than or equal to d on the sphere.

The maps:

$$\begin{array}{lccc} \tau : & V_d^3 & \longrightarrow & V_{d+1} \\ & (f, g, h)^T & \longrightarrow & xf + yg + zh \end{array}$$

$$\begin{array}{lccc} \rho : & V_{d-2} & \longrightarrow & V_d \\ & p(x, y, z) & \longrightarrow & p(x, y, z) \cdot (x^2 + y^2 + z^2 - 1) \end{array}$$

$$\begin{array}{lccc} \pi : & V_d^3 & \longrightarrow & S_d^3 \\ & p(x, y, z) & \longrightarrow & p(x, y, z)/(x^2 + y^2 + z^2 - 1) \end{array}$$

If we know the dimensions of some of these vector spaces we can use the exactness of the sequences in the diagram to compute the dimensions of the other spaces.

Recall that a sequence

$$A \xrightarrow{\phi} B \xrightarrow{\psi} C$$

is called exact if $\ker \psi = \operatorname{im} \phi$. A sequence

$$0 \longrightarrow A \xrightarrow{\phi} B \xrightarrow{\psi} C \longrightarrow 0$$

is called short exact. In this case ϕ is injective and ψ is surjective.

Looking back at the diagram we see that the short exactness of the second and third columns allows us to compute the dimensions of all the spaces in those columns. For example, the dimension of $V_{d+1} = \operatorname{rank} \pi + \operatorname{nullity} \pi$. Since π is onto, $\operatorname{rank} \pi = \dim S_{d+1}$, and since $\ker \pi = \operatorname{im} \rho$ we have that $\operatorname{nullity} \pi = \dim V_{d-1}$. Thus $\dim V_{d+1} = \dim S_{d-1} + \dim V_{d-1}$. We can make a similar argument to find the dimension of all spaces in the second column. Because the two mentioned sequences are exact and the squares in the diagram commute, by the snake lemma there is a long exact sequence from the kernels of τ to the cokernels of τ (see [AM]). The two cokernels C_{d-1} and C_{d+1} are isomorphic; therefore, $\overline{C_{d+1}} = 0$ and thus the second to last row in the diagram is also short exact. We now want to list the dimensions of all the spaces.

Dimensions of the relevant spaces:

$$\dim V_d^3 = (d+3)(d+2)(d+1)/2$$

$$\dim V_{d-2}^3 = (d+1)d(d-1)/2$$

$$\dim S_d^3 = \dim V_d^3 - \dim V_{d-2}^3 = 3(d+1)^2$$

$$\dim V_{d-1} = d(d+2)(d+1)/6$$

$$\dim V_{d+1} = (d+4)(d+3)(d+2)/6$$

$$\dim S_{d+1} = (d+2)^2$$

$$\dim \overline{K_d} = \dim S_d^3 - \dim S_{d+1} = 2d^2 + 2d - 1$$

Writing down the dimensions of the spaces of tangent vector fields on the sphere for low degrees we find the following equations:

d	dim	
1	3	$= 3$
2	11	$= 3 + 3 + 5$
3	23	$= 3 + 3 + 5 + 5 + 7$
4	39	$= 3 + 3 + 5 + 5 + 7 + 7 + 9$

Recall that we are interested in the decompositon of $\overline{K_d}$ into the direct sum of irreducible representations. Judging from the dimension of these spaces, there is a definite pattern in the way the dimensions increase from degree to degree. It seems that when we increase the degree by one we get two more irreducible representations, the one with highest degree that we had before and the one with the next higher degree.

To find out if this conjecture is correct we have to compute the decomposition of all the relevant spaces into the direct sum of irreducible representations. Fortunately we already know the decomposition of the polynomial functions of degree d on the sphere S_d from the previous sections:

$$S_d = \bigoplus_{l=0}^{d} W_{2l+1} = \bigoplus_{l=0}^{d} W_{2l+1}^{\cdot e(d,l)}.$$

where

$$e(d, l) = \begin{cases} 1, & \text{if } 0 \leq l \leq d \\ 0, & \text{if } l > d+1 \end{cases}.$$

and W_{2l+1} is the irreducible representation of dimension $2l+1$ (it contains the homogeneous polynomials of degree l).

We also know that for constant and linear polynomials it doesn't matter if the domain is the sphere or all of $\mathbb{R}^3$, since the spaces only differ once we allow quadratics or higher. Thus we know that

$$V_0 = W_1 = S_0, \quad V_1 = W_1 + W_3.$$

We now have enough information to compute the decomposition into irreducibles of all the spaces in the diagram.

Since $V_{d+1} = V_{d-1} + S_{d+1}$ we can compute V_d inductively, and we get the following relationship:

$$V_d = W_{2d+1} + W_{2d-1} + W_{2d-3}^{\oplus 2} + W_{2d-5}^{\oplus 2} + \cdots + W_1^k,$$

where $k = \frac{d+1}{2}$ if d is odd and $k = \frac{d+2}{2}$ if d is even. We can also write V_d the following way:

$$
\begin{aligned}
V_d &= \bigoplus_{l=0}^{d} W_{2l+1}^{+b(d,l)} \\
&= \bigoplus_{l=0}^{d} W_{2l+1}^{+\lfloor (d-l)/2 \rfloor + 1}.
\end{aligned}
$$

where

$$
b(d,l) = \begin{cases} \lfloor (d-l)/2 \rfloor + 1, & \text{if } 0 \le l \le d \\ 0, & \text{if } l > d \end{cases}
$$

Knowing the spaces V_d, next we can find V_d^3 using the following equation:

$$
V_d^3 = W_3 \otimes V_d.
$$

We can show that these spaces are isomorphic as representations by showing two things:

1. They are isomorphic as vector spaces.
2. The action of SO(3) is the same on both spaces.

To 1.: Comparing bases for the involved spaces we see:

- basis for V_d : f_i, polynomials in x, y, z of degree at most d
- basis for W_3 : e_1, e_2, e_3, the unit vectors in three dimensional space
- basis for $W_3 \otimes V_d$: $e_i \otimes f_j$, put f_j into ith spot
- typical element of $W_3 \otimes V_d$: $e_1 \otimes f + e_2 \otimes g + e_3 \otimes h = (f, g, h)$

Thus the spaces are isomorphic as vector spaces.

To 2.: Let $\sigma \in SO(3)$. Then on V_d^3 the action is

$$\sigma \begin{bmatrix} f(\overline{x}) \\ g(\overline{x}) \\ h(\overline{x}) \end{bmatrix} = \sigma \begin{bmatrix} f(\sigma^{-1}\overline{x}) \\ g(\sigma^{-1}\overline{x}) \\ h(\sigma^{-1}\overline{x}) \end{bmatrix}.$$

On $W_3 \otimes V_d$ the action is the following:

$$\sigma \left(\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \otimes f(\overline{x}) \right) = \sigma \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \otimes f(\sigma^{-1}\overline{x}).$$

So we see that the action on the two spaces are equivalent, and therefore the two spaces are isomorphic as representations.

To find the decomposition of V_d^3 note that $V_0^3 = W_3 \otimes W_1 = W_3$, and $W_3 \otimes W_{2l+1} = W_{2l+3} + W_{2l+1} + W_{2l-1}$, for $1 \leq l$ ([FH]).

Let

$$\begin{aligned} V_d^3 &= W_1^{\oplus a(d,1)} + W_3^{\oplus a(d,2)} + \cdots + W_{2d-1}^{\oplus a(d,d-1)} + W_{2d+1}^{\oplus a(d,d)} + W_{2d+3}^{\oplus a(d,d+1)} \\ &= \bigoplus_{l=0}^{d+1} W_{2l+1}^{\oplus a(d,l)}. \end{aligned}$$

We can list the multiplicities of the irreducible representations in the sums of V_d^3:

d	1	3	5	7	9	11	13	15
0	0	1						
1	1	2	1					
2	1	4	2	1				
3	2	5	4	2	1			
4	2	7	5	4	2	1		
5	3	8	7	5	4	2	1	
6	3	10	8	7	5	4	2	1

In general we have the following relationship between the multiplicities of the irreducible representations in the sums of V_d and V_d^3:

$$a(d,l) = \begin{cases} b(d,2), & \text{if } l = 1 \\ b(d,l-1) + b(d,l) + b(d,l+1), & \text{if } 1 < l \end{cases}$$

Thus

$$a(d,l) = \begin{cases} \lfloor (d+1)/2 \rfloor, & \text{if } l = 0 \\ \lfloor (d-l+3)/2 \rfloor + \lfloor (d-l+2)/2 \rfloor \lfloor (d-l+1)/2 \rfloor, & \text{if } 1 \le l \le d-1 \\ 2, & \text{if } l = d \\ 1, & \text{if } l = d+1 \\ 0, & \text{if } l \ge d+2 \end{cases}.$$

Now we can decompose the space of polynomial vector fields on the sphere and find:

$$S_d^3 = \bigoplus_{l=0} W_{2l+1}^{c(d,l)},$$

where $c(d,l) = a(d,l) - a(d-2,l)$. Thus

$$a(d,l) = \begin{cases} 1, & \text{if } l = 0 \\ 3, & \text{if } 1 \le l \le d-1 \\ 2, & \text{if } l = d \\ 1, & \text{if } l = d+1 \\ 0, & \text{if } l \ge d+2 \end{cases}$$

This result is not surprising. One might have guessed that the space of polynomial vector fields on the sphere splits up as the sum of three copies of the irreducible representations in each dimension. After all, the space of polynomials is the direct sum of exactly one copy in each dimension, and polynomial vector fields can be viewed as triples of polynomials. However, the situation is not quite so simple. Indeed, most of the irreducible representations in the decomposition show up with multiplicity three, except for the first and the last two dimensions.

Finally we have reached our original goal of finding the decomposition of the tangent vector fields of degree less or equal to d on the sphere $\overline{K_d}$. Let

$$\overline{K_d} = \bigoplus_{l=0} W_{2l+1}^{k(d,l)}.$$

Then

$$k(d,l) = \begin{cases} 0, & \text{if } l = 0 \\ 2, & \text{if } 1 \leq l \leq d-1 \\ 1, & \text{if } l = d \\ 0, & \text{if } l \geq d+1 \end{cases}$$

8.1.2 Examples

Now we will have a closer look at a low dimensional example of the introduced results to help us visualize the situation. We want to take a closer look at linear and polynomial vector fields that are tangent to the sphere. The only constant vector field that is tangent to the sphere is the trivial one.

Three dimensional irreducible representation, linear vector fields

A few simple computations or some thought yields the result that the linear tangent vector fields are of the form

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = S \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

where S is a three-by-three skew-symmetric matrix. Since there are exactly three free choices for entries in S, the space of linear tangent polynomial vector fields on the sphere is three dimensional, exactly one copy of W_3, as it should be, according to the theory.

The quadratic tangent polynomial vector fields are a little more interesting. We know that the space splits up into two three-dimensional, irreducible representations and one 5-dimensional, irreducible representation. The bases for the irreducible representations given below were found by projecting arbitrary vector fields onto the irreducible representation spaces. We keep one of the two three dimensional, irreducible representations from the linear vector fields. Bases of the second three-dimensional and of the 5-dimensional, irreducible representation are given below. Some selected vector fields are shown in figures 8.1-8.3. We show one vector field from each irreducible representation discussed in this section.

Three dimensional irreducible representation, quadratic vector fields

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} y^2 + z^2 \\ -xy \\ -xz \end{bmatrix}, \quad \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -xy \\ x^2 + z^2 \\ -yz \end{bmatrix}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -xz \\ -yz \\ x^2 + y^2 \end{bmatrix}.$$

Five dimensional irreducible representation

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 2yz \\ -xz \\ -xy \end{bmatrix}, \quad \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -yz \\ 2xz \\ -xy \end{bmatrix}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} xz \\ -yz \\ -x^2 + y^2 \end{bmatrix}, \quad \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -y^2 + z^2 \\ xy \\ -xz \end{bmatrix}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} xy \\ -x^2 + z^2 \\ -zy \end{bmatrix}.$$

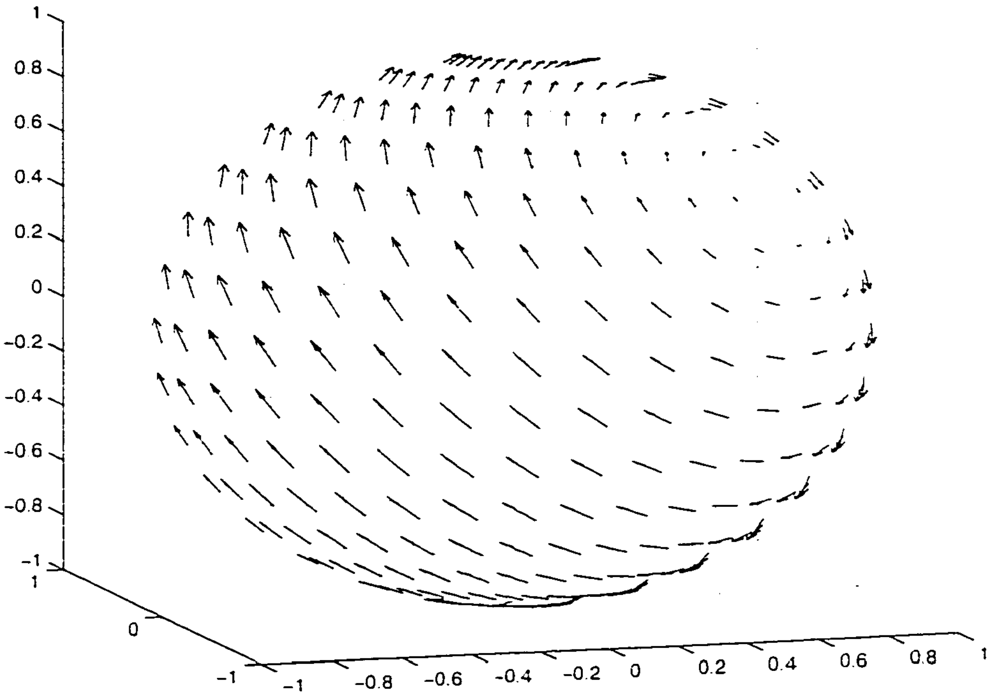

Figure 8.1: $\dot{x} = y + z$, $\dot{y} = -x + z$, $\dot{z} = -x - y$

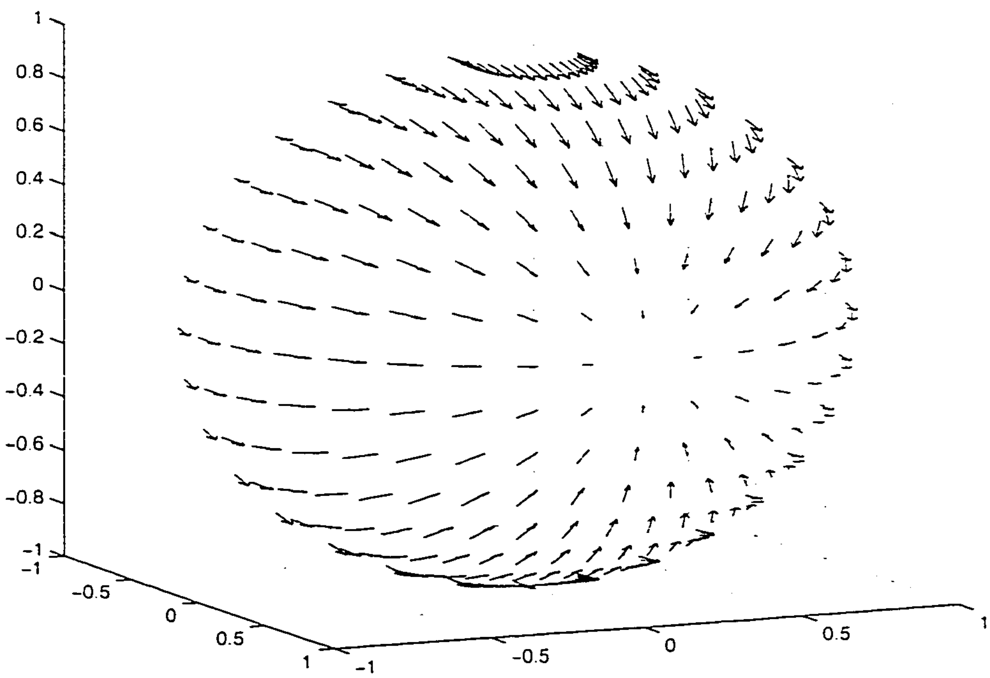

Figure 8.2: $\dot{x} = y^2 + z^2$, $\dot{y} = -xy$, $\dot{z} = -xz$

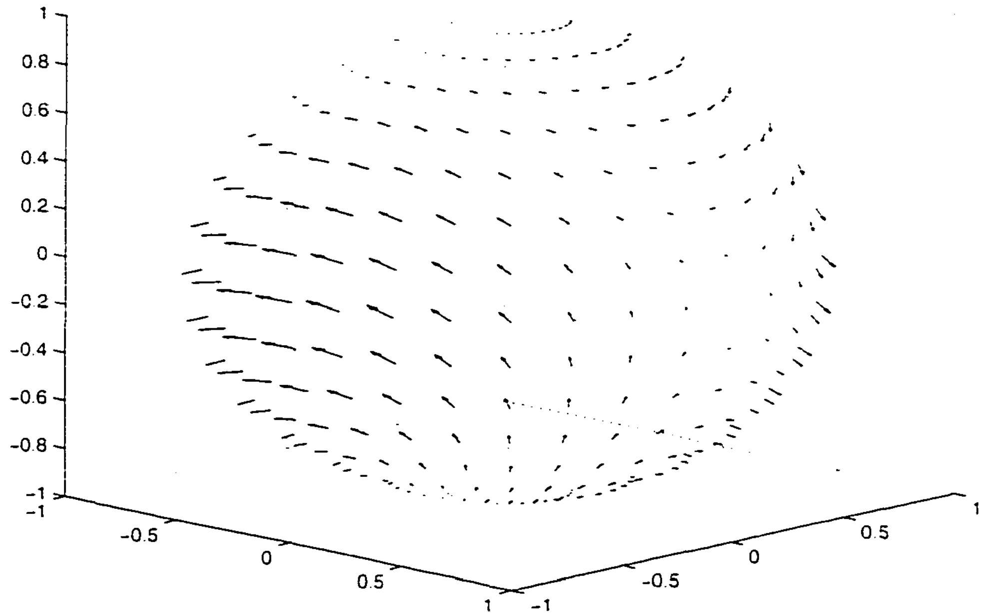

Figure 8.3: $\dot{x} = 2yz$, $\dot{y} = -xz$, $\dot{z} = -xy$

8.2 Divergence-Free Vector Fields

Divergence-free vector fields play an important role in physics, because magnetic field lines as well as streamlines of incompressible fluids are trajectories of such vector fields. Being volume preserving, they can be considered as three-dimensional versions of Hamiltonian systems.

The structure of three-dimensional chaotic volume preserving flows is, however, much less understood than for Hamiltonian systems. Some progress has been made by using the fact that they can locally be transformed to two-dimensional non-autonomous Hamiltonian vector fields. Analytical approaches to understand chaotic dynamics for such systems are based on Melnikov theory in which an autonomous, two-dimensional Hamiltonian system is perturbed near a homoclinic or heteroclinic orbit. Although this approach sheds some light into chaotic volume preserving flows, it certainly does not capture all possible phenomena, in particular the effects of fixed points of the vector field are not accessible.

Divergence-free vector fields in $\mathbb{R}^3$ which possess a conserved quantity are a generalization of autonomous Hamiltonian vector fields in the plane. Such vector fields can have fixed points and typically the phase space is foliated by families of invariant tori or cylinders. Perturbing such vector fields by an appropriate extension of Melnikov theory should give much insight in the possible types of chaotic volume preserving flows. First attempts in this direction have been made by Mezić and Wiggins in [MW]. It is, therefore, of interest to characterize divergence-free vector fields with conserved quantities.

It turns out that there is an intimate connection between certain classes of divergence free polynomial vector fields in $\mathbb{R}^3$ and on the sphere. We have investigated this connection to some extend and will present the results in this section.

Definition 8.1 *A vector field* $\begin{bmatrix} f \\ g \\ h \end{bmatrix}$ *is called divergence free if*

$$\operatorname{div}\begin{bmatrix} f \\ g \\ h \end{bmatrix} = f_x + g_y + h_z = 0$$

The question we want to answer in this section is which tangent vector fields on the sphere can be lifted to divergence free vector fields. It is easy to find examples to see that not every tangent vector field is divergence free. For example,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} y^2 + z^2 \\ -xy \\ -xz \end{bmatrix}.$$

is tangent but its divergence is $-x - x = -2x$. On the other hand, if we add a term to the vector field which is zero on the sphere, we don't change the vector field on the sphere but we succeed in making it divergence free:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} y^2 + z^2 + (x^2 + y^2 + z^2 - 1) \\ -xy \\ -xz \end{bmatrix}.$$

The divergence of this vector field is $2x - x - x = 0$.

Can this be done in general? Can each tangent vector field on the sphere be "lifted" to a divergence free vector field by adding terms that are zero on the sphere? To answer this question we want to consider the following diagram of exact sequences. Note that the spaces and maps are denoted as before and F_d^3 is the space of divergence free polynomial vector fields of degree less or equal to three.

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & V_{d-2}^3 \arrow[d, "\rho"] & & \\
0 \arrow[r] & F_{d-2}^3 \arrow[r, "\rho"] & K_d + \rho(V_d^3) \arrow[r, "\mathrm{div}"] \arrow[d, "\pi"] & V_{d-1} \arrow[r] & 0 \\
 & & \overline{K_d} \arrow[d] & & \\
 & & 0 & &
\end{tikzcd}$$

We are interested in the following map: $F_{d-2}^3 \xrightarrow{\rho} K_d + \rho(V_{d-2}^3) \xrightarrow{div} \overline{K_d}$.

We would like to know what is in the image or this map, or rather, what is not in the image of this map, because the tangent vector fields on the sphere that are not in the image of this map are the ones that cannot be lifted to divergence free vector fields. Thus we want to know what is in the cokernel of this map.

To answer this question we investigate the equivalent question: what is in the cokernel of the map that first multiplies a vector field by ρ and then takes the divergence.

$$V_{d-2}^3 \xrightarrow{\rho} K_d + \rho(V_{d-2}^3) \xrightarrow{\mathrm{div}} V_{d-1}$$

The fact that these two questions are equivalent follows from a combination of the first and third isomorphism theorems. Consider the following diagram.

$$
\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & D \arrow[d] & & \\
0 \arrow[r] & A \arrow[r] & B \arrow[r, "q"] \arrow[d, "p"] & C \arrow[r] & 0 \\
 & & E \arrow[d] & & \\
 & & 0 & &
\end{tikzcd}
$$

Since the vertical as well as the horizontal sequences are short exact it follows that $D = \ker. p$ $A = \ker q$ and both p and q are onto. Thus $p(A) = p(A+D) = (A+D)/D$ by the third isomorphism theorem. Similarly, we have $q(D) = q(D+A) = (D+A)/A$. We want to compare the cokernels of the following two maps:

$$Q : D \longrightarrow B \xrightarrow{q} C.$$

and

$$P : A \longrightarrow B \xrightarrow{p} E.$$

The cokernels of P and Q are

$$\operatorname{coker}(P) = E/p(A) = \frac{B/D}{(A+D)/D} = B/(A+D)$$

$$\operatorname{coker}(Q) = C/q(D) = \frac{C/A}{(D+A)/A} = B/(D+A).$$

The last equality in both equations follows from the third isomorphism theorem. Thus, we see that the cokernels are the same.

Now we will investigate the image and the cokernel of the map

$$V_{d-2}^3 \xrightarrow{\cdot\rho} K_d + \rho(V_{d-2}^3) \xrightarrow{\text{div}} V_{d-1}.$$

It is easy to see that constant polynomials are not in the image. It is not possible to multiply each entry in a vector field by $x^2+y^2+z^2-1$ and then take the divergence of this to obtain a constant. In fact it is true that the cokernel is one dimensional. We can show that up to a constant term, every polynomial of degree less or equal to $d-1$ is in the image. To prove this we consider a basis of monomials.

We want to show by induction that every monomial plus some constant is in the image. It is easy to show that this is true for non-mixed monomials x^n, and also for the base cases $x^k y^l z^n$ where $0 \leq k, l, n \leq 2$.

Lemma 8.2 *$x^k y^l z^n +$ lower order terms is in the image.*

Proof: We assume that $x^{k'} y^{l'} z^{n'} +$ constant is in the image for $k' \leq k$, $l' \leq l$, $n' \leq n$ where at least one of the $\leq$ is strictly less.

The candidate in V_{d-2}^3 that will map to $x^k y^l z^n$+constant is

$$\dot{x} = a x^{k-1} y^l z^n$$
$$\dot{y} = b x^{k-2} y^{l+1} z^n$$
$$\dot{z} = c x^{k-2} y^l z^{n+1}$$

where

$$a = \frac{4+l+n}{2(k+l+n+3)}, \quad b = c = \frac{1-k}{2(k+l+n+3)}.$$

To prove the claim we just have to multiply the candidate vector field by ρ and then take the divergence. The resulting polynomial is:

$$\begin{aligned}&(a(k+1)+b(l+1)+c(n+1))x^k y^l z^n\\&+(a(k-1)+b(l+3)+c(n+1))x^{k-2}y^{l+2}z^n\\&+(a(k-1)+b(l+1)+c(n+3))z^{k-2}+y^l z^{l+2}+\text{lower order terms}\end{aligned}$$

With the choices of a. b. and c as given above this polynomial is exactly $x^k y^l z^n+$ lower order terms.

qed

8.2.1 Curl and Divergence

Recall that the curl of a vector field $(P.Q.R)$ is defined as $\text{curl}(P.Q.R) = (R_y - Q_z. P_z - R_x. Q_x - P_y)$. It is easy to see that $(\text{div}\circ\text{curl})(P.Q.R) = 0$. Thus. the curl of a vector field is divergence free.

It is also true that every divergence free vector field is the curl of some other vector field.

Lemma 8.3 *Every divergence free vector field is the curl of some vector field.*

Proof: Let $(F.G.H)$ be some divergence free vector field. that is: $F_x+G_y+H_z=0$. We want to show that there exists some vector field $(P.Q.R)$ such that $(F.G.H) = (R_y - Q_z. P_z - R_x. Q_x - P_y)$.

The candidate whose curl will produce $(F.G.H)$ is

$$\begin{aligned}R&=a_1\int F dy + a_0\int G dx\\Q&=b_1\int F dz + b_0\int H dx\\P&=c_1\int G dz + c_0\int H dy.\end{aligned}$$

where $(a_0, b_0, c_0, a_1, b_1, c_1) = (k, -k, -k, k, -1, 1)$ for some arbitrary constant k.

With our choices of (P, Q, R) we get

$$R_y - Q_z = (a_1 - b_1)F + a_0 \int G_y dx - b_0 \int H_z dx$$
$$P_z - R_x = (c_1 - a_0)G + c_0 \int H_z dy - a_1 \int F_x dy$$
$$Q_x - P_y = (b_0 - c_0)H + b_1 \int F_x dz - c_1 \int G_y dz.$$

Let $a_0 = -b_0$, $c_0 = -a_1$, $b_1 = -c_1$ and recall that $F_x + G_y + H_z = 0$. Then

$$R_y - Q_z = (a_1 - b_1)F + b_0 \int F_x dx = (a_1 - b_1 + b_0)F = F$$
$$P_z - R_x = (c_1 - a_0)G + a_1 \int G_y dy = (c_1 - a_0 + a_1)G = G$$
$$Q_x - P_y = (b_0 - c_0)H + c_1 \int H_z dz = (b_0 - c_0 + c_1)H = H.$$

Note that all operations we made were legal because our vector fields are polynomial functions. qed

8.3 Example: Ginzburg Laundau Equation

In this section we are going to illustrate the KL algorithm with an example of data obtained from solving a vector partial differential equation on the sphere. First we are going to introduce the differential equation, describe how it is solved and what the data will look like. Then we are going to discuss the steps of the algorithm for exploiting symmetry in the KL computation. Finally, we will take a closer look at the results.

8.3.1 The Ginzburg Landau Equation

The Ginzburg Landau equation has been found to govern the appearance of chemical turbulences in reaction diffusion systems [KY], [Ku], [KK] as well as to describe fluid systems such as plane Poiseuille flow [SS], Raleigh-Bènard convecion [NW], Taylor-Couette flow between rotating cylinders [KD], and wind-induced water waves [Bl] ([Ke]).

The Ginzburg Landau equation has been investigated extensively with circular geometry in [Ke], [SR], [MHR], and Sibylle Geiger used their results to extend the problem to the geometry of the sphere [Ge]. Geiger showed which parameter values in the equation yield chaotic solutions. We will use her results for scalar functions on the sphere to come up with a vector-valued version of the Ginzburg Landau equation to produce a data set of vector fields that are tangent to the sphere.

Since we are not primarily interested in the behavior and characteristics of the Ginzburg Landau equation we will not go into any detail on the topic, but we refer to some papers that derive and investigate the equation thoroughly [MHR], [Ke], [SR], [Ne], [Ge].

We are mainly interested in using a data set of tangent vector fields on the sphere to demonstrate our algorithm. Numerically solving our version of the Ginzburg Landau equations provides a suitable data set.

The complex scalar valued Ginzburg Landau equation on the sphere with constant radius $r_0 = 1$ is given by:

$$\frac{\partial A}{\partial t} = q^2(i + c_0)\Delta A + \rho A + (i - \rho)|A|^2 A.$$

A is a function from the sphere into $\mathbb{R}^\infty$,

$$A : S^2 = \{(r_0, \phi, \theta) : r_0 = 1, 0 \leq \phi \leq \pi, 0 \leq \theta < 2\pi\} \rightarrow \mathbb{R}^\infty.$$

q, ρ, and c_0 are parameters and ΔA is the Laplace operator. Geiger showed that for certain parameter values, such as $q = 0.63, \rho = 0.25$, and $c_0 = 0.25$, the equation has chaotic solutions.

In spherical coordinates the Laplace operator is

$$\Delta A = \frac{\partial^2 A}{\partial r^2} + \frac{2\partial A}{r\partial r} + \frac{1}{r^2 \sin^2 \phi}\frac{\partial^2 A}{\partial^2 \theta} + \frac{1}{r^2}\frac{\partial^2 A}{\partial \phi^2} + \frac{\cot \phi}{r^2}\frac{\partial A}{\partial \phi}.$$

where $A = A(r, \phi, \theta)$. Since we work with a constant radius $r_0 = 1$, this simplifies to

$$\Delta A = \frac{1}{\sin^2 \phi}\frac{\partial^2 A}{\partial^2 \theta} + \frac{\partial^2 A}{\partial \phi^2} + \cot \phi \frac{\partial A}{\partial \phi}.$$

We are going to consider a vector valued Ginzburg Landau equation with an additional coupling term:

$$\frac{\partial A}{\partial t} = q^2(i + c_0)\left(\frac{1}{\sin^2 \phi}\frac{\partial^2 A}{\partial^2 \theta} + \frac{\partial^2 A}{\partial \phi^2} + \cot \phi \frac{\partial A}{\partial \phi}\right) + \rho A + (i - \rho)|A|^2 A + \alpha A \times \underline{x},$$

where $A = (A^1, A^2, A^3)$, and $A \times \underline{x}$ is the cross product of A with the normal vector $\underline{x} = (x, y, z)$ in $\mathbb{R}^3$.

The solutions of this equations are not necessarily tangent vector fields, however, we can produce tangent vector fields with a projection onto the tangent planes. This process can be described by

$$\begin{aligned}\frac{\partial A}{\partial t} &= P_T\left\{q^2(i + c_0)\left(\frac{1}{\sin^2 \phi}\frac{\partial^2 A}{\partial^2 \theta} + \frac{\partial^2 A}{\partial \phi^2} + \cot \phi \frac{\partial A}{\partial \phi}\right) + \rho A + (i - \rho)|A|^2 A + \alpha A \times \underline{x}\right\} \\ &= P_T\left\{q^2(i + c_0)\left(\frac{1}{\sin^2 \phi}\frac{\partial^2 A}{\partial^2 \theta} + \frac{\partial^2 A}{\partial \phi^2} + \cot \phi \frac{\partial A}{\partial \phi}\right)\right\} + \rho A + (i - \rho)|A|^2 A + \alpha A \times \underline{x}.\end{aligned}$$

Here P_T denotes the projection onto the tangent plane.

If we add a condition on the vector field and require it to be divergence-free, we can guarantee that $\Delta A = (\Delta A^1, \Delta A^2, \Delta A^3)$ is tangent. In the following Lemma we will view A as a function in t, x, y, and z.

Lemma 8.4 *If the vector field* $A = (A^1, A^2, A^3)$ *is divergence-free, i.e.* $A^1_x + A^2_y + A^3_z = 0$, *then* ΔA *is tangent.*

Proof: Let $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ and $A = (A^1, A^2, A^3)$ be tangent, i.e. $xA^1 + yA^2 + zA^3 = 0$. Then $\Delta(xA^1 + yA^2 + zA^3) = 0$. But $\Delta(xA^1 + yA^2 + zA^3) = 2(A^1_x + A^2_y + A^3_z) + x\Delta A^1 + y\Delta A^2 + z\Delta A^3 = 0$. Because of the divergence-free condition the first term in this expression is zero. Therefore we can conclude that $x\Delta A^1 + y\Delta A^2 + z\Delta A^3 = 0$, and thus ΔA is tangent. qed

8.3.2 The Data Set

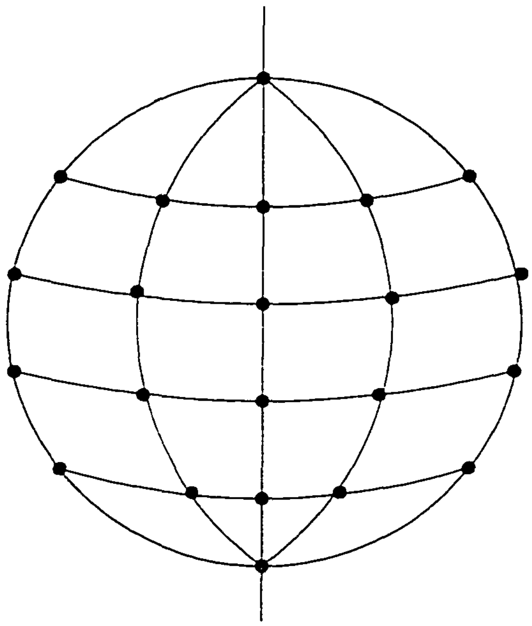

Figure 8.4: Grid on the sphere, simplified

We are going to use the same parameter values as given above, and $\alpha = 0.5$ in the PDE described in the previous section. To compute the solutions of the PDE, we use a second order Runge-Kutta method for the time derivative with stepsize $\Delta t = 0.001$, and centered differences for the spatial derivatives. For the discretization of space, as mentioned before, we change to spherical coordinates and put a 514 node grid on the sphere. The grid consists of 32 longitude lines and 16 latitude lines. This gives $32 \times 16 = 512$ grid points, and in addition we have one grid point at the north pole and one at the south pole. A simplified grid can be seen in figure 8.4.

We are computing the solutions of the PDE for 500000 timesteps and write every 100th iteration to a file, producing a data set of 5000 vectors. Each vector in the data set has 3×514 entries and represents a vector field that is tangent to the sphere. Some of the data can be seen in figures 8.6-8.8. To guarantee that the data vectors are indeed tangent, we project the updated vector onto the tangent planes at each time step, since the equations do not automatically produce tangent trajectories.

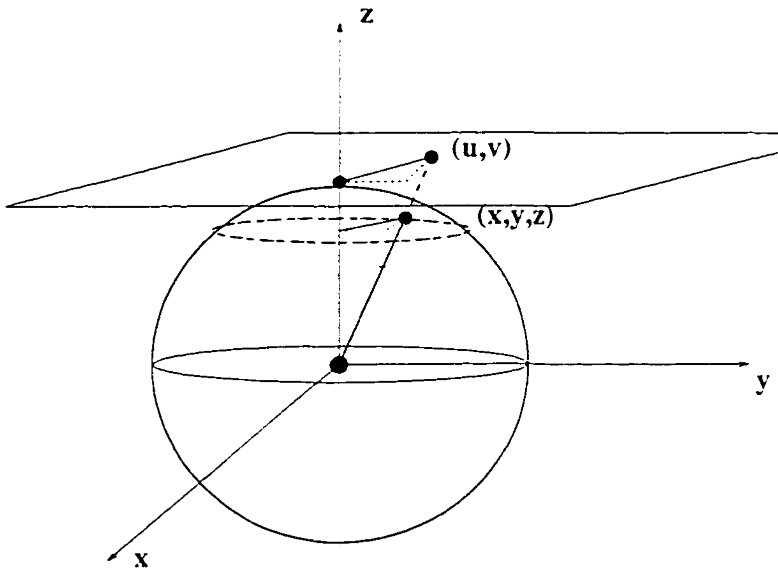

Figure 8.5: Change of coordinates at the poles

Unfortunately, changing the differential equation to spherical coordinates introduces singularities at the two poles, i.e., for $\phi = 0$ and $\phi = \pi$. To handle this problem, we project the nodes on the first and last latitude lines onto a circle in the tangent planes of the poles (see figure 8.5). This corresponds to a change of coordinates that eliminates the sigularities. The new coordinates are given by:

$$u = \frac{x}{z}, v = \frac{y}{z}, r^2 = x^2 + y^2 + z^2 = 1.$$

$$x = \frac{ru}{\sqrt{1+u^2+v^2}}, \; y = \frac{rv}{\sqrt{1+u^2+v^2}}, \; z = \frac{1}{\sqrt{1+u^2+v^2}}.$$

Then the Laplace operator becomes $\Delta A = A_{uu} + A_{vv}$, and the singularities introduced by the spherical coordinates disappear.

In figures 8.6-8.8 we show some time series of the data. Each picture represents the real part of the first coordinate of the tangent vector fields at one grid point, which were chosen more or less at random.

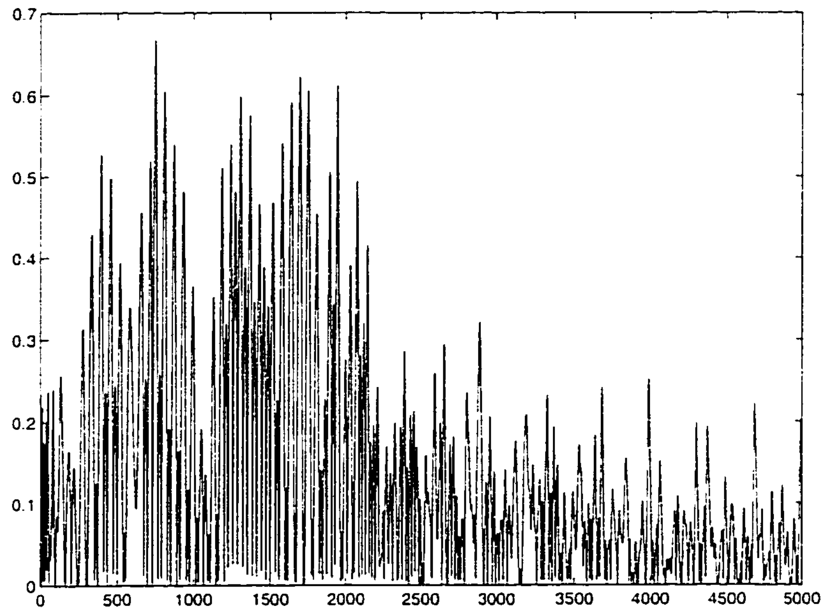

Figure 8.6: Absolute value of the first coordinate at a grid point on the 3rd latitude line

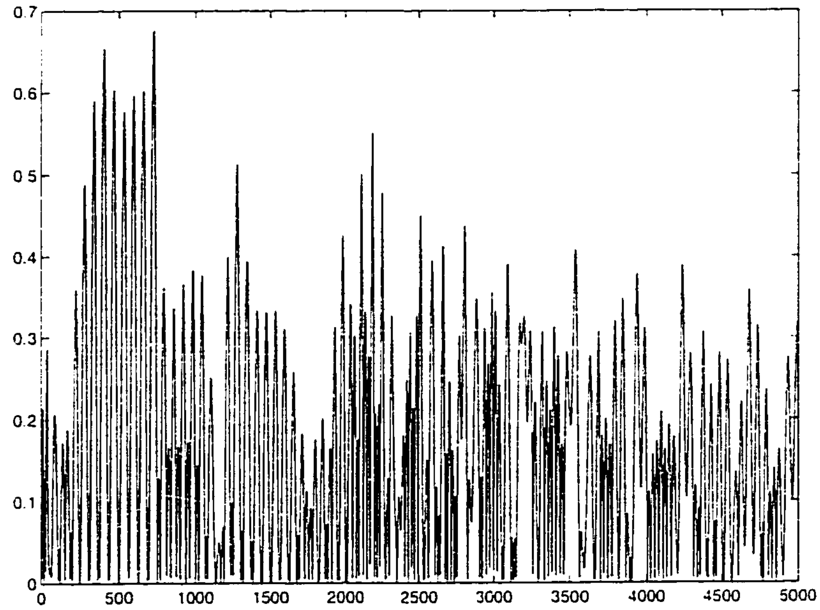

Figure 8.7: Absolute value of the first coordinate at a grid point on the 8thth latitude line

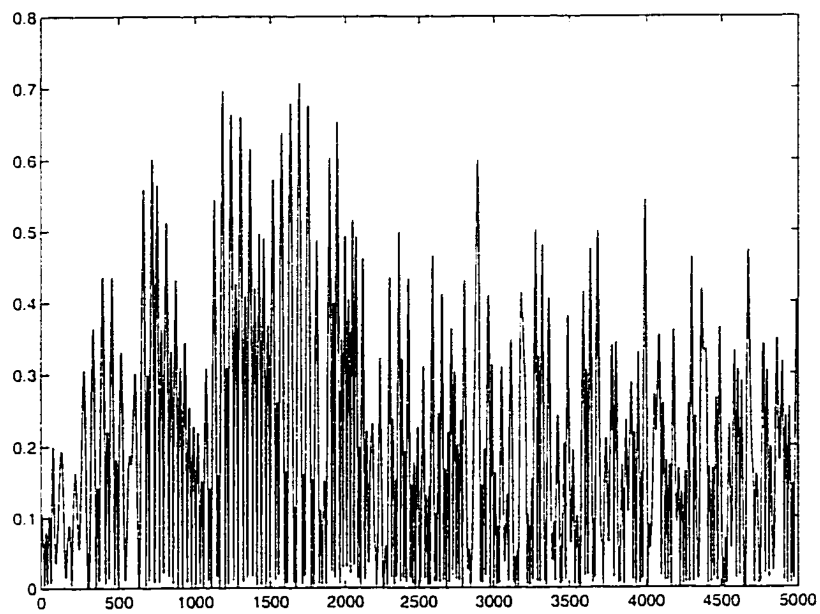

Figure 8.8: Absolute value of the first coordinate at the north pole latitude line

8.3.3 Following the Algorithm

Having collected a suitable data set, we can now proceed to follow the algorithm to find the KL eigenvectors and eigenvalues.

First we will summarize the elements involved. We are going to find a basis of eigenvectors of the ensemble average covariance matrix which optimally approximates the data set for each truncation dimension. The data set U is collected by discretizing elements of the manifold of tangent vector fields on the sphere, denoted by $\overline{K}$. The group G acting on this manifold is the special orthogonal group SO(3).

The G-invariant inner product on this manifold is defined by

$$(u, v) = \int_G \int_{S^2} (g \cdot u(\underline{x}))^T \overline{(g \cdot v(\underline{x}))} d\underline{x} dg.$$

Fortunately this integral simplifies to

$$(u, v) = \int_{S^2} u(\underline{x})^T \overline{v(\underline{x})} d\underline{x}.$$

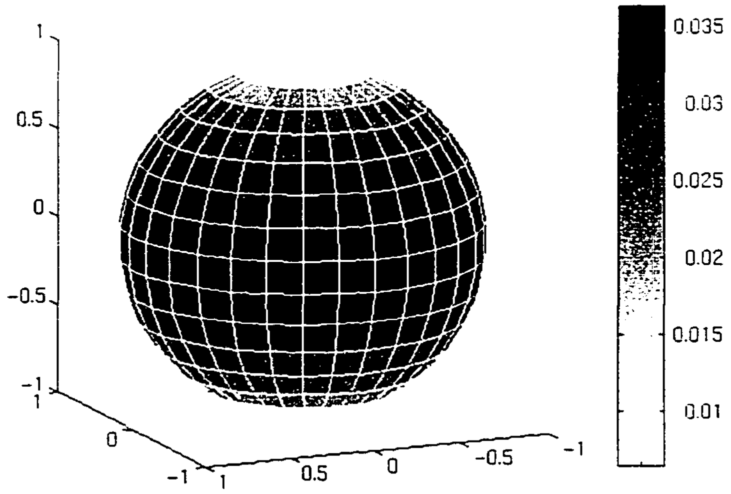

Figure 8.9: Weights for the discretization of the integral. The darker the color, the bigger the value.

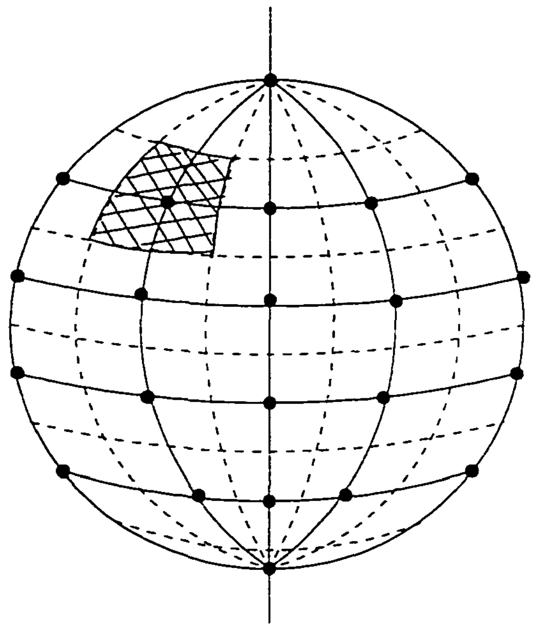

Figure 8.10: Graphical interpretation of the weights

Since we are not working with the actual functions, but with a discretization, we also have to discretize the integrals over the sphere. The discretization of the sphere is very regular with respect to the spherical coordinates. However, different points on the sphere are different distances apart, depending on where they are positioned. They are closer together near the poles than they are near the equator. Therefore it makes sense to weigh function values near the equator more heavily than those near the poles. Mathematically this weighting of gridpoints and their values happens when we discretize the integral. The integral $\int_{S^2} f(m)dm$ becomes $\sum_i f(m_i)dm_i$, where the terms dm_i are exactly the mentioned weights. Graphically, the weights are the area of little patches with one of the grid points as center, as shown in figure 8.10. Since we are integrating over the surface of the sphere it makes sense that the sum of all the weights is equal to the surface area of the sphere of radius one. Note that the weights are the same for all grid points on the same latitude line. A graphical view of the weight sizes can be seen in figure 8.9.

Now we can follow the **Algorithm for Exploiting Symmetry in the KL basis computation** step by step. Since our group SO(3) has infinitely many finite dimensional irreducible representations, we will use the algorithm to find the vectors of the KL basis that span the first 23 dimensional subspace, which form a basis for the linear, quadratic, and cubic tangent vector fields: $W = W_3^{\oplus 2} \oplus W_5^{\oplus 2} \oplus W_7$.

1. In step one of the algorithm we find the G-adapted basis for the 23 dimensional space W by first finding an orthonormal basis for the first symmetry subspace in each irreducible representation and then transferring them into the remaining symmetry subspaces. Note that the first symmetry subspace in the 3 and 5 dimensional irreducible representation is two dimensional, and there are 3 and 5 symmetry subspaces, respectively. Since we are only interested in one copy of the 7 dimensional irreducible representation, each symmetry subspace is one dimensional, and there are 7. Thus the basis is given by $\{v_{1,1}^3, \ldots, v_{3,2}^3; v_{1,1}^5, \ldots, v_{5,2}^5; v_{1,1}^7, \ldots, v_{7,1}^7\}$. The basis vectors are complex-valued, tangent vector fields where each component is a homogeneous polynomial of degree one, two, or three. The real and complex parts of selected basis vectors are shown in figures (8.12-8.21).

2. In step two we compute the functionals $\gamma_{k,\alpha}^3$, for $k = 1, 2, 3$, and $\alpha = 1, 2$; $\gamma_{k,\alpha}^5$, $k = 1, \ldots, 5$, and $\alpha = 1, 2$; $\gamma_{k,\alpha}^7$, $k = 1, \ldots, 7$, and $\alpha = 1$. These functionals are defined by $\gamma_{k,\alpha}^\rho : U \to \mathbb{C}$, $\gamma_{k,\alpha}^\rho = (v_{k,\alpha}^\rho, u)$, where $(\cdot, \cdot)$ is the G-invariant inner product.

3. For each of the three irreducible representations we find the block matrices B^ρ:

$$(B^\rho)_{ij} = \frac{1}{\dim(\rho)\,\mathrm{vol}(U)} \sum_{k=1}^{\dim(\rho)} \int_U \gamma_{ki}^\rho(u)\overline{\gamma_{kj}^\rho(u)}\, du.$$

Note that B^3 and B^5 are 2-by-2 and B^7 is 1-by-1. The matrices are:

$$B^3 = \begin{bmatrix} 0.01634 & -0.00002086 - 0.00129i \\ 0.00002086 + 0.00129i & 0.01628 \end{bmatrix}.$$

$$B^5 = 10^{-3} \cdot \begin{bmatrix} 0.22387 & -0.02451 + 0.01095i \\ -0.02451 - 0.01095i & 0.22637 \end{bmatrix}.$$

$$B^7 = 10^{-4} \cdot 0.34041.$$

4. Having computed the matrices we can now compute the eigenvalues and eigenvectors:

$$\lambda_1^3 = 0.017598, \quad \lambda_2^3 = 0.015017$$

$$w_1^3 = \begin{bmatrix} 0.71474 \\ 0.01131 - 0.6993i \end{bmatrix}, \quad w_2^3 = \begin{bmatrix} -0.69939 \\ -0.00559 - 0.71464i \end{bmatrix}.$$

$$\lambda_1^5 = 10^{-3} \cdot 0.25199, \quad \lambda_2^5 = 10^{-3} \cdot 0.19825$$

$$w_1^5 = \begin{bmatrix} 0.69047 \\ -0.66045 + 0.295061i \end{bmatrix}, \quad w_2^5 = \begin{bmatrix} 0.72336 \\ 0.63041 - 28164i \end{bmatrix}.$$

$$\lambda^7 = 10^{-5} \cdot 0.34041, \quad w^7 = 1.$$

5. Finally we can use the eigenvectors from the previous step to put together the KL basis vectors:

$$z_{k\alpha}^{\rho} = \sum_{\beta} w_{\alpha\beta}^{\rho} v_{k\beta}^{\rho}.$$

Thus the eigenvectors are linear combination of the basis vectors of the symmetry subspaces. There are two eigenvectors for each symmetry subspace of the three and five dimensional irreducible representations. The eigenvectors in the seven dimensional irreducible representation are exactly the corresponding vectors of the G-adapted basis.

8.3.4 Discussion of Results

Given below also are the computational results we received by applying the algorithm on a data set of 2500 vectors we obtained by only considering the second half of the previous data set. Note that the eigenvalues are very close to the ones in the previous case.

$$B^3 = \begin{bmatrix} 0.0163 & 0.0012i \\ -0.0012i & 0.0163 \end{bmatrix}$$

$$\lambda_1 = 0.0175, \quad \lambda_2 = 0.0151$$

$$v_1 = \begin{bmatrix} 0.70711 \\ -0.70711i \end{bmatrix}, v_2 = \begin{bmatrix} -0.70711 \\ -0.70711i \end{bmatrix}.$$

$$B^5 = 10^{-3} \cdot \begin{bmatrix} 0.2239 & -0.0245 + 0.011i \\ -0.0245 - 0.011i & 0.2264 \end{bmatrix}$$

$$\lambda_1 = 10^{-3} \cdot 0.252. \quad \lambda_2 = 10^{-3} \cdot 0.1982$$

$$v_1 = \begin{bmatrix} -0.69047 \\ 0.6599 + 0.29628i \end{bmatrix}, v_2 = \begin{bmatrix} 0.72336 \\ 0.6299 + 0.28281i \end{bmatrix}.$$

$$B^7 = \lambda = 10^{-4} \cdot 0.34041$$

In both examples we can see that the eigenvalues decrease in size rapidly from one irreducible representation to the next. That indicates that most of the energy of the data lies in the first 6 dimensional subspace, spanned by three linear and three quadratic tangent vector fields. We gain some more accuracy by adding the next 10 basis vectors in the system and the following 7 from the 7 dimensional irreducible representation. This result is supported by computing the normalized mean square error of the data.

Error of the approximation

Given below are the mean square errors of the approximated data in several subspaces for two different data sets. The spaces we are approximating in are:

- The space of all linear vector fields, which is exactly one copy of the three-dimensional irreducible representation W_3.

- The two copies of the three-dimensional irreducible representation $W_3^{\oplus 2}$. In addition to the linear vector fields it also contains certain quadratic vector fields.

- The space of all linear and quadratic vector fields, which is the direct sum of the two three dimensional irreducible representations plus one 5-dimensional representation.

- The space which is the direct sum of the two three-dimensional and the two 5-dimensional irreducible representations. It consists of all linear and quadratic vector fields, as well as some of the cubic vector fields: $W_3^{\oplus 2} + W_5^{\oplus 2}$.

- The space of all linear, quadratic, and cubic vector fields $W_3^{\oplus 2} + W_5^{\oplus 2} + W_7$.

Mean Square Error for a data set of 5000 vectors:

- the 3-dimensional subspace of linear vector fields: 52.82%,

- the 6-dimensional subspace of $W_3^{\oplus 2} = 5.74\%$,

- the 11-dimensional subspace of linear and quadratic vector fields: 4.3%,

- the 16-dimensional subspace of $W_3^{\oplus 2} + W_5^{\oplus 2} = 2.83\%$,

- the 23-dimensional subspace of $W_3^{\oplus 2} + W_5^{\oplus 2} + W_7 = 2.63\%$.

Mean Square Error for a data set of 2500 vectors:

- the 6-dimensional subspace of $W_3^{\oplus 2} = 6.43\%$.
- the 16-dimensional subspace of $W_3^{\oplus 2} + W_5^{\oplus 2} = 4.21\%$.
- the 23-dimensional subspace of linear. quadratic. and cubic vector fields: $W_3^{\oplus 2} + W_5^{\oplus 2} + W_7 = 3.97\%$.

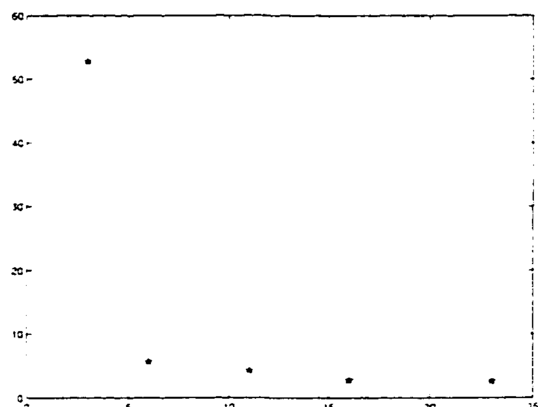

Figure 8.11: Mean square error of the 5000 element data set in the different subspaces.

It appears that most of the characteristics of the data are captured in the first irreducible representation. There is an even split in how much is captured by the linear vector fields and by the quadratic vector fields in this representation. The accuracy increases notably by adding the next 10 KL basis vectors to the approximating subspace. The 7-dimensional irreducible representation. however. only captures a fraction of the data. The accuracy is only increased by a fraction of a percent when we add those 7 KL basis vectors.

It would be interesting to repeat this experiment with a bigger data set of several tens of thousands of vectors.

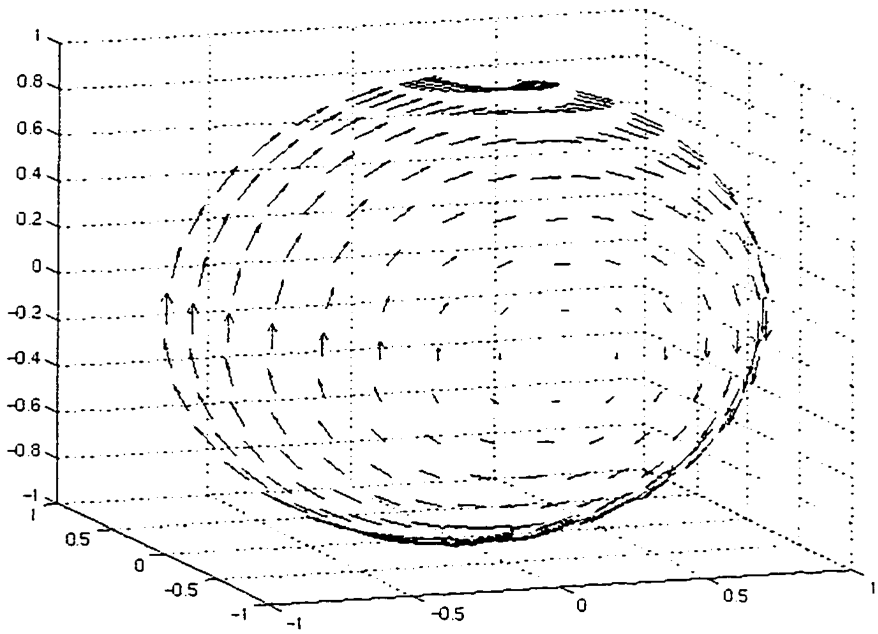

Figure 8.12: Linear basis vector. 3-dimensional irreducible representation. real part

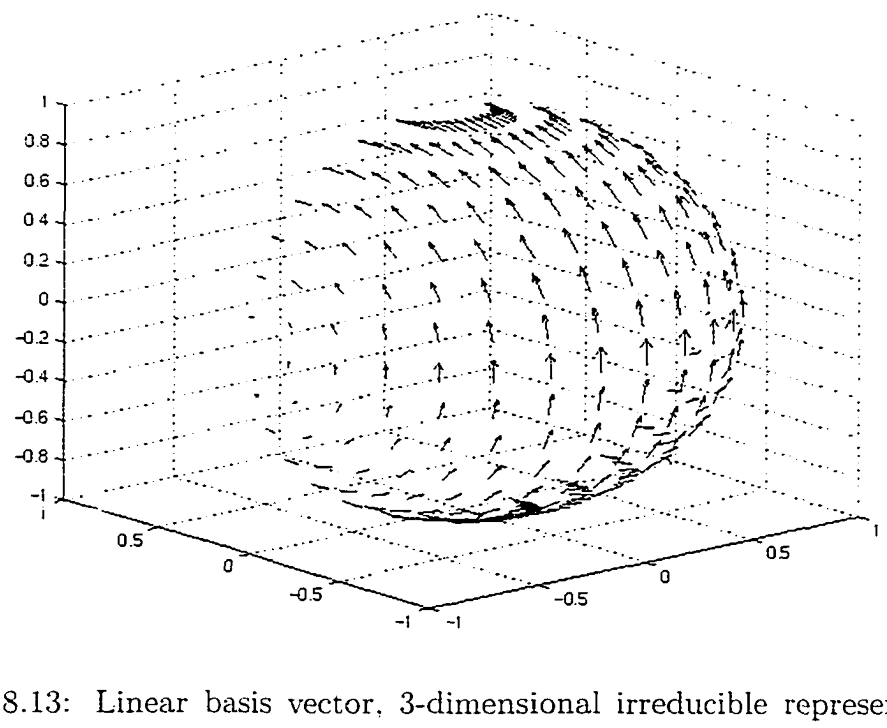

Figure 8.13: Linear basis vector, 3-dimensional irreducible representation. imaginary part

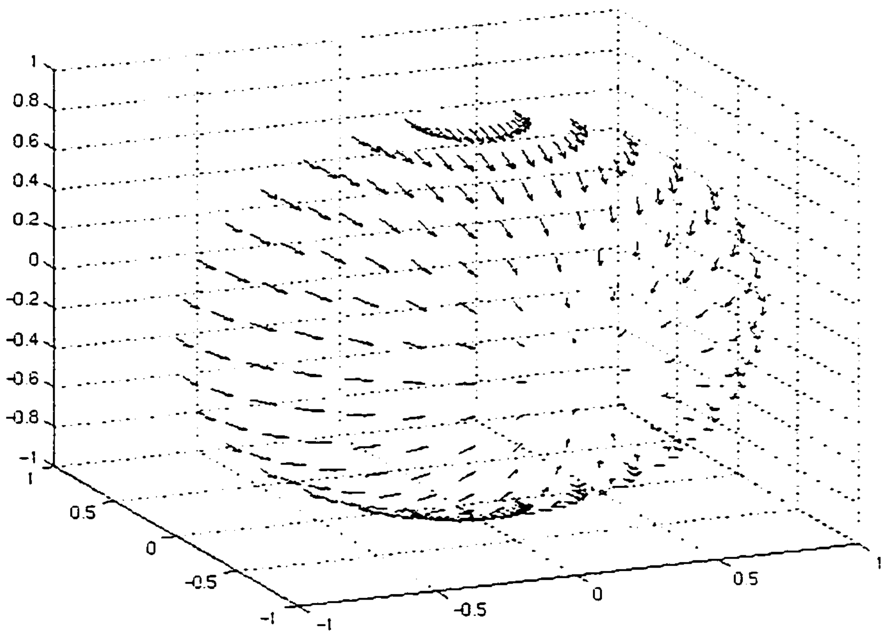

Figure 8.14: Quadratic basis vector. 3-dimensional irreducible representation. real part

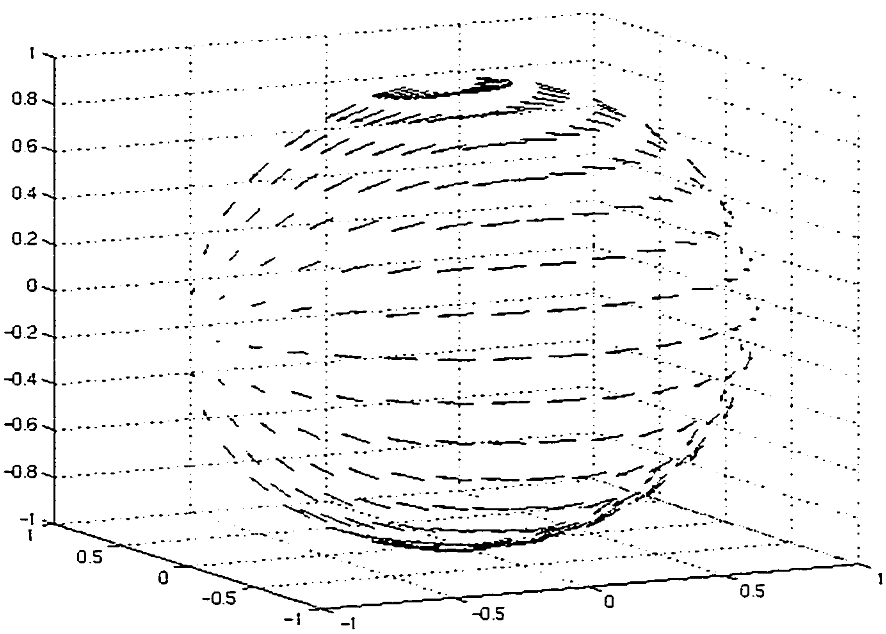

Figure 8.15: Quadratic basis vector, 3-dimensional irreducible representation, imaginary part

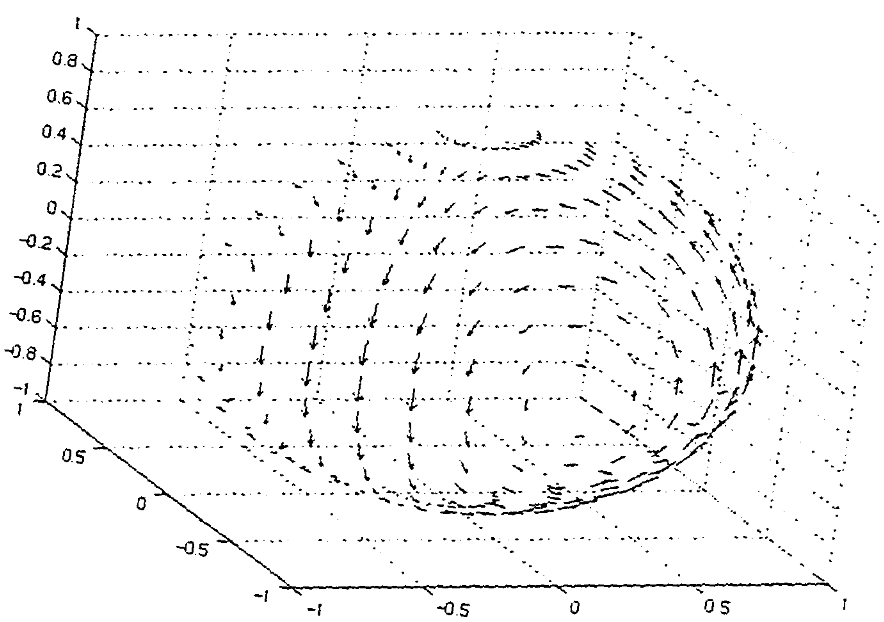

Figure 8.16: Quadratic basis vector, 5-dimensional irreducible representation, real part

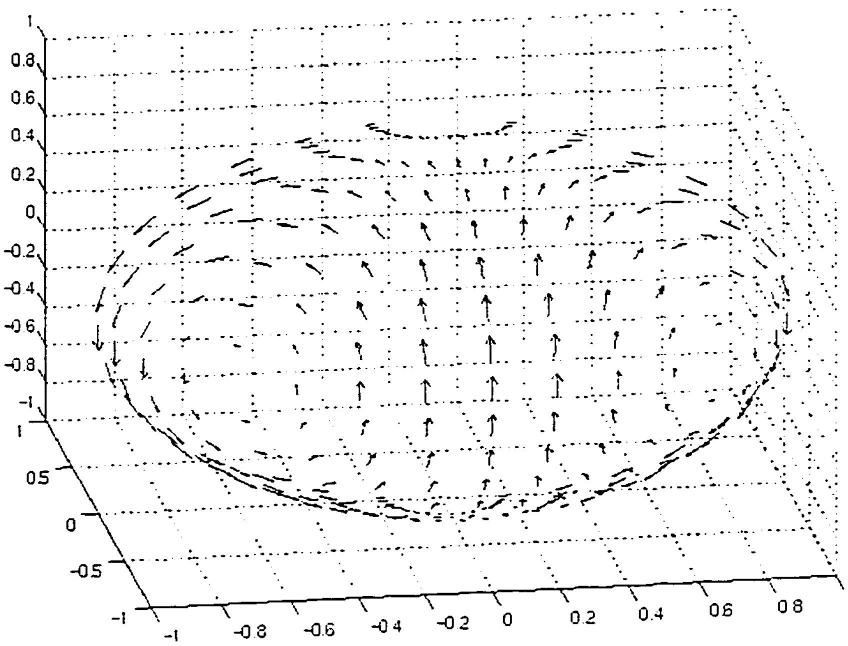

Figure 8.17: Quadratic basis vector. 5-dimensional irreducible representation. imaginary part

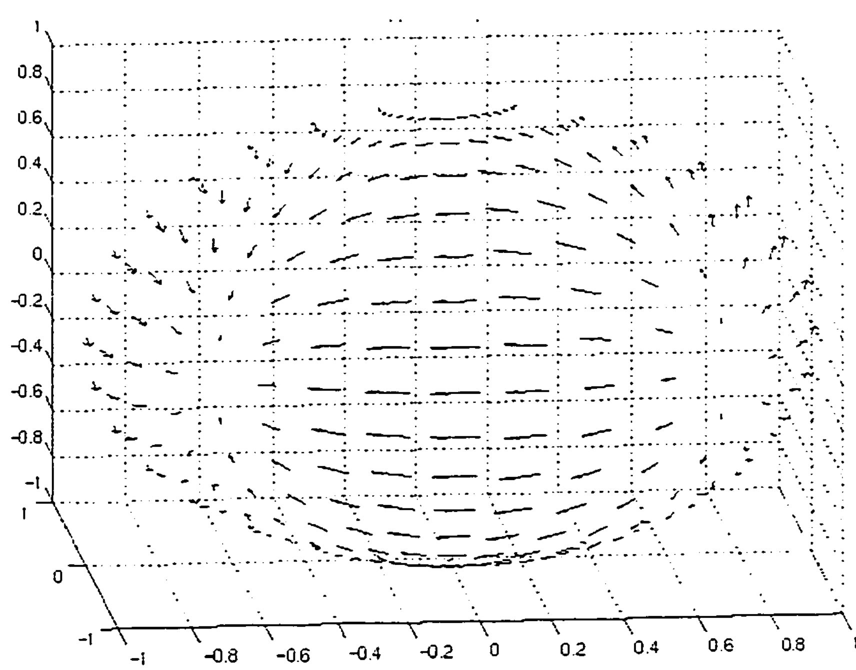

Figure 8.18: Cubic basis vector, 5-dimensional irreducible representation, real part

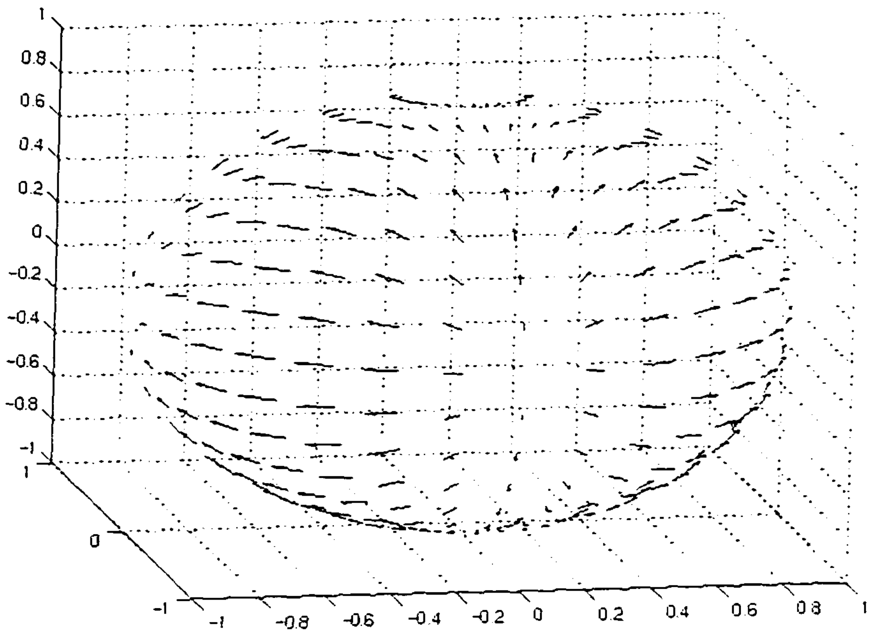

Figure 8.19: Cubic basis vector. 5-dimensional irreducible representation. imaginary part

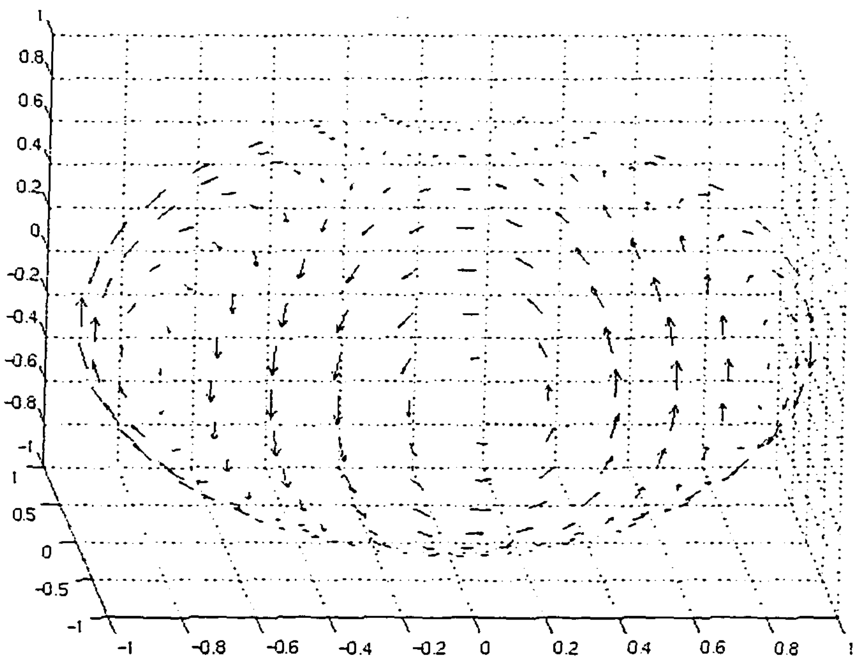

Figure 8.20: Cubic basis vector, 7-dimensional irreducible representation. real part

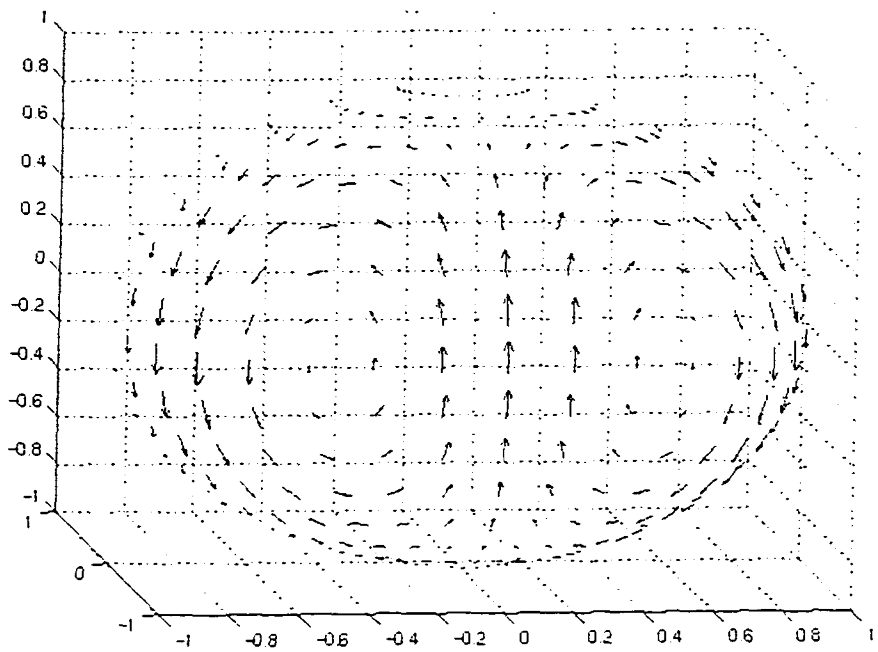

Figure 8.21: Cubic basis vector. 7-dimensional irreducible representation, imaginary part

Chapter 9

Conclusion

It was the goal of this thesis to link two apparently competing decompositions of a vector space to reduce the cost of approximating a symmetric data set in a low dimensional subspace. We also wanted to gain some insight into the structure of the data.

The two decompositions were the Karhunen-Loève decomposition and the decomposition of the vector space into symmetry subspaces with the help of group representation theory. The second decomposition was suggested by the symmetry of the data.

We began with the goal of finding a way to reduce the cost of the KL decomposition with group theoretical considerations. First we restricted our research to finite groups acting on finite dimensional vector spaces. We then succeeded in proving that we could find a basis for the vector space using group representation theory, which brought the ensemble average covariance matrix for the symmetrized data set into block form. We also found a very efficient way of computing the entries of each block, and thus managed to reduce the cost of computing the eigenvalues and eigenvectors of the covariance matrix considerably. Each block corresponded to one symmetry subspace. We succeeded in moving the eigenvalue and eigenvector computations from the high dimensional vector space into much lower dimensional subspaces.

Using these results we formulated an exact algorithm for exploiting symmetry in the KL basis computation.

We then demonstrated how much the computational cost was reduced by our algorithm compared to traditional methods. In general the ratio of a traditional method to our algorithm is

$$\alpha(G) = \frac{ad^3}{ad^3 \sum_\rho \dim(\rho)^3/|G|^3} = \frac{|G|^3}{\sum_\rho \dim(\rho)^3}.$$

where G is the symmetry group of the data and ρ denote the irreducible representations of G.

We then proceeded to illustrate the algorithm using a data set obtained by numerically solving a Lorenz type system of ordinary differential equations. This data set enjoyed S_3 symmetry. It was shown that 95% of the energy of the 7 dimensional system was concentrated in a 3 dimensional subspace. The efficiency factor in this particular example was 9.8. This example concluded the investigation of finite groups.

We then continued by extending our results to the case of compact Lie groups acting on a manifold $\mathcal{M}$ and thus on the space of continuous $\mathbb{C}^n$-valued functions $\mathcal{F}^n(\mathcal{M})$. We had to adjust our tool and theory to this situation, which mainly consisted of replacing sums in averaging processes and inner products with integrals and Haar integrals if necessary.

For this new application of the algorithm we presented several examples illustrating the procedures. The most serious and detailed example consisted of data in the manifold of vector fields tangent to the sphere which were acted upon by the special orthogonal group SO(3). We generated a 5000 element data set by numerically solving a vector valued version of the complex Ginzburg Landau equation. Using our algorithm to find the KL basis vectors we succeeded in approximating the data set with an error of 2.63% in a 23 dimensional subspace of the original vector space of tangent vector fields.

Applications for the method we described in this dissertation arise naturally in the fields of pattern analysis and image/signal processing. Whenever it is necessary to work with images it is of interest to represent them in an optimal manner with the least computational cost. The KL decomposition provides the optimal representation of a set of images. and the described algorithm reduces the computational cost of the method whenever the images exhibit symmetry. Thus our method can be applied to any kind of data enjoying symmetry. i.e. atmospheric data. readings from the earth core, solar data, images of faces. snow flakes. houses. and plants. Our method could also be used to detect symmtery in a data set. or the degree of symmetry present.

Bibliography

[AG] R.B. Ash and M.F. Gardner: *Topics in Stochastic Processes*. Academic Press. New York, 1975.

[AM] M.F. Atiyah and I.G. MacDonald: *Introduction to Commutative Algebra*. Addison-Wesley Publishing Company. 1969.

[BD] T. Bröcker. T. tom Diek: *Representations of Compact Lie Groups*. Springer Verlag, New York 1984.

[Bl] P.J. Blenerhassett: "On generation of waves by wind". Philosophical Transactions of the Royal Society of London. 298A. p- 451-494. 1980.

[Ci] P. G. Ciarlet: *Introduction to Numerical Linear Algebra and Optimization*. Cambridge Texts in Applied Mathematics. Cambridge University Press Cambridge (1982).

[CR] C. Curtis and I. Reiner: *Representation Theory of Finite Groups and Associative Algebras*. Wiley Interscience 1962.

[FH] W. Fulton and J. Harris: *Representation Theory*, Springer Verlag New York. 1991.

[Fu] K. Fukunaga: *Introduction to Statistical Pattern recognition*, Academic Press. second edition, 1990.

[FV] H. Freudenthal. H. de Vries: *Linear Lie Groups.* Academic Press. New York London 1969.

[Ge] S. Geiger: "Effiziente Simulation und Modellierung sph"arisch symmetrischer partieller Differentialgleichungen". Diplomarbeit. Institut f"ur Theoretische Physik. Eberhard-Karls-Universit"at T"ubingen. 1996.

[GMS] I.M. Gelfand. R.A. Minlos. Z.Ya. Shapiro: *Representations of the Rotation and Lorentz Groups and Their Applications.* Pergamon Press Book. The Macmillan Company. New York 1963.

[GSS] M. Golubitsky. I. Stewart. D.G. Schaeffer: *Singularities and Groups in Bifurcation Theory.* Volume 2. Springer Verlag New York. 1988.

[He] W. Hein: *Einführung in die Struktur- und Darstellungstheorie der klassischen Gruppen.* Springer Verlag Berlin Heidelberg. 1990.

[Ho] H. Hotelling: "Analysis of a complex of statistical variables into principal components". J. Educational Psychol.. Sept. 1933.

[Jo] I.T. Jolliffe: *Principle Component Analysis.* Springer-Verlag New York. 1986.

[Ka] K. Karhunen.: "Über lineare Methoden in der Wahrscheinlichkeitsrechunung". Ann. Acad. Sci. Fennicae. ser. A1. Math. Phys.. vol. 37. 1946.

[KD] S. Kogelman. R.C. DiPrima: "Stability of spatially periodic supercritical flows in hydrodynamics". Phys. Fluids 13. p. 1. 1970.

[Ke] L.A. Keefe: "Dynamics of perturbed Wavetrain Solutions to the Ginzburg-Landau Equation", Studies in Appl. Mathematics 73, p. 91-153, 1985.

[Ki] M. Kirby: *Pattern Analysis and Dimensionality Reduction*. Wiley & Sons. to appear.

[KK] Y. Kuramoto. S. Koga: "Anomalous peeriodic doubling bifurcations leading to chemical turbulence". Physics Letters A 92. p. 1. 1982.

[KS] M. Kirby. L. Sirovich: "Application of the Karhunen Loeve Procedure for the Characterization of Human Faces". IEEE Trans. of Pattern Analysis and Machine Intelligence. Vol. 12. no. 1. (1990) 103-108.

[Ku] Y. Kuramoto: "Diffusion induced chaos in reaction systems" Progress of Theoretical Physics Suppl. 64. p. 346. 1978.

[KY] Y. Kuramoto. T. Yamada: "Turbulent state in chemical reactions" Progress of Theoretical Physics 56. p. 679. 1976.

[La] S. Lang: *Complex Analysis, Third Edition*. Graduate Texts in Mathematics No. 103. Springer-Verlag New York. 1993

[Lo] M.M. Loeve: *Probability Theory*. Princeton. NJ: Van Nostrand. 1955.

[MW] I. Mezić. S. Wiggins: "On the Integrability and Perturbation of Three-Dimensional Fluid Flows with Symmetry". Journal of Nonliear Science. Vol. 4. p 157-940. 1994.

[MHR] H.T. Moon. P. Huerre. L.G. Redekopp: "Transitions to Chaos in the Ginzburg-Landau Equation". Physica D 7. p 135-150. 1983.

[Ne] A.C. Newell: "Envelope Equations", Lect. Appl. Math. 15, p.157, 1974.

[NW] A.C. Newell, J.A. Whitehead: "Finite amplitude, finite bandwidth convection". Journal of Fluid Mechanics 38, p. 279, 1969.

[Pe] K. Pearson: "On lines and planes of closest fit to systems of points in space". Phil. Mag., 6th series, 1901.

[SA] N. Smaoui and D. Armbruster: "Symmetry and the Karhunen Loève analysis" preprint 1996.

[Se] J. P. Serre: *Linear Representations of Finite Groups*. Graduate Texts in Mathematics, Vol. 42, Springer-Verlag, New York, 1977.

[SF] E. Stiefel, A. Fässler: *Gruppentheorethische Methoden und ihre Anwendungen*. Teubner Studenbücher Mathematik, Stuttgart, 1979.

[Si] L. Sirovich: "Turbulence and the dynamics of coherent structures I (Coherent Structures)". Quarterly of Applied Mathematics, Vol. XLV, No. 3, (1987) 561-571.

[SK] L. Sirovich, M. Kirby: "Low-dimensional procedure for the characterization of human faces". Journal of the Optical Society of America, Vol. 4, No. 3 (1987) 519-524.

[SR] L. Sirovich, D. Rodriguez: "Coherent structures and chaos: a model problem". Physics Letters A 120, p. 211-214, 1987.

[SS] K. Stewartson, J.T. Stuart: "Non-linear instability of plane Poiseuille flow". Journal of Fluid Mechanics 48, p 529ff, 1971.

[Vi] E. Vinberg: *Linear Representations of Groups*. Basler Lehrbücher Vol. 2, Birkhäuser Verlag, Basel 1989.

[Wa] S. Watanabe: "Karhunen Loeve expansion and factor analysis". In Trans. 4th. Prague Conf. on Inform. Theory, pages 635-660, Praque, 1965.