GEOMETRICAL TRANSFORMATIONS IN HIGHER DIMENSIONAL EUCLIDEAN SPACES

A Thesis

by

AMIT KUMAR SANYAL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2001

Major Subject: Aerospace Engineering

GEOMETRICAL TRANSFORMATIONS IN HIGHER DIMENSIONAL EUCLIDEAN SPACES

A Thesis

by

AMIT KUMAR SANYAL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

John L. Junkins (Chair of Committee)	Tom C. Pollock (Member)
Aniruddha Datta (Member)	Kyle T. Alfriend (Head of Department)

May 2001

Major Subject: Aerospace Engineering

ABSTRACT

Geometrical Transformations in Higher Dimensional Euclidean Spaces. (May 2001)

Amit Kumar Sanyal,

B.Tech., Indian Institute of Technology, Kanpur

Chair of Advisory Committee: Dr. John L. Junkins

Orientations and rotations in n-dimensional real Euclidean spaces ($\mathbf{R}^n$) are represented by proper orthogonal, or skew-symmetric matrices. A mathematical formulation that leads to these representations is presented. Orientations and rotations are indistinguishable in 2 and 3 dimensions. In higher dimensions, orientations can be achieved by a minimal set of rotations. This result is presented here as the generalization of Euler's Principal Rotation Theorem to higher dimensions. Three types of skew-symmetric orientation and rotation matrices are presented. Decompositions of orientation matrices, in terms of rotation matrices, are also presented. Comparisons are drawn between these matrix representations of rotations and orientations. The *ortho-skew* matrices, which are both orthogonal and skew-symmetric, is introduced as a special set of orientation matrices.

Symmetric matrices often arise in linear systems theory and estimation. They represent reflections and projections (both orthogonal and non-orthogonal), in Euclidean spaces. The *ortho-symmetric* matrices, which are both orthogonal and symmetric, are introduced. These matrices represent reflections in Euclidean spaces. The Householder matrices, often encountered in linear algebra problems, belong to this set and represent elementary reflections. A general symmetric matrix can be decomposed as a sum of scalar multiples of a set of Householder matrices. Elementary projections in $\mathbf{R}^n$ can be represented by a set of symmetric matrices, called the modified Householder matrices, introduced here. These

matrices are a natural choice for decomposing symmetric matrices. This decomposition closely parallels the decomposition of orientation matrices by rotation matrices.

The last part of this thesis deals with the matrix Riccati differential equation with symmetric coefficients, also known as the symplectic matrix Riccati differential equation (SRDE). This equation, along with the related but simpler Lyapunov equation, arises quite frequently in optimal control theory and estimation theory. A solution procedure, which solves the time-varying SRDE by extension to a symplectic flow field, and utilizes the properties of symplectic matrices, is presented here. This solution can be related to the analytic singular value decomposition of the time-varying symmetric matrix solution.

To my parents,

and my brother, Ashis,

for their love and support

ACKNOWLEDGEMENTS

I am thankful to my committee chairman Dr. John Junkins for providing me an opportunity to pursue this research, and providing improvements to this work during its preparatory stages. Special thanks are also due to Dr. Daniele Mortari of the University of Rome, whose work on rotations and orientations on higher dimensional Euclidean spaces began this research. I am extremely grateful to him for providing me with a topic of research.

I am thankful to all the teachers who have taught me here at Texas A&M University, from whom I have gained a lot of valuable knowledge. I especially thank Dr. Aniruddha Datta, from whom I gained most of my present knowledge of modern control theory. I am also thankful to all my teachers at the Indian Institute of Technology, Kanpur, who provided me with the basic knowledge needed to pursue higher studies in aerospace engineering. I would particularly thank Dr. Ashish Tewari for his support and belief in my abilities to become a successful researcher.

My sincere thanks to the staff at the aerospace engineering department of Texas A&M, especially Ms. Lisa Willingham and Ms. Karen Knabe, for all the secretarial and administrative help provided during my stay.

I am also grateful to my graduate school friends Malak Samaan and Sesha Sai Vaddi, for all the help, encouragement and good wishes they have provided.

I would thank my best friends, Ashish Saxena and Yogesh Shrivastava, for being there whenever I needed any form of help. I have gained immensely from our friendship since our undergraduate days, and from our fruitful discussions. Finally, I thank my parents and family for their continuous love, support and understanding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Coordinate transformations and other simple geometric transformations in Euclidean spaces are often used in many physical problems to facilitate analysis. Such transformations are represented in the form of square matrices; mainly orthogonal and symmetric matrices. Most of the standard matrix decompositions in matrix analysis are also in the form of products of orthogonal and diagonal or triangular matrices. Thus the study of these geometrical transformations in Euclidean spaces is of importance in many problems of engineering interest. This thesis first presents studies for rotations/orientations and reflections/projections in Euclidean spaces, which are represented by orthogonal/skew-symmteric and symmetric matrices respectively. This thesis then studies the matrix Riccati Differential Equation with symmetric form, also called the Symplectic Riccati Differential Equation (SRDE), and presents a novel method of solving the equation by extension of its domain.

Euclidean spaces, being linear spaces, can be represented by a single coordinate chart or set of basis vectors that span the space. The best choice for a basis vector set is an orthogonal one, in which the basis vectors are mutually orthogonal. Changes of coordinates are often useful and necessary when dealing with motions of a body in familiar three-dimensional Euclidean space or higher-dimensional dynamical systems. A change of coordinates can be achieved by an orthogonal transformation that changes only the components along the different basis vectors of a point in the space, but preserves lengths and angles between vectors. The first part of this thesis deals with rotations and re-orientations, which are the most common geometrical transformations carried out in an Euclidean space. Chapter II begins with a treatment on rotations, which are the

---

The journal model is *AIAA Journal of Guidance, Control and Dynamics.*

simplest form of orthogonal transformations, where the transformation is confined to a plane. Only the components of a vector along this plane are changed by the rotation. In 3 dimensions, there is no distinction between rotation on a plane and rotation about an axis orthogonal to this plane. For higher dimensions, it is necessary to consider rotations as occurring on a plane, as will be made evident. Representations for rotations by orthogonal and skew-symmetric matrices, and the relations between them, are detailed here. An example of a rotation in a 5-dimensional space is also presented, which shows the length and angle preserving and planar nature of rotations.

Re-orientations, which are the most general orthogonal transformations that preserve lengths and angles between vectors, are also dealt with in chapter II. The difference between rotations and re-orientations, which are more general orthogonal transformations, is stressed in this thesis. A very important and widely used result for re-orientations in three dimensions is Euler's Theorem on the motion of a rigid body with one point fixed, which was published in 1775.[1] In our familiar 3-dimensional universe, rotations cannot be distinguished from re-orientations, a fact also well-known from Euler's Theorem. However, this theorem had not been generalized satisfactorily to higher dimensions till very recently,[2,3] though there had been prior work done on this topic.[4] Although Euler's Theorem is very useful, and has given rise to many representations of a rigid body's attitude[5–7] in 3 dimensions, it does not identify the true nature of rotation. Subsequent modified statements of the theorem identified rotations with axes instead of planes, which holds true for only three dimensions, and this in fact, prevented its generalization for over two centuries. This thesis provides the generalization of Euler's Theorem to higher dimensions, as was first presented in Ref. 2. This is done by showing that the basic constituents of re-orientations are rotations, and that any general re-orientation of a body in an Euclidean space can be achieved by a set of rotations. The ortho-skew matrices, which represent a special set of orientation matrices, are introduced at the end of chapter II.

A detailed study of reflections and projections in Euclidean spaces forms the next part of this thesis. Reflections and projections are also simple geometric transformations in Euclidean spaces which, unlike rotations and re-orientations, generally do not preserve lengths or angles between vectors. The second part of this thesis deals with reflections and projections, and shows that they can be represented algebraically by symmetric matrices. Symmetric matrices are frequently encountered in problems of control and estimation of dynamical systems. Chapter III begins with a treatment of reflections, which are the simplest form of projections, where the length of a vector undergoing the projection is preserved. Unlike a rotation, which always occurs on a plane, a reflection is a transformation in an Euclidean space which reflects objects along a linear flat subspace (a hyperplane), in the space. The line joining a point and its reflection along the hyperplane, is parallel to this hyperplane and is bisected by the subspace orthogonal to it. The subspace or hyperplane along which the reflection occurs, can be of any dimension from 1 to n, where n is the dimension of the space (the n-dimensional case is a transformation which reverses the direction of every vector in the space). It is shown that reflections in Euclidean spaces can be represented by the set of ortho-symmetric matrices, which are at once both symmetric and orthogonal. The Householder matrices[8,9], which are often used in numerical linear algebra routines and are also called elementary reflection matrices, belong to this set and act as reflections along an axis (a 1-dimensional linear subspace). A decomposition of symmetric matrices by Householder matrices is also presented in this thesis.

Projections, which are more general symmetric transformations than reflections, are dealt with in the latter part of chapter III. Unlike the case of re-orientations, there is no classical result like Euler's Theorem for projections in Euclidean spaces. Reflections are a special type of projections, just like rotations are a special type of re-orientations. However, the relation between reflections and projections is not analogous to that between rotations and re-orientations.

Projections, like reflections can act along a linear subspace of any dimension from 1 to n, where n is the dimension of the Euclidean space, with the n-dimensional case shown to be represented by the most general symmetric matrix. While reflections preserve lengths (linear dimensions), projections in general may not preserve either lengths or angles between vectors. They act as transformations that change the angle between a vector and a hyperplane in an Euclidean space. The normal distance of the tip of the vector from this hyperplane, however, remains unchanged by this transformation. It is known that projections onto subspaces, also known as orthogonal projections, can be expressed by symmetric matrices.[10−13] However, orthogonal projections, in which the vectors are projected onto a subspace, are a special case of the projections discussed in this thesis. Orthogonal projections are idempotent, i.e., subsequent applications of the same projection do not have any affect. The definition of projections presented in this thesis, however, covers non-orthogonal projections as well. It is shown in this thesis that these general projections in Euclidean spaces can be represented by symmetric matrices, which confirms the fundamental association of symmetric matrices with projections and reflections.

The last part of this thesis deals with the Symplectic Riccati Differential Equation (SRDE), which has a symmetric matrix as a solution. The SRDE often arises in problems of optimal control and estimation, and related fields like dynamic programming. Chapter IV begins with an introduction to the Riccati Differential Equation, and some of the applications it arises in. Then it provides a treatment of the Symplectic Riccati Differential Equation, which has a symmetric matrix solution. It is known that the subspace of symmetric matrices in the space of $n \times n$ matrices is an invariant manifold for this equation.[14] In optimal control and estimation, the solution sought from this equation is a symmetric, positive-definite, Hurwitz matrix with all eigenvalues negative.[10,15,16] Instead of solving the SRDE by direct numerical integration, the procedure detailed in this thesis uses the flow of the extended equation. The extended equation is formulated in the natural compactification of the vector space of real symmetric $n \times n$ matrices, called the Lagrange-Grassmann manifold.[14] Radon's formula for the solution of the SRDE[14,17] is shown to be related to the spectral decomposition (eigenvector-eigenvalue decomposition) of the symmetric matrix solution, which is equivalent to its singular value decomposition.[13] The solution given by the flow of the equation remains symmetric at all times if it is symmetric initially since the vector space of real symmetric matrices is an invariant manifold for the SRDE. The flow of the SRDE is symplectic in nature, which gives the equation its name. The flow is obtained from the Hamiltonian matrix[14,17] of the equation, which is infinitesimally symplectic. The solution procedure then solves for the symplectic flow, which gives the solution at any time t with known initial conditions at an initial time t_0. However, numerical errors may accumulate during numerical integration of the Hamiltonian matrix which may make the numerically calculated flow deviate substantially from symplecticity over a large range of integration. Hence, a numerical procedure to obtain the closest symplectic matrix to the numerical solution, is developed. This procedure corrects the numerical integration of the Hamiltonian matrix so that the result is always close to Hamiltonian (infinitesimally symplectic). This ensures that the flow, given by the matrix exponential of the Hamiltonian, is always close to symplectic.[14]

This thesis was motivated by the recent work of Mortari, refs. 2 and 3, and the chapter on rotations and re-orientations is essentially a re-formulation of the contributions of refs. 2 and 3. This material is the necessary conceptual and notational foundation upon which the subsequent original contributions of this thesis on symmetric matrix decompositions, parametrizations of linear reflections and projections, and the treatment of the symplectic Riccati differential equation are based. Relations between the three main parts of this thesis are made at appropriate locations in the body of the thesis, to help connect all the material together.

Prior to the detailed developments, the special notation frequently used in this thesis is introduced here. The n-dimensional (real) Euclidean space is represented by $\mathbf{R}^n$ while $n \times m$ real matrices are sometimes denoted by $\mathbf{R}^{n\times m}$. The null space of a matrix in $\mathbf{R}^{n\times m}$ is denoted by $null(A)$ and its column space is denoted by $col(A)$. Identity matrices in $\mathbf{R}^{n\times n}$ are denoted by I_n. The orthogonal rotation and re-orientation matrices are special orthogonal matrices with determinant $+1$, which are denoted by $\mathcal{O}^+(n)$ or $\mathcal{SO}(n)$. The $n \times n$ skew-symmetric matrices are denoted by $so(n)$ and the ortho-skew matrices are denoted by the symbol $\Im$ when their dimension is clear from the context. The $n \times n$ symmetric matrices are denoted by $\mathcal{S}(n)$ and the ortho-symmetric matrices are denoted by the symbol $\Re$ when their dimension is clear from the context. The Householder matrices are denoted by the symbol H and the modified Householder matrices are denoted by the symbol $\mathcal{M}$, while their dimensions are made clear from the context. A linear subspace that is the orthogonal complement of another linear subspace $N \in \mathbf{R}^n$ is denoted by $N^\perp$. The symplectic $2m \times 2m$ matrices are denoted by $\mathcal{S}p(m)$. The infinitesimally symplectic or Hamiltonian $2m \times 2m$ matrices are denoted by $sp(m)$. Other non-standard mathematical notation used are introduced prior to their use in the thesis.

CHAPTER II

ROTATIONS AND ORIENTATIONS

The orientation of a vector in $\mathbf{R}^n$ is given by the orthogonal components of the vector. A change in orientation of the vector is usually defined in a non-singular fashion by a proper orthogonal matrix $C \in \mathcal{SO}(n)$. For rigid-body kinematics, Euler's theorem establishes the equivalence between rotation and re-orientation in $\mathbf{R}^3$. The most commonly used version of this theorem[18] is given below.

Theorem 2.1 (Euler's Principal Rotation) *Any arbitrary orientation of a rigid body with one point fixed can be obtained by a single rotation about some axis through the fixed point.*

This theorem, however, does not hold for higher dimensional spaces, where rotations and re-orientations cannot be considered identical, and one cannot describe a general orientation by a "rotation about an axis." Rotations are planar in nature and hence can only be observed in Euclidean spaces with dimension greater than or equal to 2. In 2 and 3 dimensional spaces, it is well known that any given orientation can be arrived at by just a single rotation. While this is obvious in a 2 dimensional space (a plane), Euler's Theorem generalizes this concept to $\mathbf{R}^3$. This chapter generalizes Euler's Theorem even further, to provide descriptions of rotations and orientations in higher dimensional Euclidean spaces, $\mathbf{R}^n$, where the dimension $n > 3$. To properly distinguish between the concepts of rotation and re-orientation, their definitions are given in the following section.

2.1 Basic Definitions

The concepts of rigid rotation and re-orientation are defined in this section.

Definition 2.1 *A rotation in* $\mathbf{R}^n$ *is a length-preserving, non-deforming planar geometrical transformation.*

- *A rotation preserves lengths and angles between vectors.*
- *It can be represented by proper orthogonal or skew-symmetric* $n \times n$ *matrices.*
- *The components of a vector orthogonal to the plane of rotation remains unchanged.*

The familiar notion of "rotation about an axis" holds true only in 3 dimensions, while in 2 dimensions, rotations occur about a point. In $\mathbf{R}^n$, rotations occur about a $(n-2)$-dimensional subspace. In all cases, the definition of rotation suggests that the effect of rotation is confined to a plane. Thus, in general, rotations in $\mathbf{R}^n$ can be described as occurring in planes.

Definition 2.2 *A re-orientation in* $\mathbf{R}^n$ *is a length-preserving, non-deforming geometrical transformation.*

- *A re-orientation preserves lengths and angles between vectors.*
- *It can be represented by proper orthogonal or skew-symmetric* $n \times n$ *matrices.*

Thus, rotations are special forms of re-orientations, where the transformation is confined to a plane, i.e., a 2-D subspace of $\mathbf{R}^n$. Both rotations and re-orientations can be represented by proper orthogonal matrices, which preserve the sense (right-handed or left-handed) of the coordinate system. Throughout this thesis, right-handed coordinate systems are used, in keeping with the common practice.

2.2 Infinitesimal Rotations and Two-forms

In this section, the concept of rotation is generalized and a representation for infinitesimal rotations in $\mathbf{R}^n$ is developed. The generalization to finite rotations follows from this, and is presented in the next section. Although finite rotations are non-linear with respect to coordinate changes, infinitesimal rotations change space coordinates linearly.[18,19] Let $\Delta R \in \mathcal{O}^+(n)$ be an infinitesimal rotation matrix which rotates a unit vector r by an infinitesimal angle $\Delta\phi$ to a new position $\hat{r}$. This is illustrated in Figure 2.1. Since ΔR changes the coordinates of r linearly, we have

$$\hat{x}_i = x_i + \varsigma_{i1}x_1 + \varsigma_{i2}x_2 + \cdots + \varsigma_{in}x_n, \quad i = 1, 2, \ldots, n \tag{2.1}$$

where x_i and $\hat{x}_i$ denote the coordinates of r and $\hat{r}$ respectively. The elements ς_{ij}

Figure 2.1: An infinitesimal rotation

give the infinitesimal changes in the coordinates, and in the subsequent analysis, only the first order terms in ς_{ij} are considered. In matrix-vector notation, the change in position of the vector due to rotation is given by

$$\hat{r} = \Delta R r = (I_n + \Delta S)r \tag{2.2}$$

where ΔS is the matrix whose ijth element is ς_{ij}. Since ΔR is orthogonal, this imposes a condition on ΔS. From the above equation, we can see that

$$\Delta R^{\mathrm{T}} \Delta R = (I_n + \Delta S)^{\mathrm{T}}(I_n + \Delta S) = I_n \Rightarrow \Delta S^{\mathrm{T}} = -\Delta S \tag{2.3}$$

neglecting the second order term in ΔS. This shows that the differential matrix ΔS is skew-symmetric, and it gives the small changes in coordinates due to the infinitesimal rotation. The infinitesimal rotation matrix depends on the plane of rotation and the infinitesimal angle of rotation. Clearly, the plane of rotation is the plane containing the initial and rotated vectors, r and $\hat{r}$, to which the shaded part in Figure 2.1 belongs. Let $A \in \mathbf{R}^{n\times(n-2)}$ be a matrix with orthonormal columns that span the $(n-2)$-dimensional subspace orthogonal to this plane.

Any vector $v \in col(A)$ is not affected by this rotation and remains unchanged. Thus,

$$v = \Delta R v = (I_n + \Delta S)v \Rightarrow \Delta S v = 0 \tag{2.4}$$

i.e., all vectors orthogonal to the plane of rotation belong to the null space of ΔS, or $null(\Delta S) = col(A)$. Since ΔS is skew-symmetric, all its row vectors as well as column vectors are orthogonal to v.

2.2.1 Representation of Rotations in the Grassmann Algebra

We know that in 3 dimensions, two unit vectors u and v are orthogonal to each other if u can be expressed as $v = w \times u$, where $w \in \mathbf{R}^3$ is another unit vector and '$\times$' here denotes the vector cross product. To represent the rotation ΔS, the concept of cross product needs to be generalized to $\mathbf{R}^n$. Also, from the discussion above, we can anticipate that $\Delta S \in so(n)$ can be represented by the generalized cross product of the set of orthogonal column vectors of $A = [a_1 \vdots a_2 \vdots \cdots \vdots a_{n-2}]$. The generalized cross product in $\mathbf{R}^n$ is called the exterior product[20–23], and its operator is denoted by a '$\wedge$'. The exterior product of the vectors a_i, $i = 1, 2, \ldots, n-2$ is given by

$$(a_1 \wedge a_2 \wedge \cdots \wedge a_{n-2})(1, 2, \ldots, n-2) = \sum_{\sigma} (\text{sign } \sigma) a_1(\sigma_1) a_2(\sigma_2) \cdots a_{n-2}(\sigma_{n-2}) \tag{2.5}$$

where σ denotes all bijections of the form $\sigma : \{1, 2, \ldots, n-2\} \rightarrow \{1, 2, \ldots, n-2\}$, i.e., a permutation of the first $n-2$ natural numbers. The operator σ is called the *permutation operator* and the sign of σ is $+1$ when σ is an even permutation, and -1 when σ is an odd permutation. A permutation is even(odd) if it has an even(odd) number of transpositions, where a transposition is a swap of two elements of $\{1, 2, \ldots, n-2\}$ leaving the remainder fixed. Eq. (2.5) is the n-dimensional generalization of the vector cross-product in 3 dimensions, given by $v = w \times u$, $v^k = \sum_{i,j} \epsilon_{ijk} w^i u^j$ where $i, j, k = 1, 2, 3$ and ϵ_{ijk} denote the elements of the Ricci tensor.[2]

The quantity on the left-hand side of Eq. (2.5) is a tensor of order $n-2$ in $\mathbf{R}^n$, which is skew-symmetric in its components. The components of this tensor are denoted by the indices given by σ_i, $i = 1, 2, \ldots, n-2$, and if any two indices of a component are interchanged, the resulting component changes sign. None of the indices are repeated because the permutation operator σ is a bijection (a one-to-one and onto mapping). These two properties imply that the number of unique components in this exterior product is ${}^nC_{n-2} = n!/(2!(n-2)!)$, which is the number of combinations in which $n-2$ distinct objects can be selected from n distinct objects. The exterior product of k vectors in $\mathbf{R}^n$ is known as an *exterior k-form.* Exterior k-forms in $\mathbf{R}^n$ form a vector space, as can be easily verified, and this vector space is denoted by $\Lambda^k(\mathbf{R}^n)$. The dimension of this vector space is ${}^nC_k = n!/(k!(n-k)!)$. In particular, vectors are exterior 1-forms in $\mathbf{R}^n$ and $\Lambda^1(\mathbf{R}^n) = \mathbf{R}^n$. The direct sum of all the vector spaces $\Lambda^k(\mathbf{R}^n)$ together with their structure of real vector space and multiplication induced by '$\wedge$', is called the exterior algebra or Grassmann algebra of $\mathbf{R}^n$. Hence we see that representations of rotations in $\mathbf{R}^n$ can be found in the Grassmann algebra of $\mathbf{R}^n$.

2.2.2 Exterior Two-form Representations of Rotations

Since ${}^nC_{n-2} = {}^nC_2$, the dimensions of the vector spaces $\Lambda^{n-2}(\mathbf{R}^n)$ and $\Lambda^2(\mathbf{R}^n)$ are the same. Because the $\Lambda^k(\mathbf{R}^n)$ are vector spaces, each element of $\Lambda^{n-2}(\mathbf{R}^n)$ can be associated with an element of $\Lambda^2(\mathbf{R}^n)$, and vice versa. Such a relation is called a vector space isomorphism, and this particular isomorphism is denoted by

$$* : \Lambda^{n-2}(\mathbf{R}^n) \mapsto \Lambda^2(\mathbf{R}^n), \quad e_1 \wedge e_2 \wedge \cdots \wedge e_{n-2} \mapsto e_{n-1} \wedge e_n \tag{2.6}$$

where '$*$' is called the *Hodge star operator*[20–23] and the e_i, $i = 1, 2, \ldots, n$ are the standard orthogonal basis vectors in $\mathbf{R}^n$ (the row vectors of I_n). In general, the e_i could be replaced by vectors from any orthogonal basis vector set in $\mathbf{R}^n$.

In 3 dimensions, the Hodge star operator gives us the relations

$$e_1 \wedge e_2 \mapsto e_3, \quad e_2 \wedge e_3 \mapsto e_1, \quad e_3 \wedge e_1 \mapsto e_2 \tag{2.7}$$

which are identified in standard vector cross-product notation with the familiar relations

$$e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2 \tag{2.8}$$

Generalizing this identification to $\mathbf{R}^n$, we can identify the $(n-2)$-form in Eq. (2.5) with the following 2-form

$$a_1 \wedge a_2 \wedge \cdots \wedge a_{n-2} = p_1 \wedge p_2 \tag{2.9}$$

The p_1 and p_2 are orthonormal column vectors spanning the plane of rotation orthogonal to $col(A)$ and complete a right-handed basis vector set for $\mathbf{R}^n$, i.e. $C = [A \vdots P] \in \mathcal{SO}(n)$. This 2-form representation is much easier to evaluate than the $(n-2)$-form of Eq. (2.5), and it gives a uniform representation for rotations in Euclidean spaces of any dimension ($\mathbf{R}^n$). Clearly, this generalizes nicely the main idea of Euler's Principal Rotation Theorem, and the full result will be stated in the sequel. The representations for rotation matrices presented in this chapter are developed from the 2-form representation for rotation.

The 2-form, or exterior product of two vectors (1-forms) in $\mathbf{R}^n$ can also be easily represented in vector or matrix notation. From the general representation of k-forms using the permutation operator (σ), the (1,2) component of the 2-form on the right-hand side of Eq. (2.9) is

$$(p_1 \wedge p_2)(1,2) = p_1(1)p_2(2) - p_1(2)p_2(1) \tag{2.10}$$

and the other components are obtained similarly. Since the 2-form is a second order tensor, whose components are denoted by two indices, it can also be expressed as a matrix. The matrix representation of the 2-form in Eq. (2.10) is given by

$$[p_1 \wedge p_2] = p_1 p_2^{\mathrm{T}} - p_2 p_1^{\mathrm{T}} = P J_2 P^{\mathrm{T}} = \tilde{P} \tag{2.11}$$

where

$$J_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

as can be easily verified by equating the components of the tensor in Eq. (2.10) to those of the matrix in Eq. (2.11). The matrix $\tilde{P}$ identifies the plane of rotation. Note that this matrix is skew-symmetric, like the infinitesimal skew-symmetric rotation matrix ΔS. The only other quantity the rotation depends on, is the scalar angle of rotation. For the infinitesimal rotation case, we know from the linear change of coordinates that the unit vectors spanning the plane of rotation change as

$$\hat{p_1} = p_1 + p_2 \Delta\phi, \quad \hat{p_2} = p_2 - p_1 \Delta\phi \tag{2.12}$$

where $\Delta\phi$ is the differential angle of rotation. Thus, the orthogonal rotation matrix is given by

$$\Delta R = I_n + \tilde{P}\Delta\phi \tag{2.13}$$

Comparing Eq. (2.13) with Eq. (2.2) we see that the skew-symmetric differential rotation matrix is given by

$$\Delta S = \tilde{P}\Delta\phi \tag{2.14}$$

which characterises the rotation in terms of the plane of rotation and the angle of rotation. Eq. (2.14) sums up the representation of infinitesimal rotations in terms of the rotation parameters, i.e., the plane and the angle of rotation.

2.3 Finite Rotations

The various matrix representations for finite rotations developed in this section are all based on the 2-form representation for infinitesimal rotations presented in the last section. From the representation of the skew-symmetric infinitesimal rotation matrix in Eq. (2.14), it is expected that the finite rotation skew-symmetric matrix will be given by

$$S = \tilde{P}\phi = P J_2 P^{\mathrm{T}} \phi = [p_1 p_2^{\mathrm{T}} - p_2 p_1^{\mathrm{T}}]\phi \tag{2.15}$$

where $P = [p_1 \vdots p_2]$. The column vectors of A, which are orthogonal to the plane of rotation (given by P), are all eigenvectors of S with eigenvalue 0, i.e., the algebraic multiplicity of this eigenvalue is $(n-2)$. The rank of the planar rotation skew-symmetric matrix is only 2. The orthonormal vectors p_1 and p_2 spanning the plane of rotation satisfy

$$Sp_1 = -p_2\phi, \quad Sp_2 = p_1\phi \tag{2.16}$$

with $S \in so(n)$. As can be easily verified, the eigenvectors of S in the plane of rotation are given by

$$S\tfrac{\sqrt{2}}{2}(p_1 \pm ip_2) = (\pm i\phi)\tfrac{\sqrt{2}}{2}(p_1 \pm ip_2) \tag{2.17}$$

which gives a pair of complex eigenvectors and a conjugate pair of pure imaginary eigenvalues. The change in coordinates for the finite rotation, given by the orthogonal transformation matrix, is not linear in either the coordinates or the angle of rotation. The representation of the proper orthogonal rotation matrix is developed in section 2.3.1.

2.3.1 The Proper Orthogonal Rotation Matrix

A representation of the proper orthogonal matrix for a finite rotation can be obtained from its eigenvectors and eigenvalues. Note that the $(n-2)$ orthogonal vectors a_i spanning the subspace orthogonal to the plane of rotation, are all eigenvectors of the rotation matrix with eigenvalue $+1$, i.e., the algebraic multiplicity of this eigenvalue is $(n-2)$. These eigenvectors then satisfy the relation

$$Ra_i = a_i, \quad i = 1, 2, \ldots, n-2 \tag{2.18}$$

The orthogonal vectors spanning the plane of rotation satisfy the relations

$$\left.\begin{aligned} Rp_1 &= p_1\cos(\phi) + p_2\sin(\phi) \\ Rp_2 &= p_2\cos(\phi) - p_1\sin(\phi) \end{aligned}\right\} \tag{2.19}$$

The eigenvectors spanning the plane of rotation are thus given by

$$R\tfrac{\sqrt{2}}{2}(p_1 \pm ip_2) = (\cos(\phi) \pm i\sin(\phi))\tfrac{\sqrt{2}}{2}(p_1 \pm ip_2) \tag{2.20}$$

Comparing Eq. (2.20) with Eq. (2.17), we see that R and S share the same eigenvectors. Equations (2.18) and (2.20) can be combined to give the spectral decomposition of the finite rotation orthogonal matrix. The spectral decomposition of R is hence given by

$$R = [A \;\; \breve{P}] \begin{bmatrix} I_{n-2} & 0_{n-2,2} \\ 0_{2,n-2} & \Xi \end{bmatrix} \begin{bmatrix} A^{\mathrm{T}} \\ \breve{P}^{\dagger} \end{bmatrix} = V\Lambda V^{\dagger} \tag{2.21}$$

where

$$\breve{P} = \tfrac{\sqrt{2}}{2}\left[p_1 + ip_2 \;\; p_1 - ip_2\right], \quad V = [A \;\; \breve{P}] \quad \text{and} \quad \Xi = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix}$$

and $\breve{P}^{\dagger}$ denotes the complex conjugate transpose of $\breve{P}$.

The spectral decomposition of the finite rotation proper orthogonal matrix $R \in \mathcal{SO}(n)$ leads to a relation with the skew-symmetric rotation matrix $S \in so(n)$. From Eq. (2.17), we get the spectral decomposition of S as

$$S = [A \;\; \breve{P}] \begin{bmatrix} 0_{n-2,n-2} & 0_{n-2,2} \\ 0_{2,n-2} & \Theta \end{bmatrix} \begin{bmatrix} A^{\mathrm{T}} \\ \breve{P}^{\dagger} \end{bmatrix} = V\Phi V^{\dagger} \tag{2.22}$$

where

$$\Theta = \begin{bmatrix} i\phi & 0 \\ 0 & -i\phi \end{bmatrix}$$

From Eqs. (2.21) and (2.22), it can be seen that the eigenvectors of R (which are also the eigenvectors of S) are orthogonal, and the eigenvalues of R are the exponentials of the eigenvalues of S. These two relations are simultaneously satisfied by a simple transformation, the matrix exponential map. The matrix exponential preserves the eigenvectors of matrices with orthogonal eigenvectors.

Thus $R \in \mathcal{SO}(n)$ is the matrix exponential of $S \in so(n)$, and the relation is given by

$$R = V\Lambda V^{\dagger} = \exp(S) = I_n + S + \tfrac{1}{2!}S^2 + \tfrac{1}{3!}S^3 + \ldots = V\exp(\Phi)V^{\dagger} \tag{2.23}$$

Comparing this expression with Eq. (2.2), it can be seen that for the infinitesimal rotation case, the matrix exponential was approximated with the first two terms. Using the expressions for S in Eq. (2.15), we can express the orthogonal rotation matrix in terms of the rotation parameters as

$$R = I_n + \tilde{P}\sin(\phi) + \tilde{P}^2(1-\cos(\phi)) = AA^{\mathrm{T}} + P(I_2\cos(\phi) + J_2\sin(\phi))P^{\mathrm{T}} \tag{2.24}$$

where the last relation arrives from the fact that $AA^{\mathrm{T}} + PP^{\mathrm{T}} = I_n$ since the columns of A and P together form an orthogonal basis vector set for $\mathbf{R}^n$. Eq. (2.24) is the generalization of the familiar result of refs. 18 and 19 for $\mathbf{R}^3$. Figure 2.2 is an attempt at representing a planar rotation in $\mathbf{R}^4$. In this figure, p_1, p_2, p_3 and p_4 form the original orthogonal basis vector set, and p_1, p_2 is rotated by an angle ϕ to their new positions p_1' and p_2'. The plane spanned by p_3 and p_4 remains unaffected by the rotation.

Figure 2.2: Rigid rotation in 4 dimensional space

The eigenanalysis of R can also be used to obtain the rotation parameters when R is known. However, if only the rotation angle ϕ is required, then it can be evaluated by noting that

$$\mathrm{tr}[R] = \sum_{i=1}^{n} \lambda_i = (n-2) + 2\cos(\phi) \Rightarrow \cos(\phi) = \frac{\mathrm{tr}[R] + 2 - n}{2} \tag{2.25}$$

where tr denotes the trace of a matrix. This equation can be used to obtain ϕ from a known $R \in \mathcal{O}^{+}(n)$. Note that in $\mathbf{R}^n$, the number of scalar parameters that determine a unit vector directed along one of n directions is ${}^{n}C_1 - 1 = (n-1)$, where 1 is subtracted due to the constraint of normality. The number of scalar parameters that determine another unit vector which must be linearly independent of the first vector is ${}^{(n-1)}C_1 - 1 = (n-2)$. The orthogonality between these two unit vectors gives a total of $2n-3-1 = (2n-4)$ parameters determining the plane of rotation. Taking into account the angle of rotation, the total number of scalar parameters needed to determine a rotation in $\mathbf{R}^n$ is then $(2n-3)$. However, the number of unique components in a skew-symmetric or orthogonal matrix is ${}^{n}C_2 = n(n-1)/2$. Only for $n = 2$ and $n = 3$ are these two quantities equal, and a total re-orientation can be achieved by a single rotation in $\mathbf{R}^2$ and $\mathbf{R}^3$. For $n > 3$, the orthonormal vectors spanning the plane of rotation cannot be uniquely determined from a given rotation matrix, and there are multiple solutions for these vectors. This is because the rotation can be described by $(2n-3)$ scalar parameters, and when $n > 3$, this is less than ${}^{n}C_2$, the number of unique components in $S \in so(n)$.

2.3.2 The Cayley Transform

The Cayley Transform[24,25] is an important bilinear transformation (a conformal mapping) that gives a skew-symmetric matrix from a proper orthogonal matrix. This transformation, and related transformations, have given rise to some important and useful sets of parameters for attitude representation in 3 dimensions.[24,26] The most elegant feature of the Cayley Transform is that the inverse transformation has the same form as the forward transformation. The forward transformation $\Gamma : so(n) \mapsto \mathcal{SO}(n)$ gives a general representation of proper orthogonal matrix in terms of a skew-symmetric matrix. Applying the forward transformation, we get a rotation matrix

$$R = \Gamma(T) = \begin{cases} (I_n - T)(I_n + T)^{-1} \\ (I_n + T)^{-1}(I_n - T) \end{cases} \tag{2.26}$$

where T is a skew-symmetric rotation matrix. The inverse transformation gives T in terms of R as follows

$$T = \Gamma^{-1}(R) = \begin{cases} (I_n - R)(I_n + R)^{-1} \\ (I_n + R)^{-1}(I_n - R) \end{cases} \tag{2.27}$$

Like the matrix exponential, both the forward and inverse transformations of the Cayley Transform preserve the eigenvectors of matrices with orthogonal eigenvectors (which includes orthogonal and skew-symmetric matrices). Using the spectral decomposition of R in Eq. (2.21), the inverse transformation can be written as

$$T = \begin{cases} V(I_n - \Lambda)(I_n + \Lambda)^{-1}V^{\dagger} \\ V(I_n + \Lambda)^{-1}(I_n - \Lambda)V^{\dagger} \end{cases} \tag{2.28}$$

It has been shown that the rotation matrix R has a complex conjugate pair of eigenvalues $e^{\pm i\phi}$ on the unit circle in the complex plane, corresponding to the eigenvectors spanning the plane of rotation. The eigenvalues of T corresponding to these eigenvectors can be obtained using Eq. (2.28), and they are

$$\tau_{\pm} = \frac{1 - e^{\pm i\phi}}{1 + e^{\pm i\phi}} = \mp i \tan\frac{\phi}{2} \tag{2.29}$$

The T matrix hence becomes unbounded whenever the angle of rotation is an odd multiple of π radians. Hence the attitude parameters in $\mathbf{R}^3$ developed from the Cayley Transform, called the Rodrigues parameters[24,26], also have this problem. The other $(n-2)$ eigenvalues are all 0. This matrix can also be represented in terms of the plane and angle of rotation as

$$T = -\tilde{P}\tan\left(\frac{\phi}{2}\right) = -S\frac{\tan(\phi/2)}{\phi} \tag{2.30}$$

which relates the skew-symmetric rotation matrices S and T.

Another skew-symmetric matrix representation for rotations is obtained from the standard sum decomposition of the orthogonal rotation matrix R into symmetric and skew-symmetric matrices. The skew-symmetric matrix obtained in this way is

$$W = (R - R^{\mathrm{T}})/2 = V\left(\frac{\Lambda - \Lambda^{-1}}{2}\right)V^{\dagger} \tag{2.31}$$

From the spectral decomposition of R in Eq. (2.21), the eigenvalues corresponding to the eigenvectors spanning the plane of rotation can be obtained. These are given by

$$\omega_{\pm} = \frac{e^{\pm i\phi} - e^{\mp i\phi}}{2} = \pm i\sin(\phi) \tag{2.32}$$

which are always bounded between $\pm i$, unlike those from the Cayley Transform skew-symmetric matrix, T. The other $(n-2)$ eigenvalues of W are all 0. This matrix can also be expressed in terms of the rotation plane and angle as

$$W = \tilde{P}\sin\phi = S\frac{\sin\phi}{\phi} \tag{2.33}$$

The last relation relates the skew-symmetric rotation matrices S and W.

2.3.3 Example of Rotation in Five Dimensions

An example of a rotation in $\mathbf{R}^5$ is presented here. The rotation is a full continuous rotation of 2π radians (360 deg) in a plane, i.e., ϕ varies from 0 to 2π radians. The values of the scalar (inner) products, $a_i^{\mathrm{T}} R a_i$ for $i = 1, 2, 3$

Figure 2.3: Numerical simulation of a full rotation in 5 dimensional space

and $p_1^{\mathrm{T}} R p_1$ and $p_1^{\mathrm{T}} R p_2$ and the norm $\|a_1 - Rp_1\|$ are calculated numerically at discrete angular intervals, and plotted against the angle ϕ. The plot of these scalar products is shown in Fig. 2.3. It should be noted that $p_1^{\mathrm{T}} R p_1 = p_2^{\mathrm{T}} R p_2$ throughout the rotation, and this value is given by $\cos(\phi)$ where ϕ is the current angle of rotation. The $a_i^{\mathrm{T}} R a_i$ are always equal to 1 during the rotation, since the a_i are unchanged by the rotation. Also, the distance between the tips of the unit vectors a_1 and Rp_1 remains constant at $\sqrt{2}$, since these remain orthogonal during the rotation.

2.4 Orientations

A change of orthogonal bases in $\mathbf{R}^n$ is achieved by a re-orientation. Thus re-orientations are the most general orthogonal transformations in $\mathbf{R}^n$. Unlike in a rotation, the effect of a re-orientation is not confined to a plane. But re-orientations, like rotations, preserve the sense (left-handed or right-handed) of an orthogonal basis vector set, and hence can be represented by proper orthogonal matrices. A proper orthogonal orientation matrix $C \in \mathcal{SO}(n)$ or a skew-symmetric matrix $A \in so(n)$ depends on ${}^nC_2 = n(n-1)/2$ scalar parameters in general, which is the required number of parameters to describe an orientation in $\mathbf{R}^n$. This is more than the number of parameters needed to describe a scalar rotation for $n > 3$, as has already been shown. Section 2.4.1 presents a generalization of Euler's Theorem to $\mathbf{R}^n$, and accounts for the number of parameters required to specify an orientation.

2.4.1 Generalization of Euler's Theorem

The fact that a single rotation cannot give a general orthogonal transformation (re-orientation) in $\mathbf{R}^n$ for $n > 3$ has already been shown in the preceding section. Therefore, if rotations are used to obtain a re-orientation, a generalization of Euler's Theorem to higher dimensional spaces can only be achieved by arriving at a given orientation by a *sequence* of rotations. Publications that claim to extend Euler's Theorem to higher dimensional spaces exist (see, for instance, Ref. 3), but they try to do so by generalizing the rate kinematics and arriving at expressions for mean angular velocity. This is not necessary since Euler's Theorem deals with a static geometric problem, i.e., how a final orientation can be achieved by an angular displacement about a fixed axis from an initial orientation. The concept of planar rotations, long overlooked and confused with more general orientations, provides the best method to generalize Euler's Theorem. Since the change due to a rotation is confined to a plane, the general re-orientation can only be achieved by a sequence of rotations. Only two orthogonal basis vectors (spanning the plane of rotation) may be taken to their final orientation during each successive rotation. Therefore, we anticipate a total of

$$m = \left\lfloor \frac{n}{2} \right\rfloor = \begin{cases} (n-1)/2 & \text{if } n \text{ is odd} \\ n/2 & \text{if } n \text{ is even} \end{cases} \tag{2.34}$$

planar rotations is required in order to reach a final given orientation in $\mathbf{R}^n$, where $\lfloor x \rfloor$ rounds the real positive scalar x to the nearest integer towards zero. It is apparent that in odd dimensional spaces (where n is odd), there is always a particular direction which is not changed by the re-orientation.

Theorem 2.2 (Generalized Euler's Theorem) *Any arbitrary orientation in $\mathbf{R}^n$ can be achieved by a sequence of at least $m = \lfloor n/2 \rfloor$ principal rotations, performed on a set of m principal orthogonal planes.*

Using the generalized Euler's Theorem, a decomposition of a proper orthogonal orientation matrix $C \in \mathcal{SO}(n)$ in terms of m proper orthogonal rotation matrices can be obtained. Let $R_k(P_k, \phi_k) \in \mathcal{SO}(n)$, $k = 1, 2, \ldots, m$ be m rotation matrices which carry out rotations in the principal planes P_k which are orthogonal to each other, by the principal angles ϕ_k. From Eq. (2.24), any two of the orthogonal rotation matrices R_k must satisfy

$$[R_i(P_i, \phi_i) - I_n]\,[R_j(P_j, \phi_j) - I_n] = 0_{n\times n} \ \ \forall i \neq j \tag{2.35}$$

since the P_k are orthogonal to each other. The above equation implies that the product of these two rotation matrices can also be expressed in terms of their sum, as follows

$$R_i(P_i, \phi_i) R_j(P_j, \phi_j) = R_i(P_i, \phi_i) + R_j(P_j, \phi_j) - I_n \tag{2.36}$$

As can be easily verified, this relation can be extended to a product of any combination of the m orthogonal rotation matrices to give

$$\prod_{k=1}^{l} R_k(P_k, \phi_k) = \sum_{k=1}^{m} R_k(P_k, \phi_k) - (l-1) I_n \tag{2.37}$$

Thus, *the product of any combination of these principal rotation matrices on orthogonal planes can be expressed in terms of their sum*, which also implies that these matrices commute in matrix multiplication. Since the effect of two successive orthogonal transformations can be obtained by the matrix product of the corresponding matrices, the decomposition of C is a product decomposition. By Eq. (2.37), this can also be expressed as a sum decomposition, and the is given by

$$C = \prod_{i=1}^{m} R_k(P_k, \phi_k) = \sum_{k=1}^{m} R_k(P_k, \phi_k) - (m-1)I_n \tag{2.38}$$

Thus, the orientation matrix C *can be decomposed as either a product or a sum of the same set of principal rotation matrices* in orthogonal planes, R_k, and the order of matrix multiplication of the R_k is not important since they commute. The C matrix can also be expressed directly in terms of the planes P_k and the angles. Based on whether the spatial dimension n is even or odd, we have the spectral decomposition of C as given below

$$C = \begin{cases} \sum_{k=1}^{m} P_k(I_2 \cos(\phi_k) + J_2 \sin(\phi_k))P_k^\dagger = V\Lambda V^\dagger & \text{for } n \text{ even,} \\ aa^{\mathrm{T}} + \sum_{k=1}^{m} P_k(I_2 \cos(\phi_k) + J_2 \sin(\phi_k))P_k^\dagger = V\Lambda V^\dagger & \text{for } n \text{ odd} \end{cases} \tag{2.39}$$

where a is the eigenvector with eigenvalue 1 (this vector remains unchanged) when n is odd. V is the matrix of eigenvectors of C, and it is a special unitary matrix. The above developments in Eqs.(2.35)-(2.39) follow closely the work of Mortari in Ref. 2.

Since the rotation matrices $R_k(P_k, \phi_k)$ carry out rotations in orthogonal planes, they are related by the orthogonality of the planes P_k. From the discussion in the previous section, it is known that a rotation matrix in $\mathbf{R}^n$ can be specified by $(2n-3)$ parameters. If m arbitrary rotation matrices were chosen to build up an orientation matrix, then the total number of parameters in the orientation matrix would be $m(2n-3)$, which is more than the actual number of parameters in an orientation matrix, ${}^nC_2 = n(n-1)/2$, for $n > 3$. However, since the rotations are on orthogonal planes, the total number of parameters used in representing these rotations R_k indeed add upto ${}^nC_2$. The accounting for the number of parameters is shown in Table 2.4.1. Note that the orthogonality

Table 2.1: Number of parameters in an orientation matrix

Rotation	No. of Parameters	Cumulative total
R_1	$(n-1)+(n-2)-1+1=2n-3$	$2n-3$
R_2	$(n-3)+(n-4)-1+1=2n-7$	$4n-10$
$\vdots$	$\vdots$	$\vdots$
R_m	$(n-2m+1)+(n-2m)-1+1=2n-4m+1$	$n(n-1)/2$

between the two vectors used to construct a rotation matrix imposes an extra constraint, while the angle of rotation adds a parameter. The logic followed in this accounting is the same as in section 2.3.1. This informal accounting can be used as a qualitative justification for the Generalized Euler's Theorem, as it shows that a minimum of m rotations on orthogonal planes are necessary to reach a new orientation.

2.4.2 Skew-Symmetric Orientation Matrices

There are three skew-symmetric orientation matrices which can be obtained from the proper orthogonal orientation matrix C, and which completely specify an orientation in $\mathbf{R}^n$. The first of these skew-symmetric orientation matrices can be obtained from the inverse of the matrix exponential map (the matrix logarithmic map). The orientation matrix $C \in \mathcal{SO}(n)$ can also be expressed as the matrix exponential of a skew-symmetric matrix, A. This skew-symmetric matrix can be obtained by taking the matrix logarithm (inverse matrix exponential) of C as follows

$$A = \log(C) = \log\left[\prod_{k=1}^{m} R_k(P_k, \phi_k)\right] = \sum_{k=1}^{m} \log\left[R_k(P_k, \phi_k)\right] \tag{2.40}$$

Each rotation matrix $R_k \in \mathcal{SO}(n)$ is also the matrix exponential of a skew-symmetric rotation matrix $S_k \in so(n)$. Hence, the skew-symmetric orientation matrix A can also be expressed in terms of the S_k as

$$A = \sum_{k=1}^{m} S_k(P_k, \phi_k) \tag{2.41}$$

Thus, the orientation matrix $A \in so(n)$ can be expressed as the sum of the rotation matrices $S_k \in so(n)$. Since the skew-symmetric rotation matrices S_k act along orthogonal planes, the product of any two of them gives the null matrix. All of the S_k have the same set of eigenvectors, which is also common to the skew-symmetric orientation matrix A. This set of eigenvectors is common to the proper orthogonal and skew-symmetric orientation matrices, and the rotation matrices into which they can be decomposed. Since the rank of any of the skew-symmetric rotation matrices S_k is only 2, the rank of the skew-symmetric orientation matrix A is $2m$, which is n if n is even, and $(n-1)$ if n is odd. Thus, if n is odd, then a skew-symmetric orientation matrix is always singular and has a 0 eigenvalue corresponding to the only real eigenvector. This eigenvector has an eigenvalue of 1 for the proper orthogonal orientation matrix, C.

Another skew-symmetric orientation matrix is obtained from the Cayley Transform, introduced in section 2.3.1. The inverse of this transformation gives the skew-symmetric matrix from the orthogonal matrix representation. The skew-symmetric matrix Q is given by

$$Q = \Gamma^{-1}(C) = (I_n + C)^{-1}(I_n - C) = (I_n - C)(I_n + C)^{-1} \tag{2.42}$$

Using the spectral decomposition of the orthogonal rotation matrix given by Eq. (2.21), the above equation can also be expressed as

$$Q = V(I_n + \prod_{k=1}^{m} \Lambda_k)^{-1}(I_n - \prod_{k=1}^{m} \Lambda_k)V^\dagger \tag{2.43}$$

where V is the set of eigenvectors of C and the $R_k = V\Lambda_k V^\dagger$. From the spectral decomposition of an orthogonal rotation matrix given by Eq. (2.21), we know that the R_k have $(n-2)$ eigenvalues equal to 1, corresponding to the eigenvectors orthogonal to their plane of rotation. This gives the Cayley Transform skew-symmetric orientation matrix as

$$Q = \sum_{k=1}^{m} \tilde{P}_k \frac{1-e^{i\phi}}{1+e^{i\phi}} = \sum_{k=1}^{m} \tilde{P}_k \tan(-\phi_k/2) = \sum_{k=1}^{m} T_k(P_k, \phi_k) \tag{2.44}$$

where the T_k are the Cayley Transform skew-symmetric matrices corresponding to the R_k. Obviously, the T_k have the same eigenvectors and they commute, and since they represent rotations in the principal orthogonal planes P_k, the product of any two different T_k is the null matrix. The eigenvalues of T_k corresponding to the eigenvectors spanning the plane P_k are $\pm\sin(\phi)$, while all other eigenvalues are 0. Since the rank of each of the T_k is only 2, the rank of the skew-symmetric orientation matrix Q is $2m$, as with $A \in so(n)$.

The third skew-symmetric orientation matrix can be obtained from the proper orthogonal orientation matrix by the standard sum decomposition of a square matrix into a skew-symmetric and a symmetric matrix. The skew-symmetric orientation matrix obtained in this way is given by

$$E = \frac{C - C^{\mathrm{T}}}{2} = V\left[\left(\prod_{k=1}^{m} \Lambda_k - \prod_{k=1}^{m} \Lambda_k^{\dagger}\right)/2\right] V^{\dagger} \tag{2.45}$$

and $\Lambda_k^{\dagger} = \Lambda_k^{-1}$ for all $k = 1, 2, \ldots, m$ since the Λ_k only have an unimodular complex conjugate pair of eigenvalues and the remaining $(n-2)$ eigenvalues are $+1$. Using Eq. (2.39) to represent C in terms of the P_k and ϕ_k, we get

$$E = \sum_{k=1}^{m} (P_k(2J_2 \sin(\phi_k))P_k)/2 = \sum_{k=1}^{m} \tilde{P}_k \sin(\phi_k) = \sum_{k=1}^{m} W_k \tag{2.46}$$

where the $W_k = (R_k - R_k^{\mathrm{T}})/2$ are the skew-symmetric rotation matrices formed from the R_k. It is to be shown that the E matrix gives an unique representation of orientation as a skew-symmetric matrix. This is done by showing that for a given E, we can find the orthogonal orientation matrix C, as follows. From Eq. (2.45), we get

$$\left.\begin{array}{l} 2CE = C^2 - I_n \Rightarrow V\left[\Lambda^2 - 2\Lambda\Lambda_E - I_n\right] V^{\dagger} = 0 \Rightarrow \\ \Lambda = \Lambda_E \pm \sqrt{\Lambda_E^2 + I_n} \Rightarrow C = V\left[\Lambda_E \pm \sqrt{\Lambda_E^2 + I_n}\right] V^{\dagger} \end{array}\right\} \tag{2.47}$$

It can be verified that the + sign in the above equation gives C while the − sign gives $-C^{\mathrm{T}}$ (which is not always a proper orthogonal matrix). Thus the proper orthogonal orientation matrix can be evaluated uniquely from the skew-symmetric orientation matrix E.

2.4.3 Comparisons Between the Orientation Matrices

Comparisons between the different orientation matrices in the last two sections are presented in this section. The geometrical aspects of each of these representations for orientation is also presented. Table 2.4.3 presents the forms of the eigenvalues of the C, A, Q and E matrices for even and odd-dimensional cases. It is to be noted that any skew-symmetric matrix can be converted to a

Table 2.2: Eigenvalues for the different orientation matrices

Orientation matrix	Eigenvalues	
	Even n	Odd n
C	$\exp(\pm i\phi)$	$1, \exp(\pm i\phi)$
A	$\pm i\phi$	$0, \pm i\phi$
Q	$\mp i\tan(\phi/2)$	$0, \mp i\tan(\phi/2)$
E	$\pm i\sin(\phi)$	$0, \pm i\sin(\phi)$

proper orthogonal matrix by either the exponential map, the Cayley Transform, or Eq. (2.47). Hence every skew-symmetric matrix is a representation of an orientation.

Figure 2.4: Eigenvalues of the orientation matrices on the complex plane

The eigenvalues of all the orientation matrices corresponding to the eigenvectors spanning any of the planes of rotation P_k, are shown on the complex plane in Figure 2.4. This figure also shows the eigenvalues of the symmetric matrix associated with the proper orthogonal matrix, $M = (C + C^{\mathrm{T}})/2$. This matrix has real pairs of eigenvalues of $\cos(\phi_k)$ obtained by projecting $\exp(i\phi)$ on the real axis. The eigenvalues of the inverse matrix exponential skew-symmetric A, are obtained by projecting the length of the arc along the unit circle from (1,0) to the eigenvalues of C, onto the imaginary axis. The eigenvalues of the skew-symmetric matrix Q obtained from the Cayley Transform can be obtained by stereographic projection of the eigenvalues of C on the imaginary axis with the point of projection being (-1,0), and then taking its negative. The eigenvalues of the skew-symmetric matrix $E = (C - C^{\mathrm{T}})/2$ are obtained by projecting the eigenvalues of C on the imaginary axis. While the Q and E matrices can have rank$< 2m$ for non-zero rotation angles, the A matrix has rank$= 2m$ for all non-zero rotation angles in the range $[0, 2\pi)$. Besides, the A and E matrices are always bounded. Hence, the inverse exponential skew-symmetric matrix, A, is the best skew-symmetric representation for an orientation.

2.5 The Ortho-Skew Matrices

This section deals with the set of ortho-skew matrices, first introduced in Ref. 2. As their name suggests, the ortho-skew matrices are at once both orthogonal and skew-symmetric. Since we know that odd-dimensional skew-symmetric matrices are always singular (they have at least one 0 eigenvalue), and orthogonal matrices cannot be singular, the ortho-skew matrices can only exist in even dimensions ($\mathbf{R}^{2m\times 2m}$). Ortho-skew matrices are a special set of orientation matrices, and hence are proper orthogonal. An even-dimensional proper orthogonal matrix has pairs of complex conjugate eigenvalues lying on the unit circle in the complex plane ($\lambda_k^{\pm} = \exp(\pm i\phi_k)$) where ϕ_k are the angles of rotation for the orthogonal rotations making up the orthogonal matrix. An even-dimensional skew-symmetric matrix given by the Cayley Transform has pure imaginary eigenvalues in conjugate pairs, $\lambda_k^{\pm} = \pm i \tan(\phi_k/2)$. Hence, an ortho-skew matrix $\Im \in \mathcal{SO}(n) \cap so(n),\ n = 2m$ will have only eigenvalues of $\lambda_k^{\pm} = \pm i$, corresponding to the rotation angles in all orthogonal planes P_k making up $\Im$, being odd multiples of $\pi/2$ radians. These matrices satisfy the fundamental relation

$$\Im\Im^{\mathrm{T}} = I_n, \Im + \Im^{\mathrm{T}} = 0 \ \Rightarrow\ \Im\Im = -I_n \tag{2.48}$$

arising from their definition.

The basic symplectic matrices

$$J_{2m} = \begin{bmatrix} 0_{m\times m} & I_m \\ -I_m & 0_{m\times m} \end{bmatrix},$$

are a particular subset of the set of ortho-skew matrices. The basic symplectic matrices represent rotation of all orthogonal planes in the standard orthogonal basis vector set $\{e_i\}$ by angles of $\pm\pi/2$ radians, i.e,

$$J_{2m} = \sum_{k=1}^{m} [e_{2k-1} \wedge e_{2k}] = \sum_{k=1}^{m} E_k J_2 E_k^{\mathrm{T}} \tag{2.49}$$

where the $E_k = [e_{2k-1}\ e_{2k}]$ and e_i denotes the ith row (or column) vector of I_n. A general ortho-skew matrix $\Im$ has the decomposition given in Eq. (2.49) with E_k replaced by $P_k = [p_{2k-1}\ p_{2k}]$, where the p_k belong to any set of orthonormal vectors in $\mathbf{R}^n$. The product of two ortho-skew matrices, in general, is not ortho-skew. In fact, the product may not even by skew-symmetric, and it only satisfies the relation

$$L = \Im_1\Im_2,\ LL^{\mathrm{T}} = L^{\mathrm{T}}L = I_n \text{ but } L^{\mathrm{T}} = \Im_2\Im_1 \neq L \tag{2.50}$$

i.e., the product L is only orthogonal. However, the "symplectic product" of an ortho-skew matrix with itself, given by

$$\Im_2 = \Im_1 J_{2m} \Im_1 \tag{2.51}$$

gives another ortho-skew matrix, as may be easily verified. In fact J_{2m} can be replaced in the above equation by any general ortho-skew matrix, and the resultant product will still be ortho-skew. The set of ortho-skew matrices is hence invariant under products of this form. Ortho-skew matrices are, however, not symplectic, otherwise the product in Eq. (2.51) would be equal to $-J_{2m}$.

The Cayley Transform for the ortho-skew matrices gives another interesting result, as follows

$$\left.\begin{array}{l}(I_n+\Im)^{-1}(I_n-\Im)\\(I_n-\Im)(I_n+\Im)^{-1}\end{array}\right\}=-\Im \tag{2.52}$$

Hence the Cayley Transform of an ortho-skew matrix gives back its negative, and the negative of the Cayley Transform acts as an identity map on the set of ortho-skew matrices. Since the set of ortho-skew matrices is given by the intersection of the sets of orthogonal and skew-symmetric matrices, their eigenvalues lie on the intersection of the unit circle with the imaginary axis on the complex plane, i.e., their eigenvalues are $\pm i$ as said earlier. This is shown in Figure 2.5. This

Figure 2.5: Eigenvalues of ortho-skew matrices on the complex plane

figure shows that the set of ortho-skew matrices is not a connected set, as the set of eigenvalues of $\Im$ and $-\Im$ do not form a connected set. The eigenanalysis of the $\Im$ matrix gives

$$\Im\frac{\sqrt{2}}{2}\left(p_{2k-1}\pm p_{2k}\right)=\pm i\frac{\sqrt{2}}{2}\left(p_{2k-1}\pm p_{2k}\right) \tag{2.53}$$

The ortho-skew matrices can be thought of as the extension of the imaginary unit $i=\sqrt{-1}$ to the field of real matrices. They satisfy matrix analogs of most of the complex identities that are satisfied by i.[2] Subsequent powers of i and $\Im$ follow an identical structure

$$i^k=\begin{cases}+1 & \text{for } k=4m\\+i & \text{for } k=4m+1\\-1 & \text{for } k=4m+2\\-i & \text{for } k=4m+3\end{cases}\quad\Leftrightarrow\quad \Im^k=\begin{cases}+I_n & \text{for } k=4m\\+\Im & \text{for } k=4m+1\\-I_n & \text{for } k=4m+2\\-\Im & \text{for } k=4m+3\end{cases} \tag{2.54}$$

where m is an integer. The ortho-skew matrices also satisfy a relation analogous to Euler's formula

$$e^{\vartheta+i\varphi} = e^{\vartheta}(\cos\varphi + i\sin\varphi) \quad \Leftrightarrow \quad e^{\vartheta I_n+\Im\varphi} = e^{\vartheta}(I_n\cos\varphi + \Im\sin\varphi) \tag{2.55}$$

This equation can be specialized for the case $\vartheta = 0$ to give something similar to the familiar trigonometric identities

$$\left.\begin{aligned} 2\cos\varphi &= \ e^{i\varphi} + e^{-i\varphi} \\ 2i\sin\varphi &= e^{i\varphi} - e^{-i\varphi} \end{aligned}\right\} \quad \Leftrightarrow \quad \left\{\begin{aligned} 2I_n\cos\varphi &= \ e^{\Im\varphi} + e^{-\Im\varphi} \\ 2\Im\sin\varphi &= e^{\Im\varphi} - e^{-\Im\varphi} \end{aligned}\right. \tag{2.56}$$

The polar expression for a complex number also has its corresponding expression for a real matrix, with i and 1 being replaced by $\Im$ and I_n respectively. This lead to an analogy with DeMoivre's formula in the field of real matrices. The DeMoivre formula for complex numbers is a useful expression that can give both roots and powers in the form of a polar expression. This analogy is given by

$$\begin{aligned} z^{k/j} &= \xi^{k/j}\left[\cos\left(\tfrac{k\varphi}{j}\right) + i\sin\left(\tfrac{k\varphi}{j}\right)\right] \quad \Leftrightarrow \\ Z^{k/j} &= \Xi^{k/j}\left[I_n\cos\left(\tfrac{k\varphi}{j}\right) + i\Im\left(\tfrac{k\varphi}{j}\right)\right] \end{aligned} \tag{2.57}$$

where $Z = \Xi(I_n\cos\varphi + \Im\sin\varphi)$ is a real matrix, and $z = \xi(\cos\varphi + i\sin\varphi)$ is a complex number. Note that for $k = 1$ and $j = 2, 3, \ldots$, DeMoivre's formula gives the roots of z while for $j = 1$ and $k = 2, 3, \ldots$, it gives the powers of z.

CHAPTER III

REFLECTIONS AND PROJECTIONS

Reflections and projections are common geometrical transformations in Euclidean spaces. Unlike rotations, reflections and projections in $\mathbf{R}^n$ can occur about an Euclidean subspace of any possible dimension, from a 1-dimensional axis to the entire n-dimensional space itself (ignoring the trivial case of reflection or projection along a 0-dimensional point). More generally, reflections and projections occur about or along translations of linear subspaces, or hyperplanes, in $\mathbf{R}^n$. However, since a hyperplane is always a linear translation away from a parallel hyperplane (subspace) passing through the origin, we will only consider reflections and projections along hyperplanes passing through the origin. The representation in terms of a coordinate system centered on a parallel hyperplane is then obtained by a simple linear translation. A reflection changes the coordinates of a point in $\mathbf{R}^n$ in such a manner that the initial and reflected points have equal perpendicular distance from the subspace about which the reflection takes place, and this subspace separates them. A reflection can be described as occurring along a subspace $N \subset \mathbf{R}^n$, or about the subspace $N^\perp \subset \mathbf{R}^n$, the orthogonal complement of N. The first description will be more commonly used here. Although this description is unusual, it will be found to be more compatible with the description of projections, and also the representations for reflections used here. Projections are more general geometrical transformations than reflections. The relation between projections and reflections is not entirely analogous to that between orientations and rotations, since projections can also occur along subspaces of $\mathbf{R}^n$ of any dimension from 1 to n. A projection changes the coordinates of a point in $\mathbf{R}^n$ such that the initial and projected points have the same distance from the subspace along which the projection takes place. Reflections and projections can be best described in terms of orthogonal basis vector sets, as with rotations and orientations. The geometric properties and representations of reflections and projections are dealt with in this chapter. It is shown that symmetric matrices can be used to represent reflections and projections. The definitions of the concepts of reflection and projection are given in the following section.

3.1 Basic Definitions

The concepts of reflection and projection, as used in this thesis, are defined in this section.

Definition 3.1 *A reflection along a subspace in* $\mathbf{R}^n$ *maintains the perpendicular distance from this subspace and its orthogonal complement, and the orthogonal complement separates the initial and reflected objects.*

- *A reflection preserves lengths between vectors.*
- *It can be represented by a symmetric* $n \times n$ *matrix.*
- *The components of a vector orthogonal to the subspace along which the reflection occurs, remain unchanged.*

The reflection of a point along a subspace $N \subset \mathbf{R}^n$, has the same distance from this subspace, and the subspace orthogonal to it, as the original point. The straight line joining the original and reflected points is parallel to the subspace N, i.e., it never intersects N. The subspace $N^\perp \subset \mathbf{R}^n$ which is the orthogonal complement of N, separates the original and reflected points, and bisects the straight line joining them.

Definition 3.2 *A projection is a geometrical transformation along a subspace in* $\mathbf{R}^n$ *which maintains the perpendicular distance from this subspace.*

- *A projection can be represented by a symmetric* $n \times n$ *matrix.*

- *The components of a vector orthogonal to the subspace along which the projection occurs, remain unchanged.*

Thus, projections in general do not preserve either lengths or angles between vectors, and cannot be represented by orthogonal matrices. Reflections are thus special types of projections, in which the perpendicular distance from the subspace orthogonal to the subspace along which the reflection occurs, also remain unchanged. The definition of projections presented here also includes as a special case orthogonal projections, which are more commonly known as projections. Orthogonal projections are known to be idempotent.[10] It should be noted that unlike rotations and re-orientations, reflections and projections may not change spatial coordinates in a continuous manner.

3.2 Reflections

Reflections are elementary geometrical transformations in $\mathbf{R}^n$. A reflection does not change the length of a vector, but it does change the orientation of the vector. A point and its reflection along a subspace are equidistant from every point in the orthogonal complement of this subspace. An interesting property of a reflection is that, when the same reflection is applied twice in succession on an object, it returns the object to its initial position and orientation. Thus, if $\Re$ is the matrix representation of a reflection, then it must satisfy

$$\Re^{2k} = I_n, \quad \Re^{2k+1} = \Re \tag{3.1}$$

where k is an integer. Since reflections preserve lengths (Euclidean norms) of vectors, they can be represented by orthogonal matrices. This suggests that $\Re$ in Eq. (3.1) is also orthogonal, and hence satisfies

$$\Re^{\mathrm{T}}\Re = \Re\Re^{\mathrm{T}} = I_n \tag{3.2}$$

Both Eq. (3.1) and Eq. (3.2) are satisfied if $\Re^{\mathrm{T}} = \Re$, i.e., if $\Re$ is symmetric as well. Hence, reflections can be represented by matrices that are both orthogonal and symmetric, and such matrices will be called *ortho-symmetric* matrices in this thesis. Unlike the ortho-skew matrices, ortho-symmetric matrices can exist as transformations in both even and odd dimensional spaces. The Householder transformations, also called the Householder matrices,[8,9,13] belong to the set of ortho-symmetric matrices. In this section, the representation of reflections in terms of ortho-symmetric matrices is developed and a decomposition of symmetric matrices in terms of Householder matrices is formulated.

3.2.1 Ortho-Symmetric Matrix Representation of Reflections

In our 3-dimensional world, we view reflections as having a natural symmetry. This symmetry in the nature of reflections remarkably parallels their mathematical representation, which can be carried out through the ortho-symmetric matrices. Let the reflection be along an m-dimensional subspace of $\mathbf{R}^n$ which is spanned by the orthonormal columns of the matrix $N \in \mathbf{R}^{n\times m}$. The ortho-symmetric matrix representing this reflection is given by

$$\Re = I_n - 2NN^{\mathrm{T}}, \;\; N^{\mathrm{T}}N = I_m \tag{3.3}$$

As can be easily verified, this representation satisfies the relations in Eqs. (3.1) and (3.2). If $v \in \mathbf{R}^n$ is a vector, then its reflection is given by

$$\grave{v} = \Re v = (I_n - 2NN^{\mathrm{T}})v, \; \grave{v}^{\mathrm{T}}\grave{v} = v^{\mathrm{T}}v \tag{3.4}$$

which has the same length (norm) as v. For notational ease, the column space of N, $col(N)$, will also be denoted by N and its orthogonal complement by $N^{\perp}$. Since the orthogonal complement to N is equidistant from the vectors $\grave{v}$ and v and bisects the line joining them, their orthogonal projections on N, which give their distance from $N^{\perp}$, should be the negative of each other. The component of

$\grave{v}$ on the subspace N is

$$N^{\mathrm{T}}\grave{v} = N^{\mathrm{T}}\Re v = N^{\mathrm{T}}(I_n - 2NN^{\mathrm{T}})v = -N^{\mathrm{T}}v \tag{3.5}$$

which is indeed the negative of the component of v onto N. Let M denote a matrix with orthonormal columns which spans $N^{\perp}$, i.e, $C = [N \vdots M]$ forms an orthogonal matrix. Then the orthogonality conditions for C imply that

$$\left.\begin{array}{l} N^{\mathrm{T}}N = I_m, \quad N^{\mathrm{T}}M = 0_{m\times(n-m)}, \\ M^{\mathrm{T}}M = I_{n-m}, \text{ and } NN^{\mathrm{T}} + MM^{\mathrm{T}} = I_n \end{array}\right\} \tag{3.6}$$

Again denoting $col(M)$ by M, this means that $M \equiv N^{\perp}$. The vectors $\grave{v}$ and v are also equidistant from N, which means that their orthogonal projections on $M = N^{\perp}$ should be equal. This condition can be expressed as

$$M^{\mathrm{T}}\grave{v} = M^{\mathrm{T}}\Re v = M^{\mathrm{T}}(I_n - 2NN^{\mathrm{T}})v = M^{\mathrm{T}}v \tag{3.7}$$

Equations (3.4) to (3.7) satisfy all the properties of a reflection, and show that the ortho-symmetric matrix $\Re$ is indeed a valid representation for a reflection.

The ortho-symmetric matrix $\Re$ of Eq. (3.3) is said to be of order m since it can be specified by the m orthonormal columns of N. An ortho-symmetric matrix of order m and dimension n will be denoted by $\Re(n, m)$ when the order is important, otherwise the shorthand $\Re$ will be used. It can be represented by a total of $m(n - m)$ scalar parameters. This can be accounted for as follows. The m linearly independent unit column vectors of N can be specified by a total of $m(2n - m - 1)/2$ parameters. The m columns of N, n_i, also form a set of orthonormal vectors satisfying the first relation in Eq. (3.6). This set of orthonormal vectors can be obtained from any set of m linearly independent vectors in N by Gram-Schmidt orthogonalization.[12,27] The condition $N^{\mathrm{T}}N = I_m$ then imposes $m(m - 1)/2$ additional conditions, since the n_i are already known to be unit vectors. Thus, the total number of parameters that can uniquely determine $\Re(n, m)$ is given by $m(2n - m - 1)/2 - m(m - 1)/2 = m(n - m)$.

Note that the matrix N belongs to the Stiefel manifold[20,21] $\mathcal{S}t(m,n)$, which is the space of all orthonormal m-tuple of vectors in $\mathbf{R}^n$. Using Eq. (3.7), we see that

$$\Re(N) = I_n - 2NN^{\mathrm{T}} = I_n - 2(I_n - MM^{\mathrm{T}}) = -I_n + 2MM^{\mathrm{T}} = -\Re(M) \quad (3.8)$$

where $\Re(N)$ and $\Re(M)$ are the order m and order $(n-m)$ ortho-symmetric matrices formed from N and M respectively. This shows that a reflection along a subspace N has the reverse effect of a reflection along its orthogonal complement $M = N^{\perp}$. Figure 3.1 represents this for the case $n = 4,\ m = 2$, where M and N are the two orthogonal planes spanning the space. Eq. (3.8) also implies that the number of parameters that can be used to represent $\Re(M)$ and $\Re(N)$ should be equal, which is in concordance with the fact that this number $m(n-m)$ is symmetric with respect to m and $(n-m)$. Hence, ortho-symmetric matrices $\Re(n,m)$ and $\Re(n,n-m)$ have the same number of parameters. Geometrically, $\Re(n,m)$

Figure 3.1: Reflections along orthogonal planes in 4 dimensions

is a transformation that reflects any vector v in $\mathbf{R}^n$ along the m-dimensional column space of N, which is the orthogonal complement of the $(n-m)$-dimensional column space of M. The line joining a point and its reflection along N is bisected by M, while the line joining the point and its reflection along M is bisected by N. Hence, the reflections along N and along M are just the negative of each other. Also, the product of the matrices $\Re(N)$ and $\Re(M)$ is the negative of the identity matrix, as given by

$$\Re(M)\Re(N) = (I_n - 2MM^{\mathrm{T}})(I_n - 2NN^{\mathrm{T}}) = I_n - 2MM^{\mathrm{T}} - 2NN^{\mathrm{T}} = -I_n \quad (3.9)$$

using the relations in Eq. (3.6). Thus, the composition of these two reflections, carried out in any order, on any vector, reverses its direction.

$$\Re(M)\Re(N)v = \Re(N)\Re(M)v = -v \quad (3.10)$$

Successive reflections along subspaces that are orthogonal to each other reverts the direction of a vector, since it actually amounts to rotating the vector by 180 degrees (π radians).

The eigenanalysis of an ortho-symmetric matrix $\Re(n, m)$ is easy to carry out, because of its special structure. Since it is a symmetric matrix, it has only real eigenvalues and eigenvectors. Since an ortho-symmetric matrix is also orthogonal, it can have only unimodular eigenvalues. Hence, $\Re(n, m)$ can only have eigenvalues of ± 1. Using the relations in Eq. (3.6), the spectral decomposition of $\Re(n, m)$ can be carried out as follows

$$\left.\begin{aligned} \Re(n,m) &= I_n - 2NN^{\mathrm{T}} = MM^{\mathrm{T}} - NN^{\mathrm{T}} \\ &= [N \vdots M] \begin{bmatrix} -I_m & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & I_{n-m} \end{bmatrix} \begin{bmatrix} N^{\mathrm{T}} \\ M^{\mathrm{T}} \end{bmatrix} \end{aligned}\right\} \tag{3.11}$$

The columns of N and M are the eigenvectors of $\Re(n, m)$, the eigenvalues corresponding to the column vectors of N are -1 while the eigenvalues corresponding to the column vectors of M are $+1$. Obviously, since $\Re(n, m)$ is an orthogonal matrix, its set of eigenvectors $C = [N \vdots M]$ also forms a unitary (in this case orthogonal) matrix. Since the eigenvalues of $\Re(n, m)$ can be only ± 1, its determinant, which is given by the product of its eigenvalues, is

$$|\Re(n,m)| = (-1)^m \tag{3.12}$$

which depends only on the order m of the matrix. Thus, ortho-symmetric matrices can belong to either $\mathcal{O}^+(n)$ or $\mathcal{O}^-(n)$, i.e., they may or may not be proper orthogonal matrices. The ortho-symmetric matrices, having eigenvalues of ± 1, can be thought of as the extension of the real unit to the field of real matrices, and are hence the symmetric counterparts of the ortho-skew matrices. This is represented in Figure 3.2, which shows the eigenvalues of an ortho-symmetric matrix in the complex plane.

The product of two ortho-symmetric matrices, in general, is not ortho-sym-

Figure 3.2: Eigenvalues of ortho-symmetric (reflection) matrices on the complex plane

metric, since the product is not symmetric, in general

$$(\Re_1\Re_2)^{\mathrm{T}} = \Re_2\Re_1 \neq \Re_1\Re_2 \tag{3.13}$$

No ortho-symmetric matrix of odd order is proper orthogonal, and hence cannot be represented by rotations or re-orientations. When applied to an orthogonal set of basis vectors, an ortho-symmetric matrix of odd order changes the sense (left-handed to right-handed, and vice versa) of this set of basis vectors. This change cannot be represented by a rigid body transformation like a rotation or re-orientation. An ortho-symmetric matrix of even order, however, is proper orthogonal, and hence can be represented by a rotation or sequence of rotations. The identity matrix I_n, can be thought of as a trivial reflection matrix (which does not carry out any reflection at all) or as the only ortho-symmetric matrix of zeroth order. The negative of the identity matrix, $-I_n$, is the only ortho-symmetric matrix of nth order, and it simply reverses the directions of all vectors by carrying out a reflection about the origin.

3.2.2 The Householder or Elementary Reflection Matrices

The Householder matrices, also sometimes known as the *elementary reflectors*, are first order ortho-symmetric matrices. Their effect is to reflect vectors in $\mathbf{R}^n$ along a reflection axis c. So a Householder matrix is of the form

$$H = \Re(n, 1) = I_n - 2cc^{\mathrm{T}} \tag{3.14}$$

where c is a unit vector. Note that the outer product cc^{T} represents a rank 1 perturbation, which is a special case of the rank m perturbation NN^{T} of Eq. (3.3). Since H is an odd order ortho-symmetric matrix, it has a determinant of -1, and it changes the sense of an orthogonal basis vector set by reversing the direction of one of its basis vectors. The direction of c is reversed under the transformation in Eq. (3.14) as

$$\breve{c} = Hc = (I_n - 2cc^{\mathrm{T}})c = -c \tag{3.15}$$

Any vector $v \in \mathbf{R}^n$ is reflected along the c direction, as is shown in Figure 3.3 for $\mathbf{R}^3$. In this figure, e_1, e_2, and e_3 are the standard orthogonal basis vectors in $\mathbf{R}^3$. The Householder matrices are often used in matrix computations[8,9,13] to

Figure 3.3: Elementary reflection in 3 dimensions

triangularize square matrices or reduce rectangular matrices to upper-trapezoidal form. This is because they can convert any vector in $\mathbf{R}^n$ to a scalar multiple of any of the standard basis vectors e_k, $k = 1, 2, \ldots, n$, where e_k is the kth row (or column) vector of I_n.[9] This is carried out as follows

$$(I_n - 2cc^{\mathrm{T}})s = -\sigma e_k, \quad \sigma = \|s\| \tag{3.16}$$

where c is given by

$$c = \frac{s + \sigma e_k}{\|s + \sigma e_k\|} \tag{3.17}$$

The Householder matrices, being ortho-symmetric, have real orthonormal eigenvectors, with the eigenvalue corresponding to the eigenvector c being -1, and the other eigenvectors (orthogonal to c) having eigenvalues of $+1$, i.e., the eigenvalue $+1$ has algebraic multiplicity of $(n-1)$.

The Householder matrices, being elementary reflection matrices, can also be used to construct more general reflection (ortho-symmetric) matrices. Consider the ortho-symmetric matrix $\Re$ in Eq. (3.3), and the collection of Householder matrices

$$H_k = I_n - 2n_k n_k^{\mathrm{T}}, \; k = 1, 2, \ldots, m \tag{3.18}$$

where the n_k is the kth column vector of N, i.e., the n_k are orthogonal unit vectors. From the spectral decomposition of $\Re$, we get

$$\Re = \begin{cases} I_n - 2\sum_{k=1}^{m} n_k n_k^{\mathrm{T}} = \sum_{k=1}^{m}(I_n - 2n_k n_k^{\mathrm{T}}) - (m-1)I_n \\ = \sum_{k=1}^{m} H_k - (m-1)I_n \end{cases} \tag{3.19}$$

which expresses the reflection matrix $\Re$ in terms of the sum of a collection of elementary reflection matrices. The product of any two of the Householder matrices in Eq. (3.18) gives

$$\left.\begin{array}{l} H_1 H_2 = (I_n - 2n_1 n_1^{\mathrm{T}})(I_n - 2n_2 n_2^{\mathrm{T}}) \\ = I_n - 2n_1 n_1^{\mathrm{T}} - 2n_2 n_2^{\mathrm{T}} = H_1 + H_2 - I_n \end{array}\right\} \tag{3.20}$$

which expresses their product in terms of their sum. This can actually be generalized to a product of any combination of such Householder matrices constructed from a set of orthogonal axes n_k in $\mathbf{R}^n$. The product of the m Householder matrices H_k in Eq. (3.18) can thus be represented as

$$\prod_{k=1}^{m} H_k = \sum_{k=1}^{m} H_k - (m-1)I_n \tag{3.21}$$

Thus, *the product of any combination of elementary reflection matrices reflecting along orthogonal axes, can also be expressed in terms of their sum*, which implies that these matrices commute in matrix multiplication. This result is analogous to the result in Eq. (2.37) for combinations of rotations along orthogonal planes. Using this result, the reflection matrix $\Re$ can be expressed as both a sum and a product decomposition as follows

$$\Re = \sum_{k=1}^{m} H_k - (m-1)I_n = \prod_{k=1}^{m} H_k \tag{3.22}$$

and the order of multiplication of the H_k does not matter. This suggests that any general reflection matrix reflecting along an m-dimensional subspace, can be obtained by a combination of reflections along orthogonal axes spanning the subspace, carried out in any order. Any reflection is thus a combination of elementary reflections.

3.3 Projections

A projection along a subspace in $\mathbf{R}^n$ maintains the orthogonal distance (shortest distance) from this subspace. Projections are more general geometrical transformations in Euclidean spaces than reflections, and include reflections as a special case. The definition of projections used in this thesis and given in section 3.1 also encompasses the more commonly used definition which just describes orthogonal projections.[10,12] Both reflections and orthogonal projections are special cases of the more general linear projections that are described here. A matrix representation of these projections is arrived at from their definition and geometric properties. Matrices that represent projections are simply called projection matrices in this thesis. These matrices are more general than the idempotent symmetric matrices that describe orthogonal projections, although they are symmetric as well. These projections can be used to construct hyperplanes parallel to the subspace along which the projection occurs, as shown in section 3.3.2. A projection matrix that carries out a 1-dimensional projection, i.e., it projects a vector along an axis, is called a modified Householder matrix in this thesis. This name is given because of its similarity to a Householder matrix, which carries out an elementary reflection. Since orthogonal projections are commonly used in many applications, they are dealt with separately in section 3.3.1. A matrix representation of orthogonal projections in terms of idempotent symmetric matrices is also presented.

3.3.1 Orthogonal Projections

Orthogonal projections are widely used since they occur in many applications which require the minimum distance (norm) from a vector to a linear subspace (hyperplane), which is generally known as the minimum-norm problem.[10,12,24] The main result used in solving the minimum norm problem in these applications is the projection theorem for Hilbert spaces. The statement of this theorem[12] is given below.

Theorem 3.1 (Classical Projection Theorem) *Let M be a closed subspace of a Hilbert space* $\mathbf{H}$*. Corresponding to any vector $v \in \mathbf{H}$, there is a unique vector $m_0 \in M$ such that $\|v - m_0\| \leq \|v - m\|$ for all $m \in M$. A necessary and sufficient condition that $m_0 \in M$ be the unique minimizing vector is that $v - m_0$ be orthogonal to M.*

In this case, m_0 is the orthogonal projection of v onto M, as is shown in Figure 3.4 for the 3 dimensional case. The projection theorem is known to be a spe-

Figure 3.4: Orthogonal projection in 3 dimensions

cialization of the Hahn-Banach theorem[12,28,29], which is widely used along with its corollaries in functional analysis and its applications. Orthogonal projections in Euclidean spaces have all the properties of orthogonal projections in more general Hilbert spaces.[12] They are idempotent, i.e., applying these projections more than once on a vector has no extra effect, and their operator (matrix representation) is symmetric. The orthogonal projection of v onto M is given by a projection matrix P_o such that $P_o v = m_0$. The vector m_0 lies in the subspace M, and it satisfies

$$M^{\mathrm{T}} v_o = M^{\mathrm{T}} P_o v = M^{\mathrm{T}} v, \quad N^{\mathrm{T}} v_o = N^{\mathrm{T}} P_o v = 0 \tag{3.23}$$

where $N \equiv M^{\perp}$ is the orthogonal complement of M. Here N and M are also used to denote the matrices whose orthonormal column vectors span the subspaces N and M respectively.

The vector $v - m_0$ belongs to the orthogonal complement of M, which is $M^{\perp} \equiv N$. The idempotence condition for the orthogonal projection matrix implies that

$$P_o^k = I_n, \;\; k > 1 \tag{3.24}$$

where k is an integer. One can easily verify that both Eq. (3.23) and idempotence are satisfied if

$$P_o = I_n - NN^{\mathrm{T}} = MM^{\mathrm{T}} \tag{3.25}$$

using Eq. (3.6). An orthogonal projection thus projects a vector onto the subspace M, which is the orthogonal complement of the subspace N along which the projection takes place. A subsequent application of the same orthogonal projection hence has no effect on the projected vector. In particular, if dimension of M is 1, i.e., M is an axis, then the orthogonal projection of a vector onto M is also the component of this vector along the direction of M. Hence, the most common representation of a vector $v \in \mathbf{R}^n$ in terms of Cartesian coordinates uses orthogonal projections onto each axis in a set of orthogonal axes. Also to be noted is that subsequent orthogonal projections, in any order, along subspaces that are orthogonal to (but not necessarily orthogonal complements of) each other, project a vector onto the null vector (the origin). Thus

$$P_o(N)P_o(M) = MM^{\mathrm{T}}NN^{\mathrm{T}} = 0 = NN^{\mathrm{T}}MM^{\mathrm{T}} = P_o(M)P_o(N) \tag{3.26}$$

as is geometrically intuitive.

The orthonormal columns of M and N together form a special orthogonal matrix $C = [M \,\vdots\, N]$ which also forms the set of eigenvectors for P_o. The spectral decomposition of an orthogonal projection matrix P_o is then given by

$$P_o = \left[M \,\vdots\, N \right] \begin{bmatrix} I_{n-m} & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & 0_{m\times m} \end{bmatrix} \begin{bmatrix} M^{\mathrm{T}} \\ N^{\mathrm{T}} \end{bmatrix} \tag{3.27}$$

An orthogonal projection matrix hence has only two eigenvalues, $+1$ with algebraic multiplicity $(n-m)$, and 0 with algebraic multiplicity m, where m is the dimension of the subspace along which the projection occurs. The $+1$ and 0 eigenvalues correspond to the eigenvectors spanning the subspace onto which the projection occurs, and the eigenvectors spanning the subspace along which the projection occurs, respectively. An orthogonal projection along the entire space is a trivial case, since by Eq. (3.25), this is given by the identity matrix. For all other cases, an orthogonal projection matrix is singular. Comparing Eq. (3.27) with Eq. (3.11), it can be seen that orthogonal projections are also elementary geometrical transformations as are reflections, and both are special forms of projections represented by special types of symmetric matrices.

3.3.2 Generalized Projections

The projections defined in section 3.1 may be considered as generalizations of orthogonal projections and reflections. Let P represent a projection matrix which projects a vector v along a subspace $N \subset \mathbf{R}^n$, and its projection be $\breve{v}$. Then the vector v and its projection $\breve{v}$ have the same orthogonal distance from N. This means, by the classical Projection Theorem, that their orthogonal projections on $M \equiv N^{\perp}$ are equal in length, which is given by

$$\breve{v}^{\mathrm{T}} M = v^{\mathrm{T}} P M = v\, T M \;\Rightarrow\; PM = M \tag{3.28}$$

where P is the symmetric projection matrix. However, v and its projection $\breve{v}$, have different orthogonal distances from the subspace M. Their orthogonal projections onto the subspace N give these distances, and can be represented by

$$\breve{v}^{\mathrm{T}} N = v^{\mathrm{T}} P N = v^{\mathrm{T}} N \Theta \tag{3.29}$$

where Θ is a diagonal matrix of real scalars. If N is also the matrix whose orthonormal columns span the subspace N, then any matrix of the form

$$P = I_n - 2N\Omega N^{\mathrm{T}},\; \Omega \in \mathbf{R}^{m\times m} \text{ is real and diagonal} \tag{3.30}$$

satisfies both Eq. (3.28) and Eq. (3.29) simultaneously when $I_n - 2\Omega = \Theta$. If n_i is the ith column vector of N and ω_i the ith diagonal element of Ω, then the component of $\breve{v}$ along n_i is given by

$$\breve{v}^{\mathrm{T}} n_i = v^{\mathrm{T}}(I_n - 2N\Omega N^{\mathrm{T}}) n_i = (1 - 2\omega_i) v^{\mathrm{T}} n_i \tag{3.31}$$

If $\omega_i = 1$, then the vector v is reflected along the axis n_i, while if $\omega_i = 1/2$, it is orthogonally projected along n_i. Comparing this equation with Eq. (3.3), it can be noted that the ortho-symmetric matrices representing reflections are a particular subset of the set of matrices given by Eq. (3.30), obtained by setting $\Omega = I_m$. The factor 2 in Eq. (3.30) has been retained to make this comparison easy, and also because it leads to a lucid geometrical description of the transformation induced by P. The orthogonal projections given by Eq. (3.25) are also seen to be another subset of the transformations induced by P in Eq. (3.30), obtained by setting $\Omega = (1/2)I_m$.

Figure 3.5: Illustration of a generalized projection

The projection induced by P can be obtained from a combination of a reflection and a vector addition, or a combination of an orthogonal projection and a vector addition, where the vector to be added lies in a hyperplane N' parallel to N. This transformation is much more complex than a reflection or orthogonal projection, and is represented in Figure 3.5. As can be seen from the figure, the tip of the projected vector Pv touches the hyperplane N' parallel to N. Hence, by varying Ω in Eq. (3.30), the entire hyperplane N' can be traced. The orthogonal projections of Pv and v onto $M \equiv N^{\perp}$ are equal and are denoted by $P_o v$ in the figure. Note that from the representation of the orthogonal projection P_o in Eq. (3.24), we get

$$P_o P v = M M^{\mathrm{T}} (I_n - 2N\Omega N^{\mathrm{T}}) v = M M^{\mathrm{T}} v = P_o v \tag{3.32}$$

which also shows that the orthogonal projections of v and Pv on M are equal.

The spectral decomposition of P shows that this matrix is a much more general symmetric matrix than either the orthogonal projection matrix P_o or the reflection matrix $\Re$. From the representation in Eq. (3.30), and the relations in

Eq. (3.6), we get

$$P = MM^{\mathrm{T}} + NN^{\mathrm{T}} - 2N\Omega N^{\mathrm{T}} = MM^{\mathrm{T}} + N(I_m - 2\Omega)N^{\mathrm{T}} \tag{3.33}$$

This directly leads to the following spectral decomposition for the projection matrix P

$$P = \left[M \vdots N\right] \left[\begin{array}{cc} I_{n-m} & 0_{m,n-m} \\ 0_{n-m,m} & I_m - 2\Omega \end{array} \right] \left[\begin{array}{c} M^{\mathrm{T}} \\ N^{\mathrm{T}} \end{array} \right] = C\Upsilon C^{\mathrm{T}} \tag{3.34}$$

where $C = [M \vdots N]$ has the set of orthogonal eigenvectors of P as its column vectors, and Υ is the diagonal matrix of eigenvalues of P. The projection matrix P has $(n-m)$ eigenvalues equal to 1, corresponding to the eigenvectors spanning the subspace M which is the orthogonal complement of N, the subspace along which the projection occurs. The other eigenvalues, corresponding to the subspace along which the projection occurs, are given by $\lambda_i = 1 - 2\omega_i$. Since this matrix can entirely be represented by the m orthonormal columns of N and the m eigenvalues of Ω, the total number of parameters that can be used to represent it is $m(2n-m-1)/2 + m = m(2n-m+1)/2$. It can be noted that if M is a 0-dimensional point (i.e., N is the whole space), then the projection matrix naturally gives the most general symmetric matrix with orthogonal real eigenvectors and real eigenvalues.

3.3.3 Modified Householder Matrices

The modified Householder matrices are a special set of projection matrices that carry out projections along axes (1-dimensional subspaces) in $\mathbf{R}^n$. Hence they are elementary projection matrices in the same way that the Householder matrices are elementary reflection matrices. Let $\mathcal{M}$ be a modified Householder matrix obtained by projection along the axis c. From the representation for projection matrices given in Eq. (3.30), the representation for a modified Householder matrix is obtained as

$$\mathcal{M} = I_n - 2\mu cc^{\mathrm{T}} \tag{3.35}$$

where μ is a real scalar. From Eq. (3.14) and Eq. (3.35), it can be noted that the Householder matrices are again a special form of the modified Householder matrices, as elementary reflections are a special case of elementary projections. Let $v \in \mathbf{R}^n$ be a vector projected by $\mathcal{M}$ along the direction c. Then the component of the projected vector $\tilde{v}$ along c is

$$\tilde{v}^{\mathrm{T}} v = v^{\mathrm{T}}(I_n - 2\mu cc^{\mathrm{T}})c = (1-2\mu)v^{\mathrm{T}}c \tag{3.36}$$

If the scalar factor $\mu > 1/2$ then v and $\tilde{v}$ are separated by the hyperplane $M \equiv c^{\perp}$, the orthogonal complement of c. The components (orthogonal projections) of v and $\tilde{v}$ along $M \equiv c^{\perp}$ are equal. As μ is varied in the real line, an axis parallel to c is traced by the tip of $\tilde{v}$.

The spectral decomposition of the modified Householder matrix is given by

$$\mathcal{M} = I_n - 2\mu cc^{\mathrm{T}} = C\Psi C^{\mathrm{T}} \tag{3.37}$$

where $C = [c \vdots M]$, and Ψ is a diagonal matrix of eigenvalues which are given by $\psi_1 = 1 - 2\mu$, and $\psi_i = 1$ for $i = 2, 3, \ldots, n$. Thus, an $n \times n$ modified Householder matrix has $(n-1)$ eigenvalues equal to $+1$. Note that if $\mu = 1/2$, an orthogonal projection matrix is obtained that projects along c and onto M. This matrix is also singular, and has a single 0 eigenvalue corresponding to the axis c. The general projection matrix P, described in section 3.3.2, can also be expressed in terms of a set of modified Householder matrices. From the spectral decomposition of P given in Eq. (3.34), we get

$$P = \begin{cases} I_n - 2\sum_{k=1}^{m} \omega_k n_k n_k^{\mathrm{T}} = \sum_{k=1}^{m}(I_n - 2\omega_k n_k n_k^{\mathrm{T}}) \\ \quad -(m-1)I_n = \sum_{k=1}^{m} \mathcal{M}_k - (m-1)I_n \end{cases} \tag{3.38}$$

which expresses P in terms of the sum of a collection of modified Householder matrices $\mathcal{M}_k$, $k = 1, 2, \ldots, m$. The matrices $\mathcal{M}_k$ represent elementary projections along the m orthogonal axes n_k. The product of any two of these matrices may be obtained as

$$\left.\begin{aligned} \mathcal{M}_1\mathcal{M}_2 &= (I_n - 2\omega_1 n_1 n_1^{\mathrm{T}})(I_n - 2\omega_2 n_2 n_2^{\mathrm{T}}) \\ &= I_n - 2\omega_1 n_1 n_1^{\mathrm{T}} - 2\omega_2 n_2 n_2^{\mathrm{T}} = \mathcal{M}_1 + \mathcal{M}_2 - I_n \end{aligned}\right\} \tag{3.39}$$

which expresses the product of these two matrices in terms of their sum. This can actually be generalized to a product of any number of such modified Householder matrices constructed from a set of orthogonal axes n_k in $\mathbf{R}^n$. The product of the m modified Householder matrices $\mathcal{M}_k$ in Eq. (3.38) can thus be represented as

$$\prod_{k=1}^{m} \mathcal{M}_k = \sum_{k=1}^{m} \mathcal{M}_k - (m-1) I_n \tag{3.40}$$

Thus, *the product of any combination of these elementary projection matrices along orthogonal axes, can also be expressed in terms of their sum*, which implies that these matrices commute in matrix multiplication. This result is very similar to the result in Eq. (2.37) for the rotations along orthogonal planes and for the elementary reflections along orthogonal axes. Thus, the decomposition of the general projection matrix P in terms of the modified Householder (elementary projection) matrices can also be expressed as a product decomposition

$$P = \sum_{k=1}^{m} \mathcal{M}_k - (m-1) I_n = \prod_{k=1}^{m} \mathcal{M}_k \tag{3.41}$$

and the product of the $\mathcal{M}_k$ can be carried out in any order since they commute. This also shows that a general projection along an n-dimensional subspace of $\mathbf{R}^n$ can be achieved by successive elementary projections along the n orthogonal axes spanning this subspace. This statement is analogous to the generalized Euler's Theorem relating re-orientations to successive rotations in orthogonal planes.

3.4 Symmetric Matrices and Their Decompositions

As has been shown in the preceding sections, symmetric matrices are intimately related to reflections and projections, as their representations are all in forms that characterize special types of symmetric matrices. To represent a general symmetric matrix, two decompositions for symmetric matrices are developed in this section. As is known, the spectral decomposition of a general symmetric matrix, $S \in \mathcal{S}(n)$, is given by

$$S = C\Lambda C^{\mathrm{T}} \tag{3.42}$$

where $C \in \mathcal{SO}(n)$ and Λ is a real diagonal matrix of eigenvalues. A general symmetric matrix has $n(n+1)/2$ unique components or scalar parameters, which can also be obtained from the above equation by adding the ${}^{n}C_2 = n(n-1)/2$ parameters of the orthogonal eigenvector matrix C, and the n parameters of the diagonal eigenvalue matrix Λ. The first decomposition presented is in terms of a sum of scalar multiples of a collection of Householder matrices. The second decomposition is in terms of a sum of modified Householder matrices, which is also shown to be equivalent to a product decomposition. Both the decompositions are developed from the spectral decomposition of S given in the equation above. These two decompositions are compared and their geometric descriptions discussed at the end of the section.

3.4.1 Symmetric Matrix Decomposition by Householder Matrices

The Householder matrices can also be used to represent general symmetric matrices. From the discussion in section 3.1.2, since a Householder matrix is first order ortho-symmetric, it can be specified by only $(n-1)$ scalar parameters. Clearly, more than one Householder matrix needs to be used to construct a symmetric matrix $S \in \mathcal{S}(n)$. Consider the set of Householder matrices given by

$$H_k = I_n - 2c_k c_k^{\mathrm{T}}, \quad k = 1, 2, \ldots, n, \quad C = [c_1 \vdots c_2 \vdots \cdots \vdots c_n] \tag{3.43}$$

constructed from the set of orthogonal eigenvectors of S. The spectral decomposition of the kth Householder matrix in this set is given by

$$H_k = C\Lambda_k C^T \text{ where } \Lambda_k^{i,i} = 1 - 2\delta_{ki} \tag{3.44}$$

and δ_{ki} is the Kronecker delta operator. The scalar multiple $G_k = \eta_k H_k$ of the Householder matrix H_k has its kth eigenvalue equal to $-\eta_k$, while the other eigenvalues are all equal to $+\eta_k$. Multiplying each of the Householder matrices in Eq. (3.44) by scalars η_k, and adding them together, gives a general symmetric matrix

$$S = \sum_{k=1}^{n} G_k = \sum_{k=1}^{n} \eta_k H_k = C \left(\sum_{k=1}^{n} \eta_k \Lambda_k \right) C^{\mathrm{T}} \tag{3.45}$$

Let $\lambda \in \mathbf{R}^n$ be the vector of eigenvalues of S and $\eta \in \mathbf{R}^n$ be the vector of the n scalars η_k. Comparing Eqs. (3.44) and (3.45), the eigenvalues of S are then obtained as

$$\lambda_k = \sum_{j=1}^{n} (1 - 2\delta_{kj})\eta_j \;\Rightarrow\; \lambda = D\eta \;\text{ where }\; \mathrm{D} = \begin{bmatrix} -1 \; 1 \ldots 1 \\ 1 \;\; -1 \ldots 1 \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ 1 \; 1 \ldots -1 \end{bmatrix} \tag{3.46}$$

As can be easily verified, the matrix D is non-singular, and given a symmetric matrix S, it would always be possible to find a set of coefficients (scalar multipliers) η_k for the Householder matrices H_k, such that the decomposition in Eq. (3.45) holds. Hence, it is always possible to construct a decomposition for any general symmetric matrix in terms of scalar multiples of a set of Householder matrices. Geometrically, the matrices $G_k = \eta_k H_k$ are no more reflection matrices, since they magnify the reflection along the axis c_k. This also shows that a general symmetric matrix cannot be obtained from combinations of reflections alone.

Note that the decomposition in Eq. (3.45) cannot be expressed also as a product decomposition, since the matrices $G_k = \eta_k H_k$ are not elementary reflection matrices. The accounting for the number of parameters in S is shown in Table 3.4.1. Since the c_k are orthogonal directions, the H_k are not independent of each other, and orthogonality constraints have been taken into account in Table

Table 3.1: Number of parameters in a symmetric matrix

Reflection/magnification	No. of Parameters	Cumulative Total
$\eta_1 H_1$	$(n-1)+1$	n
$\eta_2 H_2$	$(n-2)+1$	$2n-1$
$\vdots$	$\vdots$	$\vdots$
$\eta_3 H_n$	$(n-n)+1$	$n(n+1)/2$

3.4.1. Each successive orthogonal c_k has $(n-i)$ independent scalar parameters, and the multiplier η_k adds another scalar parameter. This gives the total of ${}^{n+1}C_2 = n(n+1)/2$ independent parameters that describe a general symmetric matrix.

3.4.2 Symmetric Matrix Decomposition by Modified Householder Matrices

The symmetric matrix decomposition presented in the previous section used not just a set of Householder matrices, but also a set of extra scalars, the η_k in Eq. (3.45). However, the decomposition of a symmetric matrix by a collection of modified Householder matrices does not require any extra set of parameters, as these matrices are more general than the Householder matrices. Consider the set of modified Householder matrices

$$\mathcal{M}_k = I_n - 2\omega_k c_k c_k^{\mathrm{T}}, \quad k = 1, 2, \ldots, n, \quad C = [c_1 \vdots c_2 \vdots \cdots \vdots c_n] \tag{3.47}$$

where the column vectors of C form the set of real orthogonal eigenvectors of the symmetric matrix S. The spectral decomposition of the kth Householder matrix in this collection is given by

$$\mathcal{M}_k = C\Lambda_k C^T \text{ where } \Lambda_k^{i,i} = \begin{cases} 1 & \text{when } i \neq k, \\ 1 - 2\omega_k & \text{when } i = k \end{cases} \tag{3.48}$$

As mentioned in section 3.3.2, the most general symmetric matrix $S \in \mathcal{S}(n)$ can be considered as a general projection matrix which projects along the entire space

$\mathbf{R}^n$. In section 3.3.3, a decomposition of a general projection matrix projecting along an m-dimensional hyperplane in $\mathbf{R}^n$ was also developed. This implies that the symmetric matrix S can also be decomposed in terms of the set of Householder matrices $\mathcal{M}_k$ in Eq. (3.47). From the spectral decomposition of S in Eq. (3.42), we have

$$S = \sum_{k=1}^{n} \lambda_k c_k c_k^{\mathrm{T}} = I_n - 2\sum_{k=1}^{n} \left(\frac{1-\lambda_k}{2}\right) c_k c_k^{\mathrm{T}} = I_n - 2C\Omega C^{\mathrm{T}} \tag{3.49}$$

which expresses the symmetric matrix in the same form as a projection matrix. This can be expressed as the sum of the modified Householder matrices in Eq. (3.47)

$$S = \sum_{k=1}^{n} \mathcal{M}_k - (n-1)I_n = I_n - 2C\Omega C^{\mathrm{T}} \tag{3.50}$$

in which case the eigenvalues of S are given by $\lambda_k = 1 - 2\omega_k$. Hence a decomposition of a symmetric matrix $S \in \mathcal{S}(n)$ in terms of a set of modified Householder matrices always exists. Since by Eq. (3.40), the sum of the modified Householder matrices can also be expressed as a product, we obtain

$$S = \sum_{k=1}^{n} \mathcal{M}_k - (n-1)I_n = \prod_{k=1}^{n} \mathcal{M}_k \tag{3.51}$$

in which the order of matrix multiplication does not matter. Thus, a general symmetric matrix can be expressed as both a sum and a product decomposition of the same set of modified Householder matrices. Any symmetric matrix $S \in \mathcal{S}(n)$ can hence be obtained as the result of successive elementary projections along a set of orthogonal axes spanning $\mathbf{R}^n$. It can be easily verified that the n modified Householder matrices $\mathcal{M}_k$ in Eq. (3.47) have a total of $n(n+1)/2$ unique parameters among them.

3.4.3 Householder vs Modified Householder Decompositions

On comparing the decomposition of symmetric matrices by Householder and modified Householder matrices, it is seen that the sum decomposition using the modified Householder matrices are more efficient as regards numerical evaluation. It should be noted that the decomposition using the modified Householder matrices can be expressed as either a sum or a product decomposition. The sum decomposition requires less number of floating point operations (flops) on a computing machine than the product decomposition. For a 7×7 symmetric matrix, the Householder decomposition was found to require 2785 flops more than the spectral decomposition, while the modified Householder decomposition was found to require only 1541 flops more than the spectral decomposition. The spectral decomposition itself required 3780 flops. Part of the reason why the Householder decomposition requires more flops, is the matrix inversion required to evaluate the scalars η_k from the eigenvalues λ_k using Eq. (3.46). The addition of the scalar multiples of the Householder matrices H_k in Eq. (3.45) also requires more flops than the simple addition of the modified Householder matrices in Eq. (3.50).

Figure 3.6: Combinations of projections and reflections with magnifications

The decomposition of a symmetric matrix in terms of Householder matrices is obtained from scalar magnifications of elementary reflections along orthogonal axes, while the decomposition in terms of modified Householder matrices is obtained from elementary projections along the same set of orthogonal axes. These are entirely different geometric transformations as shown in Figure 3.6, but their total effects are the same. The relation between the eigenvalues λ_k, the scalars ω_k and the scalars η_k is given by

$$\lambda_k = 1 - 2\omega_k = -\eta_k + \sum_{j=1, j \neq k}^{n} \eta_j \tag{3.52}$$

This also means that the sum of the components of the projections and the magnifications of reflections are the same for any vector, i.e., in the figure, $\eta_1 \Re_1 v_1 + \eta_2 \Re_2 v_2 = \mathcal{M}_1 v_1 + \mathcal{M}_2 v_2$. Thus, a general symmetric matrix may be considered as a representation of reflections with magnifications along a set of orthogonal axes in $\mathbf{R}^n$, or, alternatively, as a representation of a generalized projection along the entire space $\mathbf{R}^n$.

CHAPTER IV

SYMPLECTIC RICCATI DIFFERENTIAL EQUATION

The Riccati equation in its various forms has wide-ranging applications from transmission line phenomena to diffusion problems.[17] The equation gets its name from Jacopo Francesco, Count Riccati, who in 1724, considered a scalar version of the equation.[30] The Riccati differential equation, although a nonlinear differential equation, is intimately related to ordinary linear homogeneous differential equations of the second order. One of the questions which led Count Riccati to become interested in quadratic differential equations was the time-evolution of the slope of a line through the origin determined by the trajectory of a second order linear differential equation. Consider the linear (possibly time-varying) planar differential equation described by

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{bmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{4.1}$$

Let $p(t) = y(t)/x(t)$ denote the slope of the line through the origin determined by the solution $(x(t), y(t))$ to Eq. (4.1) at any time t. It can be easily verified that the slope p satisfies the differential equation

$$\dot{p} = a_{21} + (a_{22} - a_{11})p - a_{12}p^2 \tag{4.2}$$

which is a scalar Riccati differential equation. The general matrix version of the Riccati differential equation (RDE) is of the form

$$\dot{P}(t) + P(t)A(t) + D(t)P(t) + P(t)B(t)P(t) + C(t) = 0 \tag{4.3}$$

where t denotes the independent scalar variable, has a non-square matrix solution $P(t) \in \mathbf{R}^{m\times n}$ and matrix coefficients $A(t) \in \mathbf{R}^{n\times n}$, $B(t) \in \mathbf{R}^{n\times m}$, $C(t) \in \mathbf{R}^{m\times n}$, and $D(t) \in \mathbf{R}^{m\times m}$. By definition, a solution of the equation in the interval

$t \in [a, b] \equiv I$ is a matrix function $P(t)$ which is absolutely continuous and satisfies Eq. (4.3) in the interval I. The solution of the matrix RDE is also connected to the solution of a second-order linear matrix differential equation given by

$$\begin{pmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{pmatrix} = \begin{bmatrix} A(t) & B(t) \\ -C(t) & -D(t) \end{bmatrix} \begin{pmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{pmatrix} \tag{4.4}$$

where $X(t) \in \mathbf{R}^{n \times n}$ and $Y(t) \in \mathbf{R}^{m \times n}$. If $(X(t), Y(t))$ is a solution to Eq. (4.4) and $X(t)$ is non-singular in the interval I, then

$$P(t) = Y(t)X^{-1}(t) \tag{4.5}$$

is a solution to the RDE Eq. (4.3). The solution given by Eq. (4.5) is known as Radon's formula for the solution of the RDE.

In this thesis, only the particular form of the Riccati differential equation with a symmetric form that allows for the existence of a symmetric matrix solution will be considered. The special form of the RDE with such a symmetric form is called the symplectic Riccati differential equation (SRDE). It is called symplectic because of its symplectic flow, as will be shown later. This is the form of the Riccati differential equation that is encountered in optimal control, H_∞ control, and estimation theory.[10,12,14–17] The vector space $\mathcal{S}(n)$ of real symmetric matrices is an invariant manifold for the SRDE. A solution remains in this manifold if the given terminal condition is a symmetric matrix. A geometric approach to solving the SRDE is adopted here. This is done by extending the domain of the equation, $\mathcal{S}(n)$, to the domain of the solution to the associated second-order matrix differential equation. This domain is a natural compactification of $\mathcal{S}(n)$, and is an n-dimensional subspace of $\mathbf{R}^{2n}$.

4.1 The SRDE in Optimal Control

The best known occurrence of the symplectic Riccati differential equation in linear systems theory is in linear quadratic optimal control. In linear quadratic optimal control, the control law developed is for a linear system with a quadratic performance index. Consider the linear system

$$\dot{x} = A(t)x + B(t)u,\ y = C(t)x, \quad x \in \mathbf{R}^n,\ u \in \mathbf{R}^m,\ y \in \mathbf{R}^p \tag{4.6}$$

which is possibly time-varying. A control law is to be determined such that the system tracks a desired output $\bar{y}(t)$ over the time interval $I \equiv [t_0, t_f]$. We choose the performance index to be optimized as

$$J = \frac{1}{2}\int_{t_o}^{t_f} [(\bar{y} - Cx)^{\mathrm{T}} Q(\bar{y} - Cx) + u^{\mathrm{T}} Ru] dt \tag{4.7}$$

where $Q \in \mathcal{S}(p)$ and $R \in \mathcal{S}(m)$ are symmetric matrices, Q is positive semidefinite and R is positive definite. Using the usual method for solving a variational calculus problem, the Hamiltonian for this system is formed as

$$H = \frac{1}{2}[(\bar{y} - Cx)^{\mathrm{T}} Q(\bar{y} - Cx) + u^{\mathrm{T}} Ru] + \lambda^{\mathrm{T}}(Ax + Bu) \tag{4.8}$$

where $\lambda \in \mathbf{R}^n$ is a vector of Lagrange multipliers (the co-state vector). The Euler-Lagrange equations for the solution of this optimization problem are given by[15,31]

$$\dot{\lambda}^{\mathrm{T}} = -\frac{\partial H}{\partial x},\ \ \lambda(t_f) = 0,\ \ \frac{\partial H}{\partial u} = 0 \tag{4.9}$$

For the system in consideration, these equations give the relations

$$\dot{\lambda} = C^{\mathrm{T}} Q(\bar{y} - Cx) - A^{\mathrm{T}}\lambda,\ \lambda(t_f) = 0,\ u = -R^{-1}B^{\mathrm{T}}\lambda \tag{4.10}$$

Eq. (4.6) and Eq. (4.10) give a two-point boundary value problem to be solved for x and λ. Instead of solving directly for the boundary-value problem, we can solve for the flow of the $2n$ coupled linear differential equations for the n components each of the state and the co-state. Assuming the existence of transition matrices $X(t)$ and $\Lambda(t)$, we get the solutions for the state and the co-state as

$$x(t) = X(t)x(t_0),\ \ \lambda(t) = \Lambda(t)x(t_0) + g(t) = S(t)x(t) + g(t) \tag{4.11}$$

where $S(t) = \Lambda(t)X^{-1}(t)$. Substituting Eq. (4.11) in Eq. (4.10), we get the following linear differential equations

$$\left.\begin{aligned} \dot{S} &= -SA - A^{\mathrm{T}}S + SBR^{-1}B^{\mathrm{T}}S - C^{\mathrm{T}}QC, \quad S(t_f) = 0 \\ \dot{g} &= -(A^{\mathrm{T}} - SBR^{-1}B^{\mathrm{T}})g + C^{\mathrm{T}}Q\bar{y}, \quad g(t_f) = 0 \end{aligned}\right\} \tag{4.12}$$

with known terminal conditions. Note that $\bar{y}(t)$ is a known signal, since it is the desired output, which is known by the controller. The control law is then given by

$$u(t) = -K(t)x(t) + w(t), \text{ where } K = R^{-1}B^{\mathrm{T}}S, \quad w = -R^{-1}B^{\mathrm{T}}g \tag{4.13}$$

which tracks the desired output signal $\bar{y}(t)$ asymptotically. The first differential equation in Eq. (4.12), which can be written as

$$\dot{S} + SA + A^{\mathrm{T}}S - SBR^{-1}B^{\mathrm{T}}S + C^{\mathrm{T}}QC, \quad S(t_f) = 0 \tag{4.14}$$

is a symplectic Riccati differential equation, which is symmetric in form. Note that since the given terminal condition for S is the null matrix, which is symmetric, the matrix solution S at any time t is symmetric. This solution gives the gain for the control law in Eq. (4.13).

4.2 The Symmetric Solution of the SRDE

The general form of the symplectic Riccati differential equation, with a given terminal condition, is

$$\dot{S}(t) + S(t)A(t) + A^{\mathrm{T}}(t)S(t) + S(t)B(t)S(t) + C(t) = 0, \quad S(t_0) = S_0 \tag{4.15}$$

where the matrices $B(t)$ and $C(t)$ are symmetric at all times, and the terminal condition could be either an initial or a final condition. Since the SRDE is a special form of the general Riccati differential equation, it can be related to a couple of linear matrix differential equations as in Eq. (4.4). The related linear matrix differential equations are of the form

$$\dot{U}(t) = \begin{pmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{pmatrix} = \begin{bmatrix} A(t) & B(t) \\ -C(t) & -A^{\mathrm{T}}(t) \end{bmatrix} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = H(t)U(t) \tag{4.16}$$

If the solution $(X(t), Y(t))$ to Eq. (4.16) exists, and $X(t)$ is non-singular in the time interval $I \equiv [t_0\ t_f]$, then the solution to the SRDE in Eq. (4.15) also exists and is given by[14,17]

$$S(t) = Y(t)X^{-1}(t) \tag{4.17}$$

which is Radon's formula for the SRDE. Also, if the terminal value of $S = S_0$ is symmetric, then $S(t)$ is symmetric for all time t. Radon's formula for the solution of the SRDE works by extending the domain of the SRDE from the vector space of symmetric matrices $\mathcal{S}(n)$ to the space of n-dimensional subspaces of $\mathbf{R}^{2n}$ spanned by the columns of $U(t)$ in Eq. (4.16). These column vectors are guaranteed to be linearly independent if $X(t)$ is non-singular. The matrix

$$H(t) = \begin{bmatrix} A(t) & B(t) \\ -C(t) & -A^{\mathrm{T}}(t) \end{bmatrix} \tag{4.18}$$

is called the Hamiltonian of the SRDE.

4.2.1 Flow of the SRDE

The Hamiltonian of the SRDE $H(t)$ is an infinitesimally symplectic (or skew-symplectic) matrix, $H(t) \in sp(n)$. It satisfies the relation

$$JH + H^{\mathrm{T}}J = 0, \text{ where } J = \begin{bmatrix} 0_{n\times n} & I_n \\ -I_n & 0_{n\times n} \end{bmatrix} \tag{4.19}$$

as can be easily verified from the form of $H(t)$ in Eq. (4.18). The flow of the extended SRDE in Eq. (4.16) is given by

$$U(t) = \Phi(t,t_0)U(t_0) \Rightarrow \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \begin{bmatrix} \Phi_1(t,t_0) & \Phi_2(t,t_0) \\ \Phi_3(t,t_0) & \Phi_4(t,t_0) \end{bmatrix} \begin{bmatrix} X(t_0) \\ Y(t_0) \end{bmatrix} \tag{4.20}$$

where $\Phi(t, t_0)$ is the transition matrix. The transition matrix is the unique solution to the following matrix differential equation

$$\dot{\Phi}(t, t_0) = H(t)\Phi(t, t_0), \;\; \Phi(t_0, t_0) = I_{2n} \tag{4.21}$$

and is obtained from a series of time integrals of the Hamiltonian matrix, known as the Peano-Baker series.[32,33] If the Hamiltonian matrix is analytic, then so is the transition matrix obtained from it. The transition matrix obtained from the Peano-Baker series is given by

$$\begin{aligned}\Phi(t, t_0) = I_{2n} &+ \textstyle\int_{t_0}^{t} H(\sigma_1)d\sigma_1 + \int_{t_0}^{t} H(\sigma_1) \int_{t_0}^{\sigma_1} H(\sigma_2)d\sigma_2 d\sigma_1 \\ &+ \textstyle\int_{t_0}^{t} H(\sigma_1) \int_{t_0}^{\sigma_1} H(\sigma_2) \int_{t_0}^{\sigma_2} H(\sigma_3)d\sigma_3 d\sigma_2 d\sigma_1 + \dots\end{aligned} \tag{4.22}$$

which is obtained from a sequence of approximating functions to the solution,[32] and the σ_i are a decreasing sequence of time instants in the interval $I = [t_0, t]$. This is, in general, difficult to obtain either analytically or numerically. However, the transition map becomes easy to obtain in the particular case when

$$H(t)\left(\int_{t_0}^{t} H(\tau)d\tau\right) = \left(\int_{t_0}^{t} H(\tau)d\tau\right)H(t) \tag{4.23}$$

As may be verified using Eq. (4.22), the transition map in this case is given by

$$\Phi(t, t_0) = \sum_{k=0}^{\infty} \frac{1}{k!}\left[\int_{t_0}^{t} H(\tau)d\tau\right]^k = \exp\left(\int_{t_0}^{t} H(\tau)d\tau\right) \tag{4.24}$$

which is the matrix exponential map. The transition matrix satisfies the relation

$$\Phi(t, t_0)^{\mathrm{T}} J \Phi(t, t_0) = J \tag{4.25}$$

as can be easily verified by taking the time derivative of Eq. (4.25) and using Eq. (4.21) and Eq. (4.19). This shows that the transition matrix of Eq. (4.16) is symplectic, regardless of whether the Hamiltonian satisfies the property in Eq. (4.23) or not, and this gives the symplectic Riccati differential equation (SRDE) its name. Hamiltonian matrices are related to symplectic matrices in the same way that skew-symmetric matrices are related to orthogonal matrices. The matrix exponential of a Hamiltonian matrix gives a symplectic matrix. In Eq. (4.21), the time integral of $H(t)$ is also a Hamiltonian matrix, and its exponential is the symplectic transition matrix. In the numerical results presented in this chapter, the Hamiltonian matrix is chosen such that it satisfies the property in Eq. (4.23). However, most of the analytical results presented here hold even when the Hamiltonian matrix is more general.

The extension of the domain of the SRDE is achieved by an imbedding which identifies the symmetric matrix S with the n-dimensional subspace

$$col(U) \equiv col \begin{pmatrix} I_n \\ S \end{pmatrix}$$

in $\mathbf{R}^{2n}$ which is the column space of U in Eq. (4.20). The space of all n-dimensional subspaces of $\mathbf{R}^{2n}$ is called the *Grassmann manifold* of n-dimensional subspaces, and is denoted by $G(2n, n)$. It is a compact real-analytic manifold of dimension n^2.[20–23,33] A special class of Grassmann manifolds are the real projective spaces $G(m, 1) \equiv \mathbf{R}P^{m-1}$ which consist of the set of all lines through the origin in $\mathbf{R}^m$. However, due to the symmetric structure of the SRDE, the extended form is only in a subspace of the Grassmann manifold $G(2n, n)$, called the *Lagrange-Grassmann manifold*. The Lagrange-Grassmann manifold $\mathcal{L}(n)$ consists of those n-dimensional subspaces of $\mathbf{R}^{2n}$ on which a particular skew-symmetric form, ω, vanishes identically, and is given by

$$\omega(x, y) \equiv x^{\mathrm{T}} J y, \quad \mathcal{L}(n) \equiv \ \{Q \in G(2n, n) \mid \omega(x, y) = 0, \ \forall\, x, y \in Q\} \qquad (4.26)$$

as can be verified using the symplectic nature of the transition matrix in Eq. (4.25). Let the imbedding of the SRDE from the space of symmetric matrices $\mathcal{S}(n)$ to the Lagrange-Grassmann manifold $\mathcal{L}(n)$ be denoted by ψ. This imbedding is given by

$$\psi(S(t, S_0, t_0)) = col \begin{pmatrix} I_n \\ Y(t)X^{-1}(t) \end{pmatrix} = \Phi(t, t_0)\, col \begin{pmatrix} I_n \\ S_0 \end{pmatrix} \qquad (4.27)$$

which is valid as long as the inverse of $X(t)$ exists. The image in $\mathcal{L}(n)$ of this imbedding consists of those elements of $\mathcal{L}(n)$ which are complementary to the n-dimensional subspace

$$col \begin{pmatrix} 0_{n\times n} \\ I_n \end{pmatrix},$$

and we call this image space $\mathcal{L}_0(n)$. The image $\mathcal{L}_0(n)$ is an open and dense subset of $\mathcal{L}(n)$ and the imbedding ψ identifies $\mathcal{S}(n)$ with this image. The complement $\mathcal{L}(n) - \mathcal{L}_0(n)$ can be viewed as a hypersurface of points at infinity which have been added to compactify the space $\mathcal{S}(n)$. The space $\mathcal{L}(n)$ is thus a natural one-point compactification of the vector space $\mathcal{S}(n)$ and $\mathcal{S}(n) \subset \mathcal{L}(n)$. Since the transition matrix $\Phi(t, t_0)$ is symplectic, $\mathcal{L}(n)$ is an invariant manifold for the flow $S(t, S_0, t_0)$ on $G(2n, n)$. The restriction of the image to $\mathcal{L}_0(n)$ implies that solutions at "infinity" are disregarded.

4.2.2 Relating the Flow with the Spectral Decomposition

The spectral decomposition of the symmetric matrix solution of the SRDE is of the form given in Eq. (3.42). It can be recast into the form

$$S(t) = E(t)F^{-1}(t), \quad F(t) = E(t)\Lambda^{-1}(t) \tag{4.28}$$

where $E(t)$ and $\Lambda(t)$ are the matrices of eigenvectors and eigenvalues respectively. However, this does not mean that $X(t) = F(t)$ and $Y(t) = E(t)$ for the extended SRDE. Comparing this equation with Radon's formula for the solution to the SRDE, it can be seen that both equations would be satisfied if the following relations hold

$$X(t) = F(t)M, \quad Y(t) = E(t)M \tag{4.29}$$

It may be easier during numerical implementation to substitute the terminal conditions for X and Y as $X(t_0) = I$, $Y(t_0) = S_0$ instead of carrying out a spectral decomposition of the given terminal value S_0. In that case, $(X(t), Y(t))$

are related to $(F(t), E(t))$ and the given terminal condition in the following way

$$\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \begin{bmatrix} F(t)F_0^{-1} \\ E(t)F_0^{-1} \end{bmatrix} = \Phi(t, t_0) \begin{bmatrix} I_n \\ S_0 \end{bmatrix} \tag{4.30}$$

From this equation and the block partitioned form of the transition matrix in Eq. (4.20), we get

$$\left.\begin{aligned} X(t) &= \Phi_1(t, t_0) + \Phi_2(t, t_0)S_0 \\ Y(t) &= \Phi_3(t, t_0) + \Phi_4(t, t_0)S_0 \end{aligned}\right\} \tag{4.31}$$

The symplectic property of $\Phi(t, t_0)$ ensures that the following relations are satisfied by the four blocks of the matrix

$$\left.\begin{aligned} \Phi_1^{\mathrm{T}}\Phi_3 - \Phi_3^{\mathrm{T}}\Phi_1 = 0, \;\; \Phi_2^{\mathrm{T}}\Phi_4 - \Phi_4^{\mathrm{T}}\Phi_2 = 0, \\ \Phi_1^{\mathrm{T}}\Phi_4 - \Phi_3^{\mathrm{T}}\Phi_2 = I_n = -(\Phi_2^{\mathrm{T}}\Phi_3 - \Phi_4^{\mathrm{T}}\Phi_1), \end{aligned}\right\} \tag{4.32}$$

Note that these equations imply that

$$\Phi_3 = S_{31}\Phi_1, \;\; \Phi_4 = S_{42}\Phi_2, \;\; \Phi_1^{\mathrm{T}}(S_{42} - S_{31})\Phi_2 = I_n \tag{4.33}$$

where S_{31} and S_{42} are symmetric matrices. Thus, the symplectic transition matrix can be constructed entirely from the 3 matrices, Φ_1, S_{31} and S_{42}. This gives a total of $n^2 + n(n+1)/2 + n(n+1)/2 = n(2n+1)$ unique parameters that define the symplectic matrix Φ. The solution to the SRDE is then obtained from Eq. (4.31) and Radon's formula as

$$S(t) = (\Phi_3(t, t_0) + \Phi_4(t, t_0)S_0)(\Phi_1(t, t_0) + \Phi_2(t, t_0)S_0)^{-1} \tag{4.34}$$

Using the relations in Eq. (4.32) and Eq. (4.31), one can easily verify that $S(t)$ is indeed symmetric if S_0 is symmetric.

The time derivative of the spectral decomposition of a time-varying real analytic symmetric matrix reveals another interesting fact about the space of symmetric matrices $\mathcal{S}(n)$. A time-varying real analytic symmetric matrix always has an analytic spectral decomposition, in which the eigenvectors and eigenvalues are also analytic functions of time.[34] Taking the time derivative of Eq. (4.28), we get

$$\dot{S} = \dot{E}\Lambda E^{\mathrm{T}} + E\dot{\Lambda}E^{\mathrm{T}} + E\Lambda\dot{E}^{\mathrm{T}} \tag{4.35}$$

where $E(t)$ and $\Lambda(t)$ are analytic. Noting that $E(t)$ is an orthogonal matrix and is the matrix exponential of a skew-symmetric matrix, $Q(t)$, its time derivative is of the form

$$\dot{E}(t) = \frac{d}{dt}\exp\left(Q(t)\right) = \dot{Q}(t)E(t) = \Omega(t)E(t) \tag{4.36}$$

where $\Omega(t)$ is also skew-symmetric. Hence, Eq. (4.35) can also be put into the form

$$\dot{S} = \Omega E\Lambda E^{\mathrm{T}} + (E\Lambda E^{\mathrm{T}})(E\Lambda^{-1}\dot{\Lambda}\Lambda^{-1}E^{\mathrm{T}})(\Lambda E) - E\Lambda E^{\mathrm{T}}\Omega \tag{4.37}$$

which is also a symmetric form. From the spectral decomposition in Eq. (4.28), this equation can be expressed as the following SRDE

$$\dot{S} - \Omega S - S\Psi S + S\Omega = 0 \tag{4.38}$$

which shows that every time-varying real analytic symmetric matrix satisfies its own particular SRDE. Ω and Ψ give the rate of variation of the eigenvectors and eigenvalues respectively.

4.3 Numerical Solution for the SRDE

There are many procedures available for solving Riccati equations, most of them being applicable only to algebraic Riccati equations or RDEs with constant or periodic matrix co-efficients.[35–37] Numerical solutions for the general time-varying SRDE, involving numerical integration of the flow, and using the special structure of the SRDE are presented here. Two methods in which the symplectic transition matrix, $\Phi(t, t_0)$, can be obtained by numerical integration, are used and their results compared. Both these methods utilize the extension of the SRDE from the space of symmetric matrices to the Lagrange-Grassmann manifold. Hence, they integrate more variables than a direct integration of the SRDE in Eq. (4.15) would require. However, one benefit of this extension of the SRDE in numerical computation is that the numerically computed solution for the transition matrix never blows up, even though the SRDE itself may become stiff (when $X(t)$ in Radon's formula becomes singular). This is due to the fact that the Lagrange-Grassmann manifold, $\mathcal{L}(n)$ is a compact manifold, unlike the vector space of symmetric matrices $\mathcal{S}(n)$. The transition matrix, $\Phi(t, t_0)$ is solved for, and the symmetric matrix solution is obtained using Eq. (4.34). However, numerical errors caused during numerical integration accumulate and the transition matrix may not remain appreciably 'close' to symplectic. How close the matrix is to being symplectic, may be measured by a standard matrix norm, like the Frobenius norm, as follows

$$\mathsf{n} = \|\Phi^{\mathrm{T}} J \Phi - J\| \tag{4.39}$$

If the numerically evaluated transition matrix was perfectly symplectic, then the above norm would be exactly zero. When computing for the flow of the SRDE numerically, the norm n in Eq. (4.39) is a good measure for the reliability of the numerical scheme. This norm should be within the tolerances required for the desired accuracy. The symmetric matrix solution $S(t)$ may also be obtained from the flow of the extended SRDE using Eq. (4.34). However, if $X(t)$ as given by Eq. (4.31) is close to singular, then the solution $S(t)$ will blow up. Symmetricity of the solution obtained from the flow of the extended SRDE may be checked in a similar manner to the check for the symplecticity of Φ. This can be done as follows

$$\mathsf{s} = \|S - S^{\mathrm{T}}\| \tag{4.40}$$

If the norm s is small, then the solution is close to being symmetric.

4.3.1 Direct Numerical Integration for the Flow

A standard way in which to obtain the transition matrix is to numerically integrate for $\Phi(t, t_0)$ using Eq. (4.21). This method, however, does not ensure that $\Phi(t, t_0)$ will remain reasonably close to symplectic when the integration time $(t - t_0)$ is large. In this procedure, numerical integration of Eq. (4.21) is done to obtain the transition matrix Φ beginning with the initial condition $\Phi(t_0, t_0) = I_{2n}$. The symmetric matrix solution is obtained using Eq. (4.34) and the given initial condition $S(t_0) = S_0$. The result of numerically integrating the transition matrix for a particular time-varying SRDE is shown in Figure 4.1, where the norm n in Eq. (4.39) is plotted against the time t. The SRDE used in

Figure 4.1: Departure from symplecticity of flow of extended SRDE using direct numerical integration

this numerical integration is of the same form as Eq. (4.15), with time-varying coefficients

$$A(t) = \begin{bmatrix} 0.36\cos(3t) + 1.28\sin(2t) & 0.48\cos(3t) - 0.96\sin(2t) \\ 0.48\cos(3t) - 0.96\sin(2t) & 0.64\cos(3t) + 0.72\sin(2t) \end{bmatrix},$$

$$B(t) = \begin{bmatrix} 0.36(1 - e^{-2t}) - 1.28e^{-t} & 0.48(1 - e^{-2t}) + 0.96e^{-t} \\ 0.48(1 - e^{-2t}) + 0.96e^{-t} & 0.64(1 - e^{-2t}) - 0.72e^{-t} \end{bmatrix},$$

and

$$C(t) = \begin{bmatrix} -0.72e^{-2t} + 0.64(1 - e^{-t}) & 0.48(e{-}t - 1 - 2e^{-2t}) \\ 0.48(e{-}t - 1 - 2e^{-2t}) & -1.28e^{-2t} + 0.36(1 - e^{-t}) \end{bmatrix}.$$

The Hamiltonian matrix constructed from these matrices will satisfy the commutative property of Eq. (4.23). Note that all the three matrices $A(t)$, $B(t)$ and $C(t)$ are indefinite, and they also commute with each other, which is a sufficient condition for the Hamiltonian to satisfy Eq. (4.23). This figure shows that although the integration is for a short time duration (only 5 seconds), the errors accumulating in the transition matrix are quite substantial, on the order of 10^{-3}. As a result, the symmetric matrix solution evaluated using Eq. (4.34) will not be perfectly symmetric. This is shown in Figure 4.2, which shows the departure from symmetricity of the solution $S(t)$ computed from Eq. (4.34) and direct integration of $\Phi(t, t_0)$. As can be seen from this figure, the norm s is of

Figure 4.2: Departure from symmetricity of solution for SRDE using direct numerical integration

the order of 10^{-3}, which means that the matrix solution of the SRDE departs significantly from symmetricity within the time interval from 0 to 5 seconds. The matrix solution, however, does not blow up within this finite time interval.

Computing the transition matrix in this way also increases the number of variables to be integrated, from the n^2 variables if S in Eq. (4.15) was directly integrated, to $4n^2$ variables for the $2n \times 2n$ transition matrix Φ in Eq. (4.21). This procedure does not utilize the symplectic form of the transition matrix, or the special form of the Hamiltonian matrix, to obtain the solution. A better way to obtain the solution of the SRDE would be to utilize its special symmetric structure during numerical integration, which would reduce the number of variables to be integrated.

4.3.2 Solution Using the Hamiltonian Matrix

The symplectic transition matrix which determines the flow of the extended SRDE, is the matrix exponential of a Hamiltonian matrix (a matrix that is infinitesimally symplectic) for the particular case when Eq. (4.23) is satisfied. This matrix exponential is given by

$$\Phi(t, t_0) = \exp\left(F(t)\right), \quad \text{where } F(t) = \int_{t_0}^{t} H(\tau) d\tau \tag{4.41}$$

from Eq. (4.21). The matrix $F(t)$, which is the time integral of the Hamiltonian matrix $H(t)$, is also a Hamiltonian matrix itself. $F(t)$ is the inverse matrix exponential, or the matrix logarithm of the symplectic transition matrix $\Phi(t, t_0)$. Like the skew-symmetric matrix and the orthogonal matrix which are related by the matrix exponential, $F(t)$ and $\Phi(t, t_0)$ commute in matrix multiplication. An accurate way to ensure that the norm in Eq. (4.39) remains small during numerical integration, is to ensure that the matrix $F(t)$ is Hamiltonian or close to Hamiltonian. In such a procedure, the matrix $F(t)$, instead of the transition matrix $\Phi(t, t_0)$ will have to be directly evaluated. Looking at the structure of the Hamiltonian matrix $H(t)$ in Eq. (4.18), it can be noted that the off-diagonal block matrices $B(t)$ and $-C(t)$ are symmetric, and hence have a total of $2(n(n+1)/2) = n(n+1)$ independent scalar variables. The block diagonal matrices $A(t)$ and $A^{\mathrm{T}}(t)$ have a total of n^2 independent scalar variables at most. Hence the total number of variables to be integrated in obtaining $F(t)$ from $H(t)$ is $l = n^2 + n(n+1) = n(2n+1)$. This is less than the $4n^2$ variables that were numerically integrated for obtaining Φ in section 4.3.1. The transition matrix $\Phi(t, t_0)$ is then obtained from the matrix exponential of $F(t)$.

Figure 4.3: Departure from symplecticity for flow of extended SRDE using the Hamiltonian matrix

The result of obtaining the transition matrix, for the same SRDE used in the last section, by this procedure is shown in Figure 4.3, where the norm in Eq. (4.39) is plotted against time. The time duration for integration is the same as that used in the direct numerical integration for Φ (5 seconds) in section 4.3.1. This figure shows that the norm is of the order of 10^{-13}, which is 10 orders of magnitude less than that obtained from the direct numerical integration method. The matrix solution of the SRDE is obtained from Eq. (4.34), and its departure from symmetricity is plotted next. This is shown in Figure 4.4, where the norm in

Eq. (4.40) is plotted against time. As can be seen from this figure, the departure

Figure 4.4: Departure from symmetricity of solution for SRDE using the Hamiltonian matrix

from symmetricity is extremely small, on the order of 10^{-14}. This is 11 orders of magnitude smaller than that obtained from direct numerical integration of the transition matrix. This method is useful in obtaining the symmetric matrix solution of the SRDE and avoiding the problem of $S(t)$ blowing up. However, although extension of the SRDE to $\mathcal{L}(n)$ avoids the problem of the stiffness of the differential equation, evaluation of $S(t)$ using Eq. (4.34) may lead to $S(t)$ blowing up if $X(t)$ in Radon's formula (given by Eq. (4.31)) becomes close to singular. This will happen if the solution $S(t)$ of the SRDE diverges, and there is no way out of this except changing the matrix co-efficients, or the terminal condition, or the interval of integration of the SRDE. It is known that the solution of a time-invariant SRDE with B negative definite and C positive semi-definite, does not diverge.[15,16] Using the extended form of the equation, however, entirely avoids the problem of stiffness during numerical integration even if the solution $S(t)$ diverges.

4.4 Some Useful Properties of Symplectic Matrices

As remarked in section 4.2.1, symplectic and Hamiltonian matrices are related in the same way as orthogonal and skew-symmetric matrices. However, unlike skew-symmetric and orthogonal matrices, Hamiltonian and symplectic matrices can exist only in even dimensions, i.e., as $2n \times 2n$ matrices where n is a positive integer. It is known that the matrix exponential of a Hamiltonian matrix results in a symplectic matrix, so Hamiltonian and symplectic matrices are related by the exponential map and its inverse. The symplectic matrices satisfy the same defining property with the basic symplectic matrix, J_{2n}, as do orthogonal matrices with the basic orthogonal matrix, I_n

$$\Phi^{\mathrm{T}} J_{2n} \Phi = J_{2n}, \quad C^{\mathrm{T}} I_n C = I_n \tag{4.42}$$

Because of this relation, calculation of the inverse of a symplectic matrix is almost as easy as calculating the inverse of an orthogonal matrix

$$\Phi^{-1} = -J_{2n} \Phi^{\mathrm{T}} J_{2n} \tag{4.43}$$

which is a simple rearrangement of the components of Φ. In the block-partitioned form of the transition matrix in Eq. (4.20), the inverse can be expressed as[38]

$$\Phi^{-1}(t, t_0) = \Phi(t_0, t) = \begin{bmatrix} \Phi_4^{\mathrm{T}}(t, t_0) & -\Phi_2^{\mathrm{T}}(t, t_0) \\ -\Phi_3^{\mathrm{T}}(t, t_0) & \Phi_1^{\mathrm{T}}(t, t_0) \end{bmatrix} \tag{4.44}$$

which makes it easy to evaluate the inverse if required during numerical computation. Similarly, the Hamiltonian matrices satisfy the same defining property with J_{2n} as do the skew-symmetric matrices with I_n

$$H^{\mathrm{T}} J_{2n} + J_{2n} H = 0, \quad Q^{\mathrm{T}} I_n + I_n Q = 0 \tag{4.45}$$

although J_{2n} is not a Hamiltonian matrix, and I_n is not a skew-symmetric matrix. Also to be noted is the existence of Cayley Transform-like relations between Hamiltonian and symplectic matrices. These relations are given by

$$\left.\begin{array}{l} \Phi \text{ symplectic } \Rightarrow (\Phi + I_{2n})^{-1}(\Phi - I_{2n}) \text{ Hamiltonian} \\ H \text{ Hamiltonian } \Rightarrow (H + I_{2n})^{-1}(H - I_{2n}) \text{ symplectic} \end{array}\right\} \tag{4.46}$$

provided the indicated inverses exist. Thus, the Cayley Transform map also relates Hamiltonian and symplectic matrices in the same way that it relates skew-symmetric and orthogonal matrices. Both the exponential and Cayley Transform maps preserve the commutativity of the symplectic and Hamiltonian matrices related by these maps. From chapter II, we know this to be true of the orthogonal and skew-symmetric matrices related by these maps as well.

A group is a set with a binary operation $(*) : G * G \mapsto G$, called product, which satisfies some properties like associativity, existence of a unique identity element, and existence of a unique inverse of an element. The set of orthogonal matrices, $\mathcal{O}(n)$ satisfies these group properties, and hence forms a group. A group like $\mathcal{O}(n)$, which also forms a smooth manifold in which product and inverse are smooth maps, is called a *Lie group.*[33,39] The set of $2n \times 2n$ symplectic matices is also a Lie group, denoted $\mathcal{S}p(n)$, since it satisfies all the above properties. The product of two symplectic matrices is also symplectic, as may be easily verified using Eq. (4.42). The identity element in the group of symplectic matrices is I_{2n} (not J_{2n}), and the inverse of an element is given by Eq. (4.43). Obviously, the product (matrix multiplication) and the inverse are smooth maps for this group. The matrix exponential map $\exp(A) : \mathbf{R}^{n\times n} \mapsto \mathbf{R}^{n\times n}$ which maps the Hamiltonian and skew-symmetric matrices to the symplectic and orthogonal matrices respectively, is also a smooth map. The inverse of this map, called the matrix logarithmic map, $\log(X) : \mathbf{R}^{n\times n} \mapsto \mathbf{R}^{n\times n}$ is defined only for matrices close to the identity matrix I_n as

$$\log(X) = (X - I_n) - \tfrac{1}{2}(X - I_n)^2 + \tfrac{1}{3}(X - I_n)^3 + \dots \tag{4.47}$$

It turns out that the logarithmic map gives the tangent space of a Lie group at the identity.[33,39] The elements of this tangent space for $\mathcal{O}(n)$ belong to the set of skew-symmetric matrices, denoted by $so(n)$. The elements of the tangent space at the identity element I_{2n} of $\mathcal{S}p(n)$ belong to the set of Hamiltonian matrices, denoted by $sp(n)$. $\mathcal{S}p(n)$ forms a matrix Lie group of dimension $n(2n+1)$, as shown in section 4.2.2. This is also the dimension of its tangent space at identity, $sp(n)$, the space of Hamiltonian matrices, as shown in section 4.3.2. This fact was utilized in the numerical procedure used to numerically solve the SRDE in section 4.3.2. It is also known that like $\mathcal{O}(n)$, the symplectic group $\mathcal{S}p(n)$ is also a compact manifold,[20,33,39] unlike the vector space of symmetric matrices, $\mathcal{S}(n)$. This makes the flow of the extended SRDE free from problems of stiffness that are encountered during numerical computation.

CHAPTER V

CONCLUSION

This thesis dealt with elementary geometrical transformations in general Euclidean spaces, developed from the notions of 3-dimensional Euclidean spaces with which we are all familiar. The organization of this thesis was in three parts. The first part of the thesis covered rotations and re-orientations in Euclidean spaces and generalized Euler's principal rotation theorem in 3-dimensional spaces to higher dimensional Euclidean spaces. The second part dealt with reflections and projections in Euclidean spaces and generalized the concept of orthogonal projections to non-orthogonal projections on hyperplanes passing through the origin. The third part of this thesis presented an application by way of the symplectic Riccati differential equation with symmetric matrix solutions, in which extension of the domain to a compactification of the space of symmetric matrices was found to be useful.

A rotation in an n-dimensional Euclidean space is found to occur on a plane, a 2-dimensional subspace of the Euclidean space. The $(n-2)$-dimensional subspace that is the orthogonal complement of the plane of rotation in the Euclidean space, is unaffected by the rotation. Matrix representations of rotations are developed from infinitesimal rotations, which lead to a skew-symmetric exterior two-form representation for rotations. Orthogonal and skew-symmetric matrix representations for finite rotations are developed from the exterior two-form representation. It is shown that rotations can be represented by orthogonal matrices which have a pair of unimodular complex conjugate eigenvalues corresponding to a pair of complex conjugate eigenvectors, and the eigenvalue $+1$ with an algebraic multiplicity of $(n-2)$, corresponding to the $(n-2)$ eigenvectors orthogonal to the plane of rotation. Rotations can also be represented by skew-symmetric matrices which have a pair of imaginary conjugate eigenvalues, and the eigenvalue 0 with an algebraic multiplicity of $(n-2)$. The orthogonal and skew-symmetric rotation matrices are related by the matrix exponential and the Cayley Transform maps and their inverse maps. The skew-symmetric rotation matrix obtained from the Cayley Transform, is, however, not the same as that obtained from the matrix logarithmic map. All rotation matrices representing the same rotation, share the same eigenvectors. The skew-symmetric rotation matrices have $(n-2)$ zero eigenvalues and a pair of imaginary conjugate eigenvalues. All rotation matrices can be represented by $(2n-3)$ scalar parameters which give the plane and the angle of rotation.

In the second half of chapter II, it is shown that a total re-orientation in an n-dimensional Euclidean space can be obtained from $m = \lfloor n/2 \rfloor$ rotations on m orthogonal planes in the space. This result is the generalization of Euler's principal rotation theorem to higher dimensional Euclidean spaces. An orientation in an Euclidean space can be specified by orthogonal and skew-symmetric matrices, which are related to each other by the matrix exponential and Cayley Transform maps and their inverse maps. Since the $\lfloor\ \rfloor$ function rounds to the nearest integer towards zero, re-orientations in odd dimensional Euclidean spaces always leave any vector along a particular direction unchanged. Orthogonal orientation matrices have pairs of unimodular complex conjugate eigenvalues corresponding to the complex conjugate eigenvectors spanning the $m = n/2$ planes, and an extra eigenvalue of 1 if the matrix is of odd dimension. A skew-symmetric orientation matrix has m pairs of imaginary conjugate eigenvalues corresponding to the complex conjugate eigenvectors spanning the m planes and an extra 0 eigenvalue if the matrix is of odd dimension. Orthogonal and skew-symmetric orientation matrices representing the same orientation have the same eigenvectors. Orientation matrices in n-dimensional Euclidean space are shown to be specified by ${}^nC_2 = n(n-1)/2$ unique scalar parameters. It is found that the set of ortho-skew matrices, which are both orthogonal and skew-symmetric, are obtained by rotating each plane in a set of m planes in an n-dimensional space by 90 degrees.

They exist only in even dimensional spaces and behave like extensions of the imaginary unit to the even ordered matrix spaces.

A reflection in an n-dimensional Euclidean space occurs along a subspace of any dimension from 1 to n. The orthogonal complement of the subspace along which the reflection occurs, bisects the line joining a point and its reflection, and is orthogonal to this line. A point and its reflection have the same distance from the subspace along which the reflection occurs, and are not separated by it. To represent a reflection in an Euclidean space, the set of ortho-symmetric matrices, which are both orthogonal and symmetric, is introduced. The ortho-symmetric matrix which represents a reflection along an m-dimensional subspace of an n-dimensional Euclidean space, can be obtained from the m orthogonal unit vectors spanning this subspace. This ortho-symmetric matrix has m eigenvalues of -1 and $(n-m)$ eigenvalues of $+1$, corresponding to the orthonormal eigenvectors spanning the subspace along which the reflection occurred, and its orthogonal complement, respectively. It is shown that the negative of this ortho-symmetric matrix carries out a reflection along the $(n-m)$-dimensional orthogonal complement of this subspace. This ortho-symmetric matrix can be specified by $m(n-m)$ parameters. The Householder matrices which are often used in numerical linear algebra routines, are found to belong to the set of ortho-skew matrices. They carry out reflections along axes (1-dimensional subspaces) and can hence be described as elementary reflection matrices. All reflections are shown to be achieved by a combination of elementary reflections along a set of orthogonal axes in an Euclidean space.

Projections in n-dimensional Euclidean spaces, like reflections, occur along subspaces of any dimension from 1 to n. A projection of a point along an m-dimensional subspace of an n-dimensional space has the same orthogonal distance from this subspace as the point itself. However, the orthogonal complement of this subspace may not be equidistant from the point and its projection and may not even separate them. Projections include both reflections and orthogonal projections as special cases. An orthogonal projection can be represented by an idempotent symmetric matrix. This matrix has m zero eigenvalues and $(n-m)$ eigenvalues of $+1$, corresponding to the orthonormal eigenvectors spanning the subspace along which the orthogonal projection occurs, and its orthogonal complement, respectively. General projections in Euclidean spaces are represented by a set of symmetric matrices introduced in this thesis, called the projection matrices. These matrices have m real eigenvalues and $(n-m)$ eigenvalues of $+1$, corresponding to the orthonormal eigenvectors spanning the subspace along which the projection occurred, and its orthogonal complement, respectively. Such a projection matrix can be represented by $m(m+1)/2$ parameters. The modified Householder matrices, which carry out projections along an axis (a 1-dimensional subspace) in an Euclidean space, are a subset of the projection matrices. They may also be called elementary projection matrices. All projections in an Euclidean space are shown to be achieved by a combination of elementary projections along a set of orthogonal axes spanning the space. The last part of chapter III presents two decompositions of symmetric matrices in terms of the Householder and modified Householder matrices. The decomposition using the modified Householder (elementary projection) matrices can be expressed as either a sum or a product decomposition, and is shown to be more natural and more efficient to numerical computation than the decomposition using Householder matrices.

The symplectic Riccati differential equation (SRDE), which has the space of symmetric matrices as an invariant manifold, is studied in chapter IV. The extension of the domain from the space of symmetric matrices to the compact Lagrange-Grassmann manifold is found to be useful in understanding its geometric properties and also in numerical computation of its solution. The SRDE is introduced with a common application in which it arises; that of linear quadratic optimal control. The flow of the SRDE is constructed from the extended version of the SRDE. The extended SRDE has a symplectic flow, which gives the equation its name. The flow of the extended SRDE is then related to the spectral decomposition of the symmetric matrix solution and the terminal condition. The solution in the extended domain of the Lagrange-Grassmann manifold is related to the solution in the space of symmetric matrices by Radon's formula. Since the Lagrange-Grassmann manifold is a compact manifold, the solutions in this manifold correspond to finite symmetric matrix solutions of the SRDE. The solution of the extended SRDE does not blow up even when the solution of the SRDE becomes very large. This makes the numerical computation of the flow in the extended SRDE an attractive option. Two numerical methods to compute the solution of the SRDE are used in this chapter. The first method is direct numerical integration of the transition matrix representing the flow in the extended SRDE. This method does not check for the symplecticity of the transition matrix, and hence it is not accurate when integrated over appreciably large time-intervals. The second method numerically integrates for the Hamiltonian matrix, and ensures that it remains close to Hamiltonian throughout the integration. For the special case that the Hamiltonian matrix always commutes with its integral, the transition matrix is obtained from the matrix exponential map. For a more general case, the transition matrix may be obtained from numerical integration of the Hamiltonian with some form of symplectic updating. This method is numerically far more accurate than the first method, and the transition matrix remains almost symplectic throughout the integration period. The last part of this chapter presents some interesting and useful properties of symplectic and Hamiltonian matrices, which may be used in numerical procedures for solving the SRDE.

A potential application of the results in presented here for orthogonal matrices is in obtaining analytic singular value decompositions, in terms of orthogonal and real diagonal matrices, of real analytic matrices. Real analytic time-varying symmetric matrices are often encountered as inertia matrices in mechanical systems like robots, which have moving members. The decompositions of symmetric matrices presented here could be used in such applications. There are many other potential applications in the field of numerical linear algebra, where orthogonal and symmetric matrices are often used for matrix decompositions. These include the CS decomposition, which decomposes an orthogonal matrix into block-diagonal orthogonal matrices, the LU factorization which decomposes a general matrix into lower and upper triangular matrices, and the QR factorization, which decomposes a general matrix into an orthogonal and an upper triangular matrix. Extending the decompositions presented here to real analytic time-varying orthogonal and symmetric matrices, would be a natural outcome of the work presented in this thesis.

REFERENCES

[1]Euler, L., "Formulae Generales pro Trandlatione Quacunque Corporum Rigidorum," *Novi Acad. Sci. Petrop.*, Vol. 20, 1775, pp. 189-207.

[2]Mortari, D., "On the Rigid Rotation Concept in the n-Dimensional Spaces," International Conference on Non Linear Problems in Aeronautics and Astronautics, Daytona Beach, FL, May 2000.

[3]Mortari, D., "On the Rigid Rotation Concept in n-Dimensional Spaces," *Journal of the Astronautical Sciences* (to be published).

[4]Bar-Itzhack, I.Y., "Extension of Euler's Theorem to n-Dimensional Spaces," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-25, No. 6, 1989, pp. 903-909.

[5]Shuster, M.D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 439-517.

[6]Tsiotras, P., and Longuski, J.M., "A New Parametrization of the Attitude Kinematics," *Journal of the Astronautical Sciences*, Vol. 43, No. 3, 1996, pp. 342-362.

[7]Marandi, S.R., and Modi, V.J., "A Preferred Coordinate System and the Associated Orientation Representation in Attitude Dynamics," *Acta Astronautica*, Vol. 15, No. 11, 1987, pp. 833-843.

[8]Householder, A.S., *The Theory of Matrices in Numerical Analysis*, Dover Publications, New York, 1964.

[9]Stewart, G.W., *Introduction to Matrix Computations*, Academic Press, New York, 1973.

[10]Crassidis, J.L., and Junkins, J.L., *An Introduction to Optimal Estimation of*

Dynamical Systems, (draft manuscript, preprint from authors), pp. 42-45.

[11]Stewart, G.W., "On the Perturbation of Pseudo-Inverses, Projections and Linear Least Squares Problems," *SIAM Review*, Vol. 19, 1977, pp. 634-662.

[12]Luenberger, D., *Optimization by Vector Space Methods*, John Wiley and Sons, Inc., New York, 1969.

[13]Golub, G.H., and Van Loan, C.F., *Matrix Computations*, 3rd. ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[14]Bittanti, S., Laub, A.J., and Willems, J.C. (eds.), *The Riccati Equation*, Springer-Verlag, Berlin, Germany, 1991.

[15]Bryson, A.E., and Ho, Y., *Applied Optimal Control*, Hemisphere Publishing Corporation, Washington, DC, 1975.

[16]Kirk, D.E., *Optimal Control Theory: An Introduction*, Prentice Hall, Inc., Network Series, Englewood Cliffs, NJ, 1970.

[17]Reid, W.T., *Riccati Differential Equations*, Academic Press, New York, 1972.

[18]Goldstein, H., *Classical Mechanics*, 2nd. ed., Addison-Wesley Publishing Company, Inc., Reading, MA, 1980.

[19]Arnold, V.I., *Mathematical Methods of Classical Mechanics*, 2nd. ed., Springer-Verlag, New York, 1989.

[20]Abraham, R., Marsden, J.E., and Ratiu, T., *Manifolds, Tensor Analysis, and Applications*, 2nd. ed., Springer-Verlag, New York, 1988.

[21]Spivak, M., *Differential Geometry*, vols. 1-5, Publish or Perish, Inc., Waltham, MA, 1979.

[22]Flanders, H., *Differential Forms*, Academic Press, New York, 1963.

[23]Kobayashi, S., and Nomizu, K., *Foundations of Differential Geometry*, Wiley,

New York, 1963.

[24]Junkins, J.L., and Kim, Y., *Introduction to Dynamics and Control of Flexible Structures*, AIAA Education Series, AIAA, Reston, VA, 1993.

[25]Cayley, A., "On the Motion of Rotation of a Solid Body," *Cambridge Mathematics Journal*, Vol. 3, 1843, pp.224-232.

[26]Schaub, H., *Novel Coordinates for Nonlinear Multibody Motion with Applications to Spacecraft Dynamics and Control*, Ph.D. Dissertation, Texas A&M University, College Station, TX, 1998.

[27]Davis, P.J., *Interpolation and Approximation*, Blaisdell, New York, 1965.

[28]Kantorovich, L.V., and Akilov, G.P., "The Extension of Linear Operations and Functionals," *Functional Analysis in Normed Spaces*, trans. by D.E. Brown, ed. by A.P. Robertson, Macmillan, New York, 1964, pp. 141-154.

[29]Luisternik, L., and Sobolev, V., *Elements of Functional Analysis*, Frederick Ungar, New York, 1961.

[30]Riccati, Count J.F., "Animadversationes in aequationes differentiales secundi gradus," *Actorum Eruditorum quae Lipsiae publicantur, Supplementa* 8, 1724, pp. 66-73.

[31]Lanczos, C., *The Variational Principles of Mechanics*, University of Toronto Press, Toronto, Canada, 1949.

[32]Rugh, W.J., *Linear System Theory*, 2nd. ed., Prentice Hall, Inc., Upper Saddle River, NJ, 1996.

[33]Sastry, S., *Nonlinear Systems: Analysis, Stability, and Control*, Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 1999.

[34]Bunse-Gerstner, A., Byers, R., Mehrmann, V., Nichols, N.K., "Numerical Computation of an Analytic Singular Value Decomposition of a Matrix Valued Func-

tion", *Numerische Mathematik*, Vol. 60, 1991, pp. 1-39.

[35]Byers, R., and Mehrmann, V., "Symmetric Updating of the Solution of the Algebraic Riccati Equation," *Proceedings of 10th Symposium on Operations Research*, Vol. 54, Universität München, August 1985, pp. 117-125.

[36]Choi, C., and Laub, A.J., "Efficient Matrix-Valued Algorithms for Solving Stiff Riccati Differential Equations," *IEEE Transactions on Automatic Control*, Vol. AC-35, 1990, pp. 770-776.

[37]Shayman, M.A., "On the Periodic Solutions of the Matrix Riccati Equation," *Mathematical Systems Theory*, Vol. 16, 1983, pp. 267-287.

[38]Battin, R.H., *An Introduction to the Mathematics and Methods of Astrodynamics*, rev. ed., AIAA Education Series, AIAA, Reston, VA, 1999.

[39]Bredon, G.E., *Topology and Geometry*, Springer-Verlag, New York, 1993.

VITA

Amit Kumar Sanyal is the younger son of Archita and Ramen Sanyal of Asansol, India. Amit received a Bachelor of Technology in aerospace engineering from the Indian Institute of Technology, Kanpur, in May 1999. He went to Texas A&M University to begin his graduate studies in August 1999.

During his master's level studies at Texas A&M, Amit worked in adaptive control, attitude estimation and with CCD star cameras. His master's thesis was directed at understanding and applying geometrical transformations in Euclidean spaces, and was done under the supervision of Dr. John L. Junkins. After completing his M.S. degree, Amit plans to continue his graduate studies at the doctoral level. His permanent mailing address is 2, Shreshtha Apartments, Asansol, W.B. 713304, India.