KEENAN

VOLUMES IN GRASSMANN MANIFOLDS

MSc.

Volumes of Balls in Grassmann Manifolds with Applications to Coding Theory

By

Patrick Keenan, B.Math

A project thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree Master of Science.

McMaster University

MASTER OF SCIENCE (2008)

(Math)

McMaster University

Hamilton, Ontario

TITLE: Volumes of Balls in Grassmann Manifolds with Applications to Coding Theory

AUTHOR: Patrick Jordan Keenan, B.Math (University of Windsor)

SUPERVISOR: Dr. Maung Min-Oo

NUMBER OF PAGES: vii, 62

Abstract

This thesis develops the Riemannian Geometry of the real and complex Grassmann Manifolds in a notationally accessible way. The canonical volume form is related to explicit Jacobi Field calculations. The implementation of a packing algorithm based on repulsive forces is proposed. Standard packing bounds and bounds on the volumes of geodesic balls are used to test the performance of the algorithm.

Acknowledgments

A large number of people contributed directly or indirectly to the completion of this thesis. I would like to thank my supervisor Dr. M. Min-Oo for showing me his insightful approach to Geometry, and for knowing, quite literally in the case of initial tangents to geodesics, the direction in which to go with this project. Many helpful pieces of advice were contributed by Dr. M. Harada, Dr. A. Nicas, Dr. M. Wang, and Dr. R. Gohary. I would also like to thank Dr. T. Traynor, and Dr. Z. Hu for encouraging my interest in math. Most of all thank you to my parents E. Lindner-Keenan and B. Keenan. Financial support from McMaster University and the Natural Sciences and Engineering Council of Canada is gratefully acknowledged.

Contents

List of Figures

1 Introduction

Telecommunication engineers have happened upon motivation to study the Grassmann Manifold in connection with multi-antenna wireless schemes. To sketch the connection consider the case of a transmitter equipped with ℓ antennas and a receiver having ℓ antennas. Let $s_{1\times\ell}$ be a complex-valued vector representing a symbol to be sent. In the absence of noise the sent symbol s and the received symbol s^* have the relationship

$$s^* = sM$$

where $M_{\ell\times\ell}$ is called the matrix of fading coefficients between the antennas. When antennas are moving it is difficult to know exactly what the coefficients are for any extended time, so M is taken to be a matrix that is approximately valid for some time block consisting of T symbol periods. In the Rayleigh flat-fading model M is assumed to be Rayleigh distributed. When noise is considered, sending the block of symbols $S_{T\times\ell}$ results in receiving the block

$$S^* = SM + W$$

where the Gaussian distributed $W_{T\times\ell}$ is called the additive white Gaussian noise. In an important paper by [19] it was shown that there is no gain if $\ell > T/2$ so it is assumed that $\ell \leq T/2$. The ℓ -dimensional subspace $\operatorname{col} S$ is preserved by the transformation when no noise is present. This is because there is a probability of one that the matrix M is invertible, ie. it is very likely that the column spaces $\operatorname{col} SM$ and $\operatorname{col} S$ are the same. $\operatorname{col} S$ may be viewed as a point in the Grassmannian $\mathbf{Gr}_{\ell,T}(\mathbb{C})$. Consider a finite basic alphabet of signal blocks $\{S_j\}$. One way to increase the reliability of error checking is to ensure that the points $\operatorname{col} S_j$ are, in a sense to be made clear within, well spread out on $\mathbf{Gr}_{\ell,T}(\mathbb{C})$. Intuitively, if the points are well spread out before being sent then it is likely that they will still be well spread out when they are received in which case it is easier to distinguish them.

Section two defines the Grassmann and closely related Stiefel manifolds and deals with the form of tangents and canonical metrics on each manifold. Section 3 develops the necessary Riemannian-geometric tools for **Gr** such as parallel translation, covariant and Lie derivatives, curvature, and the volume form with emphasis on explicit computation. Section 4 proposes an algorithm to spread out points on **Gr** using repulsive forces and compares some preliminary results with standard packing bounds.

2 The Stiefel and Grassmann Manifolds

2.1 Definitions, Dimensions, and Coordinates

In the following $\mathbb{R}^n$ may be replaced by $\mathbb{C}^n$ and the group of orthogonal $n \times n$ matrices $\mathbf{O}_n$ by the unitary matrices $\mathbf{U}_n$ without changing the essential development. For simplicity the real case is discussed and extended later to the complex case.

Definition 2.1. The **Stiefel Manifold** $\mathbf{St}_{k,n}(\mathbb{R})$ is defined to be the set of all orthonormal (abbreviated ON) matrices of size $n \times k$, that is

$$\mathbf{St}_{k,n}(\mathbb{R}) = \{P \in \mathbb{R}^{n \times k} : P^T P = I_k\}.$$

St will stand for $\mathbf{St}_{k,n}(\mathbb{R})$ at first but will later stand for $\mathbf{St}_{k,n}(\mathbb{F})$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Writing the i^{th} column of P as $p_{\downarrow i}$, the condition

$$I = P^T P = \begin{pmatrix} p_{\downarrow 1}^T p_{\downarrow 1} & \cdots & p_{\downarrow 1}^T p_{\downarrow k} \\ \vdots & & \vdots \\ p_{\downarrow k}^T p_{\downarrow 1} & \cdots & p_{\downarrow k}^T p_{\downarrow k} \end{pmatrix}$$

shows that $P^T P = I$ represents $\frac{k(k+1)}{2}$ independent restrictions on the $n \times k$ matrix P. Since $\mathbb{R}^{n \times k} \simeq \mathbb{R}^{nk}$ this suggests that **St** is an $nk - \frac{k(k+1)}{2}$ -dimensional manifold. The proof of this will be included in Section 2.4. **St** may be equivalently defined by taking certain equivalence classes of matrices in $\mathbf{O}_n$:

$$\mathbf{St} \ni [Q] = \{S \in \mathbf{O}_n : s_{\downarrow 1} = q_{\downarrow 1}, \dots, s_{\downarrow k} = q_{\downarrow k}\}$$

$$= \left\{ P \in \mathbf{O}_n : P = Q \begin{pmatrix} I_k & 0 \\ 0 & U \end{pmatrix}, U \in \mathbf{O}_{n-k} \right\} = Q \begin{pmatrix} I_k & 0 \\ 0 & \mathbf{O}_{n-k} \end{pmatrix}$$

In other words a point $[Q] \in \mathbf{St}$ can be taken to be all ON bases of $\mathbb{R}^n$ where the first k basis vectors are identical. The matrix $\begin{pmatrix} I_k & 0 \\ 0 & \mathbf{O}_{n-k} \end{pmatrix}$ here is called an **isotropy group**. In this form it is clear that $\mathbf{St} \simeq \mathbf{O}_n / \mathbf{O}_{n-k}$.

Definition 2.2. The **Grassmannian Manifold** $\mathbf{Gr}_{k,n}(\mathbb{R})$ is defined to be the set of all k-dimensional subspaces of $\mathbb{R}^n$.

$\mathbf{Gr}$ will stand for $\mathbf{Gr}_{k,n}(\mathbb{R})$ at first but will later stand for $\mathbf{Gr}_{k,n}(\mathbb{F})$, in any case the meaning of $\mathbf{Gr}$ should be taken in context. It is assumed that $k \leq n/2$, otherwise the roles of n and k may be switched. That $k \leq n/2$ is not always assumed in the Stiefel case, for example, $\mathbf{St}_{n,n}(\mathbb{R}) = \mathbf{O}_n$.

A convenient way to represent points in $\mathbf{Gr}$ by $n \times k$ matrices is to identify matrices $P_{n\times k} \in \mathbf{St}$ whose columns span a given k-dimensional subspace of $\mathbb{R}^n$. This suggests the equivalence class $[P] = P\mathbf{O}_k$. Since the representative P will often be used to specify the point $\operatorname{span}\{p_{\downarrow 1}, \ldots, p_{\downarrow k}\}$ it will sometimes be convenient to write $\operatorname{span} P$ instead of $\operatorname{col} P = \operatorname{span}\{p_{\downarrow 1}, \ldots, p_{\downarrow k}\}$. As with the Stiefel case, points in $\mathbf{Gr}$ can be represented by equivalence classes of $n \times n$ orthogonal matrices under the identification

$$\mathbf{Gr} \ni [Q] = Q \begin{pmatrix} \mathbf{O}_k & 0 \\ 0 & \mathbf{O}_{n-k} \end{pmatrix}$$

where $\operatorname{span}\{q_{\downarrow 1}, \ldots, q_{\downarrow k}\}$ is the k-plane being specified so that $\mathbf{Gr} \simeq \mathbf{O}_n / (\mathbf{O}_k \times \mathbf{O}_{n-k})$. This identification makes intuitive sense: Let Q be partitioned as $Q = \begin{pmatrix} P & P_\perp \end{pmatrix}$, and let $M = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} \in \begin{pmatrix} \mathbf{O}_k & 0 \\ 0 & \mathbf{O}_{n-k} \end{pmatrix}$, where $P_\perp$ is any orthogonal $n \times n-k$ matrix for which $\operatorname{span} P_\perp = (\operatorname{span} P)_\perp$, or $P^T P_\perp = 0$. Then $QM = \begin{pmatrix} PM_1 & P_\perp M_2 \end{pmatrix} \in \mathbf{O}_n$, and $\operatorname{span} PM_1 = \operatorname{span} P$. Typically in both $\mathbf{St}$ and $\mathbf{Gr}$ points $[Q_{n\times n}]$ and $[P_{n\times k}]$ will be denoted Q and P. To see that $\mathbf{Gr}$ is a manifold of dimension $k(n-k)$, pick a point $Q = \begin{pmatrix} P & P_\perp \end{pmatrix} \in \mathbf{Gr}$. If $x \in \mathbb{R}^n$ lies in $\operatorname{span} P$, then $x^T P_\perp = 0$. Since $\operatorname{rank} P_\perp = n-k$, there is an invertible submatrix $P_{\perp\alpha} = \begin{pmatrix} p_{\alpha_1 \rightarrow} \\ \vdots \\ p_{\alpha_{n-k} \rightarrow} \end{pmatrix}$ of $P_\perp$. If $P_{\perp\beta}$ denotes the matrix that remains when the rows $p_{\alpha_i \rightarrow}$ are deleted, then the condition $0 = x^T P_\perp = \sum_{i=1}^{n} x_i p_{i\rightarrow} = x_\alpha^T P_{\perp\alpha} + x_\beta^T P_{\perp\beta}$

can be written $-\begin{pmatrix} x_{\beta_1} & \cdots & x_{\beta_k} \end{pmatrix} P_{\perp\beta}P_{\perp\alpha}^{-1} = \begin{pmatrix} x_{\alpha_1} & \cdots & x_{\alpha_{n-k}} \end{pmatrix}$. The $k \times n-k$ matrix $Z := -P_{\perp\beta}P_{\perp\alpha}^{-1}$ provides a coordinate system on the set $U_{\beta_1,\ldots,\beta_k}$ of all k-planes whose equation can be written in the form $\begin{pmatrix} x_{\beta_1} & \cdots & x_{\beta_k} \end{pmatrix} Z = \begin{pmatrix} x_{\alpha_1} & \cdots & x_{\alpha_{n-k}} \end{pmatrix}$. The proof that $\dim \mathbf{Gr} = k(n-k)$ will be included in Section 2.5. In the complex case the same reasoning can be used to see that $\dim_{\mathbb{R}} \mathbf{Gr}(\mathbb{C}) = 2k(n-k)$. The atlas $\{(Z, U_{\beta_1,\ldots,\beta_k})\}$ is that used in the classical developments by [3]. Although the Riemannian Geometry of **Gr** can be developed in terms of these coordinates it will be more convenient to represent k-planes by the matrices $P_{n\times k}$ or $Q_{n\times n}$ as in the relatively recent developments by [8].

Both **St** and **Gr** are of the form **G**/**K** where **G** is the compact Lie group $\mathbf{O}_n$ and **K** is the appropriate isotropy group. **G**/**K** is called a **homogeneous space** because **G** is a connected Lie group and **K** is a closed subgroup of **G**.

Gr can be defined using only oriented representations $Q \in \mathbf{SO}_n$, the special orthogonal matrices, resulting in $\mathbf{Gr} \simeq \mathbf{SO}_n/(\mathbf{SO}_k \times \mathbf{SO}_{n-k})$ but this leads to less intuitive results when defining what are called the principal angles between subspaces of $\mathbb{R}^n$ and causes complications in computation. Similarly **Gr** can also be defined with invertible matrices using $\mathbf{Gr} \simeq \mathbf{GL}_n/(\mathbf{GL}_k \times \mathbf{GL}_{n-k})$ resulting in correction factors in calculation. This is illustrated in the case of projecting a vector $a \in \mathbb{R}^n$ onto the k-dimensional subspace P.

Proposition 2.1. *Consider the matrix $P_{n\times k}$, not necessarily ON, that specifies the point* span $P \in Gr$.

i) If $a \in \mathbb{R}^n$ and the orthogonal projection of a onto span P *is denoted by $a_P = \Pi_P(a)$ then $\Pi_P = P(P^TP)^{-1}P^T$ which reduces to $\Pi_P = PP^T$ when P is an orthogonal matrix.*

ii) If $\Pi_{P_\perp}$ denotes projection onto $(\operatorname{span} P)_\perp$ *then $\Pi_{P_\perp} = (I - P(P^TP)^{-1}P^T)$ which reduces to $\Pi_{P_\perp} = (I_n - PP^T)$ when P is an orthogonal matrix.*

Proof. Assume that P is possibly not orthogonal. Let $\beta = \{b_1, \ldots, b_k\}$ be an ON basis of span P and let $B := \begin{pmatrix} b_1 & \cdots & b_k \end{pmatrix}$, then $B = PM$ where $M \in \mathbf{GL}_k$ and

$$BB^{\mathcal{T}} = B(B^{\mathcal{T}}B)^{-1}B^{\mathcal{T}} = PM((PM)^{\mathcal{T}}(PM))^{-1}(PM)^{\mathcal{T}} = PM(M^{\mathcal{T}}P^{\mathcal{T}}PM)^{-1}M^{\mathcal{T}}P^{\mathcal{T}}$$
$$= PMM^{-1}P^{-1}P^{-\mathcal{T}}M^{-\mathcal{T}}M^{\mathcal{T}}P^{\mathcal{T}} = P(P^{\mathcal{T}}P)^{-1}P^{\mathcal{T}}$$

so it may be assumed that $P = \begin{pmatrix} b_1 & \cdots & b_k \end{pmatrix}$. Now

$$a_P = \sum_{i=1}^{k} \langle a, b_i \rangle b_i = \begin{pmatrix} b_1 & \cdots & b_k \end{pmatrix} \begin{pmatrix} \langle a, b_1 \rangle \\ \vdots \\ \langle a, b_k \rangle \end{pmatrix} = \begin{pmatrix} b_1 & \cdots & b_k \end{pmatrix} \begin{pmatrix} b_1^{\mathcal{T}} \\ \vdots \\ b_k^{\mathcal{T}} \end{pmatrix} a$$
$$= PP^{\mathcal{T}}a = P(P^{\mathcal{T}}P)^{-1}P^{\mathcal{T}}a.$$

Note that if the underlying field is $\mathbb{C}$, then $\mathcal{T}$ is replaced by $\mathcal{H}$, the Hermitian conjugate, and $\langle a, b_i \rangle$ is defined as $\sum_{j=1}^{n} a_j \overline{b_{ij}}$. Since $a_{P_\perp} = a - a_P = (I_n - P(P^{\mathcal{T}}P)^{-1}P^{\mathcal{T}})a$, it must be that $\Pi_{P_\perp} = I_n - P(P^{\mathcal{T}}P)^{-1}P^{\mathcal{T}}$. □

Henceforth representative matrices will be assumed ON.

2.2 Principal Angles and Angle Directions

An important way of specifying the relationship between two k-dimensional subspaces of $\mathbb{R}^n$ is to use principal angles.

Definition 2.3. The principal angles

$$\frac{\pi}{2} = \theta_1 = \cdots = \theta_r > \theta_{r+1} \geq \cdots \geq \theta_\ell > \theta_{\ell+1} = \cdots = \theta_k = 0 \quad (r \text{ possibly } 0 \text{ , } \ell \text{ possibly } k)$$

between two k-dimensional subspaces U and V are defined by the following process:

$$\cos\theta_k = \max\{|\langle u, v\rangle| : \|u\| = \|v\| = 1, u \in U, v \in V\}, \text{ or equivalently,}$$
$$\theta_k = \min\{\cos^{-1}|\langle u, v\rangle| : \|u\| = \|v\| = 1, u \in U, v \in V\},$$
$$\vdots$$
$$\theta_{k-i} = \min\{\cos^{-1}|\langle u, v\rangle| : \|u\| = \|v\| = 1, u \in U \cap (\operatorname{span}\{u_k, u_{k-1}, \ldots, u_{k-i+1}\})_\perp,$$

$$v \in V \cap (\text{span}\{v_k, v_{k-1}, \ldots, v_{k-i+1}\})_\perp\}.$$

The angles $\{\theta_i\}_{i=1}^k$ are produced in reverse order simply so that the final list is ordered from greatest to least. Any u_i, v_i that furnish these minimums are called **angle directions** corresponding to the angle θ_i. When $\theta_i \neq 0$, u_i and v_i span a 2-dimensional plane called the i^{th} **angle 2-plane** . When $\theta_i = 0$, span $\{u_i, v_i\}$ is a line. There are ℓ angle 2-planes and $k - \ell$ angle 1-planes. This method of generating $\{(u_i, v_i, \theta_i)\}_{i=1}^k$ will be termed **method** 1.

The inconvenient convention of generating $\{\theta_i\}_{i=1}^k$ in reverse order with method 1 is justified since it is in agreement with the predominant notation in the literature when dealing with representations of points in terms of principal angles. Consider the following example where $U, V \in \mathbf{Gr}_{2,4}$.

Example 2.1.

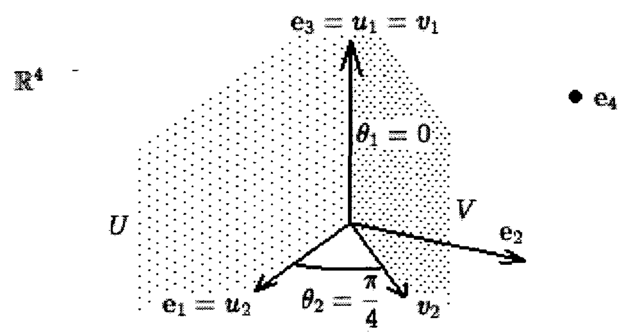

Figure 1: Principal Angles Between Subspaces of $\mathbb{R}^4$

$$U = \begin{pmatrix} u_1 & u_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \; V = \begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Since $\dim U \cap V = 1$, there is one zero angle. In general if $\dim U \cap V = k - \ell$, there are $k - \ell$ zero angles. Notice also that $u_1 \notin \text{span}\{u_2, v_2\}$ and $v_1 \notin \text{span}\{u_2, v_2\}$. When using method 1 in other dimensions, because of the conditions $u_i \in U \cap (\text{span}\{u_k, u_{k-1}, \ldots, u_{i+1}\})_\perp$ and $v_i \in V \cap (\text{span}\{v_k, v_{k-1}, \ldots, v_{i+1}\})_\perp$, it is easy to see that when

$$(u_i, v_i, \theta_i) \in \{(u_k, v_k, 0), (u_{k-1}, v_{k-1}, 0), \dots, (u_{\ell+1}, v_{\ell+1}, 0), (u_\ell, v_\ell, \theta_\ell)\},$$

$|\langle u_i, u_j \rangle| = |\langle v_i, v_j \rangle| = \delta_{ij}$, where δ_{ij} is the Kronecker delta. It is in fact possible to choose ON bases $\{u_i\}_{i=1}^k$ and $\{v_i\}_{i=1}^k$ of U and V so that $|\langle u_i, v_i \rangle| = \cos\theta_i$ and $\text{span}\{u_i, v_i\} \perp \text{span}\{u_j, v_j\}$ when $i \neq j$. With these bases there is the convenient identity $|\langle u_i, v_j \rangle| = \delta_{ij}\cos\theta_i$. This claim is easily seen once it is established that there is a rotation $R \in \mathbf{O}_n$ and representations U and V such that

$$RU = \begin{pmatrix} I_k \\ 0 \end{pmatrix} \quad \text{and} \quad RV = \begin{pmatrix} \cos\theta_1 & & 0 \\ & \ddots & \\ 0 & & \cos\theta_k \\ \sin\theta_1 & & 0 \\ & \ddots & \\ 0 & & \sin\theta_k \\ & 0 & \end{pmatrix},$$

$$\text{so that } |\langle u_{\downarrow i}, v_{\downarrow j} \rangle| = |u_{\downarrow i}^T v_{\downarrow j}| = |u_{\downarrow i}^T R^T R v_{\downarrow j}| = |(Ru_{\downarrow i})^T (Rv_{\downarrow j})| = |\mathbf{e}_i^T(\cos\theta_j\, \mathbf{e}_j + \sin\theta_j\, \mathbf{e}_{j+k})| = \delta_{ij}\cos\theta_j.$$

The existence of this popular normal form, however, is most easily proven with the identity $|\langle u_i, v_j \rangle| = \delta_{ij}\cos\theta_i$ at hand.

Claim 2.1. *There exist ON bases $\{u_i\}_{i=1}^k$ and $\{v_i\}_{i=1}^k$ of U and V satisfying $\text{span}\{u_i, v_i\} \perp \text{span}\{u_j, v_j\}$ when $i \neq j$.*

Proof. Consider the $(k - i + 1)^{th}$ step in method 1 that produces (u_i, v_i, θ_i) where the angle between

$$u_i \in U \cap (\underbrace{\text{span}\{u_k, u_{k-1}, \dots, u_{i+1}\}}_{:=S_U})_\perp$$
$$\text{and } v_i \in V \cap (\underbrace{\text{span}\{v_k, v_{k-1}, \dots, v_{i+1}\}}_{:=S_V})_\perp$$

is as small as possible. If $u_i \notin S_{V\perp}$ say $u_i = c_1 \underbrace{\alpha}_{\in U\cap S_{U\perp}\cap S_V} + c_2 \underbrace{\beta}_{\in U\cap S_{U\perp}\cap S_{V\perp}}$, where $c_1, c_2 \in \mathbb{R}$, then since $v_i \in S_{V\perp}$,

$$\pm\cos\theta_i = \langle v_i, u_i\rangle = c_1\langle v_i, \alpha\rangle + c_2\langle v_i, \beta\rangle = c_2\langle v_i, \beta\rangle.$$

So if $u_i^* := c_2\beta \in U \cap S_{U\perp} \cap S_{V\perp} \subset (\operatorname{span}\{v_k, v_{k-1}, \dots, v_{i+1}\})_\perp$ then $\frac{u_i^*}{||u_i^*||}$ has the same angle with v_i as u_i, is of length one, and is still in $S_{U\perp}$ so it may replace u_i. Similarly v_i may be chosen to have the desired properties. This shows that the required bases exist. □

Of course by replacing some of the u_i's or v_i's with $-u_i$ or $-v_i$, bases can be found so that $\langle u_i, v_j\rangle = \delta_{ij}\cos\theta_i$. The relationship between $\{u_i\}_{i=1}^k$ and $\{v_i\}_{i=1}^k$ can be clarified further.

Proposition 2.2. *If the angles θ_i and θ_j are not both $\frac{\pi}{2}$, then the angle planes* $\operatorname{span}\{u_i, v_i\}$ *and* $\operatorname{span}\{u_j, v_j\}$ *are orthogonal regardless of the choice of u_i, v_i minimizing θ_i and u_j, v_j minimizing θ_j.*

Proof. Assume that $j < i$ so that u_j and v_j are produced by method 1 later than u_i and v_i and that θ_i and θ_j are not both $\frac{\pi}{2}$. Suppose, since $\langle u_i, u_j\rangle = 0$, that $u_j = c_1\alpha + c_2 v_i$ where $\alpha \in (\operatorname{span}\{u_i, v_i\})_\perp$. Then $0 = \langle u_i, u_j\rangle = c_2\langle u_i, v_i\rangle = \pm c_2\cos\theta_i$. This implies that either $c_2 = 0$ or $\cos\theta_i = 0 \Longrightarrow \theta_i = \frac{\pi}{2} \Longrightarrow \theta_j = \frac{\pi}{2}$ contrary to the hypothesis. Therefore $c_2 = 0$ and $u_j \in (\operatorname{span}\{u_i, v_i\})_\perp$. Similarly $v_j \in (\operatorname{span}\{u_i, v_i\})_\perp$ so that the planes are orthogonal. □

Let $\Theta = \begin{pmatrix} \theta_1 & & 0 \\ & \ddots & \\ 0 & & \theta_k \end{pmatrix}$, $\cos\Theta = \begin{pmatrix} \cos\theta_1 & & 0 \\ & \ddots & \\ 0 & & \cos\theta_k \end{pmatrix}$, and $\sin\Theta = \begin{pmatrix} \sin\theta_1 & & 0 \\ & \ddots & \\ 0 & & \sin\theta_k \end{pmatrix}$.

Theorem 2.1. *Let* $\operatorname{span} U$, $\operatorname{span} V \in Gr$. *There exists representations $U_{k\times n}$ and $V_{k\times n}$ and a rotation R of $\mathbb{R}^n$ that takes*

$$U \text{ to } I_{n\times k} = \begin{pmatrix} I_k \\ 0 \end{pmatrix} \text{ and } V \text{ to } \begin{pmatrix} \cos\Theta \\ \sin\Theta \\ 0 \end{pmatrix}$$

where $\{\theta_i\}_{i=1}^k$ *are the principal angles between* $\operatorname{span} U$ *and* $\operatorname{span} V$. *In other words there exists* $R \in O_n$ *with* $RU = I_{n\times k}$ *and* $RV = \begin{pmatrix} \cos\Theta \\ \sin\Theta \\ 0 \end{pmatrix}$.

Proof. Assume that $U = \begin{pmatrix} u_{\downarrow 1} & \cdots & u_{\downarrow k} \end{pmatrix}$ and $V = \begin{pmatrix} v_{\downarrow 1} & \cdots & v_{\downarrow k} \end{pmatrix}$ where $\langle u_{\downarrow i}, v_{\downarrow j}\rangle = \delta_{ij}\cos\theta_i$. R must be of the form $R = \begin{pmatrix} U^{\mathcal{T}} \\ U_\perp^{\mathcal{T}} \end{pmatrix}$ where the columns of $U_\perp$ are ON and span $(\operatorname{span} U)_\perp$. Now

$$RU = \begin{pmatrix} U^{\mathcal{T}} \\ U_\perp^{\mathcal{T}} \end{pmatrix} U = \begin{pmatrix} I_k \\ 0 \end{pmatrix} \text{ and } RV = \begin{pmatrix} U^{\mathcal{T}} \\ U_\perp^{\mathcal{T}} \end{pmatrix} V = \begin{pmatrix} U^{\mathcal{T}}V \\ U_\perp^{\mathcal{T}}V \end{pmatrix}.$$

First, $U^{\mathcal{T}}V = \begin{pmatrix} & \vdots & \\ \cdots & u_{\downarrow i}^{\mathcal{T}} v_{\downarrow j} & \cdots \\ & \vdots & \end{pmatrix} = \cos\Theta$.

Now the freedom in choosing $U_\perp = \begin{pmatrix} u_{\downarrow k+1} & \cdots & u_{\downarrow n} \end{pmatrix}$ may be exploited.

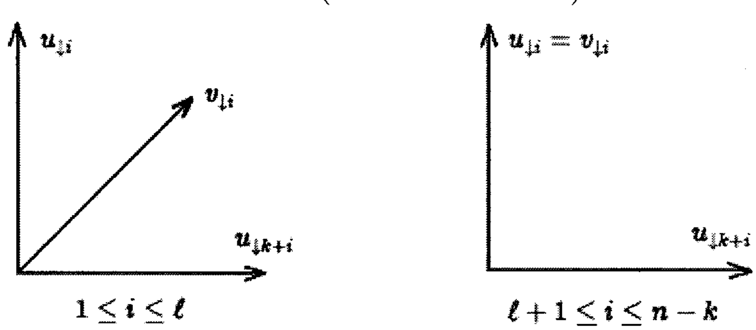

Figure 2: Choosing $u_{\downarrow k+i}$

For $i \leq \ell$ let
$$u_{\downarrow k+i} = \frac{v_{\downarrow i} - \Pi_U v_{\downarrow i}}{\|v_{\downarrow i} - \Pi_U v_{\downarrow i}\|} = \frac{v_{\downarrow i} - \sum_{\alpha=1}^{k} \langle v_{\downarrow i}, u_{\downarrow \alpha}\rangle\, u_{\downarrow \alpha}}{\|v_{\downarrow i} - \sum_{\alpha=1}^{k} \langle v_{\downarrow i}, u_{\downarrow \alpha}\rangle\, u_{\downarrow \alpha}\|}$$
$$= \frac{v_{\downarrow i} - \cos\theta_i\, u_{\downarrow i}}{\|v_{\downarrow i} - \cos\theta_i\, u_{\downarrow i}\|} = \frac{v_{\downarrow i} - \cos\theta_i\, u_{\downarrow i}}{(1 - 2\cos^2\theta_i + \cos^2\theta_i)^{1/2}} = \frac{v_{\downarrow i} - \cos\theta_i\, u_{\downarrow i}}{\sin\theta_i}.$$

If $j < k+1$ and $1 \leq i \leq \ell$, then
$$\langle u_{\downarrow k+i}, v_{\downarrow j}\rangle = \frac{\delta_{ij}\,(1 - \cos^2\theta_i)}{\sin\theta_i} = \delta_{ij}\sin\theta_i.$$

If $1 \leq i \leq k$ and $1 \leq j \leq \ell$,
$$\langle u_{\downarrow i}, u_{\downarrow k+j}\rangle = \langle u_{\downarrow i}, \frac{1}{\sin\theta_j}\,(v_{\downarrow j} - \cos\theta_j\, u_{\downarrow j})\rangle = \frac{\delta_{ij}}{\sin\theta_j}\,(\cos\theta_j - \cos\theta_j) = 0,$$

and if $1 \leq i \leq \ell$ and $1 \leq j \leq \ell$,
$$\begin{aligned}\langle u_{\downarrow k+i}, u_{\downarrow k+j}\rangle &= \frac{1}{\sin\theta_i\sin\theta_j}\langle v_{\downarrow i} - \cos\theta_i\, u_{\downarrow i}, v_{\downarrow j} - \cos\theta_j\, u_{\downarrow j}\rangle \\ &= \frac{\delta_{ij}}{\sin\theta_i\sin\theta_j}\,(1 - 2\cos^2\theta_i + \cos^2\theta_i) \; = \delta_{ij}\end{aligned}$$

which shows that the set $\{u_{\downarrow 1}, \ldots, u_{\downarrow k+\ell}\}$ is ON. Now
$$\begin{aligned}&\operatorname{span}\{u_{\downarrow 1}, \ldots, u_{\downarrow k}, u_{\downarrow k+1}, \ldots, u_{\downarrow k+\ell}\} \\ &= \operatorname{span}\{u_{\downarrow 1}, \ldots, u_{\downarrow k}, v_{\downarrow 1}, \ldots, v_{\downarrow \ell}\} \\ &= \operatorname{span}\{u_{\downarrow 1}, \ldots, u_{\downarrow k}, v_{\downarrow 1}, \ldots, v_{\downarrow k}\} \quad (\text{since } u_{\downarrow \ell+1} = v_{\downarrow \ell+1}, \ldots, u_{\downarrow k} = v_{\downarrow k}).\end{aligned}$$

Extend the set $\{u_{\downarrow 1}, \ldots, u_{\downarrow k+\ell}\}$ arbitrarily to an ON basis $\{u_{\downarrow 1}, \ldots, u_{\downarrow n}\}$ of $\mathbb{R}^n$ with the property that $u_{\downarrow k+\ell+i} \in \operatorname{span} V_\perp$ for $1 \leq i \leq n-k-\ell$. Notice for future reference that if the roles of $\begin{pmatrix} U & U_\perp \end{pmatrix}$ and $\begin{pmatrix} V & V_\perp \end{pmatrix}$ are reversed here after $u_{\downarrow k+\ell+i}$ are chosen then
$$u_{\downarrow k+\ell+i} = v_{\downarrow k+\ell+i} \qquad (1 \leq i \leq n-k-\ell)$$
would be a valid choice for $v_{\downarrow k+\ell+i}$ since then $v_{\downarrow k+\ell+i} \in \operatorname{span} U_\perp$. This gives
$$\langle u_{\downarrow k+\ell+i}, v_{\downarrow k+\ell+j}\rangle = \delta_{ij} \qquad (1 \leq i,j \leq n-k-\ell).$$

Now for $j \leq k$, $1 \leq i \leq n-k$, the identity $\langle u_{k+i}, v_j\rangle = \delta_{ij}\sin\theta_i$ still holds so that

$$U_\perp^T V = \begin{pmatrix} & \vdots & \\ \cdots & u_{\downarrow k+i}^T v_{\downarrow j} & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} \sin\Theta \\ 0 \end{pmatrix}.$$

Therefore $\begin{pmatrix} U^T V \\ U_\perp^T V \end{pmatrix} = \begin{pmatrix} \cos\Theta \\ \sin\Theta \\ 0 \end{pmatrix}$, and $R = \begin{pmatrix} U^T \\ U_\perp^T \end{pmatrix}$ is the required orthogonal matrix.

□

After a basis for $\operatorname{span} U_\perp$ is chosen, in the same way as the basis of $\operatorname{span} U_\perp$ was found above, except with the first ℓ basis vectors multiplied by -1 and the last $n-k-\ell$ set to $u_{\downarrow k+\ell+i}$, a basis $\{v_{\downarrow k+1}, \ldots, v_{\downarrow n}\}$ of $\operatorname{span} V_\perp$ can be found to satisfy $V_\perp^T U = \begin{pmatrix} -\sin\Theta \\ 0 \end{pmatrix}$. It will be convenient later that $V_\perp$ and U have this relationship.

Definition 2.4. Bases $\{u_{\downarrow 1}, \ldots, u_{\downarrow n}\}$ and $\{v_{\downarrow 1}, \ldots, v_{\downarrow n}\}$ of $\mathbb{R}^n$ satisfying

$$\begin{pmatrix} U^T \\ U_\perp^T \end{pmatrix} V = \begin{pmatrix} \cos\Theta \\ \sin\Theta \\ 0 \end{pmatrix}, \; V_\perp^T U = \begin{pmatrix} -\sin\Theta \\ 0 \end{pmatrix},$$

$$\text{and } u_{\downarrow k+\ell+i} = v_{\downarrow k+\ell+i} \text{ for } 1 \le i \le n-k-\ell$$

will be called **angle direction** (abbreviated **AD**) bases.

In this case $U^T V_\perp = (V_\perp^T U)^T = \begin{pmatrix} -\sin\Theta & 0 \end{pmatrix}$. Since $\begin{pmatrix} U & U_\perp \end{pmatrix}^T \in \mathbf{O}_n$ and $\begin{pmatrix} V & V_\perp \end{pmatrix} \in \mathbf{O}_n$, $\begin{pmatrix} U & U_\perp \end{pmatrix}^T \begin{pmatrix} V & V_\perp \end{pmatrix} \in \mathbf{O}_n$ so AD bases of this form must satisfy

$$\begin{pmatrix} U & U_\perp \end{pmatrix}^T \begin{pmatrix} V & V_\perp \end{pmatrix} = \begin{pmatrix} U^T V & U^T V_\perp \\ U_\perp^T V & U_\perp^T V_\perp \end{pmatrix} = \begin{pmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix}.$$

The above theorem is closely related to the singular value decomposition. Let $A_{k\times k} = U^T V$, where U and V are the same matrices as in the proof. In using the singular value decomposition to decompose A as $A = Q_1 \Sigma Q_2^T$ where $Q_1, Q_2 \in \mathbf{O}_k$ the first step is to look for the eigenvalues of $A^T A$.

$$A^T A = V^T U U^T V = \begin{pmatrix} \langle v_{\downarrow 1}, u_{\downarrow 1}\rangle & \cdots & \langle v_{\downarrow 1}, u_{\downarrow k}\rangle \\ \vdots & & \vdots \\ \langle v_{\downarrow k}, u_{\downarrow 1}\rangle & \cdots & \langle v_{\downarrow k}, u_{\downarrow k}\rangle \end{pmatrix} \begin{pmatrix} \langle u_{\downarrow 1}, v_{\downarrow 1}\rangle & \cdots & \langle u_{\downarrow 1}, v_{\downarrow k}\rangle \\ \vdots & & \vdots \\ \langle u_{\downarrow k}, v_{\downarrow 1}\rangle & \cdots & \langle u_{\downarrow k}, v_{\downarrow k}\rangle \end{pmatrix}$$

$$= \begin{pmatrix} \langle v_{\downarrow 1}, u_{\downarrow 1}\rangle^2 & & 0 \\ & \ddots & \\ 0 & & \langle v_{\downarrow k}, u_{\downarrow k}\rangle^2 \end{pmatrix} = \begin{pmatrix} \cos^2\theta_1 & & 0 \\ & \ddots & \\ 0 & & \cos^2\theta_k \end{pmatrix}$$

which has eigenvalues

$$0 = \cos^2\theta_1 = \lambda_1 = \cdots = \lambda_r < \lambda_{r+1} \le \cdots \le \lambda_\ell < \lambda_{\ell+1} = 1 = \cdots = \lambda_k = \cos^2\theta_k,$$

(which are real in the complex case), and singular values $\sigma_i = \sqrt{\lambda_i}$. This shows that $\Sigma = \cos\Theta$. The next step is to find an ON basis $\{w_1, \ldots, w_k\}$ of $\mathbb{R}^k$ consisting of eigenvectors of $A^T A$ and to set $Q_2 = \begin{pmatrix} w_1 & \cdots & w_k \end{pmatrix}$. Then the set $\{\frac{1}{\sigma_i} A w_i : r+1 \le i \le k\}$ is ON and once this set is extended to an ON basis $\{s_1, \ldots, s_r, s_{r+1}, \ldots, s_k\} = \{s_1, \ldots, s_r, \frac{1}{\sigma_{r+1}} A w_{r+1}, \ldots, \frac{1}{\sigma_k} A w_k\}$ of $\mathbb{R}^k$, then Q_1 is set to $Q_1 = \begin{pmatrix} s_1 & \cdots & s_k \end{pmatrix}$. In this case, because the bases of span U and span V are AD bases, a valid choice for Q_2 is $Q_2 = I_k$. This is illustrated by the equation $(\cos^2\Theta - \cos^2\theta_i I_k) w_i = 0$. When $r+1 \le i \le k$,

$$s_i = \frac{1}{\sigma_i} A w_i = \frac{1}{\sigma_i} a_{\downarrow i} = \frac{1}{\sigma_i} \begin{pmatrix} \langle u_{\downarrow i}, v_{\downarrow 1}\rangle \\ \vdots \\ \langle u_{\downarrow i}, v_{\downarrow k}\rangle \end{pmatrix} = \mathbf{e}_i$$

so that a valid choice for Q_1 is $Q_1 = I_k$. Now $U^T V = A = Q_1^T A Q_2 = \Sigma = \cos\Theta$. Next, the singular value decomposition is repeated for $B_{n-k\times k} := U_\perp^T V$ to get $B = Q_1^* \begin{pmatrix} \Sigma^* \\ 0 \end{pmatrix} Q_2^{*T}$ where $Q_1^* \in \mathbf{O}_{n-k}$ and $Q_2^* \in \mathbf{O}_k$.

$$B^{\mathcal{T}}B = (U_\perp^{\mathcal{T}}V)^{\mathcal{T}}(U_\perp^{\mathcal{T}}V) = V^{\mathcal{T}}U_\perp U_\perp^{\mathcal{T}}V = V^{\mathcal{T}}\Pi_{U_\perp}V$$

$$= V^{\mathcal{T}}\begin{pmatrix} \cdots & \Pi_{U_\perp} v_{\downarrow j} & \cdots \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & v_{\downarrow i}^{\mathcal{T}}\left(v_{\downarrow j} - \sum_{\alpha=1}^{k}\langle v_{\downarrow j}, u_{\downarrow\alpha}\rangle\, u_{\downarrow\alpha}\right) & \cdots \\ & \vdots & \end{pmatrix}$$

$$= \begin{pmatrix} & \vdots & \\ \cdots & \delta_{ij}\,(1-\cos^2\theta_i) & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} \sin^2\theta_1 & & 0 \\ & \ddots & \\ 0 & & \sin^2\theta_k \end{pmatrix}$$

which has eigenvalues

$$1 = \sin^2\theta_1 = \lambda_1^* = \cdots = \lambda_r^* > \lambda_{r+1}^* \geq \cdots \geq \lambda_\ell^* > \lambda_{\ell+1}^* = 0 = \cdots = \lambda_k^* = \sin^2\theta_k$$

and singular values $\sigma_i^* = \sqrt{\lambda_i^*}$. This shows that $\Sigma^* = \sin\Theta = (I_k - \cos^2\Theta)^{1/2} = (I_k - \Sigma^2)^{1/2}$. It is easy to show that again I_k is a valid choice for Q_2^* but even if the bases used to represent $\operatorname{span} U$ and $\operatorname{span} V$, say $\widetilde{U}$ and $\widetilde{V}$, are not AD bases, and $Q_2 \neq I_k$, it is still true that $Q_2^* = Q_2$ is a valid choice for Q_2^*. To see this assume $\widetilde{U} = UW_1$, $\widetilde{V} = VW_3$, $\widetilde{U}_\perp = U_\perp W_2$, $\widetilde{A} = \widetilde{U}^{\mathcal{T}}\widetilde{V}$, $\widetilde{B} = \widetilde{U}_\perp^{\mathcal{T}}\widetilde{V}$, and that w_i is an eigenvector of $\widetilde{A}^{\mathcal{T}}\widetilde{A}$ corresponding to the eigenvalue $\lambda_i \iff$

$$\begin{aligned}
0 &= W_3^{\mathcal{T}}(I_k - \Sigma^2 - I_k + \sigma_i^2 I_k)W_3\,w_i \\
= (\lambda_i I_k - \widetilde{A}^{\mathcal{T}}\widetilde{A})\,w_i &= W_3^{\mathcal{T}}((I_k - \Sigma^2) - (1-\sigma_i^2)I_k)W_3\,w_i \\
= (\lambda_i I_k - \widetilde{V}^{\mathcal{T}}\widetilde{U}\widetilde{U}^{\mathcal{T}}\widetilde{V})\,w_i &= W_3^{\mathcal{T}}(\Sigma^{*2} - \lambda_i^* I_k)W_3\,w_i \\
= (\lambda_i I_k - W_3^{\mathcal{T}}V^{\mathcal{T}}UW_1W_1^{\mathcal{T}}VW_3)\,w_i &= W_3^{\mathcal{T}}(V^{\mathcal{T}}U_\perp U_\perp^{\mathcal{T}}V - \lambda_i^* I_k)W_3\,w_i \\
= W_3^{\mathcal{T}}(\lambda_i I_k - V^{\mathcal{T}}UU^{\mathcal{T}}V)W_3\,w_i &= (\widetilde{V}^{\mathcal{T}}\widetilde{U}_\perp\widetilde{U}_\perp^{\mathcal{T}}\widetilde{V} - \lambda_i^* I_k)\,w_i \\
= W_3^{\mathcal{T}}(\sigma_i^2 I_k - \Sigma^2)W_3\,w_i &= (\widetilde{B}^{\mathcal{T}}\widetilde{B} - \lambda_i^* I_k)\,w_i
\end{aligned}$$

$\iff$ w_i is an eigenvector of $\widetilde{B}^{\mathcal{T}}\widetilde{B}$ corresponding to λ_i^*. In any case the set $\{\frac{1}{\sigma_i^*}Bw_i : 1 \leq i \leq \ell\}$ is ON and once this set is extended to an ON basis $\{s_1^*, \ldots, s_\ell^*, s_{\ell+1}^*, \ldots, s_{n-k}^*\} = \{\frac{1}{\sigma_1^*}Bw_1, \ldots, \frac{1}{\sigma_\ell^*}Bw_\ell, s_{\ell+1}^*, \ldots, s_{n-k}^*\}$ of $\mathbb{R}^{n-k}$, Q_1^* can be set to $Q_1^* = \begin{pmatrix} s_1^* & \cdots & s_{n-k}^* \end{pmatrix}$. The vectors completing the basis now appear on the right side of Q_1^* in contrast to Q_1

because now the eigenvalues λ_i^* are ordered from greatest to least whereas λ_i are ordered in the opposite direction. Of course in the present case $Q_1^* = I_{n-k}$ is a valid choice. It is easy to see that the singular values of $\tilde{U}^T\tilde{V}$ and $\tilde{U}_\perp^T\tilde{V}$ do not depend on the choice of ON representatives $\tilde{U}, \tilde{V}$, or $\tilde{U}_\perp$. This establishes that there exists $Q_1, Q_2 \in \mathbf{O}_k$ and $Q_1^* \in \mathbf{O}_{n-k}$ with

$$\begin{pmatrix} \tilde{U}^T\tilde{V} \\ \tilde{U}_\perp^T\tilde{V} \end{pmatrix} = \underbrace{\begin{pmatrix} Q_1 & 0 \\ 0 & Q_1^* \end{pmatrix}}_{\in \mathbf{K}} \begin{pmatrix} \cos\Theta \\ \sin\Theta \\ 0 \end{pmatrix} Q_2^T.$$

These observations indicate a computationally practical way of producing the principal angles and even AD bases with respect to span U and span V given arbitrary ON representatives $\tilde{U}$ and $\tilde{V}$, where U and V are AD representatives to be determined.

Algorithm 2.1

Step 1:

The first step is to use the singular value decomposition to get

$$\begin{pmatrix} \tilde{U}^T\tilde{V} \\ \tilde{U}_\perp^T\tilde{V} \end{pmatrix} = \begin{pmatrix} Q_1 \cos\Theta\, Q_2^T \\ Q_1^* \begin{pmatrix} \sin\Theta \\ 0 \end{pmatrix} Q_2^T \end{pmatrix},$$

which immediately yields $\theta_i = \cos^{-1}\sqrt{\lambda_i}$, where λ_i are the eigenvalues of $\left(\tilde{U}^T\tilde{V} \right)^T \left(\tilde{U}^T\tilde{V} \right)$.

Step 2:

Now, since $Q_1^T\tilde{U}^T\tilde{V}Q_2 = \cos\Theta$, $\tilde{U}Q_1 \in [U]$, and $\tilde{V}Q_2 \in [V]$, letting $U = \tilde{U}Q_1$ and $V = \tilde{V}Q_2$ gives $U^TV = \cos\Theta$. Similarly, since $\tilde{U}_\perp Q_1^* \in [U_\perp]$, letting $U_\perp = \tilde{U}_\perp Q_1^*$ gives $U_\perp^T V = Q_1^{*T}\tilde{U}_\perp^T\tilde{V}Q_2 = \begin{pmatrix} \sin\Theta \\ 0 \end{pmatrix}$. Already the AD basis $\{u_{\downarrow 1}, \ldots, u_{\downarrow n}\}$ and the partial AD basis $\{v_{\downarrow 1}, \ldots, v_{\downarrow k}\}$ have been found. The next step shows that taking $V = \tilde{V}Q_2$ is in fact unnecessary.

Step 3:

To generate $\begin{pmatrix} V & V_\perp \end{pmatrix}$ simply multiply the matrix

$$\begin{pmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix} = \begin{pmatrix} U & U_\perp \end{pmatrix}^{\mathcal{T}} \begin{pmatrix} V & V_\perp \end{pmatrix}$$

by $\begin{pmatrix} U & U_\perp \end{pmatrix}$.

Stop.

This method of generating $\{(u_i, v_i, \theta_i)\}_{i=1}^{n}$ will be called **method** 2. Even in the context of $\mathbf{Gr}(\mathbb{C})$ method 2 gives a way to produce AD bases and real θ_i.

2.3 Tangent Spaces of St and Gr

The representations of points in **St** and **Gr** using matrices in $\mathbb{R}^{n\times k}$ and $\mathbb{R}^{n\times k}$ are computationally and intuitively appealing but these so called **extrinsic coordinates** are not bona fide coordinates because the dimension of each space is less than the number of scalars used in the representation of a point. For this reason it is essential to identify which parts of the derivatives $\frac{d}{dt}P_{n\times k}(t) = \dot{P}(t)$ and $\frac{d}{dt}Q_{n\times n}(t) = \dot{Q}(t)$ are relevant tangent vectors in the usual sense. In **St** and **Gr** differentiating the condition $P^{\mathcal{T}}P = I_k$ leads to isolating $\frac{k(k+1)}{2}$ independent conditions on $\dot{P}$, leaving a $nk - \frac{k(k+1)}{2}(= \dim \mathbf{St})$ - dimensional **horizontal space**. By projecting an arbitrary $n\times k$ matrix onto this horizontal space a general form for $\dot{P}$ can be found that highlights it's $nk - \frac{k(k+1)}{2}$ -dimensional nature. In the case of **Gr** however, since $\dim \mathbf{Gr} < nk - \frac{k(k+1)}{2}$, part of this tangent must be removed corresponding to the equivalence classes of points in $\mathbf{Gr} \simeq \mathbf{St}/\mathbf{O}_k$. The remaining relevant tangent will be called the **horizontal component** of $\dot{P}$ for **Gr**. The horizontal spaces $H_P\,\mathbf{St}$ and $H_P\,\mathbf{Gr}$ or $H_Q\,\mathbf{St}$ and $H_Q\,\mathbf{Gr}$ will have $\dim H_p\,\mathbf{St} = \dim \mathbf{St}$ and $\dim H_p\,\mathbf{Gr} = \dim \mathbf{Gr}$.

Let $P_{n\times k}(t)$ be a curve in St with $P = P(0)$ and let $T = \dot{P}(0)$. Differentiating the condition $P(t)^T P(t) = I_k$ and evaluating at $t = 0$ gives $T^T P + P^T T = 0$ so that $P^T T$ is skew-symmetric. This can be expressed by the $\dfrac{k(k+1)}{2}$ independent conditions

$$t_{\downarrow i}^T p_{\downarrow j} = -p_{\downarrow i}^T t_{\downarrow j} \text{ for } i > j, \text{ and } t_{\downarrow i}^T p_{\downarrow j} = 0 \text{ for } i = j \text{ on } T.$$

This suggests the horizontal space $H_P\,\mathrm{St}$ is an $nk - \dfrac{k(k+1)}{2}$ -dimensional vector space, as expected. When $k = n$ this gives that $\dim T_P\,\mathrm{O}_n = \dfrac{n(n-1)}{2}$. Using the representations $P_{n\times k} \in \mathrm{M} = \mathbf{St}$ or $\mathbf{Gr}$ gives a natural embedding of M into $\mathbb{R}^{n\times k} \backsimeq \mathbb{R}^{nk}$. Using the identification $T_P\,\mathbb{R}^{n\times k} \backsimeq \mathbb{R}^{n\times k}$, $T_P\,\mathrm{M}$ can be thought of as a subspace of $\mathbb{R}^{n\times k}$ where the origin is at the point P. The notation $T_P\,\mathrm{M}$ here is meant to denote the space of all tangents that occur as derivatives of curves $P(t)$ and is not to be confused with the horizontal space. If $U, V \in \mathbb{R}^{n\times k}$ then using the natural inner product $\langle U, V\rangle_{n\times k} := \operatorname{tr} U^T V = \sum_{j=1}^{k} u_{\downarrow j}^T v_{\downarrow j} = \sum_{\substack{1\le i\le n\\ 1\le j\le k}} u_{ij} v_{ij}$ corresponds to the usual inner product in $\mathbb{R}^{nk}$.

Definition 2.5. The **normal space** $\perp_P \mathrm{M} \subset \mathbb{R}^{n\times k}$ is defined as $\perp_P \mathrm{M} = (T_P\,\mathrm{M})_\perp$.

Proposition 2.3. *The normal spaces $\perp_P M$ for $n \times k$ representatives have the form $\perp_P M = \{N : N = PS \text{ where } S_{k\times k} \text{ is symmetric}\}$ so that $\dim \perp_P M = \dfrac{k(k+1)}{2}$.*

Proof. Let $T \in T_P\,\mathrm{M}$ be arbitrary and assume $N = PS$ where S is symmetric, then

$$\begin{aligned}
\langle N, T\rangle_{n\times k} &= \operatorname{tr}(PS)^T T \\
&= \operatorname{tr} S^T P^T T \\
&= \operatorname{tr} S P^T T && (S \text{ is symmetric}) \\
&= -\operatorname{tr} S T^T P && (P^T T \text{ is skew}-\text{symmetric}) \\
&= -\operatorname{tr} T^T P S && (T^T P \text{ and } S \text{ are both } k \times k) \\
&= -\operatorname{tr} T^T N \quad = -\langle N, T\rangle_{n\times k}.
\end{aligned}$$

Therefore $\langle N, T\rangle_{n\times k} = 0$ and $N \in \perp_P \mathrm{M}$. $\{PS : S \text{ is symmetric}\}$ is clearly a vector space of dimension $\dfrac{k(k+1)}{2}$ which completes the proof. ☐

For $n \times n$ representatives $\perp_Q \mathbf{Gr} = \perp_Q \mathbf{St} = \perp_Q \mathbf{O}_n$ is $\dfrac{n(n+1)}{2}$ -dimensional. This is because $\perp_Q \mathbf{O}_n$ is isomorphic to the $n \times n$ symmetric matrices.

Definition 2.6. If $X \in \mathbb{R}^{k\times k}$ define $\text{symm}\,(X) := \frac{1}{2}(X+X^{\mathcal{T}})$, and $\text{skew}\,(X) := \frac{1}{2}(X-X^{\mathcal{T}})$.

Any time $S_{k\times k}$ is symmetric and $W_{k\times k}$ is skew-symmetric,

$$\langle S, W\rangle_{k\times k} = \text{tr}\, S^{\mathcal{T}} W = -\text{tr}\, S W^{\mathcal{T}} = -\text{tr}\, W^{\mathcal{T}} S = -\langle S, W\rangle_{k\times k}$$

so that $\langle S, W\rangle_{k\times k} = 0$. For any $X_{k\times k}$,

$$\begin{gathered} \text{symm}\,(X) + \text{skew}\,(X) = \frac{1}{2}(X + X^{\mathcal{T}}) + \frac{1}{2}(X - X^{\mathcal{T}}) = X \\ \text{so} \quad \mathbb{R}^{k\times k} = \text{symm}_{k\times k} \oplus \text{skew}_{k\times k} \\ \text{where } \text{symm}_{k\times k} := \text{symm}\,(\mathbb{R}^{k\times k}) \text{ and } \text{skew}_{k\times k} := \text{skew}\,(\mathbb{R}^{k\times k}). \end{gathered}$$

If $P \in \mathbb{R}_{n\times k}$ then

$$\begin{gathered} \langle PS, PW\rangle_{n\times k} = \text{tr}\, S^{\mathcal{T}} P^{\mathcal{T}} P W = \text{tr}\, S^{\mathcal{T}} W = \langle S, W\rangle_{k\times k} = 0 \\ \text{and } P\,\text{symm}\,(X) + P\,\text{skew}\,(X) = PX \\ \text{which shows } P\,\mathbb{R}^{k\times k} = P\,\text{symm}_{k\times k} \oplus P\,\text{skew}_{k\times k}. \end{gathered}$$

The following formulas, which can be found in [8], for projecting matrices $X_{n\times k}$ onto $T_P\,\mathrm{M}$ and $\perp_P \mathrm{M}$ are very simple.

Proposition 2.4. *Let $X \in \mathbb{R}^{n\times k}$, then*

$$\begin{gathered} \Pi_{\perp_P}(X) =: \Pi_{\perp_P M}(X) = P\,\text{symm}\,(P^{\mathcal{T}} X) \\ \textit{and } \Pi_{T_P}(X) =: \Pi_{T_P M}(X) = P\,\text{skew}\,(P^{\mathcal{T}} X) + \Pi_{P_\perp} X. \end{gathered}$$

Proof. If

$$E_{\alpha\beta} = \begin{pmatrix} & \vdots & \\ \cdots & \delta_{i\alpha}\delta_{j\beta} & \cdots \\ & \vdots & \end{pmatrix} = \mathrm{e}_\alpha \mathrm{e}_\beta^{\mathcal{T}},$$

then $\{E_{\alpha\beta}+E_{\beta\alpha}\}_{\alpha,\beta\in\{1,\dots,k\}}$ forms a spanning set of the $k\times k$ symmetric matrices and $\{P(E_{\alpha\beta}+E_{\beta\alpha})\}_{\alpha,\beta\in\{1,\dots,k\}}$ forms a spanning set of $\perp_P \mathrm{St}$. It is easily seen that the distinct elements of these sets are orthogonal as in the example

$$\mathrm{tr}\left(\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}^{\mathcal{T}}\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}\right)=\mathrm{tr}\left(\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}^{\mathcal{T}}P^{\mathcal{T}}P\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}\right)=0.$$

Note that

$$\begin{aligned}||P(E_{\alpha\beta}+E_{\beta\alpha})||^2_{n\times k}&=\mathrm{tr}\,(E_{\alpha\beta}+E_{\beta\alpha})^{\mathcal{T}}P^{\mathcal{T}}P(E_{\alpha\beta}+E_{\beta\alpha})\\&=\mathrm{tr}\,(E_{\beta\alpha}+E_{\alpha\beta})(E_{\alpha\beta}+E_{\beta\alpha})\\&=\mathrm{tr}\,(E_{\alpha\beta}+E_{\beta\alpha})^2\\&=\mathrm{tr}\,(\mathbf{e}_\alpha\mathbf{e}_\beta^{\mathcal{T}}+\mathbf{e}_\beta\mathbf{e}_\alpha^{\mathcal{T}})^2\\&=\underbrace{\mathrm{tr}\,\mathbf{e}_\alpha\mathbf{e}_\beta^{\mathcal{T}}\mathbf{e}_\alpha\mathbf{e}_\beta^{\mathcal{T}}}_{=\begin{cases}1,\text{ if }\alpha=\beta\\0,\text{ if }\alpha\neq\beta\end{cases}}+\underbrace{\mathrm{tr}\,\mathbf{e}_\alpha\mathbf{e}_\beta^{\mathcal{T}}\mathbf{e}_\beta\mathbf{e}_\alpha^{\mathcal{T}}}_{=1}+\underbrace{\mathrm{tr}\,\mathbf{e}_\beta\mathbf{e}_\alpha^{\mathcal{T}}\mathbf{e}_\alpha\mathbf{e}_\beta^{\mathcal{T}}}_{=1}+\underbrace{\mathrm{tr}\,\mathbf{e}_\beta\mathbf{e}_\alpha^{\mathcal{T}}\mathbf{e}_\beta\mathbf{e}_\alpha^{\mathcal{T}}}_{=\begin{cases}1,\text{ if }\alpha=\beta\\0,\text{ if }\alpha\neq\beta\end{cases}}=\begin{cases}4,\text{ if }\alpha=\beta\\2,\text{ if }\alpha\neq\beta\end{cases}.\end{aligned}$$

Now $\Pi_{\perp_P}(X)=\dfrac{1}{2}\displaystyle\sum_{\alpha,\beta=1}^{k}\frac{\langle X,P(E_{\alpha\beta}+E_{\beta\alpha})\rangle_{n\times k}\,P(E_{\alpha\beta}+E_{\beta\alpha})}{||P(E_{\alpha\beta}+E_{\beta\alpha})||^2_{n\times k}}$

$$=\frac{P}{2}\sum_{\alpha,\beta=1}^{k}\frac{(\mathrm{tr}\,(X^{\mathcal{T}}PE_{\alpha\beta})+\mathrm{tr}\,(X^{\mathcal{T}}PE_{\beta\alpha}))\,(E_{\alpha\beta}+E_{\beta\alpha})}{||P(E_{\alpha\beta}+E_{\beta\alpha})||^2_{n\times k}}$$

$$=\frac{P}{2}\left(2\sum_{\alpha=1}^{k}\frac{4\,\mathrm{tr}\,(X^{\mathcal{T}}PE_{\alpha\alpha})\,E_{\alpha\alpha}}{4}+2\sum_{\alpha<\beta}\frac{(\mathrm{tr}\,(X^{\mathcal{T}}PE_{\alpha\beta})+\mathrm{tr}\,(X^{\mathcal{T}}PE_{\beta\alpha}))\,(E_{\alpha\beta}+E_{\beta\alpha})}{2}\right)$$

$$=\frac{P}{2}\left(2\sum_{\alpha=1}^{k}(x_{\downarrow\alpha}^{\mathcal{T}}p_{\downarrow\alpha})\,E_{\alpha\alpha}+\sum_{\alpha<\beta}(x_{\downarrow\beta}^{\mathcal{T}}p_{\downarrow\alpha}+x_{\downarrow\alpha}^{\mathcal{T}}p_{\downarrow\beta})(E_{\alpha\beta}+E_{\beta\alpha})\right)$$

$$=\frac{P}{2}\left(\begin{pmatrix}&\vdots&\\\cdots&\delta_{ij}\,(p_{\downarrow i}^{\mathcal{T}}x_{\downarrow j}+x_{\downarrow i}^{\mathcal{T}}p_{\downarrow j})&\cdots\\&\vdots&\end{pmatrix}+\begin{pmatrix}&\vdots&\\\cdots&(1-\delta_{ij})\,(p_{\downarrow i}^{\mathcal{T}}x_{\downarrow j}+x_{\downarrow i}^{\mathcal{T}}p_{\downarrow j})&\cdots\\&\vdots&\end{pmatrix}\right)$$

$$= \frac{P}{2} \begin{pmatrix} & \vdots & \\ \cdots & p_{\downarrow i}^{\mathcal{T}} x_{\downarrow j} + x_{\downarrow i}^{\mathcal{T}} p_{\downarrow j} & \cdots \\ & \vdots & \end{pmatrix}$$

$$= \frac{P}{2} (P^{\mathcal{T}} X + X^{\mathcal{T}} P) = P \operatorname{symm}(P^{\mathcal{T}} X).$$

Since $P \operatorname{skew}(P^{\mathcal{T}} X) + \Pi_{\perp_P}(X) + \Pi_{P_\perp} X = P(\operatorname{symm}(P^{\mathcal{T}} X) + \operatorname{skew}(P^{\mathcal{T}} X)) + \Pi_{P_\perp} X = PP^{\mathcal{T}} X + (I_n - PP^{\mathcal{T}})X = X$, it must be that $\Pi_{T_P}(X) = P \operatorname{skew}(P^{\mathcal{T}} X) + \Pi_{P_\perp} X$. □

2.4 Vertical and Horizontal Spaces of St

The last section gave a general form of tangent vectors in $T_P\,\mathbf{St}$ and $T_P\,\mathbf{Gr}$ in terms of $n \times k$ representatives:

$$T_P\,\mathrm{M} \ni T_{n\times k} = PA + P_\perp B = \begin{pmatrix} P & P_\perp \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}.$$

More information about the character of St can be gained by using the equivalence classes $[Q]$, where $Q = \begin{pmatrix} P & P_\perp \end{pmatrix}$, and tangent vectors $T_{n\times n} \in T_Q\,\mathbf{St}$. Using the general form

$$T_{n\times k} = \begin{pmatrix} P & P_\perp \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = Q \begin{pmatrix} A \\ B \end{pmatrix}$$

where A is skew-symmetric and B is arbitrary, the general form of a tangent $T_{n\times n}$ to the curve $Q(t) = \begin{pmatrix} P(t) & P_\perp(t) \end{pmatrix}$ is easily found.

$$\operatorname{skew}_{n\times n} \ni Q^{\mathcal{T}} T_{n\times n} = Q^{\mathcal{T}} \left(Q \begin{pmatrix} A \\ B \end{pmatrix} \quad \begin{pmatrix} T_{12} \\ T_{22} \end{pmatrix} \right) := Q^{\mathcal{T}} \left(Q \begin{pmatrix} A \\ B \end{pmatrix} \quad Q \begin{pmatrix} X_1 \\ X \end{pmatrix} \right)$$

$$= \begin{pmatrix} A & X_1 \\ B & X \end{pmatrix} \Longrightarrow \begin{cases} X_1 = -B^{\mathcal{T}}, \text{ and} \\ X \in \operatorname{skew}_{n-k\times n-k} \end{cases}$$

so that $T = Q \begin{pmatrix} A & -B^{\mathcal{T}} \\ B & X \end{pmatrix}$ is the desired general form.

Definition 2.7. The **horizontal space** $H_Q\,\mathrm{St}$ is defined to be the subspace of $T_Q\,\mathrm{St}$ that is invariant under choice of representation curve $W(t) \in [Q(t)]$. The **vertical space** $V_Q\,\mathrm{St} := T_Q\,\mathrm{St} \cap (H_Q\,\mathrm{St})_\perp$.

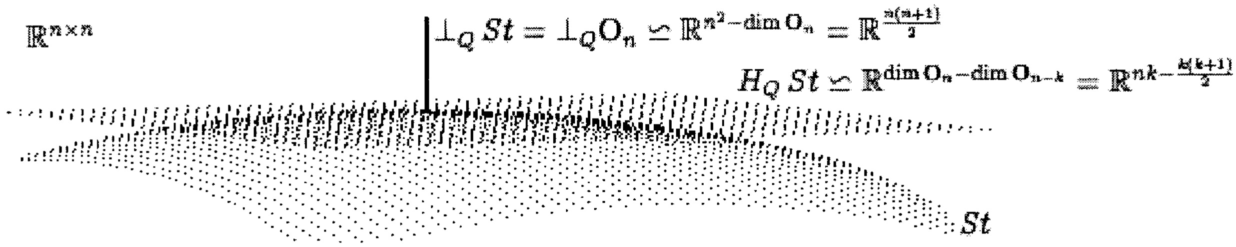

Figure 3: Horizontal and Normal Spaces of St

Theorem 2.2. *The horizontal space*

$$H_Q\,St = \left\{ Q \begin{pmatrix} A & -B^T \\ B & 0 \end{pmatrix} : A_{k\times k} \text{ skew} - \text{symmetric}, \; B_{n-k\times k} \text{ arbitrary} \right\},$$

and the vertical space $V_Q\,St = \left\{ Q \begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix} : X_{n-k\times n-k} \text{ skew} - \text{symmetric} \right\}.$

Proof. Let $W(t)$ be another representation of the curve $Q(t)$, say

$$W(t) = Q(t) \begin{pmatrix} I_k & 0 \\ 0 & M(t) \end{pmatrix}$$

where $M(t) \in \mathbf{O}_{n-k}$ for all t. Note that since $M \in \mathbf{O}_{n-k} = \mathbf{St}_{n-k,n-k}$, $\Pi_{M_\perp}\dot{M} = 0$ so $\dot{M}(0) = M(0)X_2$ for some $X_2 \in \mathrm{skew}_{n-k\times n-k}$. Now at $t = 0$

$$\begin{aligned}
\dot{W} &= \dot{Q} \begin{pmatrix} I_k & 0 \\ 0 & M \end{pmatrix} + Q \begin{pmatrix} 0 & 0 \\ 0 & \dot{M} \end{pmatrix} \\
&= Q \left(\begin{pmatrix} A & -B^T \\ B & X \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & M \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & MX_2M^T \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & M \end{pmatrix} \right) \\
&= Q \begin{pmatrix} A & -B^T \\ B & X + MX_2M^T \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & M \end{pmatrix} \in \left[Q \begin{pmatrix} A & -B^T \\ B & X + MX_2M^T \end{pmatrix} \right].
\end{aligned}$$

Depending on the choice of $M(t)$, MX_2M^T and hence $X + MX_2M^T$ may be any skew-symmetric $n-k \times n-k$ matrix but the blocks A, B, and $-B^T$ remain invariant which shows that the horizontal space is the set $\left\{Q\begin{pmatrix} A & -B^T \\ B & 0 \end{pmatrix}\right\}$. An alternative way to see this is to simply observe that the matrix X does not even appear in the general $n \times k$ form of a tangent in St. It is clear that the vertical space is the set $\left\{Q\begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix}\right\}$. □

When using $\dot{Q}$ in calculations the matrix X will be set to 0. Intuitively, movements in the vertical direction correspond to changes in representation while movements in the horizontal direction correspond to movements on the manifold. Counting the independent elements in A and B suggests that $\dim H_P \,\mathrm{St} = \dim H_Q \,\mathrm{St} = \frac{k(k-1)}{2} + k(n-k) = nk - \frac{k(k+1)}{2}$ which is the proposed dimension of St. Verifying that every tangent of this form occurs as the tangent to some curve in St confirms that $\dim \mathrm{St} = nk - \frac{k(k+1)}{2}$. To do this let $\mathcal{W}$ be arbitrary of the form $\mathcal{W} = \begin{pmatrix} A & -B^T \\ B & 0 \end{pmatrix} \in \mathrm{skew}_{n\times n}$. $\mathrm{e}^{t\mathcal{W}} \in \mathbf{O}_n$ because $(\mathrm{e}^{t\mathcal{W}})^T = \mathrm{e}^{t\mathcal{W}^T} = \mathrm{e}^{-t\mathcal{W}} = (\mathrm{e}^{t\mathcal{W}})^{-1}$. Therefore $Q\mathrm{e}^{t\mathcal{W}}$ is a curve in St passing through Q, and finally, $\frac{d}{dt}Q\mathrm{e}^{t\mathcal{W}}|_{t=0} = Q\mathcal{W}$.

2.5 Vertical and Horizontal Spaces of Gr

Since the tangent space $T_P \,\mathrm{Gr} = T_P \,\mathrm{St}$ the general form of an $n \times n$ tangent vector on **Gr** is still $T = Q\begin{pmatrix} A & -B^T \\ B & X \end{pmatrix}$. The way in which the horizontal and vertical spaces differ between St and Gr is easy to predict.

Theorem 2.3. *The horizontal space*

$$H_Q \, Gr = \left\{Q\begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} : B_{n-k\times k} \text{ arbitrary}\right\},$$

and the vertical space $V_Q\ Gr = \left\{ Q \begin{pmatrix} A & 0 \\ 0 & X \end{pmatrix} : A_{k\times k}, X_{n-k\times n-k} \text{ skew} - \text{symmetric} \right\}.$

Proof. Let $W(t)$ be another representation of the curve $Q(t)$, say

$$W(t) = Q(t) \begin{pmatrix} M_1(t) & 0 \\ 0 & M_2(t) \end{pmatrix}$$

where $M_1(t) \in \mathbf{O}_k$ and $M_2(t) \in \mathbf{O}_{n-k}$ for all t. Note that, as in the the proof of Theorem 2.2, $\dot{M}_1(0) = M(0)X_1$ for some $X_1 \in \text{skew}_{k\times k}$ and $\dot{M}_2(0) = M(0)X_2$ for some $X_2 \in \text{skew}_{n-k\times n-k}$. Differentiating and evaluating at $t = 0$ gives

$$\begin{aligned} \dot{W} &= \dot{Q} \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} + Q \begin{pmatrix} \dot{M}_1 & 0 \\ 0 & \dot{M}_2 \end{pmatrix} \\ &= Q \left(\begin{pmatrix} A & -B^T \\ B & X \end{pmatrix} \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} + \begin{pmatrix} M_1X_1M_1^T & 0 \\ 0 & M_2X_2M_2^T \end{pmatrix} \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} \right) \\ &\in \left[Q \begin{pmatrix} A + M_1X_1M_1^T & -B^T \\ B & X + M_2X_2M_2^T \end{pmatrix} \right]. \end{aligned}$$

Therefore only the blocks B and $-B^T$ remain invariant which shows that the horizontal space is the set $\left\{ Q \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \right\}$. It is then clear that the vertical space is the set $\left\{ Q \begin{pmatrix} A & 0 \\ 0 & X \end{pmatrix} \right\}$. □

This theorem shows that when dealing with $n \times k$ tangents on **Gr**,

$$H_P\ \mathbf{Gr} = \left\{ Q \begin{pmatrix} 0 \\ B \end{pmatrix} : B_{n-k\times k} \text{ arbitrary} \right\} \text{ and}$$

$$V_P\ \mathbf{Gr} = \left\{ Q \begin{pmatrix} A \\ 0 \end{pmatrix} : A_{k\times k} \text{ skew} - \text{symmetric} \right\}.$$

The matrix A in the representation $T = PA + P_\perp B$ corresponds to changes in representation and should be set to zero while the matrix B corresponds to movements on **Gr**.

The same argument as the one used in the Stiefel case shows that all such tangents occur. Counting the elements in B shows that $\dim H_P\,\mathbf{Gr} = \dim H_Q\,\mathbf{Gr} = k(n-k)$, therefore the dimension of $\mathbf{Gr}$ is $k(n-k)$.

In general a horizontal tangent T to the point $P_{n\times k} \in \mathrm{M}$ ($= \mathbf{O}_n$, $\mathbf{St}$, or $\mathbf{Gr}$) has the form

$$T = \begin{cases} PA & , \quad \mathrm{M} = \mathbf{O}_n \\ PA + P_\perp B & , \quad \mathrm{M} = \mathbf{St} \\ P_\perp B & , \quad \mathrm{M} = \mathbf{Gr}. \end{cases}$$

2.6 Canonical Metrics

Definition 2.8. The **canonical metrics** on **St** and **Gr** denoted $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbf{St}}$ and $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbf{Gr}}$ are defined at the point $P_{n\times k}$ (or $Q_{n\times n}$) as

$$\langle T_1, T_2\rangle_{\mathbf{St}} = \frac{1}{2}\langle A_1, A_2\rangle_{k\times k} + \langle B_1, B_2\rangle_{n-k\times k} \quad (T_1, T_2 \in H_P\,\mathbf{St}\ (H_Q\,\mathbf{St})) \text{ and}$$
$$\langle T_1, T_2\rangle_{\mathbf{Gr}} = \langle B_1, B_2\rangle_{n-k\times k} \quad (T_1, T_2 \in H_P\,\mathbf{Gr}\ (H_Q\,\mathbf{Gr})).$$

In particular $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbf{O}_n} = \frac{1}{2}\mathrm{tr}\,A_1^T A_2$. These Riemannian metrics correspond to the usual $\mathbb{R}^{\dim \mathbf{St}}$ and $\mathbb{R}^{\dim \mathbf{Gr}}$ inner products applied to the independent elements of tangents in $H_P\,\mathbf{St}$ and $H_P\,\mathbf{Gr}$. Some useful identities for $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbf{Gr}}$ are:

i)
$$\langle T_{1\,n\times k}, T_{2\,n\times k}\rangle_{\mathbf{Gr}} = \mathrm{tr}\,B_1^T B_2 = \mathrm{tr}\,B_1^T P^T P B_2 = \mathrm{tr}\,T_1^T T_2$$
$$= \langle T_1, T_2\rangle_{n\times k} \qquad \text{so } \langle\,\cdot\,,\,\cdot\,\rangle_{\mathbf{Gr}} = \langle\,\cdot\,,\,\cdot\,\rangle_{n\times k}.$$

ii)
$$\langle T_{1\,n\times n}, T_{2\,n\times n}\rangle_{\mathbf{Gr}} = \mathrm{tr}\,B_1^T B_2 = \frac{1}{2}\mathrm{tr}\left(\begin{pmatrix} 0 & B^T \\ -B & 0 \end{pmatrix}\begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix}\right)$$
$$= \frac{1}{2}\mathrm{tr}\left(\begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix}^T Q^T Q \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix}\right) = \frac{1}{2}\langle T_1, T_2\rangle_{n\times n}$$
so $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbf{Gr}} = \frac{1}{2}\langle\,\cdot\,,\,\cdot\,\rangle_{n\times n}$ when applied to horizontal vectors.

It is an important observation that the metric for $\mathbf{O}_n$ is the same as the metric for $\mathbf{Gr}$ when applied to vectors (or conjugates of vectors) in the horizontal space of $\mathbf{Gr}$. When the metrics are understood to be equivalent the notation $\langle \cdot, \cdot \rangle$ will replace $\langle \cdot, \cdot \rangle_{\mathbf{Gr}}$ and $\langle \cdot, \cdot \rangle_{\mathbf{O}_n}$.

The canonical metric $\langle T_1, T_2 \rangle_{\mathbf{Gr}(\mathbb{C})} = \frac{1}{2}\mathrm{tr}(T_2^{\mathcal{H}} T_1)$ where the matrix T_2 is conjugated so that $\langle \cdot, \cdot \rangle_{\mathbf{Gr}(\mathbb{C})}$ is conjugate-linear in it's second argument ([12] takes this convention).

2.7 Geodesics in $\mathbf{O}_n$, St, and Gr and Geodesic Distance in Gr

Let $\mathrm{M} = \mathbf{O}_n$, **St**, or **Gr**. Assume $C_{n\times n}$ is a geodesic in M. Differentiating the condition $C^T C = I_n$ twice gives

$$\ddot{C}^T C + 2\dot{C}^T \dot{C} + C^T \ddot{C} = 0.$$

When a Riemannian Manifold is submersed in Euclidean space the condition that the acceleration vector $\ddot{C} \in \perp_C \mathrm{M}$ characterizes geodesics (see [6] pg. 68). $\ddot{C}$ must therefore have the form

$$\ddot{C}(t) = C(t)S(t) \quad \text{where } S(t) \in \mathrm{symm}_{n\times n}$$

Substituting this into the above equation,

$$\begin{aligned} S + \dot{C}^T \dot{C} &= 0 \\ CS + C(\dot{C}^T \dot{C}) &= 0 \\ \ddot{C} + C(\dot{C}^T \dot{C}) &= 0. \end{aligned}$$

This is the geodesic equation analogous to the equation

$$\sum_k \left(\ddot{x}_k + \sum_{i,j} \Gamma_{ij}^k \, \dot{x}_i \, \dot{x}_j \right) \frac{\partial}{\partial x_k} = 0$$

in general Riemannian Manifolds where $\{x_i\}_i$ is a usual coordinate system and $\{\Gamma_{ij}^k\}_{i,j,k}$ are the Christoffel symbols (see [5] pg. 62). [8] defines a Christoffel function $\Gamma(A, A) = CA^T A$.

Theorem 2.4. *In $M = O_n$, St, and Gr the curve*

$$C(t) = C_0 e^{t\mathcal{B}_0} \qquad (\mathcal{B}_0 \in H_I\, M)$$

(modulo the appropriate isotropy group) is a geodesic emanating from C_0 in the direction $C_0\mathcal{B}_0$ with constant speed $||\mathcal{B}_0||_M$.

Proof. Substituting C into the left side of the geodesic equation,

$$\ddot{C} + C(\dot{C}^T\dot{C}) = C_0 e^{t\mathcal{B}_0}\mathcal{B}_0\mathcal{B}_0 + C(\mathcal{B}_0^T C^T C\mathcal{B}_0) = C\mathcal{B}_0\mathcal{B}_0 + C\mathcal{B}_0^T\mathcal{B}_0 = -C\mathcal{B}_0^T\mathcal{B}_0 + C\mathcal{B}_0^T\mathcal{B}_0 = 0.$$

Therefore C is a geodesic. The initial direction $\dot{C}(0) = C_0\mathcal{B}_0$. The speed of C is easily seen to be $||\mathcal{B}_0||_{\mathrm{M}}$. □

When $\mathcal{B}_0 \in H_I\,\mathrm{M}$, $\dot{C} = C\mathcal{B}_0 \in H_C\,\mathbf{M}$. In other words, in each manifold the curve $C_0 e^{t\mathcal{B}_0}$ has a tangent vector that belongs to $H_C\,\mathrm{M}$ for all t.

The orthogonal group geodesics right multiplied by the isotropy group for **Gr** are geodesics in **Gr**. This is in agreement with the general theory of homogeneous spaces (see [6] pg. 68). Suppose $C(t)$ is a geodesic in **Gr** with $C(0) = I_n$, since any representative of $C(t)$ may be used it may be assumed that the vertical components of $\dot{C}(0)$ are 0 so that

$$\dot{C}(0) = \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix}.$$

Consider a geodesic $C_1(t)$ in **Gr** with

$$C_1(0) = Q_U = \begin{pmatrix} U & U_\perp \end{pmatrix} \text{ and } C_1(t_1) = Q_V = \begin{pmatrix} V & V_\perp \end{pmatrix}$$

where $\{u_{\downarrow i}\}_{i=1}^n$ and $\{v_{\downarrow i}\}_{i=1}^n$ are AD bases and let $\{\theta_i\}_{i=1}^k$ be the principal angles between U and V. It is easy to rotate the geodesic $C_1(t)$ to a geodesic $C(t)$ with end points

$$C(0) = I_n \text{ and } C(t_1) \in \left[\begin{pmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix}\right] = [Q_U^T Q_V]$$

and vise-versa, explicitly, $C(t) = Q_U^T C_1(t)$ (see Figure 4).

When dealing with $C(t)$ the most natural choice for $\dot{C}(0)$ is the tangent

$$\dot{C}(0) = \dot{C}_{I_n} = \frac{1}{||\Theta||_{k\times k}} \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} := \Psi \text{ giving } C(t) = e^{t\Psi}.$$

Consider the geodesic $C_2(t) = \exp t \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ emanating from I_n in the direction of Ψ. The speed of $C_2(t)$ is constant since

$$\langle \dot{C}_2(t), \dot{C}_2(t)\rangle_{\mathbf{Gr}}^{1/2} = \left(\frac{1}{2}\text{tr} \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^{\mathcal{T}} \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right)^{1/2}$$

$$= \left(\frac{1}{2}\text{tr} \left(\begin{pmatrix} \Theta^2 & 0 & 0 \\ 0 & \Theta^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right) \right)^{1/2} = (\text{tr}\, \Theta^2)^{1/2} = ||\Theta||_{k\times k}.$$

Because $||\Theta||_{k\times k}$ will appear often it will be denoted $||\Theta||$. The arc length $s(t)$ along $C_2(t)$ is

$$s = \int_0^t ||\Theta|| d\tau = t\,||\Theta||$$

so $t = \frac{s}{||\Theta||}$. Re-parametrize $C_2(t)$ with the change of variable $t \longrightarrow \frac{t}{||\Theta||}$ then $C_2(t)$ becomes

$$C_2(t) = \exp \frac{t}{||\Theta||} \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = C(t)$$

so $C(t)$ is already parametrized according to arc length. The following proposition, together with the fact that the arc length along $C_1(t)$ for a given t is the same as that along $C(t)$ (because $\operatorname{tr}(\Psi^T C^T Q_U^T Q_U C \Psi) = \operatorname{tr}(\Psi^T \Psi)$), establishes the famous formula (see [3]) for the geodesic distance $d(U, V)$ between U and V;

$$d(U,V) = ||\Theta|| = \sqrt{\sum_i \theta_i^2}.$$

Engineering papers often use what is called the **distortion** or **chordal distance** defined by $d_c(U,V) = \sqrt{\sum_i \sin^2 \theta_i}$. For small θ_i the chordal distance converges to the usual distance.

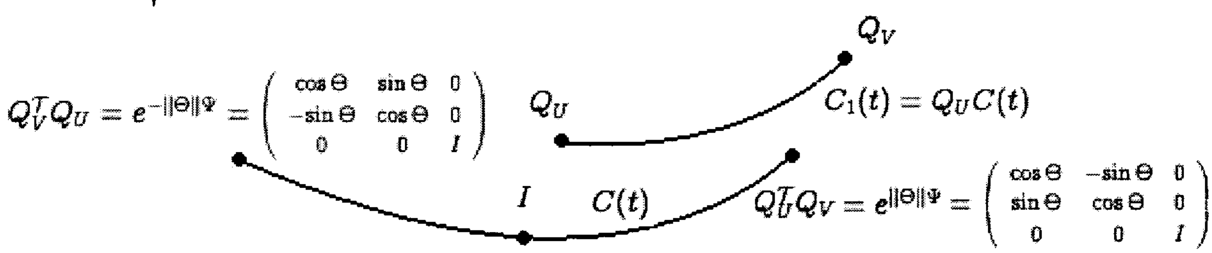

Figure 4: Translating Geodesics

Proposition 2.5. *If $\dot{C}(0)$ is chosen to be* $\frac{1}{||\Theta||}\begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ *then*

$$C(||\Theta||) = \begin{pmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix} = Q_U^T Q_V.$$

Proof. $C(||\Theta||) = \exp\begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = I_n + \sum_{j=1}^{\infty} \frac{(||\Theta||\dot{C}(0))^j}{j!}$

$$= I_n + \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} -\Theta^2 & 0 & 0 \\ 0 & -\Theta^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & \Theta^3 & 0 \\ -\Theta^3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{4!}\begin{pmatrix} \Theta^4 & 0 & 0 \\ 0 & \Theta^4 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \cdots$$

$$= \begin{pmatrix} I_k + \sum_{j=1}^{\infty} \frac{(-1)^j}{(2j)!} \Theta^{2j} & -\sum_{j=0}^{\infty} \frac{(-1)^{j+1}}{(2j+1)!} \Theta^{2j+1} & 0 \\ \sum_{j=0}^{\infty} \frac{(-1)^{j+1}}{(2j+1)!} \Theta^{2j+1} & I_k + \sum_{j=1}^{\infty} \frac{(-1)^j}{(2j)!} \Theta^{2j} & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix}$$

$$= \begin{pmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix} = Q_U^T Q_V. \qquad \square$$

Using the same calculation,

$$C(t) = \exp \frac{t}{||\Theta||} \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \cos \frac{t}{||\Theta||}\Theta & -\sin \frac{t}{||\Theta||}\Theta & 0 \\ \sin \frac{t}{||\Theta||}\Theta & \cos \frac{t}{||\Theta||}\Theta & 0 \\ 0 & 0 & I_{n-2k} \end{pmatrix}.$$

A consequence of this is that any two points Q_U and Q_V may be joined by a geodesic C having total length $\leq \left(\sum_{i=1}^{k} \left(\frac{\pi}{2}\right)^2 \right)^{1/2} = \sqrt{k}\,\frac{\pi}{2}$. By retracing some steps it can be seen that everything in this section applies as stated to the complex case.

2.8 The Cut Locus on Gr

Definition 2.9. For a Riemannian Manifold M, the **cut locus** of a point $p \in \mathrm{M}$ is defined to be the set of points

$$\mathrm{Cut}_p = \left\{ C(t_C) : C \text{ a geodesic with } ||\dot{C}|| \equiv 1, C(0) = p, t_C = \sup\{t : d_{\mathrm{M}}(C(0), C(t)) = t\} \right\}$$

(see [5] pg. 266).

The following theorem can be found without proof in [3].

Theorem 2.5. *In Gr the cut locus at I is the set*

$$\mathrm{Cut}_I = \{P : \text{ The matrix } \Theta \text{ corresponding to } I \text{ and } P \text{ has at least one } \theta_i = \pi/2\}.$$

Proof. Let $P_{n\times k}$ be a point in **Gr** and let $\{\theta_i\}_{i=1}^k$ be the principal angles between I and P. A geodesic joining I to P is

$$C(t) = \begin{pmatrix} \cos \frac{t}{||\Theta||}\Theta \\ \sin \frac{t}{||\Theta||}\Theta \\ 0 \end{pmatrix}.$$

Without loss of generality assume that only $\theta_1 = \pi/2$ and let ε be such that $\frac{||\Theta||+\varepsilon}{||\Theta||}\theta_i < \frac{\pi}{2}$ for $i > 1$.

$$\text{Let } P_1 := \begin{pmatrix} \cos \frac{||\Theta||+\varepsilon}{||\Theta||}\Theta \\ \sin \frac{||\Theta||+\varepsilon}{||\Theta||}\Theta \\ 0 \end{pmatrix}.$$

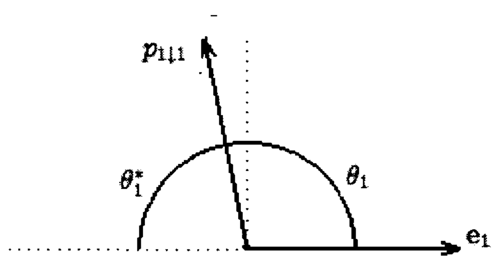

Figure 5: A Smaller Angle Between $\mathbf{e}_1$ and $p_{1\downarrow 1}$

The strategy of the proof will be to produce a geodesic $\gamma(t)$ from I_n to P_1 having length shorter than $||\Theta|| + \varepsilon$. Define $\tilde{\Theta} = \left(\frac{||\Theta||+\varepsilon}{||\Theta||}\right)\Theta$ and let

$$\theta_1^* = \pi - \left(\frac{||\Theta||+\varepsilon}{||\Theta||}\right)\theta_1 = \pi - \left(\frac{||\Theta||+\varepsilon}{||\Theta||}\right)\frac{\pi}{2}.$$

Now define Θ^* to be $\tilde{\Theta}$ but with $\left(\frac{||\Theta||+\varepsilon}{||\Theta||}\right)\theta_1$ replaced with θ_1^*.

$$\cos\theta_1^* = \cos\left(\frac{||\Theta||+\varepsilon}{||\Theta||}\right)\theta_1 \text{ and } \sin\theta_1^* = -\sin\left(\frac{||\Theta||+\varepsilon}{||\Theta||}\right)\theta_1$$

$$\text{so } \gamma(t) := \begin{pmatrix} \mathrm{e}_1 & \cdots & \mathrm{e}_k & -\mathrm{e}_{k+1} & \mathrm{e}_{k+2} & \cdots & \mathrm{e}_n \end{pmatrix}^T \begin{pmatrix} \cos \dfrac{t}{||\Theta^*||}\Theta^* \\ \sin \dfrac{t}{||\Theta^*||}\Theta^* \\ 0 \end{pmatrix}$$

is a geodesic with $\gamma(0) \in [I_n]$ and $\gamma(||\Theta^*||) = P_1$. It remains to show that $||\Theta^*|| < ||\tilde{\Theta}||$.

$$\begin{aligned} ||\Theta^*|| &= \left(\sum_{i=2}^{k} \left(\frac{||\Theta|| + \varepsilon}{||\Theta||} \right)^2 \theta_i^2 + \pi^2 - \frac{2\pi^2}{2} \left(\frac{||\Theta|| + \varepsilon}{||\Theta||} \right) \theta_1 + \left(\frac{||\Theta|| + \varepsilon}{||\Theta||} \right)^2 \theta_1^2 \right)^{1/2} \\ &< \left(\sum_{i=1}^{k} \left(\frac{||\Theta|| + \varepsilon}{||\Theta||} \right)^2 \theta_i^2 \right)^{1/2} = ||\Theta|| + \varepsilon = ||\tilde{\Theta}||. \end{aligned}$$

Therefore $\sup\{t : d(C(0), C(t)) = t\} \leq ||\Theta||$.

On the other hand the distance between I and P is $\left(\sum_i \theta_i^2 \right)^{1/2}$ which is the length of C so $\sup\{t : d(C(0), C(t)) = t\} \geq ||\Theta||$. This completes the proof. □

The following figure shows intuitively why subspaces having some principal angle between them equal to $\pi/2$ no longer have a unique minimizing geodesic joining them.

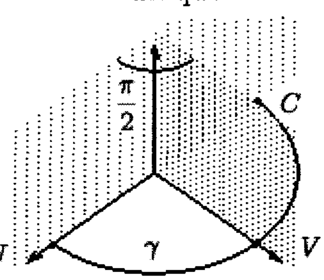

Figure 6: Subspaces With $\theta_1 = \pi/2$

Theorem 2.5 together with the fact that a unit speed geodesic between points U and V can be rotated to a unit speed geodesic between I and $U^T V$ shows that for any point P the cut locus with respect to P is the set

$$\begin{aligned} \mathrm{Cut}_P = \{U : &\text{ The matrix } \Theta \text{ corresponding to } P \text{ and } U \text{ has at least one } \theta_i = \pi/2\} \\ &\{U : \text{ The matrix } \Theta \text{ corresponding to } P \text{ and } U \text{ has } \theta_1 = \pi/2\}. \end{aligned}$$

Definition 2.10. The **injectivity radius** i_{M} of a Riemannian Manifold M is defined as

$$i_{\mathrm{M}} = \inf_{P \in \mathrm{M}} d_{\mathrm{M}}(P, \mathrm{Cut}_P).$$

This is the radius within which the exponential function is guaranteed to be injective. In other words, for all P, $\exp_P|_{B(P, i_{\mathrm{M}})}$ is injective. On **Gr** the injectivity radius is given by

$$\begin{aligned}\inf_{P \in \mathbf{Gr}} d(P, \mathrm{Cut}_P) &= d(I, \mathrm{Cut}_I) \\ &= \min\{||\Theta|| : \Theta \text{ corresponds to } I \text{ and } P, \text{ and } \theta_i = \pi/2 \text{ for some } i\} \\ &= \pi/2.\end{aligned}$$

3 Derivatives, Curvature, and Volume on Gr

3.1 The Gradient $\text{grad}_{\mathbf{Gr}} f$

Let $f : \mathbf{Gr} \longrightarrow \mathbb{R}$ be a function invariant under the choice of representation in $\mathbf{Gr}$. The gradient $\text{grad}_{\mathbf{Gr}} f$ is defined to be the tangent in $H_Q\,\mathbf{Gr}$ such that for all $T \in H_Q\mathbf{Gr}$, $\langle \text{grad}_{\mathbb{R}^{n\times n}} f, T\rangle_{n\times n} = \langle \text{grad}_{\mathbf{Gr}} f, T\rangle_{\mathbf{Gr}}$.

Proposition 3.1. *The gradient*

$$\text{grad}_{Gr} f = f_Q - Qf_Q^TQ = Q(Q^Tf_Q - f_Q^TQ) = 2Q\text{skew}\,(Q^Tf_Q)$$

$$\textit{where } f_Q = \begin{pmatrix} & \vdots & \\ \cdots & \dfrac{\partial f}{\partial q_{ij}} & \cdots \\ & \vdots & \end{pmatrix} = \text{grad}_{\mathbb{R}^{n\times n}} f.$$

Proof. Let $T = Q\mathcal{B}_0 \in H_Q\mathbf{Gr}$.

$$\begin{aligned}
&\langle f_Q, T\rangle_{n\times n} && \\
&= \text{tr}\,((QQ^Tf_Q)^TQ\mathcal{B}_0) && = \frac{1}{2}(\text{tr}\,(f_Q^TQ\mathcal{B}_0) - \text{tr}\,(Q^Tf_Q\mathcal{B}_0)) \\
&= \text{tr}\,((Q^Tf_Q)^T\mathcal{B}_0) && = \text{tr}\,(\frac{1}{2}(f_Q^TQ - Q^Tf_Q)\mathcal{B}_0) \\
&= \text{tr}\,(f_Q^TQ\mathcal{B}_0) && = \text{tr}\,((\text{skew}\,(Q^Tf_Q))^T\mathcal{B}_0) \\
&= \frac{1}{2}(\text{tr}\,(f_Q^TQ\mathcal{B}_0) + \text{tr}\,(\mathcal{B}_0f_Q^TQ)) && = 2\langle Q(\text{skew}\,(Q^Tf_Q), T\rangle_{\mathbf{Gr}} \\
&= \frac{1}{2}(\text{tr}\,(f_Q^TQ\mathcal{B}_0) + \text{tr}\,(Q^Tf_Q\mathcal{B}_0^T)) && = \langle 2Q(\text{skew}\,(Q^Tf_Q), T\rangle_{\mathbf{Gr}}
\end{aligned}$$

Therefore $\text{grad}_{\mathbf{Gr}} f = 2Q\text{skew}\,(Q^Tf_Q) = f_Q - Qf_Q^TQ$. ☐

Since the Lie Derivative $\mathcal{L}_X f = df(X) = \langle \text{grad}_{\mathbf{Gr}} f, X\rangle_{\mathbf{Gr}}$. It is now possible to easily compute Lie derivatives of functions. Defining $Xf = \mathcal{L}_X f$ gives a way to view how vector fields act on functions. Given a finite collection of vector fields $\{X_j\}_{j\in J}$ the Lie derivative

$$\left(\sum_j X_j\right) f = \langle \text{grad}_{\mathbf{Gr}} f, \sum_j X_j\rangle_{\mathbf{Gr}} = \sum_j \langle \text{grad}_{\mathbf{Gr}} f, X_j\rangle_{\mathbf{Gr}} = \sum_j (X_j f)$$

as expected. Lie derivatives of vector fields will be discussed in Section 3.5.

3.2 Parallel Translation Along Geodesics

The condition that geodesics parallel translate their own tangent vectors ($\nabla_{\dot{C}}\dot{C} = 0$) and the fact that when $C(t) = C_0 e^{t\mathcal{B}_0}$, $\dot{C} = C\mathcal{B}_0$ suggests the following proposition which can be found in [8].

Theorem 3.1. *Let $T = C_0\mathcal{W}_0$, $\mathcal{W}_0 \in \text{skew}_{n\times n}$, be a tangent in H_{C_0} St or H_{C_0} Gr and let $C(t) = C_0 e^{t\mathcal{B}_0}$. The parallel translate $\tau_{C,0,t_1}(T)$ of T along C from $t = 0$ to $t = t_1$ is given by*

$$\tau_{C,0,t_1}(T) = C_0 e^{t_1\mathcal{B}_0}\mathcal{W}_0.$$

Proof. Assume first that $T \in T_{C_0}\mathbf{O}_n$. Let $\tau(T)$ denote $\tau_{C,0,t}(T)$. At $t = 0$

$$\tau(T) = T - \Pi_\perp(T).$$

$\tau(T)$ is obtained by translating T in $\mathbb{R}^{n\times n}$ and infinitesimally removing the normal component so at $t = 0$ the formula $\left.\frac{d}{dt}\right|_{t=0}\tau(T) = -\left.\frac{d}{dt}\right|_{t=0}\Pi_\perp(T)$ holds.

$$\frac{d}{dt}\Pi_\perp(T) = \frac{d}{dt}C\frac{1}{2}(C^TT + T^TC) = \dot{C}\frac{1}{2}\underbrace{(C^TT + T^TC)}_{=0 \text{ when } t=0} + C\frac{1}{2}(\dot{C}^TT + \underbrace{C^T\dot{T}}_{=0} + \underbrace{\dot{T}^TC}_{=0} + T^T\dot{C})$$

so

$$\left.\frac{d}{dt}\right|_{t=0}\tau(T) = -C_0\frac{1}{2}(\dot{C}_0^TT_0 + T_0^T\dot{C}_0) = -C_0\frac{1}{2}(\mathcal{B}_0^TC_0^TC_0\mathcal{W}_0 + \mathcal{W}_0^TC_0^TC_0\mathcal{B}_0) = -C_0\mathcal{B}_0^T\mathcal{W}_0.$$

Since $\tau(T) \in T_C\mathbf{O}_n$ let $\tau(T) = C\mathcal{A}(t)$ where $\mathcal{A}(t) \in \text{skew}_{n\times n}$,

$$\text{then} \quad \left.\frac{d}{dt}\right|_{t=0}\tau(T) = \dot{C}_0\mathcal{A}_0 + C_0\dot{\mathcal{A}}(0) = C_0\mathcal{B}_0\mathcal{A}_0 + C_0\dot{\mathcal{A}}(0).$$

$$\text{Now} \quad -C_0\mathcal{B}_0^T\mathcal{W}_0 = C_0\mathcal{B}_0\mathcal{A}_0 + C_0\dot{\mathcal{A}}(0)$$

$$\text{so} \quad \dot{\mathcal{A}}(0) = -\mathcal{B}_0^T\mathcal{W}_0 - \mathcal{B}_0\mathcal{A}_0 = \mathcal{B}_0\mathcal{W}_0 - \mathcal{B}_0\mathcal{W}_0 = 0 \qquad (\text{since } \mathcal{W}_0 = \mathcal{A}_0).$$

The same argument may be applied anywhere along C with T replaced by the parallel translated tangent $\tau(T)$ therefore $\dot{\mathcal{A}}(t) = 0$ for all t so $\mathcal{A}(t) \equiv \mathcal{A}_0$ so that $\tau(T) = C\mathcal{W}_0$. □

Notice that if $T \in H_C\,\mathrm{M}$ where $\mathrm{M} = \mathbf{St}$ or $\mathrm{M} = \mathbf{Gr}$ then $\tau\,(T) \in H_C\,\mathrm{M}$ for all t. This shows that parallel translation along geodesics in either of these manifolds is given by the same equation.

3.3 Covariant Derivatives of Vector Fields Along Geodesics

The form of the covariant derivative of a vector field along a geodesic in **Gr** is very simple.

Theorem 3.2. *Let $C(t) = \mathrm{e}^{t\Psi}$ be a geodesic emanating from I_n and reaching $Q_U^T Q_V$ at $t = ||\Theta||$. Let*

$$H_{C(t)}\,Gr \ni Y_t = Y_{C(t)} = C(t)\mathcal{B}_t = C(t)\begin{pmatrix} 0 & -B_t^T \\ B_t & 0 \end{pmatrix}$$

be a vector field along $C(t)$. The covariant derivative $\nabla_{\dot{C}}Y(t_0) = \mathrm{e}^{t_0\Psi}\,\dot{\mathcal{B}}(t_0)$.

Proof. By definition,

$$\begin{aligned}
\nabla_{\dot{C}}Y(t_0) &= \lim_{h\to 0}\frac{1}{h}(\tau^{-1}_{C,t_0,t_0+h}(Y_{t_0+h}) - Y_{t_0}) \\
&= \lim_{h\to 0}\frac{1}{h}\left(\mathrm{e}^{t_0\Psi}\mathcal{B}_{t_0+h} - \mathrm{e}^{t_0\Psi}\mathcal{B}_{t_0}\right) \\
&= \mathrm{e}^{t_0\Psi}\lim_{h\to 0}\frac{1}{h}\left(\mathcal{B}_{t_0+h} - \mathcal{B}_{t_0}\right) \qquad = \mathrm{e}^{t_0\Psi}\left.\frac{d}{dt}\right|_{t=t_0}\mathcal{B}_t.
\end{aligned}$$

□

Note that the geodesic condition $\nabla_{\dot{C}}\dot{C} = 0$ is consistent with this result; $\nabla_{\dot{C}}\,\dot{C} = \mathrm{e}^{t_0\Psi}\left.\frac{d}{dt}\right|_{t=t_0}\Psi = 0$. The condition $\nabla_{\dot{C}}\,\tau\,(T) = 0$ on the parallel translated vector $\tau(T)$ is similarly consistent.

3.4 Normal Coordinates

An ON basis $\mathfrak{B}_1 = \{\mathrm{e}_{\alpha\beta} :\ 1 \leq \alpha \leq n-k,\ 1 \leq \beta \leq k\}$ of $\mathfrak{m}_1 := H_I\,\mathbf{Gr}(\mathbb{R})$ will now be described that will prove to be very convenient because of its relation to the eigenspace of the tensor $\langle R(\,\cdot\,,\Psi)\Psi,\,\cdot\,\rangle_{\mathbf{Gr}}$. $\langle R(\,\cdot\,,\Psi)\Psi,\,\cdot\,\rangle_{\mathbf{Gr}}$ is symmetric and positive semi-definite because of the non-negative sectional curvature of Gr which will be established in Section 3.5. This guarantees (see [13]) that the tensor is diagonalizable with real, non-negative eigenvalues. $\mathfrak{B}_1$ will consist of the natural bases of

$$\begin{pmatrix} 0 & \begin{pmatrix} -\mathrm{symm}_{k\times k}^{\mathcal{T}} & 0 \end{pmatrix} \\ \begin{pmatrix} \mathrm{symm}_{k\times k} \\ 0 \end{pmatrix} & 0 \end{pmatrix}, \begin{pmatrix} 0 & \begin{pmatrix} -\mathrm{skew}_{k\times k}^{\mathcal{T}} & 0 \end{pmatrix} \\ \begin{pmatrix} \mathrm{skew}_{k\times k} \\ 0 \end{pmatrix} & 0 \end{pmatrix},$$

$$\text{and} \begin{pmatrix} 0 & \begin{pmatrix} 0 & -\left(\mathbb{R}^{n-2k\times k}\right)^{\mathcal{T}} \end{pmatrix} \\ \begin{pmatrix} 0 \\ \mathbb{R}^{n-2k\times k} \end{pmatrix} & 0 \end{pmatrix}.$$

The indices $1 \le \beta \le \alpha \le k$ will be used to describe the matrices with symmetric submatrices in the lower left blocks, for example if $\alpha = 2$, $\beta = 1$ then define

$$\mathrm{e}_{\alpha\beta} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 1 & 0 \\ \textcircled{1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & 0 \end{pmatrix}.$$

Tedious calculations that are easily verified with a computer algebra system show that for these matrices, when $\alpha \neq \beta$, $-[[\mathrm{e}_{\alpha\beta}, \Psi], \Psi] = \dfrac{(\theta_\beta - \theta_\alpha)^2}{||\Theta||^2}\mathrm{e}_{\alpha\beta}$. For each of the matrices $\mathrm{e}_{\alpha\alpha}$ having a 1 on the diagonal in the $(\alpha\alpha)^{th}$ position in the lower left block $-[[\mathrm{e}_{\alpha\alpha}, \Psi], \Psi] = 0$. Similarly, matrices with skew-symmetric submatrices in the lower left blocks, can be defined for $1 \le \alpha < \beta \le k$ as in the following example where $\alpha = 1$ and $\beta = 3$.

$$\mathrm{e}_{\alpha\beta} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & \textcircled{1} \\ 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & 0 \end{pmatrix}.$$

Calculations show that here $-[[\mathrm{e}_{\alpha\beta}, \Psi], \Psi] = \dfrac{(\theta_\alpha + \theta_\beta)^2}{||\Theta||^2}\mathrm{e}_{\alpha\beta}$. There are $k(n-2k)$ tangent matrices $\mathrm{e}_{\alpha\beta}$ that have a 1 somewhere in the $(k+i)^{th}$ row of the lower left block and 0's elsewhere. For instance if $\alpha = k+1$ and $\beta = 2$,

$$\mathrm{e}_{\alpha\beta} = \begin{pmatrix} 0 & \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \textcircled{1} & 0 \end{pmatrix} & 0 \end{pmatrix}.$$

In these cases $-[[\mathrm{e}_{\alpha\beta}, \Psi], \Psi] = \frac{\theta_\beta^2}{||\Theta||^2}\,\mathrm{e}_{\alpha\beta}$. These relationships can be expressed by $-[[\mathrm{e}_{\alpha\beta}, \Psi], \Psi] = \lambda_{\alpha\beta}(\Theta)\,\mathrm{e}_{\alpha\beta}$ where

$$||\Theta||^2 \lambda_{\alpha\beta}(\Theta) = \begin{cases} (\theta_\beta - \theta_\alpha)^2 & , \ 1 \le \beta < \alpha \le k \\ 0 & , \ 1 \le \alpha = \beta \le k \\ (\theta_\alpha + \theta_\beta)^2 & , \ 1 \le \alpha < \beta \le k \\ \theta_\beta^2 & , \ \alpha \ge k+1. \end{cases}$$

In the complex case the role of skew-symmetric matrices is replaced with skew-Hermitian matrices. Recall that the canonical metric $\langle T_1, T_2 \rangle_{\mathbf{Gr}(\mathbb{C})} = \frac{1}{2}\mathrm{tr}(T_2^{\mathcal{H}} T_1)$, and that $\dim \mathbf{Gr}(\mathbb{C}) = 2 \dim \mathbf{Gr}(\mathbb{R})$. The ON basis $\mathfrak{B}_2$ of $\mathfrak{m}_2 := H_I\,\mathbf{Gr}(\mathbb{C})$ consisting of skew-Hermitian matrices that corresponds to $\mathfrak{B}_1 = \{\mathrm{e}_{\alpha\beta}\}$ can be described in terms of $\mathrm{e}_{\alpha\beta} := \begin{pmatrix} 0 & -B_{\alpha\beta}^{\mathcal{T}} \\ B_{\alpha\beta} & 0 \end{pmatrix}$.

$$\mathfrak{B}_2 = \left\{ \begin{pmatrix} 0 & -B_{\alpha\beta}^{\mathcal{T}} \\ B_{\alpha\beta} & 0 \end{pmatrix}, \begin{pmatrix} 0 & iB_{\alpha\beta}^{\mathcal{T}} \\ iB_{\alpha\beta} & 0 \end{pmatrix} \right\} = \left\{ \mathrm{e}_{\alpha\beta}, \begin{pmatrix} -iI_k & 0 \\ 0 & iI_{n-k} \end{pmatrix} \mathrm{e}_{\alpha\beta} \right\}$$

$$:= \{\mathrm{e}_{\alpha\beta 1}, \mathrm{e}_{\alpha\beta 2}\}$$

For this basis $-[[\mathrm{e}_{\alpha\beta\gamma}, \Psi], \Psi] = \lambda_{\alpha\beta\gamma}(\Theta)\,\mathrm{e}_{\alpha\beta\gamma}$ where

$$||\Theta||^2\lambda_{\alpha\beta\gamma}(\Theta) = \begin{cases} \gamma = 1 & \gamma = 2 & \\ (\theta_\beta - \theta_\alpha)^2 & (\theta_\beta + \theta_\alpha)^2 & , \ 1 \le \beta < \alpha \le k \\ 0 & 4\theta_\beta^2 & , \ 1 \le \alpha = \beta \le k \\ (\theta_\alpha + \theta_\beta)^2 & (\theta_\alpha - \theta_\beta)^2 & , \ 1 \le \alpha < \beta \le k \\ \theta_\beta^2 & \theta_\beta^2 & , \ \alpha \ge k+1. \end{cases}$$

Normal or Geodesic Coordinates (see [5] pg. 83) at a point Q_0 are defined on **Gr** by taking any ON basis (say $\mathfrak{B}_1$), applying the exponential to some linear combination $\sum c_{\alpha\beta}\mathrm{e}_{\alpha\beta}$ and taking $Q_0\mathrm{e}^{\Sigma c_{\alpha\beta}\mathrm{e}_{\alpha\beta}}$ as coordinates. As in a general Riemannian Manifold exponential coordinates satisfy

$$\nabla_{Q\mathrm{e}_{\alpha\beta}} Q\mathrm{e}_{ij}\Big|_Q = Q\,\frac{d}{dt}\Big|_{t=0}\mathrm{e}_{ij} = 0 \qquad \text{where } Q = Q_0\mathrm{e}^{t\mathrm{e}_{\alpha\beta}}.$$

The ON basis $\left\{ Q_0 \begin{pmatrix} 0 & -E_{ij}^T \\ E_{ij} & 0 \end{pmatrix} \right\}$ of H_{Q_0} **Gr** could also be used for exponential coordinates according to convenience.

3.5 Theory of Homogeneous Spaces

At this stage it is easiest to draw on the theory of homogeneous spaces and to interpret general results in terms of the Grassmannian. The general material in this section is developed with proof in [6] Chapter 3, and appears partially in [5] pg 187. General results will be stated without proof (labeled **Theorem**) and the application of the results to $\mathbf{O}_n/\mathbf{K}$ will be described in more detail (labeled **Claim**). The curvature sign convention is taken to be

$$R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z.$$

Definition 3.1. A **Riemannian submersion** $\pi : \mathbf{M}^{n_1+n_2} \longrightarrow \mathbf{N}^{n_1}$ between Riemannian manifolds is a differentiable map such that rank $(d\pi) = n_1$ everywhere.

Claim 3.1. *Gr is submersed onto O_n via the Riemannian submersion*

$$\pi : O_n \longrightarrow Gr = G/K \qquad \textit{where } K = \begin{pmatrix} O_k & 0 \\ 0 & O_{n-k} \end{pmatrix}.$$
$$Q \mapsto QK$$

Proof. If $\mathcal{W} \in \text{skew}_{n\times n}$ define $\mathcal{W}^h$ to be $\mathcal{W}$ with the vertical components set to zero. It has already been established that $d\pi : T_Q \mathbf{O}_n \ni Q\mathcal{W} \mapsto Q\mathcal{W}^h \in H_Q \mathbf{Gr}$ and that $\text{rank}\,(d\pi) = k(n-k)$ everywhere. Therefore π is a Riemannian submersion. □

Every $X \in H_p\,\mathbf{G/K}$ has a unique horizontal lift $\overline{X} \in T_p\,\mathbf{G}$. The unique lift of $QW \in H_Q\,\mathbf{Gr}$ is $QW \in T_Q\,\mathbf{O}_n$.

Definition 3.2. On a Lie group G a **left invariant vector field** is a vector field X such that $dL_g(X(g_1)) = X(gg_1)$ where $L_g : g_1 \mapsto gg_1$ is left multiplication by the element $g \in \mathbf{G}$. The same definition is used for **right invariant vector fields** where right multiplication by g is denoted by R_g.

There is a one to one correspondence between left invariant vector fields on $\mathbf{G}$ and tangent vectors in $\mathfrak{g} := T_e\mathbf{G}$ where e denotes the identity element of $\mathbf{G}$. $\mathfrak{g}$ is called the **Lie algebra** of $\mathbf{G}$.

In $\mathbf{O}_n$ the left invariant vector fields are fields $W_Q = Q\mathcal{W}$ where $\mathcal{W} \in T_I\,\mathbf{O}_n$ is fixed. Vector fields on $\mathbf{Gr}$ with left invariant horizontal lifts have the form $W_Q = Q\mathcal{W}$ where $\mathcal{W} \in \mathfrak{m}$ is fixed. If $\mathfrak{k} = V_I\,\mathbf{Gr}$ then $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{k}$. $e^{t\Psi}\Psi$ is both left and right invariant in $\mathbf{O}_n$ because $e^{t\Psi}$ always commutes with Ψ (see [18]).

Let X, Y, Z, and W be left invariant vector fields on $\mathbf{G}$. The Lie bracket $[X,Y] = \mathcal{L}_X Y = XY - YX$ on a Lie group has the property that if X and Y are left invariant vector fields then $dL_g[X_e, Y_e] = [dL_gX_e, dL_gY_e] = [X_g, Y_g]$, that is if X and Y are left invariant then so is $[X,Y]$. In the case of $\mathbf{O}_n$ and left invariant vector fields $\mathcal{W}_1, \mathcal{W}_2 \in T_I\,\mathbf{O}_n$, $Q[\mathcal{W}_1, \mathcal{W}_2] = [Q\mathcal{W}_1, Q\mathcal{W}_2]$. It can also be shown that on a matrix Lie group $[X,Y]$ acts on functions $f \in C^\infty(\mathbf{G})$ at e by applying the tangent matrix $X_eY_e - Y_eX_e$ to f.

Definition 3.3. Metrics in **G** invariant under left and right translation are called **bi-invariant**.

Claim 3.2. *$\langle \cdot, \cdot \rangle_{O_n}$ is bi-invariant.*

Proof. Let $\mathcal{W}_1, \mathcal{W}_2 \in \mathfrak{g}$. Left invariance is trivial. To check right invariance,
$\langle \mathcal{W}_1 Q, \mathcal{W}_2 Q \rangle_{O_n} = \frac{1}{2}\operatorname{tr}(Q^T \mathcal{W}_1^T \mathcal{W}_2 Q) = \frac{1}{2}\operatorname{tr}(\mathcal{W}_2 Q Q^T \mathcal{W}_1^T) = \langle \mathcal{W}_1, \mathcal{W}_2 \rangle_{O_n}$. □

Notice that if $X = Q\mathcal{X}$ is a left invariant vector field in $\mathbf{O}_n$ then taking the conjugate $\mathcal{X}' = Q^T \mathcal{X} Q \in \text{skew}_{n\times n}$ gives rise to a right invariant vector field X' that has the same length as X.

In both **Gr** and $\mathbf{O}_n$, $\langle [[X,Y],Z],W\rangle = \langle [X,Y],[Z,W]\rangle$ follows from the following calculation.

$$\begin{aligned}
\langle [X,Y],[Z,W]\rangle &= \langle XY - YX, ZW - WZ\rangle \\
&= \frac{1}{2}\operatorname{tr}(Y^TX^TZW - Y^TX^TWZ - X^TY^TZW + X^TY^TWZ) \\
&= \frac{1}{2}\operatorname{tr}(-Y^TX^TZ^TW + WZ^TY^TX^T + X^TY^TZ^TW - WZ^TX^TY^T) \\
&= -\langle ZXY, W\rangle + \langle XYZ, W\rangle + \langle ZYX, W\rangle - \langle YXZ, W\rangle \\
&= \langle (XY - YX)Z - Z(XY - YX), W\rangle \\
&= \langle [[X,Y],Z],W\rangle
\end{aligned}$$

The following results for **G** relate the Lie bracket to the covariant derivative and hence the curvature tensor.

Theorem 3.3. *If $\langle \cdot, \cdot \rangle_G$ is bi-invariant and $X, Y, Z,$ and W are left invariant then on G,*

i) $\nabla_X Y = \frac{1}{2}[X,Y]$

ii) $\langle R(X,Y)Z, W\rangle_G = -\frac{1}{4}\langle [[X,Y],Z],W\rangle_G$

iii) $\langle R(X,Y)Y, X\rangle_G = \frac{1}{4}\|[X,Y]\|^2$

Notice that (ii) together with $\langle [[X,Y],Z],W\rangle = \langle [X,Y],[Z,W]\rangle$ shows that the sectional curvature of $\mathbf{G}$ is always non-negative.

In general the homogeneous space inherits the metric of the original group so that $\langle\cdot,\cdot\rangle_{\mathbf{G}} = \langle\cdot,\cdot\rangle_{\mathbf{G/K}}$ when applied to horizontal vectors. It has already been observed that $\langle\cdot,\cdot\rangle_{\mathbf{O}_n} = \langle\cdot,\cdot\rangle_{\mathbf{Gr}}$ when applied to horizontal vectors. The subscripts $\mathbf{G}$, $\mathbf{G/K}$, $\mathbf{O}_n$, and $\mathbf{Gr}$ will now be dropped. The following results relate the covariant derivatives and curvature tensors on $\mathbf{G}$ and $\mathbf{G/K}$.

Theorem 3.4. *Let X,Y,Z, and W be left invariant vector fields on G/K and $\overline{X},\overline{Y},\overline{Z}$, and $\overline{W}$ be their horizontal lifts on G. Let R and $\overline{R}$ denote the curvature tensors on G/K and G, then*

i) $\overline{\nabla_X Y} = \frac{1}{2}[\overline{X},\overline{Y}]^h$

ii) $\langle \overline{R}(\overline{X},\overline{Y})\overline{Z},\overline{W}\rangle$
$$= \langle R(X,Y)Z,W\rangle + \frac{1}{4}\langle [\overline{X},\overline{Z}]^v,[\overline{Y},\overline{W}]^v\rangle - \frac{1}{4}\langle [\overline{Y},\overline{Z}]^v,[\overline{X},\overline{W}]^v\rangle + \frac{1}{2}\langle [\overline{Z},\overline{W}]^v,[\overline{X},\overline{Y}]^v\rangle.$$

On $\mathbf{Gr}$ this shows that $\nabla_X Y = \frac{1}{2}[X,Y]^h$. Now there is another way to verify Theorem 3.2.

Second proof of Theorem 3.2. Let $C = \mathrm{e}^{t\Psi}$ be a geodesic and let $X = C\mathcal{W} = C\sum a_{\alpha\beta}(t)\mathrm{e}_{\alpha\beta}$ be a vector field on $\mathbf{Gr}$.

$$\nabla_{\dot{C}}X = \sum(\nabla_{\dot{C}}a_{\alpha\beta}C\mathrm{e}_{\alpha\beta}) = \sum\left(\frac{da_{\alpha\beta}}{dt}C\mathrm{e}_{\alpha\beta} + a_{\alpha\beta}\nabla_{\dot{C}}C\mathrm{e}_{\alpha\beta}\right)$$
$$= \sum\left(\frac{da_{\alpha\beta}}{dt}C\mathrm{e}_{\alpha\beta} + a_{\alpha\beta}C\frac{1}{2}\underbrace{[\Psi,\mathrm{e}_{\alpha\beta}]^h}_{=0}\right) = C\sum\frac{da_{\alpha\beta}}{dt}\mathrm{e}_{\alpha\beta} = C\frac{d}{dt}\mathcal{W}.$$ □

Claim 3.3. *For vector fields X,Y, and Z on Gr with left invariant horizontal lifts on O_n the curvature tensor $R(X,Y)Z = -[[X,Y],Z]$.*

Proof. Without loss of generality assume that $X,Y,Z,W \in \mathfrak{m}$.

$$
\begin{aligned}
&\langle R(X,Y)Z,W\rangle \\
&= \langle \overline{R}(\overline{X},\overline{Y})\overline{Z},\overline{W}\rangle - \frac{1}{4}\langle[\overline{X},\overline{Z}]^v,[\overline{Y},\overline{W}]^v\rangle + \frac{1}{4}\langle[\overline{Y},\overline{Z}]^v,[\overline{X},\overline{W}]^v\rangle - \frac{1}{2}\langle[\overline{Z},\overline{W}]^v,[\overline{X},\overline{Y}]^v\rangle \\
&= -\frac{1}{4}\langle[[X,Y],Z],W\rangle - \frac{1}{4}\langle[X,Z],[Y,W]\rangle + \frac{1}{4}\langle[Y,Z],[X,W]\rangle - \frac{1}{2}\langle[Z,W],[X,Y]\rangle \\
&= \frac{1}{4}\langle[[Z,X],Y],W\rangle + \frac{1}{4}\langle[Y,Z],[X,W]\rangle - \frac{1}{4}\langle[[X,Z],Y],W\rangle + \frac{1}{4}\langle[[Y,Z],X],W\rangle \\
&\qquad -\frac{1}{2}\langle[[X,Y],Z],W\rangle \qquad \text{(Jacobi identity)} \\
&= \frac{1}{2}\langle[[Y,Z],X],W\rangle - \frac{1}{2}\langle[[X,Z],Y],W\rangle - \frac{1}{2}\langle[[X,Y],Z],W\rangle \qquad \text{(Jacobi identity)} \\
&= \frac{1}{2}\langle[[Y,Z],X],W\rangle + \frac{1}{2}\langle[[Z,X],Y],W\rangle + \frac{1}{2}\langle[[Y,X],Z],W\rangle \\
&= -\frac{1}{2}\langle[[X,Y],Z],W\rangle + \frac{1}{2}\langle[[Y,X],Z],W\rangle \qquad \text{(Jacobi Identity)} \\
&= -\langle\underbrace{[[X,Y],Z]}_{\in\mathfrak{m}},W\rangle.
\end{aligned}
$$

Therefore $R(X,Y)Z = -[[X,Y],Z]$. □

It will be important for finding Jacobi fields that in particular,

$$R(e^{t\Psi}e_{\alpha\beta}, e^{t\Psi}\Psi)e^{t\Psi}\Psi = -e^{t\Psi}[[e_{\alpha\beta},\Psi],\Psi].$$

3.6 Ricci, Sectional, and Scalar Curvatures

Definition 3.4. **The Ricci curvature on Gr** is given by

$$\mathrm{Ric}_I(X,Y) = \frac{1}{\nu k(n-k)-1}\sum_{\alpha,\beta,\gamma} R_I(X, e_{\alpha\beta\gamma}, e_{\alpha\beta\gamma}, Y).$$

Usually $\mathrm{Ric}(X,X)$ is written $\mathrm{Ric}(X)$.

Proposition 3.2. *On Gr,*

$$\mathrm{Ric}(X,Y) = \begin{cases} \dfrac{n-k-1}{k(n-k)-1}\langle X,Y\rangle & , \ \nu = 1 \\[2ex] \dfrac{2(n-k-1)+4}{2k(n-k)-1}\langle X,Y\rangle & , \ \nu = 2 \end{cases}.$$

Where $\nu = 1$ corresponds to the real case and $\nu = 2$ corresponds to the complex case.

Proof. It is known (see [10] Proposition 3.21) that because **Gr** is an **isotropy irreducible** homogeneous space (meaning that the isotropy representation is irreducible), $\mathrm{Ric}(\cdot,\cdot) = a\langle\cdot,\cdot\rangle$ where $a \in \mathbb{R}$ is fixed. For the real case

$$\begin{aligned}
\mathrm{Ric}(\Psi,\Psi) &= \frac{1}{k(n-k)-1}\sum R(\mathrm{e}_{\alpha\beta},\Psi,\Psi,\mathrm{e}_{\alpha\beta}) \\
&= \frac{1}{k(n-k)-1}\sum_{\alpha,\beta}\lambda_{\alpha\beta} \\
&= \frac{1}{k(n-k)-1}\left(\frac{\sum_{\alpha<\beta}((\theta_\alpha+\theta_\beta)^2+(\theta_\alpha-\theta_\beta)^2)+(n-2k)\sum_\beta\theta_\beta^2}{||\Theta||^2}\right) \\
&= \frac{(k-1)||\Theta||^2+(n-2k)||\Theta||^2}{(k(n-k)-1)||\Theta||^2} \\
&= \frac{n-k-1}{k(n-k)-1} \\
&= \frac{n-k-1}{k(n-k)-1}\langle\Psi,\Psi\rangle.
\end{aligned}$$

Therefore $a = \dfrac{n-k-1}{k(n-k)-1}$. For the complex case an analogous argument shows

$$\begin{aligned}
\mathrm{Ric}(X,Y) &= \frac{1}{2k(n-k)-1}\left(\frac{2(n-k-1)||\Theta||^2+4||\Theta||^2}{||\Theta||^2}\right)\langle X,Y\rangle \\
&= \frac{2(n-k-1)+4}{2k(n-k)-1}\langle X,Y\rangle.
\end{aligned}$$

Therefore $\mathrm{Ric}(X,Y) = a_\nu\langle X,Y\rangle$ where $a_\nu = \begin{cases} \dfrac{n-k-1}{k(n-k)-1} &, \ \nu=1 \\ \dfrac{2(n-k-1)+4}{2k(n-k)-1} &, \ \nu=2 \end{cases}$. □

Definition 3.5. On **Gr** the **sectional curvature** $K(X,Y)$ is given by

$$K(X,Y) = \frac{R(X,Y,Y,X)}{||X||||Y||-\langle X,Y\rangle^2} = \frac{\langle[X,Y],[X,Y]\rangle}{||X||||Y||-\langle X,Y\rangle^2}.$$

The maximum sectional curvature will be a useful quantity in roughly bounding the volume of a geodesic ball from below.

Proposition 3.3. *On Gr*

$$\max_{||X||=||Y||=1}\{K_Q(X,Y)\} = \begin{cases} 2 & , \nu=1 \\ 4 & , \nu=2 \end{cases}.$$

Proof. It suffices restrict attention to K_I. For any $[Q] \in \mathbf{Gr}$ there are representations Q' and I' such that $I'e^{t\Psi}$ passes through Q' where $\Psi = \frac{1}{||\Theta||}\begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. For each Ψ, if $\alpha \neq \beta$, then

$$K(\mathbf{e}_{\alpha\beta}, \Psi) = R(\mathbf{e}_{\alpha\beta}, \Psi, \Psi, \mathbf{e}_{\alpha\beta}) = \lambda_{\alpha\beta}\langle \mathbf{e}_{\alpha\beta}, \mathbf{e}_{\alpha\beta}\rangle = \lambda_{\alpha\beta}$$

which in the real case has the maximum value $\frac{(\theta_1+\theta_2)^2}{||\Theta||^2}$. Now maximizing this over all possible Θ gives

$$\max_{\Theta}\left\{\frac{(\theta_1+\theta_2)^2}{||\Theta||^2}\right\} = \frac{(\pi/2+\pi/2)^2}{2(\pi/2)^2} = 2$$

In the complex case the same argument leads to
$\max_{||X||=||Y||=1}\{K_Q(X,Y)\} = \max_{\Theta}\left\{\frac{4\theta_1^2}{||\Theta||^2}\right\} = 4\frac{(\pi/2)^2}{(\pi/2)^2} = 4.$ □

This latter number agrees with [12] pg. 3448. For future reference define

$$b_\nu = \begin{cases} 2 & , \nu=1 \\ 4 & , \nu=2 \end{cases}.$$

Definition 3.6. On **Gr** the **scalar curvature** $\mathrm{Scal}(Q)$ is given by

$$\mathrm{Scal}(Q) = \frac{1}{k(n-k)}\sum \mathrm{Ric}(Q\mathbf{e}_{\alpha\beta}).$$

The scalar curvature is constant, indeed for any $Q \in \mathbf{Gr}$,

$$\mathrm{Ric}(Q\mathbf{e}_{\alpha\beta}) = \begin{cases} \dfrac{n-k-1}{k(n-k)-1} & , \nu=1 \\ \dfrac{2(n-k-1)+4}{2k(n-k)-1} & , \nu=2 \end{cases} \quad \text{so}$$

$$\mathrm{Scal}(Q) = \begin{cases} \dfrac{n-k-1}{k(n-k)-1} & , \nu=1 \\ \dfrac{2(n-k-1)+4}{2k(n-k)-1} & , \nu=2. \end{cases}$$

3.7 Jacobi Fields

The Jacobi equation along the geodesic $C(t)$ connecting I_n to $Q_U^T Q_V$ is

$$\nabla_{\dot{C}} \nabla_{\dot{C}} J + R(J, \dot{C})\dot{C} = 0$$

as derived in [5] pg. 111. Along C a Jacobi field $J_{\alpha\beta}(t)$ in the direction of the parallel translated vector $\tau(\mathrm{e}_{\alpha\beta})$ can be written

$$J_{\alpha\beta}(t) = a_{\alpha\beta}(t)\mathrm{e}^{t\Psi}\mathrm{e}_{\alpha\beta}.$$

$$\nabla_{\dot{C}} \nabla_{\dot{C}} J_{\alpha\beta} = \nabla_{\dot{C}} \left(\frac{da_{\alpha\beta}}{dt}\mathrm{e}^{t\Psi}\mathrm{e}_{\alpha\beta} + a_{\alpha\beta} \underbrace{\nabla_{\dot{C}} \mathrm{e}^{t\Psi}\mathrm{e}_{\alpha\beta}}_{=0} \right) = \mathrm{e}^{t\Psi}\frac{d^2 a_{\alpha\beta}}{dt^2}\mathrm{e}_{\alpha\beta}.$$

It has been shown that $R(\mathrm{e}^{t\Psi}\mathrm{e}_{\alpha\beta}, \mathrm{e}^{t\Psi}\Psi)\,\mathrm{e}^{t\Psi}\Psi = \lambda_{\alpha\beta}\mathrm{e}^{t\Psi}\mathrm{e}_{\alpha\beta}$ so that the Jacobi equation becomes

$$0 = \mathrm{e}^{t\Psi}\frac{d^2 a_{\alpha\beta}}{dt^2} + a_{\alpha\beta}R(\mathrm{e}^{t\Psi}\mathrm{e}_{\alpha\beta}, \mathrm{e}^{t\Psi}\Psi)\mathrm{e}^{t\Psi}\Psi = \mathrm{e}^{t\Psi}(\frac{d^2 a_{\alpha\beta}}{dt^2} + \lambda_{\alpha\beta}a_{\alpha\beta})\mathrm{e}_{\alpha\beta}.$$

This implies that

$$\frac{d^2 a_{\alpha\beta}}{dt^2} + \lambda_{\alpha\beta}a_{\alpha\beta} = 0.$$

The condition $J_{\alpha\beta}(0) = 0$ gives $a_{\alpha\beta}(0) = 0$ so that

$$a_{\alpha\beta}(t) = \begin{cases} b_{\alpha\beta}\sin\sqrt{\lambda_{\alpha\beta}}\,t & , \quad \lambda_{\alpha\beta} \neq 0 \\ b_{\alpha\beta}\,t & , \quad \text{otherwise} \end{cases} \qquad b_{\alpha\beta} \in \mathbb{R}.$$

This leads to the volume elements used in [2] and [1]. The constants $b_{\alpha\beta}$ leading to a canonical volume form will be determined shortly. These results are summarized in the column $\gamma = 1$ in the following and the analogous complex case is described in both columns $\gamma = 1$ and 2 (assuming that $\lambda_{\alpha\beta\gamma} \neq 0$ unless $\alpha = \beta$ and $\gamma = 1$).

$$\frac{1}{b_{\alpha\beta\gamma}} a_{\alpha\beta\gamma}(t) = \begin{cases} \gamma = 1 & \gamma = 2 & \\ \sin\frac{(\theta_\beta - \theta_\alpha)}{||\Theta||}t & \sin\frac{(\theta_\beta + \theta_\alpha)}{||\Theta||}t & , \ 1 \le \beta < \alpha \le k, \\ t & \sin\frac{2\theta_\beta}{||\Theta||}t & , \ 1 \le \alpha = \beta \le k \\ \sin\frac{(\theta_\alpha + \theta_\beta)}{||\Theta||}t & \sin\frac{(\theta_\beta - \theta_\alpha)}{||\Theta||}t & , \ 1 \le \alpha < \beta \le k \\ \sin\frac{\theta_\beta}{||\Theta||}t & \sin\frac{\theta_\beta}{||\Theta||}t & , \ \alpha \ge k+1 \end{cases}$$

where the $a_{\alpha\beta\gamma}$ correspond to orthogonal Jacobi fields $J_{\alpha\beta\gamma} = a_{\alpha\beta\gamma}\, e^{t\Psi} e_{\alpha\beta\gamma}$ along C having length $|a_{\alpha\beta\gamma}(t)|$. In the cases that $\sqrt{\lambda_{\alpha\beta\gamma}} = 0$ the Jacobi fields become $J_{\alpha\beta\gamma} = b_{\alpha\beta\gamma}\, t\, e^{t\Psi} e_{\alpha\beta\gamma}$. Simple computations show that the conditions $\nabla_{\dot{C}} J_{\alpha\beta\gamma}(0) = e_{\alpha\beta\gamma}$ determine that

$$b_{\alpha\beta\gamma} = \begin{cases} 1 & , \ \lambda_{\alpha\beta\gamma} = 0 \\ \frac{1}{\sqrt{\lambda_{\alpha\beta\gamma}}} & , \ \text{otherwise} \end{cases} .$$

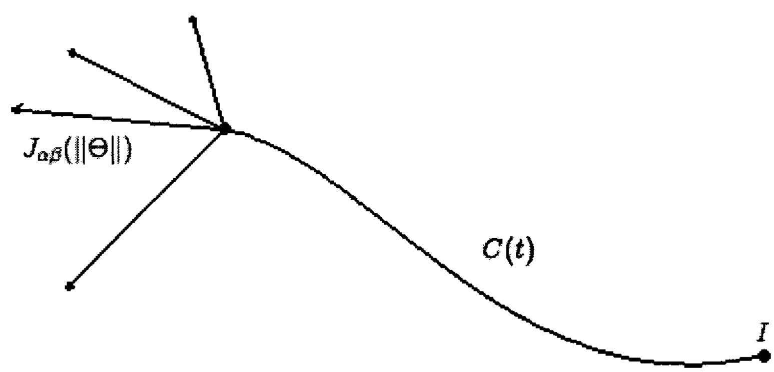

Figure 7: The Vectors $J_{\alpha\beta}(||\Theta||)$

3.8 Volumes of Geodesic Balls

Because $\mathbf{K}$ is a closed compact subset of the compact set $\mathbf{O}_n$ there exists a unique invariant density or volume form on $\mathbf{Gr}$ defined up to a multiplicative constant, this is proven in [17] pg. 168. The fact that the volume form is defined only up to a constant multiple is reflected in the fact that constant multiples of $J_{\alpha\beta\gamma}$ still satisfy the Jacobi equation. Loosely speaking the volume form is obtained by multiplying the lengths of the orthogonal Jacobi fields together. Following [10] pg. 137 and [14] pg. 412, the volume of a geodesic ball B radius $R(< \pi/2 = i_{\mathbf{Gr}})$ centered at I_n in $\mathbf{Gr}$ can be computed as follows. Let $r \in \mathbb{R}$, $r \leq R$, and let $\mathcal{U}, \mathcal{B} \in \mathfrak{m}$ be such that $r\mathcal{U} = \mathcal{B}$ and $||\mathcal{U}||_{\mathbf{Gr}} = 1$. Let $N_\nu = \dim \mathbf{Gr}(\mathbb{F})$ and relabel the $\mathrm{e}_{\alpha\beta\gamma}$ as e_i where $\mathrm{e}_1 := \mathrm{e}_{111}, \cdots, \mathrm{e}_k := \mathrm{e}_{kk1}$ and the rest are labeled in any way. Since there is a vector $\Psi = \begin{pmatrix} 0 & -\Theta & 0 \\ \Theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ having length one such that $k\mathrm{e}^{t\Psi} \in [\mathrm{e}^{t\mathcal{U}}]$ for some $k(t) \in \mathbf{K}$ it can be seen, using invariance of length under representation, that the N_ν orthogonal Jacobi fields along $\mathrm{e}^{t\mathcal{U}}$ pointing in the directions $\mathrm{e}^{t\mathcal{U}}\mathrm{e}_i$ are $J_i = a_i\, \mathrm{e}^{t\mathcal{U}}\mathrm{e}_i$, where the a_i are $a_{\alpha\beta\gamma}$ relabeled. Let $\{x_i\}_{i=1}^{N_\nu}$ be the normal coordinates at I so that $\mathcal{B} = \sum_{i=1}^{N_\nu} x_i \mathrm{e}_i$.

$$
\begin{aligned}
&\mathrm{vol}\,(B_I(R)) \\
&= \int_{B_0(R)\subset H_I \mathbf{Gr}} \exp^* \mathrm{vol}_I\,(\mathcal{B})\; dx_1 \cdots dx_{N_\nu} \\
&= \int_{\mathrm{S}^{N_\nu-1}} \int_0^R \exp^* \mathrm{vol}_I\,(r\mathcal{U})\; r^{N_\nu-1} dr d\mathrm{S}^{N_\nu-1} \\
&= \int_{\mathrm{S}^{N_\nu-1}} \int_0^R \frac{1}{r^{N_\nu}} \prod_{i=1}^{N_\nu} ||J_i(r,\mathcal{U})||\; r^{N_\nu-1} dr d\mathrm{S}^{N_\nu-1} \\
&= \int_{\mathrm{S}^{N_\nu-1}} \int_0^R \frac{r^k}{r^{N_\nu}} \prod_{i=k+1}^{N_\nu} ||J_i(r,\Psi)||\; r^{N_\nu-1} dr d\mathrm{S}^{N_\nu-1} \qquad \text{(where } \Psi \text{ is related to } \mathcal{U} \text{ as above)} \\
&= \int_{\mathrm{S}^{N_\nu-1}} \int_0^R \prod_{i=k+1}^{N_\nu} |a_i(r,\Psi)|\; r^{k-1} dr d\mathrm{S}^{N_\nu-1} \\
&= \int_{\mathrm{S}^{N_\nu-1}} \int_0^R \prod_{i=k+1}^{N_\nu} \frac{|\sin \sqrt{\lambda_i} r|}{\sqrt{\lambda_i}}\; r^{k-1} dr d\mathrm{S}^{N_\nu-1}
\end{aligned}
$$

In order to agree with [2], [1], and various other developments of the volume form, this integration over the fiber $\mathcal{U} \in \mathfrak{m}$ such that $||\mathcal{U}||_{\mathbf{Gr}} = 1$ and $\mathcal{U}$ is related to Ψ must amount to the following.

$$c_\nu \int_{\mathcal{R}_1} \int_0^R \prod_{i=k+1}^{N_\nu} |\sin\sqrt{\lambda_i} r| \; r^{k-1} dr d\mathrm{S}^{k-1}$$

(where c_ν is discussed below and $\mathcal{R}_1 := \{\pi/2 \geq \theta_1 \geq \cdots \geq \theta_k \geq 0\} \bigcap \{\sum \theta_i^2 = 1\}$)

$$= c_\nu \int_{\mathcal{R}_1} \int_0^R \prod_{i=1}^{k} (\sin r\theta_i)^{\nu(n-2k)} \prod_{i=1}^{k} (\sin 2r\theta_i)^{\nu-1} \prod_{1 \leq i < j \leq k} |\sin r(\theta_i + \theta_j) \sin r(\theta_i - \theta_j)|^\nu r^{k-1} dr d\mathrm{S}^{k-1}$$

$$= c_\nu \int_{\mathcal{R}_2} \prod_{i=1}^{k} (\sin \theta_i)^{\nu(n-2k)} \prod_{i=1}^{k} (\sin \theta_i)^{\nu-1} \prod_{1 \leq i < j \leq k} |\sin(\theta_i + \theta_j) \sin(\theta_i - \theta_j)|^\nu \; |d\Theta|$$

(where $|d\Theta| := \bigwedge_i d\theta_i$ and $\mathcal{R}_2 := \{\pi/2 \geq \theta_1 \geq \cdots \geq \theta_k \geq 0\} \bigcap \{\sum \theta_i^2 \leq R^2\}$)

$$= c_\nu \int_{\mathcal{R}_2} \prod_{i=1}^{k} (\sin \theta_i)^{\nu(n-2k)} \prod_{i=1}^{k} (\sin 2\theta_i)^{\nu-1} \prod_{1 \leq i < j \leq k} (\sin(\theta_i + \theta_j) \sin(\theta_i - \theta_j))^\nu \; |d\Theta|$$

$$= c_\nu \int_{\mathcal{R}_2} \prod_{i=1}^{k} (\sin \theta_i)^{\nu(n-2k)} \prod_{i=1}^{k} (\sin 2\theta_i)^{\nu-1} \prod_{1 \leq i < j \leq k} (\sin^2 \theta_i - \sin^2 \theta_j)^\nu \; |d\Theta| \quad (1)$$

(see Appendix 5.3)

$$= c_\nu \int_{\mathcal{R}_2} \prod_{i=1}^{k} (\sin \theta_i)^{\nu(n-2k)} \prod_{i=1}^{k} (\sin 2\theta_i)^{\nu-1} \prod_{1 \leq i < j \leq k} \frac{1}{2^\nu} (\cos 2\theta_j - \cos 2\theta_i)^\nu \; |d\Theta| \quad (2).$$

The region $\mathcal{R}_2$ is illustrated in the following figure where $k = 3$.

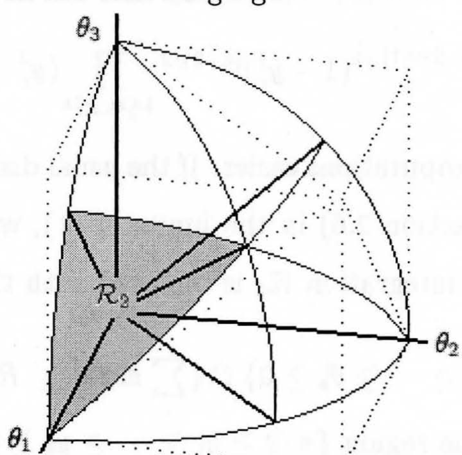

Figure 8: The Regions $\{\pi/2 \geq \theta_{\sigma_1} \geq \cdots \geq \theta_{\sigma_3} \geq 0\} \bigcap \{\sum_i \theta_i^2 \leq R^2\}$

The equivalence of (1) and (2) is shown in Appendix 5.3. (1) appears in [2] pg. 2453, and (2) appears in [1] pg. 19. For a Riemannian Manifold M^m let $|\mathrm{M}|$ denote the m -dimensional volume $\mathrm{vol}_m\,\mathrm{M}$. The conditions $\nabla_{\dot{C}} J_i(0) = \mathrm{e}_i$ were imposed to give a canonical volume form (see Appendix 5.2), meaning

$$|\mathbf{Gr}(\mathbb{R})| = \frac{|\mathbf{O}_n|}{|\mathbf{O}_k||\mathbf{O}_{n-k}|} = \begin{cases} \dfrac{1}{2}\dfrac{|\mathrm{S}^{n-1}(\mathbb{R})||\mathrm{S}^{n-2}(\mathbb{R})|\cdots|\mathrm{S}^{n-k}(\mathbb{R})|}{|\mathrm{S}^{k-1}(\mathbb{R})||\mathrm{S}^{k-2}(\mathbb{R})|\cdots|\mathrm{S}^{1}(\mathbb{R})|} & , \ k>1 \\ \dfrac{1}{2}|\mathrm{S}^{n-1}(\mathbb{R})| & , \ k=1 \end{cases}$$

$$\text{and } |\mathbf{Gr}(\mathbb{C})| = \frac{|\mathbf{U}_n|}{|\mathbf{U}_k||\mathbf{U}_{n-k}|} = \begin{cases} \dfrac{1}{2\pi}\dfrac{|\mathrm{S}^{2n-2}(\mathbb{R})||\mathrm{S}^{2n-4}(\mathbb{R})|\cdots|\mathrm{S}^{2n-2k}(\mathbb{R})|}{|\mathrm{S}^{2k-2}(\mathbb{R})||\mathrm{S}^{2k-4}(\mathbb{R})|\cdots|\mathrm{S}^{2}(\mathbb{R})|} & , \ k>1 \\ \dfrac{1}{2\pi}|\mathrm{S}^{2n-2}(\mathbb{R})| & , \ k=1 \end{cases}$$

where $|\mathrm{S}^m(\mathbb{R})|$ is the usual surface area of the m -dimensional sphere. This means that

$$c_\nu = |\mathbf{Gr}(\mathbb{F})| \Big/ \int_0^{\pi/2}\int_0^{\theta_1}\cdots\int_0^{\theta_{k-2}}\int_0^{\theta_{k-1}} \prod_{i=k+1}^{N_\nu} |\sin\sqrt{\lambda_i} r| \; d\theta_k d\theta_{k-1}\cdots d\theta_1 .$$

Using the change of variable $\sin\theta_i \longrightarrow y_i$ the volume element in (1) can be written

$$\begin{aligned} & c_\nu \prod_{i=1}^{k}(\sin\theta_i)^{\nu(n-2k)} \prod_{i=1}^{k}(\sin 2\theta_i)^{\nu-1} \prod_{1\le i<j\le k} (\sin^2\theta_i - \sin^2\theta_j)^\nu \, |d\Theta| \\ &= c_\nu\, 2^{k(\nu-1)} \prod_{i=1}^{k} y_i^{\nu(n-2k+1)-1}(1-y_i^2)^{(\nu-2)/2} \prod_{1\le i<j\le k} (y_i^2 - y_j^2)^\nu \bigwedge_{i=1}^{k} dy_i . \end{aligned}$$

This makes makes numerical computations easier. If the usual distance function is replaced by the chordal distance (see Section 2.6) in the limits of (1), which would become valid for small R, then the region of integration $\mathcal{R}_2$ is replaced with the region

$$\{1 \ge \theta_1 \ge \cdots \ge \theta_k \ge 0\} \cap \{\sum_i \sin^2\theta_i \le R^2\}$$

which in terms of y_i is the region $\{\pi/2 \ge y_1 \ge \cdots \ge y_k \ge 0\} \cap \{\sum_i y_i^2 \le R^2\}$.

The value of c_ν can be relatively easily computed in terms of y_i as

$$c_\nu = |\mathbf{Gr}(\mathbb{F})| \Big/ \int_0^1\int_0^{y_1}\cdots\int_0^{y_{k-1}} 2^{k(\nu-1)} \prod_{i=1}^{k} \frac{y_i^{\nu(n-2k+1)-1}}{(1-y_i^2)^{(2-\nu)/2}} \prod_{1\le i<j\le k} (y_i^2-y_j^2)^\nu dy_k dy_{k-1}\cdots dy_1 .$$

In the complex case the integrand is just a polynomial in y_i. Some values of c_ν for various k and n are listed in Appendix 5.4.

Another numerically useful observation is that because of symmetries in the complex case the integration in (1) may be made over the entire ball $\{\sum_i \theta_i^2 \leq R^2\}$ by introducing a factor of $\frac{1}{k!2^k}$. The factor $\frac{1}{2^k}$ corresponds to the 2^k regions $\{\pm\theta_1 \leq 0, \ldots, \pm\theta_k \leq 0\}$ and the factor $\frac{1}{k!}$ corresponds to the $k!$ regions $\{\pi/2 \geq \theta_{\sigma_1} \geq \cdots \geq \theta_{\sigma_k} \geq 0\}$. This can also be done in the real case if the absolute values are maintained in the integral. In the case where $\mathbb{F} = \mathbb{C}$, $k = n/2$, and chordal distance is used, the integral for $\operatorname{vol} B_c(R)$ takes the simplified form

$$\operatorname{vol} B_c(R) = \frac{c_2}{k!} \int_{\{\sum_{i=1}^k y_i^2 \leq R^2\}} \prod_{i=1}^k y_i \prod_{1 \leq i < j \leq k} (y_i^2 - y_j^2)^2 \bigwedge_{i=1}^k dy_i.$$

3.9 Estimates and Bounds for $\operatorname{vol} B(R)$

For anything more than small values of n and k trying to directly compute volumes is impractical. The volume of $B(R)$ in any Riemannian Manifold can be bounded from above and below based on the sectional and Ricci curvatures. [10] Theorem 3.101 states the following.

Theorem 3.5. *Let (M, g) be a complete Riemannian Manifold and $B(R)$ be a geodesic ball centered at p that does not meet the cut locus of p. Let $V^\ell(R)$ denote the volume of the ball radius R in the manifold of constant curvature ℓ and dimension $m = \dim M$. Then,*

i) **(Bishop)** *If there is a constant such that*

$$\operatorname{Ric}(X) \geq a g(X, X) \text{ then } \operatorname{vol} B(R) \leq V^a(R).$$

ii) **(Günther)** *If there is a constant b such that*

$$K(X,Y) \leq b \text{ then } \operatorname{vol} B(R) \geq V^b(R).$$

The bound involving the Ricci curvature is better because it involves an average of sectional curvatures where as the bound involving the maximum sectional curvature neglects the fact that the other sectional curvatures may be much smaller. That the upper bound on $\operatorname{vol} B(R)$ guaranteed by this theorem is the better bound will prove to be fortunate later because of the implications this has on the number of spheres that can be packed into **Gr**. It is well known that in a manifold of constant positive curvature ℓ the volume of the geodesic ball radius R is given by

$$V^{\ell}(R) = \int_{\mathbb{S}^{m-1}} \int_0^R \left(\frac{\sin\sqrt{\ell}r}{\sqrt{\ell}}\right)^{m-1} dr d\mathbb{S}^{m-1} = |\mathbb{S}^{m-1}| \int_0^R \left(\frac{\sin\sqrt{\ell}r}{\sqrt{\ell}}\right)^{m-1} dr.$$

The proof involves diagonalizing the curvature tensor. [12] applies these theorems to $\mathbf{Gr}(\mathbb{C})$ to get bounds on $B(R)$ but uses only that the sectional curvature is non - negative and takes $a = 0$. This theorem applied to the results for **Gr** in Section 3.6 together with $a < b \Longrightarrow V(R)^a > V(R)^b$ gives,

$$V^{b_\nu}(R) \leq \operatorname{vol}_{\mathbf{Gr}(\mathbb{F})} B(R) \leq V^{a_\nu}(R) < V^0(R).$$

There is also an expansion formula in terms of R and the scalar curvature for the geodesic ball $B_p(R)$ on any Riemannian Manifold $(\mathbf{M}^{n_1}, g)$ having dimension n_1. [10] Theorem 3.98 states, accounting for the difference in the definition of Scal_p used by [10], that:

Theorem 3.6. $\operatorname{vol}_M B_p(R) = R^{n_1} \operatorname{vol}_{\mathbb{R}^{n_1}} B(1) \left(1 - \dfrac{(n_1)(n_1-1)\mathrm{Scal}_p}{6(n_1+2)} R^2 + o(R^2)\right).$

In terms of **Gr**, if $B(R)$ is a ball centered at any point this formula becomes

$$\operatorname{vol} B(R) = R^{N_\nu} \left(\frac{\pi^{N_\nu/2}}{\Gamma(N_\nu/2+1)}\right) \left(1 - \frac{N_\nu(N_\nu-1)\mathrm{Scal}_\nu}{6(N_\nu+2)} R^2 + o(R^2)\right)$$

$$\text{where} \qquad \mathrm{Scal}_\nu = \frac{\nu(n-k-1) + (\nu-1)4}{N_\nu - 1}.$$

Let $\widetilde{V}_\nu(R) = R^{N_\nu}\left(\frac{\pi^{N_\nu/2}}{\Gamma(N_\nu/2+1)}\right)\left(1 - \frac{N_\nu(N_\nu-1)\mathrm{Scal}_\nu}{6(N_\nu+2)}R^2\right)$. The following graphs comparing $V^{a_1}(R), V^{b_1}(R)$, and $\widetilde{V}_1(R)$ in $\mathbf{Gr}_{3,6}(\mathbb{R})$ illustrate that $V^{a_\nu}(R)$ and $V^{b_\nu}(R)$ are good approximations when R is small and that $\widetilde{V}_\nu(R)$ is only a good approximation if R is not too small and not too large.

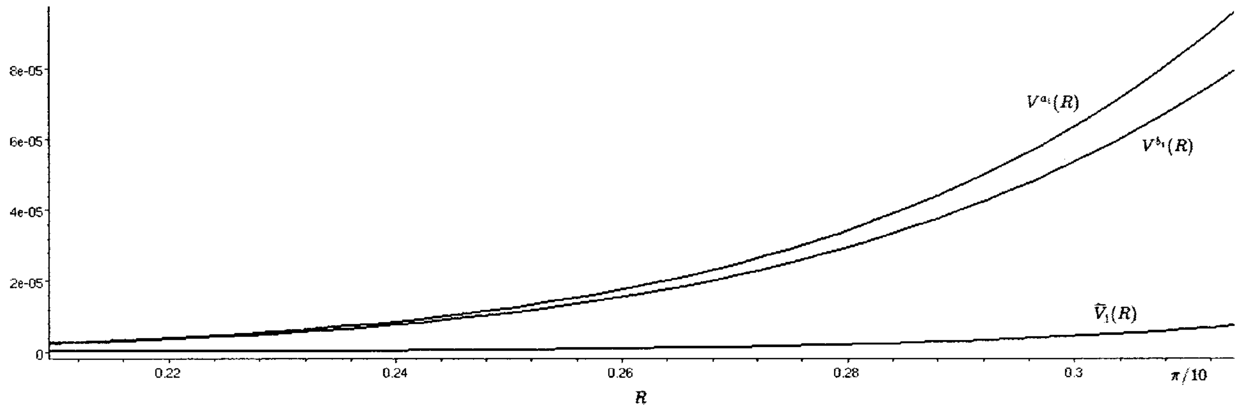

Figure 9: Comparison Between the Bounds on Volumes of Geodesic Balls $V^{a_1}(R), V^{b_1}(R)$, and the Approximation $\widetilde{V}_1(R)$

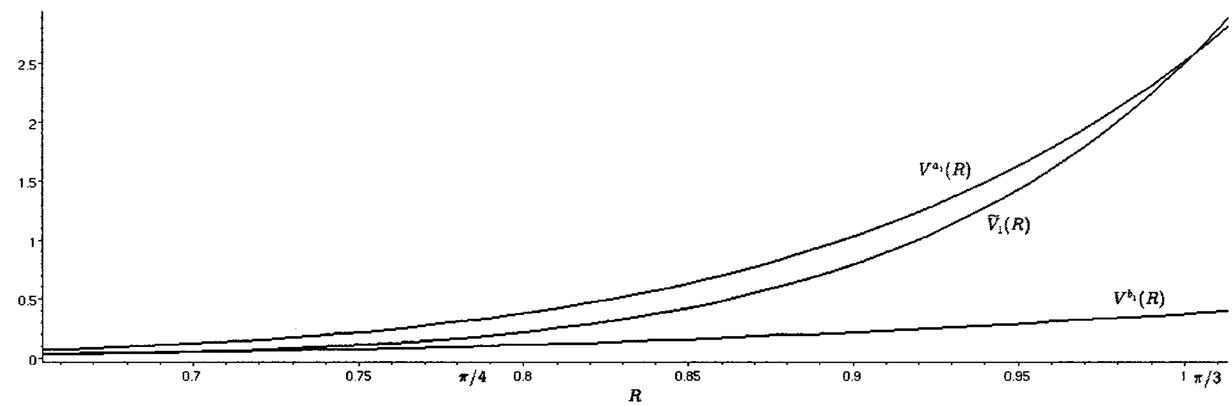

Figure 10: The Range of Validity of $\widetilde{V}_1(R)$

4 Packings

A **code** $\mathcal{C}$ in a manifold M is any finite set of points in M. For points to be well spaced out it is desirable that the **minimum distance** $d_{\min} := \min_{c_i, c_j \in \mathcal{C}} d(c_i, c_j)$ is as large as possible. Two famous bounds on arbitrary Riemannian Manifolds without boundary that relate vol M, vol $B(R)$, $d_{\min}$ and $|\mathcal{C}|$ are the **Hamming** and **Gilbert-Varshamov** bounds (see Appendix 5.5). It is clear that placing a code $\mathcal{C}$ having minimum distance $d_{\min}$ on a manifold M is equivalent to successfully packing $|\mathcal{C}|$ spheres of radius $d_{\min}/2$ in M. This is the basis of the Hamming upper bound on $|\mathcal{C}|$. The goal of this section is to place codes on **Gr** with a large minimum distance when compared to the Gilbert-Varshamov lower bound on $|\mathcal{C}|$.

Suppose arbitrary representatives $\tilde{c_1}, \tilde{c_2} \in \mathbf{O}_n$ of points $[\tilde{c_1}], [\tilde{c_2}] \in$ **Gr** are given. The essential component of the packing algorithm to be described will be finding a $\mathcal{B} \in \mathfrak{m}$ such that $\tilde{c_1} \mathrm{e}^{||\Theta||\mathcal{B}} \in [\tilde{c_2}]$. It has been shown in Section 2.2 that there exist $k_1, k_2 \in$ K and AD representatives $c_1 = \tilde{c_1} k_1$ and $c_2 = \tilde{c_2} k_2$ that have the relationship

$$\begin{aligned} c_1 \mathrm{e}^{||\Theta||\Psi} &= c_2 \\ \tilde{c_1} k_1 \mathrm{e}^{||\Theta||\Psi} &= \tilde{c_2} k_2 \\ \tilde{c_1} k_1 \mathrm{e}^{||\Theta||\Psi} k_1^{\mathcal{T}} &= \tilde{c_2} k_2 k_1^{\mathcal{T}} \\ \tilde{c_1} \mathrm{e}^{||\Theta|| k_1 \Psi k_1^{\mathcal{T}}} &= \tilde{c_2} k_2 k_1^{\mathcal{T}} \in [\tilde{c_2}] \end{aligned}$$

In Section 2.2 k_1 is written $\begin{pmatrix} Q_1 & 0 \\ 0 & Q_1^* \end{pmatrix}$. It is easy to check that $\mathcal{B} := k_1 \Psi k_1^{\mathcal{T}} \in \mathfrak{m}$ and that $||\mathcal{B}||_{\mathbf{Gr}} = 1$. Rather than try to maximize the non-differentiable functional $d_{\min}(\mathcal{C})$ over all possible $\mathcal{C}$, the following algorithm assumes that each point c_i in a code $\mathcal{C} = \{c_\alpha\}_{\alpha=1}^{|\mathcal{C}|}$ experiences an inverse square repulsive force $c_i F_{ij} = -\frac{1}{||\Theta||^2} c_i \mathcal{B}_{ij}$ from every other point c_j, $j \neq i$. It is hoped that if the points are allowed to move under these repulsive forces they will spread out and produce a large $d_{\min}$. It is computationally sensible to ignore the effect of points that are too far apart, say beyond a distance of R. During each iteration each point c_i is allowed to move a small distance in the direction of $c_i \sum_{j \neq i} F_{ij}$, that is, in each iteration $c_i \mapsto c_i e^{\varepsilon \Sigma_{j \neq i} F_{ij}}$ where $\varepsilon > 0$. ε, R, and the form of the repulsive forces may be varied to encourage convergence.

Algorithm 4.1

Step 1: Generate an approximately random code $\mathcal{C} = \{c_\alpha\}_{\alpha=1}^{|\mathcal{C}|}$ of orthogonal matrices representing points in **Gr** and pick an $\varepsilon > 0$.

Step 2: For each point c_i, if $d(c_i, c_j) < R$, use Algorithm 2.1 (found in Section 2.2) to find k_{ij}, Θ_{ij}, and Ψ_{ij} such that

$$c_i e^{||\Theta_{ij}|| k_{ij} \Psi_{ij} k_{ij}^{\mathcal{T}}} \in [c_j] \qquad (j \neq i).$$

Let $\mathcal{B}_{ij} = k_{ij} \Psi_{ij} k_{ij}^{\mathcal{T}}$ and $F_{ij} = -\frac{1}{||\Theta_{ij}||^2} \mathcal{B}_{ij}$.

Step 3: Send each point c_i to the new point $c_i e^{\varepsilon \Sigma_{j \neq i} F_{ij}}$.

Repeat.

Figure 11 compares a code consisting of 5 randomly placed points on $\mathbf{Gr}_{1,3}(\mathbb{R})$ and the initial forces on each with the perturbed code and new forces after 10 iterations.

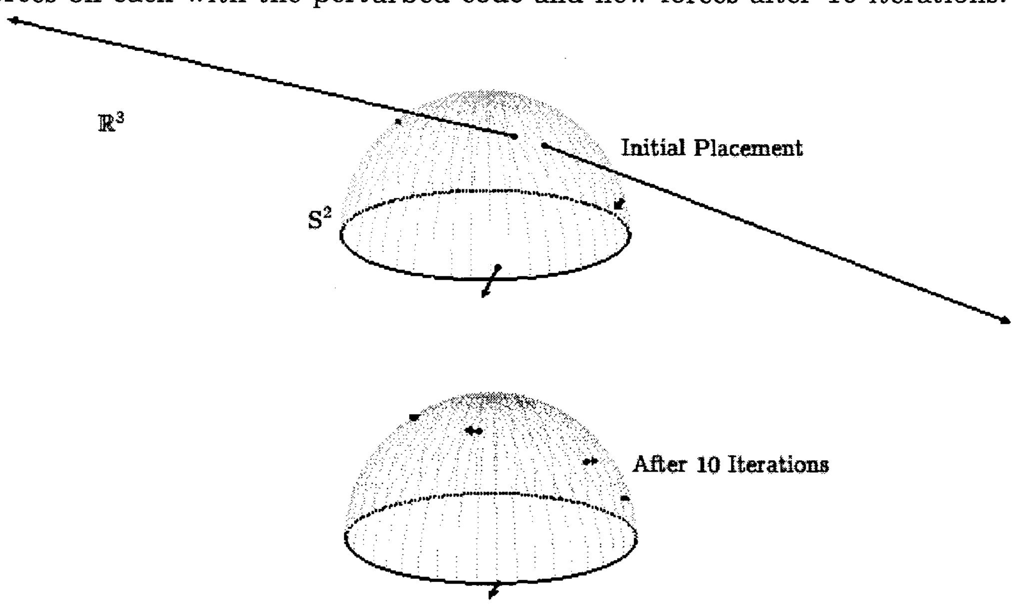

Figure 11: Points on $\mathbf{Gr}_{1,3}(\mathbb{R})$ and Repulsive Forces Between Them

Smaller values of ε should be used when $|\mathcal{C}|$ is large because in these cases the forces involved are potentially larger but the distances between points are smaller so it is important that points are not moved too far in a given iteration. The minimum distance may decrease during an iteration, this reflects the fact that in order to move the points so that the entire code is well spaced out it may be necessary for points to 'float' close to one another. It is interesting to witness the effect of setting R to a value that $d_{\min}$ can be made larger than. Algorithm 3.1 then quickly produces a packing with $d_{\min} \geq R$. Gradually increasing R is comparable to blowing up the radii of $|\mathcal{C}|$ frictionless balls within the confined space of **Gr**. Figure 12 plots $d_{\min}$ verses the number of iterations for a code consisting of 64 points on $\mathbf{Gr}_{3,6}(\mathbb{R})$ using $\varepsilon = 0.005$. $d_{\min}$ for the initial approximately random placement was 0.579834 and the best $d_{\min}$ achieved after 1000 iterations was 1.174566.

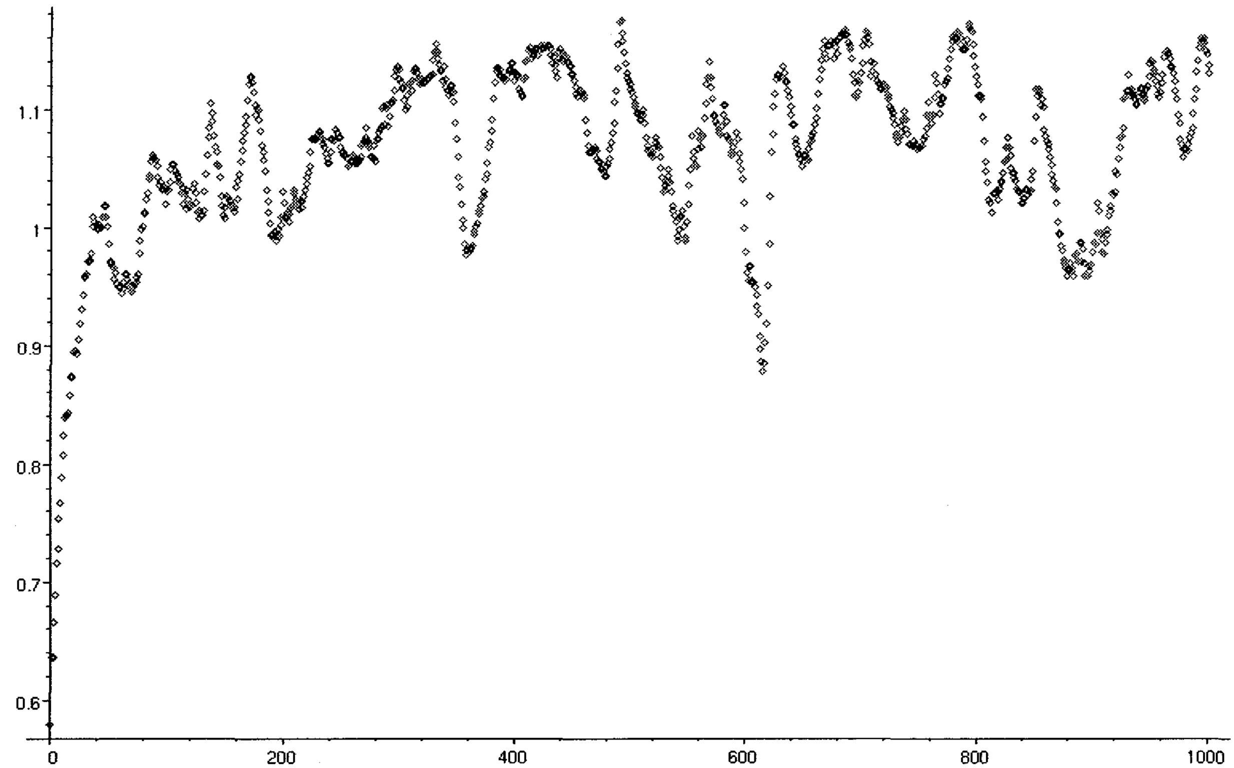

Figure 12: $d_{\min}$ Vs. Iteration for 64 Points on $\mathbf{Gr}_{3,6}(\mathbb{R})$

The Gilbert-Varshamov and Hamming bound guarantees that for a given δ there exists a code $\mathcal{C}$ on $\mathbf{M}$ having $d_{\min} \leq \delta$ satisfying $\frac{\operatorname{vol}\mathbf{M}}{\operatorname{vol}B(\delta)} < |\mathcal{C}|$. Since for any δ, $\frac{\operatorname{vol}\mathbf{M}}{V^{a_\nu}(\delta)} \leq \frac{\operatorname{vol}\mathbf{M}}{\operatorname{vol}B(\delta)}$, one way to test if a packing $|\mathcal{C}|$ having a minimum distance of $d_{\min}$ is a relatively good one is to check if $\frac{\operatorname{vol}\mathbf{M}}{V^{a_\nu}(d_{\min})} < |\mathcal{C}|$. Doing this for the current example where $d_{\min} = 1.174566$ and $|\mathcal{C}| = 64$ gives $\frac{\operatorname{vol}\mathbf{M}}{V^{a_1}(d_{\min})} = 10.9$. This lends some evidence to claim that Algorithm 4.1 produces good packings. Of course $|\mathcal{C}| \leq \frac{\operatorname{vol}\mathbf{M}}{\operatorname{vol}B(d_{\min}/2)} \leq \frac{\operatorname{vol}\mathbf{M}}{V^{b_1}(d_{\min}/2)} = 7931.5$. When Algorithm 4.1 is tested on small packings in $\mathbf{Gr}_{1,2}(\mathbb{R})$, $\mathbf{Gr}_{1,3}(\mathbb{R})$, $\mathbf{Gr}_{1,4}(\mathbb{R})$, and $\mathbf{Gr}_{2,4}(\mathbb{R})$ the graph d_{min} Vs. Iteration resembles a fly bouncing along a ceiling. This thesis does not claim that this ceiling is the global maximum for $d_{\min}$.

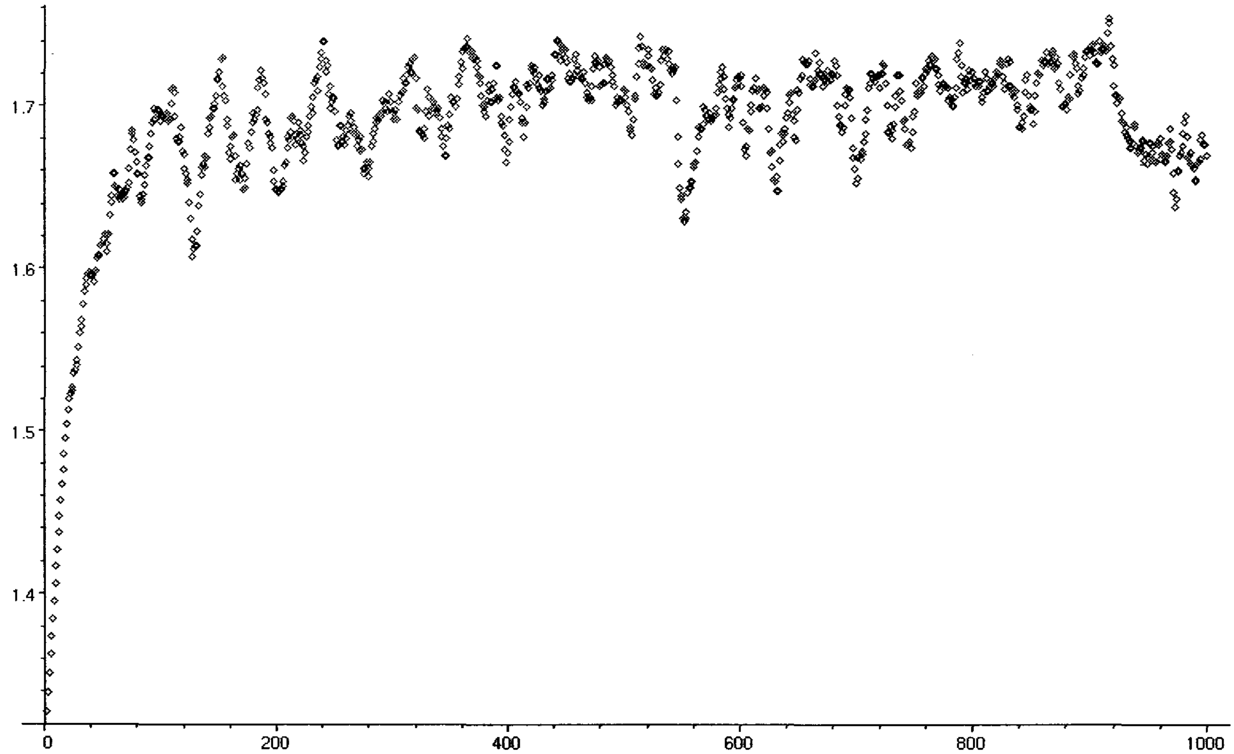

Figure 13: $d_{\min}$ Vs. Iteration for 400 Points on $\mathbf{Gr}_{5,10}(\mathbb{C})$

5 Appendix

5.1 $[\mathfrak{m},\mathfrak{m}] \subset \mathfrak{k}, [\mathfrak{m},\mathfrak{k}] \subset \mathfrak{m}$, and $[\mathfrak{k},\mathfrak{k}] \subset \mathfrak{k}$

Let $\mathcal{B}_1, \mathcal{B}_2 \in \mathfrak{m}$, then

$$[\mathcal{B}_1,\mathcal{B}_2] = \begin{pmatrix} 0 & -B_1^T \\ B_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -B_2^T \\ B_2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -B_2^T \\ B_2 & 0 \end{pmatrix}\begin{pmatrix} 0 & -B_1^T \\ B_1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -B_1^TB_2 & 0 \\ 0 & -B_1B_2^T \end{pmatrix} - \begin{pmatrix} -B_2^TB_1 & 0 \\ 0 & -B_2B_1^T \end{pmatrix}$$

$$= \begin{pmatrix} B_2^TB_1 - B_1^TB_2 & 0 \\ 0 & B_2B_1^T - B_1B_2^T \end{pmatrix} \in \mathfrak{k}.$$

Let $\mathcal{C} \in \mathfrak{k}$, then

$$[\mathcal{C},\mathcal{B}_1] = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}\begin{pmatrix} 0 & -B_1^T \\ B_1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -B_1^T \\ B_1 & 0 \end{pmatrix}\begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -C_1B_1^T \\ C_2B_1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -B_1^TC_2 \\ B_1C_1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & B_1^TC_2 - C_1B_1^T \\ C_2B_1 - B_1C_1 & 0 \end{pmatrix} \in \mathfrak{m}.$$

Let $\mathcal{A}, \mathcal{B} \in \mathfrak{k}$, then

$$[\mathcal{A},\mathcal{B}] = \begin{pmatrix} A_1B_1 & 0 \\ 0 & A_2B_2 \end{pmatrix} - \begin{pmatrix} B_1A_1 & 0 \\ 0 & B_2A_2 \end{pmatrix} \in \mathfrak{k}.$$

5.2 The Volumes of $\mathbf{S}^m(\mathbb{R}), \mathbf{O}_m, \mathbf{S}^m(\mathbb{C})$, and $\mathbf{U}_m$

It is well known, (see [9]), that the surface area of the m -dimensional unit sphere $\mathrm{S}^m(\mathbb{R}) \subset \mathbb{R}^{m+1}$ denoted $|\mathrm{S}^m(\mathbb{R})|$ where $m \geq 1$ is given by

$$|\mathrm{S}^m(\mathbb{R})| = (m+1)\frac{\pi^{(m+1)/2}}{\Gamma((m+1)/2+1)}.$$

The sphere $\mathrm{S}^{m-1}(\mathbb{R})$ is the set of all vectors of length one in $\mathbb{R}^m$ which is the definition of $\mathrm{St}_{1,m}(\mathbb{R})$. This shows that $\mathrm{S}^{m-1}(\mathbb{R}) \simeq \mathbf{O}_m/\mathbf{O}_{m-1}$. Hence the relationship

$$|\mathrm{S}^{m-1}(\mathbb{R})||\mathbf{O}_{m-1}| = |\mathbf{O}_m|.$$

Applying this relation recursively gives that

$$|\mathbf{O}_m| = |\mathrm{S}^{m-1}(\mathbb{R})||\mathrm{S}^{m-2}(\mathbb{R})|\cdots|\mathrm{S}^2(\mathbb{R})||\mathbf{O}_2| = |\mathrm{S}^{m-1}(\mathbb{R})||\mathrm{S}^{m-2}(\mathbb{R})|\cdots|\mathrm{S}^2(\mathbb{R})||\mathrm{S}^1(\mathbb{R})|\cdot 2.$$

In light of this equation and the fact that $(m+1)\dfrac{\pi^{(m+1)/2}}{\Gamma((m+1)/2+1)}$ evaluated at $m=0$ gives the value 2 it is understandable that some authors take the convention $|\mathrm{S}^0(\mathbb{R})| = 2$. This corresponds to the discrete or 0 -dimensional measure giving a value of one to each point. An excellent source that develops these relationships in the real case is [16]. For reference, the surface area of the sphere radius R is known to be $|R\mathrm{S}^m(\mathbb{R})| = R^m|\mathrm{S}^m(\mathbb{R})|$.

In the complex case

$$|\mathrm{S}^m(\mathbb{C})| = 2(m+1)\frac{\pi^{m+1}}{\Gamma(m+2)}.$$

The sphere $\mathrm{S}^{m-1}(\mathbb{C})$ is the set of all vectors of length one in $\mathbb{C}^m$ which is exactly $\mathbf{St}_{1,m}(\mathbb{C})$. So that an argument entirely similar to the real case shows

$$\begin{aligned}|\mathbf{U}_m| &= |\mathrm{S}^{m-1}(\mathbb{C})||\mathrm{S}^{m-2}(\mathbb{C})|\cdots|\mathrm{S}^1(\mathbb{C})||\mathbf{U}_1| = |\mathrm{S}^{m-1}(\mathbb{C})||\mathrm{S}^{m-2}(\mathbb{C})|\cdots|\mathrm{S}^1(\mathbb{C})|\cdot 2\pi \\ &= |\mathrm{S}^{2m-2}(\mathbb{R})||\mathrm{S}^{2m-4}(\mathbb{R})|\cdots|\mathrm{S}^2(\mathbb{R})|\cdot 2\pi.\end{aligned}$$

The surface area of the complex sphere radius R is $|R\mathrm{S}^m(\mathbb{C})| = |R\mathrm{S}^{2m}(\mathbb{R})| = R^{2m}|\mathrm{S}^{2m}(\mathbb{R})|$.

5.3 Trigonometric Identities for $\sin(\theta_i+\theta_j)\sin(\theta_i-\theta_j)$

The following calculation is useful in simplifying the volume form on **Gr**.

$$\begin{aligned}\sin(\theta_i+\theta_j)\sin(\theta_i-\theta_j) &= (\sin\theta_i\cos\theta_j+\sin\theta_j\cos\theta_i)(\sin\theta_i\cos\theta_j-\cos\theta_i\sin\theta_j) \\ &= \sin^2\theta_i\cos^2\theta_j-\cos^2\theta_i\sin^2\theta_j \\ &= (1-\cos^2\theta_i)\cos^2\theta_j-\cos^2\theta_i(1-\cos^2\theta_j) \\ &= \cos^2\theta_j-\cos^2\theta_i \quad (=\sin^2\theta_i-\sin^2\theta_j\ \geq 0 \text{ when } i<j) \\ &= \frac{1}{2}(\cos 2\theta_j+1)-\frac{1}{2}(\cos 2\theta_i+1) \\ &= \frac{1}{2}(\cos 2\theta_j-\cos 2\theta_i).\end{aligned}$$

5.4 Values of c_ν

The following tables lists values of c_ν for various k and n.

Table 1: Values of c_1

k\n	10	9	8	7	6	5	4	3	2
5	2	2π	4π	$2\pi^2$	$8\pi^2/3$	π^3	$16\pi^3/15$	$\pi^4/3$	$32\pi^4/105$
4			$4\pi^2$	$8\pi^3$	$8\pi^4$	$16\pi^5/3$	$8\pi^6/3$	$16\pi^7/15$	$16\pi^8/45$
3					$32\pi^4$	$32\pi^6$	$128\pi^7/3$	$32\pi^9/3$	$256\pi^{10}/45$
2							$64\pi^8$	$256\pi^{10}/3$	$128\pi^{12}/3$
1									$2048\pi^{12}/9$

Table 2: Values of c_2

k\n	10	9	8	7	6	5	4	3	2
5	π	π^2	$\pi^3/2$	$\pi^4/6$	$\pi^5/24$	$\pi^6/120$	$\pi^7/720$	$\pi^8/5040$	$\pi^9/40320$
4			π^4	$\pi^6/2$	$\pi^8/12$	$\pi^{10}/144$	$\pi^{12}/2880$	$\pi^{14}/86400$	$\pi^{16}/3628800$
3					$\pi^9/4$	$\pi^{12}/24$	$\pi^{15}/576$	$\pi^{18}/34560$	$\pi^{21}/4147200$
2							$\pi^{16}/144$	$\pi^{20}/3456$	$\pi^{24}/414720$
1									$\pi^{25}/82944$

5.5 Basic Packing Bounds

This section states the Hamming and Gilbert-Varshamov packing bounds.

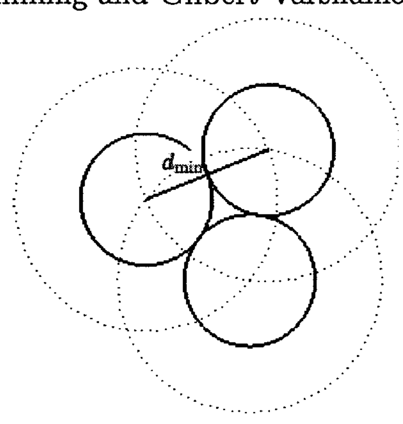

Figure 14: The Hamming Upper Bound

Theorem 5.1. (Hamming) *Let M be a Riemannian Manifold without boundary. For any code C having minimum distance $d_{\min}$,*

$$|C| \leq \frac{\operatorname{vol} M}{\operatorname{vol} B(d_{\min}/2)}.$$

The Gilbert-Varshamov bound gives a lower bound on possible $|C|$ for a given $d_{\min}$.

Theorem 5.2. (Gilbert – Varshamov) *In a Riemannian Manifold M without boundary if $\delta < i_M$ is given then there exists a code C on M having $d_{\min} \leq \delta$ satisfying*

$$|C| > \frac{\operatorname{vol} M}{\operatorname{vol} B(\delta)}$$

Sketch of proof. Given a δ if there exists an m with $m \cdot \operatorname{vol} B(\delta) \leq |\mathrm{M}|$ then there exists a code C on $\mathbf{M}$ with $|C| = m+1$ having a minimum distance $d_{\min} \geq \delta$. In other words if $m = \left\lfloor \frac{\operatorname{vol} \mathrm{M}}{\operatorname{vol} B(\delta)} \right\rfloor$ then there is a code C on $\mathbf{M}$ having $|C| = m+1$. Since $\left\lfloor \frac{\operatorname{vol} \mathrm{M}}{\operatorname{vol} B(\delta)} \right\rfloor \leq \frac{\operatorname{vol} \mathrm{M}}{\operatorname{vol} B(\delta)} < \left\lfloor \frac{\operatorname{vol} \mathrm{M}}{\operatorname{vol} B(\delta)} \right\rfloor + 1$ there exists a code C on $\mathbf{M}$ with $|C| > \frac{\operatorname{vol} \mathrm{M}}{\operatorname{vol} B(\delta)}$ and $d_{\min} \geq \delta$. □

6 Bibliography

[1] M. Adler, P. Van Moerbeke, *Integrals Over Grassmannians and Random Permutations*, Advances in Mathematics, Vol. 181, No. 1, pp. 190 - 249, Elsevier Science, 2004.

[2] Alexander Barg, Dmitry Yu. Nogin, *Bounds on The Packings of Spheres in the Grassmannian Manifold*, IEEE Transactions on Information Theory, Vol. 48, No. 9, pp. 2450 - 2454, published electronically, 2002.

[3] Yung-Chow Wong, *Differential Geometry of Grassmannian Manifolds*, Proceedings of the National Academy of Sciences of the U.S.A., Vol. 57. pp. 589 - 594, Washington DC, U.S.A, 1967.

[4] Yung-Chow Wong, *Sectional Curvatures of Grassmannian Manifolds*, Proceedings of the National Academy of Sciences of the U.S.A., Vol. 60. pp. 75 - 79, Washington DC, U.S.A., 1968.

[5] Manfredo Perdiãgo Do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[6] J. Cheeger, D. Ebin, *Comparison Theorems in Riemannian Geometry*, North Holland Co., Amsterdam, 1975.

[7] J. H. Conway, R. H. Haradin, N. J. A. Sloane, *Packing Lines, Planes, etc.: Packings in Grassmannian Spaces*, Experimental Mathematics, Vol. 5, pp. 139 - 159, 1990.

[8] Alan Edelman, Tom as A. Arias, and Steven T. Smith, *The Geometry of Algorithms with Orthogonality Constraints*, SIAM Journal on Matrix Analysis and Applications, Vol. 20, No.2, pp. 303 -353, Society for Industrial and Applied Mathematics, published electronically, www.siam.org/journals, 1998.

[9] Lawrence C. Evans, *Partial Differential Equations*, American Mathematical Society, U.S.A., 2002.

[10] Sylvestre Gallot, *Riemannian Geometry* , Springer-Verlag Berlin Heidelberg, Germany, 1987.

[11] Sigurdur Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press,

New York, U.S.A., 1962.

[12] Oliver Henkel, *Sphere Packing Bounds in the Grassmann and Stiefel Manifolds*, IEEE Transactions on Information Theory, Vol. 51, No. 10, IEEE Transactions on Information Theory Society, published electronically, 2005.

[13] Leslie Hogben, *Handbook of Linear Algebra*, CRC Press, Florida, U.S.A., 2003.

[14] Serge Lang, *Fundamentals of Differential Geometry*, Springer-Verlag, New York, U.S.A., 1999.

[15] John M. Lee, *Introduction to Smooth Manifolds*, Springer Science and Business Media, New York, U.S.A., 2006.

[16] Andrew Markoe, *Analytic Tomography*, Cambridge University Press, U.K., 2006.

[17] Luis Santaló, *Integral Geometry and Geometric Probability*, Cambridge University Press, U.K., 2004.

[18] *Matrix Exponential*, University of Utah, Department of Mathematics, http://www.math.utah.edu/~gustafso/2250matrixexponential.pdf, Dec. 2007.

[19] L. Zheng , D. Tse *Communication on the Grassmann Manifold: A Geometric Approach to Noncoherent Multiple-Antenna Channel*, IEEE Transactions on Information Theory, Vol. 48, No. 2, IEEE Transactions on Information Theory Society, published electronically, 2002.