Brownian motion on homogeneous space and control systems via Riemannian submersion using mean curvature drift

by

Huang Ching-Peng

B.Sc., National Taiwan University, Taipei, Taiwan R.O.C., 2011

M.Sc., National Taiwan University, Taipei, Taiwan R.O.C., 2013

A dissertation submitted in partial fulfillment of the

requirements for the degree of Doctor of Philosophy

in the Division of Applied Mathematics at Brown University

PROVIDENCE, RHODE ISLAND

May 2022

This dissertation by Huang Ching-Peng is accepted in its present form
by the Division of Applied Mathematics as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date________________ ________________________________

Govind Menon, Ph.D., Advisor

Recommended to the Graduate Council

Date________________ ________________________________

Brendan Hassett, Ph.D., Advisor

Date________________ ________________________________

, Ph.D., Reader

Approved by the Graduate Council

Date________________ ________________________________

Andrew G. Campbell, Dean of the Graduate School

Vita

Huang Ching-Peng completed B.Sc. and M.Sc. degrees in mathematics at National Taiwan University in 2011 and 2013 respectively. He entered the Ph.D. program at Brown University's Division of Applied Mathematics in 2016.

Acknowledgments

In my opinion, gratitude is most adequate to conclude this project.

Needless to say, I would like to thank in no particular order my advisors Prof. Brendan Hassett and Prof Govind Menon along others who gave their opinions on my work, my various support groups, either regular or irregular, including those beyond *homo sapiens* who might not read (you know I mean you), and anyone who is reading this page. Our species thrives on altruism, and it is amazing.

Last, I am reminded to thank myself. You have done great.

Contents

Chapter Zero

Introduction

0.1 Motivation: Brownian motion, submersion, mean curvature flow

It all started with a simple question: how does one write down the Brownian motion on Grassmannians and the manifolds of positive definite matrices, denoted $\mathcal{S}_+(n,n)$, which are common spaces in applications, under certain Riemannian structures?

Before going into details of what these spaces are, we see why and to whom this question is of any importance. In probably the most general terms for Riemannian manifolds, we have the equivalence of three objects:

$$\text{Brownian motion} \Longleftrightarrow \text{Laplacian/diffusion} \Longleftrightarrow \text{metric tensor.}$$

To put it simply, Brownian motion is a natural way to studying Riemannian manifolds in the language of probability.

The spaces we are interested in, Grassmannians and positive definite matrices, are common objects in various areas. Detailed and practical formulae for Brownian motion on them would certainly be of great use, for example, when one considers randomized algorithms.

From the equivalence above, to describe Brownian motion, one may first prescribe Riemannian structure. The most common metric on real Grassmannians is the quotient structure from orthogonal groups. For $\mathcal{S}_+(n,n) \simeq GL(n,\mathbb{R})/O(n)$, there are more than one common structures. The one that elicits our interest is the Bures-Wasserstein metric, which is defined as the submersion from the Euclidean metric on the general linear groups seen as an open submanifold of Euclidean spaces. The metric has various importance: it induces the Wasserstein distance on Gaussian measures and has deep connection with optimal transport; furthermore, the curvature is positive, which can be useful for optimization algorithms.

In particular, these come from Riemannian submersions under Lie group quotients. For such geometry , it is tempting to guess Brownian motion on the quotient would simply be image of that on the total space. However, stochastic calculus often behaves against intuitions from standard calculus, and this results from the second derivatives. In this case, by a calculation for the Beltrami-Laplace operator, Brownian motion gains a drift term proportional to the mean curvature of fibers with respect to the submersion.

Moreover, the Lie group quotient structure gives simpler formulae for the drift term. For Grassmannians, the submersion is totally geodesic. The same can be said for homogeneous spaces with quotient metric in general. Conversely, if a homogeneous Riemannian space of the form G/H (i.e. a homogeneous space with Riemannian metric invariant under G, which is in fact the common case) satisfies certain conditions, the metric can be lifted to G and the quotient is a Riemannian submersion that Brownian motion on G/H is easily written down as the image of G-invariant Brownian motion on G. Examples include Poincaré's upper half plane and $\mathcal{S}_+(n,n)$ with $GL(n,\mathbb{R})$-invariant metric.

On the other hand, the Bures-Wasserstein geometry is not $GL(n,\mathbb{R})$-invariant, but the $O(n)$ action is isometric along each orbit; the former causes the drift term to be nonzero and the latter makes this term autonomous on $\mathcal{S}_+(n,n)$. The non-vanishing drift results in interesting consequences–the Brownian motion, which might seem symmetric for all directions as in Euclidean cases, is combined with a mean curvature flow when projected through the submersion.

0.2 A control system on positive definite matrices

We investigate an idea to make use of the drift term. On the model example of the Bures-Wasserstein geometry, first we prove that the vertical process, X, of Brownian motion W projects through π to a deterministic process P that is exactly the mean curvature drift whose derivative is $J(P)$. Define a new metric on $\mathcal{S}_+(n,n)$,

$$\operatorname{Tr}_R VW^* := \langle V, W\rangle_R := \operatorname{Tr} RVW^*,$$

for $R \in \mathcal{S}_+(n,n)$. We may still consider the Brownian motion W_R and orthogonal projection under the new metric to obtain modified drift term $J^R(P)dt$. By varying R along the path, we have a control system. It turns out this is equivalent to a Lie-theoretical invariant control system.

Whether this first model has practical uses is unclear. But it should be an initial standpoint for future ideas

0.3 Notations and abbreviations

- Matrices, including matrix groups, are over real numbers if not specified.
- The transpose of a real matrix A is denoted as A^*. Although throughout the paper, mostly we only consider real matrices, when extending to complex numbers, AA^* is positive definite but not AA^T.

- $\mathcal{S}_+(n,n)$ denotes the space of $n \times n$ positive definite matrices of full rank. $\mathcal{S}(n)$ is the space of all $n \times n$ (real) symmetric matrices, whereas $\mathcal{S}_+(k,n)$ is the locus of semidefinite matrices of rank k.

- The differential of a smooth map, ϕ, between smooth manifolds is written as ϕ_* to avoid confusion with stochastic differentials.

- For Itô differentials we use d while for Stratonovich differential we use đ.

- *BM* is short for Brownian motions. *MCF* stands for mean curvature flow, sometimes scaled or reversed if clear from the context.

- $\vec{H}(x)$ (sometimes without the vector arrow when not emphasizing it is a vector) denotes the mean curvature vector at a point x. In most cases through out the dissertation, $\vec{H}$ is a vector field.

Chapter One

Brownian motion through Riemannian submersions on homogeneous space

In this chapter, we introduce the projection formula of the Beltrami-Laplace operator and explore by examples Brownian motion on common quotient manifolds.

1.1 Laplace-Beltrami operator through Riemannian submersions

Theorem 1.1.1. *(Theorem 1) Let $\phi : (M, g) \to (N, h)$ be a Riemannian submersion, and W_M is a Brownian motion on (M, g), then the image of W_M under ϕ can be written as*

$$d\phi(W_M) = -c\phi_*(\vec{H}_{\phi^{-1}\circ\phi(W_M)}(W_M))dt + dW_N. \tag{1.1}$$

Here W_N is a Brownian motion on (N, h). The vector $\vec{H}_{\phi^{-1}\circ\phi(W_M)}(W_M)$ is the mean curvature of the fiber $\phi^{-1}(W_M)$ as an embedded submanifold of M at W_M, and $c > 0$ is a constant depending on the dimensions of the fiber.

The proof is by a straight forward computation of projecting the Laplace-Beltrami operator on the total space to the quotient.

1.2 Mean curvature flow of group orbits as gradient flow

There are many examples of Riemannian submersions in general. We narrow down to a class of particular interest. That is, the submersion of the form

$$\phi : M \to N = M/K, \tag{1.2}$$

where K is a Lie group acting on M as a subgroup of the isometries of M, then $\phi_*(H)$ is constant along the fiber, i.e. $\phi_*(H(x)) = \phi_*(H(y)) \in T_{\phi(x)}N$ for any $y \in \phi^{-1} \circ \phi(x)$. In this case, the MCF "commutes" with the submersion. That is, the mean curvature flow x_t, in general cases described by PDEs, of the fibers in M descends to a simple ODE of $p_t := \phi(x_t)$ on N,

$$\dot{p}_t = -c\phi_*(H(x_t)) =: J(p_t), \tag{1.3}$$

determined solely by the vector field J on M/K.

Example 1.2.1. Consider the punctured Euclidean space $M = \mathbb{R}^n \setminus 0$ quotient by $K = SO(n)$. The orbits are n-spheres. The quotient space is $(0, \infty)$.

$$dX = \mathrm{Pr}_{\mathrm{ver}}(X)dW = (I - \frac{XX^*}{X^*X})dW. \tag{1.4}$$

Let $R = \sqrt{X^*X}$ and $S = R^2 = X^*X$. By Itô calculus,

$$\begin{aligned} dS &= dX^*X + X^*dX + \frac{1}{2}dX^*dX \\ &= 2X^*dX + \frac{1}{2}dW^* \operatorname{Pr}_{\text{ver}}(X)^* \operatorname{Pr}_{\text{ver}}(X)dW \\ &= 0 + \frac{1}{2}dW^*(I - \frac{XX^*}{S})dW \\ &= \frac{1}{2}(ndt - \frac{1}{S}Sdt) = \frac{n-1}{2}dt. \end{aligned} \tag{1.5}$$

So

$$S = R^2 = \frac{n-1}{2}t + S_0. \tag{1.6}$$

On the other hand, the mean curvature and R are related as

$$\dot{R} = -\frac{n-1}{2}\vec{H} = \frac{n-1}{2R}, \tag{1.7}$$

which leads to the same as Equation 1.6. We have shown in this example, the Itô vertical projection of Brownian motion is indeed the mean curvature drift predicted by the theorem 1.1.1.

A particular fact will be useful for calculating the mean curvature.

Fact 1.2.2. *(C.f. [Pac03] Proposition 1) Let M be a Riemannian manifold such that a compact Lie group K acts as a subgroup of isometries of M. The volume of orbit $vol(K \cdot m)$ is a smooth function on M, and the vector field of orbit mean curvatures is*

$$H(m) = -\operatorname{grad}\log vol(K \cdot m) \tag{1.8}$$

for $m \in M$.

We also have a recipe for the volume function.

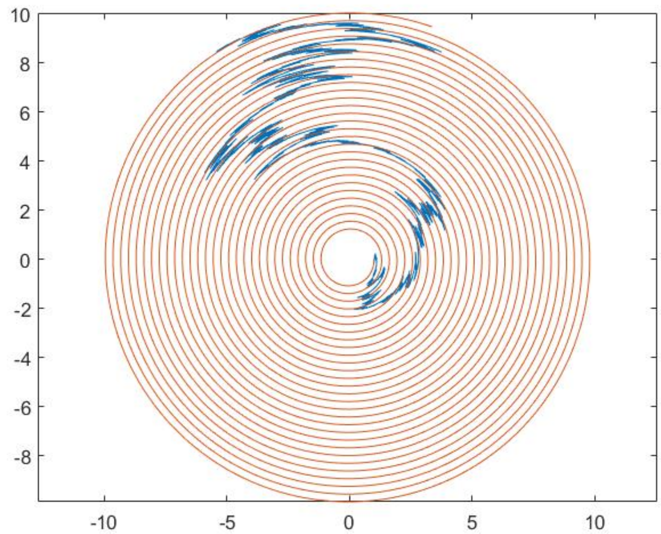

Figure 1.1: Comparing one sample path of P and the MCF at the same scale.

Lemma 1.2.3. *(C.f. [Pac03] Proof of Proposition 1) Let (M, g) be a Riemannian manifold such that a compact Lie group K acts as a subgroup of isometries of M, and let $\{Z_\alpha\}_\alpha$ be a basis of $\mathfrak{k} = Lie(K)$, the volume of an orbit is*

$$vol(K \cdot m) = \sqrt{\det i^* g} \int_K \wedge_\alpha Z^*_\alpha. \tag{1.9}$$

Here $i^* g$ is the metric pulled back from the inclusion $i : K \cdot m \hookrightarrow M$, and Z^*_α is the dual of Z_α.

1.3 Brownian motion on homogeneous Riemannian spaces

If $M = G/K$ is a G-homogeneous space with metric g, we say it is a *homogeneous Riemannian space* if the G-action on (M, g) is an isometry. Often homogeneous spaces come from Riemannian submersion from G endowed with (right, left, or bi-) invariant metric, and we are able to describe Brownian motion on M using the Lie group structures.

On the other hand, when can the invariant metric on M lift to G? This is a property shared by *symmetric spaces*. More generally, the lift exists for *reductive homogeneous spaces* whose metric satisfy certain conditions.

Most of the following examples end up coming from Lie group invariant Brownian motion.

Definition 1.3.1. Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. If $\mathfrak{g}$ is equipped with a metric, then a *left invariant Brownian motion* X on G satisfies

$$dX = X \, đ W,$$

where W is a Euclidean metric on $\mathfrak{g}$.

What is important is that for a *unimodular* Lie group G, i.e. left invariant Haar measures on G are right invariant, or equivalently $\det Ad \equiv 1$ on G or $\mathrm{Tr}\, ad \equiv 0$ on $\mathfrak{g}$, the Laplace-Beltrame operator is a "sum of squares". In particular, the general linear group $GL(n)$, special linear group $SL(n)$, and compact groups such as $O(n)$ are unimodular and the following apply.

Theorem 1.3.2. *([Ura79] §2) Let G be a Lie group and $\mathfrak{g}$ its Lie algebra equipped with an inner product. If G is unimodular and $\{Y_i\}_i$ is an orthonormal basis of $\mathfrak{g}$, then the Laplace-Beltrami operator under the left invariant metric on G induced from the inner product on $\mathfrak{g}$ is*

$$\Delta_G = \sum_i Y_i^2.$$

Corollary 1.3.3. *Let G be a Lie group and $\mathfrak{g}$ its Lie algebra equipped with an inner product. Then left invariant Brownian motion on G is exactly the Brownian motion under the left invariant metric induced from the inner product of $\mathfrak{g}$.*

1.3.1 Quotients of orthogonal groups

Real Grassmannians, flag manifolds, and Stiefel manifolds can be presented as quotients of the orthogonal group $O(n)$, and some most common metrics on them are the quotient metric from $O(n)$. The standard metric on $O(n)$, denoted g for this subsection, can be described either as the bi-invariant from the Killing form or equivalently (up to a scalar) the submanifold metric induced from the usual embedding $O(n) \hookrightarrow \mathbb{R}^{n\times n}$. Let's use the latter as the definition, which can be convenient computed as the Frobenius norm

$$g(A, B) = \operatorname{Tr} AB^*. \tag{1.10}$$

As in the introduction, the Lie algebra $\mathfrak{so}(n)$, represented as skew-symmetric matrices, has orthonormal basis

$$\{A_{ij} = \frac{E_{ij} - E_{ji}}{\sqrt{2}} | 1 \le i < j \le n\}.$$

Brownian motion, Q, on $O(n)$ satisfies

$$dQ = Q\, đA, \tag{1.11}$$

where $A = \frac{1}{\sqrt{2}} \left[W^{ij}\right]_{i,j}$ with $W^{ij} = -W^{ji}$ and $\{W^{ij} | 1 \le i < j \le n\}$ are independent standard Wiener processes.

Consider subgroups of the form

$$K = \begin{bmatrix} I_{n_0} & & & \\ & O(n_1) & & \\ & & \ddots & \\ & & & O(n_r) \end{bmatrix} \tag{1.12}$$

with $n_0 + n_1 + \cdots + n_r = n$ and $n_i \geq 0$. Right multiplication of K on $O(n)$ induces the adjoint map on $\mathfrak{so}(n)$ and is an isometry for g : at $U \in O(n)$, for $V \in K$ and $A, B \in \mathfrak{so}(n)$,

$$\mathrm{Tr}(UV(V^*AV))(UV(V^*BV))^* = \mathrm{Tr}\, AB^* = g(UA, UB). \tag{1.13}$$

Hence the quotients by K have totally geodeisc orbits.

1.3.1.1 Stiefel manifold

The compact real Stiefel manifold $Stief(k, n)$ parametrizes all orthonormal frames of k-dimensional subspaces of $\mathbb{R}^n$ is

$$\{UI_{k,n} | U \in O(n)\} \simeq O(n)/O(n-k), \tag{1.14}$$

where

$$I_{k,n} = \begin{bmatrix} I_k & \\ & 0 \end{bmatrix}.$$

BM on $Stief(k,n)$, denoted S, hence satisfies

$$dS = dQI_{k,n} = Q\,đA I_{k,n}. \tag{1.15}$$

1.3.1.2 Grassmannian

The real Grassmannian $\mathbb{G}r(k,n)$ parametrizing all k-dimensional subspaces of $\mathbb{R}^n$ is

$$\mathbb{G}r(k,n) = O(n)/O(k) \times O(n-k). \tag{1.16}$$

It is often practically and conveniently considered as projection matrix, yielding the mapping

$$\phi : U \mapsto UI_{k,n}U^*, . \tag{1.17}$$

Since the metric g is bi-invariant, the fibers of the submersion $\phi : O(n) \twoheadrightarrow \mathbb{G}r(k,n)$, i.e. the cosets of the quotient, have constant volume and therefore are totally geodesic. Consequently, ϕ projects Brownian motion on $O(n)$ to Brownian motion on $\mathbb{G}r(k,n)$. If Q is a process satisfying 1.11 , then using the projection matrix presentation, writing $P := QI_{k,n}O^*$, Brownian motion on $\mathbb{G}r(k,n)$ satisfies

$$dP = dQI_{k,n}O^* = (đQ)I_{k,n}Q^* + QI_{k,n}\,đQ^*. \tag{1.18}$$

But

$$đQ = Q\,đA = QdA + \frac{1}{2}dQdA = QdA - nQdt \tag{1.19}$$

since

$$dQdA = QdAdA = -2nQdt. \tag{1.20}$$

So now,

$$dP = QdAI_{k,n}Q^* - nQI_{k,n}Q^*dt + QI_{k,n}(dA^*)Q^* - nQI_{k,n}Q^*dt \tag{1.21}$$

$$= -2nPdt + Q(dA_{k,n} - I_{k,n}dA)Q^* \tag{1.22}$$

$$= -2nPdt + Q\sum_{i\leq k<j} dA_{ij}Q^* \tag{1.23}$$

$$Q(\sum_{i\leq k<j} dA_{ij} - 2nI_{k,n}dt)Q^*. \tag{1.24}$$

Remark 1.3.4. The author acknowledged later, in [BW20] they described BM on complex Grassmannians in inhomogeneous coordinates and are able to derive autonomous equation for the process along with detailed equation for the eigenvalue processes.

1.3.1.3 Flag manifolds

More general than Grassmannians, a real flag manifold $Flag(n_1, \cdots, n_r)$ parametrizes all *flags* of linear subspaces $V_1 < V_2 < \cdots < V_r = \mathbb{R}^n$ with $\dim_{\mathbb{R}} V_i = n_1 + \cdots + n_i$. Equivalently, a flag is an orthogonal decomposition $\mathbb{R}^n = W_1 \oplus \cdots \oplus W_r$ with $V_i = W_1 + \cdots + W_i$. Hence we can represent

$$Flag(n_1, \cdots, n_r) = O(n)/O(n_1) \times \cdots \times O(n_r). \tag{1.25}$$

To present the flag manifold in matrix form, we may consider the embedding

$$Flag(n_1, \cdots, n_r) \hookrightarrow Gr(n_1, n) \times Gr(n_r, n). \tag{1.26}$$

That is, for $Q \in O(n)$ with columns $q_1, \cdots, q_n$, let the matrix

$$Q_k = \begin{bmatrix} q_{n_{k-1}+1} & \cdots & q_k \end{bmatrix}, \tag{1.27}$$

and the image of Q in $Flag(n_1, \cdots, n_r)$ is presented as

$$(Q_1^* Q_1, \cdots, Q_r^* Q_r). \tag{1.28}$$

1.3.1.4 Normal metric of isospectral orbits

Let $\Lambda = \mathrm{Diag}(\lambda_1, \cdots, \lambda_n)$. The *adjoint orbit* of Λ by $O(n)$ is

$$Ad_{O(n)}\Lambda = \{U\Lambda U^* | U \in O(n)\}. \tag{1.29}$$

The *normal metric* on $Ad_{O(n)}$ is defined as the quotient metric from $(O(n), g)$. Indeed, WLOG, assuming $\lambda_1 = \cdots \lambda_{n_1} > \lambda_{n_1+1} = \cdots \lambda_{n_1+n_2} > \cdots > \lambda_{n_1+\cdots+n_{r-1}+1} = \cdots = \lambda_n$, we see

$$Ad_{O(n)}\Lambda \simeq Flag(n_1, \cdots, n_r). \tag{1.30}$$

1.3.2 Metric lifting of reductive spaces

Definition 1.3.5. Let G be a Lie group and H a closed subgroup. $Lie(G) = \mathfrak{g}$ and $Lie(H) = \mathfrak{h}$. The homogeneous space G/H is *reductive* if $\mathfrak{g}$ of G is a direct sum

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$$

as vector spaces (in general $\mathfrak{m}$ is not a Lie algebra) such that $\mathfrak{m}$ is stable under the adjoint of H, i.e.

$$Ad_h\mathfrak{m} \subset \mathfrak{m}$$

for any $h \in H$.

Proposition 1.3.6. *(C.f. [GQ20] §23.4) Let the space of right cosets, G/H, be a reductive space as described in the definition above.*

1. *If H is compact, then there exists a left G-invariant metric on G/H. Furthermore, if the adjoint representation of H on $\mathfrak{m}$ is irreducible, then the metric is unique up to scalars.*

2. *If $\mathfrak{m}$ has an $Ad(H)$-invariant inner product $\langle\cdot,\cdot\rangle_{\mathfrak{m}}$, then the quotient map $G \to G/H$ is a Riemannian submersion where G is endowed with any left G-invariant metric $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ extending $\langle\cdot,\cdot\rangle_{\mathfrak{m}}$ such that $\mathfrak{h}$ is orthogonal to $\mathfrak{m}$, and $\mathfrak{m}$ maps isometrically to $T_H G/H$.*

1.3.2.1 Poincaré's upper half plane

Example 1.3.7. Consider the right-quotient from the special linear group to Poincaré upper half plane,

$$q : SL(2,\mathbb{R}) \to SL(2,\mathbb{R})/SO(2) \cong \mathbb{H}_+. \tag{1.31}$$

The upper half plane $\mathbb{H}_+$ equipped with the Poincaré metric is a classical object. In fact, one can find a left invariant metric on $SL(2)$ such that q becomes a Riemannian submersion.

Define

$$X = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ Y = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},\ Z = \frac{1}{2}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{1.32}$$

These generate the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$. Let $\mathfrak{m}$ be spanned by X and Y. Then

$$\mathfrak{sl}(2,) = \mathfrak{so}(2) \oplus \mathfrak{m} \tag{1.33}$$

as vector spaces. Note $Ad_{SO(2)}\mathfrak{m} \subset \mathfrak{m}$. In other words, $\mathbb{H}_+$ is a reductive homogeneous space. Define a metric on $\mathfrak{m}$ such that X and Y are orthogonal and each have length 1. One can check this metric is $Ad_{SO(2)}$-invariant. Then any metric on $\mathfrak{sl}(2)$, i.e. left invariant metric on $SL(2)$, such that $\mathfrak{so}(2)$ is orthogonal to $\mathfrak{m}$ makes q a Riemannian submersion.

Next we calculate the mean curvature of the orbits. Using the Levi-Civita connection under a left invariant metric, the component of $\nabla_Z Z$ normal to the fiber is

$$\begin{aligned}
(\nabla_Z Z)^{\perp} &= \langle \nabla_Z Z, X\rangle + \langle \nabla_Z Z, Y\rangle \\
&= \frac{1}{2}(\langle [Z,Z], X\rangle - \langle [Z,X], Z\rangle + \langle [X,Z], Z\rangle) \\
&+ \frac{1}{2}(\langle [Z,Z], Y\rangle - \langle [Z,Y], Z\rangle + \langle [Y,Z], Z\rangle) \\
&= \langle [X,Z], Z\rangle + \langle [Y,Z], Z\rangle \\
&= 2\langle -Y, Z\rangle + 2\langle X, Z\rangle \\
&= 0
\end{aligned} \tag{1.34}$$

The mean curvature of the fibers vanishes. One consequence is that

$$dW_{\mathbb{H}_+} = dq(W_{SL(2)}). \tag{1.35}$$

$W_{SL(2)}$ here is left invariant

$$đ W_{SL(2)} = W_{SL(2)}\, đ W_{\mathfrak{sl}(2)}. \tag{1.36}$$

$W_{\mathfrak{sl}(2)}$ is a Euclidean Brownian motion according to the metric defined earlier.

In fact, the right $SO(2)$-multiplication is still an isometry by observing

$$GVQ = GQ(Q^*VQ) = GQAd_{Q^*}V, \tag{1.37}$$

where $G \in SL(2)$, $V \in \mathfrak{sl}(2)$, $Q \in SO(2)$. Then since the orbits have invariant volume, mean curvature vanishes. In addition, $Ad_{SO(2)}$ on $\mathfrak{m}$ is an irreducible representation, so the invariant metric is unique.

1.3.2.2 The Cartan-Hadamard geometry on $\mathcal{S}_+(n,n)$

Example 1.3.8. The Cartan-Hadamard or trace metric on $\mathcal{S}_+(n,n)$ is

$$\langle V, W\rangle = \mathrm{Tr}\, P^{-1}VP^{-1}W \tag{1.38}$$

at $P \in \mathcal{S}_+(n,n)$. One can check it is invariant under the action

$$GL(n) \times \mathcal{S}_+(n,n) : (G,P) \mapsto GPG^*, \tag{1.39}$$

which is induced from left multiplications of $GL(n)$ on itself through the mapping $\pi : M \mapsto MM^*$.

This quotient is reductive: let

$$\mathfrak{gl}(n) = \mathfrak{so}(n) \oplus \mathfrak{m}, \tag{1.40}$$

where (under the usual representation), $\mathfrak{m}$ consists of all symmetric $n \times n$ matrices. It is stable under Ad_Q for any $Q \in O(n)$.

Next, $\pi_* : S \in \mathfrak{m} \mapsto S + S^* = 2S$. We can easily check the lift of the trace metric is $Ad_{O(n)}$-invariant. Therefore a left $GL(n)$-invariant metric on $\mathfrak{gl}(n)$ such that $\mathfrak{so}(n)$ is orthogonal to $\mathfrak{m}$ makes π a Riemannian submersion.

Now we look at the mean curvature drift. For any $S \in \mathfrak{m}$ and $A \in \mathfrak{so}(n)$,

$$\langle \nabla_A A, S \rangle = \frac{1}{2}(\langle [A, A], S \rangle - \langle [A, S], A \rangle + \langle [S, A], A \rangle) = \langle [S, A], A \rangle = 0 \tag{1.41}$$

since $[S, A]$ is symmetric. (Or, since $Ad_{O(n)}\mathfrak{m} \subset \mathfrak{m}$, $ad_{\mathfrak{so}(n)}\mathfrak{m} \subset \mathfrak{m}$, which is orthogonal to $\mathfrak{so}(n)$.) This even proves the second fundamental form vanishes, i.e. the fibers are totally geodesic. Hence a left invariant Brownian motion on $GL(n)$ (according to the metric extending the trace metric) maps to a Brownian motion under the trace metric.

Similar to $\mathbb{H}_+ \cong SL(2)/SO(2)$, right $O(n)$-multiplication is in fact isometry and the $Ad_{O(n)}$-representation on $\mathfrak{m}$ is irreducible.

1.4 An example of nonhomogeneous metric: the Bures-Wasserstein geometry

The model example is the following

Lemma 1.4.1. *(C.f.[MA20]) The mapping*

$$\begin{array}{rccc} \pi_k: & \mathbb{R}^{nk^\times} & \to & \mathcal{S}_+(k,n) \cong \mathbb{R}^{nk^\times}/O(k) \\ & M & \mapsto & MM^* \end{array} \tag{1.42}$$

is a Riemannian submersion for $k \leq n$.

Here $\mathbb{R}^{nk}$ is the Euclidean space of real n-by-k matrices equipped with the Frobenius norm, $\langle X, Y\rangle_{Fr} := \operatorname{Tr} XY^*$, and $\mathbb{R}^{nk^\times}$ is the dense open subset of all rank k matrices. $\mathcal{S}_+(k,n)$ is the set of all n-by-n positive semidefinite matrices of rank k.

Definition 1.4.2. The induced metric on $\mathcal{S}_+(k,n)$ in the lemma above is called the **Bures-Wasserstein (B-W) metric**.

The B-W geometry has been studied in several aspects [MHA19][MA20][BJL19] and has important meanings in different areas such as physics, optimization, and information theory. Much of the content in this paper is extendable to singular matrices though we focus on the locus of nonsingular matrices $\mathcal{S}_+(n,n)$ mostly for the time being.

We are able to write down the orbit mean curvature in detail.

Proposition 1.4.3. *If $P = U\Lambda U^* \in \mathcal{S}_+(k,n)$ for $\Lambda = Diag\,(\lambda_1, \cdots, \lambda_k, 0, \cdots)$ and $U \in O(n)$, the drift term*

$$J(P) = U\tilde{\Lambda}U^*, \tag{1.43}$$

where

$$\tilde{\Lambda} = Diag\left(\sum_{j\neq 1} \frac{\lambda_1}{\lambda_1 + \lambda_j}, \cdots, \sum_{j\neq k} \frac{\lambda_k}{\lambda_k + \lambda_j}, \cdots\right) \tag{1.44}$$

Proof. Since $O(k)$ acts by isometry, it suffices to compute the mean curvature at $M = U\tilde{L}$ in a fiber, where via SVD, U is orthogonal and $\tilde{L} = \begin{bmatrix} L \\ 0 \end{bmatrix}$ is n-by-k for $L = \mathrm{Diag}(l_1, \cdots, l_k)$.

First, we need to find an orthonormal basis (o.n.b.) for $T_M MO(k) = M\mathfrak{so}(k)$: let $\{E_{ij}\}$ denote the standard basis with the (i, j)-th entry being one and zeros otherwise, and let $A_{ij} = E_{ij} - E_{ji}$. Now for $A, B \in \mathfrak{so}(k)$ (i.e. skew-symmetric),

$$< MA, MB >_{T_M MO(n)} = \mathrm{Tr}\, ULAB^* L^* U^* = \mathrm{Tr}\, LAB^* L. \tag{1.45}$$

It is easy to verify that

$$M\{\tilde{A}_{ij} := \frac{A_{ij}}{\sqrt{l_i^2 + l_j^2}} | i < j\} \tag{1.46}$$

is an o.n.b. for $T_M MO(k)$.

Each $\tilde{A}_{ij}$ has trajectory

$$c_{ij}(t) = Me^{t\tilde{A}_{ij}} \tag{1.47}$$

on $MO(n)$. Hence the second fundamental form

$$II(M\tilde{A}_{ij}, M\tilde{A}_{ij}) = (D_{M\tilde{A}_{ij}} M\tilde{A}_{ij})^\perp = (c_{ij}''(0))^\perp = (-M\frac{E_{ii} + E_{jj}}{l_i^2 + l_j^2})^\perp = -M\frac{E_{ii} + E_{jj}}{l_i^2 + l_j^2}. \tag{1.48}$$

Therefore the mean curvature is

$$\dim O(k)H = -M(\sum_{i<j}\frac{1}{l_i^2+l_j^2}(E_{ii}+E_{jj})) \qquad = -M\,\mathrm{Diag}(\cdots,\sum_{j\neq i}\frac{1}{l_i^2+l_j^2},\cdots) \tag{1.49}$$

$$= -U\,\mathrm{Diag}(\cdots,\sum_{j\neq i}\frac{l_i}{l_i^2+l_j^2},\cdots) \tag{1.50}$$

and

$$\pi_*(H) = MH^* + HM^* \tag{1.51}$$

$$= -2M(\sum_{i<j}\frac{1}{l_i^2+l_j^2}(E_{ii}+E_{jj}))M^* \tag{1.52}$$

$$= -2M\,\mathrm{Diag}(\sum_{j\neq 1}\frac{1}{l_1^2+l_j^2},\cdots,\sum_{j\neq i}\frac{1}{l_i^2+l_j^2},\cdots)M^* \tag{1.53}$$

$$= 2U\,\mathrm{Diag}(\sum_{j\neq 1}\frac{l_1^2}{l_1^2+l_j^2},\cdots,\sum_{j\neq i}\frac{l_i^2}{l_i^2+l_j^2},\cdots)U^*. \tag{1.54}$$

□

Now we apply 1.2.2 to compute the mean curvature in a different way.

Lemma 1.4.4. *The volume function for the fibers of*

$$\pi : \mathbb{R}^{nk^\times} \to \mathcal{S}_+(n,n)$$

at $M \in \pi^{-1}(P)$ *is*

$$vol(MO(k)) = \sqrt{\det g|_{MO(k)}}\cdot vol(O(k)). \tag{1.55}$$

The determinant in the formula is taken under an orthonormal basis of $\mathfrak{so}(k)$*, in which case n entry of* $g|_{MO(k)}$ *is (using Kronecker* δ*)*

$$g_{ij,ml} = \delta_{ml}P_{jl} + \delta_{jl}P_{im} - \delta_{il}P_{jm} - \delta_{jm}P_{il}. \tag{1.56}$$

and the determinant in the previous formula is taken

Proof. Applying 1.2.3, we may choose an (orthonormal) basis $\{A_{ij} = \frac{E_{ij}-E_{ji}}{\sqrt{2}} | i < j \leq k\}$ for $\mathfrak{so}(k)$. Then

$$vol(MO(k)) = \sqrt{\det g|_{MO(k)}} \cdot vol(O(k)). \tag{1.57}$$

An entry of g is

$$g_{ij,ml} = \operatorname{Tr} MA_{ij}A^*_{ml}M^* \tag{1.58}$$

$$= \delta_{ml}P'_{jl} + \delta_{jl}P'_{im} - \delta_{il}P'_{jm} - \delta_{jm}P'_{il}, \tag{1.59}$$

with $P' := M^*M$. Since $vol(MO(k))$ is constant on the orbit, we may choose $M = \sqrt{P}$ so that $P' = P$. □

We do not have a more explicit formula, but we may compute the gradient under the flat metric on $GL(n)$ and project it through the submersion.

Lemma 1.4.5. *Let $\psi : M \to N$ be a Riemannian submersion and $f : N \to \mathbb{R}$ a smooth function. Then*

$$\operatorname{grad} f = \psi_* \operatorname{grad}(f \circ \psi).$$

Proof. Since $f \circ \psi$ is constant along each fiber, on M,

$$\langle \operatorname{grad} f \circ \psi, V \rangle = V(f \circ \psi) = 0 \tag{1.60}$$

for any vertical V. Therefore $\operatorname{grad} f \circ \psi$ is horizontal.

On N, for any tangent vector W,

$$\langle \operatorname{grad} f, W\rangle = \langle \operatorname{grad} f\circ\psi, \bar{W}\rangle = \langle \psi_*(\operatorname{grad} f\circ\psi), W\rangle, \tag{1.61}$$

where $\bar{W}$ is a horizontal lift of W. □

Example 1.4.6. For $GL(2)\to\mathcal{S}_+(2,2)$, write $P = U\operatorname{Diag}(\lambda_1,\lambda_2)U^*$,

$$J(P) = U\operatorname{Diag}(\frac{\lambda_1}{\lambda_1+\lambda_2}, \frac{\lambda_2}{\lambda_1+\lambda_2})U^* = \frac{1}{\operatorname{Tr} P}P. \tag{1.62}$$

$$\det g = g_{12,12} = \operatorname{Tr} MA_{12}A_{12}^*M^* = \frac{1}{2}\operatorname{Tr} P = \frac{\lambda_1+\lambda_2}{2}. \tag{1.63}$$

The mean curvature is is

$$H = -\nabla\log\sqrt{\frac{P_{11}+P_{22}}{2}} = -\frac{1}{2\operatorname{Tr} P}\begin{bmatrix} 2M_{11} & 2M_{12} \\ 2M_{21} & 2M_{22}\end{bmatrix} = -\frac{1}{\operatorname{Tr} P}M. \tag{1.64}$$

Therefore

$$J(P) = \pi_*(H) = -\frac{\dim O(2)}{2}(MH^* + HM^*) = \frac{1}{\operatorname{Tr} P}P, \tag{1.65}$$

consistent with 1.62.

Corollary 1.4.7. *The Brownian motion on $\mathcal{S}_+(k,n)$ with the B-W metric satisfies the SDE (until leaving $\mathbb{R}^{nk^\times}$)*

$$\begin{aligned} dW_{BW} &= d(WW^*) + J(W_{BW})dt \\ &= (dW)W^* + WdW^* + \frac{1}{2}nIdt + J(W_{BW})dt, \end{aligned} \tag{1.66}$$

where W is a standard $n \times k$-dimensional Wiener process.

Remark 1.4.8. The corollary above along with the formula 1.4.3 is reminiscent of two known processes.

Dyson's Brownian motion: consider the space of $n \times n$ hermitian or real symmetric matrices as a Euclidean space over $\mathbb{R}$. If M is a BM on this space, then its eigenvalues satisfy the equation

$$d\lambda_i = dB_i + \sum_{j \neq i} \frac{dt}{\lambda_i - \lambda_j}, \tag{1.67}$$

where B_i's are standard BMs. The drift term serves as repulsion and ensures that eigenvalues do not collide. In constrast, the B-W Brownian motion 1.4.7 has an " averaging" deterministic term–if all the eigenvalues of the initial point are equal, they remain so with the flow of J.

Brownian motion of ellipsoids: (c.f. [RW00]) let G be a standard right invariant BM on $GL(n)$, i.e.

$$đG = (đW_{\mathfrak{gl}(n)})G, \tag{1.68}$$

where $W_{\mathfrak{gl}(n)}$ is a standard $n \times n$ BM. Define $X = G^*G$ and $Y = G^*G$. Let λ_i be the eigenvalues of X and $\gamma_i = \frac{1}{2} \log \lambda_i$, then

$$d\gamma_i = dW_i + \frac{1}{2} \sum_{j \neq i} \coth(\gamma_i - \gamma_j). \tag{1.69}$$

In fact, the eigenvalues never collide in this case as well.

By 1.3.8, Y as above is a BM on $\mathcal{S}_+(n, n)$ under the Cartan-Hamamard geometry, commonly referred to as Dynkin's BM. The eigenvalues do not possess descriptions as clear. However the eigenvectors have limits for $t \to \infty$.

Lemma 1.4.9. *Let $\mathcal{M}_k(n,\mathbb{R})$ be the manifold of rank k n-by-n matrices as a Riemannian submanifold of $\mathcal{M}(n,\mathbb{R})$. ($\mathcal{M}_k(n)$ is the determinantal real variety minus the singular locus, which consists of matrices of ranks lower than k.) Then the submersion*

$$\begin{aligned} \pi : \quad \mathcal{M}_k(n) &\rightarrow \mathcal{S}_+(k,n) \\ M &\mapsto MM^* \end{aligned} \tag{1.70}$$

defines the same metric as the BW metric in the previous lemma on $\mathcal{S}''(k,n)$.

Proof. $\mathbb{R}^{nk}$ is a subset (embedded?) in $\mathcal{M}_k(n)$ via the inclusion

$$i : M_0 \mapsto \left[M_0 | 0_{n(n-k)} \right]. \tag{1.71}$$

The right multiplication of $O(n)$ is an isometry, and $i(\mathbb{R}^{nk^\times})O(n) = \mathcal{M}_k(n)$. Therefore one only needs to check on $i(\mathbb{R}^{nk^\times}))$.

First, the vertical space w.r.t. π_k at M_0 is $M\mathfrak{so}(k)$, which is obviously a subset of that w.r.t. π, $M\mathfrak{so}(n)$ via the inclusion. Hence it suffices to check the horizontal space w.r.t π_k is horizontal w.r.t. to π. The horizontal space w.r.t. π_k is (c.f. [MA20])

$$\{Z \in \mathbb{R}^{nk} | Z^* M_0 \text{ is symmetric}\}. \tag{1.72}$$

For any $A \in \mathfrak{so}(n)$,

$$\operatorname{Tr} MAi(Z)^* = \operatorname{Tr}(i(Z)^* M)A = \operatorname{Tr} \begin{bmatrix} Z^* M_0 & \\ & 0 \end{bmatrix} A = 0, \tag{1.73}$$

if $Z^* M_0$ is symmetric. This concludes the proof. ◻

Chapter Two

A control system using the mean curvature drift

In this chapter, we investigate ideas to utilize the mean curvature drift into a stochastic flow. In short, on our model example of $\mathcal{S}_+(n,n)$, first we prove that the vertical process, X, of Brownian motion W projects through π a deterministic process P that is exactly the mean curvature drift whose derivative is $J(P)$. Define a new metric on $\mathcal{S}_+(n,n)$,

$$\mathrm{Tr}_R\, VW^* := \langle V, W\rangle_R := \mathrm{Tr}\, RVW^*,$$

for $R \in \mathcal{S}_+(n,n)$. We may still consider the Brownian motion W_R and orthogonal projection under the new metric to obtain modified drift term $J^R(P)dt$. By varying R along the path, we have a control system.

The main results are the following:

Proposition 2.0.1. *The following are equivalent descriptions of the same control system for the evolution $P_t \in \mathcal{S}_+(n,n) \cong GL(n)/O(n)$ in the sense that the controls give the same possible set of $\{P_t\}$ for a fixed time t and an initial condition P_0.*

1. *$\dot{P}_t = M_t C_t M_t^*$, where $M_t M_t^* = P$, and the control C_t is any matrix in $\mathcal{S}_+(n,n)$ such that the eigenvalues are of the form*

$$(\sum_{j\neq 1} \frac{1}{\lambda_1 + \lambda_j}, \cdots, \sum_{j\neq k} \frac{1}{\lambda_k + \lambda_j}, \cdots) \tag{2.1}$$

   with each $\lambda_j > 0$. Here the choice of which M_t does not change the set of possible $\dot{P}_t$.

2. *$\dot{P}_t = J^{R_t}(P_t)$, where the control is $R_t \in \mathcal{S}_+(n,n)$, and J^R denotes the drift term as previously under metric Tr_R.*

3. *$\dot{P}_t = G_t^{-*} J(G_t^* P_t G_t) G_t^{-1}$, where the control is any $G_t \in GL(n)$.*

We denote the subset in $T_P\mathcal{S}_+(n,n)$ of all possible choices of $\dot{P}$ as $\mathfrak{a}(P)$. An important observation is

Proposition 2.0.2. *$\mathfrak{a}(P)$ is the restriction at $T_P\mathcal{S}_+(n,n)$ of a collection of vector fields on $\mathcal{S}_+(n,n)$, denoted as $\mathfrak{a}$, invariant under the $GL(n)$ action descended from left $GL(n)$-multiplication on $GL(n)$.*

As in existing work in invariant control systems, we consider *piecewise constant* controls and *regulated controls*, which are thought as the limits of the former. We call the space the system can reach from the point P at some finite time t with a certain control type the *reachable set* or *attainable set*, denoted as $\mathcal{R}^{pc}(P)$ and $\mathcal{R}^{reg}(I)$ respectively. We shall introduce these concepts later. Some conclusions for the reachable set is the following.

Proposition 2.0.3. *For the control system described in 2.0.1*

1. $$P_{t_2} - P_{t_1}, \tag{2.2}$$

   is positive definite if $t_2 > t_1$.

2. *$\mathfrak{a}(I)$ is open in $T_I\mathcal{S}_+(n,n)$. Therefore the attainable set $\mathcal{R}^{pc}(I)$ has nonempty interior.*

3. *The attainable set $\mathcal{R}^{reg}(I)$ via regulated controls contains a subset*

   $$I + \{U\,Diag(\exp\lambda_1, \cdots, \exp\lambda_n)U^*|U \in O(n),\ (\lambda_1, \cdot, \lambda_n) \in cone\{e_i + e_j | i \neq j\}\}, \tag{2.3}$$

   where $\{e_j\}$ denotes the standard basis of $\mathbb{R}^n$.

First, we need some technical result for the existence of the "vertical Brownian motion".

Proposition 2.0.4. *Let* $\pi = \pi_n : GL(n) \to \mathcal{S}_+(n,n)$*. Then* $\mathrm{Pr}_{\mathrm{ver}}$ *is locally Lipschitz and hence local solutions to* $dX := \mathrm{Pr}_{\mathrm{ver}}\, dW_R$ *exist.*

Furthermore, the image process $P := \pi(X)$ *depends on* P *but not* X *and is deterministic. In other words, it satisfies an autonomous equation*

$$\dot{P} = J(P), \tag{2.4}$$

where $J(P)$ *results from the mean curvature of fibers.*

We shall see that the Lipschitz condition is provided by writing the decomposition into horizontal and vertical directions w.r.t. π as Lyapunov equations, from which numerical implementations could also benefit.

2.1 Construction

First, we prove a version of 1.1.1.

Lemma 2.1.1. *Let* $\phi : M \to N$ *be a submersion and* M *be equipped with a metric* g*. Let* $\mathrm{Pr}_{\mathrm{ver}}$ *denote the orthogonal projection to the vertical direction with respect to the submersion. Define the process* X *by*

$$dX = \mathrm{Pr}_{\mathrm{ver}}\, dW, \tag{2.5}$$

where W *is a Brownian motion on* (M, g)*. Assuming the existence of solutions, the image process,* $\phi(X) =: P \in N$*, satisfies the equation*

$$dP = -c\phi_*(H(X_t))dt, \tag{2.6}$$

where $H(X)$ is the mean curvature of the fiber $\phi^{-1}(P)$ at X in M and $c = \frac{\dim \phi^{-1}(P)}{2}$ depends on the dimension of the fiber.

Proof. Pick a normal coordinate system at $o := X_0$ such that the first m coordinate derivatives $\partial_1, \cdots, \partial_m$ span the vertical space at o.

Let W_M be a Brownian motion on M. In local coordinates,

$$dW_M^k = (\sqrt{g^{-1}}dB)^k - \frac{1}{2}g^{ij}\Gamma_{ij}^k dt. \tag{2.7}$$

Let $\xi_i = \mathrm{Pr}_{\mathrm{ver}}\,\partial_i$.

$$dX = \xi_i dW_M^i = \xi_i \circ dW_M^i - \frac{1}{2}D_{\xi_i}\xi_j d\langle W_M^i, W_M^j\rangle \tag{2.8}$$

in Stratonovich form, where D denotes the "Euclidean connection" for the coordinates.

At o, $D_{\xi_i}\xi_j$ coincides with $\nabla_{\partial_i}\partial_j$ and $\langle W_M^i, W_M^j\rangle = t$. So

$$dP = d\phi(X) = -\frac{1}{2}\sum_{i=1,\cdots,m}\phi_* \mathrm{I\!I}(\partial_i, \partial_i)dt. \tag{2.9}$$

□

The lemma indicates we can extract the drift term by orthogonal projection and achieve a mean curvature flow. If, as in the examples introduced in chapter 1, the image of mean curvature is constant along each fiber, the image process in the quotient is deterministic.

In order to create freedom for the flow, we must alter some factor in the equation.

Since the mean curvature is a consequence from the metric, we consider a "change of metric" on $GL(n)$. Define the inner product on $T_M GL(n)$:

$$\operatorname{Tr}_R VW^* := \langle V, W \rangle_R := \operatorname{Tr} RVW^*, \tag{2.10}$$

where $R > 0$ is a positive matrix in $\mathcal{S}_+(n, n)$. A new metric on $GL(n)$ induces a new horizontal space w.r.t. the submersion π and hence a different mean curvature denoted H^R. We denote the new drift on $\mathcal{S}_+(n, n)$ by J^R. Choosing different R, we have a control system

$$\begin{cases} dX = \operatorname{Pr}_{\mathrm{ver}R} dW_R \\ dP := d\pi(X) = J^R(P), \end{cases} \tag{2.11}$$

where W_R is Brownian motion under Tr_R, and $\operatorname{Pr}_{\mathrm{ver}R}$ and J^R are the orthogonal projection to vertical directions and the drift term defined as earlier in this new metric.

Remark 2.1.2. Is the metric change a reasonable enough choice to achieve the control? One may choose to project to other directions not parallel to the horizontal space; essentially it means introducing new metric on the total space. One might naturally choose other diffusion processes to replace the Brownian motion; as mentioned in introduction, this is equivalent to metric changes too.

We may require the metric change to affect W or $\operatorname{Pr}_{\mathrm{ver}}$ only instead of both. However, these do not share similar nice description.

It turns out there are more than one reasons this is a reasonable choice of new metrics. First,

Lemma 2.1.3. *The class* $\{\langle,\rangle_R \,|R \in \mathcal{S}_+(n,n)\}$ *consists of all metrics on* $GL(n)$ *(i) invariant under* $O(n)$ *right multiplication such that (ii) the metric tensors are constant in the usual* $\mathbb{R}^{n\times n}$ *coordinates. In particular, the mean curvature is still constant along the fiber under (a fixed)* Tr_R.

Proof. Notations: write E_{ab} for the $n \times n$ matrix whose (a,b)-th entry is 1 and 0 otherwise. Let e^a be the row vector with a-th entry being 1 and 0 elsewhere. $Q = [q^{ij}]$ denotes a general orthogonal matrix, of which q^i is the i-th row.

$\{E_{ab}\}$ is the standard basis for $\mathcal{M}(n,\mathbb{R}) \cong \mathbb{R}^{n\times n}$.

Let g be a metric satisfying conditions (i) and (ii). (i) means the metric tensor

$$g_{ij,kl} =< E_{ij}, E_{kl} >_g =< E_{ij}Q, E_{kl}Q >_g = \sum_{a,b} q^{ja} q^{lb} g_{ia,kb}. \tag{2.12}$$

for any $Q \in O(n)$.

Claim $g_{ij,kl} \equiv 0$ for $j \neq l$ and $g_{ij,kj}$ does not depend on j.

Take $q^j = e^{\bar{a}}$ and $q^l = e^{\bar{b}}$. Then

$$g_{ij,kl} = g_{i\bar{a},k\bar{b}} \tag{2.13}$$

for any pair $a \neq b$.

Next, for $\bar{a} < \bar{b}$, take the $(i, j) \times (\bar{a}, \bar{b})$-submatrix of Q to be

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Now

$$g_{ij,kl} = (\sum_{a \neq b} q^{ja} q^{lb}) g_{ij,kl} + \sum_{a} q^{ja} q^{la} g_{ia,ka} \tag{2.14}$$

$$= 0 + (\frac{1}{2} g_{i\bar{a},k\bar{a}} - \frac{1}{2} g_{i\bar{b},k,\bar{b}}). \tag{2.15}$$

Since $\bar{a}$ and $\bar{b}$ are arbitrary, this again implies $g_{ia,ka}$ in constant in a and $g_{ij,kl} = 0$ for $j \neq l$.

If $g = \mathrm{Tr}_R$ for $R = [r_i j]$, then

$$g_{ij,kl} = \mathrm{Tr}\, RE_{ij}E^*_{kl} = \mathrm{Tr}\, RE_{ij}E_{lk} = \delta_{jl}\, \mathrm{Tr}\, RE_{ik} = \delta_{jl} r_{ki}. \tag{2.16}$$

Hence $\{\mathrm{Tr}_R \,|R > 0\}$ is exactly the class of all metrics satisfying (i) and (ii).

$\square$

Second, it is equivalent to the pull-back metric induced by the left multiplication of $GL(n)$: for $R = G^*G$:

$$\langle V, W \rangle_R = \mathrm{Tr}\, RVW^* = \mathrm{Tr}(GV)(GW)^* = \langle GV, GW \rangle_{Fr}\,. \tag{2.17}$$

Therefore we have an isometry

$$\begin{aligned} L_{G^{-1}}: \quad (GL(n), \mathrm{Tr}_R) &\to (GL(n), Fr) \\ M &\mapsto G^{-1}M. \end{aligned} \tag{2.18}$$

In other words, we permute the Frobenius metric on $GL(n)$ by left multiplication map , L_G, $G \in GL(n)$, which are homeomorphisms naturally and commute with the quotient to left cosets. The flow under the new metric becomes

$$\dot{P} = J^R(P) = -c\phi_*(L_G^*(H(L_G(X_t)))). \tag{2.19}$$

We vary R_t to have control at all time t. Now the freedom we have is the set of all vector fields

$$\{J^R | R \in \mathcal{S}_+(n,n)\} =: \mathfrak{a} \tag{2.20}$$

It is G left-invariant on $\mathcal{S}_+(n,n)$.

2.2 Existence

Lemma 2.2.1. *The projection onto each orbit, $MO(n)$, $\mathrm{Pr}_{\mathrm{ver}}$ is locally Lipschitz, and hence the equation for X 2.5 exists and is unique locally.*

The proof is based on writing the orthogonal projection in the form of continuous algebraic Lyapunov equations, which has the following solution:

Proposition 2.2.2. *Given $P, B \in \mathcal{S}(n)$, if P is stable, i.e. its eigenvalues have negative* *real parts, then the continuous algebraic Lyapunov equation*

$$P^T X + XP + B = 0$$

has a unique solution for $X \in \mathcal{M}(n, \mathbb{R})$,

$$X = \int_0^\infty e^{tP} B e^{tP} dt.$$

Proposition 2.2.3. *Let $M \in GL(n, \mathbb{R})$ and $F = F_{MM^*} = MO(n)$ be the orbit of M under the $O(n)$ action. Let $P := MM^*$ and $P' := M^*M$.*

1. *The vertical component of $ME \in T_M F$ is*

$$MK = M \int_0^\infty e^{-tP'}(P'E - E^*P')e^{-tP'} dt \tag{2.21}$$

2. *Alternatively, a vector in the form $EM \in T_M F$ has vertical component*

$$SM = \int_0^\infty e^{-tP}(PE^* + EP)e^{-tP} dtM. \tag{2.22}$$

3. *Let $MA \in T_M MO(n)$, the second fundamental form for the orbit is*

$$II(MA, MA) = MA^2 - M \int_0^\infty e^{-tP'}(P'A^2 - A^2P')e^{-tP'} dt.$$

Proof. 1. Observe at M, a vertical vector is of the form MK, where K is skew-hermitian. The horizontal space is thus

$$\{M^{-1}S | \operatorname{Tr} MKS^*M^* = \operatorname{Tr} KS^* = 0,\ \forall K \text{ skew-symmetric}\} = \{M^{-1}S | S \text{ symmetric}\}. \tag{2.23}$$

In other words, we have the decomposition

$$T_M GL(n) = T_M F_P \oplus N_M F_P = M\mathfrak{so}(n) \oplus M^{-1*}\mathcal{S}(n), \tag{2.24}$$

where $\mathcal{S}(n)$ denotes all $n \times n$ symmetric matrices.

We need to find skew-symmetric K and symmetric S such that

$$ME = MK + M^{-1*}S. \tag{2.25}$$

$$\implies M^*ME = M^*MK + S. \tag{2.26}$$

Write $P' = M^*M$, since S is symmetric,

$$P'E - P'K = (P'E - P'K)^* = E^*P' + KP' \tag{2.27}$$

$$\implies P'E - E^*P' = P'K + KP', \tag{2.28}$$

where K is the only unknown. Let $B = P'E - E^*P'$. It becomes in the standard form of the continuous algebraic Lyapunov equation

$$-P'K - KP' + B = 0, \tag{2.29}$$

with $-P'$ negative definite.

The solution is

$$K = \int_0^\infty e^{-tP'}Be^{-tP'}dt = \int_0^\infty e^{-tP'}(P'E - E^*P')e^{-tP'}dt. \tag{2.30}$$

2. Similarly,

$$T_M GL(n) = T_M F_P \oplus N_M F_P = \mathfrak{so}(n)M^{-1*} \oplus \mathcal{S}(n)M. \tag{2.31}$$

In other words, for $EM \in T_M \mathcal{M}(n, \mathbb{R})$,

$$EM = KM^{-*} + SM, \tag{2.32}$$

with K skew-symmetric and S symmetric.

$$\implies K = EP - SP. \tag{2.33}$$

Since K is skew-symmetric,

$$\begin{aligned} -K &= P^* E^* - P^* S^* \\ &= PE^* - PS \\ &= -EP + SP. \end{aligned} \tag{2.34}$$

$$\implies SP + PS = (PE^* + EP) =: B,$$

or

$$-SP - PS + B = 0. \tag{2.35}$$

Since $-P$ is negative definite, the solution for S is

$$S = \int_0^\infty e^{-tP}(PE^* + EP)e^{-tP}dt. \tag{2.36}$$

3. For $MA \in T_M MO(n)$, with A skew-symmetric, has a trajectory on F_P

$$\alpha(t) = Me^{tA}.$$

The second fundamental form $II(MA, MA)$ is the normal component of

$$\alpha''(0) = MA^2.$$

Using 2.30,

$$II(MA, MA) = MA^2 - M \int_0^\infty e^{-tP'}(P'A^2 - A^2P')e^{-tP'} dt. \tag{2.37}$$

□

Proof. (of 2.2.1)

Since $\mathrm{Pr}_{\mathrm{ver}}(M)$ is linear, it suffices to test on $C \in T_N\mathcal{M}(n, \mathbb{R})$ with $||C|| = 1$ for N in a neighborhood of M that

$$|| \mathrm{Pr}_{\mathrm{ver}}(N)C_N - \mathrm{Pr}_{\mathrm{ver}}(M)C_M|| < k||N - M|| \tag{2.38}$$

for some local constant k.

To see this, for $G^{-1}N$ near M and $G^{-1}C \in T_N GL(n)$ with $||G^{-1}C||_R = ||C||_{Fr} = 1$, the vertical component is

$$G^{-1}SN = G^{-1} \int_0^\infty e^{-tNN^*}(NC^* + CN^*)e^{-tNN^*} dtN. \tag{2.39}$$

We see that the improper integral and its derivative w.r.t. t are uniformly convergent in a compact neighbourhood of M, and hence the projection map is continuously differentiable, which implies local Lipschitz continuity. □

Lemma 2.2.4. *If R^{-1} is a Lipschitz function in P, then $\mathrm{Pr}_{\mathrm{ver}R}$ is locally Lipschitz.*

Proof. Taking square root can be chosen to be Lipschitz. Thus by 2.18, we only need $\mathrm{Pr}_{\mathrm{ver}}$ to be locally Lipschitz and reduce the case to the standard metric. □

2.3 Invariant control system

By varying metric on $GL(n)$, we have a control system on the homogeneous space $\mathcal{S}_+(n,n)$:

$$\begin{cases} \dot{P}_t & = J^{R_t}(P_t), \\ R_t & \in \mathcal{S}_+(n,n). \end{cases} \tag{2.40}$$

Or reformulated as a left invariant control system:

$$\dot{P} \in \mathfrak{a}(P). \tag{2.41}$$

Where $\mathfrak{a}$ is an invariant set of vector fields, and $\mathfrak{a}(P)$ is its restriction to $T_P\mathcal{S}_+(n,n)$.

Invariant control systems already exist in the literature. For introduction, see [Sac00]. We shall only introduce the version tailored for our case for the time being.

Typical control theory asks whether the system is *controllable*, i.e. whether the *attainable*, or *reachable* set, the set of points in the space reachable via the controls from a point P_0, is the whole space. We shall see this is not the case for our system, in which we always increase the Loewner partial order ($A > B$ if $A - B$ is positive definite) as time increases. One other nuance is that the condition on the control could affect the controllability. Originally we considered R_t being smooth. A natural choice from the point of view of invariant system is *regulated* functions, i.e. limits of piecewise constant functions. The conclusions for the control system so far is organized in 2.0.1 and 2.0.3, which we begin proving.

Proof. (2.0.1) 2 and 3 are discussed in the previous section. We only need to derive the formula in 1.

Since $\mathfrak{a}$ is left invariant, written as symmetric matrices, $\mathfrak{a}(P) = \bar{L}_{M*}\mathfrak{a}(I) = M\mathfrak{a}(I)M^*$ for $P = MM^*$, where $\bar{L}_G$ is the left $GL(n)$ action inherited from left multiplication on $GL(n)$.

$\mathfrak{a}(I)$ is the set of vectors $J(P)$ pulled back from each $P = MM^*$ via $\bar{L}_{M*}$. Written as symmetric matrices, using 1.4.3,

$$\begin{aligned}\mathfrak{a}(I) &= \{M^{-1}J(MM^*)M^{-*}|M \in GL(n)\}\\ &= \{V\,\mathrm{Diag}(\sum_{j\neq 1}\frac{1}{\lambda_1+\lambda_j},\cdots,\sum_{j\neq i}\frac{1}{\lambda_i+\lambda_j},\cdots)V^*|\lambda_i > 0,\ V \in O(n)\}. \quad (2.42)\end{aligned}$$

□

Proof. (2.0.3 1) From 1 in 2.0.1, we see $\dot{P}_t$ is always positive definite. □

Remark 2.3.1. This proves that we can not have $\dot{P} = J^R(P) = P_\infty - P$ for arbitrary P_∞

as we initially hoped. In fact, let

$$\alpha : (\lambda_1, \cdots, \lambda_n) \mapsto (\sum_{j\neq 1} \frac{1}{\lambda_1 + \lambda_j}, \cdots, \sum_{j\neq i} \frac{1}{\lambda_i + \lambda_j}, \cdots); \tag{2.43}$$

then the image is contained in the convex cone spanned by $e_i + e_j$, where $\{e_i\}_{i=1,\cdots,n}$ is the standard basis for $\mathbb{R}^n$. This is observed via writing

$$(\sum_{j\neq 1} \frac{1}{\lambda_1 + \lambda_j}, \cdots, \sum_{j\neq i} \frac{1}{\lambda_i + \lambda_j}, \cdots) = \sum_{i<j} \frac{e_i + e_j}{\lambda_i + \lambda_j} \tag{2.44}$$

Furthermore, the image is open in $\mathbb{R}^n$.

Lemma 2.3.2. *For $n > 2$, the Jacobian matrix $d\alpha$ is (strictly) negative-definite everywhere in the positive cone. In particular, Image(α) is open in $\mathbb{R}^n$.*

Proof. The Jacobian matrix

$$d\alpha = -\begin{bmatrix} \sum_{j\neq 1} \frac{1}{(\lambda_1+\lambda_j)^2} & \frac{1}{(\lambda_1+\lambda_2)^2} & \frac{1}{(\lambda_1+\lambda_3)^2} & \cdots & \frac{1}{(\lambda_1+\lambda_n)^2} \\ \frac{1}{(\lambda_1+\lambda_2)^2} & \sum_{j\neq 2} \frac{1}{(\lambda_2+\lambda_j)^2} & \frac{1}{(\lambda_2+\lambda_3)^2} & \cdots & \frac{1}{(\lambda_2+\lambda_n)^2} \\ \vdots & & \ddots & & \vdots \\ \frac{1}{(\lambda_1+\lambda_n)^2} & \cdots & \cdots & \cdots & \sum_{j\neq n} \frac{1}{(\lambda_n+\lambda_j)^2} \end{bmatrix} \tag{2.45}$$

Claim 2.3.3. $d\alpha = -\sum_{k=1}^{n-1} M_k M_k^*$ is a sum of "squares", where $rank(M_k) = n - k$.

The is achieved by taking

$$\tilde{M}_k = \begin{bmatrix} 0 & \frac{1}{\lambda_k+\lambda_{k+1}} & \frac{1}{\lambda_k+\lambda_{k+2}} & \cdots & \frac{1}{\lambda_k+\lambda_{k+n}} \\ \frac{1}{\lambda_k+\lambda_{k+1}} & & & & \\ & \frac{1}{\lambda_k+\lambda_{k+2}} & & & \\ & & \ddots & & \\ & & & & \frac{1}{\lambda_k+\lambda_{k+n}} \end{bmatrix} \tag{2.46}$$

(zero when not indicated)

and then let

$$M_k = \begin{bmatrix} 0_{k-1\times k-1} & \\ & \tilde{M}_k \end{bmatrix} \tag{2.47}$$

We see that

$$M_k M_k^* = \begin{bmatrix} \sum_{j=k+1}^n \frac{1}{(\lambda_k+\lambda_j)^2} & \frac{1}{(\lambda_k+\lambda_{k+1})^2} & \cdots & \cdots & \frac{1}{(\lambda_k+\lambda_n)^2} \\ \frac{1}{(\lambda_k+\lambda_{k+1})^2} & \frac{1}{(\lambda_k+\lambda_{k+1})^2} & & & \vdots \\ \vdots & & \frac{1}{(\lambda_k+\lambda_{k+2})^2} & & \vdots \\ \vdots & & & \ddots & \vdots \\ \frac{1}{(\lambda_k+\lambda_n)^2} & & & \frac{1}{(\lambda_k+\lambda_n)^2} & \end{bmatrix}, \tag{2.48}$$

which sums to $-d\alpha$. Indeed, each off-diagonal entry is nonzero in $M_{min\{k,j\}}$. On the diagonal, the (k,k)-th entry is

$$\frac{1}{(\lambda_1+\lambda_k)^2} + \cdots + \frac{1}{(\lambda_{k-1}+\lambda_k)^2} + \sum_{j=k+1}^{n} \frac{1}{(\lambda_k+\lambda_j)^2} + 0 = \sum_{j\neq k} \frac{1}{(\lambda_k+\lambda_j)^2} \tag{2.49}$$

Now $d\alpha$ is negative semidefinite. If it has a zero eigenvalue, then there exists a left-eigenvector x such that

$$-\sum_{k=1}^{n-1} xM_kM_k^*x^* = 0, \tag{2.50}$$

implying

$$(xM_k)(M_k^*x^*) = 0, \tag{2.51}$$

or

$$xM_k = 0 \tag{2.52}$$

for each k. Let

$$x = \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \end{bmatrix} \tag{2.53}$$

be a left null vector of M_1, which has nullity 1. But

$$xM_{n-1} = -\frac{1}{\lambda_{n-1} + \lambda_n} \neq 0, \tag{2.54}$$

a contradiction (unless $n = 2$). Hence $d\alpha$ is nonsingular. □

Lastly, we introduce some concepts from invariant control theories and make an estimate for the explorable space by the system 2.40.

Definition 2.3.4. Let G be a Lie group. $Lie(G) = \mathfrak{g}$ is identified with left invariant vector fields. A *left invariant control system* on G is of the form

$$\begin{cases} \dot{m}_t = m_t V_t \\ V_t \in \mathfrak{c}|_{m_t} \subset \mathfrak{g}|_{m_t}. \end{cases} \tag{2.55}$$

Here $\mathfrak{c}$ is a subset of the set of invariant vector fields $\mathfrak{g}$, and "$|_{m_t}$ denotes their restriction to the tangent space $T_{m_t}G$.

Definition 2.3.5. A *piecewise constant control* is such that for $0 = t_0 < t_1 < \cdots < T$, $V_t = V_i \in \mathfrak{c}$ for $t \in [t_i, t_{i+1})$ and some V_i.

A path with piecewise constant control is of the form

$$m(t) = m(0) \exp t_1 V_0 \exp t_2 V_1 \cdots \exp(t - t_i) V_i, \tag{2.56}$$

if $t \in [t_i, t_{i+1})$.

Definition 2.3.6. A *regulated control* $V_t \in \mathfrak{c}$ for $t \in [0, T]$ is a pointwise limit of piecewise constant controls.

Definition 2.3.7. The piecewise reachable set $\mathcal{R}^{pc}(m_0)$ is the union of the paths starting from m_0 with piecewise constant controls within finite time. Similarly, the regulated reachable set $\mathcal{R}^{reg}(m_0)$ is the same with regulated controls.

The control system 2.40 can be lifted to $GL(n)$

$$\begin{cases} \dot{M} = -c\vec{H}^R \\ -\vec{H}^R \in \hat{\mathfrak{a}}, \end{cases} \tag{2.57}$$

where the supscript R means the mean curvature $\vec{H}$ of the fibers under the metric Tr_R and $\hat{\mathfrak{a}} = GL(n) \cdot (-H)$ is the orbit of negative mean curvature field $-H$ under $GL(n)$ left multiplication, which maps to $\mathfrak{a}$ surjectively.

A path with piecewise control

$$P(t) = M(0) \exp t_1 V_0 \exp t_2 V_1 \cdots \exp(t - t_i) V_i \exp(t - t_i) V_i^* \cdots \exp t_1 V_0^*, \tag{2.58}$$

if $t \in [t_i, t_{i+1})$ and $P(0) = M(0)M(0)^*$.

As we have seen in 2.42, $\hat{\mathfrak{a}}$ consists of matrices whose SVD only need to satisfy conditions on the singular values. Hence V_i's can always be chosen to be normal with same eigenspaces, and their exponentials commute. In this case, the only thing that needs to be specified is the eigenvalue.

In general, I have no knowledge of any results existing about how in general eigenvalues are affected by multiplications.

Proof. 3. As discussed, we only need to estimate what eigenvalues are allowed in the reachable set.

Since regulated functions are the limits of piecewise constant paths, we may take a boundary point $c(e_i+e_j)$, corresponding to $\lambda_i+\lambda_j = 1/c$ and $\lambda_k = \infty$ otherwise. Since diagonal matrices commute, the exponential of these commute. $\square$

Bibliography

[BJL19] Rajendra Bhatia, Tanvi Jain, and Yongdo Lim. On the bures–wasserstein distance between positive definite matrices. *Expositiones Mathematicae*, 37(2):165–191, 2019.

[BW20] Fabrice Baudoin and Jing Wang. Brownian motions and eigenvalues on complex grassmannian manifolds. 2020.

[GQ20] Jean Gallier and Jocelyn Quaintance. *Differential geometry and Lie groups: a computational perspective*, volume 12. Springer Nature, 2020.

[MA20] Estelle Massart and P-A Absil. Quotient geometry with simple geodesics for the manifold of fixed-rank positive-semidefinite matrices. *SIAM Journal on Matrix Analysis and Applications*, 41(1):171–198, 2020.

[MHA19] Estelle Massart, Julien M Hendrickx, and P-A Absil. Curvature of the manifold of fixed-rank positive-semidefinite matrices endowed with the bures–wasserstein metric. In *International Conference on Geometric Science of Information*, pages 739–748. Springer, 2019.

[Pac03] Tommaso Pacini. Mean curvature flow, orbits, moment maps. *Transactions of the American Mathematical Society*, 355(8):3343–3357, 2003.

[RW00] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales*, volume 2 of *Cambridge Mathematical Library*. Cambridge University Press, 2 edition, 2000.

[Sac00] Yu L Sachkov. Controllability of invariant systems on lie groups and homogeneous spaces. *Journal of Mathematical Sciences*, 100(4):2355–2427, 2000.

[Ura79] Hajime Urakawa. On the least positive eigenvalue of the laplacian for compact group manifolds. *Journal of the Mathematical Society of Japan*, 31(1):209–226, 1979.