ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE STATISTICHE

Ciclo 38

Settore Concorsuale: 01/A3 - ANALISI MATEMATICA, PROBABILITÀ' E STATISTICA MATEMATICA

Settore Scientifico Disciplinare: MAT/06 - PROBABILITA' E STATISTICA MATEMATICA

STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS AND COVARIANCE MATRICES: ANALYTICAL AND GEOMETRIC PERSPECTIVES

Presentata da: Leonardo Marconi

Coordinatore Dottorato

Angela Montanari

Supervisore

Enrico Bernardi

Co-supervisore

Alberto Lanconelli

Esame finale anno 2026

Stochastic Partial Differential Equations and Covariance Matrices: Analytical and Geometric Perspectives

Leonardo Marconi

Supervisor: Prof. Enrico Bernardi

PhD Thesis in Statistics

University of Bologna

Preface

The present thesis follows my personal journey across mathematics and statistics. I began this path with my Master's thesis, under the guidance of Professor Enrico Bernardi, with a focus on stochastic analysis and partial differential equations. Stochastic partial differential equations (SPDEs) occupy a remarkable position at the intersection of several major branches of mathematics — analysis, probability, geometry, and mathematical physics — and provide a natural framework to model the evolution of random phenomena in space and time. Their theory combines the analytical depth of classical partial differential equations (PDEs) with the stochastic intuition of random processes, leading to problems of both great technical difficulty and conceptual richness.

The study of SPDEs proved to be both challenging and intellectually rewarding. It allowed me to explore the delicate interplay between randomness and structure, to appreciate how analytical tools such as functional analysis, semigroup theory, and variational methods interact with probabilistic notions like martingales, stochastic integrals, and noise regularity. More importantly, it helped me to develop the rigorous and versatile toolkit that is essential to the work of an applied mathematician: the ability to move between deterministic and stochastic reasoning, between local differential structures and global geometric interpretations.

As my research evolved and I came in touch with mathematical and theoretical statistics, these analytical foundations naturally led me toward problems where geometry and statistics meet. Geometric statistics is a relatively young and rapidly growing field. What makes this area so appealing is the elegance with which it connects modern statistical challenges — arising in machine learning, signal processing, and data analysis — with some of the most beautiful and profound ideas in differential geometry and analysis. Within this framework, abstract mathematical concepts such as curvature, geodesics, and connections become powerful tools for understanding variability, uncertainty, and dependence in complex data. This confluence of theory and application provides a uniquely fertile ground where mathematical depth and practical relevance coexist, and where the intrinsic beauty of geometry translates into new insights for modern statistics.

This transition from stochastic equations to geometric statistics is underpinned by a passionate search for intrinsic structure: how randomness, geometry, and data interact to form coherent mathematical objects.

As a final note, I would like to express my deepest gratitude to my supervisor, Professor Enrico Bernardi, for his expert and invaluable guidance throughout these years. His mentorship has been instrumental in my growth — from a curious and passionate student to a researcher who loves his work and continues to learn every day.

I am also sincerely grateful to Professor Alberto Lanconelli for his thoughtful discussions and wise advice, which have been a source of clarity and encouragement during the most challenging moments of this journey.

Contents

Chapter 1

Introduction

The present thesis is devoted to the study of two distinct but conceptually related problems arising in differential geometry, matrix analysis, and the theory of partial differential equations.

The first part turns to the analysis of stochastic partial differential equations of hyperbolic type with multiple characteristics. Such equations model the propagation of random perturbations through systems governed by wave-like dynamics. While the theory of strictly hyperbolic stochastic equations is well established, the presence of multiple or non-effectively hyperbolic characteristics introduces new analytical challenges, including loss of regularity, delicate propagation of singularities, and intricate interactions between the principal symbol and the stochastic forcing. Building on deterministic microlocal techniques and recent developments in the stochastic framework, we examine three model operators that illustrate different geometric configurations of the characteristic manifold — symplectic, involutive, and mixed — and study the corresponding random-field solutions, emphasizing the influence of the geometric structure of the symbol on the probabilistic behaviour of the solution.

The second part of this thesis develops a novel geometric framework for the manifold of fixed-rank covariance matrices endowed with the Bures–Wasserstein metric. In recent years, the geometry of covariance matrices has attracted considerable attention across differential geometry, statistics, and information theory. Several metrics have been proposed to endow the cone of positive semidefinite matrices with a meaningful geometric structure, including the affine-invariant, log–Euclidean, and Bures–Wasserstein metrics, each providing different insights into data variability and statistical inference. Among these, the Bures–Wasserstein metric occupies a central role, as it naturally arises both in quantum information theory and in optimal transport between Gaussian measures. Its rich geometric properties make it particularly suitable for studying covariance evolution, interpolation, and estimation. In this work, we investigate how the Bures–Wasserstein metric adapts on a diffeomorphic model originally introduced in [27]. We derive this representation rigorously from a differential geometric standpoint. This allows us to characterize the manifold as an associated fiber bundle with completely geodesic fibers and to compute an ODE system for geodesic equations. Moreover, the framework developed here reveals a natural correspondence between the geometry of covariance matrices and that of the Grassmannian. This connection not only deepens the geometric understanding of covariance structures but also opens the way to new statistical constructions on spaces of low-rank covariances.

Although these topics belong to different mathematical domains, they are unified by a common focus on geometric and structural properties: curvature, tangent spaces, and the behaviour of geodesics or characteristics as fundamental organizing principles.

1.1 Partial differential equations

Partial differential equations describe relations between a multivariable function and its partial derivatives. They are the fundamental analytical tool for modeling propagation, diffusion, equilibrium, and oscillatory phenomena in physics, engineering, and applied sciences. A linear differential operator of order m on $\mathbb{R}^{n+1}$ with smooth coefficients is written as

$$\sum_{|\alpha|\leq m} a_\alpha(x) D^\alpha,$$

where the a_α are smooth functions and $D^\alpha := (-i\partial)^\alpha$ for any multi-index α (Appendix No. 1). We call *symbol* the function

$$p(x,\xi) = \sum_{|\alpha|\leq m} a_\alpha(x)\xi^\alpha$$

and *principal symbol* the function

$$p_m(x,\xi) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha.$$

The study of conditions for the well-posedness of the Cauchy problem is connected to the nature of the characteristic roots of the *characteristic equation*

$$p_m(x,\xi) = 0. \tag{1.1}$$

The set of points (x,ξ) for which equation (1.1) holds true is called the *characteristic manifold*

$$\Sigma = \{(x,\xi) \,|\, p_m(x,\xi) = 0\}. \tag{1.2}$$

The *singular support* of a distribution u (Appendix No. 4) is the smallest closed set outside which u is smooth. It localizes the non-smoothness of u in the physical space. The *wavefront set* $\mathcal{WF}(u) \subset T^*\mathbb{R}^{n+1} \setminus \{0\}$ (Appendix No. 17) refines this notion by recording both the location and the direction of singularities: $(x_0,\xi_0) \notin \mathcal{WF}(u)$ if there exists $\phi \in C_c^\infty$, where C_c^∞ denotes the set of smooth functions with compact support, with $\phi(x_0) \neq 0$ such that the Fourier transform $[\mathcal{F}(\phi u)](\xi)$ (Appendix No. 2) decays faster than any power of $|\xi|$ in a conic neighbourhood of ξ_0.

The *Hamiltonian vector field* of an operator is defined as

$$H_{p_m} = \sum_{j=0}^{n} \left(\frac{\partial p_m}{\partial \xi_j} \frac{\partial}{\partial x_j} - \frac{\partial p_m}{\partial x_j} \frac{\partial}{\partial \xi_j} \right),$$

and its integral curves $(x(t),\xi(t))$ satisfy

$$\dot{x} = \frac{\partial p_m}{\partial \xi}, \qquad \dot{\xi} = -\frac{\partial p_m}{\partial x}.$$

These curves are called *bicharacteristics*. The nature of the characteristic roots of the characteristic equation (1.1) distinguishes the different kinds of differential operators. The operator P is said to be *elliptic* at x if $p_m(x,\xi) \neq 0$ for all $\xi \neq 0$. Elliptic equations (such as the Laplace, Poisson, or Helmholtz equations) describe equilibrium and potential phenomena. Ellipticity ensures analytic regularity and the absence of finite-speed propagation: if $Pu = f$ and f is smooth, then u is smooth. Solutions are globally influenced by boundary data rather than propagating along characteristics ([48]).

When the symbol vanishes for some real $\xi \neq 0$, the operator is *non-elliptic*, and its behaviour is governed by the geometry of the characteristic manifold (1.2).

If the principal symbol $p_m(x,\xi)$ has only real roots in ξ, the operator is called *hyperbolic* with respect to ξ. Hyperbolic equations describe finite-speed propagation and wave phenomena ([36], [52]). When the roots have nonzero imaginary parts with a sign constraint (e.g., $\Im\lambda_j \geq 0$), the equation is *parabolic*, modeling diffusion or heat flow.

The propagation of singularities for hyperbolic equations follows the bicharacteristics of $p_m(x,\xi)$, i.e., the integral curves of the Hamiltonian system, and the geometry of the characteristic manifold Σ — symplectic, involutive, or mixed — governs the analytic behaviour of solutions, as studied in microlocal and semiclassical analysis ([52], [108]).

Furthermore, hyperbolic equations can be further classified based on the multiplicity of the roots of equation (1.1).

- If the roots are simple the operator is called *strictly* hyperbolic. This is the simplest case, as well-posedness of the Cauchy problem is guaranteed in C^∞ and clean microlocal properties hold: finite propagation speed, propagation of singularities along well-defined bicharacteristics.

- If the roots are multiple for some $(x,\xi) \neq 0$, we call the operator *non-strictly* hyperbolic. This gives rise to complications: solutions may lose derivatives, singularities may propagate in intricate ways, and lower-order terms may affect well-posedness of the Cauchy problem. In such cases, the geometry of Σ — in particular its symplectic or involutive structure — determines how singularities are transmitted, reflected, or possibly trapped.

We say that the operator has *n-tuple characteristics* if the principal *symbol* vanishes exactly to the n-th order on its characteristic manifold.

The *Hamilton map* is defined as

$$F_p(\rho) = \frac{1}{2} dH_p(\rho) = \frac{1}{2}\begin{bmatrix} \frac{\partial^2 p}{\partial x \partial \xi}(\rho) & \frac{\partial^2 p}{\partial \xi^2}(\rho) \\ -\frac{\partial^2 p}{\partial x^2}(\rho) & -\frac{\partial^2 p}{\partial \xi \partial x}(\rho) \end{bmatrix}.$$

Among non-strictly hyperbolic operators, one distinguishes those that are *effectively* hyperbolic — those for which the Hamilton map of the principal symbol has at least one pair of nonzero real eigenvalues ([52], [82], [87]). In this case, perturbations propagate with finite speed and the Cauchy problem is generally well-posed.

When all eigenvalues of the Hamilton map are purely imaginary, the operator is *non-effectively* hyperbolic. These cases typically lead to instability or ill-posedness, as singularities may fail to propagate or may amplify uncontrollably. The fine structure of the bicharacteristics and of lower-order terms then determines whether any generalized notion of well-posedness can be recovered.

1.2 Stochastic partial differential equations

Stochastic partial differential equations extend classical PDEs by incorporating random forcing terms, thus modeling systems affected by uncertainty or noise. Formally, an SPDE has the structure

$$P(x, D)u = f + \Phi \dot{F},$$

where $P(x, D)$ is a deterministic differential operator, $\dot{F}$ denotes a generalized Gaussian noise (often white in time and possibly correlated in space), and Φ is a linear operator controlling how noise enters the system. The term $\Phi \dot{F}$ is interpreted in the sense of random distributions. The class to which the deterministic differential operator belongs determines different *peculiarities* of the stochastic problem. In particular, for hyperbolic operators the stochastic term excites oscillatory modes that travel along the deterministic bicharacteristics of the symbol. For strictly hyperbolic operators, propagation of singularities remains deterministic, while randomness affects amplitudes and regularity ([32], [38]).

When the operator has multiple characteristics, stochastic perturbations interact with degeneracies in the symbol, producing distinct regimes: in the *effectively hyperbolic* case, random perturbations propagate along well-defined wavefronts; in the *non-effectively hyperbolic* case, the interplay between noise and geometric degeneracy can yield loss of predictability or explosive variance. There are two main theories one may use to tackle a stochastic differential problem. The first consists in setting the SPDE as an evolution in a Hilbert space H, namely

$$dU(t) = \big(AU(t) + F(U(t))\big)\, dt + B(U(t))\, dW_H(t),$$

where A generates a C_0 semigroup, W_H is a cylindrical Wiener process, and solutions are *mild* via the variation-of-constants formula ([37], [95], [97]). This approach is natural for parabolic problems (heat, reaction–diffusion, Navier–Stokes) and for coloured noise.

As for the second one, which is the one we use in the sequel, fundamental solutions are computed and *random field* solutions are constructed via stochastic convolutions with the Green kernel. Formally, if $E(t, x)$ *is* the fundamental solution (Appendix No. 9) of the operator P, we will say that $u(t, x)$, defined as

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} E\,(t - s, x - y)\,\dot{F}(s, y)\,ds\,dy, \tag{1.3}$$

is a *random field solution* to *the SPDE* if the stochastic integral is well defined and the map $(t, x) \mapsto u(t, x)$ is measurable. For the theory of stochastic integration, we refer to [38], where the author performs an extension and adaptation of Walsh's construction of the martingale measure stochastic integral (see [110]). We will now shortly describe it.

Let $\mathcal{B}_b(\mathbb{R}^d)$ be the σ-field of all bounded Borel sets of $\mathbb{R}^d$ and define

$$F(\phi) = \int_{\mathbb{R}^{d+1}} \phi(t, x)\dot{F}(t, x)dt\,dx.$$

The first step is to extend F to a *worthy martingale measure* (see [110]). Using (2.2) we can verify that F is L^2-continuous (Appendix No. 6) and, as such, we can extend it to a σ-finite L^2-valued measure by approximating indicator functions of sets in $\mathcal{B}_b(\mathbb{R}_+ \times \mathbb{R}^d)$ with elements of $\mathcal{D}\left(\mathbb{R}^{d+1}\right)$ (Appendix No. 3). We set

$$M_t(B) = F([0, t] \times B), \quad B \in \mathcal{B}_b(\mathbb{R}^d)$$

and

$$\mathcal{F}_t^0 = \sigma\left(M_s(B),\ s \leq t,\ B \in \mathcal{B}_b(\mathbb{R}^d)\right), \quad \mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N},$$

with $\mathcal{N}$ being the σ-field generated by the $\mathbb{P}$-null sets. We then have that, by construction, $t \mapsto M_t(B)$ is a continuous martingale and

$$F(\phi) = \int_{\mathbb{R}^+} \int_{\mathbb{R}^d} \phi(t, x)M(dt, dx), \quad \phi \in \mathcal{D}(\mathbb{R}^{d+1}).$$

We call a function $(s,x,\omega)\mapsto g(s,x,\omega)$ *elementary* if it is of the form

$$g(s,x,\omega)=1_{(a,b]}(s)\,1_A(x)\,X(\omega),\quad a,b\in\mathbb{R},\ 0\le a<b,\ A\in\mathcal{B}_b(\mathbb{R}^d),$$

where X is a $\mathcal{F}_a$-measurable r.v. We denote by $\mathcal{E}$ the space of all finite linear combinations of elementary functions and call *predictable* the σ-field on $\mathbb{R}_+\times\mathbb{R}^d\times\Omega$ generated by the elements of $\mathcal{E}$. Moreover, set

$$\|g\|_+ := E\left(\int_0^t ds\int_{\mathbb{R}^d}dy\int_{\mathbb{R}^d}dx\,|g(s,x,\cdot)|\,f(x-y)\,|g(s,y,\cdot)|\right) \tag{1.4}$$

and

$$\|g\|_0 := E\left(\int_0^t ds\int_{\mathbb{R}^d}dy\int_{\mathbb{R}^d}dx\,g(s,x,\cdot)\,f(x-y)\,g(s,y,\cdot)\right). \tag{1.5}$$

In [110], Walsh defines the martingale-measure stochastic integral on the complete space $\mathcal{P}_+$ of predictable functions g with $\|g\|_+<+\infty$.

On the other hand, Dalang in his work [38] is able to extend such a construction, defining the martingale

$$t\mapsto\int_0^t\int_{\mathbb{R}^d}v(s,x,\cdot)\,M(ds,dx)$$

for all elements v of the completion $\mathcal{P}_0$ of $(\mathcal{E},\|\cdot\|_0)$. Moreover, denoting as $\bar{\mathcal{P}}$ the space of all predictable functions $h(t,x,\omega)$ such that $x\mapsto h(t,x,\omega)\in\mathcal{S}'(\mathbb{R}^d)$ for every $(t,\omega)\in[0,T]\times\Omega$, $\mathcal{F}h(t,\cdot,\omega)(\xi)$ is a function a.s., and

$$\|h\|_0' := E\left(\int_0^t ds\int_{\mathbb{R}^d}|\mathcal{F}h(t,\cdot,\omega)(\xi)|^2\,\nu(d\xi)\right)^{1/2}<+\infty, \tag{1.6}$$

and calling $\mathcal{E}_0$ the subset of $\mathcal{P}_+$ consisting of all functions $g(s,x,\omega)$ such that $x\mapsto g(s,x,\omega)\in\mathcal{S}(\mathbb{R}^d)$, then we can identify $\mathcal{P}_0$ with the set of elements $h(t,x,\omega)$ of $\bar{\mathcal{P}}$ such that *there* exists a sequence $h_n(t,x,\omega)$ of elements of $\mathcal{E}_0$ that gives

$$\lim_{n\to\infty}\|h_n-h\|_0'=0. \tag{1.7}$$

Furthermore, we observe that this implies that, for elements $h(s,x,\omega)\in\mathcal{P}_0$,

$$\|h\|_0=\|h\|_0'.$$

Dalang's extension gives existence for spatially correlated noises beyond white noise in dimensions where the heat/wave kernels are not square-integrable ([38]). This approach matches hyperbolic SPDEs and pathwise space-time analysis ([32], [62]).

When semigroup kernels are integrable against the noise covariance, mild and random-field solutions coincide; otherwise one can still compare them via isometries in $L^2(\Omega)$ ([37], [38], [95]).

1.3 Geometric statistics and the geometry of covariance matrices

1.3.1 Geometric statistics

Classical statistics assumes that observations take values in a linear vector space, typically $\mathbb{R}^n$, and that variability can be represented through linear operations such as addition and scalar multiplication. However, many modern data types do not possess a natural linear structure. Examples include directions on a sphere, orientations in robotics, shapes up to rigid motions, positive semidefinite covariance matrices, and probability distributions endowed with transport metrics. In such cases, one must replace linear geometry with the intrinsic geometry of the underlying data space.

The study of statistical analysis on nonlinear spaces constitutes the field of *geometric statistics* ([45], [46], [61], [78], [93]). This area generalizes Euclidean methods by using concepts from differential geometry, metric geometry, and optimal transport.

Let $(\mathcal{X}, d)$ be a complete metric space and X a random variable with values in $\mathcal{X}$. The intrinsic notion of expectation is given by the *Fréchet functional* ([46])

$$F(x) = \mathbb{E}[d^2(x, X)], \qquad x \in \mathcal{X}.$$

A minimizer m of F is called a *Fréchet mean* or *intrinsic mean*. In Euclidean spaces, this reduces to the ordinary mean $\mathbb{E}[X]$. On curved spaces, existence and uniqueness of m depend on the convexity of geodesic balls, curvature bounds, and the injectivity radius ([3], [56], [60]).

The corresponding empirical Fréchet mean $\hat{m}_n$ minimizes $\sum_{i=1}^n d^2(x, X_i)$. Under regularity and curvature assumptions, a central limit theorem holds in the tangent space $T_m\mathcal{X}$:

$$\sqrt{n}\,\log_m(\hat{m}_n) \Rightarrow \mathcal{N}(0, \Sigma_{\text{eff}}),$$

where $\log_m$ denotes the Riemannian logarithm and Σ_{eff} depends on the geometry and the covariance structure of the data ([22], [23]).

When $\mathcal{X}$ is a smooth Riemannian manifold (M, g), one can exploit the differential structure to define intrinsic operations such as exponential and logarithmic maps, parallel transport, and covariant differentiation. Statistical models, estimation, and inference are formulated intrinsically on M, avoiding the distortions introduced by linear embeddings ([45], [91], [93]).

Key tasks in Riemannian statistics include:

- *Estimation of Fréchet means*;
- *Principal geodesic analysis*, the nonlinear analogue of Principal Component Analysis (PCA), where variability is decomposed along geodesics ([45], [55]);
- *Regression and dynamics on manifolds*, based on geodesic interpolation and parallel transport [83];
- *Statistical models on Lie groups and homogeneous spaces*, which exploit group structure for inference on orientations and shapes ([33]).

These tools have been applied to a wide range of data, including anatomical shapes, diffusion tensors, covariance matrices, and probability distributions on manifolds.

Many data spaces are not smooth manifolds globally but have stratified or quotient structures. Examples include shape spaces modulo rotation and scaling ([61]), Grassmannians of subspaces, and fixed-rank cones of covariance matrices ([16], [80]).

Geometric statistics extends to such spaces using tools from metric geometry and convex analysis. Spaces of nonpositive curvature (CAT(0) or RCD(K, N) metric-measure spaces) admit unique geodesics and intrinsic barycenters ([73], [102]). Stratified spaces, such as the Billera–Holmes–Vogtmann tree space ([24]), combine piecewise-Riemannian structure with combinatorial singularities, giving rise to phenomena such as "sticky" Fréchet means ([54]).

Another major framework is *information geometry*, where the Fisher–Rao metric endows families of probability distributions with a dual affine structure ([5], [6]). The resulting geometry captures natural gradients and connects differential geometry with likelihood-based inference.

1.3.2 Covariance matrices

Covariance matrices play a central role in multivariate statistics, signal processing, and diffusion tensor imaging.

Typical statistical tasks involving covariance matrices include estimation ([17], [31], [66]), shrinkage and regularization ([67], [111]), hypothesis testing on covariance structure, and inference on low-rank or structured covariance models ([13]). The smooth manifold of fixed-rank covariance *matrices* is very relevant in applications as well. Interpolation in this manifold is used in [81] for parametric order reduction and in [14] for community detection. Optimization algorithms for various problems are proposed in [26], [83] (distance learning), [76] (distance matrix completion) and [85] (role model extraction). Other studies have tackled the task of fitting a curve to points on the manifold. In [49] the authors fit a curve through covariance matrices that model a wind field. In a related vein, in [71] the authors use geodesic interpolation between Gram matrices to create intermediate protein conformations, while reference [59] employs Gram-matrix interpolation to track facial-expression changes in video data.

Geometric statistics on covariance matrices now underpins applications in diffusion tensor imaging ([43], [44], [47]), radar signal processing ([15]), finance ([29]), and quantum state estimation ([28]). Affine-invariant and Bures–Wasserstein means provide robust, congruence-invariant estimators of average covariance; geodesic PCA captures dominant modes of variability; and curvature analysis informs concentration and inference.

Ongoing research focuses on:

- extending asymptotic theory to stratified and quotient spaces ([16], [79], [106]);
- developing curvature-aware statistical tests and confidence regions;
- connecting optimal transport geometry with high-dimensional covariance estimation ([96], [109]);
- and exploring stochastic dynamics and SPDEs on covariance manifolds ([47], [93]).

The set of n-dimensional covariance matrices $\mathrm{Cov}(n)$ is defined as the set of positive semidefinite symmetric matrices

$$\mathrm{Cov}(n) := \{\Sigma \in \mathrm{Sym}(n) \,|\, \Sigma \geq 0\}\,,$$

where $\mathrm{Sym}(n)$ denotes the space of symmetric $n \times n$ matrices. The set $\mathrm{Cov}(n)$ forms a closed convex cone in the Euclidean space of square n-dimensional matrices $\mathrm{Mat}(n)$. This convex cone (Appendix No. 11) can be described as a stratified space, i.e., a topological or metric space that can be decomposed into a family of subspaces that exhibit a smooth structure. The strata of $\mathrm{Cov}(n)$ are given by the smooth manifolds (Appendix No. 15) of n-dimensional covariance *matrices* of fixed rank k (see [106]), that is

$$\mathrm{Sym}^+(n,k) := \{\Sigma \in \mathrm{Cov}(n) \,|\, \mathrm{rk}(\Sigma) = k\}\,.$$

The set of n-dimensional covariance matrices $\mathrm{Cov}(n)$ can be identified with the quotient space $\mathrm{Mat}(n)/O(n)$, where $O(n)$ denotes the orthogonal group (Appendix No. 19). The orbits of the quotient space are defined by the right action $\mathrm{Mat}(n) \times O(n) \ni (X, G) \mapsto XG \in \mathrm{Mat}(n)$ and the identification is given by the map

$$\mathrm{Mat}(n) \ni X \mapsto XX^T \in \mathrm{Cov}(n). \tag{1.8}$$

For a stratum $\mathrm{Sym}^+(n,k)$ a parallel construction to the one of $\mathrm{Cov}(n)$ can be built ([79]). In this case, we focus on the right action $\mathrm{Mat}(n,k)^* \times O(k) \ni (X,G) \mapsto XG \in \mathrm{Sym}^+(n,k)$, where $\mathrm{Mat}(n,k)^*$ denotes the manifold of full-rank $n \times k$ matrices, and the consequent quotient space $\mathrm{Mat}(n,k)^*/O(k)$ which is diffeomorphic to $\mathrm{Sym}^+(n,k)$ via the map $\mathrm{Mat}(n,k)^* \ni X \mapsto XX^T \in \mathrm{Sym}^+(n,k)$. Thanks to this identification, the Euclidean metric descends to the Bures–Wasserstein Riemannian metric on $\mathrm{Sym}^+(n,k)$. For a more in-depth analysis we refer the reader to [79].

The set of covariance matrices can be equipped with different distances. A distance on $\mathrm{Cov}(n)$ translates then to a Riemannian structure on the smooth strata. We recall the definition of the tangent space to $\mathrm{Sym}^+(n,k)$ at Σ

$$T_\Sigma \,\mathrm{Sym}^+(n,k) = \{V \in \mathrm{Sym}(n) \,|\, P_\Sigma^\perp V P_\Sigma^\perp = 0\},$$

where $P_\Sigma^\perp$ denotes the orthogonal projector onto the nullspace of Σ. We present the most relevant examples of distances and correlated metrics on the tangent spaces. For simplicity the metrics are described only on the principal stratum

$$\mathrm{SPD}(n) := \{\Sigma \in \mathrm{Sym}(n) \,|\, \Sigma > 0\}.$$

- Defined by

  $$d_{\mathrm{AI}}(\Sigma, \Lambda) = \|\log(\Sigma^{-1/2}\Lambda\Sigma^{-1/2})\|_F,$$

  the *affine-invariant distance* is invariant under congruence transformations $\Sigma \mapsto Q\Sigma Q^\top$, $\Lambda \mapsto Q\Lambda Q^\top$ for $Q \in O(n)$. The associated metric,

  $$g_\Sigma(V, W) = \mathrm{Tr}(\Sigma^{-1}V\Sigma^{-1}W), \qquad V, W \in T_\Sigma\,\mathrm{SPD}(n),$$

  makes $\mathrm{SPD}(n)$ a globally nonpositively curved manifold isometric to the symmetric space $\mathrm{GL}(n)/\mathrm{O}(n)$ ([44], [69], [92], [100], [101]). Geodesics and parallel transport admit closed forms:

  $$\gamma_{\Sigma\to\Lambda}(t) = \Sigma^{1/2}\exp\left(t\log(\Sigma^{-1/2}\Lambda\Sigma^{-1/2})\right)\Sigma^{1/2} = \Sigma^{1/2}(\Sigma^{-1/2}\Lambda\Sigma^{-1/2})^t\Sigma^{1/2}.$$

  The affine-invariant metric underlies information-geometric approaches and Riemannian gradient-based optimization on $\mathrm{SPD}(n)$ ([2], [20], [91]).

- The *log-Euclidean metric* provides a computationally simpler alternative to the affine-invariant geometry while retaining key invariance properties ([7], [20], [43], [47]). It is defined by mapping $\mathrm{SPD}(n)$ diffeomorphically to the Euclidean space of symmetric matrices via the matrix logarithm:

  $$\phi : \mathrm{SPD}(n) \to \mathrm{Sym}(n), \qquad \Sigma \mapsto \log(\Sigma).$$

  Equipping $\mathrm{Sym}(n)$ with the Frobenius inner product $\langle A, B\rangle_F = \mathrm{Tr}(AB)$ and pulling it back through ϕ yields a flat Riemannian metric on $\mathrm{SPD}(n)$:

  $$g_\Sigma(V, W) = \langle D(\log)_\Sigma[V],\, D(\log)_\Sigma[W]\rangle_F = \mathrm{Tr}((\Sigma^{-1}V)(\Sigma^{-1}W)),$$

and the corresponding geodesic distance

$$d_{\log}(\Sigma, \Lambda) = \| \log \Sigma - \log \Lambda \|_F.$$

The log–Euclidean metric has zero curvature (the manifold becomes globally flat in the logarithmic domain) and allows closed-form formulas for the mean and geodesics. It is particularly useful in diffusion tensor imaging and machine learning, where it combines invariance under congruence transformations with numerical simplicity.

- Another geometric construction exploits the *Cholesky factorization* $\Sigma = L_\Sigma L_\Sigma^\top$ with $L_\Sigma \in \mathrm{Lower}(n)$ nonsingular ([20], [42], [70], [72]). By representing $\mathrm{SPD}(n)$ through its Cholesky factors, one induces the following Riemannian metric from the Euclidean structure of the space of lower-triangular matrices:

$$g_\Sigma(V, W) = \langle L_\Sigma^{-1} V (L_\Sigma^{-1})^\top,\, L_\Sigma^{-1} W (L_\Sigma^{-1})^\top \rangle_F, \qquad \Sigma = L_\Sigma L_\Sigma^\top.$$

This metric is flat and globally parameterizes $\mathrm{SPD}(n)$ without the nonlinearity of the exponential map. Distances are computed as

$$d_{\mathrm{Chol}}(\Sigma, \Lambda) = \|L_\Sigma - L_\Lambda\|_F,$$

where L_Σ, L_Λ are the Cholesky factors of Σ and Λ. Cholesky-type coordinates preserve positive definiteness by construction and are convenient for optimization, interpolation, and statistical modeling in high dimensions. Although they lack full affine invariance, they provide efficient and numerically stable approximations to the intrinsic Riemannian distances.

We equip $\mathrm{Cov}(n)$ with the *Bures–Wasserstein distance* ([21], [41], [75], [90], [104], [105]) defined as

$$d^{BW}(\Sigma, \Lambda)^2 = \mathrm{Tr}(\Sigma) + \mathrm{Tr}(\Lambda) - 2\,\mathrm{Tr}((\Sigma^{1/2} \Lambda \Sigma^{1/2})^{1/2}).$$

With this distance, $\mathrm{Cov}(n)$ becomes a complete geodesic space (i.e., a complete metric space where the distance between two points is equal to the length of a minimizing geodesic connecting the two points (Appendix No. 26)) of nonnegative curvature ([4], [106]). This metric is full of applications and of great interest in the literature. It was introduced both in quantum information theory as the Bures distance ([30], [51]) and in optimal transport as the L^2-Wasserstein distance ([41], [90]). In particular, in the latter context, the distance acquires the meaning of a measure of dissimilarity between Gaussian distributions ([75], [109]). From a geometric point of view, the Bures–Wasserstein distance can be derived as the quotient distance via the map (1.8) if we assign the Euclidean (or Frobenius) distance to $\mathrm{Mat}(n)$ (Appendix No. 12).

The Bures–Wasserstein distance translates to a metric on the smooth strata (Appendix No. 22). Now, following what is done in [106], we set for a moment $\Sigma = QDQ^T$ with $D = \mathrm{diag}(d_1, ..., d_k)$ positive definite and $Q \in \mathrm{St}(n,k)$ and call $S_{\Sigma,V} = QS_D(Q^T VQ)Q^T$, where $S_A(B) =: S$ denotes the unique solution to the Sylvester equation $B = AS + SA$ for a symmetric positive definite matrix A and a symmetric matrix B (Appendix No. 13). For $V, W \in T_\Sigma$, the Bures–Wasserstein metric is given by

$$g_\Sigma^{BW(n,k)}(V,W) = \mathrm{Tr}(S_{\Sigma,V}\Sigma S_{\Sigma,W}) + \mathrm{Tr}(P_\Sigma^\perp V \Sigma^\dagger W),$$

where $\Sigma^\dagger$ denotes the Moore–Penrose inverse of Σ. Furthermore, the authors point out that $S_{\Sigma,W}$ and the metric are independent of the chosen decomposition.

With the above geometry constructions, the strata of $\mathrm{Cov}(n)$ have been object of study in the *literature* for some years. The first results concerned the full-rank principal stratum (see among others [21], [75], [104], [105] and [107]), where, starting from the quotient geometry, it was possible to derive equations for the exponential map ([75]), unique logarithm (Appendix No. 27) ([21], [79]), parallel transport (Appendix No. 25) ([107]) and curvature (Appendix No. 30) ([104], [105], [107]). For the geometry of the stratum $\mathrm{Sym}^+(n,k)$ of fixed-rank covariance matrices we refer to [79], [80], [106]; in particular the quotient geometry was studied in [79], a formula for the curvature tensor and the sectional curvature, as well as curvature bounds, were expressed, thanks to O'Neill's equations ([89]), in [80], and finally the exponential map, its definition domain (Appendix No. 28), cut time (Appendix No. 29) and logarithms together with their multiplicity were found in [106]. Most recently, relevant progress has been made on the geometry (and in particular on the geodesics) of the stratified space as a whole ([106]).

Geometric statistics thus provides a unified language in which Riemannian geometry, optimal transport, and information geometry converge. Within this framework, covariance matrices are not mere algebraic objects but points on a curved manifold whose structure encodes the natural symmetries of variability and uncertainty. Understanding this geometry allows statistical inference to be intrinsic, coordinate-free, and consistent with the underlying transformations of the data.

1.4 Concluding remarks and plan

Throughout the thesis, particular attention is given to the rigorous formulation of the underlying geometric and analytical tools. The Appendix collects the necessary background on linear algebra, smooth manifolds, Riemannian geometry, and variational calculus, so as to make the presentation as self-contained as possible, while maintaining the desired level of mathematical precision.

The research presented in this thesis has led to two distinct works. The results contained in Chapter 2, concerning stochastic hyperbolic partial differential equations with multiple characteristics, have been published in a joint article with my supervisor ([19]). The material developed in Chapter 3, which introduces a new geometric framework for the manifold of fixed-rank covariance matrices under the Bures–Wasserstein metric, may be found on my preprint ([77]).

In summary, the guiding idea of this work is to explore how geometric structures — whether on spaces of matrices or on the phase space of partial differential equations — govern analytic phenomena and reveal new connections between geometry, probability, and analysis.

Chapter 2

On some stochastic hyperbolic partial differential equations with multiple characteristics

Stochastic partial differential equations provide a rigorous mathematical framework for studying dynamical problems in which random perturbations are involved. They arise naturally in fields as diverse as physics, finance, biology, and engineering, where uncertainty, noise, and incomplete information play a fundamental role.

Hyperbolic equations, in particular, are utilized to study wave-like phenomena that arise commonly in mechanics and signal propagation. Applications can range from acoustics and fluid dynamics to electromagnetism and elasticity.

A number of recent papers, tapping into the powerful techniques presented in the seminal works [38], [39] and [110], have extended several classical deterministic results for solutions of hyperbolic operators of various types — linear, semilinear, and whose principal symbols have multiple involutive or symplectic characteristics — to the stochastic framework: for a current review, encompassing also other types of classical PDEs in open domains in $\mathbb{R}^n$ and Riemannian manifolds as well, one can consult, among others, [1], [8], [9], [10], [11], [12], [18], [34], [35]. Our goal in this chapter is to continue the study of possible extensions to the stochastic framework by analyzing three model linear stochastic operators, we deem not having been previously examined in the literature. In particular, we would like to understand how the symplectic geometry of the principal symbols and the conditions on the lower order terms influence the construction of the random field solutions via the corresponding colors of the noise, an analysis started in [18].

2.1 Mathematical framework and model introduction

The general problem is presented here in a compact form:

$$\begin{cases} P_i u_i = \dot{F}_i\,(t,x) \quad t \geq 0, x \in \mathbb{R}^i, i \in \{1,2,3\} \\ u_i\,(0,x) = 0 \\ \partial_t u_i\,(0,x) = 0, \end{cases} \tag{2.1}$$

where formally

$$F_i(\phi) := \int_{\mathbb{R}^{i+1}} \phi(t,x)\dot{F}_i(t,x)\,dtdx, \quad \phi \in C_0^\infty\left(\mathbb{R}^{i+1}\right),$$

is a mean zero family of normal random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with covariance

$$E(F_i(\phi)F_i(\psi)) = \int_0^\infty dt \int_{\mathbb{R}^i} dx \int_{\mathbb{R}^i} dy\, \phi(t,x) f_i(x-y)\psi(t,y), \tag{2.2}$$

for $\phi, \psi \in C_0^\infty\left(\mathbb{R}^{i+1}\right)$.

We study the operators P_1, P_2 and P_3, defined as follows

(i) $P_1 := D_t^2 - (D_x^2 + x^2)$: this is a 0-th order perturbation of the wave equation. It is strictly hyperbolic. It introduces a harmonic oscillator and is mostly used as a prototype to develop the techniques to tackle the next two operators.

(ii) $P_2 := D_t^2 - \mu(D_x^2 + x^2 D_y^2) + bD_y$, $|b| < \mu$. This operator introduces an anisotropy and degeneracy in the propagation speed through the term $x^2 D_y^2$, which leads to double characteristics and non-strict hyperbolicity. Indeed its principal symbol is

$$p_2 := \tau^2 - \mu(\xi^2 + x^2\eta^2),$$

which vanishes to the second order on the smooth manifold

$$\Sigma_2 := \{(t, x, y, \tau, \xi, \eta) \in \dot{T}^*\mathbb{R}^3, x = \tau = \xi = 0\},$$

where $\dot{T}^*\mathbb{R}^3$ denotes the phase space cotangent bundle minus the zero section (Appendix No. 17). The term bD_y introduces a drift in the y direction; although of first order, this term is crucial to the well-posedness of the Cauchy problem and gives rise to the condition $|b| < \mu$ ([52], [53], [58]). As we will see, this condition is fundamental for the existence of a random field solution for the stochastic equation.

(iii) $P_3 := D_t^2 - \mu(D_x^2 + x^2D_y^2) - aD_z^2 + bD_y$, $|b| < \mu$, $a > 0$. The final operator adds another spatial dimension and an elliptic term $-aD_z^2$. It is the most general and complete of the three models. Again, the principal symbol, that is

$$p_3 := \tau^2 - \mu(\xi^2 + x^2\eta^2) - a\zeta^2,$$

is hyperbolic and vanishes to the second order on the manifold

$$\Sigma_3 := \{(t, x, y, z, \tau, \xi, \eta, \zeta) \in \dot{T}^*\mathbb{R}^4, x = \tau = \xi = \zeta = 0\}.$$

For non-strictly hyperbolic operators the characteristic set plays an important role, as its geometry affects well-posedness and propagation of singularities. In particular, we focus on how the Hamilton map behaves on the characteristic set. For a comprehensive review of the matter we refer to [88] and [84].

Both p_2 and p_3 do not have any effectively hyperbolic couple and are hence called non-effectively hyperbolic. p_2 is fully symplectic as (x, ξ) is a symplectic couple and the couple (y, η) produces a zero eigenvalue but without a Jordan block structure (neutral couple). On the other hand, p_3 is mixed symplectic-involutive, as (z, ζ) is involutive.

The Levi-type condition $|b| < \mu$ is typical for this kind of operators, and it is necessary for the well-posedness of the Cauchy problem. For a comprehensive analysis on the matter see, e.g., [52], [53] and [58].

Let $i = 1, 2, 3$. The function $f_i : \mathbb{R}^i \to \mathbb{R}$ in (2.2) is taken to be continuous except at most at zero, and even. These requirements are necessary to guarantee that the functional expressed in (2.2) is non-negative definite. As seen in [98], this is also equivalent to the existence of a non-negative tempered measure ν_i (Appendix No. 5) whose Fourier transform is f_i, i.e., for all $\phi \in \mathcal{S}(\mathbb{R}^i)$,

$$\int_{\mathbb{R}^i} \phi(x) f_i(x) dx = \int_{\mathbb{R}^i} \mathcal{F}\phi(\xi) \nu_i(d\xi).$$

We recall that a *random field solution* is defined as

$$u_i(t,x) = \int_0^t \int_{\mathbb{R}^i} E_i\,(t-s, x-y)\,\dot{F}_i(s,y) ds\, dy,$$

where E_i denotes the fundamental solution of the operator P_i, the stochastic integral is well defined and the map $(t,x) \mapsto u_i(t,x)$ is measurable. Finally, we define the physicist's Hermite polynomials and Hermite functions as

$$\begin{aligned} H_n(x) &= (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \\ \Psi_n(x) &= \frac{1}{(2^n n! \sqrt{\pi})^{1/2}} e^{\frac{-x^2}{2}} H_n(x), \end{aligned} \tag{2.3}$$

and name

$$\rho_n\,(\eta,\zeta) := \mu|\eta|\,(2n+1) + a\zeta^2 - b\eta, \qquad \tilde{\rho}_n(\eta) := \rho_n(\eta, 0).$$

We are now ready to state our main result, which contains all the several claims proven below, itemized according to the increasing number of space variables. Recalling that our stochastic PDEs are

$$\begin{cases} P_i u_i = \dot{F}_i\,(t,x) \quad & t \geq 0, x \in \mathbb{R}^i, i \in \{1,2,3\} \\ u_i\,(0,x) = 0 & \\ \partial_t u_i\,(0,x) = 0, & \end{cases}$$

where the operators $P_i, i \in \{1,2,3\}$, are defined as

$$\begin{aligned} P_1 &:= D_t^2 - (D_x^2 + x^2), \\ P_2 &:= D_t^2 - \mu(D_x^2 + x^2 D_y^2) + bD_y, \quad |b| < \mu \\ P_3 &:= D_t^2 - \mu(D_x^2 + x^2 D_y^2) - aD_z^2 + bD_y, \quad |b| < \mu,\ a > 0, \end{aligned}$$

we have the following results.

Theorem 2.1.1. *(i) Let $\nu_1(d\xi_0) = d\xi_0$ be the Lebesgue measure. Then, the one-dimensional problem has a random field solution given by*

$$u_1(t,x) = \int_0^t \int_{\mathbb{R}} E_1\left(t-s, x; x_0\right) \dot{F}_1(s, x_0) ds\, dx_0,$$

with

$$E_1(t,x,x_0) = -H(t) \sum_{n=0}^{\infty} \frac{\sin(\sqrt{2n+1}\, t)}{\sqrt{2n+1}} \Psi_n(x_0)\Psi_n(x).$$

.

(ii) Let $\nu_2 = d\xi \hat{\nu}_2(d\hat{\eta})$ and assume that

$$\int_{\mathbb{R}} \frac{1}{|\hat{\eta}|^{1/2}} \hat{\nu}_2(d\hat{\eta}) < +\infty. \tag{2.4}$$

Then, the two-dimensional problem has a random field solution given by

$$u_2(t,x,y) = \int_0^t \int_{\mathbb{R}^2} E_2\left(t-s, x, y-y_0; x_0\right) \dot{F}_2(s, x_0, y_0) ds\, dx_0\, dy_0,$$

with

$$\begin{aligned} E_2\left(t,x,y;x_0\right) = &-\frac{H\left(t\right)}{2\pi} \\ &\times \int_{\mathbb{R}} e^{iy\eta} |\eta|^{1/2} \sum_{n=0}^{\infty} \frac{\sin\left(\tilde{\rho}_n\left(\eta\right)^{1/2} t\right)}{\tilde{\rho}_n\left(\eta\right)^{1/2}} \Psi_n\left(x_0|\eta|^{1/2}\right) \Psi_n\left(x|\eta|^{1/2}\right) d\eta. \end{aligned}$$

(iii) Let $\nu_3(d\xi, d\hat{\eta}, d\hat{\zeta}) = d\xi \hat{\nu}_3(d\hat{\eta}, d\hat{\zeta})$. Assume $\hat{\nu}_3$ is absolutely continuous with respect to the Lebesgue measure and that it admits a density of the form $(\hat{\eta}, \hat{\zeta}) \mapsto w\left(|\hat{\eta}|^2 + |\hat{\zeta}|^2\right)$ for which it exists an $\alpha < 1/3$ such that

$$\int_{\mathbb{R}^2} \frac{1}{\left(|\hat{\eta}|^2 + |\hat{\zeta}|^2\right)^{\alpha}} \hat{\nu}_3(d\hat{\eta}, d\hat{\zeta}) < +\infty. \tag{2.5}$$

Then the three-dimensional problem has a random field solution given by

$$u_3(t,x,y,z) = \int_0^t \int_{\mathbb{R}^3} E_3\left(t-s, x, y-y_0, z-z_0; x_0\right) \dot{F}_3(s, x_0, y_0, z_0) ds\, dx_0\, dy_0\, dz_0,$$

with

$$E_3\left(t,x,y,z;x_0\right) = -\frac{H\left(t\right)}{2\pi}$$
$$\times \int_{\mathbb{R}^2} e^{iy\eta + iz\zeta} |\eta|^{1/2} \sum_{n=0}^{\infty} \frac{\sin\left(\rho_n\left(\eta,\zeta\right)^{1/2} t\right)}{\rho_n\left(\eta,\zeta\right)^{1/2}} \Psi_n\left(x_0|\eta|^{1/2}\right) \Psi_n\left(x|\eta|^{1/2}\right) d\eta\, d\zeta.$$

We would like to highlight again how the color of the noise, observable through the corresponding measures, relates to the geometry of the double manifolds.

The plan of this chapter is as follows. In Section 2.2 we organize the proof for the one-dimensional case, parts of which are to be used in the later sections. In particular in Section 2.2.1 the formal fundamental solution is computed and in Section 2.2.2 the corresponding random field is constructed. Section 2.3 contains the main arguments. In Section 2.3.1 the formal fundamental solutions are explicitly calculated for both dimensions 2 and 3. Sections 2.3.2 and 2.3.3 are devoted to the estimate of the formal solutions obtained, and in the final Section 2.3.4 the existence of the random field solutions as stated in Theorem 2.1.1 is eventually proven.

As a last note, we observe that the arguments we utilize in our analysis could very well be used to deal with a more general second order hyperbolic quadratic form with symplectic characteristics. Nevertheless, the simpler nature of our model examples allows us to remove some of the complexities in [82] or [64], where only the codimension 3 was considered, and develop a sharper understanding of the structure of our fundamental solution processes. We observe that the symplectic nature of the multiple manifold is not by itself an a-priori obstacle to obtaining a closed form fundamental solution, as was proven in [18]. What makes the problem more challenging in our case is indeed the presence of the harmonic oscillators.

2.2 The case of one spatial dimension

2.2.1 Fundamental solution

We now proceed with the case of one spatial dimension. This, besides proving the related simple case, helps us collecting all the major tools needed for the subsequent parts.

We consider the problem (Appendix No. 7)

$$P_1 E_1(t, x; x_0) = \delta(t)\delta(x - x_0).$$

Performing a Hermite series expansion of $x \mapsto E_1(t, x; x_0)$ and using the well-known identities ([65], [103])

$$\sum_{n=0}^{\infty} \Psi_n(x)\Psi_n(x_0) = \delta(x - x_0), \tag{2.6}$$

$$(\partial_x^2 - x^2 + 2n + 1)\Psi_n(x) = 0 \quad \Leftrightarrow \quad (D_x^2 + x^2)\Psi_n(x) = (2n + 1)\Psi_n(x), \tag{2.7}$$

we get

$$P_1 \sum_{n=0}^{\infty} E_{1,n}(t, x_0)\Psi_n(x) = \sum_{n=0}^{\infty} (P_{1,n}E_{1,n})(t, x_0)\Psi_n(x) = \sum_{n=0}^{\infty} \delta(t)\Psi_n(x)\Psi_n(x_0), \tag{2.8}$$

where

$$P_{1,n} = D_t^2 - (2n + 1).$$

Equating term by term in (2.8) we obtain the ODE problem

$$\begin{cases} P_{1,n}E_{1,n}(t, x_0) = \delta(t)\Psi_n(x_0), \\ E_{1,n}(0, x_0) = \partial_t E_{1,n}(0, x_0) = 0, \end{cases}$$

whose solution is

$$E_{1,n}(t, x_0) = -\Psi_n(x_0) \int_0^t \delta(t - \tau) \frac{\sin(\sqrt{2n+1}\,\tau)}{\sqrt{2n+1}}\, d\tau = -H(t) \frac{\sin(\sqrt{2n+1}\,t)}{\sqrt{2n+1}} \Psi_n(x_0), \tag{2.9}$$

where $H(t)$ denotes the Heaviside function, i.e.,

$$H(t) = \begin{cases} 1, & t \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Hence,

$$E_1(t, x; x_0) = -H(t) \sum_{n=0}^{\infty} \frac{\sin(\sqrt{2n+1}\, t)}{\sqrt{2n+1}} \Psi_n(x_0)\Psi_n(x).$$

Now, the Hermite functions are eigenvectors of the Fourier transform ([103]). In particular

$$\mathcal{F}\Psi_n(\eta) = (-i)^n \Psi_n(\eta),$$

and so

$$\mathtt{F}_1(t, x; \xi_0) := \mathcal{F}E_1(t, x; \cdot)(\xi_0) = -H(t) \sum_{n=0}^{\infty} (-i)^n \frac{\sin\left(t\sqrt{2n+1}\right)}{\sqrt{2n+1}} \Psi_n(\xi_0)\Psi_n(x) \tag{2.10}$$

2.2.2 Random field solution

For the reader's convenience, we recall the expression for the candidate random field solution

$$u_1(t, x) = \int_0^t \int_{\mathbb{R}} E_1(t-s, x; x_0)\, \dot{F}_1(s, x_0) ds\, dx_0. \tag{2.11}$$

In order to prove that the expression (2.11) is indeed a real-valued process, we need to show that $E_1(t-s, x; x_0)$ is an element of the space $\mathcal{P}_0$. To do that, the first step is to show the following.

Lemma 2.2.1. *Let $\mathtt{F}_1(t, x; \xi_0)$ be defined as in (2.10). Then,*

$$I := \int_0^t \int_{\mathbb{R}} |\mathtt{F}_1(t-s, x; \xi_0)|^2 d\xi_0\, ds < +\infty.$$

Proof. By Parseval identity (Appendix No. 8), we have

$$I = \int_0^t \sum_{n=0}^{+\infty} \frac{\sin^2\left((t-s)\sqrt{2n+1}\right)}{2n+1} \Psi_n^2(x) ds \leq t \sum_{n=0}^{+\infty} \frac{1}{2n+1} \Psi_n^2(x).$$

Now, we recall that there exist constants C_1 and C_2 such that ([25])

$$C_1 n^{-1/6} < \max_{x\in\mathbb{R}} \Psi_n^2(x) < C_2 n^{-1/6}, \quad n \geq 1. \tag{2.12}$$

This implies

$$I \leq C_2 t\left(\Psi_0^2(x) + \sum_{n=1}^{+\infty}(2n+1)^{-1}n^{-1/6}\right) \leq C_2 t\left(\pi^{-1/4} + \sum_{n=1}^{+\infty}(2n+1)^{-1}n^{-1/6}\right),$$

which is bounded. □

This shows that the fundamental solution $E_1\,(t-s,x;x_0)$ belongs to the space $\bar{\mathcal{P}}$. However, to prove that $E_1 \in \mathcal{P}_0$ as well, we need to find a sequence of functions $(E_{1,m})_{m\in\mathbb{N}}$ in $\mathcal{E}_0$ converging to E_1 in $\bar{\mathcal{P}}$. Unsurprisingly, we set

$$E_{1,m}\,(t-s,x;x_0) := -H(t)\sum_{n=0}^{m}\frac{\sin\left((t-s)\sqrt{2n+1}\right)}{\sqrt{2n+1}}\Psi_n(x_0)\Psi_n(x).$$

Firstly, we observe that for each $m\in\mathbb{N}$, $x_0 \mapsto E_{1,m}\,(t,x;x_0) \in \mathcal{S}\,(\mathbb{R})$. Moreover, setting

$$\mathrm{F}_1^m\,(t-s,x;\xi_0) := -H(t)\sum_{n=m}^{\infty}(-i)^n\frac{\sin\left((t-s)\sqrt{2n+1}\right)}{\sqrt{2n+1}}\Psi_n(\xi_0)\Psi_n(x),$$

we have

$$\left\|E_{1,m}-E_1\right\|_0^2 = \int_0^t ds\int_{\mathbb{R}}\left|\mathrm{F}_1^m\,(t-s,x;\xi_0)\right|^2 d\xi_0.$$

Hence, from the proof of Lemma 2.2.1, we gather

$$\left\|E_{1,m}-E_1\right\|_0^2 \leq C_2 t\sum_{n=m}^{+\infty}(2n+1)^{-1}n^{-1/6} \xrightarrow{m\to+\infty} 0.$$

Now, it remains to prove measurability, and we do that by checking $\mathbb{L}^2(\Omega)$-continuity of the solution process in the separate variables. Denoting

$$S_1(t,x,h) = \sum_{n=0}^{+\infty}\frac{\sin^2\left(t\sqrt{2n+1}\right)}{2n+1}\left(\Psi_n(x+h)-\Psi_n(x)\right)^2,$$

the increment in space gives

$$\mathbb{E}\left[|u_1(t,x+h)-u_1(t,x)|^2\right] = \int_0^t ds \int_{\mathbb{R}} |\mathtt{F}_1(t-s,x+h;\xi_0)-\mathtt{F}_1(t-s,x;\xi_0)|^2\, d\xi_0$$
$$= \int_0^t S_1(t-s,x,h)ds.$$

In view of (2.12), we have

$$S_1(t,x,h) \le 4C_2\left(\pi^{-1/4}+\sum_{n=1}^{+\infty}(2n+1)^{-1}n^{-1/6}\right) < +\infty, \tag{2.13}$$

granting uniform convergence and consequently the continuity of $S_1(t-s,x,h)$ with respect to the variable h. Therefore, by dominated convergence,

$$\lim_{h\to 0}\mathbb{E}\left[|u_1(t,x+h)-u_1(t,x)|^2\right] = 0. \tag{2.14}$$

For the increment in time we write

$$\begin{aligned}
&\mathbb{E}\left[|u_1(t+h,x)-u_1(t,x)|^2\right]\\
\le& 2\int_0^t ds \int_{\mathbb{R}} |\mathtt{F}_1(t+h-s,x;\xi_0)-\mathtt{F}_1(t-s,x;\xi_0)|^2\, d\xi_0\\
&+2\int_t^{t+h} ds \int_{\mathbb{R}} |\mathtt{F}_1(t+h-s,x;\xi_0)|^2\, d\xi_0\\
=&2I_1(h)+2I_2(h).
\end{aligned}$$

Lemma 2.2.1 gives directly

$$\lim_{h\to 0} I_2(h) = 0.$$

Similarly to the increment in the x variable, we call

$$\tilde{S}_1(t,x,h) := \sum_{n=0}^{+\infty}\frac{\left(\sin\left((t+h)\sqrt{2n+1}\right)-\sin\left(t\sqrt{2n+1}\right)\right)^2}{\sqrt{2n+1}}\Psi_n^2(x),$$

and, by Parseval's identity, we obtain

$$I_1(h) = \int_0^t \tilde{S}_1(t-s,x,h)ds.$$

For $\tilde{S}_1(t-s,x,h)$, a bound similar to (2.13) holds. So, by dominated convergence

$$\lim_{h\to 0} I_1(h) = 0,$$

which completes the proof.

2.3 The case of two and three spatial dimensions

2.3.1 Common computations: fundamental solution

Here we explicitly derive the fundamental solution of the operator P_3. Since the structure of the calculation is much the same as in the two-dimensional case, we omit the latter and recover the corresponding fundamental solution at the end of the section.

Hence, our problem becomes now

$$P_3 E_3 = \delta(t)\delta(x-x_0)\delta(y)\delta(z).$$

Taking the Fourier transform with respect to (y,z) we get

$$\mathcal{F}(P_3E_3) = \left(D_t^2 - \mu\left(D_x^2 + x^2\eta^2\right) - a\zeta^2 + b\eta\right)\hat{E}_3\left(t,x,\eta,\zeta;x_0\right) = \delta\left(t\right)\delta\left(x-x_0\right). \tag{2.15}$$

Setting $\sigma = |\eta|^{1/2}x$ we define

$$W\left(t,\sigma,\eta,\zeta;x_0\right) = \hat{E}_3\left(t,|\eta|^{-1/2}x,\eta,\zeta;x_0\right).$$

Thus

$$\partial_\sigma^2 W\left(t,\sigma,\eta,\zeta;x_0\right) = |\eta|^{-1}\partial_x^2\hat{E}_3\left(t,|\eta|^{-1/2}x,\eta,\zeta;x_0\right),$$

and, consequently,

$$\left(D_x^2 + x^2\eta^2\right)\hat{E}_3\left(t,|\eta|^{-1/2}x,\eta,\zeta;x_0\right) = |\eta|\left(D_\sigma^2+\sigma^2\right)W\left(t,\sigma,\eta,\zeta;x_0\right).$$

Therefore, since Dirac's delta is homogeneous of degree -1, naming $\Lambda := D_t^2 - \mu|\eta|\left(D_\sigma^2+\sigma^2\right) -$

$a\zeta^2 + b\eta$, (2.15) finally becomes

$$\Lambda W(t, \sigma, \eta, \zeta; x_0) = |\eta|^{1/2} \delta(t)\, \delta\left(\sigma - x_0|\eta|^{1/2}\right). \tag{2.16}$$

We expand $\sigma \mapsto W(t, \sigma, \eta, \zeta; x_0)$ in Hermite functions

$$W(t, \sigma, \eta, \zeta; x_0) = \sum_{n=0}^{\infty} W_n(t, \eta, \zeta; x_0)\, \Psi_n(\sigma),$$

and define

$$\Lambda_n = D_t^2 - \mu|\eta|(2n+1) - a\zeta^2 + b\eta.$$

It follows by (2.6) and (2.7) that (2.16) transforms into

$$\begin{aligned} \Lambda W(t, \sigma, \eta, \zeta; x_0) =& \Lambda \sum_{n=0}^{\infty} W_n(t, \eta, \zeta; x_0)\, \Psi_n(\sigma) = \sum_{n=0}^{\infty} \Lambda_n W_n(t, \eta, \zeta; x_0)\, \Psi_n(\sigma) \\ =& \delta(t) \sum_{n=0}^{\infty} \Psi_n\left(x_0|\eta|^{1/2}\right) \Psi_n(\sigma), \end{aligned}$$

and we split the sum term by term, equating the coefficients of the expansion

$$(\Lambda_n W_n)(t, \eta, \zeta; x_0) = \delta(t)\, \Psi_n\left(x_0|\eta|^{1/2}\right) |\eta|^{1/2},$$

so that we may focus on the ODE problem

$$\begin{cases} \left(\partial_t^2 + \rho_n(\eta, \zeta)\right) W_n(t, \eta, \zeta; x_0) = -\delta(t)\, \Psi_n\left(x_0|\eta|^{1/2}\right) |\eta|^{1/2}, \\ W_n(0, \eta, \zeta; x_0) = \partial_t W_n(0, \eta, \zeta; x_0) = 0, \end{cases} \tag{2.17}$$

which is solved by

$$\begin{aligned} W_n(t, \eta, \zeta; x_0) =& -\Psi_n\left(x_0|\eta|^{1/2}\right) |\eta|^{1/2} \cdot \int_0^t \delta(t - \tau) \frac{\sin\left(\rho_n(\eta, \zeta)^{1/2} \tau\right)}{\rho_n(\eta, \zeta)^{1/2}} d\tau \\ =& -H(t)\, \Psi_n\left(x_0|\eta|^{1/2}\right) |\eta|^{1/2} \frac{\sin\left(\rho_n(\eta, \zeta)^{1/2} t\right)}{\rho_n(\eta, \zeta)^{1/2}}. \end{aligned}$$

Therefore, we find

$$W(t,\sigma,\eta,\zeta;x_0) = -H(t)\,|\eta|^{1/2}\sum_{n=0}^{\infty}\frac{\sin\left(\rho_n(\eta,\zeta)^{1/2}\,t\right)}{\rho_n(\eta,\zeta)^{1/2}}\Psi_n\left(x_0|\eta|^{1/2}\right)\Psi_n(\sigma),$$

which readily gives

$$\hat{E}_3(t,x,\eta,\zeta;x_0) = -H(t)\,|\eta|^{1/2}\sum_{n=0}^{\infty}\frac{\sin\left(\rho_n(\eta,\zeta)^{1/2}\,t\right)}{\rho_n(\eta,\zeta)^{1/2}}\Psi_n\left(x_0|\eta|^{1/2}\right)\Psi_n\left(|\eta|^{1/2}x\right),$$

so that

$$\begin{aligned}E_3(t,x,y,z;x_0) &= -\frac{H(t)}{2\pi}\\ &\times\int_{\mathbb{R}^2} e^{iy\eta+iz\zeta}|\eta|^{1/2}\sum_{n=0}^{\infty}\frac{\sin\left(\rho_n(\eta,\zeta)^{1/2}\,t\right)}{\rho_n(\eta,\zeta)^{1/2}}\Psi_n\left(x_0|\eta|^{1/2}\right)\Psi_n\left(x|\eta|^{1/2}\right)d\eta\,d\zeta.\end{aligned}$$

Now, we are interested in

$$\mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right) := \mathcal{F}E_3(t-s,x,y-\cdot,z-\cdot,\cdot)\left(\xi_0,\hat{\eta},\hat{\zeta}\right).$$

By translation and dilation properties of the Fourier transform, we see

$$\mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right) = e^{iy\hat{\eta}+iz\hat{\zeta}}\int_{-\infty}^{\infty} e^{-i\xi_0 x_0}\hat{E}_3\left(t-s,x,-\hat{\eta},-\hat{\zeta};x_0\right)dx_0.$$

Thus, we finally get

$$\begin{aligned}\mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right) &= -\frac{H(t-s)}{2\pi}e^{iy\hat{\eta}+iz\hat{\zeta}}\\ &\times\sum_{n=0}^{\infty}(-i)^n\,\frac{\sin\left(\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}(t-s)\right)}{\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}}\Psi_n\left(\frac{\xi_0}{|\hat{\eta}|^{1/2}}\right)\Psi_n\left(|\hat{\eta}|^{1/2}x\right).\end{aligned}\tag{2.18}$$

Following the above procedure along, one can also solve

$$P_2E_2 = \delta(t)\delta(x-x_0)\delta(y)$$

and obtain

$$E_2(t,x,y;x_0) = -\frac{H(t)}{2\pi} \\ \times \int_{\mathbb{R}} e^{iy\eta}|\eta|^{1/2} \sum_{n=0}^{\infty} \frac{\sin\left(\tilde{\rho}_n(\eta)^{1/2} t\right)}{\tilde{\rho}_n(\eta)^{1/2}} \Psi_n\left(x_0|\eta|^{1/2}\right)\Psi_n\left(x|\eta|^{1/2}\right) d\eta,$$

where $\tilde{\rho}_n(\eta) := \rho_n(\eta, 0)$. We conclude

$$\mathtt{F}_2(t-s,x,y;\xi_0,\hat{\eta}) = -\frac{H(t-s)}{2\pi} e^{iy\hat{\eta}} \\ \times \sum_{n=0}^{\infty} (-i)^n \frac{\sin\left(\tilde{\rho}_n(-\hat{\eta})^{1/2}(t-s)\right)}{\tilde{\rho}_n(-\hat{\eta})^{1/2}} \Psi_n\left(\frac{\xi_0}{|\hat{\eta}|^{1/2}}\right)\Psi_n\left(|\hat{\eta}|^{1/2} x\right). \tag{2.19}$$

2.3.2 Two dimensions: Integral bound

Again we rewrite here the expression for the two-dimensional candidate random field solution

$$u_2(t,x,y) = \int_0^t \int_{\mathbb{R}^2} E_2(t-s,x,y-y_0;x_0)\, \dot{F}_1(s,x_0,y_0) ds\, dx_0\, dy_0.$$

In the present section, we proceed with the study of the fundamental solutions E_2 to argue that the formal expression above is well defined as a real-valued process. We start by proving Lemma 2.3.1 below, which shows that $\mathtt{F}_2(t-s,x,y;\xi_0,\hat{\eta})$ in (2.19) is $L^2(\nu_2(d\xi_0,d\hat{\eta}))$.

Lemma 2.3.1. *Let $\mathtt{F}_2(t-s,x,y;\xi_0,\hat{\eta})$ be defined as in (2.19) and the assumptions of Theorem 2.1.1 (ii) hold. Then*

$$I_2 := \int_0^t \int_{\mathbb{R}^2} |\mathtt{F}_2(t-s,x,y;\xi_0,\hat{\eta})|^2\, \nu_2(d\xi_0,d\hat{\eta})\, ds < +\infty.$$

Proof. Let us call

$$J(t-s,x,y;\hat{\eta}) := \int_{\mathbb{R}} |\mathtt{F}_2(t-s,x,y;\xi_0,\hat{\eta})|^2\, d\xi_0$$

and employ the change of variable $\lambda = |\hat{\eta}|^{-1/2}\xi_0$. Then,

$$J(t-s,x,y;\hat{\eta}) := \int_{\mathbb{R}} \left|\mathtt{F}_2\left(t-s,x,y;|\hat{\eta}|^{1/2}\lambda,\hat{\eta}\right)\right|^2 |\hat{\eta}|^{1/2}\, d\lambda.$$

Now, if we denote

$$
\begin{aligned}
g_n(t-s,x,y;\hat{\eta}) :=& \frac{H(t-s)}{2\pi}e^{iy\hat{\eta}} \\
&\times(-i)^n \frac{\sin\left(\tilde{\rho}_n(-\hat{\eta})^{1/2}(t-s)\right)}{\tilde{\rho}_n(-\hat{\eta})^{1/2}}\Psi_n\left(|\hat{\eta}|^{1/2}x\right),
\end{aligned}
\tag{2.20}
$$

we get

$$
\mathrm{F}_2\left(t-s,x,y;|\hat{\eta}|^{1/2}\lambda,\hat{\eta}\right) = \sum_{n=0}^{\infty} g_n(t-s,x,y;\hat{\eta})\,\Psi_n(\lambda),
$$

and hence, by Parseval's identity,

$$
\int_{\mathbb{R}} \left|\mathrm{F}_2\left(t-s,x,y;|\hat{\eta}|^{1/2}\lambda,\hat{\eta}\right)\right|^2 d\lambda = \sum_{n=0}^{\infty} g_n(t-s,x,y;\hat{\eta})^2 . \tag{2.21}
$$

It follows that

$$
\begin{aligned}
I_2 =& \int_0^t \int_{\mathbb{R}} |\hat{\eta}|^{1/2} \sum_{n=0}^{\infty} g_n(t-s,x,y;\hat{\eta})^2\,\hat{\nu}_2(d\hat{\eta})\,ds \\
=& \frac{1}{4\pi^2}\int_0^t \int_{\mathbb{R}} \sum_{n=0}^{\infty} \frac{\sin^2\left(\tilde{\rho}_n^{1/2}(-\hat{\eta})(t-s)\right)}{\tilde{\rho}_n(-\hat{\eta})} |\hat{\eta}|^{1/2}\Psi_n^2\left(|\hat{\eta}|^{1/2}x\right)\hat{\nu}_2(d\hat{\eta})\,ds.
\end{aligned}
$$

Recalling (2.12) and denoting

$$
T_0 := \int_0^t \int_{\mathbb{R}} g_0(t-s,x,y;\hat{\eta})^2\,|\hat{\eta}|^{1/2}\hat{\nu}_2(d\hat{\eta})\,ds,
$$

we get

$$
\begin{aligned}
I_2 \leq& \frac{C_2}{4\pi^2}\int_0^t \int_{\mathbb{R}} \sum_{n=1}^{\infty} \frac{|\hat{\eta}|^{1/2}}{n^{1/6}\left((2n+1)\mu|\hat{\eta}|+b\hat{\eta}\right)}\hat{\nu}_2(d\hat{\eta})\,ds + T_0 \\
\leq& \frac{C_2\,t}{4\pi^2 n^{1/6}} \sum_{n=1}^{\infty} \int_{\mathbb{R}} \frac{|\hat{\eta}|^{1/2}}{\left((2n+1)\mu|\hat{\eta}|+b\hat{\eta}\right)}\hat{\nu}_2(d\hat{\eta}) + T_0.
\end{aligned}
\tag{2.22}
$$

Now, we have

$$
\int_{\mathbb{R}} \frac{|\hat{\eta}|^{1/2}}{\left((2n+1)\mu|\hat{\eta}|+b\hat{\eta}\right)}\hat{\nu}_2(d\hat{\eta}) \leq \frac{1}{n\mu}\int_{\mathbb{R}} |\hat{\eta}|^{-1/2}\,\hat{\nu}_2(d\hat{\eta}) = \frac{C}{\mu n},
$$

where the last inequality comes from the fact that the condition $|b| < \mu$ gives $\mu + \mathrm{sign}(\hat{\eta})b > 0$.

Finally, convergence of the integral T_0 under this measure is easily checked. Indeed, we can find an $\epsilon > 0$ such that $\mu + \text{sign}(\hat{\eta})b > \epsilon$ for all $\hat{\eta} \in \mathbb{R}$ and observe that $\Psi_0^2\left(|\hat{\eta}|^{1/2}x\right) = \pi^{-1/4}e^{-|\hat{\eta}|x^2/2} \leq \pi^{-1/4}$. Therefore,

$$T_0 \leq \frac{t}{4\pi^{9/4}\epsilon}\int_{\mathbb{R}} |\hat{\eta}|^{-1/2}\,\hat{\nu}_2\,(d\hat{\eta}) < +\infty.$$

□

Now, we want to find a sequence $E_{2,m}\,(t,x,y;x_0) \in \mathcal{E}_0$ so that $E_{2,m} \to E_2$ in $\bar{\mathcal{P}}$. We set

$$\begin{aligned} E_{2,m}\,(t,x,y;x_0) :=& -\frac{H\,(t)}{2\pi} \\ &\times \int_{\mathbb{R}} e^{iy\eta}|\eta|^{1/2}\sum_{n=0}^{m}\frac{\sin\left(\tilde{\rho}_n\,(\eta)^{1/2}\,t\right)}{\tilde{\rho}_n\,(\eta)^{1/2}}\Psi_n\left(x_0|\eta|^{1/2}\right)\Psi_n\left(x|\eta|^{1/2}\right)d\eta. \end{aligned}$$

Again, for each $m \in \mathbb{N}$, $(y, x_0) \mapsto E_{2,m}(t,x,y;x_0) \in \mathcal{S}\left(\mathbb{R}^2\right)$, and so setting

$$\begin{aligned} \mathrm{F}_2^m\,(t-s,x,y;\xi_0,\hat{\eta}) :=& \frac{H\,(t-s)}{2\pi}e^{iy\hat{\eta}} \\ &\times\sum_{n=m}^{+\infty}(-i)^n\,\frac{\sin\left(\tilde{\rho}_n\,(-\hat{\eta})^{1/2}\,(t-s)\right)}{\tilde{\rho}_n\,(-\hat{\eta})^{1/2}}\Psi_n\left(\frac{\xi_0}{|\hat{\eta}|^{1/2}}\right)\Psi_n\left(|\hat{\eta}|^{1/2}x\right), \end{aligned}$$

we have

$$\begin{aligned} \left\|E_{2,m} - E_2\right\|_0^2 =& \int_0^t ds\int_{\mathbb{R}^2}\left|\mathrm{F}_2^m\,(t-s,x,y;\xi_0,\hat{\eta})\right|^2\,\nu_2\,(d\xi_0,d\hat{\eta}) \\ \leq& \frac{C_2\,t}{4\pi^2}\sum_{n=m}^{+\infty}\int_{\mathbb{R}}\frac{|\hat{\eta}|^{1/2}}{n^{1/6}\,((2n+1)\,\mu\,|\hat{\eta}| + b\hat{\eta})}\hat{\nu}_2(d\hat{\eta}) \\ \leq& \tilde{C}_2\,t\sum_{n=m}^{+\infty}n^{-7/6}\xrightarrow{m\to+\infty}0, \end{aligned}$$

as desired.

2.3.3 Three dimensions: integral bound

For the reader's convenience, we recall the formal expression for the candidate three-dimensional random field solution:

$$u_3(t,x,y,z) = \int_0^t \int_{\mathbb{R}^3} E_3\left(t-s,x,y-y_0,z-z_0;x_0\right) \\ \times \dot{F}_3(s,x_0,y_0,z_0) ds\, dx_0\, dy_0\, dz_0.$$

In the present section, we would like to understand whether the above expression is a real-valued process and to check if E_3 belongs to $\mathcal{P}_0$. As always, we start with showing that $E_3 \in \bar{\mathcal{P}}$.

Lemma 2.3.2. *Let* $\mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right)$ *be defined as in (2.18) and the assumptions of Theorem 2.1.1 (iii) hold. Then,*

$$I_3 := \int_0^t \int_{\mathbb{R}^3} \left|\mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right)\right|^2 \nu_3\left(d\xi_0,d\hat{\eta},d\hat{\zeta}\right) ds < \infty.$$

Proof. The proof is very similar to the one of Proposition 2.3.1. Indeed, denoting

$$\hat{g}_n\left(t-s,x,y;\hat{\eta},\hat{\zeta}\right) := \frac{H\left(t-s\right)}{2\pi} e^{iy\hat{\eta}+iz\hat{\zeta}} \\ \times (-i)^n \frac{\sin\left(\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}(t-s)\right)}{\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}} \Psi_n\left(|\hat{\eta}|^{1/2}x\right),$$

and

$$\hat{T}_0 := \int_0^t \int_{\mathbb{R}^2} \hat{g}_0\left(t-s,x,y,z;\hat{\eta},\hat{\zeta}\right)^2 |\hat{\eta}|^{1/2} \hat{\nu}_3\left(d\hat{\eta},d\hat{\zeta}\right) ds, \\ \hat{q}_n := \int_{\mathbb{R}^2} \frac{|\hat{\eta}|^{1/2}}{\left((2n+1)\mu\,|\hat{\eta}| + a\hat{\zeta}^2 + b\hat{\eta}\right)} \hat{\nu}_3\left(d\hat{\eta},d\hat{\zeta}\right),$$

with the same strategy employed in the previous section, we obtain

$$I_3 \leq \frac{C_2\, t}{4\pi^2} \sum_{n=1}^{\infty} n^{-1/6} \hat{q}_n + \hat{T}_0. \tag{2.23}$$

Now, since $|b| < \mu$, we have $\mu + \mathrm{sign}(\hat{\eta})b > 0$ and consequently

$$\hat{q}_n \leq \int_{\mathbb{R}^2} \frac{|\hat{\eta}|^{1/2}}{n\mu\,|\hat{\eta}| + a\hat{\zeta}^2} w(\hat{\eta}^2 + \hat{\zeta}^2)\, d\hat{\eta} d\hat{\zeta} =: J(n).$$

We can then switch to polar coordinates and, recalling Young's inequality

$$\frac{x^p}{p} + \frac{y^q}{q} \geq xy, \qquad \frac{1}{p} + \frac{1}{q} = 1,$$

obtain

$$\begin{aligned}
J(n) =& 4\int_0^{+\infty} d\rho\, \rho\, w(\rho^2) \int_0^{\pi/2} \frac{(\rho\cos(\theta))^{1/2}}{n\mu\rho\cos(\theta) + a\rho^2\sin^2(\theta)} d\theta \\
=& 4\int_0^{+\infty} d\rho\, \rho^{1/2} w(\rho^2) \int_0^{\pi/2} \frac{(\cos(\theta))^{1/2}}{n\mu\cos(\theta) + a\rho\sin^2(\theta)} d\theta \\
\leq& \frac{4(1/p)^{1/p}(1/q)^{1/q}}{\mu^{1/p}a^{1/q}} n^{-1/p} \int_0^{+\infty} d\rho\, \rho^{1/2-1/q} w(\rho^2) \int_0^{\pi/2} \cos(\theta)^{1/2-1/p} \sin\theta^{-2/q} d\theta.
\end{aligned}$$

We have

$$\int_0^{\pi/2} \cos(\theta)^{1/2-1/p} \sin\theta^{-2/q} d\theta = \frac{\Gamma(3/4 - 1/2p)\,\Gamma(1/2 - 1/q)}{2\Gamma(5/4 - 1/2p - 1/q)} < +\infty,$$

with q selected as below. In fact, condition (2.5) guarantees that we can indeed select q so that

$$\int_0^{+\infty} \rho^{1/2-1/q} w(\rho^2) d\rho = \int_0^{+\infty} \rho^{1-2\alpha} w(\rho^2) d\rho < +\infty.$$

From $\frac{1}{q} = 2\alpha - \frac{1}{2}$ we have

$$\frac{1}{p} = 1 - \frac{1}{q} = \frac{3}{2} - 2\alpha > \frac{5}{6}.$$

Therefore, we get $J(n) \leq \hat{C} n^{-1/p}$, $1/p > 5/6$, which is sufficient for the convergence of the series in (2.23).

Finally, taking $\epsilon > 0$ such that $\epsilon < \mu + b\,\mathrm{sign}(\hat{\eta})$, we write

$$T_0 \leq \frac{t}{4\pi^{9/4}} \int_{\mathbb{R}^2} \frac{|\hat{\eta}|^{1/2}}{a\hat{\zeta}^2 + \epsilon\,|\hat{\eta}|} w(\hat{\eta}^2 + \hat{\zeta}^2)\, d\hat{\eta} d\hat{\zeta},$$

which is finite, by the same arguments utilized above. □

Having proved that $E_3 \in \bar{\mathcal{P}}$, in order to show that $E_3 \in \mathcal{P}_0$ we need to find a sequence of functions $E_{3,m} \in \mathcal{E}_0$ such that $E_{3,m} \to E_3$ in $\bar{\mathcal{P}}$. We set

$$E_{3,m}(t,x,y,z;x_0) := -\frac{H(t)}{2\pi} \times \int_{\mathbb{R}^2} e^{iy\eta + iz\zeta} |\eta|^{1/2} \sum_{n=0}^{m} \frac{\sin\left(\rho_n(\eta,\zeta)^{1/2} t\right)}{\rho_n(\eta,\zeta)^{1/2}} \Psi_n\left(x_0|\eta|^{1/2}\right) \Psi_n\left(x|\eta|^{1/2}\right) d\eta\, d\zeta.$$

It is readily seen that $(x_0, y, z) \mapsto E_{3,m}(t,x,y,z;x_0) \in \mathcal{S}\left(\mathbb{R}^3\right)$ for all $m \in \mathbb{N}$, and

$$\left\|E_{3,m} - E_3\right\|_0^2 \leq \frac{C_2\, t}{4\pi^2} \sum_{n=m}^{+\infty} n^{-1/6} \hat{q}_n \xrightarrow{m\to+\infty} 0.$$

2.3.4 Common computations: measurability

To prove our claim, it is left to show that the maps $(t,x,y) \mapsto u_2(t,x,y)$ and $(t,x,y,z) \mapsto u_3(t,x,y,z)$ are measurable. Our strategy in this regard is to show $L_2(\Omega)$-continuity of the solution processes. We only tackle the three-dimensional case, as the two-dimensional case is pretty much analogous.

We first consider the variable y. We call

$$\bar{F}_3(t-s,x,z;\xi_0,\hat{\eta},\hat{\zeta}) := \frac{H(t-s)}{2\pi} e^{iz\hat{\zeta}} \times \sum_{n=0}^{+\infty} (-i)^n \frac{\sin\left(\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}(t-s)\right)}{\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}} \Psi_n\left(\frac{\xi_0}{|\hat{\eta}|^{1/2}}\right) \Psi_n\left(|\hat{\eta}|^{1/2} x\right),$$

and get

$$\begin{aligned}
&\mathbb{E}\left[|u_3(t,x,y+h,z) - u_3(t,x,y,z)|^2\right] \\
&= \int_0^t ds \int_{\mathbb{R}^3} \left|\mathrm{F}_3\left(t-s,x,y+h,z;\xi_0,\hat{\eta},\hat{\zeta}\right) - \mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right)\right|^2 \nu_3(d\xi_0,d\hat{\eta},d\hat{\zeta}) \\
&= \int_0^t ds \int_{\mathbb{R}^3} \left|e^{i\hat{\eta}(y+h)} - e^{i\hat{\eta}y}\right|^2 \left|\bar{F}_3(t-s,x,z;\xi_0,\hat{\eta},\hat{\zeta})\right|^2 \nu_3(d\xi_0,d\hat{\eta},d\hat{\zeta}).
\end{aligned}$$

Now, $\bar{F}_3(t-s,x,z;\xi_0,\hat{\eta},\hat{\zeta})$ is integrable, and hence, by dominated convergence, we have

$$\lim_{h\to 0} \mathbb{E}\left[|u_3(t,x,y+h,z) - u_3(t,x,y,z)|^2\right] = 0.$$

The z variable is treated in the same way. For the increment in the x variable, we denote

$$S_3(t-s,\hat{\eta},\hat{\zeta},x,h) := \sum_{n=0}^{+\infty} \frac{\sin^2\left(\rho_n\left(-\hat{\eta},\hat{\zeta}\right)^{1/2}(t-s)\right)}{\rho_n\left(-\hat{\eta},\hat{\zeta}\right)} \left(\Psi_n(|\eta|^{1/2}(x+h)) - \Psi_n(|\eta|^{1/2}x)\right)^2,$$

and write

$$\begin{aligned}
&\mathbb{E}\left[|u_3(t,x+h,y,z)-u_3(t,x,y,z)|^2\right] \\
=&\int_0^t ds \int_{\mathbb{R}^3} \left|\mathrm{F}_3\left(t-s,x+h,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right) - \mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right)\right|^2 \nu_3(d\xi_0,d\hat{\eta},d\hat{\zeta}) \\
=&\int_0^t ds \int_{\mathbb{R}^2} \frac{|\hat{\eta}|^{1/2}}{4\pi^2} S_3(t-s,\hat{\eta},\hat{\zeta},x,h)\left(1+|\hat{\eta}|^2+|\hat{\zeta}|^2\right)^{-\alpha} d\hat{\eta}d\hat{\zeta}.
\end{aligned}$$

Moreover in view of (2.12), we have

$$S_3(t-s,\hat{\eta},\hat{\zeta},x,h) \leq \sum_{n=0}^{+\infty} \frac{4C_2}{n^{1/6}\rho_n\left(-\hat{\eta},\hat{\zeta}\right)} < +\infty, \tag{2.24}$$

granting uniform convergence and consequently the continuity of $S_3(t-s,\hat{\eta},\hat{\zeta},x,h)$ with respect to the variable h too. Furthermore, in the proof of Lemma 2.3.2 we have seen that

$$\int_0^t ds \int_{\mathbb{R}^3} |\eta|^{1/2} \sum_{n=0}^{+\infty} \frac{1}{n^{1/6}\rho_n\left(-\hat{\eta},\hat{\zeta}\right)} \nu_3(d\xi_0,d\hat{\eta},d\hat{\zeta}) < +\infty.$$

Therefore, by dominated convergence we get

$$\lim_{h\to 0} \mathbb{E}\left[|u_3(t,x+h,y,z)-u_3(t,x,y,z)|^2\right] = 0. \tag{2.25}$$

Lastly, the increment in time yields

$$\begin{aligned}
&\mathbb{E}\left[|u_3(t+h,x,y,z)-u_3(t,x,y,z)|^2\right] \\
\leq& 2\int_0^t ds \int_{\mathbb{R}^3} \left|\mathrm{F}_3\left(t+h-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right) - \mathrm{F}_3\left(t-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right)\right|^2 \nu_3(d\xi_0,d\hat{\eta},d\hat{\zeta}) \\
&+ 2\int_t^{t+h} ds \int_{\mathbb{R}^3} \left|\mathrm{F}_3\left(t+h-s,x,y,z;\xi_0,\hat{\eta},\hat{\zeta}\right)\right|^2 \nu_3(d\xi_0,d\hat{\eta},d\hat{\zeta}) \\
=& 2J_1(h) + 2J_2(h).
\end{aligned}$$

Lemma 2.3.2 immediately gives

$$\lim_{h\to 0} J_2(h) = 0.$$

Denoting

$$\tilde{S}_3(t-s,\hat{\eta},\hat{\zeta},x,h) := \sum_{n=0}^{+\infty} \frac{\left(\sin\left(\rho_n^{1/2}(\hat{\eta},\hat{\zeta})(t+h-s)\right) - \sin\left(\rho_n^{1/2}(\hat{\eta},\hat{\zeta})(t-s)\right)\right)^2}{\rho_n(\hat{\eta},\hat{\zeta})} \Psi_n^2(|\eta|^{1/2}x),$$

by Parseval's identity, we obtain

$$J_1(h) = \int_0^t ds \int_{\mathbb{R}^2} \frac{|\hat{\eta}|^{1/2}}{4\pi^2} \tilde{S}_3(t-s,\hat{\eta},\hat{\zeta},x,h)\left(1+|\hat{\eta}|^2+|\hat{\zeta}|^2\right)^{-\alpha} d\hat{\eta}d\hat{\zeta}.$$

For $\tilde{S}_3(t-s,\hat{\eta},\hat{\zeta},x,h)$, the same bound as in (2.24) holds. So, by dominated convergence, again we have

$$\lim_{h\to 0} J_1(h) = 0,$$

which concludes the proof.

Chapter 3

An associated bundle approach to the Bures–Wasserstein geometry of fixed rank covariance matrices

Covariance matrices are present and vastly utilized in many fields of study, including statistics, physics and engineering, as they can represent many different data. Their intrinsic geometry under the Bures–Wasserstein metric is central for information geometry and optimal transport. A deep geometric understanding therefore translates into stronger theoretical guarantees and more efficient algorithms across all applications.

3.1 Mathematical framework

We briefly recall the structure of the set of covariance matrices that is introduced in Chapter 1. The set of n-dimensional covariance matrices is defined as

$$\mathrm{Cov}(n) := \{\Sigma \in \mathrm{Sym}(n) \,|\, \Sigma \geq 0\}\,,$$

where $\mathrm{Sym}(n)$ denotes the space of symmetric $n \times n$ matrices. It has the structure of stratified space with smooth strata

$$\mathrm{Sym}(n, k) := \{\Sigma \in \mathrm{Cov}(n) \,|\, \mathrm{rk}(\Sigma) = k\}\,.$$

When endowed with the Bures–Wasserstein distance, defined as

$$d^{BW}(\Sigma,\Lambda)^2 = \mathrm{Tr}(\Sigma) + \mathrm{Tr}(\Lambda) - 2\,\mathrm{Tr}((\Sigma^{1/2}\Lambda\Sigma^{1/2})^{1/2}),$$

the cone $\mathrm{Cov}(n)$ is characterized as a complete geodesic space of nonnegative curvature [4].

The smooth strata are thus equipped with a resulting Riemannian structure. Denoting $\Sigma = QDQ^T$ with $D = \mathrm{diag}(d_1, ..., d_k)$ positive definite, $Q \in \mathrm{St}(n,k)$ and $S_{\Sigma,V} = QS_D(Q^TVQ)Q^T$, where $S_A(B)$ is the unique solution to the Sylvester equation $B = AS + SA$ for $A \in SPD(n)$ and $B \in \mathrm{Sym}(n)$ (Appendix No. 13), we obtain the inner product on $T_\Sigma \mathrm{Sym}^+(n,k)$, i.e.,

$$g_\Sigma^{BW(n,k)}(V,W) = \mathrm{Tr}(S_{\Sigma,V}\Sigma S_{\Sigma,W}) + \mathrm{Tr}(P_\Sigma^\perp V\Sigma^\dagger W). \tag{3.1}$$

In (3.1), V, $W \in T_\Sigma \mathrm{Sym}(n,k)$ and $\Sigma^\dagger$ denotes the Moore–Penrose inverse of Σ [86], [94]. The Bures–Wasserstein metric clearly shows two components. The first one is denoted by

$$g_\Sigma^{\mathrm{hor}}(V,W) := \mathrm{Tr}(P_\Sigma^\perp V\Sigma^\dagger W)$$

and acts on the subspace

$$\mathcal{H}_\Sigma \mathrm{Sym}(n,k) := \{V \in \mathrm{Sym}(n) \,|\, P_\Sigma^\perp V = 0\} \subset T_\Sigma \mathrm{Sym}(n,k).$$

The second one is denoted as

$$g_\Sigma^{\mathrm{ver}}(V,W) := \mathrm{Tr}(S_{\Sigma,V}\Sigma S_{\Sigma,W}),$$

and acts on the orthogonal complement of $\mathcal{H}_\Sigma \mathrm{Sym}(n,k)$. Thus, the metric is written as

$$g_\Sigma^{BW(n,k)}(V,W) = g_\Sigma^{\mathrm{hor}}(V,W) + g_\Sigma^{\mathrm{ver}}(V,W).$$

In this chapter we introduce the diffeomorphic model

$$\mathrm{Sym}(n,k) \;\cong\; M(n,k) := \mathrm{St}(n,k) \times_{O(k)} \mathrm{Sym}^+(k) = (\mathrm{St}(n,k) \times \mathrm{Sym}^+(k))/O(k),$$

where $\mathrm{St}(n,k)$ denotes the Stiefel manifold and $M(n,k)$ the bundle associated to the principal $O(k)$-bundle $\pi_G : \mathrm{St}(n,k) \to \mathrm{Gr}(k,n)$, see Subsection 3.1.2. This associated-bundle model has been used in several applied works on low-rank covariance and subspace estimation (see, e.g., [26], [27], [81]), typically endowed with product-type Riemannian metrics. However, to the best of our knowledge, a systematic analysis of this model equipped with the Bures–Wasserstein metric—and its interplay with the Grassmannian geometry of $\mathrm{Gr}(k,n)$—has not been carried out.

The main contributions of this paper can be summarized as follows.

(i) We provide a detailed analysis of the associated-bundle model

$$\mathrm{Sym}^+(n,k) \cong \mathrm{St}(n,k) \times_{O(k)} \mathrm{Sym}^+(k)$$

under the Bures–Wasserstein metric. In particular, we characterize the tangent spaces, the horizontal–vertical splitting, and the pullback metric on $M(n,k)$. This model gives a clear separation between changes and movements in the support subspace and the covariance within that subspace and it provides a standpoint to develop statistical procedures that are subjective to this split. We stress again that the diffeomorphic model is not novel per se, but it has never been studied under the Bures–Wasserstein metric and a rigorous geometric perspective helps us to develop the further results named below.

(ii) Within this model, we prove that the fibers $M_P(n,k)$, $P \in \mathrm{Gr}(k,n)$, are totally geodesic submanifolds. Equivalently, Bures–Wasserstein geodesics with vertical initial velocity remain in a fixed fiber, and geodesics joining two matrices with the same image stay in that image subspace. This gives a transparent geometric explanation of the algebraic uniqueness conditions for geodesics within a stratum observed in [106] and may prove useful in the development of algorithms aware of fiber preserving flows, e.g. in constrained optimization.

(iii) We derive a system of ordinary differential equations (ODEs) for Bures–Wasserstein geodesics in bundle coordinates on $M(n,k)$. In particular, we identify a simple matrix–valued first integral which can be interpreted as the momentum associated with the $O(k)$–symmetry of the model. This conservation law decouples part of the dynamics and will be exploited in low-dimensional examples to recover known closed-form expressions for fixed-rank Bures–Wasserstein geodesics. This ODE formulation is well suited for numerical integration and algorithms on low-rank covariance manifolds.

(iv) We establish a one-to-one correspondence between the index set of Bures–Wasserstein logarithms on $(\mathrm{Sym}^+(n,k), d^{BW})$ and the index set of Grassmannian logarithms on $(\mathrm{Gr}(k,n), d_E)$, where d_E denotes the Euclidean metric on the Grassmannian. This clarifies the precise relationship between multiplicities of minimizing geodesics in the Grassmann and fixed-rank covariance geometries.

The rest of the chapter is organized as follows. In this Section, we introduce the needed geometrical framework. In particular, in Section 3.1.1 we introduce some definitions and useful results of differential geometry, while in Section 3.1.2 we describe the principal fiber bundle $\pi_G : \mathrm{St}(n,k) \to \mathrm{Gr}(k,n)$.

In Section 3.2, we get into the heart of the topic and study the geometry of the stratum $\mathrm{Sym}(n,k)$. In Section 3.2.1, we introduce and prove the diffeomorphism $\mathrm{Sym}(n,k) \cong M(n,k)$; in Section 3.2.2 we give a characterization of the tangent spaces to $M(n,k)$; in Section 3.2.3, we study how the Bures–Wasserstein metric on $\mathrm{Sym}(n,k)$ translates to a metric on $M(n,k)$; in Section 3.2.4, we analyze the geodesics in $M(n,k)$ and find a system of geodesic equations and finally in Section 3.2.4.3 a bijection between the logarithms of $\mathrm{Sym}(n,k)$ and $\mathrm{Gr}(k,n)$ is proved.

3.1.1 Fiber bundles

Here we introduce the concepts from differential geometry that we will need in later Sections. We refer to [57], [68], [99] for a comprehensive treatment.

Definition 3.1.1. Let E, B and F be smooth manifolds and $\pi : E \to B$ a smooth surjection. The quadruple (E, B, π, F) is called a *smooth fiber bundle* with fiber F if for every $b \in B$ there exists a neighbourhood $U \subset B$ and a diffeomorphism

$$\psi_U : \pi^{-1}(U) \to U \times F$$

such that the following diagram commutes:

$$\begin{tikzcd} \pi^{-1}(U) \arrow[r, "\psi_U"] \arrow[dr, "\pi"'] & U \times F \arrow[d, "\mathrm{pr}_1"] \\ & U \end{tikzcd}$$

where pr_1 denotes the canonical projection onto U. The manifold B is called the *base space*, E the *total space*, and π the *projection*. For $b \in B$, the set $E_b := \pi^{-1}(b)$ is diffeomorphic to F and is called the *fiber* over b. The family $\{(U, \psi_U)\}$ is called a (smooth) *local trivialization* of the bundle.

Fiber bundles will often be denoted simply by (E, π) or by $\pi : E \to B$.

Definition 3.1.2. Let $\pi_1 : E_1 \to B$ and $\pi_2 : E_2 \to B$ be two fiber bundles over the same base space. A smooth map $\phi : E_1 \to E_2$ is called a *bundle morphism* if it respects the projections, i.e.,

$$\pi_1 = \pi_2 \circ \phi.$$

We observe that ϕ takes every fiber $\pi_1^{-1}(b)$ in E_1 into the corresponding fiber $\pi_2^{-1}(b)$ in E_2.

Definition 3.1.3. A bundle morphism $\phi : E_1 \to E_2$ is called an *isomorphism between fiber bundles* or a *bundle isomorphism* if it is a diffeomorphism. The two bundles $\pi_1 : E_1 \to B$ and $\pi_2 : E_2 \to B$ are called *isomorphic*.

Definition 3.1.4. Let G be a Lie group. We call *principal G-bundle* a fiber bundle $\pi : P \to B$ to which is associated a right action $R_g : P \to P$, $x \to R_g x = xg$ that satisfies the following properties.

(i) for each $g \in G$, R_g is a diffeomorphism;

(ii) the action preserves the fibers and is free and transitive on them, i.e., for every fiber $P_b = \pi^{-1}(b)$ and every pair of points $x, y \in P_b$ there exists a unique $g \in G$ such that $y = xg$.

The properties (i) and (ii) imply that every fiber P_b is diffeomorphic to G and can be identified with G itself.

Definition 3.1.5. Let $\pi : P \to B$ be a principal G-bundle. Let F be a smooth manifold equipped with a left action by G defined as $L_g : F \to F,\ x \to L_g x = g \cdot x$. Define a G action on $P \times F$ as

$$(p, x) \cdot g = (pg, g^{-1} \cdot x).$$

Denote the total space as $E = P \times_G F := (P \times F)/G$. We call *fiber bundle associated to the principal G-bundle*, or *associated G-bundle*, the fiber bundle defined by the projection

$$\tilde{\pi} : E \to B, \quad [p, x] \mapsto \pi(p),$$

where $[p, x]$ denotes an equivalence class in E.

Proposition 3.1.1. *For a $P \to B$ principal G-bundle and an associated bundle $E = P \times_G F$ with fiber F and left action $L_g f = g \cdot f$ the tangent spaces are given by*

$$T_{[p,f]}E = (T_p P \times T_f F)/\{p\xi, -L^*_\xi r \,|\, \xi \in \mathfrak{g}\},$$

where $\mathfrak{g}$ is the Lie algebra of G and $L^*_\xi := \partial_t L_{\exp(t\xi)}|_{t=0}$.

Proof. Let $q : P \times F \to E$ be the canonical projection defining the associated bundle $E = P \times_G F$. By construction, q is a submersion and its differential

$$dq_{(p,f)} : T_p P \times T_f F \to T_{[p,f]}E$$

is surjective, with kernel equal to the tangent space to the G-orbit $O_{(p,f)}$ through (p, f).

For any $\xi \in \mathfrak{g}$, consider the one-parameter subgroup $\exp(t\xi)$ and the curve

$$\gamma_\xi(t) := (p, f) \cdot \exp(t\xi) = \big(p \exp(t\xi),\ L_{\exp(-t\xi)} f\big),$$

which lies entirely in $O_{(p,f)}$. Differentiating at $t = 0$ gives

$$\partial_t \gamma_\xi(t)\Big|_{t=0} = (p\xi, -L^*_\xi f).$$

The set of such vectors, as ξ varies in $\mathfrak{g}$, spans the tangent space to the orbit at (p, f). Hence

$$\operatorname{Ker}(dq_{(p,f)}) = \{(p\xi, -L_\xi^* f) \mid \xi \in \mathfrak{g}\}.$$

Since $dq_{(p,f)}$ is surjective, we obtain

$$T_{[p,f]}E = (T_pP \times T_fF)/\operatorname{Ker}(dq_{(p,f)}) = (T_pP \times T_fF)/\{(p\xi, -L_\xi^* f) \mid \xi \in \mathfrak{g}\},$$

which is the claimed description. □

With the same notation as in Proposition 3.1.1, denoting by $\mathrm{pr}_1 : B \times F \to B$ the canonical projection onto the base space, by $q : P \times F \to E$ the quotient map, we get

$$\tilde{\pi} \circ q = \pi \circ \mathrm{pr}_1.$$

Differentiating at $(p, f) \in E$ yields

$$d\tilde{\pi}([p, f]) \circ dq(p, f) = d\pi(p) \circ d\mathrm{pr}_1(p, f)$$

and thus, for $[x, y] \in T_{[p,f]}E$

$$d\tilde{\pi}([p, f])([x, y]) = d\tilde{\pi}([p, f]) \circ dq(p, f)(x, y) = d\pi(p) \circ d\mathrm{pr}_1(p, f)(x, y) = d\pi(p)(x). \tag{3.2}$$

This gives the vertical spaces

$$\mathcal{V}_{[p,f]} := \operatorname{Ker}(d\tilde{\pi}([p, f])) = \{[x, y] \in T_{[p,f]}E \mid x \in \operatorname{Ker}(d\pi(p))\}. \tag{3.3}$$

We will also need the following Lemma, which allows us to upgrade fiberwise diffeomorphisms to a global bundle isomorphism.

Lemma 3.1.2. *Let $\pi_M : M \to B$, $\pi_N : N \to B$ be two smooth fiber bundles over the same base space. Let $\phi : M \to N, \pi_N = \pi_M \circ \phi$ be a smooth bundle morphism such that the restriction on the fibers is a diffeomorphism. Then ϕ is a global diffeomorphism.*

Proof. We first show that ϕ is a bijection.

Surjectivity. Let $y \in N$ and set $b := \pi_N(y) \in B$. The restriction

$$\phi_b := \phi\big|_{\pi_M^{-1}(b)} : \pi_M^{-1}(b) \to \pi_N^{-1}(b)$$

is a diffeomorphism by assumption. In particular, $y \in \pi_N^{-1}(b)$ admits a preimage $x \in \pi_M^{-1}(b)$ with $\phi(x) = y$.

Injectivity. Suppose $\phi(x_1) = \phi(x_2)$ for $x_1, x_2 \in M$. Then

$$\pi_M(x_1) = \pi_N(\phi(x_1)) = \pi_N(\phi(x_2)) = \pi_M(x_2),$$

so x_1 and x_2 lie in the same fiber of π_M. Since the restriction of ϕ to that fiber is a diffeomorphism, it is injective there, and thus $x_1 = x_2$.

We have shown that ϕ is a bijection.

Since M and N are smooth fiber bundles and their fibers are diffeomorphic to each other, there exists an open cover $\{U_\alpha\}$ of B and smooth local trivializations

$$\psi_\alpha^M : \pi_M^{-1}(U_\alpha) \to U_\alpha \times F, \qquad \psi_\alpha^N : \pi_N^{-1}(U_\alpha) \to U_\alpha \times F,$$

where F is a model fiber diffeomorphic to the fibers of both bundles. In these coordinates we can write

$$\widetilde{\phi}_\alpha := \psi_\alpha^N \circ \phi \circ (\psi_\alpha^M)^{-1} : U_\alpha \times F \to U_\alpha \times F.$$

Because ϕ is a bundle morphism over the identity on B, $\widetilde{\phi}_\alpha$ takes the form

$$\widetilde{\phi}_\alpha(b, x) = (b, f_\alpha(b, x)),$$

where $f_\alpha : U_\alpha \times F \to F$ is smooth and, for each $b \in U_\alpha$, the map

$$f_{\alpha,b} := f_\alpha(b, \cdot) : F \to F$$

is a diffeomorphism (since ϕ restricts to a diffeomorphism on each fiber).

The differential of $\widetilde{\phi}_\alpha$ at (b, x) has block form

$$D\widetilde{\phi}_\alpha(b, x) = \begin{bmatrix} I & 0 \\ * & D_x f_{\alpha,b} \end{bmatrix},$$

where I is the identity on $T_b U_\alpha$ and $D_x f_{\alpha,b}$ is the differential of $f_{\alpha,b}$ at x. Because each $f_{\alpha,b}$ is a diffeomorphism, $D_x f_{\alpha,b}$ is invertible, so $D\widetilde{\phi}_\alpha(b, x)$ is invertible as well. Thus $\widetilde{\phi}_\alpha$ is a local diffeomorphism, and since it is also globally bijective on $U_\alpha \times F$, it is a global diffeomorphism there, with inverse

$$\widetilde{\phi}_\alpha^{-1}(b, x) = (b, f_{\alpha,b}^{-1}(x)).$$

It follows that ϕ is smooth and has a smooth inverse in each local trivialization chart, and hence globally on M and N. Therefore ϕ is a diffeomorphism. □

3.1.2 The Grassmannian as a principal fiber bundle

We now briefly review the well-known principal $O(k)$-bundle

$$\pi_G : \mathrm{St}(n, k) \to \mathrm{Gr}(k, n), \qquad Q \mapsto \mathrm{Span}(Q),$$

where $\mathrm{St}(n, k)$ denotes the Stiefel manifold of $n \times k$ matrices with orthonormal columns, and $\mathrm{Gr}(k, n)$ the Grassmann manifold of k-dimensional subspaces of $\mathbb{R}^n$. For a comprehensive review we refer the reader to [16].

Let $Q \in \mathrm{St}(n, k)$. The tangent space to $\mathrm{St}(n, k)$ at Q can be characterized as

$$T_Q \,\mathrm{St}(n, k) = \{V \in \mathrm{Mat}(n, k) \,|\, Q^T V + V^T Q = 0\}.$$

If we fix a matrix $Q_\perp \in \mathrm{Mat}(n, n-k)$ that completes Q to an orthonormal basis of $\mathbb{R}^n$, then every $V \in T_Q \,\mathrm{St}(n, k)$ can be uniquely decomposed as

$$V = QA + Q_\perp B, \qquad A \in \mathfrak{o}(k) = \mathrm{Skew}(k),\ B \in \mathrm{Mat}(n-k, k),$$

where $\mathfrak{o}(k)$ denotes the Lie algebra of the orthogonal group $O(k)$ (see [16, Section 2.2]).

We identify the tangent space to $\mathrm{Gr}(k, n)$ at a point P with

$$T_P \,\mathrm{Gr}(k, n) \cong \mathrm{Hom}(P, P^\perp).$$

Once a basis is specified, we may further identify $\mathrm{Hom}(P, P^\perp) \cong \mathrm{Mat}(n-k, k)$.

The vertical space at Q is given by

$$\mathcal{V}_Q := \mathrm{Ker}\left(d\pi_G(Q)\right) = \{QA \,|\, A \in \mathfrak{o}(k)\}.$$

The standard construction (see [16]) is to endow both $\mathrm{St}(n, k)$ and $\mathrm{Gr}(k, n)$ with the Euclidean (Frobenius) metric, i.e.

$$\begin{aligned} g_Q^{\mathrm{St}}(V_1, V_2) &= \mathrm{Tr}(V_1 V_2^T), \qquad V_1, V_2 \in T_Q \,\mathrm{St}(n, k), \\ g_P^{\mathrm{Gr}}(B_1, B_2) &= \mathrm{Tr}(B_1 B_2^T), \qquad B_1, B_2 \in \mathrm{Mat}(n-k, k) \cong T_P \,\mathrm{Gr}(k, n), \end{aligned} \tag{3.4}$$

where the Grassmannian metric is independent of the chosen matrix representatives of P.

The horizontal space at Q is then defined as the orthogonal complement of $\mathcal{V}_Q$ in $T_Q \,\mathrm{St}(n, k)$ with respect to the Euclidean metric, namely

$$\mathcal{H}_Q := \mathcal{V}_Q^\perp = \{Q_\perp B \,|\, B \in \mathrm{Mat}(n-k, k)\}.$$

Proposition 3.1.3. *The projection* $\pi_G : (\mathrm{St}(n, k), g^{\mathrm{St}}) \to (\mathrm{Gr}(k, n), g^{\mathrm{Gr}})$ *is a Riemannian submersion, that is the restriction*

$$d\pi_G(Q)\big|_{\mathcal{H}_Q} : \mathcal{H}_Q \to T_P \,\mathrm{Gr}(k, n),$$

with $P = \mathrm{Span}(Q)$*, is an isometry for every* $Q \in \mathrm{St}(n, k)$*.*

Proof. Let $Q \in \mathrm{St}(n, k)$, and let $Q_\perp$ be a completion to an orthonormal basis of $\mathbb{R}^n$, so that $[Q \; Q_\perp] \in O(n)$. Consider a horizontal tangent vector $V \in T_Q \,\mathrm{St}(n, k)$, which by definition can be written as $V = Q_\perp B$ for some $B \in \mathrm{Mat}(n-k, k)$.

Let $Q(t) \in \mathrm{St}(n, k)$ be a smooth curve with $Q(0) = Q$ and $\dot{Q}(0) = V$. For any $x \in \mathbb{R}^k$, consider the curve $v(t) = Q(t)x \in \mathbb{R}^n$, which lies in the evolving subspace $P(t) := \pi_G(Q(t)) \in \mathrm{Gr}(k, n)$. Then

$$\dot{v}(0) = \dot{Q}(0)x = Q_\perp Bx.$$

Since $x \in \mathbb{R}^k$ is arbitrary, the derivative $d\pi_G(Q)(V)$, viewed in the coordinates induced by the orthonormal basis $[Q\ Q_\perp]$ of $\mathbb{R}^n$, is exactly the matrix $B \in \mathrm{Mat}(n-k, k)$. In particular,

$$d\pi_G(Q) : \mathcal{H}_Q \to T_{\mathrm{Span}(Q)} \mathrm{Gr}(k, n) \cong \mathrm{Mat}(n-k, k)$$

is an isomorphism, and for $V_1 = Q_\perp B_1$ and $V_2 = Q_\perp B_2$ we have

$$g_Q^{\mathrm{St}}(V_1, V_2) = \mathrm{Tr}(B_1 B_2^T) = g_{\mathrm{Span}(Q)}^{\mathrm{Gr}}(d\pi_G(Q)(V_1), d\pi_G(Q)(V_2)),$$

by the definitions in (3.4). Thus $d\pi_G(Q)$ restricts to an isometry from $\mathcal{H}_Q$ onto $T_{\mathrm{Span}(Q)} \mathrm{Gr}(k, n)$, so π_G is a Riemannian submersion. □

3.2 The Bures–Wasserstein Geometry of the fixed rank stratum

3.2.1 An associated bundle model of fixed rank covariance matrices

As recalled in Section 3.1.2, $\mathrm{St}(n, k)$ carries a principal $O(k)$-bundle structure with base space $\mathrm{Gr}(k, n)$ and structure group $O(k)$ acting by right multiplication. The projection

$$\pi_G : \mathrm{St}(n, k) \to \mathrm{Gr}(k, n), \qquad Q \mapsto \mathrm{Span}(Q),$$

is a surjective submersion and the right action of $O(k)$ on $\mathrm{St}(n, k)$ is free and proper.

We equip $\mathrm{Sym}^+(k)$ with the smooth left $O(k)$-action

$$G \cdot S := GSG^T$$

and form the associated fiber bundle

$$M(n, k) = \mathrm{St}(n, k) \times_{O(k)} \mathrm{Sym}^+(k) = (\mathrm{St}(n, k) \times \mathrm{Sym}^+(k))/O(k),$$

where the equivalence relation is

$$(Q, D) \sim (QG, G^T DG), \qquad G \in O(k),$$

and equivalence classes are denoted by $[Q, D]$. The projection of the associated bundle

$$\pi_M : M(n, k) \to \mathrm{Gr}(k, n)$$

is defined by

$$\pi_M([Q, D]) = \pi_G(Q) = \mathrm{Span}(Q),$$

and the fiber over $P \in \mathrm{Gr}(k, n)$ is

$$M_P(n, k) = \{[Q, D] \,|\, \mathrm{Span}(Q) = P\}.$$

It is a standard fact in the theory of associated fiber bundles that every fiber $M_P(n, k)$ is diffeomorphic to the model fiber $\mathrm{Sym}^+(k)$ (see, for example, [63, Chapter 1] and [57, Chapter 4]).

On the other hand, we consider the bundle

$$\pi_{\mathrm{Sym}} : \mathrm{Sym}^+(n, k) \to \mathrm{Gr}(k, n), \quad \Sigma \mapsto \mathrm{Im}(\Sigma),$$

where $\mathrm{Im}(\Sigma)$ denotes the image (column space) of Σ. The fiber over $P \in \mathrm{Gr}(k, n)$ is

$$\mathrm{Sym}^+_P(n, k) = \{\Sigma \in \mathrm{Sym}(n) \,|\, \Sigma \geq 0,\ \mathrm{rk}(\Sigma) = k,\ \mathrm{Im}(\Sigma) = P\}.$$

Thus $\mathrm{Sym}^+(n, k)$ is also a fiber bundle over $\mathrm{Gr}(k, n)$ with the same base space as $M(n, k)$.

We define the map

$$\phi : M(n, k) \to \mathrm{Sym}^+(n, k), \quad [Q, D] \mapsto QDQ^T.$$

It is readily checked that it is well defined, i.e. the value of ϕ does not depend on the representative of the class $[Q, D]$. Moreover, ϕ preserves fibers:

$$\pi_{\mathrm{Sym}}(\phi([Q, D])) = \mathrm{Im}(QDQ^T) = \mathrm{Span}(Q) = \pi_M([Q, D]).$$

Hence, ϕ is a smooth bundle morphism from $M(n,k)$ to $\mathrm{Sym}^+(n,k)$ over $\mathrm{Gr}(k,n)$.

Proposition 3.2.1. *The smooth bundle morphism* $\phi : M(n,k) \to \mathrm{Sym}^+(n,k)$ *is a fiber-wise diffeomorphism, i.e., the restriction*

$$\phi|_{M_p(n,k)} : M_p(n,k) \to Sym^+_P(n,k)$$

is a diffeomorphism.

Proof. Fix $P \in \mathrm{Gr}(k,n)$ and choose $Q_0 \in \mathrm{St}(n,k)$ such that $\mathrm{Span}(Q_0) = P$. Every class in the fiber $M_P(n,k)$ can be represented as $[Q_0, D]$ for some $D \in \mathrm{Sym}^+(k)$. Restricting ϕ to $M_P(n,k)$ and using the identification $M_P(n,k) \cong \mathrm{Sym}^+(k)$ we obtain the map

$$\phi_{M_P(n,k)} : \mathrm{Sym}^+(k) \to \mathrm{Sym}^+_P(n,k); \quad D \mapsto Q_0 D Q_0^T,$$

which is clearly a diffeomorphism with inverse $\Sigma \mapsto Q_0^T \Sigma Q_0$. □

We can now apply Lemma 3.1.2 to the bundle morphism $\phi : M(n,k) \to \mathrm{Sym}^+(n,k)$.

Theorem 3.2.2. *Let*

$$M(n,k) = \mathrm{St}(n,k) \times_{O(k)} \mathrm{Sym}^+(k) = \big(\mathrm{St}(n,k) \times \mathrm{Sym}^+(k)\big)/O(k).$$

Then the map

$$\phi : M(n,k) \xrightarrow{\cong} \mathrm{Sym}^+(n,k), \qquad [Q,D] \mapsto QDQ^T,$$

is a global diffeomorphism of fiber bundles over $\mathrm{Gr}(k,n)$*. In particular, each fiber of* $\mathrm{Sym}^+(n,k)$ *is diffeomorphic to* $\mathrm{Sym}^+(k)$*.*

Practically, we can now assign in a smooth manner a class $[Q,D] \in M(n,k)$ to any matrix $\Sigma = QDQ^T$ (with $Q \in \mathrm{St}(n,k)$ and $D \in \mathrm{Sym}^+(k)$), and conversely, any class $[Q,D]$ corresponds to the fixed-rank covariance matrix $\Sigma = QDQ^T$.

3.2.2 Tangent spaces

We now identify the tangent spaces of $M(n,k)$ by means of Proposition 3.1.1. In our case, the left action of $O(k)$ on $\mathrm{Sym}^+(k)$ is given by

$$L_G(D) = GDG^T.$$

For $A \in \mathfrak{o}(k)$ and $D \in \mathrm{Sym}^+(k)$, the infinitesimal action is

$$L_A^*(D) = \partial_t\big(\exp(tA)^T\, D\, \exp(tA)\big)\Big|_{t=0} = DA - AD = \{D, A\},$$

where we denote the matrix commutator by $\{D, A\} = DA - AD$ to avoid confusion with equivalence classes. Hence, by Proposition 3.1.1, the subspace that is quotiented out in $T_{[Q,D]}M(n,k)$ is

$$\{(QA, -\{D, A\}) \,|\, A \in \mathfrak{o}(k)\},$$

and the equivalence relation on $T_Q\,\mathrm{St}(n,k) \times T_D\,\mathrm{Sym}^+(k)$ is

$$(V, T) \sim (V + QA, T - \{D, A\}), \qquad A \in \mathfrak{o}(k).$$

Recall that every $V \in T_Q\,\mathrm{St}(n,k)$ can be uniquely decomposed as

$$V = QA + Q_\perp B, \qquad A \in \mathfrak{o}(k),\; B \in \mathrm{Mat}(n-k, k),$$

for a suitable completion $Q_\perp$ of Q to an orthonormal basis of $\mathbb{R}^n$. Using the above equivalence relation, any pair (V, T) is equivalent to a unique representative of the form $(Q_\perp B, \widetilde{T})$, obtained by absorbing the QA component into the vertical part via $T \mapsto T - \{D, A\}$.

Thus we obtain the following description:

$$\begin{aligned} T_{[Q,D]}M(n,k) =& \{[Q_\perp B, T] \,|\, B \in \mathrm{Mat}(n-k,k),\; T \in \mathrm{Sym}(k)\} \\ \cong& \{(Q_\perp B, T) \,|\, B \in \mathrm{Mat}(n-k,k),\; T \in \mathrm{Sym}(k)\}, \end{aligned}$$

where, for notational convenience, we identify each class with its representative of the form

$(Q_\perp B, T)$.

Now, Equation (3.2) can be applied to this case, yielding

$$d\pi_M([Q,D])([Q_\perp B, T]) = d\pi_G(Q)(Q_\perp B) = B.$$

This yields the following horizontal–vertical splitting:

$$\begin{aligned}
\mathcal{V}_{[Q,D]} &= \operatorname{Ker}(d\pi_M([Q,D])) = \{[0,T] \mid T \in \operatorname{Sym}(k)\}, \cong \{(0,T) \mid T \in \operatorname{Sym}(k)\},\\
\mathcal{H}_{[Q,D]} &= \{[Q_\perp B, 0] \mid B \in \operatorname{Mat}(n-k,k)\} \cong \{(Q_\perp B, 0) \mid B \in \operatorname{Mat}(n-k,k)\}.
\end{aligned}$$

In other words, in the bundle coordinates $[Q,D]$ with tangent representatives $(Q_\perp B, T)$, horizontal directions are encoded by B (moving the subspace in the Grassmannian) and vertical directions by T (changing the covariance within a fixed subspace).

3.2.3 Pullback metric

In our setting, the manifold $\operatorname{Sym}^+(n,k)$ is endowed with the Bures–Wasserstein metric $g^{BW(n,k)}$, and $M(n,k)$ is equipped with the diffeomorphism

$$\phi\colon M(n,k) \xrightarrow{\cong} \operatorname{Sym}^+(n,k), \qquad [Q,D] \mapsto QDQ^T$$

from Theorem 3.2.2. This allows us to define the Riemannian metric h on $M(n,k)$ as the pullback (Appendix No. 23)

$$h := \phi^* g^{BW(n,k)}.$$

Proposition 3.2.3. *Let $[Q,D] \in M(n,k)$ with $Q \in \operatorname{St}(n,k)$, $D \in \operatorname{Sym}^+(k)$, and let $Q_\perp \in \operatorname{St}(n,n-k)$ be such that $[Q\,Q_\perp] \in O(n)$. Moreover, let V, $W \in T_{[Q,D]}M(n,k)$ be written as*

$$V := [Q_\perp B_V, T_V], \qquad W := [Q_\perp B_W, T_W],$$

with $B_V, B_W \in \operatorname{Mat}(n-k,k)$ and $T_V, T_W \in \operatorname{Sym}(k)$. Then the pullback metric $h = \phi^* g^{BW(n,k)}$ on $M(n,k)$ satisfies

$$h_{[Q,D]}(V,W) = \operatorname{Tr}(B_V D B_W^T) + \operatorname{Tr}(S_D(T_V)\, D\, S_D(T_W)). \tag{3.5}$$

In particular, h is a well-defined Riemannian metric on $M(n,k)$, and $\phi\colon (M(n,k),h) \to (\mathrm{Sym}^+(n,k), g^{BW(n,k)})$ is a Riemannian isometry.

Proof. Take a smooth curve $[Q(t), D(t)] \in M(n,k)$ such that

$$Q(0) = Q, \quad \dot{Q}(0) = Q_\perp B, \qquad D(0) = D, \quad \dot{D}(0) = T,$$

with $B \in \mathrm{Mat}(n-k,k)$ and $T \in \mathrm{Sym}(k)$. Then the differential $d\phi_{[Q,D]}\colon T_{[Q,D]}M(n,k) \longrightarrow T_{QDQ^T}\,\mathrm{Sym}^+(n,k)$ is given by the equation

$$d\phi_{[Q,D]}(Q_\perp B, T) = \frac{d}{dt}\Big|_{t=0} Q(t)D(t)Q(t)^T = Q_\perp BDQ^T + QDB^T Q_\perp^T + QTQ^T.$$

Denote

$$\widetilde{V} := d\phi_{[Q,D]}(V), \qquad \widetilde{W} := d\phi_{[Q,D]}(W).$$

We are going to analyse separately the horizontal $g_\Sigma^{\mathrm{hor}}(\widetilde{V},\widetilde{W}) = \mathrm{Tr}\,(P_\Sigma^\perp \widetilde{V}\Sigma^\dagger \widetilde{W})$ and vertical $g_\Sigma^{\mathrm{ver}}(\widetilde{V},\widetilde{W}) = \mathrm{Tr}\left(S_{\Sigma,\widetilde{V}}\,\Sigma\,S_{\Sigma,\widetilde{W}}\right)$ parts of the metric.

Horizontal part. Using $P_\Sigma^\perp = Q_\perp Q_\perp^T$ and $\Sigma^\dagger = QD^{-1}Q^T$, we compute

$$P_\Sigma^\perp \widetilde{V} = Q_\perp Q_\perp^T(Q_\perp B_V DQ^T + QDB_V^T Q_\perp^T + QT_V Q^T) = Q_\perp B_V DQ^T,$$

because $Q_\perp^T Q = 0$ and $Q_\perp^T Q_\perp = I_{n-k}$. Thus

$$P_\Sigma^\perp \widetilde{V}\,\Sigma^\dagger = Q_\perp B_V DQ^T QD^{-1}Q^T = Q_\perp B_V Q^T.$$

Multiplying by $\widetilde{W}$ gives

$$\begin{aligned} P_\Sigma^\perp \widetilde{V}\,\Sigma^\dagger \widetilde{W} &= Q_\perp B_V Q^T(Q_\perp B_W DQ^T + QDB_W^T Q_\perp^T + QT_W Q^T) \\ &= Q_\perp B_V DB_W^T Q_\perp^T + Q_\perp B_V T_W Q^T. \end{aligned}$$

Taking the trace and using cyclicity,

$$g_\Sigma^{\mathrm{hor}}(\widetilde{V},\widetilde{W}) = \mathrm{Tr}\,(P_\Sigma^\perp \widetilde{V}\,\Sigma^\dagger \widetilde{W}) = \mathrm{Tr}\,(Q_\perp B_V DB_W^T Q_\perp^T) + \mathrm{Tr}\,(Q_\perp B_V T_W Q^T).$$

The second term vanishes because

$$\mathrm{Tr}\,(Q_\perp B_V T_W Q^T) = \mathrm{Tr}\,(Q^T Q_\perp B_V T_W) = 0,$$

so

$$g^{\mathrm{hor}}_\Sigma(\widetilde{V}, \widetilde{W}) = \mathrm{Tr}\,(Q_\perp B_V D B_W^T Q_\perp^T) = \mathrm{Tr}\,(B_V D B_W^T),$$

again by cyclicity of the trace.

Vertical part. For the vertical component, we use

$$\widetilde{V} = Q_\perp B_V D Q^T + Q D B_V^T Q_\perp^T + Q T_V Q^T,$$

and obtain

$$Q^T \widetilde{V} Q = Q^T Q_\perp B_V D Q^T Q + Q^T Q D B_V^T Q_\perp^T Q + Q^T Q T_V Q^T Q = T_V.$$

Thus

$$S_{\Sigma,\widetilde{V}} = Q S_D(T_V) Q^T,$$

and similarly $S_{\Sigma,\widetilde{W}} = Q S_D(T_W) Q^T$. Therefore,

$$\begin{aligned} g^{\mathrm{ver}}_\Sigma(\widetilde{V}, \widetilde{W}) &= \mathrm{Tr}\,(S_{\Sigma,\widetilde{V}}\, \Sigma\, S_{\Sigma,\widetilde{W}}) = \mathrm{Tr}\,(Q S_D(T_V) Q^T\; Q D Q^T\; Q S_D(T_W) Q^T) \\ &= \mathrm{Tr}\,(S_D(T_V) D S_D(T_W)), \end{aligned}$$

again by cyclicity of the trace.

Putting the horizontal and vertical contributions together, we obtain

$$h_{[Q,D]}(V, W) = g^{BW(n,k)}_\Sigma(\widetilde{V}, \widetilde{W}) = \mathrm{Tr}\,(B_V D B_W^T) + \mathrm{Tr}\,(S_D(T_V)\, D\, S_D(T_W)),$$

which proves (3.5). □

Remark 3.2.1. As can be observed in the proof of Proposition 3.2.3, the pullback metric naturally splits into an horizontal and a vertical part

$$h^{\mathrm{hor}}_{[Q,D]}(V,W) := \mathrm{Tr}\,(B_V D B_W^T),$$
$$h^{\mathrm{ver}}_{[Q,D]}(V,W) := \mathrm{Tr}\,(S_D(T_V)\,D\,S_D(T_W)).$$

Moreover, the horizontal metric retains vertical information, i.e., it is influenced by vertical position coordinate D.

3.2.4 Characterization of geodesics: totally geodesic fibers, geodesic ODE system and logarithm bijection

The present Section is subdivided into three further Subsections, namely Sections 3.2.4.1, 3.2.4.2 and 3.2.4.3, that contain three different results. Each of these Sections is independent from the others and can be taken as a standalone contribution, but together they convey a more complete characterization of geodesics in $M(n,k)$ and so in the smooth manifold $\mathrm{Sym}^+(n,k)$.

3.2.4.1 Totally geodesic fibers

We start by recalling the *Koszul formula*. Let (M,h) be a Riemannian manifold with Levi–Civita connection ∇. For vector fields X, Y, Z on M, the Koszul formula reads as

$$2h(\nabla_X Y, Z) = X(h(Y,Z)) + Y(h(Z,X)) - Z(h(X,Y)) + h(\{X,Y\},Z) - h(\{Y,Z\},X) - h(\{Z,X\},Y),$$

where $\{\cdot,\cdot\}$ denotes the Lie bracket of vector fields.

We apply this formula to the bundle $M(n,k)$ equipped with the pullback metric h from Section 3.2.3 and the horizontal–vertical splitting described in Section 3.2.2.

Proposition 3.2.4. *Let $M(n,k) = \mathrm{St}(n,k)\times_{O(k)}\mathrm{Sym}^+(k)$ with the metric h defined in Section 3.2.3. For each $P \in \mathrm{Gr}(k,n)$, the fiber*

$$M_P(n,k) = \{[Q,S] \in M(n,k) \mid \mathrm{Span}(Q) = P\}$$

is a totally geodesic submanifold of $M(n,k)$ (Appendix No. 31).

Proof. To show that each fiber $M_P(n,k)$ is totally geodesic it is enough to prove that if U and V are vertical vector fields, then $\nabla_U V$ is also vertical, i.e. its horizontal component vanishes.

Let U, V be vertical vector fields and X a horizontal vector field. We compute $h(\nabla_U V, X)$ using the Koszul formula. We use the following facts:

- $h(U,X) = h(V,X) = 0$, since vertical and horizontal directions are orthogonal.
- $\{U,X\}$ and $\{V,X\}$ are vertical; indeed, $d\pi_M(\{U,X\}) = \{d\pi_M(U), d\pi_M(X)\} = 0$ because $d\pi_M(U) = 0$, and similarly for $\{V,X\}$.
- $\{U,V\}$ is vertical, so $h(\{U,V\},X) = 0$.
- The vertical metric does not depend on the base coordinate: for vertical fields the function $\Sigma \mapsto h_\Sigma(U_\Sigma, V_\Sigma)$ depends only on the vertical coordinates, while X is horizontal, tangent to the base. Hence $X(h(U,V)) = 0$.

It follows that $h(\nabla_U V, X) = 0$ and so

$$\mathcal{H}(\nabla_U V) = 0,$$

where $\mathcal{H}(\cdot)$ denotes the horizontal projection. □

Remark 3.2.2. This is consistent with the fixed-rank Bures–Wasserstein geometry described in [106], and in fact provides a more geometric perspective on it. In particular, [106, Theorem 5.2] states that the Bures–Wasserstein geodesic connecting two matrices $\Sigma, \Lambda \in \mathrm{Sym}^+(n,k)$ is unique if and only if $\mathrm{rk}(\Sigma\Lambda) = \mathrm{rk}(\Sigma)$, i.e. if and only if Σ and Λ have the same image subspace. In our bundle picture this means that the two points lie in the same fiber $M_P(n,k)$, and Proposition 3.2.4 shows that the geodesic joining them remains inside this fiber. Thus the uniqueness criterion of [106] acquires a natural geometric interpretation, and, in addition, such instances can be studied within the full-rank Bures–Wasserstein framework on the fiber, where the geometry is technically and computationally more tractable.

3.2.4.2 ODE system for geodesic equations with examples

We now derive the geodesic equations on $M(n,k)$ via a variational approach (Appendix No. 32).

Theorem 3.2.5. *Let $M(n,k) = \mathrm{St}(n,k) \times_{O(k)} \mathrm{Sym}^+(k)$ be endowed with the pullback Bures–Wasserstein metric $h = \phi^* g^{BW(n,k)}$. Consider a smooth curve*

$$\gamma(t) = [Q(t), D(t)] \in M(n,k), \qquad t \in I \subset \mathbb{R},$$

with $Q(t) \in \mathrm{St}(n,k)$ and $D(t) \in \mathrm{Sym}^+(k)$. Let $Q_\perp(t)$ be any smooth choice of orthonormal complement of $Q(t)$ so that

$$\mathbf{Q}(t) := [Q(t) \;\; Q_\perp(t)] \in O(n),$$

and write the velocity as

$$\dot{\gamma}(t) = (Q_\perp(t)B(t), T(t)),$$

with $B(t) \in \mathrm{Mat}(n-k,k)$ and $T(t) \in \mathrm{Sym}(k)$. Then γ is a Bures–Wasserstein geodesic in $M(n,k)$ if and only if it satisfies the bundle coordinate system

$$\begin{cases} \dot{\mathbf{Q}} = \mathbf{Q}\,\mathbf{B}, & \mathbf{Q}(0) = \mathbf{Q}_0, \\ \dot{D} = DS + SD, & D(0) = D_0, \\ \dot{B} = -B(DS+SD)D^{-1}, & B(0) = B_0, \\ \dot{S} = B^T B - S^2, & S(0) = S_{D_0}(T(0)) = S_0, \end{cases} \tag{3.6}$$

where

$$\mathbf{B}(t) := \begin{bmatrix} 0 & -B(t)^T \\ B(t) & 0 \end{bmatrix} \in \mathfrak{o}(n) = \mathrm{Skew}(n).$$

In particular, the first line of (3.6) is equivalent to

$$\dot{Q} = Q_\perp B, \qquad \dot{Q}_\perp = -QB^T,$$

so that the columns of $Q(t)$ and $Q_\perp(t)$ remain orthonormal for all t.

Proof. With the metric h from Subsection 3.2.3, the kinetic energy (Lagrangian) is

$$\begin{aligned} L(B,D,T) &:= \frac{1}{2} h(\dot{\gamma}(t), \dot{\gamma}(t)) \\ &= \frac{1}{2}\mathrm{Tr}\,(BDB^T) + \frac{1}{2}\mathrm{Tr}\,(S_D(T)\,D\,S_D(T)) \end{aligned}$$

$$=: L^{\mathrm{hor}}(B, D) + L^{\mathrm{ver}}(D, T),$$

where $S_D(T)$ is the unique solution of the Sylvester equation

$$DS + SD = T.$$

A geodesic is a smooth curve γ such that L is stationary under all smooth variations with fixed endpoints. Writing L in terms of the coordinates (B, D, T) and using that L does not depend on Q, the Euler–Lagrange equations reduce to

$$\frac{d}{dt}(\partial_B L) = 0, \qquad \frac{d}{dt}(\partial_T L) - \partial_D L = 0. \tag{3.7}$$

We now compute the partial derivatives. Throughout we use the straightforward matrix-variational identity

$$\delta f(N) = \mathrm{Tr}\,(\delta N\,(\partial_N f)^T),$$

for a scalar function f of a matrix argument N, see [74] for a general reference on matrix calculus.

Derivative with respect to B. We have

$$\begin{aligned}\delta_B \,\mathrm{Tr}(BDB^T) &= \frac{d}{d\epsilon}\,\mathrm{Tr}\,((B + \varepsilon\,\delta B)D(B + \varepsilon\,\delta B)^T)\Big|_{\varepsilon=0}\\ &= \mathrm{Tr}(\delta BDB^T) + \mathrm{Tr}(BD\delta B^T)\\ &= 2\,\mathrm{Tr}(\delta BDB^T),\end{aligned}$$

so

$$\partial_B L = \partial_B L^{\mathrm{hor}} = BD.$$

The first Euler–Lagrange equation in (3.7) therefore gives

$$\frac{d}{dt}(BD) = 0 \qquad \Longrightarrow \qquad BD = K, \tag{3.8}$$

for some constant matrix $K \in \mathrm{Mat}(n-k, k)$ along the geodesic. Since D is invertible on $\mathrm{Sym}^+(k)$, we may write

$$\dot{B} = -B\,\dot{D}\,D^{-1} = -BTD^{-1}, \tag{3.9}$$

using the relation $\dot{D} = T$ (the vertical component T is precisely the derivative of D along γ). In terms of the Sylvester variable $S = S_D(T)$, we can equivalently write

$$\dot{B} = -B(DS + SD)D^{-1}.$$

Derivative with respect to T. For the vertical part we compute

$$\begin{aligned}
\delta_T \operatorname{Tr}\left(S_D(T)\,D\,S_D(T)\right) &= \operatorname{Tr}\left(S_D(\delta T)\,D\,S_D(T)\right) + \operatorname{Tr}\left(S_D(T)\,D\,S_D(\delta T)\right)\\
&= \operatorname{Tr}\left(S_D(\delta T)(DS_D(T) + S_D(T)D)\right)\\
&= \operatorname{Tr}\left(\delta T\,S_D(DS_D(T) + S_D(T)D)\right) = \operatorname{Tr}\left(\delta T\,S_D(T)\right),
\end{aligned}$$

where we used the linearity and self-adjointness of $S_D(\cdot)$ with respect to the Frobenius inner product, and the Sylvester equation

$$DS_D(T) + S_D(T)D = T.$$

Thus

$$\partial_T L = \partial_T L^{\mathrm{ver}} = \frac{1}{2} S_D(T). \tag{3.10}$$

Derivative with respect to D. For the horizontal part,

$$\delta_D \operatorname{Tr}(BDB^T) = \operatorname{Tr}(B^T B\,\delta D),$$

so

$$\partial_D L^{\mathrm{hor}} = \frac{1}{2} B^T B.$$

For the vertical part, we differentiate the Sylvester constraint

$$DS_D(T) + S_D(T)D = T$$

at fixed T, obtaining

$$D\,\delta_D S_D(T) + \delta_D S_D(T) D = -\delta D\, S_D(T) - S_D(T)\,\delta D,$$

and hence

$$\delta_D S_D(T) = -S_D(\delta D\, S_D(T) + S_D(T)\,\delta D).$$

A direct calculation then gives

$$\begin{aligned}\delta_D \operatorname{Tr}(S_D(T)\, D\, S_D(T)) &= \operatorname{Tr}(\delta_D S_D(T)\, D\, S_D(T)) \\ &\quad + \operatorname{Tr}(S_D(T)\,\delta D\, S_D(T)) + \operatorname{Tr}(S_D(T)\, D\,\delta_D S_D(T)) \\ &= \operatorname{Tr}(S_D(T)\,\delta D\, S_D(T)) - \operatorname{Tr}(\delta D\, S_D(T)^2) = -\operatorname{Tr}(\delta D\, S_D(T)^2),\end{aligned}$$

so

$$\partial_D L^{\mathrm{ver}} = -\frac{1}{2} S_D(T)^2.$$

Combining the two contributions, we obtain

$$\partial_D L = \frac{1}{2}(B^T B - S_D(T)^2). \tag{3.11}$$

Evolution equation for $S = S_D(T)$. Plugging expressions (3.10) and (3.11) into the second Euler–Lagrange equation in (3.7), we obtain

$$\frac{1}{2}\dot{S}_D(T) - \frac{1}{2}(B^T B - S_D(T)^2) = 0,$$

or, denoting $S = S_D(T)$,

$$\dot{S} = B^T B - S^2. \tag{3.12}$$

Finally, the evolution of $\mathbf{Q}$ is given by the horizontal component $Q_\perp B$ of the velocity and the skew-symmetric matrix $\mathbf{B}$ as in the statement. Conversely, any solution of (3.6) arises from a critical point of the energy functional with fixed endpoints, hence from a geodesic. This establishes the equivalence. As a final note, the system (3.6) is equivariant under the gauge change

$$(Q, D, B, S) \mapsto (QG, G^T DG, BG.G^T S G),$$

for constant $G \in O(k)$. Hence, the ODEs are well defined on the quotient $M(n,k)$. $\square$

Remark 3.2.3. As highlighted in the proof of Theorem 3.2.5, the Lagrangian

$$L(B,D,T) = \frac{1}{2}\operatorname{Tr}(BDB^{\top}) + \frac{1}{2}\operatorname{Tr}(S_D(T)\,D\,S_D(T))$$

does not depend on the base point $Q \in \mathrm{St}(n,k)$. In particular, the conjugate momentum with respect to the horizontal velocity B, that is

$$\partial_B L = BD \in \mathrm{Mat}(n-k,k),$$

is conserved along Bures–Wasserstein geodesics, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}(B(t)D(t)) = 0. \tag{3.13}$$

Equivalently, the matrix $K := BD$ is a first integral of the geodesic flow on $M(n,k)$, and can be viewed as the momentum associated with the $O(k)$-symmetry of the bundle model. This feature suggests the design of structure-preserving numerical schemes that exactly maintain BD, in analogy with angular-momentum-preserving integrators in classical mechanics. In the rank-one examples below this conservation law reduces the geodesic system to a one-dimensional problem and allows us to recover the known fixed-rank Bures–Wasserstein geodesics in closed form.

We now specialize the system to low-dimensional examples where we can solve it in closed form and explicitly recover the known Bures–Wasserstein geodesics on $\mathrm{Sym}^+(n,k)$.

Example 3.2.1. *Let* $n = 2$ *and* $k = 1$*. Then,* $\mathrm{St}(2,1) \cong S^1$ *and* $O(k) = \{\pm 1\}$*, so that* $\mathrm{Gr}(1,2) = \mathbb{RP}^1 \cong S^1/\{\pm 1\}$*, where* $\mathbb{RP}^1$ *denotes the real projective line of lines through the origin in* $\mathbb{R}^2$*. Moreover* $\mathrm{Sym}^+(1) = (0,+\infty)$ *and consequently, our manifold of study can be characterized as*

$$M(2,1) = (S^1 \times (0,+\infty))/\{\pm 1\}, \qquad \pi_M : M(2,1) \to \mathbb{R}P^1,$$

with the equivalence relation given by $(q,d) \sim (-q,d)$*. The global diffeomorphism is given by*

$$\phi : M(2,1) \xrightarrow{\cong} \mathrm{Sym}^+(2,1), \quad [q,d] \mapsto dqq^T.$$

If we write

$$q(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \qquad \theta \in (-\pi, \pi], \tag{3.14}$$

and the perpendicular vector

$$q_\perp(\theta) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix},$$

we get the bundle and tangent bundle coordinates

$$\begin{aligned} &[\theta, d], && \theta \in (-\pi, \pi],\ d \in (0, +\infty), \\ &(q_\perp(\theta) b, u), && b,\ u \in \mathbb{R}, \end{aligned}$$

where the equivalence class $[\theta, d]$ *is modulo the relation* $(\theta, d) \sim (\theta + \pi, d)$*.*

The solution to the Sylvester equation becomes

$$s := S_d(t) = \frac{u}{2d},$$

and the metric translates to

$$h_{[\theta,d]}(q_\perp b, u) = db^2 + \frac{u^2}{4d} = db^2 + ds^2.$$

The geodesic system of Theorem 3.2.5 transforms to

$$\begin{cases} \dot\theta = b, & \theta(0) = \theta_0, \\ \dot d = 2ds, & d(0) = d_0, \\ \dot b = -bd^{-1}(2ds) = -2bs, & b(0) = b_0, \\ \dot s = b^2 - s^2, & s(0) = s_0. \end{cases} \tag{3.15}$$

From (3.15), we immediately gather that the fibers are totally geodesic. Indeed $b_0 = 0$ *implies* $b \equiv 0$ *and consequently* $\dot\theta \equiv 0$*. The system reduces to*

$$\begin{cases} \dot d = 2ds, \\ \dot s = -s^2, \end{cases}$$

which gives the closed form geodesics

$$d(t) = d_0(1 + s_0 t)^2.$$

We observe that this equation matches the equation found by the authors in [106, Theorem 4.2], applied to the case $\mathrm{Sym}^+(1)$.

On the other hand, if $b_0 \neq 0$ we can still solve the system. Define $r = s/b$ and compute

$$\dot{r} = \frac{\dot{s}b - \dot{b}s}{b^2} = b(r^2 + 1)$$

and

$$\dot{b} = -2b^2 r.$$

Dividing,

$$\frac{db}{dr} = \frac{\dot{b}}{\dot{r}} = -\frac{2rb}{r^2 + 1}$$

or, equivalently,

$$\frac{d \ln b}{dr} = -\frac{2r}{r^2 + 1}.$$

Integrating, we find

$$b(r^2 + 1) = b_0(r_0^2 + 1) =: C, \quad \textit{constant}, \tag{3.16}$$

where $r_0 = s_0/b_0$. It follows that r grows linearly, $r(t) = r_0 + Ct$, and thus, by (3.16),

$$b(t) = \frac{C}{(Ct + r_0)^2 + 1},$$

and

$$s(t) = r(t)b(t) = \frac{C(Ct + r_0)}{(Ct + r_0)^2 + 1} = \frac{C^2 t + s_0(r_0^2 + 1)}{(Ct + r_0)^2 + 1}.$$

Since (3.13) gives $bd = b_0 d_0$, we get

$$d(t) = d_0 \frac{(Ct + r_0)^2 + 1}{r_0^2 + 1}. \tag{3.17}$$

Lastly, a straightforward integration of $\dot{\theta} = b$ *yields*

$$\theta(t) = \theta_0 + (\arctan(Ct + r_0) - \arctan(r_0)).$$

We see that geodesics with horizontal initial velocity ($s_0 = 0$*) do not enjoy the same nice qualities of vertical geodesics. The last equation in (3.15) shows that the vertical acceleration* $\dot{s}$ *depends on the horizontal velocity* b*. Geometrically, this is due to the horizontal metric being dependent on the vertical configuration coordinate* d*.*

Finally, we check whether the geodesic equations found match the ones computed by the authors in [106]. To start, the diffeomorphism ϕ *can be expressed in terms of the* (θ, d) *coordinates as*

$$[\theta, d] \mapsto dq(\theta)q(\theta)^T = d\begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix},$$

which is indeed a bidimensional rank one covariance matrix. Moreover, expressing $q(\theta(t))$ *in the basis* $\{q(\theta_0), q_\perp(\theta_0)\}$*, we get*

$$\begin{aligned} q(\theta(t)) &= \cos(\theta(t) - \theta_0)q(\theta_0) + \sin(\theta(t) - \theta_0)q_\perp(\theta_0) \\ &= \cos(\theta(t) - \theta_0)q_0 + \sin(\theta(t) - \theta_0)w_0, \end{aligned}$$

where, and henceforth, for simplicity of notation we denote $q_0 := q(\theta_0)$ *and* $w_0 := q_\perp(\theta_0)$*. Using cosine and sine addition formulas and the fact that* $\cos(\arctan x) = (1 + x^2)^{-1/2}$ *and* $\sin(\arctan x) = x(1 + x^2)^{-1/2}$*, we get*

$$\begin{aligned} \cos(\theta(t) - \theta_0) &= \cos(\arctan r(t) - \arctan r_0) = \frac{1 + r(t)r_0}{\sqrt{(1 + r(t)^2)(1 + r_0^2)}}, \\ \sin(\theta(t) - \theta_0) &= \sin(\arctan r(t) - \arctan r_0) = \frac{r(t) - r_0}{\sqrt{(1 + r(t)^2)(1 + r_0^2)}}. \end{aligned}$$

By (3.17), we get

$$\sqrt{d(t)} = \sqrt{d_0}\sqrt{\frac{1 + r(t)}{1 + r_0}},$$

so that

$$\begin{aligned}\sqrt{d(t)}q(\theta(t)) &= \sqrt{d_0}\sqrt{\frac{1+r(t)}{1+r_0}}\Big(\cos(\arctan r(t)-\arctan r_0)q_0+\sin(\arctan r(t)-\arctan r_0)w_0\Big)\\ &= \sqrt{d_0}\left(\frac{1+r(t)r_0}{1+r_0^2}q_0+\frac{r(t)-r_0}{1+r_0^2}w_0\right) = \sqrt{d_0}\left(\left(1+\frac{r_0Ct}{1+r_0^2}\right)q_0+\frac{Ct}{1+r_0^2}w_0\right)\\ &= \sqrt{d_0}\left((1+s_0t)\,q_0+b_0t\,w_0\right) = \sqrt{d_0}q_0+\sqrt{d_0}(s_0q_0+b_0w_0)t.\end{aligned}$$

This gives a geodesic equation on the manifold $\mathrm{Sym}^+(2,1)$*, namely*

$$\begin{aligned}\Sigma(t) =& d(t)q(\theta(t))q(\theta(t))^T\\ =&\left(\sqrt{d_0}q_0+\sqrt{d_0}(s_0q_0+b_0w_0)t\right)\left(\sqrt{d_0}q_0+\sqrt{d_0}(s_0q_0+b_0w_0)t\right)^T\\ =&d_0q_0q_0^T+d_0\left(2s_0q_0q_0^T+b_0\left(q_0w_0^T+w_0q_0^T\right)\right)t\\ &+d_0\left(s_0^2q_0q_0^T+b_0^2w_0w_0^T+s_0b_0\left(q_0w_0^T+w_0q_0^T\right)\right)t^2,\end{aligned}$$

that matches the equation found by the authors in [106] for the special case $n=2$, $k=1$ *once we notice that*

$$d_0q_0q_0^T=\Sigma(0)=\phi([\theta_0,d_0]),$$

$$\begin{aligned}d\phi_{[\theta_0,d_0]}(w_0b_0,u_0) =&u_0q_0q_0^T+d_0b_0\left(q_0w_0^T+w_0q_0^T\right)\\ =&d_0\left(2s_0q_0q_0^T+b_0\left(q_0w_0^T+w_0q_0^T\right)\right).\end{aligned}$$

With some additional effort, we can generalize Example 3.2.1 to the case $n\geq 2$, $k=1$.

Example 3.2.2. *Set* $n\in\mathbb{N}$ *and* $k=1$*. The geometric structure carries over verbatim, from the previous example, to*

$$M(n,1)=S^{n-1}\times_{\{\pm1\}}(0,+\infty),\qquad \pi_M:M(n,1)\to\mathbb{R}P^{n-1},$$

where S^{n-1} *denotes the* $n-1$*-dimensional hypersphere and* $\mathbb{R}P^{n-1}$ *indicates the real projective plane of* $n-1$ *dimensional hyperplanes in* $\mathbb{R}^n$*. The coordinates become* $([q,d],b,u)$ *with* $q\in$

S^{n-1}, $d \in (0, +\infty)$, $b \in \mathbb{R}^{n-1}$, $u \in \mathbb{R}$, and the metric simplifies to

$$h_{[q,d]}(q_\perp b, u) = d\|b\|^2 + ds^2.$$

The geodesic equations transforms to

$$\begin{cases} \dot{\mathbf{q}} = \mathbf{q}\mathbf{b}, & \mathbf{q}(0) = \mathbf{q}_0 \\ \dot{d} = 2ds, & d(0) = d_0, \\ \dot{b} = -2bs, & b(0) = b_0, \\ \dot{s} = \|b\|^2 - s^2, & s(0) = s_0, \end{cases}$$

where

$$\mathbf{q} = \begin{bmatrix} q & q_\perp \end{bmatrix} \in O(n),$$
$$\mathbf{b} = \begin{bmatrix} 0 & -b^T \\ b & 0 \end{bmatrix} \in \mathfrak{o}(n) = \mathrm{Skew}(n).$$

Defining $r = s/\|b\|$, a computation similar to Example 3.2.1 yields

$$\frac{d}{dr}\|b\| = \frac{-2r}{r^2+1},$$

and thus

$$\|b\|(r^2+1) = \tilde{C} := \|b_0\|(r_0^2+1),$$
$$b(r^2+1) = \mathbf{C} := b_0(r_0^2+1).$$

$r(t) = r_0 + \tilde{C}t$ is therefore linear and all other solution equations follow as above, namely,

$$b(t) = \frac{\mathbf{C}}{1+r(t)^2},$$
$$s(t) = \|b(t)\|r(t) = \frac{\tilde{C}r(t)}{1+r(t)},$$
$$d(t) = d_0\frac{1+r(t)^2}{1+r_0^2}.$$

To compute q, we first observe that, since $\dot{b} = -2sb$ and $s \in \mathbb{R}$ is a scalar, b has constant direction, or equivalently

$$\hat{b} = \frac{b}{\|b\|} \text{ is constant.}$$

Therefore, equation $\dot{\mathbf{q}} = \mathbf{q}\mathbf{b}$ can be written

$$\dot{\mathbf{q}}(t) = \mathbf{q}(t)\hat{\mathbf{b}}\|b(t)\| \tag{3.18}$$

with

$$\hat{\mathbf{b}} := \begin{bmatrix} 0 & -\hat{b}^T \\ \hat{b} & 0 \end{bmatrix}.$$

Equation (3.18) readily integrates to

$$\mathbf{q}(t) = \mathbf{q}_0 \exp(\hat{\mathbf{b}}\theta(t)), \tag{3.19}$$

with

$$\theta(t) := \int_0^t \|b(\tau)\| d\tau = \int_0^{r(t)} \frac{d\rho}{\rho^2 + 1} = \arctan r(t) - \arctan r_0.$$

Since $\hat{\mathbf{b}}^3 = -\hat{\mathbf{b}}$, rearranging gives

$$\begin{aligned} \exp(\hat{\mathbf{b}}\theta(t)) =& I + \sin\theta(t)\hat{\mathbf{b}} + (1 - \cos\theta(t))\hat{\mathbf{b}}^2 \\ =& \begin{bmatrix} \cos\theta(t) & -\hat{b}\sin\theta(t) \\ \hat{b}^T \sin\theta(t) & I_{n-1} - (\cos\theta(t) - 1)\hat{b}\hat{b}^T \end{bmatrix}, \end{aligned}$$

where the last equality comes from

$$\hat{\mathbf{b}}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -\hat{b}\hat{b}^T \end{bmatrix}.$$

With this, the first column of the right-hand side of (3.19) reads

$$q(t) = q_0 \cos\theta(t) + w_0 \sin\theta(t)$$

where $w_0 := q_\perp(0)\hat{b}$.

In more general cases, we can resort to numerical calculations to obtain geodesic solutions and plots.

Example 3.2.3. *Let* $n = 5$ *and* $k = 3$. *Let*

$$\mathbf{Q}_0 = I_n, \qquad B_0 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$
$$D_0 = I_3, \qquad T_0 = \frac{1}{2}I_3,$$

so that

$$S_0 = S_{D_0}(T_0) = \frac{1}{4}I_3.$$

Solving numerically the system we get the graphs shown in Figures 3.1-3.4, where the entries not plotted are constant and equal to their starting value.

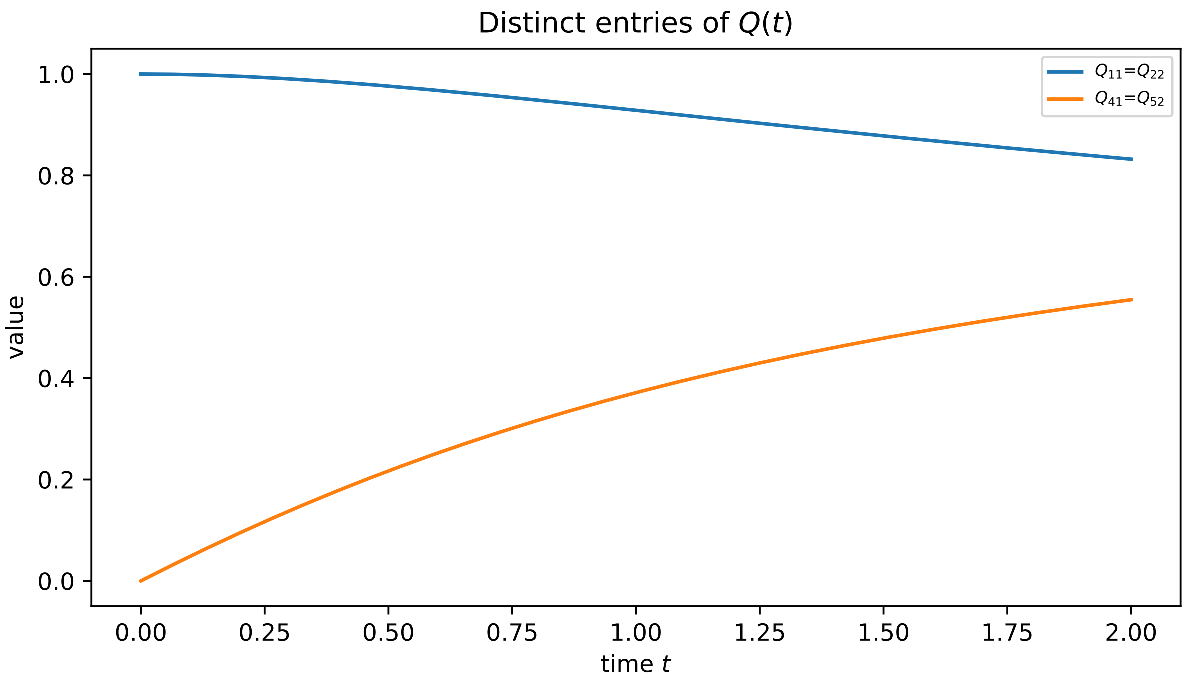

Figure 3.1: Time evolution of distinct entries of Q.

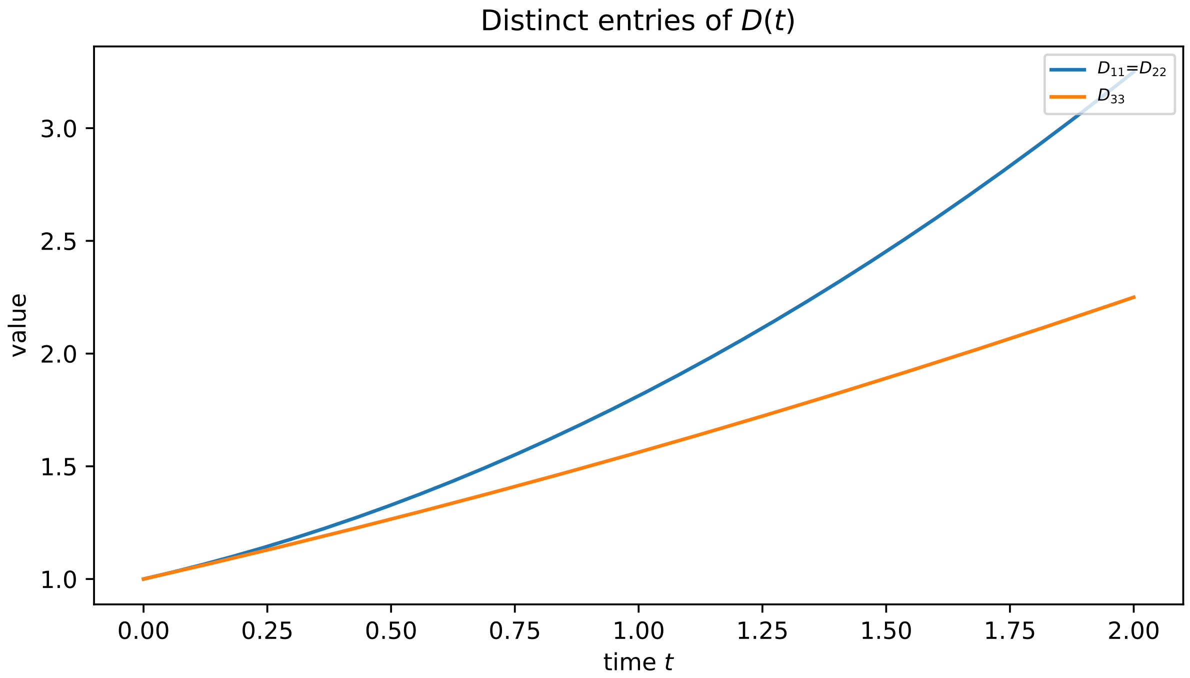

Figure 3.2: Time evolution of distinct entries of D.

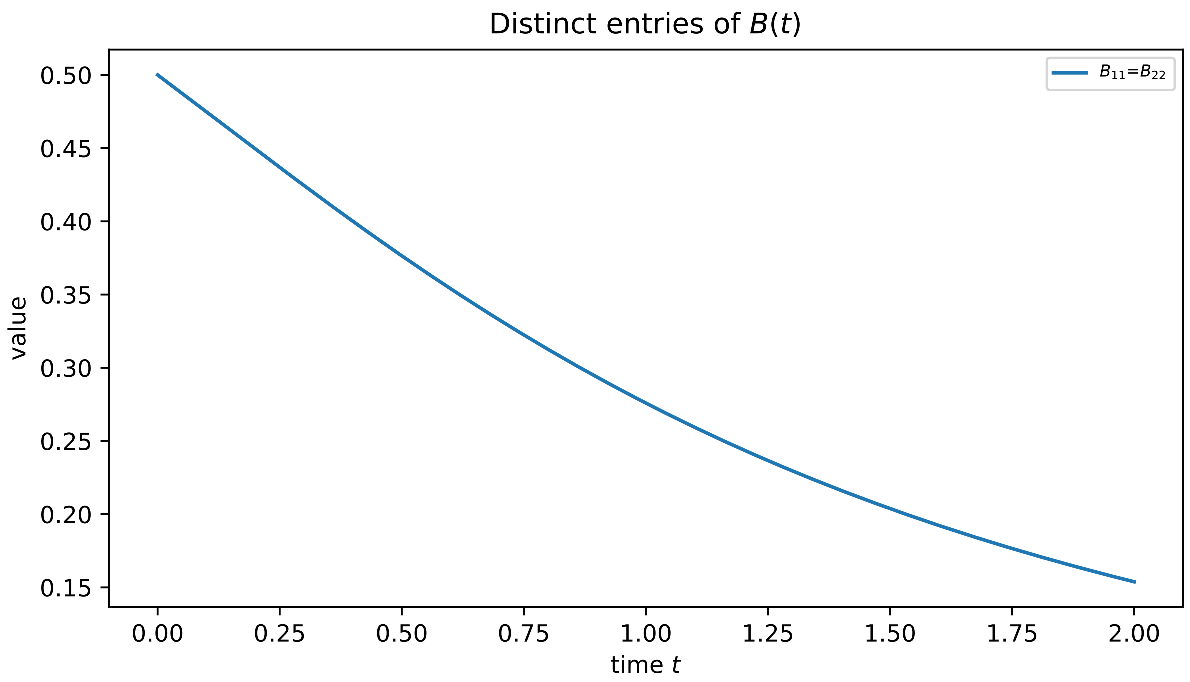

Figure 3.3: Time evolution of distinct entries of B.

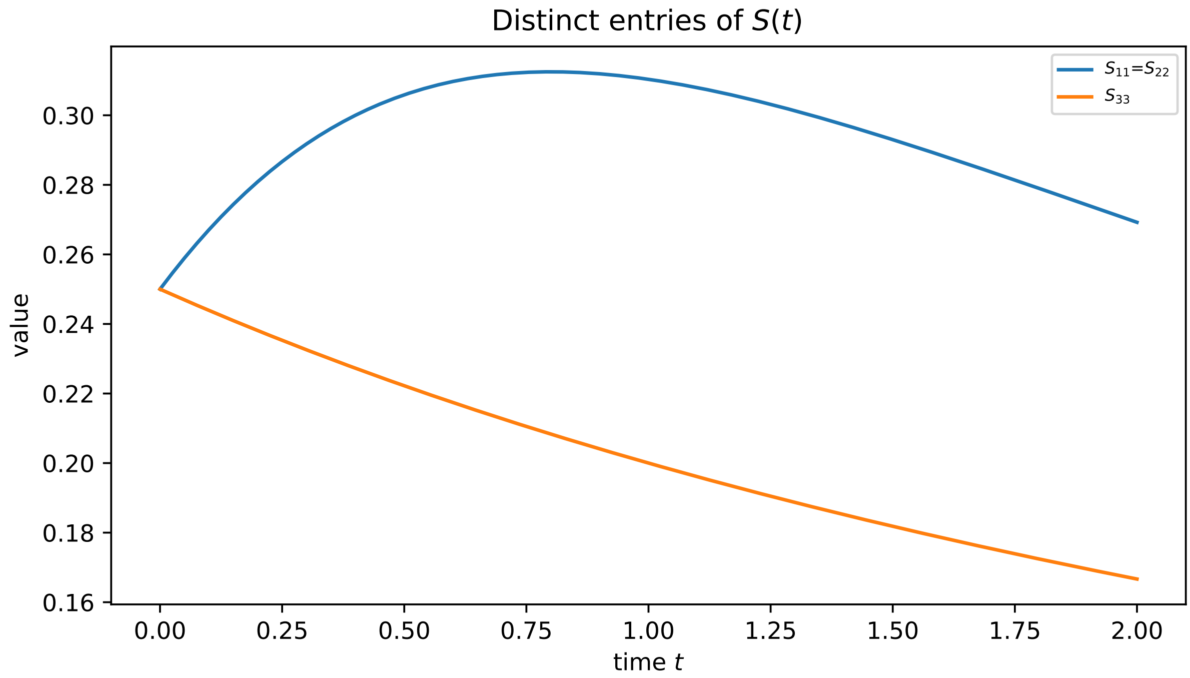

Figure 3.4: Time evolution of distinct entries of S.

With these simple initial conditions the system can be translated again to a family of systems of one dimensional equations and solved with techniques similar to the previous examples. Indeed, denoting $b_i = B_{ii}$, $d_i = D_{ii}$ *and* $s_i = S_{ii}$, *we get, for* $i = 1, 2$

$$\begin{cases} \dot{d}_i = 2 s_i d_i, & d_i(0) = 1, \\ \dot{b}_i = -2 s_i b_i, & b_i(0) = \frac{1}{2}, \\ \dot{s}_i = b_i^2 - s_i^2, & s_i(0) = \frac{1}{4}. \end{cases}$$

which is solved by

$$\begin{cases} d_i = \frac{4}{5}(1 + r^2), \\ b_i = \frac{5}{8(1+r^2)}, \\ s_i = \frac{5r}{8(1+r^2)}, \end{cases}$$

where

$$r(t) = \frac{1}{2} + \frac{5}{8} t.$$

For $i = 3$ we get the simplified system

$$\begin{cases} \dot{d}_3 = 2s_3 d_3, & d_3(0) = 1, \\ \dot{s}_3 = -s_3^2, & s_3(0) = \frac{1}{4}, \end{cases}$$

which is readily solved by

$$d_3(t) = (1 + \frac{1}{4}t)^2, \qquad s_3(t) = \frac{1/4}{1 + \frac{1}{4}t}.$$

One can show that, calling $\theta(t) = \arctan(r(t)) - \arctan(1/2)$ *and* $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, *we get*

$$\mathbf{Q}(t) = \mathrm{diag}(R(\theta(t)), R(\theta(t)), 1)$$

and

$$\begin{aligned} Q_{11}(t) &= Q_{22}(t) = \cos\theta(t), \\ Q_{41}(t) &= Q_{52}(t) = \sin\theta(t), \end{aligned}$$

while all other entries of $Q(t)$ stay constant.

The numerical machinery can be used to obtain results in cases with more complicated initial conditions. For example, let $\mathbf{Q}_0$ and D_0 be as before and

$$B_0 = \frac{1}{2}\begin{bmatrix} 1 & 1/2 & -2/5 \\ -3/5 & 1 & 0 \end{bmatrix}$$

$$T_0 = \frac{1}{2}\begin{bmatrix} 3/10 & -7/10 & 2/5 \\ 1 & -1/2 & 1/5 \\ -1 & 1 & 0 \end{bmatrix}$$

so that

$$S_0 = S_{D_0}(T_0) = \frac{1}{4}T_0.$$

In this case, closed form solutions are difficult to obtain. However, we can numerically integrate the system to obtain approximate solutions. The graphs are plotted in Figures 3.5-3.8.

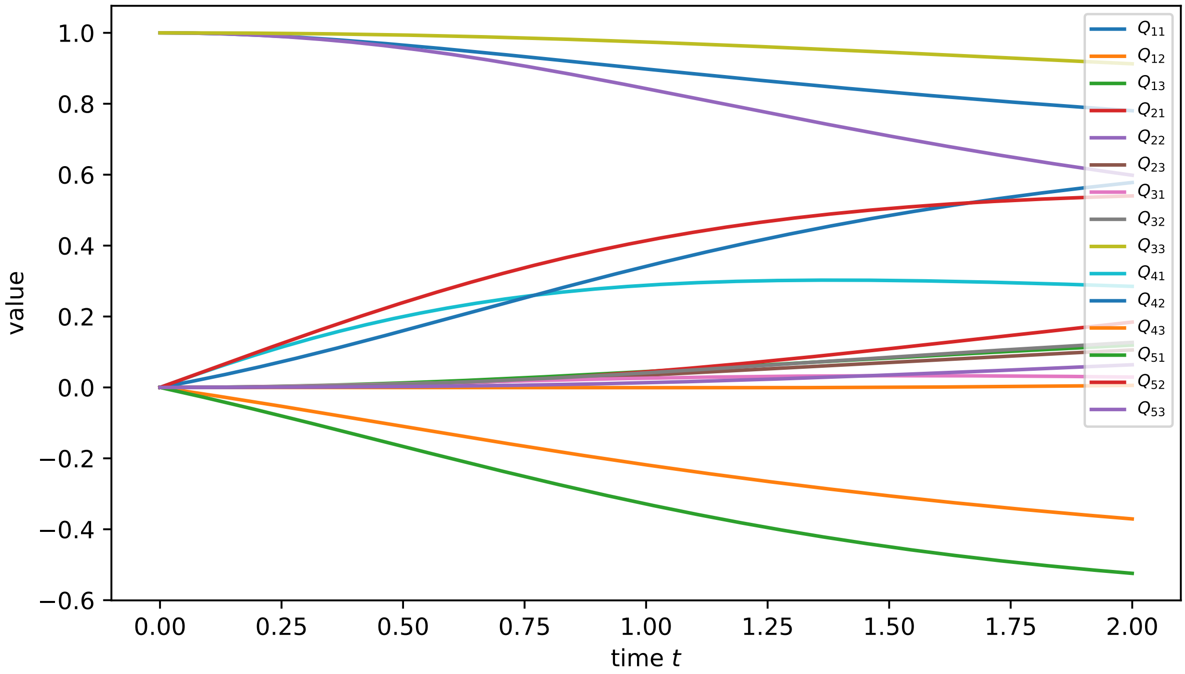

Figure 3.5: Time evolution of distinct entries of Q.

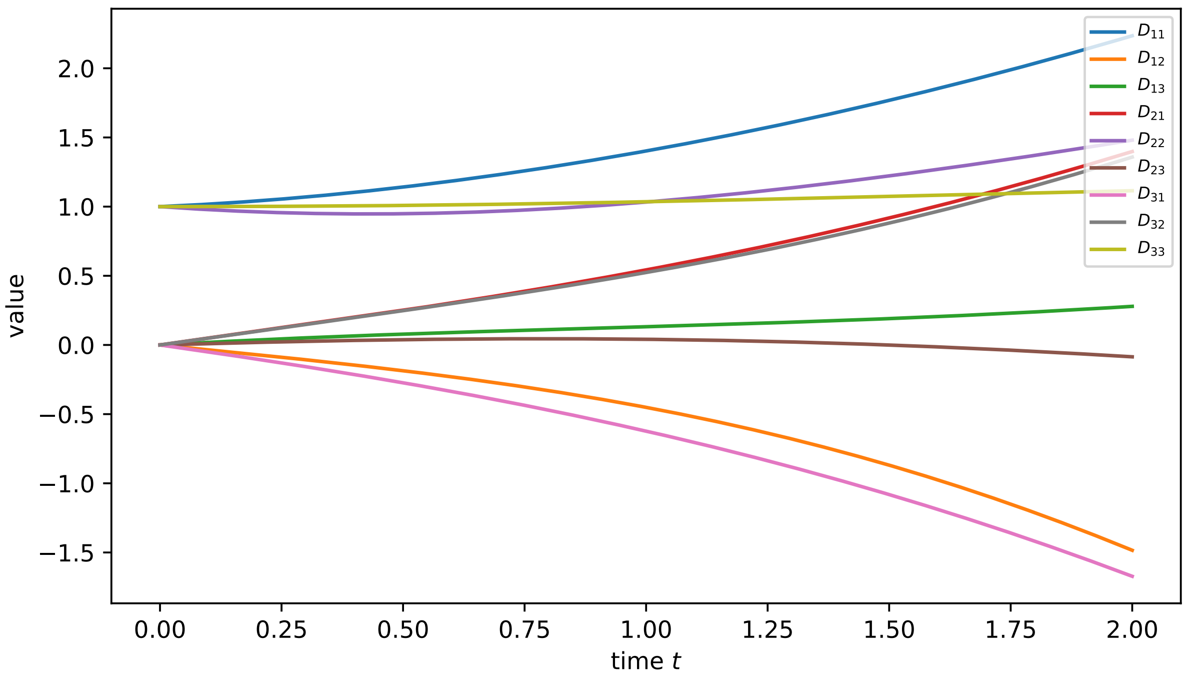

Figure 3.6: Time evolution of distinct entries of D.

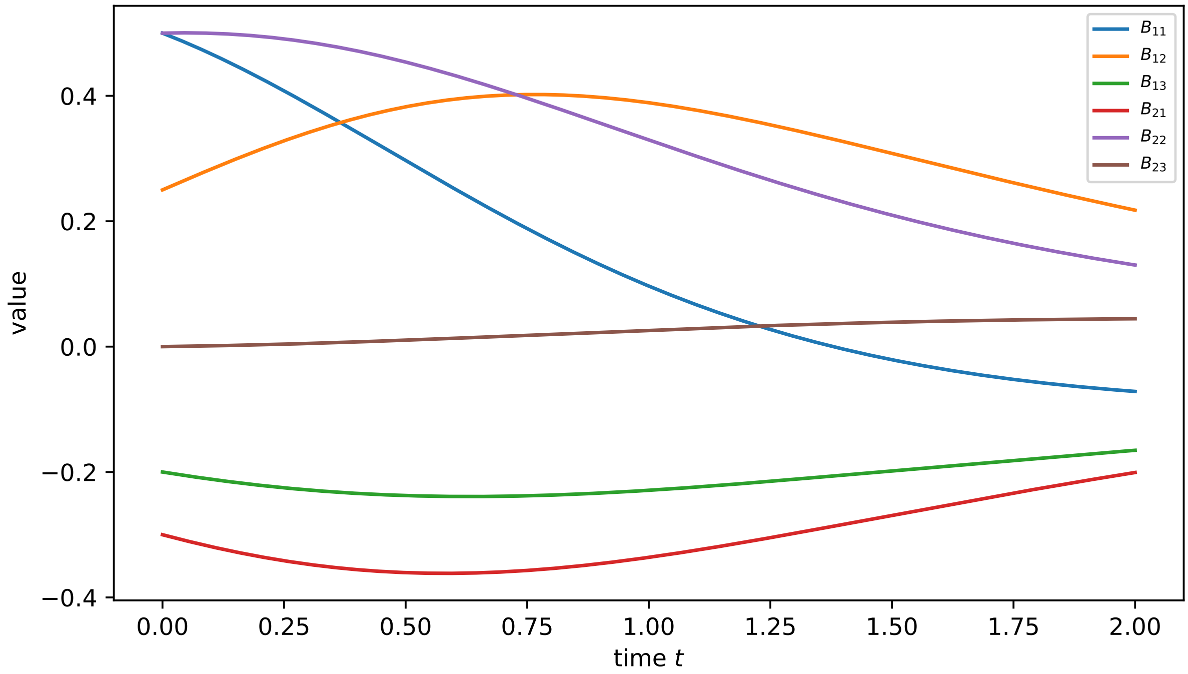

Figure 3.7: Time evolution of distinct entries of B.

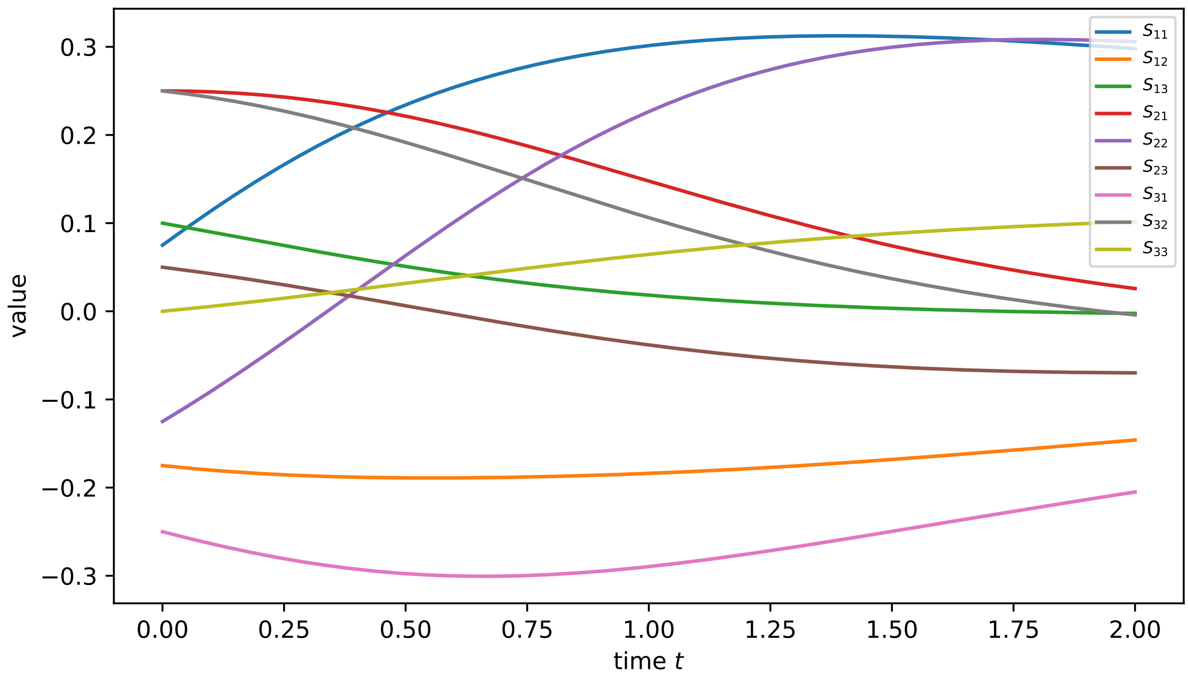

Figure 3.8: Time evolution of distinct entries of S.

3.2.4.3 Logarithm bijection

From [16, Section 5] and [106, Theorem 5.3], we notice that the number of logarithms, and hence of energy-minimizing geodesics, of $\mathrm{Gr}(k,n)$ endowed with the Frobenius metric and those of $\mathrm{Sym}^+(n,k)$ endowed with the Bures–Wasserstein metric coincides, except for the case of a vertical geodesic in $\mathrm{Sym}^+(n,k)$ that connects two points in the same fiber. Indeed, in [16], the authors link the number of minimizing geodesics between two points $P_1, P_2 \in \mathrm{Gr}(k,n)$ to the number r of orthogonal principal angles between P_1 and P_2. In particular the logarithms are indexed by set

$$I_1 = \{\mathrm{diag}(Q, I_{k-r}) \,|\, Q \in O(r)\} \cong O(r),$$

where I_{k-r} denotes the $k-r$ dimensional Identity matrix and $\mathrm{diag}(Q, I_{k-r})$ is intended as block diagonal (see [16, Theorem 5.5]). On the other hand, in [106], the authors link the number of minimizing geodesics between two matrices Σ_1, $\Sigma_2 \in \mathrm{Sym}^+(n,k)$ to the difference between k and $l := \mathrm{rk}(\Sigma_1\Sigma_2) = \mathrm{rk}(X^TY)$, where $XX^T = \Sigma_1$ and $YY^T = \Sigma_2$. In this case, the authors give the index set

$$\begin{aligned} I_2 =& \{R \in O(k) | X^TYR \in Cov(k)\} \\ =& \{R \in O(k) | X^TYR = (X^T\Sigma_2X)^{1/2}\}. \end{aligned}$$

Assume $P_1 = \mathrm{Span}(Q_1)$ and $P_2 = \mathrm{Span}(Q_2)$, and write

$$\Sigma_1 = Q_1D_1Q_1^T, \qquad \Sigma_2 = Q_2D_2Q_2^T,$$

with $D_1, D_2 \in \mathrm{Sym}^+(k)$ and $Q_1, Q_2 \in \mathrm{St}(n,k)$. The singular values of $Q_1^TQ_2$ are $\lambda_i(Q_1^TQ_2) = \cos\theta_i$, where θ_i are the principal angles between P_1 and P_2. Hence

$$\mathrm{rk}(\Sigma_1\Sigma_2) = \mathrm{rk}(Q_1^TQ_2) = \#\{i \,|\, \theta_i \neq \pi/2\}$$

and therefore

$$r := \#\{i \,|\, \theta_i = \pi/2\} = k - l.$$

We now make the correspondence between the index sets explicit.

Proposition 3.2.6. *Let $X, Y \in \mathrm{Mat}(n,k)^*$, and set $l = \mathrm{rk}(X^T Y)$ and $r = k - l$. Then there is a bijection between the index sets*

$$O(r) \quad \textit{and} \quad I_2 = \{R \in O(k) \,|\, X^T Y R = (X^T Y Y^T X)^{1/2}\}.$$

Proof. Since $\mathrm{rk}(X^T Y) = l$, we can perform the singular value decomposition

$$X^T Y = UDV^T,$$

with $U,\ V \in O(k)$,and $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_l, 0, \ldots, 0)$. In the same basis we have

$$(X^T Y Y^T X)^{1/2} = UDU^T.$$

This forces

$$R = V \begin{bmatrix} I_l & 0 \\ 0 & R_r \end{bmatrix} U^T, \quad R_r \in O(r).$$

□

Remark 3.2.4. This bijection can be readily transferred from logarithms to length-minimizing geodesics on $\mathrm{Sym}^+(n,k)$ and $\mathrm{Gr}(k,n)$: each choice of $R_r \in O(r)$ corresponds simultaneously to a minimizing Grassmannian geodesic (via [16]) and to a minimizing Bures–Wasserstein geodesic on the fixed-rank covariance manifold (via [106]). This suggests a geometric connection between $(\mathrm{Sym}^+(n,k), d_{BW})$ and $(\mathrm{Gr}(k,n), d_E)$. However, the horizontal part of the pullback metric on $M(n,k)$,

$$h^{\mathrm{hor}}_{[Q,D]}(V, W) := \mathrm{Tr}\,(B_V D B_W^\top),$$

clearly differs from the Euclidean metric on $\mathrm{Gr}(k,n)$, as it carries vertical positional information through the anisotropic weight matrix D.

This raises several questions and avenues for further investigation. How should one interpret this geodesic correspondence at the level of the metric structure? To what extent do Euclidean-geometric features of $\mathrm{Gr}(k,n)$ (such as distances, angles, curvature, cut times and locuses) encode geometric or statistical properties of $(\mathrm{Sym}^+(n,k), d_{BW})$? More speculatively, one may ask whether the Bures–Wasserstein geometry of the fixed-rank stratum can be viewed as an anisotropically weighted perturbation of the Euclidean geometry of $\mathrm{Gr}(k,n)$, coupled with the Bures–Wasserstein metric itself on the full rank fibers. We leave these as open directions for future work.

Appendix

A.1 Functional analysis

1. A *multi-index* is an n-tuple $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$. Set

$$|\alpha| := \alpha_1 + \cdots + \alpha_n, \qquad x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \qquad \partial^\alpha := \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}.$$

Multi-index notation streamlines differential expressions and symbol calculus.

2. Let $\mathcal{S}(\mathbb{R}^n)$ denote the Schwartz space of rapidly decreasing test functions. For $\phi \in \mathcal{S}(\mathbb{R}^n)$, we define the Fourier transform $\mathcal{F}\phi$ by

$$\mathcal{F}\phi(\xi) := \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\phi(x)dx$$

and

$$\mathcal{F}^{-1}\phi(x) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\phi(\xi)d\xi.$$

When it makes the notation simpler, Fourier transforms may also be denoted as $\hat{\phi} := \mathcal{F}\phi$.

3. Let $U \in \mathbb{R}^n$ be open. We denote with $\mathcal{D}(U)$ the space of functions $\phi \in C_0^\infty(U)$ endowed with the topology described by the following notion of convergence. Given a sequence $(\phi_n)_{n\in\mathbb{N}} \subset \mathcal{D}(U)$ and a function $\phi \in \mathcal{D}(U)$ we say that ϕ_n converges to ϕ in $\mathcal{D}(U)$ and write $\phi_n \to \phi$ if:

- there exists a compact set $K \subset U$ such that $\operatorname{supp}(\phi_n - \phi) \subset K$ for all $n \geq 1$,
- $\lim_{n\to\infty} D^\alpha\phi_n = D^\alpha\phi$ uniformly in K for every multi-index α.

4. Let $U \subset \mathbb{R}^n$ be open. A *distribution* on U is a continuous linear functional

$$T : \mathcal{D}(U) \to \mathbb{R},$$

where continuity is understood with respect to the topology of $\mathcal{D}(U)$. The space of all distributions on U is denoted $\mathcal{D}'(U)$.

Every locally integrable function $f \in L^1_{\mathrm{loc}}(U)$ defines a distribution via

$$T_f(\phi) = \int_U f(x)\phi(x)\,dx, \qquad \phi \in \mathcal{D}(U).$$

Derivatives of distributions are defined by duality:

$$\langle \partial^\alpha T, \phi\rangle = (-1)^{|\alpha|}\langle T, \partial^\alpha \phi\rangle, \qquad \alpha \in \mathbb{N}_0^n.$$

This framework extends differentiation and integration to non-smooth objects, and it provides the natural setting for defining generalized solutions to differential and stochastic equations.

5. A *tempered measure* on $\mathbb{R}^n$ is a Radon measure μ such that

$$\int_{\mathbb{R}^n} (1 + |\xi|^2)^{-N}\, d|\mu|(\xi) < \infty$$

for some integer $N \geq 0$. The space of tempered measures is denoted by $\mathcal{M}_{\mathrm{temp}}(\mathbb{R}^n)$.

Every tempered measure defines a *tempered distribution* by duality: for $\phi \in \mathcal{S}(\mathbb{R}^n)$, we set

$$\langle \mu, \phi\rangle := \int_{\mathbb{R}^n} \phi(\xi)\, d\mu(\xi).$$

Conversely, a distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ that is positive and continuous on bounded subsets of $\mathcal{S}(\mathbb{R}^n)$ is represented by a unique tempered measure μ_T.

Tempered measures arise naturally as *spectral measures* in the study of stationary Gaussian random fields. If $\mu \in \mathcal{M}_{\mathrm{temp}}(\mathbb{R}^n)$, the covariance functional

$$\Gamma(\phi, \psi) = \int_{\mathbb{R}^n} \mathcal{F}\phi(\xi)\,\overline{\mathcal{F}\psi(\xi)}\,\mu(d\xi)$$

defines a continuous positive semidefinite bilinear form on $\mathcal{S}(\mathbb{R}^n)$. This construction allows one to represent homogeneous Gaussian fields and random distributions with finite spectral energy in terms of their associated tempered measures.

6. Let $(X, \mathcal{A}, \mu)$ be a measure space. A family of random variables $\{F(\phi) : \phi \in \mathcal{D}(\mathbb{R}^n)\}$, or more generally a mapping $F : \mathcal{D}(\mathbb{R}^n) \to L^2(\Omega)$, is called L^2*-continuous* if

$$\lim_{\phi_k \to \phi} \mathbb{E}[|F(\phi_k) - F(\phi)|^2] = 0$$

whenever $\phi_k \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$. More generally, if I is an index space, a mapping $f : I \to L^2(\Omega)$, is said L^2-continuous if

$$\lim_{t_k \to t} \mathbb{E}[|f(s) - f(t)|^2] = 0,$$

for all sequences $\{t_k\} \subset I$ that converge to t in I.

7. The *Dirac delta distribution* δ is the distribution on $\mathbb{R}^n$ defined by

$$\langle \delta, \phi \rangle = \phi(0), \qquad \phi \in \mathcal{D}(\mathbb{R}^n).$$

For any $x_0 \in \mathbb{R}^n$, the translated delta distribution is given by

$$\langle \delta_{x_0}, \phi \rangle = \phi(x_0),$$

and satisfies $\delta_{x_0}(x) = \delta(x - x_0)$ in the sense of distributions.

The delta distribution has the *scaling property*

$$\delta(ax) = |a|^{-(n)} \delta(x), \qquad a \in \mathbb{R} \setminus \{0\},$$

and its derivatives are defined by

$$\langle \partial^\alpha \delta, \phi \rangle = (-1)^{|\alpha|} \partial^\alpha \phi(0), \qquad \alpha \in \mathbb{N}_0^n.$$

In the Fourier domain, the delta distribution is characterized by

$$\mathcal{F}\delta = 1, \qquad \mathcal{F}^{-1}(1) = \delta,$$

and thus serves as the fundamental solution of the identity operator.

8. Let $(H, \|\cdot\|)$ be a normed Hilbert space (e.g. $(L^2(\mathbb{R}), \|\cdot\|_{L^2})$). Let $\{e_i\}_{i \in I}$ denote an orthonormal basis of H for an index set I. The *Parseval identity* states that for every $f \in H$

$$\|f\|^2 = \sum_{i \in I} |\langle f, e_i \rangle|^2.$$

A.2 Partial differential equations

9. Let P be a linear partial differential operator with smooth coefficients acting on $\mathcal{D}'(\mathbb{R}^n)$. A *fundamental solution* of P is a distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ satisfying

$$PE = \delta,$$

where δ denotes the Dirac delta distribution centred at the origin. In other words, E is a right inverse of P in the sense of distributions.

For a given $f \in \mathcal{D}(\mathbb{R}^n)$, the function

$$u = E * f$$

is a distributional solution to $Pu = f$, where $*$ denotes convolution.

The existence and regularity of fundamental solutions depend on the analytic and geometric properties of P. For elliptic operators, E is smooth away from the origin, while for hyperbolic operators, E exhibits singularities propagating along the bicharacteristics determined by the principal symbol. Fundamental solutions thus play a central role in the representation of solutions to deterministic and stochastic partial differential equations.

A.3 Matrix algebra

10. Let $\mathrm{Sym}(n)$ denote the real vector space of symmetric $n \times n$ matrices. A matrix $\Sigma \in \mathrm{Sym}(n)$ is *positive semidefinite* (PSD) if $x^\top \Sigma x \geq 0$ for all $x \in \mathbb{R}^n$, and *positive definite* (PD) if $x^\top \Sigma x > 0$ for all $x \neq 0$. The notation $\Sigma \succeq 0$ (resp. $\Sigma \succ 0$) indicates PSD (resp. PD). The set of covariance matrices is

$$\mathrm{Cov}(n) := \{\Sigma \in \mathrm{Sym}(n) \ : \ \Sigma \succeq 0\}.$$

By the spectral theorem, any $\Sigma \in \mathrm{Sym}(n)$ admits an orthogonal diagonalization

$$\Sigma = QDQ^\top,$$

with $Q \in O(n)$ and $D = \mathrm{diag}(d_1, \dots, d_n)$, where $d_i \in \mathbb{R}$. Moreover, $\Sigma \in \mathrm{Cov}(n)$ if and only if $d_i \geq 0$ for all i, and $\mathrm{rank}(\Sigma)$ equals the number of strictly positive d_i.

11. The set $\mathrm{Cov}(n)$ is a closed convex cone in $\mathrm{Sym}(n)$:

- if $\lambda \geq 0$ and $\Sigma \in \mathrm{Cov}(n)$, then $\lambda\Sigma \in \mathrm{Cov}(n)$;
- if $\Sigma_m \to \Sigma$ in $\mathrm{Sym}(n)$, then $x^\top \Sigma x = \lim_m x^\top \Sigma_m x \geq 0$ for all x, hence $\Sigma \in \mathrm{Cov}(n)$;
- if $\alpha, \beta \geq 0$ and $\Sigma_1, \Sigma_2 \in \mathrm{Cov}(n)$, then $\alpha\Sigma_1 + \beta\Sigma_2 \in \mathrm{Cov}(n)$.

12. The Frobenius (Euclidean) inner product on $\mathbb{R}^{n\times n}$ is

$$\langle A, B\rangle_F := \mathrm{Tr}(AB^\top),$$

with induced norm $\|A\|_F := \sqrt{\langle A, A\rangle_F}$.

13. Let $A \in \mathrm{Sym}(n)$ be positive definite. For every $B \in \mathrm{Sym}(n)$ there exists a unique symmetric S satisfying

$$AS + SA = B.$$

The operator $S_A : \mathrm{Sym}(n) \to \mathrm{Sym}(n)$, $S_A(B) = S$, is linear and self-adjoint with respect to the Frobenius inner product. It admits the integral representation $S_A(B) = \int_0^{+\infty} e^{-tA} B e^{-tA}\, dt$, and the Kronecker form $\mathrm{vec}(S) = (I \otimes A + A \otimes I)^{-1}\mathrm{vec}(B)$.

A.4 Smooth manifolds

14. A *diffeomorphism* is a bijective smooth map whose inverse is smooth.

15. Let $d \leq n$. A subset $M \subset \mathbb{R}^n$ is a d-dimensional smooth (embedded) submanifold if, for every $p \in M$, there exist an open neighbourhood $U \subset M$, an open set $O \subset \mathbb{R}^d$, and a diffeomorphism (called *coordinate chart*) $\phi : U \to O$. Intuitively, M is locally modeled as $\mathbb{R}^d$.

For systematic references about smooth manifolds, Lie groups, and Riemannian geometry, see [50].

16. Let $M \subset \mathbb{R}^n$ be a d-dimensional embedded submanifold and let $p \in M$. A vector $v \in \mathbb{R}^n$ is *tangent* to M at p if there exists a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to M$ with

$$\gamma(0) = p, \qquad \dot{\gamma}(0) = v.$$

The set of such vectors is the tangent space T_pM, which is a d-dimensional vector subspace of $\mathbb{R}^n$.

17. Let M be a smooth manifold of dimension d. For each point $p \in M$, the tangent space T_pM is the d-dimensional vector space of tangent vectors at p. The *tangent bundle* of M is the disjoint union of all tangent spaces,

$$TM = \bigsqcup_{p \in M} T_pM,$$

together with the natural projection $\pi : TM \to M$ given by $\pi(v_p) = p$. A smooth structure can be defined on TM in such a way that π is a smooth surjective submersion and, in local coordinates $(x^1, \ldots, x^d)$ on M, each tangent vector is represented as

$$v_p = \sum_{i=1}^{d} v^i \frac{\partial}{\partial x^i}\Big|_p,$$

so that $(x^1, \ldots, x^d, v^1, \ldots, v^d)$ form local coordinates on TM. Sections of π, that is smooth maps $X : M \to TM$ such that $\pi \circ X = \mathrm{id}_M$, are called *vector fields* on M and form a module over $C^\infty(M)$, denoted $\Gamma(TM)$.

The *cotangent bundle* T^*M is defined as the dual bundle of TM, that is,

$$T^*M = \bigsqcup_{p \in M} T_p^*M,$$

where T_p^*M is the dual space of T_pM. A point of T^*M is denoted by (p, α_p), where $\alpha_p : T_pM \to \mathbb{R}$ is a linear functional. Local coordinates $(x^1, \ldots, x^d, \xi_1, \ldots, \xi_d)$ on T^*M are induced by the canonical dual basis $\{dx^i\}$, so that $\alpha_p = \sum_i \xi_i \, dx^i|_p$. Sections of the cotangent bundle are called *differential 1-forms*, and the space of all smooth 1-forms is denoted $\Gamma(T^*M)$.

The cotangent bundle carries a canonical symplectic structure, given in local coordinates by the 2-form

$$\omega = \sum_{i=1}^{d} d\xi_i \wedge dx^i,$$

which plays a fundamental role in analytical mechanics, microlocal analysis, and the geometric formulation of Hamiltonian systems.

18. Let $M, N \subset \mathbb{R}^n$ be embedded submanifolds and $f : M \to N$ a smooth map. The *differential* of f at $p \in M$ is the linear map $df(p) : T_pM \to T_{f(p)}N$ given by

$$df(p)\, v = \frac{d}{dt} f(\gamma(t))\Big|_{t=0},$$

for any smooth γ with $\gamma(0) = p$ and $\dot{\gamma}(0) = v$. The map f is a *submersion* at p if $df(p)$ is surjective, and an *immersion* at p if $df(p)$ is injective.

19. Let X be a topological space and $\sim$ an equivalence relation on X. The *quotient space* $X/\sim$ is the set of equivalence classes endowed with the quotient topology. If a group G acts on X (on the right, $x \mapsto x \cdot g$), then $[x] = \{x \cdot g : g \in G\}$ and the quotient is denoted X/G. Constructions used in this thesis include $\mathrm{Cov}(n) \cong \mathrm{Mat}(n)/O(n)$ and $\mathrm{Sym}^+(n, k) \cong \mathrm{Mat}(n, k)^*/O(k)$ (see [79]).

20. Let G be a Lie group acting smoothly on a manifold M. The action is *free* if $g \cdot p = p$ for some p implies $g = e$ (the identity), and *proper* if for every compact $K \subset M$ the set $\{g \in G : K \cap gK \neq \emptyset\}$ is compact. These hypotheses ensure good topological and geometric properties of the quotient.

21. A *Lie group* is a group that is also a smooth manifold such that the group law and inversion are smooth. Its Lie algebra $\mathfrak{g}$ is the tangent space T_eG at the identity, endowed with the Lie bracket induced by left-invariant vector fields. Standard references include [40].

A.5 Riemannian structures

22. A *Riemannian metric* on a smooth manifold M is a smoothly varying inner product g_p on T_pM for each $p \in M$.

23. Let $\phi : M \to N$ be smooth, and let g be a Riemannian metric on N. The *pullback* metric $\phi^* g$ on M is defined by

$$(\phi^* g)_p(u, v) := g_{\phi(p)}(d\phi(p)\, u,\ d\phi(p)\, v).$$

24. There exists a unique connection ∇ on (M, g), the *Levi–Civita connection*, characterized by the following properties, for $X, Y, Z \in \Gamma(TM)$ and $f \in C^\infty(M)$:

- linearity in the first argument: $\nabla_{fX} Y = f\, \nabla_X Y$;
- Leibniz rule in the second argument: $\nabla_X(fY) = X[f]\; Y + f\, \nabla_X Y$;
- torsion-free: $\nabla_X Y - \nabla_Y X = [X, Y]$;
- metric compatibility: $X[g(Y, Z)] = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$.

For embedded manifolds, ∇ agrees with the orthogonal projection of the ambient derivative onto the tangent bundle.

25. Given a smooth curve γ in M and $v \in T_{\gamma(0)}M$, the *parallel transport* of v along γ is the unique vector field X along γ such that

$$\nabla_{\dot\gamma(t)} X = 0, \qquad X(0) = v.$$

26. A smooth curve γ is a *geodesic* if

$$\nabla_{\dot\gamma} \dot\gamma = 0.$$

Equivalently, among curves with fixed endpoints, geodesics are critical points of the energy functional

$$E(\gamma) = \int_a^b g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))\, dt.$$

A minimizing geodesic between $p, q \in M$ is a (piecewise) smooth geodesic segment realizing $d(p, q)$ and has constant speed.

27. For each $p \in M$ there exists $\varepsilon > 0$ such that, for $v \in T_pM$ with $\|v\| < \varepsilon$, the geodesic with initial data (p, v) is defined on $[0, 1]$. The *exponential map* is

$$\exp_p(v) := \gamma_v(1).$$

If $q \in M$ is joined to p by a unique minimizing geodesic with initial velocity v, then v is a *logarithm* of q from p and satisfies $\|v\| = d(p, q)$. Where $\exp_p$ is a diffeomorphism, the *logarithmic map* $\log_p$ is defined as its local inverse.

28. The *injectivity radius* at $p \in M$ is the supremum of $r > 0$ such that $\exp_p$ restricts to a diffeomorphism from the open ball $B_r(0) \subset T_pM$ onto its image.

29. The *cut locus* of p is the set of points where minimizing geodesics starting at p cease to be minimizing. Equivalently, it is the image under $\exp_p$ of the boundary of the largest star-shaped domain in T_pM on which $\exp_p$ is a diffeomorphism. Along a unit vector $v \in T_pM$, the *cut time* is

$$\sup\{t > 0 \,:\, d(p, \exp_p(tv)) = t\}.$$

30. The *Riemann curvature tensor* is

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

At each $p \in M$, $R(X, Y)Z$ depends only on the values of X, Y, Z at p. If $\dim M \geq 2$, the *sectional curvature* of the plane $\Pi = \operatorname{span}\{u, v\} \subset T_pM$ is

$$K(u, v) := \frac{\langle R(u, v)v, u\rangle}{\langle u, u\rangle \langle v, v\rangle - \langle u, v\rangle^2}.$$

31. A submanifold $N \subset M$ is *totally geodesic* if every geodesic of M with initial data (p, v), where $p \in N$ and $v \in T_pN$, remains in N for as long as it is defined.

A.6 Variational calculus

32. Let

$$F(\phi) = \int_{t_1}^{t_2} L\big(t, \phi(t), \dot{\phi}(t)\big)\, dt$$

be a functional, where L is the *Lagrangian*. A smooth variation is given by $\phi_\epsilon = \phi + \epsilon\psi$, with $\epsilon \in \mathbb{R}$ and ψ smooth. A necessary condition for ϕ to be an extremal is

$$\frac{\partial}{\partial \epsilon} L\big(t, \phi_\epsilon(t), \dot{\phi}_\epsilon(t)\big)\Big|_{\epsilon=0} = 0,$$

which yields the Euler-Lagrange equations

$$\frac{\partial}{\partial \phi} L(t, \phi(t), \dot{\phi}(t)) - \frac{\partial}{\partial t}\left(\frac{\partial}{\partial \dot{\phi}} L(t, \phi(t), \dot{\phi}(t))\right) = 0.$$

Bibliography

[1] A. Abdeljawad, A. Ascanelli, and S. Coriasco, "Deterministic and stochastic Cauchy problems for a class of weakly hyperbolic operators on $\mathbb{R}^n$," *Monatsh. Math.*, vol. 192, no. 1, pp. 1–38, 2020, ISSN: 0026-9255,1436-5081. DOI: 10.1007/s00605-020-01372-0

[2] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

[3] B. Afsari, "Riemannian L^p center of mass: Existence, uniqueness, and convexity," *Proceedings of the American Mathematical Society*, vol. 139, no. 2, pp. 655–673, 2011.

[4] D. Alekseevsky, A. Kriegl, M. Losik, and P. W. Michor, "The riemannian geometry of orbit spaces. the metric, geodesics, and integrable systems," *Publicationes Mathematicae Debrecen*, vol. 62, pp. 1–30, 2001.

[5] S.-i. Amari, *Information geometry and its applications*. Springer, 2016, vol. 194.

[6] S.-i. Amari and H. Nagaoka, *Methods of information geometry*. American Mathematical Soc., 2000, vol. 191.

[7] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 411–421, 2006. DOI: 10.1002/mrm.20965

[8] A. Ascanelli, S. Coriasco, and A. Süß, "Solution theory to semilinear hyperbolic stochastic partial differential equations with polynomially bounded coefficients," *Nonlinear Anal.*, vol. 189, pp. 111574, 34, 2019, ISSN: 0362-546X,1873-5215. DOI: 10.1016/j.na.2019.111574

[9] A. Ascanelli, S. Coriasco, and A. Süß, "Solution theory to semilinear parabolic stochastic partial differential equations with polynomially bounded coefficients," *J. Pseudo-Differ. Oper. Appl.*, vol. 16, no. 1, Paper No. 20, 2025, ISSN: 1662-9981,1662-999X. DOI: 10.1007/s11868-024-00665-4

[10] A. Ascanelli, S. Coriasco, and A. Süss, "Random-field solutions of weakly hyperbolic stochastic partial differential equations with polynomially bounded coefficients," *J. Pseudo-Differ. Oper. Appl.*, vol. 11, no. 1, pp. 387–424, 2020, ISSN: 1662-9981,1662-999X. DOI: 10.1007/s11868-019-00290-6

[11] A. Ascanelli, S. Coriasco, and A. Süss, "Solution theory to semilinear stochastic equations of Schrödinger type on curved spaces I: Operators with uniformly bounded coefficients," *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM*, vol. 118, no. 2, Paper No. 60, 15, 2024, ISSN: 1578-7303,1579-1505. DOI: 10.1007/s13398-024-01554-7

[12] A. Ascanelli and A. Süß, "Random-field solutions to linear hyperbolic stochastic partial differential equations with variable coefficients," *Stochastic Process. Appl.*, vol. 128, no. 8, pp. 2605–2641, 2018, ISSN: 0304-4149,1879-209X. DOI: 10.1016/j.spa.2017.09.019

[13] J. Bai, "Inferential theory for factor models of large dimensions," *Econometrica*, vol. 71, no. 1, pp. 135–171, 2003. DOI: 10.1111/1468-0262.00470

[14] C. Bakker, M. Halappanavar, and A. Visweswara Sathanur, "Dynamic graphs, community detection, and riemannian geometry," *Applied network science*, vol. 3, pp. 1–30, 2018. DOI: 10.1007/s41109-018-0059-2

[15] F. Barbaresco, "Innovative tools for radar signal processing based on cartan's geometry of spd matrices & information geometry," in *2008 IEEE Radar Conference*, IEEE, 2008, pp. 1–6. DOI: 10.1109/RADAR.2008.4720937

[16] T. Bendokat, R. Zimmermann, and P.-A. Absil, "A grassmann manifold handbook: Basic geometry and computational aspects," *Advances in Computational Mathematics*, vol. 50, no. 1, p. 6, 2024. DOI: 10.1007/s10444-023-10090-8

[17] E. Bernardi and M. Farnè, “A log-det heuristics for covariance matrix estimation: The analytic setup,” *Stats*, vol. 5, no. 3, pp. 606–616, 2022. DOI: 10.3390/stats5030037

[18] E. Bernardi and A. Lanconelli, “On a class of stochastic hyperbolic equations with double characteristics,” *J. Fourier Anal. Appl.*, vol. 29, no. 1, Paper No. 2, 16, 2023, ISSN: 1069-5869,1531-5851. DOI: 10.1007/s00041-022-09987-7

[19] E. Bernardi and L. Marconi, “On some stochastic hyperbolic equations with symplectic characteristics,” *Journal of Fourier Analysis and Applications*, vol. 31, no. 3, pp. 1–22, 2025. DOI: 10.1007/s00041-025-10159-6

[20] R. Bhatia, “Positive definite matrices,” in *Positive Definite Matrices*, Princeton university press, 2009.

[21] R. Bhatia, T. Jain, and Y. Lim, “On the bures–wasserstein distance between positive definite matrices,” *Expositiones Mathematicae*, vol. 37, no. 2, pp. 165–191, 2019. DOI: 10.1016/j.exmath.2018.01.002

[22] R. Bhattacharya and L. Lin, “Omnibus clts for fréchet means and nonparametric inference on non-euclidean spaces,” *Proceedings of the American Mathematical Society*, vol. 145, no. 1, pp. 413–428, 2017.

[23] R. Bhattacharya and V. Patrangenaru, “Large sample theory of intrinsic and extrinsic sample means on manifolds—ii,” 2005.

[24] L. J. Billera, S. P. Holmes, and K. Vogtmann, “Geometry of the space of phylogenetic trees,” *Advances in Applied Mathematics*, vol. 27, no. 4, pp. 733–767, 2001. DOI: 10.1006/aama.2001.0759

[25] S. S. Bonan and D. S. Clark, “Estimates of the Hermite and the Freud polynomials,” *J. Approx. Theory*, vol. 63, no. 2, pp. 210–224, 1990, ISSN: 0021-9045,1096-0430. DOI: 10.1016/0021-9045(90)90104-X

[26] S. Bonnabel, G. Meyer, and R. Sepulchre, “Adaptive filtering for estimation of a low-rank positive semidefinite matrix,” in *19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010)*, 2010.

[27] S. Bonnabel and R. Sepulchre, "Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1055–1070, 2010.

[28] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," *Physical Review Letters*, vol. 72, no. 22, p. 3439, 1994. DOI: 10.1103/PhysRevLett.72.3439

[29] A. Bucci, M. Palma, and C. Zhang, "Geometric deep learning for realized covariance matrix forecasting," 2024. arXiv: 2412.09517.

[30] D. Bures, "An extension of kakutani's theorem on infinite product measures to the tensor product of semifinite w*-algebras," *Transactions of the American Mathematical Society*, vol. 135, pp. 199–212, 1969. DOI: 10.1090/S0002-9947-1969-0236719-2

[31] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, "Latent variable graphical model selection via convex optimization," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2010, pp. 1610–1613.

[32] L. Chen and R. C. Dalang, "Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions," 2015.

[33] G. S. Chirikjian, *Stochastic models, information theory, and Lie groups, volume 2: Analytic methods and modern applications*. Springer Science & Business Media, 2011, vol. 2.

[34] S. Coriasco, S. Pilipović, and D. Seleši, "Solutions of hyperbolic stochastic PDEs on bounded and unbounded domains," *J. Fourier Anal. Appl.*, vol. 27, no. 5, Paper No. 77, 42, 2021, ISSN: 1069-5869,1531-5851. DOI: 10.1007/s00041-021-09858-7

[35] S. Coriasco, S. Pilipović, and D. Seleši, "Chaos expansion solutions of a class of magnetic schrödinger wick-type stochastic equations on $\mathbb{R}^d$," *Journal of Mathematical Analysis and Applications*, vol. 550, no. 2, p. 129 620, 2025. DOI: 10.1016/j.jmaa.2025.129620

[36] R. Courant and D. Hilbert, *Methoden der mathematischen Physik*. Springer-Verlag, 2013.

[37] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*. Cambridge university press, 2014, vol. 152.

[38] R. C. Dalang, "Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s," *Electron. J. Probab.*, vol. 4, no. 6, 29, 1999, ISSN: 1083-6489. DOI: 10.1214/EJP.v4-43

[39] R. C. Dalang and N. E. Frangos, "The stochastic wave equation in two spatial dimensions," *Ann. Probab.*, vol. 26, no. 1, pp. 187–212, 1998, ISSN: 0091-1798,2168-894X. DOI: 10.1214/aop/1022855416

[40] M. P. Do Carmo, *Riemannian geometry*. Springer, 1992, vol. 2.

[41] D. C. Dowson and B. V. Landau, "The fréchet distance between multivariate normal distributions," *Journal of multivariate analysis*, vol. 12, no. 3, pp. 450–455, 1982. DOI: 10.1016/0047-259X(82)90077-X

[42] I. L. Dryden, A. Koloydenko, and D. Zhou, "Non-euclidean statistics for covariance matrices, with applications to diffusion tensor imaging," *The Annals of Applied Statistics*, pp. 1102–1123, 2009. DOI: 10.1214/09-AOAS249

[43] P. Fillard, X. Pennec, V. Arsigny, and N. Ayache, "Clinical dt-mri estimation, smoothing, and fiber tracking with log-euclidean metrics," *IEEE transactions on medical imaging*, vol. 26, no. 11, pp. 1472–1482, 2007. DOI: 10.1109/TMI.2007.899173

[44] P. T. Fletcher and S. Joshi, "Riemannian geometry for the statistical analysis of diffusion tensor data," *Signal Processing*, vol. 87, no. 2, pp. 250–262, 2007. DOI: 10.1016/j.sigpro.2005.12.018

[45] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE transactions on medical imaging*, vol. 23, no. 8, pp. 995–1005, 2004. DOI: 10.1109/TMI.2004.831793

[46] M. Fréchet, "Les éléments aléatoires de nature quelconque dans un espace distancié," in *Annales de l'institut Henri Poincaré*, vol. 10, 1948, pp. 215–310.

[47] J. Gallier and J. Quaintance, "The log-euclidean framework applied to spd matrices," in *Differential Geometry and Lie Groups: A Computational Perspective*, Springer, 2020, pp. 665–673.

[48] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*. Springer, 1977, vol. 224.

[49] P.-Y. Gousenbourger et al., “Piecewise-bézier c1 smoothing on manifolds with application to wind field estimation.,” in *Esann*, 2017.

[50] N. Guigui, N. Miolane, X. Pennec, et al., “Introduction to riemannian geometry and geometric statistics: From basic theory to implementation with geomstats,” *Foundations and Trends® in Machine Learning*, vol. 16, no. 3, pp. 329–493, 2023. DOI: 10.1561/2200000098

[51] C. W. Helstrom, “Minimum mean-squared error of estimates in quantum statistics,” *Physics letters A*, vol. 25, no. 2, pp. 101–102, 1967. DOI: 10.1016/0375-9601(67)90366-0

[52] L. Hörmander, “The Cauchy problem for differential equations with double characteristics,” *J. Analyse Math.*, vol. 32, pp. 118–196, 1977, ISSN: 0021-7670,1565-8538. DOI: 10.1007/BF02803578

[53] L. Hörmander, “Quadratic hyperbolic operators,” in *Microlocal Analysis and Applications*, L. Cattabriga and L. Rodino, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, pp. 118–160.

[54] T. Hotz and S. Huckemann, “Intrinsic means on the circle: Uniqueness, locus and asymptotics,” *Annals of the Institute of Statistical Mathematics*, vol. 67, no. 1, pp. 177–193, 2015.

[55] S. Huckemann and H. Ziezold, “Principal component analysis for riemannian manifolds, with an application to triangular shape spaces,” *Advances in Applied Probability*, vol. 38, no. 2, pp. 299–319, 2006.

[56] S. F. Huckemann, “Intrinsic inference on the mean geodesic of planar shapes and tree discrimination by leaf growth,” 2011.

[57] D. Husemöller, *Fibre bundles*. Springer, 1966, vol. 5.

[58] V. J. Ivriĭ and V. M. Petkov, “Necessary conditions for the correctness of the Cauchy problem for non-strictly hyperbolic equations,” *Uspehi Mat. Nauk*, vol. 29, no. 5(179), pp. 3–70, 1974, ISSN: 0042-1316.

[59] A. Kacem, M. Daoudi, B. B. Amor, S. Berretti, and J. C. Alvarez-Paiva, "A novel geometric framework on gram matrix trajectories for human behavior understanding," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 1, pp. 1–14, 2018. DOI: 10.1109/TPAMI.2018.2872564

[60] H. Karcher, *Riemannian comparison constructions*. SFB 256, 1987.

[61] D. G. Kendall, "Shape manifolds, procrustean metrics, and complex projective spaces," *Bulletin of the London mathematical society*, vol. 16, no. 2, pp. 81–121, 1984. DOI: 10.1112/blms/16.2.81

[62] D. Khoshnevisan, *Analysis of stochastic partial differential equations*. American Mathematical Soc., 2014, vol. 119.

[63] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, Volume I* (Interscience Tracts in Pure and Applied Mathematics 15). New York: Interscience Publishers, 1963, Reprinted by Wiley Classics Library, 1996.

[64] B. Lascar, "Propagation des singularités pour des équations hyperboliques à caractéristique de multiplicité au plus double et singularités masloviennes," *Amer. J. Math.*, vol. 104, no. 2, pp. 227–285, 1982, ISSN: 0002-9327,1080-6377. DOI: 10.2307/2374160

[65] N. N. Lebedev, *Special functions and their applications*, Revised, R. A. Silverman, Ed. Dover Publications, Inc., New York, 1972, pp. xii+308, Unabridged and corrected republication.

[66] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of multivariate analysis*, vol. 88, no. 2, pp. 365–411, 2004. DOI: 10.1016/S0047-259X(03)00096-4

[67] O. Ledoit and M. Wolf, "Shrinkage estimation of large covariance matrices: Keep it simple, statistician?" *Journal of Multivariate Analysis*, vol. 186, p. 104 796, 2021. DOI: 10.1016/j.jmva.2021.104796

[68] J. M. Lee, *Introduction to topological manifolds*. Springer, 2000.

[69] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, "Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor mri processing," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 3, pp. 423–444, 2006. DOI: 10.1007/s10851-006-6897-z

[70] P. Li, Q. Wang, H. Zeng, and L. Zhang, "Local log-euclidean multivariate gaussian descriptor and its application to image classification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 4, pp. 803–817, 2016. DOI: 10.1109/TPAMI.2016.2560816

[71] X.-B. Li and F. J. Burkowski, "Conformational transitions and principal geodesic analysis on the positive semidefinite matrix manifold," in *Bioinformatics Research and Applications: 10th International Symposium, ISBRA 2014, Zhangjiajie, China, June 28-30, 2014. Proceedings 10*, Springer, 2014, pp. 334–345.

[72] Z. Lin, "Riemannian geometry of symmetric positive definite matrices via cholesky decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 40, no. 4, pp. 1353–1370, 2019. DOI: 10.1137/18M1221084

[73] J. Lott and C. Villani, "Ricci curvature for metric-measure spaces via optimal transport," *Annals of Mathematics*, pp. 903–991, 2009. DOI: 10.4007/annals.2009.169.903

[74] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons, 2019.

[75] L. Malagò, L. Montrucchio, and G. Pistone, "Wasserstein riemannian geometry of gaussian densities," *Information Geometry*, vol. 1, pp. 137–179, 2018. DOI: 10.1007/s41884-018-0014-4

[76] M. Marchand, W. Huang, A. Browet, P. Van Dooren, and K. A. Gallivan, "A riemannian optimization approach for role model extraction," in *Proceedings of the 22nd International Symposium on Mathematical Theory of Networks and Systems*, 2016, pp. 58–64.

[77] L. Marconi, "An associated bundle approach to the bures–wasserstein geometry of fixed rank covariance matrices," 2025. arXiv: 2507.05873.

[78] K. V. Mardia and P. E. Jupp, *Directional statistics*. John Wiley & Sons, 2009.

[79] E. 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*, vol. 41, no. 1, pp. 171–198, 2020. DOI: 10.1137/18M1231389

[80] E. Massart, J. 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*, Springer, 2019, pp. 739–748. DOI: 10.1007/978-3-030-26980-7_77

[81] E. M. Massart, P.-Y. Gousenbourger, N. T. Son, T. Stykel, and P.-A. Absil, “Interpolation on the manifold of fixed-rank positive-semidefinite matrices for parametric model order reduction: Preliminary results.,” in *ESANN*, vol. 2019, 2019, pp. 281–286.

[82] R. B. Melrose, “The wave equation for a hypoelliptic operator with symplectic characteristics of codimension two,” *Journal d’Analyse Mathématique*, vol. 44, no. 1, pp. 134–182, Dec. 1984, ISSN: 1565-8538. DOI: 10.1007/BF02790194

[83] G. Meyer, S. Bonnabel, and R. Sepulchre, “Regression on fixed-rank positive semidefinite matrices: A riemannian approach,” *The Journal of Machine Learning Research*, vol. 12, pp. 593–625, 2011.

[84] K. R. Meyer and D. C. Offin, *Introduction to hamiltonian dynamical systems and the n-body problem*. Springer, 2017.

[85] B. Mishra, G. Meyer, and R. Sepulchre, “Low-rank optimization for distance matrix completion,” in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, IEEE, 2011, pp. 4455–4460.

[86] E. H. Moore, “On the reciprocal of the general algebraic matrix (abstract),” *Bulletin of the American Mathematical Society*, vol. 26, pp. 394–395, 1920.

[87] T. Nishitani, “Local energy integrals for effectively hyperbolic operators i,” *Journal of Mathematics of Kyoto University*, vol. 24, no. 4, pp. 623–658, 1984.

[88] T. Nishitani, *Cauchy Problem for Differential Operators with Double Characteristics: Non-effectively Hyperbolic Characteristics*. Springer, 2017, vol. 2202.

[89] B. O’Neill, “The fundamental equations of a submersion.,” *Michigan Mathematical Journal*, vol. 13, no. 4, pp. 459–469, 1966.

[90] I. Olkin and F. Pukelsheim, "The distance between two random vectors with given dispersion matrices," *Linear Algebra and its Applications*, vol. 48, pp. 257–263, 1982. DOI: 10.1016/0024-3795(82)90112-4

[91] X. Pennec, "Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 1, pp. 127–154, 2006. DOI: 10.1007/s10851-006-6228-4

[92] X. Pennec, P. Fillard, and N. Ayache, "A riemannian framework for tensor computing," *International Journal of computer vision*, vol. 66, no. 1, pp. 41–66, 2006. DOI: 10.1007/s11263-005-3222-z

[93] X. Pennec, S. Sommer, and T. Fletcher, *Riemannian geometric statistics in medical image analysis*. Academic Press, 2019.

[94] R. Penrose, "A generalized inverse for matrices," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3, pp. 406–413, 1955. DOI: 10.1017/S0305004100030401

[95] S. Peszat and J. Zabczyk, *Stochastic partial differential equations with Lévy noise: An evolution equation approach*. Cambridge University Press, 2007, vol. 113.

[96] G. Peyré, M. Cuturi, et al., "Computational optimal transport: With applications to data science," *Foundations and Trends® in Machine Learning*, vol. 11, no. 5-6, pp. 355–607, 2019. DOI: 10.1561/2200000073

[97] C. Prévôt and M. Röckner, *A concise course on stochastic partial differential equations*. Springer, 2007.

[98] L. Schwartz, *Théorie des distributions* (Publications de l'Institut de Mathématique de l'Université de Strasbourg). Hermann, Paris, 1966, vol. IX-X, pp. xiii+420, Nouvelle édition, entiérement corrigée, refondue et augmentée.

[99] R. W. Sharpe, *Differential geometry: Cartan's generalization of Klein's Erlangen program*. Springer Science & Business Media, 2000, vol. 166.

[100] C. L. Siegel, *Symplectic geometry*. Elsevier, 2014.

[101] L. T. Skovgaard, "A riemannian geometry of the multivariate normal model," *Scandinavian journal of statistics*, pp. 211–223, 1984.

[102] K.-T. Sturm, *Probability measures on metric spaces of nonpositive curvature*. SFB 611, 2003.

[103] G. Szegő, *Orthogonal polynomials* (American Mathematical Society Colloquium Publications), Fourth. American Mathematical Society, Providence, RI, 1975, vol. Vol. XXIII, pp. xiii+432.

[104] A. Takatsu, "On wasserstein geometry of gaussian measures," *Probabilistic approach to geometry*, vol. 57, pp. 463–472, 2010. DOI: 10.2969/aspm/05710463

[105] A. Takatsu, "Wasserstein geometry of gaussian measures," 2011. DOI: 10.18910/4973

[106] Y. Thanwerdas and X. Pennec, "Bures–wasserstein minimizing geodesics between covariance matrices of different ranks," *SIAM Journal on Matrix Analysis and Applications*, vol. 44, no. 3, pp. 1447–1476, 2023. DOI: 10.1137/22M149168X

[107] Y. Thanwerdas and X. Pennec, "O (n)-invariant riemannian metrics on spd matrices," *Linear Algebra and its Applications*, vol. 661, pp. 163–201, 2023. DOI: 10.1016/j.laa.2022.12.009

[108] F. Trèves, *Introduction to pseudodifferential and Fourier integral operators Volume 2: Fourier integral operators*. Springer Science & Business Media, 1980, vol. 2.

[109] C. Villani et al., *Optimal transport: old and new*. Springer, 2008, vol. 338.

[110] J. B. Walsh, "An introduction to stochastic partial differential equations," in *École d'été de probabilités de Saint-Flour, XIV—1984*, ser. Lecture Notes in Math. Vol. 1180, Springer, Berlin, 1986, pp. 265–439, ISBN: 3-540-16441-3. DOI: 10.1007/BFb0074920

[111] D. I. Warton, "Penalized normal likelihood and ridge regularization of correlation and covariance matrices," *Journal of the American Statistical Association*, vol. 103, no. 481, pp. 340–349, 2008. DOI: 10.1198/016214508000000021