LIMBURGS
UNIVERSITAIR
CENTRUM

RIEMANNIAN SUBMERSIONS AND SKEW-PRODUCT DECOMPOSITIONS OF BROWNIAN MOTION

Eric J. Pauwels

LIMBURGS UNIVERSITAIR CENTRUM

RIEMANNIAN SUBMERSIONS AND SKEW-PRODUCT DECOMPOSITIONS OF BROWNIAN MOTION

proefschrift
tot het bekomen van de graad van
doctor in de wetenschappen
door

ERIC J. PAUWELS

Promotor

Prof. dr. L.C.G. Rogers

November 1989

Acknowledgments

It's a pleasure to express my gratitude to the many, many people who made this work possible.

First of all I would like to thank my supervisor prof. Chris Rogers without whose constant help and encouragement this thesis would never have been written. On innumerable occasions when negotiating a difficult corner, I found his expertise and insights invaluable and he patiently helped me groping my way towards an understanding of the matters at hand.

I am also very grateful to prof. David Williams for creating a stimulating and hospitable environment at the Cambridge Statistical Laboratory. His interest in my work meant a lot to me.

Next I want to express my appreciation for the senior members of the Statistical Research Unit at L.U.C.: profs. Herman Callaert, Paul Embrechts, Paul Janssen and Noel Veraverbeke who allowed me plenty of leeway so that I could venture into unknown territory.

I would also like to thank prof P. Verstraelen (K.U.Leuven) for some very helpful conversations about geometry and Theo Moons who never cringed away when I dropped by to discuss another algebraic problem.

A lot of credit is due to Gudrun Gorissen whose efficient and skillful handling of the wordprocessor saved me hours of painstaking proofreading.

To my friends and colleagues at both the Cambridge Statistical Laboratory and Limburgs Universitair Centrum I feel very much indebted for creating a pleasant and relaxed atmosphere in which it was good to dwell. In particular I'd like to mention Els Goetghebeur, Irene Gijbels and Marc Aerts. Their friendship and caring made all the difference.

Finally, I would like to express my gratitude to my parents and family and dedicate this work to them. Their love and support have always been of vital importance to my well-being and they never doubted that—eventually—I would produce something worthwile. To them I can only say: I'll keep on trying!

Table of Contents

Chapter 1: Introduction

1.1. Introduction and summary
1.2. Some important definitions and results
1.2.1. Differential geometry
1.2.2. Stochastic analysis

Chapter 2: Riemannian Submersions of Brownian Motion

2.1. The Submersion Theorem
2.2. Brownian motion on Liegroups
2.3. Canonical and bi-invariant metrics on real Liegroups
Appendix: Some definitions in Lie-algebra theory

Chapter 3: Applications of the Submersion Theorem

3.1. Actions of a Liegroup on a manifold
3.2. Bessel-processes on Euclidean space
3.3. Bessel-processes on the hyperbolic space
3.4. The action of the Euclidean group on Euclidean space
3.5. The action of the Lorentz-group on the Euclidean plane
3.6. The action of $SL(2, \mathbb{R})$ on the hyperbolic plane
3.7. Triangular shape-diffusion on $\mathbb{R}^n$

Chapter 4: Skew-product Decompositions of Brownian Motion

4.1. Introduction
4.2. The Decomposition Theorem

4.3. Some applications of the Decomposition Theorem
4.3.1. The hyperbolic plane
4.3.2. Surfaces of revolution
4.3.3. Symmetric Matrices
4.3.4. Normal Matrices

References

Samenvatting

Chapter 1

Introduction

1.1 Introduction

Ever since in 1827 the English botanist Robert Brown observed the apparantly chaotic motion of small pollen grains suspended in water, diffusion has been a constant source of inspiration to physicists and mathematicians alike.
However, it wasn't until 1905 that Einstein was able to explain this so called Brownian motion as the macroscopic manifestation of random vibrations at a molecular level. Einstein's contribution was purely phenomological: he expressed the mean quadratic displacement of a Brownian particle as a function of the elapsed time.
A sound mathematical foundation was provided by Wiener who in 1923 introduced the notion of an integral with respect to integrators of unbounded variation. This idea was taken up by K. Itô who developed the theory of stochastic integrals and stochastic differential equations (SDE). He worked out a calculus which now bears his name and which is based on the famous Itô-formula. The simplest version of this equation can be formulated as follows:

Let $x(t)$ be a real-valued continuous semimartingale on a filtered probability-space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ which supports a Brownian motion β_t.
If x_t satisfies the SDE:

$$dx_t = \sigma(x_t)d\beta_t + b(x_t)dt \tag{1.1}$$

and if

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

is C^2-differentiable,
then $f(x_t)$ is a continuous semimartingale which satisfies the SDE:

$$df(x_t) = f'(x_t)\sigma(x_t)d\beta_t + (Gf)(x_t)dt \tag{1.2}$$

where

$$G = \frac{1}{2}\sigma^2(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}$$

is the generator of x_t.

This result has been generalised in many ways and Itô's formula has been extended to include multi-dimensional general semimartingales, functions f that are convex but not necessarily C^2, processes living on abstract state-spaces, etc .... We refer to Rogers & Williams [22] and the references therein for a comprehensive discussion of these important matters.

Relatively recently people have become interested in the interplay between stochastic processes and the geometry of the state-space on which they live. Natural problems that suggest themselves concern the existence and construction of Brownian motion on Riemannian manifolds, first-hitting probabilities for small geodesical spheres,transience and recurrence of diffusions, etc .... Most of the main results that emerged from this research in stochastic differential geometry can be found in Rogers & Williams [22], Ikeda & Watanabe [8] and Pinsky [20].

In this thesis we look at a specific example of Itô's transformation formula in stochastic geometry. We investigate how a Brownian motion on a Riemannian manifold is transformed when we submerge this manifold into another.
We can reformulate the problem in more precise terms as follows:
Let (M, g) and (B, h) be two (finite-dimensional) Riemannian manifolds and

$$\pi : M \longrightarrow B$$

a C^∞-submersion. (For a short review of some important concepts in differential geometry, we refer to §2).

Submersion Problem:

Denoting Brownian motion on M by β_t we want to determine under what conditions $\pi(\beta_t)$ will be an autonomous diffusionprocess on B.

There are two different situations where this problem arises naturally.
Firstly, there are diffusion-problems which, although formulated in a high-dimensional space, are essentially low-dimensional. A typical example is the calculation of the first-hitting probability for (small) geodesical spheres.
The dimension of the manifold on which the diffusion lives is not essential: we are only interested in the radial distance of the diffusing particle with respect to the centre of the sphere. So we (at least locally) project the manifold on a one-dimensional space and solve the problem there.
There is a second area of application which is slightly more involved. Some problems become more tractable if we expand the statespace. Loosely speaking one can say that adding dimensions provides extra degrees of freedom which make the problem easier to handle. The idea then is: reformulate the question in a higher-dimensional space, solve it there and then project the solution to the original space. We come across a prime example of this technique when we study Brownian motion on a homogeneous space $M \cong G/H$ with respect to a Liegroup G. The manifold M is a clumsy object for which in general there isn't a nice parametrisation. The Liegroup G on the other hand has a lot of nice features which make it much more manageable. So in order to construct Brownian motion on a homogeneous space we define it on the corresponding Liegroup and then project it down via the canonical submersion:

$$\pi : G \longrightarrow G/H \cong M$$

For another but closely related example of this solutionstrategy we look at Brownian motion on the orthonormal framebundle $O(M)$ of a Riemannian manifold M. Instead of working on M itself we transfer the problem to $O(M)$. The point in doing this is that $O(M)$ carries global canonical vectorfields which can be used to construct Brownian motion globally. Once we have done this the Brownian motion on M can be retrieved by applying the submersion:

$$\pi : O(M) \longrightarrow M$$

This method avoids the cumbersome patching together of locally defined diffusions. For a more detailed account of this technique we refer to Rogers & Williams [22], and Ikeda & Watanabe[8].

Let us now retrace our steps to give an outline of the results in this thesis.
In Chapter 2 we return to the submersionproblem and determine necessary and

sufficient conditions for $\pi(\beta_t)$ to be an autonomous diffusion. It turns out that two elements are important : first of all π has to be a Riemannian submersion. Secondly and most importantly, the fibres of π have to be of constant mean curvature in the sense that for each fibre the projection under $d\pi$ of the mean curvature vectorfield must be the same for every point in that fibre. If this is the case then the projected curvature field determines a vector field on B and $\pi(\beta_t)$ is seen to be a Brownian motion subject to this additional drift. It is possible to give a very natural and intuitively appealing explanation of this result if we interpet the Brownian motion β_t as a (weak) limit of a geodesical random walk. In order to do this we need to give a new interpretation of the mean curvature vector in terms of the deviation of geodesics on a submainfold with respect to the corresponding geodesics in the ambient space.

Next we turn our attention to Liegroups. The natural way to define Brownian motion on a Liegroup is by looking at Brownian motion on the corresponding Liealgebra and using the exponential map to carry the process to the group.

This however is only possible in some very special cases.

In general we have to add an additional drift to the process on the algebra to compensate for the curvature of the one-parameters subgroups. This again can be easily explained using the random walk approximation.

Finally we consider the problem of canonical an bi-invariant Brownian motions on Liegroups : does there exist a natural choice of metric on a Liegroup, i.e. a choice completely determined by the algebraic properties of the group and is this result invariant under left- and right-translations ? The Killing-metric provides an answer to the first question. As for the bi-invariance, this requires more care and is linked to the relative compactness of the adjoint Liegroup. This result implies that most Liegroups don't support a bi-invariant Brownian motion. The rotationgroups $SO(n)$ are rather exceptional in that they do support a bi-invariant metric (and hence Brownian motion). Moreover, we get a rather unexpected result as a bonus : these bi-invariant metrics are unique except when $n = 4$, for then we have a two-parameter family of such metrics. So, as far as bi-invariant metrics are concerned, $SO(4)$ is the odd one out among the rotationgroups.

In Chapter 3 we concentrate on some applications of the submersion theorem. Besselprocesses are considered on the Euclidean and hyperbolic space. The interpretation of the driftterm in terms of curvature of the fibres yields a simple explanation for the transience of Brownian motion on the hyperbolic plane.

Next we look at the action of some Liegroups on manifolds. Most of these do not produce any surprises but the action of $SL(2, \mathbb{R})$ on the hyperbolic plane turns out to be unexpectedly complicated.

Finally we take a quick glance at shape-diffusion as introduced by D.G. Kendall [10]. The submersion theorem allows us to give a short proof of some basic results in this area.

In Chapter 4 we take another look at the submersion theorem and study how we can retrieve some of the information we lost when we submerged the high-dimensional manifold into a lower-dimensional one. More precisely we consider product-manifolds of the form

$$M = X \times Y$$

together with the canonical projections π_1 and π_2 on the components. We assume that π_1 satisfies the conditions specified in the submersion theorem. We then prove the decomposition theorem which states that if the above product is orthogonal with respect to the metric on M, then the projection $\pi_2(\beta)$ is an nonautonomous Brownian motion on Y.
Such a decomposition is called a skew-product decomposition, in analogy with the well-known description of $BM(\mathbb{R}^2)$ in terms of polar coordinates.
This result is then applied to two interesting matrix-classes : the symmetric and the normal matrices. Elements of these manifolds can be interpreted as suitable products of diagonal- and unitary matrices. The decomposition then tells us that the eigenvalues perform autonomous diffusions whereas the eigenvectors are driven by independent white noise the clock of which depends on the distance between the eigenvalues. It is quite surprising to discover that the eigenvalues behave as diffusing electrically-charged particles on the real line (in the case of symmetric matrices) or in the complex plane (for the normal matrices) that repel each other with a force that is inversely proportional to their separation.
We can try and balance this repulsion by adding an eleastic restoring force which is directed towards the origin, thus obtaining an Ornstein-Uhlenbeck-process on these matrixmanifolds. It then becomes possible to compute the corresponding stationary probability measure.

1.2 Some important definitions and results

In this paragraph we recall some important definitions and results which we will use throughout this thesis.

1.2.1 Differential geometry

Let M be a topological space which is Hausdorff. An **open chart** on M is a pair (U, ϕ) where U is an open subset of M and ϕ is a homeomorphism of U onto an open subset of $\mathbb{R}^n$.

A **differentiable structure** on M of dimension n is a collection of open coordinate-charts $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ on M where $\phi_\alpha(U_\alpha)$ is an open subset of $\mathbb{R}^n$ such that the following conditions are satisfied:

1. $M = \bigcup\limits_{\alpha \in A} U_\alpha$;
2. For each pair $\alpha, \beta \in A$ the mapping $\phi_\beta \circ \phi_\alpha^{-1}$ is a differentiable mapping of $\phi_\alpha(U_\alpha \bigcap U_\beta)$ onto $\phi_\beta(U_\alpha \bigcap U_\beta)$;
3. The collection $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ is a maximal family of open charts for which (1) and (2) hold.

A **differentiable manifold** (or C^∞- manifold or simply manifold) of dimension n is a Hausdorff space with an differentiable structure of dimension n.
A **Riemannian structure** on M is a tensor field g of type (0,2) which satisfies:

1. $g(X, Y) = g(Y, X)$ for all tangentvectors X, Y;
2. for each $p \in M, g_p$ is a positive definite bilinear form on $T_pM \times T_pM$.

A **Riemannian manifold** is a connected C^∞-manifold with a Riemannian structure.

A differentiable curve $\gamma : I \longrightarrow M$ is called a **geodesic** if the family of tangent vectors $\dot{\gamma}(t)$ is parallel with respect to γ; i.e. if ∇ is the specified affine connection on M then

$$\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0$$

Let M and B be two manifolds and $\pi : M \longrightarrow B$ a C^∞-surjection. The map π is called a **submersion** if at each $x \in M$ the derivative $d\pi(x)$ is a surjection. It then follows from the inverse function theorem that for each $b \in B$ the corresponding

fibre $\pi^{-1}\{b\}$ is a submanifold of M.

A Riemannian manifold M is said to be **complete** if every Cauchy sequence in M is convergent.

From now on we will always assume, unless otherwise indicated, that all manifolds are complete. We have the following important theorem (cfr. Helgason [7], p. 56):

Theorem 1

Let M be a Riemannian manifold. The following conditions are equivalent:

1. M is complete;
2. Each bounded closed subset of M is compact;
3. Each maximal geodesic in M has the form $\gamma_X(t)(-\infty < t < \infty)$, where γ_X is the unique geodesic emanating from $\gamma_X(0)$ with velocityvector X. Loosely speaking this condition says that each maximal geodesic has infinite length.

In a complete Riemannian manifold M with metric d, each pair $p, q \in M$ can be joined by a geodesic of length $d(p,q)$.

□

Let (M, g) be a n-dimensional Riemannian manifold equipped with local coordinate-charts $(x^1, \ldots, x^n)$. The Laplace-Beltrami operator Δ associated with this metric can be defined in terms of the local coordinates as (using the summation convention):

$$\Delta(f) = \frac{1}{\sqrt{detg}} \frac{\partial}{\partial x^i}(g^{ik} \sqrt{detg} \frac{\partial f}{\partial x^k})$$

where

$$\begin{aligned} g_{ik} &= g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^k}) \\ g^{ik} &= (g^{-1})_{ik} \\ detg &= det(g_{ik}) \end{aligned}$$

More information on Riemannian geometry can be found in Helgason [7].

1.2.2 Stochastic analysis

Throughout this thesis we will work with the basic probabilistic set-up $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ where $(\Omega, \mathcal{F}, P)$ is a **probability-space** and $\{\mathcal{F}_t\}_{t\geq 0}$ is a **filtration** satisfying the 'usual conditions' (cfr. Williams [27], II. 40).

A **stochastic process** $X : I\!R^+ \times \Omega \longrightarrow S : (t, \omega) \longmapsto X_t(\omega)$ on a measurable state-space S, is said to be **adapted** if X_t is $\mathcal{F}_t$- measurable for every $t \geq 0$.

We will assume that all considered processes are **R-processes** by which we mean that all paths are right-continuous on $I\!R^+$ with limits from the left on $I\!R_0^+$.

Definition 1 (Rogers & Williams [22])

A process X is a **semi-martingale** (relative to $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ if X is an adapted R-process which may be written in the form:

$$X = X_0 + M + A$$

where

i) M is a local martingale null at 0,

ii) A is a process null at 0 with paths of finite variation,

iii) M_0 is $\mathcal{F}_0$-measurable.

The following theorem introduces the important concept of quadratic variation for a martingale.

Theorem 2 (Rogers & Williams [22] p. 53)

Let $M \in c\mathcal{M}_0^2$, i.e.M is a continuous martingale null at 0 and bounded in L^2. Then there exists a unique continuous increasing process $[M]$ null at 0 such that $M^2 - [M]$ is a uniformly integrable martingale.

□

Definition 2

The process [M] as defined in Theorem 2 is called the **quadratic-variation process** of M.
If $M, N \in c\mathcal{M}_0^2$ then we define the **quadratic covariation** of M and N by polarisation:

$$[M, N] := \frac{1}{4}([M+N] - [M-N])$$

Theorem 3 : Itô's formula

Let $f : I\!R^n \longrightarrow I\!R$ be C^2 and let $X_t = (X_t^1 \dots X_t^n)$ be a continuous semimartingale in $I\!R^n$ (that is: each X^i is a continuous semimartingale).
Then

$$f(X_t) - f(X_0) = \int_0^t (\partial_i f)(X_s) dX_s^i + \frac{1}{2} \int_0^t (\partial_{ij}^2 f)(X_s) d[X^i, X^j]_s$$

□

<u>Remarks:</u>

i) Once again we used the summation convention.

ii) From now on we will write Itô's formula in its differential-form:

$$df(X_t) = (\partial_i f)(X_t) dX_t^i + \frac{1}{2}(\partial_{ij}^2 f)(X_t) d[X^i, X^j]_t$$

Also, we will follow standard practice and use the suggestive notation:

$$d[X,Y] = dXdY$$

Definition 3 (Rogers & Williams [22] p. 110)

Let $a : I\!R^n \longrightarrow S_n^+$ and $b : I\!R^n \longrightarrow I\!R^n$ be measurable functions, where

$$S_n^+ := \{\text{real } n \times n \text{ nonnegative-definite symmetric matrices}\}$$

A **diffusion** in $I\!R^n$ with **covariance** a and **drift** b is a continuous $I\!R^n$-valued semimartingale $X = (X^1, \ldots, X^n)$ defined on some filtered probability space (satisfying the usual conditions) such that, for each $i = 1, \ldots, n$,

$$M_t^i := X_t^i - X_0^i - \int_0^t b^i(X_s)ds$$

is a continuous local martingale for which:

$$[M^i, M^j]_t = \int_0^t a^{ij}(X_s)ds$$

For brevity we refer to a diffusion with covariance a and drift b as an (a,b)-diffusion.

Application of Itô's formula yields that if X is an (a,b)-diffusion, then

A. for each $f \in C^\infty$, the process

$$C_t^f := f(X_t) - f(X_0) - \int_0^t (Gf)(X_s)ds$$

where G is the **generator** of X given by

$$G := \frac{1}{2}a^{ij}\partial_i\partial_j + b^i\partial_i,$$

is a local martingale;

B. if in addition a and b are locally bounded, then for each $f \in C_K^\infty$ (smooth functions with compact support), C^f is a martingale.

In fact we have the following important theorem

Theorem 4 (Rogers & Williams [22] p. 111)

Let X be a continuous semimartingale in $\mathbb{R}^n$, and suppose that $a : \mathbb{R}^n \longrightarrow S_n^+$ and $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ are locally bounded. Then the following are equivalent:

1. X is an (a, b)-diffusion;
2. X satisfies condition A;
3. X satisfies condition B.

□

The connection between diffusions and stochastic differential equations is highlighted by the following simple observation. Let us assume that a covariance coefficient a and a driftterm b are given and let us assume furthermore that for some positive integer d we can find a (measurable) map;

$$\sigma : \mathbb{R}^n \longrightarrow L(\mathbb{R}^d, \mathbb{R}^n) \cong \mathbb{R}^{n\times d}$$

such that for each $x \in \mathbb{R}^n$

$$\sigma(x)\sigma(x)^T = a(x)$$

Moreover suppose that we can find a probabilistic set-up $(\Omega, \{\mathcal{F}_t\}, P)$ carrying both an $\mathbb{R}^n$-valued semimartingale X and an $\mathbb{R}^d$-valued Brownian motion B relative to $(\{\mathcal{F}_t\}, P)$ such that:

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt$$

which is shorthand for

$$X_t^i = X_0^i + \int_0^t \sigma_q^i(X_s)dB_s^q + \int_0^t b^i(X_s)ds$$

Then:

$$dM_t := dX_t - b(X_t)dt = \sigma(X_t)dB_t$$

defines a local martingale and since $dBdB^T = Idt$ we obtain:

$$dMdM^T = \sigma(X)dBdB^T\sigma(X)^T = a(X)dt$$

Hence X is an (a, b)-diffusion with generator:

$$G = \frac{1}{2}(\sigma\sigma^T)_{ij}\partial_{ij}^2 + b^i\partial_i$$

For an comprehensive study of the existence and uniqueness of SDE (and corresponding diffusions) we refer to Rogers & Williams [22] and Ikeda & Watanabe[8].

Notice that the above-mentioned conditions A and B can be generalized to manifolds by looking at smooth functions $f \in C^\infty_{(K)}(M, I\!R)$. This proves to be an elegant way to define diffusions on manifolds.
In particular we define **Brownian motion** on the Riemannian manifold (M, g) to be a diffusion with a generator (extending) $\frac{1}{2}\Delta$ on $C^\infty_K(M)$.
Again we refer to Ikeda & Watanabe [8] (chapter V) for an in depth study of all the technicalities that are involved.

We conclude this introductory chapter by pointing out that we will use the symbol ∂X to denote the Stratonovich-differential of X.
Recall that we can switch from Itô - to Stratonovichdifferentials by using the equation:

$$X\partial Y = X dY + \frac{1}{2} dX dY$$

where of course $dX dY = d[X, Y]$.

Chapter 2

Riemannian Submersions of Brownian Motion

2.1 The Submersion Theorem

As announced in chapter 1 we are interested in submersions

$$\pi : M \longrightarrow B$$

of Riemannian manifolds (M, g) and (B, h) and we want to determine when the projection $\pi(\beta)$ of Brownian motion β on M (notation: $\beta = BM(M)$) is an autonomous diffusion. Before we embark on the formulation and the proof of the main result in this chapter we want to look at some concrete examples which illustrate the problem we want to study and the kind of results we might expect. Since even the simplest examples illustrate the essential features of the problem, we will restrict our attention to low-dimensional spaces.

Example 1 :

Let $M = \mathbb{R}^2$ and $B = \mathbb{R}$ be equipped with the usual Euclidian metric, and let π be the standard-projection on the first component :

$$\pi : \mathbb{R}^2 \rightarrow \mathbb{R} : (x, y) \mapsto x$$

If $\beta = BM(\mathbb{R}^2)$ then $\beta = (\beta_1, \beta_2)$ where β_1 and β_2 are independent standard

$BM(\mathbb{R})$ and hence $\pi(\beta) = \beta_1$ is $BM(\mathbb{R})$.

$\diamond$

Example 2 :

We take another look at the above example; first however we observe that :

$$\forall x \in \mathbb{R}^2 \backslash \{0\} : P_x\{\beta_t = 0 \text{ for some } t \geq 0\} = 0$$

Hence $\mathbb{R}^2 \backslash \{0\} \equiv \mathbb{R}_0^2$ is stochastically complete.
Introducing polar coordinates (r, θ) on $\mathbb{R}_0^2$ we get the canonical projection :

$$\begin{aligned} \pi : \quad & \mathbb{R}_0^2 \to \mathbb{R}_0^+ \\ & (r, \theta) \mapsto r \end{aligned}$$

The projection $r_t = \pi(\beta_t)$ is a BESSEL(2)-process and therefore satisfies the SDE:

$$dr_t = db_t + \frac{1}{2r_t} dt \qquad \text{(b is } BM(\mathbb{R}))$$

This means that the projected process is a Brownian motion plus an additional drift.

$\diamond$

In both the above examples the projection of Brownian motion is a diffusion. However, it is easy to think of nontrivial situations where this isn't the case.

Example 3 :

Let M be the helix in $\mathbb{R}^3$ given by the following explicit parametrisation :

$$x = cost, y = sint, z = e^t, \qquad \text{(where } t \in \mathbb{R})$$

The metric on M is the induced Euclidean $\mathbb{R}^3$-metric and let β_t be the associated Brownian motion on M.
Furthermore, we define

$$B = \{x, y, 0) \in \mathbb{R}^3 | x^2 + y^2 = 1\}$$

and let $\pi = M \to B$ be the standard projection on the (x, y)-plane.
Now it's easy to see that because the vertical velocity dr/dt depends on t, the projection $\pi(\beta_t)$ cannot be a Markov-process, for it is obvious that if A is a Borelset in B and p, q are two different points in M such that $\pi(p) = \pi(q)$, then in general:

$$P\{\beta_t \in \pi^{-1}(A)|\beta_0 = p\} \neq P\{\beta_t \in \pi^{-1}(A)|\beta_0 = q\}$$

◇

From this example it is clear that we will at least need conditions on π which will ensure that the result $\pi(\beta_t)$ to be Markovian.
Such conditions can be found in Dynkin [5]:

Theorem 1 (Dynkin)

Let $X = (x_t, \mathcal{F}_t, P_x)$ be a Markov process on the statespace S with σ-algebra $\mathcal{S}$ and transition function $P(t, x, A)$.
Let f be a measurable transformation of $(S, \mathcal{S})$ into the statespace $(S', \mathcal{S}')$ such that $f(S) = S'$ and $f(\mathcal{S}) \subset \mathcal{S}'$, and such that the following condition is satisfied :

$$\forall A \in \mathcal{S}', \text{and for any } x, y \in S \;\text{ such that } f(x) = f(y) : \qquad$$
$$P(t, x, f^{-1}(A)) = P(t, y, f^{-1}(A)) \tag{2.1}$$

Set $x'_t = f(x_t)$ and denote by $\mathcal{N}'_0$ the σ-algebra generated by the sets $\{X'_t \in A\}(t \geq 0, A \in \mathcal{S}')$.
Further let $\mathcal{F}'_t = \mathcal{F}'_t \bigcap \mathcal{N}'_0$ and construct a transition probability:

$$P'_{f(x)}(A) = P_x(A) \qquad \text{for } A \in \mathcal{N}'_0 \tag{2.2}$$

Then the system $X' = (x'_t, \mathcal{F}'_t, P'_x)$ defines a Markov process on the statespace $(S', \mathcal{S})$ with transition function

$$P'(t, x', A) = P(t, x, f^{-1}(A)) \tag{2.3}$$

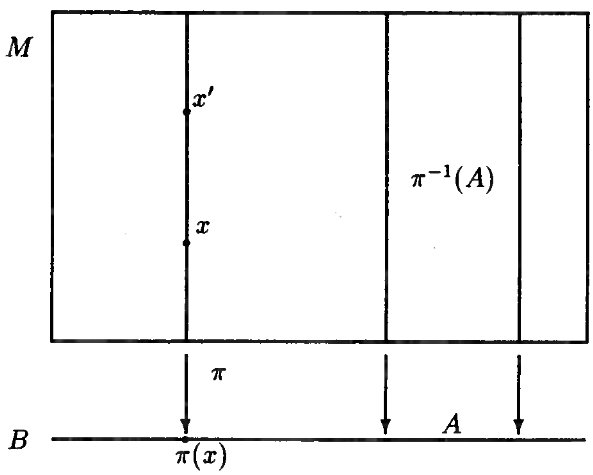
If X is a strong Markov process, then so is the process X'.

□

In the case of submersions, the inverse images $\pi^{-1}(A)$ (A in the σ-algebra on B) are cylindersets.
Since Brownian motion in a (Riemannian) manifold is intricately tied up with the metric we can, for a fixed set A and a fixed time t, interpret the transition probability $P(t, x, \pi^{-1}(A))$ as a sort of distance from x to $\pi^{-1}(A)$. Looked at in this light, condition (2.1) in Thm 1, suggests that in order for $\pi(\beta_t)$ to be a Markov process, all the points in the fibre $\pi^{-1}\{\pi(x)\}$ must have the same distance to the cylinderset $\pi^{-1}(A)$.
It is therefore not surprising that an essential condition on which the main result in this chapter hinges, turns out to be that π must be a **Riemannian** submersion. This notion captures precisely the idea that the distance in M of a point x to an arbitrary but fixed cylinderset $\pi^{-1}(A)$ is a fibre-invariant : i.e.,

$$\forall\, x' \in \pi^{-1}\{\pi(x)\} : \forall A \subset B : d_M(x, \pi^{-1}(A)) = d_M(x', \pi^{-1}(A)) \tag{2.4}$$

In thm. 7 we will prove that this characterization of a Riemannian submersion is equivalent to the usual definition.



Finally we notice that if the conditions of thm 1 hold then $\pi(\beta_t)$ is strongly Markovian and continuous since β_t enjoys these two properties.

To formulate the submersion theorem we recall the following notions from differential geometry; let (M,g) and (B,h) be two (finite-dimensional) Riemannian manifolds and $\pi : M \to B$ a C^∞-surjection.

Tangent vectors of M which are tangent to a fibre are called **vertical.** At each point $x \in M$ we can decompose the tangent-space T_xM as the direct sum of the subspace of vertical vectors $\mathbf{V}_x$ and its orthogonal complement (with respect to g): the horizontal subspace $\mathbf{H}_x$: $T_xM = \mathbf{V}_x \oplus \mathbf{H}_x$.

A submersion $\pi : M \to B$ is said to be a Riemannian submersion if for each $x \in M$ the horizontal subspace $\mathbf{H}_x$ is mapped isometrically by $d\pi(x)$ to $T_{\pi(x)}B$ (cf. Poor [21], p. 217).

We will also need the concept of the mean-curvature vector-field of a submanifold which is defined as follows: If N is a k-dimensional submanifold of the n -dimensional Riemannian manifold (M,g) with (Levi-Civita) connection ∇, then for each $x \in N$ it is always possible to construct — at least in a small enough neighbourhood U of x — orthonormal vectorfields $E_1, \ldots, E_n$ such that

$$\forall\, y \in U \cap N : E_1(y), \ldots, E_k(y) \in T_yN.$$

Using these vectorfields the mean curvature vector of N at x is defined by:

$$H := \frac{1}{k} \sum_{i=1}^{k} \sum_{j=k+1}^{n} g(\nabla_{E_i} E_i, E_j) E_j \tag{2.5}$$

(cf. Chen[3], p. 113). We can now formulate the main result: (cfr. Pauwels [16])

Theorem 2 : Submersion theorem

Let M be a n-dimensional and B a k-dimensional Riemannian manifold and let $\pi : M \to B$ be a Riemannian submersion. For every $x \in M$, let H_x be the mean curvature vector of the fibre through x.

If the projection of the mean-curvature vector H is constant along each fibre,

i.e.: $\forall\, b \in B, \forall\, x \in \pi^{-1}(b), d\pi(H_x)$ depends on b only,

then Brownian motion β_t on M projects down to a diffusion-process on B which can be decomposed into a Brownian motion with respect to the metric on B and an additional drift V given by:

$$V = -\frac{n-k}{2} d\pi(H) \tag{2.6}$$

Proof

Since the Brownian motion β_t is completely characterized by the Laplacian Δ_M on M, it suffices to study the projection of Δ_M. Take an arbitrary point $x \in M$ and denote $b = \pi(x) \in B$. In a neighbourhood U of x we choose a set of orthornormal vectorfields $E_1, \ldots, E_n$ as follows: on $\pi(U) \subset B$ we pick orthonormal vectorfields $F_1, \ldots, F_k$ which we then lift horizontally to $E_1, \ldots, E_k$ on U.Notice that because π is a Riemannian submersion the vectorfields $E_1, \ldots, E_k$ are orthonormal. We complete these framefields by adding orthonormal vertical vectors $E_{k+1}, \ldots, E_n$.

If we now introduce the functions K^i_{jk} defined by:

$$\nabla_{E_j} E_k = \sum_i K^i_{jk} E_i \tag{2.7}$$

then we can write the Laplacian as (cf. Rogers and Williams [22] eq. (V.34.76)):

$$\begin{aligned} \Delta_M &= \sum_i E_i^2 - \sum_i (\sum_j K^i_{jj}) E_i \\ &= \sum_{p=1}^{k} E_p^2 - \sum_{p=1}^{k} (\sum_{q=1}^{k} K^p_{qq}) E_p - \sum_{p=1}^{k} (\sum_{u=k+1}^{n} K^p_{uu}) E_p \\ &\quad + \sum_{u=k+1}^{n} E_u^2 - \sum_{u=k+1}^{n} (\sum_{j=1}^{n} K^u_{jj}) E_u \end{aligned} \tag{2.8}$$

Since the last two terms only involve vertical vectors their projection will vanish; moreover since π is a Riemannian submersion the vectorfields $F_1 = d\pi(E_1), \ldots, F_k = d\pi(E_k)$ satisfy (cfr. Crampin & Pirani [30], p.287)

$$d\pi(\nabla_{E_i} E_j) = \tilde{\nabla}_{F_i} F_j$$

where $\tilde{\nabla}$ is the (Levi-Civita) connection on B; therefore the first two terms in (2.8) will project down to the Laplacian Δ_B on B. By (2.5), we can write the formula for the mean curvature of the fibre through x as:

$$H = \frac{1}{n-k} \sum_{i=1}^{k} \sum_{j=k+1}^{n} K_{jj}^{i} E_i \tag{2.9}$$

from which we see that the third term in (4) is a multiple of the mean curvature H of the fibre (at that point). Hence projecting Δ_M down on B yields

$$\Delta_B - (n-k) d\pi(H) \tag{2.10}$$

and since the generator of β_t is $\frac{1}{2}\Delta_M$ the theorem follows.

□

Although the proof of theorem 2 is short and straightforward, it does not really explain why the result is true: more precisely it does not explain where the extra drift V comes from.

We will therefore give the outline of an alternative proof which is — although much more involved— intuitively more appealing.

Instead of looking at $BM(M)$ we consider a geodesic random walk on M which will converge (weakly) to $BM(M)$ (cf. Jørgensen [9]). Take an arbitrary point $x \in M$ and construct an orthonormal basis $E_1, \ldots, E_n$ at this point such that $E_1, \ldots, E_k$ are horizontal and $E_{k+1}, \ldots, E_n$ vertical (we will use the same notation as in theorem 2).

Next we construct the geodesics γ_i emanating from x with velocity E_i. To construct the geodesic random walk we parametrise these geodesics using the arclength-parameter s such that $\gamma_i(0) = x$. Then we can define $2n$ points $a_{\pm i} := \gamma_i(\pm\epsilon)$ (where ϵ is sufficiently small). The random walk is now obtained by jumping to each of these points $a_{\pm i}$ with probability $\frac{1}{2n}$ (using an appropiate clock).

Because the submersion π is Riemannian, the 'horizontal' geodesics $\gamma_p (p = 1, \ldots, k)$ will be mapped on geodesics in B under π (cf. O'Neill [14], cor 46, p. 212). This implies that jumps to 'horizontal' points $a_{\pm p}$ will result (under π) in a random-walk-process on B, which is invariant under geodesic reflection about x. We will call such a process symmetric for short.

As for the 'vertical' geodesics there are two possibilities: if the fibre $F_x = \pi^{-1}(\pi(x))$ is totally geodesic then (by definition) the geodesics $\gamma_u(t)(u = k+1, \ldots, n)$ will remain in F_x (at least for t sufficiently small). This implies that $a_{\pm u} \in F_x$ and hence $\pi(a_{\pm u}) = \pi(x)$ so that the projection of the random walk remains symmetric.

If however the fibre is not totally geodesic, the emanating geodesic γ_u will immediately leave F_x which implies that in general $\pi(a_{\pm u}) \neq \pi(x)$. This will result in an asymmetric random walk on B and therefore (in the limit) in a Brownian motion plus extra drift.

The case where F_x is minimal is interesting because although the geodesics leave the fibre, they do so in a symmetric way so that the contributions of the points $\pi(a_{\pm u})$ $(u = k+1, \dots, n)$ do not destroy the symmetry of $\pi(a_{\pm p})$.

To make all this precise we need to prove that the way in which the geodesics $\gamma_u(t)$ leave the fibre F_x is measured by the mean curvature vector H at x.
Let us formulate this in a slightly more general setting and introduce some more notation:

Let $N \subset \mathbb{R}^n$ be any p-dimensional submanifold of $\mathbb{R}^n (p < n)$. On $\mathbb{R}^n$ we assume an arbitrary connection ∇ and metric g and on N the induced connection $\tilde{\nabla}$ and metric $\tilde{g}$.

Take an arbitrary point $x \in N \subset \mathbb{R}^n$ and let $E_1, \dots, E_n$ be orthonormal vectorfields in a neighbourhood U of x such that if $y \in U \cap N$ then $E_1(y), \dots, E_p(y) \in T_yN$.

For $a = 1, \dots, p$, let $\gamma_a(t)$ (resp. $\delta_a(t)$) be the unique geodesic in $(\mathbb{R}^n, g)$ (resp. in $(N, \tilde{g})$) emanating from x with velocity-vector $E_a(x)$. Since we are working in $\mathbb{R}^n$ we can define the following vectorfields $K_a(t)$ along the geodesics $\gamma_a(t)$:

$$K_a(t) := \delta_a(t) - \gamma_a(t) \qquad \text{(for t sufficiently small)}$$

and

$$K_a(x) := 2 \lim_{t \to 0} \frac{K_a(t)}{t^2}$$

We can now formulate the following theorem:

Theorem 3

If

$$H(x) = \frac{1}{p} \sum_{j=p+1}^{n} g\Big(\sum_{a=1}^{p} \nabla_{E_a} E_a, E_j\Big) E_j$$

is the mean curvature vector (at x),
then

$$K(x) := \frac{1}{p} \sum_{a=1}^{p} K_a(x) = H(x)$$

Proof

Since we assume that all vectorfields are smooth we can apply l'Hôpital's rule to the definition of K_a:

$$\begin{aligned} K_a(x) : &= 2\lim_{t\to 0}\frac{\delta_a(t)-\gamma_a(t)}{t^2} = \lim_{t\to 0}(\ddot{\delta}_a(t)-\ddot{\gamma}_a(t)) \\ &= \ddot{\delta}_a(0)-\ddot{\gamma}_a(0). \end{aligned}$$

Since γ_a and σ_a are geodesics we get (using the summation convention):

$$\ddot{\gamma}_a^i(0) = -\Gamma^i_{jk}(0)\dot{\gamma}_a^j(0)\dot{\gamma}_a^k(0) = -\Gamma^i_{jk}E_a^jE_a^k \text{ (evaluated at } x \text{)}$$

where

$$E_a = E_a^i\partial_i$$

and similarly

$$\ddot{\sigma}_a^i(0) = -\tilde{\Gamma}^i_{jk}\dot{\sigma}_a^j(0)\dot{\sigma}_a^k(0) = -\tilde{\Gamma}^i_{jk}E_a^jE_a^k$$

Hence we can rewrite K_a as (suppressing x)

$$\begin{aligned} K_a &= \sum_i\sum_{jk}(\Gamma^i_{jk}-\tilde{\Gamma}^i_{jk})E_a^jE_a^k\partial_i \\ &= \sum_{jk}(\nabla_j\partial_k-\tilde{\nabla}_j\partial_k)E_a^jE_a^k \end{aligned}$$

Next we use Gauss' formula (cf. Chen[3] p. 109)

$$\nabla_XY-\tilde{\nabla}_XY = h(X,Y)$$

(where h is the second fundamental form of N) from which we conclude that:

$$K_a = \sum_{jk}E_a^jE_a^kh(\partial_j,\partial_k) = h(E_a,E_a)$$

and since

$$H = \frac{1}{p}\sum_{a=1}^{p}h(E_a,E_a)$$

we see that:

$$K = \frac{1}{p}\sum_{a=1}^{p} K_a = H$$

$\square$

Remarks

1. Notice that this theorem explains why the mean curvature always crops up in the drift of Brownian motion on submanifolds (cf. Rogers and Williams [22] (V.31.22) and Lewis [12]).It is not enough to project the Brownian motion of the ambient space on the submanifold, we have to add an extra drift which counteracts the curvature to keep it there.
Moreover, the above considerations make it clear why on a general Riemannian manifold the SDE for Brownian motion in terms of the local coordinates can be quite complicated. In general the (local) integralcurves of the coordinate-vectorfields $\partial_1, \dots, \partial_n$ will not be geodesics and therefore will have a nontrivial mean curvature. This introduces an extra drift-term in the SDE for the Brownian motion, messing up its simple structure. We will encounter an example of this when we have a closer look at Brownian motion on Liegroups.

2. If the ambient space is a general manifold instead of $\mathbb{R}^n$, the vectorfields $K_a(t)$ are no longer well-defined intrinsically. To use the same definition of $K_a(t)$ one has to imbed the manifold and its submanifold isometrically into a large enough $\mathbb{R}^N$ (this is always possible, cf. Rogers & Williams [22] p.233). The proof of theorem 2 remains essentially unchanged because the component of the mean curvature vectors of the manifold and its submanifold with respect to the ambient $\mathbb{R}^N$ are the same and therefore cancel.

Remarks on Theorem 2

1. The conditions in theorem 2 are independent of each other : the fact that, because π is a Riemannian submersion, all horizontal spaces along a fixed fibre are isometric, does not imply that the mean curvature of these fibres is constant along the fibres. This can easily be seen in the following example :

Example 1 :

Consider $M > I\!R^2$ equipped with the metric $g = \begin{pmatrix} 1 & 0 \\ 0 & 1+x^2y^2 \end{pmatrix}$ and $B = I\!R$ with the standard Euclidean metric.
The usual submersion $\pi : I\!R^2 \to I\!R$ is obviously Riemannian. The fibres are straight lines parallel to the y-axis but the have a nontrivial (mean) curvature :
take vectorfields $E_1 = \partial_x$ and $E_2 = \frac{1}{\sqrt{1+x^2y^2}}\partial_y$
then :

$$\nabla_{E_2}E_2 = \frac{1}{\sqrt{1+x^2+y^2}}\{\frac{\partial}{\partial y}(\frac{1}{\sqrt{1+x^2+y^2}}\partial_y) + \frac{1}{\sqrt{1+x^2+y^2}}\nabla_y\partial_y\}$$

But since we are only interested in the horizontal component of $\nabla_{E_2}E_2$ we get;

$$(\nabla_{E_2}E_2)_H = \frac{1}{\sqrt{1+x^2+y^2}}\Gamma^x_{yy}\partial_x$$

where $\Gamma^x_{yy} = -xy^2$
Hence :

$$(\nabla_{E_2}E_2)_H = \frac{-xy^2}{\sqrt{1+x^2+y^2}}E_1$$

which is a function of both x and y.

◇

On the other hand, the following is an example where the fibres have zero mean curvature, but the induced metric on the horizontal subspaces is not a fibre-invariant.

Example 2 :

Let $M = I\!R^2$ with the metric tensor

$$g = \begin{pmatrix} 1+y^2 & 0 \\ 0 & 1 \end{pmatrix}$$

and again

$$\pi : I\!R^2 \to I\!R : (x, y) \mapsto x$$

Since the metric on the horizontal spaces is given by:

$$ds^2 = (1 + y^2)dx^2$$

we see that horizontal spaces along a fixed fibre are not isometric. However a simple calculation similar to the one in the previous example yields that the mean curvature of each fibre is zero, for take :

$$E_1 = \frac{1}{\sqrt{1+y^2}}\partial_x \quad \text{and} \quad E_2 = \partial_y$$

then since $\Gamma^x_{yy} = \Gamma^y_{yy} = 0$ we get $\nabla_{E_2} E_2 = 0$. ◇

2. On the other hand, the conditions on the geometric structure of M are not so stringent as to exclude all but a number of essentially trivial cases where the metric on M factorizes.
   In particular, the following is an example of a manifold (M, g) where

   (i) all the horizontal space are isometric,

   (ii) the mean-curvaturevector is a fibre-invariant

   but the metric g is not invariant along fibres.

Example 3 :

Let $M = \mathbb{R}^2$ be equipped with the metric : $g = \begin{pmatrix} 1 & 0 \\ 0 & e^{2(x+y)} \end{pmatrix}$

We consider the submersion : $\pi : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x$.

Global orthonormal vectorfields are given by :

$$E_1 = \partial_x \qquad\qquad E_2 = e^{-(x+y)}\partial_y$$

Computing the relevant Christoffelsymbols we see that,

$$\Gamma^x_{yy} = -e^{2(x+y)} \quad \text{and} \quad \Gamma^y_{yy} = 1$$

Hence : $\nabla_{E_2} E_2 = -\partial_x = -E_1$

which yields : $H(x, y) = -\partial_x$ (constant !)

However it is easy to see that the Lie-derivative $L_{\partial_y} g(\partial_y, \partial_y) = 2e^{2(x+y)}$and hence :

$$L_{\partial_y} g \neq 0$$

◇

3. Theorem 2 gives us sufficient conditions to obtain a diffusion on B, but are they also necessary.
   Reexamining examples 1, 2 in remark 1 shows us that neither of them on its own is sufficient for that purpose.
   The easiest way to see this is by computing the Laplacians associated with the metrics on $\mathbb{R}^2$: in general we have :

$$\Delta = \frac{1}{\sqrt{detg}}\partial_i(g^{ik}\sqrt{detg}\partial_k)$$

   Hence in ex. 1 :

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{1}{1+x^2y^2}\frac{\partial^2}{\partial y^2} + \frac{xy^2}{1+x^2y^2}\frac{\partial}{\partial x} - \frac{x^2y}{(1+x^2y^2)^2}\frac{\partial}{\partial y}$$

   and therefore the projected process X_t satisfies the SDE :

$$dx_t = db_t + \frac{xy^2}{2(1+x^2y^2)}dt \qquad b_t \text{is } BM(\mathbb{R})$$

   which is obviously non-Markovian.

   Similarly in ex. 2 :

$$\Delta = \frac{1}{1+y^2}\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial y^2} + \frac{y}{1+y^2}\frac{\partial}{\partial y}$$

   and hence x_t satisfies :

$$dx_t = \frac{1}{\sqrt{1+y^2}}db_t \qquad \text{(where } b_t \text{ is } BM(\mathbb{R}))$$

   again, non-Markovian.

◇

In what follows we will prove that if we want the projected process to be a diffusion with a (Markovian) generator of the form (2.10) both conditions in thm 2 are necessary.

Firstly however we recall a very useful technical result which tells us that by a judicious choice of coordinate-system we can (at least locally) make a submersion look like a standard-projection in $\mathbb{R}^n$.
More precisely we have the following theorem
(cfr. Spivak [25] vol I, p. 2 - 24):

Theorem 4

Let $\pi : M \longrightarrow B$ be a C^∞-submersion, $dimM = n + m$, $dimB = n$.
Take $p \in M, q = \pi(p) \in B$.
Then there exist local charts

$$\varphi : p \in U \longrightarrow \mathbb{R}^{n+m}$$
$$\psi : q \in V \longrightarrow \mathbb{R}^n$$

and a decomposition $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m \ni (x, y)$
such that

$$\psi \circ \pi \circ \varphi^{-1}(x, y) = x$$

□

Let us now return to the problem at hand.
It's obvious that the projection of the mean curvature vector $d\pi(H)$ can only depend on $b \in B$, since if this wasn't the case the operator (2.7) in thm 2 would have a first-order part that involved coordinates within the fibre.
To prove that we also need the Riemannian character of the submersion to arrive at a generator of the form (2.7) we proceed as follows.

Invoking thm 4 we see that without loss of generality we can restrict our attention to the following situation: $M = \mathbb{R}^{n+m}, B = \mathbb{R}^n$ and

$$\pi : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^n : (x^1 \ldots x^n, x^{n+1} \ldots x^{n+m}) \longmapsto (x^1 \ldots x^n)$$

For convenience we introduce the following generic notation for the indices:

$$p, q : 1 \longrightarrow n,$$

$$u, v : n+1 \longrightarrow n+m$$

Furthermore we factorize the metric g on M as follows:

$$g = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{m\times m}$.
The metric on B will be denoted by h.

We then have the following lemma:

Lemma 5:

(i) If $\pi : M \longrightarrow B$ is a Riemannian submersion, then the explicit form of h is given by:

$$h = A - BC^{-1}B^T \tag{2.11}$$

(ii) Conversely, if the metric h on B is given by (2.11) (which a fortiori implies that the RHS of (2.11) is independent of $(x^u : u = n+1, \ldots, n+m)$), then $\pi : (M, g) \longrightarrow (B, h)$ is a Riemannian submersion.

Proof:

Using the aforementioned indexconvention, any tangentvector X in $\mathbb{R}^{n+m}$ can be written as:

$$X = \xi^p \partial_p + \eta^u \partial_u$$

and this vector X will be horizontal if and only if

$$\forall\, \zeta \in \mathbb{R}^{m\times 1} : (0 \quad \zeta^T) \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = 0$$

or again

$$\forall\, \zeta \in \mathbb{R}^{m\times 1} : \zeta^T (B^T \xi + C\eta) = 0$$

Since $B^T\xi + C\eta \in I\!R^{m\times 1}$, it follows that:

$$X \text{ is horizontal } \underline{\text{iff}}\ B^T\xi + C\eta = 0$$

Notice that since C is symmetric and positive-definite, it is invertible and we therefore can construct η explicitely for any given ξ:

$$\eta = -C^{-1}B^T\xi \tag{2.12}$$

In particular, we can choose $\xi = \xi(x^1, \ldots, x^n)$ (independent of $x^{n+1}, \ldots, x^{n+m}$) so that the projection $d\pi(X) = \xi^p\partial_p$ is constant along each fibre.
Using such vectorfields, we can compute an explicit formula for the metric h on $I\!R^n$. For if

$$X = \xi^p\partial_p + \eta^u\partial_u$$

is horizontal and $\xi^p = \xi^p(x^p)$, then the fact that π is Riemannian implies:

$$h(d\pi(X), d\pi(X)) = g(X, X)$$

The LHS is given by:

$$LHS = h_{pq}\xi^p\xi^q = \xi^T h\xi$$

whereas:

$$\begin{aligned} RHS &= \begin{pmatrix} \xi^T & \eta^T \end{pmatrix} \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \\ &= \xi^T A\xi + \xi^T B\eta + \eta^T B^T\xi + \eta^T C\eta \\ &= \xi^T(A - BC^{-1}B^T)\xi \end{aligned}$$

for C, and hence C^{-1}, is symmetric.
Comparing the LHS and the RHS and observing the equality holds for all ξ, proves the first part of the lemma.
As for the converse all we have to do is to check that for an arbitrary horizontal vector

$$X = \xi^p\partial_p + \eta^u\partial_u$$

we have that

$$h(d\pi(X), d\pi(X)) = g(X, X)$$

But this is a trivial consequence of (2.12).

$\square$

Remark: Risking the scorn and indignation of purists we will introduce the following notation for the metric (2.11):

$$h = \pi(g)$$

We are now in a position to prove the following result.

Theorem 6

If using the notation of thm 2, the submersion $\pi(\beta)$ of β is a diffusionprocess with generator of the form (2.10) then π must be an Riemannian submersion.

Proof:

Again we involve thm 4 to construct local coordinate-neighbourhoods such that the submersion

$$\pi : (M, g) \longrightarrow (B, h)$$

can be expressed as ($dimM = n$, $dimB = k$)

$$\pi(x^1, \ldots, x^n) = (x^1, \ldots, x^k)$$

We assume that the metric g can be written as

$$g = \begin{pmatrix} S & W \\ W^T & T \end{pmatrix} \tag{2.13}$$

where $S \in \mathbb{R}^{k \times k}, W \in \mathbb{R}^{k \times (n-k)}$ and $T \in \mathbb{R}^{(n-k) \times (n-k)}$. To facilitate the use of the summationconvention, we will use generic indices:

$$i, j : 1 \longrightarrow n$$
$$u, v : \longrightarrow k$$

Let β be Brownian motion on M and hence β has a generator equal to $\frac{1}{2}\Delta_M$ where

$$\begin{aligned}\Delta_M &= \frac{1}{\sqrt{detg}}\partial_i(g^{ij}\sqrt{detg}\partial_j) \\ &= g^{ij}\partial^2_{ij} + [\partial_i g^{ij} + \frac{1}{2}g^{ij}\partial_i(lndetg)]\partial_j\end{aligned}$$

In particular it follows from this that

$$d\beta^i d\beta^j = g^{ij}dt \qquad (2.14)$$

On the other hand, let $\gamma_t = \pi(\beta_t)$ be the projection of β and let's assume that γ_t is a diffusion with generator

$$G = \frac{1}{2}a^{uv}\partial_u\partial_v + b^u\partial_u$$

Since

$$\gamma^u(t) = \beta^u(t) \qquad u = 1, \ldots, k$$

we obtain from (2.14):

$$d\gamma^u d\gamma^v = d\beta^u d\beta^v = g^{uv}dt$$

from which it follows that for all u,v:

$$a^{uv} = g^{uv} \qquad (2.15)$$

If we now introduce a decomposition of g^{-1} analogous to (2.13):

$$g^{-1} = \begin{pmatrix} P & Q \\ Q^T & R \end{pmatrix} \qquad (2.16)$$

then it is easy to check that

$$P = (S - WT^{-1}W^T)^{-1} \tag{2.17}$$

(inverte (2.17) and multiply both sides by P).
Inverting the equality we obtain by combining (2.15) and (2.17) we see that:

$$(a_{uv}) = S - WT^{-1}W^T \tag{2.18}$$

In particular the RHS of (2.18) can only depend on $(x^1 \dots x^k)$.

Next if we want the diffusion on B to have a generator of the form

$$G = \frac{1}{2}\Delta_B + \ (1^{\text{st}} \text{ - order operator})$$

then we must insist on the equality:

$$(a_{uv}) = (h_{uv}) \tag{2.19}$$

and hence combining (2.18) and (2.19) we see:

$$h = S - WT^{-1}W^T$$

which is independent of $(x^{k+1}, \dots, x^n)$.
In view of lemma 5 this completes the proof.

□

We end this section with a characterization of Riemannian submersions which consolidates the intuitive explanation on page 18 as to why the Riemannian character of the submersion is necessary.

Theorem 7 (Characterization of Riemannian Submersions)
Let M and B be two complete Riemannian manifolds and $\pi : M \longrightarrow B$.
Then:

$$\pi \text{ is a Riemannian submersion}$$
$$\underline{\text{iff}}$$
$$(A) : \forall\, x, y \in M : d_B(\pi(x), \pi(y)) = d_M(x, F_y)$$

Proof

$\Downarrow$: First we observe that if $\{\gamma(t) \mid t \in [a,b]\}$ is a (differentiable) curve in M and

$$l_M(\gamma) = \int_a^b ||\dot{\gamma}(t)||_M dt \tag{2.20}$$

its length, then

$$l_B(\pi(\gamma)) \leq l_M(\gamma), \tag{2.21}$$

and equality holds iff γ is horizontal.
This follows immediately from the realisation that in view of (2.20) it suffices to show that

$$\forall x \in M,\ \forall v \in T_xM,\quad ||d\pi(v)||_B \ \leq\ ||v||_M$$

But this is trivial since for every tangent vector v there is a unique decomposition $v = v_M + v_V$ in a horizontal and a vertical component ($d\pi(v_V) = 0$).
Hence:

$$\begin{aligned} ||d\pi(v)||_B &= ||d\pi(v_M)||_B \\ &= ||v_H||_M \qquad\qquad (\pi \text{ is Riemannian submersion}) \\ &\leq ||v||_M \end{aligned}$$

Obviously equality holds iff v is horizontal.

Next, to prove this first part of the theorem, fix $x, y \in M$ such that $x \notin F_y$. Now take a new curve $\gamma = \{\gamma(t) \mid 0 \leq t \leq 1\}$ in B such that $\gamma(0) = \pi(x), \gamma(1) = \pi(y)$ and

$$l_B(\gamma) = d_B(\pi(x), \pi(y))$$

This is possible since B is metrically complete.
We then lift this curve horizontally to a curve $\tilde{\gamma}$ in M such that $\tilde{\gamma}(0) = x$. Clearly since $\pi(\tilde{\gamma}(1)) = \pi(y)$ we have (by definition): $\tilde{\gamma}(1) \in F_y$.
Since $\tilde{\gamma}$ is horizontal we know that

$$l_M(\tilde{\gamma}) = l_B(\gamma) = d_B(\pi(x), \pi(y))$$

On the other hand, $\tilde{\gamma}$ connects x and F_y and hence it follows that

$$d_M(x, F_y) \leq l_M(\tilde{\gamma})$$

Combining these two inequalities yields:

$$d_M(x, F_y) \leq d_B(\pi(x), \pi(y)) \tag{2.22}$$

The reversed inequality is obvious since, because F_y is closed there is a curve $\tilde{\sigma} = \{\tilde{\sigma}(t) \mid 0 \leq t \leq 1\}$ such that:

$$\tilde{\sigma}(0) = x, \quad \tilde{\sigma}(1) \in F_y \quad \text{and} \quad l_M(\tilde{\sigma}) = d_M(x, F_y)$$

This can be seen as follows: take an arbitrary point z in F_y and consider (in M) the closed sphere S with radius equal to $r := d_M(x, z)$. S is compact since M is complete (cfr. thm 1 in chapter 1) and hence we can conclude that the intersection $F_y \cap S$ is compact. It therefore follows that the continous function

$$\varphi(p) := d_M(x, p)$$

attains its minimum on $F_y \cap S$ (in q say).
Again invoking the completeness of M we know that there is a geodesic segment $\tilde{\sigma}$ connecting X and q, the length of which is equal to the distance between thee two points.
The projection $\sigma := \pi(\tilde{\sigma})$ satisfies: $\sigma(0) = \pi(x)$ and $\sigma(1) = \pi(y)$ and hence

$$d_B(\pi(x), \pi(y)) \leq l_B(\sigma)$$

But on the other hand, it follows from equation (2.21) that

$$l_B(\sigma) \leq l_M(\tilde{\sigma}) = d_M(x, F_y)$$

Combining these two inequalities yields:

$$d_B(\pi(x), \pi(y)) \leq d_M(x, F_y) \tag{2.23}$$

Hence (2.22) and (2.23) imply:

$$d_B(\pi(x), \pi(y)) = d_M(x, F_y)$$

$\Uparrow$: Invoking the polarization-equality we see that all we have to do is to prove that if $TM = HM \oplus VM$ is the decomposition of the tangent bundle in a horizontal and a vertical part, then assumption A implies:

$$\forall x \in M,\ \forall v \in H_x(M) : ||d\pi(v)||_B = ||v||_M \tag{2.24}$$

Notice that since the LHS of A only depends on $\pi(x)$ and not on x itself, so does the RHS.

For the sake of clarity we split this proof into different lemmas:

Lemma 8:

If (A) holds then $\forall\ v \in TM : ||d\pi(v)||_B \leq ||v||_M$.

Proof:

First we note that because $d\pi$ is a linear map it suffices to prove the above inequality for 'small' vectors for we can always adjust the norm of a vector by multiplying it by an appropriate (positive) factor.
Now fix $p \in B$ and take an arbitrary $w \in T_pB$.
If $||w||_B$ is sufficiently small we can define:

$$q := Exp_p w$$

(where Exp is the exponential map on a Riemannian manifold), and then we know that:

$$d_B(p,q) = l_B(\gamma) = ||w||_B \tag{2.25}$$

where γ is the geodesical segment connecting p and q (cfr. Poor [21] p. 131).

By assumption:

$$d_B(p,q) = d_M(\pi^{-1}(p), \pi^{-1}(q)). \tag{2.26}$$

Pick an arbitrary $x \in \pi^{-1}(p)$ and let

$$\lambda_x : T_{\pi(x)}B \longrightarrow H_x(M)$$

denote the horizontal lift to x of vectors and, slightly abusing the notation, of curves (starting at $\pi(x)$); then we can define:

$$v = \lambda_x(w) \qquad \text{and} \quad \hat{\gamma} = \lambda_x(\gamma)$$

Since $\hat{\gamma}$ connects x and $\pi^{-1}(q)$ it follows that:

$$d_M(x, \pi^{-1}(q)) = d_M(\pi^{-1}(p), \pi^{-1}(q)) \leq l_M(\hat{\gamma}) \tag{2.27}$$

Combining (2.25), (2.26) and (2.27) we see:

$$l_B(\gamma) = d_B(p,q) = d_M(\pi^{-1}(p), \pi^{-1}(q)) \leq l_M(\hat{\gamma}). \tag{2.28}$$

Since the curves γ and $\hat{\gamma}$ are smooth the integral-definition of curvelength in combination with inequality (2.28) implies that:

$$||v||_M = ||\lambda_x(w)||_M \geq ||w||_B. \tag{2.29}$$

Since for an arbitrary horizontal vector v we have:

$$v = \lambda_x(d\pi_x(v))$$

we see that inequality (2.29) can be rewritten as:

$$\forall\ v \text{ horizontal} : ||d\pi(v)||_B \leq ||v||_M \tag{2.30}$$

If $v \in TM$ isn't horizontal then we can make the unique decomposition: $v = v_H + v_V$
and

$$||v||_M^2 = ||v_H||_M^2 + ||v_V||_M^2$$

Therefore for general $v \in TM$:

$$||d\pi(v)||_B = ||d\pi(v_H)||_B \leq ||v_H||_M \leq ||v||_M \tag{2.31}$$

□

Notice that in order to prove the theorem we must show that the inequality (2.31) is in fact an equality.
Notice moreover that this lemma also implies that the submersion reduces curve-lengths: $l_B(\pi(\gamma)) \leq l_M(\gamma)$.

Lemma 9:

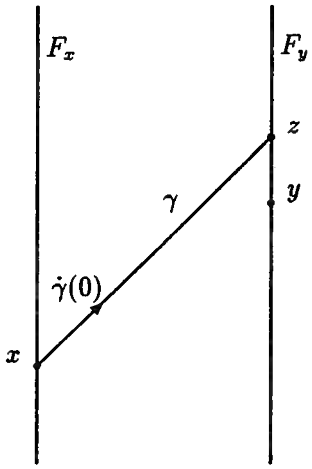
If $x, y \in M$ and γ is a curve in M such that $l_M(\gamma) = d(x, F_y)$, then (A) implies that γ is (a segment of) a horizontal geodesic.

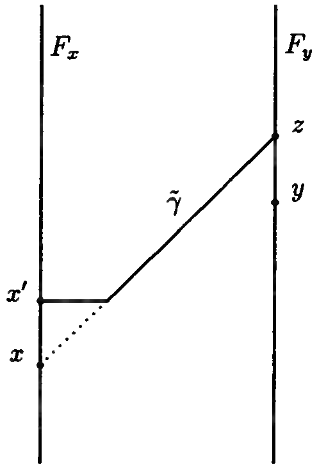
Proof:

Let $\gamma = \{\gamma(t) \mid o \leq t \leq 1\}$ be such that $\gamma(0) = x$ and $\gamma(1) = z \in F_y$. (We'll assume that $x \notin F_y$).
Obviously γ must be a geodesic segment, otherwise there would be another curve connecting x and z which has a shorter length.
Intuitively it is obvious that the vector $\dot{\gamma}(0)$ must be horizontal for if it were not, it would be easy to construct a shorter curve $\tilde{\gamma}$ connecting the fibres F_x and F_y, thereby contradicting the assumption that γ realises the distance $d(x, F_y) = d(F_x, F_y)$.
More precisely this is an immediate consequence of the following:

Proposition (Cheeger & Ebin [29], prop. 1.5 p.5)

Let N and $\bar{N}$ be two submanifolds of the manifold M, without boundary and let $\gamma : [0,1] \longrightarrow M$ be a geodesic such that $\gamma(0) \in N, \gamma(1) \in \bar{N}$ and γ is the shortest curve from N to $\bar{N}$. Then $\dot{\gamma}(0)$ is perpendicular to $T_{\gamma(0)}N$ and $\dot{\gamma}(1)$ is perpendicular to $T_{\gamma(1)}\bar{N}$. $\diamond$





We can now apply the same argument to every point of the curve, for suppose p is a point on the geodesic γ where the velocity vector isn't horizontal: then by assumption $\gamma|_{[p,y]}$ is the shortest connection between p and F_y, for if there was a shorter arc γ' connecting p and $y' \in F_y$ then of course $\tilde{\gamma} := \gamma|_{[x,p]} \cup \gamma'|_{[p,y']}$ would be a shorter arc connecting x and F_y which contradicts our assumption of minimality of the original arc γ.
But since $\gamma|_{[p,z]}$ establishes the shortest connection between p and F_y it follows (from assumption A) that it also realises the shortest distance between F_p and F_y. Hence we can again apply the abovementioned lemma to deduce that the velocity vector at p is horizontal. This proves lemma 6.

$\square$

It's important to realise that lemma 6 establishes the existence of horizontal geodesics if assumption (A) is satisfied. If we drop this assumption then it is easy to give counterexamples of submersions for which horizontal geodesics (in general) do not exist.

Counterexample
Let $M = \mathbb{R}^2$ and $B = \mathbb{R}$ both equipped with the standard Euclidean metric. Consider the submersion:

$$\pi : \mathbb{R}^2 \longrightarrow \mathbb{R} \ : \ (x,y) \longmapsto x - y^2$$

for which it's easy to see that (A) doesn't hold. It's obvious that horizontal vectors are proportional to the gradients of $F(x,y) = x - y^2$, hence a horizontal vector at (x,y) is proportional to

$$\nabla F = (1, -2y)$$

If $\gamma(t) = (\gamma_x(t), \gamma_y(t))$ is a curve in $\mathbb{R}^2$ then $\gamma(t)$ will be horizontal if there is a scalar function $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ such that:

$$\forall\, t : \dot\gamma(t) = \alpha(t)(1, -2\gamma_y(t))$$

for which an explicit solution is given by:

$$\gamma(t) = (A(t), ce^{-2A(t)}) \qquad (c \in \mathbb{R})$$

where $\quad dA/dt = \alpha(t)$.
However since $M = \mathbb{R}^2$ is equipped with the usual metric, geodesics are straight lines and hence we see that $\gamma(t)$ will only be a geodesic if : $c = 0$ and $A(t) = at + b$. So if we choose a starting-point (p,q) where $q \neq 0$ then it is impossible to construct a horizontal geodesic emanating from (p,q).

$\diamond$

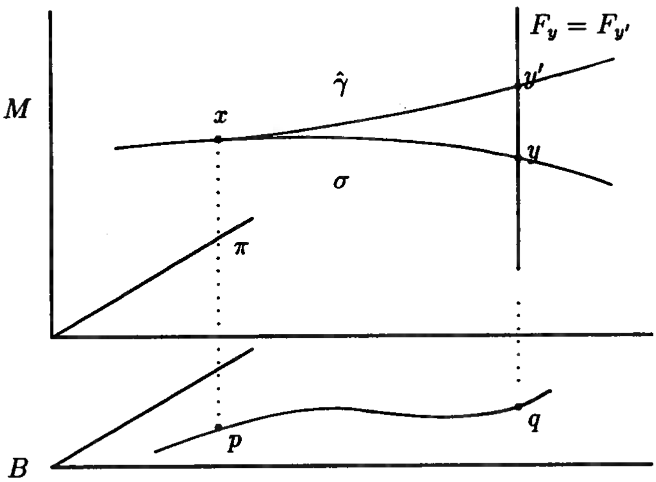
This counterexample illustrates that the following lemma is nontrivial:

Lemma 10:

If the assumption (A) holds then the horizontal lift of a geodesic in B is a horizontal geodesic in M.

Proof:

Let γ be an arbitrary geodesic in B and pick a point p on γ. Choose an arbitrary point $x \in \pi^{-1}(p)$, and let $\hat{\gamma} = \lambda_x(\gamma)$ be the (unique) horizontal lift of γ, passing through x. To show that $\hat{\gamma}$ is a geodesic we will prove that if we take a point y' on $\hat{\gamma}$ (close enough to x), then the segment $\hat{\gamma}_{xy'}$ will have a length equal to the distance $d_M(x, y')$. Since a geodesic is characterized by the fact that it is the shortest curve between any two of its points (which lie in a small enough neighbourhood of each other) (cfr.: O'Neill [14] p. 137) it follows that $\hat{\gamma}$ is a geodesic.
Pick points q on γ and y' on $\hat{\gamma}$ such that $\pi(y') = q$ and both q and y' lie in a spherical normal neighbourhood of p resp. x.



Let σ be the horizontal geodesical segment (geosedic for short) such that: (cfr. lemma 9)

$$l_M(\sigma) = d(x, F_{y'})$$

and let $y \in F_{y'}$ be the point where σ 'hits' the fibre $F_{y'}$.
Hence, since $F_{y'} = F_y$ we have

$$\begin{aligned} l_M(\sigma) = d(x, F_y) &= d(\pi(x), \pi(y)) \\ &= d(p, q) = l_B(\gamma) \qquad \text{(where } \gamma \equiv \gamma_{pq}\text{: segment)} \quad (2.32) \end{aligned}$$

On the other hand, if we define $\bar{\sigma} := \pi(\sigma)$ then it follows from lemma 8 :

$$l_B(\bar{\sigma}) \le l_M(\sigma) \tag{2.33}$$

But $\bar{\sigma}$ connects the points p and q and therefore:

$$l_B(\gamma) \equiv d_B(p,q) \le l_B(\bar{\sigma}) \tag{2.34}$$

Combining (2.32), (2.33) and (2.34) we obtain:

$$l_B(\gamma) \le l_B(\bar{\sigma}) \le l_M(\sigma) = l_B(\gamma)$$

So: $l_B(\bar{\sigma}) = l_B(\gamma)$ and since γ is a geodesical segment in a (small) spherical normal neighbourhood of p it follows that $\bar{\sigma} = \gamma$ and hence we can conclude that $\bar{\sigma}$ is a geodesical segment.
But:

$$\lambda_x(\bar{\sigma}) = \lambda_x(\pi(\sigma)) = \sigma \quad \text{(for } \sigma \text{ is horizontal)}$$

hence:

$$\lambda_x(\gamma) = \lambda_x(\bar{\sigma}) = \sigma$$

and we see that the horizontal lift of the geodesical segment γ is a (horizontal) geodesical segment. □

Proof of second part of characterization-theorem

It is now trivial to prove that (A) implies that the submersion π is in fact Riemannian. For from lemma 10 it follows that
since

$$\bar{\sigma} = \gamma \Longrightarrow \lambda_x(\bar{\sigma}) = \lambda_x(\gamma)$$

or again

$$\sigma = \hat{\gamma}$$

But:

$$l_M(\hat{\gamma}) = l_M(\sigma) = l_B(\gamma).$$

Therefore the horizontal lift of the geodesical segment γ produces a horizontal geodesical segment of the same length.

Using the integral-definition of curvelength this implies that the submersion of horizontal vectors is norm-preserving, thus proving equation(2.24).

□

2.2 Brownian motion on Liegroups

The most natural way to define Brownian motion on a Liegroup G is by means of the so-called method of product-integral injection (cfr. Rogers and Williams [22],V.35); essentially this amounts to taking an appropriate diffusion on the Lie-algebra g of G (where g is interpreted as a vectorspace) and then using the exponential map exp: $g \longrightarrow G$ to transfer this diffusion on g to a (left or right) Brownian motion on G.

Because Brownian motions on Liegroups will feature in several subsequent examples we look at them in a bit more detail.
Let us start by introducing some notation: in this section G will denote a general (finite-dimensional) connected Liegroup and g its corresponding Lie-algebra. Each element $\gamma \in G$ gives rise to two canonical isomorphisms of the Liegroup G: left- and right-translation:

$$\lambda_\gamma : G \longrightarrow G \; : \; \xi \longmapsto \gamma\xi \tag{2.35}$$

$$\rho_\gamma : G \longrightarrow G \; : \; \xi \longmapsto \xi\gamma \tag{2.36}$$

the derivatives of which can be used to transport an arbitrary tangent-vector at the identity $e \in G$ around the group, thus defining left- and right-invariant vectorfields on G (L/R-IVF).
So if $v \in T_eG$ then the associated LIVF and RIVF are given by:

$$\forall \xi \in G \quad : \quad L_v(\xi) := d\lambda_\xi(\xi)(v) \tag{2.37}$$
$$R_v(\xi) := d\rho_\xi(\xi)(v) \tag{2.38}$$

Adopting standard-practice we will identify g and T_eG by identifying the LIVF L_v and its value v at e.

Next we recall that with every LIVF L_v there is associated an one-parametersubgroup $\{q_v(t) \mid t \in \mathbb{R}\}$ which is the integralcurve of L_v passing through e; i.e.: $q_v(t)$ is the (unique) solution of the following differential equation:

$$\begin{cases} \dot{q}_v(t) = L_v(q_v(t)) \\ q_v(0) = e \end{cases} \tag{2.39}$$

Notice that it follows immediately from the definition that the left-cosets of $q_v(t)$ constitute the integralcurves of the LIVF L_v: for let $\xi \in G$ be an arbitrary point and $\tilde{q}_v(t) = \xi q_v(t) = \lambda_\xi(q_v(t))$ the left coset passing through ξ;
then:

$$\begin{aligned}
\frac{d}{dt}(\tilde{q}_v(t)) &= \frac{d}{dt}(\lambda_\xi(q_v(t)) \\
&= d\lambda_\xi(\frac{d}{dt}q_v(t)) \\
&= d\lambda_\xi(L_v(q_v(t)) \\
&= d\lambda_\xi(d\lambda_{q_v(t)}(v)) \\
&= d\lambda_{\xi q_v(t)}(v) = L_v(\tilde{q}_v(t))
\end{aligned}$$

Using these one-parameter subgroups we can construct the exponential map

$$\begin{aligned}
exp \ : \ & g \longrightarrow G \\
& v \longmapsto exp(v) := q_v(1)
\end{aligned}$$

From this it immediately follows that:

$$\forall t \in I\!R: \ exp(tv) = q_v(t) \tag{2.40}$$

Finally we recall that if $L_1, \ldots, L_n$ form a basis for g, then the corresponding structure constants C_{ij}^k are defined by:

$$[L_i, L_j] = C_{ij}^k L_k \tag{2.41}$$

Notice that we can introduce similar constants for the RIVF R_i:

$$[R_i, R_j] = \tilde{C}_{ij}^k R_k \tag{2.42}$$

The relation between these two sets of constants can be obtained by looking at the inversionmap:

$$i : G \longrightarrow G : \gamma \longmapsto \gamma^{-1}$$

which is a Liegroup-isomorphism and therefore induces a Lie-algebra-isomorphism:

$$di : g \longrightarrow g$$

Hence if L_v is the left-invariant field generated by v at e then for any $\gamma \in G$

$$L_v(\gamma) = \frac{d}{dt}(\gamma \cdot exp(tv))|_{t=0}$$

and consequently:

$$\begin{aligned} di(L_v(\gamma)) &= \frac{d}{dt}(i(\gamma \cdot exp(tv)))|_{t=0} \\ &= \frac{d}{dt}(exp(-tv)\gamma^{-1})|_{t=0} \\ &= R_{-v}(\gamma^{-1}) = -R_v(\gamma^{-1}) \end{aligned}$$

Therefore applying di to (2.42) we get:

$$[-L_i, -L_j] = \tilde{C}^k_{ij}(-L_k)$$

or again:

$$[L_i, L_j] = -\tilde{C}^k_{ij} L_k$$

and hence:

$$\tilde{C}^k_{ij} = -C^k_{ij}$$

Notice moreover that left- and right-invariant vectorfields commute:

Lemma 11

If L is a left-invariant vectorfield on G and R is a right-invariant vectorfield on G, then:

$$[L, R] = 0$$

Proof:

The easiest way to see that the above equality holds is to use the following well-known characterization of the Lie-bracket (cfr. Helgason [7] p. 89, ex 8):

$$\forall \gamma \in G: \ [X, Y]_\gamma = \lim_{t \longrightarrow 0} \dot{\sigma}(t)$$

where

$$\sigma(t) = \psi_{-\sqrt{t}}(\varphi_{-\sqrt{t}}(\psi_{\sqrt{t}}(\varphi_{\sqrt{t}}(\gamma))))$$

and φ, ψ are integralcurves of the vectorfields X en Y.
But using the fact that the integralcurves of L and R are the left- (resp. right-) cosets of $\exp(tL_e)$ and $\exp(tR_e)$ we get:

$$\forall t \in I\!R: \ \sigma(t) = exp(-\sqrt{t}R_e)exp(\sqrt{t}R_e)\gamma exp(\sqrt{t}L_e)exp(-\sqrt{t}L_e) = \gamma$$

□

So we come to the following conclusion:

Proposition 12

If G is a Liegroup and $L_i, R_i (i = 1, \dots, n)$ are the left- (resp. right) invariant vectorfields corresponding to a basis $v_1, \dots, v_n$ of T_eG.
Then there are constants C_{ij}^k such that:

$$\begin{aligned}
&\text{(i)} \ \ [L_i, L_j] = C_{ij}^k L_k \\
&\text{(ii)} \ \ [R_i, R_j] = -C_{ij}^k R_k \\
&\text{(iii)} \ \ [L_i, R_j] = 0 \quad \forall i, j
\end{aligned}$$

□

To be able to define Brownian motion on the Liegroup G we must specify a Riemannian metric on the manifoldstructure of G. There are of course many ways to do this but we can only expect meaningful results if we exploit the underlying algebraic structure. We will therefore concentrate on metrics which are invariant under left or right translation and consequently compatible with the groupactions. So in order to define the inner product of two tangentvectors at an arbitrary $\gamma \in G$, we use either left- or right-translation to transport these vectors to the identity $e \in G$ and evaluate them there.

Therefore every Euclidean inner product $<,>$ on $T_eG \cong g$ will give rise to both a left- and right-invariant metric (g_L or g_R) on G:
if $X_\gamma, Y_\gamma \in T_\gamma G$:

$$g_L(X_\gamma, Y_\gamma) := < d\lambda_{\gamma^{-1}}(X_\gamma), d\lambda_{\gamma^{-1}}(Y_\gamma) > \tag{2.43}$$

$$g_R(X_\gamma, Y_\gamma) := < d\rho_{\gamma^{-1}}(X_\gamma), d\rho_{\gamma^{-1}}(Y_\gamma) > \tag{2.44}$$

A metric which is both left- and right-invariant is called bi-invariant. By its very definition, if X and Y are LIVFs then $g_L(X_\gamma, Y_\gamma)$ will be independent of γ and a similar result holds for g_R. This shows that an equivalent (but more elegant) way of writing down left-invariant metrics is given:

$$g_L = g_{ij}\theta^i \otimes \theta^j \tag{2.45}$$

where (g_{ij}) is a symmetric, positive definite, <u>constant</u> matrix and $\theta^1, \ldots, \theta^n$ are the dual 1-forms associated with a basis $L_1, \ldots, L_n$ of LIVF i.e.:

$$\forall i,j : \quad \theta^i(L_j) = \delta^i_j$$

Of course a similar statement holds for a right-invariant metric. For the sake of convenience we introduce the "inverse" metric corresponding to (2.45):

$$g_L^{-1} = g^{ij} L_i \otimes L_j \tag{2.46}$$

where $(g^{ij}) = (g_{ij})^{-1}$.
Notice that $g(g^{-1}) = \dim G$

The advantage of characterizing (left) invariant metrics as in equation (2.45) lies in the fact that on matrix-Liegroups there is a simple and explicit way to construct left- or right-invariant 1-forms:

Proposition 13 (Sattinger & Weaver [24] p. 90)
Let $(x^1, \ldots, x^n)$ be a smooth system of local coordinates on a domain V of a matrix-Liegroup and $B(x_1, \ldots, x_n)$ the corresponding matrix element.
Then

$$\Omega_L = B^{-1}dB \quad (\text{resp. } \Omega_R = (dB)B^{-1})$$

is a matrix of n linearly independent left (resp. right) invariant one-forms on V.

□

Example

Consider the affine group :

$$\mathbb{G} = \{\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mid y > 0, x \in \mathbb{R}\}$$

Take $B = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}$,

then:

$$\begin{aligned}
\Omega_L &= B^{-1}dB = \frac{1}{y}\begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix}\begin{pmatrix} dy & dx \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{dy}{y} & \frac{dx}{y} \\ 0 & 0 \end{pmatrix} \\
\Omega_R &= (dB)B^{-1} = \frac{1}{y}\begin{pmatrix} dy & dx \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix} = \begin{pmatrix} \frac{dy}{y} & \frac{ydx - xdy}{y} \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Therefore:

$$\theta^1 = \frac{dx}{y}, \ \theta^2 = \frac{dy}{y}$$

are left-invariant 1-forms, whereas:

$$\omega^1 = \frac{ydx - xdy}{y} \quad \omega^2 = \frac{dy}{y}$$

are right-invariant 1-forms.
From this we can construct left- and right- invariant metrics:

$$\begin{aligned}
g_L &= \theta^1 \otimes \theta^1 + \theta^2 \otimes \theta^2 = \frac{dx^2 + dy^2}{y^2}, \\
g_R &= \omega^1 \otimes \omega^1 + \omega^2 \otimes \omega^2 = \frac{y^2dx^2 - 2xydxdy + (1 + x^2)dy^2}{y^2}.
\end{aligned}$$

For the sake of completeness we point out that the representation equations (2.45 and 2.46) make it obvious that RIVFs are Killingfields for the left-invariant metric and vice versa, for if R is a right-invariant vectorfield then it is easy to compute the Lie derivative

$$\begin{aligned}
\mathcal{L}_R(g^{-1}) &= g^{ij}\mathcal{L}_R(L_i \otimes L_j) \\
&= g^{ij}(\mathcal{L}_R(L_i) \otimes L_j + L_i \otimes \mathcal{L}_R(L_j)) \\
&= g^{ij}([R, L_i] \otimes L_j + L_i \otimes [R, L_j]) \\
&= 0 \quad \text{(because of (iii) in prop (2))}
\end{aligned}$$

At this point, two questions immediately suggest themselves;
firstly: is there a natural choice of inner product on g and
secondly: given a particular inner product on g, it is possible to extend this over the rest of the Liegroup in such a way that the resulting metric is bi-invariant (i.e.: both left- and right-invariant). In order not to break the continuity of this résumé we will defer the discussion of these problems till the end of this paragraph, and for the moment assume that the Liegroup G has been equipped with an (arbitrary) left- (or right-) invariant Riemannian metric g.

Definition 1:

A stochastic process X_t on a Liegroup G equipped with a left (or right) invariant metric g is called a left (or right) Brownian motion on G if it is a Brownian motion on the Riemannian manifold (G, g).

Notice that this definition is more restrictive than the one used in Rogers & Williams [22], p. 225, (35.2). This can most easily be seen by observing that for the additive Liegroup $(\mathbb{R}, +)$ equipped with the usual Euclidean metric, every X_t of the form

$$X_t = \sigma B_t + \mu t \quad (\sigma > 0, \mu \in \mathbb{R} \text{ const}, B_t = BM(\mathbb{R}))$$

would be a Brownian motion in the sense of Rogers & Williams, but in order to qualify under definition (1) we must impose the restrictions $\sigma = 1$ and $\mu = 0$.

Although the above definition makes the notion of a Brownian motion on a Liegroup precise, one interesting question remains to be solved: how does this tie in with

the idea (propounded at the beginning of this paragraph) that we should be able to construct the process by taking a diffusion on the Liealgebra g (which is a vectorspace) and using the exponential map to project this diffusion on G.
It is obvious that simply applying the exponential map $exp : g \longrightarrow G$ to a Brownian motion B_t on g will not do: in general exp is not surjective even if G is connected (eg. take $G = SL(2, \mathbb{R})$, cfr. Sattinger & Weaver[24] p.39 ex. 4).
Again it is quite instructive to look at what happens if we study the random walk approximations of the considered diffusions.
Suppose we start off with a Brownian motion B_t on g; since g is a (flat) vectorspace, its geodesics are straight lines.
Therefore the (geodesical) random walk approximation R_t of B_t will be a random walk on a straight-line grid in g. The exponential map will carry this grid (at least locally) to a grid of one-parameter subgroups of G (and their cosets), and hence $exp(R_t)$ is a randomwalk on a grid of one-parameter subgroups and their cosets.
On the other hand, the (geodesical) random walk corresponding to $BM(G)$ will live on a geodesical grid. In general however these two grids will be very different because one-parametersubgroups and geodesics emanating from e with the same velocity vector will not coincide (unless the metric is bi-invariant,cfr. O'Neill [14], p.304 ff.). Since – by definition – geodesics are the straight lines on the manifolds, we can reformulate this last sentence by saying that in general the one parameter subgroups of a Liegroup are curved. This suggests that exp (B_t) will be different from $BM(G)$ and that this is basically due to the fact that the oneparametergroups are curved.

Let us now cast the above considerations in a more precise form.
First of all we pick an arbitrary basis $v_1, \ldots, v_n$ of the Liealgebra $g \cong T_eG$ and we construct the corresponding LIVFs $L_1, \ldots, L_n$ on G.
Let $\beta_1, \ldots, \beta_n$ be independent $BM(\mathbb{R})$, then for every $v \in g \cong T_eG$ the process

$$A_t = \sum_{i=1}^{n} \beta_i(t)v_i + tv \tag{2.47}$$

is a drifting Brownian motion on g.
Now we want to map this diffusion to G to obtain a Brownian motion-process $\gamma(t)$. The obvious way to do this is by (formally) defining this process by means of a Stratonovich SDE:

$$\partial\gamma(t) = d\lambda_{\gamma(t)}(\partial A_t) \tag{2.48}$$

which says that the increment at γ_t is obtained by shifting the corresponding increment of $BM(g)$ using left-translation, thus capturing the idea of the exponential

map.
Substituting (2.47) into (2.48) we obtain (still formally):

$$\begin{aligned} \partial\gamma(t) &= d\lambda_{\gamma(t)}(\sum v_i(\partial\beta_i(t)) + v\partial_t) \\ &= \sum_i L_i\partial\beta_i(t) + L_v\partial_t \end{aligned} \tag{2.49}$$

when of course L_v is the LIVF generated by $v \in g$.

To make rigorous sense of (2.49) (and hence of (2.48)) we let both sides act on an arbitrary smooth (realvalued) function and hence we see that (2.49) means that:

$$\begin{aligned} \forall\, f \in C^\infty(G, \mathbb{R}) &\;: \\ \partial f(\gamma(t)) &= \sum_i (L_i f)(\gamma(t))\partial\beta_i + (L_v f)(\gamma(t))\partial t \end{aligned} \tag{2.50}$$

The corresponding generator is given by (cfr. Rogers & Williams [22] p. 185):

$$\mathcal{L} = \frac{1}{2}\sum_i L_i^2 + L_v \tag{2.51}$$

If the resulting process is to be $BM(G)$ then this generator has to be equal to half the Laplacian Δ associated with the left-invariant metric g_L determined by our choice of basis.
A straightforward calculation, using the connectioncoefficients K_{ij}^k for the left-invariant fields which are defined by

$$\nabla_{L_i} L_j = K_{ij}^k L_k \tag{2.52}$$

then shows that (cfr.Rogers & Williams [22], p.217):

$$\Delta = \sum_{i=1}^n L_i^2 - \sum_{i=1}^n (\sum_{k=1}^n K_{kk}^i) L_i$$

Hence: $\mathcal{L} = \frac{1}{2}\Delta$ <u>iff</u> $L_v = -\frac{1}{2}\sum_{i=1}^n(\sum_{k=1}^n K_{kk}^i)L_i$

Notice that since the connectioncoefficients K^i_{kk} are constant we can reformulate the above condition in terms of elements of $g = T_eG$

$$v = -\frac{1}{2}\sum_{i=1}^{n}(\sum_{k=1}^{n} K^i_{kk})v_i \tag{2.53}$$

So we have proved the following proposition:

Proposition 14 (cfr. Rogers & Williams [22] , eq. (35.57))
Let $v_1, \ldots, v_n$ be a basis for the Liealgebra $g \cong T_eG$ of the Liegroup G and $L_1, \ldots, L_n$ (resp. g_L) the corresponding LIVF (resp. left-invariant metric).
Then the process γ_t defined on G via (2.50) is $BM(G, g_L)$ if and only if:

$$v = -\frac{1}{2}\sum_i \sum_k K^i_{kk} v_i \tag{2.54}$$

□

All that remains to be done is to explain this additional constant drift in terms of the curvature of the one-parameter subgroups (at the identity $e \in G$).
This is easy if we realise that the one-parametersubgroups are integralcurves of the left-invariant vectorfields.
So let $q_k(t) \subset G$ be the subgroup generated by $v_k \in g$:

$$q_k(t) = \{exp\ tv_k \mid t \in \mathbb{R}\}$$

Then we see that the mean curvature of q_k (at e) is given by:

$$\begin{aligned} H_e(q_k) &= \sum_{i \neq k} g(\nabla_{L_k} L_k, L_i) L_i(e) \\ &= \sum_{i \neq k} g(\sum_j K^j_{kk} L_j, L_i) v_i \\ &= \sum_{i \neq k} \sum_j K^j_{kk} g(L_j, L_i) v_i \end{aligned}$$

Hence: $H_e(q_k) = \sum_i K^i_{kk} v_i$
(Notice that $K^k_{kk} = 0$).

Since the curvature of each different subgroup $q_1, \ldots, q_n$ gives rise to an extra drift term, we get the total drift by summing over all subgroups (the factor $\frac{1}{2}$ appears because the generator of Brownian motion is half the Laplacian):

$$v = \frac{1}{2}\sum_k H_e(q_k) = -\frac{1}{2}\sum_{k=1}^{n}(\sum_i K^i_{kk}v_i) = -\frac{1}{2}\sum_i(\sum_k K^i_{kk})v_i$$

which is exactly equal to (2.54).

Remark:

The relation between the constants K^i_{jk} and the structural constants C^i_{jk} of the Liegroup is given by (cfr.Rogers & Williams [22] eq. 34.74).

$$2K^i_{jk} = C^i_{jk} + C^j_{ik} + C^k_{ij}$$

Hence it follows that:

$$\sum_j K^i_{jj} = \sum_j C^j_{ij}.$$

2.3 Canonical and bi-invariant metrics on real Liegroups

In this paragraph we want to take up two questions which we left unanswered in the preceding section:

(i) is there a natural choice of inner-product on g;

(ii) given a particular innerproduct on g, is it possible to extend this over the rest of the Liegroup in such a way that the resulting metric is bi-invariant (i.e.: both left- and right-invariant).

To address these questions we have to introduce some more notation and results from standard Liegroup theory.
Recall that associated with every $\gamma \in G$ we have an inner automorphism

$$\sigma_\gamma : G \longrightarrow G \; : \; \xi \longmapsto (\lambda_\gamma \circ \rho_{\gamma^{-1}})(\xi) = \gamma\xi\gamma^{-1} \tag{2.55}$$

Identifying g and T_eG we see that the derivative of this map at e defines an automorphism of the Lie-algebra g:

$$\forall \gamma \in G : d\sigma_\gamma(e) : g \longrightarrow g$$

which we will denote by $Ad(\gamma) := d\sigma_\gamma(e)$

The mapping:

$$Ad : G \longrightarrow GL(g) : \gamma \longmapsto Ad(\gamma) \tag{2.56}$$

is called the adjoint representation of G on g and $Ad(G)$ the adjoint group of the Liealgebra g.
Notice that if G is a matrix-Liegroup then:

$$\forall \gamma \in G : Ad(\gamma)(X) = \gamma X \gamma^{-1} \quad \text{(matrix-multiplication)}$$

Taking the derivative of Ad at $e \in G$ we obtain the adjoint representation

$$ad := d(Ad)(e)$$

of the Liealgebra g on $gl(g)$ which makes the following diagramme commutative:

$$\begin{array}{ccc} g & \xrightarrow{ad} & gl(g) \\ \downarrow{\scriptstyle exp} & & \downarrow{\scriptstyle exp} \\ G & \xrightarrow{Ad} & GL(g) \end{array}$$

It is easy to prove (cfr. Poor [21]) that

$$\begin{aligned} ad : g &\longrightarrow gl(g) \\ X &\longmapsto ad(X) \quad : \quad g \longrightarrow g \\ & \qquad\qquad\qquad\quad y \longmapsto ad(X)(Y) = [X,Y] \end{aligned}$$

The importance of the adjoint group $Ad(G)$ becomes clear if we realise that a metric g on G is bi-invariant if and only if its restriction $<,>$ to the Liealgebra g is $Ad(G)$-invariant:
i.e. for all $X, Y \in g,\ \gamma \in G$:

$$< Ad(\gamma)X, Ad(\gamma)Y > = < X, Y > \tag{2.57}$$

Now, recall that a compact and connected (real) Liegroup has a (up to a constant factor) unique bi-invariant measure (the Haar-measure). Therefore, if the adjoint group $Ad(G)$ is compact, we can use its Haar-measure to 'average out' a given innerproduct on g over $Ad(G)$, thus obtaining a new innerproduct satisfying (2.57). It is therefore not surprising that we have the following result:

Theorem 15

A Liegroup G admits a bi-invariant metric if and only if the adjoint group $Ad(G)$ has compact closure in $GL(g)$.

A short and elegant proof can be found in (Poor[21], prop. 6.8, p.190). An alternative and more constructive demonstration of this theorem is given below:

Proof:

(i) Let us first assume that G supports a bi-invariant metric, or equivalently that there exists an innerproduct $<,>$ on g which is $Ad(G)$-invariant.
This means that $Ad(G)$ is a subgroup of $SO(g) < GL(g)$ where, of course, the orthogonality on g is defined with respect to the innerproduct $<,>$. Since $SO(g)$ is compact it follows that $Ad(G)$ has compact closure in $GL(g)$.

(ii) Conversely, let's assume that $Ad(G)$ has compact closure in $GL(g)$. For the sake of simplicity we pick a basis of g and identify the aforementioned groups with their matrixrepresentation. Since the maximal compact, connected subgroups of $Gl(g)$ are conjugacy-classes of $SO(g)$ (cfr. R. Gillmore [6], p. 199) it follows that maximal compact, connected subgroups can be written as:

$$P^{-1}SO(g)P$$

where $P \in GL(g)$.
This implies that if $Ad(G)$ has compact closure, it must be a subgroup of one of the above conjugacy-classes and hence for any $\gamma \in G$:

$$PAd(\gamma)P^{-1}$$

will be orthogonal.
Therefore:

$$(PAd(\gamma)P^{-1})^T(PAd(\gamma)P^{-1}) = I$$

or equivalently:

$$(Ad(\gamma))^T Q Ad(\gamma) = Q \tag{2.58}$$

where $Q = P^T P$ is a positive definite symmetric (constant) matrix.
But condition (2.58) states in fact that Q can be extended to a bi-invariant metric on G.

$\square$

Remarks:

1. From elementary grouptheory we know that all Liesubgroups of $SO(g) \cong SO(n)$ are compact. Hence, the fact that $Ad(G)$ is a subgroup of $SO(g)$ (or one of its conjugacyclasses) implies that $Ad(G)$ itself must be compact, which is a slight strengthening of theorem (12).

2. Since the kernel of the adjoint representation is equal to the centre of the group (cfr. Warner [26] 3.50)

$$ker\ Ad = Z(G) := \{\gamma \in G \mid \forall\ \alpha \in G : \alpha\gamma = \gamma\alpha\}$$

it follows that:

$$Ad(G) \cong G/Z(G)$$

This means that the following table from (M.L. Curtis [4] p. 103) gives us valuable information about the adjoint groups:

Group G	dimension (over $\mathbb{R}$)	Centre Z(G)
$U(n)$	n^2	$\{e^{i\theta}I \mid \theta \in \mathbb{R}\} \cong S^1$
$SU(n)$	$n^2 - 1$	$\{wI \mid w^n = 1\} \cong Z_n$
$SO(2n+1)$	$2n^2 + n$	$\{I\}$
$SO(2n)$ $(n \geq 2)$	$2n^2 - n$	$\{+I, -I\}$
$GL(n, \mathbb{R})$	n^2	$\mathbb{R}_o I$
$SL(n, \mathbb{R})$	$n^2 - 1$	$\begin{cases} \{I\} & \text{if } n \text{ is odd} \\ \{+I, -I\} & \text{if } n \text{ is even} \end{cases}$
$Sp(2n, \mathbb{R})^{(*)}$	$2n^2 + n$	$\{+I, -I\}$

(*) The symplectic group (over $\mathbb{R}$) is defined by:

$$Sp(2n, \mathbb{R}) = \{X \in \mathbb{R}^{2n \times 2n} \mid X^T J X = J\}$$

where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

The corresponding Lie-algebra is given by:

$$Sp(2n, \mathbb{R}) = \{\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix} \mid A, B, C \in \mathbb{R}^{n \times n}, B^T = B, C^T = C\}$$

From this list it is obvious that the rotationgroups $SO(n)$ and $SU(n)$ support a bi-invariant metric, whereas $SL(n, \mathbb{R})$ and $GL(n, \mathbb{R})$ don't.

3. In some cases an explicit representation of the adjoint group is called for. The following remarks provide us with a simple and efficient way to obtain such a representation. Pick a basis $v_1, \ldots, v_n$ of the Liealgebra g; to compute the matrixrepresentation of $Ad(G)$ with respect to this basis, we construct the corresponding LIVF L_i and RIVF R_i such that:

$$L_i(e) = v_i = R_i(e)$$

By construction, we have for each $\gamma \in G$:

$$\begin{aligned} L_i(\gamma) &= d\lambda_\gamma(e)(v_i) \\ &= d\lambda_\gamma(e)(d\rho_{\gamma^{-1}}(\gamma) \circ d\rho_\gamma(e))(v_i) \\ &= d\sigma_\gamma(\gamma)(R_i(\gamma)) \\ &= \sum_j a_{ij}(\gamma) R_j(\gamma) \end{aligned} \tag{2.59}$$

where, of course, $(a_{ij}(\gamma))$ is the matrixrepresentation of the linear map $d\sigma_\gamma(\gamma)$ with respect to the basis $R_i(\gamma)$. Again, we use the definition of invariant vectorfields to rewrite the above equation as:

$$d\lambda_\gamma(e)(v_i) = \sum_i a_{ij}(\gamma) d\rho_\gamma(e)(v_j)$$

and applying $d\rho_{\gamma^{-1}}(\gamma)$ at both sides we get:

$$d\sigma_\gamma(e)(v_i) = \sum_j a_{ij}(\gamma) v_j$$

from which it follows that $(a_{ij}(\gamma))$ as defined in (2.59) is the matrixrepresentation of

$$d\sigma_\gamma(e) \equiv Ad(\gamma)$$

with respect to the basis $v_1, \ldots, v_n$.
Notice that if we write down the equivalent formula of (2.59) for the corresponding dual invariant 1-forms θ^i and ω^j we get:

$$\theta^k(\gamma) = \sum_l b_{kl}(\gamma)\omega^l(\gamma) \tag{2.60}$$

Applying the 1-forms to the vectorfields yields:

$$\delta_i^k = \sum_j \sum_k a_{ij} b_{kl} \delta_j^l$$

Hence, $(b_{kl}(\gamma))$ is a matrixrepresentation of

$$(Ad(\gamma))^{-1} = Ad(\gamma^{-1}).$$

Introducing the notation

$$\theta(\gamma) = \begin{pmatrix} \theta^1(\gamma) \\ \cdot \\ \cdot \\ \cdot \\ \theta^n(\gamma) \end{pmatrix} \quad \text{and} \quad \omega(\gamma) = \begin{pmatrix} \omega^1(\gamma) \\ \cdot \\ \cdot \\ \cdot \\ \omega^n(\gamma) \end{pmatrix}$$

and identifying the linear transformation with its matrixrepresentation, we can rewrite equation (2.60) as:

$$\omega(\gamma) = Ad(\gamma)\theta(\gamma) \tag{2.61}$$

As an illustration let's have another look at the affine group

$$\mathbb{G} = \{\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R},\ y > 0\}$$

Earlier on (cfr p. 49) we computed a basis of invariant 1-forms:

<u>left:</u> <u>right:</u>

$$\theta^1 = \frac{dx}{y} \qquad \omega^1 = \frac{ydx - xdy}{y}$$

$$\theta^2 = \frac{dy}{y} \qquad \omega^2 = \frac{dy}{y}$$

Since they coincide at $e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ it follows from

$$\begin{pmatrix} \omega^1 \\ \omega^2 \end{pmatrix} = \begin{pmatrix} y & -x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix}$$

that we have the following matrixrepresentation:

$$Ad \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} y & -x \\ 0 & 1 \end{pmatrix}$$

Since $Z(\mathbb{G}) = 1$ this tallies with the observation

$$Ad(\mathbb{G}) \cong \mathbb{G}/Z(\mathbb{G}) = \mathbb{G}$$

◇

It is important to realise that in general, even if G supports a bi-invariant metric, there will be a large class of left-invariant metrics which are not right-invariant and vice-versa. The easiest way to see this is by observing that there is a bijective correspondence between the inner-products on g and the left-invariant metrics on G, but in order to generate a bi-invariant metric, the inner-product must be $Ad(G)$-invariant. So unless $Ad(G) = 1$, i.e. G is abelian, (recall that since G is connected, so is $Ad(G)$) there will be left-invariant metrics which fail to be bi-invariant.
On the other hand, if $Ad(G)$ is compact then there is the question of how many bi-invariant metrics there exist. Picking a basis in g and identifying transformations with their matrixrepresentations the above problem can be reformulated as: can we characterize the positive-definite symmetric matrices S which satisfy

$$Ad(\gamma)^T S Ad(\gamma) = S \qquad \text{for all } \gamma \in G \tag{2.62}$$

Let us first look at the simplest case: $Ad(G)$ is a conjugacy-class of $SO(g)$: i.e. there is an invertible matrix $P \in GL(g) = GL(n)$ such that:

$$Ad(G) = PSO(g)P^{-1} \tag{2.63}$$

Using (2.63) condition (2.62) can be reformulated as:

$$\forall\, U \in SO(n) : (PUP^{-1})^T S (PUP^{-1}) = S \tag{2.64}$$

Introducing the positive-definite symmetric matrix

$$R = P^T S P \tag{2.65}$$

we see that condition (2.62) becomes:
characterize all positive-definite,symmetric matrices R such that

$$\forall\, U \in SO(n) : U^T R U = R \tag{2.66}$$

Using the factorisation:

$$R = V^T D V$$

(V orthogonal, D positive-definite diagonal), condition (2.66) can be rewritten as:
characterise all D for which:

$$\forall\, U \ \in SO(n) : \ [D, U] = 0 \tag{2.67}$$

Condition (2.67) implies that $D = \lambda I$ for some $\lambda > 0$, and hence

$$R = \lambda I \tag{2.68}$$

Combining eqs (2.65) & (2.68) we see that

$$S = \lambda (P P^T)^{-1} \qquad (\lambda > 0) \tag{2.69}$$

This implies that (up to the scalar factor λ) S is unique, for if $Q \in GL(n)$ such that

$$Ad(G) = P SO(g) P^{-1} = Q SO(g) Q^{-1}$$

then it follows that

$$\forall\, U \in SO(n), \exists\, V \in SO(n) : \quad P U P^{-1} = Q V Q^{-1}$$

In particular it follows that if we denote $M = Q^{-1}P$ then:

$$\forall\, U \in SO(n) : (MUM^{-1})^T(MUM^{-1}) = I$$

or again:

$$\forall\, U \in SO(n) : \quad U^T(M^TM)U = M^TM$$

Observing that M^TM is both symmetric and positive-definite (M is invertible!) we get

$$M^TM = \alpha I \qquad\qquad \text{for some } \alpha > 0$$

or equivalently:

$$PP^T = \alpha QQ^T$$

This proves the uniqueness of S.

Finally, referring back to equation (2.67), we notice that

$$[D, V] = 0 \quad \text{implies} \quad [D, \lambda V] = 0 \qquad \forall \lambda \in \mathbb{R}$$

And since

$$\lambda V \in SO(n) \quad \text{iff} \begin{cases} \lambda = 1 & \text{if n odd} \\ \lambda = \pm 1 & \text{if n even} \end{cases}$$

it follows that if dim G is even, it suffices to check eq (2.67) for all $V \in SO(n)/\pm 1$. Thus we have proved:

Proposition 16:

If G is a real, connected Liegroup such that

$$Ad(G) \cong SO(g) \quad \text{or } Ad(G) \cong SO(g)/\pm 1$$

<u>then</u> G supports an (up to a scalar factor) unique bi-invariant metric.

□

The case where $Ad(G)$ is isomorphic to a nontrivial subgroup of $SO(g)$ seems more difficult.
However if we restrict our attention to the case of simple real Liegroups, it is possible to give explicit representrations of the bi-invariant metrics they support.
Recall that a Lie-group is simple iff its algebra is (cfr. Bourbaki [2], 9.8).

Moreover the bi-invariance-condition (2.57) on the Liegroup can be expressed in terms of the adjoint representation of the algebra as follows (cfr. Poor [21])

$$\forall X, Y, Z \in g :< ad(Z)X, Y > + < X, ad(Z)Y >= 0 \tag{2.70}$$

Since a complete classification of simple Lie-algebras is known, all we have to do is to determine how many solutions of (2.70) each class supports.
We restrict ourselves to the non-exceptional algebras:

<u>Non-exceptional simple real Lie-algebras</u>

Type	Algebra g	Group G	Ad(G) ≅ G/Z(G)	Supports bi-inv. metric
A_{n-1}	$Sl(n)$	$SL(n, \mathbb{R})$	non-compact	no
B_n	$SO(2n+1)$	$SO(2n+1)$	compact	yes
C_n	$Sp(2n)$	$Sp(2n)$	non-compact	no
D_n	$SO(2n)$	$SO(2n)$	compact	yes

From this table it is obvious that only the rotationgroups $SO(n)$ are of interest to us. So let's study them in more detail.

Since $SO(2) \cong S^1$ is one-dimensional and abelian its structure is trivial.
On the other hand, for $n > 2$:

$$Ad(G) \cong SO(n)/Z(SO(n)) = \begin{cases} SO(n) & \text{if n odd} \\ SO(n)/\pm 1 & \text{if n even} \end{cases}$$

whereas

$$SO(g) = SO(so(n)) \cong SO(\frac{n(n-1)}{2})$$

In particular this means that:

$$Ad(SO(3)) \cong SO(3) \cong SO(so(3))$$

and hence it follows from proposition 16 that (up to a scalar constant) the bi-invariant metric on $SO(3)$ is unique.
However, for $n > 3$

$$Ad(SO(n)) \subset SO(so(n))$$

and it is therefore impossible to apply the above proposition.
To get round this difficulty we proceed as follows.

First of all we notice that because of the linearity of (2.70) it suffices to let Z range over a basis in g.

On the Lie-algebra

$$so(n) = \{A \in I\!R^{n \times n} \mid A + A^T = 0\}$$

we use the natural (lexicographically ordered) basis:

$$\{H_{ij} \in I\!R^{n \times n} \mid 1 \leq i < j \leq n\} \tag{2.71}$$

where $H_{ij} = E_{ij} - E_{ji}$

and the matrices E_{ij} constitute the standardbasis on $I\!R^{n \times n}$:

$$(E_{ij})_{kl} = \begin{cases} 1 & \text{if } i = k, j = l \\ 0 & \text{otherwise} \end{cases}$$

A straightforward calculation then shows that:

$$ad(H_{ij})(H_{kl}) = [H_{ij}, H_{kl}] = \delta_{jk} H_{il} - \delta_{jl} H_{ik} - \delta_{ik} H_{jl} + \delta_{il} H_{jk} \tag{2.72}$$

which enables us to make the following simple but relevant observations:

(i) Since $ad(X)(Y) = [X, Y]$ it follows immediately that

$$y \in ker\ ad(X) \quad \underline{\text{iff}} \quad X \in ker\ ad(Y) \tag{2.73}$$

(ii) Since $i \neq j$ and $k \neq l$, inspection of (2.72) reveals that

$$ad(H_{ij})(H_{kl}) = 0 \quad \underline{\text{iff}} \quad \begin{array}{l} \text{either} \{k, l\} = \{i, j\} \\ \text{or } \{k, l\} \cap \{i, j\} = \emptyset \end{array} \tag{2.74}$$

Consequently,

$$H_{kl} \notin ker\ ad(H_{ij}) \ \underline{\text{iff}}\ \sharp[\{k, l\} \cap \{i, j\}] = 1 \tag{2.75}$$

(iii) Since $H_{ij} = -H_{ji}$ it is obvious that the most general solution of equation (2.75) is given by:

$$k = i \quad \text{and} \quad l \neq j \quad (\text{and of course } \ l \neq i, \text{ for } l \neq k) \tag{2.76}$$

In this case we get:

$$ad(H_{ij})(H_{kl}) = ad(H_{ij})(H_{il}) = -H_{jl}$$

but also:

$$ad(H_{ij})(H_{jl}) = H_{il} = H_{kl}$$

which shows that, with respect to the standardbasis, the matrixrepresentation of $ad(H_{ij})$ is skewsymmetric.

Let us now return to our original problem: we want to characterize all inner-products on the Liealgebra $so(n)$ which satisfy condition (2.70). Representing the inner-product by a positive-definite, symmetric matrix P and identifying $ad(Z)$ and its matrixrepresentation with respect to the basis (2.71)

$$ad(Z)(H_\alpha) = \sum_\beta (ad(Z))_\alpha^\beta H_\beta$$

conditions (2.70) becomes: determine all positive-definite symmetric matrices P such that:

$$\forall\, 1 \leq i < j \leq n : ad(H_{ij})^T P + P ad(H_{ij}) = 0 \tag{2.77}$$

or equivalently (using the skew-symmetry established in (iii)):

$$\forall\, 1 \leq i < j \leq n : [ad(H_{ij}), P] = 0 \tag{2.78}$$

Moreover, since P is both symmetric and positive-definite we know:

$$\begin{aligned} P &= VDV^T && \text{where } V \text{ orhtogonal}, D = \text{diag}(d_i \mid d_i > 0) \\ &= Ve^\Lambda V^T && \qquad\qquad \Lambda = \text{diag}(\lambda_i \mid \lambda_i \in I\!R) \\ &= e^{V\Lambda V^T} && \\ &= e^S && \text{where } S \text{ is symmetric} \end{aligned}$$

It is then easy to prove (see lemma 18) that (2.78) is equivalent to:

Condition (S):
characterize all symmetric matrices S such that

$$\forall\, 1 \leq i < j \leq n : [ad(H_{ij}),\ S] = 0$$

To solve this problem we observe that every matrix S satisfying condition (S) must leave the kennel of each $ad(H_{ij})$ $(1 \leq i < j \leq n)$ invariant.
For if $x \in ker\ ad(H_{ij})$, then:

$$ad(H_{ij})S(x) = S(ad(H_{ij})(x)) = 0$$

and hence

$$S(x) \in ker\ ad(H_{ij})$$

This puts severe restrictions on the structure of S. To see this we proceed as follows: first of all we observe that the intersection of two invariant subspaces is itself invariant.
Secondly, it is trivial that (using the multi-index $\alpha = (ij)$)

$$H_\alpha \in ker\ ad(H_\alpha).$$

The idea now is to get the smallest invariant subspace containing a particular H_α by intersecting all the kernels to which H_α belongs:

$$D_\alpha := \bigcap_\beta \{ker\ ad(H_\beta) \mid H_\alpha \in ker\ ad(H_\beta)\} \tag{2.79}$$

Recall that the low-dimensional cases $n = 2$ and $n = 3$ have already been looked at.
We therefore concentrate on the case $n \geq 4$ for which we will prove the following result: except when $n = 4$, the intersection (2.79) is equal to the one-dimensional vectorspace generated by H_α. For convenience we introduce the incidence-matrix L which is defined by:
$\forall\, 1 \leq \alpha, \beta \leq n(n-1)/2$

$$L_{\alpha\beta} = \begin{cases} 1 & \text{if } H_\beta \in ker\ ad(H_\alpha) \\ 0 & \text{otherwise} \end{cases}$$

Notice that because of (2.73) L is symmetric.
Moreover, from (2.72) it is easy to see that:

$$\forall\, 1 \leq i < j \leq n : \ \dim\, Im\ ad(H_{ij}) = 2(n-2)$$

from which it follows that

$$\forall\, 1 \leq i < j \leq n : \ \dim\, ker\ ad(H_{ij}) = \frac{n(n-1)}{2} - 2(n-2) = \frac{n^2-5n+8}{2}$$

This implies that in the incidence-matrix L each row (and hence each column) contains the same number of ones (and hence zeros). Let us now assume that β and γ are two different multi-indices such that

$$\forall\, \alpha = 1, \ldots, n(n-1)/2 \ : \ L_{\alpha\beta} = L_{\alpha\gamma}$$

This means that:

$$\forall\, \alpha : H_\beta \in ker\ ad(H_\alpha) \ \ \underline{\text{iff}} \ \ H_\gamma \in ker\ ad(H_\alpha) \tag{2.80}$$

In terms of equation (2.79) this would mean that both D_β and D_γ contain (at least) two basisvectors.
We now proceed to show that if $n \geq 5$ then (2.80) can only hold if $\beta = \gamma$; for suppose that:

$$\begin{aligned} \beta &= (i,j) \qquad (i<j) \\ \gamma &= (k,l) \qquad (k<l) \\ & \quad \text{and} \beta \neq \gamma. \end{aligned}$$

We then have the following two possibilities:

(i) $\{i,j\} \cap \{k,l\} = \emptyset$;
since $n \geq 5$ it is possible to pick $p \notin \{i,j,k,l\}$ and from (2.74) it follows that:

$$H_{ip} \in ker\ ad(H_{kl}) \quad \text{but } H_{ip} \notin ker\ ad(H_{ij});$$

(ii) $\sharp[\{i,j\} \cap \{k,l\}] = 1$;
without loss of generality we can assume: $i = k, j \neq l$; then again because of (2.74) we have:

$$H_{ij} \in ker\ ad(H_{ij}) \quad \text{but } H_{ij} \notin ker\ ad(H_{kl})$$

This shows that if $n \geq 5$, kernels of different maps $ad(H_{ij})$ are different spaces, or equivalently, that L has no two identical rows. Because of the symmetry of L this implies that likewise L has no two identical columns. This observation, combined with the fact that all columns contain the same amount of zeros and ones proves that each D_α (as defined by (2.79)) is the one-dimensional vectorspace spanned by H_α.
But if each basisvector generates a one-dimensional invariant subspace of S, then S must be a diagonal matrix.
Substituting this back into condition (S) yields that in fact S is a multiple of the identity and hence the most general solution of (2.78) is given by:

$$P = \lambda I \qquad (\lambda > 0)$$

The above argument does not apply to the case $n = 4$.
In fact using the lexicographically ordered multi-indices:

$$1 = (1,2) \quad 2 = (1,3) \quad 3 = (1,4) \quad 4 = (2,3) \quad 5 = (2,4) \quad 6 = (3,4)$$

a simple inspection of (2.73) yields:

$$\begin{aligned} ker\ ad(H_1) &= span[H_1, H_6] = ker\ ad(H_6) \\ ker\ ad(H_2) &= span[H_2, H_5] = ker\ ad(H_5) \\ ker\ ad(H_3) &= span[H_3, H_4] = ker\ ad(H_4) \end{aligned}$$

implying that S has a 'crossed-diagonal' structure. Using this it is straightforward to figure out that the most general solution of condition (S) is of the form:

$$S = aI + bJ \qquad (a, b \in \mathbb{R}) \tag{2.81}$$

where I is (6×6)-identity and

$$J = \begin{pmatrix} 0 & K \\ K & 0 \end{pmatrix} \quad \text{where } K = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{2.82}$$

Notice that since $K^2 = I$ it follows that $J^2 = I$ and hence we see that the corresponding positive-definite symmetric matrix is given by:

$$P = exp(aI + bJ) = e^a(cosh(b)I + sinh(b)J) \qquad (a, b \in \mathbb{R}) \tag{2.83}$$

Summarizing all this we get the following result:

Proposition 17

Consider the orthogonal groups $SO(n) \quad (n \geq 2)$:

If $\underline{n \neq 4}$, then the bi-invariant metric on $SO(n)$ is (up to a constant multiplicative factor) unique (and equal to the Killing-metric; see next paragraph)

If $\underline{n = 4}$, then there is a 2-parameter family of bi-invariant metrics the matrix of which with respect to the (ordered) basis $\{H_{12}, H_{13}, H_{14}, H_{23}, H_{24}, H_{34}\}$ is given by:

$$P = e^a(cosh(b)I + sinh(b)J)$$

$(I, J$ defined by (2.82))

□

We conclude this paragraph by proving a little technical lemma which we needed in prop. 17.

Lemma 18

Let $A, S \in \mathbb{R}^{n\times n}$ and S symmetric,
then:

$$[A, S] = 0 \Longleftrightarrow [A, e^S] = 0$$

Proof:

As the first implication is trivial, we need only consider the second.
Since S is symmetric we can decompose it as

$$S = U\Lambda U^T \tag{2.84}$$

(Λ diagonal, U orthogonal) and hence:

$$0 = [A, e^S] = [A, Ue^\Lambda U^T] = [\tilde{A}, e^\Lambda] \tag{2.85}$$

where

$$\tilde{A} = U^T AU$$

The diagonality of Λ reduces (2.85) to a particularly simple form:

$$\forall\, i,j = 1 \ldots n: \quad \tilde{a}_{ij}(e^{\lambda_i} - e^{\lambda_j}) = 0 \tag{2.86}$$

Hence,

(i) either $\tilde{a}_{ij} = 0$ from which it follows that

$$\tilde{a}_{ij}(\lambda_i - \lambda_j) = 0$$

(ii) or $\tilde{a}_{ij} \neq 0$ and then $e^{\lambda_i} - e^{\lambda_j} = 0$ which again implies that $\tilde{a}_{ij}(\lambda_i - \lambda_j) = 0$.

Thus (2.85) implies

$$[\tilde{A}, \Lambda] = 0$$

whence

$$[A, S] = 0$$

□

Let us now tackle the other question we want to answer in this paragraph: is there an inner product on g which is basically determined by the algebraic structure of g. Again it is the adjoint representation that provides the solution; this time it is used to introduce the Killing form of the Lie-algebra.

This is a symmetric bilinear form on $g \times g$ defined by:

$$K(X,Y) := Tr(adX \circ adY)$$

where the trace is obtained by looking at an arbitrary matrixrepresentation of the endomorphism. Since the Killingform is a symmetric bilinear form which is unique and completely determined by the algebraic structure of the Lie-algebra, it suggests itself as the natural candidate for a metric on g (and hence on G). What makes it even more attractive is the fact that it is $Ad(G)$-invariant. This can be seen as follows: since

$$Ad(\gamma) \in Aut(g) := \{F \in GL(g) \mid F[X,Y] = [F(X), F(Y)], \forall\ X,Y \in g\}$$

we see that: $\forall\ \gamma \in G,\ \forall\ X,Y \in g$:

$$\begin{aligned} ad(Ad(\gamma)X)(Y) &= [Ad(\gamma)(X), Y] \\ &= [Ad(\gamma)(X), Ad(\gamma)Ad(\gamma^{-1})(Y)] \\ &= Ad(\gamma)[X, Ad(\gamma^{-1})(Y)] \\ &= Ad(\gamma)(ad(X)(Ad(\gamma^{-1})(Y)) \\ &= Ad(\gamma) \circ ad(X) \circ Ad(\gamma^{-1})(Y) \end{aligned}$$

Therefore:

$$\begin{aligned} K(Ad(\gamma)(X), Ad(\gamma)(Y)) &= Tr(ad(Ad(\gamma)(X) \circ ad(Ad(\gamma)(Y)) \\ &= Tr(Ad(\gamma) \circ ad(X) \circ Ad(\gamma^{-1}) \circ Ad(\gamma) \circ ad(Y) \circ Ad(\gamma^{-1})) \\ &= Tr(ad(X) \circ ad(Y)) \\ &= K(X,Y) \end{aligned}$$

The only problem is that, in general, K isn't necessarily non-degenerate, let alone positive-definite. The following famous result by Cartan settles this issue.

Theorem 19 **(Cartan's Criterion)**

The Killing-form of a Lie-algebra g is non-degenerate
<u>iff</u>
the Lie-algebra g is semi-simple.

□

(For a proof and more details see eq. Sagle & Walde [23] p.236 ff.).

If we are assured of the non-degeneracy of K we can define a positive-definite inner-product on g as follows:

$$\begin{cases} < X,X >:= |K(X,X)| \\ < X,Y >:= \frac{1}{2}(< X+Y,X+Y > - < X,X > - < Y,Y >) \end{cases}$$

A special case of the above result is obtained when we restrict our attention to <u>compact</u> semi-simple Lie-algebras (a compact Liealgebra is the Liealgebra of a compact Liegroup). In that case we can apply the following result:

Theorem 20 (Weyl)(cfr. Sattinger & Weaver [24] thm II.3)

A (real) Liegroup is compact <u>iff</u> its Killingform is negative-definite.

□

Moreover, since the Liegroup G is compact, so is its adjoint representation $Ad(G)$ and therefore we can extend the Killingform to a bi-invariant metric on G.
Hence we could define the standard metric on a (connected) compact, semi-simple Liegroup to be minus (the extension of) the Killingform of its Lie-algebra.

In view of the above theorems it becomes important to be able to compute the Killingform. Since it is a bilinear form it suffices to calculate its matrix with respect to a basis $A_1, \ldots, A_n$ of g; so let

$$K_{ij} = K(A_i, A_j) = Tr(adA_i \circ adA_j)$$

then applying the transformation to the basisvectors yield:

$$\begin{aligned}(adA_i \circ adA_j)(A_k) &= adA_i([A_j, A_k]) \\ &= \sum_l C^l_{jk}[A_i, A_l] \\ &= \sum_l \sum_p C^l_{jk} C^p_{il} A_p\end{aligned}$$

Therefore:

$$Tr(adA_i \circ adA_j) = \sum_k \sum_l C^l_{jk} C^k_{il}$$

or again:

$$k_{ij} = \sum_k \sum_l C^k_{il} C^l_{jk}$$

Recall that:

$$(adA_i)(A_j) = [A_i, A_j] = \sum_k C^k_{ij} A_k$$

So that:

$$C_i := ((C_i)_{jk}) = (C^k_{ij})$$

is the matrix of adA_i with respect to the basis $A_1, \ldots, A_n$.
Hence if all adA_i are skewsymmetric, then:

$$C^k_{ij} = -C^j_{ik}$$

and hence:

$$k_{ij} = \sum_k \sum_l C^k_{il} C^l_{jk} = -\sum_k \sum_l C^k_{il} C^k_{jl}$$

and therefore:

$$k_{ii} = -\sum_{k,l}(C_{il}^{k})^2 \leq 0.$$

Appendix: Some definitions in Lie-algebra theory

A Lie-algebra is **simple** if it is non-abelian and has no proper ideals.
A Lie-algebra is **semi-simple** if it is the **direct sum** of simple ideals,i.e.:

$g = g_1 \oplus g_2 \oplus \ldots \oplus g_k$
where the above equality is a vectorspace equality
plus : $\forall i, j : \; [g_i, g_j] = 0$

A Liegroup is **(semi)simple** iff its Lie-algebra is (semi)simple.

Proposition (Poor [21] p 191)
If $g = g_1 \oplus \ldots \oplus g_k$ is a simple-ideal-decomposition of a semi-simple Lie-algebra, then:

(i) This ideal-decomposition is unique (up to a permutation).

(ii) Every ideal of g is the direct sum (unique up to order) of some of the g_i.

(iii) Every non-zero ideal of g is the direct sum complement of an ideal of g.

(iv) The derived algebra $[g, g]$ is equal to g.

(v) The centre $z(g) = 0$.

(vi) dim $g \geq 3$.

Chapter 3

Applications of the Submersion-theorem

In this chapter we will look at some situations where processes arise in a natural way as the submersion of a Brownian motion on a higher-dimensional manifold. The most important example of this occurs when we are looking at the (transitive) action of a Liegroup on a manifold.

3.1 Actions of a Liegroup on a manifold

Let M be a (finite-dimensional) manifold and G a Liegroup.

Definition 1

A C^∞-map

$$\begin{aligned} \theta \ : \ & G \times M \longrightarrow M \\ & (\gamma, x) \longmapsto \theta(\gamma, x) \equiv \theta_\gamma(x) \equiv \gamma \cdot x \end{aligned}$$

is called a (smooth) **left-action** of G on M if:

(i) $\forall x \in M : \theta(e, x) = x$ $\quad$ (e is the identity element of G)

(ii) $\forall x \in M, \forall \gamma_1, \gamma_2 \in G : \theta_{\gamma_1}(\theta_{\gamma_2}(x)) = \theta_{\gamma_1\gamma_2}(x)$

The orbit generated by $x \in M$ is defined as:

$$G_x := \{\theta_\gamma(x) |\ \gamma \in G\}$$

The action θ will be called a (smooth) **right-action** if condition (ii) is changed to:
(ii)' : $\forall x \in M, \forall \gamma_1, \gamma_2 \in G : \theta_{\gamma_1}(\theta_{\gamma_2}(x)) = \theta_{\gamma_2\gamma_1}(x)$

From now on, unless stated otherwise, action will mean: (smooth) left-action.

Definition 2:

The action θ of G on M is called

(i) **transitive** if:

$$\forall x, y \in M, \exists\ \gamma \in G : y = \theta_\gamma(x)$$

(ii) **regular** if:

1. all the orbits have the same dimension as submanifolds of M;
2. for each point $x \in M$, there exist arbitrarily small neighbourhoods U of x with the property that each orbit of G intersects U in a pathwise connected subset.

Notice that if θ is transitive then there is only one orbit which coincides with the whole manifold.
Consider as an example the action θ of the rotationgroup $SO(2, I\!R)$ on $I\!R^2$

$$\forall R \in SO(2, I\!R), \forall \begin{pmatrix} x \\ y \end{pmatrix} \in I\!R^2\ :\ \theta_R \begin{pmatrix} x \\ y \end{pmatrix} := R \begin{pmatrix} x \\ y \end{pmatrix}$$

The orbits are given by:

$$G_{(x,y)} = \begin{cases} \{(0,0)\} & \text{if } (x,y) = (0,0) \\ \{R\begin{pmatrix} x \\ y \end{pmatrix} \mid R \in SO(2)\} & \text{if } (x,y) \neq (0,0) \\ \text{(circle with centre } (0,0) \text{ and radius } \sqrt{x^2+y^2}) & \end{cases}$$

Hence this action is neither transitive nor regular.
However the same action θ is regular on $M = \mathbb{R}^2 \setminus \{(0,0)\}$.
A less trivial example of a non-regular action is obtained by looking at the following action of the additive Lie-group $\mathbb{R}$ on the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$: choose $0 < \alpha < 1$ irrational and let

$$\theta : \mathbb{R} \times T^2 \longrightarrow T^2$$

be defined by

$$\theta_t(x,y) = (x + \alpha t, y + t)$$

Then it is well-known that each orbit will intersect every open set infinitely often.

If a Liegroup G acts on a manifold M through an action θ, then it establishes an equivalencerelation on M, the equivalence - classes of which are the orbits. If θ is non-transitive then the orbitspace:

$$M/G := \{G_x \mid x \in M\}$$

will be a manifold in its own right provided that the action is regular.
If however θ is transitive, then M/G is a singleton indicating that G is 'larger' than M and that there exists a natural projection $\pi : G \longrightarrow M$. In that case the manifold M is called homogeneous (with respect to θ).

The following propositions make these notions precise.

Theorem 1 (cfr. Olver[13] p. 214)

Let M be a smooth n-dimensional manifold. Suppose G is a Liegroup which acts regularly on M with k-dimensional orbits. Then there exists a smooth $(n-k)$-dimensional manifold (possibly non-Hausdorff !) together with a projection

$$\pi : M \longrightarrow M/G$$

which satisfies the following properties:

(i) The projection π is a smooth map between manifolds;

(ii) The points x and y lie in the same orbit of G in M if and only if $\pi(x) = \pi(y)$;

(iii) If g denotes the Lie-algebra of G, then for every $x \in M$, the linear map

$$d\pi(x) : T_x M \longrightarrow T_{\pi(x)}(M/G)$$

is surjective, and the kernel is isomorphic to g.

□

Remarks:

1) Notice that the three results (i), (ii), (iii) imply that the projection

$$\pi : M \longrightarrow M/G$$

is in fact a submersion.

2) The quotientmanifold M/G need not be Hausdorff.For a simple explicit example see Olver[13], p. 217. However we can always remove the non-Hausdorff singularities in M/G by restricting attention to a smaller open submanifold $\tilde{M} \subset M$ such that $\tilde{M}/G$ is an open Hausdorff submanifold of M/G.

In the case the action is transitive we have the following results: first we introduce the set of all left cosets corresponding to a closed subgroup H of G :

$$G/H := \{\gamma H | \ \gamma \in G\}$$

with the canonical projection

$$\pi : G \longrightarrow G/H : \gamma \longrightarrow \gamma H$$

From the next proposition it then follows that G/H is a manifold in its own right:

Proposition 2 (cfr. B. O'Neill [14] p. 307)

If H is a closed subgroup of the Liegroup G, then there is a unique way to make G/H a manifold so that the projection $\pi : G \longrightarrow G/H$ is a submersion.

$\square$

Finally we show that we can choose the subgroup H in such a way that the coset-manifold G/H is diffeomorphic to M.
To this end we introduce the isotropy-group H_x of an element $x \in M$:

$$H_x := \{\gamma \in G | \ \theta_\gamma(x) = x\}$$

Since the action θ is smooth (and hence a fortiori continuous) H_x is closed in G and hence the coset-manifold G/H_x is well-defined. Moreover the transitivity of θ entails that isotropygroups corresponding to different points are isomorphic. We then have the following proposition:

Proposition 3 (cfr.B. O'Neill [14] p. 307)

Let $\theta : G \times M \longrightarrow M$ be a transitive action of the Liegroup G on the manifold M and let H be its isotropy-group at a point $x \in M$
Then the natural map:

$$\begin{aligned} j : G/H &\longrightarrow M \\ \gamma H &\longrightarrow \theta_\gamma(x) \equiv \gamma \cdot x \end{aligned}$$

is a diffeomorphism.
Hence, in particular, the projection

$$\begin{aligned} \pi : G &\longrightarrow M \\ \gamma &\longrightarrow \gamma \cdot x \end{aligned}$$

is a submersion.

□

Remark

Notice that if we consider the subgroup H as a Liegroup in its own right, then we can construct a natural left- and right- action of H on G (considered as a manifold) by using the group-multiplication:

left-action: $\theta^L : H \times G \longrightarrow G : (\alpha, \gamma) \longrightarrow \theta^L_\alpha(\gamma) := \alpha\gamma$

right-action: $\theta^R : H \times G \longrightarrow G : (\alpha, \gamma) \longrightarrow \theta^R_\alpha(\gamma) := \gamma\alpha$

It is obvious that the orbits of the left-action are the right-cosets $H\gamma \in H\backslash G$ whereas the right-action gives rise to the left-cosets $\gamma H \in G/H$.

In most of the cases we will be interested in, the manifold M under consideration will be Riemannian and the Liegroup G will be group of isometries acting on M. In order to be able to use the results established in the preceding chapter we need to have some information on the mean curvature of the fibres of the submersion:

$$\pi : M \longrightarrow M/G$$

The following proposition gives us a sufficient condition for the mean curvature to be constant along the orbits.

Proposition 4

If $\theta : G \times M \longrightarrow M$ is a regular action on the Riemannian manifold (M, g) such that $\forall \gamma \in G : \ \theta_\gamma$ is an isometry of M
then the mean curvature vectorfield is constant along the orbits; more precisely: if $x, y \in G_x$ and $\gamma \in G$ such that $y = \theta_\gamma(x)$ and $H(x)$ is the mean curvaturevector of G_x at x,
then:

$$H(y) = d\theta_\gamma(H(x))$$

Proof:

Let x be an arbitrary point of M and consider the corresponding orbit G_x (assume: dim $G_x = k < n = dim\ M$).
To compute the mean curvature vector H at x we choose a small enough neighbourhood V of x in M and pick vectorfields $E_1, \ldots, E_n$ in V such that:

(i) $E_1, \ldots, E_n$ orthonormal;

(ii) $\forall\, p \in V$: if $p \in G_x$ then $E_1(p), \ldots, E_k(p)$ are tangent to G_x.

Then we have by definition:

$$\forall\, p \in (G_x \cap V) : H(p) = \sum_{j=k+1}^{n} g((\sum_{i=1}^{k} \nabla_{E_i} E_i)_p, E_j(p))E_j(p)$$

Next take an arbitrary $y \in G_x$ and let $\gamma \in G$ such that $y = \theta_\gamma(x)$. Because each θ_γ is an isometry of M we can get a set of orthonormal vectorfields $F_1, \ldots, F_n$ in a neighbourhood of y by applying the derivative

$$d\theta_\gamma : T_x M \longrightarrow T_y M$$

to $E_1, \ldots, E_n$; hence $F_i := d\theta_\gamma(E_i)$.
Clearly, in the neighbourhood $\theta_\gamma(V)$ of y these vectorfields F_i will satisfy the above mentioned conditions (i) and (ii).
Therefore:

$$H(y) = \sum_{j=k+1}^{n} g((\sum_{i=1}^{k} \nabla_{F_i} F_i)_y, F_j(y))F_j(y)$$

.

But since θ_γ is an isometry:

$$\nabla_{F_i} F_i = \nabla_{d\theta_j(E_i)}(d\theta_\gamma(E_i)) = d\theta_\gamma(\nabla_{E_i} E_i)$$

Therefore:

$$\begin{aligned}
H(y) &= \sum_{j=k+1}^{n} g(d\theta_\gamma(\sum_{i=1}^{k} \nabla_{E_i} E_i)_x, d\theta_\gamma(E_j))d\theta_\gamma(E_j) \\
&= \sum_{j=k+1}^{n} g(\sum_{i=1}^{k}(\nabla_{E_i} E_i)_x, E_j(x))d\theta_\gamma(E_j(x)) \\
&= d\theta_\gamma(\sum_{j=k+1}^{n} g(\sum_{i=1}^{k}(\nabla_{E_i} E_i)_x, E_j(x))E_j(x)) \\
&= d\theta_\gamma(H_x)
\end{aligned}$$

□

Corollary 5

If G is a Liegroup of isometries acting regularly on a Riemannian manifold M and $\pi : M \longrightarrow M/G$ is the canonical projection on the quotient-manifold, then:

$$\forall x, y \in G_x : \; d\pi(H(x)) = d\pi(H(y))$$

□

Notice that proposition 4 is only nontrivial if the action θ is non-transitive, for otherwise there is just one orbit which coincides with the manifold itself, the mean curvature of which is (by construction) identically equal to zero.

However, as shown in proposition 2 and 3, the transitive action of a Liegroup G on a homogeneous manifold M gives rise to a natural submersion,

$$\pi : \; G \longrightarrow G/M \cong M$$

and in order to apply the submersion-theorem we must be able to show that we can construct metrics on G and M such that π becomes a Riemannian submersion with fibres of constant mean curvature.

It is clear that proposition 4 furnishes us with part of the answer, for if we look at the submersion

$$\pi : \; G \longrightarrow G/M$$

then the fibres of π are the left-cosets of H, or equivalently: the orbits of the right-action of H on G.

So if we construct a Riemannian metric g_G on G which is invariant under right-translation ρ by elements of H, i.e.:

$$\forall \gamma \in G, \forall x \in H, \forall X, Y \in T_\gamma G:$$
$$g_G(\gamma\alpha)(d\rho_\alpha(X), d\rho_\alpha(Y)) = g_G(\gamma)(X, Y) \tag{3.1}$$

then the action of H will generate isometries on G and therefore it follows from proposition 4 that the mean curvature vector fields of the orbits (i.e.: the left cosets) are invariant under this action. Notice in particular that equation (3.1) implies that right-translation generates isometries between the horizontal spaces along the same fibre.
Hence we can induce a metric on the quotientspace G/H which turns π into a Riemannian submersion.
So we can conclude:

Proposition 6

If g_G is a Riemannian metric on a Liegroup G which is invariant under right-translation ρ by a closed Liesubgroup H (in the sense of equation (3.1)),
then it is possible to induce a Riemannian metric g_Q on the quotientspace G/H such that the canonical projection

$$\pi: \; G \longrightarrow G/M$$

becomes a Riemannian submersion with fibres of ρ_H−invariant mean curvature.

□

Remark:

A metric g_G on G which satisfies equation (3.1) will be called a H-right-invariant metric and is an obvious generalisation of an ordinary right-invariant metric (which is short for G-right-invariant).
Of course, we have an analogous definition for H-left-invariance. Metrics which are H-left (right-invariant (when H $\neq$ $\{e\}$)) are also called 'dimensionally reducible metrics' because they are completely determined by their values on lower-dimensional submanifolds. Eg. a left-invariant metric is completely determined by its values at a single point.
A simple example of a metric which is only invariant under a subgroup is obtained

by looking at the affine group:

$$G = \{\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mid y > 0, x \in \mathbb{R}\}$$

Two linearly independent left-invariant 1-forms on G are given by:

$$\theta_1 = \frac{dx}{y} \quad \text{and} \quad \theta_2 = \frac{dy}{y}$$

Hence,

$$g = \theta_1 \otimes \theta_1 + \theta_2 \otimes \theta_2 = \frac{1}{y^2}(dx^2 + dy^2)$$

is a left-invariant metric, but it is easy to check that the metric:

$$\tilde{g} = \theta_1 \otimes \theta_1 + y^2\theta_2 \otimes \theta_2 = \frac{dx^2 + ydy^2}{y^2}$$

is only H-left-invariant, where H is the translationsubgroup:

$$h = \{\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R}\} \subset G$$

. $\diamond$

Proposition 6 shows us how to use the metric g_G to induce a metric on the quotientspace $G/H \cong M$. In most applications however, the situation is slightly more complicated because a typical set-up will consist of a given Riemannian manifold (M, g_M) on which a Liegroup G of isometrics acts transitively (such a metric g_M is called G-left-invariant).
Therefore, we must construct, starting from g_M, a metric g_G on G such that

$$\pi : \ (G, g_G) \longrightarrow (M, g_M)$$

is a Riemannian submersion.
The key-result is given in the next proposition; recall that the action of a Liegroup G on a manifold M is called **effective** if the identity element $e \in G$ is the only element of G which acts trivially on M:

$$\text{if } \forall\, x \in M : \theta_\gamma(x) = x, \ \text{then } \gamma = e$$

.

Proposition 7 (cfr. Poor [21] p. 213)

The following conditions are equivalent if G acts effectively on a connected homogeneous space $M = G/H$:

(i) M admits a G-left-invariant Riemannian metric;

(ii) G admits a left-invariant metric which is bi-invariant under H;

(iii) g admits an Ad_H-invariant inner product;

(iv) $Ad(H)$ has compact closure in $Gl(g)$.

□

For our purposes it is important to notice that the implication $(i) \Rightarrow (ii)$ can be slightly strengthened: if a G-left-invariant Riemannian metric g_M on M is given, then we can construct a left-invariant metric on G (bi-invariant under H) which actually projects down (under the canonical projection $\pi : G \longrightarrow G/H \cong M$) to g_M. To see this we proceed as follows: let $\ll \cdot, \cdot \gg$ be an Ad_H -invariant inner product on g (the existence of which is assured by (iii)).Since the Lie-algebra h of H is Ad_H-invariant, so is its orthogonal complement (with respect to $\ll, \gg$) $h^\perp$. Let P denote the orthogonal projection of g onto h. We know that the derivative of the canonical projection $d\pi \equiv d\pi(e)$ is non-singular on $h^\perp$, so we define for $X, Y \in g$:

$$< X, Y > := \ll P(X), P(Y) \gg \; + g_M(d\pi(X), d\pi(Y)).$$

Notice that with respect to this new inner-product the spaces h and $h^\perp$ are still orhogonal.
As always we extend this inner product to a metric g_G over the rest of the group by using left-translation, thus obtaining a left-invariant metric.
Moreover g_G is bi-invariant under H because $<,>$ is invariant under Ad_H: the only thing to check is that the restriction of $<,>$ to the ($Ad(H)$-invariant) subspace $h^\perp$ is invariant under Ad_H.
But if $X, Y \in h^\perp$, then $< X, Y > = \; g_M(d\pi(X), d\pi(Y))$
and $\forall \alpha \in H$:

$$< Ad(\alpha)X, Ad(\alpha)Y >= g_M(d\pi(Ad(\alpha)X), d\pi(Ad(\alpha)Y))$$

$$\begin{aligned}
&= g_M(d\pi \circ d\lambda_\alpha \circ d\rho_{\alpha^{-1}}(X), d\pi \circ d\lambda_\alpha \circ d\rho_{\alpha^{-1}}(Y)) \\
&= g_M(d\theta_\alpha \circ d\pi \circ d\rho_{\alpha^{-1}}(X), d\theta_\alpha \circ d\pi \circ d\rho_{\alpha^{-1}}(Y)) \quad (\text{since } \pi \circ \lambda_\alpha = \theta_\alpha \circ \pi) \\
&= g_M(d\theta_\alpha \circ d\pi(X), d\theta_\alpha \circ d\pi(Y)) \quad (\text{since } \pi \circ \rho_{\alpha^{-1}} = \pi) \\
&= g_M(d\pi(X), d\pi(Y)) \qquad \text{by } G\text{-left-invariance of } g_M \\
&= < X, Y >
\end{aligned}$$

Obviously, since g_M is G-left-invariant and is equal to the projection of g_G at $\pi(e)$, this equality will hold on the entire manifold.
Therefore it is evident that

$$\pi : (G, g_G) \longrightarrow (M, g_M)$$

is a Riemannian submersion.

The bi-invariance of g_G under H will ensure that the mean curvature of the fibres are ρ_H-invariant (cfr. prop. 4).
In fact it is easy to show that these mean-curvature vectorfields vanish identically. First of all we observe that we can isometrically transform any fibre γ_H to H by using the left-translation λ_γ^{-1}. This implies that we only need to compute the mean curvature of H at e. To do this we decompose the Lie-algebra g with respect to the $Ad(H)$-invariant innerproduct $<,>$ which is induced by g_G:

$$g = h \oplus h^\perp \tag{3.2}$$

(dim $h = p$, dim $h^\perp = q$, $p+q=n$). Next we pick an orthonormal basis for g such that $E_a \in h$ $(a = 1,\ldots,p)$, and $E_u \in h^\perp (u = p+1,\ldots,n)$. Using left-translation we then extend these vectors to vectorfields (same notation) in a neighbourhood of e. Thus we obtain orthonormal framefields of which we know that at points belonging to H, the first vectorfields $E_a(a = 1,\ldots,p)$ are tangent to H. This means that we can use these vectorfields to compute the mean curvature. In particular we have to evaluate the inner-products:

$$g_G(\nabla_{E_a} E_a, E_u) \qquad \text{for } a = 1\ldots p, u = p+1\ldots n$$

But since the connection is always assumed to be Riemannian, we have (cfr. Helgason [7], p.48):

$$\begin{aligned} 2g_G(\nabla_{E_a} E_a, E_u) &= 2E_a(g_G(E_u, E_a)) - E_u(g_G(E_a, E_a)) \\ &\quad -2g_G(E_a, [E_a, E_u]) \end{aligned}$$

By construction of the vectorfields, all the terms in the RHS vanish identically. Only the last one need some explanation. Since the inner-product on g is $Ad(H)$-invariant it follows that the decomposition (3.2) is reductive by which we mean:

$$Ad(H)(h^\perp) \subset h^\perp$$

$$ds^2 = \sum_i dx_i^2 = (\sum_i f_i^2)dr^2 + 2rdr(\sum_i f_i df_i) + r^2(\sum_i df_i^2) \tag{3.8}$$

But since $\sum_i x_i^2 = r^2$ it follows from (3.6) that:

$$\sum_i f_i^2 = 1 \tag{3.9}$$

and hence:

$$\sum_i f_i df_i = 0 \tag{3.10}$$

Substituting (3.9) and (3.10) in (3.8) yields:

$$ds^2 = dr^2 + r^2(\sum_i df_i^2)$$

which proves that π becomes a Riemannian submersion when both $\mathbb{R}^n \setminus \{0\}$ and $\mathbb{R}_0^+$ are equipped with the usual Euclidean metric.
Next, using thm 3 in chapter 2 it is obvious that the mean curvature of spheres depends only on their radius and it is given by:

$$H(r) = -\frac{1}{r}\partial_r$$

Therefore the submersiontheorem states that the projected process will be a Brownian motion plus an additional drift given by

$$V = -\frac{n-1}{2}d\pi(H) = -(\frac{n-1}{2})(-\frac{1}{r})d\pi(\partial_r)$$

Since

$$d\pi(\partial_r) = \partial_r$$

this is equivalent to saying that $R_t = \pi(\tilde{B}_t)$ satisfies the SDE(3.4).

3.3 Bessel-processes on Hyperbolic Space

Consider the disc-model of n-dimensional hyperbolic space:

$$\mathbb{H}_n(\lambda) := \{x \in \mathbb{R}^n \mid ||x||_E < \lambda\} \tag{3.11}$$
$$(\lambda > 0, ||\cdot||_E : \text{ Euclidean norm})$$

equipped with the metric:

$$ds^2 = \frac{\sum dx_i^2}{(1 - \frac{1}{\lambda^2}\sum x_i^2)^2} \tag{3.12}$$

This is a space of constant (sectional) curvature

$$K = -\frac{4}{\lambda^2} \tag{3.13}$$

Notice that the restriction to $\{x_1, \ldots, x_p, x_{p+1} = 0, \ldots, x_n = 0\}$ gives a p-dimensional hyperbolic space with the same curvature.
Again we introduce spherical coordinates r and $\theta = (\theta_1, \ldots, \theta_{n-1})$ such that

$$x_i = r f_i(\theta) \tag{3.14}$$

and $r^2 = \sum x_i^2$ (or equivalently: $\sum f_i^2 = 1$).

If we denote the hyperbolic distance-function specified in (3.12) by d_H then we get:

$$\rho := d_H(0, x) = \int_0^r ds = \int_0^r \frac{du}{(1 - \frac{u^2}{\lambda^2})}$$

and therefore

$$\rho = \frac{\lambda}{2} ln(\frac{\lambda + r}{\lambda - r}) \qquad (0 < r < \lambda) \tag{3.15}$$

Hence

$$r = \lambda tgh(\frac{\rho}{\lambda}) \tag{3.16}$$

from which it follows that:

$$dr = sech^2(\frac{\rho}{\lambda})d\rho \tag{3.17}$$

Since

$$\sum dx_i^2 = dr^2 + r^2(\sum df_i^2) \tag{3.18}$$

we can substitute (3.16) and (3.17) into (3.18) to rewrite the metric (3.12) in terms of the radial distance ρ:

$$ds^2 = d\rho^2 + \frac{\lambda^2}{4} sinh^2(\frac{2\rho}{\lambda})(\sum df_i^2) \tag{3.19}$$

In particular, it follows from (3.19) that

$$\pi : (\mathbb{H}_n(\lambda), d_H) \longrightarrow (\mathbb{R}_o^+, g_E)$$

is a Riemannian submersion.
Therefore, in order to study the radial part of Brownian motion on $\mathbb{H}_n(\lambda)$, all we have to do is to compute the mean curvature of the spheres centered at the origin.
Let's look at the two-dimensional case first.
Then we know that:

$$f_1(\theta) = cos\theta \qquad f_2(\theta) = sin\theta$$

and therefore:

$$\sum df_i^2 = d\theta^2$$

Hence equation (3.19) becomes:

$$ds^2 = d\rho^2 + \frac{\lambda^2}{4} sh^2(\frac{2\rho}{\lambda})d\theta^2$$

From this it follows that horizontal and vertical unit-vectors are given by resp.

$$\begin{aligned} E_1 &= \partial_\rho \\ E_2 &= \frac{2}{\lambda sh(\frac{2\rho}{\lambda})}\partial_\theta \end{aligned}$$

Therefore

$$\nabla_{E_2}E_2 = \frac{4}{\lambda^2 sh^2(\frac{2\rho}{\lambda})}\nabla_\theta\partial_\theta$$

and since

$$\begin{aligned}
\Gamma^\theta_{\theta\theta} &= 0 \\
\Gamma^\rho_{\theta\theta} &= \Gamma^1_{22} = \frac{1}{2}g^{11}(\partial_2 g_{21} + \partial_2 g_{21} - \partial_1 g_{22}) \\
&= -\frac{1}{2}\lambda sh(\frac{2\rho}{\lambda})ch(\frac{2\rho}{\lambda})
\end{aligned}$$

we obtain:

$$\nabla_{E_2}E_2 = -\frac{2}{\lambda}coth(\frac{2\rho}{\lambda})\partial_\rho \qquad (\rho > 0)$$

Hence it follows from the submersion theorem that the radial part ρ_t of $BM(\mathbb{H}_2(\lambda))$ satisfies the SDE:

$$d\rho_t = db_t + \frac{1}{\lambda}coth(\frac{2\rho}{\lambda})dt \qquad (b\, is\; BM(\mathbb{R})) \tag{3.20}$$

The generalisation to higher-dimensional hyberbolic spaces is straightforward because just as in the Euclidean case the extrinsic curvature of the spheres is independent of the direction this curvature is measure in. Therefore, the mean curvature vector of an $(n-1)$ dimensional sphere in $\mathbb{H}_n$ is still equal

$$H = -\frac{2}{\lambda}coth(\frac{2\rho}{\lambda})\partial_\rho \tag{3.21}$$

Substituting this into the submersion-theorem we can conclude that the Bessel-process in n-dimensional hyperbolic space satisfies the SDE:

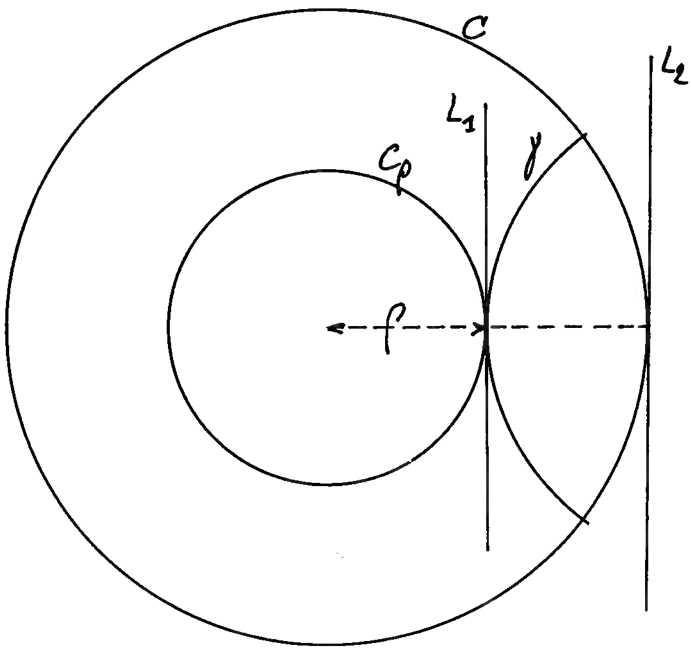
$$d\rho_t = db_t + \frac{n-1}{\lambda}coth(\frac{2\rho}{\lambda})dt \quad (\text{ } b \text{ is } BM(\mathbb{R})) \tag{3.22}$$

Remark

The interpretation of the mean curvature in terms of the separation of geodesics (cfr. thm 3 in chapter 2) provides us with an easy and intuitively appealing explanation of the difference between Euclidean and hyperbolic Bessel-processes. This difference is most striking in the two-dimensional case. The drift-term in the SDE which govern the Bessel-process is proportional to the curvature of a circle of radius ρ;
in the Euclidean space this is given by ρ^{-1} which tends to 0 as ρ increases. In the hyperbolic plane however this same curvature is given by

$$coth(\frac{2\rho}{\lambda}) \longrightarrow 1 \quad \text{as} \quad \rho \longrightarrow \infty$$

This introduces a constant outward drift which makes Brownian motion on the hyperbolic plane transient.
Looking at the following figure it is easy to see that the curvature of circles will tend to a non-vanishing constant value.



For example, to determine the mean curvature of the circle C_ρ at the point p we have to measure the rate at which C_ρ curves away from the geodesic γ. Recall that geodesics in the disc-model of $I\!H_2$ are circle-segments that meet the boundary $\partial I\!H_2$ at the right angles. But it is obvious that we can obtain a lower bound for this by looking at the rate of seperation between C_ρ and the straight line L_1, which in its turn is larger than the curvature of the boundary C of the disc with respect to L_2. This last quantity however is independent of ρ and therefore provides us with a lower bound for the curvature of every circle in the disc.

3.4 The action of the Euclidean group on Euclidean space

The action of the Euclidean group $E(n) \cong SO(n) \times I\!R^n$ on $I\!R^n$ is certainly one of the simplest nontrivial examples of a Liegroup acting on a manifold. By definition $E(n)$ is the group of the orientationpreserving isometries on $I\!R^n$.
A useful representation of this group is given by:

$$E(n) = \{\begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix} \in SL(n+1) \mid R \in SO(n) \text{ } and \text{ } \xi \in I\!R^{n\times 1}\} \tag{3.23}$$

and hence a natural (smooth) left-action on $I\!R^n$ is defined by:

$$\begin{aligned} \eta : E(n) \times I\!R^n &\longrightarrow I\!R^n \\ (\begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix}, p) &\longmapsto Rp + \xi \end{aligned} \tag{3.24}$$

which can be rewritten, using matrixnotation and by identifying p and $\begin{pmatrix} p \\ 1 \end{pmatrix}$ as:

$$\begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p \\ 1 \end{pmatrix} = \begin{pmatrix} Rp+\xi \\ 1 \end{pmatrix}$$

For any $p \in I\!R^n$, the isotropy-group H_p is given by:

$$H_p : \; = \; \{G \in E(n) \mid Gp = p\}$$

$$
\begin{aligned}
&= \{\begin{pmatrix} R & (I-R)p \\ 0 & 1 \end{pmatrix} | R \in SO(n)\} \\
&= \begin{pmatrix} I & p \\ 0 & 1 \end{pmatrix}\begin{pmatrix} SO(n) & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I & -p \\ 0 & 1 \end{pmatrix} \cong SO(n) \qquad (3.25)
\end{aligned}
$$

The action η gives rise to the following family of canonical projections:

$$
\begin{aligned}
\pi_p : E(n) &\longrightarrow I\!R^n \\
G &\longmapsto Gp \qquad (3.26)
\end{aligned}
$$

Since the action η is transitive it follows from proposition 3 that $I\!R^n$ is a homogeneous space diffeomorphic to $E(n)/SO(n)$ and that π_p is a submersion.

Next we want to construct a left (right) metric on $E(n)$: for this purpose we write:

$$
G = \begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix} \qquad G^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}\xi \\ 0 & 1 \end{pmatrix}
$$

which we then use to construct the following matrices of left - and right - invariant 1-forms:

$$
\begin{aligned}
\Omega_L &= G^{-1}dG = \begin{pmatrix} R^{-1} & -R^{-1}\xi \\ 0 & 1 \end{pmatrix}\begin{pmatrix} dR & d\xi \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R^{-1}dR & R^{-1}d\xi \\ 0 & 0 \end{pmatrix} \qquad (3.27) \\
\Omega_R &= (dG)G^{-1} = \begin{pmatrix} (dR)R^{-1} & d\xi - (dR)R^{-1}\xi \\ 0 & 0 \end{pmatrix} \qquad (3.28)
\end{aligned}
$$

Since the higher-dimensional case is a straightforward but cumbersome generalization of the two-dimensional situation, we shall restrict our attention to $E(2)$. Introducing the notation:

$$
R \equiv R_\theta = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}, \quad \xi = \begin{pmatrix} x \\ y \end{pmatrix} \qquad (3.29)
$$

we see that:

$$R^{-1}(dR) = R^T(dR) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d\theta$$

$$(dR)R^{-1} = (dR)R^T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d\theta$$

Hence:

$$\Omega_L = \begin{pmatrix} \theta & -d\theta & cos\theta dx + sin\theta dy \\ d\theta & 0 & -sin\theta dx + cos\theta dy \\ 0 & 0 & 0 \end{pmatrix} \tag{3.30}$$

$$\Omega_R = \begin{pmatrix} 0 & -d\theta & dx + yd\theta \\ d\theta & 0 & dy - xd\theta \\ 0 & 0 & 0 \end{pmatrix} \tag{3.31}$$

Hence we get the following set of independent invariant 1-forms:

$$\begin{aligned} \text{left-invariant} : \theta^1 &= cos\theta dx + sin\theta dy \\ \theta^2 &= -sin\theta dx + cos\theta dy \\ \theta^3 &= d\theta \end{aligned} \tag{3.32}$$

$$\begin{aligned} \text{right-invariant} : \omega^1 &= dx + yd\theta \\ \omega^2 &= dy - xd\theta \\ \omega^3 &= d\theta \end{aligned} \tag{3.33}$$

and the associated metrics:

$$g_L = \sum_i \theta^i \otimes \theta^i = d\theta^2 + dx^2 + dy^2 \tag{3.34}$$

$$g_R = \sum_i \omega^i \otimes \omega^i = (1 + x^2 + y^2)d\theta^2 + dx^2 + dy^2 + 2(ydxd\theta - xdyd\theta) \tag{3.35}$$

Notice that at the identity $e = (\theta = 0, x = 0, y = 0)$ we have

$$\theta(e) = \omega(e) = (dx, dy, d\theta)$$

and therefore expressing ω in terms of θ gives us an explicit parametrisation of the adjoint group: from (3.32) it follows:

$$\begin{pmatrix} dx \\ dy \\ d\theta \end{pmatrix} = \begin{pmatrix} cos\theta & -sin\theta & 0 \\ sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta^1 \\ \theta^2 \\ \theta^3 \end{pmatrix}$$

and hence this in (3.33) yields:

$$\omega = \begin{pmatrix} 1 & 0 & y \\ 0 & 1 & -x \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx \\ dy \\ d\theta \end{pmatrix} = \begin{pmatrix} cos\theta & -sin\theta & y \\ sin\theta & cos\theta & -x \\ 0 & 0 & 1 \end{pmatrix} \theta$$

from which we can conclude that $Ad(E(2)) \cong E(2)$ and therefore non-compact. Hence $E(2)$ does not support a bi-invariant metric. This in its turn means that the Killingform is degenerate and therefore $E(2)$ cannot be semi-simple. (Of course, this also follows from the elementary observation that

$$T = \{\begin{pmatrix} I & \xi \\ 0 & 1 \end{pmatrix} \mid \xi \in \mathbb{R}^{2\times 1}\}$$

is an abelian normal subgroup.).

Recall that we can use the metric to 'raise the indices' of the 1-forms and thus obtain the corresponding invariant vectorfields: for if the bases dx^i and ∂_i are dual then we can establish a natural relation between the vectorfield $X = X^i\partial_i$ and the form $\theta = \theta_i dx^i$ by means of the formula:

$$X^i = g^{ij}\theta_j. \tag{3.36}$$

Since

$$g_L^{-1}(x,y,\theta) = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \text{ and } g_R^{-1}(x,y,\theta) = \begin{pmatrix} 1+y^2 & -xy & -y \\ -xy & 1+x^2 & x \\ -y & x & 1 \end{pmatrix} \tag{3.37}$$

we get by combining (3.32),(3.33),(3.36) and (3.37):

LIVF at (x, y, θ)	RIVF at (x, y, θ)
$L_1 = cos\theta\frac{\partial}{\partial x} + sin\theta\frac{\partial}{\partial y}$	$R_1 = \frac{\partial}{\partial x}$
$L_2 = -sin\theta\frac{\partial}{\partial y} + cos\theta\frac{\partial}{\partial y}$	$R_2 = \frac{\partial}{\partial y}$
$L_3 = \frac{\partial}{\partial \theta}$	$R_3 = \frac{\partial}{\partial \theta} - y\frac{\partial}{\partial \theta} + x\frac{\partial}{\partial y}$

To construct the corresponding left- and right-invariant Brownian motions on $E(2)$ compute the Laplacians associated with g_L and g_R; recall that:

$$\Delta = \frac{1}{\sqrt{detg}}\partial_i(\sqrt{detg}\, g^{ij}\partial_j)$$

A straightforward calculation yields:

$$\Delta_L = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial \theta^2} \tag{3.38}$$

$$\Delta_R = (1+y^2)\frac{\partial^2}{\partial x^2} + (1+x^2)\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial \theta^2} - 2xy\frac{\partial^2}{\partial x \partial y} - 2y\frac{\partial^2}{\partial x \partial \theta} + 2x\frac{\partial^2}{\partial y \partial \theta} - x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y} \tag{3.39}$$

which are (twice) the generators of the left (resp. right) Brownian motion $\beta_L(t)$ (resp. $\beta_R(t)$) on $E(2)$.
By letting $E(2)$ act on a fixed point $0 \in I\!R^2$ we obtain the following projection:

$$\begin{aligned} \pi \equiv \pi_o : E(2) &\longrightarrow I\!R^2 \\ G &\longmapsto \pi(G) = \begin{pmatrix} x \\ y \end{pmatrix} \cong \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = G \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

Using lemma 5 in chapter 2 it is straightforward to compute the induced metrics $h_L = \pi(g_L)$ and $h_R = \pi(g_R)$ and their Laplacians on $I\!R^2$:

$$h_L(x,y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad h_R(x,y) = \frac{1}{1+x^2+y^2}\begin{pmatrix} 1+x^2 & xy \\ xy & 1+y^2 \end{pmatrix} \tag{3.40}$$

and

$$\Delta_L = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{3.41}$$

$$\Delta_R = (1+y^2)\frac{\partial^2}{\partial x^2} - 2xy\frac{\partial^2}{\partial x \partial y} + (1+y^2)\frac{\partial^2}{\partial y^2} - x(\frac{2+x^2+y^2}{1+x^2+y^2})\frac{\partial}{\partial x} - y(\frac{2+x^2+y^2}{1+x^2+y^2})\frac{\partial}{\partial y} \tag{3.42}$$

To obtain the structure of the projected process we have to compute the mean curvature vector of the fibres $F_{(x,y)}$ over (x,y).
Since $E(2)$ equipped with the left-invariant metric is isometric to the cylinder $I\!R^2 \times S^1$ with the induced Euclidean $I\!R^3$-metric, a moment's thought convinces us that the mean curvature of the fibres vanishes.
Hence $\pi(\beta_L)$ is the standard Brownian motion on $I\!R^2$.
The case of the right-invariant Brownian motion β_R is slightly more complicated. Picking a vertical field:

$$E_3(x,y,\theta) = \frac{1}{\sqrt{1+x^2+y^2}}\frac{\partial}{\partial \theta}$$

and completing this with a choice of (orthonormal) horizontal fields

$$E_1(x,y,\theta) = \sqrt{\frac{1+x^2+y^2}{1+x^2}}(\partial_x - \frac{y}{1+x^2+y^2}\partial_\theta)$$

and

$$E_2(x,y,\theta) = \sqrt{\frac{1+x^2+y^2}{1+y^2}}(\partial_y + \frac{x}{1+x^2+y^2}\partial_\theta)$$

we get after some simple calculations:

$$H(x,y,\theta) = -\frac{x\partial_x + y\partial_y}{1+x^2+y^2} \tag{3.43}$$

(notice: H is independent of θ)
Combining (3.42) and (3.43) we see that the generator of $\pi(\beta_R)$ is given by:

$$\frac{1}{2}(\Delta_R - d\pi(H)) = \frac{1}{2}((1+y^2)\partial^2_{xx} - 2xy\partial^2_{xy} + (1+x^2)\partial^2_{yy} - x\partial_x - y\partial_y)$$

Of course the same result could be obtained by looking at the corresponding $SDEs$. To do this we have to construct the appropriate process on the Lie-algebra. First of all we pick a basis of the Lie-algebra $e(2)$ and to make the results readily interpretable we make sure that this basis agrees with the left-(and right-) invariant vectorfields at the identity:

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.44}$$

A simple calculation shows that the only nonvanishing structural constants are given by:

$$C^2_{31} = 1 = -C^2_{13}$$
$$C^1_{23} = 1 = -C^1_{32}$$

and hence

$$\sum_j K^i_{jj} = \sum_j C^j_{ij} = 0 \quad \text{for all i = 1,2,3}$$

Therefore Brownian motion on $E(2)$ is generated by a non-drifting Brownian motion.

$$B(t) = A_1 b_1(t) + A_2 b_2(t) + A_3 b_3(t) \quad (b_i\text{: indep. BM}(\mathbb{R})) \tag{3.45}$$

on the Liealgebra $e(2)$.
This means that the corresponding left-and right-invariant $BM(E(2))$ are defined by the stratovitch SDE:

$$\partial\beta_L = \beta_L \partial B \tag{3.46}$$

and

$$\partial\beta_R = (\partial B)\beta_R \tag{3.47}$$

or the Ito-formulae:

$$d\beta_L = \beta_L dB + \frac{1}{2}\beta_L (dB)^2 \tag{3.48}$$

and

$$d\beta_R = (dB)\beta_R + \frac{1}{2}(dB)^2\beta_R \tag{3.49}$$

But since it follows from (3.44) and (3.45) that

$$dB^2 = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 0 \end{pmatrix} dt$$

we get:

$$d\beta_L = \beta_L dB - \frac{1}{2}\beta_L \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix} dt \tag{3.50}$$

and

$$d\beta_R = (dB)\beta_R - \frac{1}{2} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix} \beta_R dt \tag{3.51}$$

Using the representation (3.29) a straigthforward calculation yields the following SDEs for the components x, y, θ:

Left-invariant BM(E(2))

$$dx = cos\theta db_1 - sin\theta db_2$$
$$dy = sin\theta db_1 + cos\theta db_2$$
$$d\theta = db_3$$

from which it follows that x,y and θ perform independent BM($I\!R$), which agrees with Laplacian;

Right-invariant BM(E(2))

$$dx = -y db_3 + db_1 - \frac{1}{2}x dt$$

$$dy = x db_3 + db_2 - \frac{1}{2} y dt$$
$$d\theta = db_3$$

Again, the covariationstructure of these processes agrees with the corresponding Laplacian.

3.5 The action of the Lorentz-group on the Euclidean plane

Whereas the Euclidean group $E(2)$ leaves the usual inner product on $\mathbb{R}^2$ invariant, the Lorentz-group preserves the Lorentz-metric:

$$ds^2 = dx^2 - dy^2.$$

More explicitely, the Lorentzgroup can be represented by

$$L(2) = \{\begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix} \mid R \in SO(1,1),\ \xi \in \mathbb{R}^{2\times 1}\}$$

where

$$SO(1,1) = \{\begin{pmatrix} ch\theta & sh\theta \\ sh\theta & ch\theta \end{pmatrix} \mid \theta \in \mathbb{R}\}$$

Similar to the Euclidean case there is a natural (smooth) left-action of $L(2)$ on $\mathbb{R}^2$, given by:

$$\begin{array}{rcl} \eta : \xi(2) \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (\begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix}, p) & \longmapsto & Rp + \xi \cong \begin{pmatrix} Rp + \xi \\ 1 \end{pmatrix} = \begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ 1 \end{pmatrix} \end{array}$$

and the associated family of projections:

$$\begin{array}{rl} \pi_p : & L(2) \longrightarrow \mathbb{R}^2 \\ & \begin{pmatrix} R & \xi \\ 0 & 1 \end{pmatrix} \longmapsto Rp + \xi \end{array}$$

the isotropy-group of which is isomorphic to $SO(1,1)$.
Since the action is transitive it follows that $\mathbb{R}^2$ is a homogeneous space diffeomorphic to $L(2)/SO(1,1)$, and that π_p is a submersion.
As it is easy to see that

$$Z(L(2)) = 1$$

it follows that

$$Ad(L(2)) \cong L(2)$$

and therefore $L(2)$ does not support a bi-invariant metric.
Hence, the next step in our programme is to compute a left- and a right-invariant metric. For this purpose we return to equations (3.27) and (3.28),but now of course:

$$R^{-1} = \begin{pmatrix} ch\theta & sh\theta \\ sh\theta & ch\theta \end{pmatrix}^{-1} = \begin{pmatrix} ch\theta & -sh\theta \\ -sh\theta & ch\theta \end{pmatrix}$$

and hence

$$R^{-1}dR = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} d\theta = (dR)R^{-1}$$

This yields:

$$\begin{aligned} \Omega_L &= \begin{pmatrix} 0 & d\theta & ch\theta dx - sh\theta dy \\ d\theta & 0 & -sh\theta dx + ch\theta dy \\ 0 & 0 & 0 \end{pmatrix} \\ \Omega_R &= \begin{pmatrix} 0 & d\theta & dx - y d\theta \\ d\theta & 0 & dy - x d\theta \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Hence we get the following set of independent and invariant 1-forms:
<u>left-invariant:</u>

$$\begin{aligned} \theta^1 &= ch\theta dx - sh\theta dy \\ \theta^2 &= -sh\theta dx + ch\theta dy \\ \theta^3 &= d\theta \end{aligned}$$

<u>right-invariant:</u>

$$\begin{aligned} \omega^1 &= dx - y d\theta \\ \omega^2 &= dy - x d\theta \\ \omega^3 &= d\theta \end{aligned}$$

and the associated metrics:

$$\begin{aligned} g_L &= \sum(\theta^i)^2 = ch2\theta(dx^2 + dy^2) + d\theta^2 - 2sh2\theta dx dy & (3.52) \\ g_R &= \sum(\omega^i)^2 = dx^2 + dy^2 + (1 + x^2 + y^2)d\theta^2 - 2x dy d\theta - 2y dx d\theta & (3.53) \end{aligned}$$

In contrast to the Euclidean case, it is actually impossible to construct a metric on $I\!R^2$ such that the projection π_p of $(L(2), g_L)$ on $I\!R^2$ would become a Riemannian submersion.
Therefore we restrict our attention to g_R which induces the following metric on $I\!R^2$ (cfr. lemma 5, chapter 2); since:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -y \\ -x \end{pmatrix} \quad C = 1 + x^2 + y^2$$

we get:

$$\begin{aligned} h : &= \pi(g_R) = A - BC^{-1}B^T \\ &= \frac{1}{1+x^2+y^2} \begin{pmatrix} 1+x^2 & -xy \\ -xy & 1+y^2 \end{pmatrix} \end{aligned}$$

Therefore, the corresponding Laplacian is given by:

$$\begin{aligned} \triangle_{I\!R^2} &= (1+y^2)\partial^2_{xx} + 2xy\partial_{xy} + (1+x^2)\partial^2_{yy} \\ &+ x(\frac{x^2-y^2}{1+x^2+y^2})\partial_x + y(\frac{y^2-x^2}{1+x^2+y^2})\partial_y \end{aligned}$$

On the other hand, the mean curvature H of the fibre

$$F_{(x,y)} = \{ \begin{pmatrix} ch\theta & sh\theta & x \\ sh\theta & ch\theta & y \\ 0 & 0 & 1 \end{pmatrix} \mid \theta \in I\!R \}$$

is given by:

$$H = -\frac{x(1+2y^2)\partial_x + y(1+2x^2)\partial_y + 2xy\partial_\theta}{1+x^2+y^2}$$

Therefore $d\pi(H)$ does not depend on θ and we have that the generator of the projected process is given by:

$$\frac{1}{2}(\triangle_{I\!R^2} - d\pi(H)) = \frac{1}{2}\{(1+y^2)\partial^2_{xx} + 2xy\partial^2_{xy} + (1+x^2)\partial^2_{yy} + x\partial_x + y\partial_y\}$$

Compare this with the similar equation for the Euclidean group.

Remark: There is of course, a more straightforward way of proving that there does not exist a Riemannian metric on $\mathbb{R}^2$ which is invariant under the left-action of $L(2)$.
For if such a metric existed then its matrix G should satisfy the equation:

$$R^T G R = G \quad \text{for all} \quad R = \begin{pmatrix} ch\theta & sh\theta \\ sh\theta & ch\theta \end{pmatrix} \tag{3.54}$$

It is easy to see that R has eigenvalues e^{θ} and $e^{-\theta}$ so that (3.54) can be rewritten as:

$$DSD = S \qquad \forall D = diag(e^{\theta}, e^{-\theta}) \tag{3.55}$$

where

$$D = \begin{pmatrix} e^{\theta} & 0 \\ 0 & e^{-\theta} \end{pmatrix} \quad S = VGV^T \, , \, V \in SO(2)$$

Now it is obvious that the most general symmetric solution of (3.55) is of the form:

$$S = \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}$$

but this matrix is not positive definite.
So there doesn't exist a positive-definite symmetric solution of (3.54).

3.6 The special linear group acting on the hyperbolic plane

The action of $SL(2, \mathbb{R})$ on the hyperbolic plane provides us with a more complicated and slightly puzzling application of the submersion theorem. In this case the most convenient representation of the hyperbolic plane is given by the Poincaré-model:

$$\mathbb{H} = \{(x,y) \in \mathbb{R}^2 \mid y > 0\} = \{z \in \mathbb{C} \mid Im\, z > o\} \tag{3.56}$$

equipped with the metric:

$$ds^2 = \frac{dx^2 + dy^2}{y^2} \tag{3.57}$$

There is a natural left-action of $SL(2,\mathbb{R})$ on $\mathbb{H}$ given explicitly by:

$$\begin{array}{rcl} SL(2,\mathbb{R}) \times \mathbb{H} & \longrightarrow & \mathbb{H} \\ (\begin{pmatrix} a & b \\ c & d \end{pmatrix}, z) & \longmapsto & \dfrac{az+b}{cz+d} \end{array} \tag{3.58}$$

This action is obviously transitive so we can use it to generate a submersion of $SL(2,\mathbb{R})$ onto $\mathbb{H}$ by applying it to a fixed point $z_o \in \mathbb{H}$. For convenience we choose $z_o = i = (0,1)$.
In terms of coordinates this projection becomes:

$$\pi : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \frac{ai+b}{ci+d} = \frac{(ac+bd)+i}{c^2+d^2} = \frac{(ac+bd,1)}{c^2+d^2} \tag{3.59}$$

The isotropy-group H_i of i must therefore be a subgroup of $SL(2,\mathbb{R})$ satisfying the equations:

$$ac + bd = 0 \tag{3.60}$$

$$c^2 + d^2 = 1 \tag{3.61}$$

From (3.60) it follows that there exists a value λ in $\mathbb{R}$ such that:

$$a = \lambda d \quad \text{and} \quad b = -\lambda c \tag{3.62}$$

But by definition of $SL(2,\mathbb{R})$ we know

$$ad - bc = 1 \tag{3.63}$$

Substituting (3.62) into (3.63) and using (3.60) yields:

$$a^2 + b^2 = 1 \tag{3.64}$$

Collecting (3.60), (3.61) and (3.64) shows that the isotropygroup is equal to:

$$H_i = SO(2)$$

Hence it follows from proposition 3 that $\mathbb{H}$ is a homogeneous space isomorphic to $SL(2,\mathbb{R})/SO(2)$ and that π (as defined by (3.59)) is a submersion.
Next we use this result to find an appropriate parametrisation of $SL(2,\mathbb{R})$.
Take any $(x,y) \in \mathbb{H}$, then it is easy (using (3.59) again) to construct a matrix in $SL(2,\mathbb{R})$ which projects down to (x,y);
in fact

$$\frac{1}{\sqrt{y}}\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \stackrel{\pi}{\longmapsto} (x,y)$$

and therefore the corresponding fibre $F_{(x,y)} = \pi^{-1}(x,y)$ is obtainded by multiplying (on the right) by $SO(2)$.
This yields the following explicit representation of $SL(2,\mathbb{R})$:

$$SL(2,\mathbb{R}) = \{\frac{1}{\sqrt{y}}\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \mid \theta \in \mathbb{R}, x \in \mathbb{R}, y > 0\} \tag{3.65}$$

This parametrisation can then be used to construct left- and right-invariant 1-forms on $SL(2,\mathbb{R})$. Notice that we have already established that the group cannot support a bi-invariant metric since: (cfr. remark 2 p.55)

$$Ad(SL(2,\mathbb{R})) \cong SL(2,\mathbb{R})/Z(SL(2,\mathbb{R})) = SL(2,\mathbb{R})/\pm 1$$

and is therefore non-compact.
To simplify some of the formulae, let us denote a general element $G \in SL(2,\mathbb{R})$ as:

$$G = \frac{1}{\sqrt{y}} XR \tag{3.66}$$

where of course:

$$X = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \quad \text{and } R = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \tag{3.67}$$

Hence we obtain that:

$$G^{-1} = \sqrt{y} R^{-1} X^{-1} \quad \text{where } X^{-1} = \frac{1}{y} \begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix} \tag{3.68}$$

and

$$dG = \frac{1}{\sqrt{y}}(-\frac{dy}{2y} XR + dXR + XdR) \tag{3.69}$$

Therefore a matrix of left-invariant 1-forms is given by:

$$\Omega_L = G^{-1}(dG) = -\frac{dy}{2y} I + R^T(X^{-1}dX)R + R^T dR \tag{3.70}$$

whereas

$$\Omega_R = (dG)G^{-1} = -\frac{dy}{2y} I + (dX)X^{-1} + X(dR)R^T X^{-1} \tag{3.71}$$

yields right-invariant 1-forms.
Since

$$R^T dR = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d\theta = (dR)R^T \tag{3.72}$$

and

$$X^{-1}dX = \frac{1}{y} \begin{pmatrix} dy & dx \\ 0 & 0 \end{pmatrix} \tag{3.73}$$

$$(dX)X^{-1} = \frac{1}{y} \begin{pmatrix} dy & ydx - xdy \\ 0 & 0 \end{pmatrix} \tag{3.74}$$

Substituting all this in (3.70) and (3.71) yields:

$$(\Omega_L)_{11} \quad = \quad -\frac{dy}{2y} + \frac{1}{y}(cos\theta sin\theta dx + cos^2\theta dy)$$

$$
\begin{aligned}
(\Omega_L)_{12} &= \frac{1}{y}(cos^2\theta dx - cos\theta sin\theta dy) - d\theta \\
(\Omega_L)_{21} &= -\frac{1}{y}(sin^2\theta dx + cos\theta sin\theta dy) + d\theta \\
(\Omega_L)_{22} &= -\frac{dy}{2y} + \frac{1}{y}(sin^2\theta dy - cos\theta sin\theta dx)
\end{aligned}
$$

From this we choose the following three independent left-invariant 1-forms:

$$\theta^1 = 2(\Omega_L)_{11} = \frac{sin2\theta dx + cos2\theta dy}{y} \tag{3.75}$$

$$\theta^2 = (\Omega_L)_{12} + (\Omega_L)_{21} = \frac{cos2\theta dx - sin2\theta dy}{y} \tag{3.76}$$

$$\theta^3 = (\Omega_L)_{12} - (\Omega_L)_{21} = \frac{dx - 2yd\theta}{y} \tag{3.77}$$

The corresponding left-invariant metric is given by:

$$g_L = \sum(\theta^i)^2 = \frac{2dx^2 + dy^2 + 4y^2 d\theta^2 - 4ydxd\theta}{y^2} \tag{3.78}$$

Partitioning the metric tensor as:

$$A = \begin{pmatrix} 2/y^2 & 0 \\ 0 & 1/y^2 \end{pmatrix} \quad B = \begin{pmatrix} -2/y \\ 0 \end{pmatrix} \quad C = 4$$

we see that the induced metric is given by:

$$\pi(g_L) = A - BC^{-1}B^T = \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix} \tag{3.79}$$

But this is exactly the metric (3.57) on the hyperbolic plane, which shows us that the projection π of $(SL(2,\mathbb{R}), g_L)$ on the hyperbolic plane is in fact a Riemannian submersion.
The next thing to do is to compute the mean curvature of the fibres. But it follws from thm 8 in chapter 3 that these must vanish identically. An explicit check of

this can be obtainde by looking at the covariant derivative of the orthonormal vertical vectorfield $E_1 = \frac{1}{2}\partial_\theta$:

$$\begin{aligned}\nabla_{E_1}E_1 &= \frac{1}{4}\nabla_\theta\partial_\theta \\ &= \frac{1}{4}(\Gamma^x_{\theta\theta}\partial_x + \Gamma^y_{\theta\theta}\partial_y + \Gamma^\theta_{\theta\theta}\partial_\theta)\end{aligned}$$

Since the inverse metric is given by

$$g_L^{-1} = \begin{pmatrix} y^2 & 0 & \frac{y}{2} \\ 0 & 1 & 0 \\ \frac{y}{2} & 0 & \frac{1}{2} \end{pmatrix} \tag{3.80}$$

a straightforward calculation using the equation

$$\Gamma^i_{jk} = \frac{1}{2}g^{il}(\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jh})$$

yields

$$\Gamma^x_{\theta\theta} = \Gamma^y_{\theta\theta} = \Gamma^\theta_{\theta\theta} = 0$$

which confirms that the mean curvature of the fibres of π vanish identically.
Hence we can conclude that the left-invariant Brownian motion on $SL(2,\mathbb{R})$ projects down to the Brownian motion on the hyperbolic plane.
Let us now look at the right-invariant Brownian motion which is the example David Williams considers in Rogers & Williams [22], p.245.
Here the situation becomes far more complicated.
First of all we need construct right-invariant 1-forms : substituting (3.72) and (3.74) into (3.71) gives us:

$$\begin{aligned}(\Omega_R)_{11} &= \frac{dy + 2xd\theta}{2y} \\ (\Omega_R)_{12} &= \frac{ydx - xdy - (x^2+y^2)d\theta}{y} \\ (\Omega_R)_{21} &= \frac{d\theta}{y} \\ (\Omega_R)_{22} &= -\frac{dy + 2xd\theta}{y}\end{aligned}$$

From this we pick the following basis of right-invariant 1-forms:

$$\omega^1 = 2(\Omega_R)_{11} = \frac{dy + 2xd\theta}{y} \tag{3.81}$$
$$\omega^2 = (\Omega_R)_{12} = \frac{ydx - xdy - (x^2+y^2)d\theta}{y} \tag{3.82}$$
$$\omega^3 = (\Omega_R)_{21} = \frac{d\theta}{y} \tag{3.83}$$

The corresponding metric is equal to:

$$\begin{aligned} g_R &= \sum(\omega^i)^2 \\ &= \frac{1}{y^2}\{y^2dx^2 + (1+x^2)dy^2 + (1+4x^2+(x^2+y^2)^2)d\theta^2 \\ &- 2xydxdy - 2y(x^2+y^2)dxd\theta + 2x(2+x^2+y^2)dyd\theta\} \end{aligned} \tag{3.84}$$

Because the coordinatization of $SL(2,\mathbb{R})$ we use, means that the submersion π is just a projection on the first two coordinates we can easily compute the induced metric $h = \pi(g_R)$ by partitioning g_R into:

$$g_R(x,y,\theta) = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

where

$$A = \begin{pmatrix} 1 & -\frac{x}{y} \\ -\frac{x}{y} & \frac{1+x^2}{y^2} \end{pmatrix} \quad B = \begin{pmatrix} -\frac{x^2+y^2}{y} \\ \frac{x(2+x^2+y^2)}{y^2} \end{pmatrix}$$
$$C = \frac{1+4x^2+(x^2+y^2)^2}{y^2}$$

Then it follows from lemma 5 in chapter 2 that:

$$h = \begin{pmatrix} \frac{1+4x^2}{1+4x^2+(x^2+y^2)^2} & \frac{-x(1+2x^2-2y^2}{y(1+4x^2+(x^2+y^2)^2)} \\ \frac{-x(1+2x^2-2y^2)}{y(1+4x^2+(x^2+y^2)^2} & \frac{1+x^2+(x^2-y^2)^2}{y^2(1+4x^2+(x^2+y^2)^2} \end{pmatrix}$$

Since

$$det\, h = y^{-2}(1 + 4x^2 + (x^2 + y^2)^2)^{-1}$$

and

$$h^{-1} = \begin{pmatrix} 1 + x^2 + (x^2 - y^2)^2 & xy(1 + 2(x^2 - y^2)) \\ xy(1 + 2(x^2 - y^2)) & y^2(1 + 4x^2) \end{pmatrix}$$

the corresponding Laplacian has the following (revolting) form:

$$\begin{aligned}
\triangle_h &= (1 + x^2 + (x^2 - y^2)^2)\frac{\partial^2}{\partial x^2} + 2xy(1 + 2(x^2 - y^2))\frac{\partial^2}{\partial x \partial y} + y^2(1 + 4x^2)\frac{\partial^2}{\partial y^2} \\
&+ \frac{2x}{1 + 4x^2 + (x^2 + y^2)^2}(x^6 - x^4y^2 + 6x^4 - 5x^2y^4 - 12x^2y^2 \\
&+ 3x^2 - 3y^6 - 2y^4 - 5y^2 - 1)\frac{\partial}{\partial x} \\
&+ \frac{2y}{1 + 4x^2 + (x^2 + y^2)^2}(3x^6 + 5x^4y^2 + 16x^4 + x^2y^4 \\
&+ 7x^2 - y^6 - y^2 + 1)\frac{\partial}{\partial y} \qquad (3.85)
\end{aligned}$$

To compute the mean curvature of the fibres we look at the orthonormal vertical vectorfield

$$E_1 = \frac{y}{\sqrt{1 + 4x^2 + (x^2 + y^2)^2}}\partial_\theta$$

for which we compute the covariant derivative:

$$\begin{aligned}
\nabla_{E_3} E_3 &= \frac{y^2}{1 + 4x^2 + (x^2 + y^2)^2}\nabla_\theta \partial_\theta \\
&= \frac{y^2}{1 + 4x^2 + (x^2 + y^2)^2}(\Gamma^1_{33}\partial_x + \Gamma^2_{33}\partial_y + \Gamma^3_{33}\partial_z)
\end{aligned}$$

Since the inverse of the right-invariant metric is given by:

$$
g_R^{-1} = \begin{pmatrix} 1+x^2+(x^2-y^2)^2 & xy(1+2(x^2-y^2) & y(y^2-x^2) \\ xy(1+2(x^2-y^2)) & y^2(1+4x^2) & -2xy \\ y(y^2-x^2) & -2xy & y^2 \end{pmatrix}
$$

a straightforward albeit messy calculation yields:

$$
\begin{aligned}
\Gamma^1_{33} &= \frac{x(3x^4-2x^2y^2-5y^4-4y^2-3)}{y^2} \\
\Gamma^2_{33} &= \frac{7x^4+6x^2y^2+4x^2-y^4+1}{y} \\
\Gamma^3_{33} &= \frac{-2x(1+2(x^2+y^2))}{y}
\end{aligned}
$$

From this it follows that

$$
\begin{aligned}
\nabla_{E_1} E_1 &= \frac{1}{1+4x^2+(x^2+y^2)^2}\{x(3x^4-2x^2y^2-5y^4-4y^2-3)\partial_x \\
&+ y(7x^4+6x^2y^2+4x^2-y^4+1)\partial_y - 2xy(1+2(x^2+y^2))\partial_\theta\}
\end{aligned}
$$

and since $\nabla_{E_1} E_1$ is orthogonal to E_1 we have that the mean curvature vectorfield H is given by:

$$
H = \nabla_{E_1} E_1
$$

(independent of θ !). Hence:

$$
\begin{aligned}
d\pi(H) &= \frac{1}{1+4x^2+(x^2+y^2)^2}\{x(3x^4-2x^2y^2-5y^4-4y^4-3)\partial_x \\
&+ y(7x^4+6x^2y^2+4x^2-y^4+1)\partial_y\}
\end{aligned} \tag{3.86}
$$

Combining (3.85) and (3.86) we are relieved to see that most of the terms cancel so that (twice) the generator of the projected process takes on the following form:

$$
\begin{aligned}
\Delta_h - d\pi(H) &= (1+x^2+(x^2-y^2)^2)\partial^2_{xx} + 2xy(1+2(x^2-y^2))\partial^2_{xy} \\
&+ y^2(1+4x^2)\partial^2_{yy} \\
&+ x(1+2x^2-6y^2)\partial_x + y(1+6x^2-2y^2)\partial_y
\end{aligned} \tag{3.87}
$$

an expression still sufficiently horrible to elicit a cry of vexed contempt from David Williams in Rogers & Williams [22] p. 245.)

Because of the rather chaotic structure of this last generator, it is of some (independent) interest to see whether these results can be obtained more efficiently by looking at the corresponding SDEs.
This means that we start off by defining $G_t = BM(SL,(2,\mathbb{R}))$ by means of a (Stratonovich) SDE:

$$\text{left-invariant} \;:\; \partial G = G(\partial B) \tag{3.88}$$
$$\text{right-invariant} \;:\; \partial G = (\partial B)G \tag{3.89}$$

where B is the appropriate (drifting) Brownian motion on $SL(2,\mathbb{R})$. Notice that the projection π (as defined by equation (3.59)) essentially amounts to a multiplication (on the right) by a fixed element $z_o \in \mathbb{H}$:

$$G = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \longmapsto \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} z_o \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma_{11} z_o + \gamma_{12} \\ \gamma_{21} z_o + \gamma_{22} \end{pmatrix} \cong \begin{pmatrix} \dfrac{\gamma_{11} z_o + \gamma_{12}}{\gamma_{21} z_o + \gamma_{22}} \\ 1 \end{pmatrix} \tag{3.90}$$

$(\gamma_{21} z_o + \gamma_{22} \neq 0 ::$ since $Im\, z_o > 0)$.

It is easy to write down an autonomous SDE for the projected right-invariant $BM(SL,(2,\mathbb{R}))$: multiplication (on the right) of (3.89) by z_o yields:

$$\partial(Gz_o) = (\partial B)(Gz_o)$$

or equivalently (denoting $z = Gz_o$):

$$\partial z = (\partial B)z \tag{3.91}$$

To get a similar SDE for the projection of the left-invariant BM seems not so easy, although the submersiontheorem actually tells us that the result is a very simple autonomous process.
One way to get round this difficulty is to use equation (3.88) to find explicit SDEs for the matrixelements which we then can substitute in (3.90).

To work out this programme we first pick a suitable basis for the Lie-algebra $SL(2,\mathbb{R})$. Since we want to be able to compare these results to the generator obtained earlier, we make sure that the vectorfields are dual to the one-forms θ. Recall that

$$X = X^i \partial_i \quad \text{and} \quad \theta = \theta_i dx^i$$

are dual iff:

$$X^i = g^{ij}\theta_j \tag{3.92}$$

Using the inverse-metric g_L^{-1} (as given in (3.80)) substitution of (3.75),(3.76),(3.77) in (3.92) yields the following orthonormal LIVFs:

$$\begin{aligned}
X_1 &= y sin2\theta\partial_x + \frac{1}{y} cos2\theta\partial_y + \frac{1}{2} sin2\theta\partial_\theta && (3.93)\\
X_2 &= y cos2\theta\partial_x - \frac{1}{y} sin2\theta\partial_y + \frac{1}{2} cos2\theta\partial_\theta && (3.94)\\
X_3 &= -\frac{1}{2}\partial_\theta && (3.95)
\end{aligned}$$

At the identity $I = (x = 0, y = 1, \theta = 0)$ they generate the following basis for $SL(2,\mathbb{R})$:

$$\begin{aligned}
A_1 = X_1(I) &= \frac{\partial}{\partial y}\left(\frac{1}{\sqrt{y}}\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}\right)\Big|_{(x=0,y=1,\theta=0)} && (3.96)\\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} && (3.97)
\end{aligned}$$

Similarly:

$$A_2 = X_2(I) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.98}$$

$$A_3 = X_3(I) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.99}$$

A straightforward calculation (using this basis) then shows that:

$$\sum K^1_{ii} = \sum K^2_{ii} = \sum K^3_{ii} = 0$$

which implies that the B-process (in (3.88)) has no drift and is therefore just $BM(sl(2,\mathbb{R}))$:

$$\begin{aligned} B &= b_1A_1 + b_2A_2 + b_3A_3 \quad (b_i\text{: indep. } BM(\mathbb{R})) \\ &= \frac{1}{2}\begin{pmatrix} b_1 & b_2+b_3 \\ b_2-b_3 & -b_1 \end{pmatrix} \\ &\equiv \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & -\beta_1 \end{pmatrix} \end{aligned} \tag{3.100}$$

where we have the following quadratic (co-)variations:

$$d\beta_1^2 = \frac{1}{4}dt \tag{3.101}$$

$$d\beta_2^2 = d\beta_3^2 = \frac{1}{2}dt \tag{3.102}$$

$$d\beta_1 d\beta_2 = d\beta_1 d\beta_3 = d\beta_2 d\beta_3 = 0 \tag{3.103}$$

Rewriting (3.88) in Itô-notation we obtain:

$$dG = GdB + \frac{1}{2}dGdB = GdB + \frac{1}{2}GdB^2$$

but since it follows from (3.100) that:

$$dB^2 = \frac{1}{4}Idt$$

this becomes:

$$dG = GdB + \frac{1}{8}Gdt \tag{3.104}$$

Next, we use equation (3.59) in combination with (3.104) to write down explicit SDEs for the projected process: for, since

$$\begin{cases} x = y(\gamma_{11}\gamma_{21} + \gamma_{12}\gamma_{22}) \\ y = (\gamma_{21}^2 + \gamma_{22}^2)^{-1} \end{cases} \tag{3.105}$$

differentiation of (3.105) yields:

$$\begin{aligned} dy &= -2y^2(\gamma_{21}d\gamma_{21} + \gamma_{22}d\gamma_{22}) + y^2(4y\gamma_{21}^2 - 1)d\gamma_{21}^2 \\ &+ y^2(4y\gamma_{22}^2 - 1)d\gamma_{22}^2 + 8y^3\gamma_{21}\gamma_{22}d\gamma_{21}d\gamma_{22} \end{aligned} \tag{3.106}$$

Separating the martingale- from the finite-variation-part we get:

$$dy \doteq -2y^2\{(\gamma_{21}^2 - \gamma_{22}^2)d\beta_1 + \gamma_{21}\gamma_{22}(d\beta_2 + d\beta_3)\} \tag{3.107}$$

(where $\doteq$ denotes equality of martingale-parts). Squaring this expression we obtain:

$$dy^2 = y^2 dt \tag{3.108}$$

As for the finite-variationpart of (3.106) a straightforward calculation shows that

$$dy \overset{FV}{=} 0 \tag{3.109}$$

Similarly, differentiating (3.105) and using (3.107) and (3.109) yields:

$$dx \doteq \frac{x}{y}dy + y\{2(\gamma_{21}\gamma_{11} - \gamma_{12}\gamma_{22})d\beta_1 + (\gamma_{12}\gamma_{21} + \gamma_{11}\gamma_{22})(d\beta_2 + d\beta_3)\} \tag{3.110}$$

and

$$dx \overset{FV}{=} 0 \tag{3.111}$$

Squaring (3.110) we obtain:

$$dx^2 = y^2 dt \tag{3.112}$$

Moreover, multiplying (3.107) and (3.110) shows that

$$dxdy = 0 \tag{3.113}$$

Combining (3.108), (3.109), (3.111), (3.112) and (3.113) we can conclude that

$$dx = yd\omega_1 \tag{3.114}$$

$$dy = yd\omega_2 \tag{3.115}$$

where ω_1, ω_2 are independent $BM(\mathbb{R})$.
This tallies with (3.79).
Similar, but cumbersome calculations show that the projection of the right-invariant $BM(SL(2,\mathbb{R}))$ is governed by the following SDEs:

$$dx = xd\omega_1 + d\omega_2 + (y^2 - x^2)d\omega_3 + x(\frac{1}{2} + x^2 - 3y^2)dt \tag{3.116}$$

$$dy = yd\omega_1 - 2xyd\omega_3 + y(\frac{1}{2} + 3x^2 - y^2)dt \tag{3.117}$$

where of course ω_1, ω_2 and ω_3 are independent $BM(\mathbb{R})$.
This confirmes equation (3.87).

3.7 Triangular Shape-diffusion on $\mathbb{R}^n$

In this paragraph we will apply submersiontheorem to a simple example of shape-diffusion as introduced by D.G. Kendall [10]. A triangle Δxyz in $\mathbb{R}^n$ is completely determined by the three non-collinear points $x, y, z \in \mathbb{R}^n$.
Therefore

$$M := (\mathbb{R}^n)^3 \setminus \{(x,y,z) \in (\mathbb{R}^n)^3 | \ \exists \ \lambda \in \mathbb{R} : z = \lambda x + (1-\lambda)y\}$$

will be the manifold on which we concentrate our attention.
The usual Euclidean metric on $(\mathbb{R}^n)^3$ will be denoted by g_E. The shape-manifold is obtained by making the quotient of M over the group

$$G = \mathbb{R}_0^+ \times E(n) = \{(\lambda, R, v) | \ \lambda \in \mathbb{R}_0^+, R \in SO(n), v \in R^{n\times 1}\}$$

which acts componentwise on $(\mathbb{R}^n)^3$ as follows:

$$\theta : G \times M \to M$$

$$((\lambda, R, v), (x, y, z)) \mapsto (\lambda Rx + v, \lambda Ry + v, \lambda Rz + v) \qquad (3.118)$$

((x, y, z) are interpreted as columns).
As this action θ is regular and nontransitive, the quotientspace $\Sigma_n^3 = M/G$ is a well-defined manifold called the shape-manifold (more detailed information can be found in Kendall [10]). We want to show that the canonical projection $\pi : M \to \Sigma_n^3$ maps Brownian motion $BM(M, g_E)$ to an autonomous diffusion on Σ_n^3 (the so-called shape-diffusion).

It is obvious that for each $R \in SO(n)$ and $v \in \mathbb{R}^{n\times 1}$ the map $\theta_{(1,R,v)} : M \to M$ is an isometry (with respect to g_E). However this is no longer true if $\lambda \neq 1$. We therefore introduce a new metric g on M which is conformally equivalent to g_E:

$$g(x,y,z) := \frac{1}{S(x,y,z)} g_E(x,y,z) \qquad (3.119)$$

where

$$S(x,y,z) := ||x-y||^2 + ||y-z||^2 + ||z-x||^2$$

measures the size of the triangle. It is now evident that θ generates isometries with respect to g.
Since all the horizontal spaces are isometric we can induce a metric on Σ_n^3 by projecting down g using the canonical projection

$$\pi : M \to \Sigma_n^3$$

Moreover, since the projection of the mean-curvature vector is constant along each orbit (cfr. prop. 4) we know that the projection of $BM(M,g)$ will give us an autonomous diffusion on Σ_n^3.

However, we started off with $BM(M,g_E)$ and not with $BM(M,g)$. This means that if X_t is $BM(M,g_E)$ it will be a more complicated diffusion with respect to the metric g; so the problem really is to determine the form of the diffusion X_t with respect to g and see whether it will still project down to an autonomous diffusion on the quotient manifold.

To do this we make use of the fact that the metrics g_E and g are conformally related. Using (3.119) a straightforward calculation yields the following relation between the corresponding Laplacians Δ_E and Δ:

$$\Delta_E = \frac{1}{S}\Delta + \frac{3n-2}{2}\sum_k \partial_k(lnS)\partial_k \qquad (3.120)$$

where of course $\partial_k = \partial_{x_1}, \ldots, \partial_{x_n}$
Therefore, up to a time-change depending on the size S, the diffusion X_t is a Brownian motion with respect to g plus an additional drift proportional to:

$$V = \sum_k (\partial_k S)\partial_k$$

But it is easy to check that $\forall(\lambda, R, v) \in G$, if we denote $\theta \equiv \theta_{(\lambda,R,v)}$ then:

$$d\theta_{(x,y,z)}(V) = V(\theta(x,y,z))$$

Therefore the projection of this vectorfield is constant along each fibre and hence $BM(M,g_E)$ projects down to an autonomous diffusion on Σ_n^3.

Remark
If we start off with the OU-process on M instead of Brownian motion, the projected result on Σ_n^3 will still be the same for the only change is the addition of a vectorfield

$$\xi(x,y,z) = -k(x,y,z) \quad (k > 0)$$

on M; however since this vectorfield is tangent to the fibre its contribution will vanish once we apply the projection (cf. Kendall [11]).

Chapter 4

Skew-product Decompositions of Brownian Motions

4.1 Introduction

Up till now we have been interested in the projection of a Brownian motion on a Riemannian manifold onto a lower-dimensional space. By the very nature of this mapping, some information on the original Brownian motion is lost: it becomes impossible, using the image only, to reconstruct the process we started off with. It is instructive to have another look at the simple examples which furnished the motivation for the submersiontheorem.

Example 1 (cfr. p. 13)

The decomposition $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is symmetric in its arguments and the projection

$$\pi_2 \; : \; \mathbb{R}^2 \longmapsto \mathbb{R} \; : \; (x, y) \longmapsto y$$

enjoys the same properties as the projection on the x-axis. In particular, it is a Riemannian submersion and the image $\pi_2(\beta) = \beta_2$ is standard $BM(\mathbb{R})$.

Example 2 (cfr. p. 14) In this case we are looking at the decomposition:

$$\mathbb{R}^2_o = \mathbb{R}^+_o \times S^1 \qquad (S^1 = \mathbb{R}/2\pi\mathbb{Z})$$

here the projection on the second component (θ) has a less trivial structure. Basically, this is due to the fact that the symmetry between the two projections breaks down; for whereas

$$\pi_1 \; : \; \mathbb{R}^2_o \longrightarrow \mathbb{R}^+_o \; : \; (x, \theta) \longmapsto r$$

is a Riemannian submersion, this is no longer true for

$$\pi_2 \; : \; \mathbb{R}^2_o \longrightarrow S^1 \; : (r, \theta) \longrightarrow \theta$$

(both spaces are assumed to be equipped with the usual Euclidean metric).
Yet, a straightforward application of Itô's formula to the transformationrules

$$x = r cos\theta$$
$$y = r sin\theta$$

yields, together with the autonomous SDE for the radial process:

$$dr_t = db_t + \frac{1}{2r_t}dt, \tag{4.1}$$

the following simple SDE for the angular part:

$$d\theta_t = \frac{1}{r_t}d\tilde{b}_t \tag{4.2}$$

where $\tilde{b} = BM(\mathbb{R})$ independent of b.
This equation tells us that $\theta_t = \pi_2(\beta_t)$ is itself a Brownian motion, albeit one whose clock depends on the radial process r_t. All in all, the equations (4.1) &(4.2) give a surprisingly transparant and appealing representation of $BM(\mathbb{R}^2)$.

The aim of this chapter is to investigate conditions which will guarantee the existence of such a decomposition in a more general setting.

More precisely we are interested in Riemannian manifolds (M, g) which have the product-form:

$$M = X \times Y$$

Associated with this topological decomposition are two canonical projections:

$$\begin{aligned} \pi_1 &: M \longrightarrow X \\ \pi_2 &: M \longrightarrow Y \end{aligned}$$

Let β be the Brownian motion on M determined by the metric g. We want to know what conditions will allow a skew-product of β with respect to X and Y, by which we mean:

(i) The projection $\pi_1(\beta)$ is an autonomous diffusion on X

(ii) at time t, the process $\pi_2(\beta)$ is Brownian motion independent of $\pi_1(\beta)$.

4.2 The Decomposition-theorem

Throughout this chapter we will assume that the statespace of Brownian motion is a Riemannian manifold (M, g) which has a product form:

$$M = X \times Y \tag{4.3}$$

where X and Y are connected manifolds.
The corresponding projection on the components are denoted by π_1 and π_2.

If we want to ensure that $\beta = BM(M)$ allows a skew-product decomposition with respect to X and Y, then we must put rather severe restrictions on the metric g. The following simple considerations bear this out:

Let $M = \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ $(n = p + q)$ be equipped with an arbitrary Riemannian metric $g = (g_{ij})$. The corresponding Brownian motion is generated by the Laplacian:

$$
\begin{aligned}
\Delta &= \frac{1}{\sqrt{detg}}\partial_i(\sqrt{detg}g^{ij}\partial_j) \\
&= g^{ij}\partial^2_{ij} + (\partial_i g^{ij} + \frac{1}{2}g^{ij}\partial_i(log\ det\ g))\partial_j
\end{aligned} \tag{4.4}
$$

Since we want the processes on the two components to be independent, we must insist on:

$$
g^{ij} = 0 = g^{ij} \quad \text{if } 1 \leq i \leq p \text{ and } p+1 \leq j \leq n
$$

which of course implies that the metric g on M factorizes in

$$
g = \begin{pmatrix} \rho(x,y) & 0 \\ 0 & \gamma(x,y) \end{pmatrix} \qquad \rho \in I\!R^{p\times p}, \gamma \in I\!R^{q\times q}. \tag{4.5}
$$

Furthermore, since we want $\pi_1(\beta)$ to be an autonomous diffusion we must make sure that the induced metric $\pi_1(g) = \rho$ on X coincides with the original metric g_X on X. In particular this implies that ρ must be independent of $y = (x_{p+1}, \ldots x_n)$. These two considerations cut down the possible forms of the metric g to:

$$
g = \begin{pmatrix} \rho(x) & 0 \\ 0 & \gamma(x,y) \end{pmatrix} \tag{4.6}
$$

Introducing generic notation for the indices:

$$
k, l, \ldots = 1 \ldots p \qquad (x\text{- component})
$$

$$
u, v, \ldots = p+1 \ldots n \qquad (y\text{-component})
$$

we see that in view of (4.6) we get:

$$
log\ det\ g(x,y) = log\ det\ g(x) + log\ det\ \gamma(x,y)
$$

Hence:

$$
\partial_u(log\ det\ g) = \partial_u(log\ det\ \gamma). \tag{4.7}
$$

Substituting (4.6) and (4.7) in (4.4) we see that we get the following decomposition for the Laplacian Δ_M:

$$\begin{aligned}\Delta_M &= \rho^{kl}\partial^2_{kl} + (\partial_k\rho^{kl})\partial_l + \frac{1}{2}\rho^{kl}\partial_k(log\ det\ g)\partial_l \\ &+ \gamma^{uv}\partial^2_{uv} + (\partial_u\gamma^{uv})\partial_v + \frac{1}{2}\gamma^{uv}\partial_u(log\ det\ \gamma)\partial_v \end{aligned} \tag{4.8}$$

It is interesting to note that the last three terms in (4.8) formally correspond to the Laplacian associated with the metric

$$\gamma_x(y) = \gamma(x, y)$$

(where x is interpreted as a parameter). However, the first three terms fail to specify an autonomous differential operator on the X-manifold since, in general, the terms involving the determinant will depend on y:

$$\partial_u\partial_k(log\ det\ g) \neq 0 \tag{4.9}$$

The most general C^∞-solution to the equation:

$$\partial_u\partial_k(log\ det\ g) = 0 \quad \forall u, k \tag{4.10}$$

is of the form

$$log\ det\ g = \varphi(x) + \psi(y) \tag{4.11}$$

where $\varphi : I\!R^p \longrightarrow I\!R$ and $\psi : I\!R^q \longrightarrow I\!R$ are arbitrary C^∞- functions.
But eq (4.11) implies that the determinant of the metric factorizes:

$$det\ g = e^{\varphi(x)}e^{\psi(x)} \tag{4.12}$$

Therefore we come to the conclusion that starting with a metric of the form (4.6) the factorizationcondition (4.12) is a necessary and sufficient condition to obtain an autonomous diffusion on the X-manifold.
But from the submersiontheorem we recall that $\pi_1(\beta)$ will be an autonomous diffusion if in addition to (4.6) we impose the condition that π_1 has fibres the mean curvature of which have a constant projection along each fibre. It therefore transpires that for metrics of the form (4.6) this proprety of π_1 must be equivalent to

the factorizationcondition (4.12). A direct proof of this will be given in proposition 2.

Finally we notice that by the very definition of a productmanifold it is possible to construct for each point $p \in M$ (open) coordinate-patches U and V such that:

(i) $$p \in U \times V$$

(ii) there exists coordinatemaps

$$\begin{aligned} x : \ U \subset X &\longrightarrow \mathbb{R}^p \\ y : \ V \subset Y &\longrightarrow \mathbb{R}^q. \end{aligned}$$

Hence the above analysis also applies to the general case.

Let us therefore summarize these facts in the following theorem:(cfr. Pauwels & Rogers [19])

Theorem 1

Let $M = X \times Y$ be a n-dimensional connected product manifold equipped with a Riemannian metric g (dim $X = p$, dim $Y = q$).
Moreover we assume that X (resp. Y) is covered with (open) coordinatepatches (U, x) (resp (V, y)) as defined by (i) and (ii) above.
Furthermore, we assume that in terms of the above coordinates the metric tensor g takes the form:

$$g = \begin{pmatrix} \rho(x) & 0 \\ 0 & \gamma(x,y) \end{pmatrix} \tag{4.13}$$

which is equivalent to saying that:

(i) at each point $p = (x, y) \in M$ we have an orthogonal decomposition of the tangentspace:

$$T_pM = T_pX_y \oplus T_pY_x$$

where of course $X_y = \{(x, y) \in M \mid x \in X\}$
and $\quad Y_x = \{(x, y) \in M \mid y \in Y\}$

(ii) The canonical projection on the first component

$$\pi_1 \; : \; M \longrightarrow X$$

is a Riemannian submersion with respect to the metric ρ on X.

Then, the Laplacian Δ_M is of the form:

$$\Delta_M = G_x + \Delta_{\gamma_x(y)} \tag{4.14}$$

where G_x is an autonomous second-order differential operator on X and $\Delta_{\gamma_x(y)}$ is the Laplacian associated with the metric $\gamma_x(y) = \gamma(x, y)$

if and only if

the volume-element associated with the metric g factorizes, i.e. : there exist (C^∞)functions:

$$\xi : \mathbb{R}^p \longrightarrow \mathbb{R} \qquad \text{and} \quad \eta : \mathbb{R}^q \longrightarrow \mathbb{R}$$

such that

$$det\ g(x, y) = \xi(x)\eta(y) \tag{4.15}$$

□

Remark: Because of the orthogonality condition (4.13) we have that

$$det\ g(x, y) = det\ \rho(x) \cdot det\ \gamma(x, y) \tag{4.16}$$

and hence condition (4.15) is actually equivalent to the factorization of $det\ \gamma$:

$$det\ \gamma(x, y) = \mu(x)\eta(y) \qquad \text{for some } \mu : \mathbb{R}^p \longrightarrow \mathbb{R} \tag{4.17}$$

Hence:

$$log\ det\ g(x,y) = log\ det\ \rho(x) + log\ det\ \mu(x) + log\ det\ \eta(y)$$

from which it follows that we can rewrite (4.8) as:

$$\begin{aligned} \Delta_M &= \rho^{kl}\partial^2_{kl} + (\partial_k \rho^{kl})\partial_l + \frac{1}{2}\rho^{kl}\partial_k(log\ det\ \rho)\partial_l + \frac{1}{2}\rho^{kl}\partial_k(log\ det\ \mu)\partial_l \\ &+ \Delta_{\gamma(x,y)} \end{aligned} \tag{4.18}$$

Comparison with (4.14) tells us that:

$$G_x = \Delta_X + \frac{1}{2}\rho^{kl}\partial_k(log\ det\ \mu)\partial_l \tag{4.19}$$

Let us now return to the observation we made earlier: namely that the factorizationcondition (4.15) must be equivalent to the fact that the mean curvature must be a fibre-invariant for the π_1-fibres.
In fact we can prove the following result:

Proposition 2

Let (M, g) be a Riemannian manifold of productform:

$$M = X \times Y$$

where dim $X = p$, dim $Y = q$ and dim $M = p + q = n$, and

$$\bar{u} : M \longrightarrow X : p = (x,y) \longmapsto x$$

the canonical projection on the first component.
Furthermore, we assume that the metric g is of the form:

$$g(x,y) = \begin{pmatrix} \rho(x) & 0 \\ 0 & \gamma(x,y) \end{pmatrix} \tag{4.20}$$

Then the condition that $det\ \gamma$ factorizes; i.e.:

$$det\ \gamma(x,y) = \varphi(x)\psi(y) \tag{4.21}$$

for some appropriate C^∞-functions $\varphi : \mathbb{R}^p \longrightarrow \mathbb{R}$ and $\psi : \mathbb{R}^q \longrightarrow \mathbb{R}$, is equivalent with the statement that the mean curvature vectorfields on the fibres

$$\pi^{-1}\{x\} = \{(x,y) \in M \mid y \in Y\} \equiv Y_x$$

are independent of y.

Before embarking on the proof of proposition 2 we formulate a simple lemma which we will need:

Lemma 3

Let $S = S(t)$ be a positive-definite, symmetric (and hence invertible) matrix, which is a smooth function of the variable t.
Then:

$$\frac{d}{dt}(ln\ det\ S) = Tr(S^{-1}\frac{dS}{dt}) \tag{4.22}$$

Proof:
Rewriting S as:

$$S = Ue^\Lambda U^T \qquad (\Lambda \text{ diagonal, } U \text{ orthogonal})$$

we get:

$$\frac{d}{dt}(ln\ det\ S) = \sum \frac{d\lambda_i}{dt}$$

On the other hand, the RHS of (4.22) is equal to:

$$
\begin{aligned}
&Tr[Ue^{-\Lambda}U^T(\dot{U}e^{\Lambda}U^T + U\dot{\Lambda}e^{\Lambda}U^T + Ue^{\Lambda}\dot{U}^T)] \\
&= Tr[U^T\dot{U} + \dot{\Lambda} + U\dot{U}^T] \\
&= Tr\dot{\Lambda} \qquad \text{since} \quad U^TU = I
\end{aligned}
$$

$\square$

Proof (of proposition 2):

$\Downarrow$: To make notation more transparent we will again adopt the convention that:

$i, j, k, \ldots$: are 'general' indices ranging from 1 through n,
$a, b, c, \ldots$: are 'horizontal' indices ranging from 1 through p,
$u, v, w, \ldots$: are 'vertical' indices ranging from p+1 through n.

Next, we choose in the neighbourhood of each point a basis of orthonormal vectorfields as follows:

(i) $E_1, \ldots, E_p$: horizontal, i.e. linear combinations of the vectorfields $\partial_1, \ldots, \partial_p$

(ii) $E_{p+1}, \ldots, E_n$: vertical, i.e. linear combinations of $\partial_{p+1}, \ldots, \partial_n$

If we collect the vectorfields in columns:

$$\partial = \begin{pmatrix} \partial_1 \\ \cdot \\ \cdot \\ \cdot \\ \partial_n \end{pmatrix} \quad \text{and} \quad E = \begin{pmatrix} E_1 \\ \cdot \\ \cdot \\ \cdot \\ E_n \end{pmatrix} \tag{4.23}$$

this means that there is an invertible matrix

$$S = \begin{pmatrix} U & 0 \\ 0 & R \end{pmatrix} \qquad U \in \mathbb{R}^{p \times p}, R \in \mathbb{R}^{q \times q}, \tag{4.24}$$

such that:

$$E = S\partial \tag{4.25}$$

Notice that since the E-vectorfields are orthonormal it follows from

$$\partial = S^{-1}E$$

that

$$g_{ij} = \sum_k (S^{-1})_{ik}(S^{-1})_{jk} = ((S^{-1})(S^{-1})^T)_{ij}$$

In particular:

$$U^T U = \rho^{-1} \quad \text{and} \quad R^T R = \gamma^{-1} \tag{4.26}$$

Next we recall that the mean curvature vector of the π_1-fibres is given by:

$$H(x,y) = \sum_a g(\sum_u \nabla_{E_u} E_u, E_a) E_a \tag{4.27}$$

Because of (4.25) this becomes:

$$H = \sum_a \sum_u \sum_{v,w} \sum_{b,c} R_{uv} R_{uw} U_{ab} U_{ac} g(\nabla_v \partial_w, \partial_b) \partial_c$$

Summing over a and u and using (4.27) we get:

$$H = \sum_{v,w} \sum_{b,c} (\gamma^{-1})_{vw} (\rho^{-1})_{bc} g(\nabla_v \partial_w, \partial_b) \partial_c \tag{4.28}$$

and since:

$$\nabla_v \partial_w = \Gamma^a_{vw} \partial_a + \Gamma^u_{vw} \partial_u \qquad \text{(sum. conv)}$$

we obtain (using the standardnotation $(\gamma^{-1})_{vw} = \gamma^{vw}$ etc ...)

$$H = \sum_a \sum_{v,w} \sum_{b,c} \gamma^{vw} \rho^{bc} \Gamma^a_{vw} \rho_{ab} \partial_c$$

or again (after successive summation over b and c):

$$H = \sum_a \sum_{v,w} \gamma^{vw} \Gamma^a_{vw} \partial_a . \tag{4.29}$$

Finally, using the well-known equation for the Christoffel-symbols:

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} (\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk})$$

it follows immediately that:

$$\Gamma^a_{vw} = -\frac{1}{2}\rho^{ab}\partial_b\gamma_{vw} \qquad \text{(sum. conv)}, \tag{4.30}$$

whence:

$$H = -\frac{1}{2}\sum_{a,b}\rho^{ab}(\sum_{v,w}\gamma^{vw}\partial_b\gamma_{vw})\partial_a. \tag{4.31}$$

All that remains to be done to complete the proof of this first part, is to show that:

$$\sum_{v,w}\gamma^{vw}(\partial_b\gamma_{vw})$$

is independent of y (for all b). But this follows easily from the observation that:

$$\begin{aligned}
\sum_{v,w}\gamma^{vw}(\partial_b\gamma_{vw}) &= Tr[\gamma^{-1}(\partial_b\gamma^T)] \\
&= Tr[\gamma^{-1}(\partial_b\gamma)] \qquad (\gamma \text{ is symmetric}) \\
&= \partial_b(ln\ det\gamma) \qquad \text{(lemma 3)} \\
&= \partial_b(ln\ \varphi(x) + ln\ \psi(y)) \qquad \text{(by assumption)} \\
&= \partial_b ln\varphi(x),
\end{aligned}$$

which is a function of x only.

$\Uparrow$: Since the computation of the mean curvature (as in (4.31)) did not involve the factorizationcondition, the expression (4.31) still remains valid. In particular, the assumption that H is constant in the y-variable implies that:

$$\forall\ a = 1,\ldots,p : \sum_b \rho^{ab}\sum_{v,w}\gamma^{vw}(\partial_b\gamma_{vw})$$

depends on x only; or again (using lemma 3):

$$\partial_u(\sum_b \rho^{ab}\partial_b(ln\ det\gamma)) = 0 \qquad \forall\ a = 1\ldots p, \forall\ u = p+1\ldots n$$

Since $\rho = \rho(x)$ this condition becomes:

$$\sum_b \rho^{ab} \partial_u \partial_b (ln\ det\gamma) = 0 \qquad \forall\ a, \forall\ u. \tag{4.32}$$

Interpreting equation (4.32) as a matrixequation and multiplying both sides by ρ, we conclude:

$$\forall a, \forall u : \partial_a \partial_u ln\ det\gamma = 0.$$

The most general solution to this equation is given by:

$$ln\ det\gamma = \varphi(x) + \psi(y)$$

where φ and ψ are arbitrary smooth functions, which proves that:

$$det\gamma(x, y) = e^{\varphi(x)} e^{\psi(y)} \tag{4.33}$$

□

Let us summarize these results in the following theorem, which can be interpreted as a further specialization of the sumbmersiontheorem:

Theorem 4 (Decomposition - theorem)

Let $M = X \times Y$ be a n-dimensional product-manifold equipped with a Riemannian metric g.
Let β be the associated Brownian motion

(i) If $\pi_1 : M \longrightarrow X$ is a Riemannian submersion such that the mean curvature is a fibre-invariant then $\pi_1(\beta)$ is an autonomous diffusion on X;

(ii) If in addition the natural decomposition of tangentspaces:

$$T_p M = T_p X_y \oplus T_p Y_x \qquad (p = (x, y) \in M) \tag{4.34}$$

is an orthogonal decomposition, then β allows a skew-product decomposition of the form (4.14).

(iii) If the decomposition (4.34) is orthogonal and the projection π_1 is a Riemannian submersion, which means that the metric tensor g on M is of the form:

$$g = \begin{pmatrix} \rho(x) & 0 \\ 0 & \gamma(x,y) \end{pmatrix}$$

then the mean curvature is constant along fibres if and only if $det\gamma(x,y)$ factorizes (in the sense of equation (4.33)).

$\square$

Remark:
In many concrete examples it is rather straightforward to obtain the metric tensor. Condition (iii) provides us with a quick way to inspect the invariance of the mean curvature.

4.3 Some applications of the decomposition theorem

4.3.1 Hyperbolic Plane

Let us return for a moment to the disc-model for the hyperbolic plane:

$$\mathbb{H}_2(\lambda) = \{x \in \mathbb{R}^2 \mid \|x\|_E < \lambda\}$$

The introduction of geodesical polar coordinates (ρ, θ) allows us to interpret $\mathbb{H}_2^\circ(\lambda) := \mathbb{H}_2(\lambda)\backslash\{0\}$ as a product-manifold

$$\mathbb{H}_2^\circ(\lambda) = \mathbb{R}_\circ^+ \times S^1$$

equipped with the Riemannian metric:

$$g(\rho,\theta) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\lambda^2}{4} sh^2(\frac{2\rho}{\lambda}) \end{pmatrix}$$

Obviously this metric satisfies the conditions of theorem 4 and therefore $BM(\mathbb{H}_2)$ has a skew-product decomposition with respect to $\mathbb{R}^+$ and S^1.
A simple calculation shows that the associated Laplacian is given by

$$\triangle = \partial^2_{\rho\rho} + \frac{2}{\lambda} coth(\frac{2\rho}{\lambda})\partial_\rho + \frac{4}{\lambda^2 sh^2(\frac{2\rho}{\lambda})}\partial^2_{\theta\theta}$$

Hence, in addition to the autonomous diffusion for the radial part, governed by the SDE:

$$d\rho = db_1 + \frac{1}{\lambda} coth(\frac{2\rho}{\lambda})dt$$

we have an independent Brownian motion driving the angular process:

$$d\theta = \frac{2}{\lambda sh(\frac{2\rho}{\lambda})} db_2$$

(b_1 and b_2 are independent $BM(\mathbb{R})$).

4.3.2 Surfaces of Revolution

Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ be a strictly-positive C^∞-function and define the corresponding surface of revolution S by the parameter equations:

$$S : \begin{cases} x = \varphi(u) cos\theta & u \in \mathbb{R}, \theta \in S^1 \equiv \mathbb{R}/2\pi\mathbb{Z} \\ y = \varphi(u) sin\theta \\ z = u \end{cases}$$

The parameters u and θ provide us with a global coordinatesystem on the surface which we can interpret as a productmanifold:

$$S = \mathbb{R} \times S^1 \tag{4.35}$$

The ambient Euclidean space induces a metric on S:

$$ds^2 = (1 + (\dot{\varphi}(u))^2)du^2 + \varphi^2(u)d\theta^2 \quad (\dot{\varphi}(u) = \frac{d\varphi}{du}) \tag{4.36}$$

Since

$$detg = (1 + (\dot{\varphi}(u))^2)\varphi^2(u) \tag{4.37}$$

the metric g satisfies the conditions of theorem (4) and therefore we have a skewproduct-decomposition:

(i) the process $u(t)$ performs an autonomous diffusion governed by the SDE:

$$du(t) = \frac{1}{\sqrt{1+\dot{\varphi}^2}}db_1(t) + \frac{1}{2}\frac{\dot{\varphi}}{\varphi}\frac{(1+\dot{\varphi}-\varphi\ddot{\varphi})}{(1+\dot{\varphi}^2)^{\frac{3}{2}}}dt$$

(ii) the angular process θ_t is driven by an independent Brownian motion the clock of which depends on $\varphi(u_t)$:

$$d\theta(t) = \frac{1}{\varphi(u_t)}db_2(t)$$

4.3.3 Symmetric Matrices

Consider the vectorspace of symmetric matrices:

$$Sym(n) := \{S \in \mathbb{R}^{n \times n} \mid S^T = S\}$$

which is an $\frac{n(n+1)}{2}$-dimensional subspace of $\mathbb{R}^{n\times n}$.
From elementary linear algebra we know that it is possible to obtain an eigenvalue/eigenvector-decomposition of these matrices of the form:

$$S = U\Lambda U^T \tag{4.38}$$

where $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$ and $U \in SO(n)$.
For our purposes it suffices to restrict our attention to the case when all the eigenvalues are different:

$$M_1 = \{S \in Sym(n) \mid S \text{ has } n \text{ different eigenvalues}\} \tag{4.39}$$

Since this is an open subset of a manifold, it is a manifold in its own right. Moreover, the decomposition (4.38) of $S \in M_1$ is unique, up to two, essentially trivial, transformations: permutations of the eigenvalues (and corresponding eigenvectors), and changing the sign of an even number of eigenvectors. We can therefore interpret M_1 as a product manifold in the following sense: if

$$D_n = \{(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n \mid \lambda_1 < \lambda_2 < \ldots < \lambda_n\}$$

then

$$M = D_n \times SO(n) \tag{4.40}$$

is a multiple covering of M_1 (in the same sense as S^3 is a double covering of $SO(3)$), with the associated covering-map:

$$\begin{aligned} j \ : \ D_n \times SO(n) &\longrightarrow M \\ (\Lambda, U) &\longmapsto U\Lambda U^T \end{aligned} \tag{4.41}$$

This map can then be used to transport the usual Riemannian metric on $\mathbb{R}^{n\times n}$ to the productspace M. To do this we consider a differentiable curve

$$S(t) = U(t)\Lambda(t)U^T(t)$$

from which we get:

$$dS = U[U^T(dU)\Lambda + d\Lambda + \Lambda(dU^T)U]U^T \tag{4.42}$$

Next we define a matrixfunction $A(t)$ (for t sufficiently small) by the system of ODEs:

$$dA = U^T dU \tag{4.43}$$

with an initial condition such that:

$$e^{A(0)} = U(0)$$

Notice that since $U^T U = I$ it follows that

$$dA + dA^T = 0$$

and hence the skewsymmetry of $A(0)$ implies that $A(t)$ is skewsymmetric for all values of t (for which it is defined).
Using the standard innerproduct on $\mathbb{R}^{n \times n}$:

$$< A, B > := Tr(AB^T),$$

the squared norm of (4.42) becomes:

$$Tr(dSdS^T) = Tr[d\Lambda^2 + 2\Lambda dA\Lambda dA - 2\Lambda^2(dA)^2] \tag{4.44}$$

A straightforward computation shows that the last two terms of the RHS of (4.44) are equal to:

$$\sum_i \sum_k \lambda_k(\lambda_k - \lambda_i) da_{ik}^2$$

Interchanging the rôle of i and k we get a similar equation and the sum of these two shows that:

$$2Tr[\Lambda dA\Lambda dA - \Lambda^2 dA^2] = \sum_i \sum_k (\lambda_k - \lambda_i)^2 da_{ik}^2 \tag{4.45}$$

Substituting (4.45) into (4.44) shows that the induced metric on M can be written as:

$$\begin{aligned} ds^2 &= \sum_i d\lambda_i^2 + \sum_{i \neq k} (\lambda_k - \lambda_i)^2 da_{ik}^2 \\ &= \sum_i d\lambda_i + 2\sum_{i<k} (\lambda_k - \lambda_i)^2 da_{ik}^2 \end{aligned} \tag{4.46}$$

One look at the metric tensor (4.46) convinces us that Brownian motion on M must have a skew-product decomposition.

A simple calculation then shows that the Laplacian can be split in two parts: one part referring to the λ_i-coordinates which is of the form:

$$\sum_{i=0}^{n} [\frac{\partial^2}{\partial \lambda_i^2} + \sum_{k \neq i} \frac{1}{(\lambda_i - \lambda_k)} \frac{\partial}{\partial \lambda_i}] \tag{4.47}$$

and a second part which pertains to the a_{ik}-variables ($i < k$) and is of the form:

$$\forall i < k : \qquad \frac{1}{2(\lambda_i - \lambda_k)^2} \frac{\partial^2}{\partial a_{ik}^2} \tag{4.48}$$

In terms of SDEs this gives us the following picture:

$$d\lambda_i = db_i + \frac{1}{2} \sum_{k \neq i} \frac{dt}{(\lambda_i - \lambda_k)} \tag{4.49}$$

$$da_{ik} = \frac{1}{\lambda_i - \lambda_k} d\tilde{b}_{ik} \quad (i < k) \tag{4.50}$$

$$da_{ki} = -da_{ik} \tag{4.51}$$

$$da_{ii} = 0 \tag{4.52}$$

where all the b_i and $\tilde{b}_{ij}$ are independent $BM(\mathbb{R})$.
Because of (4.43) we can recover the eigenvectorprocess $U(t)$ by solving the Stratonovich SDE

$$\partial U = U \partial A$$

For another derivation of (4.49) & (4.50) and a proof of the non-explosion of (4.49) see Williams [28].

4.3.4 Normal matrices

The striking and rather unexpected skew-product decomposition that was obtained for the symmetric matrices prompts us to try and generalize the results to larger classes of matrices, such as the Hermitian, the unitary or better still the normal matrices.
Recall that by definition a matrix $N \in \mathbb{C}^{n\times n}$ is normal if it commutes with its (complex) adjoint:

$$NN^* = N^*N \tag{4.53}$$

A well-known result in linear algebra tells us that the normal matrices constitute the largest class of matrices that are diagonalizable with respect an orthonormal basis.

Hence

$$\begin{gathered}\text{N is normal}\\ \underline{\text{iff}}\\ \exists\ \Lambda = diag(\lambda_i|\lambda_i \in \mathbb{C}), \exists\ U \in U(n):\ \ N = U\Lambda U^*\end{gathered} \tag{4.54}$$

Again it is easy to see that this decomposition isn't unique: even if we rule out permutations of the eigenvalues, we still have quite some leeway left as far as the eigenvectors are concerned:
for if

$$N = U\Lambda U^* = V\Lambda V^* \qquad (U, V \in U(n)) \tag{4.55}$$

then

$$(V^*V)\Lambda = \Lambda(V^*V) \tag{4.56}$$

Again we will restrict our attention to the case where all the eigenvalues are different and then (4.56) implies that V^*V must be diagonal. However since this matrix is also unitary we find that

$$V = U\begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n}\end{pmatrix} \quad \text{for some } \theta_1, \ldots, \theta_n \in \mathbb{R} \tag{4.57}$$

This means that the matrix of eigenvectors U is determined up to an element in the n-dimensional torus:

$$T^n := \{diag(e^{i\theta_1} \dots e^{i\theta_n}) \mid \theta_1 \dots \theta_n \in \mathbb{R}\} \cong S^1 \times \dots \times S^1 \tag{4.58}$$

which obviously is a (closed) subgroup of $U(n)$.
Hence if we want to interpret

$$M = \{N \in \mathbb{C}^{n\times n} \mid N \text{ is normal with } n \text{ different eigenvalues}\} \tag{4.59}$$

as a product manifold, then we end up with:

$$M = D_n(\mathbb{C}) \times (U(n)/T^n) \tag{4.60}$$

where of course

$$D_n(\mathbb{C}) = \{diag(\lambda_1 \dots \lambda_n) \in \mathbb{C}^{n\times n} \mid \lambda_1 \prec \dots \prec \lambda_n\}$$

and $\prec$ is some way of ordering a finite set of complex numbers,
eg:

$$\begin{array}{lll} \lambda \prec \mu & \underline{\text{iff}} & Re\lambda < Re\mu \\ & & \text{or } Re\lambda = Re\mu \ \text{ but } \ Im\lambda < Im\mu \end{array}$$

(If we dispense with the above ordering, the product in (4.60) becomes a multiple covering of M).

A skew-product decomposition on M will therefore involve the construction of a (nonautonomous) Brownian motion on the homogeneous space $U(n)/T^n$. Although this approach doesn't pose any conceptual problems, to actually work this out in detail is going to be quite messy. The reason for this is that an homogeneous space is a very clumsy object; there is no nice global chart and indeed the only tidy way to write it down is as a quotient.
Therefore the clean way to handle this situation is to look at the manifold

$$N = D_n(\mathbb{C}) \times U(n) \tag{4.61}$$

instead, which we obtained by replacing the homogenous space by its Lie-group. The point in doing this is given by theorem 8 in chapter 3 : if the metric on $U(n)/T^n$ is invariant under the (left)action of $U(n)$ then we can construct a left-invariant metric on $U(n)$ (bi-invariant under T^n) such that $BM(U(n))$ will project down to $BM(U(n)/T^n)$. Hence the skewproduct decomposition on N will give rise to a skewproduct decomposition on M.
All we have to do is to check that the action of $U(n)$ on $U(n)/T^n$ is effective; but this is trivial since $U(n)$ acts effectively on itself.

Let us go back to the example at hand: to put a Riemannian structure on M we take an arbitrary element

$$(\Lambda, U) \in D_n(\mathbb{C}) \times U(n)$$

and consider a differentiable curve through this point:

$$N(t) = U(t)\Lambda(t)U^\star(t)$$

Differentiating with respect to t yields:

$$dN = U(d\Lambda + \Lambda dA - dA\Lambda)U^\star$$

where again $A(t)$ is a skew-Hermitian matrixfunction defined by

$$dA = U^\star dU$$

with the appropriate initial condition.
Using the standard innerproduct on $\mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2}$ we obtain

$$\begin{aligned} ds^2 &= Re(Tr(dNdN^\star)) \\ &= \sum_i |d\lambda_i|^2 + 2\sum_{i<j}\sum |\lambda_i - \lambda_j|^2 |da_{ij}|^2 \end{aligned} \tag{4.62}$$

Notice that unlike in the case of the symmetric matrices (4.62) does not specify a Riemannian metric on N since it does not involve the diagonal elements a_{ij}.

Still the form of (4.62) makes it obvious that Brownian motion on the normal matrices allows a skew-product decomposition in terms of Λ and A.

All we have to do is to compute the Laplacian associated with (4.62). Since we interpret M as a real manifold we get that the corresponding determinant to (4.62) equals:

$$2\prod_{i<j}|\lambda_i - \lambda_j|^4$$

From this it follows that the Laplacian becomes:

$$\begin{aligned} \triangle &= \sum_{i=1}^{n}\{\frac{\partial^2}{\partial\lambda_i\partial\bar{\lambda}_i} + 4\sum_{k\neq i}\frac{\lambda_i - \lambda_k}{|\lambda_i - \lambda_k)|^2}\frac{\partial}{\partial\lambda_i}\} && (4.63) \\ &+ \sum_{i\neq k}\frac{1}{|\lambda_i - \lambda_k|^2}\frac{\partial^2}{\partial a_{ik}\partial\bar{a}_{ik}} && (4.64) \end{aligned}$$

when we use the standard notation:

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial x} - i\frac{\partial}{\partial y};$$
$$\frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial x} + i\frac{\partial}{\partial y}.$$

So again we come to the conclusion that Brownian motion of normal matrices forces the eigenvalues to perform a diffusion the driftterm of which is determined by a repulsive interaction between the individual eigenvalues.

Hence the eigenvalues behave as diffusing electrically charged particles which repel each other with a force inversely proportional to the distance between them:

$$d\lambda_i = d\beta_i + 2\sum_{k\neq i}\frac{\lambda_i - \lambda_k}{|\lambda_i - \lambda_k|^2}dt \qquad (4.65)$$

(β_i independent $BM(\mathbb{C})$).

As for the eigenvectors, they are governed by the Stratonovich SDE:

$$\partial U = U \partial A \tag{4.66}$$

where

$$da_{ij} = \frac{db_{ij}}{|\lambda_i - \lambda_j|} \quad (i < j) \tag{4.67}$$
$$da_{ij} = -d\bar{a}_{ji} \quad (i > j)$$
$$da_{ij} = 0 \tag{4.68}$$

(b_{ij} are independent $BM(\mathbb{C})$, independent of β_i).
It's interesting to look at a related diffusionprocess on the normal matrices which we obtain by changing the dynamics of the eigenvalues: in (4.65) we introduce an elastic restoring force which pulls all the eigenvalues towards the origin:

$$d\lambda_i = d\beta_i - k\lambda_i dt + 2\sum_{k \neq i} \frac{\lambda_i - \lambda_k}{|\lambda_i - \lambda_k|^2} dt \tag{4.69}$$

The addition of this extra vectorfield has as a consequence the existence of a stationary probability measure for the eigenvalues.

Proposition

The eigenvalueprocess defined by (4.69) has a stationary probability measure which is defined (up to a normalisationconstant) by

$$f(z_1, \ldots, x_n) = \prod_{i<j} |z_i - z_j|^4 exp(-k \sum_{i=1}^{n} |z_i|^2)$$

Proof :

We introduce the notation:

$$\alpha(z_1, \ldots, z_n) = \prod_{i<j} |z_i - z_j|^4 = \prod_{i<j} ((x_i - x_j)^2 + (y_i - y_j)^2)^2 \tag{4.70}$$

$$\beta(z_1, \ldots, z_n) = exp(-k \sum |z_i|^2) = exp(-k \sum (x_i^2 + y_i^2)) \tag{4.71}$$

which we will interpret as real functions in the variables x_i, y_i. All we have to do is to show that

$$G^* f = 0 \tag{4.72}$$

where G^* is the adjoint generator of the process (4.69) and is given by:

$$\begin{aligned} G^* f &= \frac{1}{2}\sum_i \left(\frac{\partial^2 f}{\partial x_i^2} + \frac{\partial^2 f}{\partial y_i^2}\right) + k\sum_i \left[\frac{\partial}{\partial x_i}(x_i f) + \frac{\partial}{\partial y_i}(y_i f)\right] \\ &- 2\sum_i \sum_{k \neq i} \left[\frac{\partial}{\partial x_i}\left(\frac{(x_i - x_k) f}{|z_i - z_k|^2}\right) + \frac{\partial}{\partial y_i}\left(\frac{(y_i - y_k) f}{|z_i - z_k|^2}\right)\right] \end{aligned} \tag{4.73}$$

Straightforward calculations yield:

$$\frac{\partial \alpha}{\partial x_i} = 4\alpha \sum_{k \neq i} \frac{(x_i - x_k)}{|z_i - z_k|^2}$$

$$\frac{\partial \alpha}{\partial y_i} = 4\alpha \sum_{k \neq i} \frac{(y_i - y_k)}{|z_i - z_k|^2}$$

$$\frac{\partial \beta}{\partial x_i} = -2k x_i \beta$$

$$\frac{\partial \beta}{\partial y_i} = -2k y_i \beta$$

Substituting this into (4.73) proves (4.72).

□

A proof of the nonexplosion of the eigenvalue-process and similar results for the Hermitian and unitary matrices can be found in Pauwels [15] and Pauwels & Rogers[19].

References

[1] W.M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, New York,1975.

[2] N. Bourbaki, *Eléments de Mathématique: Groupes et algebres de Lie*, Herman, Paris,1960-1975.

[3] Bang-yen Chen, *Total Mean Curvature and Submanifolds of Finite Type*, Series in Pure Mathematics - volume 1, World Scientific Publishing Company, 1984.

[4] M.L. Curtis, *Matrixgroups*, Springer-Verlag, 1984.

[5] E.B. Dynkin, *Markov processes I*, Springer-Verlag, Berlin, 1965.

[6] R. Gillmore, *Liegroups, Liealgebras and some of their Applications*,

[7] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press,1978.

[8] N. Ikeda & S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publishing Cy, 1981.

[9] E. Jørgensen, *The Central Limit Theorem for Geodesic Random Walks*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 32 (1975), 1 - 64.

[10] D.G. Kendall, *Shape Manifolds, Procrustean Metrics and Complex Projective Spaces*, Bull. London. Mathematical Society 16 (1984), 81-121.

[11] W.S. Kendall, *Symbolic Computation and the Diffusion of Shapes of Triads*, preprint.

[12] J.T. Lewis, *Brownian Motion on a Submanifold of Euclidean Space*, Bull. London Math. Soc. 18, 616-20, 1986.

[13] P.J. Olver, *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics 107, Springer-Verlag, New York, 1986.

[14] B. O'Neill, *Semi-Riemannian Geometry*, Pure and Applied Mathematics 103, Academic Press, 1983.

[15] E.J. Pauwels, *Some Diffusion Processes on the Hermitian and Unitary Matrices*, Proc. 4th EYSM, Varna,1983.

[16] E.J. Pauwels, *Riemannian Submersions of Brownian Motions*, to appear in Stochastics.

[17] E.J. Pauwels, *An elementary characterization of Riemannian submersions*, in preparation.

[18] E.J. Pauwels, *Bi-invariant Brownian motions on the rotationgroups*, in preparation.

[19] E.J. Pauwels & L.C.G. Rogers, *Skew-product Decompositions of Brownian motions*, Contemporary Mathematics, vol 73,1988.

[20] M. Pinsky, *Stochastic Riemannian Geometry*, in: Probabilistic Analysis and Related Topics 1, ed.: Bharucha-Reid, Academic Press, New York, 1978.

[21] W.A. Poor, *Differential Geometric Structures*, McGraw-Hill,1981.

[22] L.C.G. Rogers & David Williams, *Diffusions, Markov Processes and Martingales, volume 2: Itô Calculus*, Wiley, 1987.

[23] A. Sagle & R. Walde, *Introduction to Liegroups and Liealgebras*, Academic Press, New York, 1973.

[24] D.H. Sattinger & O.L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry and Mechanics*, Applied Mathematical Sciences vol. 61, Springer-Verlag, 1986.

[25] M. Spivak, *A Comprehensive Introduction to Differential Geometry I*, Publish or Perish,1970.

[26] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer-Verlag,1983.

[27] D. Williams, *Diffusions, Markov Processes and Martingales, volume 1: Foundations*, Wiley,1979.

[28] D. Williams, *Brownian Motion on the Symmetric Matrices*, preprint.

[29] J. Cheeger & D.G. Ebin, *Comparison Theorems in Riemannian Geometry*, North-Holland Mathematical Library vol.9, North-Holland,1975.

[30] M.Crampin & F.A.E. Pirani,*Applicable Differential Geometry*, London Mathematical Society Lecture Notes Series 59, Cambridge University Press, 1986.

Samenvatting

In deze thesis bestuderen we een specifiek voorbeeld van de Itô transformatie formule in stochastische meetkunde: we onderzoeken hoe een Brownse beweging op een Riemannse varieteit getransformeerd wordt wanneer we die varieteit door middel van een submersie afbeelden op een andere. Meer bepaald zijn we geinteresseerd in het volgende probleem.

Zij (M, g) en (B, h) twee eindig-dimensionale Riemannse varieteiten en

$$\pi : M \longrightarrow B$$

een C^∞-submersie.
We formuleren dan het

Submersie-probleem :

Als we de Brownse beweging op M noteren met β_t, dan wensen we de voorwaarden te bepalen waaronder $\pi(\beta_t)$ een autonoom diffusieproces op B zal zijn.

In hoofdstuk 1 herhalen we enkele belangrijke begrippen uit de differentiaalmeetkunde en de stochastische analyse.
In hoofdstuk 2 keren we dan terug naar het submersie-probleem en bepalen we nodige en voldoende voorwaarden opdat $\pi(\beta_t)$ een autonome diffusie zou zijn. Het blijkt dat twee factoren hierbij essentieel zijn: ten eerste moet π een Riemannse submersie zijn. Bovendien moeten de vezels van π een gemiddelde kromming hebben die zodanig is dat de projectie van dit vectorveld onder $d\pi$ steeds hetzelfde resultaat oplevert voor de verschillende punten binnen eenzelfde vezel. Als dat het geval is dan bepaalt de projectie van de gemiddelde kromming een vectorveld op de basisruimte B en $\pi(\beta_t)$ blijkt een Brownse beweging te zijn op B met dit vectorveld als drift.
Indien we Brownse beweging interpreteren als de limiet van een stochastische wandeling op de varieteit is het mogelijk om dit resultaat op een intuitieve wijze uit te leggen. We dienen hiervoor echter wel een nieuwe interpretatie van de gemiddelde krommingsvector in termen van deviatiesnelheid van geodeten af te leiden, alsook een andere karakterisatie van Riemannse submersies.
Vervolgens bekijken we Liegroepen. De meest voor de hand liggende manier om hierop Brownse beweging te definieren bestaat erin dat we het proces construeren op de corresponderende (lineaire) Lie-algebra en het vervolgens overdragen naar de groep door middel van de exponentiele afbeelding. Dit is echter slechts mogelijk in een zeer beperkt aantal gevallen. Meestal moeten we immers een bijkomende drift toevoegen om te compenseren voor de kromming van de 1-parameter deelgroepen. Dit kan ook weer geinterpreteerd worden aan de hand van de overeenkomstige stochastische wandeling.
Tenslotte beschouwen we het probleem van de kanonieke en bi-invariante Brownse

beweging op Lie-groepen: bestaat er een natuurlijke keuze voor de metriek (en dus Brownse beweging) op een Lie-groep? Hiermee bedoelen we een keuze die geheel bepaald wordt door de algebraische karakteristieken van de groep en is dit resultaat invariant onder linkse en rechtse translaties? De Killing-metriek levert ons een antwoord op de eerste vraag. Wat de tweede vraag betreft zijn de zaken een stuk ingewikkelder: hiervoor hebben we nodig dat de toegevoegde Liegroep relatief compact is. Dit impliceert dat de 'meeste' Liegroepen geen bi-invariante metriek supporteren. De rotatiegroepen $SO(n)$ vormen hierop een uitzondering aangezien zij dit wel doen. Bovendien krijgen we nog een onverwachte bonus: deze bi-invariante metrieken zijn uniek behalve wanneer $n = 4$; de groep $SO(4)$ beschikt over een 2-parameter familie van zulke metrieken.

In hoofdstuk 3 concentreren we ons op enkele toepassingen van het submersietheorema. Besselprocessen worden bekeken zowel in de Euclidische als in de hyperbolische ruimte. De interpretatie van de driftterm in termen van de kromming van de vezels geeft een eenvoudige uitleg voor het transient gedrag van de Brownse beweging in het hyperbolisch vlak.
Daarna bekijken we de werking van enkele Lie-groepen op varieteiten. De meeste hiervan leveren geen verrassingen op, maar de werking van $SL(2, \mathbb{R})$ in het hyperbolisch vlak blijkt onverwacht ingewikkeld te zijn.
Tenslotte bekijken we in het kort de vorm-diffusie zoals die voorgesteld werd door D.G. Kendall. Het submersietheorema laat ons toe een kort bewijs te geven van enkele basisresultaten op dit gebied.

In hoofdstuk 4 bekijken we opnieuw het submersietheorema en bestuderen we hoe we de informatie kunnen terugkrijgen die we verloren bij de projectie van de hoogdimensionele varieteit naar een varieteit van een lagere dimensie. We beschouwen meer bepaald, produkt-varieteiten van de vorm

$$M = X \times Y$$

tesamen met de kanonieke projecties π_1 en π_2 van de componenten. We nemen hierbij aan dat π_1 voldoet aan de voorwaarden gespecifieerd in het submersietheorema. We bewijzen vervolgens het decompositietheorema dat stelt dat als het bovenstaande produkt orthogonaal is met betrekking tot de metriek op M, dan is de projectie $\pi_2(\beta)$ een nonautonome Brownse beweging op Y.
Zo'n decompositie wordt een *skew-product decomposition* genoemd, naar analogie met de welbekende beschrijving van $BM(\mathbb{R}^2)$ in termen van poolcoordinaten.
Het resultaat wordt dan toegepast op twee interessante matrix-klassen: de symmetrische en de normale matrices. Elementen van deze varieteiten kunnen gezien worden als produkten van diagonale en unitaire matrices. De decompositie zegt ons dat de eigenwaarden autonome diffusies vervullen terwijl de eigenvectors worden

gedreven door onafhankelijke witte ruis waarvan de klok afhangt van de afstand tussen de eigenwaarden. Het is verrassend te ontdekken dat de eigenwaarden zich gedragen als elektrisch geladen deeltjes op de reele lijn (in het geval van symmetrische matrices) of in het complex vlak (voor normale matrices) die elkaar afstoten met een kracht die omgekeerd evenredig is met hun onderlinge afstand. We kunnen proberen deze afstoting in evenwicht te brengen door een elastisch herstellende kracht toe te voegen die gericht is naar de oorsprong zodat een Ornstein-Uhlenbeck-proces verkregen wordt op deze matrixvarieteiten. Het wordt dan mogelijk de overeenkomstige stationaire waarschijnlijkheidsmaat te berekenen.