Stochastic calculus on manifolds

Part I: SDEs on Euclidean space / Basic notions for smooth manifolds

Geelon So, agso@eng.ucsd.edu
Sampling/optimization reading group — August 17, 2021

I. Review of SDEs on Euclidean space

Ordinary calculus

We can compute the value of X_t if we know its rate of change, by solving the ODE:

$$dX_t = b(t, X_t)\,dt.$$

Stochastic calculus

If there is white noise throughout the process, we have the SDE:

$$dX_t = \underbrace{a(t, X_t)\, dB_t}_{\text{white noise}} + \underbrace{b(t, X_t)\, dt}_{\text{deterministic drift}}$$

where B_t is a model of white noise (here and throughout, Brownian motion).

- Need to define an appropriate notion of a **stochastic integral**:

$$\int_0^t f(s)\, dB_s.$$

Stochastic integral

A natural way to define the integral is as a limit of simple (step) functions:

$$\int_0^t f(s)\, dB_s = \lim_{N\to\infty} \sum_{j=1}^{N} f(\tau_j) \cdot \left(B_{t_j} - B_{t_{j-1}}\right).$$

- Construct a step function by holding it constant at $f(\tau_j)$ over the interval $[t_{j-1}, t_j)$.

Ito and Stratonovich integrals

Unlike the Riemann-Stieltjes integral, the choice of τ_j makes a difference:

- When $\tau_j = t_{j-1}$, we obtain the **Ito integral**, $\int f(t)\, dB_t$.
  - Since it does not 'look ahead into the future', it has the intuitive stochastic property of being a martingale:
$$\mathbb{E}\left[\int_0^t f(t)\, dB_t\right] = 0.$$
  - However, the chain rule operates differently:
$$df(X_t) = \nabla f(X_t)\, dX_t + \frac{1}{2} dX_t^{\mathsf{T}}\, \mathbf{H} f(X_t)\, dX_t$$
- When $\tau_j = \frac{t_j - t_{j-1}}{2}$, we obtain the **Stratonovich integral**, $\int f(t) \circ dB_t$.
  - It is no longer a martingale, but the chain rule obeys the rules of ordinary calculus.
  - It is possible to transform these different integrals into each other.

Semimartinagles

We know how to take ordinary integrals and integrals driven by Brownian motion:

$$\int f(t)\,dt \qquad \text{and} \qquad \int f(t)\,dB_t.$$

The largest class of integrators that the Ito or Stratonovich integrals can be defined are called **semimartingales**.

- If Z_t is a semimartinagle, then we can define the stochastic integral:

$$\int f(t)\,dZ_t.$$

Diffusion process

Definition (Hsu (2002))

*A **diffusion process** X_t on $\mathbb{R}^N$ is given by:*

- *a locally-Lipschitz diffusion coefficient $\sigma : \mathbb{R}^N \to \mathbb{R}^{N\times\ell}$,*
- *a driving $\mathbb{R}^\ell$-semimartingale Z_t,*

where the integral form of X_t is given by the Ito integral:

$$X_t = X_0 + \int_0^t \sigma(X_s)\, dZ_s.$$

Usual form of diffusion

Example

The following stochastic process:

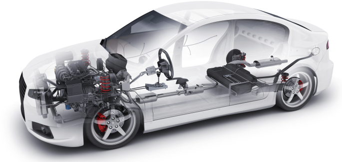
$$dX_t = a(X_t)\, dB_t + b(X_t)\, dt$$

is given by $\sigma(X_t) = \big(a(X_t), b(X_t)\big)$ *and* $Z_t = (B_t, t)^\mathsf{T}$.

An intuitive analogy

$$X_t = X_0 + \int_0^t \sigma(X_s)\, dZ_s$$

- X_t is the position of the car at time t.
- Z_t is the input to the steering wheel/pedal at time t.
- $\sigma(X_t)$ specifies how the input is converted into an instantaneous change in position.



Looking ahead: stochastic calculus on manifolds

What happens if we're not driving on a flat surface but a curved surface?

- Define calculus on manifolds (i.e. nonlinear spaces).
- Carry out the same driving analogy to define stochastic processes on manifolds.

II. Basic notions regarding smooth manifolds

Calculus on nonlinear spaces?

The **directional derivative** $D_u f$ on Euclidean space E is defined:

$$D_u f(x) = \lim_{t \downarrow 0} \frac{f(x + tu) - f(x)}{h}, \quad \forall x, u \in E.$$

- Notice that $x + tu$ assumes the existence of some linear structure.
- On a nonlinear space, how to specify a direction? What should replace "$x + tu$"?

Smooth manifolds

> Generally speaking, a manifold is a topological space that locally resembles Euclidean space. A smooth manifold is a manifold $\mathcal{M}$ for which this resemblance is sharp enough to permit the establishment of partial differential equation—in fact, all the essential features of calculus—on $\mathcal{M}$.
>
> *O'Neill (1983)*

General sketch

Introducing a smooth structure onto a space $\mathcal{S}$:

- Relate $\mathcal{S}$ to $\mathbb{R}^n$ by assigning coordinates $\xi(p) \in \mathbb{R}^n$ to points $p \in \mathcal{S}$.
- Define smoothness of functions on $\mathcal{S}$ with respect to the coordinate functions.

Example

Consider the upper half of the unit circle $S^\top$ in $\mathbb{R}^2$. Parametrize it by $\theta : (0, \pi) \to S^\top \subset \mathbb{R}^2$,

$$\theta(u) = \begin{bmatrix} \cos u \\ \sin u \end{bmatrix}.$$

We can say that $f : S^\top \to \mathbb{R}$ is smooth iff $f \circ \theta : (0, \pi) \to \mathbb{R}$ is smooth.

Coordinate systems

Definition (O'Neill (1983))

*A **coordinate system** or **chart** on a topological space $\mathcal{S}$ is a continuous map $\xi : U \to \xi(U)$ with continuous inverse, where $U \subset \mathcal{S}$ and $\xi(U) \subset \mathbb{R}^n$ are open sets.*

- If for each $p \in U$, we write:

$$\xi(p) = \left(x^1(p), \ldots, x^n(p)\right),$$

we say that $x^1, \ldots, x^n$ are the **coordinate functions** of ξ.

Coordinate systems

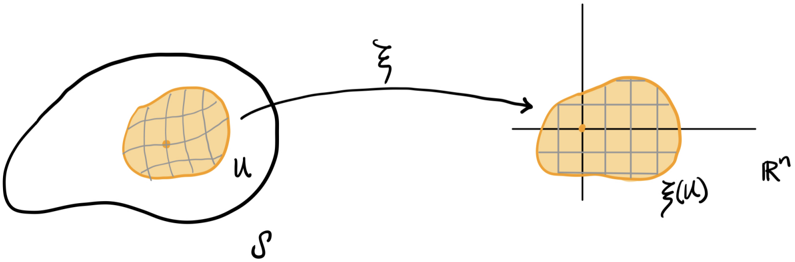


Figure 1: A chart ξ on the space $\mathcal{S}$ assigns coordinates to points on $\mathcal{S}$.

Smooth compatibility

Let $\xi : U \to \xi(U)$ and $\eta : V \to \eta(V)$ be two charts. Their **transition maps** are defined:

$$\eta \circ \xi^{-1} : \xi(U \cap V) \to \eta(U \cap V)$$
$$\xi \circ \eta^{-1} : \eta(U \cap V) \to \xi(U \cap V).$$

We say that ξ and η are **smoothly compatible** if their transition maps are smooth.

Smooth compatibility

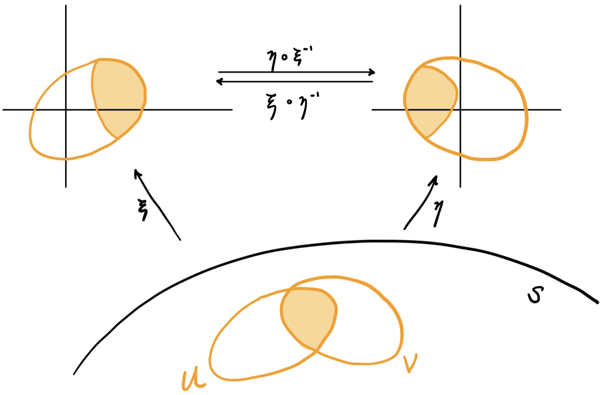


Figure 2: Two charts ξ and η are smoothly compatible if their transition maps are smooth (in the Euclidean sense).

Smooth manifold

Definition

An ***atlas*** $\mathcal{A}$ *on a topological space* $\mathcal{S}$ *is a collection of smoothly compatible charts such that for each* $p \in \mathcal{S}$*, there is a chart* (ξ, U) *such that* U *contains* p*.*

Definition

A ***smooth manifold*** $\mathcal{M}$ *is a (Hausdorff) space equipped with an atlas* $\mathcal{A}$*.*

- If the charts in $\mathcal{A}$ map to open sets in $\mathbb{R}^n$, we say that $\mathcal{M}$ is n-dimensional.

Smooth mappings

Smooth functions can then be defined with respect to the coordinate functions.

Smooth maps to $\mathbb{R}$

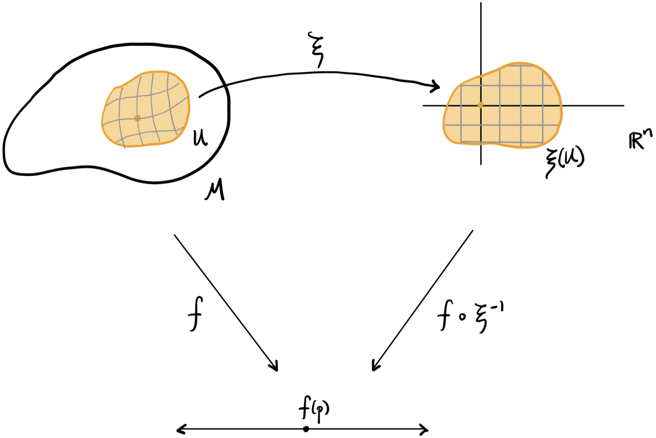


Figure 3: A map $f : \mathcal{M} \to \mathbb{R}$ is smooth if $f \circ \xi^{-1}$ is smooth for all charts $\xi \in \mathcal{A}$.

Smooth maps between manifolds

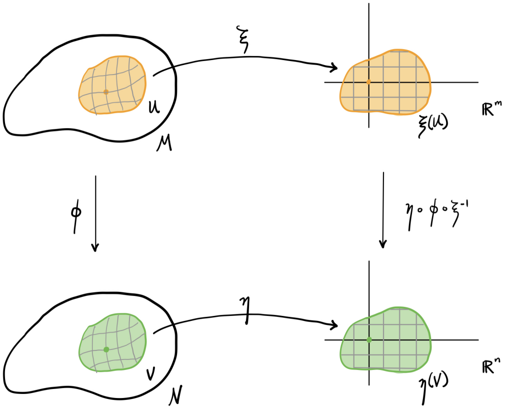


Figure 4: A map $\phi : \mathcal{M} \to \mathcal{N}$ is smooth if $\eta \circ f \circ \xi^{-1}$ is smooth for all charts $\xi \in \mathcal{A}_{\mathcal{M}}$ and $\eta \in \mathcal{A}_{\mathcal{N}}$. We say that ϕ is a **diffeomorphism** if it is smooth and has smooth inverse.

Tangent vectors

- Let $p \in \mathcal{M}$. Want to define $T_p(\mathcal{M})$ to be directions we can travel along on $\mathcal{M}$ at p.
- Equivalently, we can ask how fast do the values of smooth functions $f \in C^\infty(\mathcal{M})$ change at p along certain directions?
  - We can define **tangent vectors** as **directional derivatives**.

Axiomatizing the directional derivative

We can axiomatize the derivative as a linear operator satisfying the product rule:

Definition (O'Neill (1983))

*Let p be a point on $\mathcal{M}$. A **tangent vector** to $\mathcal{M}$ at p is a real-valued function $v : C^\infty(\mathcal{M}) \to \mathbb{R}$ that is:*

- *$\mathbb{R}$-linear:* $v(af + bg) = av(f) + bv(g)$
- *Leibnizian:* $v(fg) = v(f)g(p) + f(p)v(g)$

for all $a, b \in \mathbb{R}$ and $f, g \in C^\infty(\mathcal{M})$.

*The set of all tangent vectors $T_p(\mathcal{M})$ at p is called the **tangent space** at p, and is made a vector space by usual functional addition and scalar multiplication.*

Read: $v(f)$ is the directional derivative of f along v.

Directional derivatives via coordinates

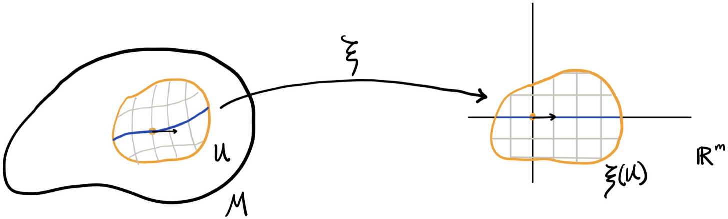


Figure 5: We can take directional derivatives by mapping to the coordinates. Define $\partial_i\big|_p$ as the directional derivative:

$$\partial_i f(p) = \frac{\partial f}{\partial x^i}(p) = \frac{\partial (f \circ \xi^{-1})}{\partial u^i}(\xi p),$$

where $u^1, \ldots, u^n$ are the natural coordinate functions of $\mathbb{R}^n$.

The basis theorem

Theorem (O'Neill (1983))

If $\xi = (x^1, \ldots, x^n)$ *is a coordinate system in* $\mathcal{M}$ *at* p*, then its coordinate vectors* $\partial_1\big|_p, \ldots, \partial_n\big|_p$ *form a basis for the tangent space* $T_p(\mathcal{M})$*.*

Tangent space

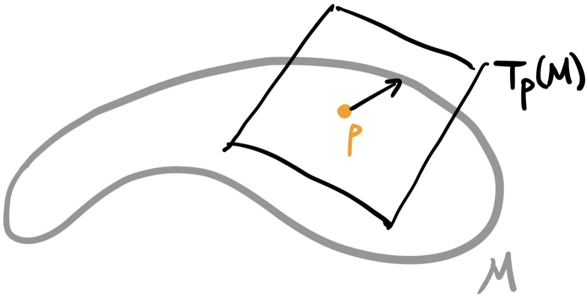


Figure 6: If we embed $\mathcal{M}$ into an ambient Euclidean space, the coordinate vectors $\partial_i\big|_p$ are the instantaneous direction we are moving at p when traveling at unit speed along the coordinate lines.

A spaceship analogy

Imagine flying a spaceship through curved space $\mathcal{M}$.

- When you are at position $p \in \mathcal{M}$, the input into the joystick is a choice of tangent vector $v \in T_p(\mathcal{M})$.

Smooth vector fields

A **vector field** X is an assignment of each $p \in \mathcal{M}$ to a tangent vector $X(p) \in T_p(\mathcal{M})$.

- X is **smooth** if for all $f \in C^\infty(\mathcal{M})$, the function Xf is smooth, $Xf \in C^\infty(\mathcal{M})$.

$$Xf(p) = \text{“derivative of } f \text{ along the } X(p) \text{ direction at } p\text{”}$$

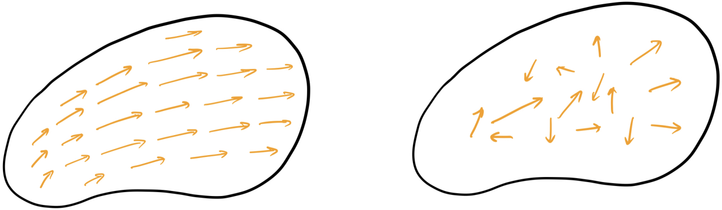


Figure 7: Two vector fields, one smooth (left) and one not (right).

Connecting the tangent spaces

In Euclidean space, all of its tangent spaces $T_p(\mathbb{R}^n)$ are canonically isomorphic to $\mathbb{R}^n$.

- If we lived on a flat map, we can identify at any point on the map a *north* direction.

But on a general manifold $\mathcal{M}$, there isn't a natural isomorphism between different tangent spaces $T_p(\mathcal{M})$ and $T_q(\mathcal{M})$.

- If we're traveling on a smooth manifold, we could say whether we are traveling on a smooth path, but there is no canonical way to quantify *how smooth* that path is.
- For example, you couldn't say you've been traveling along a fixed direction (i.e. smooth to the point that there is no change in your direction).

Connections, or covariant derivatives

We would like to define a notion of taking directional derivatives of vector fields.

- In Euclidean space, if X and Y are vector fields, then the derivative of Y with respect to X is given:

$$\nabla_X Y(p) = \lim_{t \downarrow 0} \frac{Y\big(p + tX(p)\big) - Y(p)}{t}.$$

  - Issue one: $p + tX(p)$ is not sensible on manifolds, since $p \in \mathcal{M}$ but $tX(p) \in T_p(\mathcal{M})$.
  - Issue two: $Y(p') - Y(p)$ is not defined since $T_p(\mathcal{M})$ and $T_{p'}(\mathcal{M})$ are two linear spaces.
- For manifolds, we need to make a choice for how to take derivatives of vector fields.
  - That is, we need to choose which vector fields correspond to ‘constant’ vector fields.

Axiomatizing the covariant derivative

Similar to how we can define the directional derivative for real-valued functions, we axiomatize the covariant derivative. Let $\Gamma(T\mathcal{M})$ be the set of smooth vector fields.

Definition

A ***connection*** *or a* ***covariant derivative*** $\nabla : \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \to \Gamma(T\mathcal{M})$ *is an* $\mathbb{R}$*-linear map that is:*

- $C^\infty(\mathcal{M})$*-linear in the first argument:* $\nabla_{fX} Y = f \nabla_X Y$
- *Leibnizian:* $\nabla_X(fY) = f \nabla_X Y + X(f) Y$.

Covariant derivatives via coordinates

Fix a coordinate system ξ at p and let $\partial_1\big|_p, \ldots, \partial_n\big|_p \in T_p(\mathcal{M})$ be the coordinate vectors. Recall that the coordinate vectors form a basis on $T_p(\mathcal{M})$.

- Since $\nabla\big|_p : T_p(\mathcal{M}) \times T_p(\mathcal{M}) \to T_p(\mathcal{M})$ is linear, we can describe a connection by its **Christoffel symbol** Γ, which is defined:[1]

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k.$$

[1] We'll use the Einstein summation convention, so that:

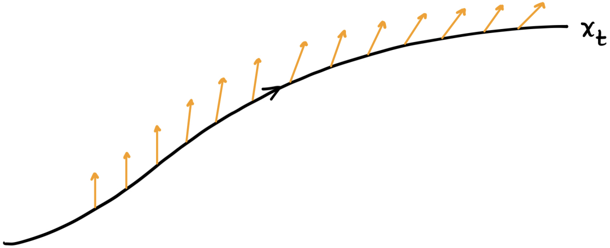
$$\Gamma_{ij}^k \partial_k \equiv \sum_{k=1}^{n} \Gamma_{ij}^k \partial_k.$$

Parallel

Definition

*A vector field X is said to be **parallel** along the curve x_t on $\mathcal{M}$ if:*

$$\nabla_{\dot{x}_t} X(x_t) = 0.$$



Parallel transport

Proposition (Gallier (2018))

Let $\mathcal{M}$ be a smooth manifold and let ∇ be a connection on $\mathcal{M}$. For every C^1-curve x_t in $\mathcal{M}$ and for every $v \in T_{x_t}(\mathcal{M})$, there is a unique parallel vector field X along x such that

$$X(t) = v.$$

Geodesics

Definition (Hsu (2002))

A curve x_t on $\mathcal{M}$ is called a ***geodesic*** *if its tangent vector field is parallel to itself. That is,*

$$\nabla_{\dot{x}_t} \dot{x}_t = 0.$$

Looking back: the big picture

1. We would like to do **calculus on nonlinear spaces**.
2. The spaces we direct our attention to are **smooth manifolds**.
   - Coordinate systems give us ways to locally map back to Euclidean space.
3. A basic question of calculus is how to take **directional derivatives**.
   - We could compute using local coordinates, $\partial_1 f(p), \ldots, \partial_n f(p)$.
   - Basis theorem implies we can take an algebraic definition, giving us tangent spaces.
4. A next question was how to take **derivatives of vector fields**.
   - A covariant derivative is a choice we make to define how directional vectors in different tangent spaces are related.

Revisiting the spaceship analogy

When flying a spaceship with a joystick, the tangent direction indicated on the joystick dashboard corresponds to the instantaneous direction that the spaceship travels.

- If we introduce a connection, that intuitively means that we can say that 'very smooth curves' are those on which we don't need to shift our joysticks very much.
- Parellel vector fields along a curve correspond to the direction we would turn if we pointed the joystick in that direction.
- Traveling along geodesics mean just pointing the joystick forward.

Frames of reference

We can make the relationship between the joystick and the instantaneous direction more explicit by introducing a *frame*.

Definition

A ***frame*** *u at $p \in \mathcal{M}$ is a linear isomorphism,*

$$u : \mathbb{R}^n \to T_p(\mathcal{M}).$$

Denote the set of frames at p by $\mathcal{F}(\mathcal{M})_p$.

- We can think of $\mathbb{R}^n$ as the possible inputs to the spaceship. A frame is a choice of how those inputs are translated into directions on the manifold.
- If $e_1, \ldots, e_n$ is the coordinate basis of $\mathbb{R}^n$, then $ue_1, \ldots, ue_n$ is a basis of $T_p(\mathcal{M})$.

Frame bundle

Definition

The ***frame bundle*** $\mathcal{F}(\mathcal{M})$ *is the disjoint union of frames over* $\mathcal{M}$*,*

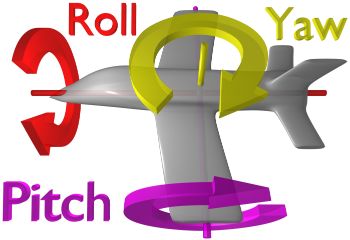
$$\mathcal{F}(\mathcal{M}) = \bigsqcup_{p \in \mathcal{M}} \mathcal{F}(\mathcal{M})_p.$$

- Intuitively, this is a choice of position $p \in \mathcal{M}$ and an orientation of the spaceship.
- The frame bundle can be made into a manifold of dimension $n + n^2$.
  - The first n dimensions correspond to p.
  - The latter n^2 dimensions correspond to a choice of frame $u \in \mathcal{F}(\mathcal{M})_p$.

Smooth paths in the (orthonormal) frame bundle

Let's fix a frame of reference $\mathbb{R}^n$ inside the spaceship traveling on $\mathcal{M}$.

- A smooth path on $\mathcal{F}(\mathcal{M})$ corresponds to the spaceship traveling on a smooth path through $\mathcal{M}$ while possibly spinning/twisting about its center smoothly.
- Given a smooth path x_t on $\mathcal{M}$ and an initial orientation $u_0 \in \mathcal{F}(\mathcal{M})_{x_0}$, there is a unique way to produce a smooth path u_t on $\mathcal{F}(\mathcal{M})$ such that the spaceship does not twist as it travels on x_t.
  - This curve u_t is called the **horizontal lift** of x_t.



Anti-developments

Imagine two spaceships: (i) the first travels on $\mathcal{M}$, while (ii) the second travels on $\mathbb{R}^n$.

- Whatever inputs are fed in the first ship are also transmitted to the second.
- As the first spaceship travels on a smooth curve $x_t \in \mathcal{M}$, the second travels on the analogous curve $z_t \in \mathbb{R}^n$.
- Assuming neither spaceships twists as they travel, then the curves x_t and z_t are in one-to-one correspondence through $u_t \in \mathcal{F}(\mathcal{M})$.
  - The curve z_t is called the **anti-development** of u_t.
  - The curve u_t is called the **development** of z_t.

III. Looking ahead to stochastic calculus

Stochastic developments

From the theory of smooth manifolds, we know how to develop a smooth curve $z_t \in \mathbb{R}^n$ to a smooth curve $x_t \in \mathcal{M}$.

- We can extend this procedure to semimartingales on $\mathbb{R}^n$, leading to a **stochastic development** on $\mathcal{F}(\mathcal{M})$.
- The projection of the stochastic development on $\mathcal{F}(\mathcal{M})$ down to $\mathcal{M}$ leads to the notion of a semimartingale on $\mathcal{M}$.

Martingales on manifolds

Definition (Hsu (2002))

Let $\mathcal{M}$ be a smooth manifold equipped with a connection ∇. An $\mathcal{M}$-valued semimartingale X_t is called a ∇-martingale if its antidevelopment Z_t with respect to ∇ on $\mathbb{R}^n$ is a (local) martingale.

Brownian motion on manifolds

Definition (Hsu (2002))

An $\mathcal{M}$-valued stochastic process X_t is called a (Riemannian) Brownian motion if its anti-development is standard Brownian motion on Euclidean space.

References

Jean H Gallier. Lecture notes for CIS 610 Advanced Geometric Methods in Computer Science: Connections on manifolds. *https://www.cis.upenn.edu/~cis610/cis610-18-sl12.pdf*, 2018.

Elton P Hsu. *Stochastic analysis on manifolds*. Number 38. American Mathematical Soc., 2002.

Barrett O'Neill. *Semi-Riemannian geometry with applications to relativity*. Academic press, 1983.