1
Riemannian geometry

We introduce the basic tools of Riemannian geometry that we shall need assuming that readers already know basic facts about manifolds. Only smooth manifolds are discussed unless otherwise mentioned. The use of a local chart (local coordinates) will be convenient for readers at the beginning. Tensor calculus is used to introduce geodesics, parallelism, covariant derivative and the Riemannian curvature tensor. In Sections 1.1 to 1.3 the Einstein convention will be adopted without further mention. However, it is not convenient for later discussion, for example, the second variation formula or Jacobi fields. To avoid confusion we shall use vector fields and connection forms to discuss such matters as Jacobi fields and conjugate points.

1.1 The Riemannian metric

Let M be an n-dimensional, connected and smooth manifold and (U, x) local coordinates around a point $p \in M$. A point $q \in U$ is expressed as $x(q) = (x^1(q), \dots, x^n(q)) \in x(U) \subset \mathbf{R}^n$. The tangent space to M at p is denoted by T_pM or M_p and $TM := \bigcup_{p \in M} T_pM$ denotes the tangent bundle over M. We denote by $\pi : TM \to M$ the projection map. Let $\mathcal{X}(M)$ and $\mathcal{X}(U)$ be the spaces of all smooth vector fields over M and U respectively and $C^\infty(M)$ the space of all smooth functions over M. A positive definite smooth symmetric bilinear form $g : \mathcal{X}(M) \times \mathcal{X}(M) \to C^\infty(M)$ is by definition a *Riemannian metric* over M. The metric g is locally expressed in (U, x) as follows. Let $X_i \in \mathcal{X}(U)$ be the ith basis-vector field tangent to the ith coordinate curve, i.e., $X_i := d(x^{-1})(\partial/\partial x^i)$, where $\partial/\partial x^i$ is the canonical vector field in $x(U) \subset \mathbf{R}^n$ parallel to the ith coordinate axis of $\mathbf{R}^n$. Then T_qM for every $q \in U$ is spanned by $X_1(q), \dots, X_n(q)$. If $X, Y : U \to TU$ are local vector fields expressed as

$X = \sum_i \phi^i X_i$ and $Y = \sum_j \psi^j X_j$ and if $g_{ij} := g(X_i, X_j)$ then

$$g(X, Y) = \sum_{i,j=1}^{n} \phi^i \psi^j g_{ij}, \tag{1.1.1}$$

where $g_{ij} = g_{ji}$ and the g_{ij} are smooth functions on U. Thus the length (or norm) $\|v\|$ of a vector $v \in T_pM$ is defined by $\|v\| := g_p(v, v)^{1/2} = (\sum g_{ij}(p)v^i v^j)^{1/2}$, where $v := \sum v^i X_i(p)$. The *angle* $\angle(u, v)$ between two vectors u and v in T_pM is thus defined by

$$\cos \angle(u, v) := \frac{\langle u, v\rangle}{\|u\|\,\|v\|}. \tag{1.1.2}$$

Here the angle $\angle(u, v)$ takes values in $[0, \pi]$ and $\langle u, v\rangle = g_p(u, v)$.

The volume of a parallel n-cube in T_pM spanned by $X_1(p), \dots, X_n(p)$ is given by $|X_1(p) \wedge \dots \wedge X_n(p)| = (\det g_{ij}(p))^{1/2}$. Thus the *volume element* $\mathrm{d}M$ of M is expressed as follows:

$$\mathrm{d}M = (\det g_{ij})^{1/2}\, \mathrm{d}x^1 \wedge \dots \wedge \mathrm{d}x^n. \tag{1.1.3}$$

The manifold M equipped with a Riemannian metric g is called a *Riemannian manifold* and denoted by (M, g) or simply by M.

By a smooth curve $c : [\alpha, \beta] \to M$ we always mean that there exists an open interval $I \supset [\alpha, \beta]$ such that c is defined over I and is regular at all points on I. A piecewise-smooth curve $c : [\alpha, \beta] \to M$ is a continuous map consisting of finitely many smooth curves. Namely, there are finitely many points $t_0 = \alpha < t_1 < \dots < t_k = \beta$ such that $c|_{[t_i, t_{i+1}]}$, for every $i = 0, \dots, k-1$, is a smooth curve. A piecewise-smooth vector field X along a curve $c : [\alpha, \beta] \to M$ is by definition a piecewise-smooth map $X : [\alpha, \beta] \to T_cM$ such that $\pi \circ X(t) = c(t)$ for all $t \in [\alpha, \beta]$, where $T_cM := \bigcup_{t\in[\alpha,\beta]} T_{c(t)}M$. A curve $c : [\alpha, \beta] \to U$ in a coordinate neighborhood is expressed by $x \circ c(t) = (x^1(t), \dots, x^n(t))$, for $t \in [\alpha, \beta]$. Its velocity vector field is

$$\dot{c}(t) = \sum_{i=1}^{n} \frac{\mathrm{d}x^i}{\mathrm{d}t} X_i \circ c(t).$$

The length $L(c)$ of a curve $c : [\alpha, \beta] \to U$ is given by

$$L(c) = \int_\alpha^\beta \|\dot{c}\|\, \mathrm{d}t = \int_\alpha^\beta \sqrt{\sum_{i,j=1}^{n} g_{ij}(c(t)) \frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}x^j}{\mathrm{d}t}}\, \mathrm{d}t. \tag{1.1.4}$$

If $s(t)$ for $t \in [\alpha, \beta]$ is the length of the subarc $c|_{[\alpha,t]}$ of c then $\mathrm{d}s(t)/\mathrm{d}t = \|\dot{c}(t)\| > 0$, and hence $s(t)$ has an inverse, $t = t(s)$. Thus c can be parameterized by its arc length $s \in [0, L]$, where $L = L(c)$ is the total length of c. From the relation $\mathrm{d}s(t) = \|\dot{c}\|\,\mathrm{d}t$ we derive a quadratic differential form

$$\mathrm{d}s^2 = \sum_{i,j=1}^{n} g_{ij}\mathrm{d}x^i\,\mathrm{d}x^j,$$

which we call the *line element* of M.

1.2 Geodesics

From now on the Einstein convention is used. Once the length of a curve in M has been established, we discuss a curve with the special property of having a local minimum in length among the neighboring curves with the same endpoints. This is referred to as the locally minimizing property. Such a curve, a *geodesic*, is obtained as the solution of a nonlinear second-order ordinary differential equation, (1.2.1) below, whose coefficients depend only on the g_{ij} and their partial derivatives. Thus geodesics are defined as the solutions of (1.2.1). The set of all solutions with a fixed starting point p corresponds to a domain in T_pM that is star-shaped with respect to the origin. Thus the exponential map and its injectivity radius at that point is introduced here.

Definition 1.2.1. A unit-speed curve $c : [\alpha, \beta] \to M$ is said to have the *locally minimizing property* iff there exists for every $s \in [\alpha, \beta]$ a positive number δ with $[s - \delta,\ s + \delta] \subset I$ and a neighborhood $\mathcal{N}$ around $c([s - \delta,\ s + \delta])$ such that $c([s - \delta,\ s + \delta])$ has the minimum length among all the curves in $\mathcal{N}$ joining $c(s - \delta)$ to $c(s + \delta)$.

We define the *Christoffel symbols*

$$\Gamma^i_{jk} := \tfrac{1}{2} g^{i\ell}(\partial_j g_{\ell k} + \partial_k g_{j\ell} - \partial_\ell g_{jk}),$$

where $\partial_j g_{\ell k} := \partial g_{\ell k}/\partial x^j$ and (g^{ij}) is the inverse matrix of (g_{ij}), i.e.,

$$g^{i\ell} g_{\ell k} := \delta^i_k.$$

Here δ^i_j is the *Kronecker delta*, i.e., $\delta^i_k = 1$ for $i = k$ and $\delta^i_k = 0$ for $i \neq k$.

With this notation we prove

Theorem 1.2.1. *If a unit-speed curve $c : [\alpha, \beta] \to M$ has the locally minimizing property, then the local expression*

$$x \circ c(s) = (x^1(s), \ldots,\ x^n(s))$$

of c in a coordinate neighborhood U satisfies

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0. \tag{1.2.1}$$

Proof. Since the discussion is local, we may restrict ourselves to the case where $c([\alpha, \beta])$ is contained entirely in a coordinate neighborhood U and write $x \circ c(s) = (x^1(s), \dots, x^n(s))$. A *variation along* $c : [\alpha, \beta] \to U$ is by definition a (piecewise-smooth) map $V : (-\varepsilon_0, \varepsilon_0) \times [\alpha, \beta] \to U$ such that

$$V(0, s) = c(s) \qquad \text{for all } s \in [\alpha, \beta]$$

and $V_\varepsilon(s) := V(\varepsilon, s)$ for every $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ is a curve $V_\varepsilon : [\alpha, \beta] \to U$. If $x \circ V(\varepsilon, s) = (x^1(\varepsilon, s), \dots, x^n(\varepsilon, s))$ for $(\varepsilon, s) \in (-\varepsilon_0, \varepsilon_0) \times [\alpha, \beta]$ then the variational vector field $Y : [\alpha, \beta] \to T_cM$ associated with V is given by

$$Y(s) = dV_{(0,s)}\left(\frac{\partial}{\partial\varepsilon}\right) = \frac{\partial x^i}{\partial\varepsilon}(0, s)\, X_i \circ c(s).$$

The length $L(\varepsilon)$ of each variation curve V_ε is given by

$$L(\varepsilon) := L(V_\varepsilon) = \int_\alpha^\beta \sqrt{g_{ij}\frac{\partial x^i(\varepsilon, s)}{\partial s}\frac{\partial x^j(\varepsilon, s)}{\partial s}}\, ds.$$

Thus we have, by taking the arc length parameter $0 \le s \le L =: L(c)$,

$$\begin{aligned}\frac{dL}{d\varepsilon}\bigg|_{\varepsilon=0} &= \int_0^L \frac{\partial}{\partial\varepsilon}\sqrt{g_{ij}\frac{\partial x^i(\varepsilon, s)}{\partial s}\frac{\partial x^j(\varepsilon, s)}{\partial s}}\,\Bigg|_{\varepsilon=0} ds \\ &= \frac{1}{2}\int_0^L \left((\partial_k g_{ij})\frac{\partial x^k}{\partial\varepsilon}\frac{\partial x^i}{\partial s}\frac{\partial x^j}{\partial s} + 2g_{ij}\frac{\partial^2 x^i}{\partial s\partial\varepsilon}\frac{\partial x^j}{\partial s}\right)\bigg|_{\varepsilon=0} ds.\end{aligned}$$

Since the second term in the integrand can be expressed as

$$\frac{\partial}{\partial s}\left(g_{ij}\frac{\partial x^i}{\partial\varepsilon}\frac{\partial x^j}{\partial s}\right) - \frac{\partial}{\partial s}\left(g_{kj}\frac{\partial x^j}{\partial s}\right)\frac{\partial x^k}{\partial\varepsilon}$$

we have

$$\begin{aligned}\frac{dL(0)}{d\varepsilon} &= g_{ij}\frac{\partial x^i}{\partial\varepsilon}\frac{\partial x^j}{\partial s}(0, s)\bigg|_0^L \\ &\quad - \int_0^L \left(\partial_\ell g_{kj}\frac{\partial x^\ell}{\partial s}\frac{\partial x^j}{\partial s} + g_{kj}\frac{\partial^2 x^j}{\partial s^2} - \frac{1}{2}\partial_k g_{ij}\frac{\partial x^i}{\partial s}\frac{\partial x^j}{\partial s}\right)\frac{\partial x^k}{\partial\varepsilon}\bigg|_{(0,s)} ds.\end{aligned}$$

By setting $2[ij;k] := \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}$, we see that $\Gamma^k_{ij} g_{k\ell} = [ij;\ell]$, and then the integrand can be rewritten as

$$\left(\frac{1}{2}(\partial_\ell g_{kj} + \partial_j g_{k\ell} - \partial_k g_{\ell j}) \frac{\mathrm{d}x^\ell}{\mathrm{d}s}\frac{\mathrm{d}x^j}{\mathrm{d}s} + g_{kj}\frac{\mathrm{d}^2 x^j}{\mathrm{d}s^2}\right)\frac{\partial x^k}{\partial \varepsilon}$$
$$= g_{kj}\left(\frac{\mathrm{d}^2 x^j}{\mathrm{d}s^2} + \Gamma^j_{\ell m}\frac{\mathrm{d}x^\ell}{\mathrm{d}s}\frac{\mathrm{d}x^m}{\mathrm{d}s}\right)\frac{\partial x^k}{\partial \varepsilon}.$$

Therefore we obtain

$$L'(0) = g(\dot{c}, Y)|_0^L - \int_0^L g_{kj}\left(\frac{\mathrm{d}^2 x^j}{\mathrm{d}s^2} + \Gamma^j_{\ell m}\frac{\mathrm{d}x^\ell}{\mathrm{d}s}\frac{\mathrm{d}x^m}{\mathrm{d}s}\right) Y^k\,\mathrm{d}s. \quad (1.2.2)$$

The locally minimizing property of c then implies that $L'(0) = 0$ for every variation V with $V(\varepsilon, 0) = c(0)$ and $V(\varepsilon, L) = c(L)$ for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$. Therefore $g(\dot{c}, Y)|_0^L = 0$ follows from $Y(0) = Y(L) = 0$. The proof is concluded since the variation vector field Y can be taken as arbitrary. □

Now the differential equation (1.2.1) will be discussed. Changing the parameter via $s = at$ for a constant $a > 0$, we observe that (1.2.1) becomes

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + \Gamma^i_{jk}\frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\mathrm{d}x^k}{\mathrm{d}t} = 0.$$

Geodesics are always parameterized proportionally to arc lengths. The equation (1.2.1) is equivalent to the following system of first-order differential equations:

$$v^i(s) = \frac{\mathrm{d}x^i}{\mathrm{d}s}, \qquad \frac{\mathrm{d}v^i}{\mathrm{d}s} + \Gamma^i_{jk} v^j v^k = 0. \quad (1.2.3)$$

Because the Γ^i_{jk} are smooth functions, the above differential equations satisfy the Lipschitz condition and hence have a unique solution for given initial conditions. Let $p \in U$ and $\xi \in T_pM$ be expressed as $x(p) = (p^1, \ldots, p^n)$ and $\xi = \xi^i X_i(p)$. Then (1.2.1) has locally a unique solution for the initial conditions $x^i(0) = p^i$ and $\mathrm{d}x^i/\mathrm{d}s(0) = \xi^i$ for $i = 1, \ldots, n$. If $\gamma(s) := \gamma(p, \xi; s)$ for $s \in [0, s_0)$ is the maximal solution of (1.2.1) with $\gamma(0) := \gamma(p, \xi; 0) = p$ and $\dot{\gamma}(0) = \xi$ then $\gamma(p, \xi; t) = \gamma(p, a\xi; t/a) = \gamma(p, t\xi; 1)$. If we set

$$\widetilde{M}_p := \{u \in T_pM; \gamma(p, u; 1) \text{ makes sense}\}$$

then $\widetilde{M}_p$ is a domain in T_pM that is star-shaped with respect to the origin of T_pM.

Definition 1.2.2. The *exponential map* at $p \in M$ is a map defined on $\widetilde{M}_p$ such that

$$\exp_p u := \gamma(p, u; 1).$$

Clearly $\exp_p$ is a smooth map.

Theorem 1.2.2. *There exists an open set $U_p \subset \widetilde{M}_p$ around the origin such that $\exp_p|_{U_p} : U_p \to M$ is an embedding. In particular, there exists an open set $V_p \subset U_p$ such that any two points in $\exp_p V_p$ can be joined by a geodesic.*

Proof. It is clear from the definition of the exponential map at p that

$$d(\exp_p)|_o = E_n,$$

where E_n is the $n \times n$ identity matrix and $u = (u^1, \ldots, u^n)$. Therefore we can find a small neighborhood U_p around the origin of T_pM as desired. Let $\widetilde{TM} := \bigcup_{p\in M} \widetilde{M}_p$ and $\phi := (\pi, \exp) : \widetilde{TM} \to M \times M$. Then the above discussion shows that, at each zero section $o \in \widetilde{TM}$,

$$d\phi|_o = \begin{pmatrix} E_n & E_n \\ 0 & E_n \end{pmatrix}.$$

Therefore we can find an open set $W \subset \widetilde{TM}$ around the set of zero sections such that $\phi|_W$ is an embedding. Thus there exists an open set $V_p \subset U_p$ around p such that $V_p \times V_p \subset \phi(W)$. This proves Theorem 1.2.2. □

Lemma 1.2.1 (The Gauss lemma). *If $u \in \widetilde{M}_p$ and if $A \in T_uT_pM$ is orthogonal to u then*

$$\langle d(\exp_p)_u A, d(\exp_p)_u u\rangle = 0.$$

Proof. The conclusion is obvious if $d(\exp_p)_u A = 0$. Assume that $d(\exp_p)_u A \neq 0$. We then choose a geodesic variation $V : (-\varepsilon_0, \varepsilon_0) \times [0, \ell] \to M$ along the geodesic $\gamma : [0, \ell] \to M$ with $\gamma(0) = p$, $\dot\gamma(0) = u/\|u\|$ and $\ell = \|u\|$ such that

$$V(\varepsilon, t) := \exp_p t(u/\|u\| + \varepsilon A).$$

Then $V_\varepsilon : [0, \ell] \to M$ for every $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ is a geodesic with length $\ell\sqrt{1 + \varepsilon^2\|A\|^2}$. If $Y(t) := dV_{(0,t)}(\partial/\partial\varepsilon)$ is the variational vector field associated with V then $Y(0) = 0$, $Y(\ell) = d(\exp_p)_u A$ and (1.2.2) implies that

$$L'(0) = \langle Y(t), \dot\gamma(t)\rangle\,|_0^\ell = 0.$$ □

A *geodesic polar coordinate system around a point* p is defined by the embedding $\exp_p|_{U_p} : U_p \to M$. Let $B(0, r) := \{u \in \mathbf{R}^n;\ \|u\| < r\}$ and $\mathbf{S}^{n-1} := \{u \in \mathbf{R}^n;\ \|u\| = 1\}$. They are placed in T_pM by a trivial identification. If $(\theta^1, \ldots, \theta^{n-1})$ is a local coordinate system of $\mathbf{S}^{n-1}$ around a point $u \in \mathbf{S}^{n-1}$ then $\exp_p|_{U_p}$ is expressed locally by $(\exp_p|_{U_p})^{-1}(q) = (r(q), \theta^1(q), \ldots, \theta^{n-1}$

$(q)) \in U_p$. By setting $u^1 := r, u^2 := \theta^1, \dots, u^n := \theta^{n-1}$, we see from the Gauss lemma that the metric g can be expressed as

$$g_{ij}\, \mathrm{d}u^i\, \mathrm{d}u^j = \mathrm{d}r^2 + h_{ab}\, \mathrm{d}\theta^a\, \mathrm{d}\theta^b, \qquad (1.2.4)$$

where (h_{ab}) is a positive definite symmetric $(n-1) \times (n-1)$ matrix.

Definition 1.2.3. The *injectivity radius* of $\exp_p$ at p is defined by

$$i(p) := \sup\{r > 0\,;\ \exp_p|_{B(0,r)} \text{ is an embedding}\}.$$

Lemma 1.2.2. *If $r < i(p)$ then every point $q \in \exp_p B(0,r)$ is joined to p by a unique geodesic whose length attains the infimum of all the lengths of curves joining p to q. In particular every geodesic has the locally minimizing property.*

Proof. It follows from Theorem 1.2.2 that in $\exp_p B(0, i(p))$ there exists a unique geodesic joining p to q with length $r(q)$. Let $c : [0,1] \to \exp_p B(0, i(p))$ be a (piecewise-smooth) curve with $c(0) = p$ and $c(1) = q$. There exists a lift $\psi : [0,1] \to B(0, i(p)) \subset T_pM$ of c such that $c(t) = \exp_p \circ \psi(t)$ for all $t \in [0,1]$. Then the lift can be expressed by $\psi(t) = (r(t), \theta^1(t), \dots, \theta^{n-1}(t))$, and hence

$$\dot c(t) = \mathrm{d}(\exp_p)_{\psi(t)}\dot\psi(t) = \mathrm{d}(\exp_p)_{\psi(t)}\{\dot r\psi(t)/\|\psi\| + rA(t)\},$$

where $r(t) = \|\psi(t)\|$, $\psi(t)/\|\psi(t)\| = (\theta^1(t), \dots, \theta^{n-1}(t)) \in S_p^{n-1}$ and $A(t)$ is the component of $\dot\psi(t)$ tangential to S_p^{n-1} at $\psi(t)/\|\psi(t)\|$. Then we have

$$\begin{aligned} L(c) &= \lim_{\varepsilon\to 0}\int_\varepsilon^1 \|\dot c\|\, \mathrm{d}t = \lim_{\varepsilon\to 0}\int_\varepsilon^1 \sqrt{g_{ij}\dot u^i \dot u^j}\,\mathrm{d}t \\ &= \lim_{\varepsilon\to 0}\int_\varepsilon^1 \sqrt{\dot r^2 + h_{ab}\dot\theta^a\dot\theta^b}\,\mathrm{d}t \ge \int_\varepsilon^1 \left|\frac{\mathrm{d}r}{\mathrm{d}t}\right| \mathrm{d}t \\ &\ge \int_0^1 \mathrm{d}r = r(q) - r(p) = r(q). \end{aligned}$$

We now consider a (piecewise-smooth) curve $c : [0,1] \to M$ joining p to q that is not contained in $\exp_p B(0, i(p))$. Then there exists a point $c(t_0)$ such that the subarc $c|_{[0,t_0)}$ is contained entirely in $\exp_p B(0, i(p))$. The previous discussion now implies that $L(c|_{[0,t_0]}) \ge i(p) > r(q)$. This means that such a curve has length greater than $i(p) > r(q)$. □

1.3 The Riemannian curvature tensor

Newton's first law states that the motion of a particle is in a straight line at constant speed if no force acts upon it. Geodesics on a Riemannian manifold are understood to be subject to Newton's first law: it is considered that the straightness of a geodesic is equivalent to the existence of a parallel velocity vector field along it. In view of equation (1.2.1), this idea leads us to the definition of parallel fields along a curve.

Lemma 1.3.1. *Let $Z(t) = \xi^i(t)X_i \circ \gamma(t)$ be a vector field along a curve γ, where $\gamma(t) = (x^1(t), \dots, x^n(t))$ in a coordinate neighborhood U. Then the map*

$$t \mapsto \left(\frac{\mathrm{d}\xi^i}{\mathrm{d}t} + \Gamma^i_{kj}\xi^k\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) X_i \circ \gamma(t)$$

is independent of the choice of local coordinates, and hence is a vector field along γ.

Proof. Let (V, y) be another set of local coordinates such that $U \cap V$ contains a subarc of γ, which is expressed by $x \circ \gamma(t) = (x^1(t), \dots, x^n(t))$ and $y \circ \gamma(t) = (y^1(t), \dots, y^n(t))$. The line element $\mathrm{d}s^2$ is expressed in $U \cap V$ as $\mathrm{d}s^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j = h_{ab}\mathrm{d}y^a\mathrm{d}y^b$. We denote by $(A^a_i) = (\partial y^a/\partial x^i)$ the Jacobian matrix and by $(B^i_a) = (\partial x^i/\partial y^a)$ its inverse matrix. Clearly we have $Z(t) = \xi^i X_i \circ \gamma(t) = \eta^a Y_a \circ \gamma(t)$, where $\xi^i = B^i_a\eta^a$. By differentiating the relation $g_{ij} = A^a_i A^b_j h_{ab}$ with respect to x^k, we obtain

$$\partial_k g_{ij} = A^a_i A^b_j A^c_k \frac{\partial h_{ab}}{\partial y^c} + \left(\frac{\partial^2 y^a}{\partial x^k \partial x^i}A^b_j + \frac{\partial^2 y^b}{\partial x^k \partial x^j}A^a_i\right) h_{ab}.$$

Also, by setting

$$\bar{\Gamma}^a_{bc} := \tfrac{1}{2}h^{ad}(\partial_b h_{dc} + \partial_c h_{bd} - \partial_d h_{bc}),$$
$$\overline{[bc, a]} := h_{da}\bar{\Gamma}^d_{bc},$$

we get

$$[ij, k] = A^b_i A^c_j A^a_k\,\overline{[bc, a]} + \frac{\partial^2 y^d}{\partial x^i \partial x^j}A^e_k h_{ed}.$$

It can be shown that

$$\Gamma^i_{jk} = A^b_j A^c_k B^i_a \bar{\Gamma}^a_{bc} + \frac{\partial^2 y^d}{\partial x^j \partial x^k}B^i_d.$$

Therefore we have

$$\begin{aligned}
&\left(\frac{d\xi^i}{dt} + \Gamma^i_{jk}\xi^j\frac{dx^k}{dt}\right) X_i \\
&= \left\{\frac{d}{dt}(B^i_a\eta^a) + \left(A^b_j A^c_k B^i_a \bar{\Gamma}^a_{bc} + \frac{\partial^2 y^d}{\partial x^j \partial x^k} B^i_d\right)\xi^j\frac{dx^k}{dt}\right\} A^e_i Y_e \\
&= \left\{\left(\frac{\partial^2 x^i}{\partial y^a \partial y^c}\frac{dy^c}{dt}\eta^a A^e_i + B^i_a\frac{d\eta^a}{dt}A^e_i\right) + \eta^b\frac{dy^c}{dt}\delta^e_a\bar{\Gamma}^a_{bc}\right. \\
&\quad \left. + \frac{\partial^2 y^d}{\partial x^j \partial x^k}\delta^e_d B^j_a B^k_c \eta^a \frac{dy^c}{dt}\right\} Y_e \\
&= \left\{\frac{d\eta^e}{dt} + \eta^a\frac{dy^c}{dt}\left(\frac{\partial^2 x^i}{\partial y^a \partial y^c}A^e_i + \frac{\partial^2 y^e}{\partial x^j \partial x^k}B^j_a B^k_c + \bar{\Gamma}^e_{ac}\right)\right\} Y_e.
\end{aligned}$$

It follows from

$$0 = \frac{\partial}{\partial y^a}\left(\frac{\partial y^e}{\partial x^k}\frac{\partial x^k}{\partial y^c}\right) = \frac{\partial^2 y^e}{\partial x^k \partial x^j}B^j_a B^k_c + A^e_k \frac{\partial^2 x^k}{\partial y^c \partial y^a}$$

that

$$\left(\frac{d\xi^i}{dt} + \Gamma^i_{jk}\xi^j\frac{dx^k}{dt}\right) X_i \circ c(t) = \left(\frac{d\eta^e}{dt} + \bar{\Gamma}^e_{bc}\eta^b\frac{dy^c}{dt}\right) Y_e \circ c(t). \qquad \square$$

The above reasoning shows that if Z is a vector field defined in a neighborhood around p and if its restriction to a curve c emanating from $p = c(0)$ with a given initial tangent vector $v = \dot{c}(0) \in M_p$ is expressed by $Z \circ c(t) = \xi^i(t)X_i \circ c(t)$ then the vector $(d\xi^i/dt + \Gamma^i_{jk}\xi^j dx^k/dt)X_i(p)$ is independent of the choice of c and depends only on v and Z. We write

$$\nabla_v Z := \left(\frac{d\xi^i}{dt} + \Gamma^i_{jk}\xi^j\frac{dx^k}{dt}\right) X_i(p) \tag{1.3.1}$$

and call $\nabla_v Z$ the *covariant derivative* of Z in v.

The covariant derivative also defines a map $\nabla : \mathcal{X}(U) \times \mathcal{X}(U) \to \mathcal{X}(U)$ as follows. If $Y, Z \in \mathcal{X}(U)$ then a vector field $\nabla_Y Z \in \mathcal{X}(U)$ is obtained by setting $Y = \phi^i X_i$ and $Z = \psi^i X_i$:

$$\nabla_Y Z := \phi^i\left(\frac{\partial\psi^k}{\partial x^i} + \Gamma^k_{ij}\psi^j\right) X_k.$$

Exercise 1.3.1. Show that the covariant derivative of a vector field has the following properties. If $Y, Z : U \to TU$ and α, β are functions defined on U then

$$\begin{cases} \nabla_{\alpha Y} Z = \alpha \nabla_Y Z, \\ \nabla_Y(\beta Z) = Y(\beta) Z + \beta \nabla_Y Z, \\ X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \\ \nabla_X Y - \nabla_Y X = [X, Y]. \end{cases} \tag{1.3.2}$$

Here $[X, Y]$ is the vector field defined by $[X, Y]f := X(Yf) - Y(Xf)$ for a smooth function f.

A vector field $Z : [0, \ell] \to T_c M$ along a unit-speed curve $c : [0, \ell] \to U$ is by definition *parallel* iff $\nabla_{\dot c} Z = 0$. The *geodesic curvature vector* $\mathbf{k}(s)$ of c at $c(s)$ is defined by

$$\mathbf{k}(s) := \left(\frac{d^2 x^i}{ds^2} + \Gamma^i_{jk} \frac{dx^j}{ds} \frac{dx^k}{ds} \right) X_i \circ c(s),$$

where s is the arc length of c. The *geodesic curvature* $\kappa(s)$ of c is defined by

$$\kappa(s) := \|\mathbf{k}(s)\|.$$

Remark 1.3.1. The geodesic curvature vector $\mathbf{k}(s)$ at a point $c(s)$ of a unit-speed smooth curve c has the following property. Let $\gamma_\pm : [0, a_\pm] \to M$ for a sufficiently small $h > 0$ be a unit-speed minimizing geodesic with $\gamma_\pm(0) = c(s)$ and $\gamma_\pm(a_\pm) = c(s \pm h)$ and let $\tau : TM \to T_{c(s)}M$ be the parallel translation along the minimizing geodesics at $c(s)$. Then

$$\lim_{h \to 0} \frac{\tau \circ \dot c(s+h) - \dot c(s)}{h} = \mathbf{k}(s). \tag{1.3.3}$$

To see this we define *parallel fields* $\xi_\pm$ along $\gamma_\pm$ that are generated by $\xi_\pm(a_\pm) := \dot c(s \pm h)$. If $x \circ c(s) = (x^1(s), \ldots, x^n(s))$ and $x \circ \gamma_\pm(t) = (x^1_\pm(t), \ldots, x^n_\pm(t))$ and if $\xi_\pm(t) = \xi^i_\pm(t) X_i \circ \gamma_\pm(t)$ for $0 \le t \le a_\pm$ are local expressions then

$$\xi^i_\pm(a_\pm) - \xi^i_\pm(0) = a_\pm \frac{d\xi^i_\pm}{dt}(\tilde a^i_\pm), \qquad i = 1, \ldots, n$$

for some $\tilde a^i_\pm \in (0, a_\pm)$. The ith component of the left-hand side of (1.3.3) is expressed as

$$\frac{1}{h}\left(\xi^i_\pm(0) - \frac{dx^i(s)}{ds} \right) = \frac{1}{h}\left(\frac{dx^i(s \pm h)}{ds} - \frac{dx^i(s)}{ds} \right) + \frac{1}{h}\left(a_\pm \Gamma^i_{jk} \xi^j_\pm \frac{dx^k_\pm}{dt} \right).$$

Taking the limit as $h \to 0$, we see that $\dot\gamma_\pm(a_\pm)$ converges to $\pm \dot c(s)$ and that $\lim_{h\to 0}(a_\pm/h) = 1$. This proves (1.3.3).