Extrinsic Differential Geometry

J.W.R

Most Recent Revision: April 18, 2008

Euclidean Space

This is the arena of Euclidean geometry; i.e. every figure which is studied in Euclidean geometry is a subset of Euclidean space. To define it one could proceed axiomatically as Euclid did; one would then verify that the axioms characterized Euclidean space by constructing "Cartesian Co-ordinate Systems" which identify the n-dimensional Euclidean space E^n with the n-dimensional numerical space $\mathbf{R}^n$. This program was carried out rigorously by Hilbert. We shall adopt the mathematically simpler but philosophically less satisfying course of taking the characterization as the definition.

We shall use three closely related spaces: n-dimensional Euclidean space E^n, n-dimensional Euclidean vector space $\mathbf{E}^n$; and the space $\mathbf{R}^n$ of all n-tuples of real numbers. The distinction among them is a bit pedantic, especially if one views as the purpose of geometry the interpretation of calculations on $\mathbf{R}^n$.

We can take as our model of E^n any n-dimensional affine subspace of some numerical space $\mathbf{R}^k$ $(k > n)$; the vector space $\mathbf{E}^n$ is then the unique fector subspace of $\mathbf{R}^k$ for which:

$$E^n = p + \mathbf{E}^n$$

for $p \in E^n$. (Note that $\mathbf{E}^n$ contains the "preferred" point 0 while E^n has no preferred point; each $p \in E^n$ determines a different bijection $v \to p + v$ from $\mathbf{E}^n$ onto E^n.) Any choice of an origin $p_0 \in E^n$ and an orthonormal basis $e_1, \ldots, e_n$ for $\mathbf{E}^n$ gives a bijection:

$$\mathbf{R}^n \to E^n : (x^1, \ldots, x^n) \mapsto p_0 + \sum_i x^i e_i$$

(the inverse of which is) called a Cartesian co-ordinate system on E^n.

Such space E^n and $\mathbf{E}^n$ would arise in linear algebra by taking E^n to be the space of solutions of $k - n$ independent inhomogeneous linear equations in k unknowns while $\mathbf{E}^n$ is the space of solutions of the corresponding homogeneous equations.

We now give precise definitions.

Definition 1. *The* **orthogonal group** $O(n)$ *of* $\mathbf{R}^n$ *is the group of all* $n \times n$ *matrices* a *whose transpose is their inverse:*

$$a \in O(n) \Longleftrightarrow aa^* = e$$

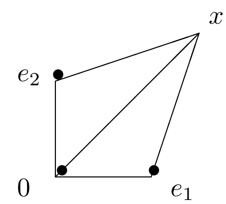
where e *is the identity matrix. An equivalent characterization is that:*

$$\langle ax, ay \rangle = \langle x, y \rangle$$

for all $x, y \in \mathbf{R}^n$. *A* **group** $R(n)$ **of rigid motions of** $\mathbf{R}^n$ *is the group generated by* $O(n)$ *and the group of translation, thus for* $\tilde{a} : \mathbf{R}^n \to \mathbf{R}^n$ *we have:*

$$\tilde{a} \in R(n) \Longleftrightarrow \tilde{a}(x) = ax + v \qquad (\forall x \in \mathbf{R}^n)$$

for some $a \in O(n)$, $v \in \mathbf{R}^n$.



Exercise 2. Show that $\tilde{a} : \mathbf{R}^n \to \mathbf{R}^n$ is a rigid motion, i.e. an element of $R(n)$, if and only if it is an isometry:

$$\|\tilde{a}(x) - \tilde{a}(y)\| = \|x - y\|$$

for all $x, y \in \mathbf{R}^n$. Hint: You will need the following key lemma: Let e_i $(i = 1, \ldots, n)$ be the point given by

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

with 1 in the i-th slot. Assume $x, y \in \mathbf{R}^n$ satisfy the $n+1$ equations:

$$\|x - 0\| = \|y - 0\|, \quad \|x - e_i\| = \|y - e_i\| \qquad i = 1, \ldots, n.$$

Then $x = y$.

Let E^n be a set. Call two bijections $x : E^n \to \mathbf{R}^n$ and $y : E^n \to \mathbf{R}^n$ equivalent iff the transformation $yx^{-1} : \mathbf{R}^n \to \mathbf{R}^n$ is a rigid motion, i.e. $yx^{-1} \in R(n)$.

Definition 3. *An* **(n-dimensional) Euclidean space** *is a set E^n together with an equivalence class of bijections $x : E^n \to \mathbf{R}^n$; the elements of the equivalence class are called* **Cartesian Co-ordinate** *systems of the Euclidean space. (Note that by definition any Euclidean space is a manifold diffeomorphic to $\mathbf{R}^n$.)*

Definition 4. *Call two pairs of points (p_1, q_1), (p_2, q_2) from E^n equivalent iff for some (and hence every) cartesian co-ordinate system we have:*

$$x(q_1) - x(p_1) = x(q_2) = x(p_2)\,.$$

Denote the equivalence class of the pair (p, q) by $q - p$ and call it **the vector from** *p to q. Denote the set of all such vectors by $\mathbf{E}^n$ and call it the* **Euclidean vector space** *associated to E^n. The cartesian co-ordinate system $x : E^n \to \mathbf{R}^n$ induces a map $x_* : \mathbf{E}^n \to \mathbf{R}^n$ by:*

$$x_*(q - p) = x(q) - x(p)\,.$$

Define the operations of vector addition, scalar multiplication, and inner product in $\mathbf{E}^n$ by declaring that x_* intertwine these operations with the corresponding ones on $\mathbf{R}^n$; the operations are well-defined (i.e. independent of x) since $yx^{-1} = \tilde{a} \in R(n)$ implies that $y_* x_*^{-1} = a \in O(n)$.

Definition 5. *If we fix $p \in E^n$ we obtain a bijection:*

$$E^n \to \mathbf{E}^n : q \to q - p\,;$$

denote the inverse bijection by:

$$\mathbf{E}^n \to E^n : v \mapsto p + v\,.$$

The set of all transformations:

$$E^n \to E^n : p \to p + v$$

form a group called the **translation group** *of E^n; it is naturally isomorphic to the additive group of the vector space $\mathbf{E}^n$ and will be denoted by the same notation.*

Definition 6. *The groups*

$$R(E^n) = x^{-1}R(n)x, \quad O(\mathbf{E}^n) = x_*^{-1}O(n)x_*$$

are independent of the choice of cartesian co-ordinates $x : E^n \to \mathbf{R}^n$. The group $R(E^n)$ is also set of all transformations $\tilde{a} : E^n \to E^n$ which preserve the distance function

$$(p, q) \mapsto \|p - q\| = \sqrt{\langle p - q, p - q\rangle};$$

it is called the **group of rigid motions** *of E^n. The group $O(\mathbf{E}^n)$ is the groups of* **orthogonal linear transformations** *of $\mathbf{E}^n$. For $p \in E^n$ let $R(E^n)_p$ denote the isotropy group:*

$$R(E^n)_p = \{\tilde{a} \in R(E^n) : \tilde{a}(p) =\}\,.$$

Exercise 7. The translation subgroup $\mathbf{E}^n$ is a normal subgroup on $R(E^n)$:

$$\tilde{a}\mathbf{E}^n\tilde{a}^{-1} = \mathbf{E}^n$$

for $\tilde{a} \in R(E^n)$. For each fixed $p \in E^n$ the bijection:

$$\mathbf{E}^n \to E^n : v \mapsto p + v$$

intertwines the group $O(\mathbf{E}^n)$ acting on $\mathbf{E}^n$ with the group $R(E^n)_p$. The subgroup $R(E^n)_p$ is *not* normal in $R(E^n)$ but every element of $R(E^n)$ is uniquely expressible as the product of an element in $R(E^n)_p$ and an element of $\mathbf{E}^n$:

$$R(E^n) = \mathbf{E}^n \cdot R(E^n)_p, \quad \mathbf{E}^n \cap R(E^n)_p = \{\text{identity}\}.$$

Remark 8. One summarizes this exercise by saying that $R(E^n)$ is the semi-direct product of $O(\mathbf{E}^n)$ and $\mathbf{E}^n$ and writing:

$$R(E^n) \cong O(\mathbf{E}^n) \bowtie \mathbf{E}^n\,.$$

Choosing co-ordinates gives the analogous assertion:

$$R(n) \cong O(n) \bowtie \mathbf{R}^n$$

which is nothing more than a fancy way of saying that $\tilde{a} \in R(n)$ has form $\tilde{a}(x) = ax + v$ for some $a \in O(n)$ and $v \in \mathbf{R}^n$. There is an important point to be made here. The embeddings of $O(\mathbf{E}^n)$ into $R(E^n)$ depends on the choice of the "origin" p while the embedding of $\mathbf{E}^n$ into $R(E^n)$ is independent of any such choice. The subgroup $\mathbf{E}^n$ of $R(E^n)$ has an "invariant interpretation" and this accounts for why it is normal. (It also explains why normal subgroups are sometimes called "invariant subgroups"; their definitions are independent of the choice of co-ordinates.)

Definition 9. *An m-dimensional affine subspace of E^n is a space E^m of form*

$$E^m = p + \mathbf{E}^m$$

where $\mathbf{E}^m$ is an m-dimensional vector subspace of $\mathbf{E}^n$. It is again a Euclidean space in the obvious way: an (inverse of) a cartesian co-ordinate system is a map:

$$\mathbf{R}^m \to E^m : (x^1, \ldots, c^m) \mapsto p + \sum_i x^i e_i$$

where p is any point of E^m and $e_1, \ldots, e_m$ is any orthonormal basis for $\mathbf{E}^m$.

Notation

Let $M = M^m$ be an m-dimensional submanifold of n-dimensional Euclidean space E^n. Denote the space of vectors of E^n by $\mathbf{E}^n$. Identify the tangent space T_pM and the normal space $T_P^\perp M$ to M at p with vector Subspaces of $\mathbf{E}^n$:

$$T_pM = \{\dot{\gamma}(0) | \gamma : \mathbf{R} \to M, \gamma(0) = p\} \subseteq \mathbf{E}^n$$

$$T_p^\perp M = \{u \in \mathbf{E}^n | \langle u, T_pM \rangle = 0\}.$$

Denote by $\mathcal{X}(M)$ and $\mathcal{X}^\perp(M)$ the space of (tangent) vector fields and normal (vector) fields on M respectively. Thus for $X, U : M \to \mathbf{E}^n$ we have:

$$X \in \mathcal{X}(M) \Longleftrightarrow X(p) \in T_pM \qquad (\forall p \in M)$$

$$U \in \mathcal{X}^\perp(M) \Longleftrightarrow U(p) \in T_p^\perp M \qquad (\forall p \in M)$$

More generally for any map $\varphi : N \to M$ of a manifold N into M we denote by $\mathcal{X}(\varphi)$ (resp. $\mathcal{X}^\perp(\varphi)$) the space of all **vector fields** (resp. **normal fields**) along φ. Thus for:

$$X, U : N \to \mathbf{E}^n$$

we have:

$$X \in \mathcal{X}(\varphi) \Longleftrightarrow X(q) \in T_{\varphi(q)}M \qquad (\forall q \in N)$$

and:

$$U \in \mathcal{X}^{\perp}(\varphi) \Longleftrightarrow U(q) \in T^{\perp}_{\varphi(q)}M \qquad (\forall q \in N)\,.$$

Such fields are especially important when φ is a curve (i.e. N is an open interval in $\mathbf{R}$).

Note that $\mathcal{X}(M) = \mathcal{X}(\varphi)$ and $\mathcal{X}^{\perp}(M) = \mathcal{S}^{\perp}(\varphi)$ when $\varphi : M \to M$ is the identity. Further note that $X \circ \varphi \in \mathcal{X}(\varphi)$ when $X \in \mathcal{X}(M)$ and $\varphi : N \to M$.

We denote by $\tilde{T}_pM \subset E^n$ the **affine tangent space** at $p \in M$:

$$\tilde{T}_pM = p + T_pM\,.$$

Note the natural isomorphism:

$$T_pM \to \tilde{T}_pM : v \mapsto p + v\,.$$

We introduce $\tilde{T}_pM$ because it is more natural to draw tangent vectors to M at p with their tails at p rather than translated to some artificial origin.

The **first fundamental form** is the field which assigns to each $p \in M$ the bilinear map:

$$g_p \in L^2(T_pM;\mathbf{R})$$

given by:

$$g_p(v,w) = \langle v, w\rangle$$

for $v, w \in T_pM \subset \mathbf{E}^n$.

Second Fundamental Form

For each $p \in M$ denote by $\Pi(p) \in L(\mathbf{E}^n)$ the orthogonal projection of $\mathbf{E}^n$ on T_pM. It is characterized by the three equations:

$$\Pi(p)^2 = \Pi(p), \quad \Pi(p)\mathbf{E}^n = T_pM, \quad \Pi(p)^* = \Pi(p).$$

Similarly denote by $\Pi^{\perp}(p)$ the orthogonal projection on $T_p^{\perp}M$:

$$\Pi^{\perp}(p) = I - \Pi(p)$$

where I is the identity transformation of $\mathbf{E}^n$.

It should be emphasized that $\Pi(p)$ is to be considered as a linear map:

$$\Pi(p) : \mathbf{E}^n \to \mathbf{E}^n$$

with "target" $\mathbf{E}^n$ (but image T_pM). Thus Π itself is a "matrix" valued function:

$$\Pi : M \to L(\mathbf{E}^n)\,.$$

Example 10. Take rectangular co-ordinates (x, y, z) on E^3 and let $M = S^2$ be the sphere with equation:

$$x^2 + y^2 + z^2 = 1\,.$$

Then the formula for $\Pi(p)$ is:

$$\Pi(p) = \begin{bmatrix} 1 - x^2 & -xy & -xz \\ -yx & 1 - y^2 & -yz \\ -zx & -zy & 1 - z^2 \end{bmatrix}.$$

More generally for any hypersurface $M^m \quad (n = m + 1)$ with unit normal U we have:

$$\Pi(p)v = v - \langle v, U(p) \rangle U(p)$$

for $p \in M^m, \ \ v \in \mathbf{E}^n$.

Lemma 11. *For $p \in M$ and $v \in T_pM$ we have:*

$$D\Pi^{\perp}(p)v = -D\Pi(p)v;$$

$$\{D\Pi(p)v\}T_pM \subset T_p^{\perp}M;$$

$$\{D\Pi(p)v\}T_p^{\perp}M \subset T_pM.$$

Proof. The first equation arises by differentiating the definition $\Pi^{\perp} = I - \Pi$. For the second differentiate the identity $\Pi = \Pi^2$ to obtain:

$$D\Pi(p)v = \Pi(p)\{D\Pi(p)v\} + \{D\Pi(p)v\}\Pi(p)\,.$$

For $w \in T_pM$ we have $\Pi(p)w = w$ so the last equation yields:

$$0 = \Pi(p)\{D\Pi(p)v\}w$$

or

$$\{D\Pi(p)v\}w \in T_p^{\perp}M$$

as required. The same argument (reading $\Pi^{\perp}$ for Π) proves the third equation of the lemma. □

Definition 12. *The field h which assigns to each $p \in M$ the linear map:*

$$h_p : T_pM \to L(T_pM, T_p^{\perp}M)$$

defined by:

$$h_p(v)w = \{D\Pi(p)v\}w$$

for $v, w \in T_pM$ is called the **second fundamental form** *of M.*

Proposition 13. *For $p \in M$, $v \in T_pM$ the adjoint $h_p(v)^* : T_p^{\perp}M \to T_pM$ of the linear map $h_p(v) : T_pM \to T_p^{\perp}M$ is given by:*

$$h_p(v)^*u = \{D\Pi(p)v\}u$$

for $u \in T_p^{\perp}M$.

Proof. Differentiate the formula

$$\langle \Pi(p)w, u\rangle = \langle w, \Pi(p)u\rangle$$

(which holds for all $w, u \in \mathbf{E}^n$ and all $p \in M$) in the direction v and use the definition and the lemma. □

Exercise 14. Choose rectangular co-ordinates $(x, y) \in \mathbf{R}^m \times \mathbf{R}^{n-m} = \mathbf{R}^n$ on E^n so that $(x(p), y(p)) = (0, 0)$ and:

$$T_pM = \mathbf{R}^m \times 0, \quad T_p^{\perp}M = 0 \times \mathbf{R}^{n-m}.$$

By the implicit function theorem there is a map $f : \mathbf{R}^m \to \mathbf{R}^{n-m}$ such that near p the equation of M is:

$$y = f(x)\,.$$

Note that:

$$f(0) = 0, \quad Df(0) = 0,$$

by the choice of co-ordinates. Show that:

$$h_p(v)w = D^2 f(0)vw$$

for $v, w \in T_pM = \mathbf{R}^m$. Thus the second fundamental form is the unique quadratic form whose graph has second-order contact with M at the point in question. One might call this graph the "osculating quadric" but to my knowledge no one has.

Exercise 15. Let $p \in M$ and $v \in T_pM$ with $\|v\| = 1$. For $r > 0$ let L denote the ball of radius r about p in the $(n - m + 1)$ dimensional affine subspace of E^n through p and parallel to the vector space $\mathbf{R} \cdot v + T_p^{\perp}M \subseteq \mathbf{E}^n$:

$$L = \{p + tv + u : u \in T_pM, t^2 + \|u\|^2 < r^2\}\,.$$

Show that for r sufficiently small, $L \cap M$ is a one-dimensional manifold (curve) with curvature vector $\ddot{\gamma}(0)$ at p given by:

$$\ddot{\gamma}(0) = h_p(v)\,.$$

(Here $\gamma : J \to L \cap M$ is the parametrization of $L \cap M$ by arclength ($\|\dot{\gamma}\| = 1$) determined by the initial conditions $\gamma(0) = p$ and $\dot{\gamma}(0) = v$.)

Covariant Derivative

Let $\gamma : \mathbf{R} \to M$, $X \in \mathcal{X}(\gamma)$, and $U \in \mathcal{X}^{\perp}(\gamma)$. The derivatives:

$$\dot{X}, \dot{U} : \mathbf{R} \to \mathbf{E}^n$$

will in general be neither tangent nor normal (example: $M =$ circle).

Figure 1: A normal plane section

Definition 16. *The fields* $\nabla X \in \mathcal{X}(\gamma)$ *and* $\nabla^{\perp} U \in \mathcal{X}^{\perp}(\gamma)$ *given by:*

$$(\nabla X)(t) = \Pi(\gamma(t))\dot{X}(t)$$

$$(\nabla^{\perp} U)(t) = \Pi^{\perp}(\gamma(t))\dot{U}(t)$$

are called the **covariant derivatives** *of* X *and* U *respectively.*

Theorem 17 (Gauss-Weingarten Equations)**.** *For* $X \in \mathcal{X}(\gamma)$ *and* $U \in \mathcal{X}^{\perp}(\gamma)$ *the formulas:*

$$\dot{X} = \nabla X + h(\dot{\gamma})X$$

$$\dot{U} = -h(\dot{\gamma})^* U + \nabla^{\perp} U$$

resolve $\dot{X}$ *and* $\dot{U}$ *into tangential and normal components respectively.*

Proof. The conditions $X \in \mathcal{X}(\gamma)$ and $U \in \mathcal{X}^{\perp}(\gamma)$ may be written:

$$X(t) = \Pi(\gamma(t))X(t)$$

$$U(t) = \Pi^{\perp}(\gamma(t))U(t).$$

Differentiate and use the definitions. □

We use notations for covariant derivatives analogous to the notations used for ordinary derivatives. Thus for $\varphi : N \to M$ and $X \in \mathcal{X}(M)$, $U \in \mathcal{X}^{\perp}(M)$ we write:

$$\nabla X(q)w = \Pi(\varphi(q))DX(q)w$$

$$\nabla^{\perp} U(q)w = \Pi^{\perp}(\varphi(q))DU(q)W$$

for $q \in N$, $w \in T_q N$ and if $x, y, \ldots, z$ are co-ordinates on N we write:

$$\nabla_x X = \Pi \cdot \partial_x X$$

$$\nabla_x^{\perp} U = \Pi^{\perp} \cdot \partial_x U$$

etc. where X and U are evaluated at $q = q(x, y, \dots, z)$ and Π and $\Pi^\perp$ are evaluated at $\varphi(q)$.

Finally when $X, Y \in \mathcal{X}(M)$ and $U \in \mathcal{X}^\perp(M)$ we define $\nabla_Y X \in \mathcal{X}(M)$ and $\nabla_Y U \in \mathcal{X}^\perp(M)$ by:

$$\begin{aligned} \nabla_Y X &= \Pi \cdot D_Y X \\ \nabla_Y^\perp U &= \Pi^\perp \cdot D_Y U \end{aligned}$$

where for any vector-valued function F defined on M we define $D_Y F$ by:

$$(D_Y F)(p) = DF(p)Y(p)$$

for $p \in M$.

In this notation the Gauss-Weingarten equations take The form:

$$D_Y X = \nabla_Y X + h(Y)X$$

$$D_Y U = \nabla_Y^\perp U - h(Y)^* U$$

for $X, Y \in \mathcal{X}(M)$ and $U \in \mathcal{X}^\perp(M)$.

Remark 18. If $X \in \mathcal{X}(M)$, $U \in \mathcal{X}^\perp(M)$ and $\gamma : \mathbf{R} \to M$, then $X \circ \gamma \in \mathcal{X}(\gamma)$ and $U \circ \gamma \in \mathcal{X}^\perp(\gamma)$ and we have by the chain rule:

$$(\nabla(X \circ \gamma))(t) = (\nabla_Y X)(\gamma(t))$$

$$(\nabla^\perp(U \circ \gamma))(t) = (\nabla_Y^\perp U)(\gamma(t))$$

for $t \in \mathbf{R}$ and where $Y \in \mathcal{X}(M)$ is any vector field such that

$$Y(\gamma(t)) = \dot{\gamma}(t)$$

for the particular value of t in question. These formulas are useful in calculations establish the relation between covariant differentiation of a vector field in the direction of another vector field and covariant differentiation of a vector field along a curve.

Let (t, s) be co-ordinates on $\mathbf{R}^2$ and $\gamma : \mathbf{R}^2 \to M$. Note that the partial derivatives of γ are vector fields along γ:

$$\partial_t \gamma, \partial_s \gamma \in \mathcal{X}(\gamma)\,.$$

Proposition 19. *We have:*

$$\nabla_t \partial_s \gamma = \nabla_s \partial_t \gamma$$

$$h(\partial_t \gamma)\partial_s \gamma = h(\partial_s \gamma)\partial_t \gamma\,.$$

Proof. By Gauss-Weingarten:

$$\partial_t \partial_s \gamma = \nabla_t \partial_s \gamma + h(\partial_t \gamma)\partial_s \gamma\,.$$

But $\partial_t \partial_s = \partial_s \partial_t$ so interchange s and t and equate tangential and normal components. □

Corollary 20. *The second fundamental form is symmetric:*

$$h_p(v)w = h_p(w)v$$

for $p \in M$; $v, w \in T_pM$.

Proof. Choose γ so that $\gamma(0,0) = p$, $\partial_t\gamma(0,0) = v$, $\partial_s\gamma(0,0) = w$. □

Parallel Transport

The various tangent spaces T_pM are all mutually isomorphic (since they have the same dimension) but there is no natural choice of isomorphism between two of them. In this section we define isomorphisms for the tangent spaces along a curve $\gamma : \mathbf{R} \to M$.

Lemma 21. *Let* $t_0 \in \mathbf{R}$ *and* $v \in T_{\gamma(t_0)}M$. *Then there is a unique* $X \in \mathcal{X}(\gamma)$ *such that*

$$\nabla X = 0, \quad X(t_0) = v.$$

Proof. Let $E_1, \ldots, E_m \in \mathcal{X}(M_0)$ be a moving frame defined in a neighborhood M_0 of $\gamma(t_0)$ and define the **Christofel symbols** with respect to the frame by:

$$\nabla_{E_j} E_j = \sum_k \Gamma_{ij}^k E_k \, .$$

(The m^3 components Γ_{ij}^k are functions on M_0.) Resolve X and $\dot{\gamma}$ into components with respect to the frame:

$$\begin{aligned} X(t) &= \sum_i \xi^i(t) E_i(\gamma(t)) \\ \dot{\gamma}(t) &= \sum_j \eta^j(t) E_j(\gamma(t)). \end{aligned}$$

The functions ξ^i and η^j are defined for t sufficiently near t_0 that $\gamma(t) \in M_0$. Applying ∇ to the equation for X yields:

$$\nabla X = \sum_k \left\{ \cdot\xi^k + \sum_{i,j} \Gamma_{ij}^k \xi^i \eta^j \right\} E_k$$

so that the equation $\nabla X = 0$ reduces to a system of m linear differential equations in the unknowns ξ^k $(k = 1, \ldots, m)$:

$$\dot{\xi}^k(t) + \sum_{i,j} \Gamma_{ij}^k(\gamma(t)) \xi^i(t) \eta^j(t) = 0 \, .$$

This (by the existence and uniqueness theorem for differential equations) establishes local existence and uniqueness.

The equations are linear so that the solution is defined on any interval on which the coefficients are defined; i.e. on any interval J for which $\gamma(J) \subset M_0$. Hence the global solution may be constructed by covering the curve γ by sets M_0 and piecing together. $\square$

Definition 22. *For $t_0, t_1 \in \mathbf{R}$ the map:*

$$\tau(\gamma, t_1, t_0) : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$$

given by:

$$\tau(\gamma, t_1, t_0)v = X(t_1)$$

(where X is given by the lemma) is called **parallel transport**. *A vector field $X \in \mathcal{X}(\gamma)$ along γ is called* **parallel** *iff $\nabla X = 0$; i.e. iff:*

$$X(t) = \tau(\gamma, t, t_0)X(t_0).$$

The vector $X(t)$ is called the **parallel transport** *of $X(t_0) \in T_{\gamma(t_0)}M$ to $T_{\gamma(t)}M$ along γ.*

Exercise 23. Show that if $E_1, \ldots, E_m$ on the co-ordinate vector fields of a local co-ordinate system $x^1, x^2, \ldots, x^m$ then:

$$\eta^j(t) = \frac{d}{dt}x^j(\gamma(t))$$

and the Christoffel symbols are symmetric in (i, j):

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

Remark 24. In case M is an affine subspace of E^n the tangent space T_pM is independent of p and we have $\dot{X} = \nabla X$. Thus in this case a parallel vector field is constant; i.e. the vector $X(t)$ (when drawn with their tails at the points $\gamma(t)$) are parallel. This partially explains the terminology; we shall see a deeper interpretation below.

Proposition 25. *Parallel transport is linear, orthogonal, and respects the operations of reparametrization, inversion and composition:*

(1) $\tau(\gamma, t_1, t_0) \in L(T_{\gamma(t_0)}M, T_{\gamma(t_1)}M)$,

(2) $\langle \tau(\gamma, t_1, t_0)v, \tau(\gamma, t_1, t_0)w \rangle = \langle v, w \rangle$,

(3) $\tau(\gamma \circ \sigma, t_1, t_0) = \tau(\gamma, \sigma(t_0), \sigma(t_1)))$,

(4) $\tau(\gamma, t_0, t_1) = \tau(\gamma, t_1, t_0)^{-1}$,

(5) $\tau(\gamma, t_2, t_1)\tau(\gamma, t_1, t_0) = \tau(\gamma, t_2, t_0)$.

for $t_0, t_1, t_2 \in \mathbf{R}$; $v, w \in T_{\gamma(t_0)}M$; and $\sigma : \mathbf{R} \to \mathbf{R}$ a diffeomorphism.

Figure 2: Parallel transport

Proof. Linearity is clear since the equation $\nabla X = 0$ is linear. The last two equations follow immediately from the definition of τ (and the existence and uniqueness theorem). For the second we must show that $\langle X(t), Y(t)\rangle$ is constant when $X, Y \in \mathcal{X}(\gamma)$ are parallel. Differentiating:

$$\begin{aligned} \frac{d}{dt}\langle X, Y\rangle &= \langle \dot{X}, Y\rangle + \langle X, \dot{Y}\rangle \\ &= \langle \nabla X, Y\rangle + \langle X, \nabla Y\rangle \\ &= 0 \end{aligned}$$

as required. (We have used the fact that:

$$\langle w, v\rangle = \langle \Pi(p)w, v\rangle$$

for $w \in \mathbf{E},\ \ v \in T_pM$. This is obvious geometrically; alternatively one may use $\Pi(p)v = v$ and $\Pi(p)^* = \Pi(p)$.) □

Everything in this section carries over word for word if the tangent field $X \in \mathcal{X}(\gamma)$ is replaced by a normal field $U \in \mathcal{X}^{\perp}(M)$ and ∇ is replaced by $\nabla^{\perp}$. For the record we give the definitions.

Definition 26. *The normal field $U \in \mathcal{X}^{\perp}(\gamma)$ along γ is called parallel if $\nabla^{\perp}U =$ 0.* **Parallel transport** *from $\gamma(t_0)$ to $\gamma(t_1)$ along γ) is the map:*

$$\tau^{\perp}(\Gamma, t_1, t_0) : T^{\perp}_{\Gamma(t_0)}M \to T^{\perp}_{\gamma(t_1)}M$$

given by:

$$\tau^{\perp}(\gamma, t_1, t_0)u = U(t_1)$$

for $u \in T^{\perp}_{\gamma(t_0)}M$ where U is the unique parallel normal field along γ satisfying the initial condition:

$$U(t_0) = u\,.$$

Figure 3: Orthogonal projection onto T_pM.

Covariant and Parallel

In the last section we defined parallel transport in terms of the covariant derivative; in this section we do the reverse. Thus ∇ and τ determine one another. They may be viewed as different incarnations of the same object.

Let X be a vector field along the curve γ. Note that:

$$\tau(\gamma, t_0, t)X(t) \in T_{\gamma(t_0)}M\,;$$

i.e. as t varies with t_0 fixed the expression on the left lies in a non-varying vector subspace of $\mathbf{E}$. Hence its derivative lies in that subspace. In fact, more is true:

Proposition 27. *Covariant differentiation may be recovered from parallel transport via the formula:*

$$(\nabla X)(t_0) = \frac{d}{dt}\tau(\gamma, t_0, t)X(t)\bigg|_{t=t_0}$$

for $X \in \mathcal{X}(\gamma)$ and $t_0 \in \mathbf{R}$. An analogous formula holds for normal fields:

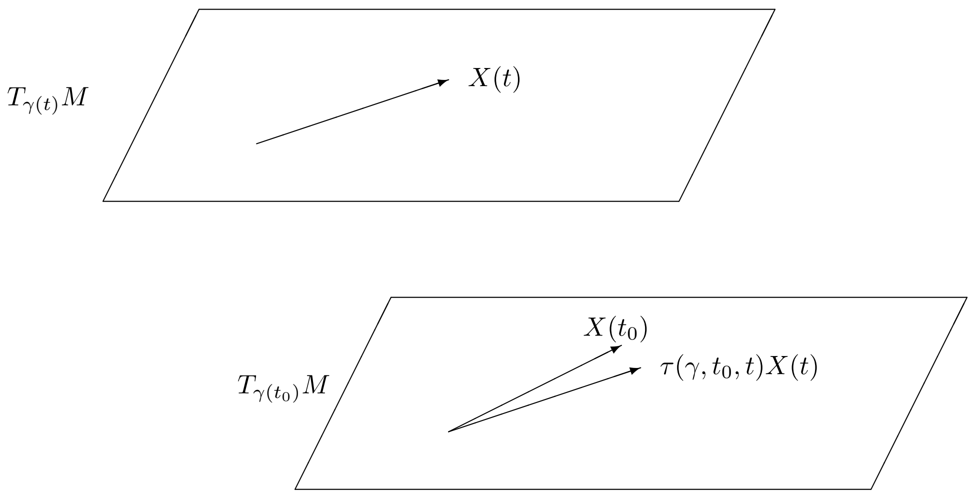
$$(\nabla^{\perp}U)(t_0) = \frac{d}{dt}\tau^{\perp}(\gamma, t_0, t)U(t)\bigg|_{t=t_0}$$

for $U \in \mathcal{X}^{\perp}(\gamma)$.

Proof. We prove the first formula; the same argument works for the second.

Let $E_1, \ldots, E_m \in \mathcal{X}(\gamma)$ be a parallel moving frame along γ; i.e.

$$E_i(t) = \tau(\gamma, t, t_0)v_i$$



where $v_1, \ldots, v_m$ form a basis for $T_{\gamma(t_0)}M$. Thus:

$$\nabla E_i = 0\,.$$

Resolve X into components:

$$X(t) = \sum_i \xi^i(t) E_i(t)\,.$$

Apply ∇:

$$\nabla X = \sum_i \dot{\xi}^i E_i\,.$$

But:

$$\begin{aligned} \tau(\gamma, t_0, t)X(t) &= \sum \xi^i(t)\tau(\gamma, t_0, t)E_i(t) \\ &= \sum_i \xi^i(t) E_i(t_0)\,. \end{aligned}$$

The proposition follows on differentiating the last equation at $t = t_0$. □

Motions

Our immediate aim in the next few sections is to define motion without sliding, twisting, or wobbling. This is the motion that results when a heavy object is rolled, with a minimum of friction, along the floor. It is also the motion of the large snowball a child creates as E rolls it into the bottom part of a snowman.

We shall eventually justify mathematically the physical intuition that either of The curves of contact in such ideal rolling may be specified arbitrarily; the other is then determined uniquely. Thus for example the heavy object may be rolled along an arbitrary curve on the floor; if that curve is marked in wet ink another curve will be traced in the object. Conversely if a curve is marked in wet ink on the object, the object may be rolled so as to trace a curve on the floor. However, if both curves are prescribed, it will be necessary to slide the object as it is being rolled if one wants to keep the curves in contact.

We shall denote a typical rigid motion of E^n by:

$$\tilde{a} \in R(E^n)$$

and denote the induced map on vectors by:

$$a \in O(\mathbf{E}^n)\,.$$

Thus in rectangular co-ordinates we have:

$$\tilde{a}(x) = ax + b$$

for $x \in \mathbf{R}^n$ where $a \in O(n)$ is an orthogonal matrix and $b \in \mathbf{R}^n$.

Let M' be another m-dimensional submanifold of E^n. Objects on M shall be denoted by the same letters as M with primes affixed. Thus for example, $\Pi(p)$ is the orthogonal projection of $\mathbf{E}^n$ on T_pM for $p \in M$.

Definition 28. *a* **motion of** M **along** M' *is a triple* $(\tilde{a}, \gamma, \gamma')$ *where*

$$\tilde{a} : \mathbf{R} \to R(E^n), \quad \gamma : \mathbf{R} \to M, \quad \gamma' : \mathbf{R} \to M'$$

such that:

$$\tilde{a}(t)(\gamma(t)) = \gamma'(t), \quad a(t)T_{\gamma(t)}M = T_{\gamma'(t)}M'$$

for $t \in \mathbf{R}$*. The curves* γ *and* γ *are called the* **curves of contact** *of the motion in* M *and* M *respectively.*

Note that a motion also matches normal vectors:

$$a(t)T^{\perp}_{\gamma(t)}M = T^{\perp}_{\gamma(t)}M$$

(as $a(t)$ *is an orthogonal transformation) and it matches affine tangent spaces:*

$$\tilde{a}(t)\tilde{T}_{\gamma(t)}M = \tilde{T}_{\gamma(t)}M$$

(by adding the two equations in the definition).

We define three operators on motions:

Definition 29 (Reparameterization)**.** *If* $\sigma : \mathbf{R} \to \mathbf{R}$ *is a diffeomorphism, and* $(\tilde{a}, \gamma, \gamma)$ *is a motion of* M *along* M'*, then* $(\tilde{a} \circ \sigma, \gamma \circ \sigma, \gamma \circ \sigma)$ *is a motion of* M *along* M*.*

Definition 30 (Inversion)**.** *If $(\tilde{a}, \gamma, \gamma)$ is a motion of M along M' then $(\tilde{a}, \gamma, \gamma)$ is a motion of M' along M where $\tilde{a}(t) = \tilde{a}(t)^{-1}$.*

Definition 31 (Composition)**.** *If $(\tilde{a}, \gamma, gamma)$ is a motion of M along M' and $(\tilde{a}, \gamma, \gamma'')$ is a motion of M' along M'' (note: same γ) then $(\tilde{a}'', \gamma, \gamma'')$ is a motion of M along M'' where $a''(t) = \tilde{a}(t) \circ \tilde{a}(t)$.*

We now give the three simplest examples of "bad" motions; i.e. motions which do not satisfy the concepts we are about to define. In all three of these examples, p is a point of M and M' is the affine tangent space to M at p:

$$M' = \tilde{T}_p M\ .$$

Example 32 (Pure Sliding). Take a non-zero vector $v \in T_p M$ (i.e. parallel to M) and let

$$\gamma(t) = p, \quad \gamma(t) = p + tv, \quad \tilde{a}(t)(q) = q + tv$$

for $q \in E^n$, $t \in \mathbf{R}$. Note that a is the identity, $\dot{\gamma} = 0$, $\dot{\gamma} = v \neq 0$, so that:

$$a\dot{\gamma} \neq \dot{\gamma}\ .$$

Example 33 (Pure Twisting). Take any curve of rotations which which acts as the identity on the affine normal space $p + T_p^{\perp} M$; thus for all t:

$$\gamma(t) = \gamma(t) = p, \quad a(t)(p) = p, \quad a(t)u = u \qquad (u \in T_p^{\perp} M).$$

Note that the derivative $\dot{a}$ maps tangent vectors to tangent vectors:

$$\dot{a}(t) T_{\gamma(t)} M \subset T_{\gamma(t)} M\ .$$

When $m \doteq 2$ and $n = 3$ the motion is a rotation about the axis through p normal to the plane M.

Example 34 (Pure Wobbling). This is the same as pure twisting except that motion is the identity on the affine tangent space M. Note that:

$$\dot{a}(t) T_{\gamma(t)}^{\perp} M \subset T_{\gamma(t)}^{\perp} M.$$

For example when $m = 1$ and $n = 3$ the motion is a rotation about the axis M. If moreover M is a plane curve, $a(t)M$ will be in a (rotating) plane which contains the fixed line M.

Sliding

When a train slides on the track (e.g. in the process of stopping suddenly), there is a terrific screech. Since we usually do not hear a screech, this means that the wheel moves along without sliding. In other words the velocity of the point of contact in the train wheel M equals the velocity of the point of contact in the track M'. But the track is not moving; hence the point of contact in the wheel is not moving. One may explain the paradox this way: the train is moving forward and the wheel is rotating around the axle. The velocity of a point on the wheel is the sum of these two velocities. When the point is on the bottom of the wheel, the two velocities cancel.

Definition 35. *A motion* $(\tilde{a}, \gamma, \gamma')$ *is* **without sliding** *iff for all* $t_0 \in \mathbf{R}$ *it satisfies the two equivalent conditions:*

$$\frac{d}{dt}\tilde{a}(t)(\gamma(t_0))\bigg|_{t=t_0} = 0, \quad a(t_0)\dot{\gamma}(t_0) = \dot{\gamma}'(t_0).$$

To see the equivalence of these two conditions differentiate the equation $\tilde{a} \cdot \gamma = \gamma'$ to obtain:

$$\frac{d}{dt}\tilde{a}(t)(\gamma(t_0))\bigg|_{t=t_0} + a(t_0)\dot{\gamma}(t_0) = \dot{\gamma}'(t_0)\,.$$

Any rigid motion which is not a translation will have fixed points; hence it is not surprising that for some $p \in E^n$ the curve $t \to \tilde{a}(t)(p) \in E^n$ has vanishing velocity at $t = t_0$ while for some $q \neq p$ the curve $t \to \tilde{a}(t)(q)$ has non-vanishing velocity. In particular, take $p = \gamma(t_0)$. The curve $t \to \tilde{a}(t)(\gamma(t_0)) \in E^n$ will in general be non-constant, but (when the motion is without sliding) its velocity will vanish at the instant $t = t_0$; i.e. at the instant when it becomes the point of contact. Thus *a motion is without sliding if and only if the point of contact is motionless.*

We remark that if the motion is without sliding we have:

$$\|\dot{\gamma}'\| = \|a\dot{\gamma}\| = \|\dot{\gamma}\|$$

so that the curves γ and γ' have the same arclength:

$$\int_{t_0}^{t_1} \|\dot{\gamma}'(t)\| dt = \int_{t_0}^{t_1} \|\dot{\gamma}(t)\| dt\,.$$

Hence any motion where any $\dot{\gamma} = 0$ and $\dot{\gamma}' \neq 0$ (e.g. the example of pure sliding above) is *not* without sliding.

Exercise 36. Give an example of a motion where $\|\dot{\gamma}'\| = \|\dot{\gamma}\|$ but the motion is *not* without sliding.

Example 37. We describe mathematically the motion of the train wheel. Let the center of the wheel move right along the x-axis and the wheel have radius one and make one revolution in 2π units of time. Then the track M' has equation $y = -1$ and we take $x^2 + y^2 = 1$ as the equation for M. Take

$$\begin{aligned} \gamma(t) &= \left(\cos\left(t - \frac{\pi}{2}\right), \sin\left(t - \frac{\pi}{2}\right)\right) \\ &= (\sin t, -\cos t)\,; \\ \gamma'(t) &= (t, -1)\,; \end{aligned}$$

and $\tilde{a}(t)$ to be given by:

$$\begin{aligned} x' &= (\cos t)x + (\sin t)y + t \\ y' &= -(\sin t)x + (\cos t)y \end{aligned}$$

for $p = (x, y)$, $\tilde{a}(t)p = (x', y')$. The reader can easily verify that this is a motion without sliding. A fixed point p_0 on M, say $p_0 = (1, 0)$ sweeps out a cycloid with parametric equations:

$$\begin{aligned} x &= (\cos t) + t \\ y &= -\sin t. \end{aligned}$$

(Check that $(\dot{x}, \dot{y}) = (0, 0)$ when $y = -1$; i.e. for $t = \left(2n + \frac{1}{2}\right)\pi$.)

Remark 38. These same formulas give a motion of a sphere M rolling without sliding along a straight line in a plane M'. Namely in rectangular co-ordinates (x, y, z) the sphere has equation $x^2 + y^2 + z^2 = 1$, the plane is $y = -1$ and the line is $y = -1$, $z = 0$. The z-co-ordinate of a point is unaffected by the motion. Note that the γ' traces out a straight line in the plane M' and the curve γ traces out a great circle on the sphere M.

Remark 39. The operations of reparametrization, inversion, and composition respect motion without sliding; i.e. if $(\tilde{a}, \gamma, \gamma')$ and $(\tilde{a}', \gamma', \gamma'')$ are motions without sliding and $\sigma : \mathbf{R} \to \mathbf{R}$ is a diffeomorphism, then the motions $(\tilde{a} \circ \sigma, \gamma \circ \sigma, \gamma' \circ \sigma)$, $(\tilde{a}(\cdot)^{-1}, \gamma', \gamma)$ and $(\tilde{a}'(\cdot)\tilde{a}(\cdot), \gamma, \gamma'')$ are also without sliding. (The proof is immediate from the definition.)

Twisting and Wobbling

A motion $(\tilde{a}, \gamma, \gamma')$ of M along M' transforms a vector field along γ into a vector field along γ' via the formula:

$$(aX)(t) = a(t)X(t)$$

for $t \in \mathbf{R}$, $X \in \mathcal{X}(\gamma)$ (so $aX \in \mathcal{X}(\gamma')$).

Definition 40. *The motion $(\tilde{a}, \gamma, \gamma')$ is* **without twisting** *iff it satisfies the following four equivalent conditions:*

(1) The instantaneous velocity of each tangent vector is normal:

$$\dot{a}(t)(T_{\gamma(t)}M) \subset T^{\perp}_{\gamma'(t)}M'.$$

(2) It transforms parallel vector fields along γ into parallel vector fields along γ':

$$\nabla X = 0 \Longrightarrow \nabla'(aX) = 0.$$

(3) It intertwines parallel transport:

$$a(t_1)\tau(\gamma, t_1, t_0) = \tau'(\gamma', t_1, t_0)a(t_0).$$

(4) It intertwines covariant differentiation:

$$\nabla'(aX) = a\nabla X.$$

We prove the conditions equivalent. Note that we have the equation:

$$a(t)\Pi(\gamma(t)) = \Pi'(\gamma'(t))a(t)$$

as this is merely a restatement of the second equation in the definition of motion. Differentiate the equation $aX = X'$ to obtain:

$$\dot{a}X + a\dot{X} = \dot{X}'.$$

Apply $\Pi'(\gamma'(t))$:

$$\Pi' \cdot (\dot{a}X) + a\nabla X = \nabla' X'.$$

Thus $\dot{a}X$ is normal iff $a\nabla X = \nabla' X'$; this establishes the equivalence (1) $\iff$ (4).

The implication (2) $\implies$ (3) is a restatement of the definition of parallel transport, while the implication (3) $\implies$ (4) follows from the definition of ∇ in terms of τ. The implication (4) $\implies$ (2) is obvious.

Definition 41. *The motion $(\tilde{a}, \gamma, \gamma')$ is* **without wobbling** *iff it satisfies the following four equivalent conditions:*

(1) The instantaneous velocity of each normal vector is tangent:

$$\dot{a}(t)(T^{\perp}_{\gamma(t)}M) \subset T_{\gamma'(t)}M'.$$

(2) It transforms parallel normal fields along γ into parallel normal fields along γ':

$$\nabla^{\perp}U = 0 \Longrightarrow {\nabla'}^{\perp}(aU) = 0.$$

(3) It intertwines normal parallel transport:

$$a(t_1)\tau^{\perp}(\gamma, t_1, t_0) = {\tau'}^{\perp}(\gamma', t_1, t_0)a(t_0)$$

(4) It intertwines normal covariant differentiation:

$${\nabla'}^{\perp}(aU) = a\nabla^{\perp}U.$$

(Here $U \in \mathcal{X}^{\perp}(M)$ so $aU \in \mathcal{X}^{\perp}(M')$ and $t, t_1, t_2 \in \mathbf{R}$.)

The proof that the four conditions are equivalent is word for word the same as before.

In summary *a motion is without twisting iff tangent vectors at the point of contact are rotating towards the normal space and it is without wobbling iff normal vectors at the point of contact are rotating towards the tangent space.* In case $m = 2$ and $n = 3$ motion without twisting means that the instantaneous axis of rotation is parallel to the tangent plane.

Remark 42. The operations of reparametrization, inversion and composition respect motion without twisting (resp. without wobbling); i.e. if $(\tilde{a}, \gamma, \gamma')$ and $(\tilde{a}', \gamma', \gamma'')$ are motions without twisting (resp. without wobbling) and $\sigma : \mathbf{R} \to \mathbf{R}$ is a diffeomorphism; then the motions $(\tilde{a} \circ \sigma, \gamma \circ \sigma, \gamma' \circ \sigma)$, $(\tilde{a}(\cdot)^{-1}, \gamma', \gamma)$, and $(\tilde{a}'(\cdot)\tilde{a}(\cdot), \gamma, \gamma'')$ are also without twisting (resp. without wobbling).

Remark 43. Given curves $\gamma : \mathbf{R} \to M$ and $\gamma' : \mathbf{R} \to M'$ and a rigid motion $\tilde{a}_0 \in R(E^n)$ satisfying:

$$\tilde{a}_0(\gamma(0)) = \gamma'(0), \quad a_0(T_{\gamma(0)}M) = T_{\gamma'(0)}M'$$

there exists a unique motion $(\tilde{a}, \gamma, \gamma')$ of M along M' (with the given γ and γ') **without twisting or wobbling** satisfying the initial condition:

$$\tilde{a}(0) = \tilde{a}_0 \, .$$

Indeed $a(t)$ is defined uniquely by conditions (3) in the definitions so that $\tilde{a}(t)$ is determined by the additional condition that $\tilde{a}(t)(\gamma(t)) = \gamma'(t)$. We prove below a somewhat harder result where the motion is without twisting, wobbling, **or sliding**. It is in this situation that γ and γ' determine one another (up to an initial condition).

Remark 44. We can now give another interpretation of parallel transport. Given $\gamma : \mathbf{R} \to M$ and $v \in T_{\gamma(0)}M$ take M' to be an affine subspace of the same dimension as M. Let $(\tilde{a}, \gamma, \gamma')$ be a motion of M along M' without twisting (and, if you like, without sliding or wobbling). Let $X' \in \mathcal{X}(\gamma')$ be the constant vector field along γ' (so that $\nabla' X' = 0$) with value $a(0)v$ and let $X \in \mathcal{X}(\gamma)$ be the corresponding vector field along γ:

$$X'(t) = a_0 v, \quad a(t)X(t) = x'(t) \, .$$

Then $X(t) = \tau(\gamma, t_0, t)v$.

To put it another way, imagine that M is a ball. To define parallel transport along a given curve γ roll the ball (without sliding etc.) along a plane M' keeping the curve γ in contact with the plane M'. Let γ' be the curve traced out in M'. If a constant vector field in the plane M' is drawn in wet ink along the curve γ' it will mark off a (covariant) parallel vector field along γ in M.

Exercise 45. Describe parallel transport along a great circle in a sphere.

Development

Development is the intrinsic version of motion without sliding or twisting.

Definition 46. *A* **development** *of M along M' is a triple (b, γ, γ') where $\gamma : \mathbf{R} \to M$ and $\gamma' : \mathbf{R} \to M'$ are curves and b is a function which assigns to each $t \in \mathbf{R}$ an orthogonal linear isomorphism:*

$$b(t) : T_{\gamma(t)}M \to T_{\gamma'(t)}M'$$

and such that the following equivalent conditions are satisfied:

(1) There is a motion $(\tilde{a}, \gamma, \gamma')$ *without sliding, twisting, or wobbling such that*

$$b(t) = a(t) \mid T_{\gamma(t)}M$$

for $t \in \mathbf{R}$*;*

(2) there is a motion exactly as in (1) except possibly not without wobbling;

(3) b *intertwines parallel transport:*

$$b(t_1)\tau(\gamma, t_1, t_0) = \tau'(\gamma', t_1, t_0)b(t_0)$$

and satisfies:

$$b(t)\dot{\gamma}(t) = \dot{\gamma}'(t).$$

We prove the equivalence of three conditions. The implications (1) $\Longrightarrow$ (2) and (2) $\Longrightarrow$ (3) are obvious. For (3) $\Longrightarrow$ (1) choose any $\tilde{a}_0 \in R(E^n)$ such that $\tilde{a}_0(\gamma(0)) = \gamma'(0)$ and $a_0 \mid T_{\gamma(0)}M = b(0)$. There is a unique motion $(\tilde{a}, \gamma, \gamma')$ without twisting or wobbling such that $\tilde{a}(0) = \tilde{a}_0$. As a and b both intertwine τ and τ' we have $a(t) \mid T_{\gamma(t)}M = b(t)$. Hence $a\dot{\gamma} = b\dot{\gamma} = \dot{\gamma}'$ so the motion $(\tilde{a}, \gamma, \gamma')$ is also without sliding as required.

Remark 47. The operations of reparametrization, inversion, and composition yield developments when applied to developments; i.e. if (b, γ, γ') is a development of M along M', (b', γ', γ'') is a development of M' along M'', and $\sigma : \mathbf{R} \to \mathbf{R}$ is a diffeomorphism, then $(b \circ \sigma, \gamma \circ \sigma, \gamma' \circ \sigma)$, $(b'(\cdot)^{-1}, \gamma', \gamma)$, and $(b'(\cdot)b(\cdot), \gamma, \gamma'')$ are all developments.

Theorem 48 (Developments along affine subspaces)**.** *Assume* $M' = E^m$ *is an affine subspace of* E^n *and that we are given* $\gamma : \mathbf{R} \to M$*,* $o' \in M'$*,* $t_0 \in \mathbf{R}$*, and an orthogonal isomorphism:*

$$b_0 : T_{\gamma(t_0)}M \to T_{o'}E^m = \mathbf{E}^m\,.$$

Then there exists a unique development (b, γ, γ') *satisfying the initial conditions:*

$$b(t_0) = b_0, \quad \gamma'(t_0) = o'.$$

Corollary 49. *Assume (*$M' = E^m$ *and) that we are given* $\gamma : \mathbf{R} \to M$*,* $t_0 \in \mathbf{R}$*, and* $\tilde{a}_0 \in R(E^n)$ *such that:*

$$\tilde{a}_0(\tilde{T}_{\gamma(t_0)}M) = E^m\,.$$

Then there exists a unique motion $(\tilde{a}, \gamma, \gamma')$ *without sliding, twisting or wobbling satisfying the initial condition:*

$$\tilde{a}(t_0) = \tilde{a}_0\,.$$

Proof. We first prove uniqueness. Let the development (b, γ, γ') be given and choose a basis $e'_1, \ldots, e'_m$ for $\mathbf{E}^m$. Define a moving frame $E_1, \ldots, E_m$ along γ by:

$$b(t)E_i(t) = e'_i\,.$$

Since (b, γ, γ') is a development the E_i are parallel:

$$\nabla E_i = 0$$

and satisfy the initial condition:

$$b_0 E_i(t_0) = e'_i\,.$$

This determines $B(t) : T_{\gamma(t)}M \to \mathbf{E}^m$ uniquely. Define $\xi^i = \xi^i(t)$ by:

$$\dot{\gamma}(t) = \sum_i \xi^i(t) E_i(t);$$

since $b\dot{\gamma} = \dot{\gamma}'$ the curve γ' is uniquely determined by:

$$\dot{\gamma}'(t) = \sum_i \xi^i(t) e'_i, \quad \gamma'(t_0) = o'.$$

For existence note that the formulas just derived can be taken as definitions of b and γ'. □

Remark 50. Below we prove a theorem which asserts the existence of γ given γ'.

Remark 51. Any two developments (b_1, γ, γ'_1) and (b_2, γ, γ'_2) of the same curve γ in M are related by:

$$b_2(t) = cb_1(t), \quad \gamma'_2(t) = c\gamma'_1(t) + v$$

where c is a constant linear orthogonal isomorphism and v is a constant vector. (This follows from uniqueness since one easily checks that if (b_2, γ, γ'_2) is defined from (b_1, γ, γ'_1) by these formulas then it is a development.)

Affine Parallel Transport

Now let (b, γ, γ') be any development of M along an affine space $M' = E^m$ and define for each $t \in \mathbf{R}$ an affine map:

$$\tilde{b}(t) : \tilde{T}_{\gamma(t)}M \to M' = E^m = \gamma'(t) + \mathbf{E}^m$$

of affine tangent spaces by imposing the conditions that $\tilde{b}(t)$ induce the linear map:

$$b(t) : T_{\gamma(t)}M \to \mathbf{E}^m$$

on vectors and that $b(t)$ carries the point $\gamma(t) \in M$ to the point $\gamma'(t) \in M' = E^m$:

$$b(t)\gamma(t) = \gamma'(t)\,.$$

In other words $\tilde{b}(t)$ is the restriction to the affine tangent space of $\tilde{a}(t)$ where $(\tilde{a}, \gamma, \gamma')$ is a motion of M along $M' = E^n$ without sliding or twisting.

Definition 52. *For $t_0, t_1 \in \mathbf{R}$ the affine isometry of affine tangent spaces:*

$$\tilde{\tau}(\gamma, t_1, t_0) : \tilde{T}_{\gamma(t_0)}\tilde{M} \to T_{\gamma(t_1)}M$$

given by:

$$\tilde{\tau}(\gamma, t_1, t_0) = \tilde{b}(t_1)^{-1}\tilde{b}(t_0)$$

is called **affine parallel transport** $\gamma(t_0)$ *to* $\gamma(t_1)$ *along* γ*.*

Affine parallel transport is independent of the choice of the development (b, γ, γ') of M along E^m but depends only on the curve γ. As a notation indicates, the affine map $\tilde{\tau}$ induces the (linear) parallel transport map:

$$\tau(\gamma, t_1, t_0) : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$$

on the (vector) tangent spaces.

To see the independence of the development note that by the last remark b_1 and b_2 arising from different developments are related by:

$$\tilde{b}_2(t) = \tilde{c} \cdot \tilde{b}_1(t)$$

where $\tilde{c}$ is an affine isomorphism independent of t. Thus:

$$\tilde{b}_2(t_1)^{-1} \cdot \tilde{b}_2(t_0) = \tilde{b}_1(t_1)^{-1} \cdot \tilde{b}_1(t_0)$$

as required.

We next prove the analog for affine parallel transport of the formula:

$$\left.\frac{d}{dt}\tau(\gamma, t_0, t)X(t)\right|_{t=t_0} = \nabla X(t_0)$$

for $X \in \mathcal{X}(\gamma)$. Choose a vector field X along γ and form the the “affine field”:

$$\tilde{X}(t) = \gamma(t) + X(t)\,.$$

Thus $\tilde{X}(t)$ lies in the affine tangent space at $\gamma(t)$:

$$\tilde{X}(t) \in \tilde{T}_{\gamma(t)}M\,.$$

Apply $\tilde{\tau}(\gamma, t_0, t)$ to get a curve in a fixed affine space:

$$\tilde{\tau}(\gamma, t_0, t)(\tilde{X}(t)) \in \tilde{T}_{\gamma(t_0)}M\,.$$

Proposition 53. *We have:*

$$\frac{d}{dt}\tilde{\tau}(\gamma, t_0, t)(\tilde{X}(t))\Big|_{t=t_0} = \nabla X(t_0) + \dot{\gamma}(t_0) \in T_{\gamma(t_0)}M\,.$$

Proof. Choose a development (b, γ, γ') with:

$$E^m = m' = \tilde{T}_{\gamma(t_0)}M, \quad \tilde{b}(t_0) = \text{ identity.}$$

Then:

$$\tilde{\tau}(\gamma, t_0, t) = \tilde{b}(t)$$

so:

$$\begin{aligned} \tilde{\tau}(\gamma, t_0, t)(\tilde{X}(t)) &= \tilde{b}(t)(\gamma(t)) + b(t)X(t) \\ &= \gamma'(t) + \tau(\gamma, t_0, t)X(t)\,. \end{aligned}$$

Hence:

$$\begin{aligned} \frac{d}{dt}\tilde{\tau}(\tilde{X})\Big|_{t=t_0} &= \dot{\gamma}'(t_0) + \nabla X(t_0) \\ &= \dot{\gamma}(t_0) + \nabla X(t_0) \end{aligned}$$

as required. □

Corollary 54. *Affine parallel transport is given by:*

$$\tilde{\tau}(\gamma, t, t_0)(\gamma(t_0) + v_0) = \gamma(t) + X(t) + \tau(\gamma, t, t_0)v_0$$

for $v_0 \in T_{\gamma(t_0)}M$ *where* $X \in \mathcal{X}(\gamma)$ *is the solution of the (linear inhomogeneous) ordinary differential equation:*

$$\nabla X + \dot{\gamma} = 0, X(t_0) = 0.$$

Exercise 55. Consider the case where $m = 1$ and $n = 2$; i.e. M is a plane curve. Show that in this case the curve $\alpha : \mathbf{R} \to E^2$ given by:

$$\alpha(t) = \tilde{\tau}(\gamma, t, t_0)(\gamma(t_0))$$

is an **involute** of γ. This means that

$$\gamma(t) - \alpha(t) = s(t)X(t)$$

where X is the unit tangent vector:

$$X(t) = \|\dot{\gamma}(t)\|^{-1}\dot{\gamma}(t)$$

and s is the arclength:

$$s(t) = \int_{t_0}^{t} \|\dot{\gamma}\| dt\,.$$

Figure 4: An Evolute

Verify that γ is the **evolute** of α: the locus of centers of curvature:

$$\gamma - \alpha = \left\| \frac{d^2\alpha}{dr^2} \right\|^{-2} \frac{d^2\alpha}{dr^2}$$

where r is the arclength parameter of α:

$$r(t) = \int_{t_0}^{t} \|\alpha\| dt\,.$$

Thus α is generated from γ by unwiding a string wound around γ while γ is the locus of intersections of infinitessimally near normal lines to α.

The Frame Bundle

A frame for an m-dimensional vector space V is an ordered basis $e = (e_1, \ldots, e_m)$ for V. This is the same as a vector space isomorphism $e : \mathbf{R}^m \to V^m$:

$$e(\xi) = \sum_i \xi^i e_i$$

for $\xi = (\xi^1, \ldots, \xi^m) \in \mathbf{R}^m$. Denote by $L_{\mathrm{is}}(\mathbf{R}^m, V)$ the set of all frames for V:

$$L_{\mathrm{is}}(\mathbf{R}^m, V) = \{e \in V^{\times m} : e_1, \ldots, e_m \ \text{ a basis}\}\,.$$

Note that the group $GL(m)$ of all invertible $m \times m$ matrices acts on $L_{\mathrm{is}}(\mathbf{R}^m, V)$ via the formula:

$$(a^* e)_j = \sum_i a^i_j e_i$$

for $a \in GL(m)$ and $e \in L_{\mathrm{is}}(\mathbf{R}^m, V)$; in terms of isomorphisms:

$$(a^* e)(\xi) = e(a\xi)$$

for $a : \mathbf{R}^m \to \mathbf{R}^m$, $e : \mathbf{R}^m \to V$, and $\xi \in \mathbf{R}^m$. The action is called a right action because $a \mapsto a^*$ is an antihomomorphism:

$$(a_1 a_2)^* = a_2^* a_1^* \,;$$

this law looks like an associative law if we write the group element on the right:

$$(ea_1)a_2 = e(a_1 a_2)$$

where $ea = a^* e$. The action is free:

$$a^* e = e \Longrightarrow a = \text{ identity}$$

and transitive:

$$e, e' \in L_{\mathrm{is}}(\mathbf{R}^m, V) \Longrightarrow \exists a \in LG(m) : e' = a^* e$$

so that each choice of $e \in L_{\mathrm{is}}(\mathbf{R}^m, V)$ determines a diffeomorphism:

$$GL(m) \to L_{\mathrm{is}}(\mathbf{R}^m, V) : La \mapsto a^* e$$

from the (open) set of invertible matrices onto the open subset

$$L_{\mathrm{is}}(\mathbf{R}^m, V) \subset V^{\times m}.$$

Note however that $L_{\mathrm{is}}(\mathbf{R}^m, V)$ is *not* naturally a group (although it is diffeomorphic to one) for the diffeomorphism $a \mapsto a^* e$ depends on the choice of e.

Definition 56. *The* **frame bundle** $F(M)$ *of the submanifold* $M \subset E^n$ *is the set:*

$$F(M) = \{(p, e) : p \in M, e \in F(M)_p\}$$

where $F(M)_p$ *is the space of frames of the tangent space at* p*:*

$$F(M)_p = L_{\mathrm{is}}(\mathbf{R}^m, T_p M)\,.$$

Define a right action of $GL(m)$ *on* $F(M)$ *by:*

$$a^*(p, e) = (p, a^* e)$$

for $a \in GL(m)$ *and* $(p, e) \in F(M)$*.*

Proposition 57. *The frame bundle $F(M)$ is a submanifold of $E^n \times (\mathbf{E}^n)^{\times m}$ of dimension $m + m^2$ and the projection:*

$$F(M) \to M : (p, e) \to p$$

is locally trivial. The orbits of the $GL(m)$ action are the fibers of this projection:

$$GL(m)^*(p, e) = F(M)_p$$

for $(p, e) \in F(M)$.

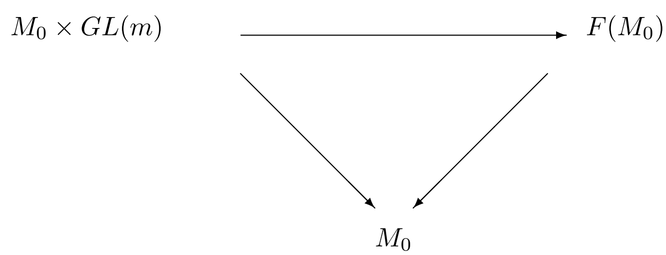
Proof. Any moving frame $E_1, \ldots, E_m \in \mathcal{X}(M_0)$ defined over an open subset M_0 of M gives a bijection:

$$M_0 \times GL(m) \to F(M_0) : (p, a) \to a^*E(p)$$

where:

$$E(p) = (E_1(p), \ldots, E_m(p)) \in F(M)_p\,.$$

This bijection (when composed with a parametrization of M_0) gives a parametrization of the open subset $F(M_0)$ of $F(M)$. The diagram:



clearly commutes so the assertions of the proposition are evident. □

Exercise 58. Denote by $O(M)$ the **orthonormal** frame bundle of M:

$$O(M) = \{(p, e) \in F(M) : \langle e_i, e_j \rangle = \delta_{ij}\}\,.$$

Show that $O(M)$ is a submanifold of $F(M)$, that the projection:

$$O(M) \to M : (p, e) \mapsto p$$

is locally trivial, and that the action of $GL(m)$ on $F(M)$ restricts to an action of $O(m)$ on $O(M)$ whose orbits are the fibers:

$$O(M)_p = O(M) \cap F(M)_p$$

of this projection.

Horizontal Lifts

We have previously used the idea of moving frame along a curve $\gamma : \mathbf{R} \to M$; i.e. vector fields $E_1, \ldots, E_m \in \mathcal{X}(\gamma)$ which give a basis for $T_{\gamma(t)}M$ for each $t \in \mathbf{R}$. Such a frame can be viewed as a curve $\beta : \mathbf{R} \to F(M)$ in the frame bundle:

$$\beta(t) = (\gamma(t), E_1(t), \ldots, E_m(t))$$

for $t \in \mathbf{R}$.

Definition 59. We call such a curve β a **lift** of γ; thus $\beta : \mathbf{R} \to F(M)$ lifts $\gamma : \mathbf{R} \to M$ iff $\pi \circ \beta = \gamma$ where $\pi : F(M) \to M$ is the projection. When the vector fields E_i are parallel along γ the curve β is called **horizontal**; thus a horizontal curve is one of form:

$$\beta(t) = (\gamma(t), \tau(\gamma, 0, t)e)$$

where $e = (e_1, \ldots, e_m) \in F(M)_{\gamma(0)}$. For $(p, e) \in F(M)$ we define two subspaces of the tangent space $T_{(p,e)}F(M)$. These are the **vertical space** $V_{(p,e)}$ which is simply the tangent space to the fiber:

$$V_{(p,e)} = T_{(p,e)}F(M)_p$$

and the **horizontal space** $H_{(p,e)}$ consisting of all tangent vectors to horizontal curves through (p, e):

$$H_{(p,e)} = \{\dot\beta(0) \mid \beta(0) = (p, e), \beta : \mathbf{R} \to F(M) \text{ horizontal}\}.$$

Proposition 60. *The horizontal space is a vector space complement to the vertical space:*

$$T_{(p,e)}F(M) = V_{(p,e)} \oplus H_{(p,e)}.$$

The vertical space is the kernel of the derivative of the projection $\pi : F(M) \to M$:

$$V_{(p,e)} = \ker D\pi(p, e), \quad D\pi(p, e) : T_{(p,e)}F(M) \to T_pM$$

so that this derived projection restricts to an isomorphism:

$$D\pi(p, e) \mid H_{(p,e)} : H_{(p,e)} \to T_pM.$$

Finally a curve $\beta : \mathbf{R} \to F(M)$ is horizontal if and only if its tangent vector lies in the horizontal space:

$$\dot\beta(t) \in H_{\beta(t)} \qquad (\forall T \in \mathbf{R}).$$

If $\beta : \mathbf{R} \to F(M)$ is horizontal and $a \in GL(m)$ then $a^*\beta : \mathbf{R} \to F(M)$ is also horizontal

Proof. Much of this is a trivial restatement of the definition. One sees that:

$$V_{(p,e)} = \{(\dot{p}, \dot{e}) \in T_{(p,e)}F(M) : \dot{p} = 0\}.$$

While

$$H_{(p,e)} = \{(\dot{p}, \dot{e}) : \dot{e}_i = h_p(\dot{p})e_i \quad i = 1, \ldots, m\}.$$

(This last equation is because according to the Gauss-Weingarten equations, the equation $\nabla E_i = 0$ for $E_i \in \mathcal{X}(\gamma)$ takes the form:

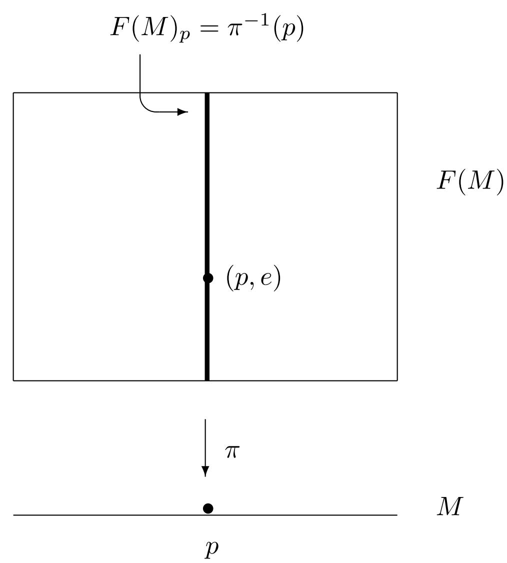
$$\dot{E}_i = h(\dot{\gamma})E_i$$

where h is the second fundamental form.) The formula for $H_{(p,e)}$ shows that it intersects $V_{(p,e)}$ only in the zero vector and that any vector $(\dot{p}, \dot{e})$ can be resolved into components:

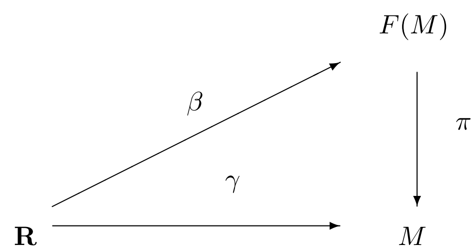
$$(\dot{p}, \dot{e}) = (0, \dot{e} - h(\dot{p})e) + (\dot{p}, h(\dot{p})e)$$

as required. The formula for $H_{(p,e)}$ also shows that for $\beta : \mathbf{R} \to F(M)$ the condition that β have the form $\beta = (\gamma, \tau e)$ is the same as the condition that $\dot{\beta} \in H_\beta$. □

Remark 61. The reason for the terminology is that one draws the following extremely crude picture of the frame bundle:



One thinks of $F(M)$ as "lying over" M. One would then represent the equation $\gamma = \pi \circ \beta$ by a commutative diagram:



hence the word "lift". The vertical space is tangent to the vertical line in the picture while the horizontal space is transverse to the vertical space. This crude imagery can be extremely helpful.

Exercise 62. Recall the orthonormal frame bundle $O(M) \subset F(M)$ defined above. Show that the horizontal space $H_{(p,e)}$ is tangent to $O(m)$:

$$H_{(p,e)} \subset T_{(p,e)}O(M)$$

and the proposition remains true if $O(M)$ is read for $F(M)$ and:

$$V'_{(p,e)} = T_{(p,e)}O(M)_p$$

is read for $V_{(p,e)}$.

The Development Theorem

We now use the frame bundle to prove the existence of developments. For the first time it is necessary to talk about developments (b, γ, γ') or motions $(\tilde{a}, \gamma, \gamma')$ where the curves are not necessarily defined for all $t \in \mathbf{R}$ but possibly only on some interval $J \subset \mathbf{R}$. The definitions are unchanged. The concept of completeness used in the formulation of the theorem will be defined in the proof and discussed at greater length below. As usual M and M' denote arbitrary m-dimensional submanifolds of Euclidean space E^n.

Theorem 63 (Existence and Uniqueness of Developments)**.** *Let $\gamma' : \mathbf{R} \to M'$, $o \in M$, $t_0 \in \mathbf{R}$, and an orthogonal isomorphism:*

$$b_0 : T_oM \to T_{\gamma'(t_0)}M'$$

be given. Then there exists a development $t \to (b(t), \gamma(t), \gamma'(t))$ of M along M' defined for $t \in J$ where J is an open interval containing t_0 satisfying the initial conditions:

$$b(t_0) = b_0, \quad \gamma(t_0) = o.$$

Any two such developments agree wherever both are defined. If M is complete (in the sense defined below), there is a globally defined development, i.e. one with $J = \mathbf{R}$.

Corollary 64. *Let $\gamma' : \mathbf{R} \to M'$, $0 \in M$, $t_0 \in \mathbf{R}$, and a rigid motion $\tilde{a}_0 \in R(E^n)$ satisfying:*

$$\tilde{a}(o) = \gamma'(t_0), \quad a_0 T_o M = T_{\gamma'(t_0)} M'$$

be given. Then there exists a motion without sliding, twisting or wobbling of M along M' defined on an open interval J containing t_0 and satisfying the initial condition:

$$\tilde{a}(t_0) = \tilde{a}_0 \, .$$

Any two such agree wherever both are defined and if M is complete there is one with $J = \mathbf{R}$.

Definition 65. As any two developments (b_1, γ_1, γ') and (b_2, γ_2, γ') with $\gamma_1(t_0) = \gamma_2(t_0)$ and $b_1(t_0) = b_2(t_0)$ agree on $J_1 \cap J_2$ there is a development defined on $J_1 \cup J_2$; hence there is a unique maximally defined development (b, γ, γ') defined on J; any development (b_1, γ_1, γ') with $\gamma_1(t_0) = \gamma(t_0)$ and $b_1(t_0) = b(t_0)$ satisfies $J_1 \subset J$. We call this the *(maximally defined) development corresponding to the given γ' and the given initial conditions $\gamma(t_0) = 0$ and $b(t_0) = b_0$.*

Remark 66. The statment of the theorem is essentially symmetric in M and M' as the operation of inversion carries developments to developments. Hence given $\gamma : \mathbf{R} \to M$, $o' \in M'$, $t_0 \in \mathbf{R}$, and $b_0 : T_{\gamma(t_0)} M \to T_{o'} M'$, we may speak of the development (b, γ, γ') corresponding to γ with initial conditions $\gamma'(t_0) = o'$ and $b(t_0) = b_0$.

Proof of theorem. We have already proved the theorem in case M is a Euclidean space. Since the operations of composition and inversion yield developments when applied to developments, we may assume w.l.o.g. that M' is a Euclidean space E^m.

Our first step is to find a differential equation whose solutions correspond to developments.

Definition 67. *Given a smooth map $\xi : \mathbf{R} \to \mathbf{R}^m$ the time dependent vectorfield $B = B(t, p, e)$ on $F(M)$ characterized by the conditions that $B(t, p, e)$ be horizontal:*

$$B(t, p, e) \in H_{(p,e)} \subset T_{(p,e)} F(M)$$

and project to $e(\xi(t)) \in T_p M$:

$$D\pi(p, e) B(t, p, e) = e(\xi(t))$$

where $\pi : F(M) \to M$ is the projection ($\pi(p, e) = p$) is called the **basic vectorfield** *corresponding to ξ.*

We remark that B can be given explicitly via the formulas:

$$B(t, p, e) = (\hat{p}, \hat{e})$$

where

$$\begin{aligned} e &= (e_1, \dots, e_m) \in F(M)_p \\ \hat{p} &= \sum_i \xi^i(t) e_i \\ \hat{e} &= (\hat{e}_1, \dots, \hat{e}_m) \\ \hat{e}_i &= h_p(\hat{p}) e_i \end{aligned}$$

for $t \in \mathbf{R}$ and $(p, e) \in F(M)$.

As B is horizontal any integral curve $\beta = \beta(t) \in F(M)$:

$$\dot{\beta} = B(t, \beta)$$

has form:

$$\beta(t) = (\gamma(t), \tau(\gamma, t, t_0) e_0)$$

where $e_0 \in F(M)_{\gamma(t_0)}$. As the parallel transport map:

$$\tau(\gamma, t, t_0) : T_{\gamma(t_0)} M \to T_{\gamma(t)} M$$

is orthogonal it follows that B is tangent to the orthonormal frame bundle; i.e. that $\beta(t) \in O(M)$ whenever $\beta(t_0) \in O(M)$ (i.e. $e_0 \in O(M)_{\gamma(t_0)}$).

Now choose an orthonormal frame $e' = (e'_1, \dots, e'_m)$ for $\mathbf{E}^m = T_{o'} E^m$. The formulas:

$$b(t) E_i(t) = e'_i$$

$(i = 1, \dots, m)$ establish a bijective correspondence between curves $\beta : J \to F(M)$:

$$\beta(t) = (\gamma(t), E_1(t), \dots, E_m(t))$$

and triples (b, γ, γ') where $\gamma : J \to M$ and for each $t \in J$, $b(t) : T_{\gamma(t)} M \to \mathbf{E}^m$ is a linear isomorphism. Let $\xi : \mathbf{R} \to \mathbf{R}^m$ be given by:

$$\dot{\gamma}'(t) = \sum_i \xi^i(t) e'_i .$$

Denote by B the corresponding basic vectorfield.

Claim . . *The curve $\beta : J \to O(M)$ is an integral curve for B:*

$$\dot{\beta}(t) = B(t, \beta(t))$$

if and only if the corresponding triple (b, γ, γ') is a development.

Indeed the equation $\dot{\beta} = B(t, \beta)$ can be written:

$$\dot{\beta} \in H_\beta, \quad \dot{\gamma} = \Sigma \xi^i E_i$$

which in turn takes the form:

$$b(t) \tau(\gamma, t, t_0) = b(t_0), \quad b(t) \dot{\gamma}(t) = \dot{\gamma}'(t)$$

for $t \in J$ and these are precisely the equations for a development.

The development theorem is now an immediate consequence of the existence and uniqueness theorem for ordinary differential equations. □

Definition 68. *The manifold M is* **complete** *iff it satisfies the following two equivalent conditions:*

(1) for every curve $\gamma' : \mathbf{R} \to E^m$ in Euclidean space, every $t_0 \in \mathbf{R}$, every point $o \in M$, and every orthogonal isomorphism $b_0 : T_0 M \to \mathbf{E}^m$ the maximally defined development (b, γ, γ') determined by the initial conditions $\gamma(t_0) = 0$ and $b(t_0) = b_0$ is defined for all $t \in \mathbf{R}$;

(2) for every curve $\xi : \mathbf{R} \to \mathbf{R}^m$ the corresponding basic vectorfield B is complete; i.e. for each $(t_0, o, e_0) \in \mathbf{R} \times F(M)$ the initial value problem:

$$\dot{\beta}(t) = B(t, \beta(t)), \quad \beta(t_0) = (o, e_0)$$

has a globally solution $\beta : \mathbf{R} \to F(M)$.

Remark 69. We have already noted that a basic vectorfield B is tangent to the orthonormal frame bundle. Now note that if $\beta(t) = (\gamma(t), E(t))$ is an integral curve to B so is $a^*\beta(t) = (\gamma(t), a^*E(t))$ where $a \in GL(m)$ is a (constant) invertible matrix. Since any frame at $\gamma(t_0)$ is carried to any other by a suitable matrix $a \in GL(m)$ it follows that if *one* integral curve β with $\gamma(t_0) = 0$ is defined for all t then *every* integral curve β with $\gamma(t_0) = 0$ is defined for all t. In particular, the vectorfield B is complete if and only if its restriction $B \mid \mathbf{R} \times O(M)$ to the orthonormal frame bundle is complete.

It is of course easy to give an example of a manifold which is *not* complete; if (b, γ, γ') is any development then $M \backslash \{\gamma(t_1)\}$ is not complete as the given development is only defined for $t < t_1$. Below we give equivalent characterizations of completeness; we will see that any closed submanifold of E^n is complete.

Geodesics

The concept of a geodesic in a manifold generalizes that of a straight line in Euclidean space. A straight line has parametrizations of form $t \to p + \sigma(t)v$ where $\sigma : \mathbf{R} \to \mathbf{R}$ is a diffeomorphism, $p \in E^n$, $v \in \mathbf{E}^n$; different σ yield different parametrizations of the same line. Certain parametrizations are preferred; viz. those parametrizations which are "proportional to arclength", i.e. where $\sigma(t) = at + b$ (for constants $a, b \in \mathbf{R}$) so that the tangent vector $\dot{\sigma}(t)v$ has constant length. The same distinctions can be made for geodesics. Some authors use the term geodesic to include all parametrizations of a geodesic while others restrict the term to cover only geodesics parametrized proportional to arclength. We follow the latter course, referring to the more general concept as a "reparametrized geodesic". (Thus a reparametrized geodesic need not be a geodesic.)

Definition 70. *Let $J = [a, b]$ be a bounded closed interval in* $\mathbf{R}$ *and $\gamma : J \to M$*

be a smooth curve in M. We define the **length** $L(\gamma)$ *and* **energy** $E(\gamma)$ *of* γ *by:*

$$\begin{aligned} L(\gamma) &= \int_a^b \|\dot\gamma(t)\| dt \\ E(\gamma) &= \int_a^b \|\dot\gamma(t)\|^2 dt. \end{aligned}$$

A **variation** *of* $\gamma : J \to M$ *is a family* $\gamma_\lambda : J \to M$ *of curves where* λ *ranges over an open interval* $I \subset \mathbf{R}$ *about* 0 *such that the map:*

$$I \times J \to M : (\lambda, t) \mapsto \gamma_\lambda(t)$$

is smooth and:

$$\gamma_0 = \gamma\,.$$

The variation **has fixed endpoints** *iff*

$$\gamma_\lambda(a) = \gamma(a), \ \ \gamma_\lambda(b) = \gamma(b)$$

for all $\lambda \in I$.

We shall generally suppress notation for the endpoints of J. When $\gamma(a) = p$ and $\gamma(b) = q$ we say γ is a **curve from** p to q. We can always compose γ with an affine reparametrization:

$$t' = (b-a)t + a$$

to get a new curve:

$$\gamma'(t) = \gamma(t')$$

defined for $0 \le t \le 1$. Note that:

$$L(\gamma') = L(\gamma)$$

and

$$E(\gamma') = (b-a)E(\gamma)\,.$$

More generally note that the arclength integral $L(\gamma)$ (but *not* the energy integral $E(\gamma)$) is invariant under **reparametrization**; i.e. if $\sigma : J' \to J$ is a diffeomorphism) then:

$$L(\gamma \circ \sigma) = L(\gamma)\,.$$

This is simple the change of variables formula for the integral of a real valued function of a real variable.

Definition 71. *A curve* $\gamma : J \to M$ *defined on a closed bounded interval is called a* **geodesic** *iff it satisfies the following equivalent conditions:*

(1) It is an **extremal of the energy integral***:*

$$\frac{d}{d\lambda} E(\gamma_\lambda)|_{\lambda=0} = 0$$

for every variation $\{\gamma_\lambda\}$ *of* γ *with fixed endpoints;*

(2) It is **parametrized proportional to arclength***:*

$$\|\dot{\gamma}(t)\| = \quad constant$$

and is either a constant (i.e. $\dot{\gamma} \equiv 0$*) or is and* **extremal of the arclength integral***:*

$$\frac{d}{d\lambda} L(\gamma_\lambda)|_{\lambda=0} = 0$$

for every variation $\{\gamma_\lambda\}$ *of* γ *with fixed endpoints.*

(3) Its acceleration vector is normal to M*:*

$$\ddot{\gamma}(t) \in T^{\perp}_{\gamma(t)} M \qquad (\forall t \in J).$$

(4) Its velocity vector is parallel:

$$\nabla \dot{\gamma}(t) = 0 \qquad (\forall t \in J).$$

(5) Some (and hence every) development (b, γ, γ') *of* γ *along an affine space* $M' = E^m$ *is a straight line parametrized proportional to arclength:*

$$\gamma'(t) = \gamma'(t_0) + (t - t_0)\dot{\gamma}'(t_0) \qquad (\forall t \in J)$$

for some (each) $t_0 \in J$*.*

When the interval J is infinite conditions (1) and (2) are meaningless since the integrals $E(\gamma)$ and $L(\gamma)$ are infinite. But note that (e.g. using (3) or (4) as the definition) $\gamma \mid J'$ is a geodesic whenever γ is a geodesic and J' is a closed subinterval of J. Hence for an infinite interval J (say $J = \mathbf{R}$) we call $\gamma : J \to M$ a **geodesic** iff $\gamma \mid J'$ is a geodesic for every closed bounded subinterval $J' \subset J$; this is equivalent to the equivalent conditions (3), (4), and (5) above.

Proof that the conditions are equivalent.

(3) $\iff$ (4). This is immediate from the Gauss-Weingarten equations: $\nabla\dot{\gamma}$ is the tangential component of $\ddot{\gamma}$ and vanishes $\iff$ $\ddot{\gamma}$ is normal.

(1) $\iff$ (3). Define $X \in \mathcal{X}(\gamma)$ by:

$$X(t) = \frac{d}{d\lambda}\gamma_\lambda(t)|_{\lambda=0}\,.$$

Then

$$\begin{aligned}
\frac{d}{d\lambda} E(\gamma_\lambda)|_{\lambda=0} &= \int \frac{d}{d\lambda}\|\dot{\gamma}_\lambda(t)\|^2|_{\lambda=0} dt \\
&= 2\int \langle \dot{\gamma}(t), \dot{X}(t)\rangle dt \\
&= -2\int \langle \ddot{\gamma}(t), X(t)\rangle dt
\end{aligned}$$

where in the last equation we have used integration by parts (the boundary terms vanish since X vanishes at the endpoints since $\{\gamma_\lambda\}$ is a variation with endpoints fixed). Clearly then (3) $\Longleftrightarrow$ (1) since:

$$X(t) \in T_{\gamma(t)}M$$

so that

$$\langle \ddot{\gamma}(t), X(t) \rangle = 0$$

if (3) holds. Conversely assume (3) fails, i.e. (by (3) $\Longleftrightarrow$ (4)) that $\nabla\gamma(t_0) \neq 0$ for some t_0 in the interval. We must find a variation γ_λ such that $(dE/d\lambda) \neq 0$. By continuity we may assume that $\nabla\gamma(t) \neq 0$ for t in an open interval J' containing t_0. Let $\rho(t)$ be a smooth real valued function which is non-negative, supported in J' but not identically vanishing. Define $X \in \mathcal{X}(\gamma)$ by:

$$X(t) = \rho(t)\nabla\dot{\gamma}(t)\,.$$

By construction X vanishes at the endpoints of J, is supported in J, and

$$\langle \ddot{\gamma}(t), X(t) \rangle \geq 0$$

with strict inequality on an open set. Hence

$$-\int \langle \ddot{\gamma}(t), X(t) \rangle dt < 0\,;$$

we must find a variation $\{\gamma_\lambda\}$ with $X = d\gamma_\lambda/d\lambda$. For this note that we may assume J' was chosen so small that $\gamma(t) \in M_0$ for $t \in J'$ where M_0 is the domain of some local co-ordinate system $(x^1, \ldots, x^m)$ on M. Let $E_1, \ldots, E_m \in \mathcal{X}(M_0)$ be the co-ordinate vectorfield

$$D_{E_i} f = \frac{\partial f}{\partial x^i}$$

for any function f on M_0 and resolve X into components

$$X(t) = \sum_i \xi^i(t) E_i(\gamma(t))$$

for $t \in J'$. Define $\gamma_\lambda(t)$ for λ near 0 by

$$x^i(\gamma_\lambda(t)) = x^i(\gamma(t)) + \lambda\xi^i(t)$$

for $t \in J'$ and by

$$\gamma_\lambda(t) = \gamma(t)$$

for $t \in J \backslash J'$. Then γ_λ is a variation of γ with fixed endpoints and $d\gamma_\lambda \mid d\lambda = X$ at $\lambda = 0$ as required.

(1) $\Longleftrightarrow$ (2). First note that we may assume that γ is parametrized proportional to arclength:

$$\|\dot{\gamma}(t)\| = c$$

where c is constant. Indeed this is explicitly part of (2) whereas if we assume (1), then by (1) $\iff$ (3) we have

$$\frac{d}{dt}\|\dot{\gamma}(t)\|^2 = 2\langle\ddot{\gamma}(t), \dot{\gamma}(t)\rangle = 0\,.$$

Next calculate

$$\begin{aligned}
\frac{d}{d\lambda}L(\gamma_\lambda)|_{\lambda=0} &= \int \frac{d}{d\lambda}\|\dot{\gamma}_\lambda(t)\||_{\lambda=0}dt \\
&= \int \|\dot{\gamma}(t)\|^{-1}\langle\dot{\gamma}(t), \dot{X}(t)\rangle dt \\
&= \frac{1}{2c}\int \frac{d}{d\lambda}\|\dot{\gamma}_\lambda(t)\|^2|_{\lambda=0}dt \\
&= \frac{1}{2c}\frac{d}{d\lambda}E(\gamma_\lambda)|_{\lambda=0}
\end{aligned}$$

so (1) $\iff$ (2) is clear.

(4) $\iff$ (5). Let (b, γ, γ') be a development of γ in a Euclidean space $M' = E^m$. By the definition of development the vectorfield $\dot{\gamma} \in \mathcal{X}(\gamma)$ is parallel if and only if the vectorfield $\dot{\gamma}' \in \mathcal{X}(\gamma')$ is. But in affine space ordinary differentiation and covariant differentiation coincide: the condition that $\dot{\gamma}'$ be parallel is simply that $\ddot{\gamma} \equiv 0$ which is clearly equivalent to (5).

Remark 72. Denote by $\Omega(J, p, q)$ the space of all smooth curves $\gamma : J \to M$ from p to q. The energy and arclength integrals can be viewed as real valued functions

$$E, L : \Omega(J, p, q) \to \mathbf{R}$$

while a variation $\{\gamma_\lambda\}$ with fixed endpoints can be viewed as a curve:

$$I \to \Omega(J, p, q) : \lambda \mapsto \gamma_\lambda$$

in $\Omega(J, p, q)$. Hence according to (1) in the definition we may say that $\gamma \in \Omega(J, p, q)$ a geodesic if and only if it is a critical point of E. In fact one can make $\Omega(J, p, q)$ (or rather a certain completion of it) into a smooth infinite dimensional manifold; E will be a smooth function on this manifold and the last assertion becomes literally true.

Now we noted above that reparametrization $\sigma : J' \to J$ yields a map

$$\Omega(J, p, q) \to \Omega(J', p, q) : \gamma \mapsto \gamma \circ \sigma$$

which intertwines L:

$$L(\gamma \circ \sigma) = L(\gamma)\,.$$

Thus if $\gamma \circ \sigma$ is an extremal of L (critical point) then (by the chain rule), γ should be a critical point of L. Moreover if $\sigma : J' \to J$ is a diffeomorphism, then the map $\gamma \mapsto \gamma \circ \sigma$ is bijective. Finally, if the tangent vector $\dot{\gamma}$ vanishes nowhere, then γ may be parametrized by arclength; i.e. one can find a diffeomorphism $\sigma : J' \to J$ or that the tangent vector to $\gamma \circ \sigma$ is a unit vector. (One defines $\sigma(t')$ by

$$\sigma(t') = t \Longleftrightarrow t' = \int^t \|\dot{\gamma}\| dt.)$$

These remarks suggest the following:

Exercise 73. Assume $\gamma : J \to M$ has nowhere vanishing tangent vector. Show that for any variation $\{\gamma_\lambda\}$ of γ with endpoints fixed we have

$$\frac{d}{d\lambda} L(\gamma_\lambda)|_{\lambda=0} = -\int \langle \dot{V}(t), X(t) \rangle dt$$

where V is the unit tangent vector to γ:

$$V(t) = \|\dot{\gamma}(t)\|^{-1} \dot{\gamma}(t)$$

and X is the infinitesimal variation

$$X(t) = \frac{d}{d\lambda} \gamma_\lambda(t)|_{\lambda=0} \,.$$

Conclude that γ is an extremal of the arclength:

$$\frac{d}{d\lambda} L(\gamma_\lambda)|_{\lambda=0} = 0$$

if and only if there is a geodesic $\gamma' : J \to M$ and a diffeomorphism $\sigma : J' \to J$ such that

$$\gamma' = \gamma \circ \sigma \,.$$

Below we shall generalize this exercise to cover the case where $\dot{\gamma}$ is allowed to vanish. For the moment note that $L(\gamma_\lambda)$ need note even be differentiable. (Suppose e.g. that γ is constant on some subinterval.) However we have the following

Exercise 74. Continue the notation of the last exercise but drop the hypothesis that $\dot{\gamma}$ vanishes nowhere. Show that the one-sided derivative of $L(\dot{\gamma}_\lambda)$ exists at $\lambda = 0$ and satisfies

$$-\int \|\dot{X}(t)\| dt \leq \frac{d}{d\lambda} L(\gamma_\lambda)|_{\lambda=0} \leq \int \|\dot{X}(t)\| dt \,.$$

Exercise 75. Show that the development of a geodesic is a geodesic; i.e. if (b, γ, γ') is a development of M along M' and γ is a geodesic in M then γ' is a geodesic in M'. Hint: When $M' = E^m$ this is the definition.

Distance

Two points p and q of the manifold M are of distance $\|p-q\|$ apart in the ambient Euclidean space E^n. In this section we define a distance function which is more intimately tied to M. We must assume that M is connected.

Definition 76. *The* **intrinsic distance** *between two points p and q of M is defined by*

$$d(p,q) = \inf L(\gamma)$$

where the infimum is over all curves γ from p to q. (As M is connected there exist such curves; if M is not connected one might define $d(p,q) = \infty$ if p and q lie in distinct connected components.)

Remark 77. It is natural to ask if the infimum is always attained. This is easily seen not to be the case in general. Let M result from a Euclidean space by removing a point o. The distance $d(p,q)$ is equal to length of the line segment from p to q and any other curve from p to q is longer. Hence if o is in the interior of this line segment the infimum is not attained. We shall prove below that the infimum *is* attained when M is complete.

Exercise 78. The function $d : M \times M \to \mathbf{R}$ is a **metric** on M; i.e. it satisfies the following conditions for $o, p, q \in M$:

(i) $d(p,q) \geq 0$;

(ii) $d(p,q) = 0 \iff p = q$;

(iii) $d(p,q) = d(q,p)$;

(iv) $d(o,q) \leq d(o,p) + d(p,q)$.

Definition 79. *For $p \in M$ and $r \geq 0$ denote by $B^r(p) = B^r(p,M)$, $\overline{B}^r(p) = \overline{B}^r(p,M)$, $S^r(p) = S^r(p,M)$ respectively the* **open ball, closed ball, and sphere** *or radius r centered at p:*

$$\begin{aligned} B^r(p) &= \{q \in M : d(p,q) < r\} \\ \overline{B}^r(p) &= \{q \in M : d(p,q) \leq r\} \\ S^r(p) &= \{q \in M : d(p,q) = r\}. \end{aligned}$$

The following two propositions assert that (locally) the metric d is roughly Euclidean.

Proposition 80. *Given $o \in M$ we have:*

$$\lim_{p,q \to 0} \frac{d(p,q)}{\|p-q\|} = 1\,.$$

Proposition 81. *Given $o \in M$ and local co-ordinates $x = (x^1, \ldots, x^m)$ defined near o and such that the linear map:*

$$Dx(o) : T_0 M \to \mathbf{R}^m$$

is orthogonal (this may always be accomplished by a further linear change of variables) we have

$$\lim_{p,q\to 0} \frac{d(p,q)}{\|x(p) - x(q)\|} = 1\,.$$

Remark 82. The propositions imply that the topology M inherits as a subset of E^n, the topology on M determined by the metric d, and the topology on M induced by the local co-ordinate systems of M are all the same; i.e. given a sequence $\{p_n\}_n \subset M$ and a point $p \in M$ the following three equations are equivalent:

$$\begin{aligned} \lim_{n\to\infty} \|p_n - p\| &= 0 \\ \lim_{n\to\infty} d(p_n, p) &= 0 \\ \lim_{n\to\infty} \|x(p_n) - x(p)\| &= 0 \end{aligned}$$

(where x is any local co-ordinate system defined near p).

Proof of 80. The triangle inequality:

$$\| \int \dot\gamma dt\| \leq \epsilon \|\dot\gamma\| dt$$

together with the fundamental theorem of calculus gives the obvious inequality

$$\|q - p\| \leq d(q,p)\,.$$

(A straight line is the shortest distance between two points.) The proposition asserts another inequality as well; viz. given $\epsilon > 0$ we have

$$\|q - p\| \leq d(q,p) \leq (1+\epsilon)\|q - p\|$$

for $q, p \in M$ sufficiently near o. Let $\Pi(o)$ and $\Pi^\perp(o)$ be the orthogonal projections on $T_0 M$ and $T_0^\perp M$ respectively and define $c : E^n \to T_0 M$ and $y : E^n \to T_0 M$ by:

$$x(p) = \Pi(o)(p - o), \quad y(p) = \Pi^\perp(o)(p - o).$$

By the implicit function theorem there is a smooth map $f : T_o M \to T_o^\perp M$ whose graph near o is M:

$$p \in M \iff y(p) = f(x(p))$$

for $p \in E^n$ sufficiently near o. Note that $f(0) = 0$ (as $o \in M$) and

$$Df(0) = 0$$

Figure 5: Locally, M is the graph of f.

as the x-axis is the tangent space to M at o. Hence for $v \in T_o M$ and $x = x(p)$ sufficiently near 0 we have:

$$\|df(x)v\| \leq \epsilon \|v\| \, .$$

Given $p, q \in M$ let γ be the curve in M from p to q whose projection on the x-axis is a straight line

$$\begin{aligned} x(\gamma(t)) &= x(p) + t(x(q) - x(p)) = x(t) \\ y(\gamma(t)) &= f(x(\gamma(t)) = y(t). \end{aligned}$$

Then

$$\begin{aligned} L(\gamma) &= \int_0^1 \|\dot{x}(t) + \dot{y}(t)\| dt \\ &= \int_0^1 \|\dot{x}(t) + Df(x(t))\dot{x}(t)\| dt \\ &\leq (1+\epsilon)\|x(q) - x(p)\| \\ &= (1+\epsilon)\|\Pi(o)(q-p)\| \\ &\leq (1+\epsilon)\|q-p\| \end{aligned}$$

so $d(p, q) \leq (1+\epsilon)\|q - p\|$ as required. $\square$

Figure 6: A non-convex neighborhood of o.

Proof of 81. By assumption we have

$$\|Dx(o)v\| = \|v\|$$

for $v \in T_o M$; hence (by continuity) given $\epsilon > 0$ there is a neighborhood M_0 of o such that for $p \in M_0$ and $v \in T_p M$ we have

$$(1-\epsilon)\|Dx(p)v\| \leq \|v\| \leq (1+\epsilon)\|Dx(p)v\| \,.$$

Thus for any curve γ lying wholly in M_0 we have

$$(1-\epsilon)L(x \circ \gamma) \leq L(\gamma) \leq (1+\epsilon)L(x \circ \gamma) \,.$$

We are tempted to take the infumum over all γ in M_0 from $p \in M_0$ to $q \in M_0$ to obtain:

$$(1-\epsilon)\|x(q) - x(p)\| \leq d(q,p) \leq (1+\epsilon)\|x(q) - x(p)\| \,.$$

We must justify this however; the infimum over γ in M_0 need not be the same as the infimum over γ in M.

It suffices to show that the inequalities hold on a smaller neighborhood M_1 of o. If $x(M_1)$ is convex then the right-hand inequality obtains, for the curve γ from p to q such that $x \circ \gamma$ is a straight line will lie in M_0. For the left inequality, find $r > 0$ and M_1 such that

$$o \in M_1 \subset B^r(o) \subset B^{3r}(o) \subset M_0 \,.$$

Then for $p, q \in M_1$ we have $d(p, q) \leq 2r$ while $4r \leq L(\gamma)$ for any curve γ from p to q which leaves M_0. Hence in calculating $d(p, q)$ we may take the infimum over curves which lie in M_0; this (together with the inequality $\|x(q) - x(p)\| \leq L(x \cdot \gamma)$), establishes the left inequality. $\square$

The Geodesic Spray

We now describe the differential equation governing geodesics.

Definition 83. *The* **tangent bundle** TM *of* M *is the submanifold of* $E^n \times \mathbf{E}^n$ *defined by*

$$TM = \{(p, v) : p \in M, v \in T_pM\}.$$

(To see that TM *is indeed a submanifold of* $E^n \times \mathbf{E}^n$ *choose a moving frame* $E_1, \ldots, E_n \in \mathcal{X}(M_0)$ *defined over an open subset* M_0 *of* M*; the map*

$$M_0 \times \mathbf{R}^m \to TM_0 \subset TM : (p, \xi) \mapsto (p, \Sigma\xi^i E_i(p))$$

gives (when composed with a suitable parametrization of M_0*) a local parametrization.)*

Definition 84. *Let* $S = (S_1, S_2) : TM \to \mathbf{E}^n \times \mathbf{E}^n$ *be a vectorfield on* TM*; i.e.* $S \in \mathcal{X}(TM)$*. Call* S *a* **second order differential equation** *iff*

$$S_1(p, v) = v$$

for $(p, v) \in TM$*. This means that every integral curve* $\beta : \mathbf{R} \to TM$ *of* S *(i.e.* $\dot{\beta} = S(\beta)$*) is of form* $\beta = (\gamma, \dot{\gamma})$ *for some curve* $\gamma : \mathbf{R} \to M$ *is a base integral curve iff it satisfies the differential equation*

$$\ddot{\gamma} = S_2(\gamma, \dot{\gamma})$$

(which accounts for the terminology). Call S *a* **spray** *iff, in addition,* S_2 *is homogeneous of degree 2 in* v*:*

$$S_2(p, \lambda v) = \lambda^2 S_2(p, v)$$

for $(p, v) \in TM$ *and* $\lambda \in \mathbf{R}$*. This means that an affine reparametrization* $t' \mapsto \gamma(\lambda t' + t_0)$ *of a base integral curve* $t \mapsto \gamma(t)$ *is again a base integral curve.*

Definition 85. *There is a (necessarily unique) spray* S *on* TM *whose base integral curves are the geodesics of* M*; it is called the* **geodesic spray** *of* M*. It is given by* $S = (S_1, S_2)$ *where:*

$$S_1(p, v) = v, \quad S_2(p, v) = h_p(v)v$$

for $(p, v) \in TM$ *where* h *is the second fundamental form.*

To justify the definition note that by the Gauss-Weingarten equations a curve γ in M is a geodesic if and only if it satisfies the ordinary differential equation

$$\ddot{\gamma}(t) = h_{\gamma(t)}(\dot{\gamma}(t))\dot{\gamma}(t)\,.$$

Since (by the development theorem for example) we can always find a geodesic γ satisfying the initial conditions:

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

for any $(p, v) \in TM$ it follows that

$$S(p, v) = (\dot{\gamma}(0), \ddot{(0)}) \in T_{(p,v)}(TM)$$

i.e. that $S \in \mathcal{S}(TM)$. It is immediate that S is a spray as required.

We recall that the geodesic equation $\nabla\dot{\gamma} = 0$ takes the form

$$\dot{\xi}^k(t) + \sum_{i,j} \Gamma^k_{ij}(\gamma(t))\xi^i(t)\xi^j(t) = 0$$

where

$$\dot{\gamma}(t) = \sum_k \xi^k(t)E_k(\gamma(t))\,.$$

If the moving frame $E_1, \ldots, E_m$ consists of the co-ordinate vector fields for some co-ordinate system $x^1, \ldots, x^m$; i.e.

$$D_{E_i} f = \frac{\partial f}{\partial x^i}$$

for any function f, then

$$\xi^i(t) = \dot{x}^i(t)$$

where we have abbreviated

$$x^i(t) = x^i(\gamma(t))\,.$$

We shall thus write the geodesic equation in the abbreviated form:

$$\ddot{x} + \Gamma(x)\dot{x}^2 = 0\,.$$

Remark 86. In fact any spray has this form in local co-ordinates. The second order differential equation $\ddot{x} + \Gamma(x, \dot{x}) = 0$ is a spray if and only if:

$$\Gamma(x, \lambda\dot{x}) = \lambda^2\Gamma(x, \dot{x})\,.$$

Apply $\left(\frac{d}{d\lambda}\right)^2|_{\lambda=0}$ to this last equation to obtain the desired form for $\Gamma(x, \dot{x})$.

Definition 87. *There is an open set N in TM and a map:*

$$N \to M : (p, v) \mapsto \mathrm{Exp}_p(v)$$

called the **exponential map** *such that any geodesic $\gamma : J \to M$ has form:*

$$\gamma(t) = \mathrm{Exp}_p(tv)$$

where

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

(The existence of the exponential map follows easily from the existence and uniqueness theorem for ordinary differential equations together with the fact that for $\lambda \in \mathbf{R}$ the curve $t \to \gamma(\lambda t)$ is a geodesic (with tangent vector $\lambda\dot{\gamma}(0)$ at $t = 0$) whenever γ is a geodesic. Hence $\gamma(1)$ is defined if $\dot{\gamma}(0)$ is sufficiently small so we may define $\mathrm{Exp}_p(v)$ by taking $t = 1$ above.)

Remark 88. By definition the map

$$\mathrm{Exp}_p : N_p = N \cap T_pM \to M$$

satisfies

$$\mathrm{Exp}_p(0) = p. \quad D\mathrm{Exp}_p(0)v = v$$

for $p \in M$, $v \in T_pM$. Hence for sufficiently small neighborhoods V_p of 0 in T_pM the map $\mathrm{Exp}_p \mid V_p$ maps V_p diffeomorphically onto an open neighborhood of p in M (Inverse Function Theorem). This map can thus be used to introduce local co-ordinates in M (compose with linear co-ordinates on T_pM). The result co-ordinate system has the property that straight lines represent geodesics. One calls such co-ordinates **geodesic normal** at p.

Convexity

A subset of an affine space is called **convex** iff it contains the line segment joiningany two of its points. The definition carries over to a submanifold M of Euclidean space (or indeed more generally to any manifold M equipped with a spray) once we reword the definition so as to confront the difficulty that a geodesic joining two points might not exist nor, if it does, need it be unique.

Definition 89. *A subset M_0 of M is* **convex** *iff for any $p, q \in M_0$ there exists a unique (up to an affine reparametrization) geodesic from p to q which lies wholly in M_0. (It is not precluded that there be other geodesics from p to q which leave and then re-enter M_0.)*

Theorem 90. *Each point $o \in M$ has arbitrarily small convex neighborhoods. In fact, if $x^1, \ldots, x^m$ are local co-ordinates defined on a neighborhood M_0 of o in M and if U_r is defined by:*

$$U_r = \{p \in M_0 : \delta(p) < r^2\}$$

where

$$\delta(p) = \sum_k (x^k(p) - x^k(o))^2 \,,$$

then U_r *is convex for suffciently small* $r > 0$.

Proof. For $p \in M_0$ let Q_p be the quadratic form defined by:

$$Q_p(v) = \sum_k (v^k)^2 - \sum_{i,j,k} (x^k(p) - x^k(o)) \Gamma^k_{ij}(p) v^i v^j$$

where Γ^k_{ij} are the christofel symbols and where $v^1, \ldots, v^m$ are the components of $v \in T_pM$ with respect to the basis given by the co-ordinate vector fields evaluated at p. For $p = o$ we have

$$Q_p(v) = \sum_k (v^k)^2$$

so by shrinking M_0 if necessary we may assume that Q_p is positive definite for $p \in M_0$.

If $\gamma : J \to M_0$ is any non-constant geodesic

$$\frac{d}{dt^2} \delta(\gamma(t)) = 2 Q_{\gamma(t)}(\dot\gamma(t)) > 0$$

so that $t \to \delta(\gamma(t))$ is a convex real-valued function; hence

$$\gamma(t) \in U_r$$

for $t \in J$ where $r = \max(\delta(p), \delta(q))$, p and q being the endpoints of γ.

We show that for sufficiently small $r > 0$ and $p, q \in U_r$ there is at *least* one geodesic from p to q lying in U_r. Indeed since $p = q = o$, $v = 0$ is a solution of

$$\mathrm{Exp}_p(v) = q$$

and since $D\,\mathrm{Exp}_o(0) = \text{identity}$ the implicit function theorem yields a smooth solution $v = h(p, q)$ defined for p, q near o and satisfying $h(o, o) = 0$. Since

$$\mathrm{Exp}_o(\mathrm{t}h(o, o)) = 0$$

continuity gives

$$\mathrm{Exp}_p(\mathrm{t}h(p, q)) \in M_0$$

for $0 \le t \le 1$ and $p, q \in U_r$ with $r > 0$ sufficiently small. But as we have seen above this (if r is small enough) implies

$$\mathrm{Exp}_p(\mathrm{t}h(p, q)) \in U_r$$

for $0 \le t \le 1$ as required.

We show that for sufficiently small $r > 0$ and $p, q \in U_r$ there is at *most* one geodesic from p to q lying in U_r. By the implicit function theorem choose $\epsilon, r > 0$ such that for $p, q \in U_r$ and $\|v\| < \epsilon$ we have

(i) $\|h(p, q)\| < \epsilon$, and

(ii) $\mathrm{Exp}_p(v) = q \Longleftrightarrow v = h(p, q)$.

Next choose $p, q \in U_r$ and let $\mathrm{Exp}_p(tv)$ $(0 \le t \le 1)$ be any geodesic from p to q which remains in U_r. We must show that $\|v\| < \epsilon$ for then $v = h(p, q)$ by (ii) which establishes the desired uniqueness. Suppose the contrary, i.e. that $\|v\| \ge \epsilon$. Then by (ii) we have

$$h(p, \mathrm{Exp}_p(tv)) = tv$$

and hence

$$\|h(p, \mathrm{Exp}_p(tv)\| = t\|v\|$$

and hence

$$\|h(p, \mathrm{Exp}_p(tv)\| = t\|v\|$$

for $0 \le t < \bar{t} = \epsilon/\|v\|$. In the last equation let t approach $\bar{t}$ to obtain

$$\|h(p, \mathrm{Exp}_p(\bar{t}v))\| = \epsilon$$

which contradicts (i). This completes the proof that U_r is convex for sufficiently small $r > 0$. □

Remark 91. These arguments work for *any* spray.

Minimal Geodesics

A straight line is the shortest distance between two points in Euclidean space. The analogous assertion for geodesics in a manifold M is false; consider e.g. an arc which is more than half of a great circle on a sphere. In this section we consider curves which realize the shortest distance between their endpoints.

Definition 92. *A curve* $\gamma : J \to M$ *from* p *to* q *is called a* **minimal geodesic** *iff it satisfies the two following equivalent condtions:*

(1) it is parametrized proportional to arclength

$$\|\dot{\gamma}(t)\| = \text{ constant}$$

for $t \in J$*, and it minimizes arclength integral:*

$$L(\gamma) \leq L(\gamma')$$

for all curves γ' *from* p *to* q*;*

(2) it minimizes the energy integral:

$$E(\gamma) \leq E(\gamma')$$

for all curves $\gamma' : J \to M$ *from* p *to* q*.*

Remark 93. Condition (1) says that ($\|\dot{\gamma}\|$ is constant and) $L(\gamma) = d(p, q)$; i.e. that γ is a shortest curve from p to q. It is not precluded that there be more than one such γ; consider for example the case where M is a sphere and p and q are antipodal.

Condition (2) implies that:

$$\frac{d}{d\lambda} E(\gamma_\lambda)\,|_{\lambda=0} \; = 0$$

for every variation $\{\gamma_\lambda\}$ of γ with endpoints fixed. Hence a minimal geodesic is a geodesic.

Finally, $L(\gamma')$ (but *not* $E(\gamma')$) is independent of the parametrization of γ'. Hence if γ is a minimal geodesic $L(\gamma) \leq L(\gamma')$ for *every* γ' (from p to q) whereas $E(\gamma) \leq E(\gamma')$ for those γ' defined on (an interval the same length as) J.

Proof of the equivalence of the conditions. We prove (1) $\iff$ (2). Let c be the (constant) value of $\|\dot{\gamma}(t)\|$ and $|J|$ the length of the interval J. Thus

$$L(\gamma) = c|J|, \quad E(\gamma) = c^2|J|.$$

Then for $\gamma' : J \to M$ from p to q we have:

$$\begin{aligned} E(\gamma)^2 &= c^2 L(\gamma)^2 \\ &\leq c^2 L(\gamma')^2 \\ &\leq c^2 E(\gamma')|J| \\ &= E(\gamma')E(\gamma) \end{aligned}$$

so (dividing by $E(\gamma)$) $e(\gamma) \leq E(\gamma')$ as required. (The inequality $L(\gamma')^2 \leq E(\gamma')|J|$ used above is the Schwartz inequality:

$$\left(\int fg dt\right)^2 \leq \left(\int f^2 dt\right)\left(\int g^2 dt\right)$$

with $f(t) = \|\dot\gamma'(t)\|$ and $g(t) = 1$.)

We prove (2) $\iff$ (1). We have already shown that γ is a geodesic. It is easy to dispose of the case where M is one-dimensional: Any γ minimizing $E(\gamma)$ or $L(\gamma)$ must be monotonic onto a subarc; otherwise it could be altered so as to make the integral smaller. Hence suppose M is of dimension at least two. Suppose $L(\gamma') < L(\gamma)$ for some curve γ' from p to q; we will derive a contradiction. Since the dimension of M is not one we may approximate γ' by a curve whose tangent vector nowhere vanishes; i.e. we may assume w.l.o.g. that $\dot\gamma'(t) \neq 0$ for all t. Hence we may reparametrize γ' proportional to arclength and assume w.l.o.g. that $\gamma' : J \to M$ and $\|\dot\gamma'(t)\| = c'$ for $t \in J$. But then

$$L(\gamma) = c|J|; \quad L(\gamma') = c'|J|;$$

$$E(\gamma) = c^2|J|; \quad E(\gamma') = {c'}^2 J|;$$

so that $L(\gamma') < L(\gamma)$ implies $E(\gamma') < E(\gamma)$ which is the desired contradiction. □

Our next theorem asserts the existence of minimal geodesics. For $p \in M$ and $r > 0$ recall the definitions:

$$\begin{aligned} B^r(p) &= \{q \in M : d(p,q) < r\} \\ S^r(p) &= \{q \in M : d(p,q) = r\},; \end{aligned}$$

define ball and sphere in the tangent space by:

$$\begin{aligned} V^r(p) &= \{v \in T_pM : \|v\| < r\} \\ \Sigma^r(p) &= \{v \in T_pM : \|v\| = r\}. \end{aligned}$$

Theorem 94. *Suppose $p \in M$ and $r > 0$ is sufficiently small that Exp_p maps $V^r(p)$ diffeomorphically onto a neighborhood of p in M. Let $v \in V^r(p)$, $q = \mathrm{Exp}_p(v) \in M$, and $\gamma : J \to M$ be a curve from p to q. Then $L(\gamma) = d(p,q)$ if and only if there exists a monotonic surjective map $\sigma : J \to [0,1]$ such that*

$$\gamma(t) = \mathrm{Exp}_p(\sigma(t)v)$$

for $t \in J$. Hence

$$\mathrm{Exp}_p(V^r(p)) = B^r(p).$$

Proof. First note that any curve of the indicated form has length $L(\gamma) = \|v\|$; indeed such a curve is a reparametrization of the geodesic:

$$[0,1] \to M : s \mapsto \mathrm{Exp}_p(sv)$$

and this geodesic has length $\|v\|$ as its initial tangent vector is V. Thus $d(p,q) \leq \|v\| < r$. We will show that any curve *not* of the desired form has length $L(\gamma) > \|v\|$ if it remains in $\mathrm{Exp}_p(V^r(p))$ and length $L(\gamma) \geq r$ if it exists. This will show that $\|v\| = d(p,q)$ and prove the theorem.

Suppose for convenience that $J = [0,1]$. Thus $\gamma(0) = p$ and $\gamma(1) = q$. We may assume that $\gamma(t) \neq p$ for $t \in (0,1]$ for otherwise we could restrict γ to $[t_0, 1]$ where $t_0 = \sup\{t \in J : \gamma(t) = p\}$. Assume for the moment that $\gamma(t)$ remains in $\mathrm{Exp}_p(V^r(p))$; then we may introduce "geodesic polar co-ordinates"; i.e. write γ in the form:

$$\gamma(t) = \mathrm{Exp}_p(\sigma(t)\mathcal{O}(t))$$

for $t \in (0,1]$ where $\sigma(t) \in (0, r/\|v\|]$ and $\mathcal{O}(t) \in \Sigma^{\|v\|}(p)$ (thus $\sigma(1) = 1$ and $\mathcal{O}(1) = v$). Introduce the auxiliary map:

$$\alpha(s,t) = \mathrm{Exp}_p(s\mathcal{O}(t))$$

for $s \in (0, r/\|v\|]$ and $t \in (0,1]$; thus

$$\gamma(t) = \alpha(\sigma(t), t)\,.$$

Claim . $\langle \partial_s\alpha, \partial_t\alpha \rangle = 0$.

Indeed:

$$\begin{aligned}
\partial_s\langle \partial_s\alpha, \partial_t\alpha \rangle &= \langle \partial_s^2\alpha, \partial_t\alpha \rangle + \langle \partial_s\alpha, \partial_s\partial_t\alpha \rangle \\
&= \langle \nabla_s\partial_s\alpha, \partial_t\alpha \rangle + \frac{1}{2}\partial_t\|\partial_s\alpha\|^2 \\
&= 0 + 0
\end{aligned}$$

since $s \mapsto \alpha(s,t)$ is a geodesic with initial tangent vector of length $\|\mathcal{O}(t)\| = \|v\|$. Thus $\langle \partial_s\alpha, \partial_t\alpha \rangle$ is constant in s. But for $s = 0$ we have

$$\alpha(0,t) = \mathrm{Exp}_p(0) = p$$

hence $\partial_t\alpha = 0$ at $s = 0$ so $\langle \partial_s\alpha, \partial_t\alpha \rangle = 0$ at $s = 0$ so $\langle \partial_s\alpha, \partial_t\alpha \rangle = 0$ identically.

Now using the claim we write

$$\dot\gamma = \dot\sigma(\partial_s\alpha) + (\partial_t\alpha)$$

where the vectors on the right are orthogonal. Hence

$$\|\dot\gamma\|^2 = \dot\sigma^2\|v\|^2 + \|\partial_t\alpha\|^2$$

so

$$\begin{aligned}
L(\gamma) &= \int_0^1 \|\dot\gamma(t)\| dt \\
&\geq \int_0^1 |\dot\sigma(t)| dt \|v\| \\
&\geq \int_0^1 \dot\sigma(t) dt \|v\| \\
&= \|v\|
\end{aligned}$$

with strict inequality unless $\mathcal{O}(t)$ is constant (so $\partial_t\alpha = 0$) and σ is monotonic (so $|\dot\sigma| = \dot\sigma$). To remove the assumption that γ remains in $\mathrm{Exp}_p(V^r(p))$ note that if it exits, say at time t_1, then the above argument shows that $L(\gamma) \geq L(\gamma \mid [0, t_1]) \geq r$ as required. □

Figure 7: Orthogonals to Geaodesics

Remark 95. We can now conclude that

$$S^s(p) = \operatorname{Exp}_p(\Sigma^s(p))$$

for $0 \le s \le r$. The proof of the claim above shows that the geodesics

$$[0,1] \to M : s \mapsto \operatorname{Exp}_p(sw)$$

emanating from p (with $w \in \Sigma^1(p)$) are the orthogonal trajectories to the concentric spheres $S^s(p)$.

Exercise 96. Let $M \subset E^3$ be of dimension two and suppose that (orthogonal) reflection in some plane E^2 preserves M. Show that E^2 intersects M in a geodesic. (Hint: Otherwise there would be points $p, q \in M$ very close to one another joined by two distinct minimal geodesics.) Conclude for example that the co-ordinate planes intersect the ellipsoid $(x/a)^2 + (y/b)^2 + (z/c)^2 = 1$ in geodesics.

Exercise 97. Introduce geodesic normal co-ordinates near p via

$$q = \operatorname{Exp}_p\left(\sum_i x^i(q) e_i\right)$$

where $e_1, \ldots, e_m$ is an orthonormal basis for T_pM. Then $x^i(p) = 0$ and

$$B^r(p) = \{q \in M : \Sigma(x^i(q) - x^i(p))^2 < r^2\}.$$

Hence by the convexity theorem $B^r(p)$ is convex for r sufficiently small.

(1) Show that it can happen that a geodesic in $B^r(p)$ is not minimal. (Hint: Take M to be essentially the hemisphere $x^2+y^2+z^2=1$, $z>0$ together with the disk $x^2+y^2<1$, $z=0$ (but smooth the corner $x^2+y^2=1$, $z=0$). Let $p=(0,0,1)$ and $r=\pi/2$.)

(2) Show that if r is sufficiently small then the unique geodesic γ in $B^r(p)$ joining two points $q,q'\in B^r(p)$ *is* minimal and that in fact *any* curve γ' from q to q' which is not a reparametrization of γ is strictly longer

$$L(\gamma') > L(\gamma) = d(q,q')\,.$$

Exercise 98. Let $\gamma : J \to M$ be a curve from p to q with nowhere vanishing tangent vectors

$$\dot\gamma(t) \neq 0$$

for $t\in J$. Then the following three conditions are equivalent:

(1) The curve γ is an **extremal of the arclength integral** L:

$$\frac{d}{d\lambda}L(\gamma_\lambda)\,|_{\lambda=0}\; = 0$$

for every variation $\{\gamma_\lambda\}$ of γ with fixed endpoints.

(2) The curve γ is a **reparametrized geodesic**; i.e. there exists a smooth surjective $\sigma : J \to [0,1]$ with $\dot\sigma \geq 0$ and $v\in T_pM$ such that $q = \operatorname{Exp}_p(v)$ and

$$\gamma(t) = \operatorname{Exp}_p(\sigma(t)v)\,.$$

(3) The curve γ **minimizes arclength locally**: i.e. there exists $\epsilon>0$ such that for every closed subinterval $[t_1,t_2]\subset J$ of length less than ϵ we have

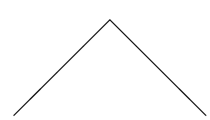
$$L(\gamma \mid [t_1,t_2]) = d(\gamma(t_1),\gamma(t_2))\,.$$

We remark that the hypothesis $\dot\gamma(t)\neq 0$ implies that the reparametrization σ in (2) is in fact a diffeomorphism:

$$\dot\sigma(t) \neq 0$$

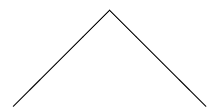
for $t\in J$.

It is often convenient to consider curves γ where $\dot\gamma(t)$ is allowed to vanish for some values of t; then γ cannot (in general) be parametrized by arclength. Such a curve $\gamma : J\to M$ can be smooth (as a map) yet its image may have corners (where $\dot\gamma$ necessarily



A smooth curve

Note that a curve with corners can never minimize distance even locally.



Exercise 99. Show that conditions (2) and (3) of the last exercise are equivalent even without the assumption that $\dot\gamma$ is nowhere vanishing. Conclude that if $\gamma : J \to M$ is a shortest curve from p to q:

$$L(\gamma) = d(p,q)$$

then γ is a reparametrized geodesic.

Show by example that one can have a variation $\{\gamma_\lambda\}$ with fixed points of a reparametrized geodesic for which the map $\lambda \to L(\gamma_\lambda)$ is not even differentiable at $\lambda = 0$. (Hint: Take γ to be constant.) Show that however that conditions (1), (2) and (3) of the previous exercise remain equivalent if the hypothesis that $\dot\gamma$ is nowhere vanishing is weakened to the hypothesis that $\dot\gamma(t) \neq 0$ for all but finitely many $t \in J$. Conclude that a broken geodesic is a reparametrized geodesic if and only if it minimizes arclength locally. (A broken geodesic is a continuous map $\gamma : J \to M$ for which there exist $t_0 < t_1 < \cdots < t_n$ with $J = [t_0, t_n]$ and $\gamma \mid [t_{i-1}, t_i]$ a geodesic for $i = 1, \ldots, n$. It is thus a geodesic if and only if $\dot\gamma$ is continuous at the break points:

$$\dot\gamma(t_{i^-}) = \dot\gamma(t_{i^+})$$

for $i = 1, \ldots, n-1$.)

Completeness

Intuitively a manifold is complete if there are no points missing. Note that the second part of the following definition is word for word the same as the definition of completeness given on page 34.

Definition 100. *The manifold $M \subset E^n$ is called complete iff it satisfies the following equivalent conditions:*

(1) the geodesic spray is complete as a vectorfield on TM: i.e. for every $(p,v) \in TM$ there exists a geodesic $\gamma : \mathbf{R} \to M$ (defined for all time t) such that $\dot\gamma(0) = p$, $\dot\gamma(0) = v$;

(2) for every curve $\gamma' : \mathbf{R} \to E^m$ in a Euclidean space, every $t_0 \in \mathbf{R}$, every $o \in M$, and every orthogonal isomorphism $b_0 : T_0M \to \mathbf{E}^m$, the maximally defined development (b, γ, γ') with initial conditions $\gamma(t_0) = o$ and $b(t_0) = b$ is defined for all $t \in \mathbf{R}$;

(3) *the metric d is complete: i.e. given points $p_k \in M$ with:*

$$\lim_{j,k\to\infty} d(p_j, p_k) = 0$$

there exists $p \in M$ with:

$$\lim_{k\to\infty} d(p, p_k) = 0\,;$$

(4) *every closed and bounded (with respect to the metric d) subset of M is compact.*

We shall prove the equivalence of these conditions and at the same time prove the following:

Theorem 101 (Hopf-Rinow)**.** *If M is complete and $p, q \in M$ then there exists a minimal geodesic γ in M from p to q.*

The pattern of proof of the equivalence of conditions (1) - (4) in the definition of completeness is

$$(4) \Longrightarrow (3) \Longrightarrow (2) \Longrightarrow (1) \Longrightarrow (4).$$

We shall need the Hopf-Rinow theorem (where (1) is used as the definition of completeness).

Proof of (4) $\Longrightarrow$ (3). Assume $\lim d(p_j, p_k) = 0$. Then the point set $\{p_k : k = 1, 2, \ldots\}$ is clearly bounded so by (4) its closure is compact. Hence the sequence $\{p_k\}$ contains a convergent subsequence. As $\lim d(p_j, p_k) = 0$ the whole sequence must converge to the limit of this subsequence. □

Proof of (3) $\Longrightarrow$ (2). We recall that, according to the proof of the development theorem, the development (b, γ, γ') is essentially a solution of a certain differential equation on the orthonormal frame bundle $O(M)$. Hence, by the continuation theorem in the theory of ordinary differential equations, it is enough to show that if $b(t)$ and $\gamma(t)$ are defined for $t < t_+ < \infty$ then for some sequence t_n increasing to t_+ the corresponding $b(t_n)$ and $\gamma(t_n)$ converge; it then follows that the solution can be defined for $t < t_+ + \epsilon$ where ϵ is some small positive number.

For this note that the equation:

$$b(t)\gamma(t) = \gamma'(t)$$

which is part of the definition of development implies that for any subinterval $[t, t']$ we have:

$$d(\gamma(t), \gamma(t')) = L(\gamma \mid [t, t']) = L(\gamma' \mid [t, t'])$$

so that:

$$\lim_{t,t'\uparrow t_+} d(\gamma(t), \gamma(t')) = 0\,.$$

Hence by (3) there exists $p_+ \in M$ with:

$$\lim_{t \uparrow t_+} \gamma(t) = p_+ \,.$$

Now choose an orthonormal moving frame $E_1, \ldots, E_m \in \mathcal{X}(M_0)$ defined in a neighborhood M_0 of p_+. Using this frame (and a fixed orthonormal frame $e'_1, \ldots, e'_m$ for $\mathbf{E}^m$) we may identify the orthogonal isomorphism $b(t) : T_{\gamma(t)}M \to \mathbf{E}^m$ with an orthogonal matrix. But the space $O(m)$ of orthogonal matrices is compact; hence $b(t_n)$ converges for a suitable sequence $t_n \uparrow t_+$ as required. □

Proof of (2) $\Longrightarrow$ (1). Given $(p, v) \in TM$ choose any orthogonal isomorphism $b_0 : T_pM \to \mathbf{E}^m$ where $\mathbf{E}^m$ is the vector space of some Euclidean space E^m. Then choose any straight line $\gamma' : \mathbf{R} \to E^m$ with tangent vector $b_0(v) \in \mathbf{E}^m$:

$$\gamma'(t) = \gamma'(0) + tb_0(v) \,.$$

By (2) there is a development (b, γ, γ') with $\gamma(0) = p$, $b(0) = b_0$; then $b_0\dot{\gamma}(0) = \dot{\gamma}'(0) = b_0(v)$ so γ is a geodesic (its development is a straight line) with the desired initial conditions. □

Remark 102. In the language used in the proof of the development theorem the implication (2) $\Longrightarrow$ (1) is even more obvious: If the basic vectorfield B corresponding to $\xi : \mathbf{R} \to \mathbf{R}^m$ is always complete, it is in particular complete when ξ is constant.

Proof of (1) $\Longrightarrow$ (4). Let $K \subset M$ be closed and bounded. Fix $p \in M$. Then for r sufficiently large:

$$K \subset \overline{B}^r(p) \,.$$

By the theorem (proved next):

$$\overline{B}^r(p) \subset \operatorname{Exp}_p(\overline{V}^r(p))$$

where $\overline{V}^r(p)$ is the closed ball of radius r in T_pM. Now $\overline{V}^r(p)$ is certainly compact and hence so is its image by the continuous map Exp_p. Thus K is a closed subset of a compact set and hence is itself compact. □

Proof of Theorem. We shall assume that every geodesic leaving p is defined for all time i.e. that Exp_p is defined on all of T_pM. Set:

$$r = d(p, q) \,.$$

We must find $v \in T_pM$ with $\|v\| = 1$ and

$$\operatorname{Exp}_p(rv) = q \,.$$

Lemma 103. *Assume that $s > 0$ is sufficiently small that the map*

$$\exp_p : \Sigma^s(p) \to S^s(p)$$

(from the sphere of tangent vectors of length s to the sphere of radius s in the metric d) is a diffeomorphism. Assume also that $d(p, q) > s$ and let $z \in S^s(p)$ be any point at which the function:

$$S^s(p) \to \mathbf{R} : z \mapsto d(z, q)$$

achieves its minimum. Then

$$d(p, q) = d(p, z) + d(z, q) \, .$$

The lemma is obvious since any curve from p to q must pass through $S^s(p)$.

By lemma 103 choose a unit vector $v \in T_pM$ so that $d(\mathrm{Exp}_p(sv), q)$ is minimized (by compactness). Let $\gamma : [0, r] \to M$ be the geodesic given by:

$$\gamma(t) = \mathrm{Exp}_p(tv)$$

for $t \in [0, r]$. We must prove that $\gamma(r) = q$. We shall in fact prove a stronger statement:

$$d(\gamma(t), q) = r - t$$

for $t \in [s, r]$. For these we need the following obvious:

Lemma 104. *If $p', q' \in M$ satisfy:*

$$d(p, p') + d(p', q') + d(q', q) = d(p, q)$$

then:

$$d(p, p') + d(p', q') = d(p, q') \, .$$

Now set:

$$I = \{t \in [s, r] : d(\gamma(t), q) = r - t\} \, ;$$

our objective is to prove that $I = [s, r]$; i.e. that I is non-empty, closed, and open in $[s, r]$.

Now I is non-empty as $s \in I$ by construction and lemma 103. By continuity I is closed. Hence we need only prove $t' + s' \in I$ for sufficiently small s' under the assumption that $[s, t'] \subset I$.

Read $\gamma(t')$ for p', s' for s in lemma 103. We obtain a point $q' \in S^{s'}(p')$ with:

$$d(p', q') + d(q', q) = d(p', q) \, .$$

But:

$$d(p', q) = d(\gamma(t'), q) = r - t'$$

as $t' \in I$ and:

$$d(p, p') = d(p, \gamma(t')) = t'$$

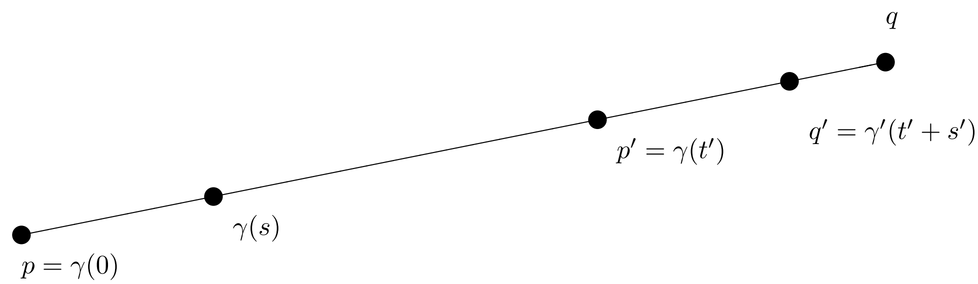


Figure 8: The Hopf-Rinow Theorem

as $t' = L(\gamma \mid [0,t']) \geq d(p,\gamma(t')) = d(p,p')$ and strict inequality would yield

$$d(p',q) = r - t' < r - d(p,p')$$

which contradicts $d(p,q) = r$. Hence

$$d(p,p') + d(p',q') + d(q',q) = d(p,q)$$

so by lemma 104

$$t' + s' = d(p,p') + d(p',q') = d(p,q') .$$

Now let $\gamma' : [t', t'+s'] \to M$ be a minimal geodesic from $p' = \gamma(t')$ to $q' \in S^{s'}(p')$. Then

$$\gamma' = \gamma \mid [t', t' + s']$$

for otherwise there would be a broken geodesic ($= \gamma$ from 0 to $t' = \gamma'$ from t' to $t' + s'$) from p to q' of minimal length which is impossible. Hence

$$d(\gamma(t' + s'), q) = d(q', q) = (r' - t') - s'$$

so $t' + s' \in I$ as required. □

Remark 105. The proof of the Hopf-Rinow theorem did not use the full strength of the hypothesis: It was not necessary to assume that *all* geodesics in M are defined for all time but only those which pass through the given point p. Also in the proof that (1) $\Longrightarrow$ (4) of the definition it was not necessary to know that

$$B^r(p) \subset \text{Exp}_p(V^r(p))$$

for all p but only for some p. Hence we have proved that if a manifold M has the property that for *some* point $p \in M$ every geodesic passing through p is defined for all t, then this is true for *every* point $p \in M$; i.e. M is complete.

Remark 106. If M is a closed submanifold of Euclidean space, then (as $\|p-q\| \leq d(p,q)$) M is complete. In particular, a compact manifold is always complete. The topologist's sine curve with equation (in $\mathbf{R}^2$):

$$y = \sin\frac{1}{x} \qquad (x > 0)$$

is complete but *not* closed.

Isometries

Let M and M' be submanifolds of Euclidean space E^n. A isometry is an isomorphism of the "intrinsic" geometries of M and M'.

Definition 107. *An* **isometry** *from M onto M' is a bijective transformation $\varphi : M \to M'$ which satisfies the following equivalent conditions:*

(1) the map φ is a diffeomorphism and for each $p \in M$ the linear map:

$$D\varphi(p) : T_pM \to T_{\varphi(p)}M'$$

is an orthogonal isomorphism:

(2) the map φ is a diffeomorphism and preserves the lengths of curves:

$$L(\gamma) = L(\varphi \circ \gamma)$$

for every curve $\gamma : J \to M$;

(3) the map φ intertwines the intrinsic metrics d and d' on M and M':

$$d(p,q) = d'(\varphi(p), \varphi(q))$$

for all $p, q \in M$.

We prove the equivalence of conditions (1) - (3). As

$$\begin{aligned} L(\dot{\gamma}) &= \int \|\dot{\gamma}(t)\| dt \\ &= \int \|D\varphi(\gamma(t))\dot{\gamma}(t)\| dt \\ &= L(\varphi \circ \gamma), \end{aligned}$$

the implication (1) $\Longrightarrow$ (2) is immediate. The implication (2) $\Longrightarrow$ (3) follows immediately from the definition of the distance as the infimum of the lengths of curves. For (3) $\Longrightarrow$ (1) fix $p \in M$ and define $\Phi_p(v) \in T_{\varphi(p)}M'$ for sufficiently small $v \in T_pM$ by:

$$\text{(i)} \qquad \mathrm{Exp}_{\varphi(p)}(\Phi_p(v)) = \varphi(\mathrm{Exp}_p(v)).$$

Since $\mathrm{Exp}_p(tv)$ (for $0 \le t \le 1$) is characterized by the equations:

$$\begin{aligned} d(p,\ \mathrm{Exp}_p(v)) &= d(p,\ \mathrm{Exp}_p(tv)) + d(\mathrm{Exp}_p(tv),\ \mathrm{Exp}_p(v)) \\ d(p,\ \mathrm{Exp}_p(tv)) &= td(p,\ \mathrm{Exp}_p(v)) \end{aligned}$$

it follows that Φ_p extends (uniquely) to a map

$$\Phi_p : T_pM \to T_{\varphi(p)}M'$$

which is homogeneous of degree one

$$\Phi_p(tv) = t\Phi_p(v)$$

for $v \in T_pM$, $t > 0$. It follows that

$$(ii) \qquad \Phi_p(v) = \frac{d}{dt}\varphi(\mathrm{Exp}_p(tv))\,|_{t=0}$$

and hence that

$$(iii) \qquad \|\Phi_p(v)\| = \|v\|$$

(since $d(p,\ \mathrm{Exp}_p(tv)) = t\|v\|$ for sufficiently small $t > 0$). Thus it suffices to prove that Φ_p is linear; then equation (i) shows that φ is smooth and equation (ii) shows that

$$(iv) \qquad Phi_p(v) = D\varphi(p)v$$

so equation (ii) implies (1) as required.

Hence choose $v, w \in T_pM$. By polarization

$$\begin{aligned} 2\langle v, w\rangle &= \|v\|^2 + \|w\|^2 - \|v - w\|^2 \\ &= \|v\|^2 + \|w\|^2 - \lim_{t\to 0} \frac{d(\mathrm{Exp}_p(tv),\ \mathrm{Exp}_p(tw))^2}{d^t} \end{aligned}$$

since:

$$\lim_{t\to 0} \frac{d(\mathrm{Exp}_p(tv),\ \mathrm{Exp}_p(tw))}{\|tv - tw\|} = 1$$

by an earlier theorem. The same equation holds when v, w replaced by $\Phi_p(v), \Phi_p(w)$ so (applying φ) this give:

$$\langle \Phi_p(v), \Phi_p(w)\rangle = \langle v, w\rangle$$

for $v, w \in T_pM$. But $v = v_1 + v_2$ is characterized by the equation

$$\langle v, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$$

for all $w \in T_pM$; hence

$$\Phi_p(v_1 + v_2) = \Phi_p(v_1) + \Phi_p(v_2)$$

as required.

Remark 108. If $\tilde{a} : E^n \to E^n$ is an isometry of the ambient Euclidean space with $\tilde{a}(M) = M'$ then certainly $\varphi = \tilde{a} \mid M$ is an isometry from M onto M'. On the other hand the map $\varphi : M \to M'$ given by

$$\varphi(0, \mathcal{O}, z) = (\cos \mathcal{O}, \sin \mathcal{O}, z)$$

where M is a plane set

$$M = \{(x, y, z) \in \mathbf{R}^3; x = 0, 0 < y < \pi/2\}$$

and M' is cylindrical

$$M' = \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 = 1; x, y > 0\}$$

is an isometry which is *not* of form $\varphi = \tilde{a} \mid M$. Indeed, an isometry of form $\varphi = \tilde{a} \mid M$ necessarily preserves the second fundamental form (as well as the first) in the sense that

$$ah_p(v)w = h'_{\tilde{a}(p)}(av)aw$$

for $v, w \in T_pM$ but in the example h vanishes identically while h' does not.

We may thus distinguish two fundamental question:

I. Given M and M' when are they extrinsically isomorphic? i.e. when is there a rigid motion $\tilde{a} \in R(E^n)$ with $\tilde{a}(M) = M'$?

II. Given M and M' when are they intrinsically isomorphic? i.e. when is there an isometry $\varphi : M \to M'$ from M onto M'?

As we have noted, both the first and second fundamental forms are preserved by extrinsic isomorphisms while only the first fundamental form need be preserved by an intrinsic isomorphism (i.e. an isometry).

A question which occurred to Gauss (who worked for a while as a cartographer) is this: Can one draw a perfectly accurate map of a portion of the earth? (i.e. a map for which the distance between points on the map is proportional to the distance between the corresponding points on the surface of the earth). We can now pose this question as follows: Is there an isometry from an open subset of a sphere to an open subset of a plane? Gauss answered this question negatively by associating an invariant, the Gauss curvature $K : M \to \mathbf{R}$, to a surface $M \subset E^3$. According to his "theorem egregium":

$$K'(\varphi(p)) = K(p)$$

for an isometry $\varphi : M \to M'$. The sphere has positive curvature; the plane has zero curvature; hence the perfectly accurate map does not exist. Our immediate aim is to explain these ideas.

We shall need a concept slightly more general then that of "isometry".

Definition 109. *A map $\varphi : M \to M'$ is called a* **local isometry** *iff it satisfies the following equivalent conditions*

(1) For $p \in M$ the linear map

$$D\varphi(p) : T_pM \to T_{\varphi(p)}M'$$

is an orthogonal linear isomorphism;

(2) Given $p \in M$ there are neighborhoods U of p in M and U' of $\varphi(p)$ in M' such that $\varphi \mid U$ is an isometry from U onto U'.

As $\varphi \mid U$ and φ have the same derivative at p it is clear that (2) $\Longrightarrow$ (1). On the other hand (1) implies that $D\varphi(p)$ is invertible so that (2) follows by the inverse function theorem. The simplest example of a local isometry which is not an isometry is the covering of the circle, by the line:

$$\varphi : \mathbf{R} \to S^1. \quad \varphi(\theta) = \cos\theta, \sin\theta).$$

The Riemann Curvature Tensor

Let $\gamma : \mathbf{R}^2 \to M$, $Z \in \mathcal{X}(\gamma)$ and (s, t) denote co-ordinates on $\mathbf{R}^2$. Recall the formula

$$\nabla_s \partial_t \gamma - \nabla_t \partial_s \gamma = 0$$

which says that ordinary partial differentiation and covariant partial differentiation commute. The analogous formula (which results on replacing ∂ by ∇ and γ by Z) is in general false. Instead we have the following

Definition 110. *The* **Riemann curvature tensor** *is the field which assigns to each $p \in M$ the multilinear map*

$$R_p \in L_a^2(T_pM, L(T_pM, T_pM))$$

characterized by the equation

$$R_p(u, v)w = (\nabla_s \nabla_t Z - \nabla_t \nabla_s Z)_{s=t=0} \tag{1}$$

for $u, v, w \in T_pM$ where $\gamma : \mathbf{R}^2 \to M$ and $Z \in \mathcal{X}(\gamma)$ satisfy

$$\gamma(0,0) = p, \quad (\partial_s\gamma)(0,0) = u, \quad (\partial_t\gamma)(0,0) = v, \quad Z(0,0) = w.$$

The fact that R is well-defined (i.e. depends only on the values of γ, $\partial_s\gamma$, $\partial_t\gamma$, and Z at $(0,0)$ and not on any higher derivatives) follows from the Gauss-Codaizzi formula established below. We shall give another proof which generalizes easily.

Suppose that the field R is well-defined and replace ∂_s, ∂_t, and Z by vectorfields $X, Y, Z \in \mathcal{X}(M)$. We obtain

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z + \nabla_{[X,Y]} Z \tag{1'}$$

The fact that $[X, Y] \in \mathcal{X}(M)$ (i.e. is tangent) follows immediately from Gauss-Weingarten equations and the symmetry of the second fundamental form. Formula (1′) clearly implies formula (1) for we may specialize to the case where $x_1 = s$, $x_2 = t$, $x_3, \ldots, x_m$ are local co-ordinates on M and X, Y are the co-ordinate vectorfields corresponding to s, t. Hence it is enough to show that $R(X, Y)Z$ is well-defined. For this it is enough to show that $(X, Y, Z) \to R(X, Y)Z$ is multilinear over the ring $\mathcal{F}(M)$ of smooth real-valued functions on M:

$$R(fX, Y)Z = R(X, fY)Z = R(X, Y)fZ = fR(X, Y)Z$$

for $f \in \mathcal{F}(M)$. Indeed, once we have this $\mathcal{F}(M)$-multilinearity we may write

$$R(X, Y)Z = \Sigma \xi^i \eta^j \rho^k R(E_i, E_j)E_k$$

where $E_1, \ldots, E_m \in \mathcal{X}(M)$ is a local moving frame and ξ^i, η^j, ζ^k are the components of X, Y, Z:

$$\begin{aligned} X &= \Sigma \xi^i E_i \\ Y &= \Sigma \eta^j E_j \\ Z &= \Sigma \zeta^k E_k \,. \end{aligned}$$

If $X'(p) = X(p)$ then $\xi^i(p) = {\xi'}^i(p)$ so if $X(p) = X'(p)$ $Y(p) = Y'(p)$, $Z(p) = Z'(p)$ then

$$(R(X, Y)Z((p) = (R(X', Y')Z')(p)$$

as required. The multilinearity follows by an easy computation using the two formulas:

$$\begin{aligned} \nabla_{fX} Y &= f \nabla_X Y \\ \nabla_X (fY) &= f \nabla_X Y + (D_X f)Y. \end{aligned}$$

Theorem 111 (Gauss-Codaizzi Equations)**.** *For $p \in M$ and $u, v \in T_pM$*

$$R_p(u, v) = h_p(u)^* h_p(v) - h_p(v)^* h_p(u)$$

$$(\nabla h)_p(u, v) = (\nabla h)_p(v, u)$$

where

$$h_p \in L(T_pM, L(T_pM, T_p^\perp M))$$

is the second fundamental form and

$$(\nabla h)_p \in L^2(T_pM, L(T_pM, T_p^\perp M))$$

is defined below.

Proof. Choose $\gamma : \mathbf{R}^2 \to M$ and $Z \in \mathcal{X}(\gamma)$ satisfying

$$\begin{aligned} \gamma(0, 0) = p; &\qquad Z(0, 0) = W; \\ (\nabla_s \gamma)(0, 0) = u; &\qquad (\nabla_t \gamma)(0, 0) = v. \end{aligned}$$

By Gauss-Weingarten

$$\begin{aligned}
\partial_t Z &= \nabla_t Z + h(\partial_t\gamma)Z; \\
\partial_s \nabla_t Z &= \nabla_s \nabla_t Z + h(\partial_s\gamma)\nabla_t Z; \\
\partial_s\{h(\partial_t\gamma)Z\} &= \nabla_s^\perp\{h(\partial_t\gamma)Z\} - h(\partial_s\gamma)^* h(\partial_t\gamma Z.
\end{aligned}$$

Hence

$$\begin{aligned}
\partial_s\partial_t Z &= \nabla_s\nabla_t Z - h(\partial_s\gamma)^* h(\partial_t\gamma)Z \\
&\quad +\nabla_s^\perp\{h(\partial_t\gamma)Z\} + h(\partial_s\gamma)\nabla_t Z\,.
\end{aligned}$$

Now let

$$\begin{aligned}
(\nabla h)(\partial_s\gamma, \partial_t\gamma)Z &= \nabla_s^\perp\{h(\partial_t\gamma)Z\} \\
&\quad -h(\nabla_s\partial_t\gamma)Z - h(\partial_t\gamma)\nabla, Z
\end{aligned}$$

and note that

$$\begin{aligned}
\partial_s\partial_s Z &= \partial_t\partial_s Z \\
\nabla_s\partial_t\gamma &= \nabla_t\partial_s\gamma
\end{aligned}$$

so interchanging s and t and subtracting gives

$$\begin{aligned}
0 &= \{R(\partial_s\gamma, \partial_t\gamma) - h(\partial_s\gamma)^* h(\partial_t\gamma) + h(\partial_t\gamma)^* h(\partial_s\gamma)\}Z \\
&\quad +\{(\nabla h)(\partial_s\gamma, \partial_t\gamma) - \nabla h(\partial_t\gamma, \partial_s\gamma)\}Z\,.
\end{aligned}$$

Now equating tangential and normal components gives the theorem provided $(\nabla h)_p$ is defined by

$$(\nabla h)_p(u,v)w = ((\nabla h)(\partial_s\gamma, \partial_t)Z)(0,0)\,.$$

As with R we must check $(\nabla h)_p$ is well-defined. The same argument works but we must use a normal moving frame $E_{m+1}, \ldots, E_n \in \mathcal{X}^\perp(M)$. The role of $(1')$ is played by the formula

$$(\nabla h)(X,Y)Z = \nabla_X^\perp\{h(Y)Z\} - h(\nabla_X Y)Z - h(Y)\nabla_X Z\,.$$

□

Remark 112. Equation $(1')$ can be written succinctly as

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

This can be contrasted with the identity

$$[D_X, D_Y] + D_{[X,Y]} = 0$$

where s$D_X : \mathcal{F}(M) \to B(M)$ is defined by

$$(D_X f)(p) = Df(p)X(p)$$

for a real-valued function $f \in \mathcal{F}(M)$.

Exercise 113. Show that the field which assigns to each $p \in M$ the multilinear map

$$R_p^\perp \in L_a^2(T_pM, L(T_p^\perp M, T_p^\perp M))$$

characterized by

$$\nabla_s^\perp \nabla_t^\perp U - \nabla_t^\perp \nabla_s^\perp U = R^\perp(\partial_s\gamma, \partial_t\gamma)U$$

for $\gamma : \mathbf{R}^2 \to M$ and $U \in \mathcal{X}^\perp(\gamma)$ satisfies the equation

$$R_p^\perp(u, v) = h_p(u)h_p(v)^* - h_p(v)h_p(u)^*$$

for $p \in M$, $u, v \in T_pM$.

Generalized Theorem Egregium

We will now show that geodesics, covariant differentiation, parallel transport, and the Riemann curvature tensor are all intrinsic; i.e. intertwined by isometries. These results are somewhat surprising since these objects are all defined using the second fundamental form, whereas isometries need not preserve the second fundamental form in any sense but only the first fundamental form.

Below we shall give a formula expressing the Gaussian curvature of a surface M^2 in E^3 in terms of the Riemann curvature tensor and the first fundamental form. It follows that the Gaussian curvature is also intrinsic. This fact was called by Gauss the "Theorema Egregium" which explains the title of this section.

Let $\varphi : M \to M'$ be a diffeomorphism. Using φ we can move objects from M to M': *a curve* $\gamma : J \to M$ *induces a curve* $\varphi \circ \gamma : J \to M'$. The first fundamental form g which assigns to each $p \in M$ the bilinear map $g_p \in L^2(T_pM; \mathbf{R})$ given by

$$g_p(v, w) = \langle v, w \rangle$$

induces a similar form $\varphi_* g$ on M' via the formula

$$(\varphi_* g)_{\varphi(p)}(D\varphi(p)v, D\varphi(p)w) = g_p(v, w)$$

for $p \in M$, $v, w \in T_pM$ (recall that $D\varphi(p) : T_pM \to T_{\varphi(p)}M'$ is invertible);

A vectorfield $X \in \mathcal{X}(M)$ induces a vectorfield $\varphi_* X \in \mathcal{X}(M')$ via the formula

$$(\varphi_* X)(\varphi(p)) = D\varphi(p)X(p)$$

for $p \in M$; A vectorfield $X \in \mathcal{X}(\gamma)$ along γ induces a vectorfield $\varphi_* X \in \mathcal{X}(\varphi \circ \gamma)$ along $\varphi \circ \gamma$ via the formula:

$$(\varphi_* X)(t) = D\varphi(\gamma(t))X(t)\,;$$

The Riemann curvature tensor R which assigns to each $p \in M$ a multilinear map $R_p \in L^2(T_pM; L(T_pM, T_pM))$ induces a similar tensor $\varphi_* R$ on M' via the formula

$$(\varphi_* R)_{\varphi(p)}(D\varphi(p)u, D\varphi(p)v)D\varphi(p)w = D\varphi(p)R_p(u, v)w$$

for $u, v, w \in T_pM$.

Proposition 114. *The first fundamental form is intrinsic; i.e. if $\varphi : M \to M'$ is an isometry then:*

$$\varphi_* g = g' .$$

Proof. This merely a restatement of the definition of isometry; viz. that $D\varphi(p) : T_pM \to T_{\varphi)p)}M$ be an orthogonal isomorphism. □

Proposition 115. *Geodesics are intrinsic; i.e. if $\varphi : M \to M'$ is an isometry and $\gamma : J \to M$ is a geodesic in M, then $\varphi \circ \gamma : J \to M'$ is a geodesic in M'.*

Proof. Clearly φ preserves the energy integral

$$E(\varphi \circ \gamma) = E(\gamma)$$

and hence also its extremal. □

Proposition 116. *The covariant derivative is intrinsic; i.e. if $\varphi :\to M'$ is an isometry, $\gamma : J \to M$, and $X \in \mathcal{X}(\gamma)$, then*

$$\nabla'(\varphi_* X) = \varphi_*(\nabla X) .$$

Proof. Choose a local parametrization $\psi : U \to M_0 \subset M$ on M, let $x^1, \ldots, x^m$ be the corresponding local co-ordinate system with co-ordinate vectorfields $E_1, \ldots, E_m \in \mathcal{X}(M_0)$, and Christoffel symbols Γ^k_{ij}:

$$\begin{aligned} &\psi(x^1(p), \ldots, x^m(p)) = p, \\ &E_i \circ \psi = \partial_i \psi, \\ &\nabla_i E_j = \sum_k \Gamma^k_{ij} E_k. \end{aligned}$$

Use the transformation φ to induce a local parametrization $\varphi \circ \psi : U \to M'_0 = \varphi(M_0) \subset M'$ so that

$$\begin{aligned} &{x'}^i(\varphi(o)) = x^i(p), \\ &\varphi_* E_i = E'_i, \\ &\nabla'_i E'_j = \sum_k {\Gamma'}^k_{ij} E'_k. \end{aligned}$$

Now let ξ^i be the components of X in the frame E_i:

$$X(t) = \sum_i \xi^i(t) E_i(\gamma(t)).$$

They are also the components of $\varphi_* X$ in the induced frame:

$$(\varphi_* X)(t) = \sum_i \xi^i(t) E'_i(\varphi \circ \gamma(t)).$$

Now compute the covariant derivatives of X and $\varphi_* X$ in theses frames:

$$\nabla X = \sum_{i,j,k} \left(\dot{\xi}^k + \Gamma^k_{ij} v^i \xi^j \right) E_k$$

$$\nabla' \varphi_* X = \sum_{i,j,k} \left(\dot{\xi}^k + {\Gamma'}^k_{ij} v^i \xi^j \right) E'_k$$

where v^i are the components of the velocity vector to the curve:

$$v^i = \frac{d}{dt} x^i(\gamma(t)) = \frac{d}{dt} {x'}^i(\varphi \circ \gamma(t)).$$

Thus to prove that $\nabla' \varphi_* X = \varphi_* \nabla X$ we must prove that φ preserves the Christoffel symbols, that is,

$${\Gamma'}^k_{ij}(\varphi(p)) = \Gamma^k_{ij}(p)$$

for $p \in M_0$.

We must show that the equation

$$(*) \qquad \sum_{ij} {\Gamma'}^k_{ij}(\varphi(p)) v^i \xi^j = \sum_{ij} \Gamma^k_{ij}(p) v^i \xi^j$$

holds identically in (v, ξ). Since φ maps geodescis to geodesics it must preserve the geodesic equation

$$\dot{v}^k = \sum_{ij} \Gamma^k_{ij} v^i v^j.$$

This means that $(*)$ holds when $v = \xi$. Hence (by polarization) to show that $(*)$ holds identically it is enough to show that the Christoffel symbols are symmetric in (i, j). Here's the calculation which proves that:

$$\begin{aligned} \sum_k \Gamma^k_{ij} E_k &= \nabla_i \partial_j \psi \\ &= \nabla_j \partial_i \psi \\ &= \sum_k \Gamma^k_{ji} E_k \end{aligned}$$

□

Second Proof. The proposition can be reformulated in terms of vectorfields:

$$\nabla^p r_{\varphi_* Y} \varphi_* X = \varphi_* \nabla_Y X$$

for $X, Y \in \mathcal{X}(M)$. For this it suffices to show that

$$\langle \nabla'_{\varphi_* Y} \varphi_* X, \varphi_* Z \rangle = \langle \nabla_Y X, Z \rangle$$

for $X, Y, Z \in \mathcal{X}(M)$. But

$$D_Y\langle X, Z\rangle = \langle \nabla_Y X, Z\rangle + \langle X, \nabla_Y Z\rangle .$$

In the last equation cyclically permute X, Y, Z to get two further equations. Add two of them and subtract the third to obtain:

$$\begin{aligned} D_Y\langle X, Z\rangle + D_X\langle Y, Z\rangle - D_Z\langle X, Y\rangle &= 2\langle \nabla_Y X, Z\rangle + \langle [Y, X], Z\rangle \\ &\quad + \langle [Z, X], Y\rangle + \langle [Z, Y], X\rangle \end{aligned}$$

where we recall that the Lie brackets satisfy

$$[Y, X] = \nabla_X Y - \nabla_X Y .$$

Since we have

$$\begin{aligned} \varphi_*[X, Y] &= [\varphi_* X, \varphi_* Y] \\ \varphi_*(D_X f) &= D_{\varphi_* X}\varphi_* f \end{aligned}$$

for any diffeomorphism and any $X, Y \in \mathcal{X}(M)$, $f \in \mathcal{F}(M)$ the desired formula is now established. □

Proposition 117. *Parallel transport is intrinsic; i.e. if $\varphi : M \to M'$ is an isometry, and $\gamma : \mathbf{R} \to M$, then*

$$D\varphi(p)\tau(\gamma, t, t_0) = \tau'(\varphi \circ \gamma, t, t_0) D\varphi(p_0)$$

where $p = \gamma(t)$, $p_0 = \gamma(t_0)$. Also affine parallel transport is intrinsic

$$\tilde{D}\varphi(p)\tilde{\tau}(\gamma, t, t_0) = \tilde{\tau}'(\varphi \circ \gamma, t, t_0)\tilde{D}\varphi(p_0)$$

where $\tilde{D}\varphi(p) : \tilde{T}_p M \to \tilde{T}_{\varphi(p)} M'$ is given by

$$\tilde{D}\varphi(p)(p + v) = \varphi(p) + D\varphi(p)v$$

for $v \in T_p M$.

Proof. The fields $X(t) = \tau(\gamma, t, t_0)v_0$ and $\tilde{Y}(t) = \tilde{\tau}(\gamma, t, t_0)(p_0 + v_0)$ (for $v_0 \in T_{p_0} M$) are characterized by the equations

$$\begin{aligned} \nabla X &= 0, \\ X(t_0) &= v_0, \\ \nabla(\tilde{Y} - \dot{\gamma}) + \dot{\gamma} &= 0, \\ \tilde{Y}(t_0) &= p_0 + v_0. \end{aligned}$$

□

Exercise 118. Formulate and prove a proposition which asserts that developments are intrinsic.

Proposition 119. *The Riemann curvature tensor R is intrinsic; i.e. if $\varphi : M \to M'$ is an isometry, then*

$$\varphi_* R = R'.$$

Proof. R is characterized by

$$R(\partial_s \gamma, \partial_t \gamma) X = \nabla_s \nabla_t X - \nabla_t \nabla_s X\,.$$

□

Exercise 120. Given local co-ordinates $x^1, \ldots, x^m : M_0 \to \mathbf{R}$ as above define the **coefficients** $g_{ij} : M_0 \to \mathbf{R}$ **of the first fundamental form** by:

$$g_{ij} = g(E_i, E_j)$$

and define the coefficients $R^\ell_{ijk} : M_0 \to \mathbf{R}$ of the curvature tensor by

$$R(E_i, E_j) E_k = \sum_\ell R^\ell_{ijk} E_\ell\,.$$

Show that Γ^k_{ij} and R^ℓ_{ijk} can be expressed in terms of the g_{ij} thereby deriving the more traditional proofs of the propositions in this section.

Hint. The desired formulas are:

$$2 \sum_\ell g_{k\ell} \Gamma^\ell_{ij} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}$$

and

$$\begin{aligned} R^\ell_{ijk} &= \partial_i \Gamma^\ell_{jk} - \partial_j \Gamma^\ell_{ik} \\ &\quad + \sum_h (\Gamma^\ell_{ih} \Gamma^h_{jk} - \Gamma^\ell_{jh} \Gamma^h_{ik})\,. \end{aligned}$$

To prove the former cyclically permute the indices in the formulas

$$\partial_i g(E_j, E_k) = g(\nabla_i E_j, E_k) + g(E_j, \nabla_i E_k)$$

to get two similar formulas. Add two of them and subtract the third. □

The Cartan-Ambrose-Hicks Theorem

In this section we address what might be called the "fundamental problem of intrinsic differential geometry": viz. when are two manifolds isometric? In some sense the theorem of this section answers that question (at least locally) although the equivalent conditions given there are probably more difficult to verify in most examples than the condition that there exist an isometry. We begin with a proposition which asserts that there cannot be too many isometries.

Proposition 121. *Let M and M' be connected manifolds, $p \in M$, and $\varphi_1, \varphi_2 : M \to M'$ be local isometries. Assume*

$$\varphi_1(p) = \varphi_2(p) = p', \quad D\varphi_1(p) = D\varphi_2(p) = b_0.$$

Then $\varphi_1 = \varphi_2$.

Proof. The formulas

$$\varphi_i(\mathrm{Exp}_p(v)) = \mathrm{Exp}'_{p'}(b_0 v)$$

for $i = 1, 2$ and sufficiently small $v \in T_pM$ show that $\varphi_1 = \varphi_2$ in a small neighborhood of p (viz. a neighborhood which is the diffeomorphic image via Exp_p of a small ball about the origin in T_pM). This argument shows more generally that the set of points $q \in M$ at which $\varphi_1(q) = \varphi_2(q)$ and $D\varphi_1(q) = D\varphi_2(q)$ is open. This set is clearly closed by continuity, so the proposition is proved. □

Definition 122. *A* **homotopy** *of maps from J to M is a map*

$$\gamma : [0, 1] \times J \to M.$$

We often write

$$\gamma_\lambda(t) = \gamma(\lambda, t)$$

for $\lambda \in [0, 1]$ and $t \in J$ and call γ a **homotopy between γ_0 and γ_1**. *When $J = [a, b]$ is a closed finite interval we say the homotopy has* **fixed endpoints** *if*

$$\gamma_\lambda(a) = \gamma_0(a), \quad \gamma_\lambda(b) = \gamma_0(b)$$

for all $\lambda \in [0, 1]$.

We remark that a homotopy and a variation are essentially the same thing; viz. a curve of maps (curves). The difference is pedagogical. We used the word "variation" to describe a curve of maps through a given map; when we use this word we are going to differentiate the curve t find a tangent vector (field) to the given map. The word "homotopy" is used to describe a curve joining two maps; it is a global rather than a local (infinitesimal) concept.

Definition 123. *The manifold M is* **simply connected** *iff for any two curves $\gamma_0, \gamma_1 : [a, b] \to M$ with*

$$\gamma_0(a) = \gamma_1(a), \quad \gamma_0(b) = \gamma_1(b)$$

there is a homotopy from γ_0 to γ_1 with endpoints fixed. (The idea is that the space $\Omega(p, q)$ of curves from p to q is connected.)

Example 124. A Euclidean space is simply connected

$$\gamma(\lambda, t) = \gamma_0(t) + \lambda(\gamma_1(t) - \gamma_0(t))$$

while the punctured plane $\mathbf{C}\backslash\{0\}$ is not, for the curves

$$\gamma_n(t) = e^{2\pi int} \qquad (0 \le t \le 1)$$

are not homotopic for distinct n.

Theorem 125 (Global C-A-H Theorem)**.** *Let M and M' be connected, simply connected and complete, $p \in M$, $p' \in M'$, and*

$$b_0 : T_pM \to T_{p'}M'$$

an orthogonal linear isomorphism. Then the following conditions are equivalent

(1) there is an isometry

$$\varphi : M \to M'$$

satisfying

$$\varphi(p) = p', \quad D\varphi(p) = b_0, ;$$

(2) for any development (b, γ, γ') satisfying the initial condition

$$(b(0), \gamma(0), \gamma'(0)) = (b_0, p, p')$$

we have $\gamma'(1) = p'$ and $b(1) = b_0$ whenever $\gamma(1) = p$;

(3) for any pair of developments $(b_i, \gamma_i, \gamma_i')$ $(i = 0, 1)$ satisfying the initial condition

$$(b_i(0), \gamma_i(0), \gamma_i'(0)) = (b_0, p, p')$$

we have $\gamma_0'(1) = \gamma_1'(1)$ whenever $\gamma_0(1) = \gamma_1(1)$;

(4) for any development (b, γ, γ') with $(b(0), \gamma(0), \gamma'(0)) = (b_0, p, p')$ we have

$$b(t)_* R_{\gamma(t)} = R'_{\gamma'(t)}$$

for all t.

Moreover, when these equivalent conditions hold we have:

$$\begin{aligned} \gamma'(t) &= \varphi(\gamma(t)) \\ b(t) &= D\varphi(\gamma(t)) \end{aligned}$$

for any development (b, γ, γ') satisfying the initial condition in (2).

Example 126. Before giving the proof let us interpret the conditions in case M and M' are two-dimensional spheres of radius r and r' respectively in three-dimensional Euclidean space E^3. Imagine that the spheres are tangent at $p = p'$. Clearly the spheres will be isometric exactly when $r = r'$. Condition (2) says that if the spheres are rolled along one another without sliding or twisting then the endpoint $\gamma'(1)$ of one curve of contact depends only on the endpoint $\gamma(1)$ of the other and not on the intervening curve $\gamma(t)$. Finally we shall see below that the Riemann curvature of a two manifold at a point q is determined by a number $K(q)$ called the Gauss curvature; and that for spheres we have $K(q) = 1/r$.

Exercise 127. Let γ be the closed curve which bounds the first octant as shown in the diagram for example 126. Find γ'.

Figure 9: Diagram for example 126

Exercise 128. Show that in case M is two-dimensional, the condition $b(1) = b_0$ may be dropped from (2).

Proof of theorem. We first prove a slightly different theorem; viz. we drop the condition that M' be simply connected (but continue to assume that M is) and weaken (1) to assert that φ is a local isometry (i.e. not necessarily bijective).

We prove (1) $\Longrightarrow$ "Moreover ...". Indeed if $\gamma' = \varphi \circ \gamma$ and $b = (D\varphi) \circ \gamma$ as in "Moreover" then $\dot{\gamma}' = b\dot{\gamma}$ by the chain rule and $b\tau = \tau' b$ by the last section so (b, γ, γ') is the development with initial condition $(b(0), \gamma(0), \gamma'(0)) = (b_0, p, p')$.

We prove (1) $\Longrightarrow$ (2). Given a development as in (2) we have (by "moreover")

$$\begin{aligned} \gamma'(1) &= \varphi(\gamma(1)) = \varphi(p) = p' \\ b(1) &= D\varphi(\gamma(1)) = D\varphi(p) = b_0 \end{aligned}$$

as required.

We prove (2) $\Longrightarrow$ (3). Choose developments $(b_i, \gamma_i, \gamma_i')$ $i = 0, 1$ as in (3). Define a curve $\gamma : [0, 1] \to M$ by "composition"

$$\begin{aligned} \gamma(t) &= \gamma_0(2t) && 0 \le t \le \tfrac{1}{2} \\ &= \gamma_1(2 - 2t) && \tfrac{1}{2} \le t \le 1 \end{aligned}$$

so that γ is continuous and piecewise smooth and $\gamma(1) = p$. Let (b, γ, γ') be the development determined by the initial condition $(b(0), \gamma(0), \gamma'(0)) = (b_0, p, p')$ so that by (2) we have that $\gamma'(1) = p$ and $b(1) = b_0$. By the uniqueness of developments and the invariance under reparametrization

$$(b(t), \gamma(t), \gamma'(t)) = \begin{cases} (b_0(2t), \gamma_0(2t), \gamma_0'(2t)) & 0 \le t \le \frac{1}{2} \\ (b_1(2 - 2t), \gamma_1(2 - 2t), \gamma_1'(2 - 2t)) & \frac{1}{2} \le t \le 1 \end{cases}$$

hence $\gamma_0'(1) = \gamma'\left(\frac{1}{2}\right) = \gamma_1'(1)$ as required.

We prove (3) $\Longrightarrow$ (1). By (3) we may define $\varphi : M \to M'$ by

$$\varphi(q) = \gamma'(1)$$

where $\gamma : [0, 1] \to M$ is any curve from p to q and (b, γ, γ') is the development with $b(0) = b_0$. (According to (3), φ is well-defined; i.e. independent of the choice of γ from p to q.) Now

$$\varphi(\gamma(t)) = \gamma'(t)$$

for $0 \le t \le 1$ so applying (d/dt) gives:

$$D\varphi(\gamma(t))\gamma(t) = \dot{\gamma}'(t) = b(t)\gamma(t)\,.$$

But each $b(t)$ is an orthogonal transformation so

$$\|D\varphi(\dot\gamma(t))\dot\gamma(t)\| = \|b(t)\dot\gamma(t)\| = \|\dot\gamma(t)\|\,.$$

Given $q \in M$ and $v \in T_qM$ we may always choose $\gamma(1) = q$ and $\dot\gamma(1) = v$ so that:

$$\|D\varphi(q)v\| = \|v\|$$

i.e. φ is a local isometry as required.

We prove (1) $\Longrightarrow$ (4). By "moreover"

$$b(t) = D\varphi(\gamma(t))$$

for any development as in (4); hence (4) follows from the previous section.

We prove (4) $\Longrightarrow$ (3). Choose developments $(b_i, \gamma_i, \gamma_i')$ as in (3). Since M is simply connected choose a homotopy γ from γ_0 to γ_1 with endpoints fixed. Let $(b_\lambda, \gamma_\lambda, \gamma_\lambda')$ be the development with initial conditions:

$$\gamma_\lambda(0) = p, \quad b_\lambda(0) = b_0$$

so to show $\gamma_1'(1) = \gamma_0'(1)$ we can show $\partial_\lambda \gamma_\lambda'(1) = 0$. In fact we will show that for each fixed t the curve

$$\lambda \mapsto (b_\lambda(t), \gamma_\lambda(t), \gamma_\lambda'(t))$$

is a development; then by the definition of development we have

$$\partial_\lambda \gamma_\lambda'(1) = b_\lambda(1)\partial_\lambda\gamma_\lambda(1) = 0$$

as required.

Choose a basis $E_{l0}, \ldots, E_{m0}$ for T_pM and extend to get a moving frame $E_i \in \mathcal{X}(\gamma)$ along the homotopy γ by imposing the conditions that the vectorfields $t \mapsto E_i(\lambda, t)$ be parallel

$$\nabla_t E_i = 0\,.$$

Let

$$\begin{aligned}\partial_t\gamma &= \sum_i \xi^i E_i \\ E_i' &= bE_i\end{aligned}$$

so that by the definition of development we have

$$\begin{aligned}\partial_t\gamma' &= \sum_i \xi^i E_i' \\ \nabla_t E_i' &= 0.\end{aligned}$$

By definition of curvature we have

$$\begin{aligned}\nabla_t\nabla_\lambda E_i &= R(\partial_t\gamma, \partial_\lambda\gamma)E_i \\ \nabla_t'\nabla_\lambda' E_i' &= R'(\partial_t\gamma', \partial_\lambda'\gamma)E_i'\end{aligned}$$

and applying b to the former equation gives

$$\nabla'_t\{b\nabla_\lambda E_i\} = R'(\partial_t\gamma', b(\partial_\lambda\gamma))E'_i$$

where we have used the hypothesis $b_*R = R'$ and the fact that $b\nabla_t = \nabla'_t b$ which holds since we have developments for fixed λ. Now

$$\begin{aligned}\nabla'_t\partial_\lambda\gamma' &= \nabla'_\lambda\partial_t\gamma' \\ &= \sum_i\{(\partial_\lambda\xi^i)E'_i + \xi^i(\nabla'_\lambda E'_i)\}\end{aligned}$$

while

$$\begin{aligned}\nabla'_t(b\partial_\lambda\gamma) &= b\nabla_t\partial_\lambda\gamma \\ &= b\nabla_\lambda\partial_t\gamma \\ &= \sum_i\{(\partial_\lambda\xi^i)E'_i + \xi^i(b\nabla_\lambda E_i)\}\,.\end{aligned}$$

We have shown that the equations

$$\begin{aligned}\nabla'_t Y'_i &= R'(\partial_t\gamma', X')E'_i \\ \nabla'_t X' &= \sum_i\{(\partial_\lambda\xi^i)E'_i + \xi^i Y'_i\}\end{aligned}$$

are satisfied by both

$$\begin{aligned}X' &= \partial_\lambda\gamma' \\ Y'_i &= \nabla'_\lambda E'_i\end{aligned}$$

and also by

$$\begin{aligned}X' &= b\partial_\lambda\gamma \\ Y'_i &= b\nabla_\lambda E_i\,.\end{aligned}$$

Both solutions satisfy the initial conditions that they vanish identically at $t = 0$ so they are equal. But this says that $\partial_\lambda\gamma' = b\partial_\lambda\gamma$ and $b\nabla_\lambda = \nabla'_\lambda b$ which says that $\lambda \to (b_\lambda(t), \gamma_\lambda(t), \gamma'_\lambda(t))$ is a development as required.

Now the modified theorem (where φ is a local isometry) is proved. The original theorem follows immediately. Condition (4) is symmetric in M and M'; hence if we assume (4) we have local isometries $\varphi : M \to M'$, $\psi : M' \to M$ with

$$\begin{aligned}\varphi(p) = p', &\qquad \psi(p') = p, \\ D\varphi(p) = b_0, &\qquad D\psi(p') = b_0^{-1}.\end{aligned}$$

But then $\psi\circ\varphi$ is a local isometry with $\psi\circ\varphi(p) = p$ and $D(\psi\circ\varphi)(p) =$ identity. Hence $\psi\circ\varphi$ is the identity. Similarly $\varphi\circ\psi$ is the identity so φ is bijective (and $\psi = \varphi^{-1}$) as required. □

Theorem 129 (Local C-A-H Theorem)**.** *Suppose given $p \in M$, $p' \in M'$, and an orthogonal linear isomorphism $b_0 : T_pM \to T_{p'}M'$. Suppose $r > 0$ is sufficiently small that Exp_p and $\mathrm{Exp}_{p'}$ map the balls of radius r about the origin diffeomorphically onto $B^r(p)$ and $B^r(p')$ respectively. Then the following are equivalent:*

(1) There is an isometry $\varphi : B^r(p) \to B^r(p')$ with $\varphi(p) = p'$ and $D\varphi(p) = b_0$.

(2) For every development (b, γ, γ') with $\gamma([0, 1]) \subset B^r(p)$, $\gamma'([0, 1]) \subset B^r(p')$, and satisfying the initial conditions $(b(0), \gamma(0), \gamma'(0)) = (b_0, p, p')$ we have that $\gamma'(1) = p'$ and $b(1) = b_0$ whenever $\gamma(1) = p$.

(3) For every pair of developments $(b_0, \gamma_0, \gamma_0')$, $(b_1, \gamma_1, \gamma_1')$ as in (2) we have that $\gamma_0'(1) = \gamma_1(1)$ whenever $\gamma_0(1) = \gamma_1(1)$;

(4) For each $v \in T_pM$ with $\|v\| < r$ we have

$$b(t)_* R_{\gamma(t)} = R'_{\gamma'(t)} \qquad (0 \le t \le 1)$$

where

$$\begin{aligned} \gamma(t) &= \mathrm{Exp}_p(tv) \\ \gamma'(t) &= \mathrm{Exp}_{p'}(tb_0 v) \\ b(t) &= \tau'(\gamma', t, 0) b_0 \tau(\gamma, 0, t)\,. \end{aligned}$$

Moreover when these equivalent conditions hold, φ is given by

$$\varphi(\mathrm{Exp}_p(v)) = \mathrm{Exp}_{p'}(b_0 v)\,.$$

Proof. The proofs (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (1) $\Longrightarrow$ (4) are as before; the reader might note that when $L(\gamma) \le r$ we also have $L(\gamma') \le r$ for any development so that there are plenty of developments with $\gamma : [0, 1] \to B^r(p)$ and $\gamma' : [0, 1] \to B^r(p')$. The proof that (4) $\Longrightarrow$ (1) is a little different since (4) here is somewhat weaker than (4) of the global theorem: the equation $b_* R = R'$ is only assumed for certain developments.

Hence assume (4), choose $v, w \in T_pM$ with $\|v\| < r$ and define $\alpha(\lambda, t)$ by

$$\alpha(\lambda, t) = \mathrm{Exp}_p(t(v + \lambda w))$$

for $0 \le t \le 1$ and λ near 0.

Define $X = X(t)$ by

$$X(t) = \frac{\partial}{\partial \lambda} \alpha(\lambda, t)\,|_{t=0}\ .$$

Claim . *X is a "Jacobi field" along $\gamma(t) = \alpha(0, t)$; i.e. X satisfies the differential equation*

$$\nabla^2 X = R(\dot\gamma, X)\dot\gamma\,.$$

Moreover X satisfies the initial conditions

$$X(0) = 0, \quad \nabla X(0) = w.$$

To prove the claim we calculate

$$\begin{aligned}\nabla X(0) &= \nabla_t \partial_\lambda \alpha(0,0)\\ &= \nabla_\lambda \partial_t \alpha(0,0)\\ &= \frac{d}{d\lambda}(v+\lambda w)\,|_{\lambda=0}\\ &= w\end{aligned}$$

and

$$\begin{aligned}\nabla^2 X &= \nabla_t \nabla_t \partial_\lambda \alpha\\ &= \nabla_t \nabla_\lambda \partial_t \alpha\\ &= \nabla_\lambda \nabla_t \partial_t \alpha + R(\partial_t \alpha, \partial_\lambda \alpha)\partial_t \alpha\\ &= R(\dot\gamma, X)\dot\gamma\,.\end{aligned}$$

Proving the claim.

Now define φ as in "moreover" so that

$$\varphi(\alpha(\lambda,t)) = \operatorname{Exp}_{p'}(t(b_0 v + \lambda b_0 w))$$

by definition. Hence reading M for M' in the claim gives

$$\nabla^2 X' = R'(\dot\gamma', X')\dot\gamma'$$

where

$$X'(t) = \frac{\partial}{\partial\lambda}\varphi(\alpha(\lambda,t))\,|_{\lambda=0}\ .$$

On the other hand, if $b(\lambda t)$ is defined by demanding that $b(\lambda,0) = b_0$ and that for each λ the curve

$$t \mapsto (b(\lambda,t), \alpha(\lambda,t), \varphi(\alpha(\lambda,t)))$$

be a development then by the hypothesis of the theorem, $b_* R = R'$ so applying b to the claim gives

$$\nabla'^2(bX) = R'(\gamma', bX)\gamma'$$

at $\lambda = 0$. Hence X' and bX satisfy the same differential equation (with the same initial condition) and so are equal. Thus

$$D\varphi(q)u = X'(t) = bu$$

with $q = \alpha(0,t)$, $b = b(0,t)$, $u = X(t)$ so that

$$\|D\varphi(q)u\| = \|bu\| = \|u\|$$

as required. □

Flat Spaces

Our aim in the next few sections is to give applictions of the Cartan-Ambrose-Hicks Theorem. It is clear that the hypothesis $b_*R = R'$ for *all* developments will be difficult to verify without drastic hypotheses on the curvature. The most drastic such hypothesis is that the curvature vanishes identically.

Definition 130. *The manifold M is called* **flat** *iff it satisfies the following equivalent conditions:*

(1) Every point has a neighborhood which is isometric to an open subset of Euclidean space: i.e. at every point p there exist local co-ordinates $x^1, \dots, x^m$ such that the co-ordinate vectorfields $E_1, \dots, E_m$ are orthonormal.

(2) The Riemann curvature tensor R vanishes identically: $R \equiv 0$

Theorem 131. *A complete connected, and simply connected manifold M is flat if and only if there is an isometry $\varphi : M \to R^m$ onto a Euclidean space.*

(The equivalence of (1) and (2) and the theorem both follow immediately from the C-A-H Theorem.)

Exercise 132. A one-dimensional manifold (curve) is always flat.

Exercise 133. If $M_1 \subset E^{n_1}$ and $M_2 \subset E^{n_2}$ are flat so is $M = M_1 \times M_2 \subset E^{n_1} \times E^{n_2} \simeq E^{n_1+n_2}$.

Exercise 134. Let a, b be positive and c non-negative and define $M \subset \mathbf{C}^3 = \mathbf{R}^6 = E^6$ by the equations

$$|u| = a, \quad |v| = b \quad w = cuv$$

where $u, v, w \in \mathbf{C}$. Then M is diffeomorphic to a torus (a product of two circles), M is flat, and if M' is similarly defined from numbers a', b', c' then there is an isometry $\varphi : M \to M'$ if and only if $a = a'$, $b = b'$, and $c = c'$ (i.e. $M = M'$). (Hint: each Circle $u = u_0$ is a geodesic as well as each circle $v = v_0$; the numbers a, b, c can be computed from the length of the circle $u = u_0$ the length of the circle $v = v_0$, and angle between them.)

Exercise 135. Let $n = m + 1$ and let $\tilde{E}(t)$ be a one-parameter family of hyperplanes in E^n; more precisely

$$\tilde{E}(t) = \gamma(t) + U(t)^\perp$$

where $\gamma : \mathbf{R} \to E^n$, $U : \mathbf{R} \to \mathbf{E}^n \backslash \{0\}$, and

$$U^\perp = \{v \in \mathbf{E}^n : \langle U^\perp, v \rangle = \circ\,\}\,.$$

Show that if $U(t_0)$ and $\dot{U}(t_0)$ are linearly independent, then the limit

$$\tilde{L}(t_0) = \lim_{t \to t_0} \tilde{E}(t) \cap \tilde{E}(t_0)$$

exists and is a Euclidean space of dimension $n-2=m-1$; moreover if $\dot{\gamma}(t_0)$ is not parallel to $\tilde{L}(t_0)$, then the affine spaces $\tilde{L}(t)$ fit together to form a manifold near $\gamma(t_0)$; more precisely the set

$$M_0 = \bigcup_{|t-t_0|<\epsilon} \tilde{L}(t)_\rho$$

where

$$\tilde{L}(t)_\rho = \{P \in \tilde{L}(t) \mid \|p-\gamma(t)\| < \rho\}$$

is a manifold of dimension $m = n-1$ for $\rho, \epsilon > 0$ sufficiently small. A manifold which arises this way is called **developable**. Show that the affine tangent spaces to M_0 are the original hyperplanes $\tilde{E}(t)$

$$\tilde{T}_p M_0 = \tilde{M}_{\gamma(t)} M_0 = \tilde{E}(t)$$

for $p \in \tilde{L}(t)_\rho$ and $|t-t_0| < \epsilon$. (One therefore calls M_0 the "envelope" of the family of hyperplanes $\tilde{E}(t)$.) Show that M_0 is flat (hint: use Gauss-Codaizzi) and conclude that any development (b, γ, γ') of M_0 along a Euclidean space $E^m = m'$ "unrolls" M in the sense that the map $\varphi : M_0 \to E^m$ given by

$$\varphi(p) = \gamma'(t) + b(t)(p - \gamma(t))$$

for $|t-t_0| < \epsilon$ and $p \in \tilde{L}(t)_\rho$ is an isometry. When $n=3$, $m=2$ one can visualize M as a twisted sheet of paper as in the figure.

Remark 136. Given a hypersurface $M \subset E^{m+1}$ and a curve $\gamma : \mathbf{R} \to M^*$ we may form the **osculating developable** M_0 to M^* along γ by taking

$$\tilde{E}(t) = \tilde{T}_{\gamma(t)} M\,.$$

This developable has common tangent spaces with M along γ:

$$\tilde{T}_{\gamma(t)} M = \tilde{T}_{\gamma(t)} M_0\,.$$

This gives a nice interpretation of parallel transport: M_0 may be unrolled onto a hyperplane where parallel transport has an obvious meaning and the identification of the tangent spaces thereby defines parallel transport in M.

Exercise 137. Show that each of the following is a developable surface in E^3

(1) A cone on a plane curve $\Gamma \subset E^2$

$$M = \{tp + (1-t)q : t > 0, q \in \Gamma\}$$

where p is a fixed point not on the plane E^2.

(2) A cylinder on a plane curve $\Gamma \subset R^2$

$$M = \{q + tv : q \in \Gamma, t \in \mathbf{R}\}$$

where v is a fixed vector not parallel to E^2 (this is a cone with the cone point p at infinity).

Figure 10: A developable manifold

(3) The tangent developable to a space curve $\gamma : \mathbf{R} \to E^3$

$$M = \{\gamma(t) + s\dot{\gamma}(t) : |t - t_0| < \epsilon, 0 < s < \epsilon\}$$

where $\dot{\gamma}(t_0)$ and $\ddot{\gamma}(t_0)$ are linearly independent and $\epsilon > 0$ is sufficiently small.

(4) The normal developable to a space curve γ

$$M = \{\gamma(t) + s\ddot{\gamma}(t) : |t - t_0| < \epsilon, |s| < \epsilon\}$$

where $\|\dot{\gamma}(t)\| = 1$ for all t, $\gamma(t_0) \neq 0$, and $\epsilon > 0$ is sufficiently small.

Symmetric Space

In the last section we applied the Cartan-Ambrose-Hicks Theorem in the flat case; the hypothesis $b_* R = R'$ was easy to verify since both sides vanish. To find more general situations where we can verify this hypothesis note that for any development (b, γ, γ') we have

$$b(t) = \tau'(\gamma', t, 0) b(0) \tau(\gamma, 0, t)$$

so that the hypothesis $b_* R = R'$ is certainly implied by the three hypotheses

$$\begin{aligned}
\tau(\gamma, t, 0)_* R_p &= R_{\gamma(t)} \\
\tau'(\gamma', t, 0)_* R'_{p'} &= R'_{\gamma'(t)} \\
b_{0*} R_p &= R'_{p'}
\end{aligned}$$

where (b, γ, γ') is a development with $(b(0), \gamma(0), \gamma'(0)) = (b_0, p, p')$. The last hypothesis is a condition on the initial linear isomorphism:

$$b_0 : T_pM \to T_{p'}M'$$

while the former hypothesis are conditions on M and M' respectively; viz. that the Riemann curvature tensor is invariant by parallel transport. It is rather amazing that this condition is equivalent to a rather simple geometric condition as we now show.

Definition 138. *The manifold M is* **symmetric about the point** $p \in M$ *iff there is a (necessarily unique) isometry $\varphi : M \to M$ such that*

$$\varphi(p) = p, \quad D\varphi(p) = -I,$$

where I denotes the identity transformation of T_pM; M is a **symmetric space** *iff it is symmetric about each of its points.*

Exercise 139. A symmetric space is complete. (Hint: If $\gamma : J \to M$ is a geodesic and $\varphi : M \to M$ is a symmetry about $\gamma(t_0)$ then $\varphi(\gamma(t)) = \gamma(2t_0 - t)$.)

Definition 140. *The manifold M is called a* **locally symmetric space** *iff it satisfies the following equivalent conditions*

(1) *For each $p \in M$ the open ball $B^r(p)$ with center p and radius r is symmetric about p for sufficiently small $r \geq 0$.*

(2) *The covariant derivative ∇R, (defined below) vanishes identically*

$$(\nabla R)_p(v)(v_1, v_2)w = 0$$

for all $p \in M$ and all $v, v_1, v_2, w \in T_pM$.

(3) *The curvature tensor R is invariant under parallel transport*

$$\tau(\gamma, t_1, t_0)_* R_{\gamma(t_0)} = R_{\gamma(t_1)}$$

for all curves $\gamma : \mathbf{R} \to M$ and all $t_0, t_1 \in \mathbf{R}$.

Before proving the equivalence of (1), (2), (3) we note some immediate corollaries.

Corollary 141. *Let M and M' be locally symmetric spaces, $p \in M$, and $p' \in M'$. Then there is an isometry $\varphi : B^r(p) \to B'^r(p')$ with $\varphi(p) = p'$ if and only if there is an orthogonal linear isomorphism $b_0 : T_pM \to T_{p'}M'$ intertwining R_p and $R'_{p'}$: $b_{0*}R_p = R'_{p'}$.*

Corollary 142. *Let M and M' be simply connected symmetric spaces. Then there is an isometry $\varphi : M \to M'$ if and only if there is an orthogonal linear isomorphism $b_0 : T_pM \to T_pM'$ with $b_{0*}R_p = R'_{p'}$ for some $p \in M$ and $p' \in M'$.*

Corollary 143. *A complete, simply connected, locally symmetric space is symmetric.*

Corollary 144. *A connected, simply connected symmetric space M is* **homogeneous***; i.e. given $p, q \in M$ there exists an isometry $\varphi : M \to M$ with $\varphi(p) = q$.*

All the corollaries follow immediately from the appropriate version (local and global) of the C-A-H Theorem. For the third corollary it is helpful to note that $b_0 : T_pM \to T_pM$ given by $b_0 = -I$ intertwines R with itself: $b_0 * R_p = R_p$. For the last corollary note that a symmetric space is obviously locally symmetric so that one can take $b_0 = \tau(\gamma, 1, 0) : T_pM \to T_qM$ where γ is a curve from p to q.

Exercise 145. Show that the hypothesis that M be simply connected is not needed in the last corollary. (Hint: If $p, q \in M$ are sufficiently close consider the symmetry about the point $\gamma\left(\frac{1}{2}\right)$ where $\gamma : [0, 1] \to M$ is the unique minimal geodesic from p to q.)

We now prove the equivalence of (1), (2), (3) in the definition. The implication (3) $\implies$ (1) is immediate from the local C-A-H Theorem. For the rest we need the following.

Definition 146. *The* **covariant derivative of the curvature tensor** *is the field ∇R which assigns to each $p \in M$ the multilinear map*

$$(\nabla R)_p \in L(T_pM; L^2(T_pM; L(T_pM, T_pM)))$$

given by

$$(\nabla R)_p(v) = \frac{d}{dt}\tau(\gamma, 0, t)_* R_{\gamma(t)} \,|_{t=0}$$

where $\gamma : \mathbf{R} \to M$ is any curve such that

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

Note that

$$R_{\gamma(t)} \in L^2(T_{\gamma(t)}M; L(T_{\gamma(t)}, T_{\gamma(t)}M))$$

so that $\tau_* R$ is a curve in a "constant" vector space

$$\tau(\gamma, 0, t)_* R_{\gamma(t)} \in L^2(T_pM; L(T_pM; L(T_pM, T_pM)) \,.$$

Exercise 147. Justify this definition; i.e. show that $(\nabla R)_p(v)$ is independent of the choice of γ satisfying $\gamma(0) = p$ and $\dot{\gamma}(0) = v$. Hint: Consider the expression

$$\begin{aligned}(\nabla R)(X)(X_1, X_2)Y &= \nabla_X(R(X_1, X_2)Y) - R(\nabla_X X_1, X_2)Y \\ &\quad -R(X_1, \nabla_X X_2)Y - R(X_1, X_2)\nabla_X Y\end{aligned}$$

for vectorfields $X, X_1, X_2, Y \in \mathcal{X}(M)$.

Now (2) $\Longrightarrow$ (3) is immediate from the definition of ∇R; i.e. the condition that ∇R vanishes identically is equivalent to the condition that $\tau(\gamma, 0, t)_* R_{\gamma(t)}$ is independent of t which is equivalent to the condition that it equals its value for $t = 0$ and this is condition (3).

As for (1) $\Longrightarrow$ (2) we clearly have

$$(\varphi_*(\nabla R))_p = (\nabla R)_{\varphi(p)}$$

for any isometry $\varphi : M \to M$ and any $p \in M$. (This is by the generalized theorem eregium.) Take φ to be a symmetry about p so that $\varphi(p) = p$ and $D\varphi(p) = -I$. This becomes

$$-(\nabla R)_p = (\nabla R)_p$$

which entails $(\nabla R)_p = 0$ as required.

Remark 148. Note that

$$R_p \in L^3(T_pM; T_pM)$$

while

$$(\nabla R)_p \in L^4(T_pM; T_pM)\,.$$

Thus if $b_0 = -I : T_pM \to T_pM$ we have

$$b_{0*}R_p = R_p, \;\; b_{0*}(\nabla R)_p = -(\nabla R)_p.$$

The former equation was used to prove (3) $\Longrightarrow$ (1); the latter to prove (1) $\Longrightarrow$ (2).

Example 149. A flat manifold is locally symmetric.

Example 150. If $M_1 \subset E^{n_1}$ and $M_2 \subset E^{n_2}$ are (resp. locally) symmetric, so is $M = M_1 \times M_2 \subset E^n$ where $n = n_1 + n_2$.

Example 151. The flat tori of the previous section are symmetric (but not simply connected). This shows that the hypothesis of simply connectivity cannot be dropped in the second corollary.

Example 152. Below we define manifolds of constant curvature and show that they are locally symmetric. The simplest example, after a flat space, is a sphere. The symmetry φ of the sphere $x_0^2 + x_1^2 + \cdots + x_m^2 = r$ about the point $x = (r, 0, \ldots, 0)$ is given by

$$\varphi(x_0, x_1, \ldots, x_m) = (x_0, -x_1, \ldots, -x_m)\,.$$

Example 153. A compact two-dimensional manifold of constant negative curvature is locally symmetric (as its universal cover is symmetric) but not homogeneous (as closed geodesics of a given period are isolated) and hence not symmetric. This shows that the hypothesis that M be simply connected cannot be dropped in the third corollary.

Example 154. The simplest example of a symmetric space which is not of constant curvature is the orthogonal group $M = O(k)$. Its defining equation is:

$$O(k) = \{a \in \mathbf{R}^{k\times k} aa^* = I\}$$

where $\mathbf{R}^{k\times k}$ denotes the space of all k by k matrices, a^* denotes the transpose of a, and I is the identity matrix. Thus $n = k^2$ and I is the identity matrix. Thus $n = k^2$ and $m = k(k-1)/2$. The inner product in the ambient Euclidean space $\mathbf{R}^{k\times k}$ is given by:

$$\langle A, B\rangle = tr(AB^*)$$

for $A, B \in \mathbf{R}^k_k$; the symmetry φ about the point $p = I$ is given by:

$$\varphi(a) = a^{-1}\,.$$

(Exercise: Prove all this.)

Gaussian Curvature

In this section $M \subset E^n$ shall be a hypersurface; i.e. $n = m + 1$ where m is the dimension of M. We shall suppose that M is endowed with a unit normal; i.e. a smooth map $U : M \to \mathbf{E}^n$ with:

$$\|U(p)\| = 1$$

$$T_pM = U(p)^\perp$$

for $p \in M$ (where $u^\perp = \{v \in \mathbf{E}^n : \langle u, v\rangle = 0\}$). These conditions determine $U(p)$ up to a sign. We denote by $S = S^m$ the unit sphere in $\mathbf{E}^n$:

$$S = \{u : \mathbf{E}^n : \|u\| = 1\}\,.$$

Note that the unit normal can be viewed as a map:

$$U : M \to S\,.$$

When so viewed it is called the **Gauss map**. If $\gamma : \mathbf{R} \to S$ then $\|\gamma(t)\|^2 = 1$ so $\langle\gamma(t), \dot\gamma(t)\rangle = 0$ which shows that

$$T_uS = u^\perp$$

for $u \in S$. Hence for $p \in M$

$$T_pM = T_{U(p)}S$$

so that the derivative

$$DU(p) : T_pM \to T_{U(p)}S$$

of the Gauss map is a linear map from a vector space to itself. Hence we make the following

Definition 155. *The* **Gaussian curvature** *of the hypersurface M is the real-valued function $K : M \to \mathbf{R}$ defined by*

$$K(p) = \det(DU(p))$$

for $p \in M$. (Note that replacing U by $-U$ has the effect of replacing K by $(-1)^m K$ so that K is independent of the choice of the unit normal U when m is even.)

Remark 156. Given a subset $B \subset M$ the set $U(B) \subset S$ is often called the **spherical image** of B. If U is a diffeomorphism on a neighborhood of B the change of variables formula for an integral gives

$$\int_{U(B)} \mu_S(du = \int_B |K(p)| \mu_M(dp)$$

where μ_M and μ_S denote the volume elements on M and S respectively. Introducing the notation

$$\mathrm{Area}_M(B) = \int_B \mu_M(dp)$$

we obtain immediately the formula

$$|K(p)| = \lim_{B \to p} \frac{\mathrm{Area}_S(U(B))}{\mathrm{Area}_M(B)}$$

which says that the curvature at p is roughly the ratio of the (m-dimensional) area of the spherical image $U(B)$ to the area of B where B is a very small open neighborhood of p in M. The sign of $K(p)$ is positive when U preserves orientation at p and negative when it reverses orientation.

We see that the Gaussian curvature is a natural generalization of Euler's curvature for a plane curve. Indeed for $m = 1$ and $n = 2$ we have

$$K = \frac{d\varphi}{ds}$$

where s is the arclength parameter and φ is the angle made by the normal (or the tangent) with some constant line.

Example 157. The Gaussian curvature of a sphere of radius r is r^{-m} (constant).

Example 158. Show that the Gaussian curvature of the surface $z = x^2 - y^2$ is $-4 < 0$ at the origin.

Remark 159. It is often quite easy to see that the Gauss curvature K vanishes. For example consider the developable surface defined in the last section. As the unit normal is constant along the lines of the ruling, the spherical image is one dimensional; hence $\mathrm{Area}_S(U(B)) = 0$ so $K = 0$. The following theorem relates K and R (and may be applied to give another proof that a developable surface is flat).

(highly curved)

(slightly curved)

(negatively curved)

Theorem 160. *Assume that M is a surface; i.e. that $m = 2$ and $n = 3$. Choose $p \in M$ and let $v, w \in T_pM$ be a basis. Then*

$$K(p) = \frac{-\langle R_p(v,w)v, w\rangle}{A(v,w)}$$

where $A(v,w)$ is the square of the area of the parallelogram spanned by v and w

$$A(v,w) = \|v\|^2\|w\|^2 - \langle v, w\rangle^2 .$$

Corollary 161. *(Theorema Egregium of Gauss) Gaussian curvature is intrinsic; i.e. if $\varphi : M \to M'$ is an isometry of surfaces then*

$$K(p) = K'(\varphi(p))$$

for $p \in M$.

Proof. Define $h_p(v_1, v_2) \in \mathbf{R}$ for $v_1, v_2 \in T_pM$ by

$$h_p(v_1, v_2) = \langle h_p(v_1)v_2, U(p)\rangle$$

so that

$$h_p(v_1)v_2 = h_p(v_1, v_2)U(p) .$$

Since $U(p)$ is a unit vector it is orthogonal to its derivative in any direction; whence by the Gauss-Weingarten equation

$$\begin{aligned} \langle DU(p)v_1, v_2\rangle &= -\langle h_p(v_1)^*U(p), v_2\rangle \\ &= -h_p(v_1, v_2) . \end{aligned}$$

Note also that

$$h_p(v_1, v_2) = h_p(v_2, v_1)$$

by the symmetry of the second fundamental form. Hence by the Gauss Codaizzi equations

$$\begin{aligned} -\langle R_p(v,w)v, w\rangle &= -\langle h_p(v)^*h_p(w)v - h_p(w)^*h_p(v)v, w\rangle \\ &= \langle h_p(v)v, h_p(w)w\rangle - \|h_p(v)w\|^2 \\ &= h_p(v,v)h_p(w,w) - h_p(v,w)^2 \\ &= \langle DU(p)v, v\rangle\langle DU(p), w, w\rangle \\ &\quad -\langle DU(p)v, w\rangle\langle DU(p)w, v\rangle \end{aligned}$$

so the theorem follows on reading T_pM for $\mathbf{R}^2$ and $DU(p)$ for Q in the following:

Lemma 162. *Let v, w be a basis for a Euclidean vector space $\mathbf{E}^2$ and let $Q : \mathbf{R}^2 \to \mathbf{E}^2$ be a linear transformation. Then*

$$\det(Q) = \frac{\langle Qv, v\rangle\langle Qw, w\rangle - \langle Qw, v\rangle\langle Qv, w\rangle}{\langle v, v\rangle\langle w, w\rangle - \langle w, v\rangle\langle v, w\rangle} .$$

Proof of lemma. If $C(v,w)$ denotes either the numerator or the denominator of the right then

$$C(v,w) = (ab - bc)^2 C(v', w')$$

whenever

$$\begin{aligned} v &= av' + bw' \\ w &= cv' + dw' . \end{aligned}$$

Hence the right-hand side is independent of the choice of basis so we may assume without loss of generality that v, w are orthonormal. But in that case we have

$$\begin{aligned} Qv &= \langle Qv, v\rangle v + \langle Qv, w\rangle w \\ Qw &= \langle Qw, v\rangle v + \langle Qw, w\rangle w \end{aligned}$$

so that the numerator is the determinant of the coefficients. As the denominator is one for an orthonormal basis the lemma is proved. □

□

Exercise 163. For $m = 1$, $n = 2$ the curvature is clearly *not* intrinsic as any two curves are locally isometric (parameterized by arclength). Show that the curvature $K(p)$ is intrinsic for even m while its absolute value $|K(p)|$ is intrinsic for odd $m \geq 3$. Hint: We still have the equation

$$-\langle R_p(v,w)v, w\rangle = \det \begin{bmatrix} \langle DU(p)v, v\rangle & \langle DU(p)v, w\rangle \\ \langle DU(p)w, v\rangle & \langle DU(p)w, w\rangle \end{bmatrix}$$

for $v, w \in T_pM$. Thus the 2 by 2 minors of the matrix

$$(\langle DU(p)v_i, v_j\rangle)_{i,j}$$

(where $v_1, \ldots, v_m$ form an orthonormal basis for T_pM) are intrinsic. Thus everything reduces to the following

Lemma 164. *The determinant of an m by m matrix is an expression in its 2 by 2 minors if m is even; the absolute value of the determinant is an expression in the 2 by 2 minors if m is odd and ≥ 3. The lemma is proved by induction on m. For the absolute value, note the formula*

$$\det(A)^m = \det(\det(A)I) = \det(AB) = \det(A)\det(B)$$

for an m by m matrix A where B is the transposed matrix of cofactors.

Spaces of Constant Curvature

In the last section we saw that the Gaussian curvature of a two-dimensional surface is intrinsic: we gave a formula for it in terms of the Riemann curvature tensor and the first fundamental form. We may use this formula to define the Gauss curvature for *any* two-dimensional manifold (even if its co-dimension is greater than one). We make a slightly more general definition.

Definition 165. *Let p be a point of the manifold M and $\mathbf{F} \subset T_pM$ a two-dimensional vector subspace of the tangent space. The* **sectional curvature** *of M at $(p, \mathbf{R})$ is the number*

$$K(p, \mathbf{F}) = \frac{\langle R_p(v,w)v, w\rangle}{A_p(v,w,v,w)}$$

where $\mathbf{F} =$ span(v, w) and for $v_1, v_2, v_3, v_4 \in T_pM$

$$A_p(v_1, v_2, v_3, v_4) = \langle v_1, v_4\rangle\langle v_2, v_3\rangle - \langle v_1, v_3\rangle\langle v_2, v_4\rangle\,.$$

(By an argument in the previous section, $K(p, \mathbf{F})$ is independent of the choice of the basis v, w of $\mathbf{F}$ and is equal to the Gauss curvature when M has dimension two so that $\mathbf{F} = T_pM$.)

Exercise 166. Suppose $r > 0$ is sufficiently small that the restriction of Exp_p to $V^r(p) = \{v \in T_pM : \|v\| < r\}$ is a diffeomorphism onto $B^r(p)$ and let N be the two-dimensional manifold given by

$$N = \mathrm{Exp}_p(\mathbf{F} \cap V^r(p))\,.$$

Show that the sectional curvature $K(p, \mathbf{F})$ of M at $(p, \mathbf{F})$ is the same as the Gauss curvature of N at p (i.e. the sectional curvature of N at $(p, \mathbf{F}) = (p, T_pN)$).

Exercise 167. Let $p \in M \subset E^n$ and let $\mathbf{F} \subset T_pM$ be a two-dimensional vector subspace. For $r > 0$ let L denote the ball of radius r in the $(n - m + 2)$ dimensional affine subspace of E^n through p and parallel to the vector subspace $\mathbf{F} + T_p^\perp M$:

$$L = \{p + v + u : v \in \mathbf{F}, u \in T_p^\perp M, \|v\|^2 + \|u\|^2 < r^2\}\,.$$

Show that for r sufficiently small, $L \cap M$ is a two-dimensional manifold with Gauss curvature $K_{L\cap M}(p)$ at p given by

$$K_{L\cap M}(p) = K(p, F)\,.$$

Definition 168. *A* **space of constant curvature** $k \in \mathbf{R}$ *is a manifold M which satisfies the two equivalent conditions*

(1) for each $p \in M$ and each two-dimensional subspace $F \subset T_pM$:

$$K(p, \mathbf{F}) = k;$$

(2) for each $p \in M$ and all $v_1, v_2, v_3, v_4 \in T_pM$:

$$\langle R_p(v_1, v_2,)v_3, v_4\rangle = kA_p(v_1, v_2, v_3, v_4)\,.$$

We prove the conditions are equivalent. For (2) $\Longrightarrow$ (1) simply take $v_1 = v_3 = v$ and $v_2 = v_4 = w$ and use the definitions. For (1) $\Longrightarrow$ (2) we must "polarize". Note that both sides of (2) have form $cQ(v, w, v, w)$ where

$$Q(v_1, v_2, v_3, v_4) = \langle h(v_1, v_4), h(v_2, v_3)\rangle - \langle h(v_1, v_3), h(v_2, v_4)\rangle$$

where h is a symmetric bilinear map from T_pM to an inner produce space $\mathbf{E}$:

$$h : T_pM \times T_pM \to \mathbf{E}, \quad h(v, w) = h(w, v).$$

(For the left side take $\mathbf{E} = T_p^{\perp}M$ and $h =$ second fundamental form and use the Gauss-Codaizzi equations; for the right side take $\mathbf{E} = \mathbf{R}$ and $h(v, w) = \langle v, w\rangle$.) Hence both sides of (2) exhibit the following symmetries:

(i) $Q(v_1, v_2, v_3, v_4) + Q(v_2, v_1, v_3, v_4) = 0$;

(ii) $Q(v_1, v_2, v_3, v_4) + Q(v_1, v_2, v_4, v_3) = 0$;

(iii) $Q(v_1, v_2, v_3, v_4) + Q(v_2, v_3, v_1, v_4) + Q(v_3, v_1, v_2, v_4) = 0$.

The difference of the two sides of (2) also has these symmetries and by polarization of (1) satisfies the additional symmetry:

(iv) $Q(v_1, v_2, v_3, v_4) + Q(v_3, v_2, v_1, v_4) + Q(v_1, v_4, v_3, v_2) + Q(v_3, v_4, v_1, v_2) = 0$.

Thus we must show that (i)-(iv) imply that $Q = 0$.

In (iii) interchange v_4 with v_1, v_2, and v_3 successively and add the four resulting equations. Use (i) and (ii) and divide by 2 to obtain

$$Q(v_1, v_4, v_3, v_2) + Q(v_4, v_2, v_3, v_1) + Q(v_4, v_3, v_1, v_2) = 0\,.$$

By the permutation of (iii):

$$q(v_3, v_1, v_4, v_2) + Q(v_1, v_4, v_3, v_2) + Q(v_4, v_3, v_1, v_2) = 0\,.$$

Hence

(v) $q(v_4, v_2, v_3, v_1) = Q(v_3, v_1, v_4, v_2)$.

Use (v) twice in (iv) and divide by two

$$Q(v_1, v_2, v_3, v_4) + Q(v_3, v_2, v_1, v_4) = 0\,.$$

This equation together with (i) shows that $Q(v_1, v_2, v_3, v_4)$ is skew-symmetric in (v_1, v_2, v_3) so it must vanish by (iii) as required.

Remark 169. The symmetric group S_4 on four symbols acts naturally on $L^4(T_pM; \mathbf{R})$ and conditions (i)-(iv) say that the four elements

$$\begin{aligned} a &= (\;) + (12) \\ b &= (\;) + (34) \\ c &= (\;) + (123) + (132) \\ d &= (\;) + (13) + (24) + (13)(24) \end{aligned}$$

of the group ring annihilate Q. This subbests an alternate proof of (1) $\implies$ (2). A representation of a finite group is completely reducible so one can prove that $Q = 0$ by showing that any vector in any irreducible representation of S_4 which is annihilated by the four elements a, b, c and d must necessarily be zero. This can be checked case by case for each irreducible representation. (The group S_4 has 5 irreducible representations: two of dimension 1, two of dimension 3, and one of dimension 2.)

Clearly any isometry $b : T_pM \to T_{p'}M'$ intertwines A_p and $A'_{p'}$ as they are defined from the first fundamental forms g_p and $g'_{p'}$ which are intertwined by b. Hence by the appropriate version (local or global) of the C-A-H theorem we have the following corollaries of the definition.

Corollary 170. *Let M and M' be spaces of the same constant curvature k and let $p \in M$ and $p' \in M'$. Then there is an isometry $\varphi : B^r(p) \to {B'}^r(p')$ for sufficiently small $r > 0$.*

Corollary 171. *Any two connected, simply connected and complete spaces of the same constant curvature are isometric.*

Corollary 172. *Let M be a space of constant curvature k, $p \in M$ and M' a manifold. Then M' is of constant curvature k if and only if for every $p' \in M'$ and every orthogonal isomorphism $b_0 : T_pM \to T_{p'}M'$ there is an isometry $\varphi : B^r(p) \to {B'}^r(p')$ with*

$$\varphi(p) = p', \quad D\varphi(p) = b_0.$$

Corollary 173. *A connected, simply connected, complete manifold M is of constant curvature iff its isometry group acts transitively on its orthonormal frame bundle $O(M)$; i.e. if and only if for any $p, p' \in M$ and each orthogonal isomorphism $b_0 : T_pM \to T_{p'}M$ there exists a (necessarily unique) isometry $\varphi : M \to M$ with*

$$\varphi(p) = p', \quad D\varphi(p) = b_0.$$

(To prove the converses of the last two corollaries make the observation that

$$K'(p', \mathbf{F}') = K(p, \mathbf{F})$$

if φ is an isometry with $\varphi(p) = p'$ and $D\varphi(p)\mathbf{F} = \mathbf{F}'$. Also observe that given $\mathbf{F}$ and $\mathbf{F}'$ we can always find b_0 with $b_0\mathbf{F} = \mathbf{F}'$. Then use (1) in the definition.)

Example 174. Any flat space is of constant curvature $k = 0$.

Example 175. Euclidean space E^m is the unique (up to isometry) connected, simply connected, and complete space of constant curvature $k = 0$. A model of E^m is

$$E^m = \mathbf{R}^m$$

the space of m-tuples of real numbers with first fundamental form

$$g_p(v, w) = \langle v, w \rangle$$

for $p \in E^m$; $v, w \in T_pE^n = \mathbf{R}^m$ where

$$\langle v, w\rangle = v^1w^1 + v^2w^2 + \cdots + v^mw^m\,.$$

The isometry group is

$$I(E^m) = R(m) = O(m) \times \mathbf{R}^m$$

the group of rigid motions and geodesics are straight lines

$$\gamma(t) = p + tv$$

for $p \in E^m$, $v \in T_pE^m$, $t \in \mathbf{R}$.

Example 176. The unit sphere S^m is the unique (up to isometry) connected, simply connected, and complete space of constant curvature $k = 0$. A model of S^m is

$$S^m = \{p \in \mathbf{R}^{m+1} : \langle p, p\rangle = 1\}$$

with first fundamental form

$$g_p(v, w) = \langle v, w\rangle$$

for $p \in S^m$; $v, w \in T_pS^m$ where

$$\langle v, w\rangle = v^0w^0 + v^1w^1 + \cdots + v^mw^m\,.$$

The isometry group

$$I(S^m) = O(m + 1)$$

is the orthogonal group

$$P(m + 1) = \{a \in GL(m + 1) : \langle av, aw\rangle = \langle v, w\rangle \forall v, w \in \mathbf{R}^{m+1}\}$$

acting by restriction: $a \mid S^m$ is an isometry of S^m for $a \in O(m + 1)$. The geodesics are given by

$$\gamma(t) = \cos(t)p + \sin(t)v$$

for $p \in S^m$; $v \in T_pS^m$ with $\langle v, v\rangle = 1$ and $t \in \mathbf{R}$.

The proofs of these assertions are left to the reader, but note that the unit normal to S^m is given by

$$U(p) = p$$

so that

$$T_pS^m = \{v \in \mathbf{R}^{m+1} : \langle p, v\rangle = 0\}\,.$$

Hence if $v_1, \ldots, v_m$ is an orthonormal bases of T_pS^m then $p, v_1, \ldots, v_m$ is an orthonormal basis of $\mathbf{R}^{m+1}$. This proves that $I(S^m) = O(m + 1)$ and that $I(S^m)$ acts transitively on $O(S^m)$ the orthonormal frame bundle.

Exercise 177. Show that a product of spheres is *not* a space of constant curvature (but it *is* a symmetric space).

Example 178. Hyperbolic space H^m is the unique (up to isometry) connected, simply connected, and complete space of constant curvature $k = -1$. A model of H^m is

$$H^m = \{p \in \mathbf{R}^{m+1} : Q(p,p) = -1, p_0 < 0\}$$

with first fundamental form

$$g_p(v,w) = Q(v,w)$$

for $p \in H^m$; $v, w \in T_pH^m$ where

$$Q(v,w) = -v^0w^0 + v^1w^1 + \cdots + v^mw^m \, .$$

The isometry group

$$I(H^m) \subseteq O(m,1)$$

is the subgroup of the pseudo-orthogonal group

$$O(m,1) = \{a \in GL(m+1) : Q(av,aw) = Q(v,w) \forall v,w \in \mathbf{R}^{m+1}\}$$

which preserves H^m:

$$I(H^m) = \{a \mid H^m : a \in O(m,1); a(H^m) = H^m\} \, .$$

The graph of $Q(p,p) = -1$ has two components one of which is H^m; an element of $O(m,1)$ can interchange them. The geodesics of H^m are given by

$$\gamma(t) = \cosh(t)p + \sinh(t)v$$

for $p \in H^m$, $v \in T_pH^m$ with $Q(v,v) = 1$, and $t \in \mathbf{R}$.

We remark that the manifold H^m does not quite fit into the framework of this chapter as it is not exhibited as a submanifold of Euclidean space but rather of "pseudo-Euclidean space": the positive definite innter produce $\langle v,w\rangle$ of the ambient space $\mathbf{R}^{m+1}$ is replaced by a non-degenerate symmetric bilinear form $Q(v,w)$. All the theory developed thus far goes through (reading $Q(v,w)$ for $\langle w,w\rangle$) provided we make the additional hypothesis (true in the example $M = H^m$) that the first fundamental form $g_p = Q \mid T_pM$ is positive definite. For then $Q \mid T_pM$ is nondegenerate and we may define orthogonal projection $\Pi(p)$ onto T_pM by the following:

Lemma 179. *Let Q be a symmetric bilinear form on a vector space $\mathbf{E}$ and for each subspace $\mathbf{F}$ of $\mathbf{E}$ define its orthogonal complement by*

$$\mathbf{F}^\perp = \{u \in \mathbf{E} : Q(v,u) = 0 \; \forall v \in \mathbf{F}\} \, .$$

Assume Q is non-degenerate

$$\mathbf{E}^\perp = \{0\} \, .$$

Then $\mathbf{F}^\perp$ is a vectorspace complement to $\mathbf{F}$

$$\mathbf{E} = \mathbf{F} \oplus \mathbf{F}^\perp$$

if and only if $Q \mid \mathbf{F}$ is non-degenerate

$$\mathbf{F} \cap \mathbf{F}^\perp = \{0\} \, .$$

The proofs of the various properties of H^m are entirely analogous to the corresponding proofs for S^m. Thus the unit normal field to H^m is given by

$$U(p) = p$$

for $p \in H^m$ although the "square of its length" is $Q(p,p) = -1$. We have

$$T_pH^m = \{v \in \mathbf{R}^{m+1} : Q(v,p) = 0\}$$

so that orthonormal projection $\Pi(p) : \mathbf{R}^{m+1} \to T_pH^m$ is given by

$$\Pi(p)w = w + Q(w,p)p\,.$$

If $v_1, \ldots, v_m$ is an othronormal basis for T_pH^m then $p, v_1, \ldots, v_m$ is an orthonormal basis for $\mathbf{R}^{m+1}$ (with respect to Q) so $O(m,1)$ acts transitively on the orthonormal frame bundle $O(H^m)$. For the curve γ we have

$$\ddot{\gamma}(t) = \gamma(t), \quad \gamma(0) = p, \;\; \dot{\gamma}(0) = v, \quad Q(\dot{\gamma}(t), \dot{\gamma}(t)) = -1$$

so that it is a geodesic as required. Hence the second fundamental form is given by

$$h_p(v)v = Q(v,v)p$$

(as both sides are $\gamma(0)$) so

$$g_p(R_p(v,w)v,w) = \{Q(w,v)^2 - Q(w,w)Q(v,v)\}Q(p,p)$$

by (the analog of) Gauss-Weingarten which, as $Q(p,p) = -1$, shows that $k = -1$ as required.

Covariant Differentiation of Tensors

The formulas of §5 show how to define covariant differentiation of tensor fields more general than vectorfields and normal fields. Given such a field S along γ, $S(t)$ will lie in the tensor space associated to the point $\gamma(t)$ and so may be "parallel translated" to the point $\gamma(t_0)$. This gives a curve

$$t \to \tau(\gamma, t_0, t)_* S(t)$$

in a fixed vector space. The derivative of this curve at $t = t_0$ will be called the covariant derivative of S:

$$\nabla S(t_0) = \frac{d}{dt}\tau(\gamma, t_0, t)_* S(t)\,|_{t=t_0}\;\;.$$

Note that the Leibnitz product rule

$$\nabla(S_1 \otimes S_2) = (\nabla S_1) \otimes S_2 + s_1 \otimes (\nabla S_2)$$

follows immediately; some authors simply define covariant differentiation of tensors to be the (unique) operation which agrees with ordinary differentiation on functions, covariant differentiation on vectorfields, and satisfies the product rule (for tensor product and contraction).

Remark 180. Note that

$$S(t) = \tau(\gamma, t, 0)_* S(0)$$

for all t if and only if

$$\nabla S \equiv 0\,.$$

To see this calculate

$$\begin{aligned}\frac{d}{dt}\tau(\gamma, t, 0)_* S(t)\,|_{t=t_0} &= \frac{d}{dt}\tau(\gamma, 0, t_0)_*(\gamma, t, t_0)_* S(t)\,|_{t=t}\ , \\ &= \tau(\gamma, 0, t_0)_* \frac{d}{dt}\tau(\gamma, t_0, t)_* S(t)\,|_{t=t_0} \\ &= \tau(\gamma, 0, t_0)_*(\nabla S)(t_0)\,.\end{aligned}$$

Hence the curve

$$t \to \tau(\gamma, 0, t)_* S(t)$$

(which takes values in the tensor space associated to $\gamma(0)$) has vanishing derivative iff ∇S vanishes identically.

To illustrate consider a field g which assigns to each point $p \in M$ a bilinear map g_p on the tangent space at p

$$g_p \in L^2(T_pM, \mathbf{R})\,.$$

Exercise 181. Given a tensor field $S \in \mathcal{T}_s^r(M)$ there is a unique tensor field $\nabla S \in \mathcal{T}_s^{r+1}(M)$ such that for any $p \in M$, $v \in T_pM$, and $\gamma : \mathbf{R} \to M$ with $\gamma(0) = p$, $\gamma(0) = v$ we have

$$(\nabla S)_p(v) = \frac{d}{dt}\tau(\gamma, 0, t) S_{\gamma(t)}\,|_{t=0}\ .$$

Such a field is by definition a tensor field of type $(0, 2)$.) Composition with the curve γ gives

$$g_{\gamma(t)} \in L^2(T_{\gamma(t)}M, \mathbf{R})\,.$$

The linear transformation

$$\tau = \tau(\gamma, t_0, t) : T_{\gamma(t)}M \to T_{\gamma(t_0)}M$$

acts on $g_{\gamma(t)}$ via

$$(\tau_* g_{\gamma(t)})(v, w) = g_{\gamma(t)}(\tau^{-1}v, \tau^{-1}w)$$

for $v, w \in T_{\gamma(t_0)}M$; thus

$$\tau(\gamma, t_0, t)_* g_{\gamma(t)} \in L^2(T_{\gamma(t_0)}M, \mathbf{R})\,.$$

Differentiating the identity

$$g_{\gamma(t)}(X(t), Y(t)) = (\tau_* g_{\gamma(t)})(\tau X(t), \tau Y(t))$$

at $t = t_0$ gives the Leibnitz product formula

$$\frac{d}{dt} g(X,Y) = (\nabla g)(X,Y) + g(\nabla X, Y) + g(X, \nabla Y)$$

for $X, Y \in \mathcal{X}(\gamma)$. In the alternate notation

$$D_Z(g(X,Y)) = (\nabla_Z g)(X,Y) + g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

for $X, Y, Z \in \mathcal{X}(M)$.

Now consider the tensor g given by

$$g_p(v, w) = \langle v, w \rangle$$

for $p \in M$; $v, w \in T_p M$. This tensor is called the **first fundamental form**. Then the fact that parallel transport τ is an orthogonal transformation can be expressed compactly in the equation $\nabla g = 0$. In summary we have

Proposition 182. *The first fundamental form is parallel along any curve.*

Since we can also parallel transport normal fields we can also covariantly differentiate tensor fields which involve both tangential and normal fields. For example, consider a field α which, like the second fundamental form h, assigns to each p a bilinear map from the tangent space at p to the normal space at p:

$$\alpha_p \in L_p = L(T_p M, L(T_p M, T_p^{\perp} M)) \,.$$

The linear transformations

$$\begin{aligned} \tau &= \tau(\gamma, t_0, t) : T_{\gamma(t)} M \to T_{\gamma(t_0)} M \\ \tau^{\perp} &= \tau^{\perp}(\gamma, t_0, t) : T^{\perp}_{\gamma(t)} \to M T^{\perp}_{\gamma(t_0)} M \end{aligned}$$

induce a transformation

$$\tau^{\#} = \tau^{\#}(\gamma, t, t_0) : L_{\gamma(t)} \to L_{\gamma(t_0)}$$

defined by

$$(\tau^{\#} \alpha_{\gamma(t)})(v) w = \tau^{\perp} \cdot \alpha_{\gamma(t)}(\tau^{-1} v) \tau^{-1} w$$

for $v, w \in T_{\gamma(t_0)} M$. Define the covariant derivative $\nabla^{\#} \alpha$ along γ by

$$(\nabla^{\#} \alpha)(t_0) = \frac{d}{dt} \tau^{\#} \alpha_{\gamma(t)} \,|_t = t_0 \,.$$

Differentiating the identity

$$\tau^{\perp}(\alpha(X)Y) = (\tau^{\#} \alpha)((\tau X))(\tau Y)$$

gives the formula

$$\nabla^{\perp}(\alpha(X)Y) = (\nabla^{\#} \alpha)(X) Y + \alpha(\nabla X) Y + \alpha(X) \nabla Y$$

for $X, Y \in \mathcal{X}(\gamma)$ or equivalently

$$\nabla^{\perp}_Z(\alpha(X)Y) = (\nabla^{\#}_Z \alpha)(\nabla_Z X) Y + \alpha(X) \nabla_Z Y$$

for $X, Y, Z \in \mathcal{X}(M)$.