Math 396. Covariant derivative, parallel transport, and General Relativity

1. Motivation

Let M be a smooth manifold with corners, and let (E, ∇) be a C^∞ vector bundle with connection over M. Let $\gamma : I \to M$ be a smooth map from a nontrivial interval to M (a "path" in M); keep in mind that γ may not be injective and that its velocity may be zero at a rather arbitrary closed subset of I (so we cannot necessarily extend the standard coordinate on I near each $t_0 \in I$ to part of a local coordinate system on M near $\gamma(t_0)$). In pseudo-Riemannian geometry $E = TM$ and ∇ is a specific connection arising from the metric tensor (the Levi-Civita connection; see §4).

A very fundamental concept is that of a (smooth) section along γ for a vector bundle on M. Before we give the official definition, we consider an example.

Example 1.1. To each $t_0 \in I$ there is associated a velocity vector

$$\gamma'(t_0) = \mathrm{d}\gamma(t_0)(\partial_t|_{t_0}) \in \mathrm{T}_{\gamma(t_0)}(M) = (\gamma^*(TM))(t_0).$$

Hence, we get a set-theoretic section of the pullback bundle $\gamma^*(TM) \to I$ by assigning to each time t_0 the velocity vector $\gamma'(t_0)$ at that time. This is not just a set-theoretic section, but a smooth section.

Indeed, this problem is local, so pick $t_0 \in I$ and an open $U \subseteq M$ containing $\gamma(J)$ for an open neighborhood $J \subseteq I$ around t_0, with J and U so small that U admits a C^∞ coordinate system $\{x_1, \dots, x_n\}$. Let $\gamma_i = x_i \circ \gamma|_J$; these are smooth functions on J since γ is a smooth map from I into M. By the usual rules for computing velocity vectors (i.e., the Chain Rule), for $t \in J$ we have

$$\gamma'(t) = \sum_i \gamma_i'(t)\partial_{x_i}|_{\gamma(t)} = \sum_i \gamma_i'(t)(\gamma^*(\partial_{x_i}))(t)$$

in $\mathrm{T}_{\gamma(t)}(M) = (\gamma^*(TM))(t)$. In other words, as set-theoretic sections of $\gamma^*(TM)$ over J we have $\gamma' = \sum_i \gamma_i' \cdot \gamma^*(\partial_{x_i})|_J$. Since the $\gamma^*(\partial_{x_i})$'s constitute a local frame for the bundle $\gamma^*(TM)$ over J (as the ∂_{x_i}'s are a local frame for the bundle TM over U), the desired smoothness over J is exactly the fact that the coefficient functions γ_i' with respect to this frame are smooth functions on J.

Definition 1.2. If $E \to M$ is a vector bundle, a (smooth) *section of E along* γ is an element of the space $(\gamma^*(E))(I)$ of smooth global sections of the pullback bundle $\gamma^*(E) \to I$.

By the universal property of pullback bundles, we can identify smooth sections $s : I \to \gamma^*(E)$ with smooth maps of manifolds $\widehat{s} : I \to E$ fitting into a commutative diagram

(1.1)
$$\begin{tikzcd} & E \arrow[d] \\ I \arrow[ur, "\widehat{s}"] \arrow[r, "\gamma"'] & M \end{tikzcd}$$

where the right vertical map is the structure map of the bundle over M.

Remark 1.3. In the classical literature with $E = TM$ one sees the notion of "vector field along a parametric curve $\gamma : I \to M$": a family $\{\vec{v}(t)\}_{t \in I}$ of tangent vectors $\vec{v}(t) \in \mathrm{T}_{\gamma(t)}(M)$ such that for local coordinates $\{x_i\}$ on M near any $\gamma(t_0)$ the local basis expansion $\vec{v}(t) = \sum v_i(t)\partial_{x_i}|_{\gamma(t)}$ for t near t_0 has smooth coefficient functions v_i. This is exactly a concrete description of the preceding definition. In natural examples the path may cross itself, stop at some time, or change direction (i.e., γ may be non-injective or not a local immersion), so time along γ may fail to be part of a local coordinate system on M around points of $\gamma(I)$. By working with the bundle $\gamma^*(TM)$ over I, or with diagrams like (1.1), the apparent complications of working "in M" with vector fields along a self-intersecting or non-immersive γ are eliminated.

Consider our bundle with connection (E, ∇) over M. For any path $\gamma : I \to M$ and any $t_0 \in I$, there is a distinguished class of sections of E along γ, namely the ones that are flat with respect to the pullback connection $\gamma^*(\nabla)$ on $\gamma^*(E)$. More precisely, recall from our study of connections on bundles over I that for each element of a fiber there is a unique global section extending that element such that the global section is flat for the connection. In our case, for each $t_0 \in I$ and $s_0 \in E(\gamma(t_0)) = (\gamma^* E)(t_0)$ there exists a unique $\gamma^*(\nabla)$-flat section $\widetilde{s}_0 \in (\gamma^*(E))(I)$ such that $\widetilde{s}_0(t_0) = s_0$. Specializing at any $t_1 \in I$ then defines a linear isomorphism

$$P_{t_1,t_0,\gamma} : E(\gamma(t_0)) = (\gamma^*(E))(t_0) \simeq (\gamma^*(E))(t_1) = E(\gamma(t_1))$$

via $s_0 \mapsto \widetilde{s}_0(t_1)$. This is called *parallel transport*. (Of course, this concept depends heavily on the particular connection being used on E and on the particular path γ linking $m_0 = \gamma(t_0)$ to $m_1 = \gamma(t_1)$.) Observe that $P_{t_0,t_0,\gamma}$ is the identity (why?), and since a global flat section over I is uniquely determined by its value in one fiber it follows that for any $t_0, t_1, t_2 \in I$ we have the transitivity law $P_{t_2,t_1,\gamma} \circ P_{t_1,t_0,\gamma} = P_{t_2,t_0,\gamma}$. For example, $P_{t_0,t_1,\gamma} = P_{t_1,t_0,\gamma}^{-1}$. This is all "physically obvious".

Example 1.4. Let us give the example that explains the reason we use the word "parallel". Let M be an open subset in a finite-dimensional vector space, so there is a canonical trivialization $TM \simeq M \times V$. There is a unique connection ∇ on $E = TM$ for which the locally constant vector fields are the flat sections. Indeed, pick an ordered basis $\{e_1, \dots, e_n\}$ of V and let $\{x_1, \dots, x_n\}$ be the dual basis of $V^\vee$ and $\underline{e}_1, \dots, \underline{e}_n$ the associated frame of constant vector fields over M (such that $\underline{e}_i(m) = e_i$ under the canonical isomorphism $\mathrm{T}_m(M) \simeq V$ for all $m \in M$). A smooth vector field over an open $U \subseteq M$ has the unique form $\vec{v} = \sum a_i \underline{e}_i$ for smooth functions a_i on U, and we define

$$\nabla(\vec{v}) = \sum \mathrm{d}a_i \otimes \underline{e}_i \in (T^*M \otimes TM)(U).$$

One readily checks that this is a connection on TM, and that the flat sections $\vec{v}$ over an open are those vector fields for which all $\mathrm{d}a_i$ vanish, which is to say that $\vec{v}$ has locally constant coefficients with respect to the frame $\{\underline{e}_i\}$. By the Leibnitz Rule and the existence of a global frame consisting of constant vector fields, this is the only possible connection that kills the constant vector fields, and so we have both uniqueness and existence.

Now with respect to this canonical connection ∇, what is the parallel transport isomorphism $P_{t_1,t_0,\gamma}$ along a path $\gamma : I \to M$? Put another way, what are the $\gamma^*(\nabla)$-flat sections of $\gamma^*(TM)$? Since $\nabla(\underline{e}_i) = 0$ for all i, by the characterization of pullback connections we have $(\gamma^*(\nabla))(\gamma^*(\underline{e}_i)) = 0$ for all i. Thus, $\{\gamma^*(\underline{e}_i)\}$ is a global frame of flat sections along γ and a vector field $\sum a_i \gamma^*(\underline{e}_i)$ is $\gamma^*(\nabla)$-flat if and only if $a_i' = 0$ on I for all i. Hence, over the connected I all functions a_i must be constant, so the flat sections of $\gamma^*(TM)$ are precisely the vector fields $\gamma^*(\underline{v})$ along γ, where $v \in V$ and $\underline{v}$ is the associated constant vector field on M. That is, its value in each fiber $\mathrm{T}_{\gamma(t)}(M) \simeq V$ is v. Since we visualize the canonical isomorphisms $\mathrm{T}_m(M) \simeq V$ as corresponding to "parallel translation" from the origin to m, the picture of $\gamma^*(\underline{v})$ is as the collection of parallel translates of v with "initial endpoint" moving along γ. It looks like something that deserves to be called parallel transport along γ!

In practice it is convenient to permit time reparameterization of our path, which is to say that we precompose γ with a smooth isomorphism between intervals in $\mathbf{R}$. Fortunately, such time reparameterization does not affect parallel transport:

Lemma 1.5. *With notation and hypotheses as above, let $\gamma : I \to M$ be a smooth path and let $\varphi : J \simeq I$ be a C^∞ isomorphism from an interval $J \subseteq \mathbf{R}$. Let $\tau_i = \gamma^{-1}(t_i)$ for $t_0, t_1 \in I$ and let $m_i = \gamma(t_i)$. The parallel transport isomorphisms*

$$P_{t_1,t_0,\gamma}, P_{\tau_1,\tau_0,\gamma\circ\varphi} : E(m_0) \rightrightarrows E(m_1)$$

coincide.

Proof. Since pullback (for bundles, connections, and sections of bundles) is compatible with composition of smooth maps, we may replace (M, E, ∇) with $(I, \gamma^*(E), \gamma^*(\nabla))$ to reduce to the special case that $M = I$ is a nontrivial interval over which we have a bundle with connection (E, ∇). By the very definition of parallel transport in terms of flat global sections across an entire interval (and its specialization at pairs of points), we are thereby reduced to showing that if $s \in E(I)$ is a ∇-flat section then for any smooth isomorphism $\varphi : J \simeq I$ the pullback section $\varphi^*(s) \in (\varphi^*E)(J)$ is $\varphi^*(\nabla)$-flat. But this is immediate from the local identities that uniquely characterize the pullback connection: in our situation, for any $\tau \in J$ and $t := \varphi(\tau) \in I$ we have

$$((\varphi^*(\nabla))(\varphi^*(s)))(\tau) = (\mathrm{d}\varphi(\tau)^\vee \otimes 1)((\nabla(s))(t)) = 0$$

since $\nabla(s) = 0$. Hence, $\varphi^*(s)$ is indeed $\varphi^*(\nabla)$-flat since τ was arbitrary. ∎

In Example 1.4 there is a very special property: if we let $m_0 = \gamma(t_0)$ and $m_1 = \gamma(t_1)$ then $P_{t_1,t_0,\gamma} : E(m_0) \simeq E(m_1)$ only depends on m_0 and m_1, *not* on γ. Indeed, in that example we shows that $P_{t_1,t_0,\gamma}$ is the composite $E(m_0) = \mathrm{T}_{m_0}(M) \simeq V \simeq \mathrm{T}_{m_1}(M) = E(m_1)$ of isomorphisms that have nothing to do with γ. In general, if $\gamma, \widehat{\gamma} : I \rightrightarrows M$ are two paths (over a common time interval) in a smooth manifold with corners M and if $\widehat{\gamma}(t_0) = \gamma(t_0) = m_0$ and $\widehat{\gamma}(t_1) = \gamma(t_1) = m_1$, then for a bundle with connection (E, ∇) over M the two resulting parallel transport isomorphisms

$$P_{t_1,t_0,\gamma}, P_{t_1,t_0,\widehat{\gamma}} : E(m_0) \simeq E(m_1)$$

are usually *very* different. An important special case is when $t_0 \neq t_1$ but $m_0 = m_1$ and $\widehat{\gamma}$ is the constant map ($t \mapsto m_0$ for all t), in which case $P_{t_1,t_0,\widehat{\gamma}}$ is the identity on $E(m_0)$ (why?) but $P_{t_1,t_0,\gamma} : E(m_0) \simeq E(m_0)$ may not be the identity. That is, parallel transport in E with respect to ∇ along the windy path γ from m_0 back to itself through the time interval between t_0 and t_1 may be a nontrivial automorphism of $E(m_0)$. Such a phenomenon is called *nontrivial holonomy*, and if we restrict attention to paths $\gamma : [0, 1] \to M$ with $\gamma(0) = \gamma(1) = m_0$ then the image of the map $\gamma \mapsto P_{1,0,\gamma} \in \mathrm{Aut}(E(m_0))$ sending "(piecewise) smooth loops" at m_0 to parallel transport automorphisms of $E(m_0)$ is a subgroup of $\mathrm{Aut}(E(m_0))$ called the *holonomy group* of (E, ∇) at m_0. (So the point is that the holonomy group may be rather non-trivial.) Strictly speaking, to make this definition work we have to take care of the problem that a concatenation of smooth loops may be just piecewise smooth (e.g., one loop may end with a different velocity than that with which the next loop begins); Remark 1.7 addresses the important (but ultimately trivial) technical issue of allowing γ to be piecewise smooth.

Natural geometric examples of nontrivial holonomy with $E = TM$ are given on the unit sphere $S^2 \subseteq \mathbf{R}^3$ in §6, with γ any latitude circle (with constant-speed parameterization) away from the north and south poles. If p is a point on such a circle, then the circle can (in a sense that is physically obvious, and not hard to define rigorously) be "smoothly contracted" to p in S^2 without moving p. Thus, parallel transport $P_{t_1,t_0,\gamma} \in \mathrm{Isom}(E(m_0), E(m_1))$ is extremely sensitive to deformation of the path γ (keeping the positions $m_0 = \gamma(t_0)$ and $m_1 = \gamma(t_1)$ fixed). The absence of this sensitivity in Example 1.4 (with $E = TM$ and ∇ there actually determined by the metric tensor in a sense to be explained in Example 4.3) reflects the flat nature of Euclidean geometry, as will be made a bit more precise in the discussion immediately following Theorem 3.11.

By the definition of parallel transport, the $\gamma^*(\nabla)$-flat sections of E along γ are precisely those generated by parallel transport along γ from the fiber $E(\gamma(t_0)) = (\gamma^*E)(t_0)$ over a point $t_0 \in I$. We would like to give a description of parallel transport as the kernel of a suitable differentiation process of classical flavor on the interval I, generalizing Example 1.4 in which the vanishing of partials $\partial a_i/\partial x_j$ of all local coefficient functions a_i characterized the flat sections. This sought-after process will be called *covariant differentiation* (along γ with respect to ∇).

Example 1.6. For any pseudo-Riemannian manifold with corners $(M, \mathrm{d}s^2)$, the tangent bundle $E = TM$ admits a certain canonical connection ∇ determined by the metric tensor; this connection is called the Levi-Civita connection (see §4). Thus, it makes sense to ask whether a path γ in M has its associated vector field γ' along γ given by parallel transport with respect to this connection. That is, does $P_{t_1,t_0,\gamma}(\gamma'(t_0)) = \gamma'(t_1)$ for all t_0, t_1? Such γ are called *geodesics* in $(M, \mathrm{d}s^2)$ provided that γ is not a constant path concentrated at a point of M. (By Example 3.5, non-constancy of such a γ is equivalent to the condition that $\gamma'(t) \neq 0$ for all t, or even for one t.)

A very important example of geodesics arises in General Relativity, according to which gravity is not a force but rather is a manifestation of geometry. Let us review some basic terminology to discuss General Relativity from the viewpoint of the mathematician. (The reader uninterested in such things, if such a reader can possibly exist, may safely skip this and all subsequent digressions into physics.) A *spacetime* is a 4-dimensional oriented and connected smooth Lorentzian manifold $(\mathbf{U}, \mathrm{d}s^2)$ endowed with a time orientation (in the sense of §5 in the handout on orientations of manifolds). This Lorentzian manifold is equipped with its associated Levi-Civita connection (as mentioned in Example 1.6 and developed in §4), and for each $u \in \mathbf{U}$ a connected component of the *local time cone*

$$\{v \in \mathrm{T}_u(\mathbf{U}) \,|\, \langle v, v\rangle_u < 0\}$$

is selected in a "continuously varying" manner by the time orientation; we call the chosen component the *future half-cone* at u (and the other connected component is of course called the *past half-cone* at u). A *particle* is a path $\gamma : I \to \mathbf{U}$ such that for all $t \in I$ two conditions hold: the vector $\gamma'(t) \in \mathrm{T}_{\gamma(t)}(\mathbf{U})$ is a *nonzero* point in the closure of the future half-cone at $\gamma(t)$ (everything moves into the future, even if we see it "at rest" in a 3-dimensional sense!) and $\langle \gamma'(t), \gamma'(t)\rangle_{\gamma(t)} = -m^2$ for some $m \geq 0$. The constant m is called the *rest mass* of the particle. The vector field γ' in the rank-4 bundle $\gamma^*(T\mathbf{U})$ is called the *energy-momentum* vector field: it encodes data corresponding to both energy and momentum in classical physics, and generally there is *no* canonical (orthogonal) decomposition of $T\mathbf{U}$ into "spacelike" and "timelike" subbundles, so one *cannot* intrinsically extract something like classical absolute energy or absolute momentum from γ' (but something can be done in a relative sense, as we will see in Example 3.6). In particular, a particle has no way to detect anything like its own absolute velocity akin to the classical case (unless the particle is massless; see Example 3.6). Of course, in some special models of $\mathbf{U}$ (such as Minkowski space) there is a such an orthogonal decomposition given for $T\mathbf{U}$, but there is no reason to beleive that these special models reflect reality so one should keep more general possibilities in mind for the structure of the spacetime manifold.

Following Einstein, the particle is said to be in *free fall* during a nontrivial time interval $J \subseteq I$ if $\gamma|_J$ is a geodesic. In general, the failure of γ to be a geodesic during a nontrivial time interval $J \subseteq I$ is physically understood to indicate the presence of non-gravitational effects. We will prove in Remark 2.7 that if we specify the position $\gamma(t_0)$ and velocity $\gamma'(t_0)$ of a geodesic $\gamma : I \to M$ in a pseudo-Riemannian manifold with corners at a time $t_0 \in I$ then γ is uniquely determined (up to possibly extending its interval of definition). This fits well with our physical picture of the deterministic nature of free fall motion in a gravitational field, and is essentially the reason that Einstein identified geodesic trajectories in $\mathbf{U}$ (having "velocity" vector field, or rather energy-momentum vector field, that is non-vanishing and lying in the closure of the future half-cone at all points) with the classical idea of free fall motion. Though this uniqueness principle for geodesics sounds analogous to the one for integral curves for vector fields in any smooth manifold (without specifying any metric structure at all), there is a fundamental difference: for the integral curve we have to specify a velocity field over the entire manifold to get the ODE, whereas in the geodesic case the metric tensor determines the ODE. Thus, though the theories of (non-constant) integral curves to vector fields on smooth manifolds and geodesic paths in pseudo-Riemannian manifolds do share some superficial similarities and flavor, neither logically follows from the other.

Remark 1.7. Strictly speaking, one should really define parallel transport $P_{t_1,t_0,\gamma} : E(\gamma(t_0)) \simeq E(\gamma(t_1))$ not only for smooth paths $\gamma : I \to M$ but also for paths γ that are *piecewise smooth* in the sense that I is a locally finite union of closed nontrivial subintervals I_j that meet only at endpoints and such that $\gamma|_{I_j}$ is smooth for all j. The reason we require this is that when we concatenate paths it is rare for the resulting path to be smooth; rather, it is usually just piecewise smooth. Of course, to have a flexible theory for concatenating paths we make essential use of Lemma 1.5 to give us flexibility in shifting time parameterizations.

To do parallel transport from t_0 to t_1 in the piecewise smooth case, we consider the finitely many adjacent subintervals I_j with the first containing t_0 and the last containing t_1. Applying the preceding theory of "smooth" parallel transport to go from t_0 to t_1 step by step through the endpoints of these I_j's then defines the desired isomorphism $P_{t_1,t_0,\gamma}$, and by the associativity condition in the smooth case it is clear that the choice of I_j's does not affect this definition (and more importantly that this extended notion of parallel transport still satisfies the desired associativity conditions as in the smooth case). We will not comment on this issue again, but the interested reader will be easily able to adapt all that follows to the case of piecewise smooth paths by essentially copying the above procedure of chopping up the time interval into such I_j's as above (and checking that this choice never matters).

2. Covariant differentiation and geodesics

Let V_γ be the finite-dimensional vector space of $\gamma^*(\nabla)$-flat sections of $\gamma^*(E)$ over I, so V_γ maps isomorphically to each fiber $(\gamma^*(E))(t) = E(\gamma(t))$ under specialization at t (the composite isomorphism $E(\gamma(t)) \simeq V_\gamma \simeq E(\gamma(t'))$ for $t, t' \in I$ is exactly the parallel transport isomorphism $P_{t',t,\gamma}$ mentioned above). Hence, the natural map of C^∞ vector bundles $I \times V_\gamma \to \gamma^*(E)$ over I is an isomorphism (as it is so on fibers over I). Using the inverse isomorphism, we may therefore describe smooth sections s of $\gamma^*(E)$ over I as smooth maps $\widetilde{s} : I \to V_\gamma$ in the sense of multivariable calculus. In the language of parallel transport, if we *fix* some $t_0 \in I$ and identify V_γ with $E(\gamma(t_0))$ via specialization at t_0 then $\widetilde{s}(t) \in V_\gamma \simeq E(\gamma(t_0))$ is $P_{t_0,t,\gamma}(s(t))$ for all $t \in I$.

For the smooth map $\widetilde{s} : I \to V_\gamma$ from an interval to a finite-dimensional vector space, it makes sense to differentiate it! This gives another smooth map $\widetilde{s}' : I \to V_\gamma$. Converting back into the language of sections of E along γ, this "is" a smooth section $I \to \gamma^*(E)$. We call this section the *covariant derivative* of s with respect to ∇ along γ, and it is denoted via the notation $\frac{Ds}{\mathrm{d}t}$ that suppresses both the path and the connection. (In an evident manner, covariant differentiation commutes with shrinking the interval I and the target M.) The operation $\frac{D}{\mathrm{d}t}$ is a self-map of $(\gamma^*(E))(I)$, and in the language of parallel transport with respect to $\gamma^*(\nabla)$ if we identify V_γ as above with $E(\gamma(t_0))$ via specialization at t_0 then $\frac{Ds}{\mathrm{d}t}(t_1) \in E(\gamma(t_1))$ is equal to

$$P_{t_1,t_0,\gamma}(\widetilde{s}'(t_1)) = P_{t_1,t_0,\gamma}(\frac{\mathrm{d}}{\mathrm{d}t}|_{t_1} P_{t_0,t}(s(t))) \in E(\gamma(t_1)).$$

In words, if we fix some $t_0 \in I$ then all fibers of $\gamma^*(E)$ are identified with $(\gamma^*(E))(t_0) = E(\gamma(t_0))$ via parallel transport by $\gamma^*(\nabla)$ and we may thereby form "difference quotients" in $E(\gamma(t_0))$ between values of s in fibers at different times. For each $t_1 \in I$ we pass to the limit on such difference quotients in the finite-dimensional vector space $E(\gamma(t_0))$ as $t \to t_1$ and then parallel-transport the result from $E(\gamma(t_0)) = (\gamma^*(E)(t_0)$ back to the fiber $(\gamma^*(E))(t_1) = E(\gamma(t_1))$. This dynamic description may make the covariant derivative look uncomputable, since parallel transport in general requires solving linear ODE's and this is usually impossible to do in closed form. However, we shall show that locally along the path where we have local coordinates on M and a local frame for the bundle E there is a very simple formula for $\frac{Ds}{\mathrm{d}t}$ in terms of the associated Christoffel symbols of the connection (pulled back along γ to smooth functions on small opens in I).

Remark 2.1. There is one immediate observation that should be made: the equation $\frac{Ds}{\mathrm{d}t} = 0$ for $s \in (\gamma^*(E))(I)$ is exactly the equation of $\gamma^*(\nabla)$-flatness, or in more classical language $\frac{D}{\mathrm{d}t} = 0$ is the equation governing parallel transport in E (along paths) with respect to ∇. Indeed, by definition $\frac{Ds}{\mathrm{d}t} = 0$ says that the old-fashioned vector space derivative $\widetilde{s}' : I \to V_\gamma$ is zero, which is to say that the map $\widetilde{s} : I \to V_\gamma$ is constant. But this latter map encodes exactly the original section s of the bundle $\gamma^*(E) \to I$ trivialized via the space V_γ of global flat sections (via the specialization isomorphisms $V_\gamma \simeq E(\gamma(t))$ for all t), with respect to which the constant maps to V_γ are precisely the elements of $V_\gamma \subseteq (\gamma^*(E))(I)$, which is to say the flat sections. In more dynamic (but equivalent!) terms, since we have seen that $\widetilde{s}(t) = P_{t_0,t,\gamma}(s(t))$ in $V_\gamma = E(\gamma(t_0))$ for all t, the constancy of $\widetilde{s}$ says $P_{t_0,t,\gamma}(s(t)) = \widetilde{s}(t_0)$ for all t, and since $\widetilde{s}(t_0) = s(t_0)$ this says exactly (upon applying $P_{t,t_0,\gamma} = P_{t_0,t,\gamma}^{-1}$) that $s(t) = P_{t,t_0,\gamma}(s(t_0))$ for all t: s is generated by parallel transport.

In order to give a local formula for the covariant derivative, we first observe one basic property of Leibniz type:

Lemma 2.2. *The operator $\frac{D}{\mathrm{d}t}$ on the space of sections of E along γ is $\mathbf{R}$-linear, and for $f \in C^\infty(I)$ and $s \in (\gamma^*(E))(I)$ we have*

$$\frac{D(fs)}{\mathrm{d}t} = f' \cdot s + f \cdot \frac{Ds}{\mathrm{d}t}.$$

Proof. Passing to the language of smooth maps $\widetilde{s} : I \to V_\gamma$, this is just the classical fact that ordinary differentiation of such maps is $\mathbf{R}$-linear and satisfies the Leibniz Rule with respect to multiplication against smooth functions. ■

Remark 2.3. As a quick application of Lemma 2.2, let us explain in more succinct terms how the covariant differentiation of sections of E along γ with respect to ∇ is really just another way to describe the pullback connection $\gamma^*(\nabla)$ acting on $(\gamma^*(E))(I)$ via the canonical trivialization of Ω^1_I by the section $\mathrm{d}t$. More precisely, we claim that $(\gamma^*(\nabla))(s) = \mathrm{d}t \otimes \frac{Ds}{\mathrm{d}t}$ in $(\Omega^1_I \otimes \gamma^*(E))(I)$. The entire construction of $\frac{D}{\mathrm{d}t}$ was given in terms of the pullback bundle $\gamma^*(E)$ and the pullback connection $\gamma^*(\nabla)$ (which defines the process of parallel transport along γ), so we lose nothing and simplify our notation considerably by renaming this pullback data as E and ∇ and work over I. The problem is really to show that if (E, ∇) is a smooth vector bundle with connection over the interval I and if $s \in E(I)$ is a global smooth section then $\nabla(s) = \mathrm{d}t \otimes \frac{Ds}{\mathrm{d}t}$ in $(\Omega^1_I \otimes E)(I)$ or equivalently $\nabla_{\partial_t}(s) = \frac{Ds}{\mathrm{d}t}$ in $E(I)$. Observe that both sides of this proposed identity are additive in s, and both have the same Leibniz-type behavior with respect to multiplication of s against a smooth function f on I (here we use Lemma 2.2). Hence, if the desired identity holds for $s_1, \dots, s_n \in E(I)$ then it holds for any $C^\infty(I)$-linear combination $\sum a_j s_j$.

By the classification of bundles with connection over an interval, the vector space of ∇-flat sections in $E(I)$ is finite-dimensional and if $\{s_1, \dots, s_n\}$ is a basis of this space then these give a global trivialization of the bundle E. Hence, it is enough to prove the desired identity for these s_j's, or more specifically for ∇-flat $s \in E(I)$. These satisfy $\nabla(s) = 0$, and by Remark 2.1 (with γ the identity map) they also satisfy $\frac{Ds}{\mathrm{d}t} = 0$.

We need one final lemma, which is "general nonsense" for pullback connections with respect to smooth maps between manifolds with corners:

Lemma 2.4. *Let $f : M' \to M$ be a smooth map between smooth manifolds with corners, and let (E, ∇) be a smooth vector bundle with connection over M. Let (E', ∇') be the pullback bundle with connection over M'. For any open $U \subseteq M$ with preimage U', section $s \in E(U)$, and point $m' \in U'$, the values $\nabla'(f^*(s))(m') \in \mathrm{T}_{m'}(M')^\vee \otimes E'(m') = \mathrm{T}_{m'}(M')^\vee \otimes E(f(m'))$ and $\nabla(s)(f(m')) \in \mathrm{T}_{f(m')}(M)^\vee \otimes E(f(m'))$ satisfy*

$$\nabla'(f^*(s))(m') = (\mathrm{d}f(m')^\vee \otimes 1)(\nabla(s)(f(m'))). \tag{2.1}$$

In particular, if $\gamma : I \to M$ is a smooth path and $s \in E(U)$ is a section over an open $U \subseteq M$ containing $\gamma(I)$ then $\frac{D(\gamma^* s)}{\mathrm{d}t}(t_0) \in E(\gamma(t_0))$ is the pairing of $\nabla(s)(\gamma(t_0)) \in \mathrm{T}_{\gamma(t_0)}(M)^\vee \otimes E(\gamma(t_0))$ against the velocity vector $\gamma'(t_0) \in \mathrm{T}_{\gamma(t_0)}(M)$.

Proof. The first part is just the evaluation at m' for the local identities that uniquely characterize the pullback connection $\nabla' = f^*(\nabla)$. (Loosely speaking, $\nabla'(f^*(s)) = f^*(\nabla(s))$.) The second part is the special case $f = \gamma$ because of Remark 2.3. ■

The identity at the end of Lemma 2.4 is classically written as "$\frac{Ds}{\mathrm{d}t} = \nabla_{\gamma'}(s)$", which is terribly confusing notation at first glance because the s on the left is really $\gamma^*(s) \in (\gamma^*(E))(I)$ and more seriously the velocity field γ' along γ is not a section of Vec_M over an open in M (nor does it locally extend to one if γ is not an immersion) and it is not a section of Vec_I over an open in I either, so the notation $\nabla_{\gamma'}$ does not make sense as an operator on sections of E over opens in M (whence it must not be confused with the operator $\nabla_{\vec{v}}$ on sections of $E|_U$ for smooth vector fields $\vec{v}$ over opens U in M) nor does it make sense as an operator on sections of $\gamma^*(E)$ over I (whence it must not be confused with the operator $\gamma^*(\nabla)_{\vec{w}}$ on $(\gamma^*(E))(I)$ for $\vec{w} \in \mathrm{Vec}_I(I)$). The *content* of the classical statement "$\frac{Ds}{\mathrm{d}t} = \nabla_{\gamma'}(s)$", or even of (2.1), is that the left side involves a differential operator over opens in the source manifold and the right side involves a differential operator ∇ over opens in the target manifold.

We have finally reached the point where we can prove the classical local formula for covariant differentiation in terms of the connection coefficients Γ_{ij}^k:

Theorem 2.5. *Let (E, ∇) be a smooth vector bundle with connection over a smooth manifold with corners M. Let $\gamma : I \to M$ be a smooth path. Let $U \subseteq M$ be an open set admitting smooth coordinates $\{x_1, \dots, x_n\}$ and assume $E|_U$ admits a trivializing frame $\{e_1, \dots, e_r\}$. Let Γ_{ij}^k be the associated Christoffel symbols, which is to say $\nabla(e_j) = \sum_{i,k} \Gamma_{ij}^k \mathrm{d}x_i \otimes e_k$. On $J = \gamma^{-1}(U) \subseteq I$, let $\gamma_i = x_i \circ \gamma$ be the local coordinates of the path.*

For any section s of E along γ, with $s|_J = \sum a_j \gamma^*(e_j)$, we have

$$\frac{Ds}{\mathrm{d}t} = \sum_k (a_k' + \sum_{i,j} a_j \gamma_i' \cdot (\Gamma_{ij}^k \circ \gamma))\gamma^*(e_k)$$

in $(\gamma^* E)(J)$.

As promised, this theorem gives a simple explicit local formula for covariant differentiation in terms of three pieces of data: (i) the local coefficients of the section $s \in (\gamma^*(E))(I)$ with respect to a pullback frame, (ii) the local coordinates of the path γ (with respect to local coordinates on the target), and (iii) the pullback along γ of the Christoffel symbols with respect to these local choices of frame and coordinates on E and M. Note, for example, that if we were so lucky as to have these Christoffel symbols vanish then (with such choices of local frame and local coordinates!) covariant differentiation would be given over the open $J \subseteq I$ by componentwise differentiation with respect to the pullback frame for all γ; that is, $\sum a_j \gamma^*(e_j) \mapsto \sum a_j' \gamma^*(e_j)$. One is almost never so lucky.

Proof. The problem is intrinsic to $\gamma|_J$ since covariant differentiation commutes with shrinking I and M, so we can assume $U = M$ and $J = I$. In particular, $\gamma^*(E)$ has the trivializing frame $\{\gamma^*(e_j)\}$. Both sides of the proposed identity are additive in s, and by Lemma 2.2 (and a simple explicit calculation on the right side) both sides have the same behavior with respect to multiplication of s against a smooth function on I. Hence, since s is a $C^\infty(I)$-linear combination of the $\gamma^*(e_j)$'s, we are reduced to the case $s = \gamma^*(e_{j_0})$ for some j_0. In this case, the coefficient functions a_i are equal to δ_{ij_0}, so in particular all a_i' vanish. By Lemma 2.4, the value of $\frac{Ds}{\mathrm{d}t}$ in $(\gamma^*(E))(t_0) = E(\gamma(t_0))$ is the pairing of $\nabla(e_{j_0})(\gamma(t_0)) \in \mathrm{T}_{\gamma(t_0)}(M)^\vee \otimes E(\gamma(t_0))$ against the tangent vector $\gamma'(t_0) \in \mathrm{T}_{\gamma(t_0)}(M)$.

Since $\gamma'(t_0) = \sum \gamma_i'(t_0)\partial_{x_i}|_{\gamma(t_0)}$, by using the definition of the Christoffel symbols we see that pairing $\nabla(e_{j_0})(\gamma(t_0))$ against $\gamma'(t_0)$ gives the output

$$\sum_{i,k} \Gamma^k_{ij_0}(\gamma(t_0))\gamma_i'(t_0)e_k(\gamma(t_0)) = \sum_k (\sum_i \gamma_i'(t_0)\cdot(\Gamma^k_{ij_0}\circ\gamma)(t_0))\gamma^*(e_k)(t_0).$$

This is exactly the desired formula in the special case $s = \gamma^*(e_{j_0})$. ∎

Example 2.6. Consider the special case $E = TM$, and suppose that $\gamma : I \to M$ is a path. As we have seen in Example 1.1, the velocity γ' is a section of E along γ. It therefore make sense, given *any* connection ∇ on TM, to form the covariant derivative of the velocity; this is the notion of acceleration "determined" by ∇. (The velocity vector field γ' along γ makes sense without the data of a connection, but the acceleration vector field γ'' along γ cannot be defined until we define a concept of differentiation for sections of TM along γ; this is exactly what ∇ does.) The vanishing of $\gamma'' := \frac{D\gamma'}{\mathrm{d}t}$, which is to say the parallelism of γ' along γ with respect to ∇, is akin to the classical concept of "motion with constant velocity" (i.e., in a straight line at constant speed); cf. Example 1.6.

In terms of local coordinates $\{x_1, \dots, x_n\}$ on an open $U \subseteq M$ and the associated trivializing frame $\{\partial_{x_1}, \dots, \partial_{x_n}\}$ of $E|_U = TM|_U$, we get Christoffel symbols Γ^k_{ij} and on $\gamma^{-1}(U)$ we have $\gamma'(t) = \sum_i \gamma_i'(t)\partial_{x_i}|_{\gamma(t)}$ with $\gamma_i = x_i \circ \gamma$ the component functions of the path of motion. What is the equation of parallelism of γ' along γ with respect to ∇ in terms of the γ_i's and Γ^k_{ij}'s? Well, by Theorem 2.5 (with $a_j = \gamma_j'$) the equation is

$$0 = \sum_k (\gamma_k'' + \sum_{i,j} \gamma_j'\gamma_i' \cdot (\Gamma^k_{ij}\circ\gamma))\gamma^*(\partial_{x_k}),$$

or in other words

$$\gamma_k'' = -\sum_{i,j} \gamma_j'\gamma_i' \cdot (\Gamma^k_{ij}\circ\gamma) \tag{2.2}$$

for all k. If we are in the special case that the Christoffel symbols all vanish then the system (2.2) for all k is the classical system of equations $\gamma_k'' = 0$ (for all k) that defines a *linearly parameterized* straight line (or point!) $\gamma : t \mapsto v_0 + tv_1$ in an open in a finite-dimensional inner product space (i.e., constant speed motion in the Newtonian sense). When M is endowed with a pseudo-Riemannian metric $\mathrm{d}s^2$, it is a basic result of Levi-Civita (see §4) that there is a *unique* connection on TM

that satisfies two nice properties encoded in terms of parallel transport and covariant differentiation. This distinguished connection determined solely by the metric tensor (the "geometry" of the situation) is called the *Levi-Civita connection* on the tangent bundle TM of $(M, \mathrm{d}s^2)$, and it gives rise to a concept of "zero acceleration paths" in M: the paths γ satisfying the condition $\frac{D\gamma'}{\mathrm{d}t} = 0$. Such paths that are *non-constant* are called *geodesics* for the metric tensor. (See Remark 2.1 for the consistency with the definition of geodesics in Example 1.6, and see Example 3.5 for the fact that $\langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)}$ is independent of t, so it is nowhere zero in the Riemannian case. In the general pseudo-Riemannian case this constant can vanish, as happens for massless particles in General Relativity.) Since the Levi-Civita connection is determined by the geometry (the metric tensor), so is the concept of geodesic in pseudo-Riemannian geometry.

Remark 2.7. Let ∇ be *any* connection on $E = TM$ for a smooth manifold with corners M. Let $\gamma, \widehat{\gamma} : I \rightrightarrows M$ be two paths in M that have the same position m_0 and the same velocity $\vec{v}_0 \in \mathrm{T}_{m_0}(M)$ at a common time $t_0 \in I$. Assume moreover that γ and $\widehat{\gamma}$ are ∇-geodesic in the sense that $\gamma' \in (\gamma^*(TM))(I)$ is $\gamma^*(\nabla)$-flat and $\widehat{\gamma}' \in (\widehat{\gamma}^*(TM))(I)$ is $\widehat{\gamma}^*(\nabla)$-flat (or equivalently, γ' is parallel along γ and $\widehat{\gamma}'$ is parallel along $\widehat{\gamma}$, with parallelism defined by ∇). We claim that $\gamma = \widehat{\gamma}$. (Thus, for example, a geodesic on a pseudo-Riemannian manifold is uniquely determined by specifying its position and velocity at a single time, up to possibly extending its interval of definition.) By local compactness arguments on I much like in the proof of global uniqueness for solutions to first-order ODE initial-value problems, we may reduce to the case when $\gamma(I)$ is contained in a coordinate domain on M, so we can replace M with this coordinate domain to reduce to showing that the system of equations (2.2) has *at most* one solution with a specified initial value.

That is, given smooth functions $h_{ij}^k : I \to \mathbf{R}$ for $1 \le i, j \le n$ and $1 \le k \le N$, we seek to prove that any two solutions $I \to \mathbf{R}^N$ to the system of *non-linear* ODE's

$$y_k'' + \sum_{i,j} h_{ij}^k y_j' y_i' = 0$$

for $1 \le k \le N$ are equal if they agree at a point $t_0 \in I$ *and* have the same first-derivatives at t_0 too. (Note that we do *not* claim existence across all of I, but merely uniqueness upon specification of an "initial" value at some $t_0 \in I$, and in practice geodesics may fail to propogate for all time. The situation is analogous to the case of intervals of definition of integral curves for vector fields; for example, in a compact Riemannian manifold it is a serious theorem that geodesics *do* propogate for all time, much as we have seen is the case for integral curves of smooth vector fields on compact manifolds.) We simply use the old engineering trick to drop the order of the non-linear ODE: for a smooth map $\phi : I \to \mathbf{R}^{2N}$ with component functions $\phi_1, \dots, \phi_{2N}$, we consider the system of first-order ODE's $\phi_k' = \phi_{k+N}$ for $1 \le k \le N$ and

$$\phi_k' = -\sum_{1 \le i,j \le N} h_{ij}^{k-N} \phi_{i+N} \phi_{j+N}$$

for $N + 1 \le k \le 2N$. This can be written in the form $\phi'(t) = A(t, \phi(t))$ where $A : I \times \mathbf{R}^{2N} \to \mathbf{R}^{2N}$ is the smooth map

$$A(t, c_1, \dots, c_{2N}) = (c_{N+1}, \dots, c_{2N}, -\sum_{i,j} h_{ij}^1(t) c_{i+N} c_{j+N}, \dots, -\sum_{i,j} h_{ij}^N(t) c_{i+N} c_{j+N}).$$

It is clear that to give $y_1, \dots, y_N$ solving the given system of second-order non-linear ODE's with initial values $y_i(t_0) = a_i \in \mathbf{R}$ and initial derivatives $y_i'(t_0) = b_i$ at some $t_0 \in I$ for all $1 \le i \le N$ is equivalent to giving a solution ϕ to the new first-order $\mathbf{R}^{2N}$-valued non-linear ODE with initial value $\phi(t_0) = (a_1, \dots, a_N, b_1, \dots, b_N) \in \mathbf{R}^{2N}$. Thus, the global uniqueness theorem for vector-valued first-order (possible non-linear!) initial-value ODE's on an interval gives the required result.

To illustrate the simplest geodesics of all, consider the case that M is open in a finite-dimensional vector space V. We have the *canonical* trivialization $TM \simeq M \times V$ and this allows us to use any non-degenerate quadratic form q on V to define a pseudo-Riemannian smooth metric tensor $\mathrm{d}s_q^2$ on TM via the canonical isomorphisms $\mathrm{T}_m(M) \simeq V$ for all $m \in V$. We will see in Example 4.3 that the Levi-Civita connection for this example is the connection ∇ as in Example 1.4, with respect to which we have seen that the locally constant vector fields are the flat sections and the Γ_{ij}^k's for any *linear* coordinate system restricted to $M \subseteq V$ are all identically 0. Thus, in such cases the geodesics are precisely the non-empty connected opens in $M \cap L$ for affine lines L in V (i.e., lines possibly displaced from the origin by a translation) equipped with *linear parameterization*.

In general, geodesics are the key to pseudo-Riemannian geometry: they play the role of straight lines in classical Euclidean geometry. Though one can say that (in local coordinates) geodesics are simply solutions to the second-order system of ODE's given brutally by the system (2.2) for all k (these are highly non-linear when the Γ_{ij}^k's aren't all zero), this viewpoint sheds very little light on the remarkable geometrical properties of geodesics. One has to do some serious geometric work to unlock the power of the concept of geodesics. For example, it is not obvious but true in the Riemannian case that geodesics are characterized by a certain kind of local "length-minimizing" property, akin to the familiar feature of straight lines in a finite-dimensional inner product space, and it is true but not at all evident that for any $m \in M$ and any nonzero $\vec{v} \in \mathrm{T}_m(M)$ there is a unique geodesic γ in M through m at time 0 (with maximal interval of definition) such that the velocity $\gamma'(0)$ at time 0 is equal to $\vec{v}$. This latter feature of geodesics generalizes the classical observation in a finite-dimensional vector space V that there is a unique linearly parameterized ("constant speed") affine line L passing through a point $m_0 \in V$ at time 0 with a specified nonzero velocity $v \in \mathrm{T}_{m_0}(V) = V$: the parametric line $t \mapsto m_0 + tv$ in V.

Remark 2.8. It is perhaps worth noting that a connection ∇ on E is uniquely determined by its associated covariant differentiation operators on the spaces $(\gamma^*(E))(I)$ of sections along all paths $\gamma : I \to M$. Indeed, if $\{x_i\}$ is a local coordinate system on an open $U \subseteq M$ and $\{e_j\}$ is a local frame for E over U then by taking γ to be the x_i-coordinate axis through a point $m \in U$ the covariant derivative $\frac{D(\gamma^*(e_j))}{\mathrm{d}t}$ along γ has value at m equal to $\sum_k \Gamma_{ij}^k(m)\gamma^*(e_k)$ by Theorem 2.5. Hence, this procedure determines the values $\Gamma_{ij}^k(m)$ at the arbitrary point $m \in U$, so it determines $\nabla|_U$ for any open coordinate domain U, whence it determines ∇ globally (by locality of connections).

3. Metric compatibility and symmetry

We now assume that our vector bundle E is endowed with extra structure, namely a pseudo-Riemannian metric that we shall denote $\langle \cdot, \cdot \rangle$. A linear isomorphism of fibers $E(m) \simeq E(m')$ is called an *isometry* if it carries $\langle \cdot, \cdot \rangle_m$ to $\langle \cdot, \cdot \rangle_{m'}$. The following lemma links several natural properties:

Lemma 3.1. *Let (E, ∇) be a smooth vector bundle with connection over a smooth manifold with corners M, and suppose E is endowed with a pseudo-Riemannian metric tensor $\langle \cdot, \cdot \rangle$. The following three conditions are equivalent:*

1. *for all smooth paths $\gamma : I \to M$, parallel transport $P_{t_1,t_0,\gamma} : E(\gamma(t_0)) \simeq E(\gamma(t_1))$ is an isometry,*
2. *for all open $U \subseteq M$, sections $s_1, s_2 \in E(U)$, and smooth vector fields $\vec{v} \in \mathrm{Vec}_M(U)$, the Leibnitz-style identity*

$$\vec{v}(\langle s_1, s_2 \rangle) = \langle \nabla_{\vec{v}}(s_1), s_2 \rangle + \langle s_1, \nabla_{\vec{v}}(s_2) \rangle \tag{3.1}$$

holds as smooth functions on U,

(3) *for all open $U \subseteq M$ and sections $s_1, s_2 \in E(U)$,*

$$\mathrm{d}(\langle s_1, s_2\rangle) = \langle \nabla(s_1), s_2\rangle + \langle s_1, \nabla(s_2)\rangle \tag{3.2}$$

in $\Omega^1_M(U)$, where the $C^\infty(U)$-valued pairings on the right between $E(U)$ and $(T^*M \otimes E)(U)$ are defined in the evident manner by carrying the 1*-form on the outside.*

When these conditions hold, the connection ∇ is *compatible* with the metric. Condition (1) has the most geometric meaning, though condition (2) is very useful for calculations, especially in the context of covariant differentiation. Condition (3) is largely including for technical purposes (though it is algebraically very appealing for its own sake): it arises as a handy intermediate step for proving the equivalence of (1) and (2).

Remark 3.2. The reader will observe that non-degeneracy of the metric tensor is used nowhere in the proof; it could be any symmetric bilinear form. In particular, it is reasonable to consider the pullback of $(E, \nabla, \langle \cdot, \cdot \rangle)$ with respect to any smooth map $f : M' \to M$ (the point is that $f^*(\langle \cdot, \cdot \rangle)$ may not be non-degenerate, even if $\langle \cdot, \cdot \rangle$ is Riemannian), and the proof of Lemma 3.1 shows that conditions (1)–(3) are inherited by such pullback.

Proof. (of Lemma 3.1). We first prove that (1) implies (2). Since (2) is a local statement, we can assume $M = U$ admits global coordinates $\{x_1, \dots, x_n\}$ and that E has a trivializing frame $\{e_1, \dots, e_r\}$. Both sides of the identity (3.1) are $C^\infty(M)$-linear in $\vec{v}$, so it suffices to treat the case $\vec{v} = \partial_{x_i}$ for some i. By the Leibnitz Rule for connections, it is easy to check that both sides have the same behavior with respect to $C^\infty(M)$-linear combinations in s_1 for fixed s_2 and in s_2 for fixed s_1. Thus, it is enough to treat the case when s_1 and s_2 are members of the arbitrary initial choice of global frame $\{e_j\}$ for E.

For $m_0 \in M$ let $\gamma : I = (-\varepsilon, \varepsilon) \to M$ be the parametric x_i-axis through m_0, so $\gamma(0) = m_0$ and $\gamma'(t) = \partial_{x_i}|_{\gamma(t)}$; in particular, $\gamma'(0) = \partial_{x_i}|_{m_0}$. By Lemma 2.4,

$$\nabla_{\partial_{x_i}}(s)(m_0) = (\gamma^*(\nabla))_{\partial_t}(\gamma^*(s))(0)$$

in $E(m_0) = (\gamma^*(E))(0)$ for any $s \in E(M)$. Since (1) ensures that parallel transport for sections of $\gamma^*(E)$ over I with respect to $\gamma^*(\nabla)$ is an isometry with respect to the pullback metric tensor $\gamma^*(\langle \cdot, \cdot \rangle)$, the identity (3.1) can be formulated in terms of the pullback situation and this situation does satisfy (1) for the identity parameterization of the interval I. Hence, passing to the γ-pullback bundle equipped with its γ-pullback pseudo-Riemannian metric tensor and γ-pullback connection reduces us to proving (3.1) for global sections of a pseudo-Riemannian vector bundle with connection E over an interval I in the case that parallel transport along I is an isometry.

Arguing exactly as in the preceding reduction steps reduces us to the special case when $\vec{v} = \partial_t$ and s_1 and s_2 are members of a single choice of global frame over I (if one exists). Pick a frame of flat sections (as we always may over an interval)! In this case the right side of (3.1) is zero (since $\nabla(s_1)$ and $\nabla(s_2)$ vanish), so we just want $\partial_t(\langle s_1, s_2\rangle) = 0$. That is, we want the pairing $\langle s_1(t), s_2(t)\rangle_t$ to be a constant function of t. The flat sections s_1 and s_2 are generated by parallel transport along I, and by hypothesis such parallel transport is an isometry. This gives the desired constancy.

Now we prove that (2) is equivalent to (3). Again, the problem is local, so we may assume that there are global coordinates $\{x_1, \dots, x_n\}$. In particular, the tangent bundle and cotangent bundle are trivialized by the ∂_{x_i}'s and $\mathrm{d}x_i$'s repsectively. Hence, (3.2) holds if and only if it holds after pairing both sides against each of the ∂_{x_i}'s. But the output of such pairings are (3.1) for $\vec{v}$ equal to any of the ∂_{x_i}'s, and by linearity in $\vec{v}$ over $C^\infty(U)$ such special cases of (2) imply (2) in general. Thus, (2) and (3) are in fact equivalent.

Finally, we prove that (3) implies (1). Note that (1) is a local assertion along the interval I. I claim that property (3) is inherited under pullback. Indeed, to check (3) it is enough to work locally to get to the case when there is a trivializing frame, and by linearity arguments with the Leibnitz Rule it suffices to check for s_1 and s_2 members of a global frame. Thus, to prove that (3) is preserved under pullback we just have to check the pullback situation for the pullbacks of a global frame. Since d commutes with pullback, by using the characterization of the pullback connection and the fiberwise formula for the pullback metric tensor we see that pulling back the identity in (3) for a pair of sections s_1 and s_2 of E gives the analogous identity for the pullback sections of the pullback bundle (endowed with pullback connection and pullback metric). ■

We can restate the compatibility of the connection and the pseudo-Riemannian metric tensor in more classical terms via the language of covariant differentiation: it is necessary and sufficient that for every smooth path $\gamma : I \to M$ and sections s_1 and s_2 of E along γ,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle s_1, s_2\rangle_\gamma = \langle \frac{Ds_1}{\mathrm{d}t}, s_2\rangle_\gamma + \langle s_1, \frac{Ds_2}{\mathrm{d}t}\rangle_\gamma \tag{3.3}$$

as smooth functions on I, where $\langle\cdot,\cdot\rangle_\gamma$ denotes the pullback metric tensor on $\gamma^*(E)$. Here is the simple proof. The condition (3.3) implies the criterion in Lemma 3.1(1) by taking s_1 and s_2 to be generated by parallel transport along γ. Conversely, compatibility with the metric is preserved under pullback (Remark 3.2), so by Remark 2.3 it follows that using pullback by γ identifies (3.3) as an instance of Lemma 3.1(2) with global sections of the γ-pullback and the vector field $\vec{v} = \partial_t$ on the parameter interval.

Example 3.3. Suppose M is open in a finite-dimensional vector space V and $E = TM$ is endowed with the "constant" metric tensor $\mathrm{d}s_q^2$ coming from a non-degenerate quadratic form q on V via the canonical trivialization $TM \simeq M \times V$. For the connection ∇ as in Example 1.4 it is trivial to check Lemma 3.1(1) in this case, so ∇ is compatible with the metric tensor $\mathrm{d}s_q^2$.

Example 3.4. Suppose ∇ is a metric-compatible connection on TM for a pseudo-Riemannian manifold $(M, \mathrm{d}s^2)$. Let $\vec{v} \in \mathrm{Vec}_M(U)$ be a smooth vector field over an open set $U \subseteq M$. For each $m \in U$ we get a unique integral curve $\gamma : I \to U$ for $\vec{v}$ through m at time 0: $\gamma(0) = m$ and $\gamma'(t) = \vec{v}(\gamma(t))$ for all $t \in I$ (and I maximal as such). Recall also that γ is non-constant if and only if $\vec{v}(m) \neq 0$, in which case $\gamma'(t) \neq 0$ for all t. Can we encode in terms of $\vec{v}$ the condition that the non-constant integral curves for $\vec{v}$ are geodesics for the Levi-Civita connection? To simplify the discussion, let us remove from U the closed subset where $\vec{v}$ vanishes, so we suppose that $\vec{v}(m) \neq 0$ for all $m \in U$ or equivalently that all integral curves are non-constant. Hence, the problem is to encode the property that $\frac{D\gamma'}{\mathrm{d}t} = 0$ for all such γ (this covariant derivative being computed along γ; i.e., this $\frac{D}{\mathrm{d}t}$ operation depends on γ!). I claim that this is equivalent to the condition $\nabla_{\vec{v}}(\vec{v}) = 0$. Indeed, if $\gamma : I \to U$ is an integral curve for $\vec{v}$ then for any $t_0 \in I$ we have $\nabla_{\vec{v}}(\vec{v})(\gamma(t_0)) = \frac{D\gamma^*(\vec{v})}{\mathrm{d}t}(t_0)$ by Lemma 2.4, and the "integral curve" condition with respect to $\vec{v}$ says exactly $\gamma^*(\vec{v}) = \gamma'$ in $(\gamma^*(TM))(I)$. This concludes the proof (since every point in U lies on the image of a unique integral curve for $\vec{v}$).

Example 3.5. Let ∇ be a connection on TM for a pseudo-Riemannian manifold with corners M, and assume that ∇ is compatible with the metric. If $\gamma : I \to M$ is a path along which two vector fields $\vec{v}_1, \vec{v}_2 \in (\gamma^*(TM))(I)$ are flat, which is to say $\frac{D\vec{v}_1}{\mathrm{d}t} = 0 = \frac{D\vec{v}_2}{\mathrm{d}t}$, then $\langle \vec{v}_1(t), \vec{v}_2(t)\rangle_{\gamma(t)}$ is independent of t. (In particular, in the case of a Riemannian metric and $\vec{v}_1 = \vec{v}_2 = \gamma'$ with $\frac{D\gamma'}{\mathrm{d}t} = 0$, the speed $\|\gamma'(t)\|_{\gamma(t)}$ is constant.) To see this, we just use (3.3), which gives $\frac{\mathrm{d}}{\mathrm{d}t}\langle \vec{v}_1(t), \vec{v}_2(t)\rangle_{\gamma(t)} = 0$.

For example, if γ is a geodesic in M in the equivalent senses of Examples 1.6 and 2.6, then $\langle\gamma'(t),\gamma'(t)\rangle_{\gamma(t)}$ is a constant $c_\gamma \in \mathbf{R}$ that is independent of t (but may depend on γ). In the Riemannian case this constant must be nonzero, for otherwise γ' is identically zero and so γ is constant. In general, let $\gamma : I \to M$ be a path for which $\langle\gamma'(t),\gamma'(t)\rangle_{\gamma(t)}$ is a nonzero constant c_γ. (This includes particles in spacetime with nonzero rest mass, in the sense of Example 1.6.) If φ is a reparameterization of the path (i.e., a smooth isomorphism between the interval I and a second interval in $\mathbf{R}$), under what conditions does $\gamma\circ\varphi$ retain this constancy property? Since $(\gamma\circ\varphi)'(t) = \varphi'(t)\cdot\gamma'(\varphi(t))$, the pointwise calculation of the inner products changes by a scaling factor that depends on t:

$$\langle(\gamma\circ\varphi)'(t),(\gamma\circ\varphi)'(t)\rangle_{(\gamma\circ\varphi)(t)} = \varphi'(t)^2\langle\gamma'(\varphi(t)),\gamma'(\varphi(t))\rangle_{\gamma(\varphi(t))} = \varphi'(t)^2 c_\gamma.$$

Since $c_\gamma \neq 0$, in order that $\gamma\circ\varphi$ retain the constancy condition it is necessary and sufficient that the non-vanishing $|\varphi'|$ be constant, or equivalently (by connectivity of the parameter interval I) that φ' be a nonzero constant.

Thus, when considering geodesics, the *only* t-reparameterizations that can possibly preserve the geodesicity are linear reparameterizations $\varphi(t) = at+b$ with $a \neq 0$, in which case $c_{\gamma\circ\varphi} = a^2 c_\gamma$. For such φ it is easy to check that $\gamma\circ\varphi$ is in fact again a geodesic when φ is. In the case of a particle γ with positive rest mass $m > 0$, a time reparameterization that *preserves* the rest mass must have the form $\varphi(t) = at+b$ with $a = \pm 1$. But if $a = -1$ then since the energy-momentum vector $\gamma'(\varphi(t))$ in the time cone at $\gamma(\varphi(t))$ lies in the future half-cone (by definition of γ being a particle in spacetime), the energy-momentum vector $(\gamma\circ\varphi)'(t) = -\gamma'(\varphi(t))$ lies in the past half-cone at $\gamma(\varphi(t))$. Thus, if $\gamma\circ\varphi$ is to be a particle with the same rest mass as γ then it is necessary and sufficient that $\varphi(t) = t+b$ for some $b \in \mathbf{R}$. We say that two particles $\gamma_1 : I_1 \to \mathbf{U}$ and $\gamma_2 : I_2 \to \mathbf{U}$ are *physically indistinguishable* if their rest masses m_1 and m_2 coincide and there exists a C^∞ isomorphism $\varphi : I_1 \simeq I_2$ such that $\gamma_1 = \gamma_2\circ\varphi$. We physically interpret the preceding argument as saying that physically indistinguishable particles have the same sense of "proper time" up to additive translation. That a particle determines *its own* sense of time (not in any absolute universal sense that is coordinated with anything else) is very non-Newtonian!

Example 3.6. Let $\gamma : I \to \mathbf{U}$ be a particle with rest mass $m \geq 0$ in the sense of Example 1.6, so $\langle\gamma',\gamma'\rangle_\gamma$ is a constant. By (3.3), the vector field $\frac{D\gamma'}{\mathrm{d}t}$ is therefore orthogonal to the energy-momentum vector field γ' along the entire interval I. Hence, $\frac{D\gamma'}{\mathrm{d}t}$ at time t lies in the hyperplane in $\mathrm{T}_{\gamma(t)}(\mathbf{U})$ orthogonal to the line $\mathbf{R}\cdot\gamma'(t)$. The 3-dimensional hyperplane $(\mathbf{R}\cdot\gamma'(t))^\perp$ is called the *local rest space* of the particle at "proper time t" and the vector $\frac{D\gamma'}{\mathrm{d}t}$ in this local rest space is interpreted as the acceleration felt by γ at time t due to non-gravitational effects. The local rest space is the 3-dimensional vector space of velocities that can be perceived by the particle at time t. (This is *not* a subset of spacetime, but rather is a subspace of the tangent space to spacetime at $\gamma(t)$.) Note that $\gamma'(t)$ is not in the local rest space, except of course if the particle is massless ($m = 0$). Since the Lorentz metric has signature $(3,1)$, if γ has positive rest mass then the Lorentz metric has negative-definite restriction to $\mathbf{R}\cdot\gamma'(t)$ and so we have an orthogonal decomposition $\mathrm{T}_{\gamma(t)}(\mathbf{U}) = \mathbf{R}\cdot\gamma'(t)\oplus(\mathbf{R}\cdot\gamma'(t))^\perp$ with the Lorentz metric having positive-definite restriction to the local rest space. Moreover, since $\mathbf{U}$ is oriented and $\mathbf{R}\cdot\gamma'(t)$ is oriented (by declaring the velocity vector $\gamma'(t)$ to be in the positive half-line), it follows that for particles with positive mass the local rest space is also canonically oriented.

Let us now explain how to make relative observations of 3-dimensional velocity from the viewpoint of another particle at the same point of spacetime (ignore Heisenberg's uncertainty principle!). Let $\gamma_1 : I_1 \to \mathbf{U}$ and $\gamma_2 : I_2 \to \mathbf{U}$ with respective rest masses $m_1 \geq 0$ and $m_2 > 0$ be two particles that lie at $u_0 \in \mathbf{U}$ at some times $t_1 \in I_1$ and $t_2 \in I_2$. Typically the lines $\mathbf{R} \cdot \gamma_1'(t_1), \mathbf{R} \cdot \gamma_2'(t_2) \subseteq \mathrm{T}_{u_0}(\mathbf{U})$ are not the same, so the local rest spaces of the particles at these times are likewise generally rather different hyperplanes in $\mathrm{T}_{u_0}(\mathbf{U})$. Let

$$\vec{w}_{12} = \gamma_1'(t_1) - \langle \gamma_1'(t_1), \frac{\gamma_2'(t_2)}{m_2} \rangle_{u_0} \cdot \frac{\gamma_2'(t_2)}{m_2}$$

be the projection of the energy-momentum vector $\gamma_1'(t_1)$ into the local rest space $(\mathbf{R} \cdot \gamma_2'(t_2))^\perp$ of γ_2 at its time t_2 (recall $m_2 > 0$ by hypothesis), and define

$$\vec{p}_{12} = \frac{m_2 \vec{w}_{12}}{-\langle \gamma_1'(t_1), \gamma_2'(t_2) \rangle_{u_0}} \tag{3.4}$$

to be the *relative momentum* of γ_1 (at its time t_1) from the viewpoint of γ_2 at its time t_2. The denominator in (3.4) is nonzero because otherwise the nonzero vector $\gamma_1'(t_1)$ with non-positive self inner product would lie in the local rest space $(\mathbf{R} \cdot \gamma_2'(t_2))^\perp$ that is Riemannian for the Lorentz metric (as $m_2 > 0$). Of course, the local rest spaces for these two particles at their respective times t_1 and t_2 agree in $\mathrm{T}_{u_0}(\mathbf{U})$ if and only if $\vec{w}_{12} = 0$, or equivalently $\vec{p}_{12} = 0$. In general, since $m_2 > 0$, we can uniquely write the orthogonal decomposition of $\gamma_1'(t_1)$ in $\mathrm{T}_{u_0}(\mathbf{U}) = \mathbf{R} \cdot \gamma_2'(t_2) \oplus (\mathbf{R} \cdot \gamma_2'(t_2))^\perp$ as

$$\gamma_1'(t_1) = \frac{E_{12}}{m_2} \cdot \gamma_2'(t_2) + \vec{w}_{12}$$

for some $E_{12} \in \mathbf{R}$. We have $E_{12} \neq 0$ since $\gamma_1'(t_1)$ does not lie in the local rest space of γ_2 at its time t_2, and since $\gamma_1'(t_1)$ and $\gamma_2'(t_2)$ lie in the closure of the same connected component of the time cone (namely, the closure of the future half-cone selected by the time orientation) we must also have $E_{12} \geq 0$ (use Lemma 5.1 in the handout on orientations on manifolds). Thus, $E_{12} > 0$; this is called the *relative energy* of γ_1 (at its time t_1) from the viewpoint of γ_2 at its time t_2. Since the local rest spaces of γ_2 are Riemannian (as $m_2 > 0$), we have

$$0 \leq \langle \vec{w}_{12}, \vec{w}_{12} \rangle_{u_0} = E_{12}^2 - m_1^2$$

and $\langle \gamma_1'(t_1), \gamma_2'(t_2) \rangle_{u_0} = -E_{12} m_2$. (In particular, $m_1 \leq E_{12}$ with equality if and only if the local rest spaces of γ_1 at t_1 and γ_2 at t_2 coincide.) Thus, $\vec{p}_{12} = \vec{w}_{12}/E_{12}$ and if we define $\vec{v}_{12} = \vec{p}_{12}/E_{12}$ to be the *relative velocity* of γ_1 from the viewpoint of γ_2 (at the usual local times) we compute the corresponding *relative speed* to be

$$\sqrt{\langle \vec{v}_{12}, \vec{v}_{12} \rangle_{u_0}} = \sqrt{1 - (m_1/E_{12})^2} \leq 1$$

with equality if and only if $m_1 = 0$. The physical interpretation is that from the viewpoint of a particle with positive rest mass all observed motion is at most the speed of light (which is 1 in our dimensionless system of measuring speed), and a particle is observed travelling at light speed if and only if it is massless. Moreover, in case $m_1 > 0$ we have

$$\vec{p}_{12} = E_{12} \vec{v}_{12} = \frac{m_1 \vec{v}_{12}}{\sqrt{1 - \|\vec{v}_{12}\|_{u_0}}}.$$

If $m_1 > 0$ then the classical definition of momentum in Newtonian mechanics suggests that we should interpret the relative energy $E_{12} = m_1/\sqrt{1 - \|\vec{v}_{12}\|_{u_0}} \geq m_1$ as the *relative mass* of γ_1 (at its time t_1) from the viewpoint of γ_2 (at its time t_2), thereby equating relative energy and relative mass (i.e., $E = mc^2$ with $c = 1$). Of course, one has to work out low-speed examples to "justify" these definitions of relative energy, relative momentum, and relative velocity.

Let us now return to our general development by introducing a new property of connections in the special case $E = TM$. Suppose ∇ is a connection on TM, so for a pair of smooth vector fields $\vec{v}$ and $\vec{w}$ over an open $U \subseteq M$ we get several new vector fields: $\nabla_{\vec{v}}(\vec{w})$, $\nabla_{\vec{w}}(\vec{v})$, and $[\vec{v}, \vec{w}]$.

Lemma 3.7. *With notation as above, the vector field*

$$(\nabla_{\vec{v}}(\vec{w}) - \nabla_{\vec{w}}(\vec{v})) - [\vec{v}, \vec{w}] \tag{3.5}$$

depends in an alternating bilinear manner on $\vec{v}$ and $\vec{w}$ over the ring $C^\infty(U)$.

Proof. The additivity and skew-symmetry aspects are trivial, so the only real issue is to replace $\vec{v}$ with $f\vec{v}$ for $f \in C^\infty(U)$ and to compute what happens (the same will then go for $\vec{w}$ by swapping the roles of the vector fields and negating). The first term depends linearly on $\vec{v}$ over $C^\infty(U)$, so we just have to prove that

$$\nabla_{\vec{w}}(f\vec{v}) + [f\vec{v}, \vec{w}] \stackrel{?}{=} f \cdot (\nabla_{\vec{w}}(\vec{v}) + [\vec{v}, \vec{w}]).$$

Using the Leibnitz Rule for $\nabla_{\vec{w}}$ and the trivial identity $[f \cdot \vec{v}, \vec{w}] = f[\vec{v}, \vec{w}] - \vec{w}(f)\vec{v}$, the nuisance term $\vec{w}(f)\vec{v}$ cancels out and so the desired result is obtained. ■

Since the formation of (3.5) in this lemma is compatible with shrinking U, by the dictionary between $\mathscr{O}$-modules and vector bundles we conclude that (3.5) defines an alternating bilinear pairing of vector bundles $TM \times TM \to TM$, or equivalently a linear map of bundles

$$\wedge^2(TM) \to TM.$$

This map may be identified with a global section $\mathbf{T}_\nabla$ of the vector bundle $\wedge^2(T^*(M)) \otimes TM$ (or the $\mathscr{O}$-module $\Omega^2_M \otimes \mathrm{Vec}_M$), called the *torsion tensor* of the connection ∇ on $E = TM$. Over an open $U \subseteq M$ admitting smooth coordinates $\{x_1, \ldots, x_n\}$ the torsion tensor has a unique expression as a $C^\infty(U)$-linear combination $\sum_k \sum_{i<j} h_{ijk}(\mathrm{d}x_i \wedge \mathrm{d}x_j) \otimes \partial_{x_k}$. What are the h_{ijk}'s? Using ∂_{x_i}'s as the local frame for the bundle TM, consider the associated Christoffel symbols defined by

$$\nabla(\partial_{x_j}) = \sum_{i,k} \Gamma^k_{ij} \mathrm{d}x_i \otimes \partial_{x_k},$$

or equivalently $\nabla_{\partial_{x_i}}(\partial_{x_j}) = \sum_k \Gamma^k_{ij} \partial_{x_k}$. Since the commutators among the ∂_{x_i}'s are all zero, we thereby easily compute the classical local formula for the torsion tensor in terms of the Christoffel symbols:

$$\mathbf{T}_\nabla|_U = \sum_k \sum_{i<j} (\Gamma^k_{ij} - \Gamma^k_{ji})(\mathrm{d}x_i \wedge \mathrm{d}x_j) \otimes \partial_{x_k}.$$

In terms of the Christoffel symbols Γ^k_{ij} (for a choice of local coordinates $\{x_i\}$ and the associated local frame $\{\partial_{x_k}\}$ of $E = TM$), we see that $\mathbf{T}_\nabla(m) = 0$ for some $m \in U$ if and only if

$$\Gamma^k_{ij}(m) = \Gamma^k_{ji}(m) \tag{3.6}$$

for all i, j, k. (In particular, this latter condition is *independent* of the choice of local coordinates since $\mathbf{T}_\nabla$ is an intrinsic global object.) This motivates the following terminology:

Definition 3.8. The connection ∇ on TM is *symmetric* if $\mathbf{T}_\nabla = 0$.

Keep in mind that the simple-minded (but useful!) Christoffel-symbol symmetry criterion (3.6) for the vanishing of $\mathbf{T}_\nabla(m)$ only works for Christoffel symbols associated to the classical local trivialization $\{\partial_{x_k}\}$ of TM associated to the choice of local coordinate system around m. We also emphasize that, unlike metric compatibility, symmetry is only meaningful for connections on TM (not on more general vector bundles) and it has nothing to do with any metric data. In the case $\dim M \leq 1$ symmetry is automatic since $\Omega^2_M = 0$. The interesting case is therefore $\dim M \geq 2$.

Example 3.9. Consider $M \subseteq V$ and ∇ on TM as in Example 1.4 and Example 3.3. Since $\Gamma_{ij}^k = 0$ when using (the restriction of) a linear coordinate system, ∇ is symmetric.

We now give another symmetry criterion in terms of parametric surfaces in M, the 2-dimensional analogue of a path. We define a *parametric surface* in M to be a smooth map $\sigma : A \to M$ where $A \subseteq \mathbf{R}^2$ is an open subset; as with the case of paths, we do not require σ to be a local immersion or even to be injective. For such a "surface", we define the notion of a (smooth) section of a bundle on M along σ to simply be a (smooth) global section of the pullback bundle $\sigma^*(TM)$ over A. For each point $p = (x_0, y_0) \in A$, let $A_{x_0} = A \cap \{x = x_0\}$ and $A_{y_0} = A \cap \{y = y_0\}$ be the horizontal and vertical "lines" through p in A. Near p these are open intervals, so $\sigma|_{A_{x_0}}$ and $\sigma|_{A_{y_0}}$ are two paths in M through p. They each have velocity vectors at p, which we denote $(\partial_x\sigma)(p), (\partial_y\sigma)(p) \in \mathrm{T}_{\sigma(p)}(M) = (\sigma^*(TM))(p)$. A key point is:

Lemma 3.10. *The two set-theoretic sections of $\sigma^*(TM) \to A$ defined by $\partial_x\sigma$ and $\partial_y\sigma$ are smooth.*

Proof. This is a trivial calculation in local coordinates $\{u_1, \dots, u_n\}$ near $\sigma(p)$: the local component functions $\sigma_i = u_i \circ \sigma$ are smooth and the usual velocity formula along the path gives

$$(\partial_x\sigma)(p) = \sum_i (\partial_x\sigma_i)(p)\partial_{u_i}|_{\sigma(p)} = \sum_i (\partial_x\sigma_i)(p)(\sigma^*(\partial_{u_i}))(p),$$

so the local coefficient functions of $\partial_x\sigma$ with respect to the local frame of $\sigma^*(\partial|_{u_i})$'s are the functions $\partial_x\sigma_i$ that are smooth. Hence, $\partial_x\sigma$ is a smooth section. The case of $\partial_y\sigma$ goes the same way. ■

The two sections $\partial_x\sigma$ and $\partial_y\sigma$ encode the horizontal and vertical velocities of the 2-dimensional parameteric surface σ in M. Now the question arises: does horizontal covariant differentiation of the vertical velocities equal vertical covariant differenation of the horizontal velocities (using the pullback connection on $\sigma^*(TM)$)? Here it is understood that we compute covariant derivatives with respect to *specific* paths in M, namely the ones arise from restriction of σ to vertical and horizontal "lines" in $A \subseteq \mathbf{R}^2$ endowed with their natural coordinate (i.e., structure of path in M via σ). For notational purposes, the horizontal and vertical covariant derivative operators on sections of $\sigma^*(TM)$ along the horizontal and vertical lines in A are denoted $\frac{D}{\partial x}$ and $\frac{D}{\partial y}$ respectively.

Theorem 3.11. *A connection ∇ on TM is symmetric if and only if for any smooth parametric surface $\sigma : A \to M$,*

$$\frac{D}{\partial x}(\partial_y\sigma) = \frac{D}{\partial y}(\partial_x\sigma)$$

as sections of TM along σ.

Proof. Since the entire problem is of local nature, we can work locally on M and A to reduce to the case that M admits global coordinates $\{u_1, \dots, u_n\}$. Let Γ_{ij}^k be the Christoffel symbols of ∇ with respect to the coordinates $\{u_1, \dots, u_n\}$ on M and the associated frame $\{\partial_{u_1}, \dots, \partial_{u_n}\}$ of TM. Let $\sigma_k = u_k \circ \sigma$ be the smooth coefficient functions of σ and let $e_k = \sigma^*(\partial_{u_k})$ be the associated trivializing frame of the vector bundle $E = \sigma^*(TM)$ over A that is equipped with the connection $\sigma^*(\nabla)$. Since

$$\partial_x\sigma = \sum_k \frac{\partial\sigma_k}{\partial x} e_k,$$

the final formula for covariant differentiation in Theorem 2.5 (applied to the restriction to σ to vertical lines in A) gives

$$\frac{D}{\partial y}(\partial_x \sigma) = \sum_k \left(\frac{\partial^2 \sigma_k}{\partial y \partial x} + \sum_{i,j} \frac{\partial \sigma_j}{\partial x} \cdot \frac{\partial \sigma_k}{\partial y} \cdot (\Gamma_{ij}^k \circ \sigma) \right) e_k$$

pointwise on A. (This is really to be viewed as an equality in fibers of the pullback of TM to vertical lines in A.)

Doing the same calculation with the roles of x and y reversed and comparing answers in the fiber at a point of A, the general equality of sums involving the second-order partials is automatic from equality of mixed partials. Now if the connection is symmetric, then $\Gamma_{ij}^k = \Gamma_{ji}^k$ so the equality of the sums of the remaining terms drops out. Conversely, suppose such equalities always hold. We wish to infer that ∇ is symmetric. As has already been noted, symmetry is automatic if $n = \dim M \leq 1$. Thus, we may assume $n \geq 2$. Take the case when σ is the parametric $u_i u_j$-coordinate plane through a point $m \in M$ with $i < j$. In this case the equality $\frac{D}{\partial y}(\partial_x \sigma) = \frac{D}{\partial x}(\partial_y \sigma)$ is the identity $\sum_k (\Gamma_{ij}^k \circ \sigma) e_k = \sum_k (\Gamma_{ji}^k \circ \sigma) e_k$, or in other words $\Gamma_{ij}^k \circ \sigma = \Gamma_{ji}^k \circ \sigma$ for all k. Evaluating at the point of this planar slice corresponding to m gives $\Gamma_{ij}^k(m) = \Gamma_{ji}^k(m)$ for all k. Since $m \in M$ and $i < j$ were arbitrary we get $\Gamma_{ij}^k = \Gamma_{ji}^k$ on M for all i, j, k. This is the criterion (3.6) for vanishing of the torsion tensor of the connection. ■

Finally, we come to a third interpretation of symmetry for connections ∇ on TM that is perhaps the most geometric of all (the criterion in Theorem 3.11 looks vaguely geometric, but it is hard to really say what is geometric about it). We wish to consider the concept of "flatness" for a connection on TM. This is a notion to be studied later for connections ∇ on *any* smooth vector bundle $E \to M$ by means of the concept of a curvature tensor $R_\nabla \in (\wedge^2(T^*M) \otimes E)(M)$, but in the special case that $E = TM$ where $(M, \mathrm{d}s^2)$ is a pseudo-Riemannian manifold with corners and ∇ is the Levi-Civita connection (see §4), it will be a consequence of the Frobenius integrability theorem that the following three conditions are equivalent: $R_\nabla = 0$, locally there exist coordinates with respect to which the metric tensor has constant coefficients (classical "flat" geometry), and locally there exist coordinates with respect to which the Christoffel symbols (associated to the coordinate system and the associated classical trivialization of $E = TM$) all vanish. In fact, further work (theorem of Cartan–Ambrose–Singer) shows that these are all also equivalent to the condition that *all* parallel transport isomorphisms $P_{m',m,\gamma} : \mathrm{T}_m(M) \simeq \mathrm{T}_{m'}(M)$ are unaffected by "continuous deformation" of the path γ. (In more precise topological terminology, we have to allow γ to be piecewise smooth and not merely smooth, and the invariance under continuous deformation means that the $P_{m',m,\gamma}$'s only depend on the homotopy class of γ. The key case is that of parallel transport around all small piecewise smooth loops that begin and end at an arbitrary point $m' = m$.) Under these equivalent conditions, the geometry of the pseudo-Riemannian manifold $(M, \mathrm{d}s^2)$ is said to be *flat*. (None of the equivalences just mentioned will be used here, except briefly for motivational purposes.)

One can ask if there is a reasonable pointwise version of flatness. That is, what should it mean to say that $(M, \mathrm{d}s^2)$ is "flat at $m \in M$"? One possible definition is that the value of the (as yet undefined) curvature tensor R_∇ at m vanishes: $R_\nabla(m) \in \wedge^2(\mathrm{T}_m(M)^\vee) \otimes \mathrm{T}_m(M)$ is equal to 0. A more vivid (but *not* equivalent) definition is inspired by the Christoffel symbol criterion: can we find a coordinate system $\{x_i\}$ around m such that for the Christoffel symbols Γ_{ij}^k associated to $\{x_i\}$ and the local frame $\{\partial_{x_k}\}$ of TM around m the values $\Gamma_{ij}^k(m)$ all vanish? That is, we ask that $\nabla_{\partial_{x_i}}(\partial_{x_j})(m) = 0$ in $\mathrm{T}_m(M)$ for all i and j. This is a good notion of flatness at m in the sense that the global notion of flatness for ∇ on TM over all of M as discussed above is equivalent to the local existence of coordinate systems $\{x_i\}$ such that the Γ_{ij}^k's all vanish on the entire coordinate domain. The vanishing property at m is *not* coordinate-independent, but the existence of *some* such coordinate system is equivalent to symmetry:

Theorem 3.12. *Let ∇ be a connection on TM, and let $m \in M$ be a point. Let $\mathbf{T}_\nabla$ be the torsion tensor of this connection. There exists a local coordinate system $\{x_i\}$ around m such that all $\nabla_{\partial_{x_i}}(\partial_{x_j})(m)$ vanish (or equivalently, all $\Gamma_{ij}^k(m)$ vanish, where the Γ_{ij}^k's are associated to the local coordinates $\{x_i\}$ and the associated trivialization $\{\partial_{x_k}\}$ of TM) if and only if $\mathbf{T}_\nabla(m) = 0$.*

Proof. Since in local coordinates $\mathbf{T}_\nabla = \sum_k \sum_{i<j} (\Gamma_{ij}^k - \Gamma_{ji}^k)(\mathrm{d}x_i \wedge \mathrm{d}x_j) \otimes \partial_{x_k}$, it is equivalent to prove that there exists a coordinate system $\{y_k\}$ around m for which the associated Christoffel symbols vanish at m if and only if there exists a coordinate system $\{x_k\}$ around m such that the associated Christoffel symbols satisfy $\Gamma_{ij}^k(m) = \Gamma_{ji}^k(m)$ for all i, j, k. The first condition obviously implies the second (take $x_k = y_k$), and for the converse we suppose that we are given a coordinate system $\{x_k\}$ near m such that $\Gamma_{ij}^k(m) = \Gamma_{ji}^k(m)$ for all i, j, k. Letting $y_k = x_k + \sum_{ij} c_{ij}^k (x_i - x_i(m))(x_j - x_j(m))$ for constants c_{ij}^k to be determined, the y_k's consistute a coordinate system around m (why?). From an earlier calculation in class, the Christoffel symbols for the connection in this new coordinate system (using the trivialization ∂_{y_k} for TM around m, of course) are given by $\Gamma_{ij}^k - (c_{ij}^k + c_{ji}^k)$ if $i \neq j$ and $\Gamma_{ii}^k - c_{ii}^k$ otherwise. Thus, it suffices to find constants c_{ij}^k such that $\Gamma_{ij}^k(m) = c_{ij}^k + c_{ji}^k$ if $i \neq j$ and $\Gamma_{ii}^k(m) = c_{ii}^k$ for all i. (Note that this is impossible without the symmetry of $\Gamma_{ij}^k(m)$ in i and j!) We simply take $c_{ij}^k = \Gamma_{ij}^k(m)/2$ if $i \neq j$ and $c_{ii}^k = \Gamma_{ii}^k(m)$; the symmetry hypothesis on the $\Gamma_{ij}^k(m)$'s ensures that these constants work. ■

Here is a nice reformulation of the condition "flatness at m" in Theorem 3.12:

Corollary 3.13. *Let ∇ be a connection on TM and let $m \in M$ be a point. Then $\mathbf{T}_\nabla(m) = 0$ if and only if there exists a local coordinate system $\{x_i\}$ around m such that for any path $\gamma : I \to M$ through $m = \gamma(t_0)$ and any vector field $\vec{v} = \sum v_k \gamma^*(\partial_{x_k}) \in (\gamma^*(TM))(I)$ along γ the covariant derivative $\frac{D\vec{v}}{\mathrm{d}t}$ at time t_0 is given by componentwise differentiation: $\frac{D\vec{v}}{\mathrm{d}t}(t_0) = \sum v_k'(t_0) \partial_{x_k}|_m$.*

Proof. By Theorem 2.5, the final condition is equivalent to the condition that for each k the sum $\sum_{i,j} v_j(t_0)\gamma_i'(t_0)\Gamma_{ij}^k(m)$ vanishes for any γ, where $\gamma_i = x_i \circ \gamma$. Taking the coordinate vector fields $\vec{v} = \partial_{x_j}|_\gamma$ for each j, it is equivalent that $\sum_i \gamma_i'(t_0)\Gamma_{ij}^k(m) = 0$ for all j, k. Taking γ to be a coordinate axis through m, it is equivalent that $\Gamma_{ij}^k(m) = 0$ for all i, j, k, and by Theorem 3.12 this is equivalent to $\mathbf{T}_\nabla(m) = 0$. ■

4. Levi-Civita connection

The classical "fundamental theorem of Riemannian geometry" is a basic calculation due to Levi-Civita:

Theorem 4.1. *Let $(M, \mathrm{d}s^2)$ be a pseudo-Riemannian manifold with corners. There exists a unique symmetric connection ∇ on TM that is compatible with the metric tensor. Explicitly, if $\mathrm{d}s^2 = \sum g_{ij} \mathrm{d}x_i \otimes \mathrm{d}x_j$ in local coordinates and $(g^{\alpha\beta}) = (g_{ij})^{-1}$ then $\nabla_{\partial_{x_i}}(\partial_{x_j}) = \sum_k \Gamma_{ij}^k \partial_{x_k}$ where*

$$\Gamma_{ij}^k = \frac{1}{2} \cdot \sum_\ell (\partial_{x_i} g_{j\ell} + \partial_{x_j} g_{i\ell} - \partial_{x_\ell} g_{ij}) g^{\ell k}. \tag{4.1}$$

This connection ∇ is called the *Levi-Civita connection* of $(M, \mathrm{d}s^2)$.

Remark 4.2. Note that $(g^{\alpha\beta})$, and hence (4.1), makes sense because (g_{ij}) is pointwise invertible: $\mathrm{d}s^2$ is a pseudo-Riemannian metric tensor (i.e., pointwise a *non-degenerate* symmetric bilinear form). This non-degeneracy will of course also be crucial in the proof of the existence and uniqueness (which fortunately will not require any use of coordinates). Observe also that, due to (4.1), when the metric tensor coefficients g_{ij} for the chosen coordinate system are *locally constant* then all Γ_{ij}^k vanish. The converse is also true: if for some coordinate system all connection coefficients Γ_{ij}^k of the Levi-Civita connection vanish then the g_{ij} for that coordinate system are locally constant (or equivalently constant, if the coordinate domain is connected). Indeed, in such cases by Theorem 2.5 the parallelism equation $\frac{D\vec{v}}{\mathrm{d}t} = 0$ for a vector field $\vec{v} = \sum v_k \gamma^*(\partial_{x_k})$ along a path γ in the coordinate domain says $v_k'(t) = 0$ for all k, which is to say that each $v_k(t)$ is constant. That is, the vector fields along γ that have constant coefficients with respect to the $\{\partial_{x_k}\}$-frame along γ are the ones generated by parallel transport along *any* γ. Thus, by Example 3.5, $g_{ij} = \langle \partial_{x_i}, \partial_{x_j} \rangle$ is *constant* along all paths in the coordinate domain. By local path-connectivity of M (for which we may use smooth paths), the g_{ij}'s are therefore locally constant (i.e., constant on each connected component of the coordinate domain).

Proof. We write $\langle \cdot, \cdot \rangle$ to denote the symmetric bilinear metric tensor pairing on smooth vector fields. If such ∇ exists then for any open $U \subseteq M$ and smooth vector fields $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathrm{Vec}_M(U)$ we have

$$\vec{v}_1(\langle \vec{v}_2, \vec{v}_3 \rangle) = \langle \nabla_{\vec{v}_1}(\vec{v}_2), \vec{v}_3 \rangle + \langle \vec{v}_2, \nabla_{\vec{v}_1}(\vec{v}_3) \rangle, \tag{4.2}$$

$$\vec{v}_2(\langle \vec{v}_3, \vec{v}_1 \rangle) = \langle \nabla_{\vec{v}_2}(\vec{v}_3), \vec{v}_1 \rangle + \langle \vec{v}_3, \nabla_{\vec{v}_2}(\vec{v}_1) \rangle, \tag{4.3}$$

$$\vec{v}_3(\langle \vec{v}_1, \vec{v}_2 \rangle) = \langle \nabla_{\vec{v}_3}(\vec{v}_1), \vec{v}_2 \rangle + \langle \vec{v}_1, \nabla_{\vec{v}_3}(\vec{v}_2) \rangle, \tag{4.4}$$

by metric compatibility of ∇. Thus, using that $\langle \cdot, \cdot \rangle$ is symmetric bilinear we compute

$$\begin{aligned} \vec{v}_1(\langle \vec{v}_2, \vec{v}_3 \rangle) + \vec{v}_2(\langle \vec{v}_3, \vec{v}_1 \rangle) - \vec{v}_3(\langle \vec{v}_1, \vec{v}_2 \rangle) \quad = \quad & \langle \nabla_{\vec{v}_1}(\vec{v}_2) - \nabla_{\vec{v}_2}(\vec{v}_1), \vec{v}_3 \rangle + \langle \nabla_{\vec{v}_2}(\vec{v}_3) - \nabla_{\vec{v}_3}(\vec{v}_2), \vec{v}_1 \rangle \\ & - \langle \nabla_{\vec{v}_3}(\vec{v}_1) - \nabla_{\vec{v}_1}(\vec{v}_3), \vec{v}_2 \rangle + 2 \langle \nabla_{\vec{v}_2}(\vec{v}_1), \vec{v}_3 \rangle, \end{aligned}$$

and by symmetry of ∇ this is equal to

$$\langle [\vec{v}_1, \vec{v}_2], \vec{v}_3 \rangle + \langle [\vec{v}_2, \vec{v}_3], \vec{v}_1 \rangle - \langle [\vec{v}_3, \vec{v}_1], \vec{v}_2 \rangle + 2 \langle \nabla_{\vec{v}_2}(\vec{v}_1), \vec{v}_3 \rangle.$$

Thus, necessarily we must have

$$\langle \nabla_{\vec{v}_2}(\vec{v}_1), \vec{v}_3 \rangle = \frac{1}{2} \cdot (\vec{v}_1(\langle \vec{v}_2, \vec{v}_3 \rangle) + \vec{v}_2(\langle \vec{v}_3, \vec{v}_1 \rangle) - \vec{v}_3(\langle \vec{v}_1, \vec{v}_2 \rangle) - \langle [\vec{v}_1, \vec{v}_2], \vec{v}_3 \rangle - \langle [\vec{v}_2, \vec{v}_3], \vec{v}_1 \rangle + \langle [\vec{v}_3, \vec{v}_1], \vec{v}_2 \rangle). \tag{4.5}$$

Due to the non-degeneracy of $\langle \cdot, \cdot \rangle$, this local formula uniquely determines $\nabla_{\vec{v}_2}(\vec{v}_1)$, and so it uniquely determines ∇. Conversely, *define* set-theoretic maps $\nabla_U : \mathrm{Vec}_M(U) \to (\Omega_M^1 \otimes \mathrm{Vec}_M)(U) = \mathrm{Hom}_U(\mathrm{Vec}_U, \mathrm{Vec}_U)$ by the requirement that for any $\vec{v}_1 \in \mathrm{Vec}_M(U)$, open $U' \subseteq U$, and $\vec{v}_2 \in \mathrm{Vec}_M(U')$, the vector field evaluation $\nabla_{U,\vec{v}_2}(\vec{v}_1)$ of $\nabla_U(\vec{v}_1)$ against $\vec{v}_2$ satisfies (4.5). Such ∇_U's are uniquely determined in this way and are compatible with shrinking U (i.e., $\nabla_U(\vec{v})|_{U'} = \nabla_{U'}(\vec{v})$ for any open $U' \subseteq U$ and $\vec{v} \in \mathrm{Vec}_M(U)$), and by uniqueness and the formula it is immediate that the resulting collection of maps $\nabla = \{\nabla_U\}_U$ is a symmetric connection on the vector bundle TM associated to the $\mathscr{O}$-module Vec_M. The criterion in Lemma 3.1(2) gives the metric compatibility.

Finally, we derive the explicit coordinate formula. Let $\{x_i\}$ be a local coordinate system on an open $U \subseteq M$ and let $\vec{v}_1 = \partial_{x_j}$, $\vec{v}_2 = \partial_{x_i}$, and $\vec{v}_3 = \partial_{x_\ell}$, so

$$\langle \nabla_{\vec{v}_2}(\vec{v}_1), \vec{v}_3 \rangle = \langle \sum_k \Gamma_{ij}^k \partial_{x_k}, \partial_{x_\ell} \rangle = \sum_{k,\ell} \Gamma_{ij}^k g_{k\ell}.$$

By (4.5), we conclude

$$\sum_{k,\ell} g_{k\ell}\Gamma_{ij}^k = \frac{1}{2}\cdot(\partial_{x_j} g_{i\ell} + \partial_{x_i} g_{j\ell} - \partial_{x_\ell} g_{ij})$$

since the commutators among ∂_{x_i}, ∂_{x_j}, and ∂_{x_k} all vanish. For fixed i and j this is a system of linear equations for the Γ_{ij}^k's with coefficients in the symmetric invertible matrix $(g_{k\ell})$ (either pointwise in U, or over the ring $C^\infty(U)$), so we can solve these equations using the (symmetric!) inverse matrix $(g^{\alpha\beta})$. This gives the asserted formula for the Christoffel symbols. ■

Example 4.3. By Example 3.3 and Example 3.9, the connection ∇ in Example 1.4 is the Levi-Civita connection with respect to the metric tensor $\mathrm{d}s_q^2$ arising from any non-degenerate quadratic form q on V via the canonical bundle isomorphism $TM \simeq M \times V$. Hence, for any $m_0, m_1 \in M$ we can identify the canonical isomorphism $\mathrm{T}_{m_0}(M) \simeq V \simeq \mathrm{T}_{m_1}(M)$ as parallel transport with respect to the Levi-Civita connection arising from any "flat" metric tensor (i.e., one arising from q on V as above). Beware that this process cannot be used to *define* these isomorphisms of tangent spaces since the very definition of $\mathrm{d}s_q^2$ uses the canonical isomorphisms $\mathrm{T}_m(M) \simeq V$ in the first place to actually use q to put the metric structure on the tangent spaces at all points of M. Rather, it simply puts these isomorphisms into a more general framework, as elementary examples of a structure that we have on any pseudo-Riemannian manifold at all.

In classical Riemannian geometry, the basic objects of study were surfaces in $\mathbf{R}^3$ or smooth hypersurfaces in a finite-dimensional inner product space $(V, \langle\cdot,\cdot\rangle)$, using the induced metric tensor from the "flat" one on the manifold V induced by $\langle\cdot,\cdot\rangle$ as in the preceding Example. That is, classically the length of a tangent vector at a point on a submanifold $i : M \hookrightarrow V$ was computed by using the given inner product on V: this is exactly the pullback metric tensor via the inclusions $\mathrm{T}_m(M) \subseteq \mathrm{T}_{i(m)}(V) \simeq V$ (the final isomorphism being the canonical one as for any finite-dimensional vector space). The general concept of connection was not known; rather, given a section $\vec{v}$ of TM along a curve $\gamma : I \to M$, the covariant derivative was defined by first computing the old-fashioned derivative of the V-valued map

$$I \stackrel{\vec{v}}{\to} TM \stackrel{\mathrm{d}i}{\to} TV \simeq V \times V \stackrel{\mathrm{pr}_2}{\to} V$$

(sending each t to $\vec{v}(t) \in \mathrm{T}_{\gamma(t)}(M) \subseteq \mathrm{T}_{i(\gamma(t))}(V) = V$) and then projecting this orthogonally into the subspace $\mathrm{T}_{\gamma(t)}(M)$. The resulting set-theoretic section $I \to \gamma^*(TM)$ was taken to be the "derivative with respect to M for $\vec{v}$ along γ." In fact this extrinsic operation is intrinsic to M equipped with its (induced) metric tensor: it is the covariant derivative $\frac{D\vec{v}}{\mathrm{d}t}$ with respect to the Levi-Civita connection on TM with its induced metric tensor from the "flat" one on TV. The reason is that by Theorem 2.5 (with all Γ_{ij}^k equal to 0) the classical calculus-style derivative of coeffcient functions with respect to a frame of "constant" vector fields computes the covariant derivative of $\vec{v} \in (i\circ\gamma)^*(TV)$ along $i\circ\gamma$ in V with respect to the "flat" metric structure on the manifold V, and so to justify the asserted intrinsic nature of the classical procedure we need a compatibility condition between orthogonal projection and covariant differentiation with respect to the Levi-Civita connection:

Theorem 4.4. *Let $i : M \hookrightarrow M'$ be an isometric immersion between Riemannian manifolds, and let ∇' and ∇ be the associated Levi-Civita connections on TM' and TM. For any smooth map $\gamma : I \to M$ and vector field $\vec{v} \in \gamma^*(TM)(I)$ along γ, the covariant derivative $\frac{D\vec{v}}{\mathrm{d}t}$ with respect to ∇ along γ is the orthogonal projection of the covariant derivative $\frac{D(\mathrm{d}i\circ\vec{v})}{\mathrm{d}t}$ with respect to ∇' along $i\circ\gamma$. Here we use the orthogonal projection from $i^*(TM')$ onto its subbundle TM via the metric tensor.*

Rather than prove Theorem 4.4 directly, it is more elegant to recast it in a general setting that is unrelated to tangent bundles because this makes the situation better-suited to arguments resting on pullback to an interval. We first give a general procedure that relates the Levi-Civita connection on TM with the pullback to $i^*(TM')$ of the Levi-Civita connection on TM'. Note that TM is a subbundle of $i^*(TM')$ with metric tensor given by restriction of the i-pullback of the one on TM'. This motivates us to consider rather generally a vector bundle with connection (E', ∇') over a smooth manifold with corners M, and let $E \subseteq E'$ be a subbundle admitting a choice of complementary bundle $E_1 \subseteq E'$ (so $E \oplus E_1 = E'$). For example, as in our case of interest, if E' is endowed with a Riemannian metric then a canonical E_1 is the orthogonal complement to E in E'. We can define a connection ∇ on E induced by the data of ∇' and the projection $p : E' \to E$ as follows. For open $U \subseteq M$ and $s \in E(U) \subseteq E'(U)$, we get the section $\nabla'(s) \in (\Omega^1_M \otimes E')(U)$. Thus, $\nabla(s) := (1 \otimes p)((\nabla')(s)) \in (\Omega^1_M \otimes E)(U)$ is a smooth section over U. This operation ∇ is compatible with shrinking on U and is clearly additive in s. Also, for $f \in C^\infty(U)$ it follows from the definition that $\nabla(s) = \mathrm{d}f \otimes s + f\nabla(s)$. Hence, ∇ is a connection on E determined by ∇' and p. Observe a very important property of this construction: *it is compatible with pullback*. That is:

Lemma 4.5. *With notation as above, if $f : N \to M$ is a smooth map and we endow $f^*(E') = f^*(E) \oplus f^*(E_1)$ and $f^*(E)$ with the pullback connections $f^*(\nabla')$ and $f^*(\nabla)$ then $f^*(\nabla')$ and the projection $f^*(p) : f^*(E') \to f^*(E)$ induce $f^*(\nabla)$.*

Proof. This is an exercise in the characterization of the pullback connection. The asserted identity is of local nature on M (i.e., if $\{U_i\}$ is an open cover of M then it suffices to prove the lemma for the restricted situation $f^{-1}(U_i) \to U_i$ for all i). Hence, we can assume E and E_1 are trivial bundles, say with trivializing frames $\{s_j\}$ and $\{t_i\}$ that together give a frame for E'. Then $\{f^*(s_j)\}$ and $\{f^*(t_i)\}$ serve the analogous purpose for the pullback bundles on N. To justify the desired equality of connections on $f^*(E)$ it suffices (by locality and the Leibnitz rule) to compare values on the $f^*(s_j)$'s. More generally, if s is any section in $E(M)$ we will check equality after evaluation on $f^*(s)$. By definition, the connection on $f^*(E)$ induced by projection of $f^*(\nabla')$ has value on $f^*(s)$ equal to the $f^*(E)$-component of $f^*(\nabla')(f^*(s))$. Thus, we just have to show that the projection $f^*(E') \to f^*(E)$ carries $f^*(\nabla')(f^*(s))$ to $f^*(\nabla)(f^*(s))$. By definition of the pullback connection, $f^*(\nabla')(f^*(s)) = f^*(\nabla'(s))$ in $(f^*(E'))(N)$ and $f^*(\nabla)(f^*(s)) = f^*(\nabla(s))$ in $(f^*(E))(N)$. Hence, we want $(f^*(p))(f^*(\nabla'(s))) = f^*(\nabla(s))$. Since $\nabla(s) = p(\nabla'(s))$ by definition of ∇ in terms of ∇', we are reduced to checking that $(f^*(p))(\xi) = f^*(p(\xi))$ in $(f^*(E))(N)$ for any $\xi \in E'(M)$. This is obvious by passing to fibers over points of N. ■

Our interest in this projection construction with connections is due to:

Lemma 4.6. *Let $i : M \to M'$ be an immersion of manifolds with corners, and endow M' with a Riemannian metric. Consider the pullback metric tensor on $E' = i^*(TM')$, and the pullback $i^*(\nabla')$ of the Levi-Civita connection. Combining this pullback with orthogonal projection onto the subbundle TM defines a connection ∇ on TM in accordance with the above construction. This connection on TM is the Levi-Civita connection for the induced metric tensor on M.*

Note that it is essential for the metric tensor to be definite rather than perhaps indefinite, as otherwise the restriction to TM may fail to have the non-degeneracy properties that are necessary for the existence and uniqueness of a Levi-Civita connection.

Proof. By the uniqueness of the Levi-Civita connection, we just have to prove that ∇ is symmetric and is compatible with the induced metric tensor on TM. For metric compatibility, we use the criterion in Lemma 3.1(1): for a smooth path $\gamma : I \to M$ we want parallel transport in TM along

γ to be a fiberwise isometry. This problem is intrinsic to $\gamma^*(\nabla)$ on $\gamma^*(TM)$ equipped with the γ-pullback of the induced metric tensor on TM. But $\gamma^*(TM)$ is a subbundle of $(i \circ \gamma)^*(TM') = \gamma^*(i^*(TM))$, so the metric compatibility becomes the following problem. We have a vector bundle with connection (E', ∇') over a manifold with corners (such as I) and a subbundle E as well as a metric tensor $\langle \cdot, \cdot \rangle$ on E' with respect to which ∇' is compatible. Using orthogonal projection induces a connection ∇ on E and we want to prove that ∇ is compatible with the restriction of the metric tensor. That is, by Lemma 3.1(3), if s_1 and s_2 are local sections of E over an open then we can consider these as sections of E' and we want the 1-form $\mathrm{d}(\langle s_1, s_2 \rangle)$ to equal $\langle \nabla(s_1), s_2 \rangle + \langle s_1, \nabla(s_2) \rangle$. By definition, under the decomposition $E' = E \oplus E^{\perp}$ we have $\nabla'(s) = \nabla(s) + \nabla^{\perp}(s)$ for any local section s of E', so we just need $\nabla^{\perp}(s_1)$ to be orthogonal to s_2 and $\nabla^{\perp}(s_2)$ to be orthogonal to s_1 as sections of E'. But s_1 and s_2 are sections of the subbundle E that is fiberwise orthogonal to the subbundle $E^{\perp}$, so the metric compatibility problem is settled.

Next we consider the symmetry problem. It seems unpleasant to attack this via Christoffel symbols; Theorem 3.11 provides the right symmetry criterion. For a smooth parametric surface $\sigma : A \to M$ with an open $A \subseteq \mathbf{R}^2$, we seek to prove $\frac{D}{\partial x}(\partial_y \sigma) = \frac{D}{\partial y}(\partial_x \sigma)$ as A-sections of $\sigma^*(TM)$, where we compute covariant derivatives along the x and y-lines of A with respect to the connection ∇. By symmetry of ∇', we have the equality of covariant derivatives

$$\frac{D'}{\partial x}(\partial_y(i \circ \sigma)) = \frac{D'}{\partial y}(\partial_x(i \circ \sigma)) \tag{4.6}$$

as A-sections of $(i \circ \sigma)^*(TM') = \sigma^*(i^*(TM'))$, where $\frac{D'}{\partial x}$ and $\frac{D'}{\partial y}$ denote covariant derivatives along $i \circ \gamma$ with respect to ∇'. At $p = (x_0, y_0) \in A$, (4.6) is an equality in $\mathrm{T}_{i(\sigma(p))}(M')$ and so we likewise get an equality of orthogonal projections in $\mathrm{T}_{\sigma(p)}(M) = (\sigma^*(TM))(p)$. Hence, it suffices to prove in general that the orthogonal projection from $\sigma^*(i^*(TM'))$ to $\sigma^*(TM)$ at p carries $\frac{D'}{\partial x}(\partial_y(i \circ \sigma))(p)$ to $\frac{D}{\partial x}(\partial_y \sigma)(p)$ (and then the same holds with x and y swapped). Since $\partial_y(i \circ \sigma) = \mathrm{d}i \circ \partial_y \sigma$, by Lemma 4.5 (applied to $f = \sigma(\cdot, y_0)$) and the relationship between covariant derivatives and connections on vector bundles over an interval (as in Remark 2.3) we are reduced to the following problem. Consider a bundle with connection (E', ∇') over an interval I such that ∇' is compatible with a given Riemannian metric on E'. This induces a connection ∇ on a subbundle $E \subseteq E'$ via orthogonal projection. For a section $s \in E(I) \subseteq E'(I)$ we seek to show that $\nabla'_{\partial_t}(s) \in E'(I)$ has orthogonal projection $\nabla_{\partial_t}(s)$ in $E(I)$. By the definition of ∇ in terms of ∇' this is a tautology. ■

To prove Theorem 4.4, the preceding lemma allows us to leave the restrictive setting of Levi-Civita connections on tangent bundles and to work more generally with connections on arbitrary vector bundles. More specifically, Theorem 4.4 is now a special case of the following general "algebraic" situation that is susceptible to pullback arguments (whereas tangent bundles are too special for such tricks). We consider a bundle with connection (E', ∇') over a manifold with corners M, and we let $E' = E \oplus E_1$ be a direct sum decomposition into subbundles. By projection $E' \to E$, ∇' induces a connection ∇ on E. For any path $\gamma : I \to M$ and section $s \in (\gamma^*(E))(I) \subseteq (\gamma^*(E'))(I)$, we must show that the projection of $\frac{D's}{\mathrm{d}t} \in (\gamma^*(E'))(I)$ (covariant derivative with respect to ∇') into $(\gamma^*(E))(I)$ is equal to the covariant derivative $\frac{Ds}{\mathrm{d}t}$ with respect to ∇. If we rename $\gamma^*(E')$, $\gamma^*(E)$, $\gamma^*(\nabla')$, and $\gamma^*(\nabla)$ as E', E, ∇', and ∇ then (thanks to Lemma 4.5 with $f = \gamma$) we retain the basic relationship that ∇ is constructed from ∇' via the projection. Hence, our problem is as at the end of the proof of Lemma 4.6: does the projection $E'(I) \to E(I)$ carry $\nabla'_{\partial_t}(s)$ to $\nabla_{\partial_t}(s)$? This again is a tautology, due to the relationship between ∇' and ∇ with respect to bundle projection (now over I), and so completes the proof of Theorem 4.4.

Here is a variant on Theorem 4.4 that describes flatness along a path in a submanifold (with induced metric tensor) in terms of ambient covariant differentiation.

Corollary 4.7. *Let $i : M \to M'$ be an isometric immersion between Riemannian manifolds with corners. Let $\gamma : I \to M$ be a path, and $\vec{v} \in (\gamma^*(TM))(I) \subseteq ((i \circ \gamma)^*(TM'))(I)$ a vector field on M along the path. This vector field is parallel for the Levi-Civita connection on M if and only if its covariant derivative $\frac{D'\vec{v}}{\mathrm{d}t}$ along $i \circ \gamma$ with respect to ∇' is fiberwise orthogonal to the tangent space $\mathrm{T}_{\gamma(t)}(M) \subseteq \mathrm{T}_{i(\gamma(t))}(M')$.*

Proof. We have shown above that the covariant derivative $\frac{D\vec{v}}{\mathrm{d}t}$ along γ with respect to ∇ is the orthogonal projection of $\frac{D'\vec{v}}{\mathrm{d}t}$. Since the vanishing of $\frac{D\vec{v}}{\mathrm{d}t}$ is the parallelism condition, we conclude that it is equivalent to the vanishing of the orthogonal projection of $\frac{D'\vec{v}}{\mathrm{d}t}$ into the tangent spaces of M along γ. This is precisely the claim in the corollary. ■

5. The Fermi–Walker construction

In this optional section (which the reader uninterested in physics can safely omit) we discuss the problem of how to define a general time-invariant notion of *rest frame* of a particle with positive mass in General Relativity, and how on any pseudo-Riemannian manifold M the same idea associates to any (possibly non-geodesic!) path γ for which $\langle \gamma', \gamma' \rangle_\gamma$ is a nonzero constant a metric-compatible connection on $\gamma^*(TM)$ with respect to which γ is geodesic! (In the case that γ is a geodesic for the Levi-Civita connection on M, this construction will recover the γ-pullback of the Levi-Civita connection.)

For motivational purposes, consider a spacetime manifold $\mathbf{U}$ (equipped with its Levi-Civita connection ∇) and a particle $\gamma : I \to \mathbf{U}$ in the sense of Example 1.6. We do not assume this particle to be in free fall, which is to say that γ' may not be a parallel vector field along γ with respect to ∇. However, we do assume it has positive rest mass, so by Example 3.6 at each time $t \in I$ there is associated a canonical 3-dimensional subspace $(\mathbf{R} \cdot \gamma'(t))^\perp \subseteq \mathrm{T}_{\gamma(t)}(\mathbf{U})$, the so-called local rest space at time t, and it is Riemannian and oriented because of the hypothesis of positive rest mass. Since γ' generates a (trivial) line subbundle L of the tangent bundle to spacetime along γ, by forming the rank-3 orthogonal complement subbundle $L^\perp \subseteq \gamma^*(T\mathbf{U})$ (whose t-fiber is the local rest space at time t), we get a "smooth family" of local rest spaces along γ. Clearly $L^\perp$ is an oriented Riemannian rank-3 subbundle of $\gamma^*(T\mathbf{U})$.

Our problem is to find a way to *canonically* identify the local rest spaces with each other (as oriented inner product spaces!) at different times along the path so as to enable the particle to have a sense of being in a *fixed* 3-dimensional oriented inner product space (rest frame!) as time changes. I claim that it suffices to find a canonical metric-compatible connection ∇_F^γ on the vector bundle $\gamma^*(T\mathbf{U})$ with respect to which the section γ' is parallel. Suppose we have such a connection. The resulting isometric parallel transport isomorphisms

$$\mathscr{P}_{t_1,t_0,\gamma} : \mathrm{T}_{\gamma(t_0)}(\mathbf{U}) = (\gamma^*(T\mathbf{U}))(t_0) \simeq (\gamma^*(T\mathbf{U}))(t_1) = \mathrm{T}_{\gamma(t_1)}(\mathbf{U})$$

are orientation-preserving, due to continuity/connectivity reasons and the fact that parallel transport is the identity for $t_1 = t_0$. These isomorphisms also carry $\mathbf{R} \cdot \gamma'(t_0)$ to $\mathbf{R} \cdot \gamma'(t_1)$ respecting the canonical orientations of these lines (as $\mathscr{P}_{t_1,t_0,\gamma}(\gamma'(t_0)) = \gamma'(t_1)$ by the parallelism hypothesis for γ' with respect to ∇_F^γ). Thus, this parallel transport will have to carry the local rest space $(\mathbf{R} \cdot \gamma'(t_0))^\perp$ at time t_0 to the local rest space $(\mathbf{R} \cdot \gamma'(t_1))^\perp$ at time t_1 as oriented inner product spaces. In fact, by the theory of connections on bundles over an interval, such parallel transports for varying $t_1 \in I$ and fixed $t_0 \in I$ canonically identify the Riemannian oriented vector bundle $L^\perp$ of local rest spaces over I with the constant bundle $I \times (\mathbf{R} \cdot \gamma'(t_0))^\perp$ (equipped with its Riemannian structure and orientation induced by the ones on $(\mathbf{R} \cdot \gamma'(t_0))^\perp$). For a fixed choice of orthonormal positive basis of the local rest space at time t_0 this oriented Riemannian bundle trivialization gives a "consistent" selection of positive basis in the local rest space of the particle at all times. This is exactly what physicists call a "lab frame". We prefer to avoid bases and non-canonical choices of $t_0 \in I$, so once ∇_F^γ is constructed we define the *rest frame* of the particle to be the 3-dimensional vector space of ∇_F^γ-flat sections of the bundle $L^\perp$ of local rest spaces along γ. (It has a canonical orientation and Riemannian structure, namely the common one induced by its specialization isomorphism onto the local rest space of γ at any $t \in I$.)

This is all good and well, but where is the required metric-compatible connection ∇_F^γ making γ' parallel along γ to come from? In the case of free fall we can of course use the γ-pullback of the Levi-Civita connection over spacetime. In general, one applies the following interesting construction. Let $(M, \mathrm{d}s^2)$ be a pseudo-Riemannian manifold with corners, and let $\gamma : I \to M$ be a path in M with velocity field γ' such that $\langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)}$ is equal to a *nonzero* constant c_γ. (This corresponds to the positivity requirement on the rest mass of the particle above.) Let $L \subseteq \gamma^*(TM)$ be the trivial line subbundle spanned by γ', and let $L^\perp$ be its orthogonal complement; we have $\gamma^*(TM) = L \oplus L^\perp$ since $c_\gamma \neq 0$. Let $p : \gamma^*(TM) \to L$ and $p^\perp : \gamma^*(TM) \to L^\perp$ denote the two orthogonal projections. The *Fermi-Walker connection* ∇_F^γ on $\gamma^*(TM)$ is the one whose associated covariant derivative operator $\frac{D_F^\gamma}{\mathrm{d}t}$ over an open $J \subseteq I$ is

$$\vec{v} \mapsto p(\frac{D(p(\vec{v}))}{\mathrm{d}t}) + p^\perp(\frac{D(p^\perp(\vec{v}))}{\mathrm{d}t}) \in L(J) \oplus L^\perp(J) = (\gamma^*(TM))(J).$$

In other words, ∇_F^γ on $\gamma^*(TM) = L \oplus L^\perp$ is the direct sum of the connections induced on L and $L^\perp$ by ∇ and the projections p and $p^\perp$ via the construction considered below Theorem 4.4. Here is an alternative useful formula for the Fermi–Walker connection:

$$\vec{v} \mapsto \frac{D\vec{v}}{\mathrm{d}t} - \frac{\langle \gamma', \vec{v} \rangle_\gamma}{c_\gamma} \cdot \frac{D\gamma'}{\mathrm{d}t} + \frac{\langle \frac{D\gamma'}{\mathrm{d}t}, \vec{v} \rangle_\gamma}{c_\gamma} \cdot \gamma', \tag{5.1}$$

or equivalently we claim

$$\nabla_F^\gamma(\vec{v}) = \nabla(\vec{v}) - c_\gamma^{-1} \langle \gamma', \vec{v} \rangle_\gamma \cdot \nabla(\gamma') + c_\gamma^{-1} \langle \frac{D\gamma'}{\mathrm{d}t}, \vec{v} \rangle_\gamma \cdot \mathrm{d}t \otimes \gamma',$$

where the right side is visibly a connection (check!). It suffices to compare the connections on local frames, due to the Leibnitz Rule. By working with the global frame $\{\gamma'\}$ of L and local frames of $L^\perp$, it suffices to compare the two formulas on γ' and on a vector field $\vec{v}$ (over an open in I) that is orthogonal to γ'. The case of γ' works because $\frac{D\gamma'}{\mathrm{d}t}$ is orthogonal to γ' (thanks to the constancy hypothesis and (3.3)), and the case of $\vec{v}$ orthogonal to γ' works because $\langle \frac{D\gamma'}{\mathrm{d}t}, \vec{v} \rangle_\gamma = -\langle \gamma', \frac{D\vec{v}}{\mathrm{d}t} \rangle_\gamma$ (apply (3.3) with $s_1 = \vec{v}$ and $s_2 = \gamma'$).

The initial term in (5.1) is the Levi-Civita covariant derivative along γ, and the other terms can be nonzero only over those time intervals when the Levi-Civita acceleration $\frac{D\gamma'}{\mathrm{d}t}$ is not identically zero, which is to say that γ is not a geodesic (so for positive-mass particles in General Relativity this means it is not in free fall). Hence, the Fermi-Walker connection agrees with the Levi-Civita connection over opens in I where γ is a geodesic. The converse holds too: since $\frac{D\gamma'}{\mathrm{d}t}$ and γ' are orthogonal with γ' nowhere zero (as $c_\gamma \neq 0$), if the Fermi–Walker and Levi-Civita connections agree over an open $J \subseteq I$ then both extra terms in (5.1) must vanish identically for any $\vec{v}$ over an open subset of J. Looking at the final term in (5.1), this forces $\frac{D\gamma'}{\mathrm{d}t}$ over J to be orthogonal to all sections of $\gamma^*(TM)$ over opens in J, and so by non-degeneracy of the metric tensor we deduce the identity $\frac{D\gamma'}{\mathrm{d}t}|_J = 0$ that is the geodesic equation for γ (over J) with respect to the Levi-Civita connection. In the language of General Relativity, the rest frame of a particle with positive mass is carried forward across a time interval by the Levi-Civita connection if and only if the particle is in free fall during that time interval.

We now check the features of the Fermi–Walker connection that are crucial for its usefulness in solving the initial problem of defining the rest frame of a particle with positive rest mass: it is compatible with the pseudo-Riemannian metric tensor and $\nabla_F^\gamma(\gamma') = 0$ (i.e., γ' becomes parallel). The parallelism condition is obvious from (5.1) and the orthogonality of γ' and $\frac{D\gamma'}{\mathrm{d}t}$. As for metric compatibility, by the criterion (3.3) we just need to prove

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \vec{v}, \vec{w}\rangle_\gamma = \langle \frac{D_F^\gamma \vec{v}}{\mathrm{d}t}, \vec{w}\rangle_\gamma + \langle \vec{w}, \frac{D_F^\gamma \vec{w}}{\mathrm{d}t}\rangle_\gamma$$

for all sections $\vec{v}, \vec{w}$ of $\gamma^*(TM)$ over opens in I. The formula (5.1) for Fermi–Walker covariant differentiation and the fact that (3.3) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \vec{v}, \vec{w}\rangle_\gamma = \langle \frac{D\vec{v}}{\mathrm{d}t}, \vec{w}\rangle_\gamma + \langle \vec{v}, \frac{D\vec{w}}{\mathrm{d}t}\rangle_\gamma$$

(since the Levi-Civita connection is metric compatible) reduces us to proving (after multiplying through by c_γ) that

$$\left\langle \langle \gamma', \vec{v}\rangle_\gamma \cdot \frac{D\gamma'}{\mathrm{d}t} - \langle \frac{D\gamma'}{\mathrm{d}t}, \vec{v}\rangle_\gamma \cdot \gamma', \vec{w}\right\rangle_\gamma + \left\langle \vec{v}, \langle \gamma', \vec{w}\rangle_\gamma \cdot \frac{D\gamma'}{\mathrm{d}t} - \langle \frac{D\gamma'}{\mathrm{d}t}, \vec{w}\rangle_\gamma \cdot \gamma'\right\rangle_\gamma$$

vanishes. This vanishing is obvious by inspection.

6. A computation of flat frames

Suppose $H \subseteq \mathbf{R}^{n+1}$ is a smooth hypersurface and $\gamma : I \to H$ is a path. Let $\vec{v}$ be a vector field on H along γ (that is, $\vec{v} \in (\gamma^*(TH))(I)$). We give H the induced metric tensor from $\mathbf{R}^{n+1}$. Covariant differentiation on H may look painful, but on the flat space $\mathbf{R}^{n+1}$ it is trivial: we just do the classical differentiation of the smooth map $\vec{v}_0 : I \to \mathbf{R}^{n+1}$ that "is" $\vec{v}$ (viewing $\mathbf{R}^{n+1}$ as $\mathrm{T}_{\gamma(t)}(\mathbf{R}^{n+1})$ for all t). By Theorem 4.4, $\frac{D\vec{v}}{\mathrm{d}t}$ at a point $\gamma(t_0) \in H$ is the orthogonal projection of the derivative vector $\vec{v}_0'(t_0)$ from $\mathbf{R}^{n+1} = \mathrm{T}_{\gamma(t_0)}(\mathbf{R}^{n+1})$ into the hyperplane $\mathrm{T}_{\gamma(t_0)}(H)$. In particular, by locally computing the orthogonal bundle projection $T(\mathbf{R}^{n+1})|_H \to TH$ we can write down the differential equation $\frac{D}{\mathrm{d}t} = 0$ that expresses parallelism for vector fields along paths in H. (Of course, writing down an equation explicitly doesn't always necessarily constitute genuine progress.)

We shall work out this formalism in the first non-trivial case, the sphere in 3-space. Pick $0 < r_0 \leq 1$ and let $S^2 \subseteq \mathbf{R}^3$ be the unit sphere with induced metric tensor. Let $\gamma : \mathbf{R} \to S^2$ be the path

$$\gamma(t) = \left(r_0 \cos(t), r_0 \sin(t), \sqrt{1 - r_0^2}\right)$$

that parameterizes the circle $S^2 \cap \{z = \sqrt{1 - r_0^2}\}$ on S^2 with radius r_0 that is parallel to and above (or in) the xy-plane. We wish to compute a flat frame for $\gamma^*(TS^2)$ with respect to (the pullback of) the Levi-Civita connection ∇ on S^2. Let $\{\rho, \theta, \phi\}$ be the usual spherical "coordinates" on $\mathbf{R}^3$.

We will give two ways to write down the equation $\frac{D\vec{v}}{\mathrm{d}t} = 0$ for $\vec{v}(t) = f(t)\partial_\theta|_{\gamma(t)} + g(t)\partial_\phi|_{\gamma(t)} \in \mathrm{T}_{\gamma(t)}(S^2)$: the "classical" extrinsic method indicated above for hypersurfaces (orthogonal projection of a covariant derivative in the ambient flat space $\mathbf{R}^3$), and an intrinsic method by computing the Christoffel symbols from the metric tensor in $\{\theta, \phi\}$ "coordinates" on S^2. (There is a third way that is much more insightful and geometric, and requires virtually no calculation beyond high school trigonometry; this rests on studying the cone C tangent to the sphere along γ. The point is that $\gamma^*(TC) = \gamma^*(T(S^2))$ inside of $\gamma^*(T(\mathbf{R}^3))$ and the metric tensor of C induces a Levi-Civita connection over the cone that is flat in the sense of the discussion following Theorem 3.11. Thus, shifting covariant differentiation on $\gamma^*(T(S^2)) = \gamma^*(TC)$ from the viewpoint of the sphere to the viewpoint of the cone via several applications of Theorem 4.4 makes parallel transport very easy to visualize without any calculations. Of course, when γ is the equator this cone is really a cylinder, or better is a cone in the real projective plane with vertex on the line at infinity.)

The basic problem is to compute $\frac{D}{\mathrm{d}t}(\gamma^*(\partial_\theta))$ and $\frac{D}{\mathrm{d}t}(\gamma^*(\partial_\phi))$. First we will use the extrinsic method, so we have to shift to the standard flat frame $\{\partial_x, \partial_y, \partial_z\}$ of $T(\mathbf{R}^3)$. Recall that in $T(\mathbf{R}^3)|_{S^2}$ we have (by easy calculation from the Chain Rule)

$$\partial_\theta = -\sin\theta\sin\phi\,\partial_x + \cos\theta\sin\phi\,\partial_y,\ \ \partial_\phi = \cos\theta\cos\phi\,\partial_x + \sin\theta\cos\phi\,\partial_y - \sin\phi\,\partial_z$$

$$\partial_\rho = \cos\theta\sin\phi\,\partial_x + \sin\theta\sin\phi\,\partial_y + \cos\phi\,\partial_z.$$

Since $\gamma(t)$ has $\{\theta,\phi\}$-coordinates $\theta(t) = t$ and $\phi(t) = \sin^{-1}(r_0) \in (0, \pi/2]$ (as $\sqrt{1-r_0^2} \geq 0$), this gives the 1-parameter formulas

$$\partial_\theta|_{\gamma(t)} = -r_0\sin t\,\partial_x|_{\gamma(t)} + r_0\cos t\,\partial_y|_{\gamma(t)},\ \ \partial_\phi|_{\gamma(t)} = \sqrt{1-r_0^2}(\cos t\,\partial_x|_{\gamma(t)} + \sin t\,\partial_y|_{\gamma(t)} - r_0\,\partial_z|_{\gamma(t)}),$$

$$\partial_\rho|_{\gamma(t)} = r_0\cos t\,\partial_x|_{\gamma(t)} + r_0\sin t\,\partial_y|_{\gamma(t)} + \sqrt{1-r_0^2}\,\partial_z|_{\gamma(t)}.$$

Computing covariant derivatives along γ in $\gamma^*(T(\mathbf{R}^3))$ (we use the notation $D_{\mathbf{R}^3}/\mathrm{d}t$ to distinguish this from covariant differentation $D/\mathrm{d}t$ in $\gamma^*(TS^2)$) amounts to just t-differentiation of the coefficients in these formulas since the standard frame $\{\partial_x, \partial_y, \partial_z\}$ of $T(\mathbf{R}^3)$ is flat with respect to the Levi-Civita connection from $\mathbf{R}^3$. This gives

$$\frac{D_{\mathbf{R}^3}}{\mathrm{d}t}(\gamma^*(\partial_\theta)) = -r_0\cos t\,\partial_x|_{\gamma(t)} - r_0\sin t\,\partial_y|_{\gamma(t)},$$

$$\frac{D_{\mathbf{R}^3}}{\mathrm{d}t}(\gamma^*(\partial_\phi)) = \sqrt{1-r_0^2}(-\sin t\,\partial_x|_{\gamma(t)} + \cos t\,\partial_y|_{\gamma(t)}).$$

For $p \in S^2$, $\partial_\rho|_p$ is a unit vector in the line $\mathrm{T}_p(S^2)^\perp \subseteq \mathrm{T}_p(\mathbf{R}^3)$, so orthogonal projection from $\mathrm{T}_p(\mathbf{R}^3)$ onto $\mathrm{T}_p(S^2)$ is $\vec{v} \mapsto \vec{v} - \langle\vec{v}, \partial_\rho|_p\rangle_p\partial_\rho|_p$. We use this formula with $p = \gamma(t)$ to compute the orthogonal projection of the $\mathbf{R}^3$-covariant derivatives of ∂_θ and ∂_ϕ along γ into $\mathrm{T}_{\gamma(t)}(S^2)$; these projections are respectively equal to the covariant derivatives within S^2 along γ. With a little algebra, the answer is

$$\text{(6.1)}\qquad \frac{D}{\mathrm{d}t}(\gamma^*(\partial_\theta)) = -r_0\sqrt{1-r_0^2}\cdot\gamma^*(\partial_\phi),\ \ \frac{D}{\mathrm{d}t}(\gamma^*(\partial_\phi)) = \frac{\sqrt{1-r_0^2}}{r_0}\cdot\gamma^*(\partial_\theta).$$

Let us now derive these same equations by working intrinsically with the induced metric tensor on the sphere. Computing the inner product on $\mathrm{T}_p(S^2)$ by working in $\mathrm{T}_p(\mathbf{R}^3)$ for $p \in S^2$ gives that the vector fields $\partial_\theta, \partial_\phi$ on S^2 (away from the north and south poles) are pairwise perpendicular and ∂_ϕ is a unit vector field but ∂_θ has length $\sin\phi$. That is, the metric tensor is $\sin^2\phi\,\mathrm{d}\theta^{\otimes 2} + \mathrm{d}\phi^{\otimes 2}$. Hence, with local coordinates $x_1 = \theta$ and $x_2 = \phi$ away from the poles we have

$$(g_{ij}) = \begin{pmatrix}\sin^2\phi & 0\\ 0 & 1\end{pmatrix},\ \ (g^{\alpha\beta}) = (g_{ij})^{-1} = \begin{pmatrix}\sin^{-2}\phi & 0\\ 0 & 1\end{pmatrix}.$$

In particular, $\partial_{x_k}g_{ij} = 0$ except for perhaps the case $i = j = 1$ and $k = 2$: $\partial_\phi g_{11} = 2\sin\phi\cos\phi$ (this vanishes on the equator). The general formula for Christoffel symbols in terms of the metric tensor as in Theorem 4.1 therefore simplifies considerably:

$$\nabla_{\partial_\theta}(\partial_\theta) = \Gamma^1_{11}\partial_\theta + \Gamma^2_{11}\partial_\phi,\ \ \nabla_{\partial_\phi}(\partial_\theta) = \nabla_{\partial_\theta}(\partial_\phi) = \Gamma^1_{12}\partial_\theta + \Gamma^2_{12}\partial_\phi,\ \ \nabla_{\partial_\phi}(\partial_\phi) = \Gamma^1_{22}\partial_\theta + \Gamma^2_{22}\partial_\phi,$$

and $\Gamma^k_{11} = (-1/2)(\partial_\phi g_{11})g^{2k}$, $\Gamma^k_{12} = (1/2)(\partial_\phi g_{11})g^{1k}$, and $\Gamma^k_{22} = 0$. (Recall that the general equation $\nabla_{\partial_{x_i}}(\partial_{x_j}) = \nabla_{\partial_{x_j}}(\partial_{x_j})$ holds for any symmetric connection ∇, since $[\partial_{x_i}, \partial_{x_j}] = 0$.) Thus:

$$\nabla_{\partial_\theta}(\partial_\theta) = -\sin\phi\cos\phi\,\partial_\phi,\ \ \nabla_{\partial_\phi}(\partial_\theta) = \cot\phi\,\partial_\theta,\ \ \nabla_{\partial_\phi}(\partial_\phi) = 0.$$

This gives

$$\nabla(\partial_\theta) = -\sin\phi\cos\phi\,\mathrm{d}\theta\otimes\partial_\phi + \cot\phi\,\mathrm{d}\phi\otimes\partial_\theta,\ \ \nabla(\partial_\phi) = \cot\phi\,\mathrm{d}\theta\otimes\partial_\theta.$$

Pulling this back along the path γ given by $t \mapsto (t, \sin^{-1}(r_0))$ in $\{\theta,\phi\}$-coordinates, and noting that $\cot\phi_0 = \sqrt{1-r_0^2}/r_0$ when $\sin\phi_0 = r_0$ (with $\phi_0 \in (0, \pi/2]$ forcing the non-negative square root), we get

$$(\gamma^*(\nabla))(\gamma^*(\partial_\theta)) = \gamma^*(\nabla(\partial_\theta)) = -r_0\sqrt{1-r_0^2}\,\mathrm{d}t\otimes\gamma^*(\partial_\phi),$$

$$(\gamma^*(\nabla))(\gamma^*(\partial_\phi)) = \gamma^*(\nabla(\partial_\phi)) = \frac{\sqrt{1-r_0^2}}{r_0}\,\mathrm{d}t\otimes\gamma^*(\partial_\theta).$$

Dividing out the $\mathrm{d}t$ (i.e., using $\frac{Ds}{\mathrm{d}t} = \nabla_{\partial_t}(s)$ for a connection on a vector bundle over an interval; cf. Remark 2.1), we obtain exactly (6.1) once again.

Let us now use (6.1) to write down and solve the system of linear ODE's satisfied by the coefficient functions for a flat vector field along γ in S^2. Writing $\vec{v}(t) = f(t)\partial_\theta|_{\gamma(t)} + g(t)\partial_\phi|_{\gamma(t)}$, the above covariant derivative calculation coupled with the Leibnitz rule for covariant differentiation gives (after some algebra)

$$\frac{D\vec{v}}{\mathrm{d}t} = (f'(t) + r_0^{-1}\sqrt{1-r_0^2}g(t))\partial_\theta|_{\gamma(t)} + (g'(t) - r_0\sqrt{1-r_0^2}f(t))\partial_\phi|_{\gamma(t)}.$$

Hence, the flatness condition is the system of linear ODE's $f' = -(\sqrt{1-r_0^2}/r_0)g$ and $g' = r_0\sqrt{1-r_0^2}f$ on the entire real line. In the special case $r_0 = 1$, which is to say the equator γ_0, these just say $f' = 0$ and $g' = 0$. Hence, along the equator γ_0 a flat frame is given by $\{\partial_\theta|_{\gamma_0}, \partial_\phi|_{\gamma_0}\}$ (or any invertible constant linear combination of these). For $0 < r_0 < 1$, we have $g = -(r_0/\sqrt{1-r_0^2})f'$ and

$$\frac{-r_0}{\sqrt{1-r_0^2}}f'' - r_0\sqrt{1-r_0^2}f = 0.$$

This latter equation is $f'' + (1-r_0^2)f = 0$ with $1 - r_0^2 > 0$, and a basis for its 2-dimensional space of solutions is given by the functions $\sin(t\sqrt{1-r_0^2})$ and $\cos(t\sqrt{1-r_0^2})$ on $\mathbf{R}$. Using each of these as f and solving for g in each case, we get the flat frame

$$\left\{\sin(t\sqrt{1-r_0^2})\,\partial_\theta|_\gamma - r_0\cos(t\sqrt{1-r_0^2})\,\partial_\phi|_\gamma,\ \ \cos(t\sqrt{1-r_0^2})\,\partial_\theta|_\gamma + r_0\sin(t\sqrt{1-r_0^2})\,\partial_\phi|_\gamma\right\}. \tag{6.2}$$

Note, as a safety check, that these two vector fields along γ do have constant speed (both r_0, in fact) and constant inner product against each other (namely, $1 - r_0^2$), as we know must be the case for parallel transport with respect to the Levi-Civita connection (Example 3.5). Observe also that in the "limit" $r_0 \to 1^-$ (so γ "approaches" the equator γ_0) the common speed r_0 tends to a nonzero limit and the frames (6.2) "coverge" to the pair $\{-\partial_\phi|_{\gamma_0}, \partial_\theta|_{\gamma_0}\}$ that is a flat frame along the equator γ_0. That there is limiting behavior to a flat frame on the equator is of course not a coincidence, and more precisely it can be theoretically predicted without explicit global calculations by computing on *one* fiber (say $t = 0$) and appealing to the theorem on smooth dependence of solutions to first-order ODE's under variation of parameters and initial conditions. Do you see what the argument is?

Remark 6.1. The flat vector fields $t \mapsto \vec{v}(t)$ on S^2 along γ are constant linear combinations of the flat frame in (6.2). Clearly $\gamma(t_1) = \gamma(t_0)$ if and only if $t_1 - t_0 \in 2\pi\mathbf{Z}$ (by inspection of the definition of γ), but when does it happen that there is a t_0 and a nonzero $n_0 \in \mathbf{Z}$ such that $\vec{v}(t_0 + 2\pi n_0) = \vec{v}(t_0)$ in $\mathrm{T}_{\gamma(t_0)}(S^2) = \mathrm{T}_{\gamma(t_0+2\pi n_0)}(S^2)$? Of course, by parallel transport the same identity must then hold with any $t \in \mathbf{R}$ in the role of t_0 (using the same n_0), so it is the same to ask when parallel transport along γ is a periodic process. The problem is to determine when every constant linear combination of the frame in (6.2) is invariant under $t \mapsto t + 2\pi n_0$ for some nonzero $n_0 \in \mathbf{Z}$. (The best scenario is $n_0 = \pm 1$, which means that parallel transport once around the circle is the identity.) It is easy to check (do it!) that for $B \neq 0$ in $\mathbf{R}$ and $A_1, A_2 \in \mathbf{R}$ not both zero, the function $t \mapsto A_1 \sin(Bt) + A_2 \cos(Bt)$ is invariant under $t \mapsto t + c$ if and only if $c \in (2\pi/B)\mathbf{Z}$. Thus, γ admits a *nonzero* periodic flat vector field on S^2 if and only if the radius r_0 satisfies the arithmetic condition $\sqrt{1 - r_0^2} \in \mathbf{Q}$, in which case *every* nonzero flat vector field on S^2 along γ is periodic with minimal period equal to 2π times the denominator of the rational number $\sqrt{1 - r_0^2} \in \mathbf{Q} \cap [0, 1)$. In particular, the period is 2π only along the equator ($r_0 = 1$). This analysis of periodicity can actually be carried out by a much simpler and more geometric (and insightful) method, requiring no calculus whatsoever, by staring at the cone tangent to S^2 along γ.