# Affine Connections

Kevin Ostrowski

**Outline**: We are starting a new chapter on affine and Riemannian connections. These concepts, which enable us to differentiate vector fields, will bring us closer to a definition of curvature that is intrinsic to our manifold. Before I start, though, I would like to say a little bit to motivate the developments to come. In particular, since my background is in physics, I would like to briefly touch on some applications of the material we have covered so far. Then, in this context, I will give a few reasons for why we might want a definition of curvature that does not rely on some ambient space. Finally, I will work through the definitions of affine connections, covariant derivatives, and parallel-transported vectors.

**Motivation**: In simple problems in mechanics, we use Newton's 2nd law, $F(x(t), v(t)) = ma(t)$, to relate the acceleration of a particle to the force exerted on it. Given the initial position and velocity, finding the trajectory of the particle amounts to solving a second order initial value problem. Although physicists usually start thinking about velocities in a particular basis, we can reinterpret them from our geometric viewpoint. In particular, we think of them as living in the tangent space at each point in $\mathbb{R}^3$. Now, this does not seem particularly helpful, since $\mathbb{R}^3$ is a flat space. However, in more advanced formulations of mechanics, this perspective is useful.

In Lagrangian mechanics, the differential equations that govern the evolution of a system (called the equations of motion) are obtained from the Lagrangian function, $\mathcal{L} : TM \to \mathbb{R}$, where $M$ is the so-called configuration space of the system. Explicitly, $\mathcal{L} = T - V$, where $T$ is the kinetic energy and $V$ is the potential energy. Points in the configuration space are in one-to-one correspondence with the allowed positions of the particles in the system. Therefore, specifying a point in $TM$ is like specifying all of the positions and velocities. In classical mechanics, this information is assumed to completely determine the future behavior.

Examples of the correspondence between manifolds and physical systems abound. For an object swinging on a pendulum with a fixed hinge, we could describe its positions with points in the circle $S$. If we allow the hinge to slide along a wire, we would instead use the product manifold $S \times \mathbb{R}$. The orientation of a rigid body with one point fixed is given by the Lie group $SO(3)$, which we identify with rotations of $\mathbb{R}^3$. As you can see, manifolds naturally enter the discussion any time objects have their motions constrained. They also have a place in discussions of systems of particles. For instance, the double pendulum has as its configuration space the product manifold of two circles, or the torus $T$. For $n$ generalized coordinates and $n$ associated velocities, we will have $n$ equations of motion. The advantage to the Lagrangian approach is that, no matter what coordinates we choose to use, the equations of motion (Euler-Lagrange equations) take the same form. Although this is a nice result, the story does not end there. A third formulation of mechanics offers other benefits.

In Hamiltonian mechanics, we work with the Legendre transform of the Lagrangian with respect to the velocities, the Hamiltonian function $\mathcal{H} : T^*M \to \mathbb{R}$. Typically, $\mathcal{H} = T + V$, or the total energy. We can also obtain the equations of motion (Hamilton's equations) from this function. Except, this time, the equations are first order in time, and they can be viewed as defining a vector field on the cotangent bundle. Then, the trajectories that the system follows for different initial conditions are just the integral curves. The cotangent bundle is a symplectic manifold, known as the phase space, and we are interested in transformations that preserve its structure, called canonical transformations (by physicists) or symplectomorphisms (by geometers). If you want to learn more about this topic, there is a mathematics course offered next quarter on symplectic geometry. In any case, it turns out that clever transformations can enormously simplify some problems in mechanics, and in a sense, these transformations are the pinnacle of the increasing levels of abstraction we have covered here.

We started with Newtonian mechanics, in which the position and velocity have obvious, physical meaning. In Lagrangian mechanics, we use generalized coordinates on manifolds, which, although different than usual Cartesian coordinates, still often have a physical interpretation in terms of angles, displacements, etc. Finally, in Hamiltonian mechanics, we treat coordinates and momenta on equal footing by allowing transformations that mix the two sets of variables. Although the transformed coordinates are not easily interpreted in terms of the physical problem, they may more quickly lead to solutions, and the transformation can be inverted to recover the physics.

Although it may seem like Hamiltonian mechanics is the most powerful approach, I would like to emphasize that all three formulations make the same physical predictions, and each has its own advantages. In particular, the different formulations relate to other subfields of physics. The Hamiltonian approach to classical mechanics ties to both statistical mechanics and nonrelativistic quantum mechanics (although it appears as an operator instead of a function). The Lagrangian approach connects more naturally to quantum field theory, where one often speaks of the Lagrangian density, related to the energies of the fields.

To summarize the development so far, we have seen that two important functions in classical mechanics are naturally defined on differentiable manifolds. But, in those cases, we had the option of considering the motion of the system in a larger, ambient space, namely $\mathbb{R}^{3n}$, where $n$ is the number of particles. Physicists do not always have this choice. In situations where gravitational effects are large, we can no longer pretend that time and space are independent of one another, or that spacetime is flat. In short, we need general relativity. Because spacetime is curved by the presence of mass/energy, we have to apply the techniques of differential geometry to understand motion within it. And, unlike in classical mechanics, we do not have a readily available ambient space to retreat to. After all, what can be taken to be outside of all spacetime? The question makes no sense from a physical perspective, so, as promised, we have a strong incentive to find an intrinsic definition for curvature. We will continue making progress towards that goal by talking about how to differentiate vector fields. Then, we can talk about curvature as a measure of how much a vector changes when it is transported around a small loop at a point. Before I talk about the intrinsic definitions, though, I will start by giving an extrinsic definition we can build from.

**Definition (Extrinsic)**: Suppose that we have a surface, $S \subset \mathbb{R}^3$, a parametrized curve, $c : I \subset \mathbb{R} \to S$, and a vector field, $V : I \subset \mathbb{R} \to \mathbb{R}^3$. Then, the derivative of that vector field does not, in general, belong to $T_{c(t)}S$. However, because we are working in $\mathbb{R}^3$, we can project out the component of $\frac{dV}{dt}(t)$ in this plane. We call this orthogonal projection the *covariant derivative of $V$ along $c$* and we denote it by $\frac{DV}{dt}(t)$.

We think of the covariant derivative as the derivative of V from the viewpoint of the surface. As a way of seeing this, you are invited to consider an object moving with constant speed along the equator of a sphere. At an intuitive level, it makes sense that the velocity vector is not changing along this curve from the viewpoint of the surface, even though it is changing directions in $\mathbb{R}^3$. An explicit calculation readily verifies that, if we define a vector field along this curve to be its velocity at each point, then the covariant derivative vanishes.

**Definition**: The *acceleration of a curve* is the covariant derivative of its velocity.

**Definition**: The zero acceleration curves are the *geodesics* of a surface.

In light of these definitions, the curve moving along the equator of a sphere at a constant speed is a geodesic of the sphere. In particular, it is a great circle of the sphere.

**Definition**: Let $\mathfrak{X}(M)$ be the set of all $C^\infty$ vector fields on $M$, and let $D(M)$ be the ring of real-valued $C^\infty$ functions defined on $M$. Then, an *affine connection*, $\nabla$, is a map $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ denoted by $(X, Y) \xrightarrow{\nabla} \nabla_X Y$, that satisfies the properties,

- $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$
- $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$
- $\nabla_X (fY) = f\nabla_X Y + X(f)Y$

for $X, Y, Z \in \mathfrak{X}(M)$ and $f, g \in D(M)$.

Note that the first bullet is a linearity property, the second is distributivity over addition, and the third is a type of Leibniz rule. How do we think of the affine connection? Informally, it is like taking the covariant derivative of $Y$ along integral curves of $X$, giving a new vector field on $M$.

**Definition (Intrinsic)/Proposition**: If $M$ has an affine connection, then, given a differentiable curve $c : I \to M$, and a vector field $V : I \to T_{c(t)}M$, we can uniquely define another vector field $\frac{DV}{dt}$ along c, called the *covariant derivative of $V$ along $c$*, which satisfies the following properties, for $W$ a vector field along $c$ and $f$ a differentiable function on $I$,

- $\frac{D}{dt}(V + W) = \frac{DV}{dt} + \frac{DW}{dt}$
- $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$
- If $V$ is induced by $Y \in \mathfrak{X}(M)$, i.e., $V(t) = Y(c(t))$, then $\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} Y$

The first property is distributivity over addition, and the second is a type of Leibniz rule. The third property is perhaps the most important, as it establishes that the affine connection gives rise to a covariant derivative on the manifold. We can show that the covariant derivative depends upon the values of the two vector fields at a point and how the differentiated field changes along the curve defined by the other.

**Proof**: Once we have shown existence and uniqueness on one coordinate patch, uniqueness on any overlapping coordinate patch forces the definition of the covariant derivative to be consistent on the regions of overlap. Therefore, we just need an explicit representation on one coordinate patch. The strategy is to use the third property to write the covariant derivative as the application of the affine connection in some set of coordinates. Then, applying the properties of the connection, we arrive at the following expression for the covariant derivative, $\frac{DV}{dt} = \sum_j \frac{dv_j}{dt} e_j + \sum_{ij} \frac{dx_i}{dt} v_j \nabla_{e_i}(e_j)$. Uniqueness comes from the fact that property three gives this explicit representation, and existence comes from observing that this representation also obeys properties one and two.

**Definition**: Let $M$ be a differentiable manifold with an affine connection, $\nabla$. A vector field along a curve $c : I \to M$ is called *parallel* when $\frac{DV}{dt}(t) = 0$ for all $t \in I$.

**Proposition**: Let $c : I \to M$ be a differentiable curve in $M$ and let $V_0$ be a vector tangent to $M$ at $c(t_0)$, $t_0 \in I$. In other words, $V_0 \in T_{c(t_0)}M$. Then, there exists a unique parallel vector field $V$ along $c$, such that $V(t_0) = V_0$.

**Definition**: The vector field in the proposition above is called the *parallel transport* of the vector along the curve.

**Proof**: The strategy here is similar to the other proof. First, suppose that the theorem holds when the curve is contained in a single coordinate neighborhood. Then, any compact interval mapped to a curve on the manifold can be covered with a finite number of coordinate neighborhoods, and we can define a unique vector field on each. But then, by uniqueness, the definitions have to agree on the intersections, so we can define the vector field over the whole interval. Therefore, we only have to prove the proposition on a single coordinate neighborhood. The existence and uniqueness on a single coordinate neighborhood comes from the coordinate representations of the covariant derivative and the vector being transported. Satisfying $V(t_0) = V_0$ and $\frac{DV}{dt} = 0$ simultaneously amounts to solving a linear initial value problem. From the theory of differential equations, we know that the solution exists and is unique, so the proposition is proved.

**Takeaways**: In this section, we have seen that,

- Choosing an affine connection uniquely defines a covariant derivative, that is, a way of differentiating vector fields along curves.
- Also, given an affine connection, we can uniquely extend a vector at a point on a curve to a vector field along that curve, which we call its parallel transport.

Where are we headed? Any more physics connections?

- In the next section, we see that requiring the affine connection to be compatible with the Riemannian metric in a certain sense gives a unique affine connection for every Riemannian manifold.
- In general relativity, the metric does not have the positive definiteness property, so technically it is a pseudo-Riemannian manifold. However, even in this case it is possible to construct a unique affine connection that respects the structure of the metric. The details are the content of exercise 9 of Do Carmo.

**References**: I followed the text of Do Carmo pretty closely when writing the mathematical definitions and propositions above; all credit for the logical developement goes to him. The physics material comes from V. I. Arnold's "Mathematical Methods of Classical Mechanics", Theodore Frankel's "The Geometry of Physics", Brian C. Hall's "Quantum Theory for Mathematicians" and "Lie Groups, Lie Algebras, and Representations", Gerald B. Folland's "Quantum Field Theory: A Tourist Guide for Mathematicians", and Bernard Schutz's "A First Course in General Relativity".