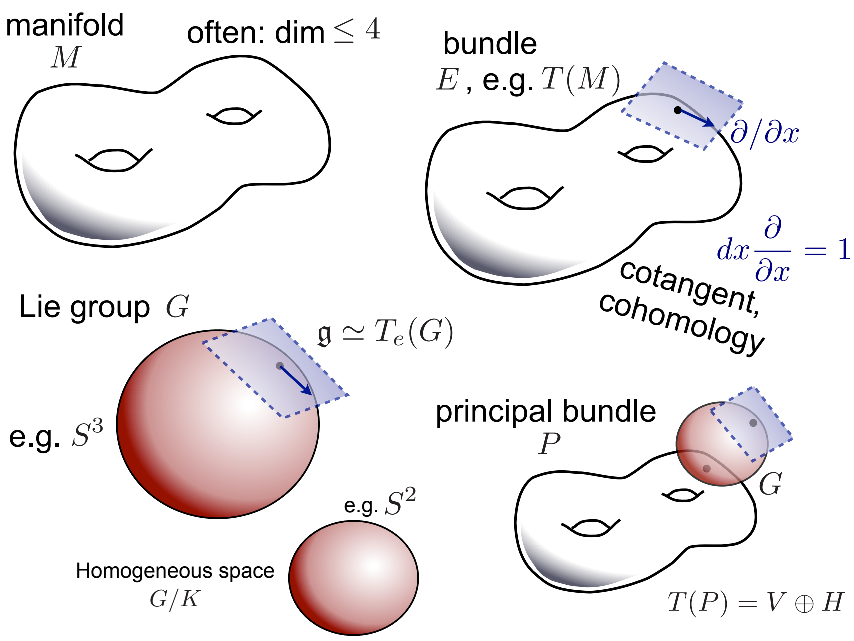
Notes for Conlon study questions: Answers

Marcus Berg, 2018

$\boxed{\textbf{WRITE DOWN ANSWERS TO THE QUESTIONS FIRST!}}$

Give yourself $0 - 2$ points per question, and only self-grade the chapter you actually looked at.



1 Chapter 1: Topological Manifolds

1. **Basic topology**. Wikipedia has some good individual entries, but it's a little hit-or-miss and sometimes too long-winded; there is a Wikibook on topology. Encyclopedia of Mathematics (encyclopediaofmath.org) is good, or Mathonline. A more idiosyncratic site is nLab. Even with all of these, I find it's good to have a well-tested standard paper book, like probably a topology you had as an undergraduate, if you want to increase the chance that it's actually correct. Another way to get a short summary is to look in a physics book like Wald's, that summarizes some of this in Appendix A. Some words are slightly different in Wald, like Conlon (p.87) calls "chart" the pair (U, φ), for Wald it's just φ. But U is implicitly associated with φ, and I think early on all these words are essentially compatible. There are some slightly nontrivial differences later (like with "connection" and "index notation") that we might talk more about.

2. **Example 1.1.3**. The two points 0 and 0′ are mutually exclusive, by the second bullet point in the example. So it's like having two copies of zero, which makes zero impossible to separate. Incidentally, Example 1.1.3 returns in Example 1.3.16, where Conlon makes the conceptually useful point that although you might not want to consider examples with double points that from our point of view are pathological (i.e. excluded from our definition of topological manifold in Chapter 1.2), you might *inadvertently* make a double point by some otherwise useful method (quotienting), and then you need to recognize it as potentially causing problems.

3. **Point of Chapter 1.1?**. The point of Chapter 1.1 is to make the definition of manifold in Chapter 1.2 seem reasonable. It is very useful to demand Hausdorff and 2nd countable. If the *dimension* had varied across the space, things would have become much more complicated. (Incidentally, there could be applications where dimension-varying is interesting, but we have to start somewhere.)

4. **Coordinate notation**. A vector is naturally "contravariant", in the language used in physics: the vector $\vec{v} = \sum_i v^i \vec{e}_i$, so the index is upstairs on the vector coordinates v^i. "Contra" is the opposite of "covariant" which varies along with something (see the vector X on p.65 and the co-vector ω on p. 187). What we would really like in the end is someting *invariant*, the idea roughly being "contravariant $\times$ covariant = invariant". For more on "contravariance", see below.

5. **Tangential direction**. This just means the direction *from the origin* to w is perpendicular to the tangent plane, which is obvious for a sphere. The origin is not part of the sphere (a sphere is of course always hollow, unlike a ball), so this definition uses "extrinsic" as opposed to "intrinsic" information, which Conlon will work hard to get away from. (To some students this seems like a "trick question" since the answer is obvious, but it lets me emphasize this distincion.)

6. **Bundle**. Yes, but he means that the triple $(S^n, T(S^n), p)$ (or more generally $(M, T(M), p)$) and the relation between them "is" the bundle, the bundle is not "just" the total space by itself.

7. **Parallelizable**. If at a point p tangent vectors *are* in fact linearly independent, the differential map is one-to-one (do Carmo, sec. 2-2, "regularity condition"). If not, the point is a singular point of the vector field of tangent vectors. (That doesn't mean it's uninteresting, see e.g. the discussion on p.74-75!)

8. **Klein bottle**. As you could have guessed this is done on Wikipedia. It is not too surprising that adding more dimensions to embed in makes things easier, but *time* makes it particularly easy: consider the path you take to work. It might be topologically nontrivial viewed as a traced-out "trajectory", but since you trace it out in *time*, you never self-intersect, even when you cross a path you went before. (For relativity fans: if you did encounter yourself, you accidentally went down a *closed timelike curve* of your manifold. Solutions of Einstein's equations that have such curves, like the Gödel solution, are often thrown out on grounds of being unphysical, but nothing prevents them from occurring in the first place.)

9. **Connected sum**. An empty tube, i.e. you cut out two disks and join their boundaries along a 2-dimensional surface, an empty tube.

10. **Classification theorem**. The value of the Euler characteristic χ (an integer), and whether it's orientable (yes/no). In general, a topological classification gives very limited information about (and is not "intended" to give detailed information about) curvature, like the Euler characteristic might tell you the value of an integral over the curvature over the entire surface, nothing how curvature might be distributed on the surface.

    Looking forward, topology naturally becomes more constraining if you add additional structure like that the surface is a Riemann surface (a surface that can be thought of as *complex*; we'll get back to that, but for now notice this is not the same as *Riemannian manifold*, which is a much more general concept, though it's of course the same person: Bernhard Riemann). For Riemann surfaces there is the Uniformization theorem: you can in fact always choose the curvature of a Riemann surface to be everywhere *constant*, in which case topological information about integrals over the curvature of course fix the curvature completely, given the total area of the surface (and the area can be thought of as a normalization, e.g. to unity, or 4π). But apart from very special situations like this, you should think of topology as giving no *detailed* (local) information about differential-geometry properties like curvature.

11. **Genus**. The connection is $\chi = 2 - 2g$ for boundaryless, oriented surfaces. With b boundaries and c number of P^2's attached ("c" for "crosscap", though Conway has argued that a better word for a sphere with a P^2 would be "cross surface"), it's $\chi = 2 - 2g - b - c$. (This is part of the answer to Exercise 1.3.14.) Check e.g. on Wikipedia that the Euler characteristic of P^2 is indeed $\chi = 2 - 1 = 1$. Note: Conlon calls it P^2, it's often called $\mathbb{R}P^2$ to distinguish it from $\mathbb{C}P^2$.

12. **Orientability, do Carmo vs. Conlon**. Conlon's definition by triangulation is "intrinsic", in the sense that it does not make reference to a direction of a normal vector pointing out into the space surrounding the surface, like in do Carmo. (The Gauss map is also defined by a normal vector, yet the Gauss curvature K – the determinant of the differential of the Gauss map – turns out to be intrinsic. But from that point of view it's a *theorem*, not an "obvious fact", and it's generally nice if an important feature is built in from the start, rather than seeming like miracle.). One example of a triangulation of the Möbius strip with 6 triangles (with orientations drawn in, to show there is no single globally consistent orientation) is given in Nakahara fig. 3.8.

13. **What is $\mathfrak{A}$?** To some of you this was completely obvious, for the rest of us it becomes a little clearer on p.18-19. (In general, sometimes one needs a little patience in this book.) The symbol $\mathfrak{A}$ is an index set, i.e. just a set of indices that labels a basis for a topology. In the TeX font $\mathfrak{Fraktur}$ (German "Fraktur"), the symbol $\mathfrak{A}$ is `\mathfrak A`.

14. **Cones**. Do Carmo at the end of section 2-2 uses the cone as a counterexample to a surface being everywhere regular. Note that it is *not* enough to show that the tip looks singular in *one* parametrization; it could have been regular in another (but in this case, isn't). For some purposes it's OK to just remove the tip. If you have do Carmo, read Example 1 in Sec 4-4 at this point: parallel transport around the base of a cone. Quotient spaces often have singularities, but one representative example that doesn't is the n-torus formed by dividing out an integer lattice: $T^n = \mathbb{R}^n/\mathbb{Z}^n$ (Exercise 1.3.24).

15. **Coset**. Conlon's statement is equivalent to $gH = \{gh : h$ an element of $H\}$ is the left coset of H in G with respect to g. Note that (as it nicely emphasizes on the Wikipedia page) it has to be stated with respect to which subgroup something is a left coset, since with respect to another subgroup in the same group it could be a right coset. (That subgroup is the conjugate of the previous subgroup.) Incidentally, all this is important in Ch.5 for homogeneous spaces that are then used in Ch.11 on principal bundles.

16. **Partition of Unity**. The sum is itself (by definition) constantly equal to one (unity), so its differential vanishes. The idea of splitting up "1" into pieces is so simple that it supposedly has many applications, like the Linkwitz-Riley audio filter, but it's not clear to me whether those people actually got the idea from topology.

17. **The topologist's sine curve**. First (again, for the physicist), "onto" means surjective, and "one-to-one" means injective.[1]

    Any sufficiently small neighborhood in $\mathbb{R}^2$ of a point p on the y-axis (the leftmost vertical part of the curve) intersects the curve in infinitely many neighboorhoods that are disjoint on the curve (components). Loosely, different "turns" of the sine curve close to the y-axis are infinitely "close" to each other and to the y-axis, but "close" is not well-defined since we haven't introduced a metric yet: in topology, it is sufficient to talk about neighborhoods overlapping as sets, or not.

18. **Other plane curves**. a) the cusp *is* a topological manifold (whether it's *smooth embedding* in the plane is not a question we're asking at the moment; that comes in Ch.2 and 3, so let's return to that one we defined this concept; for now we're asking for less.) b) The self-intersection of the "figure eight" lemniscate prevents it from being locally Euclidean (since it's 1-dimensional, "locally Euclidean" for the manifold itself means "locally $\mathbb{R}$", I just emphasize this since it's easy to start thinking it means "locally $\mathbb{R}^2$", but then you're starting to think about embedding, which is an equally important but separate question.)

    What's wrong with the figure eight? Think about whether "small" overlapping neighborhoods near the self-intersection are necessarily "small" overlapping neighborhoods in $\mathbb{R}$. (A test of this is: if you remove the actual self-intersection point, how many topological sets, or components, does the figure eight split into, compared to how many components $\mathbb{R}$ splits into?) So topologically speaking, while the cusp above was a topological manifold, a self-intersection (as in ∞) is not. To be clear, I'm not saying we're not *interested* in self-intersections — Conlon introduces the "self-intersection number" on p.281 — just that self-intersecting curves are not themselves 1-dimensional topological manifolds under Conlon's (very reasonable) definition. Finally, since the figure eight is not a 1-dimensional topological manifold it cannot globally be embedded into $\mathbb{R}^2$ (Definition 1.5.1) but it *can* be locally embedded, or *immersed* in $\mathbb{R}^2$ (Definition 1.5.2). Clearly immersion is a weaker requirement than embedding. Conlon gives the example of the Klein bottle, that can be immersed into $\mathbb{R}^3$, but not embedded due to self-intersection.

19. **Functor**. As it discusses there, for functors it is *opposite* meaning of for vectors, in the meaning given there. In a sense the meaning for vectors arises because the basis vector is $\mathbf{e}_i$, but the components are $\mathbf{x} = x^i\mathbf{e}_i$, so the *components* already transform "backwards" ("contra"), before going to the dual space. Just for fun, also take a look at the page about contravariance in computer science.

[1] There is an important potential confusion about this: physicists sometimes by "one-to-one" mean bijective (injective *and* surjective). In fact, literally I would argue this is what "one-to-one" should mean! But I will avoid this here: mathematicians usually define *function* as *strictly single-valued function*, whereas physicists tend to be more, let's say "flexible" on this point. If we do restrict our functions to be single-valued, then if you think about it (or look at the set diagrams at the Wikipedia link above), "one-to-one" should in fact mean "injective" as it does in mathematics. Note to any mathematicians getting worked up about physicists being "flexible": you're also actually flexible, for example at least most people I know don't shy away from speaking of the "Jacobi theta function" even though it's certainly a "multivalued function". Of course you can blame history there, but another example of this flexibility is "one-to-one-correspondence": is it surjective or bijective? By the way, since this footnote is already ridiculously long, I might as well add historical reference: the set-theory-style words *injective* and so on were introduced in the 1950s or so by Nicolas Bourbaki, a French mathematician who never existed. See also New Math and the French Mathématiques modernes (described as a *bourbakisme*), and a 1965 Peanuts cartoon about the frustration of American children with New Math.

2 Chapter 2: The Local Theory of Smooth Functions

1. **Definition of derivative**. In basic calculus you can, but that's because the scope is more limited. From the back blurb of Conlon's book: *The themes of linearization, (re-)integration and global versus local calculus are emphasized throughout*. It is enough to define the derivative like this. Incidentally, the differential map (that comes next) is a linear map, because it's not to be thought of as a general global derivative, but locally evaluated at a point p (mapping the tangent space).

2. **Differential map in basic calculus**. The composition of the curve $\alpha(t) = (u(t), v(t))$ and the map takes a little bit of writing, but it's a good thing to do:

$$\beta(t) = (F \circ \alpha)(t) = (x(u(t)), v(t)), y(u(t)), v(t)), z(u(t)), v(t)))$$

, we have

$$\begin{aligned} \beta'(0) &= \left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt}\right)\mathbf{e}_x + \left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt}\right)\mathbf{e}_y + \left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial z}{\partial v}\frac{dv}{dt}\right)\mathbf{e}_z \\ &= \begin{pmatrix} \partial_u x & \partial_v x \\ \partial_u y & \partial_v y \\ \partial_u z & \partial_v z \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} \equiv dF_q(w) \end{aligned}$$

because the 3×2 matrix is the Jacobian matrix, and the vector $(\dot{u}, \dot{v})^{\mathrm{T}}$ is clearly w. I wrote out the long first line because it makes it particularly evident that the curve could be changed to another curve $\tilde{\alpha}(t) = (\tilde{u}(t), \tilde{v}(t))$ without affecting the differential map itself. (To be clear, this means the matrix and the column vector separately change, but the map as a whole does not.)

3. **Chain rule in basic calculus**. The composition $G \circ F : \mathbb{R}^2 \to \mathbb{R}^2$, and a 2×2 matrix is expressed as 2×3 times 3×2:

$$d(G \circ F)_p = \begin{pmatrix} \partial_u \tilde{u} & \partial_v \tilde{u} \\ \partial_u \tilde{v} & \partial_v \tilde{v} \end{pmatrix} = \begin{pmatrix} \partial_x \tilde{u} & \partial_y \tilde{u} & \partial_z \tilde{u} \\ \partial_x \tilde{v} & \partial_y \tilde{v} & \partial_z \tilde{v} \end{pmatrix} \begin{pmatrix} \partial_u x & \partial_v x \\ \partial_u y & \partial_v y \\ \partial_u z & \partial_v z \end{pmatrix} = dG_{F(p)} \circ dF_p$$

Indeed, this last 3×2 matrix is the Jacobian matrix in dF_p above.

4. **Jacobian**. Conlon does allow $m = 1$, as evidenced for example in Lemma 6.2.3 on p. 186. The case $m = n = 1$ is the derivative of a single function of a single variable, the case $n = 1, m > 1$ is the gradient, and $n > 1, m > 1$ is a generic Jacobian matrix.

5. **Differential**. "Pushforward", in fact this is the main name on the Wikipedia page, to distinguish it from "total derivative", and "infinitesimal". I suggest to quickly browse through these three pages to orient yourself.

6. **Germ**. It's "cereal germ" in plant biology, which is the embryo of a *seed*, so it's something very basic from which other things develop. The metaphor is continued as "sheaf" with stalks. We will not use this language directly, but it is good at this point to have heard those words.

7. **Definition of $\mathcal{L}_X Y$**. The Wikipedia page says: "If X and Y are both vector fields, then the Lie derivative of Y with respect to X is also known as the Lie bracket of X and Y, and is sometimes denoted $[X, Y]$". So it's by definition. Many other references also do this. Conlon's approach (to be clear: the one he describes, he didn't invent it) is more formal, but more logical.

8. **Sard's Theorem**. It is used! As he says, he uses it in the next chapter. First in the proof of the Whitney embedding theorem (3.7.12), but also section 2.9B on nondegenerate critical values follows naturally after 2.9A on Sard's theorem, and 2.9B leads into Morse theory in the next chapter. (The details of both Morse and Whitney in Chapter 3 can be thought of as "advanced", but it is reasonable for a course like this to at least introduce them.)

9. **Example 2.9.4** Koch's snowflake is an example.

10. **The Hessian**. Actually no, Conlon says he makes a clever choice of coordinates (*geodesic normal coordinates*) that gives $g_{ij}(p) = \delta_{ij}$ and hence zero Christoffel symbol at that point. (In particular, the Christoffel symbol is a first derivative of the metric, but unlike what you might think from $g_{ij}(p) = \delta_{ij}$, the second derivative is still meant to be nonzero, since as we see in Ch.10, the Riemann curvature cannot be transformed away by a coordinate choice.) If we had not made this clever choice and kept the Christoffel symbols, there is an extra cross term (first derivative of f), as on the Wikipedia page. This looks suspiciously like the geodesic equation, and there is a connection, but remember f is a function, not the velocity vector of a curve. Conlon returns to this at the end of Ch.3.

3 Chapter 3: The Global Theory of Smooth Functions

1. **Bookkeeping**. As he says in Ch. 3.4, it will nicely fit with the discussion of Principal bundles in Ch.11. Also, as he says right away, the word "cocycle" is an attempt to connect to the discussion of Čech cohomology (p.285; It's not really the same since there it's addition, here in Ch.3 it's composition, but it's the same rough idea, and they agree in detail for 1 dimension!). So basically this way of presenting things is meant to maximally connect to the other things. In other presentations, often the other things (cohomology, but even bundles) look like completely separate things from manifolds, whereas really they are closely related. As Conlon emphasizes, he first linearizes (local theory, tangent space) then by "reintegration" (which involves cohomology) we will recover the whole space. Also, since a bundle is a manifold, Conlon emphasizes that it's fair to view a bundle over an n-dimensional manifold as a "special case" of an $n+m$-dimensional manifold where the m-dimensional sub-part $\pi^{-1}(p)$ of the "total" manifold (for points p on the manifold) is a *vector space*, which is usually not required of a general manifold. This is not the "usual view" of the relationship manifold/bundle at least how I originally learned it: the usual view is that a bundle is a structure *on top* a manifold (literally people use the terms "horizontal" and "vertical", see Ch.10), in which case it would certainly feel like a generalization, not a special case.

2. n**-plane bundle**. It's a vector bundle over M of fiber dimension n, i.e. there is a vector space $\pi^{-1}(x)$ over each point $x \in M$, and there is a trivializing neighborhood where the bundle of fibers looks like just a product. The point is that this may not globally be true and the typical example is Ex. 3.3.13 (Möbius bundle).

3. G**-structures**. There are many points, but one point is that this way to describe manifolds is pretty general and includes most things we might be interested in for this course. (Conlon calls the section "philosophical" but I would say he means he can only discuss things superficially at this point, not that the section is somehow out of the main line of argument, instead it is completely central.) Skim through the many examples on the Wikipedia page for orientation.

4. **Something simple**. It vanishes in the other chart, so the vector field has a "singular" (special) point. ("Stationary" is perhaps a better word.) (In fact, by the Poincaré-Hopf theorem, the total index is +2, which says something slightly more detailed about this singular point, but for now all we need to know is that there does not seem to be an everywhere (globally) nonzero vector field on S^2. Which you knew from Theorem 1.2.12 (Def. 3.3.10).

5. k**-plane distribution**. It's a subbundle with a block matrix ($\mathrm{Gl}(k, n-k)$) structure for the G-structure.

6. **Immersions.** If the torus is flat (and it can always be chosen flat), then it shouldn't have nonzero curvature anywhere. Why compact implies positive curvature is slightly more subtle but also explained there. Even such a torus can be isometrically immersed in $\mathbb{R}^4$ though! (We don't require a smooth manifold to have a metric, as below, whereas do Carmo usually always does.)

7. **Metric.** Well, in general in mathematics, it pays off to separate things that are logically distinct; we were able to talk about both topological manifolds and differentiable manifolds without making any assumptions about the existence of a metric, and then we should. A more concrete reason: if we want to generalize to less differentiable situations (non-smooth). Another example: existence of a metric doesn't mean it's known explicitly. For example, for compact Calabi-Yau manifolds there is no known metric, but there is a proof that they exist. So those of us intersted in them need to be able to work without knowing them. Which might not make a huge difference in practice, because it *is* known how many parameters it would depend on, which for some examples is 100 or more, so the metric could be too complicated to use explicitly anyway. (For noncompact Calabi-Yau manifolds, several examples are known, as we will explore a little.)

8. **The Schwarzschild metric.** It's static and spherically symmetric, so not really applicable to astrophysical black holes that *rotate*, given how they form, cf. how the solar system formed, by gravitational attraction from some original gas cloud. Rotating black holes are described by something more ellipsoidal-looking and (slightly) time-dependent, like the Kerr metric. (By "slightly" I mean what's called "stationary" in physics, technically there is a timelike Killing vector field.) But: the Schwarzschild metric just describes a static and spherically symmetric point mass, so if the physical surface of that point mass is outside the "event horizon" (Schwarzschild radius), then by the Birkhoff-Jebsen theorem, it describes the gravitational field of any localized static and mass distribution, such as the Sun, the Earth, or yourself! In fact, when calculating the 1.75 arc seconds bending of light by the Sun, as observed by Eddington in 1919 in the first (direct) test of general relativity, the path of light is a Schwarzschild geodesic. (Sternberger's Harvard lecture notes on how to calculate this, section 3.21.2, are reasonably useful.)

9. **The Friedmann-Lemaitre-Robertson-Walker (FRW) metric.** It's *the universe*. Or at least the standard ΛCDM model of the universe. Incidentally, this Howard Robertson was quite a tough referee for the *Physical Review*: as we'll get back to in Ch.7, he rejected a 1936 paper on general relativity submitted by Albert Einstein, the theory's sole creator.

10. **Winding number.** That's the turning number. I would have said that the Gauss map is by definition the map of the normal vector of a surface to a fixed origin, so that there is no "tangential Gauss map", but it's still clear what is meant; transport the tangent vector of a curve to the origin. In Do Carmo, this is the *Theorem of Turning Tangents*. A related result is Hopf's *Umlaufsatz*. Compare also physics concepts like Topological quantum number and Dislocations, we'll eventually get back to these.

11. **Induced Metric.** The 2nd fundamental form (a projection of the Weingarten map or "shape operator" L to a basis, related to the Gauss map of normal vector fields, see do Carmo and Ch.10) is induced by a normal vector field, the 1st fundamental form by a tangent vector field. So they are complementary (one could even say "orthogonal") views of an embedding (or immersion).

12. **Bad joke**. A lime bundle.

4 Chapter 4: Flows and Foliations

1. **Conlon's little introductions**. Global flows are related to global ($t \in \mathbb{R}$, not some subset) ordinary (i.e. single-variable) differential equations (ODEs), and foliations are related to global partial (i.e. multivariable) differential equations (PDEs). This distinction is nontrivial; to a differential equations person, PDEs are infinitely more difficult than ODEs. To appreciate the difference here, we need to be patient, but at least it should be clear from Ch.2. that flows as he defined them correspond to ODEs, in fact only 1st order ODE systems, but they can be (highly) nonlinear.

2. **First order?**. No. If you have an equation $y'' + y = 0$, just introduce $z = y'$, then you have that 1st order ODE and now the original 2nd order ODE looks like 1st order: $z' + y = 0$, i.e. by doubling the number of ODEs we get from 2nd to 1st order, and so on. (Some particularly creative student might already have noticed that this "doubling" sounds suspiciously like the "doubling" of dimensions when we go from a manifold (like S^2, which is of course 2-dimensional) to its tangent bundle (like $T(S^2)$, which is 4-dimensional). Indeed it is related.

3. **Local flows in Ch. 4 vs. local flows in Ch.2** . It's the exact same concept, but with one crucial new piece of notation (that we will later get rid of again!): the flow now has a superscript α as in $\Phi_t^\alpha(x)$. This is conceptually crucial: in Ch.2 we didn't yet have the definition of manifold as a patching together of coordiante charts $(U_\alpha, \varphi_\alpha)$. Now we do, so we should concretely think of each local flow as being defined in a local coordinate chart. A global flow on an abstract manifold will be the central idea, even though we often like to work in local coordinates, we should think what it means to change coordinates, i.e. change α on $\Phi_t^\alpha(x)$.

4. **Flows vs. flow lines**. For the flow line (also called "orbit" which will turn out to be related to "orbit of group action" in group theory), we emphasize the "time" dependence, but in the flow, we emphasize the dependence on initial conditions (initial point) q. Note that for 1st order ODE's, there's only one initial condition per variable. (For 2nd order, like Newton's laws $m\ddot{x} = F$ there's both position and velocity, but by the above answer, that can always be converted to double the number of 1st order ODEs).

5. **Complete vector field**. Probably not. (If you did, let me know!) It's described more intuitively (from my point of view!) in Ch. 10.4 (Complete Riemannian Manifolds), but it's true that the concept is more general than the context of Ch.10.

6. **Commuting flows**. Pick a small parallellogram and flow around it both ways; see Nakahara. Conlon does return to this in subsection 4.4 (Commuting Flows).

7. **Matrices**. By the exponential of a matrix, since $e^x = 1 + \ldots$, the exponential does not have zero determinant even if the matrix in the exponent does.

8. **Example 4.5.4**. As it says, it's overdetermined, i.e. it may not have a solution. (Analogy: a linear system of 4 equations in 3 variables may not have a solution), at least we need the integrability condition to avoid inconsistency. On the other hand, integrability does not guarantee nontrivial solutions, only that they have a chance to exist. The foliation here is by the surfaces $z = f(x, y)$, or equivalently, level sets of $f(x, y)$. (These two statements are the same, since $f(x, y)$ was defined by differential equations that had no undifferentiated $f(x, y)$, so there is always a constant of integration. For fixed $f(x, y)$, we can think of a family of surfaces $\ldots, f(x, y) - \epsilon$, $f(x, y)$ $f(x, y) + \epsilon, \ldots$. Alternatively, we ask for level sets of $f(x, y) = c$ for any c where such level sets exist.) Note that the discussion here is *local*; we certainly have no guarantee we can foliate globally by nice integral surfaces; for example our surfaces may miss a point, or hit some point twice. But that's sort of the point of manifolds; even if there is no single formula globally, we can still hope to piece together things in local charts. For a general definition of foliations, the

Mathworld is better (shorter) than Wikipedia (consider codimension $c = 1$ in that definition, for easy visualization).

The term comes from geology: rocks have layers that show the geologist the evolution in *time*, i.e. the foliation is by time. Another somewhat similar analogy is tree rings, that tell you how the tree grew as a function of time; some years there's a thick ring (leaf of a foliation, where "ring" doesn't mean it's circular, it can be very deformed), some years there's a thin ring. And, there is a one-to-one correspondence between time and the foliation of the tree trunk by rings: every piece of wood in the tree is on *some* ring, and every piece of wood in the tree is only on *one* ring. [2] (Can you be sure that a later ring cannot intersect an earlier ring?)

9. **Figure 4.5.1**. (p.143). Time "steps through" the level sets. So as time goes on, you switch topology from the cylinder to the plane, or conversely. By the way, foliations even have their own webpage, where a version of this example is featured: `foliations.org`, and some normal vectors are drawn, that point in all kinds of different directions.

10. **Penrose-Hawking**. Those theorems were the main part of Hawking's thesis on geodesic completeness (or non-completeness) as applied to *cosmology*, the study of the universe as a whole. In popular science terms (as used in the aforementioned movie), Penrose and Hawking were trying to prove the existence of a beginning of time and explore whether time could end. Since then, new data changed cosmology in two important (and logically distinct, but in some accounts similar-sounding) ways: inflationary cosmology (first fraction of a second after Big Bang) and dark energy (7 billion years to present, and to future, it is believed), that both violate the (positive-energy) assumptions of those theorems. The experimental/observational foundation for those two advances were explored starting in the early 1990s. Nowadays, the theorems are perhaps more interesting for black hole formation in numerical relativity, where the assumptions presumably are valid.

11. **Gradient**. Again, Jones is talking about the gradient of a function f that is not necessarily the function g that defined our surface (in Conlon, "manifold") as a level set $g = 0$. In Jones, "intrinsic" gradient of f means ∇f *projected to the tangent plane* of this surface (figure on p.5, Ch.5), by using the ambient inner product (the usual in calculus). In this language, Conlon's choice is the obvious choice $f = g$, which means the "intrinsic" gradient is the usual gradient. (So we don't need to talk about this any more for Chapter 4.2, but it could come in handy later to have this in mind: on a curved surface, the gradient of a function f that is not "adapted" to the level sets of g has no reason to be perpendicular to them, unlike the gradient of the surface-defining g itself, which is perpendicular, but this in itself assumes the ambient metric is Euclidean.)

12. **The "clock" foliation**. In principle yes (this is discussed in detail in for example Henneaux & Teitelboim's book, but this is way more advanced than we need here). The point is that any theory that is reparametrization invariant has the possibility of choosing $H = 0$, but it's not automatic, it would need to be arranged. In practice (e.g. in simulations in numerical relativity) the Hamiltonian constraint is not necessarily evolved (it's a *constraint*, not an equation of motion) but needs to be monitored at all times to see that numerical instabilities don't develop.

[2] This assumes the tree never temporarily literally *stops* growing at any point in its life, in which case it's not a good foliation by time, since several points in time would then correspond to the same ring in the tree. Apparently real trees do sometimes such have "missing rings", but it seems a reasonable definition of this simple analogy to exclude such cases and assume monotonously increasing growth.

5 Chapter 5: Lie Groups and Lie Algebras

1. **Lie algebra**. This only works for matrix groups; for more general groups, define left-invariant vector fields $L_{gh} = gh$ (that mimic how it works for matrices).

2. **Exponential map**. This only works for matrix groups; for more general groups, the homomorphism $c(s+t) = c(s)c(t)$ works.

3. **Hilbert's Problems**. Smooth manifold; depending on how you count, it was solved by Yamabe in the 1950s (with additional statements making it more precise).

4. **Cartan classification**. The classification of algebras lead to the List of simple Lie groups. Four families A_n, B_n, C_n, and D_n – with five exceptions E_6, E_7, E_8, F_4, and G_2. Look up what they are in more ordinary terms, like which is SO(n)? What is Sp(n)?

5. **Dynkin.** Since so much about the correspondence between Lie groups and Lie algebras relies on exponentiation, I find it very important to have a constructive formula showing that the Lie group multiplication can be written purely in terms of commutators. In fact, if you do the calculation naively, it looks like there are some amazing cancellation that need to happen; for example in the quadratic term in the group multiplication $e^X e^Y$, there are a priori terms like X^2 and Y^2, not just $[X, Y]$. The answer was provided by Eugene Dynkin in 1947: the BCH formula. (Obviously, B, C, and H had some ideas about this before Dynkin, but not in explicit generality.)

6. **What's special about special orthogonal?** If the determinant is $+1$, for n odd that precludes matrices that are not connected to the identity, like $-\mathbf{1}$, a reflection of all coordinates. (For n even, it still precludes subsets being -1.) You might say, a reflection for sure has nothing to do with a rotation so we should think of them as different things and say only SO(n) has to do with rotations, but by the Cartan-Dieudonné theorem, you'd be wrong: rotations can be built out of reflections.

7. **Isotropy/stabilizer/little group**. It's either already travelling at the speed of light, or never will. The former (photon, gluon, or possibly graviton) has a velocity given by a massless/lightlike/null vector that points 45° up to the right. Because there is no free pseudoangle φ to specify when $v = 1$ (or $v = c$ in ordinary units), only the energy, this is the first class of base vector. For a massive/timelike vector, the base vector is different: the pseudoangle φ is freely specifiable (and in fact specifies the velocity). It is true that we can place ourselves in the coordinate system of a massive particle, where time is called "proper" time and the particle stands still – velocity vector straight up in a spacetime diagram – but then the freedom φ is "hidden" in that coordinate transformation and can be "noticed" by other observers travelling by, whereas the photon looks like it is travelling at the speed of light to all observers. So these are the two possible classes of base vectors, and they have different stabilizers. (They're SE(2) and SO(3), the former for the photon, the latter for massive particles like the "Z and W" particles, but you don't need to know that; that's typically discussed in an introductory course on relativistic quantum mechanics or quantum field theory.)

8. **Is S^2 a Lie group?** No. SO(n) is a normal subgroup of O(n), but there are no other continuous ones, only discrete ones (like Conlon's Γ). So we can't make S^2 as a Lie group, but Conlon does the second best: he makes S^{n-1} as homogeneous space dividing out by a non-normal stabilizer, which is pretty good too. But as we expect from the Poincaré-Hopf theorem for S^2, there will be some stationary point of a flow on S^2 (fixed point of a map), we'll see this in more detail later. It's fair to think of the Hairy ball theorem as related to the non-normal stabilizer. This means that also for higher-dimensional spheres S^n for $n \geq 3$, the non-normal stabilizer prevents them from being Lie groups, only homogeneous spaces – except for the "lucky" case of S^3, by its relationship with quaternions. (To be precise, I'm not saying that not being a Lie

group implies it's not parallelizable, as in the Hairy ball theorem: S^7 is also not a Lie group, but unlike S^2, it's parallelizable. Apart from S^7, all the parallelizable spheres S^0, S^1 and S^3 are Lie groups, so parallelizable is *weaker* than being a Lie group, at least for spheres. It's fair – if imprecise – to say that S^7 is the odd one out among higher-dimensional spheres; it's the "unit octonions", which aren't associative, so it's not a group at all, but it's a division algebra, which lets S^7 be parallelizable. It's the odd one out since as you go higher, this connection never recurs above S^7; the sedenions are not a division algebra. Conversely, among the lowest-dimensional spheres S^n for $n \leq 3$, what singles out S^2 as the odd one out is that it doesn't get a "cheap ride to parallelizability" by being a unit element of some division algebra, like quaternions, that are then used to show S^3 is a Lie group. Not just S^3 but also S^1 gets a cheap ride, as unit complex numbers, and trivially S^0, as unit real numbers (-1 and $+1$).)

9. **Stiefel manifold**. The source of all truth lists Stiefel spaces $V_{n,1}(\mathbb{R}^n) = S^{n-1}$, $V_{n,1}(\mathbb{C}^n) = S^{2n-1}$. (Conlon also says this, but he has some TeX typo for the complex so it's a little hard to read!). The "highest" Stiefel spaces are $V_{n,n}(\mathbb{R}^n) = \mathrm{O}(n)$ and $V_{n,n}(\mathbb{C}^n) = \mathrm{U}(n)$. These adjectives just mean that the lowest-dimensional frame you can have (and actually, kind of boring) in an n-manifold is $k = 1$, just a single vector going around. The highest is n, naturally, that's what we might normally ask about. But there's also an interesting answer for $n - 1$. (What will this be useful for later? It ties in naturally with a "frame bundle", or generally a principal bundle, which is Ch.11. More specifically, Stiefel spaces naturally lead to the .)

6 Chapter 6: Covectors and 1-Forms

1. **Egg crates**. It counts the number of times the vector w pierces the set of planes ("egg crate"). In "Gravitation" (affectionately known as MTW), they go completely crazy and add *sound effects*: each piercing makes a bell ring.

2. **Covector vs. cotangent vector**. Yes, I think he uses them interchangeably, p. 183-186, e.g. in Corollary 6.2.5 I would have expected "cotangent vector"; in general "covector" could refer to the dual of some other space than the tangent space.

3. **Corollary 6.2.5**. The physicist is missing that Conlon is saying clearly (by using germs and specifying the point x) that this is a local statement at a point. That's why he says "more generally" shortly thereafter. If you are concerned, immediately check Example 6.3.12 for a counterexample to the (wrong) idea that the local statement always generalizes.

4. **Exterior derivative**. It's true that $A^1(M) = \Gamma(T^*(M))$ is the space of smooth sections of the cotangent bundle (Definition 6.2.6), and in general not every function f_i in local coordinates $\omega = \sum_i f_i dx^i$ is the derivative of a single function f (cf. previous question), but here we are in fact acting with d on a single function f, so as long as we are working with 1-forms, it's the same idea. Usually the point of the exterior derivative d is that it generalizes the differential map to p-forms for $p > 1$, but that's Chapter 7 and 8.

5. **Leibniz rule for 1-forms**. It doesn't really specify how "big" *dt* has to be for this to make sense. Here we don't really need to talk about the "size" of *df* as an abstract object, but we haven't completely avoided the question; Conlon's original definition of derivative as a linear map specifically means that it is defined by neglecting nonlinearities, which we then reintroduce by integration.

6. **Monotonic change of parameter?**. I think they're consistent, they both split into two cases (like Lemma 6.3.2) but Lee makes a bigger deal of it.

7. **Fundamental theorem of calculus?** We just used it. We do it again more generally in Theorem 8.2.9.

8. **Path-independence**. One example (there are infinitely many) is ydx. Consider a rectangle with corners at $p = (0, 0)$ and $q = (1, 1)$. Calculate the line integral from p to q by going first up, then right, then do it again by going first right, then up. This is done in the freely downloadable thermodynamics book by Gould & Tobochnik, Ch.2, p.90. By Theorem 6.3.10, path-dependent is equivalent to inexact, so this gives an example of an inexact differential. Checking inexact directly means checking the obstruction to integrability, i.e. $a(x, y)dx + b(x, y)dy$ should satisfy $\partial a/\partial y = \partial b/\partial x$, which ydx doesn't: $a = y$ but $b = 0$. Checking failure of path-independence means exhibiting two paths along which the line integral gives two different values, as above. Geometrically, it's obvious that these two conditions are the same since they are both implied by the existence of a height map. (Note this is all local at this point in Conlon. The first example of a locally exact but globally inexact form is Example 6.3.12, which is very important and leads the way to cohomology in Ch.8.)

9. **Path-independence in thermodynamics**. To indicate that it's "infinitesimal", not that it's exact. We have separated those two parts of Leibniz's differential: for us, 1-forms are not infinitesimal anyway, there is no problem to write simply Q for a 1-form describing heat instead of δQ or $đQ$. As I mention in the video on differential forms, the distinction exact/inexact is important in thermodynamics since path-dependence of work W and heat Q means that neither work nor heat can be used to characterize the "state" of a system, they are only energy *transfers*. (The difference between the two is that heat is energy transfer by temperature gradient, and work is not heat.) If work and heat were exact differentials, a "cyclic process" that returns to its initial

state could not generate any work at all (e.g. produce electricity). In fact, most power plants do operate cyclically, typically on the Rankine cycle.

10. **Inexact differential made exact?**. For example, if we make the ansatz $c(x, y) = c(x)$, then the differential equation for c is $c\partial_y a = \partial_x(c)b + c\partial_x(b)$. Restricting even more, we could try a power, i.e. $c(x) = x^\alpha$, in which case $\partial_y a = \alpha b/x + \partial_x b$. Of course, no ansatz like this always works, but it seems reasonably plausible that in some case there could be a simple solution to a slightly generalized equation like $\partial_y a = \alpha b/x + \partial_x b$ even if there was no simple solution to $\partial_y a = \partial_x b$. Bonus question: by the above, I will write Q for the inexact heat 1-form, not some crazy thing like $đQ$. So now we can write $Q = TdS$, where T is temperature and S is *entropy*, as for example in Fitzpatrick's lectures. Entropy is a property, and can be used to characterize the state of a thermodynamic system. Of course, sometimes I would want to emphasize that there was a small heat transfer, then I would probably write δQ. (To first launch into a diatribe why something is terrible, then admit I sometimes do it myself, is of course the prerogative of every instructor.)

11. **Arctan and polar angle**. If you just use a single formula $\theta = \arctan y/x$, then $x = 1, y = 1$ manifestly gives the same answer as $x = -1, y = -1$ (where by "single formula" I mean you really picked a specific branch of arctan – see below). But intuitively we would like to mean $\theta = 45^\circ$ (1st quadrant) by the former and $\theta = 135^\circ$ (3rd quadrant) by the latter. Computer science people sometimes make a big deal that they invented a new multivalued function to deal with that: `atan2`. As you can see from the plots in that link[3], that function changes continuously as you go around. An even better is the complex log that also allows continuing above 2π, as long as you stay away from the origin. So the original problem of $\theta = 45^\circ$ vs. and $\theta = 135^\circ$ came both from the ratio y/x being insensitive to the reflection $(x, y) \mapsto (-x, -y)$, and from arctan by itself only giving values between $-\pi$ and π. One way to express a "multivalued arctan", is to give a case-by-case formula that depends on where (x, y) actually is and not just the ratio, so give two different formulas that each uses the single-valued version of arctan. This is what Conlon does[4]. The only remaining issue is the inevitable branch cut that we're used to from the complex logarithm or "Arg" (complex argument) function. (In fact, what we often do in polar coordinates is to flip back from $\theta = 2\pi$ to $\theta = 0$. We can now express that as forming a *quotient* of the "staircase-shaped" complex-logarithm space. But as usual with quotients, once we realize that's what we're doing, we might as well lift to the covering space, i.e. just allow the angle θ to continue increasing above 2π.)

[3] An example of the aforementioned big deal: the `atan2` Wikipedia page currently reads *"The atan2 function was first introduced in computer programming languages, but now it is also common in other fields of science and engineering. It dates back at least as far as the FORTRAN programming language..."*. It is of course true that computer scientists were the first to introduce it in computer science, but the formulation makes it sound like mathematicians learned how to solve the "mystery of $\theta = \arctan y/x$" from computer scientists, but needless to say, Riemann know how to take complex logarithms, and probably Argand before him.

[4] In more detail (and this is probably too elementary for most of you!): the original arctan only covers -90 to $+90$ degrees (i.e. 1st and 4th quadrants counting counterclockwise), and what Conlon then does is to switch $x \mapsto y$ and $y \mapsto -x$, i.e rotate by +90 degrees. Then you cover what was previously the 2nd and 3rd quadrants, but not the original 1st and 4th.

7 Chapter 7: Multilinear Algebra and Tensors

1. **Decomposable**. The 2-form $v_1 \wedge v_2 + v_3 \wedge v_4$ in $\mathbb{R}^4$ is not decomposable, i.e. it is *in*decomposable. This is easiest to see by looking at the opposites, the decomposables. In $\mathbb{R}^3$ or below, there are only decomposable p-forms; for example a 2-form $v_1 \wedge v_2$ in $\mathbb{R}^3$ is already decomposed, whereas $v_1 \wedge v_2 + v_2 \wedge v_3 = v_1 \wedge v_2 - v_3 \wedge v_2 = (v_1 - v_3) \wedge v_2$. So we see that to *not* be able to decompose, we need at least four independent basis objects. In general, decomposition is done easily with the `DifferentialGeometry` package `DecomposeForm`, as discussed on its Maplesoft help page. (This is a little clearer on the Mathworld page than on the Wikipedia page, by the way.)

2. **The metric tensor**. Not directly, in the sense that it is a "gravitational potential" like the mgh you learned about in high school, but gradients in space or time, for example time-dependent variations are physical. One obvious example are gravitational waves, the Nobel prize for physics 2017 (The Popular and Advanced information in the left-hand menu of that link are both read-worthy). (By the way, it's sometimes stated that gravitational waves "confirm Einstein's theory", which is certainly morally true, but curiously Einstein wrote a paper in 1936 saying gravitational waves don't exist. That paper by Einstein may not be the best piece of science ever, but it might be the best example of the peer review process review ever: the referee for *Physical Review* rejected a paper on general relativity by Albert Einstein![5] In the first 2 mins of this video I mention how it took until a meeting my supervisor organized in 1957 to recover from Einstein's confusion.)

3. **Exterior products of vectors**. Vectors can be wedged together, in fact that will be used in the next chapter. It's somewhat like the cross product in vector calculus. On the other hand (at least in my experience) wedging is a more "natural" operation on 1-forms than on vectors; the "natural" antisymmetric relation on vectors is the Lie algebra $[X, Y] = -[Y, X]$. There is a close relation between those two concepts, as we will see more of soon.

4. **Covariant and contravariant tensors, forms, symmetric tensors**.
Covariant tensors are sections of tensor products of the cotangent bundle. In local coordinates, $T = T_{ij}(dx^i \otimes dx^j)$. (Note that without further assumptions, here T_{ij} has no special symmetry, so it has n^2 independent components in n dimensions. It transforms in a particular way – as a product of 1-forms – under coordinate transformations, by the construction in this chapter.)
Contravariant tensors are sections of tensor products of the tangent bundle, i.e. vector fields. In local coordinates, $T = T^{ij}(\partial/\partial x^i \otimes \partial/\partial x^j)$. Recall that for the summation convention to work, we should think of 1/"down" = "up", so $\partial/\partial x^i$ has index "down". Again, also this T^{ij} has no special symmetry.
Now, k-forms are sections of the exterior power of the cotangent bundle, which are then identified with the exterior power of 1-forms. In local coordinates, $\omega = \omega_{ij} dx^i \wedge dx^j$, where $\omega_{ij} = -\omega_{ji}$. Due to antisymmetry, ω_{ij} are $n(n-1)/2$ independent components in n dimensions. (This is the number of components of an antisymmetric matrix; try it for $n = 2$ and $n = 3$.) This loss of independent components due to symmetry causes a little bit of complication how to use the summation convention efficiently while avoiding overcounting, see next chapter. Finally, symmetric tensors are $g = g_{ij}(dx^i \otimes dx^j)$, like the metric tensor (Ch.10, but it has appeared before.). Sometimes this is written $g = g_{ij}(dx^i \otimes dx^j + dx^j \otimes dx^i)$ to make the symmetry manifest. In any case, $g_{ij} = +g_{ji}$, and there are $n(n+1)/2$ independent components.

   Note that the things we started with – vector fields – ended up being "contra"-(opposite)-variant (components have index up, $X = f^i \partial_i$). As mentioned earlier, this is essentially a historical accident that comes from writing a vector in linear algebra as $\mathbf{v} = v^i \mathbf{e}_i$, so the components v^i of $\mathbf{v}$ transform "oppositely" (to the obvious basis vectors $\mathbf{e}_i$, that is).

[5] It is now known that the referee was American mathematician Howard Robertson, of the Robertson-Walker metric in cosmology, mentioned in an earlier section.

8 Chapter 8: Integration of Forms and de Rham Cohomology

1. **Pre-Poincaré Lemma**. It wasn't a question.

2. **Summation**. It's to avoid overcounting, as discussed in the previous chapter answers. Due to antisymmetry of the wedge, ω_{ij} are $n(n-1)/2$ independent components in n dimensions, not n^2 as for a generic object with 2 indices in n dimension. This causes a little bit of complication how to use the summation convention efficiently while avoiding overcounting, as commented upon above. The indices in Conlon's sum are strictly ordered, which is sometimes called "strict components". You can also consider unrestricted sums that do overcount, and correct for it by an explicit factor in front.

3. **Singular vs. simplicial homology**. There are several, one is that one vertex in singular homology as used by Conlon is special, as you can see in his Fig.8.2.1 and 8.2.2. This is a subcase on the Wikipedia page, under "Simplexes with an "orthogonal corner"". Compare Gram determinant, a linear algebra concept that finds one application in the quantum field theory of particle physics: it gives you the combinations of momenta that occur in scattering amplitudes in certain situations, sort of a baby version of what later became the amplituhedron, that is of interest to some mathematicians.

4. **Ch.8.4-8.5**. $H^0(S^2) = \mathbb{R}$, $H^1(S^2) = 0$, $H^2(S^2) = \mathbb{R}$. The first just means we can write functions on the sphere. $H^1(S^2) = 0$ means that all loops (closed curves) on the 2-sphere can be contracted to a point. (Here we should be careful, because this statement sounds like *homotopy*, which is similar to but in general not identical to *homology*; the latter is an equivalence class of cycles modulo boundaries. But, there are no cycles on the sphere that cannot be deformed into each other, so here the distinction doesn't matter. On the torus there are two obvious cycles that cannot be deformed into each other. Homology is "designed to detect holes", as Conlon beautifully described on p.254.) The top cohomology $H^2(S^2) = \mathbb{R}$ we can think of in homology as wrapping a plastic bag around a sphere, the bag cannot be contracted to a point. A natural representative is the volume form on the sphere, though in this chapter, we don't usually assume we have a metric that lets us talk about lengths, areas or volumes, but we might call it the "pre-volume" form or something like that.

5. **Degree theory**. It tries to construct a proper map F (inverse images of compact subsets are compact) that deforms a tangent vector to a normal vector. If you assume S^2 ($n = 2$), this attempt runs into a contradiction, since the degree of the antipodal exchange map $\alpha(\vec{v}) = -\vec{v}$ is supposed to be both $+1$ (you can deform α to the identity map by F) and -1 (the reflection matrix $-I_3$ in the ambient $\mathbb{R}^3$ has determinant -1, so it flips orientation), but it can't be both! For S^3 (or S^n for any odd n), the determinant of a reflection matrix $-I_4$ is +1 (orientation-preserving), so there is no contradiction, and it is possible to define a nowhere zero tangent vector field.

6. **Poincaré duality**. For compact manifolds, $H^{n-k}(M) \simeq H^k(M)$, i.e. there is an isomorphism, but it's not canonical.

7. **Poincaré-Hopf theorem**. The relationship becomes somewhat clearer in Exercise 8.9.14, part 4, where he states that $\chi(M) = \iota(\Delta_M, \Delta_M)$, the *algebraic self-intersection number* of a manifold. (Not to be confused with M actually having self-intersections! This is a more abstract notion: take a copy of each point x and write the "diagonal" $\Delta = \{(x, x), x \in P\}$, which is used for example to find fixed points of a map f geometrically: intersect the graph of f with the diagonal Δ.) And then, the statement is simply a subset of Poincaré-Hopf: if e.g. $\chi > 0$, as for S^2, there must be a source (index 1) or sink (index 1) or saddle (index 2) somewhere, and there the vector field has a stationary point, i.e. vanishes. So in particular, the algebraic self-intersection number of the 2-sphere in this context equal to $\chi = 2$, but S^2 obviously doesn't self-intersect in the naive sense.

(In algebraic geometry, another way to think of a self-intersection is discussed here, for which the answer is -1 for the holomorphic sphere. Often in intersection theory one wants to think of slightly shifting one of the two copies, but this goes beyond the scope of this course. If you do try to skim through that link, you might want to know that Conlon does define the cup product in Appendix D, but he suppresses the actual cup symbol! A related topic is *linking number* – see links above.)

8. **BRST cohomology**. It's an idea of "selecting physical objects in a geometric way" that applies to any gauge theory (as in particle physics) or string theory. The idea that physical states are invariant under some symmetry whose generator we can suggestively call "d", in which case applying this generator means the physical object, thought of as a 1-form, is closed: $d\omega = 0$. (In other words, since it doesn't "shift" under applying this symmetry transformation, it is invariant.) Physical objects are also gauge invariant, $\omega \sim \omega + d\lambda$, i.e. we want identify forms ω that differ by an exact form. These two properties (invariance under symmetry, identification under gauge transformations) mean that physical states form a cohomology.

   The nontrivial claim is that Yang-Mills theory (defined classically by Lagrangian function and the calculus of variations, as should be discussed together with the topics in Ch.11) in fact has such an anticommuting (graded) symmetry, sometimes initially called "s" instead of "d", that when applied twice gives zero, $s^2 = 0$. Such an s is the BRST symmetry. Showing $s^2 = 0$ uses the fact that Yang-Mills theory is characterized by Lie algebra (Nakahara Ch. 13.4), and lets us identify $s = d$ and write $d^2 = 0$ that we are becoming used to for differential forms.

9 Chapter 9: Forms and Foliations

1. **Ideal**. Specifically, a differential ideal.

2. **More equivalences in Frobenius theorem**. $\omega([X,Y]) = 0$ means that $[X,Y] = \zeta$ for some $\zeta \in \Gamma(E)$. This is the definition of involutive (which then implies integrable, by the earlier version of the theorem). One could worry that we only know that $\omega(\zeta) = 0$, but that we can't exclude that also $\omega(\eta) = 0$ for some $\eta \notin \Gamma(E)$? In other words, can we really "un-take" the application of ω? But it doesn't matter if there is such an η, since X, Y and ω are all arbitrary as objects of each kind, so ζ is the most general ζ we need.

3. **A simple example**. First $d\theta = d(xdy + ydz + zdx) = dx \wedge dy + dy \wedge dz + dz \wedge dx$, since $d^2 = 0$. Therefore, $\theta \wedge d\theta = (xdy + ydz + zdx \wedge (dx \wedge dy + dy \wedge dz + dz \wedge dx) = xdy \wedge (dz \wedge dx) + ydz \wedge (dx \wedge dy) + zdx \wedge (dy \wedge dz) = (x + y + z)dx \wedge dy \wedge dz$. (There are no signs, either from permutation, or observing they're all three already in cyclic order $x - y - z$.) Since we removed the origin, this three-form never vanishes.

   Let's set $z = f(t)$; we believe $f(t) = e^{-t/c}$ is correct, but let's see. $\theta = xdy + ydz + zdx = td(c) + cdf(t) + f(t)dt = 0 + cf'(t)dt + f(t)dt$. So we see that if $f'(t) = -f(t)/c$, then $\theta = 0$ along s, so we see that $f(t) = de^{-t/c}$ for any constant d.

   Finally, what's the general theory? If θ is a nowhere vanishing 1-form, and $d\theta = \eta \wedge \theta$ for some η (9.2.13, part 2), we would have a 1-plane normal bundle Q on which a normal 1-form $\omega = \theta$ could be defined, with a 2-plane integrable distribution E of vectors v for which $\theta(v) = 0$ (we would say θ is integrable, and E is the nullspace of ω). Now if $d\theta = \eta \wedge \theta$ were true, then $\theta \wedge d\theta = \theta \wedge (\eta \wedge \theta) = -\theta \wedge \theta \wedge \eta = 0$ for any 1-form η, but we checked that in fact $\theta \wedge d\theta \neq 0$, so this is a contradiction.

   For an integral curve (yielding a 1-plane integrable distribution) in 3-space, the normal bundle is a 2-plane bundle, transverse to the curve. The general discussion of Ch. 9.2 doesn't really apply for the "degenerate" case of integral curves as opposed to higher-dimensional integral manifolds: we could try to put $\omega = d\theta$ as the normal bundle nonvanishing 2-form, and try to define E as $\omega(v \wedge v_1) = 0$ for all v_1 in tangent space.[6] But since E is 1-dimensional, $\mathbf{v}_1 \propto \mathbf{v}$, so $v \wedge v_1 = 0$ automatically, and the condition gives no information. Instead of working with the normal bundle we should simply do the "opposite" (and more obvious) thing: demand that $\theta = 0$ directly, which in fact works, as above.

   A reasonable analogy of what is going on here is the "Theorem of Turning Tangents" (do Carmo) vs. the Gauss map. A curve in the plane can be nicely characterized by its tangent vector (field), and in the Theorem of Turning Tangents, the vector that is transported to the origin is the *tangent* vector. For the Gauss map, it is a little easier to picture the *normal* vector being transported to the origin and sweeping out a curve on a sphere, rather than a whole tangent plane being transported to the origin and flopping around there. As a rule of thumb, if you are given the choice of working with a tangent bundle or a normal bundle, and one of them has fewer dimensions, it might be the easier choice, but the above integral curve example is a counterexample to this rule of thumb, since our method became degenerate below 2 dimensions. If everything is nondegenerate, ultimately you should obtain the same results from either one.

4. **A more interesting example**. If $\omega^1 = df \wedge dt - gdx \wedge dt = 0$ and $\omega^2 = dg \wedge dt + df \wedge dx = 0$, from the first we see that since $df = \partial_x f dx + \partial_t f dt$ we in fact have $\partial_x f dx \wedge dt - gdx \wedge dt = 0$ so $g = \partial_x f$ (actually this was given, but it's a good check). We have $d\omega^1 = 0$ automatically, but for $d\omega^2 = \partial_x g dx \wedge dt + \partial_t f dt \wedge dx = (\partial_x g - \partial_t f) dx \wedge dt$ to vanish we should have $\partial_x g = \partial_t f$, or $\partial_x^2 f = \partial_t f$, which is the heat equation. (It is often given with a relative dimensionful constant between the two terms, but this can be absorbed in a rescaling of t, for example).

[6] The 2-form $d\theta$ acts on $v = (v_x, v_y, v_z)$ as $(dx \wedge dy + dy \wedge dz + dz \wedge dx)(v \wedge v_1) = v_x v_{1y} + v_y v_{1z} + v_z v_{1x}$, and in principle we could try to set this to zero.

5. **A closed, nonsingular 1-form?**. There is one if the manifold can be viewed as a fibering over a circle (Corollary 9.3.16). Note that "nonsingular" means "nowhere zero" as he states at the beginning of section 9.3. This is actually fairly intuitive; if there is a circle embedded throughout the manifold, we might expect to have our friend, the cohomologically nontrivial 1-form $\eta = d\theta$, that we know from $H^1(S^1) = \mathbb{R}$. But as usual, intuition should not be applied too lightly; for example, $\mathbb{R}^2$ in polar coordinates looks nicely foliated by circles, everywhere except at the origin. So if you didn't check the assumptions carefully enough, you might incorrectly conclude that $\mathbb{R}^2$ is fibered over S^1 and therefore by Corollary 9.3.16 has a closed, nonsingular 1-form, which actually $\mathbb{R}^2$ does not.

   By the way, Conlon's definition of fibration as locally trivial is not the most general one. But it captures many examples, like the Hopf fibration of S^3, with base S^2 and fibers being circles. When you see the Hopf fibration for the first time, it seems surprisingly complicated for connecting such simple things as spheres; can't we just make a trivial bundle, like $S^3 \simeq S^2 \times S^1$? Answer: No! As we saw in Chapter 5, there kind of has to be something complicated about viewing S^3 as circles over S^2: as it turned out, S^3 has both a natural group structure and a natural manifold structure (it is a *Lie group* manifold), whereas S^2 is certainly a nice smooth manifold, but not a Lie group. And when Conlon introduced bundles in Chapter 3.4, we learned that there was a $\mathrm{Gl}(n)$ group structure to the transition functions (geometric cocycles; for general fiber bundles, this group structure is called the structure group). So from this perspective, in Hopf's view of S^3 as an S^1 bundle over an S^2 base, we are trying to express something simple (S^3, Lie group) as a fibration over something more complicated from the group-theory perspective (S^2, not a Lie group) then the fibering could be expected to be a little complicated.

10 Chapter 10: Riemannian Geometry

1. **Gauss map**. The normal vector is not globally defined.

2. **Gauss curvature**. The one with Christoffel symbols is pretty enlightening, and Christoffel introduced them after Gauss's death. Of course K is for *Krümmung*, "curvature" in German. I guess R is for Riemann. Wald in his GR book does use C for "connection". Brioschi's formula could also qualify as direct proof of the "Totally Awesome Theorem" (as Lee translates *Theorema Egregium*).

3. **Second fundamental form**. As Conlon says above (10.3), the open set $U \subset \mathbb{R}^3$, i.e. flat. So (10.3) is a special case.

4. **Sectional curvature**. It's a generalization in the sense that it reduces to Riemann curvature for orthogonal vectors u and v, but it's a special case in the sense that it is some (possibly complicated) subset of the ambient Riemann tensor. See for example Toponogov's theorem.

5. **Jacobian cocycles and integrably parallellizable**. If you have $Jg_{\alpha\beta} =$ identity for all transition functions, then the manifold is flat. I tend to say that it's then Euclidean, but if it has some compact component, it that component is diffeomorphic to a (flat) torus, and any remaining noncompact component is indeed diffeomorphic to Euclidean space. As we know, the covering space of a flat torus is Euclidean space, so that distinction is only global, not local as for curvature.

6. **Flat = integrable**. $\mathcal{M}(n)$ is more abstract (Ch.7): the *category of R-modules and R-linear maps, where R is a fixed commutative ring with unity 1* (p.209). On the other hand, $\mathfrak{M}(n)$ are simply general real $n \times n$ matrices (i.e. they can have determinant zero, so more general than $\mathrm{Gl}(n)$, with nonzero determinant).[7] Our main example will be $B \in \mathrm{O}(n)$, orthogonal $n \times n$ matrices.

   Conlon says after the proof what it's about: *curvature is the obstruction to integrability of a certain n-plane distribution*. For Riemannian manifolds, this is equivalent to the $\mathrm{O}(n)$ structure given by the metric, as he shows in examples afterwards.

   What about other structures? For example other Lie groups acting on the vector space of a vector bundle with some $\gamma_{ij} \in \mathrm{Gl}(n)$, as in section 3.4, other than $\mathrm{O}(n)$? One such example is the Lie group $\mathrm{Sp}(n)$, in which case you don't get a Riemannian but a symplectic manifold (i.e. the tangent space is acted upon by symplectic matrices $\gamma_{ij} \in \mathrm{Sp}(n)$), that Conlon doesn't talk about. Another can be complex matrices $\mathrm{Gl}(n, \mathbb{C})$, that he also doesn't emphasize, but does discuss in Example 3.4.23. More generally, what n-plane distribution is it that is integrable, if it's not $\mathrm{O}(n)$ structure on tangent space as for Riemannian manifolds? In Ch. 11, it will be the distribution H of "horizontal vectors" on a principal bundle, that makes sense for any Lie group – we'll define these words in Ch. 11. If you want to peek ahead, it's property (1) in Theorem 11.4.19 on p. 365.

7. **Einstein's equation**. "Only nontrivial contraction" means that up to using symmetries of the pairs of indices on the Riemann tensor expressed in local coordinates (as Conlon does discuss, e.g. Exercise 10.6.5: those symmetries follow from the original expression of $R(X, Y)Z$), there is essentially only one way to contract index pairs on $R^i{}_{jkl}$ without getting zero. The Ricci scalar curvature is given as $R = g^{ij}R_{ij}$ as on that list, and it is *not* the Riemann tensor $R(X, Y)Z$ that takes two vectors X and Y, acts on a vector Z and produces another vector W, so it should have four indices as a tensor in local coordinates. But the scalar curvature R is an object without tensor indices, i.e. it's simply a scalar function over the manifold M, that doesn't "act" on anything. This means that it is "more" independent of coordinate choices than the Riemann tensor $R^i{}_{jkl}$, that transforms contravariantly in the i index (i.e. like components of a vector

[7] Actually, I'm not sure this sudden generality is intentional, he talks about matrices in the proof, so he might well just have mistyped the font command and meant $\mathfrak{M}(n)$ instead of $\mathcal{M}(n)$ everywhere in Lemma 10.6.4! But it gave me the opportunity to distinguish them.

$X = f^i\partial_i$, it's "index up") and covariantly in the jkl indices (i.e. like components of a co-vector $\theta = f_j dx^j$, "index down"). The Ricci scalar curvature doesn't transform at all, it's a *curvature invariant*. Even this statement can be slightly misleading: if you express it in local coordinates, you still need to say how it is a function of those coordinates (unless R=constant, a very special case, but one of some interest). The scalar curvature is in some cases related to the Gauss curvature (that Conlon calls κ, but is sometimes called K).

11 Chapter 11: Principal Bundles*

1. **Yang-Mills theory**. It's an unsolved problem to rigorously construct *quantum* Yang-Mills theory. I don't know anyone who has a problem with the level of rigor of *classical* Yang-Mills theory, in fact the theory of principal bundles is due to Cartan, who was certainly a mathematician. Perhaps his brand of mathematics is now considered somewhat oldfashioned, but the young generation after him (including his son) put it on a solid modern foundation.

   (Incidentally, here I used "classical" to mean "non-quantum", this can be somewhat confusing to students. In undergraduate teaching, "modern physics" is usually the complement of the set of "classical physics", and "modern" includes special relativity (which can be pinpointed in time very precisely: it's from 1905[8]), therefore "classical" does not. In graduate-level physics, "classical" definitely includes not only special but also general relativity (from 1915 – there was a lot of centennial celebration in 2015), as evidenced by Thorne & Blandford's new book "Modern Classical Physics", where they define "classical" (in the foreword of the book) as "non-quantum". And since Thorne is a general-relativist, this is not swagger by the "other team" (people working in quantum physics, like particle physics). In physics, typically researchers don't want their own work to be called "classical" (i.e. oldfashioned), so it shouldn't be lost on the reader that Thorne is taking a stance there, in fact the foreword sounds strangely apologetic about calling the book "classical", which shows that this is a sensitive issue.[9])

2. **Holonomy**. Berger's list is a classification of what holonomy groups are possible for (simply connected) Riemannian manifolds, under some conditions. In particular, $SU(3)$ is a Calabi-Yau threefold (6 real dimensions). The quote I referred to is *curvature gives the infinitesimal holonomy over a closed loop (the infinitesimal parallelogram). More formally, the curvature is the differential of the holonomy action at the identity of the holonomy group.* Curvature as the obstruction to the holonomy group being identity is worked out in detail in Nakahara section 7.3.2.

3. **Homogeneous spaces**. For Gl, the maximal compact subgroup is O, for Gl_+, the maximal compact subgroup is SO. So if you have a global orientation, you can have an SO bundle reduction. A principal bundle has a "whole manifold" as fiber $p^{-1}(x)$, not just a vector space, so it is natural to view a homogeneous space G/H as a principal bundle with fiber the Lie group manifold G. This doesn't fit naturally in the idea of a vector bundle, when the fiber is supposed to be a (linear) vector space. A connection between the two ideas comes when you think of the tangent space $T(P)$ of the principal bundle (to be distinguished from the tangent bundle $T(M)$ of the manifold M!); this is again a vector space, and can be split up into "vertical" and "horizontal" as $T(P) = V \oplus H$ (p.359). See my little summary cartoon at the beginning.

4. **A "trivial" example**. I put the solution at the end of Presentation 5 (summary), but think about it first: the ODE you get from imposing $\omega(X) = 0$ is $\dot{x}y/r^2 - \dot{y}x/r^2 + \dot{f} = 0$, which simplifies further for a unit circle $(\cos 2\pi t, \sin 2\pi t)$. Nakahara gives some more detail: $\phi : ((x, y), f) \mapsto u$ is a local trivialization. We're implicitly using that $\omega(A^\sharp) = A$, where $A^\sharp = A\partial/\partial f$ is a vertical vector, and since the group G is abelian, $R_{g*}\omega = g^{-1}\omega g = \omega$. All this "bookkeeping" is hard to keep track of in a course at this level, but the formulation in the study question (solving $\omega(X) = 0$) should be reasonable to figure out with the given information, from what we did

[8] Einstein's "miracle year" as historians of science like to call it, due to the insane number of things he revolutionized in 1905. The paper he got the Nobel prize for was his first paper that year, but it was not relativity (what was his Nobel prize for?). Paper 2 that year was also not relativity, but it "made atoms real", i.e. allowed a direct experimental determination of Avogadro's number, which is conceptually at least as important as the other two. Einstein's 3rd paper in 1905 was relativity, and his 4th paper was $E = mc^2$, probably his most famous equation.

[9] Anything *quantum* is considered somewhat tangential for this course, but if you're interested, think for a moment about this: when did quantum mechanics begin then, i.e. where's the cutoff between classical and modern in the "graduate" sense of these words? From my rambling discussion of the meaning of classical, you should get the idea that the "quantum revolution" should be after 1915 (the year of general relativity), but when was it? Misleading hint: isn't Bohr's model of the atom from before 1915?

earlier. As should be clear once you find a solution, the space curve is not horizontal in any naive sense. Motion along the fiber (action by G) is literally vertical here, however, so that deserves the name.

5. **Magnetic monopoles and notation**. In Conlon $s^\sharp$ in a principal G-bundle is vertical lift, and $s^\flat$ is horizontal lift. You can think of the vertical space V as representing "symmetries" (in Choquet-Bruhat et al, they can be Killing vectors), whereas horizontal H space represents actual "motion" along the bundle (e.g. parallel transport, think "h" for holonomy). For a principal G-bundle, they are related by group multiplication $A \in G$ on the right: $s^\sharp(t) = s^\flat(t)A(t)$. Corollary 11.4.6 is similar to how I would approach this, namely write the tangent vector of a section (the vertical lift) in terms of a tangent vector of the horizontal lift and an extra term, that comes from the product rule applied to $s^\sharp(t) = s^\flat(t)A(t)$, evaluated at $t = 0$.

   Lemma 11.4.5 says that $\dot{A}(0) = \tilde{\omega}(\dot{s}(0))$, where $\tilde{\omega} = \sigma^*(\omega)$ represents the "provisional local forms" at the beginning of section 11.4 (these are the forms from Ch 10.6, where they had no tildes). The key input is at the bottom of p.361: the (local) connection form $\tilde{\omega}$ tells you how to decompose $\nabla_{\dot{s}}$ in terms of frame fields X_i. So if you know the "initial direction" $\tilde{\omega}(\dot{s}(0))$ that a lift should go in (this is abstract, since $\tilde{\omega}$ is a matrix; when applied to a basis vector X_k as at the bottom of p.361, it becomes a concrete vector, generically different from X_k if $\Gamma^k{}_{ij} \neq 0$), then you know what (equally abstract) "initial direction" $\dot{A}(0)$ the matrix $A \in G$ should go in.

   The gauge transformation $A \mapsto A + d\lambda$ comes from changing charts *on the bundle*, which changes the section σ. I put "on the bundle" in italics, since although we have become used to the idea that a change of charts on the base manifold M can "induce" a change of charts on a bundle (for example, a diffeomorphism $\varphi : M \to M$ can be pushed forward to the tangent bundle: φ_*), that is true even in simple cases like a trivial (i.e. product) vector bundle, whereas here for principal G-bundles, the change of charts is really nontrivial only if the bundle itself is nontrivial. This is related to the fact, as discussed on the Wikipedia page, that unlike vector bundles, principal bundles only admit global sections if they are actually trivial as bundles, which we usually don't want. (This is stated at the top of p.352 in Conlon, but he doesn't emphasize "global", so it's not so clear at that point.)

   So gauge transformations are most interesting in the case of nontrivial bundles, as the magnetic monopole; this is a principal $G = U(1)$-bundle over S^2 with the north-pole and south-pole charts on S^2 giving two inequivalent sections σ_1 and σ_2, and if you work with e.g. the south-pole chart, you can allow a Dirac string from the origin of the ambient space through the north pole, that "carries away excess curvature" (of the connection A, not of spacetime) as it is expressed on that Wikipedia link, so that the "curvature of the connection" $dF = 0$ (Bianchi identity) everywhere except at Dirac's magnetic monopole.[10] That, to be clear, hasn't been discovered yet in the real world. Which doesn't prevent us discussing its mathematics.

   Compare this to the more trivial example above, where we imposed that $\omega(X) = 0$, i.e. demanded that X is horizontal, and it was easy to ensure that it remains horizontal at all times since in Euclidean space, $\Gamma^k{}_{ij} \equiv 0$. In general, there will be some matrix action that changes the meaning of horizontal as you traverse various lifts of the curve $s(t)$.

   Since $G = U(1)$ for the magnetic monopole (of electromagnetism, an abelian gauge theory), this example is still almost trivial. A nontrivial example would be monopole solutions for nonabelian G. In fact there are such solutions of Yang-Mills theory, and they are interesting. [11]

[10] Polchinski used this kind of argument in 1995 in the discovery of "D-branes", important objects in string theory. His argument can be thought of as a higher-dimensional generalization of Dirac's ideas about the Dirac string, but applied to detailed information that was known from the quantum mechanics of relativistic 1-dimensional objects ("strings"). This is discussed in Polchinski's string theory book, e.g. fig. 13.3 shows a Dirac string in higher dimensions.

[11] If you think all this sounds too much like physics, some mathematicians would disagree:
M.Atiyah, N. Hitchin, "The Geometry and Dynamics of Magnetic Monopoles" (1988).

12 Appendix B: Inverse function theorem

1. **Banach space**. For example, L^2 functions on a proper subset $U \subset \mathbb{R}^n$ are not complete. This example is given in "Analysis, Manifolds and Physics". Another is to violate the Parallelogram law by for example C^0 functions with the supremum norm.

2. **Bad joke**. A Bananach space.