Chapter 8
Some Geometrical Apparatus

Up to now our treatment of differential geometry has been rather informal. The reader is assumed to have the modicum of background knowledge of manifolds and tensors needed to follow the discussion at the modest level of rigor appropriate to the occasion. In the study of renormalization in the next chapter, however, covariant derivatives and curvature tensors will enter in an essential and technical way. It seems prudent, therefore, to pause for a comparatively thorough and precise discussion of those concepts, which are central to contemporary physics — not only in general relativity, but also in the gauge theories of elementary particles.

We are concerned here only with purely "local" matters. Global topology, which leads in physics to monopoles, the Aharonov–Bohm effect, instantons, and so on, has been the subject of a vast amount of recent expository writing and is beyond the scope of this book.

Covariant derivatives

Let's start with a conceptual introduction to, or review of, the two most familiar and elementary instances of covariant differentiation in physics. They correspond to the fundamental coupling of matter to the electromagnetic and gravitational fields, respectively.

In nonrelativistic quantum mechanics the basic dynamical object is the wave function, $\psi(\underline{x}) = \psi(t, \mathbf{x})$. Insofar as the operational significance of ψ lies in the probability density, $|\psi(\underline{x})|^2$, the *phase* of ψ is irrelevant. At least, the overall phase is. At first sight it might appear that even the *relative* phase of the wave function at different points is unobservable:

$$\tilde{\psi}(\underline{x}) \equiv e^{i\theta(\underline{x})}\, \psi(\underline{x}) \equiv [U\psi](\underline{x})$$

contains the same information as ψ. This is not quite correct, however, as soon as one considers the momentum observable (or anything else which does not commute with position). A component of the momentum operator is

$$p^j = -i\partial_j \equiv -i\frac{\partial}{\partial x^j} \qquad (j = 1, \ldots, d),$$

so its expectation value in the state ψ is

$$\langle p^j \rangle = -i \int \psi^* \partial_j \psi \, d^d x.$$

But

$$-i\partial_j \psi = e^{-i\theta}[-i\partial_j \tilde{\psi} - (\partial_j \theta)\tilde{\psi}],$$

and because of the extra term,

$$\int \psi^* \partial_j \psi \, d^d x \neq \int \tilde{\psi}^* \partial_j \tilde{\psi} \, d^d x.$$

Thus ψ and $\tilde{\psi}$ are not physically equivalent. A Schrödinger equation of the elementary form is not invariant under U:

$$i\partial_0 \psi = -\frac{1}{2m} \sum_{j=1}^{d} \partial_j{}^2 \psi + V(\mathbf{x})\psi$$

transforms to

$$i\partial_0 \tilde{\psi} = -\frac{1}{2m} \sum_{j=1}^{d} [\partial_j - i(\partial_j \theta)]^2 \, \tilde{\psi} + [V + (\partial_0 \theta)] \, \tilde{\psi}.$$

Suppose, however, that each differentiation operator in the Schrödinger equation was already accompanied by a function:

$$i\partial_0 \psi = -\frac{1}{2m} \sum_{j=1}^{d} [\partial_j + iA_j(\mathbf{x})]^2 \psi + V(\mathbf{x})\psi. \qquad (8.1)$$

Then the equation satisfied by $\tilde{\psi}$ will be of the same form (8.1), but with different functions $\mathbf{A}$ and V, because

$$U[\partial_\mu + iA_\mu]U^{-1} = \partial_\mu + i(A_\mu - \partial_\mu \theta) \equiv \partial_\mu + i\tilde{A}_\mu \, .$$

(If we allow θ to depend on t, then $\mathbf{A}$ and V must also.) Of course, $\mathbf{A}$ really does exist; it is the electromagnetic vector potential. In this way, a theoretical physicist who had never heard of magnetism might be led to predict its existence, on the basis of the purely aesthetic requirement that quantum mechanics be invariant under the local phase transformations, $U = U\{\theta(\underline{x})\}$.

This observation makes possible the following new point of view. Refusing to commit ourselves to any local phase convention, we regard the wave function in an abstract way: ψ maps each point $\underline{x}$ onto some point in a certain space — call it $F_{\underline{x}}$ — about which we can say only that it has the same linear and metric structure as the complex plane. We specifically deny that this space can be identified with $\mathbf{C}$ in a fixed way; any such identification would be tantamount to an arbitrary phase convention. On these abstract wave functions there is defined an operation of *covariant differentiation*,

$$\psi \mapsto \nabla_\mu \psi.$$

Given any particular local phase convention, this operation is expressible by a concrete formula among complex-valued functions:

$$\nabla_\mu \psi(\underline{x}) = [\partial_\mu + iA_\mu(\underline{x})]\psi(\underline{x}). \tag{8.2}$$

If the phase convention is changed, $\tilde{\psi} \equiv e^{i\theta(\underline{x})}\psi$, then the *connection coefficients*, A_μ, change according to

$$\tilde{A}_\mu = A_\mu - \partial_\mu \theta. \tag{8.3}$$

A particular local phase convention is called "a gauge" (or "a choice of gauge"). The expectation value

$$\langle p^j - A^j \rangle = \int \psi^* (-i\nabla_j \psi)\, d^d x$$

is gauge-invariant. (In the Lagrangian formalism of particle mechanics, $\mathbf{p} - \mathbf{A}$ equals $m\mathbf{v}$, which is indeed the quantity with direct physical significance.) Similarly, the gauge-covariant Schrödinger equation is

$$i\nabla_0 \psi = -\frac{1}{2m} \sum_{j=1}^{d} \nabla_j{}^2 \psi, \tag{8.4}$$

which is (8.1). The Klein–Gordon and Dirac equations can be treated in the same way. ("*Co*variant" means that the *form* remains unchanged, while the particular coefficients change according to (8.3).)

Notational remarks:

(1) Our sign conventions are those of Messiah 1961, Chap. 20. In particular, $A_0 = +V$ and the components of $\mathbf{A}$ are $A^j = -A_j$. The quantum correspondence rule is $p_\mu \mapsto +i\partial_\mu$ for $\mu = 0, 1, \ldots, d$. Because of our metric-signature convention, the components of both spatial vectors, $\mathbf{p}$ and $\mathbf{A}$, change sign when indices are raised and lowered. People who use the metric signature $(-+++)$ also usually reverse the relative sign of ∂_μ and iA_μ in (8.2).

(2) Units are chosen so that c and $\hbar$ equal 1. Moreover, the unit electrical charge, e, is absorbed into A_μ. (It will come back eventually in the denominator of the kinetic Lagrangian of the electromagnetic field, $e^{-2}F_{\mu\nu}F^{\mu\nu}$.)

The second example is the differentiation of vector fields on a manifold. Any definition of such a derivative which can be covariantly expressed in all coordinate systems and satisfies some basic formal properties (to be reviewed briefly below) leads in each particular coordinate system to a formula of the type

$$\nabla_\mu v^\nu \equiv v^\nu{}_{;\mu} = \partial_\mu v^\nu + \Gamma^\nu_{\rho\mu} v^\rho. \tag{8.5}$$

where the connection coefficients or *Christoffel symbols* $\Gamma^\nu_{\rho\mu}$ are functions of $\underline{x}$. When the manifold is Riemannian or pseudo-Riemannian, there is a preferred connection whose coefficients are constructed from first-order derivatives of the metric tensor, but our preliminary remarks here are valid more generally.

To see the near inevitability of (8.5) and to give the Christoffel symbols a more intrinsic meaning, suppose that a basis has been chosen for the space of (contravariant) vectors at each point in the manifold: $\{\mathbf{e}_\alpha(\underline{x})\}_{\alpha=1}^n$. (The α labels different vectors in the basis, not the components of a single vector $\mathbf{e}$.) Suppose also that $\mathbf{e}_\alpha$ varies smoothly with $\underline{x}$, so that it makes sense to differentiate it. Each vector field $\mathbf{v} \equiv \{v^\mu(\underline{x})\}$ can be written as a linear combination, $\mathbf{v} = v^\alpha \mathbf{e}_\alpha$. We assume that covariant differentiation of the product of a scalar function and a vector field satisfies the Leibniz rule. Then $\nabla\mathbf{v}$, the tensor whose components are standardly written $\nabla_\mu v^\nu$ or $v^\nu{}_{;\mu}$, must be calculable by

$$\nabla_\mu(v^\alpha \mathbf{e}_\alpha) = \big(\nabla_\mu(v^\alpha)\big)\mathbf{e}_\alpha + v^\alpha \nabla_\mu \mathbf{e}_\alpha ,$$

where $\nabla_\mu(v^\alpha)$ is the covariant derivative of the *scalar function* v^α, assumed to reduce to the ordinary derivative, $\partial_\mu v^\alpha \equiv v^\alpha{}_{,\mu}$. *Define* the

Γ's by

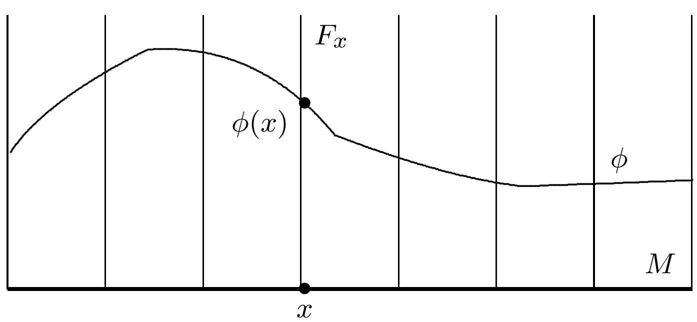
$$\nabla_\mu \mathbf{e}_\alpha \equiv \Gamma^\beta_{\alpha\mu} \mathbf{e}_\beta \,. \tag{8.6}$$

Then our calculation shows that

$$\partial_\mu \mathbf{v} = (v^\alpha{}_{,\mu} + \Gamma^\alpha_{\beta\mu} v^\beta) \mathbf{e}_\alpha \,,$$

which is a restatement of (8.5).

Now let's pass from these examples to a general framework. Let M be a manifold of dimension n. Usually for us it will be equipped with a metric tensor, $g_{\mu\nu}(\underline{x})$. We want to consider derivatives of objects $\phi(\underline{x})$, which are examples of some type of (classical) "field" on M; technically, they are sections of a vector bundle over M. Each ϕ is a kind of function, whose value at a point $\underline{x}$ belongs to some vector space $F_{\underline{x}}$ (the *fiber* at $\underline{x}$). Physically, the elements of this space may be numbers or vectors, as in the two examples, or tensors, spinors, isovectors (of a non-Abelian gauge theory), etc. $F_{\underline{x}}$ is *isomorphic* to any other fiber, $F_{\underline{y}}$, but not *canonically identified* with it. Let the common dimension of the fiber spaces be r. They may be real or complex vector spaces, depending on application. Let F be a "fiducial" vector space isomorphic to the fibers. (Thus $F = \mathbf{C}$ in our first example, $\mathbf{R}^N$ (N = dimension of manifold) in the second.)



Ordinarily I use an index-free vector and matrix notation in F, but a classical index notation for tensors over M. For the latter, we will always be using a *coordinate basis*, rather than a more general "moving frame" (see, e.g., Schutz 1985, Chap. 5).

To do calculations with ϕ it helps to represent $\phi(\underline{x})$ by numbers. Therefore, we introduce a basis in each $F_{\underline{x}}$, so that

$$\phi(\underline{x}) = \phi^j \mathbf{e}_j \equiv \sum_{j=1}^{r} \phi^j(\underline{x}) \mathbf{e}_j(\underline{x}).$$

Relative to the basis $\{\mathbf{e}_j\}$, each field ϕ is identified with a sequence $\{\phi^j(\underline{x})\}_{j=1}^r$ of ordinary real- or complex-valued functions. A change of basis, or *gauge transformation*, is specified by an equation of the form

$$\tilde{\mathbf{e}}_j(\underline{x}) \equiv \sum_{k=1}^{r} \mathbf{e}_k(\underline{x})[U(\underline{x})^{-1}]^k{}_j \tag{8.7a}$$

or, equivalently,

$$\tilde{\phi}^j(\underline{x}) = \sum_{k=1}^{r} U(\underline{x})^j{}_k \phi^k(\underline{x}) \equiv (U\phi)^j. \tag{8.7b}$$

We assume that all "admissible" bases belong to an equivalence class such that the matrices $U^j{}_k$ relating them are smooth functions.

Now to the main business at hand, defining the derivative of ϕ. The literal partial derivatives $\partial\phi_j/\partial x^\mu$ do not fit together as the components of an intrinsically meaningful object, because they do not include information on the $\underline{x}$-dependence of the basis vectors. (In general there are no "constant" $\mathbf{e}$'s against which other things may be compared.) Consequently, $\partial_\mu\phi^j$ has a complicated, inhomogeneous transformation law involving derivatives of U. As the derivative of a scalar function is a covariant vector field, one would prefer the derivative of a section ϕ to be a covector-valued section (or $F_{\underline{x}}$-valued covector), $\nabla_\mu\phi$, which continues to behave like (8.7b) under gauge transformations (and also like an ordinary covector field under coordinate transformations in M).

Therefore, we define *a covariant differentiation* to be any mapping of ordinary sections into covector-valued sections which satisfies

$$\nabla_\mu(\phi_1 + \phi_2) = \nabla_\mu\phi_1 + \nabla_\mu\phi_2$$

and

$$\nabla_\mu(f\phi) = (\partial_\mu f)\phi + f\nabla_\mu\phi$$

(where f is any ordinary, scalar-valued function). Applying these axioms to the expansion $\phi = \phi^j\mathbf{e}_j$, we get

$$\begin{aligned}\nabla_\mu\phi &= (\partial_\mu\phi^j)\mathbf{e}_j + \phi^j\nabla_\mu\mathbf{e}_j\\ &\equiv (\partial_\mu\phi^j)\mathbf{e}_j + \phi^j(\nabla_\mu\mathbf{e}_j)^k\mathbf{e}_k\\ &= [\partial_\mu\phi^j + \phi^k(\nabla_\mu\mathbf{e}_k)^j]\mathbf{e}_j,\end{aligned}$$

where we had to do some index relabeling in the last step. Therefore, if $w_\mu(\underline{x})$ (which is, for each $\underline{x}$ *and each* μ, a matrix) is defined by

$$\nabla_\mu(\mathbf{e}_k) \equiv [w_\mu]^j{}_k \mathbf{e}_j, \tag{8.8}$$

there follows

$$(\nabla_\mu\phi)^j(\underline{x}) = \left(\partial_\mu\phi^j\right)(\underline{x}) + [w_\mu(\underline{x})]^j{}_k\phi^k(\underline{x}),$$

usually abbreviated to

$$\nabla_\mu\phi = \partial_\mu\phi + w_\mu\phi. \tag{8.9}$$

Conversely, any derivative defined by such a formula satisfies the linearity and Leibnitz conditions from which we started.

Note that the covariant derivative (or the associated *connection form*, w_μ), is *extra structure* — it is not uniquely determined by the manifold and vector bundle.

Exercise 29: What happens to w under a gauge transformation,

$$\tilde{\mathbf{e}} = \mathbf{e}U^{-1}, \qquad \tilde{\phi} = U\phi\,?$$

(a) Show that

$$\tilde{w}_\mu = U[w_\mu - U^{-1}\partial_\mu U]U^{-1} = Uw_\mu U^{-1} - (\partial_\mu U)U^{-1}$$

and hence

$$\tilde{\nabla}_\mu = U\nabla_\mu U^{-1}.$$

(I.e., $\tilde{\nabla}$ and ∇ represent the same geometrical operation with respect to different bases.)

(b) Since we use coordinate bases, a coordinate transformation in M determines a gauge transformation, in the present extended sense, on the vector fields on M. Show that the application of (a) to contravariant vector fields yields the transformation law of Christoffel symbols found in classical texts on differential geometry,

$$\tilde{\Gamma}^\rho_{\mu\nu} = \frac{\partial\tilde{x}^\rho}{\partial x^\gamma}\,\frac{\partial x^\alpha}{\partial\tilde{x}^\mu}\,\frac{\partial x^\beta}{\partial\tilde{x}^\nu}\,\Gamma^\gamma_{\alpha\beta} - \frac{\partial^2\tilde{x}^\rho}{\partial x^\alpha\partial x^\beta}\,\frac{\partial x^\alpha}{\partial\tilde{x}^\mu}\,\frac{\partial x^\beta}{\partial\tilde{x}^\nu}\,.$$

The transformation of the index μ under the coordinate change must also be taken into account here.

Gauge theories in physics involve two additional elements which are not emphasized here. First, the connection form itself becomes a dynamical field, satisfying its own field equations (Maxwell, Einstein, or Yang–Mills equations). This gives the theory a true gauge *in*variance, not just a covariance. Second, in the gauge theories of particle physics, not all smooth choices of local basis are admissible; or, as it is more often expressed, the group of allowed gauge transformations at a point is not the entire general linear group of dimension r. Rather, one considers an equivalence class of bases related among themselves by gauge transformations whose local values belong to a subgroup that (typically) preserves some quadratic form. The group, such as U(1), SU(3), SO(10), etc., is prescribed before the vector bundles supporting it as a gauge group are constructed, and, regardless of the bundle, the connection matrices w_μ always belong to some representation of the corresponding Lie algebra. We have already seen this situation in the electromagnetic case, where only multiplication by functions of modulus 1 (the group U(1)) was considered, and the connection cofficients consequently were purely imaginary. Since the choice of group determines the possible multiplets of elementary particles, particle physicists place great emphasis on the group in expounding gauge theories. For our purposes, however, the group is of secondary importance.

Curvature

The presence of a nonvanishing electromagnetic field means that there is *no* basis with respect to which $A_\mu = 0$. Similarly, the presence of a nontrivial gravitational field means that there is no coordinate system in which the Christoffel symbols are identically zero. (Either of these conditions could be enforced at a single point, but not throughout an open region.) Both of these fields are manifestations of *curvature* of vector bundles (more precisely, curvature of particular connections on bundles).

There are a few more facts about connections that need to be placed on record before we proceed.

First, if F^* is the dual space of F, then the covariant derivative of F^*-valued fields is defined by postulating that a Leibnitz rule holds for differentiation of the pairing of an F^*-valued field with an F-valued field (that is, a scalar function $\Psi(\phi) \equiv \Psi_j \phi^j$, where $\Psi(\underline{x})(\,\cdot\,) \in F_{\underline{x}}{}^*$ and $\phi(\underline{x}) \in F_{\underline{x}}$). This amounts to saying that the connection matrices for F^*-valued fields are the negatives of the transposes of those for F-valued fields. For example, the formula for differentiation of a covector field is

$$\nabla_\mu v_\nu \equiv v_{\nu;\mu} = \partial_\mu v_\nu - \Gamma^\rho_{\nu\mu} v_\rho\,;$$

in the comparison with (8.5), the transposition is on the indices (ν, ρ).

Exercise 30: Prove the assertion ("This amounts to saying . . . ").

Second, similar reasoning establishes that the connection form for the derivative of any kind of tensor should be the sum of the connection forms for the various factors of the tensor-product space in question. That is, the formula for the covariant derivative of a tensor contains, besides the literal derivative, one term for each index of the tensor, and the form of that term is the same as it would be if that were the only index the tensor possessed. This refers both to contravariant and covariant space-time indices (Greek superscripts and subscripts) and to primordial and dual bundle indices (referring to components of vectors in F and F^*). Example:

$$\nabla_\mu B^j_{\nu k} \equiv \partial_\mu B^j_{\nu k} - \Gamma^\rho_{\nu\mu} B^j_{\rho k} + [w_\mu]^j{}_l B^l_{\nu k} - [w_\mu]^l{}_k B^j_{\nu l}\,.$$

Charged scalar fields (for which $w_\mu = iA_\mu$) are included in this formalism, although the "index" involved has a range of only one value and would not normally be written as such.

Exercise 31: Let $M(\underline{x})$ be a "matrix-valued" field — i.e., its value at each point $\underline{x}$ is a linear operator mapping $F_{\underline{x}}$ into itself. Show that the covariant derivative of such an object is

$$M_{;\mu} = M_{,\mu} + [w_\mu, M].$$

Everything that has been said about differentiation on bundles applies, of course, to these tensor bundles. That is, the index j in the formalism may (and often does) stand for a whole string of indices, some of which are Greek and others are the bundle indices of one or more elementary field types such as spinors, charged scalars, isovectors, etc. (It should also be kept in mind that the elementary field type could simply be the ordinary vector field; inclusion of a connection in the formalism does not necessarily imply presence of a gauge field in the narrowest physical sense.) The derivative of a tensor is a tensor with one additional covariant space-time index.

Knowing how to differentiate a general tensor, we are now able to calculate higher-order derivatives of the original sections. The second derivative is

$$\begin{aligned}\phi_{;\mu\nu} &\equiv \nabla_\nu \nabla_\mu \phi \qquad \text{[sic!]} \\ &= (\phi_{;\mu})_{,\nu} + w_\nu \phi_{;\mu} - \Gamma^\alpha_{\mu\nu}\phi_{;\alpha} \\ &= (\phi_{,\mu\nu} + w_{\mu,\nu}\phi + w_\mu \phi_{,\nu}) + (w_\nu \phi_{,\mu} + w_\nu w_\mu \phi) \\ &\qquad - \Gamma^\alpha_{\mu\nu}\phi_{,\alpha} - \Gamma^\alpha_{\mu\nu} w_\alpha \phi.\end{aligned}$$

(Some mathematicians consider the classical index notation particularly offensive or misleading in this context; see Remark below.) We note that this expression is not symmetric in μ and ν; covariant differentiations in different directions need not commute. To be precise,

$$\phi_{;\nu\mu} - \phi_{;\mu\nu} = (w_{\nu,\mu} - w_{\mu,\nu} + [w_\mu, w_\nu])\phi + (\Gamma^\alpha_{\mu\nu} - \Gamma^\alpha_{\nu\mu})\phi_{;\alpha}\,.$$

Since the values of ϕ, $\phi_{;\alpha}$, and the left side of the equation at any particular $\underline{x}$ are all tensors, and since the values of ϕ and $\phi_{;\alpha}$ could be anything in the respective fiber spaces, the objects

$$Y_{\mu\nu}(\underline{x}) \equiv w_{\nu,\mu} - w_{\mu,\nu} + [w_\mu, w_\nu] \tag{8.10}$$

and

$$T^\rho_{\mu\nu}(\underline{x}) \equiv \Gamma^\alpha_{\mu\nu} - \Gamma^\alpha_{\nu\mu} \tag{8.11}$$

must be tensors themselves, being linear maps from one tensor space into another. Y is called the *curvature* tensor of the bundle to which ϕ belongs. Note that for *each* μ and ν, $Y_{\mu\nu}(\underline{x})$ is a linear map from $F_{\underline{x}}$ into itself, hence, in component language, an $r \times r$ matrix, or a tensor with one contravariant bundle index and one covariant bundle index (both suppressed in our notation). T is the *torsion* tensor; it has nothing specifically to do with the bundle with fiber F, but rather is completely determined by the connection defining covariant derivatives of vector fields on M. If the Christoffel symbols of that connection are symmetric in their two subscripts, as is usually assumed, then the torsion is zero. To get the curvature of this manifold connection, let the sections ϕ in our general formalism be the contravariant vector fields; then the w's in (8.10) are the Christoffel symbols, and one has

$$[Y_{\mu\nu}]^\alpha{}_\beta = R^\alpha{}_{\beta\mu\nu} \equiv \Gamma^\alpha_{\beta\nu,\mu} - \Gamma^\alpha_{\beta\mu,\nu} + \Gamma^\alpha_{\gamma\mu}\Gamma^\gamma_{\beta\nu} - \Gamma^\alpha_{\gamma\nu}\Gamma^\gamma_{\beta\mu}\,. \tag{8.12}$$

This is the *Riemann curvature tensor*. (See also (8.14).) The curvature tensor for covectors is the same with a minus sign (and a transposition, implemented automatically by the rule that a subscript can be contracted (summed over) only with a superscript). Finally, note that in the electromagnetic case, where $w_\mu = iA_\mu$,

$$Y_{\mu\nu} = i(A_{\nu,\mu} - A_{\mu,\nu}) \equiv iF_{\mu\nu}$$

is essentially the electromagnetic field-strength tensor.

Exercise 32: Verify directly from (8.10) that $Y_{\mu\nu}$ transforms under gauge transformations as a tensor with one covariant and one contravariant bundle index:

$$\tilde{Y}_{\mu\nu}(\underline{x}) = U(\underline{x})Y_{\mu\nu}(\underline{x})U(\underline{x})^{-1}.$$

The equation that resulted from our second-derivative calculation can be rewritten

$$\phi_{;\nu\mu} = \phi_{;\mu\nu} + Y_{\mu\nu}\phi + T^{\rho}_{\mu\nu}\phi_{;\rho}\,. \tag{8.13}$$

This *Ricci identity* tells one how to unscramble repeated covariant derivatives into any desired order. Once again, tensor fields of any type are covered by this formula if $Y\phi$ is suitably interpreted; in general it will be a sum of terms, some involving Riemann tensors and some involving the Y of some elementary type of field (the *gauge curvature* or *gauge field strength* of a physical gauge theory). For example, let G_μ be the object $M_{;\mu}$ of Exercise 31; then

$$G_{\mu;\rho\nu} = G_{\mu;\nu\rho} - R^{\alpha}{}_{\mu\nu\rho}G_\alpha + [Y_{\nu\rho}, G_\mu],$$

if there is no torsion and if Y now denotes the elementary gauge field strength.

Remark: A tensor identity such as

$$\nabla_\mu\nabla_\nu v^\rho = \nabla_\nu\nabla_\mu v^\rho + R^{\rho}{}_{\sigma\mu\nu}v^\sigma \tag{8.14}$$

can be interpreted either as an equation between abstract tensors, or as an equation between the concrete components of the tensors with respect to some basis. In the latter case, the correct interpretation of the left-hand side, for instance, is

$$(\mathbf{e}_\mu)^\alpha(\mathbf{e}_\nu)^\beta\nabla_\alpha\nabla_\beta\mathbf{v}$$

and *not*

$$(\mathbf{e}_\mu)^\alpha\nabla_\alpha\left((\mathbf{e}_\nu)^\beta\nabla_\beta\mathbf{v}\right).$$

There is thus a dangerous ambiguity in the notation if one insists on regarding "∇_μ" as an abbreviation for a directional derivative along a particular concrete basis-vector field, since one will then be led to the second interpretation, which differs from the first by a term involving the derivative of $\mathbf{e}_\nu$. Many modern texts on differential geometry give the Ricci identity in a directional-derivative form, in which an extra term must be subtracted off to account for the derivatives of the vector fields along which one is taking the directional derivatives. With the notation $\mathbf{u}\cdot\nabla$ for $u^\mu\nabla_\mu$, etc., that equation is (in the case of zero torsion)

$$\begin{aligned}R(\mathbf{v},\mathbf{u},\mathbf{w}) &\equiv \{R^\alpha{}_{\beta\mu\nu}v^\beta u^\nu w^\nu\}\\ &= [\mathbf{u}\cdot\nabla,\mathbf{w}\cdot\nabla]\mathbf{v} - ([\mathbf{u},\mathbf{w}]\cdot\nabla)\mathbf{v},\end{aligned}$$

where $[\mathbf{u},\mathbf{w}]$ is the vector field with components $u^\mu w^\alpha{}_{,\mu} - w^\mu u^\alpha{}_{,\mu}$ (which is, incidentally, the commutator of $\mathbf{u}\cdot\nabla$ and $\mathbf{w}\cdot\nabla$ when acting on *scalars*). For a *coordinate* basis, $[\mathbf{e}_\mu,\mathbf{e}_\nu]$ will always be zero, and so the form of the Ricci identity is unchanged after all; for more general basis-vector fields, the commutator term is nontrivial. In any case, however, the classical Ricci identity (8.14) (or (8.13)) is a valid equation when properly understood as referring to components of tensors with respect to an arbitrary basis *at a point*, not to basis-vector fields; and that is how we shall always understand it.

The curvature and torsion tensors have certain symmetries, which must be taken into account when attempting to simplify expressions involving them into a unique and compact form. From the definition, it's obvious that

$$Y_{\nu\mu} = -Y_{\mu\nu} \tag{8.15a}$$

and hence

$$R^\alpha{}_{\beta\nu\mu} = -R^\alpha{}_{\beta\mu\nu}\,. \tag{8.15b}$$

The torsion, also, is antisymmetric in its two subscripts. The *Bianchi identity* is

$$Y_{\alpha\beta;\gamma} + Y_{\beta\gamma;\alpha} + Y_{\gamma\alpha;\beta} = -T^{\rho}_{\alpha\beta} Y_{\gamma\rho} - T^{\rho}_{\beta\gamma} Y_{\alpha\rho} - T^{\rho}_{\gamma\alpha} Y_{\beta\rho}\,; \qquad (8.16\text{a})$$

as a special case,

$$R^{\mu}{}_{\nu\alpha\beta;\gamma} + R^{\mu}{}_{\nu\beta\gamma;\alpha} + R^{\mu}{}_{\nu\gamma\alpha;\beta} = \text{torsional terms.} \qquad (8.16\text{b})$$

There is also the so-called *cyclic identity*,

$$\begin{aligned} &R^{\mu}{}_{\alpha\beta\gamma} + R^{\mu}{}_{\beta\gamma\alpha} + R^{\mu}{}_{\gamma\alpha\beta} \\ &\quad = -T^{\mu}_{\alpha\beta;\gamma} - T^{\mu}_{\beta\gamma;\alpha} - T^{\mu}_{\gamma\alpha;\beta} + T^{\rho}_{\alpha\beta} T^{\mu}_{\rho\gamma} + T^{\rho}_{\beta\gamma} T^{\mu}_{\rho\alpha} + T^{\rho}_{\gamma\alpha} T^{\mu}_{\rho\beta}\,. \end{aligned} \qquad (8.17)$$

This may be regarded as a Bianchi identity for the torsion; but it is more familiar in the context where the torsion is zero, in which case (8.17) with (8.15b) simply says that R is annihilated by complete antisymmetrization on its last three indices. Many authors call (8.17) the *first Bianchi identity*; (8.16b) is then the *second Bianchi identity*. Finally, the Ricci identity (8.13), applied to any of these tensors or any covariant derivative (to any order) of one of them, belongs in the list of symmetries.

Proof of the Bianchi identities: Covariant derivatives satisfy the *Jacobi identity*

$$[\nabla_\mu, [\nabla_\nu, \nabla_\rho]]\phi + [\nabla_\nu, [\nabla_\rho, \nabla_\mu]]\phi + [\nabla_\rho, [\nabla_\mu, \nabla_\nu]]\phi = 0.$$

(This is a purely formal, or metamathematical, consequence of the antisymmetry of the definition of a commutator. The commutator here is just an abbreviation for the difference of two mixed second-order covariant derivatives in opposite order. It is not necessary to interpret the ∇ symbols as directional-derivative operations, so as to apply the Jacobi identity for *operators*. Therefore, it is correct to carry out this proof in the abstract index formalism.) Writing out all the derivatives and performing a few index manipulations, one arrives at

$$\begin{aligned} 0 = &Y_{\nu\rho;\mu}\phi + T^{\lambda}_{\nu\rho;\mu}\phi_{;\lambda} + T^{\lambda}_{\nu\rho} Y_{\mu\lambda}\phi + T^{\sigma}_{\nu\rho} T^{\lambda}_{\mu\sigma} u_{;\lambda} - R^{\lambda}_{\mu\nu\rho} u_{;\lambda} \\ &+ \text{cyclic permutations on the free indices.} \end{aligned}$$

Since ϕ and $\phi_{;\lambda}$ are arbitrary (and independent at any one $\underline{x}$), their respective coefficients may be set equal to 0. The results are versions of (8.16a) and (8.17).

In Riemannian geometry and gravitational theory one almost always deals only with connections which are *metric-compatible*:

$$\nabla_\rho g_{\mu\nu} = 0. \tag{8.18}$$

This condition has the happy consequence that differentiation commutes with raising and lowering of indices. It also implies a large number of additional symmetries for the Riemann tensor. Differentiating (8.18) once more and antisymmetrizing leads to the conclusion that R is antisymmetric in its *first* two indices:

$$R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\mu\nu}\,. \tag{8.19}$$

(The first index here is the erstwhile superscript, lowered by the metric tensor.) This, together with the cyclic identity, implies the *pair symmetry*:

$$R_{\gamma\delta\alpha\beta} = R_{\alpha\beta\gamma\delta} + \text{torsion terms}. \tag{8.20}$$

In addition, one usually takes the torsion to be zero. This uniquely determines the metric-compatible connection to be

$$\Gamma^\rho_{\mu\nu} = \tfrac{1}{2} g^{\rho\tau}(g_{\nu\tau,\mu} + g_{\mu\tau,\nu} - g_{\mu\nu,\tau})\,; \tag{8.21}$$

the resulting Riemann tensor (8.12) is a sum of terms linear in second (literal) derivatives of $\underline{g}$ and terms quadratic in its first derivatives. I have included the torsional terms in most of the foregoing formulas simply to make them more useful for reference. However, it turns out that the techniques to be described in the remainder of this chapter also apply to theories with torsion, provided that the torsion tensor is *totally antisymmetric*. By convention, the superscript of T is regarded as the last index, so that

$$T_{\alpha\beta\gamma} \equiv T^\rho_{\alpha\beta} g_{\rho\gamma}\,.$$

The condition in question is then that this totally covariant tensor be antisymmetric under all permutations of the six indices.

Exercise 33: Consider a generic derivative of the Riemann tensor,

$$\nabla^p R = \{R_{\mu_1\mu_2\mu_3\mu_4;\mu_5\ldots\,\mu_{4+p}}\}.$$

(Assume that the connection is metric-compatible.) Show that after all the symmetries are taken into account, there are $\frac{1}{2}(p+1)(p+4)$ independent permutations of the indices, modulo torsion and lower-order derivatives of R. Namely, let $(\alpha, \beta, \gamma, \delta, \epsilon, \dots)$ be a canonical ordering of the formal index list; then a nonredundant list of independent components of the tensor $\nabla^p R$ is

(a) $R_{\alpha\beta\gamma\delta;\epsilon\dots}$

(b) $R_{\alpha\beta\gamma\mu;\delta\dots}$ for p choices of μ

(c) $R_{\alpha\gamma\beta\delta;\epsilon\dots}$

(d) $R_{\alpha\gamma\beta\mu;\epsilon\dots}$ for p choices of μ

(e) $R_{\alpha\mu\beta\nu;\gamma\dots}$ for $\frac{1}{2}p(p+1)$ choices of pairs $(\mu\nu)$ of canonically ordered distinct indices chosen from $(\delta, \epsilon, \dots)$.

Parallel transport and geodesic distance; covariant power series

Let $x(\tau)$ be a curve in M, and let $\phi(\underline{x})$ be a section, defined at least on the [image of the] curve. One says that the values of ϕ are *parallel along* $x(\tau)$ if (with respect to any gauge)

$$\frac{d}{d\tau}\phi\big(x(\tau)\big) + w_\mu\big(x(\tau)\big)\phi\big(x(\tau)\big)\,\frac{dx^\mu}{d\tau} = 0. \tag{8.22}$$

The left-hand side is called the *absolute derivative* of ϕ along the curve. If ϕ is defined in an open neighborhood of the curve (at least), then its covariant derivative is defined and the absolute derivative can be expressed in terms of it as

$$\dot{x}(\tau)\cdot\nabla\phi \equiv \frac{dx^\mu}{d\tau}\,\nabla_\mu\phi\big(x(\tau)\big).$$

Suppose that $x(0) \equiv \underline{x}'$, $x(1) \equiv \underline{x}$, and we are given $\phi_0 \in F_{\underline{x}'}$. Then the result of *parallel transport* of ϕ_0 to $\underline{x}$ (along this particular curve) is defined to be $\phi(1)$, where $\phi(\tau)$ is the solution of the differential equation (8.22) with initial value $\phi(0) = \phi_0$. ($\phi(\tau)$ is in $F_{x(\tau)}$ for each τ.) Following DeWitt 1965, we write

$$\phi(1) \equiv I_x(\underline{x}, \underline{x}')\phi_0\,. \tag{8.23}$$

I_x is a one-to-one, linear mapping from $F_{\underline{x}'}$ onto $F_{\underline{x}}$; in concrete terms, it is a matrix-valued function of the two points, whose left index pertains to the fiber at $\underline{x}$ while its right index lives at $\underline{x}'$. (Consequently, when I is covariantly differentiated with respect to $\underline{x}$, for example, a connection form $w_\mu(\underline{x})$ attaches to the left index, but no connection form is needed for the right index.)

What is parallel transport in our two standard, elementary examples? In the electromagnetic case, (8.22) is

$$\frac{d\phi}{d\tau} + i\dot{x}^\mu A_\mu \phi = 0,$$

which is solvable in closed form:

$$\phi_1 = \exp\left(-i\int_{\underline{x}'}^{\underline{x}} A_\mu(\underline{y})\,dy^\mu\right)\phi_0\,,$$

where the line integral is along the curve $\underline{y} = x(\tau)$. Thus

$$I_x(\underline{x},\underline{x}') = \exp\left(-i\int_{\underline{x}'}^{\underline{x}} A_\mu(\underline{y})\,dy^\mu\right).$$

(In non-Abelian gauge theories A_μ is a matrix, (8.22) is a first-order differential *system*, and hence it generally can't be solved explicitly. Physicists sometimes denote the parallel-transport matrix in that case by

$$I_x(\underline{x},\underline{x}') = \mathcal{P}\exp\left(-i\int_{\underline{x}'}^{\underline{x}} A_\mu(\underline{y})\,dy^\mu\right),$$

where $\mathcal{P}$ indicates a *path-ordered exponential* summarizing the solution of the equation by perturbation theory (i.e., as a power series in the magnitude of A).) In the gravitational case, the equation of parallel transport for a contravariant vector field is, from (8.5),

$$\frac{dv^\nu}{d\tau} = -\Gamma^\nu_{\rho\mu} v^\rho \dot{x}^\mu. \tag{8.24}$$

DeWitt 1965 and Christensen 1976 denote the resulting parallel-transport matrix by $g^\mu{}_{\nu'}$, where primed indices refer to $\underline{x}'$ and unprimed ones to $\underline{x}$.

Intuitively, one might prefer to take parallel transport as the fundamental notion and define covariant differentiation from it: In the context of a vector bundle, the problem with defining a derivative by a difference quotient,

$$\left.\frac{d\phi}{d\tau}\right|_{\tau=0} \equiv \lim_{\Delta\tau\to 0}\frac{\phi(x(\Delta\tau)) - \phi(x(0))}{\Delta\tau}\,,$$

is that we don't know how to subtract values of ϕ at two different $\underline{x}$'s. But I_x gives us an identification of $F_{\underline{y}}$ with $F_{\underline{x}'}$, for each $\underline{y}$ on the curve x. (This identification depends on the curve!) Therefore, it is possible to modify the formula so that it makes sense: the absolute derivative at 0 is

$$\lim_{\Delta\tau\to 0} \frac{\phi(x(\Delta\tau)) - I_x(x(\Delta\tau), x(0))\phi(x(0))}{\Delta\tau}\,.$$

Exercise 34: Verify this formula for the absolute derivative within the framework of the definitions we've adopted. Would it matter if we put the inverse of I on the first term instead of I on the second?

The curve $x(\tau)$ is a *geodesic* if its own tangent vector is parallel along it. From (8.24), this condition is

$$\frac{d^2x^\nu}{d\tau^2} = -\Gamma^\nu_{\rho\mu}\,\frac{dx^\rho}{d\tau}\,\frac{dx^\mu}{d\tau}\,. \tag{8.25}$$

This is a nonlinear second-order differential equation, which may be treated like the equation of motion of a classical mechanical system.

Remark: The antisymmetric part of Γ, which is the torsion tensor, makes no contribution to (8.25). Therefore, it is often said that the geodesics of a manifold are independent of the torsion. There is a complication, however: If a connection is metric-compatible and has a torsion tensor which is *not* totally antisymmetric, then the symmetric part of that connection is *not* the metric-compatible connection (8.21). In such a situation there are two sets of geodesics, those defined by the metric through (8.21), and those defined by the torsional, metric-compatible connection. This is the reason why our discussions below of Synge–DeWitt tensors and asymptotic expansions of Green functions are limited to totally antisymmetric torsions.

The *arc length* along $x(\tau)$ from $\underline{x}'$ to $\underline{x}$ is

$$s \equiv \int_0^1 \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}\,d\tau.$$

(To avoid complications, assume momentarily that the metric is positive definite.) We are interested in this quantity almost exclusively in the case that $x(\tau)$ is a geodesic. In practice we shall almost never use the integral expression just given; certain formal properties of the geodesic arc length suffice in all the many calculations involving it.

We wish to regard the geodesic arc length (or *geodesic distance*) as a function of two arguments, $\underline{x}$ and $\underline{x}'$. Some questions of principle need to be disposed of first. Two points in M may be joined by more than one geodesic. (For example, if M is a 2-sphere with the usual metric, then antipodal points are joined by infinitely many geodesics of equal length, and a generic pair of points is joined by two arcs of a great circle, one relatively short and one long.) Moreover, if the manifold is incomplete, two points may not be joined by any geodesic at all. (Think of two points on opposite sides of a "hole" cut out of the plane.) In fact, the same is true for some *complete* pseudo-Riemannian manifolds; the anti-DeSitter space described in Chapter 6 is an example. But it is intuitively plausible that all $\underline{x}$ *sufficiently close* to a given $\underline{x}'$ can be joined to $\underline{x}'$ by a geodesic, which can be uniquely and continuously defined by characterizing it as the *shortest* geodesic between those two points. In fact, more is true [Whitehead 1932; Friedlander 1975, Sec. 1.2]:

Theorem: Every point of M has a neighborhood D which is *geodesically convex* (or *normal*); that is, any two points in D are joined by precisely one geodesic segment which lies entirely in D.

(Note that a comma after "segment" would turn this proposition into a falsehood!)

Henceforth we tacitly restrict our discussion to a normal neighborhood. Define

$$\sigma(\underline{x}, \underline{x}') \equiv \frac{1}{2}\, s^2.$$

This quantity was dubbed the *world function* by Synge 1960. Hadamard 1952, Synge 1960, and DeWitt 1965 have made marvelously productive use of it. See also the expositions of Garabedian 1964, esp. Sec. 2.6 and Chap. 5, and Friedlander 1975, esp. Sec. 1.2 (both of whom write 2Γ rather than σ). The world function is a C^∞ function (given a smooth metric to start from), unlike s, which has a conical singularity at 0. Furthermore, σ has a natural extension to manifolds of indefinite metric, in which case its values are (in my sign convention) positive for timelike separation, negative for spacelike separation, and zero on the light cone. In flat space, σ reduces to $\frac{1}{2}\|\underline{x} - \underline{x}'\|^2$.

Differentiating this last expression, we see that

$$\eta^{\mu\nu}\partial_\nu \sigma = (\underline{x} - \underline{x}')^\mu$$

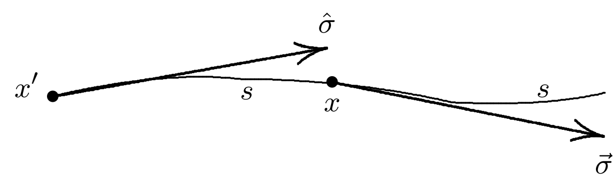
in flat space. The analogous statement in curved space is that $-g^{\mu\nu}\nabla_\nu\sigma$ *is the tangent vector* (to the geodesic) *at* $\underline{x}$ *pointing toward* $\underline{x}'$ *with length equal to* s. Since the vector $\nabla\sigma$ is even more important than the scalar σ, it is customary to omit the semicolon in writing covariant derivatives of σ:

$$\sigma_\mu \equiv \sigma_{;\mu}, \quad \text{etc.}$$

Another significant object is

$$\hat{\sigma}^\mu \equiv -\sigma^{\mu'},$$

which is *the tangent vector at* $\underline{x}'$ *pointing toward* $\underline{x}$ *with length* s. (The primed index denotes a derivative with respect to x'.)



Thus the geodesics set up a one-to-one correspondence between a neighborhood of $\underline{x}'$ in M and a neighborhood of 0 in the tangent space $T_{\underline{x}'}(M)$:

$$\underline{x} \longleftrightarrow \hat{\sigma}(\underline{x}, \underline{x}').$$

The components of $\hat{\sigma}$ form a special coordinate system, the *Riemann normal coordinates at* $\underline{x}'$. (Riemann normal coordinates can be constructed without reference to the world function; see, e.g., Parker 1979 or many textbooks on general relativity.) Alternatively, one may stick to general coordinates, or to coordinate-free methods, and think of $\hat{\sigma}$ as just a certain vector field. This latter approach combines the best of both worlds: the advantages of a special coordinate system in a manifestly covariant, or geometrically intrinsic, formalism.

From the geometric interpretation of the quantities involved, it is obvious that

$$\sigma = \frac{1}{2}\,\sigma_\mu\sigma^\mu. \tag{8.26}$$

This first-order, nonlinear partial differential equation is also the Hamilton–Jacobi equation associated with the mechanical system whose "Newtonian" equation of motion is the geodesic equation, (8.25).

Restricting attention to normal neighborhoods and to geodesics within them renders the parallel-transport operator unique, so that we may write $I(\underline{x}, \underline{x}')$ without ambiguity. Recall that I is, in effect, defined by the differential equation $\dot{x} \cdot \nabla I = 0$ and the initial condition $I(\underline{x}', \underline{x}') = 1$. In the present context, $\dot{x}$ is proportional to the σ vector, and so the equation can be written

$$\sigma^\mu I_{;\mu} = 0. \tag{8.27}$$

In dealing with functions of two variables, $f(\underline{x}, \underline{x}')$, a convenient shorthand is

$$[f] \equiv f(\underline{x}, \underline{x}).$$

Such a quantity is often called a *coincidence limit* in the physics literature, although the word "limit" is misleading — the function is simply to be *evaluated* at $\underline{x} = \underline{x}'$, with no commitment as to its continuity there. (The preferred jargon in the mathematical literature is "value on the diagonal".) When f is a covariant derivative of something, primes and the lack thereof on the indices are used to indicate which argument was the subject of the differentiation before the arguments were set equal:

$$[B_{;\mu\nu'}] \equiv \nabla_\mu \nabla_{\nu'} B(\underline{x}, \underline{x}')\big|_{\underline{x}=\underline{x}'}.$$

Note that $[B]_{;\mu}$ is not the same thing as $[B_{;\mu}]$, since the former implicitly involves differentiation with respect to $\underline{x}'$ as well as $\underline{x}$. From the multivariable chain rule (traditionally called **Synge's theorem** in this context), we have

$$[B]_{;\mu} = [B_{;\mu}] + [B_{;\mu'}]. \tag{8.28}$$

This principle can be used to obtain primed (and mixed) derivatives easily from unprimed ones, and vice versa. Note that primed and unprimed derivatives involve independent connections, so their order with respect to each other is arbitrary. Therefore, the derivative to be eliminated can always be assumed to act last. In particular,

$$[\hat{\sigma}^\mu{}_{\nu\rho\tau\ldots}] \equiv -[\sigma^{\mu'}{}_{\nu\rho\tau\ldots}] = -[\sigma_{\nu\rho\tau\ldots}{}^{\mu'}] = \nabla^\mu[\sigma_{\nu\rho\tau\ldots}] - [\sigma_{\nu\rho\tau\ldots}{}^{\mu}]. \tag{8.29}$$

Consider now a scalar function, $f(\underline{x}, \underline{x}')$ (where "scalar" means not only that the dependent variable has only one — real or complex — component, but also that the bundle connection is trivial). By the multivariable version of Taylor's theorem, f (to whatever order it is smooth) possesses a power-series expansion in the Riemann normal coordinates of $\underline{x}$ about $\underline{x}'$. The coefficients in the series are the coincidence limits of the *literal* partial derivatives of f with respect to $\underline{x}$. I claim, however, that they may be replaced by the covariant derivatives:

$$f(\underline{x}, \underline{x}') \sim [f] + [f_{;\mu}]\hat{\sigma}^{\mu} + \frac{1}{2!}[f_{;\mu\nu}]\hat{\sigma}^{\mu}\hat{\sigma}^{\nu} + \cdots . \tag{8.30}$$

Equivalently, the totally symmetric part of the covariant derivative is equal to the literal derivative in normal coordinates. (In (8.30) and the following discussion, the coincidence limits are evaluated at $\underline{x}'$. Equivalently, one can reverse the roles of the two points, putting primes on all the derivatives.)

To see this, regard $f(\underline{x}, \underline{x}')$ as a function of the arc-length parameter, s, of the geodesic from $\underline{x}'$ to $\underline{x}$. Apply the single-variable Taylor theorem and the chain rule:

$$\begin{aligned} f(\underline{x}, \underline{x}') &= f(x(s), \underline{x}') \\ &\sim \cdots + \frac{1}{2}\frac{d^2}{d\tilde{s}^2}\left(f(x(\tilde{s}), \underline{x}')\right)\Big|_{\tilde{s}=0} s^2 + \cdots \\ &= \cdots + \frac{1}{2}\frac{d}{d\tilde{s}}\left[\nabla_{\mu} f(x(\tilde{s}), \underline{x}')\frac{dx^{\mu}}{d\tilde{s}}\right]_{\tilde{s}=0} s^2 + \cdots . \end{aligned}$$

Absolute differentiation along the curve obeys the Leibnitz rule; we use that fact to calculate the needed second derivative. The absolute derivative of $dx^{\mu}/d\tilde{s}$ vanishes because the curve is a geodesic [(8.25)]. Thus the term is

$$\frac{1}{2}\nabla_{\nu}\nabla_{\mu} f(x(\tilde{s}), \underline{x}')\frac{dx^{\nu}}{d\tilde{s}}\frac{dx^{\mu}}{d\tilde{s}} s^2 \qquad (\tilde{s} = 0).$$

But $dx^{\mu}/d\tilde{s}$ is the *unit* tangent vector to the geodesic, which becomes $\hat{\sigma}$ upon absorbing a factor s. So we have reproduced the second-order term in (8.30). The argument continues by induction to higher orders.

If f is replaced by a section of a vector bundle, (8.30) as it stands does not make sense, because its two sides belong to different fibers. However, inserting a parallel-transport operator makes it correct:

$$\phi(\underline{x}, \underline{x}') \sim \sum_{p=0}^{\infty}\frac{1}{p!} I(\underline{x}, \underline{x}')[\phi_{;\mu_1\ldots\mu_p}]\hat{\sigma}^{\mu_1}\cdots\hat{\sigma}^{\mu_p} . \tag{8.31}$$

The proof is the same as for (8.30), except that it must be applied to $I(\underline{x},\underline{x}')^{-1}\phi(\underline{x},\underline{x}')$ and that the absolute derivative makes its appearance already in the first order. The absolute derivative of I vanishes [(8.27)].

Now take a covariant derivative of (8.30) with respect to $\underline{x}$, recalling that the coincidence limits there are independent of $\underline{x}$. Take the coincidence limit of the result, noting that $[\hat{\sigma}^\mu] = 0$:

$$[f_{;\nu}] = [f_{;\mu}][\hat{\sigma}^\mu{}_\nu]\,.$$

Since ∇f is arbitrary, it must be true that $[\hat{\sigma}^\mu{}_\nu] = \delta^\mu_\nu$, or, equivalently, $[\sigma_{\mu\nu}] = g_{\mu\nu}$. Considering higher derivatives of (8.30) in the same way, one concludes that *the totally symmetric* (in the subscripts) *part of* $[\hat{\sigma}^\mu{}_{\nu_1\ldots\nu_p}]$ *equals* 0 *whenever* $p \neq 1$. Considering derivatives of (8.31), one concludes in the same way that *the totally symmetric part of* $[I_{\mu_1\ldots\mu_p}]$ *is* 0 *whenever* $p \neq 0$. These facts about the diagonal values of derivatives of σ and I also follow easily from the explicit calculations in the next section; reversal of the present argument then provides a more concrete proof of (8.30) and (8.31) [cf. Widom 1980, reinterpreted following Drager 1978].

Recursive calculations of the Synge–DeWitt tensors

The objects $[\hat{\sigma}^\lambda{}_{\mu\nu\ldots}]$, $[\sigma_{\mu\nu\ldots}]$, and $[I_{;\mu\nu\ldots}]$ are of the utmost importance in computing asymptotic expansions of Green functions of partial differential operators, both by the classical HaMiDeW methods described in the next chapter and by the more modern and more general method of pseudodifferential operators [Widom 1980, Fulling & Kennedy 1988]. As we've seen in (8.29), the first class of these is easy to obtain from the second. Let us refer to the other two classes as the *Synge–DeWitt tensors*. DeWitt 1965 gave an algorithm for calculating them from the basic properties (8.26) and (8.27). It has been pushed to high order, with various degrees of computer assistance, by Schimming 1981, Rodionov & Taranov 1987, Fulling 1989 and unpublished, and Christensen & Parker 1989.

Remark: Widom 1980 gives a different algorithm, which, unlike DeWitt's, applies to connections which are not metric-compatible (or have incompletely antisymmetric torsion). (In that context σ is not defined, but $\hat{\sigma}^\mu$ is.) However, that approach requires consideration of many more index permutations than does DeWitt's ($p!$ versus p).

We already know that

$$[\sigma] = 0, \qquad [\sigma_\mu] = 0, \qquad [\sigma_{\mu\nu}] = g_{\mu\nu}\,. \tag{8.32}$$

To obtain the next item in the list by DeWitt's method, differentiate (8.26) three times:

$$\sigma_{\mu\nu\rho} = \sigma^\alpha{}_{\nu\rho}\sigma_{\alpha\mu} + \sigma^\alpha{}_\nu\sigma_{\alpha\mu\rho} + \sigma^\alpha{}_\rho\sigma_{\alpha\mu\nu} + \sigma^\alpha\sigma_{\alpha\mu\nu\rho}\,.$$

Take the coincidence limit, using (8.32):

$$[\sigma_{\mu\nu\rho}] = [\sigma_{\mu\nu\rho}] + [\sigma_{\nu\mu\rho}] + [\sigma_{\rho\mu\nu}].$$

Subtract $3[\sigma_{\mu\nu\rho}]$ and use the Ricci identity:

$$\begin{aligned}
-2[\sigma_{\mu\nu\rho}] &= ([\sigma_{\nu\mu\rho}] - [\sigma_{\mu\nu\rho}]) + ([\sigma_{\rho\mu\nu}] - [\sigma_{\mu\nu\rho}]) \\
&= [\sigma_{\nu\mu\rho}] - [\sigma_{\mu\nu\rho}] + ([\sigma_{\rho\mu\nu}] - [\sigma_{\mu\rho\nu}] + [\sigma_{\mu\rho\nu}] - [\sigma_{\mu\nu\rho}]) \\
&= [(T^\alpha_{\mu\nu}\sigma_\alpha)_{;\rho}] + [(T^\alpha_{\mu\rho}\sigma_\alpha)_{;\nu}] + [T^\alpha_{\nu\rho}\sigma_{\mu\alpha}] - [\sigma_\alpha R^\alpha{}_{\mu\nu\rho}] \\
&= T_{\mu\nu\rho} + T_{\mu\rho\nu} + T_{\nu\rho\mu} \\
&= T_{\mu\nu\rho}\,,
\end{aligned}$$

since T must be assumed antisymmetric. Therefore,

$$[\sigma_{\mu\nu\rho}] = -\tfrac{1}{2}T_{\mu\nu\rho}\,. \tag{8.33}$$

In particular, $[\sigma_{\mu\nu\rho}] = 0$ in a torsion-free theory, the case of usual interest.

Exercise 35:

(a) Show that if there is no torsion, then

$$[\sigma_{\mu\nu\rho\tau}] = \tfrac{1}{3}(R_{\rho\mu\nu\tau} + R_{\tau\mu\nu\rho}). \tag{8.34}$$

(b) Find the torsion terms left out of (8.34). *Answer:*

$$-\tfrac{1}{3}(T_{\mu\nu\rho;\tau} + T_{\mu\nu\tau;\rho}) + \tfrac{1}{4}T^\alpha_{\mu\nu}T_{\alpha\rho\tau} - \tfrac{1}{12}(T^\alpha_{\mu\rho}T_{\alpha\nu\tau} + T^\alpha_{\mu\tau}T_{\alpha\nu\rho}).$$

Exercise 36:

(a) Show that the pth derivative of (8.26) is

$$\sigma_{\mu_1\ldots\mu_p} = \sum \sigma^\alpha{}_{\nu_1\ldots\nu_k}\sigma_{\alpha\mu_1\rho_2\ldots\rho_{p-k}}\,,$$

where the sum is over all 2^{p-1} partitions of the formal index sequence $(\mu_1, \ldots, \mu_p)$ into two subsequences $(\nu_1, \ldots, \nu_k)$ and $(\mu_1, \rho_2, \ldots, \rho_{p-k})$ with $\mu_1 \equiv \rho_1$ belonging to the second sequence. (Both subscript lists (not including α) remain in their natural order within $(\mu_1, \ldots, \mu_p)$.)

(b) Assuming no torsion, show in analogy with our treatment of the case $p = 3$ that

$$[\sigma_{\mu_1 \ldots \mu_p}] = \frac{1}{p-1} \Biggl\{ \sum_{i=2}^{p} \sum_{j=1}^{i-1} \sum_{\substack{\text{subseqs.,}\\ 0 \le k \le p-j-1}} \sum_{l=1}^{j-1} [\sigma_{\mu_1 \ldots \mu_{l-1} \alpha \mu_{l+1} \ldots \mu_{j-1} \nu_1 \ldots \nu_k}] R^{\alpha}{}_{\mu_l \mu_j \mu_i ; \rho_1 \ldots \rho_{p-j-1-k}} - \sum_{\substack{\text{subseqs.,}\\ 2 \le k \le [n/2]}} [\sigma_{\alpha \mu_1 \rho_2 \ldots \rho_{p-k}}][\sigma^{\alpha}{}_{\nu_1 \ldots \nu_k}] \Biggr\}.$$

In the first term the ν's are a subsequence (of length k) of the sequence $(\mu_{j+1}, \ldots, \mu_p)$ with μ_i omitted. In the other term the ν's are any subsequence of the original subscript list of length $k \le [n/2]$, with the proviso (to prevent double counting) that $\nu_1 = \mu_1$ if $k = n/2$, and the ρ's are the complementary sequence.

The formula in part (b) of this exercise is not yet a *solution* of the recursion relation for $[\nabla^p \sigma]$, because it still involves $[\nabla^q \sigma]$ with $q < p$ on the right-hand side. But it is a start! The number of terms in the explicit formula for $[\nabla^p \sigma]$ is extremely large if $p \ge 8$. Research is continuing on the use of computers to evaluate quantities formed from $[\nabla^p \sigma]$ (such as the terms of the HaMiDeW series in the next chapter) in an efficient way.

The derivatives of I can be computed similarly. We already know that

$$[I] = 1 \quad \text{(the identity operator in the fiber).} \tag{8.35}$$

By differentiating (8.27), setting $\underline{x}' = \underline{x}$, taking (8.32) into account, and using the Ricci identity to isolate the derivative of highest order, one finds

$$[I_\mu] = 0, \qquad [I_{;\mu\nu}] = -\tfrac{1}{2} Y_{\mu\nu}\,, \tag{8.36}$$

and higher-order formulas involving the Riemann (and torsion) tensors as well as Y and its derivatives.

The formulas for the tensors $[\nabla^p\sigma]$ and $[\nabla^p I]$ have the following important qualitative properties (which can easily be proved by induction, absent an actual solution of the recursions):

(1) Each term is a product of covariant derivatives of the tensors R, T, and (in the case of I) Y.

(2) Each term in $[\nabla^{p+2}\sigma]$ or $[\nabla^p I]$ has *order* exactly p, in the following sense: Count 1 for each explicit covariant differentiation, 2 for each R, 1 for each T, and 2 for each Y, and add up. (Recall that T and R respectively involve 1 and 2 differentiations of the metric tensor.) The order of a quantity is the same as its physical dimension in powers of [mass] or [length]$^{-1}$, if $\hbar = c = 1$.

(3) There are no internal contractions (e.g., yielding Ricci tensors). In fact, there are not even any closed loops of contractions involving more than one factor. On the other hand, all the factors are linked by contractions into a single structure, if the entire product of Y-factors found in a term of $[\nabla^p I]$ is regarded as a single factor. That is, in the terminology of graph theory [Harary 1969], the contraction structure of a term is a *tree* — a graph which is both acyclic and connected. The number of contractions is exactly one fewer than the number of factors, as is clear from the way the terms are built up [see Exercise 36(b)]. (According to Harary, this is precisely the condition under which acyclicity and connectedness become equivalent.)

Exercise 37: Show that any term of the tree form has the correct relationship between *order* and *number of free indices* required for it to be a potential term in the Synge–DeWitt tensor of the appropriate order. (This does not mean that all possible tree terms actually appear.)