A QUICK INTRODUCTION TO THE EINSTEIN EQUATIONS

DOUGLAS N. ARNOLD
INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

Abstract. These notes, written for the IMA workshop on Numerical Relativity, June 2002, are meant to give a fairly self-contained, quite formal, and very succinct introduction to the Einstein equations, including the required differential geometry and a choice of notational conventions. The equations are first presented in an entirely coordinate-free manner emphasizing their geometric content, and then the corresponding PDEs satisfied by the metric components with respect to some coordinate system are derived. The gauge freedom in the equations is discussed both in the coordinate-free and the coordinatized context.

1. Tensors on manifolds

Let M denote a smooth n-dimensional manifold. At each point $p \in M$ the tangent space T_pM to M at p is an n-dimensional vectorspace. The tangent bundle TM is the disjoint union of all the T_pM, $p \in M$, which is itself a smooth $2n$-dimensional manifold. A vectorfield on M is a smooth section on this bundle, i.e., a smooth function $v : M \to TM$ such that $v(p) \in T_pM$. Henceforth we will usually write v_p rather than $v(p)$ when evaluating sections.

To any finite-dimensional vectorspace V we may associate the dual vectorspace V^* of linear functionals on V. We have $\dim V^* = \dim V$, but without further structure there is no canonical isomorphism between V and V^*. For integers $k, l \geq 0$, we may form the tensor product

$$\underbrace{V \otimes \cdots \otimes V}_{k} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{l},$$

which is a vectorspace of dimension n^{k+l}. It is canonically identified with the space of $(k+l)$-linear maps $\underbrace{V^* \times \cdots \times V^*}_{k} \times \underbrace{V \times \cdots \times V}_{l} \to \mathbb{R}$.

In the case $V = T_pM$, we write $T_p^{(k,l)}M$ for this tensor product space, and $T^{(k,l)}M$ for the associated bundle. A (k,l) tensor is a smooth section of this bundle. That is, a (k,l) tensor is a map v on M for which v_p is a $(k+l)$-linear map acting on k covectors and l vectors at p. For example, a $(1,0)$ tensor is a vectorfield on M and a $(0,1)$ tensor is a covectorfield, and a $(0,0)$ tensor is simply a real-valued smooth function on M. The set of all (k,l) tensors is an (infinite-dimensional) vectorspace which we denote as $\mathcal{T}_{k,l}M$. The pair (k,l) is called the *variance* of the tensor.

2. Metrics

Let $g \in \mathcal{T}_{0,2}M$. Then at each point p, g_p is a bilinear map $T_pM \times T_pM \to \mathbb{R}$. If for each p this map is symmetric (that is $g_p(v,w) = g_p(w,v)$), and non-degenerate (which means that $g_p(v,w) = 0$ for all $w \in T_pM$ only if $v = 0$), then g is called a *pseudo-Riemannian metric* or simply a metric on M. That is, a metric is a map which associates to each point

p of M a pseudo-innerproduct $T_pM \times T_pM \to \mathbb{R}$ on the tangent space at the point given by $(v, w) \mapsto g_p(v, w)$. The prefix pseudo is used because the product is only assumed to be non-degenerate, not positive definite. If the metric gives a true, positive definite innerproduct at each point, it is called a *Riemannian metric*.

Now a pseudo-innerproduct on a vector space V sets up a natural isomorphism $V \to V^*$, and so there is a natural identification of $T^{(k,l)}M$ and $T^{(k+l,0)}M$ once a metric is specified on M.

3. Notational conventions

So far we have had no need of a coordinate system on the manifold, and we shall continue for a while without introducing one. However it is useful at this point to introduce a notation, called the *abstract index notation*, which, while remaining coordinate-free, has the appearance of coordinate notation. The notation is simply this. We use symbols with a single superscript, like v^a to represent vectors or vectorfields. The subscript a has no particular meaning. It is just an adornment like the arrow in the notation $\vec{v}$. v^a and v^b are two ways of writing precisely the same object (in these notes we use the letters a through h for abstract indices and the letters i through m as actual indices). We use subscripts for covector(fields): e.g., v_a. For a (k, l) tensor we use k distinct superscripts and l distinct subscripts: a metric might be written g_{ab} and a $(1, 3)$ tensor R^d_{abc}. Now g_{ab} acts on two vectors v^a and w^a to produce a real number. We write this as $g_{ab}v^a w^b$. Unless g_{ab} is symmetric this is different from $g_{ab}w^a v^b$ (which we could as well write $g_{ab}v^b w^a$ or $g_{ba}v^a w^b$). This exemplifies the use of different subscripts and superscripts to indicate how one tensor is operating on another. As another use, we can write $g_{ab}v^a$. This is a covector with the single index b—indices that are repeated as subscript and superscript in a product expression "contract" away and are not counted in determining the type of the expression. In non-index notation this would be the covector $g(v, \cdot)$. Some other useful quantities that are expressed easily in abstract index notation are the tensor product of two tensors $v^{c\cdots d}_{a\cdots b}$ and $w^{g\cdots h}_{e\cdots f}$, which is simply $v^{c\cdots d}_{a\cdots b}w^{g\cdots h}_{e\cdots f}$; and the trace of a $(1, 1)$ tensor[1] v^b_a, which is simply v^a_a. Given a (k, l) tensor with $k, l > 0$, we can take the trace with respect to any one of the k covector arguments and any one of the l vector arguments. This is hard to notate without indices, but in abstract index notation, we simply write $v^{d\cdots b\cdots e}_{a\cdots b\cdots c}$ for this trace, with the repeated index serving to signal the variables with respect to which the trace is taken.

Another notation used with abstract index notation is the use of parentheses about indices to indicate the symmetric part of a tensor with respect to those indices, and brackets to indicate the antisymmetric part. Thus, for example,

$$v_{(ab)c} := \frac{1}{2}(v_{abc} + v_{bac}), \quad v_{[ab]c} := \frac{1}{2}(v_{abc} - v_{bac}).$$

By $v_{(abc)d}$ we would mean the sum of the 6 terms obtained by permuting the indices a, b, and c, divided by 6. The antisymmetric part $v_{[abc]d}$ would be the same sum except that each term would be multiplied by the sign of the corresponding permutation. The equation $v_{abcd} = v_{(abc)d}$ means that the tensor v is symmetric with respect to its first three arguments.

[1]If V is a vectorspace, the trace operator $V \otimes V^* \to \mathbb{R}$ is well-defined. By contrast, there is no natural way to define a trace on $V \otimes V$ or $V^* \otimes V^*$ unless a metric (or other additional structure) is given.

When a metric is given there is another useful notational convention employed together with abstract index notation. In this situation we may identify a vector with a covector. The convention is to use the same letter for these, only moving the index. Thus the covector associated to v^a would be written as v_a (or v_b). The formula is simply $v_b = g_{ab}v^a$, i.e., tensor with the metric and then take the trace. In the same way we use the metric to lower multiple indices in a higher order tensor: if v_{ab}^{cde} is a $(3,2)$ tensor, v_{abcd}^{e} is defined as the tensor $g_{cf}g_{dh}v_{ab}^{fhe}$. The inverse metric, g^{ab} is defined by $g^{ab}v_b = v^a$ for all v^a, and may be used to raise indices in tensors of arbitrary order.

4. The covariant derivative operator

A metric on M uniquely determines a covariant derivative operator ∇ as shall now be explained. The covariant derivative acts on a tensor of any order (k,l) and returns one of order $(k,l+1)$. In index-free notation we write $v \mapsto \nabla v$, while in abstract index notation it would be $v_{a\cdots b}^{c\cdots d} \mapsto \nabla_e v_{a\cdots b}^{c\cdots d}$. The covariant derivative operator has the following properties:

(1) (linearity) ∇ is a linear operator from $\mathcal{T}_{k,l}M \to \mathcal{T}_{k,l+1}M$ for all k, l.
(2) (Leibniz rule) $\nabla(v \otimes \mu) = (\nabla v) \otimes \mu + v \otimes (\nabla \mu)$ for all tensors v and μ.
(3) (commutativity with traces) $\nabla_b(v_{\cdots a \cdots}^{\cdots a \cdots}) = \nabla_b v_{\cdots a \cdots}^{\cdots a \cdots}$. (i.e., ∇ applied to the $(k-1,l-1)$ tensor obtained by taking a trace of an (k,l) tensor v is the same as the corresponding trace of the $(k,l+1)$ tensor ∇v.
(4) (differentiation of scalarfields) If $f \in \mathcal{T}_{0,0}M = C^\infty(M)$, then $(\nabla f)_p : T_pM \to \mathbb{R}$ is simply the usual differential or exterior derivative df_p.
(5) (symmetry) If $f \in C^\infty(M)$, then the $(0,2)$ tensor $\nabla_a \nabla_b f$ is symmetric.
(6) (compatibility with the metric) If g_{ab} is the metric, then $\nabla_a g_{bc} = 0$.

It is a (fairly easy) theorem, that, given the metric g_{ab} there is a unique covariant derivative operator satisfying these properties.

Because of the Leibniz rule, commutativity with traces, and compatibility with the metric, the abstract index notation for covariant derivatives is unambiguous. For example, given a covectorfield v_b, the notation $\nabla_a v^b$ might mean the result of raising the index on the covectorfield to obtain a vectorfield v^b and then differentiating to obtain a $(1,1)$ tensor, or it might mean the result of differentiating the covectorfield to obtain a $(0,2)$ tensor and then raising the second index of that. Because of the properties just mentioned, these in fact give the same result.

5. The Riemann curvature tensor

If f is a smooth function, then the symmetry $\nabla_{[a}\nabla_{b]}f = 0$ is one of the defining conditions of the covariant derivative. But if we apply ∇ twice to a vectorfield, covectorfield, or higher order tensor, such commutativity will not generally hold. For example, for a covector v_c, $\nabla_{[a}\nabla_{b]}v_c$ will not generally vanish, but it can be shown that at any point it depends linearly on v_c at that point, i.e.,

$$\nabla_{[a}\nabla_{b]}v_c = \frac{1}{2}R_{abc}^{d}v_d \tag{1}$$

for some $(1,3)$ tensor called the *Riemann curvature tensor*. The corresponding $(0,4)$ tensor R_{abcd} is also called the Riemann curvature tensor and has the following symmetries

$$R_{(ab)cd} = 0, \quad R_{abcd} = R_{cdab}, \quad R_{[abc]d} = 0.$$

These symmetries imply that the Riemann tensor is determined by $n^2(n^2-1)/12$ numbers at each point; that is, 1 in 2 dimensions (the Gaussian curvature), 6 in 3 dimensions, and 20 in 4 dimensions.

In addition, the Riemann tensor satisfies a differential identity, the Bianchi identity:

$$\nabla_{[a} R_{bc]de} = 0.$$

The trace $R_{ab} := R^d_{adb}$ is called the *Ricci tensor*. It is a symmetric $(0,2)$ tensor, like the metric. Its trace, $R := R^a_a = g^{ab} R_{ab}$, is called the *scalar curvature*. The *Einstein tensor* is then given by the formula

$$G_{ab} = R_{ab} - \frac{1}{2} R g_{ab}.$$

In 4-dimensions this means that the Einstein tensor has the same trace-free part as the Ricci tensor, but the opposite trace: it is the trace-reversed Ricci tensor. From the Bianchi identity, after taking traces twice, we find that G_{ab} is not an arbitrary symmetric $(0,2)$ tensorfield, but rather it is always divergence-free:

$$\nabla^a G_{ab} := g^{ac} \nabla_c G_{ab} = 0.$$

6. The Einstein equations

The vacuum Einstein equations are simply $G_{ab} = 0$. That is, a solution of the vacuum Einstein equations consists of a manifold M and a metric g_{ab} on M for which the Einstein tensor vanishes (or, equivalently, the metric is *Ricci-flat*, that is, the Ricci tensor vanishes). More precisely, we are only interested in manifolds of dimension 4 and Lorentzian metrics, i.e., metrics with signature $(-,+,+,+)$ (this means that at each point p, the metric is not positive definite on T_pM but it is positive definite on a 3-dimensional subspace of T_pM). The full Einstein equations replace the vanishing of the Einstein tensor with the equation $G_{ab} = T_{ab}$ where T_{ab} is the *stress-energy tensor*, a symmetric $(0,2)$ tensor computed from a *matter model*. Here we consider only the vacuum Einstein equations, but a brief discussion of matter models and the stress-energy tensor will be given in Alan Rendall's introduction to the Cauchy problem. Note that there the stress-energy tensor surely cannot be arbitrary if there is to exist a solution to the Einstein equations. By the contracted Bianchi identity, it must be divergence-free.

7. Gauge freedom

If M and N are any smooth manifolds and $\phi : M \to N$ a smooth mapping, then for each point $p \in M$, the differential $d\phi_p$ is a linear map from T_pM to $T_{\phi(p)}N$. We can combine these for the different points $p \in M$ to define the *push-forward* map $\phi_* : T^{(1,0)}M \to T^{(1,0)}N$ which associates to a vector $v_p \in T_pM$ the vector $d\phi_p\, v_p \in T_{\phi(p)}N$.

Now, covectors map the other way. Namely the adjoint map $(d\phi_p)^*$ maps the covector space $(T_{\phi(p)}N)^* \to (T_pM)^*$. We cannot in general combine these to create a map $T^{(0,1)}N \to T^{(0,1)}M$ between the covector bundles. Indeed, if ϕ is not onto, then not every element of the covector bundle of N belongs to some $(T_{\phi(p)}N)^*$, and if ϕ is not one-to-one, then some elements will map to two different covector spaces of M.

Now suppose that ϕ is a *diffeomorphism*, i.e., a smooth one-to-one and onto map $M \to N$. Then the problem just mentioned disappears, and we can define the *pull-back* $\phi^* : T^{(0,1)}N \to T^{(0,1)}M$. But we also have $\phi^{-1} : N \to M$ at our disposal, we can also define the push-forward of covectors by $\phi_* := (\phi^{-1})^* : T^{(0,1)}M \to T^{(0,1)}N$. Taking tensor products, in the case of a diffeomorphism, we have the push-forward on all the tensor bundles $\phi_* : T^{(k,l)}M \to T^{(k,l)}N$ (and, of course, a pull-back as well). Finally, given a (k,l) tensor v on M (so $v_p \in T^{(k,l)}M$), we can define $\phi_* v$ as the (k,l) tensor on N taking $\phi(p)$ to $\phi_* v_p$. Thus we overload the notation so that $\phi_* : \mathcal{T}^{(k,l)}M \to \mathcal{T}^{(k,l)}N$ as well.

Take the case of a manifold M with a metric $g \in \mathcal{T}^{(0,2)}M$. The diffeomorphism ϕ sets up a one-to-one correspondence between points p of M and $q = \phi(p)$ of N, and also a linear isomorphism between the space T_pM and T_qN. Untangling the definitions, we find that $\phi_* g$ assigns to two vectors in T_qN exactly the same value as g assigns to the corresponding vectors in T_pM. In other words, by construction, $\phi_* g$ is a metric on N which is *isometric* to the metric g on M.

It also follows directly from the definitions that $\phi_*(\nabla v) = \nabla(\phi_* v)$ where on the left ∇ denotes the covariant derivative associated with g and on the right it denotes the covariant derivative associated with $\phi_* g$ (and v is a tensor of arbitrary variance on M). From this we get that the Riemann tensor associated to $\phi_* g$ on N is just the push-forward of the Riemann tensor associated to g on M, and similarly for the Ricci tensor, scalar curvature, and Einstein tensor.

This brings us to an important, if in some sense trivial, conclusion: if g is a metric on a manifold M satisfying the vacuum Einstein equations ($G_{ab} = 0$), then for any diffeomorphism $\phi : M \to N$, $\phi_* g$ is a metric on N satisfying the vacuum Einstein equations. Thus the Einstein equations admit a great deal of non-uniqueness. A manifold can be expected to admit many diffeomorphisms onto itself, and so if it admits a single solution to the vacuum Einstein equations, it admits many. If the manifold has a boundary, we can choose the diffeomorphisms which reduce to the identity in a neighborhood of the boundary. This shows that boundary conditions can never be sufficient to suppress the non-uniqueness. Instead, if we are to get unique solutions we will need to look for equivalence classes of metrics, where two metrics are equivalent if they are related by push-forward through a diffeomorphism, i.e., if they are isometric.

The ability to generate many solutions from a single solution through diffeomorphisms is called the *gauge freedom* of the Einstein equations. Any procedure to fix a particular solution in the equivalence class is called *gauge fixing*.

8. Frames and coordinates

Up until now our discussion has been entirely coordinate-free. However if we wish to view the Einstein equations as partial differential equations, for example in order to use standard numerical methods, we need to introduce coordinates.

First we discuss frames. A frame is an ordered basis in each of the tangent spaces T_pM which depends smoothly on p. Otherwise stated a frame consists of n vectorfields $X_1^a, \dots, X_n^a$ which are linearly independent at each point. (Note that the subscripts here are actual indices, not abstract indices.) If a metric is given we may talk about orthogonal frames ($g_{ab}X_i^aX_j^b = 0$ if $i \neq j$) and orthonormal frames (where also $g_{ab}X_i^aX_i^b = \pm 1$ for each i).

Given a basis for a vectorspace V, by duality we obtain a basis for V^* (the dual basis), and then by taking tensor products, a basis for all of the tensor product spaces $V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$. Thus a frame determines a smoothly varying basis for all the spaces $T_p^{(k,l)}M$. We can then define the components of any (l, k) tensor at p as the numbers obtained by evaluating the tensor at the basis functions for $T_p^{(k,l)}M$. Thus, for example, the components of a metric g_{ab} would be the $n \times n$ array $(g_{ab}X_i^aX_j^b)_{1 \leq i,j \leq n}$ whose entries depend on p.

Each choice of a frame on a manifold with metric leads to a triply-indexed array of numbers a_{ij}^k called the *connection coefficients* with respect to the frame. The defining conditions are

$$X_i^a \nabla_a X_j^b = \sum_{k=1}^n a_{ij}^k X_k^b, \quad i, j = 1, \ldots n.$$

Now suppose that we have a system of coordinates on M (typically we will need to decompose M into overlapping patches and set up different coordinate systems on each, but for simplicity of notation, let us assume that we can find global coordinates on M). By a *coordinate system* we mean a set on n smooth functions $x^i : M \to \mathbb{R}$ such that $(x^1, \ldots, x^n) : M \to \mathbb{R}^n$ is an diffeomorphism of M onto an open subset Ω of $\mathbb{R}^n$. Let $\Phi : \Omega \to M$ denote the inverse map. If $\Phi(z) = p$, then $d\Phi_z$ is a linear isomorphism of $\mathbb{R}^n$ (the tangent space to Ω at z) onto T_pM. Thus the standard basis for $\mathbb{R}^n$ maps, via $d\Phi_z$, to a basis for T_pM. In this way a coordinate system determines a frame, called a *coordinate frame.* Note that on most manifolds there are no coordinates for which the coordinate frame is orthonormal. So we usually have to choose between computing with coordinates or computing with orthonormal frames.

For a coordinate frame, the connection coefficients are almost always denoted Γ_{jk}^i and are called the *Christoffel symbols* of the metric with respect to the particular coordinates. They satisfy $\Gamma_{jk}^i = \Gamma_{kj}^i$ and may be computed explicitly:

$$\Gamma_{jk}^i = \frac{1}{2} g^{il} \left(\frac{\partial g_{lk}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^l} \right). \tag{2}$$

Here the g_{ij} are the components of the metric (which may be viewed as function on Ω as so differentiated), and the g^{ij} are the components of the inverse metric, which, as a matrix, is inverse to the matrix of components of the metric. We are of course employing the Einstein summation convention that repeated indices are summed from 1 to n.

The Christoffel symbols allow us to explicitly compute the components of the covariant derivative of a tensor in terms of the component of the tensor. Writing $\nabla_i v_{j \cdots k}^{l \cdots m}$ for the components of the covariant derivative of $v_{a \cdots b}^{c \cdots d}$, we have

$$\nabla_i v = \frac{\partial v}{\partial x^i}, \quad \nabla_i v^j = \frac{\partial v^j}{\partial x_i} + \Gamma_{ik}^j v^k, \quad \nabla_i v_j = \frac{\partial v_j}{\partial x_i} - \Gamma_{ij}^k v_k, \quad \nabla_i v_k^j = \frac{\partial v_k^j}{\partial x_i} + \Gamma_{il}^j v_k^l - \Gamma_{ik}^m v_m^j.$$

The pattern continues for higher order tensors: for each index one term involving a Christoffel symbol is added (for raised indices) or subtracted (for lowered indices).

From the defining equation of the Riemann curvature tensor, (1), we can then express its components in terms of the Christoffel symbols and their partial derivatives:

$$R^l_{ijk} = \frac{\partial \Gamma^l_{jk}}{\partial x^i} - \frac{\partial \Gamma^l_{ik}}{\partial x^j} + \Gamma^m_{jk}\Gamma^l_{mi} - \Gamma^m_{ik}\Gamma^l_{mj} \tag{3}$$

Of course the components of the Ricci tensor, scalar curvature, and Einstein tensor are obtained as

$$R_{ij} = R^l_{ilj}, \quad R = g^{ij}R_{ij}, \quad G_{ij} = R_{ij} - \frac{1}{2}Rg_{ij}. \tag{4}$$

Combining (2)–(4), we have reduced the vacuum Einstein equations to a set of 10 second order quasilinear partial differential equations on the 4-dimensional domain Ω, with 10 unknowns, namely the components g_{ij} of the metric tensor. Because of all the summations in the equations above, the partial differential equations involve many terms.

9. Gauge freedom and coordinates

We have just shown that if g is a metric on M and the functions $g_{ij} : \Omega \to \mathbb{R}$ are its components with respect to some coordinate system $M \to \Omega$, then g is Ricci-flat if and only if the symmetric matrixfield (g_{ij}) satisfies the system of 10 PDEs just discussed. The difficulty with this approach is that, given a metric, there is no natural way to associate a system of functions, but rather there are many different ways, one for each choice of coordinate system.

For a different coordinate system $M \to \Omega'$, we get a different set of components $g'_{ij} : \Omega' \to \mathbb{R}$. The two sets of components are simply related by the usual formula for change of coordinates in a tensor:

$$g_{ij}(x) = \frac{\partial \psi^k}{\partial x^i}(x)\frac{\partial \psi^l}{\partial x^j}(x)g'_{kl}(x'), \tag{5}$$

where $x' = \psi(x)$ is the composition $\Omega \to M \to \Omega'$, a diffeomorphism of two domains in $\mathbb{R}^4$. The g'_{ij} satisfy the vacuum Einstein PDEs if and only if the g_{ij} do (since each statement is true if and only if the metric g is Ricci-flat). This is the gauge freedom in terms of the PDEs: from one solution of the PDEs you can generate many others by applying the transformation (5) for an arbitrary diffeomorphism ψ. Note that for a specific metric we obtain a specific set of component functions by picking a particular coordinate system. In this sense, fixing the gauge comes down to choosing coordinates.

Let us define two symmetric matrixfields (g_{ij}) on Ω and (g'_{ij}) on Ω' as equivalent if there is a diffeomorphism $\psi : \Omega \to \Omega'$ such that (5) holds. Then, to any given metric on a manifold (always admitting a global coordinate chart) we can associate an equivalence class of functions. If two metrics on two manifolds are equivalent in the sense discussed earlier (isometry), then they are associated to the same equivalence class of matrixfields. In short, there is a natural equivalence between equivalence classes of metrics and equivalence classes of symmetric matrixfields. However, there is no natural equivalence between specific metrics in the former equivalence class and specific matrixfields in the latter. Most numerical approaches calculate a single matrixfield, and so involve some gauge fixing. This is clearly one of the sources of difficulty in computations.