BATTELLE RENCONTRES

1967 LECTURES IN MATHEMATICS AND PHYSICS

EDITED BY

Cecile M. DeWitt
University of North Carolina at Chapel Hill

AND

John A. Wheeler
Princeton University

3 THE ABSTRACT INDEX NOTATION

When performing calculations in general relativity it is frequently necessary to operate with tensors of quite high valence.[10] Even such a basic quantity as the curvature tensor has valence four, and it possesses the familiar somewhat complicated symmetries. This makes it practically imperative that an index notation be employed for many calculations, so that the different connections between the quantities involved may be easily kept track of. It seems that there is a common feeling among mathematicians that such notations are to be avoided, presumably because of the connotation that their use entails explicit reference to a particular basis frame. However, when a physicist refers to "g_{ab}" or "$R^{a}{}_{bcd}$," I do not think that he usually means to be referring to a set of frame-dependent components but rather to a physical, frame-independent object which these components represent. But, the index notation allows a very convenient set of algebraic operations to be

[10] The term "valence" is used here in preference to "rank" because it is more descriptive and because the word "rank" has other connotations in the case of matrices.

applied to such objects, which produce new objects—these operations being actually completely frame-independent. The algebraic operations are, in essence, extremely simple, but they also allow great flexibility in the building up of more complicated operations out of simple ones. It would seem a great pity to forbid oneself the use of such a powerful and flexible notation merely because of some uneasy feelings about summation conventions and dependence on special basis frames. What I shall present here is an entirely frame-independent algebra which allows one to calculate with indexed quantities exactly as before (but now with a clear conscience!) and which, by use of a notational device, even permits a greater freedom than before, when it comes to introducing coordinate systems and basis frames (compare Schouten [96a]). The advantages will be particularly apparent when we come to consider spinors in the next section.

Let us not be completely formal, so we may be able to save some time and complication. I hope the essential ideas will be clear. Consider a vector space $\mathbf{V}^\bullet$ over a field $\mathbf{F}$—or, more generally, we allow $\mathbf{V}^\bullet$ to be a module,[11] where $\mathbf{F}$ is a ring of suitable type (for example, the elements of $\mathbf{V}^\bullet$ could be vector fields[12] and those of $\mathbf{F}$, C^∞ functions, on a manifold). The idea will be to construct what is essentially the usual tensor product of $\mathbf{V}^\bullet$ a number of times with its dual[13] $\mathbf{V}_\bullet$ a number of times, but where, by use of indices, we can keep track of the effect of symmetries and contractions easily. This is done by simply mirroring the usual index notation (with summation conventions, etc.) but where now the indices $a, b, c, \ldots$ are *not* to be thought of merely as generic symbols standing, say, for $0, 1, 2, \ldots, N$, but as *abstract labels.* We shall require an infinite supply

$$a, b, c, \ldots, a_0, b_0, \ldots, a_1, b_1, \ldots, a_2, \ldots \tag{3.1}$$

of abstract labels, so that expressions of arbitrary length can be built up. Let $\mathbf{L}$ denote the set of labels (3.1). For any element $\boldsymbol{\xi}$ of $\mathbf{V}^\bullet$ and any label $x \in \mathbf{L}$, we shall allow ourselves to write a symbol ξ^x. As $\boldsymbol{\xi}$ ranges over the elements of $\mathbf{V}^\bullet$, the associated object ξ^x ranges over a corresponding set $\mathbf{V}^x$. It should be emphasized here that ξ^x is an entity in its own right and *not* the set of components of $\boldsymbol{\xi}$ in some frame. Now, since we wish to mirror the usual tensor rules for indexed quantities, we are not permitted to write $\xi^a + \eta^b$, but $\xi^a + \eta^a$ and $\xi^b + \eta^b$ will be both allowable. (We must think of ξ^a and ξ^b as *different* objects.) For any $\lambda \in \mathbf{F}$, we shall also be permitted to write $\lambda\xi^a$. Thus, each of $\mathbf{V}^a, \mathbf{V}^b, \ldots, \mathbf{V}^{a_0}, \ldots$ is a vector space or module canonically isomorphic with $\mathbf{V}^\bullet$.

---

[11] A module differs from a vector space in that the scalars form a ring with identity rather than a field. A ring differs from a field in that division by nonzero elements is not always possible.

[12] Here a "field" means a cross-section of the appropriate vector bundle over $\mathscr{M}$.

[13] The space of all linear mappings of the module $\mathbf{V}^\bullet$ into the ring $\mathbf{F}$.

It may be felt that it is unnatural to introduce an infinite number of isomorphic spaces, when actually we only have one space. But we may view the situation in a slightly different way. Each element of $\mathbf{L}$ is really just a kind of organizational marker which keeps a tag on a particular vector (etc) irrespective of where it may occur in an expression. Thus, ξ^x is just a *pair*, consisting of $\boldsymbol{\xi}$ together with the marker x. That is to say, it is an element of $\mathbf{V}^\bullet \times \mathbf{L}$. We then have $\mathbf{V}^a = \mathbf{V}^\bullet \times (a)$, $\mathbf{V}^b = \mathbf{V}^\bullet \times (b)$, etc. The vector space or module axioms will, of course, apply to each $\mathbf{V}^x$:

$$\begin{aligned} \xi^x + (\eta^x + \zeta^x) &= (\xi^x + \eta^x) + \zeta^x \\ \lambda(\xi^x + \eta^x) &= \lambda\xi^x + \lambda\eta^x \\ (\lambda + \mu)\xi^x &= \lambda\xi^x + \mu\xi^x \\ \lambda(\mu\xi^x) &= (\lambda\mu)\xi^x \\ 1\xi^x &= \xi^x \\ 0\xi^x &= 0\eta^x \end{aligned} \tag{3.2}$$

Here $\lambda, \mu, 1, 0 \in \mathbf{F}$ with 1 and 0 being the multiplicative and additive identities, respectively. We also have $\xi^x + \eta^x = \eta^x + \xi^x$ (expanding $(1+1)(\xi^x + \eta^x)$) and $\xi^x + (-\xi^x) = 0$ (writing $-\xi^x$ for $(-1)\xi^x$ and 0 for $0\eta^x$).

The *dual* space $\mathbf{V}_\bullet$ will also have an infinite number of canonically isomorphic copies: $\mathbf{V}_a, \mathbf{V}_b, \ldots, \mathbf{V}_{a_0}, \ldots$. We may think of $\mathbf{V}_x$ as being the dual space of $\mathbf{V}^x$ for each $x \in \mathbf{L}$. The elements of $\mathbf{V}_x$ are linear mappings of $\mathbf{V}^x$ into $\mathbf{F}$. Thus, for $\theta_x \in \mathbf{V}_x$ we have

$$\theta_x(\xi^x + \eta^x) = \theta_x\xi^x + \theta_x\eta^x \tag{3.3}$$

$$\theta_x(\lambda\xi^x) = \lambda(\theta_x\xi^x) \tag{3.4}$$

where the effect of the mapping θ_x on ξ^x is written simply $\theta_x\xi^x$. We shall also allow this to be written in the reverse order: $\theta_x\xi^x = \xi^x\theta_x$. We require

$$\theta_a\xi^a = \theta_b\xi^b = \cdots = \theta_x\xi^x = \cdots \tag{3.5}$$

Now each of $\mathbf{V}_a, \mathbf{V}_b, \ldots$ will be vector space or module where $\lambda\theta_x$ and $\theta_x + \phi_x$ are defined by

$$(\lambda\theta_x)\xi^x = \lambda(\theta_x\xi^x) \qquad (\theta_x + \phi_x)\xi^x = \theta_x\xi^x + \phi_x\xi^x \tag{3.6}$$

The idea will be to use the elements of $\mathbf{F}, \mathbf{V}^a, \mathbf{V}^b, \ldots, \mathbf{V}_a, \mathbf{V}_b, \ldots$ to generate our algebra. To see what the rules of this algebra should be, we must recall what the rules are for the ordinary tensor index notation. We note, for instance, that products such as $\xi^a\eta^b$ will be permitted, whereas $\xi^a\eta^a$ will not. Furthermore, the allowable products must be *commutative*:

$$\xi^a\eta^b = \eta^b\xi^a \tag{3.7}$$

but in general $\xi^a \eta^b \neq \eta^a \xi^b$. The requirement (3.7) shows us that while *in essence* $\xi^a \eta^b$ is just the tensor product of elements $\xi^a \otimes \eta^b$, we cannot simply *identify* $\xi^a \eta^b$ with $\xi^a \otimes \eta^b$. For tensor products, according to the strict technical definition, are not commutative. Here we are allowed to define a commutative version of a tensor product, essentially for the reason that $\xi^a \eta^a$ is not defined. In a product $\xi^a \eta^b$, it is the *labels* a and b that tell us which factor is which, not the ordering of the factors. One method (suggested to me by S. Mac Lane) of precisely defining the type of product used here is to take the *symmetric algebra* [54a] on the direct sum $\mathbf{V}^a \oplus \mathbf{V}_a \oplus \mathbf{V}^b \oplus \mathbf{V}_b \oplus \cdots$ and then, for each pair of disjoint (finite) sets of elements of $\mathbf{L}$, say a, p, r and b, m, we select the corresponding subspace $\mathbf{V}^{apr}_{bm}$ spanned by the elements of the form

$$\xi^a \eta^p \zeta^h \theta_b \phi_m \tag{3.8}$$

The general element of $\mathbf{V}^{abr}_{bm}$ will be a linear combination of expressions like (3.8). The construction so made, ensures that each product (3.8) is fully commutative and also that the various distributive laws hold (for example, $\psi^{apr}_m(\theta_b + \chi_b) = \psi^{apr}_m \theta_b + \psi^{apr}_m \chi_b$ with $\psi^{apr}_m \in \mathbf{V}^{apr}_m$). There is no significance in the ordering of a, p, r in $\mathbf{V}^{apr}_{bm}$ or of b, m. Thus $\mathbf{V}^{apr}_{bm} = \mathbf{V}^{par}_{bm} = \mathbf{V}^{arp}_{mb}$, etc. However, the ordering of the indices for an element ρ^{apr}_{bm} *is* significant.

Every element $\rho^{apr}_{bm} \in \mathbf{V}^{apr}_{bm}$ is a linear combination of commutative products of the type (3.8):

$$\rho^{apr}_{bm} = \sum_{i=1}^{M} \overset{(i)}{\lambda}\, \overset{(i)}{\xi^a}\, \overset{(i)}{\eta^p}\, \overset{(i)}{\zeta^r}\, \overset{(i)}{\theta_b}\, \overset{(i)}{\phi_m} \tag{3.9}$$

but there will be many ways of expressing ρ^{apr}_{bm} as such. A convenient criterion for the equality of two expressions (3.9) for the modules that we are interested in is that for *every* choice $\alpha_a \in \mathbf{V}_a$, $\beta_p \in \mathbf{V}_p$, $\gamma_r \in \mathbf{V}_r$, $\sigma^b \in \mathbf{V}^b$, $\tau^m \in \mathbf{V}^m$, the scalar

$$\rho^{apr}_{bm} \alpha_a \beta_p \gamma_r \sigma^b \tau^m = \sum_{i=1}^{M} \overset{(i)}{\lambda} (\overset{(i)}{\xi^a} \alpha_a)(\overset{(i)}{\eta^p} \beta_p)(\overset{(i)}{\zeta^r} \gamma_r)(\overset{(i)}{\theta_b} \sigma^b)(\overset{(i)}{\phi_m} \tau^m) \tag{3.10}$$

should be the same for both expressions. From this all the algebraic properties will follow. (Note, that in general $\rho^{arp}_{bm} \neq \rho^{par}_{bm} \neq \rho^{apr}_{mb}$, etc.)

The entire tensor system $\{\mathbf{V}\}$ consists of all of the $\mathbf{V}^{x \ldots z}_{u \ldots w}$, including $\mathbf{V} = \mathbf{F}$:

$$\{\mathbf{V}\} = (\mathbf{V}, \mathbf{V}^a, \mathbf{V}^b, \ldots, \mathbf{V}_a, \mathbf{V}_b, \ldots, \mathbf{V}^{ab}, \ldots, \mathbf{V}^{x \ldots z}_{u \ldots w}, \ldots)$$

There are four basic operations on $\{\mathbf{V}\}$, namely

ADDITION:	$\mathbf{V}^{x \ldots z}_{u \ldots w} \times \mathbf{V}^{x \ldots z}_{u \ldots w} \to \mathbf{V}^{x \ldots z}_{u \ldots w}$	(3.11)
MULTIPLICATION:	$\mathbf{V}^{a \ldots d}_{p \ldots r} \times \mathbf{V}^{x \ldots z}_{u \ldots w} \to \mathbf{V}^{a \ldots dx \ldots z}_{p \ldots ru \ldots w}$	(3.12)
INDEX SUBSTITUTION:	$\mathbf{V}^{x \ldots z}_{u \ldots w} \to \mathbf{V}^{f \ldots h}_{k \ldots m}$	(3.13)
(a, b)-CONTRACTION:	$\mathbf{V}^{ax \ldots z}_{bu \ldots w} \to \mathbf{V}^{x \ldots z}_{u \ldots w}$	(3.14)

In (3.11), (3.12), and (3.14), the differently denoted index letters appearing are all assumed to be different elements of $\mathbf{L}$. In (3.13), the elements $x, \ldots, z, u, \ldots, w$ of $\mathbf{L}$ are all distinct and so are $f, \ldots, h, k, \ldots, m$. Otherwise they are unrestricted except that $x, \ldots, z$ and $f, \ldots, h$ are equal in number and that $u, \ldots, w$ and $k, \ldots, m$ are equal in number. Addition and multiplication are defined the obvious way. Index substitution is induced simply by a permutation applied to $\mathbf{L}$. (The validity of any equation is unaffected by a permutation of the elements of $\mathbf{L}$.) To define contraction, we consider, for example, the (p, b)-contraction: $\mathbf{V}^{par}_{bm} \to \mathbf{V}^{ar}_m$ as applied to the element $\rho^{apr}_{bm} \in \mathbf{V}^{par}_{bm}$ given by (3.9). The result is

$$\rho^{axr}_{xm} = \sum_{i=1}^{M} \overset{(i)}{\lambda} (\overset{(i)}{\eta^x} \overset{(i)}{\theta_x}) \overset{(i)}{\xi^a} \overset{(i)}{\zeta^r} \overset{(i)}{\phi_m} \tag{3.15}$$

We have $\rho^{axr}_{xm} \in \mathbf{V}^{ar}_m$, so the x labels are "dummies" which do not contribute to the total valence type.

It may be verified algebraically that all the usual tensor rules[14] follow from the above constructions. Thus, addition gives an Abelian group structure to each $\mathbf{V}^{x \ldots z}_{u \ldots w}$. Multiplication is commutative and distributive over addition. Contraction appropriately commutes with addition, with multiplication, and with other contractions. The contraction of a zero element is again a zero element. (If we use equality between the scalars (3.10) as the definition of equality between formal expressions (3.9), then this last property follows from the property that any matrix over $\mathbf{F}$ whose square vanishes also has vanishing trace. This property holds in the cases which interest us here, but would not be true for certain rings of finite characteristic.)

So far, the question of a basis frame for $\mathbf{V}^\bullet$ has not even arisen. However, it is often convenient to work with basis frames and we shall need a notation to be able to distinguish basis indices from the abstract labels. I shall adopt the convention that *German* indices $\mathfrak{a}, \mathfrak{b}, \ldots, \mathfrak{a}_0, \ldots$ will denote the numbering of basis elements in the standard way, that is, each of $\mathfrak{a}, \mathfrak{b}, \ldots$ denotes one of the integers $0, 1, \ldots, N$ [for an $(N+1)$-dimensional space]. The use of German indices will be to remind us of two things; first that a choice of a (possibly arbitrary) basis frame is involved in any expression containing such indices, with a consequent loss of covariance; and secondly that the Einstein *summation convention* is being used whenever repeated indices occur in a term in an expression. Now, let $\boldsymbol{\delta}_0, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_N \in \mathbf{V}^\bullet$ be a basis for $\mathbf{V}^\bullet$ (assuming finite-dimensionality) and let $\boldsymbol{\delta}^0, \boldsymbol{\delta}^1, \ldots, \boldsymbol{\delta}^N \in \mathbf{V}_\bullet$ be the corresponding dual basis. For $x \in \mathbf{L}$ we have canonical images in $\mathbf{V}^x$ and $\mathbf{V}_x$:

$$\delta^x_0 \ldots, \delta^x_N \in \mathbf{V}^x, \qquad \delta^0_x, \ldots, \delta^N_x \in \mathbf{V}_x \tag{3.16}$$

[14] See any standard work on classical tensor calculus, for example, [100].

We may use the generic symbols $\delta^x_{\mathfrak{x}} \in \mathbf{V}^x$, $\delta^{\mathfrak{x}}_x \in \mathbf{V}_x$, so the basis orthogonality relation takes the form

$$\delta^x_{\mathfrak{y}}\delta^{\mathfrak{x}}_x = \delta^{\mathfrak{x}}_{\mathfrak{y}} \tag{3.17}$$

where $\delta^{\mathfrak{x}}_{\mathfrak{y}}$ is the ordinary Kronecker delta symbol. (No relation between x and $\mathfrak{x}$ is to be implied by the notation.) We can also define an element δ^x_y of $\mathbf{V}^x_y$ by

$$\delta^{\mathfrak{x}}_y\delta^x_{\mathfrak{x}} = \delta^x_y \tag{3.18}$$

The quantity δ^x_y satisfies the usual properties

$$\xi^{\cdots y \cdots}_{\cdots\;\cdots}\delta^x_y = \xi^{\cdots x \cdots}_{\cdots\;\cdots} \qquad \eta^{\cdots\;\cdots}_{\cdots x \cdots}\delta^x_y = \eta^{\cdots\;\cdots}_{\cdots y \cdots} \tag{3.19}$$

and is clearly actually independent of the choice of basis, by virtue of this. The *components*, with respect to the basis, of any element $\zeta^{x\cdots z}_{u\cdots w} \in \mathbf{V}^{x\cdots z}_{u\cdots w}$ are given by

$$\zeta^{\mathfrak{x}\cdots\mathfrak{z}}_{\mathfrak{u}\cdots\mathfrak{w}} = \zeta^{x\cdots z}_{u\cdots w}\delta^u_{\mathfrak{u}}\cdots\delta^w_{\mathfrak{w}}\delta^{\mathfrak{x}}_x\cdots\delta^{\mathfrak{z}}_z \tag{3.20}$$

Conversely, to express $\zeta^{x\cdots z}_{u\cdots w}$ in terms of its components, we simply write

$$\zeta^{x\cdots z}_{\mathfrak{u}\cdots\mathfrak{w}} = \zeta^{\mathfrak{x}\cdots\mathfrak{z}}_{\mathfrak{u}\cdots\mathfrak{w}}\delta^{\mathfrak{u}}_u\cdots\delta^{\mathfrak{w}}_w\delta^x_{\mathfrak{x}}\cdots\delta^z_{\mathfrak{z}} \tag{3.21}$$

One aspect of this notation which is an advantage in certain contexts is that we may convert some indices into component form and leave others as abstract labels:

$$\rho^{a\mathfrak{p}r}_{\mathfrak{b}m} = \rho^{apr}_{bm}\delta^{\mathfrak{p}}_p\delta^b_{\mathfrak{b}} \in \mathbf{V}^{ar}_m \tag{3.22}$$

Generally, all *algebraic* relations will be unaffected by whether or not an index is in German type. But when we consider covariant derivatives in the next section, we shall see that an important *formal* difference arises between the ways the two types of index are treated—in addition to the present merely conceptual difference.

An elementary, but important, property of the type of algebraic structure that we have built up here, is that we can sometimes embed one such structure in another by the device of *grouping together indices*. Thus, we may consider a new labeling set $\hat{\mathbf{L}}$, say, whose elements are (disjoint) subsets of elements of $\mathbf{L}$. For example, we could put $\alpha = abc$, $\beta = def$, $\gamma = ghi$, etc., where $\mathbf{L}$ is divided exhaustively into disjoint triplets, these triplets being the elements of $\hat{\mathbf{L}}$. It is clear that, in this case, the sets $\mathbf{V}^\alpha = \mathbf{V}^{abc}$, $\mathbf{V}^\beta = \mathbf{V}^{def}, \ldots, \mathbf{V}_\alpha = \mathbf{V}_{abc}, \ldots, \mathbf{V}^{\alpha\beta}_\gamma = \mathbf{V}^{abcdef}_{ghi}, \ldots$ will satisfy just the same rules as before. More complicated groupings are also possible.

We may also consider systems where the labeling set contains elements of different "types." The only modifications of our scheme would then come in index substitution (where only labels of the *same* type may be substituted for one another) and in contraction, where only an upper and

lower indices of the *same* type may be contracted together. When we consider spinors in the next section we shall see an example of this kind of system. We shall have labels of two different types (related to each other via an operation of complex conjugation).