

CHAPTER II

TENSOR THEORY IN FOUR-SPACE

In this chapter, we will provide the sufficient details of tensor theory in four-dimensional space, or, for brevity, called four-space which will be useful for our further studies. However, the generalization of the theory for n -dimensional space can be treated in the same manner. The concepts of tensor principles, which are presented in this chapter, are collected from some textbooks on tensor theory and special theory of relativity, for examples, Sokolnikoff(1964), Borisenko(1968), Jackson(1975), and Rindler(1991).

Form Invariance of Tensor Equations

Tensor analysis is concerned with a study of abstract objects called *tensors*, whose properties are independent of the reference frames used to describe the object. A tensor is represented in a particular reference frame by a set of functions, termed its *components*. Whether a given set of transformations represents a tensor depends on the law of transformation of these functions from one coordinate system to another. The *permissible* transformations for tensors should be non-singular; i.e., that the equations which express the $X^{V'}$ in terms of the X^V can be solved uniquely for the X^V in terms of the $X^{V'}$, and should be differentiable as often as may be required.

If the natural laws are valid in wide class of reference systems, they suggested to formulate in the form of *tensor equations* which are invariant with respect to a given category of coordinate transformations. For clarity at the beginning, we consider the following explanation. Let

$$F(\Phi, A^V, B^{V\mu}, \dots) = 0 \quad (2.1)$$

be an equation of tensors (of various ranks) in four-dimensional space. If the equation F is transformed into another system to be

$$G(\Phi', A^{\nu'}, B^{\nu'\mu'}, \dots) = 0. \quad (2.2)$$

In general, Eqs.(2.1) and (2.2) are not necessary of the same form; i.e., $F \neq G$. But if F contains the same form shown as:

$$F(\Phi', A^{\nu'}, B^{\nu'\mu'}, \dots) = 0. \quad (2.3)$$

it is said to be invariant under the given transformation.

Covariant and Contravariant Transformations

We have known that tensor equations must be invariant under some kinds of permissible transformations called *tensor transformations*. The tensor transformations themselves can be generally demonstrated in two forms named *covariant* and *contravariant* transformations. Their concepts in three-dimensional space can be shown as follows.

Let us first begin with vector, the first-rank tensor in three-space. The concept of vector in Cartesian coordinates may not be complicated because its always linear and orthogonal. But vectors are not always this simple especially when expressed in non-Cartesian coordinates systems such as linear (but not orthogonal), curvilinear (orthogonal), and curvilinear (but not orthogonal). In order to reduce the concepts of covariance and contravariance to their simplest terms, consider three-dimensional linear coordinates, which have three noncoplanar vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, which are in general neither orthogonal nor of unit length, be their basis vectors. Then vector \mathbf{A} can be described so that

$$\mathbf{A} = A^1\mathbf{e}_1 + A^2\mathbf{e}_2 + A^3\mathbf{e}_3 \quad (2.4)$$

To define a scalar product in a nonorthogonal coordinate systems, it is convenience to define what is called a *reciprocal lattice system*, which has three basis vectors, $\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3$ which satisfy the conditions

$$\mathbf{e}^1 = \frac{\mathbf{e}_2 \times \mathbf{e}_3}{\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)}, \quad \mathbf{e}^2 = \frac{\mathbf{e}_3 \times \mathbf{e}_1}{\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)}, \quad \mathbf{e}^3 = \frac{\mathbf{e}_1 \times \mathbf{e}_2}{\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)} \quad (2.5)$$

Then we can say that two bases $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and $\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3$ are said to be *reciprocal* if they exhibit as:

$$\mathbf{e}^i \cdot \mathbf{e}_k = \begin{cases} 0, & \text{if } i \neq k \\ 1 & \text{if } i = k \end{cases} \quad (2.6)$$

or $\mathbf{e}^i \cdot \mathbf{e}_k = \delta_k^i$ where δ_k^i is called the *kroncker delta function* whose values are defined as above.

Now vector \mathbf{A} can be expanded in the reciprocal lattice as

$$\mathbf{A} = A_1\mathbf{e}^1 + A_2\mathbf{e}^2 + A_3\mathbf{e}^3. \quad (2.7)$$

The numbers A^i are called the *contravariant* components of \mathbf{A} , while the numbers A_i are called the *covariant* components of \mathbf{A} . We also notice that

$$\mathbf{A} \cdot \mathbf{e}^i = \sum_k A^k \mathbf{e}_k \cdot \mathbf{e}^i = A^i \mathbf{e}_i \cdot \mathbf{e}^i = A^i \quad (i=1,2,3), \quad (2.8)$$

and it can be proved similarly for $\mathbf{A} \cdot \mathbf{e}_i = A_i$. Thus Eqs.(2.4) and (2.7) can be rewritten as

$$\mathbf{A} = (\mathbf{A} \cdot \mathbf{e}^1)\mathbf{e}_1 + (\mathbf{A} \cdot \mathbf{e}^2)\mathbf{e}_2 + (\mathbf{A} \cdot \mathbf{e}^3)\mathbf{e}_3, \quad (2.9a)$$

and

$$\mathbf{A} = (\mathbf{A} \cdot \mathbf{e}_1)\mathbf{e}^1 + (\mathbf{A} \cdot \mathbf{e}_2)\mathbf{e}^2 + (\mathbf{A} \cdot \mathbf{e}_3)\mathbf{e}^3, \quad (2.9b)$$

respectively. The norm of \mathbf{A} can be shown as

$$|\mathbf{A}| = (\mathbf{A} \cdot \mathbf{A})^{1/2} = \sum_{i,k} (A^i \mathbf{e}_i \cdot A_k \mathbf{e}^k) = (\sum_i A^i A_i)^{1/2}. \quad (2.10)$$

Hence, the scalar product of two vectors \mathbf{A} and \mathbf{B} are readily defined as follows:

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i,k} (A^i \mathbf{e}_i \cdot B_k \mathbf{e}^k) = \sum_i A^i B_i \quad (2.11a)$$

$$= \sum_{i,k} (A_i \mathbf{e}^i \cdot B^k \mathbf{e}_k) = \sum_i A_i B^i. \quad (2.11b)$$

Note that the Cartesian system has three base vectors \mathbf{i} , \mathbf{j} , \mathbf{k} , all are mutually orthogonal to each other then, according to Eq.(2.6), its reciprocal system is the same as the inertial. As a result, the covariant and contravariant components of vectors in Cartesian coordinate system are *undistinguishable* so it does not matter whether the superscript or subscript is used.

The relation between covariant and contravariant components can be determined from the Eq.(2.8),

$$A^i = \mathbf{A} \cdot \mathbf{e}^i = \sum_k A_k (\mathbf{e}^k \cdot \mathbf{e}^i), \quad (2.12a)$$

$$A_i = \mathbf{A} \cdot \mathbf{e}_i = \sum_k A^k (\mathbf{e}_k \cdot \mathbf{e}_i). \quad (2.12b)$$

Then introducing the notations

$$\mathbf{e}_i \cdot \mathbf{e}_k = g_{ik}, \quad \mathbf{e}^i \cdot \mathbf{e}^k = g^{ik}, \quad \mathbf{e}^i \cdot \mathbf{e}_k = g^i_k = \delta^i_k, \quad (2.13)$$

therefore Eq.(2.12) becomes

$$A_i = \mathbf{A} \cdot \mathbf{e}_i = \sum_k A^k (\mathbf{e}_k \cdot \mathbf{e}_i) = \sum_k g_{ik} A^k, \quad (2.14a)$$

$$A^i = \mathbf{A} \cdot \mathbf{e}^i = \sum_k A_k (\mathbf{e}^k \cdot \mathbf{e}^i) = \sum_k g^{ik} A_k. \quad (2.14b)$$

These designations of the components of a vector in Eqs.(2.4) and (2.7) lead to the law of transformations:

$$A_{i'} = \sum_k a^{k}_{i'} A_k, \quad (2.15a)$$

for covariant components and

$$A^{i'} = \sum_k a^{i'}_k A^k, \quad (2.15b)$$

for contravariant components where $a^{k}_{i'}$ and $a^{i'}_k$ denote $\mathbf{e}^k \cdot \mathbf{e}_{i'}$ and $\mathbf{e}^{i'} \cdot \mathbf{e}_k$ respectively. To prove Eq.(2.15a) we observe that

$$A_{i'} = \mathbf{A} \cdot \mathbf{e}_{i'} = \sum_k A_k (\mathbf{e}^k \cdot \mathbf{e}_{i'}) = \sum_k a^{k}_{i'} A_k,$$

which satisfies Eq.(2.15a) as wanted, and Eq.(2.15b) can be proved in the same manner. Notice that a point vector \mathbf{x} in three-space, or $\mathbf{x} = \mathbf{x}(x^1, x^2, x^3)$, will have components which are transformed as contravariant components,

$$dx^{i'} = \sum_k (\partial x^{i'} / \partial x^k) dx^k = \sum_k a^{i'}_k dx^k. \quad (2.16)$$

Tensor Theory in Four-space

After we have studied the concepts of the covariant and contravariant transformations in three-dimensional space, as presented in the previous section. Now, we will generalize these concepts to the four-dimensional space as follows.

Consider a set of four real variables $\{X^0, X^1, X^2, X^3\}$, which we may write as $\{X^v\}$ ($v = 0, 1, 2, 3$). It is useful, though not necessary, to regard these variables as coordinates in *four-dimensional space* V_4 , and to adopt a geometric language. Any non-singular transformation of the X s to a new set of variables $\{Y^v\}$ can then be regarded as a re-coordinatization of V_4 . We may associate with the space V_4 a system of tensors. Tensors in four-space are objects that can be described by their components, which form an ordered set of, say M , real numbers T_1, \dots, T_m . (We may think of familiar *three-vectors* of classical physics associated with Euclidean three-space.) These components generally change when the coordinates of V_4 are changed. In fact, we may define a tensor T as a *map* $T: S \rightarrow R_m$ from the set S of permissible coordinates system $\{X^v\}$ to the space R_m of real-number M -tuplets (T_1, \dots, T_m) . Like a three-vector, a tensor in four-space can be defined at *one point only* (*point tensor*), in which case its components are just numbers, or it can be defined on some larger subspace of V_4 (*field tensor*), in which case its components are functions of position in V_4 .

In four-dimensional space a tensor of *rank* r has $M=4^r$ components. In particular, a tensor of rank zero has one component T and is called a *scalar*. A tensor of rank one has 4 components (T_1, \dots, T_4) and is called a *four-vector*. A tensor of rank two has $4^2 = 16$ components, which can be exhibited in matrix form thus:

$$\begin{bmatrix} T_{00} & T_{01} & \dots & T_{03} \\ T_{10} & T_{11} & \dots & T_{13} \\ \dots & \dots & \dots & \dots \\ T_{30} & T_{31} & \dots & T_{33} \end{bmatrix}$$

Such a tensor is said to be *singular* if the determinant of this matrix vanishes. Tensors of higher ranks cannot be exhibited in such convenience forms, but tensors of all ranks are usually represented by a typical component, e.g. we may loosely speak of the tensor $T_{\nu\alpha\beta}$, the tensor of rank three, the tensor $T_{\nu\alpha\beta\gamma}$, the tensor of rank four, etc. A more logical notation would be $\{T_{\nu\alpha\beta}\}$ for the entire tensor and $T_{\nu\alpha\beta}$ for a single component; but this can become tedious and we shall not use it. In general, the order of the indices is significant, e.g. $T_{123} \neq T_{132}$, though specific tensors of all ranks may possess various symmetry properties such as $T_{\nu\alpha\beta} = T_{\nu\beta\alpha}$. The indices ν, α, β will always be understood to range from 0 to 3.

For reasons that we have shown above, we will use the subscripts for the covariant components and superscripts for contravariant components. Therefore, typical tensor components may look thus: $A_{\nu\alpha}$ (rank two), B^{ν}_{α} (rank two), $C^{\nu\alpha\beta}$ (rank three), etc. When we exhibit the components of a second-rank tensor ($T_{\nu\alpha}$, T^{ν}_{α} , $T^{\nu\alpha}$) as a metric, the first or upper index will always refer to the row and the another to the column. When the off-diagonal components vanish, we may write

$$T_{\nu\alpha} = \text{diag}(T_{00}, T_{11}, T_{22}, T_{33})$$

We shall find it convenient to use *Einstein's summation convention*, namely: if any index appears twice in a given term, once as a subscript and once as a superscript, a summation over the range of that index is implied. Thus, for example, we write $A^{\nu}A_{\nu} \equiv \sum_{\nu} A^{\nu}A_{\nu}$ or $A_{\nu\alpha\beta}B^{\alpha\beta} \equiv \sum_{\alpha, \beta} A_{\nu\alpha\beta}B^{\alpha\beta}$, etc. The repeated indices signaling summation are called *dummy indices* while a non-repeated index

is called *free index*. An obvious but important principle is that a dummy index pair can be replaced by any other: e.g. $A^\nu A_\nu = A_\beta A^\beta$. Such a replacement is often necessary to avoid the triple occurrence of an index which would lead to ambiguities.

Note that from now on, if not state otherwise, we will use Greek indices $\mu, \nu, \alpha, \beta, \dots$ for the range 0,1,2,3, while on occasion we shall use Latin indices i, j, k, \dots for the range 1,2,3. The four-tensors will be denoted by capital letters ($A, B^\alpha, C^{\alpha\beta}$, etc.) and three-tensors by lowercase letters (a, b_i, c_{ij} , etc.)

Coordinate Transformations

To distinguish between various coordinate systems for our space V_4 , we shall use primed and multiply primed indices. All range from 0 to 3. Thus :

$$\mu, \nu, \alpha, \dots; \mu', \nu', \alpha', \dots; \mu'', \nu'', \alpha'', \dots; \dots = 0, 1, 2, 3.$$

No special relation is implied between, say, μ and μ' : they are as independent as μ and ν' . A first system of coordinates can then be denoted by $\{X^\nu\} = \{X^0, X^1, X^2, X^3\}$, a second by $\{X^{\nu'}\} = \{X^{0'}, X^{1'}, X^{2'}, X^{3'}\}$, etc. Similarly the components of a given tensor in different coordinate systems are distinguished by the primes on their indices. Thus, for example, the components of some third-rank tensor may be denoted by $A_{\nu\alpha\beta}$ in the $\{X^\nu\}$ system, by $A_{\nu'\alpha'\beta'}$ in the $\{X^{\nu'}\}$ system, etc. When primed indices take particular numerical values, we can prime these, so as not to lose sight of the relevant coordinate system. Thus, for example, when $\nu'=1, \alpha'=2, \beta'=3$, $A_{\nu'\alpha'\beta'}$ becomes $A_{1'2'3'}$. This will already have been noted for the case of the coordinates above. (However, sometimes we will adopt the simpler device of priming the kernel as A'_{123} .)

When we make a coordinate transformation from one set of coordinates X^{ν} to another $X^{\nu'}$ (we often drop the brace), it will be assumed that the transformation is non-singular, i.e. that the equations which express the $X^{\nu'}$ in term of the X^{ν} can be solved uniquely for the X^{ν} in term of the $X^{\nu'}$. We also assume that the functions specifying a transformation are differentiable as often as may be required. For convenience, we write, similarly to Eq.(2.16),

$$\frac{\partial X^{\nu'}}{\partial X^{\nu}} = a^{\nu'}_{\nu}, \quad \frac{\partial X^{\nu}}{\partial X^{\nu'}} = a^{\nu}_{\nu'}, \quad \frac{\partial^2 X^{\nu'}}{\partial X^{\nu} \partial X^{\alpha}} = a^{\nu'}_{\nu\alpha}, \quad (2.17)$$

and use a similar notation for other such derivatives. We observe that, by the chain rule of differentiation,

$$a^{\nu}_{\nu'} a^{\nu'}_{\nu''} = a^{\nu}_{\nu''}, \quad a^{\nu}_{\nu'} a^{\nu'}_{\alpha} = \delta^{\nu}_{\alpha}, \quad (2.18)$$

where δ^{ν}_{α} (the *Kronecker delta*) equals 1 or 0 according as $\nu = \alpha$ or $\nu \neq \alpha$. It is important to note the 'index-substitution' action of δ^{ν}_{α} exemplified by

$$A_{\nu\alpha\beta} \delta^{\nu}_{\chi} = A_{\chi\alpha\beta}.$$

Informal Definition of Tensors

We are now ready to give an informal definition of tensors, which allow us to recognize a tensor when we see one. A mathematically formal definition of tensors is given as follows:

An object which having components $A^{\nu\alpha\dots\beta}$ in the X^{ν} system of coordinates and $A^{\nu'\alpha'\dots\beta'}$ in the $X^{\nu'}$ system is said to behave as a *contravariant* tensor under the transformation $\{X^{\nu}\} \rightarrow \{X^{\nu'}\}$ if

$$A^{v'\alpha'\dots\beta'} = a^{v'\alpha'\dots\beta'}_{v\alpha\dots\beta} A^{v\alpha\dots\beta} \quad (2.19)$$

Similarly, $A_{v\alpha\dots\beta}$ is said to behave as a *covariant* tensor under $\{X^v\} \rightarrow \{X^{v'}\}$ if

$$A_{v'\alpha'\dots\beta'} = a_{v'\alpha'\dots\beta'}^{v\alpha\dots\beta} A_{v\alpha\dots\beta} \quad (2.20)$$

Lastly, $A^{\mu\dots\nu}_{\alpha\dots\beta}$ is said to behave as a *mixed* tensor (contravariant in $\mu\dots\nu$ and covariant in $\alpha\dots\beta$) under $\{X^v\} \rightarrow \{X^{v'}\}$ if

$$A^{\mu'\dots\nu'}_{\alpha'\dots\beta'} = a^{\mu'\dots\nu'}_{\mu\dots\nu} a_{\alpha'\dots\beta'}^{\alpha\dots\beta} A^{\mu\dots\nu}_{\alpha\dots\beta} \quad (2.21)$$

Note that Eq.(2.21) evidently subsumes both Eqs.(2.19) and (2.20) as special cases.

At a given point in four-space V_4 the a s are pure numbers. Thus the tensor transformation Eqs.(2.19)-(2.21) are linear: the components in the new coordinate system are linear functions of the components in the old system, the coefficients being the product of the a s. Contravariant tensors involve derivatives of the of the new coordinates $X^{v'}$ with respect to the old, X^v , covariant tensors involve the derivatives of the old coordinates with respect to the new, and mixed tensors involve both type of derivatives. The convention of using subscripts for covariance superscripts for contravariance, together with the requirement that the free indices on both sides of the equations must balance, serve as a perfect mnemonic for reproducing Eqs.(2.19)-(2.21).

If we say that an object is a tensor it is understood that the object *behaves* as a tensor under *all* non-singular differentiable transformations of the coordinates of V_4 . An object which behaves as a tensor only under a certain subgroup of non-singular differentiable coordinate transformations, like the Lorentz transformations

Eq.(1.22), may be called a *qualified* tensor, and its name should be qualified by an adjective recalling the subgroup in question, as in 'Lorentz tensor', more commonly called 'four-tensor'. These tensors are, as a matter of fact, the (qualified) tensors used in special relativity. But we shall occasionally lapse from this strict terminology by omitting the adjective 'qualified' when no confusion seems likely.

The above definitions, when applied to a tensor of rank zero (a scalar) imply $A' = A$ (no *as!*), whence a scalar is a function of position in V_4 only, i.e. it is independent of the coordinate system. A scalar is therefore often called an *invariant*. The *zero tensor* of any type $A^{\mu\dots\nu}_{\alpha\dots\beta}$ is defined as having all its components zero in all coordinate systems. It is clear from Eq.(2.21) that it is tensor. For brevity it is usually written as 0, with the indices omitted.

Evidently we must call two tensors *equal* if they constitute the same map $S \rightarrow R_m$, in other words if they have the same components in all coordinate systems. Now the main theorem of the tensor calculus—trivial in its proof, profound in its implications—is this; if two tensors of the same *valence* [we shall say that a tensor has valence (s,t) if its components have s contravariant and t covariant indices) have equal components in any one coordinate system then they are equal. This is an immediate consequence of the definition Eq.(2.21). It implies that tensor-(component) equations always express physical or geometric facts, i.e. facts transcending the coordinate system used to describe them.

The simplest example of a *contravariant* vector is provided by the differentials of the coordinates, dX^ν . For,

$$dX^{\nu'} = (\partial X^{\nu'} / \partial X^\nu) dX^\nu = a^{\nu'}_\nu dX^\nu. \quad (2.22)$$

Under *linear* transformations $X^{\nu'} = A^{\nu'}_{\nu} X^{\nu} + B^{\nu'}$ ($A^{\nu'}_{\nu}, B^{\nu'} = \text{constant}$) the coordinate differences ΔX^{ν} transform like the differentials dX^{ν} and thus constitute a 'qualified' vector—usually called the *displacement vector*. Because of this, the displacement vector can then serve to represent any contravariant vector. (Recall the 'direct line segments' of elementary vector analysis!) The coordinates X^{ν} themselves behave as vectors only under linear homogeneous transformations ($B^{\nu'} = 0$). A case in point is the homogeneous Lorentz transformation group.

The simplest example of a *covariant* vector is provided by the gradient of a function of position $\Phi = \Phi(X^0, X^1, X^2, X^3)$. For, if we write

$$\partial_{\nu} \Phi = \partial \Phi / \partial X^{\nu},$$

we have

$$\partial_{\nu'} \Phi = \partial \Phi / \partial X^{\nu'} = (\partial \Phi / \partial X^{\nu}) (\partial X^{\nu} / \partial X^{\nu'}) = a^{\nu}_{\nu'} \partial_{\nu} \Phi \quad (2.23)$$

An important example of a mixed second-rank tensor is provided by the Kronecker delta introduced after Eq.(2.18). For, by use of an analogue of Eq.(2.18) we have

$$a^{\nu'}_{\nu} a^{\alpha}_{\alpha'} \delta^{\nu}_{\alpha} = a^{\nu'}_{\alpha} a^{\alpha}_{\alpha'} = \delta^{\nu'}_{\alpha'}. \quad (2.24)$$

The Group Properties of Tensors

It follows readily from their definitions that the tensor component transformation in Eq.(2.21) satisfy the two *group properties* of *symmetry* and *transitivity*. In other words, if an object behaves as a tensor under $\{X^{\nu}\} \rightarrow \{X^{\nu'}\}$ then it also behaves so under $\{X^{\nu'}\} \rightarrow \{X^{\nu}\}$ (symmetry); and if an object behaves as a tensor under $\{X^{\nu}\} \rightarrow \{X^{\nu'}\}$ and under $\{X^{\nu'}\} \rightarrow \{X^{\nu''}\}$ then it also behaves so under $\{X^{\nu}\} \rightarrow \{X^{\nu''}\}$ (transitivity). The general method to proof will be sufficiently indicated

by dealing with an object of type A^μ_ν . Note the use of substitute dummies α and β in the first of the following equations to avoid the triple occurrences of μ and ν , and the use of Eq.(2.18)(or analogous of it):

$$a^\mu_{\mu'} \nu'_{\nu} A^{\mu'}_{\nu'} = a^\mu_{\mu'} \nu'_{\nu} (a^{\mu'}_{\alpha} \beta_{\nu'} A^{\alpha}_{\beta}) = \delta^\mu_{\alpha} \delta^{\beta}_{\nu} A^{\alpha}_{\beta} = A^\mu_{\nu};$$

$$A^{\mu''}_{\nu''} = a^{\mu''}_{\mu'} \nu'_{\nu''} A^{\mu'}_{\nu'} = a^{\mu''}_{\mu'} \nu'_{\nu''} (a^{\mu'}_{\mu} \nu_{\nu'} A^{\mu}_{\nu}) = a^{\mu''}_{\mu} \nu_{\nu''} A^{\mu}_{\nu}.$$

As a result of these group properties, we can construct a tensor by specifying its components arbitrarily in *one* coordinate system, say $\{X^\nu\}$, and then using the transformation law Eq.(2.21) to define its components in all other systems, or, in a case of qualified tensor, in all those systems which are mutually connected by transformations belonging to the chosen subgroup. The group properties then ensure that any two sets of components will be related tensorially. For if A^μ , say, is related tensorially to $A^{\mu'}$ and to $A^{\mu''}$, then $A^{\mu'}$ is so related to A^μ (by symmetry) and consequently to $A^{\mu''}$ (by transitivity).

The group properties also allow us to make precise and formal our preliminary definitions of tensors given in the last two sections. Note that the permissible coordinate system $\{X^\nu\}$ on V_4 form an equivalent class, two such systems being equivalent if the transformation from the one to the other is non-singular, i.e. invertible. (In the case of a qualified tensor under a specific transformation group, any two coordinate systems related by a member of that group are to be regarded as equivalent.) Now, because of the group properties, if a tensor T associates the set of components $A^{\mu\dots\alpha\dots}$ with the system $\{X^\nu\}$, the pairs $\{\{X^\nu\}, A^{\mu\dots\alpha\dots}\}$ also form an equivalence class, two such pairs being equivalent if the coordinate systems are equivalent, and if the corresponding A s are tensorially related according to Eq.(2.21). *The tensor T can then be defined as the equivalence class of all these pairs.*



Tensor Algebra

The algebra of tensors consist of four basic operations- *sum*, *outer product*, *contraction*, and *index permutation*-which all have properties of producing tensors from tensors. All can be defined by the relevant operations on the tensor components, but must then be checked for tensor character.

The *sum* $C^{\mu\dots\alpha\dots}$ of two tensors $A^{\mu\dots\alpha\dots}$ and $B^{\mu\dots\alpha\dots}$ of the same valence is defined thus:

$$C^{\mu\dots\alpha\dots} = A^{\mu\dots\alpha\dots} + B^{\mu\dots\alpha\dots}.$$

Trivially, it is a tensor (we exhibit the proof for a particular case):

$$\begin{aligned} C^{\mu'}_{\alpha'} &= A^{\mu'}_{\alpha'} + B^{\mu'}_{\alpha'} = a^{\mu'}_{\mu\alpha'} A^{\mu}_{\alpha} + a^{\mu'}_{\mu\alpha'} B^{\mu}_{\alpha} = a^{\mu'}_{\mu\alpha'} (A^{\mu}_{\alpha} + B^{\mu}_{\alpha}) \\ &= a^{\mu'}_{\mu\alpha'} C^{\mu}_{\alpha}. \end{aligned}$$

Note, however, that the sum of tensor at *different* points of V_4 is not general a tensor since in the third step above we could not generally pull out the *as*. But under *linear* coordinate transformations the *as* are constant, and then the sum of tensors even at different points is a tensor. Analogous remarks apply to the product of tensors to be defined next.

If A^{\dots} and B^{\dots} are tensors of arbitrary valences, the juxtaposition of their components defines their *outer product*. Thus, for example,

$$C^{\mu\nu}_{\alpha\beta} = A^{\mu}_{\alpha} B^{\nu}_{\beta}.$$

is the tensor of the valence indicated by its indices. As a particular case, A^{\dots} could be a scalar. In conjunction with sum, therefore, we see that any linear

combination of tensors of equal valence is a tensor. The outer product of two vectors $\mathbf{A}(=A^\mu)$ and $\mathbf{B}(=B^\mu)$ is sometimes written $\mathbf{A} \otimes \mathbf{B}$.

Contraction of a tensor of valence (s,t) consists in the replacement of one superscript and one subscript by a dummy index pair, and results in a tensor of valence $(s-1, t-1)$. For example, if $A^{\mu\nu}_{\alpha\beta}$ is a tensor, then

$$B^\mu_{\alpha} = A^{\mu\nu}_{\alpha\nu}$$

is a tensor of valence indicated by its indices. Contraction in conjunction with outer product results in an *inner product*, e.g. $C_{\alpha\beta} = A^\mu_{\alpha} B_{\mu\beta}$. A most important particular case of contraction or inner multiplication arises when no free indices remain: the result is an *invariant*. E.g. A^μ_{μ} , $A^{\mu\nu}_{\mu\nu}$, $A^{\mu\nu} A_{\mu\nu}$ are invariant if the A s are tensors. (A particular case $\delta^\mu_{\mu} = 1$.)

The last of the algebraic tensor operations is *index permutation*. For example, if tensor components $A_{\alpha\beta}$ are exhibited in matrix form, $B_{\beta\alpha} = A_{\alpha\beta}$ denotes the components of the transposed matrix, and those components form a tensor, as is immediately obvious from Eq.(2.20). Index permutations of all order are permissible among *either* the subscripts *or* the superscripts of a tensor. Thus we can form such tensor sums as $A_{\beta\alpha} + A_{\alpha\beta}$ or $A^{\mu\nu}_{\alpha\beta} + A^{\nu\mu}_{\alpha\beta}$, and such tensor equations as $A_{\beta\alpha} = A_{\alpha\beta}$. It follows, in particular, that the symmetry (or antisymmetry) of a tensor is an invariant property, i.e. is preserved under coordinate transformations.

Symmetrical Properties of Tensors

A contravariant or covariant tensor of second or higher rank is said to be *symmetrical* if two components, which are obtained the one from the other by the interchange of two indices, are equal. The covariant tensor $A^{\mu\nu}$, or the contravariant

tensor $A_{\mu\nu}$, is thus symmetrical if for any combination of the indices μ, ν , we have $A^{\mu\nu} = A^{\nu\mu}$, or $A_{\mu\nu} = A_{\nu\mu}$, respectively. In contrast, a covariant or a contravariant tensor of the second, third, and fourth-rank is said to be *antisymmetrical* if two components, which are obtained the one from the other by the interchange of two indices, are equal and of opposite sign, or, for the second-rank tensor, $A^{\mu\nu} = -A^{\nu\mu}$ and $A_{\mu\nu} = -A_{\nu\mu}$ respectively.

For the antisymmetric case of $A^{\mu\nu}$, of the sixteen components, the four components $A^{\mu\mu}$ vanish; the rest are equal and of opposite sign in pairs, so that there are only six components numerically different. Similarly, we see that the antisymmetrical tensor of the third-rank $A^{\mu\nu\alpha}$ has only four numerically different components, while the antisymmetrical fourth-rank has only one. There are no antisymmetrical tensors of higher rank than the fourth in a continuum four-space.

Any second-rank tensor $A^{\mu\nu}$ can be presented as a sum of a symmetric tensor and an antisymmetric tensor. In fact,

$$\begin{aligned} A^{\mu\nu} &= (1/2)[A^{\mu\nu} + A^{\nu\mu}] + (1/2)[A^{\mu\nu} - A^{\nu\mu}] \\ &= B^{\mu\nu} + C^{\mu\nu}, \end{aligned}$$

then obviously $B^{\mu\nu}$ is called symmetric part of $A^{\mu\nu}$ and $C^{\mu\nu}$ is called antisymmetric part of $A^{\mu\nu}$, respectively.

Differentiation of Tensors

We shall write

$$\frac{\partial(A^{\mu\dots\nu}_{\alpha\dots\beta})}{\partial x^\varepsilon} = A^{\mu\dots\nu}_{\alpha\dots\beta,\varepsilon}$$

Then if $A^{\mu\dots\nu}_{\alpha\dots\beta}$ is a (field-) tensor, differentiation of the general tensor component transformation Eq.(2.21) yields (by use of $\partial/\partial\chi^{\varepsilon'} = a^{\varepsilon}_{\varepsilon'}\partial/\partial\chi^{\varepsilon}$):

$$A^{\mu'\dots\nu'}_{\alpha'\dots\beta',\varepsilon'} = a^{\mu'}_{\mu\dots\nu}\alpha'\dots\beta'\alpha\dots\beta a^{\varepsilon}_{\varepsilon'} A^{\mu\dots\nu}_{\alpha\dots\beta,\varepsilon} + B_1 + B_2 + \dots,$$

where B_s are terms involving derivatives of the a s. [It should be noted that a product with implied summations—like the right-hand side of Eq.(2.21)—can be differentiated with complete disregard of these summations, since sum and derivative commute.] Under general coordinate transformations, therefore, $A^{\mu\dots\nu}_{\alpha\dots\beta,\varepsilon}$ is not a tensor. But under *linear* coordinate transformations (a s constant) $A^{\mu\dots\nu}_{\alpha\dots\beta,\varepsilon}$ behaves as a tensor of the valence indicated by all its indices, including ε , since then the B s vanish. By a repetition of the argument, all higher-order partial derivatives,

$$A^{\mu\dots\nu}_{\alpha\dots\beta,\varepsilon\eta} = \frac{\partial^2 (A^{\mu\dots\nu}_{\alpha\dots\beta})}{\partial\chi^{\varepsilon}\partial\chi^{\eta}}$$

etc. also behave as tensors under linear transformations, each partial differentiation adding a new covariant index.

Consider a curve in space defined by the equations $X^{\mu} = X^{\mu}(s)$, where s is a scalar (invariant) parameter. Then dX^{μ}/ds is a vector under *all* transformations, the proof being similar to that for dx^{μ} . That the scalar derivative of any field tensor, $(d/ds)A^{\mu\dots\nu}_{\alpha\dots\beta}$, behave as a tensor under *linear* transformations follow at once from the differentiation of Eq.(2.21). We may note how the four basic vectors of classical mechanics—velocity, acceleration, momentum, force—are all built up from the operations of differentiation and multiplying by scalars: dx^i/dt , d^2x^i/dt^2 , mdx^i/dt , md^2x^i/dt^2 ($t =$ Newtonian time and $i = 1,2,3$).

The Quotient Rule

Although we cannot usually form a 'quotient' of tensors, an object like $C^{\mu\nu}$ in the equation $A^\mu = C^{\mu\nu}B_\nu$, where A^μ and B_ν are tensors, can be formally regarded as a kind of quotient of A^μ and B_μ . This gives the name to a most useful rule for recognizing tensors, the *quotient rule*, which roughly says that the quotient of tensors is itself a tensor. Accurately stated, it reads thus: *If a set of components, when combined by a given type of multiplication with an arbitrary tensor of a given valence yields a tensor, then the set constitutes a tensor.* The general method of proof can be sufficiently indicated by considering the above special case. Suppose we know of the components $C^{\mu\nu}$ that for an arbitrary tensor B_ν the product $C^{\mu\nu}B_\nu$ is a tensor. Let $C^{\mu\nu}$ and $C^{\mu'\nu'}$ be the components of our object in two arbitrary coordinate system $\{X^\mu\}$ and $\{X^{\mu'}\}$. Then first by the hypothesis and second by the tensor character of B_μ , we have

$$C^{\mu'\nu'}B_{\nu'} = a^{\mu'}_{\mu}(C^{\mu\nu}B_\nu) = a^{\mu'}_{\mu}a^{\nu'}_{\nu}C^{\mu\nu}B_{\nu'},$$

whence for all μ' ,

$$(C^{\mu'\nu'} - a^{\mu'}_{\mu}a^{\nu'}_{\nu}C^{\mu\nu})B_{\nu'} = 0.$$

But $B_{\nu'}$ in one coordinate system is quite arbitrary, so we can successively give it values $(1,0,0,0)$, $(0,1,0,0)$, etc. and thereby find that the above expression in parentheses vanishes for all ν' . This shows that $C^{\mu\nu}$ is a tensor.

The Metric

For the special structure of V_4 in which tensors play a role are *metric spaces*, i.e., they possess a rule which assigns 'distances' to pairs of neighbouring points. In

particular, one calls a space (*pseudo-*)*Riemannian* if there exists a quadratic differential form

$$ds^2 = g_{\mu\nu}dX^\mu dX^\nu, \quad (2.22)$$

where the g s are generally functions of position, and are subject only to the restriction $\det(g_{\mu\nu}) \neq 0$. They may, without loss of generality, be assumed to be symmetric: $g_{\mu\nu} = g_{\nu\mu}$. If $ds^2 > 0$ for all $dX^\mu \neq 0$, the space is called *strictly Riemannian*. Euclidean N -space, which has

$$ds^2 = (dX^1)^2 + (dX^2)^2 + \dots + (dX^N)^2$$

is only one example. Since we require ds^2 to be invariant, it follows from a simple variation of the quotient rule that $g_{\mu\nu}$ must be a tensor. We call it the *metric tensor*, and Eq.(2.22) the *metric*.

In Riemannian four-space one often adopts a notations for vectors analogous to that in the three-space of which we have learned earlier. A scalar product of two four-vectors is then defined by, similar to Eq.(2.11),

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= A^\mu B_\nu = g^{\mu\nu} A_\mu B_\nu \\ &= A_\mu B^\nu = g_{\mu\nu} A^\mu B^\nu. \end{aligned} \quad (2.23)$$

This clearly satisfies the relations

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}, \quad \mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$$

Two vectors are said to be *orthogonal* if their scalar product vanishes. A particular case of scalar product is the *square* of a vector (which in pseudo-Riemannian spaces can be positive or negative):

$$\mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A} = g_{\mu\nu} A^\mu A^\nu. \quad (2.24)$$

From it one defines the (non-negative) *magnitude* $|\mathbf{A}|$, or simply A , by the equation $A = |\mathbf{A}|^{1/2} \geq 0$. Thus the metric ds^2 itself can be regarded as the square of the differential displacement vector $d\mathbf{s} = dX^\mu$. For its magnitude we write ds .

In Riemannian spaces there exists a fifth basic algebraic tensor operation, namely the *raising* and *lowering* of indices. For this purpose we define $g^{\mu\nu}$ as the elements of the inverse of the matrix $(g_{\mu\nu})$. Because of the symmetry of $(g_{\mu\nu})$, its inverse $(g^{\mu\nu})$ is also symmetric. The $g^{\mu\nu}$ are defined uniquely by the equations

$$g^{\mu\nu} g_{\nu\alpha} = \delta^\mu_\alpha. \quad (2.25)$$

If $g^{\mu'\nu'}$ denote the tensor transform of $g^{\mu\nu}$ in the $X^{\mu'}$ system [according to Eq.(2.19)], then by the form-invariance of tensor component equations (since $g^{\mu\nu}$ and δ^μ_ν are tensors), we have from Eq.(2.26)

$$g^{\mu'\nu'} g_{\mu'\alpha'} = \delta^{\mu'}_{\alpha'}. \quad (2.26)$$

But these are also the equations that uniquely define the inverse $(g^{\mu'\nu'})$ of the matrix $(g_{\mu'\nu'})$. Hence the $g^{\mu\nu}$ defined by equations like Eq.(2.27) in all coordinate systems constitute a contravariant tensor said to be *conjugate* to $g_{\mu\nu}$.

Now the operations of raising and lowering indices consist in forming inner products of a given tensor with $g_{\mu\nu}$ or $g^{\mu\nu}$. For examples,

$$\begin{aligned} A^\mu &= g^{\mu\nu} A_\nu, \\ A_\mu &= g_{\mu\nu} A^\nu. \end{aligned} \quad (2.27)$$

As can easily be verified, these operations are consistent, in that the raising of a lowered index, and vice versa, leads back to the original component. They can of course be extended to raise or lower any or all of free indices of any given tensor: e.g. if $A_{\mu\nu}{}^\alpha$ is a tensor we can define $A^\mu{}_{\nu\alpha}$ by the equations

$$A^\mu{}_{\nu\alpha} = g^{\mu\beta}g_{\nu\rho}A_{\beta\alpha}{}^\rho.$$

It may be mentioned that there is another important use of the metric tensor $g_{\mu\nu}$, namely in the construction of the so-called 'covariant derivative' of tensor under general coordinate transformations. But this is not needed in case of flat space, and so will not be dealt with here.

Pseudo tensors and Dual Tensors

The only kind of transformation that we have discussed so far are called the *proper* rotation transformations which can be built from a succession of infinitesimal transformations. In contrast, sometimes we have to deal with the improper rotation transformation which are discrete and cannot be obtained by compounding infinitesimal transformation. Operations such as parity or time reversal are examples of the improper transformations. Their transformation matrices, denoted by (a), may be characterized by the condition

$$\det (a) = -1, \quad (2.28a)$$

while the proper transformation matrices satisfy

$$\det (a) = +1. \quad (2.28b)$$

The *parity* transformation corresponds to a rotation by 180° plus a reflection through a plane, while the time-reversal transformation reverses the time but not the spatial coordinates. We will find it useful to classify tensors with respect to their properties under both proper and improper transformations.

For example, we consider the transformation of vector \mathbf{b} in three-space. Under proper transformation \mathbf{b} transforms as

$$b^{i'} = a^{i'}_j b^j, \quad (2.29)$$

where $a^{i'}_j$ is the coefficients of the transformation. But special mention must be made if we define $\mathbf{b} = \mathbf{c} \times \mathbf{d}$ because, in component form, this vector shorthand reads

$$b^i = \varepsilon_{ijk} c^j d^k, \quad (2.30)$$

where ε_{ijk} is the so-called Levi-Civita symbol in three-space which is defined as

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } i,j,k \text{ is an even cyclic permutation of } 1,2,3, \\ -1 & \text{if } i,j,k \text{ is an odd cyclic permutation of } 1,2,3, \\ 0 & \text{if } i,j,k \text{ is not a cyclic permutation of } 1,2,3. \end{cases} \quad (2.31)$$

Because of the presence of two vectors on the right-hand side, the cross product has some attributes of a traceless antisymmetric second-rank tensor. Since such a tensor, in three-space, has only three independent components, we treat it as a vector. This has justification, of course, only in so far as it transforms under proper transformation law Eq.(2.29). In actual fact, the transformation law for the cross product is

$$b^{i'} = \det(a) a^{i'}_j b^j, \quad (2.32)$$

For proper rotations, we have $\det(a)=+1$, Eq.(2.32) is thus in agreement with the basic coordinate transformation Eq.(2.29). Thus, under proper transformations, the cross product transforms as a vector.

We now consider some discrete transformations called parity transformation, or spatial inversion. Space inversion corresponds to reflection of all three components of every coordinate vector through the origin, $\mathbf{x}_i \rightarrow \mathbf{x}'_i = -\mathbf{x}_i$. This transformation correspond to $\det(a)=-1$. It follows that vector change sign under spatial inversion, but cross products, which behave according to Eq.(2.31), do not. We are therefore force to distinguished two kinds of vectors (under proper transformation):

Polar vectors (or just vector) that transform according to Eq.(2.29) and under spatial inversion, $\mathbf{x}_i \rightarrow \mathbf{x}'_i = -\mathbf{x}_i$, behave as $\mathbf{b} \rightarrow \mathbf{b}' = -\mathbf{b}$.

Axial vectors or pseudovectors that transform according to Eq.(2.32) and under spatial inversion $\mathbf{x}_i \rightarrow \mathbf{x}'_i = -\mathbf{x}_i$, $\mathbf{b} \rightarrow \mathbf{b}' = \mathbf{b}$.

Similar distinction must be made for scalar under the transformation. We speak of scalars or pseudoscalars, depending on whether the quantities do not or do change sign under spatial inversion. The triple scalar product $\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})$ is an example of a pseudoscalar quantity, provided \mathbf{b} , \mathbf{c} , and \mathbf{d} are all polar vectors. The transformation of higher rank tensors under spatial inversion can be deduced directly if they are built up by taking products of components of polar or axial vectors. If a tensor of rank r transforms under spatial inversion with a factor $(-1)^r$, we call it a true tensor or just a tensor, while if the factor is $(-1)^{r+1}$ we call it a *pseudotensor* of rank r . For example, we consider the pseudotensor of rank two in four-space, it is a sixteen-component quantity in every coordinate system with the transformation law

$$A^{\mu'\nu'} = (\det a) a^{\mu'}_{\mu} a^{\nu'}_{\nu} A^{\mu\nu} \quad (2.33a)$$

while the ordinary tensors will transform as

$$A^{\mu'\nu'} = a^{\mu'}_{\mu} a^{\nu'}_{\nu} A^{\mu\nu} \quad (2.33b)$$

The values of $\det(a)$ depend on type of the transformation as have mentioned earlier.

We have already stated that, by means of the Levi-Cavita symbol Eq.(3.31), we can associate a second-rank antisymmetric tensor e_{jk} with a pseudovector (axial vector) \mathbf{b} by

$$b^i = \varepsilon^{ijk} e_{jk}, \quad (2.34)$$

or,

$$\mathbf{b} = (b^1, b^2, b^3) = (e_{23}, e_{31}, e_{12}).$$

The components of \mathbf{b} are thus considered as the components of an axial vector dual to tensor e_{jk} . Similarly, in four-space we can associate an antisymmetrical tensor of rank $r \leq 4$ with a pseudotensor of rank $(4-r)$ by means of the Levi-Cavita symbol in four-space $\varepsilon^{\mu\nu\alpha\beta}$ defined as

$$\varepsilon^{\mu\nu\alpha\beta} = \begin{cases} +1 & \text{if } \mu\nu\alpha\beta \text{ is an even cyclic permutation of } 0,1,2,3, \\ -1 & \text{if } \mu\nu\alpha\beta \text{ is an odd cyclic permutation of } 0,1,2,3, \\ 0 & \text{if } \mu\nu\alpha\beta \text{ is not a cyclic permutation of } 0,1,2,3. \end{cases} \quad (2.35)$$

Note that the tensor is totally antisymmetric fourth-rank tensor and is a pseudotensor under spatial inversion. Then we can say that for any four-vector \mathbf{A} ,

any antisymmetry tensor \mathbf{B} , and any third-rank antisymmetric tensor \mathbf{C} , we can construct new tensors defined by

$${}^*A_{\mu\nu\alpha} = \varepsilon_{\mu\nu\alpha\beta} A^\beta, \quad {}^*B_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\alpha\beta} B^{\alpha\beta}, \quad {}^*C_\mu = \frac{1}{3!} \varepsilon_{\mu\nu\alpha\beta} B^{\nu\alpha\beta}, \quad (2.36)$$

where $\varepsilon_{\mu\nu\alpha\beta} = -\varepsilon^{\mu\nu\alpha\beta}$. We call *A , *B , and *C , the *dual tensors* of A , B , and C respectively. By the definition of $\varepsilon^{\mu\nu\alpha\beta}$, we can verify that (aside from sign) any completely antisymmetric tensor \mathbf{D} can be found from its dual ${}^*\mathbf{D}$ by taking the dual once again,

$${}^{**}\mathbf{D} = (-1)^{r+1} \mathbf{D}. \quad (2.37)$$

This shows that \mathbf{D} and ${}^*\mathbf{D}$ contain precisely the same information. Then the dual of the dual of antisymmetric second-rank tensor \mathbf{B} is the negative of the original one or,

$${}^{**}\mathbf{B} = -\mathbf{B}.$$

In this sense $*$ has the same property as the imaginary number i : $** = ii = -1$. Thus we can write

$$e^{*\theta} = \cos\theta + {}^*\sin\theta.$$

This operation, applied to \mathbf{B} , carries attention from the generic second-rank tensor \mathbf{B} to another "duality rotated tensor" $e^{*\theta}\mathbf{B}$. If \mathbf{B} presents as a tensor field which is satisfied field equations in empty space, so does its dual, $e^{*\theta}\mathbf{B}$, with suitable choice of θ . If we take $\theta = \pi/2$ then the definition ${}^*B_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\alpha\beta} B^{\alpha\beta}$ is appeared.