A Gentle Introduction to Tensors

Thomas C. Mangan

Abstract

We extend the familiar concepts of scalar and vector quantities in order to arrive at a definition of tensor. In this manner, we make the definition and ideas less intimidating for readers who are not ready to tackle a typical textbook introduction to tensors.

1 Introduction

Many undergraduate students complete a Bachelors degree having never been exposed to the concept of tensors. Perhaps this is understandable, as the topic is both broad and deep, requiring knowledge of multivariate calculus, vector analysis, linear algebra, and differential equations. For those with a basic understanding of these foundational topics, however, an introduction to tensors should be within reach, and may alleviate some of the panic which may otherwise accompany a student's first exposure to tensor equations.

To illustrate the usual nature of that first exposure, we present the first definition of tensor found on page 11 of Synge's *Tensor Calculus*[9]:

> A set of quantities T^{rs} are said to be the components of a contravariant tensor of the second order if they transform according to the equation
>
> $$T'^{rs} = T^{mn} \frac{\partial x'^r}{\partial x^m} \frac{\partial x'^s}{\partial x^n}. \tag{1.1}$$

This is the type of definition that's used by somebody who already knows what a tensor is, and needs to determine if the thing she's looking at is, in fact, a tensor. It doesn't offer any understanding for a student who is initially trying to find out what a tensor *is*. A tensor, we read, is a set of quantities called components. The rest is a jumble of unfamiliar terms and notation. What does "contravariant" mean? If we don't know what a tensor is, how do we know what it means for its components to "transform"? Is that really T to the mn power? For the student who has resisted the urge to close the book and run away screaming, delving into these questions can be a long and complicated process, the result of which is the barest readiness to be introduced to the concept of a tensor.

In this paper, we hope to guide the reader through more familiar concepts of vector notation and transformation, to end up with a definition of tensor, its notational conventions, and a motivation for expending the required effort. Some knowledge of vector analysis is assumed, as these leading sections will be presented in the manner of "reminding" the reader of the vector concepts upon which we will build. As our intent is to prevent panic, rather than induce it, we have avoided rigorous proof where it does not contribute to understanding. In this manner, students of science and engineering may take away an understanding of the basic ideas, while those interested in pure mathematics should be less intimidated by the concept and notation, and may delve more deeply into the rigorous mathematics involved.

2 Invariant Quantities

There are certain things we can measure that do not depend upon our frame of reference. For instance, we can impose a coordinate system with a fixed origin on the street outside, and then measure the temperature of the asphalt at a specific point on a hot day. Our friend can then drive past in a car, where she has defined her own coordinate system with origin inside (and moving with) the car. When she measures the temperature of the asphalt at that same point, she will obtain the same result as we did. A quantity that behaves like this, unchanging no matter what coordinate system we use when measuring, we call an invariant.

To put this concept into symbols, we have a transformation rule $T = \bar{T}$, where T is the temperature measured in frame S, and $\bar{T}$ is the temperature measured in frame $\bar{S}$. It should also be clear that the rule holds during a transformation between *any* two arbitrary coordinate systems. These are the conditions that must be met for a mathematical object to be a tensor. We have a collection of quantities (only one in this case), associated with a point in space, which transform according to an unchanging rule between any two coordinate systems. Thus temperature, which we recognize as a *scalar* quantity, can also be considered a *tensor of rank 0*. We'll come back to the meaning of "rank" after we've seen a few more examples. While every rank-0 tensor is a scalar, not every scalar is a rank-0 tensor.

3 Vectors

3.1 New Notation

The next logical step is to add more quantities to the collection we're examining. In so doing we will obtain another familiar object: a vector. This object is a collection of quantities, again associated with a point in space, but now we need to be more careful about how these quantities are measured. Before exploring further, we need to introduce some new notation which will greatly simplify matters as we add more quantities.

We are quite used to seeing a vector described as a sum of several terms, each term the product of a coefficient and a unit vector, such as

$$\mathbf{r} = \alpha\hat{\mathbf{x}} + \beta\hat{\mathbf{y}} + \gamma\hat{\mathbf{z}}.$$

In this familiar notation, each hatted bold symbol is a unit vector in the direction of a basis vector for the space in which we're working. For simplicity we've chosen as our basis the three orthogonal axes of a Cartesian coordinate system, but this restriction is by no means necessary. In three dimensions, this is a perfectly fine way to represent the unit vectors, but if we wish to describe vectors in more dimensions, we'll rapidly run out of letters to put hats on. Let us instead represent those same unit vectors using a letter and a superscripted index, so that our vector (the same vector as above) is now represented by:

$$\mathbf{r} = \alpha_1 x^1 + \alpha_2 x^2 + \alpha_3 x^3.$$

This is an improvement over using different letters for each unit vector, but this notation is still going to get tedious when we start adding dimensions. Fortunately, we have a symbol to express a sum of similar terms:

$$\mathbf{r} = \sum_{i=1}^{3} \alpha_i x^i.$$

This is certainly more compact than the expression we started with, but thanks to Einstein, we can reduce our notation yet further. The good professor noted that each term in our expression has a repeated index, and we can take it as implied that, when we see an index exactly twice in a single term, once as a superscript and once as a subscript, we will sum that term over all possible values of that index. Thus our vector, in this new notation, is simply

$$\mathbf{r} = \alpha_i x^i. \tag{3.1}$$

The index we're summing over in the Einstein notation is called a *dummy index*, and can be replaced with any variable, so long as it is used exactly twice, in the manner prescribed.

Note that (3.1) can represent a vector in an arbitrary number of dimensions in a very compact form. The values allowed for these indices will usually be made clear by context. It has become customary to use lower-case Latin letters to represent indices which take on the values 1, 2, ..., N, whilst we use lower-case Greek letters to denote indices which take the values 0, 1, 2, ..., N. Thus a four-vector, commonly found in relativity, can be easily expressed as

$$a_\mu x^\mu = a_0 x^0 + a_1 x^1 + a_2 x^2 + a_3 x^3.$$

The left-hand side of this equation will clearly be more pleasant to write during our calculations than the right, so we see that it is worth the effort to become familiar with the compact notation.

Armed with our new shorthand, we can now examine a few types of vectors and see how they might be measured by observers in different reference frames.

3.2 Position Vectors

Recall that a vector is a mathematical construction which describes both a magnitude and a direction. These quantities can be interpreted in a number of ways: a position, a displacement, a flux, an area, and so forth. When we study a vector, we generally choose our coordinate system such that measuring and describing the vector is most convenient.

Other times, we may know the components of a vector in one reference frame, but find that our calculations will be more easily performed in another. In those cases, we need to understand how to manipulate the components from one coordinate system such that they yield the components in the second. We also need to know if there's some indicator that a vector will be well-behaved when transforming to *any* arbitrary system. Let's look at a representative vector and how it might appear in and transform between two systems.

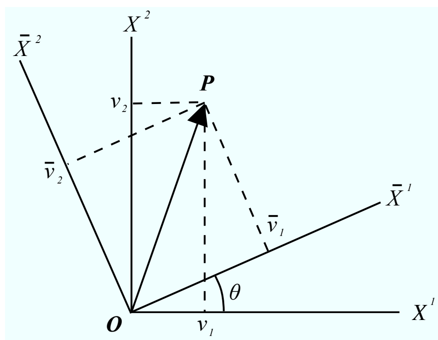

Figure 3.1: Vector OP measured in two coordinate systems [4]

In Figure 3.1, we have a position vector OP. When we measure it in the Cartesian X^i-coordinate system, we obtain the vector v_i with coordinates (v_1, v_2). But consider another Cartesian coordinate system, $\bar{X}^i$, rotated by an angle θ with respect to X^i. Measuring the vector OP in this system yields a vector $\bar{v}_i = (\bar{v}_1, \bar{v}_2)$. In general, $v_1 \neq \bar{v}_1$ and $v_2 \neq \bar{v}_2$. Although our vector has not changed at all, the measurements of its components are different.

Examining the situation a little bit more closely, the point P has coordinates (x^1, x^2) in the initial coordinate system, and coordinates $(\bar{x}^1, \bar{x}^2)$ in the rotated system. The components of our vector v^i in X^i are $(v^1, v^2) = (x^1, x^2)$, and the components of our vector $\bar{v}^i$ in $\bar{X}^i$ are $(\bar{v}_1, \bar{v}_2) = (\bar{x}_1, \bar{x}_2)$. To introduce some new terminology, the components of v_i are *covariant* with the coordinates under a rotation of axes.

On the other hand, if we measure OP in a coordinate system whose origin O' is *not* coincident with that of X^i, we will find the components of a vector $O'P$, which is not the same vector we started with. We can't find out how the components of an object transform into another coordinate system if we don't measure the same object in both. Clearly, we can't speak of *a* tensor when we're looking at two different entities. Although it looked promising when our coordinate systems differed only by a rotation, the transformation of OP did not generalize to an *arbitrary* change of systems. Since $OP \neq O'P$, a position vector is *not* a tensor.

3.3 Displacement Vectors

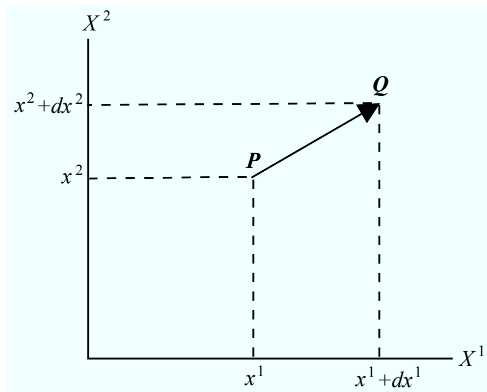

Figure 3.2: Infinitesimal displacement vector OP

Let us now consider the infinitesimal displacement vector PQ in Figure 3.2. It clearly has components dx^1 and dx^2 in the two-dimensional Cartesian coordinate system with x^i as basis vectors. According to convention, let us call this vector dx^p, where in this case p can equal 1 or 2. Now suppose we want to describe that vector using a different coordinate system with basis $\bar{x}^i$. Before we learned our new notation, we would have performed this transformation by writing out the terms of the chain rule:

$$d\bar{x}^1 = \frac{\partial \bar{x}^1}{\partial x^1} dx^1 + \frac{\partial \bar{x}^1}{\partial x^2} dx^2 \tag{3.2a}$$

$$d\bar{x}^2 = \frac{\partial \bar{x}^2}{\partial x^1} dx^1 + \frac{\partial \bar{x}^2}{\partial x^2} dx^2, \tag{3.2b}$$

which is straightforward, but tedious. It takes a lot of writing to get a result, and we're still only dealing with two dimensions. Writing out this transformation in three dimensions would involve three equations, each with three terms on the right-hand side. The transformation in n dimensions requires n equations of n terms each.

Now, thanks to some shorthand, we can describe this transformation as

$$d\bar{x}^q = \frac{\partial \bar{x}^q}{\partial x^p} dx^p. \tag{3.3}$$

This looks suspiciously like our initial definition (1.1) of a tensor, and indeed the displacement vector dx^p is a *rank-1 contravariant tensor*. It is a collection of quantities, associated with a point

P in space, which transform according to the definition into any other arbitrary coordinate system. Many textbooks describe a displacement vector as the *prototype* of a contravariant tensor.

Now that we've seen a rank-0 tensor, and a rank-1 tensor, we can also note that we use the word *rank* to denote how many indices it takes to specify our tensor. When we come to rank-2 tensors, we will see that they require two indices, and so forth into higher ranks.

3.4 Gradient Vectors

Consider next a set of functions in the x^r coordinates $\phi(x^1, x^2, \ldots, x^N)$. Taking the partial derivatives of our functions with respect to a new set of coordinates $\bar{x}^p$, we obtain

$$\frac{\partial \phi}{\partial \bar{x}^p} = \frac{\partial \phi}{\partial x^r} \frac{\partial x^r}{\partial \bar{x}^p}.$$

To simplify this equation, we can recognize two of these terms as gradients, and rewrite our equation as

$$\bar{\nabla} \phi_p = \nabla \phi_r \frac{\partial x^r}{\partial \bar{x}^p}.$$

Since gradients are vectors, let's go ahead an put them in the notation we've been using up to now, and we see that

$$\bar{T}_p = \frac{\partial x^r}{\partial \bar{x}^p} T_r.$$

which looks very much like our definition of a contravariant vector, only the partial derivative is the inverse of the previous one. This formula is actually the definition of a *covariant* vector, and we can say that the gradient vector is a *rank-1 covariant tensor*. Many textbooks, in fact, will say that a gradient is the prototype of a covariant tensor.

3.5 Final Vector Notes

Although we found our contravariant vector by examining an infinitesimal displacement, a vector need not be either a displacement or infinitesimal to be a rank-1 contravariant tensor. Satisfaction of the transformation rule is the necessary and sufficient condition for tensor character. Similarly, a gradient is merely an example of a covariant rank-1 tensor, and any vector which satisfies the definition can rightly bear the name.

Sharp-eyed reader may have noted another notational difference between our contravariant and covariant definitions. The contravariant tensors were indexed using a superscript, while the covariant tensors bore a subscript index. This is the customary way to indicate which rule the tensor satisfies when undergoing a transformation. Tensors of higher rank can have components of both types. We will call these *mixed* tensors, and their symbols will include both subscripted and superscripted indices.

4 Rank-N Tensors

4.1 Manipulating Rank-1 Tensors

In our previous interactions with vectors, we have learned to perform certain operations on them. We have multiplied a vector by a scalar to form a new vector with the same direction as the first. We have taken an inner product of two vectors to produce a scalar. We have taken a cross-product of two vectors to form a third vector orthogonal to both. We have added two vectors to obtain a third vector.

This vector addition rule still holds for rank-1 tensors, provided we are adding tensors of the same type. Two covariant tensors add to form a new covariant tensor: $a_p + b_p = c_p$. Similarly for contravariant tensors, $t^m + u^m = v^m$. This addition generalizes to same-type tensors of higher rank.

It's comforting to see that vector addition still holds, at least for vectors with the same type of components. Now let's examine what vector multiplication looks like. With the benefit of knowing the outcome, we're going to start by multiplying two vectors which are *not* of the same type, but share the same index.

Consider the product $u_i v^i$. Since we have the summation notation fresh in our head, we'll notice that we have a repeated index, and that it appears once in a subscript and once in a superscript. That means we're going to sum the term over all allowed values of the index,

$$u_i v^i = u_1 v^1 + u_2 v^2 + u_3 v^3.$$

That's the product of the coefficients of the first unit vector, plus the product of the coefficients of the second unit vector, plus the product of the coefficients of the third unit vector. In the old notation, we would have written $\mathbf{u} \cdot \mathbf{v} = u_x v_x + u_y v_y + u_z v_z$, and we recognize that we've found the dot-product, or inner-product, of our two vectors. Both notations tell us the same thing, but the old way doesn't scale well to more dimensions.

4.2 The Inner Product

It's worth a minor digression here to note that the dot-product of two rank-1 tensors has an interesting property. Let's first remind ourselves of the definitions. Remember that a contravariant tensor transforms according to the rule

$$\bar{a}^i = \frac{\partial \bar{x}^i}{\partial x^j} a^j, \tag{4.1}$$

and covariant tensors transform according to their own rule,

$$\bar{t}_i = \frac{\partial x^j}{\partial \bar{x}^i} t_j. \tag{4.2}$$

Claim: The inner-product of two rank-1 tensors is an invariant scalar.

Proof. Let u_i and v^i be a covariant and a contravariant vector, respectively, and let x^n and $\bar{x}^n$ represent the basis vectors in two coordinate systems. Then the inner product of our vectors is expressed as

$$u_i v^i = \frac{\partial \bar{x}^j}{\partial x^i} \bar{u}_j \frac{\partial x^i}{\partial \bar{x}^k} \bar{v}^k = \frac{\partial \bar{x}^j}{\partial x^i} \frac{\partial x^i}{\partial \bar{x}^k} \bar{u}_j \bar{v}^k = \frac{\partial \bar{x}^j}{\partial \bar{x}^k} \bar{u}_j \bar{v}^k$$

But $\bar{x}^j$ and $\bar{x}^k$ are basis vectors, thus they are independent by definition. Therefore the quotient $\frac{\partial \bar{x}^j}{\partial \bar{x}^k}$ vanishes unless $j = k$, in which case it is unity. Now the only term surviving on the right-hand side is $\bar{u}_j \bar{v}^j$, and we recall that our summation variable is a dummy index, which we can replace with any letter we choose. That is, $u_i v^i = \bar{u}_i \bar{v}^i$. Thus $u_i v^i$ is an invariant. To put it another way, the dot-product of two rank-1 tensors is a rank-0 tensor. □

We glossed over something interesting on our way to completing that proof. There is a special symbol in tensor language to describe that entity which is unity when the indices are equal, and zero otherwise.

$$\delta^j_k = \frac{\partial \bar{x}^j}{\partial \bar{x}^k} = \begin{cases} 1, j = k \\ 0, j \neq k \end{cases} \tag{4.3}$$

The symbol δ^r_q in (4.3) is called the Kronecker delta after 19th century German mathematician Leopold Kronecker, and it finds use in many advanced operations we won't cover in this introductory paper.

4.3 Rank-2 Tensors

Note that when we took the inner product of two rank-1 tensors, we had to use the covariant components of one and the contravariant components of the other, and use the same index for both. What happens if we use the same type of components for the two tensors? Let's multiply two contravariant rank-1 tensors a^j and b^k. Since we no longer have a summation indicated, this is clearly not going to be an inner product. Instead we have

$$a^j b^k = (a^1 x^1 + a^2 x^2 + a^3 x^3)(b^1 x^1 + b^2 x^2 + b^3 x^3),$$

and our basic algebra classes suggest that the appropriate multiplication method here is term-by-term. That is

$$\begin{aligned} a^j b^k = {} & a^1 b^1 (x^1)^2 + a^1 b^2 x^1 x^2 + a^1 b^3 x^1 x^3 \\ & + a^2 b^1 x^2 x^1 + a^2 b^2 (x^2)^2 + a^2 b^3 x^2 x^3 \\ & + a^3 b^1 x^3 x^1 + a^3 b^2 x^3 x^2 + a^3 b^3 (x^3)^2. \end{aligned}$$

We end up with something which still looks like what we've been calling a tensor. It's a collection of terms, and each term is a product of coefficients and basis vectors, but we haven't seen that it actually *is* a tensor. Using our rules (4.1) and (4.2),

$$\bar{u}^r \bar{v}^s = \frac{\partial \bar{x}^r}{\partial x^m} u^m \frac{\partial \bar{x}^s}{\partial x^n} v^n = \frac{\partial \bar{x}^r}{\partial x^m} \frac{\partial \bar{x}^s}{\partial x^n} u^m v^n.$$

The quantity $u^m v^n$ transforms according to rule (1.1) into $\bar{u}^m \bar{v}^n$, thus we are justified in calling the quantity $w^{mn} = u^m v^n$ a contravariant tensor of rank 2. Although we found this tensor by multiplying two others, we should note again that satisfaction of the transformation rule is both necessary and sufficient to indicate a tensor. It needn't have been formed as a product of two (or more) other tensors. As mentioned above, we also see that our rank-2 tensor has 2 indices.

Unfortunately, our geometrical intuition will not carry us through higher order tensors as it has with rank 1. It's hard to picture $(x^1)^2$ or $x^2 x^3$ as unit vectors in some given direction. We can note that in orthogonal coordinate systems, $x^i x^j = \delta_i^j$, so that our mixed terms vanish, but the same is not true for non-orthogonal coordinates. In either case, we're still left with a difficult geometric interpretation, and it fails completely when we add another rank. This may be part of the reason textbooks seem to have such difficulty giving a basic definition of a tensor. They can't appeal to our geometric instincts in the same way as is possible for vectors.

Moving past our difficulty in forming a mental picture, we can define rank-2 covariant tensors

$$\bar{g}_{\mu\nu} = \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} g_{\alpha\beta}, \tag{4.4}$$

and rank-2 mixed tensors

$$\bar{t}_j^k = \frac{\partial x^m}{\partial \bar{x}^j} \frac{\partial \bar{x}^k}{\partial x^n} t_m^n.$$

Extending the concept even further, the components of a rank-3 contravariant tensor transform according to the rule

$$\bar{A}^{def} = \frac{\partial \bar{x}^d}{\partial x^p} \frac{\partial \bar{x}^e}{\partial x^q} \frac{\partial \bar{x}^f}{\partial x^r} A^{pqr},$$

and so forth up to n dimensions.

Since the notation may still seem daunting, let's remind ourselves that a rank-2 tensor in four-dimensional spacetime represents 16 equations, each with 16 terms. Here is the first of the 16 equations represented by the metric tensor (4.4):

$$\begin{aligned}\bar{g}_{00} = {} & \frac{\partial x^0}{\partial \bar{x}^0}\frac{\partial x^0}{\partial \bar{x}^0}g_{00} + \frac{\partial x^0}{\partial \bar{x}^0}\frac{\partial x^1}{\partial \bar{x}^0}g_{01} + \frac{\partial x^0}{\partial \bar{x}^0}\frac{\partial x^2}{\partial \bar{x}^0}g_{02} + \frac{\partial x^0}{\partial \bar{x}^0}\frac{\partial x^3}{\partial \bar{x}^0}g_{03} \\ & + \frac{\partial x^1}{\partial \bar{x}^0}\frac{\partial x^0}{\partial \bar{x}^0}g_{10} + \frac{\partial x^1}{\partial \bar{x}^0}\frac{\partial x^1}{\partial \bar{x}^0}g_{11} + \frac{\partial x^1}{\partial \bar{x}^0}\frac{\partial x^2}{\partial \bar{x}^0}g_{12} + \frac{\partial x^1}{\partial \bar{x}^0}\frac{\partial x^3}{\partial \bar{x}^0}g_{13} \\ & + \frac{\partial x^2}{\partial \bar{x}^0}\frac{\partial x^0}{\partial \bar{x}^0}g_{20} + \frac{\partial x^2}{\partial \bar{x}^0}\frac{\partial x^1}{\partial \bar{x}^0}g_{21} + \frac{\partial x^2}{\partial \bar{x}^0}\frac{\partial x^2}{\partial \bar{x}^0}g_{22} + \frac{\partial x^2}{\partial \bar{x}^0}\frac{\partial x^3}{\partial \bar{x}^0}g_{23} \\ & + \frac{\partial x^3}{\partial \bar{x}^0}\frac{\partial x^0}{\partial \bar{x}^0}g_{30} + \frac{\partial x^3}{\partial \bar{x}^0}\frac{\partial x^1}{\partial \bar{x}^0}g_{31} + \frac{\partial x^3}{\partial \bar{x}^0}\frac{\partial x^2}{\partial \bar{x}^0}g_{32} + \frac{\partial x^3}{\partial \bar{x}^0}\frac{\partial x^3}{\partial \bar{x}^0}g_{33}. \end{aligned} \tag{4.5}$$

Remember, (4.5) is only one sixteenth of a rank-2 tensor in 4 dimensions. It's among the smaller entities we'll encounter in tensor analysis. The rank-3 tensor in 3 dimensions A^{pqr} mentioned above represents 27 equations, each with 27 terms.

In his PBS NOVA special "The Elegant Universe", Columbia University string theorist Brian Greene briefly shows us the Einstein Field Equations,

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}, \tag{4.6}$$

which again represents 16 equations, and this time each of the 3 tensors in the equation has 16 terms per equation. We're not going to write this one out. Clearly, operations with tensors are made possible by the summation notation. Computation without it would be prohibitively time-consuming and error-prone.

4.4 Tensor as Matrix

In the process of learning relativity, physicists sometimes write the metric tensor (4.4) as a matrix [6], and indeed its expanded form (4.5) suggests a matrix form with its indices.

$$g_{\mu\nu} = \begin{bmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{bmatrix} \tag{4.7}$$

Substituting (4.7) into (4.6), along with its counterpart matrices for the Ricci Tensor $R_{\mu\nu}$ and the stress-energy tensor $T_{\mu\nu}$, we come up with a matrix equation. This should be more familiar to us, as we've encountered similar formulae in Linear Algebra.

$$\begin{bmatrix} R_{00} & R_{01} & R_{02} & R_{03} \\ R_{10} & R_{11} & R_{12} & R_{13} \\ R_{20} & R_{21} & R_{22} & R_{23} \\ R_{30} & R_{31} & R_{32} & R_{33} \end{bmatrix} - \frac{1}{2}R \begin{bmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{bmatrix} = 8\pi G \begin{bmatrix} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & T_{11} & T_{12} & T_{13} \\ T_{20} & T_{21} & T_{22} & T_{23} \\ T_{30} & T_{31} & T_{32} & T_{33} \end{bmatrix}$$

Matrices stop at 2 dimensions, though, so they're not useful for representing tensors of higher order than 2. Computer programmers who deal with multi-dimensional array structures may have some success in visualizing a 3- or 4- or n-dimensional matrix, but even then our matrix operations don't scale up. We again see the summation notation to be the best choice available.

4.5 Some Basic Tensor Operations

Just as we could add rank-1 tensors of similar type, we can add rank-n tensors of the same covariant, contravariant, or mixed nature:

$$C^{ab} = A^{ab} + B^{ab}$$

$$T^{\mu\nu}_{\lambda} = R^{\mu\nu}_{\lambda} + S^{\mu\nu}_{\lambda}$$

The multiplication of rank-1 tensors also generalizes to higher orders, and in so doing loses the condition that the tensors be of the same type. Therefore we can multiply any two tensors to obtain another tensor, with the indices stacking in an intuitive manner.

$$D^{ab}_{pqr} = E^{a}_{pq} F^{b}_{r}$$

There are myriad more advanced operations beyond the scope of this paper, whose intent is merely to ease the undergraduate into an introduction to tensors. Topics for further study include tensor symmetries and anti-symmetries, the Christoffel symbols, differentiation of tensors, and the derivative notations (including the comma derivative and semi-colon derivative). The bibliography provides a starting point for exploring these subjects, the next most accessible reference being Kolecki's freely available paper [3].

5 Conclusion

We began with frustration over the choice of many tensor textbooks and learning materials to start with a definition which is of no use to the beginner. That definition was more frightening than helpful; thus we set out to discover the nature of a tensor, so we could properly apply the textbook definition.

We found that vectors which obey certain transformation rules can justifiably be called tensors, which offered us an accessible route into the more difficult concept of higher-rank tensors. By building on our knowledge of what a vector is, i.e. a collection of components describing a direction and a magnitude, we discovered that a tensor is a collection of components as well. For higher rank tensors, our geometric interpretations break down, but we can determine that a tensor is a set of quantities which provide information about a point in space, or spacetime, or some generalized coordinate system. Most importantly, that collection of components obeys very specific rules that enable us to study it in any convenient coordinate system.

Therein lies the power of tensors. Although we've avoided some of the details in order to keep this introduction gentle, all of the gymnastics we've been through are designed to make sure tensors transform correctly into *any* coordinate system, be it orthonormal, non-orthogonal, curvilinear, etc. This is especially useful in the field of relativity, where spacetime itself is not flat in the presence of a gravitational field.

We should include a caveat here that physicists and mathematicians may be talking about different but related things when they use the word tensor. Our description has been of a physical tensor, which a mathematician, when asked, would be likely to call a tensor field. This distinction is important at higher levels, but we have chosen to downplay it to avoid confusion.

Perhaps most importantly at this level, we saw how the tensor notation is built up from more familiar symbology, why a new notation is necessary, and how to use it to perform some basic tensor operations. Tensors find application in physics, engineering, and many other sciences, so those possessing a knowledge of and ability to work with tensors are at a distinct advantage in many of these fields. While the Einstein summation notation can be intimidating, and may make tensors seem too advanced in the minds of many students, we hope that removing some of the notational and conceptual mystery will put the topic within reach.

References

[1] W. G. Bickley and R. E. Gibson, *Via Vector to Tensor*, John Wiley and Sons, Inc., New York, 1962.

[2] Brian Greene, *"NOVA: The Elegant Universe"*, 2003, WGBH Educational Foundation `http://www.pbs.org/wgbh/nova/elegant/`.

[3] Joseph C. Kolecki, *An Introduction to Tensors for Students of Physics and Engineering*, Tech. report, NASA Glenn Research Center, Cleveland, September 2002, `http://gltrs.grc.nasa.gov/`.

[4] Wasley S. Krogdahl, *Tensor Analysis: Fundamentals and Applications*, University Press of America, Washington, D.C., 1978.

[5] A. Lichnerowicz, *Elements of Tensor Calculus*, John Wiley and Sons, Inc., New York, 1962.

[6] David McMahon, *Relativity Demystified*, McGraw Hill, New York, 2006.

[7] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler, *Gravitation*, W. H. Freeman and Company, San Francisco, 1973.

[8] Barry Spain, *Tensor Calculus*, Interscience Publishers, Inc., New York, 1960.

[9] J. L. Synge and A. Schild, *Tensor Calculus*, University of Toronto Press, Toronto, 1949.