19.6: Appendix - Tensor Algebra

Tensors

Mathematically scalars and vectors are the first two members of a hierarchy of entities, called **tensors**, that behave under coordinate transformations as described in appendix 19.4. The use of the tensor notation provides a compact and elegant way to handle transformations in physics.

A scalar is a rank 0 tensor with one component, that is invariant under change of the coordinate system.

$$\phi(x'y'z') = \phi(xyz) \tag{19.6.1}$$

A vector is a rank 1 tensor which has three components, that transform under rotation according to matrix relation

$$\mathbf{x}' = \boldsymbol{\lambda} \cdot \mathbf{x} \tag{19.6.2}$$

where $\boldsymbol{\lambda}$ is the rotation matrix. Equation 19.6.2 can be written in the suffix form as

$$x_i' = \sum_{j=1}^{3} \lambda_{ij} x_j \tag{19.6.3}$$

The above definitions of scalars and vectors can be subsumed into a class of entities called tensors of rank n that have 3^n components. A scalar is a tensor of rank $r=0$, with only $3^0 = 1$ component, whereas a vector has rank $r=1$, that is, the vector $\mathbf{x}$ has one suffix i and $3^1 = 3$ components.

A second-order tensor T_{ij} has rank $r=2$ with two suffixes, that is, it has $3^2 = 9$ components that transform under rotation as

$$T_{ij}' = \sum_{k=1}^{3} \sum_{l=1}^{3} \lambda_{ik} \lambda_{jl} T_{kl} \tag{19.6.4}$$

For second-order tensors, the transformation formula given by Equation 19.6.4 can be written more compactly using matrices. Thus the second-order tensor can be written as a 3×3 matrix

$$\mathbf{T} \equiv \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} \tag{19.6.5}$$

The rotational transformation given in Equation 19.6.4 can be written in the form

$$T_{ij}' = \sum_{l=1}^{3} \left(\sum_{k=1}^{3} \lambda_{ik} T_{kl} \right) \lambda_{jl} = \sum_{l=1}^{3} \left(\sum_{k=1}^{3} \lambda_{ik} T_{kl} \right) \lambda_{lj}^T \tag{19.6.6}$$

where λ_{lj}^T are the matrix elements of the transposed matrix $\boldsymbol{\lambda}^T$. The summations in 19.6.6 can be expressed in both the tensor and conventional matrix form as the matrix product

$$\mathbf{T}' = \boldsymbol{\lambda} \cdot \mathbf{T} \cdot \boldsymbol{\lambda}^T \tag{19.6.7}$$

Equation 19.6.7 defines the rotational properties of a spherical tensor.

Tensor products

Tensor outer product

Tensor products feature prominently when using tensors to represent transformations. A second-order tensor $\mathbf{T}$ can be formed by using the **tensor product**, also called **outer product**, of two vectors $\mathbf{a}$ and $\mathbf{b}$ which, written in suffix form, is

$$\mathbf{T} \equiv \mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix} \tag{19.6.8}$$

In component form the matrix elements of this matrix are given by

$$T_{ij} = a_i b_j \tag{19.6.9}$$

This second-order **tensor product** has a rank $r = 2$, that is, it equals the sum of the ranks of the two vectors. Equation 19.6.8 is called a *dyad* since it was derived by taking the dyadic product of two vectors. In general, multiplication, or division, of two vectors leads to second-order tensors. Note that this second-order tensor product completes the triad of tensors possible taking the product of two vectors. That is, the scalar product $\mathbf{a} \cdot \mathbf{b}$, has rank $r = 0$, the vector product $\mathbf{a} \times \mathbf{b}$, rank $r = 1$ and the tensor product $\mathbf{a} \otimes \mathbf{b}$ has rank[1] $r = 2$.

Higher-order tensors can be created by taking more complicated tensor products. For example, a rank-3 tensor can be created by taking the tensor outer product of the rank-2 tensor T_{ij} and a vector c_k which, for a dyadic tensor, can be written as the tensor product of three vectors. That is,

$$T_{ijk} = T_{ij} c_k = a_i b_j c_k \tag{19.6.10}$$

In summary, the rank of the tensor product equals the sum of the ranks of the tensors included in the tensor product.

Tensor Inner Product

The lowest rank tensor product, which is called the **inner product**, is obtained by taking the tensor product of two tensors for the special case where one index is repeated, and taking the sum over this repeated index. Summing over this repeated index, which is called **contraction**, removes the two indices for which the index is repeated, resulting in a tensor that has rank r equal to the sum of the ranks minus 2 for one contraction. That is, the product tensor has rank $r = r_1 + r_2 - 2$.

The simplest example is the inner product of two vectors which has rank $r = 1 + 1 - 2 = 0$, that is, it is the scalar product that equals the trace of the inner product matrix, and this inner product is commutative.

An especially important case is the inner product of a rank-2 dyad $\mathbf{a} \otimes \mathbf{b}$, given by Equation 19.6.8, with a vector $\mathbf{c}$, that is, the inner product $\mathbf{T} = \mathbf{a} \otimes \mathbf{b} \cdot \mathbf{c}$. Written in component form, the inner product is

$$\sum_i^3 a_i b_i c_j = \left(\sum_i^3 a_i b_i \right) c_j = (\mathbf{a} \cdot \mathbf{b}) c_j \tag{19.6.11}$$

The scalar product $\mathbf{a} \cdot \mathbf{b}$ is a scalar number, and thus the inner-product tensor is the vector $\mathbf{c}$ renormalized by the magnitude of the scalar product $\mathbf{a} \cdot \mathbf{b}$. That is, it has a rank $r = 2 + 1 - 2 = 1$. Thus the inner product of this rank-2 tensor with a vector gives a vector. The inner product of a rank-2 tensor with a rank-1 tensor is used in this book for handling the rotation matrix, the inertia tensor for rigid-body rotation, and for the stress and the strain tensors used to describe elasticity in solids.

Example 19.6.1: Displacement gradient tensor

The displacement gradient tensor provides an example of the use of the matrix representation to manipulate tensors. Let $\boldsymbol{\phi}(x_1, x_2, x_3)$ be a vector field expressed in a cartesian basis. The definition of the gradient $G = \boldsymbol{\nabla} \boldsymbol{\phi}$ gives that

$$d\boldsymbol{\phi} = \mathbf{G} \cdot d\mathbf{x}$$

Calculating the components of $d\boldsymbol{\phi}$ in terms of $\mathbf{x}$ gives

$$d\phi_1 = \frac{\partial \phi_1}{\partial x_1} dx_1 + \frac{\partial \phi_1}{\partial x_2} dx_2 + \frac{\partial \phi_1}{\partial x_3} dx_3$$
$$d\phi_2 = \frac{\partial \phi_2}{\partial x_1} dx_1 + \frac{\partial \phi_2}{\partial x_2} dx_2 + \frac{\partial \phi_2}{\partial x_3} dx_3$$
$$d\phi_3 = \frac{\partial \phi_3}{\partial x_1} dx_1 + \frac{\partial \phi_3}{\partial x_2} dx_2 + \frac{\partial \phi_3}{\partial x_3} dx_3$$

Using index notation this can be written as

$$d\phi_i = \frac{\partial \phi_i}{\partial x_j} dx_j$$

The second-rank gradient tensor $\mathbf{G}$ can be represented in the matrix form as

$$\mathbf{G}=\begin{vmatrix} \frac{\partial\phi_1}{\partial x_1} & \frac{\partial\phi_1}{\partial x_2} & \frac{\partial\phi_1}{\partial x_3} \\ \frac{\partial\phi_2}{\partial x_1} & \frac{\partial\phi_2}{\partial x_2} & \frac{\partial\phi_2}{\partial x_3} \\ \frac{\partial\phi_3}{\partial x_1} & \frac{\partial\phi_3}{\partial x_2} & \frac{\partial\phi_3}{\partial x_3} \end{vmatrix}$$

Then the vector $\boldsymbol{\phi}$ can be expressed compactly as the inner product of $\mathbf{G}$ and $\mathbf{x}$, that is

$$d\boldsymbol{\phi}=\mathbf{G}\cdot d\mathbf{x}$$

Tensor Properties

In principle one must distinguish between a 3×3 square matrix, and the tensor component representations of a rank-2 tensor. However, as illustrated by the previous discussion, for orthogonal transformations, the tensor components of the second rank tensor transform identically with the matrix components. Thus functionally, the matrix formulation and tensor representations are identical. As a consequence, all the terminology and operations used in matrix mechanics are equally applicable to the tensor representation.

The tensor representation of the rotation matrix provides the simplest example of the equivalence of the matrix and tensor representations of transformations. Appendix $19.4.2$ showed that the unitary rotation matrix $\boldsymbol{\lambda}$, acting on a vector $\mathbf{x}$ transforms it to the vector $\mathbf{x}'$ that is rotated with respect to $\mathbf{x}$. That is, the transformation is

$$\mathbf{x}'=\boldsymbol{\lambda}\cdot\mathbf{x} \tag{19.6.12}$$

where

$$\mathbf{x}'\equiv\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} \quad \mathbf{x}\equiv\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \boldsymbol{\lambda}\equiv\begin{pmatrix} \hat{\mathbf{e}}_1'\cdot\hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_1'\cdot\hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_1'\cdot\hat{\mathbf{e}}_3 \\ \hat{\mathbf{e}}_2'\cdot\hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2'\cdot\hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_2'\cdot\hat{\mathbf{e}}_3 \\ \hat{\mathbf{e}}_3'\cdot\hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_3'\cdot\hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3'\cdot\hat{\mathbf{e}}_3 \end{pmatrix} \tag{19.6.13}$$

Appendix $19.4.2$ showed that the rotation matrix $\boldsymbol{\lambda}$ requires 9 components to fully specify the transformation from the initial 3-component vector $\mathbf{x}$ to the rotated vector $\mathbf{x}'$. The rotation tensor is a dyad as well as being unitary and dimensionless. Note that Equation 19.6.12 is an example of the inner product of a rank-2 rotation tensor acting on a vector leading to a another vector that is rotated with respect to the first vector.

In general, rank-2 tensors have dimensions and are not unitary. For example, the angular velocity vector $\boldsymbol{\omega}$ and the angular momentum vector $\mathbf{L}$ are related by the inner product of the inertia tensor $\{\mathbf{I}\}$ and $\boldsymbol{\omega}$. That is

$$\mathbf{L}=\{\mathbf{I}\}\cdot\boldsymbol{\omega} \tag{19.6.14}$$

The inertia tensor has dimensions of $mass\times length^2$ and relates two very different vector observables. The stress tensor and the strain tensor, discussed in chapter 15, provide another example of second-order tensors that are used to transform one vector observable to another vector observable analogous to the case of the rotation matrix or the inertia tensor.

Note that pseudo-tensors can be used to make a rotational transformation plus a change in the sign. That is, they lead to a parity inversion.

The tensor notation is used extensively in physics since it provides a powerful, elegant, and compact representation for describing transformations.

Contravariant and covariant tensors

In general the configuration space used to specify a dynamical system is not a Euclidean space in that there may not be a system of coordinates for which the distance between any two neighboring points can be represented by the sum of the squares of the coordinate differentials. For example, a set of cartesian coordinate does not exist for the two-dimension motion of a single particle constrained to the curved surface of a fixed sphere. Such curved spaces need to be represented in terms of Riemannian geometry rather than Euclidean geometry. Curved configuration spaces occur in some branches of physics such as Einstein's General Theory of Relativity.

Tensors have transformation properties that can be either contravariant or covariant. Consider a set of generalized coordinates q' that are a function of the coordinates q. Then infinitessimal changes dq^m will lead to infinitessimal changes dq'^n where

$$dq'^n = \sum_m \frac{\partial q'^n}{\partial q^m} dq^m \tag{19.6.15}$$

Contravariant components of a tensor transform according to the relation

$$\lambda'^n = \sum_m \frac{\partial q'^n}{\partial q^m} \lambda^m \tag{19.6.16}$$

Equation 19.6.16relates the contravariant components in the unprimed and primed frames.

Derivatives of a scalar function ϕ, such as

$$\lambda'_n = \frac{\partial \phi}{\partial q^n} = \sum_m \frac{\partial \phi}{\partial q^m}\frac{\partial q^m}{\partial q^n} = \sum_m \frac{\partial q^m}{\partial q^n}\lambda^m \tag{19.6.17}$$

That is, **covariant** components of the tensor transform according to the relation

$$\lambda'_n = \sum_m \frac{\partial q^m}{\partial q^n}\lambda^m \tag{19.6.18}$$

It is important to differentiate between contravariant and covariant vectors. The superscript/subscript convention for distinguishing between these two flavours of tensors is given in table 19.6.1

Table 19.6.1: Einstein notation for tensors.

x^μ	denotes a contravariant vector
x_ν	denotes a covariant vector

In linear algebra one can map from one coordinate system to another as illustrated in appendix 19.4. That is, the tensor $\mathbf{x}$ can be expressed as components with respect to either the unprimed or primed coordinate frames

$$\mathbf{x} = \hat{\mathbf{e}}'_1 x'_1 + \hat{\mathbf{e}}'_2 x'_2 + \hat{\mathbf{e}}'_3 x'_3 = \hat{\mathbf{e}}_1 x_1 + \hat{\mathbf{e}}_2 x_2 + \hat{\mathbf{e}}_3 x_3 \tag{19.6.19}$$

For a $n-$dimensional manifold the unit basis column vectors $\hat{\mathbf{e}}$ transform according to the transformation matrix $\boldsymbol{\lambda}$

$$\hat{\mathbf{e}}' = \boldsymbol{\lambda} \cdot \hat{\mathbf{e}} \tag{19.6.20}$$

Since the tensor $\mathbf{x}$ is independent of the coordinate basis, the components of $\mathbf{x}$ must have the opposite transform

$$\mathbf{x}' = \left(\boldsymbol{\lambda}^{-1}\right)^T \cdot \mathbf{x} \tag{19.6.21}$$

This normal vector $\mathbf{x}$ is called a “contravariant vector” because it transforms contrary to the basis column vector transformation.

The inverse of Equation 19.6.21gives that the column vector element

$$x_\mu = \sum_\nu \boldsymbol{\lambda}_{\mu\nu} x'_\nu \tag{19.6.22}$$

Consider the case of a gradient with respect to the coordinate $\mathbf{x}$ in both the unprimed and primed bases. Using the chain rule for the partial derivative then the component of the gradient in the primed frame can be expanded as

$$(\nabla f)'_\mu = \frac{\partial f}{\partial x'_\mu} = \sum_\nu \frac{\partial f}{\partial x_\nu}\frac{\partial x_\nu}{\partial x'_\mu} = \sum_\nu \frac{\partial f}{\partial x_\nu}\boldsymbol{\lambda}_{\nu\mu}\delta_{\mu\nu} = \lambda_{\mu\mu}\frac{\partial f}{\partial x_\mu} \tag{19.6.23}$$

That is, the gradient transforms as

$$\boldsymbol{\nabla}' f = \boldsymbol{\lambda} \cdot \boldsymbol{\nabla} f \tag{19.6.24}$$

That is, *a gradient transforms as a covariant vector, like the unit vectors, whereas a vector x is contravariant under transformation.*

Normally the basis is orthonormal, $\left(\boldsymbol{\lambda}^{-1}\right)^{T}=\boldsymbol{\lambda}$, and thus there is no difference between contravariant and covariant vectors. However, for curved coordinate systems, such as non-Euclidean geometry in the General Theory of Relativity, the covariant and contravariant vectors behave differently.

The Einstein convention is extended to apply to matrices by writing the elements of the matrix $\mathbf{A}$ as A_{ν}^{μ} while the elements of the transposed matrix $\mathbf{A}^{-1}$ are written as A_{μ}^{ν}. The matrix product for $\mathbf{A}$ with a contravariant vector $\mathbf{X}$ is written as

$$X'^{\mu}=\sum_{\nu} A_{\nu}^{\mu} X^{\nu} \tag{19.6.25}$$

where the summation over ν effectively cancels the identical superscript and subscript ν.

Similarly a covariant vector, such as a gradient, is written as,

$$(\boldsymbol{\nabla}' f)_{\mu}=\sum_{\nu}\left(A^{-1}\right)_{\mu}^{T\nu}(\boldsymbol{\nabla} f)_{\nu}=\sum_{\nu}\left(A^{-1}\right)_{\mu}^{\nu}(\boldsymbol{\nabla} f)_{\nu} \tag{19.6.26}$$

Again the summation cancels the ν superscript and subscript. The Kronecker delta symbol is written as

$$\sum_{\nu} \delta_{\nu}^{\mu} X^{\nu}=X^{\mu} \tag{19.6.27}$$

Generalized inner product

The generalized definition of an *inner product* is

$$S=\sum_{\mu\nu} g_{\mu\nu} X^{\mu} Y^{\nu} \tag{19.6.28}$$

where $g_{\mu\nu}$ is a unitary matrix called a covariant metric. The covariant metric transforms a contravariant to a covariant tensor. For example the matrix element of a covariant tensor X_{ν} can be written as

$$X_{\nu}=\sum_{\mu} g_{\mu\nu} X^{\mu} \tag{19.6.29}$$

By association of the *covariant metric* with either of the vectors in the inner product gives

$$S=\sum_{\mu\nu} g_{\mu\nu} X^{\mu} Y^{\nu}=\sum_{\nu} X_{\nu} Y^{\nu}=\sum_{\mu} X^{\mu} Y_{\mu} \tag{19.6.30}$$

Similarly it can be defined in terms of an *orthogonal contravariant metric* $g^{\mu\nu}$ where

$$S=\sum_{\mu\nu} g^{\mu\nu} X_{\mu} Y_{\nu} \tag{19.6.31}$$

Then

$$X^{\nu}=\sum_{\mu} g^{\mu\nu} X_{\mu} \tag{19.6.32}$$

Association of the contravariant metric with one of the vectors in the inner product gives the inner product

$$S=\sum_{\mu\nu} g^{\mu\nu} X_{\mu} Y_{\nu}=\sum_{\nu} X^{\nu} Y_{\nu}=\sum_{\mu} X_{\mu} Y^{\mu} \tag{19.6.33}$$

For most situations in this book the metric $g_{\mu\nu}$ is diagonal and unitary.

Transformation Properties of Observables

In physics, observables can be represented by spherical tensors which specify the angular momentum and parity characteristics of the observable, and the tensor rank is independent of the time dependence. The transformation properties of these tensors, coupled with their time-reversal invariance, specify the fundamental characteristics of the observables.

Table $19.6.2$ summarizes the transformation properties under rotation, spatial inversion and time reversal for observables encountered in classical mechanics and electrodynamics. Note that observables can be scalar, vector, pseudovector, or second-order tensors, under rotation, and even or odd under either space inversion or time inversion. For example, in classical mechanics the inertia tensor $\mathbf{I}$ relates the angular velocity vector $\boldsymbol{\omega}$ to the angular momentum vector $\mathbf{L}$ by taking the inner product $\mathbf{L}=\mathbf{I}\cdot\boldsymbol{\omega}$. In general $\mathbf{I}$ is not diagonal and thus the angular momentum is not parallel to the angular velocity $\boldsymbol{\omega}$. A similar example in electrodynamics is the dielectric tensor $\mathbf{K}$ which relates the displacement field $\mathbf{D}$ to the electric field $\mathbf{E}$ by $\mathbf{D}=\mathbf{K}\cdot\mathbf{E}$. For anisotropic crystal media $\mathbf{K}$ is not diagonal leading to the electric field vectors $\mathbf{E}$ and $\mathbf{D}$ not being parallel.

As discussed in chapter 7, Noether's Theorem states that symmetries of the transformation properties lead to important conservation laws. The behavior of classical systems under rotation relates to the conservation of angular momentum, the behavior under spatial inversion relates to parity conservation, and time-reversal invariance relates to conservation of energy. That is, conservative forces conserve energy and are time-reversal invariant.

Table 19.6.2: Transformation properties of scalar, vector, pseudovector, and tensor observables under rotation, spatial inversion, and time reversal[2]

Physical Observable		Rotation (Tensor rank)	Space inversion	Time reversal	Name
1) Classical Mechanics					
Mass density	ρ	0	Even	Even	Scalar
Kinetic energy	$p^2/2m$	0	Even	Even	Scalar
Potential energy	$U(r)$	0	Even	Even	Scalar
Lagrangian	L	0	Even	Even	Scalar
Hamiltonian	H	0	Even	Even	Scalar
Gravitational potential	ϕ	0	Even	Even	Scalar
Coordinate	$\mathbf{r}$	1	Odd	Even	Vector
Velocity	$\mathbf{v}$	1	Odd	Odd	Vector
Momentum	$\mathbf{p}$	1	Odd	Odd	Vector
Angular momentum	$\mathbf{L}=\mathbf{r}\times\mathbf{p}$	1	Even	Odd	Pseudovector
Force	$\mathbf{F}$	1	Odd	Even	Vector
Torque	$\mathbf{N}=\mathbf{r}\times\mathbf{F}$	1	Even	Even	Pseudovector
Gravitational field	$\mathbf{g}$	1	Odd	Even	Vector
Inertia tensor	$\mathbf{I}$	2	Even	Even	Tensor
Elasticity stress tensor	$\mathbf{T}_{ik}$	2	Even	Even	Tensor
2) Electromagnetism					
Charge density	ρ	0	Even	Even	Scalar
Current density	$\mathbf{j}$	1	Odd	Odd	Vector
Electric field	$\mathbf{E}$	1	Odd	Even	Vector
Polarization	$\mathbf{P}$	1	Odd	Even	Vector
Displacement	$\mathbf{D}$	1	Odd	Even	Vector
Magnetic B field	$\mathbf{B}$	1	Even	Odd	Pseudovector

Physical Observable		Rotation (Tensor rank)	Space inversion	Time reversal	Name
Magnetization	$\mathbf{M}$	1	Even	Odd	Pseudovector
Magnetic H field	$\mathbf{H}$	1	Even	Odd	Pseudovector
Poynting vector	$\mathbf{S}=\mathbf{E}\times\mathbf{H}$	1	Odd	Odd	Vector
Dielectric tensor	$\mathbf{K}$	2	Even	Even	Tensor
Maxwell stress tensor	$\mathbf{T}_{ik}$	2	Even	Even	Tensor

References

[1]The common convention is to denote the scalar product as $\mathbf{a}\cdot\mathbf{b}$, the vector product as $\mathbf{a}\times\mathbf{b}$, and tensor product as $\mathbf{a}\otimes\mathbf{b}$.

[2]Based on table 6.1 in *"Classical Electrodynamics"* 2^{nd} edition, by J.D. Jackson [Jac75]