6. Spaces, Bases, Vectors

A vector space $\mathcal{V}$ over a field $\mathcal{S}$ of elements, called scalars, is a set of elements $\vec{\mathbf{v}}$ called vectors together with two binary operations that obey the usual rules for vector addition and for multiplication of vectors by scalars. A vector space may have other structure, such as a norm or an inner product.

6.1 Change of Basis

Consider a vector space $\mathcal{V}$ of dimension n over scalars $\mathcal{S}$, and let

$$\mathbf{f} = (\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_n) \quad \text{and} \quad \mathbf{f}' = (\vec{\mathbf{k}}_1, \dots, \vec{\mathbf{k}}_n) \tag{1}$$

each be a basis of $\mathcal{V}$, the $\vec{\mathbf{e}}_i$ and $\vec{\mathbf{k}}_i$ being basis vectors. We shall regard f and k as row vectors of matrix theory, but it is their elements, $\vec{\mathbf{e}}_i$ and $\vec{\mathbf{k}}_i$, that are vectors of the vector space $\mathcal{V}$.

A change of basis from f to f′ is described by

$$\mathbf{f} \mapsto \mathbf{f}' = (\textstyle\sum_i a_{i1}\vec{\mathbf{e}}_i,\ \dots,\ \sum_i a_{in}\vec{\mathbf{e}}_i) = \mathbf{f}\,\mathbf{A} \tag{2}$$

for some non-singular (invertible) $n \times n$ matrix A with elements a_{ij}. Sums are over 1 to n.

That is, each vector $\vec{\mathbf{k}}_j$ of the f′ basis is a linear combination of the vectors $\vec{\mathrm{e}}_i$ of the f basis:

$$\vec{\mathbf{k}}_j = \textstyle\sum_i a_{ij}\vec{\mathbf{e}}_i \,. \tag{3}$$

Any vector $\vec{\mathrm{v}}$ in $\mathcal{V}$ is expressed uniquely as a linear combination of the elements of the f basis as

$$\vec{\mathbf{v}} = \textstyle\sum_i v^i[\mathbf{f}]\,\vec{\mathbf{e}}_i \,. \tag{4}$$

The v^i[f] are scalars in $\mathcal{S}$ called the components of $\vec{\mathrm{v}}$ in the f basis; they can be formed into a column vector of matrix theory:

$$\mathbf{v}[\mathbf{f}] = \mathrm{col}\{\, v^1[\mathbf{f}]\,,\, v^2[\mathbf{f}]\,,\, \ldots\,,\, v^n[\mathbf{f}]\,\}\,. \tag{5}$$

So the f-basis expansion (4) can be written as a matrix product:

$$\vec{\mathbf{v}} = \mathbf{f}\,\mathbf{v}[\mathbf{f}]\,. \tag{6}$$

The vector $\vec{\mathrm{v}}$ may also be expressed in the f′ basis: the relation

$$\mathbf{f}\,\mathbf{v}[\mathbf{f}] = \vec{\mathbf{v}} = \mathbf{f}'\,\mathbf{v}[\mathbf{f}'] \tag{7}$$

expresses invariance of a vector under choice of basis.

6.2 Contravariance and Covariance

A. Vectors

Invariance of $\vec{v}$ plus the the relation $\mathbf{f}' = \mathbf{f}\,\mathbf{A}$ between bases imply that $\mathbf{f}\,\mathbf{v}[\mathbf{f}] = \mathbf{f}\,\mathbf{A}\,\mathbf{v}[\mathbf{f}\,\mathbf{A}]$. This gives, via matrix multiplication:

$$\mathbf{v}[\mathbf{f}\,\mathbf{A}] = \mathbf{A}^{-1}\mathbf{v}[\mathbf{f}] \tag{8}$$

or

$$v^i[\mathbf{f}\,\mathbf{A}] = \textstyle\sum_j \tilde{a}^{ij} v^j\,[\mathbf{f}]\,, \tag{9}$$

with $\tilde{a}^{ij}$ denoting the elements of $\mathbf{A}^{-1}$. This is the vector component transformation rule, expressing each component of $\vec{v}$ in the new basis as a homogeneous linear combination of its components in the original basis.

We note that:

- **The components of a vector transform with the inverse of the matrix that transforms the basis elements: vector components are said to transform contravariantly under a basis change.**

B. Covectors

A covector or linear functional (meaning, roughly, a function of a function – see below) α on $\mathcal{V}$ is expressed uniquely in terms of its components (scalars in $\mathcal{S}$) in the f basis as

$$\alpha_i[\mathbf{f}] = \alpha(\vec{\mathbf{e}}_i), \quad i = 1, 2, \ldots, n\,. \tag{10}$$

These represent the action of α on the basis vectors $\vec{\mathbf{e}}_i$ of f.

Under the basis change from f to $\mathbf{f}' = \mathbf{f}\,\mathbf{A}$, the α_i transform via

$$\alpha_i[\mathbf{f}\,\mathbf{A}] = \alpha(\vec{\mathbf{k}}_i) = \alpha\left(\sum_k a_{ki}\vec{\mathbf{e}}_k\right) = \sum_k a_{ki}\alpha(\vec{\mathbf{e}}_k)$$
$$= \sum_j a_{ki}\alpha_k[\mathbf{f}]\,; \tag{11}$$

we have used: Eq (10) in the f′ basis, the expansion of $\vec{\mathbf{k}}_i$ in the f basis, the linear property of α, and Eq (10) in reverse.

Let's denote the row vector of components of α by:

$$\alpha[\mathbf{f}] = (\,\alpha_1[\mathbf{f}]\,, \alpha_2[\mathbf{f}]\,, \ldots\,, \alpha_n[\mathbf{f}]\,)\,. \tag{12}$$

Eq (11) can be written as a matrix product:

$$\alpha[\mathbf{f}\,\mathbf{A}] = \alpha[\mathbf{f}]\,\mathbf{A}\,. \tag{13}$$

If a column vector representation were used instead, the transformation law would be the transpose:

$$\alpha^{\mathrm{T}}[\mathbf{f}\,\mathbf{A}] = \mathbf{A}^{\mathrm{T}}\alpha^{\mathrm{T}}[\mathbf{f}]\,. \tag{14}$$

We note that:

- **The components of a linear functional transform with the matrix that transforms the basis elements: the functional's components are said to transform covariantly under a basis change.**

§6.1 & §6.2 have been adapted (Apr 2012) from http://en.wikipedia.org/wiki/ Covariance and contravariance of vectors

Note: The terms 'covariant' and 'contravariant' were introduced by J.J. Sylvester in 1853 in algebraic invariant theory. A set of simultaneous equations is said to be contravariant in the variables.

6.3 Some Mathematical Definitions

(1) A functional is a map from a vector space to its field of scalars: a functional acts on a vector as its argument and produces a scalar. If the vector space is a space of functions, a functional acts on a function to produce a scalar: it is then a functional in the older sense – a function of a function. Example: a definite integral of a function is a functional.

(2) A linear map between two vector spaces preserves the operations of vector addition and multiplication of a vector by a scalar:

$$f(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = f(\vec{\mathbf{v}}) + f(\vec{\mathbf{w}}) \quad \text{for all} \;\; \vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathcal{V}$$

$$f(a\vec{\mathbf{v}}) = af(\vec{\mathbf{v}}) \quad \text{for all} \;\; \vec{\mathbf{v}} \in \mathcal{V}, \; a \in \mathcal{S}.$$

(3) A linear functional, also called a covector or one-form, is a linear map from a vector space to its field of scalars.

(4) For any positive integer n, the set of all n-tuples of real numbers forms an n-dimensional vector space, $\mathcal{R}^n$, called real coordinate space. An element is written $\vec{\mathbf{x}} = (x_1, x_2, \ldots, x_n)$ with each x_i a real number.

The vector space operations on $\mathcal{R}^n$ are defined by

$$\vec{\mathbf{x}} + \vec{\mathbf{y}} = (x_1 + y_1 \,, \, \ldots \,, x_n + y_n) \,,$$
$$a \, \vec{\mathbf{x}} = (ax_1 \,, \, \ldots \,, ax_n) \,.$$

The standard basis is:

$$\vec{\mathbf{e}}_1 = (1, 0, \ldots, 0) \,, \quad \vec{\mathbf{e}}_2 = (0, 1, \ldots, 0) \,,$$
$$\ldots \,, \quad \vec{\mathbf{e}}_n = (0, 0, \ldots, 1) \,,$$

so an arbitrary vector in $\mathcal{R}^n$ can be written as $\vec{\mathbf{x}} = \Sigma_{i=1}^{n} x_i \vec{\mathbf{e}}_i$.

If vectors in $\mathcal{R}^n$ are represented as column vectors, then linear functionals are represented as row vectors, and their action on vectors is represented by a matrix product with the row vector on the left and the column vector on the right.

$\mathcal{R}^n$ **is the prototype of a real n-dimensional vector space. Euclidean space has the additional structure of the inner (or dot) product, leading to the pythagorean metric.**

7. Coordinates

We begin by noting Eqs (4) & (9) of §6 above:

$$\vec{\mathbf{v}} = \textstyle\sum_i v^i[\mathbf{f}]\,\mathbf{e}_i\,, \tag{1}$$

– expressing an arbitrary vector $\vec{\mathbf{v}}$ in terms of its components $v^i[\mathrm{f}]$ in the f basis – and

$$v^i[\mathbf{f\,A}] = \textstyle\sum_j \tilde{a}^{ij} v^j\,[\mathbf{f}] \tag{2}$$

for the vector component transformation rule, with the $\tilde{a}^{ij}$ denoting the elements of $\mathbf{A}^{-1}$, the inverse basis transformation matrix.

7.1 Introducing Coordinates

Equation (1) shows that choice of basis f on $\mathcal{V}$ defines a set of coordinate functions by means of

$$x^i\,[\mathbf{f}](\vec{\mathbf{v}}) = v^i[\mathbf{f}] \tag{3}$$

– the functions $x^i\,[\mathrm{f}]$ act on any $\vec{\mathbf{v}}$ to generate its components $v^i[\mathbf{f}]$.

Equation (2) shows that coordinate functions ("coordinates") on $\mathcal{V}$ are contravariant in that

$$x^i\,[\mathbf{f\,A}] = \textstyle\sum_{j=1}^{n} \tilde{a}^{ij} x^j[\mathbf{f}]\,. \tag{4}$$

Conversely, a set of n quantities v^i that transform like the coordinates x^i on $\mathcal{V}$ forms a contravariant vector.

A set of n quantities that transform oppositely to the coordinates forms a covariant vector. Thus we interpret contravariance and covariance in terms of coordinate changes rather than basis changes as in §6 above.

This coordinate approach to contravariance / covariance is widely used in physical applications set on a manifold (roughly: a space that, at every point, is locally like Euclidean space). Given a local coordinate system x^k on a manifold, the reference axes

$$\vec{\mathbf{e}}_1 = \frac{\partial}{\partial x^1}, \quad \ldots, \quad \vec{\mathbf{e}}_n = \frac{\partial}{\partial x^n} \tag{5}$$

provide a frame f $= (\vec{\mathbf{e}}_1, \ldots, \vec{\mathbf{e}}_n)$ at all points of a local coordinate system (coordinate patch).

For a different coordinate system y^k, the

$$\vec{\mathbf{k}}_1 = \frac{\partial}{\partial y^1}, \quad \dots, \quad \vec{\mathbf{k}}_n = \frac{\partial}{\partial y^n} \tag{6}$$

form a frame f′ related to f by the inverse of the Jacobian matrix of the coordinate transformation:

$$\mathbf{f}' = \mathbf{f} J^{-1}, \qquad J = \left(\frac{\partial y^i}{\partial x^j}\right)_{i,j=1}^{n}. \tag{7}$$

In index form:

$$\frac{\partial}{\partial y^i} = \sum_{j=1}^{n} \frac{\partial x^j}{\partial y^i} \frac{\partial}{\partial x^j}. \tag{8}$$

A tangent vector is, by definition, a vector that is a linear combination of the coordinate partial derivatives $\partial/\partial x^i$:

$$\vec{\mathbf{v}} = \textstyle\sum_{i=1}^{n} v^i[\mathbf{f}]\vec{\mathbf{e}}_i = \mathbf{f}\ \mathbf{v}[\mathbf{f}]. \tag{9}$$

So tangent vectors are contravariant under a frame change. Let's now look at tangent vectors under coordinate changes.

Under changes in the coordinate system, one has

$$\mathbf{v}[\mathbf{f}'] = \mathbf{v}[\mathbf{f}\,J^{-1}] = J\,\mathbf{v}[\mathbf{f}]\,. \tag{10}$$

That is, the components of a tangent vector transform via

$$v^i[\mathbf{f}'] = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j} v^j[\mathbf{f}]\,. \tag{11}$$

A set of n quantities v^i that transform in this way between coordinate systems forms a contravariant vector.

7.2 Euclidean Space: Covariant-Contravariant Link

In a Euclidean space $\mathcal{E}$, the distinction between covariant and contravariant vectors is blurred, because the inner (dot) product allows covariant vectors (covectors) to be identified with vectors.

A vector $\vec{\mathbf{v}}$ determines uniquely a covector α via

$$\alpha(\vec{\mathbf{w}}) = \vec{\mathbf{v}} \cdot \vec{\mathbf{w}}\,, \quad \textbf{for all vectors } \vec{\mathbf{w}}\,, \tag{12}$$

and each covector α determines a unique vector $\vec{\mathbf{v}}$ by this equation. So we see that:

- **In Euclidean space, vectors have covariant components and contravariant components: they are two representations of the same vector using reciprocal bases.**

For any basis $\mathbf{f} = (\vec{\mathbf{e}}_1, \ldots, \vec{\mathbf{e}}_n)$ of $\mathcal{E}$, there is a unique reciprocal basis $\mathbf{f}^\star = (\vec{\varepsilon}^{\,1}, \ldots, \vec{\varepsilon}^{\,n})$ of $\mathcal{E}$, determined by

$$\vec{\varepsilon}^{\,i} \cdot \vec{\mathbf{e}}_j = \delta^i_j\,. \tag{13}$$

Any vector $\vec{\mathbf{v}}$ can be written in terms of either basis:

$$\vec{\mathbf{v}} = \textstyle\sum_i v^i[\mathbf{f}]\vec{\mathbf{e}}_i = \mathbf{f}\,\mathbf{v}[\mathbf{f}] = \textstyle\sum_i v_i[\mathbf{f}]\vec{\varepsilon}^{\,i} = \mathbf{f}^{\star}\mathbf{v}^{\star}[\mathbf{f}]\,; \qquad (14)$$

we have used the component and matrix forms [Eqs (4) & (6) of Section §6 above] of the expansion of $\vec{\mathbf{v}}$ in both bases.

The $v^i[\mathbf{f}]$ are the contravariant components of $\vec{\mathbf{v}}$ in the basis f and the components $v_i[\mathbf{f}]$ are the covariant components of $\vec{\mathbf{v}}$ in the basis f. This terminology reflects the behaviour under any change of basis. For a basis transformation matrix A:

$$\mathbf{v}[\mathbf{f}\,\mathbf{A}] = \mathbf{A}^{-1}\mathbf{v}[\mathbf{f}]\,, \qquad \mathbf{v}^{\star}[\mathbf{f}\,\mathbf{A}] = \mathbf{A}^{\mathrm{T}}\mathbf{v}^{\star}[\mathbf{f}]\,. \qquad (15)$$

§7.1 & §7.2 have been adapted (Apr 2012) from http://en.wikipedia.org/wiki/ Covariance and contravariance of vectors

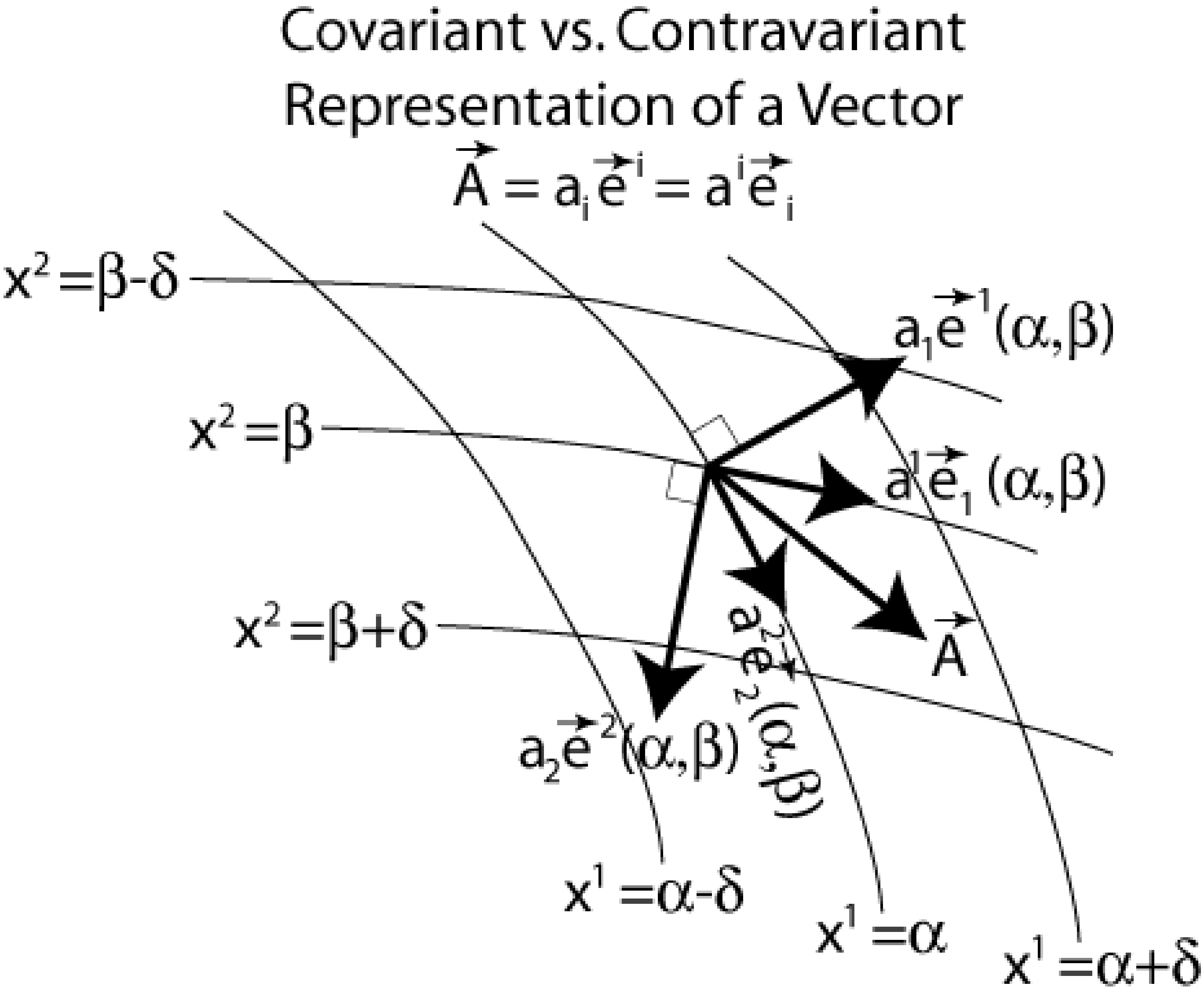

Figure by anon.

The terms 'covariance' and 'contravariance' describe how the algebraic description of certain geometric or physical entities changes under a change of basis from one coordinate system to another.

When one coordinate system is a rotation of the other, the covariant-contravariant distinction disappears. When more general coordinate systems, such as non-orthogonal and curvilinear coordinates, are used the distinction is apparent. Covariant vectors generally arise when we take a gradient of a function – in a sense dividing by a vector.

7.3 Dual Vector Spaces

(i) Any vector space $\mathcal{V}$ over a field $\mathcal{F}$, has a dual space, $\mathcal{V}^*$, over the same field, defined as the set of all linear functionals (linear maps) $\phi : \mathcal{V} \to \mathcal{F}$. Elements of the dual space are called covectors or one-forms.

(ii) The dual space becomes a dual vector space when equipped with the usual linear operations of addition of vectors and multiplication of vectors by scalars.

Dual vector spaces defined on finite-dimensional vector spaces are used in defining tensors in a *coordinate-free* fashion. For vector spaces of functions (typically of infinite dimensions, such as Hilbert space), dual spaces are used in defining measure.

For a (true) vector – such as the position of a point relative to an origin – to be invariant, its components must "contra-vary" with a change of basis to compensate: the components must vary via the inverse transformation of the basis change. So vectors (as opposed to dual vectors) are said to be *contravariant*.

For a dual vector, such as the gradient of a scalar, to be invariant, its components must "co-vary" with a change of basis to maintain the dual vector (e.g. gradient) nature: the components must vary by the same transformation as the change of basis. So dual vectors (as opposed to vectors) are said to be *covariant*.

These last two sections of these notes on tensors, §6 and §7, are intended to provide an initial link to the coordinate-free and index-free approach to tensors, as used widely in mathematics courses and in some advanced physics courses.

Key Points – Vector Space Formalism

- A vector space $\mathcal{V}$ over a field $\mathcal{S}$ of elements, called scalars, is a set of elements $\vec{v}$ called vectors together with the binary operations of vector addition and multiplication of vectors by scalars.

- The components of a vector transform with the inverse of the matrix that transforms the basis elements: vector components are said to transform contravariantly under a basis change.

- The components of a linear functional transform with the matrix that transforms the basis elements: the functional's components are said to transform covariantly under a basis change.

- In Euclidean space, vectors have covariant components and contravariant components: they are two representations of the same vector using reciprocal bases.

- A functional is a map from a vector space $\mathcal{V}$ to its field of scalars, acting on a vector to produce a scalar. If $\mathcal{V}$ is a space of functions, a functional is a function of a function.

- **A linear map between two vector spaces preserves the operations of vector addition and multiplication of a vector by a scalar. A linear functional, also called a covector or one-form, is a linear map from a vector space to its field of scalars.**

- **For a manifold, as in physical applications: given a local coordinate system x^k (coordinate patch), the reference axes $\vec{\mathrm{e}}_1 = \partial/\partial x^1 \ldots$ form a frame $\mathrm{f} = (\vec{\mathrm{e}}_1, \ldots)$ at all points of the patch.**

- **Any vector space $\mathcal{V}$ over a field $\mathcal{F}$, has a dual space, $\mathcal{V}^*$, over the same field, defined as the set of all linear functionals (linear maps) $\phi : \mathcal{V} \to \mathcal{F}$. Elements of the dual space are called one-forms or covectors. It becomes a dual vector space when equipped with the operations of addition of vectors and multiplication of vectors by scalars. Dual vector spaces defined on finite-dimensional vector spaces are used in defining tensors in a *coordinate-free* fashion.**

- **For a (true) vector – such as the position of a point relative to an origin – to be invariant, its components must "contra-vary" with a change of basis to compensate: the components must vary via the inverse transformation of the basis change. So vectors (as opposed to dual vectors) are said to be *contravariant*.**

- **For a dual vector, such as the gradient of a scalar, to be invariant, its components must "co-vary" with a change of basis to maintain the dual vector (*e.g.* gradient) nature: the components must vary by the same transformation as the change of basis. So dual vectors (as opposed to vectors) are said to be *covariant*.**

Bibliography

Billings, A. R., *Tensor Properties of Materials: Generalized Compliance and Conductivity* **(Wiley-Interscience, London & New York, 1969).**

Holt, E.H, & Haskell, R.E., *Foundations of Plasma Dynamics* **(Macmillan, New York, 1965); Chapter 2: The notation of Cartesian tensors (pp 25–45); Appendix B: Transformation properties of Cartesian tensors (pp 478–81).†**

Jeffreys, H., *Cartesian Tensors* **(Cambridge University Press, 1931).**

Rindler, W., *Introduction to Special Relativity* **(Clarendon Press, Oxford, 1st ed. 1982, 2nd ed. 1991); Appendix: Tensors for special relativity (pp 177–89, 1st ed.).†**

Svozil, K., *Mathematical Methods of Theoretical Physics* **(Edition Funzl, 1st ed. 2011, 2nd ed. 2013); available from arXiv:1203.4558v3 [math-ph] 15 Feb 2013. See: Chapter 4, Finite-dimensional vector spaces, pp 3383; Chapter 5, Tensors, pp 85111.**

Van Bladel, J.G., *Electromagnetic Fields* **(IEEE Press and Wiley- Interscience, Hobokken, 2nd ed. 2007); Appendix B: Dyadic analysis (pp 213–17).†**

† Recommended reading.

From Svozil's *Introduction* (p 15):

"I kindly ask the perplexed to please be patient, do not panic under any circumstances, and do not allow themselves to be too upset with mistakes, omissions & other problems of this text. At the end of the day, everything will be fine, and in the long run we will be dead anyway."

End of tensor notes.