Chapter 10

Vectors and Tensors

In this chapter we explain how a vector space V gives rise to a family of associated tensor spaces, and how mathematical objects such as linear maps or quadratic forms should be understood as being elements of these spaces. We then apply these ideas to physics. We make extensive use of notions and notations from the appendix on linear algebra, so it may help to review that material before we begin.

10.1 Covariant and contravariant vectors

When we have a vector space V over $\mathbb{R}$, and $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ are both bases for V, then we may expand each of the basis vectors $\mathbf{e}_\mu$ in terms of the $\mathbf{e}'_\mu$ as

$$\mathbf{e}_\nu = a^\mu_\nu \mathbf{e}'_\mu. \tag{10.1}$$

We are here, as usual, using the Einstein summation convention that repeated indices are to be summed over. Written out in full for a three-dimensional space, the expansion would be

$$\begin{aligned}
\mathbf{e}_1 &= a^1_1\mathbf{e}'_1 + a^2_1\mathbf{e}'_2 + a^3_1\mathbf{e}'_3,\\
\mathbf{e}_2 &= a^1_2\mathbf{e}'_1 + a^2_2\mathbf{e}'_2 + a^3_2\mathbf{e}'_3,\\
\mathbf{e}_3 &= a^1_3\mathbf{e}'_1 + a^2_3\mathbf{e}'_2 + a^3_3\mathbf{e}'_3.
\end{aligned}$$

We could also have expanded the $\mathbf{e}'_\mu$ in terms of the $\mathbf{e}_\mu$ as

$$\mathbf{e}'_\nu = (a^{-1})^\mu_\nu \mathbf{e}'_\mu. \tag{10.2}$$

As the notation implies, the matrices of coefficients a^μ_ν and $(a^{-1})^\mu_\nu$ are inverses of each other:

$$a^\mu_\nu (a^{-1})^\nu_\sigma = (a^{-1})^\mu_\nu a^\nu_\sigma = \delta^\mu_\sigma. \tag{10.3}$$

If we know the components x^μ of a vector $\mathbf{x}$ in the $\mathbf{e}_\mu$ basis then the components x'^μ of $\mathbf{x}$ in the $\mathbf{e}'_\mu$ basis are obtained from

$$\mathbf{x} = x'^\mu \mathbf{e}'_\mu = x^\nu \mathbf{e}_\nu = (x^\nu a^\mu_\nu)\,\mathbf{e}'_\mu \tag{10.4}$$

by comparing the coefficients of $\mathbf{e}'_\mu$. We find that $x'^\mu = a^\mu_\nu x^\nu$. Observe how the $\mathbf{e}_\mu$ and the x^μ transform in "opposite" directions. The components x^μ are therefore said to transform *contravariantly*.

Associated with the vector space V is its *dual space* V^*, whose elements are *covectors*, *i.e.* linear maps $\mathbf{f} : V \to \mathbb{R}$. If $\mathbf{f} \in V^*$ and $\mathbf{x} = x^\mu \mathbf{e}_\mu$, we use the linearity property to evaluate $\mathbf{f}(\mathbf{x})$ as

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x^\mu \mathbf{e}_\mu) = x^\mu \mathbf{f}(\mathbf{e}_\mu) = x^\mu\, f_\mu. \tag{10.5}$$

Here, the set of numbers $f_\mu = \mathbf{f}(\mathbf{e}_\mu)$ are the components of the covector $\mathbf{f}$. If we change basis so that $\mathbf{e}_\nu = a^\mu_\nu \mathbf{e}'_\mu$ then

$$f_\nu = \mathbf{f}(\mathbf{e}_\nu) = \mathbf{f}(a^\mu_\nu \mathbf{e}'_\mu) = a^\mu_\nu \mathbf{f}(\mathbf{e}'_\mu) = a^\mu_\nu f'_\mu. \tag{10.6}$$

We conclude that $f_\nu = a^\mu_\nu f'_\mu$. The f_μ components transform in the same manner as the basis. They are therefore said to transform *covariantly*. In physics it is traditional to call the the set of numbers x^μ with upstairs indices (the components of) a *contravariant vector*. Similarly, the set of numbers f_μ with downstairs indices is called (the components of) a *covariant vector*. Thus, contravariant vectors are elements of V and covariant vectors are elements of V^*.

The relationship between V and V^* is one of mutual duality, and to mathematicians it is only a matter of convenience which space is V and which space is V^*. The evaluation of $\mathbf{f} \in V^*$ on $\mathbf{x} \in V$ is therefore often written as a "pairing" $(\mathbf{f}, \mathbf{x})$, which gives equal status to the objects being put togther to get a number. A physics example of such a mutually dual pair is provided by the space of displacements $\mathbf{x}$ and the space of wave-numbers $\mathbf{k}$. The units of $\mathbf{x}$ and $\mathbf{k}$ are different (meters *versus* meters^{-1}). There is therefore no meaning to "$\mathbf{x} + \mathbf{k}$," and $\mathbf{x}$ and $\mathbf{k}$ are not elements of the same vector space. The "dot" in expressions such as

$$\psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}} \tag{10.7}$$

cannot be a true inner product (which requires the objects it links to be in the same vector space) but is instead a pairing

$$(\mathbf{k},\mathbf{x}) \equiv \mathbf{k}(\mathbf{x}) = k_\mu x^\mu. \tag{10.8}$$

In describing the physical world we usually give priority to the space in which we live, breathe and move, and so treat it as being "V". The displacement vector $\mathbf{x}$ then becomes the contravariant vector, and the Fourier-space wave-number $\mathbf{k}$, being the more abstract quantity, becomes the covariant covector.

Our vector space may come equipped with a *metric* that is derived from a non-degenerate inner product. We regard the inner product as being a bilinear form $\mathbf{g} : V \times V \to \mathbb{R}$, so the length $\|\mathbf{x}\|$ of a vector $\mathbf{x}$ is $\sqrt{\mathbf{g}(\mathbf{x},\mathbf{x})}$. The set of numbers

$$g_{\mu\nu} = \mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) \tag{10.9}$$

comprises the (components of) the *metric tensor*. In terms of them, the inner of product $\langle \mathbf{x},\mathbf{y}\rangle$ of pair of vectors $\mathbf{x} = x^\mu \mathbf{e}_\mu$ and $\mathbf{y} = y^\mu \mathbf{e}_\mu$ becomes

$$\langle \mathbf{x},\mathbf{y}\rangle \equiv \mathbf{g}(\mathbf{x},\mathbf{y}) = g_{\mu\nu} x^\mu y^\nu. \tag{10.10}$$

Real-valued inner products are always symmetric, so $\mathbf{g}(\mathbf{x},\mathbf{y}) = \mathbf{g}(\mathbf{y},\mathbf{x})$ and $g_{\mu\nu} = g_{\nu\mu}$. As the product is non-degenerate, the matrix $g_{\mu\nu}$ has an inverse, which is traditionally written as $g^{\mu\nu}$. Thus

$$g_{\mu\nu} g^{\nu\lambda} = g^{\lambda\nu} g_{\nu\mu} = \delta^\lambda_\mu. \tag{10.11}$$

The additional structure provided by the metric permits us to identify V with V^*. The identification is possible, because, given any $\mathbf{f} \in V^*$, we can find a vector $\widetilde{\mathbf{f}} \in V$ such that

$$\mathbf{f}(\mathbf{x}) = \langle \widetilde{\mathbf{f}}, \mathbf{x}\rangle. \tag{10.12}$$

We obtain $\widetilde{\mathbf{f}}$ by solving the equation

$$f_\mu = g_{\mu\nu} \widetilde{f}^\nu \tag{10.13}$$

to get $\widetilde{f}^\nu = g^{\nu\mu} f_\mu$. We may now drop the tilde and identify $\mathbf{f}$ with $\widetilde{\mathbf{f}}$, and hence V with V^*. When we do this, we say that the covariant components f_μ are related to the contravariant components f^μ by *raising*

$$f^\mu = g^{\mu\nu} f_\nu, \tag{10.14}$$

or *lowering*

$$f_\mu = g_{\mu\nu} f^\nu, \tag{10.15}$$

the index μ using the metric tensor. Bear in mind that this $V \cong V^*$ identification depends crucially on the metric. A different metric will, in general, identify an $\mathbf{f} \in V^*$ with a completely different $\widetilde{\mathbf{f}} \in V$.

We may play this game in the Euclidean space $\mathbb{E}^n$ with its "dot" inner product. Given a vector $\mathbf{x}$ and a basis $\mathbf{e}_\mu$ for which $g_{\mu\nu} = \mathbf{e}_\mu \cdot \mathbf{e}_\nu$, we can define two sets of components for the same vector. Firstly the coefficients x^μ appearing in the basis expansion

$$\mathbf{x} = x^\mu \mathbf{e}_\mu, \tag{10.16}$$

and secondly the "components"

$$x_\mu = \mathbf{e}_\mu \cdot \mathbf{x} = \mathbf{g}(\mathbf{e}_\mu, \mathbf{x}) = \mathbf{g}(\mathbf{e}_\mu, x^\nu \mathbf{e}_\nu) = \mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) x^\nu = g_{\mu\nu} x^\nu \tag{10.17}$$

of $\mathbf{x}$ along the basis vectors. These two set of numbers are then respectively called the contravariant and covariant components of the vector $\mathbf{x}$. If the $\mathbf{e}_\mu$ constitute an orthonormal basis, where $g_{\mu\nu} = \delta_{\mu\nu}$, then the two sets of components (covariant and contravariant) are numerically coincident. In a non-orthogonal basis they will be different, and we must take care never to add contravariant components to covariant ones.

10.2 Tensors

We now introduce tensors in two ways: firstly as sets of numbers labelled by indices and equipped with transformation laws that tell us how these numbers change as we change basis; and secondly as basis-independent objects that are elements of a vector space constructed by taking multiple tensor products of the spaces V and V^*.

10.2.1 Transformation rules

After we change basis $\mathbf{e}_\mu \to \mathbf{e}'_\mu$, where $\mathbf{e}_\nu = a^\mu_\nu \mathbf{e}'_\mu$, the metric tensor will be represented by a new set of components

$$g'_{\mu\nu} = \mathbf{g}(\mathbf{e}'_\mu, \mathbf{e}'_\nu). \tag{10.18}$$

These are be related to the old components by

$$g_{\mu\nu} = \mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) = \mathbf{g}(a_\mu^\rho \mathbf{e}'_\rho, a_\nu^\sigma \mathbf{e}'_\sigma) = a_\mu^\rho a_\nu^\sigma \mathbf{g}(\mathbf{e}'_\rho, \mathbf{e}'_\sigma) = a_\mu^\rho a_\nu^\sigma \, g'_{\rho\sigma}. \qquad (10.19)$$

This transformation rule for $g_{\mu\nu}$ has both of its subscripts behaving like the downstairs indices of a covector. We therefore say that $g_{\mu\nu}$ transforms as a *doubly covariant tensor*. Written out in full, for a two-dimensional space, the transformation law is

$$\begin{aligned}
g_{11} &= a_1^1 a_1^1 g'_{11} + a_1^1 a_1^2 g'_{12} + a_1^2 a_1^1 g'_{21} + a_1^2 a_1^2 g'_{22},\\
g_{12} &= a_1^1 a_2^1 g'_{11} + a_1^1 a_2^2 g'_{12} + a_1^2 a_2^1 g'_{21} + a_1^2 a_2^2 g'_{22},\\
g_{21} &= a_2^1 a_1^1 g'_{11} + a_2^1 a_1^2 g'_{12} + a_2^2 a_1^1 g'_{21} + a_2^2 a_1^2 g'_{22},\\
g_{22} &= a_2^1 a_2^1 g'_{11} + a_2^1 a_2^2 g'_{12} + a_2^2 a_2^1 g'_{21} + a_2^2 a_2^2 g'_{22}.
\end{aligned}$$

In three dimensions each row would have nine terms, and sixteen in four dimensions.

A set of numbers $Q^{\alpha\beta}{}_{\gamma\delta\epsilon}$, whose indices range from 1 to the dimension of the space and that transforms as

$$Q^{\alpha\beta}{}_{\gamma\delta\epsilon} = (a^{-1})^\alpha_{\alpha'} (a^{-1})^\beta_{\beta'} \, a_\gamma^{\gamma'} \, a_\delta^{\delta'} \, a_\epsilon^{\epsilon'} \, Q'^{\alpha'\beta'}{}_{\gamma'\delta'\epsilon'}, \qquad (10.20)$$

or conversely as

$$Q'^{\alpha\beta}{}_{\gamma\delta\epsilon} = a^\alpha_{\alpha'} a^\beta_{\beta'} (a^{-1})_\gamma^{\gamma'} (a^{-1})_\delta^{\delta'} (a^{-1})_\epsilon^{\epsilon'} Q^{\alpha'\beta'}{}_{\gamma'\delta'\epsilon'}, \qquad (10.21)$$

comprises the components of a *doubly contravariant, triply covariant* tensor. More compactly, the $Q^{\alpha\beta}{}_{\gamma\delta\epsilon}$ are the components of a tensor of type $(2,3)$. Tensors of type (p,q) are defined analogously. The total number of indices $p+q$ is called the *rank* of the tensor.

Note how the indices are wired up in the transformation rules (10.20) and (10.21): free (not summed over) upstairs indices on the left hand side of the equations match to free upstairs indices on the right hand side, similarly for the downstairs indices. Also upstairs indices are summed only with downstairs ones.

Similar conditions apply to equations relating tensors in any particular basis. If they are violated you do not have a valid tensor equation — meaning that an equation valid in one basis will not be valid in another basis. Thus an equation

$$A^\mu{}_{\nu\lambda} = B^{\mu\tau}{}_{\nu\lambda\tau} + C^\mu{}_{\nu\lambda} \qquad (10.22)$$

is fine, but

$$A^{\mu}{}_{\nu\lambda} \stackrel{?}{=} B^{\nu}{}_{\mu\lambda} + C^{\mu}{}_{\nu\lambda\sigma\sigma} + D^{\mu}{}_{\nu\lambda\tau} \tag{10.23}$$

has something wrong in each term.

Incidentally, although not illegal, it is a good idea not to write tensor indices directly underneath one another — *i.e.* do not write Q^{ij}_{kjl} — because if you raise or lower indices using the metric tensor, and some pages later in a calculation try to put them back where they were, they might end up in the wrong order.

Tensor algebra

The sum of two tensors of a given type is also a tensor of that type. The sum of two tensors of different types is not a tensor. Thus each particular type of tensor constitutes a distinct vector space, but one derived from the common underlying vector space whose change-of-basis formula is being utilized.

Tensors can be combined by multiplication: if $A^{\mu}{}_{\nu\lambda}$ and $B^{\mu}{}_{\nu\lambda\tau}$ are tensors of type $(1,2)$ and $(1,3)$ respectively, then

$$C^{\alpha\beta}{}_{\nu\lambda\rho\sigma\tau} = A^{\alpha}{}_{\nu\lambda} B^{\beta}{}_{\rho\sigma\tau} \tag{10.24}$$

is a tensor of type $(2,5)$.

An important operation is *contraction*, which consists of setting one or more contravariant index index equal to a covariant index and summing over the repeated indices. This reduces the rank of the tensor. So, for example,

$$D_{\rho\sigma\tau} = C^{\alpha\beta}{}_{\alpha\beta\rho\sigma\tau} \tag{10.25}$$

is a tensor of type $(0,3)$. Similarly $\mathbf{f}(\mathbf{x}) = f_{\mu} x^{\mu}$ is a type $(0,0)$ tensor, *i.e.* an *invariant* — a number that takes the same value in all bases. Upper indices can only be contracted with lower indices, and *vice versa*. For example, the array of numbers $A_{\alpha} = B_{\alpha\beta\beta}$ obtained from the type $(0,3)$ tensor $B_{\alpha\beta\gamma}$ is *not* a tensor of type $(0,1)$.

The contraction procedure outputs a tensor because setting an upper index and a lower index to a common value μ and summing over μ, leads to the factor $\ldots (a^{-1})^{\mu}_{\alpha} a^{\beta}_{\mu} \ldots$ appearing in the transformation rule. Now

$$(a^{-1})^{\mu}_{\alpha} a^{\beta}_{\mu} = \delta^{\beta}_{\alpha}, \tag{10.26}$$

and the Kronecker delta effects a summation over the corresponding pair of indices in the transformed tensor.

Although often associated with general relativity, tensors occur in many places in physics. They are used, for example, in elasticity theory, where the word "tensor" in its modern meaning was introduced by Woldemar Voigt in 1898. Voigt, following Cauchy and Green, described the infinitesimal deformation of an elastic body by the *strain tensor* $e_{\alpha\beta}$, which is a tensor of type (0,2). The forces to which the strain gives rise are described by the *stress tensor* $\sigma^{\lambda\mu}$. A generalization of Hooke's law relates stress to strain via a tensor of elastic constants $c^{\alpha\beta\gamma\delta}$ as

$$\sigma^{\alpha\beta} = c^{\alpha\beta\gamma\delta} e_{\gamma\delta}. \tag{10.27}$$

We study stress and strain in more detail later in this chapter.

Exercise 10.1: Show that $g^{\mu\nu}$, the matrix inverse of the metric tensor $g_{\mu\nu}$, is indeed a doubly contravariant tensor, as the position of its indices suggests.

10.2.2 Tensor character of linear maps and quadratic forms

As an illustration of the tensor concept and of the need to distinguish between upstairs and downstairs indices, we contrast the properties of matrices representing linear maps and those representing quadratic forms.

A linear map $M : V \to V$ is an object that exists independently of any basis. Given a basis, however, it is represented by a matrix $M^\mu{}_\nu$ obtained by examining the action of the map on the basis elements:

$$M(\mathbf{e}_\mu) = \mathbf{e}_\nu M^\nu{}_\mu. \tag{10.28}$$

Acting on $\mathbf{x}$ we get a new vector $\mathbf{y} = M(\mathbf{x})$, where

$$y^\nu \mathbf{e}_\nu = \mathbf{y} = M(\mathbf{x}) = M(x^\mu \mathbf{e}_\mu) = x^\mu M(\mathbf{e}_\mu) = x^\mu M^\nu{}_\mu \mathbf{e}_\nu = M^\nu{}_\mu x^\mu \, \mathbf{e}_\nu. \tag{10.29}$$

We therefore have

$$y^\nu = M^\nu{}_\mu x^\mu, \tag{10.30}$$

which is the usual matrix multiplication $\mathbf{y} = \mathbf{M}\mathbf{x}$. When we change basis, $\mathbf{e}_\nu = a^\mu_\nu \mathbf{e}'_\mu$, then

$$\mathbf{e}_\nu M^\nu{}_\mu = M(\mathbf{e}_\mu) = M(a^\rho_\mu \mathbf{e}'_\rho) = a^\rho_\mu M(\mathbf{e}'_\rho) = a^\rho_\mu \mathbf{e}'_\sigma M'^\sigma{}_\rho = a^\rho_\mu (a^{-1})^\nu_\sigma \mathbf{e}_\nu M'^\sigma{}_\rho. \tag{10.31}$$

Comparing coefficients of $\mathbf{e}_\nu$, we find

$$M^{\nu}{}_{\mu} = a_\mu^\rho (a^{-1})_\sigma^\nu M'^{\sigma}{}_{\rho}, \tag{10.32}$$

or, conversely,

$$M'^{\nu}{}_{\mu} = (a^{-1})_\mu^\rho a_\sigma^\nu M^{\sigma}{}_{\rho}. \tag{10.33}$$

Thus a matrix representing a linear map has the tensor character suggested by the position of its indices, *i.e.* it transforms as a type $(1,1)$ tensor. We can derive the same formula in matrix notation. In the new basis the vectors $\mathbf{x}$ and $\mathbf{y}$ have new components $\mathbf{x}' = \mathbf{A}\mathbf{x}$, and $\mathbf{y}' = \mathbf{A}\mathbf{y}$. Consequently $\mathbf{y} = \mathbf{M}\mathbf{x}$ becomes

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \mathbf{A}\mathbf{M}\mathbf{x} = \mathbf{A}\mathbf{M}\mathbf{A}^{-1}\mathbf{x}', \tag{10.34}$$

and the matrix representing the map M has new components

$$\mathbf{M}' = \mathbf{A}\mathbf{M}\mathbf{A}^{-1}. \tag{10.35}$$

Now consider the quadratic form $Q : V \to \mathbb{R}$ that is obtained from a symmetric bilinear form $Q : V \times V \to \mathbb{R}$ by setting $Q(\mathbf{x}) = Q(\mathbf{x}, \mathbf{x})$. We can write

$$Q(\mathbf{x}) = Q_{\mu\nu} x^\mu x^\nu = x^\mu Q_{\mu\nu}\, x^\nu = \mathbf{x}^T \mathbf{Q}\mathbf{x}, \tag{10.36}$$

where $Q_{\mu\nu} \equiv Q(\mathbf{e}_\mu, \mathbf{e}_\nu)$ are the entries in the symmetric matrix $\mathbf{Q}$, the suffix T denotes transposition, and $\mathbf{x}^T\mathbf{Q}\mathbf{x}$ is standard matrix-multiplication notation. Just as does the metric tensor, the coefficients $Q_{\mu\nu}$ transform as a type $(0,2)$ tensor:

$$Q_{\mu\nu} = a_\mu^\alpha a_\nu^\beta Q'_{\alpha\beta}. \tag{10.37}$$

In matrix notation the vector $\mathbf{x}$ again transforms to have new components $\mathbf{x}' = \mathbf{A}\mathbf{x}$, but $\mathbf{x}'^T = \mathbf{x}^T\mathbf{A}^T$. Consequently

$$\mathbf{x}'^T\mathbf{Q}'\mathbf{x}' = \mathbf{x}^T\mathbf{A}^T\mathbf{Q}'\mathbf{A}\mathbf{x}. \tag{10.38}$$

Thus

$$\mathbf{Q} = \mathbf{A}^T\mathbf{Q}'\mathbf{A}. \tag{10.39}$$

The message is that linear maps and quadratic forms can both be represented by matrices, but these matrices correspond to distinct types of tensor and transform differently under a change of basis.

A matrix representing a linear map has a basis-independent determinant. Similarly the *trace* of a matrix representing a linear map

$$\operatorname{tr}\mathbf{M} \stackrel{\text{def}}{=} M^{\mu}{}_{\mu} \tag{10.40}$$

is a tensor of type $(0,0)$, i.e. a scalar, and therefore basis independent. On the other hand, while you can certainly compute the determinant or the trace of the matrix representing a quadratic form in some particular basis, when you change basis and calculate the determinant or trace of the transformed matrix, you will get a different number.

It *is* possible to make a quadratic form out of a linear map, but this requires using the metric to lower the contravariant index on the matrix representing the map:

$$Q(\mathbf{x}) = x^\mu g_{\mu\nu} Q^\nu{}_\lambda x^\lambda = \mathbf{x} \cdot \mathbf{Q}\mathbf{x}. \tag{10.41}$$

Be careful, therefore: the matrices "$\mathbf{Q}$" in $\mathbf{x}^T\mathbf{Q}\mathbf{x}$ and in $\mathbf{x}\cdot\mathbf{Q}\mathbf{x}$ are representing different mathematical objects.

Exercise 10.2: In this problem we will use the distinction between the transformation law of a quadratic form and that of a linear map to resolve the following "paradox":

- In quantum mechanics we are taught that the matrices representing two operators can be simultaneously diagonalized only if they commute.
- In classical mechanics we are taught how, given the Lagrangian

$$L = \sum_{ij}\left(\frac{1}{2}\dot{q}_i M_{ij}\dot{q}_j - \frac{1}{2}q_i V_{ij} q_j\right),$$

to construct normal co-ordinates Q_i such that L becomes

$$L = \sum_i \left(\frac{1}{2}\dot{Q}_i^2 - \frac{1}{2}\omega_i^2 Q_i^2\right).$$

We have apparently managed to simultaneously diagonize the matrices $M_{ij} \to \mathrm{diag}\,(1,\ldots,1)$ and $V_{ij} \to \mathrm{diag}\,(\omega_1^2,\ldots,\omega_n^2)$, even though there is no reason for them to commute with each other!

Show that when $\mathbf{M}$ and $\mathbf{V}$ are a pair of symmetric matrices, with $\mathbf{M}$ being positive definite, then there exists an invertible matrix $\mathbf{A}$ such that $\mathbf{A}^T\mathbf{M}\mathbf{A}$ and $\mathbf{A}^T\mathbf{V}\mathbf{A}$ are simultaneously diagonal. (Hint: Consider $\mathbf{M}$ as defining an inner product, and use the Gramm-Schmidt procedure to first find a orthonormal frame in which $M'_{ij} = \delta_{ij}$. Then show that the matrix corresponding to $\mathbf{V}$ in this frame can be diagonalized by a further transformation that does not perturb the already diagonal M'_{ij}.)

10.2.3 Tensor product spaces

We may regard the set of numbers $Q^{\alpha\beta}{}_{\gamma\delta\epsilon}$ as being the components of an object $\mathbf{Q}$ that is element of the vector space of type $(2,3)$ tensors. We denote this vector space by the symbol $V\otimes V\otimes V^*\otimes V^*\otimes V^*$, the notation indicating that it is derived from the original V and its dual V^* by taking *tensor products* of these spaces. The tensor $\mathbf{Q}$ is to be thought of as existing as an element of $V\otimes V\otimes V^*\otimes V^*\otimes V^*$ independently of any basis, but given a basis $\{\mathbf{e}_\mu\}$ for V, and the dual basis $\{\mathbf{e}^{*\nu}\}$ for V^*, we expand it as

$$\mathbf{Q} = Q^{\alpha\beta}{}_{\gamma\delta\epsilon}\,\mathbf{e}_\alpha\otimes\mathbf{e}_\beta\otimes\mathbf{e}^{*\gamma}\otimes\mathbf{e}^{*\delta}\otimes\mathbf{e}^{*\epsilon}. \tag{10.42}$$

Here the tensor product symbol "$\otimes$" is distributive

$$\begin{aligned}\mathbf{a}\otimes(\mathbf{b}+\mathbf{c}) &= \mathbf{a}\otimes\mathbf{b}+\mathbf{a}\otimes\mathbf{c},\\ (\mathbf{a}+\mathbf{b})\otimes\mathbf{c} &= \mathbf{a}\otimes\mathbf{c}+\mathbf{b}\otimes\mathbf{c},\end{aligned} \tag{10.43}$$

and associative

$$(\mathbf{a}\otimes\mathbf{b})\otimes\mathbf{c} = \mathbf{a}\otimes(\mathbf{b}\otimes\mathbf{c}), \tag{10.44}$$

but is not commutative

$$\mathbf{a}\otimes\mathbf{b}\neq\mathbf{b}\otimes\mathbf{a}. \tag{10.45}$$

Everything commutes with the field, however,

$$\lambda(\mathbf{a}\otimes\mathbf{b}) = (\lambda\mathbf{a})\otimes\mathbf{b} = \mathbf{a}\otimes(\lambda\mathbf{b}). \tag{10.46}$$

If we change basis $\mathbf{e}_\alpha = a_\alpha^\beta\mathbf{e}'_\beta$ then these rules lead, for example, to

$$\mathbf{e}_\alpha\otimes\mathbf{e}_\beta = a_\alpha^\lambda a_\beta^\mu\,\mathbf{e}'_\lambda\otimes\mathbf{e}'_\mu. \tag{10.47}$$

From this change-of-basis formula, we deduce that

$$T^{\alpha\beta}\mathbf{e}_\alpha\otimes\mathbf{e}_\beta = T^{\alpha\beta}a_\alpha^\lambda a_\beta^\mu\,\mathbf{e}'_\lambda\otimes\mathbf{e}'_\mu = T'^{\lambda\mu}\,\mathbf{e}'_\lambda\otimes\mathbf{e}'_\mu, \tag{10.48}$$

where

$$T'^{\lambda\mu} = T^{\alpha\beta}a_\alpha^\lambda a_\beta^\mu. \tag{10.49}$$

The analogous formula for $\mathbf{e}_\alpha\otimes\mathbf{e}_\beta\otimes\mathbf{e}^{*\gamma}\otimes\mathbf{e}^{*\delta}\otimes\mathbf{e}^{*\epsilon}$ reproduces the transformation rule for the components of $\mathbf{Q}$.

The meaning of the tensor product of a collection of vector spaces should now be clear: If $\mathbf{e}_\mu$ consititute a basis for V, the space $V\otimes V$ is, for example, the space of all linear combinations[1] of the abstract symbols $\mathbf{e}_\mu \otimes \mathbf{e}_\nu$, which we declare by *fiat* to constitute a basis for this space. There is no geometric significance (as there is with a vector product $\mathbf{a} \times \mathbf{b}$) to the tensor product $\mathbf{a} \otimes \mathbf{b}$, so the $\mathbf{e}_\mu \otimes \mathbf{e}_\nu$ are simply useful place-keepers. Remember that these are *ordered* pairs, $\mathbf{e}_\mu \otimes \mathbf{e}_\nu \neq \mathbf{e}_\nu \otimes \mathbf{e}_\mu$.

Although there is no *geometric* meaning, it is possible, however, to give an *algebraic* meaning to a product like $\mathbf{e}^{*\lambda} \otimes \mathbf{e}^{*\mu} \otimes \mathbf{e}^{*\nu}$ by viewing it as a multilinear form $V \times V \times V :\to \mathbb{R}$. We define

$$\mathbf{e}^{*\lambda} \otimes \mathbf{e}^{*\mu} \otimes \mathbf{e}^{*\nu} (\mathbf{e}_\alpha, \mathbf{e}_\beta, \mathbf{e}_\gamma) = \delta^\lambda_\alpha\, \delta^\mu_\beta\, \delta^\nu_\gamma. \tag{10.50}$$

We may also regard it as a linear map $V \otimes V \otimes V :\to \mathbb{R}$ by defining

$$\mathbf{e}^{*\lambda} \otimes \mathbf{e}^{*\mu} \otimes \mathbf{e}^{*\nu} (\mathbf{e}_\alpha \otimes \mathbf{e}_\beta \otimes \mathbf{e}_\gamma) = \delta^\lambda_\alpha\, \delta^\mu_\beta\, \delta^\nu_\gamma \tag{10.51}$$

and extending the definition to general elements of $V \otimes V \otimes V$ by linearity. In this way we establish an isomorphism

$$V^* \otimes V^* \otimes V^* \cong (V \otimes V \otimes V)^*. \tag{10.52}$$

This multiple personality is typical of tensor spaces. We have already seen that the metric tensor is simultaneously an element of $V^* \otimes V^*$ and a map $\mathbf{g} : V \to V^*$.

Tensor products and quantum mechanics

When we have two quantum-mechanical systems having Hilbert spaces $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$, the Hilbert space for the combined system is $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$. Quantum mechanics books usually denote the vectors in these spaces by the Dirac "bra-ket" notation in which the basis vectors of the separate spaces are denoted by[2] $|n_1\rangle$ and $|n_2\rangle$, and that of the combined space by $|n_1, n_2\rangle$. In this notation, a state in the combined system is a linear combination

$$|\Psi\rangle = \sum_{n_1,n_2} |n_1, n_2\rangle\langle n_1, n_2|\Psi\rangle, \tag{10.53}$$

[1]Do not confuse the tensor-product space $V \otimes W$ with the Cartesian product $V \times W$. The latter is the set of all ordered pairs $(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in V$, $\mathbf{y} \in W$. The tensor product includes also *formal sums* of such pairs. The Cartesian product of two vector spaces can be given the structure of a vector space by defining an addition operation $\lambda(\mathbf{x}_1, \mathbf{y}_1) + \mu(\mathbf{x}_2, \mathbf{y}_2) = (\lambda\mathbf{x}_1 + \mu\mathbf{x}_2, \lambda\mathbf{y}_1 + \mu\mathbf{y}_2)$, but this construction does not lead to the tensor product. Instead it defines the *direct sum* $V \oplus W$.

[2]We assume for notational convenience that the Hilbert spaces are finite dimensional.

This is the tensor product in disguise. To unmask it, we simply make the notational translation

$$\begin{aligned}
|\Psi\rangle &\rightarrow \boldsymbol{\Psi}\\
\langle n_1, n_2|\Psi\rangle &\rightarrow \psi^{n_1,n_2}\\
|n_1\rangle &\rightarrow \mathbf{e}^{(1)}_{n_1}\\
|n_2\rangle &\rightarrow \mathbf{e}^{(2)}_{n_2}\\
|n_1, n_2\rangle &\rightarrow \mathbf{e}^{(1)}_{n_1} \otimes \mathbf{e}^{(2)}_{n_2}. \qquad (10.54)
\end{aligned}$$

Then (10.53) becomes

$$\boldsymbol{\Psi} = \psi^{n_1,n_2}\, \mathbf{e}^{(1)}_{n_1} \otimes \mathbf{e}^{(2)}_{n_2}. \qquad (10.55)$$

Entanglement: Suppose that $\mathcal{H}^{(1)}$ has basis $\mathbf{e}^{(1)}_1, \ldots, \mathbf{e}^{(1)}_m$ and $\mathcal{H}^{(2)}$ has basis $\mathbf{e}^{(2)}_1, \ldots, \mathbf{e}^{(2)}_n$. The Hilbert space $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ is then nm dimensional. Consider a state

$$\boldsymbol{\Psi} = \psi^{ij} \mathbf{e}^{(1)}_i \otimes \mathbf{e}^{(2)}_j \in \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}. \qquad (10.56)$$

If we can find vectors

$$\begin{aligned}
\boldsymbol{\Phi} &\equiv \phi^i \mathbf{e}^{(1)}_i \in \mathcal{H}^{(1)},\\
\mathbf{X} &\equiv \chi^j \mathbf{e}^{(2)}_j \in \mathcal{H}^{(2)}, \qquad (10.57)
\end{aligned}$$

such that

$$\boldsymbol{\Psi} = \boldsymbol{\Phi} \otimes \mathbf{X} \equiv \phi^i \chi^j \mathbf{e}^{(1)}_i \otimes \mathbf{e}^{(2)}_j \qquad (10.58)$$

then the tensor $\boldsymbol{\Psi}$ is said to be *decomposable* and the two quantum systems are said to be *unentangled*. If there are no such vectors then the two systems are *entangled* in the sense of the Einstein-Podolski-Rosen (EPR) paradox.

Quantum states are really in one-to-one correspondence with *rays* in the Hilbert space, rather than vectors. If we denote the n dimensional vector space over the field of the complex numbers as $\mathbb{C}^n$, the space of rays, in which we do not distinguish between the vectors $\mathbf{x}$ and $\lambda\mathbf{x}$ when $\lambda \neq 0$, is denoted by $\mathbb{C}P^{n-1}$ and is called *complex projective space*. Complex projective space is where *algebraic geometry* is studied. The set of decomposable states may be thought of as a subset of the complex projective space $\mathbb{C}P^{nm-1}$, and, since, as the following excercise shows, this subset is defined by a finite number of homogeneous polynomial equations, it forms what algebraic geometers call a *variety*. This particular subset is known as the *Segre variety*.

Exercise 10.3: The Segre conditions for a state to be decomposable:

i) By counting the number of independent components that are at our disposal in $\mathbf{\Psi}$, and comparing that number with the number of free parameters in $\mathbf{\Phi}\otimes\mathbf{X}$, show that the coefficients ψ^{ij} must satisfy $(n-1)(m-1)$ relations if the state is to be decomposable.

ii) If the state is decomposable, show that

$$0 = \begin{vmatrix} \psi^{ij} & \psi^{il} \\ \psi^{kj} & \psi^{kl} \end{vmatrix}$$

for all sets of indices i, j, k, l.

iii) Assume that ψ^{11} is not zero. Using your count from part (i) as a guide, find a subset of the relations from part (ii) that constitute a necessary and sufficient set of conditions for the state Ψ to be decomposable. Include a proof that your set is indeed sufficient.

10.2.4 Symmetric and skew-symmetric tensors

By examining the transformation rule you may see that if a pair of upstairs or downstairs indices is *symmetric* (say $Q^{\mu\nu}{}_{\rho\sigma\tau} = Q^{\nu\mu}{}_{\rho\sigma\tau}$) or *skew-symmetric* ($Q^{\mu\nu}{}_{\rho\sigma\tau} = -Q^{\nu\mu}{}_{\rho\sigma\tau}$) in one basis, it remains so after the basis has been changed. (This is **not** true of a pair composed of one upstairs and one downstairs index.) It makes sense, therefore, to define symmetric and skew-symmetric tensor product spaces. Thus skew-symmetric doubly-contravariant tensors can be regarded as belonging to the space denoted by $\bigwedge^2 V$ and expanded as

$$\mathbf{A} = \frac{1}{2}A^{\mu\nu}\,\mathbf{e}_\mu\wedge\mathbf{e}_\nu, \tag{10.59}$$

where the coefficients are skew-symmetric, $A^{\mu\nu} = -A^{\nu\mu}$, and the *wedge product* of the basis elements is associative and distributive, as is the tensor product, but in addition obeys $\mathbf{e}_\mu\wedge\mathbf{e}_\nu = -\mathbf{e}_\nu\wedge\mathbf{e}_\mu$. The "1/2" (replaced by $1/p!$ when there are p indices) is convenient in that each independent component only appears once in the sum. For example, in three dimensions,

$$\frac{1}{2}A^{\mu\nu}\,\mathbf{e}_\mu\wedge\mathbf{e}_\nu = A^{12}\,\mathbf{e}_1\wedge\mathbf{e}_2 + A^{23}\,\mathbf{e}_2\wedge\mathbf{e}_3 + A^{31}\,\mathbf{e}_3\wedge\mathbf{e}_1. \tag{10.60}$$

Symmetric doubly-contravariant tensors can be regarded as belonging to the space $\mathrm{sym}^2 V$ and expanded as

$$\mathbf{S} = S^{\alpha\beta}\,\mathbf{e}_\alpha\odot\mathbf{e}_\beta \tag{10.61}$$

where $\mathbf{e}_\alpha \odot \mathbf{e}_\beta = \mathbf{e}_\beta \odot \mathbf{e}_\alpha$ and $S^{\alpha\beta} = S^{\beta\alpha}$. (We do not insert a "1/2" here because including it leads to no particular simplification in any consequent equations.)

We can treat these symmetric and skew-symmetric products as symmetric or skew multilinear forms. Define, for example,

$$\mathbf{e}^{*\alpha} \wedge \mathbf{e}^{*\beta}\,(\mathbf{e}_\mu, \mathbf{e}_\nu) = \delta^\alpha_\mu \delta^\beta_\nu - \delta^\alpha_\nu \delta^\beta_\mu, \tag{10.62}$$

and

$$\mathbf{e}^{*\alpha} \wedge \mathbf{e}^{*\beta}\,(\mathbf{e}_\mu \wedge \mathbf{e}_\nu) = \delta^\alpha_\mu \delta^\beta_\nu - \delta^\alpha_\nu \delta^\beta_\mu. \tag{10.63}$$

We need two terms on the right-hand-side of these examples because the skew-symmetry of $\mathbf{e}^{*\alpha} \wedge \mathbf{e}^{*\beta}(\ ,\ )$ in its slots does not allow us the luxury of demanding that the $\mathbf{e}_\mu$ be inserted in the exact order of the $\mathbf{e}^{*\alpha}$ to get a non-zero answer. Because the p-th order analogue of (10.62) form has $p!$ terms on its right-hand side, some authors like to divide the right-hand-side by $p!$ in this definition. We prefer the one above, though. With our definition, and with $\mathbf{A} = \frac{1}{2}A_{\mu\nu}\mathbf{e}^{*\mu} \wedge \mathbf{e}^{*\nu}$ and $\mathbf{B} = \frac{1}{2}B^{\alpha\beta}\mathbf{e}_\alpha \wedge \mathbf{e}_\beta$, we have

$$\mathbf{A}(\mathbf{B}) = \frac{1}{2}A_{\mu\nu}B^{\mu\nu} = \sum_{\mu<\nu} A_{\mu\nu}B^{\mu\nu}, \tag{10.64}$$

so the sum is only over independent terms.

The wedge ($\wedge$) product notation is standard in mathematics wherever skew-symmetry is implied.[3] The "sym" and $\odot$ are not. Different authors use different notations for spaces of symmetric tensors. This reflects the fact that skew-symmetric tensors are extremely useful and appear in many different parts of mathematics, while symmetric ones have fewer special properties (although they are common in physics). Compare the relative usefulness of determinants and permanents.

Exercise 10.4: Show that in d dimensions:

i) the dimension of the space of skew-symmetric covariant tensors with p indices is $d!/p!(d-p)!$;
ii) the dimension of the space of symmetric covariant tensors with p indices is $(d+p-1)!/p!(d-1)!$.

[3] Skew products and abstract vector spaces were introduced simultaneously in Hermann Grassmann's *Ausdehnungslehre* (1844). Grassmann's mathematics was not appreciated in his lifetime. In his disappointment he turned to other fields, making significant contributions to the theory of colour mixtures (Grassmann's law), and to the philology of Indo-European languages (another Grassmann's law).

Bosons and fermions

Spaces of symmetric and skew-symmetric tensors appear whenever we deal with the quantum mechanics of many indistinguishable particles possessing Bose or Fermi statistics. If we have a Hilbert space $\mathcal{H}$ of single-particle states with basis $\mathbf{e}_i$ then the N-boson space is $\mathrm{Sym}^N\mathcal{H}$ which consists of states

$$\mathbf{\Phi} = \Phi^{i_1 i_2 \ldots i_N}\mathbf{e}_{i_1} \odot \mathbf{e}_{i_2} \odot \cdots \odot \mathbf{e}_{i_N}, \tag{10.65}$$

and the N-fermion space is $\bigwedge^N \mathcal{H}$, which contains states

$$\mathbf{\Psi} = \frac{1}{N!}\Psi^{i_1 i_2 \ldots i_N}\,\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_N}. \tag{10.66}$$

The symmetry of the Bose wavefunction

$$\Phi^{i_1 \ldots i_\alpha \ldots i_\beta \ldots i_N} = \Phi^{i_1 \ldots i_\beta \ldots i_\alpha \ldots i_N}, \tag{10.67}$$

and the skew-symmetry of the Fermion wavefunction

$$\Psi^{i_1 \ldots i_\alpha \ldots i_\beta \ldots i_N} = -\Psi^{i_1 \ldots i_\beta \ldots i_\alpha \ldots i_N}, \tag{10.68}$$

under the interchange of the particle labels α, β is then natural.
Slater Determinants and the Plücker Relations: Some N-fermion states can be decomposed into a product of single-particle states

$$\begin{aligned} \mathbf{\Psi} &= \boldsymbol{\psi}_1 \wedge \boldsymbol{\psi}_2 \wedge \cdots \wedge \boldsymbol{\psi}_N \\ &= \psi_1^{i_1}\psi_2^{i_2}\cdots\psi_N^{i_N}\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_N}. \end{aligned} \tag{10.69}$$

Comparing the coefficients of $\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_N}$ in (10.66) and (10.69) shows that the many-body wavefunction can then be written as

$$\Psi^{i_1 i_2 \ldots i_N} = \begin{vmatrix} \psi_1^{i_1} & \psi_1^{i_2} & \cdots & \psi_1^{i_N} \\ \psi_2^{i_1} & \psi_2^{i_2} & \cdots & \psi_2^{i_N} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^{i_1} & \psi_N^{i_2} & \cdots & \psi_N^{i_N} \end{vmatrix}. \tag{10.70}$$

The wavefunction is therefore given by a single *Slater determinant*. Such wavefunctions correspond to a very special class of states. The general many-fermion state is not decomposable, and its wavefunction can only be expressed as a sum of many Slater determinants. The Hartree-Fock method of quantum chemistry is a variational approximation that takes such a single Slater determinant as its trial wavefunction and varies only the one-particle wavefunctions $\langle i|\psi_a\rangle \equiv \psi_a^i$. It is a remarkably successful approximation, given the very restricted class of wavefunctions it explores.

As with the Segre condition for two distinguishable quantum systems to be unentangled, there is a set of necessary and sufficient conditions on the $\Psi^{i_1 i_2 \ldots i_N}$ for the state $\mathbf{\Psi}$ to be decomposable into single-particle states. The conditions are that

$$\Psi^{i_1 i_2 \ldots i_{N-1}[j_1}\Psi^{j_2 j_3 \ldots j_{N+1}]} = 0 \tag{10.71}$$

for any choice of indices $i_1, \ldots i_{N-1}$ and $j_1, \ldots, j_{N+1}$. The square brackets $[\ldots]$ indicate that the expression is to be antisymmetrized over the indices enclosed in the brackets. For example, a three-particle state is decomposable if and only if

$$\Psi^{i_1 i_2 j_1}\Psi^{j_2 j_3 j_4} - \Psi^{i_1 i_2 j_2}\Psi^{j_1 j_3 j_4} + \Psi^{i_1 i_2 j_3}\Psi^{j_1 j_2 j_4} - \Psi^{i_1 i_2 j_4}\Psi^{j_1 j_2 j_3} = 0. \tag{10.72}$$

These conditions are called the *Plücker relations* after Julius Plücker who discovered them long before before the advent of quantum mechanics.[4] It is easy to show that Plücker's relations are necessary conditions for decomposability. It takes more sophistication to show that they are sufficient. We will therefore defer this task to the exercises as the end of the chapter. As far as we are aware, the Plücker relations are not exploited by quantum chemists, but, in disguise as the *Hirota bilinear equations*, they constitute the geometric condition underpinning the many-soliton solutions of the Korteweg-de-Vries and other soliton equations.

10.2.5 Kronecker and Levi-Civita tensors

Suppose the tensor δ^μ_ν is defined, with respect to some basis, to be unity if $\mu = \nu$ and zero otherwise. In a new basis it will transform to

$$\delta'^\mu_\nu = a^\mu_\rho (a^{-1})^\sigma_\nu \delta^\rho_\sigma = a^\mu_\rho (a^{-1})^\rho_\nu = \delta^\mu_\nu. \tag{10.73}$$

In other words the Kronecker delta symbol of type $(1,1)$ has the same numerical components in all co-ordinate systems. This is not true of the Kronecker delta symbol of type $(0,2)$, *i.e.* of $\delta_{\mu\nu}$.

[4] As well as his extensive work in algebraic geometry, Plücker (1801-68) made important discoveries in experimental physics. He was, for example, the first person to observe the deflection of cathode rays — beams of electrons — by a magnetic field, and the first to point out that each element had its characteristic emission spectrum.

Now consider an n-dimensional space with a tensor $\eta_{\mu_1\mu_2\dots\mu_n}$ whose components, in some basis, coincides with the Levi-Civita symbol $\epsilon_{\mu_1\mu_2\dots\mu_n}$. We find that in a new frame the components are

$$\begin{aligned}\eta'_{\mu_1\mu_2\dots\mu_n} &= (a^{-1})^{\nu_1}_{\mu_1}(a^{-1})^{\nu_2}_{\mu_2}\cdots(a^{-1})^{\nu_n}_{\mu_n}\epsilon_{\nu_1\nu_2\dots\nu_n}\\ &= \epsilon_{\mu_1\mu_2\dots\mu_n}\,(a^{-1})^{\nu_1}_{1}(a^{-1})^{\nu_2}_{2}\cdots(a^{-1})^{\nu_n}_{n}\epsilon_{\nu_1\nu_2\dots\nu_n}\\ &= \epsilon_{\mu_1\mu_2\dots\mu_n}\det\mathbf{A}^{-1}\\ &= \eta_{\mu_1\mu_2\dots\mu_n}\det\mathbf{A}^{-1}.\end{aligned} \tag{10.74}$$

Thus, unlike the δ^μ_ν, the Levi-Civita symbol is not quite a tensor.

Consider also the quantity

$$\sqrt{g} \stackrel{\text{def}}{=} \sqrt{\det[g_{\mu\nu}]}. \tag{10.75}$$

Here we assume that the metric is positive-definite, so that the square root is real, and that we have taken the positive square root. Since

$$\det[g'_{\mu\nu}] = \det[(a^{-1})^\rho_\mu(a^{-1})^\sigma_\nu g_{\rho\sigma}] = (\det\mathbf{A})^{-2}\det[g_{\mu\nu}], \tag{10.76}$$

we see that

$$\sqrt{g'} = |\det\mathbf{A}|^{-1}\sqrt{g} \tag{10.77}$$

Thus $\sqrt{g}$ is also not quite an invariant. This is only to be expected, because $\mathbf{g}(\ ,\ )$ is a quadratic form and we know that there is no basis-independent meaning to the determinant of such an object.

Now define

$$\varepsilon_{\mu_1\mu_2\dots\mu_n} = \sqrt{g}\,\epsilon_{\mu_1\mu_2\dots\mu_n}, \tag{10.78}$$

and assume that $\varepsilon_{\mu_1\mu_2\dots\mu_n}$ has the type $(0,n)$ tensor character implied by its indices. When we look at how this transforms, and restrict ourselves to *orientation preserving* changes of of bases, *i.e.* ones for which $\det\mathbf{A}$ is positive, we see that factors of $\det\mathbf{A}$ conspire to give

$$\varepsilon'_{\mu_1\mu_2\dots\mu_n} = \sqrt{g'}\,\epsilon_{\mu_1\mu_2\dots\mu_n}. \tag{10.79}$$

A similar exercise indictes that if we define $\epsilon^{\mu_1\mu_2\dots i_n}$ to be numerically equal to $\epsilon_{i_1 i_2\dots\mu_n}$ then

$$\varepsilon^{\mu_1\mu_2\dots\mu_n} = \frac{1}{\sqrt{g}}\,\epsilon^{\mu_1\mu_2\dots\mu_n} \tag{10.80}$$

also transforms as a tensor — in this case a type $(n, 0)$ contravariant one — provided that the factor of $1/\sqrt{g}$ is always calculated with respect to the current basis.

If the dimension n is even and we are given a skew-symmetric tensor $F_{\mu\nu}$, we can therefore construct an invariant

$$\varepsilon^{\mu_1\mu_2\dots\mu_n} F_{\mu_1\mu_2} \cdots F_{\mu_{n-1}\mu_n} = \frac{1}{\sqrt{g}} \epsilon^{\mu_1\mu_2\dots\mu_n} F_{\mu_1\mu_2} \cdots F_{\mu_{n-1}\mu_n}. \tag{10.81}$$

Similarly, given an skew-symmetric covariant tensor $F_{\mu_1\dots\mu_m}$ with m $(\leq n)$ indices we can form its *dual*, denoted by F^*, a $(n-m)$-contravariant tensor with components

$$(F^*)^{\mu_{m+1}\dots\mu_n} = \frac{1}{m!} \varepsilon^{\mu_1\mu_2\dots\mu_n} F_{\mu_1\dots\mu_m} = \frac{1}{\sqrt{g}} \frac{1}{m!} \epsilon^{\mu_1\mu_2\dots\mu_n} F_{\mu_1\dots\mu_m}. \tag{10.82}$$

We meet this “dual” tensor again, when we study differential forms.

10.3 Cartesian tensors

If we restrict ourselves to Cartesian co-ordinate systems having orthonormal basis vectors, so that $g_{ij} = \delta_{ij}$, then there are considerable simplifications. In particular, we do not have to make a distinction between co- and contravariant indices. We shall usually write their indices as roman-alphabet suffixes.

A change of basis from one orthogonal n-dimensional basis $\mathbf{e}_i$ to another $\mathbf{e}'_i$ will set

$$\mathbf{e}'_i = O_{ij} \mathbf{e}_j, \tag{10.83}$$

where the numbers O_{ij} are the entries in an *orthogonal* matrix $\mathbf{O}$, *i.e.* a real matrix obeying $\mathbf{O}^T\mathbf{O} = \mathbf{O}\mathbf{O}^T = \mathbf{I}$, where T denotes the transpose. The set of n-by-n orthogonal matrices constitutes the *orthogonal group* $\mathrm{O}(n)$.

10.3.1 Isotropic tensors

The Kronecker δ_{ij} with both indices downstairs is unchanged by $\mathrm{O}(n)$ transformations,

$$\delta'_{ij} = O_{ik} O_{jl} \delta_{kl} = O_{ik} O_{jk} = O_{ik} O^T_{kj} = \delta_{ij}, \tag{10.84}$$

and has the same components in any Cartesian frame. We say that its components are *numerically invariant*. A similar property holds for tensors made up of products of δ_{ij}, such as

$$T_{ijklmn} = \delta_{ij}\delta_{kl}\delta_{mn}. \tag{10.85}$$

It is possible to show[5] that any tensor whose components are numerically invariant under all orthogonal transformations is a sum of products of this form. The most general O(n) invariant tensor of rank four is, for example.

$$\alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{lj} + \gamma\delta_{il}\delta_{jk}. \tag{10.86}$$

The determinant of an orthogonal transformation must be ± 1. If we only allow orientation-preserving changes of basis then we restrict ourselves to orthogonal transformations O_{ij} with $\det \mathbf{O} = 1$. These are the *proper* orthogonal transformations. In n dimensions they constitute the group SO(n). Under SO(n) transformations, both δ_{ij} and $\epsilon_{i_1 i_2 \ldots i_n}$ are numerically invariant and the most general SO(n) invariant tensors consist of sums of products of δ_{ij}'s and $\epsilon_{i_1 i_2 \ldots i_n}$'s. The most general SO(4)-invariant rank-four tensor is, for example,

$$\alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{lj} + \gamma\delta_{il}\delta_{jk} + \lambda\epsilon_{ijkl}. \tag{10.87}$$

Tensors that are numerically invariant under SO(n) are known as *isotropic tensors*.

As there is no longer any distinction between co- and contravariant indices, we can now contract any pair of indices. In three dimensions, for example,

$$B_{ijkl} = \epsilon_{nij}\epsilon_{nkl} \tag{10.88}$$

is a rank-four isotropic tensor. Now $\epsilon_{i_1 \ldots i_n}$ is *not* invariant when we transform via an orthogonal transformation with $\det \mathbf{O} = -1$, but the product of two ϵ's *is* invariant under such transformations. The tensor B_{ijkl} is therefore numerically invariant under the larger group O(3) and must be expressible as

$$B_{ijkl} = \alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{lj} + \gamma\delta_{il}\delta_{jk} \tag{10.89}$$

for some coefficients α, β and γ. The following exercise explores some consequences of this and related facts.

---

[5] The proof is surprisingly complicated. See, for example, M. Spivak, *A Comprehensive Introduction to Differential Geometry* (second edition) Vol. V, pp. 466-481.

Exercise 10.5: We defined the n-dimensional Levi-Civita symbol by requiring that $\epsilon_{i_1 i_2 \ldots i_n}$ be antisymmetric in all pairs of indices, and $\epsilon_{12\ldots n} = 1$.

a) Show that $\epsilon_{123} = \epsilon_{231} = \epsilon_{312}$, but that $\epsilon_{1234} = -\epsilon_{2341} = \epsilon_{3412} = -\epsilon_{4123}$.
b) Show that
$$\epsilon_{ijk}\epsilon_{i'j'k'} = \delta_{ii'}\delta_{jj'}\delta_{kk'} + \text{five other terms},$$
where you should write out all six terms explicitly.
c) Show that $\epsilon_{ijk}\epsilon_{ij'k'} = \delta_{jj'}\delta_{kk'} - \delta_{jk'}\delta_{kj'}$.
d) For dimension $n = 4$, write out $\epsilon_{ijkl}\epsilon_{ij'k'l'}$ as a sum of products of δ's similar to the one in part (c).

Exercise 10.6: *Vector Products*. The vector product of two three-vectors may be written in Cartesian components as $(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k$. Use this and your results about ϵ_{ijk} from the previous exercise to show that

i) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$,
ii) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$,
iii) $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$.
iv) If we take $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$, with $\mathbf{d} \equiv \mathbf{b}$, to be unit vectors, show that the identities (i) and (iii) become the sine and cosine rule, respectively, of spherical trigonometry. (Hint: for the spherical sine rule, begin by showing that $\mathbf{a} \cdot [(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{c})] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.)

10.3.2 Stress and strain

As an illustration of the utility of Cartesian tensors, we consider their application to elasticity.

Suppose that an elastic body is slightly deformed so that the particle that was originally at the point with Cartesian co-ordinates x_i is moved to $x_i + \eta_i$. We define the (infinitesimal) *strain tensor* e_{ij} by

$$e_{ij} = \frac{1}{2}\left(\frac{\partial \eta_j}{\partial x_i} + \frac{\partial \eta_i}{\partial x_j}\right). \tag{10.90}$$

It is automatically symmetric: $e_{ij} = e_{ji}$. We will leave for later (exercise 11.3) a discussion of why this is the natural definition of strain, and also the modifications necessary were we to employ a non-Cartesian co-ordinate system.

To define the *stress tensor* σ_{ij} we consider the portion Ω of the body in figure 10.1, and an element of area $dS = \mathbf{n}\, d|S|$ on its boundary. Here, $\mathbf{n}$ is the unit normal vector pointing out of Ω. The force $\mathbf{F}$ exerted on this surface element by the parts of the body exterior to Ω has components

$$F_i = \sigma_{ij} n_j \, d|S|. \tag{10.91}$$

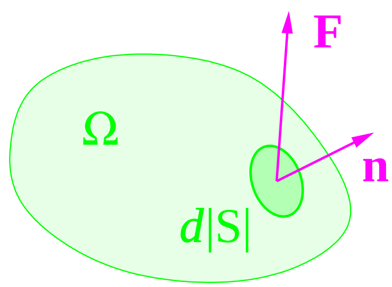

Figure 10.1: *Stress forces.*

That $\mathbf{F}$ is a linear function of $\mathbf{n}\, d|S|$ can be seen by considering the forces on an small tetrahedron, three of whose sides coincide with the co-ordinate planes, the fourth side having $\mathbf{n}$ as its normal. In the limit that the lengths of the sides go to zero as ϵ, the mass of the body scales to zero as ϵ^3, but the forces are proprtional to the areas of the sides and go to zero only as ϵ^2. Only if the linear relation holds true can the acceleration of the tetrahedron remain finite. A similar argument applied to torques and the moment of inertia of a small cube shows that $\sigma_{ij} = \sigma_{ji}$.

A generalization of Hooke's law,

$$\sigma_{ij} = c_{ijkl} e_{kl}, \tag{10.92}$$

relates the stress to the strain via the tensor of *elastic constants* c_{ijkl}. This rank-four tensor has the symmetry properties

$$c_{ijkl} = c_{klij} = c_{jikl} = c_{ijlk}. \tag{10.93}$$

In other words, the tensor is symmetric under the interchange of the first and second pairs of indices, and also under the interchange of the individual indices in either pair.

For an isotropic material — a material whose properties are invariant under the rotation group SO(3) — the tensor of elastic constants must be an isotropic tensor. The most general such tensor with the required symmetries is

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{10.94}$$

As isotropic material is therefore characterized by only two independent parameters, λ and μ. These are called the *Lamé* constants after the mathematical engineer Gabriel Lamé. In terms of them the generalized Hooke's law becomes

$$\sigma_{ij} = \lambda\delta_{ij}e_{kk} + 2\mu e_{ij}. \tag{10.95}$$

By considering particular deformations, we can express the more directly measurable *bulk modulus*, *shear modulus*, *Young's modulus* and *Poisson's ratio* in terms of λ and μ.

The bulk modulus κ is defined by

$$dP = -\kappa\frac{dV}{V}, \tag{10.96}$$

where an infinitesimal isotropic external pressure dP causes a change $V \to V + dV$ in the volume of the material. This applied pressure corresponds to a surface stress of $\sigma_{ij} = -\delta_{ij}\, dP$. An isotropic expansion displaces points in the material so that

$$\eta_i = \frac{1}{3}\frac{dV}{V}x_i. \tag{10.97}$$

The strains are therefore given by

$$e_{ij} = \frac{1}{3}\delta_{ij}\frac{dV}{V}. \tag{10.98}$$

Inserting this strain into the stress-strain relation gives

$$\sigma_{ij} = \delta_{ij}(\lambda + \frac{2}{3}\mu)\frac{dV}{V} = -\delta_{ij}dP. \tag{10.99}$$

Thus

$$\kappa = \lambda + \frac{2}{3}\mu. \tag{10.100}$$

To define the shear modulus, we assume a deformation $\eta_1 = \theta x_2$, so $e_{12} = e_{21} = \theta/2$, with all other e_{ij} vanishing.

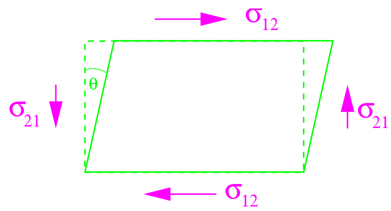

Figure 10.2: *Shear strain. The arrows show the direction of the applied stresses. The* σ_{21} *on the vertical faces are necessary to stop the body rotating.*

The applied shear stress is $\sigma_{12} = \sigma_{21}$. The shear modulus, is defined to be σ_{12}/θ. Inserting the strain components into the stress-strain relation gives

$$\sigma_{12} = \mu\theta, \tag{10.101}$$

and so the shear modulus is equal to the Lamé constant μ. We can therefore write the generalized Hooke's law as

$$\sigma_{ij} = 2\mu(e_{ij} - \tfrac{1}{3}\delta_{ij}e_{kk}) + \kappa e_{kk}\delta_{ij}, \tag{10.102}$$

which reveals that the shear modulus is associated with the traceless part of the strain tensor, and the bulk modulus with the trace.

Young's modulus Y is measured by stretching a wire of initial length L and square cross section of side W under a tension $T = \sigma_{33}W^2$.

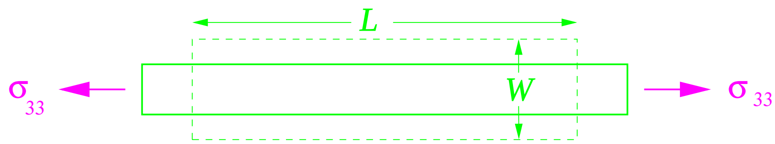

Figure 10.3: *Forces on a stretched wire.*

We define Y so that

$$\sigma_{33} = Y\frac{dL}{L}. \tag{10.103}$$

At the same time as the wire stretches, its width changes $W \to W + dW$. Poisson's ratio σ is defined by

$$\frac{dW}{W} = -\sigma\frac{dL}{L}, \tag{10.104}$$

so that σ is positive if the wire gets thinner as it gets longer. The displacements are

$$\begin{aligned} \eta_3 &= z\left(\frac{dL}{L}\right), \\ \eta_1 &= x\left(\frac{dW}{W}\right) = -\sigma x\left(\frac{dL}{L}\right), \\ \eta_2 &= y\left(\frac{dW}{W}\right) = -\sigma y\left(\frac{dL}{L}\right), \end{aligned} \tag{10.105}$$

so the strain components are

$$e_{33} = \frac{dL}{L}, \quad e_{11} = e_{22} = \frac{dW}{W} = -\sigma e_{33}. \tag{10.106}$$

We therefore have

$$\sigma_{33} = (\lambda(1-2\sigma) + 2\mu)\left(\frac{dL}{L}\right), \tag{10.107}$$

leading to

$$Y = \lambda(1-2\sigma) + 2\mu. \tag{10.108}$$

Now, the side of the wire is a free surface with no forces acting on it, so

$$0 = \sigma_{22} = \sigma_{11} = (\lambda(1-2\sigma) - 2\sigma\mu)\left(\frac{dL}{L}\right). \tag{10.109}$$

This tells us that[6]

$$\sigma = \frac{1}{2}\frac{\lambda}{\lambda+\mu}, \tag{10.110}$$

and

$$Y = \mu\left(\frac{3\lambda + 2\mu}{\lambda+\mu}\right). \tag{10.111}$$

Other relations, following from those above, are

$$\begin{aligned} Y &= 3\kappa(1-2\sigma), \\ &= 2\mu(1+\sigma). \end{aligned} \tag{10.112}$$

[6]Poisson and Cauchy erroneously believed that $\lambda = \mu$, and hence that $\sigma = 1/4$.

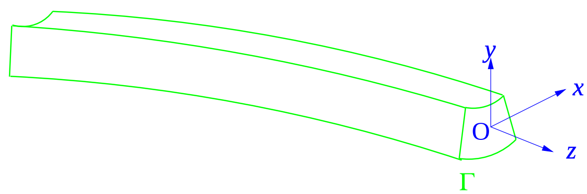

Figure 10.4: *Bent beam.*

Exercise 10.7: Show that the symmetries

$$c_{ijkl} = c_{klij} = c_{jikl} = c_{ijlk}$$

imply that a general homogeneous material has 21 independent elastic constants. (This result was originally obtained by George Green, of Green function fame.)

Exercise 10.8: A steel beam is forged so that its cross section has the shape of a region $\Gamma \in \mathbb{R}^2$. When undeformed, it lies along the z axis. The centroid O of each cross section is defined so that

$$\int_\Gamma x\,dxdy = \int_\Gamma y\,dxdy = 0,$$

when the co-ordinates x, y are taken with the centroid O as the origin. The beam is slightly bent away from the z axis so that the line of centroids remains in the y, z plane. (See figure 10.4) At a particular cross section with centroid O, the line of centroids has radius of curvature R.

Assume that the deformation in the vicinity of O is such that

$$\begin{aligned}\eta_x &= -\frac{\sigma}{R}xy,\\ \eta_y &= \frac{1}{2R}\left\{\sigma(x^2-y^2)-z^2\right\},\\ \eta_z &= \frac{1}{R}yz.\end{aligned}$$

Observe that for this assumed deformation, and for a positive Poisson ratio, the cross section deforms *anticlastically* — the sides bend *up* as the beam bends *down*. This is shown in figure 10.5.

Compute the strain tensor resulting from the assumed deformation, and show that its only non-zero components are

$$e_{xx} = -\frac{\sigma}{R}y, \qquad e_{yy} = -\frac{\sigma}{R}y, \qquad e_{zz} = \frac{1}{R}y.$$

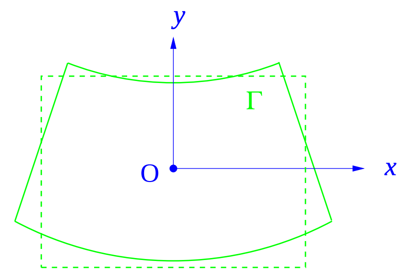

Figure 10.5: *The original (dashed) and anticlastically deformed (full) cross-section.*

Next, show that

$$\sigma_{zz} = \left(\frac{Y}{R}\right) y,$$

and that all other components of the stress tensor vanish. Deduce from this vanishing that the assumed deformation satisfies the free-surface boundary condition, and so is indeed the way the beam responds when it is bent by forces applied at its ends.

The work done in bending the beam

$$\int_{\text{beam}} \frac{1}{2} e_{ij} c_{ijkl} e_{kl} \, d^3x$$

is stored as elastic energy. Show that for our bent rod this energy is equal to

$$\int \frac{YI}{2}\left(\frac{1}{R^2}\right) ds \approx \int \frac{YI}{2}(y'')^2 dz,$$

where s is the arc-length taken along the line of centroids of the beam,

$$I = \int_\Gamma y^2 \, dxdy$$

is the moment of inertia of the region Γ about the x axis, and y'' denotes the second derivative of the deflection of the beam with respect to z (which approximates the arc-length). This last formula for the strain energy has been used in a number of our calculus-of-variations problems.

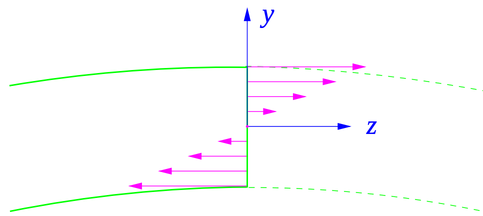

Figure 10.6: *The distribution of forces σ_{zz} exerted on the left-hand part of the bent rod by the material to its right.*

10.3.3 Maxwell stress tensor

Consider a small cubical element of an elastic body. If the stress tensor were position independent, the external forces on each pair of opposing faces of the cube would be equal in magnitude but pointing in opposite directions. There would therefore be no net external force on the cube. When σ_{ij} is *not* constant then we claim that the total force acting on an infinitesimal element of volume dV is

$$F_i = \partial_j \sigma_{ij}\, dV. \tag{10.113}$$

To see that this assertion is correct, consider a finite region Ω with boundary $\partial\Omega$, and use the divergence theorem to write the total force on Ω as

$$F_i^{\text{tot}} = \int_{\partial\Omega} \sigma_{ij} n_j d|S| = \int_{\Omega} \partial_j \sigma_{ij} dV. \tag{10.114}$$

Whenever the force-per-unit-volume f_i acting on a body can be written in the form $f_i = \partial_j \sigma_{ij}$, we refer to σ_{ij} as a "stress tensor," by analogy with stress in an elastic solid. As an example, let $\mathbf{E}$ and $\mathbf{B}$ be electric and magnetic fields. For simplicity, initially assume them to be static. The force per unit volume exerted by these fields on a distribution of charge ρ and current $\mathbf{j}$ is

$$\mathbf{f} = \rho \mathbf{E} + \mathbf{j} \times \mathbf{B}. \tag{10.115}$$

From Gauss' law $\rho = \operatorname{div} \mathbf{D}$, and with $\mathbf{D} = \epsilon_0 \mathbf{E}$, we find that the force per unit volume due the electric field has components

$$\begin{aligned} \rho E_i = (\partial_j D_j) E_i &= \epsilon_0 \Big(\partial_j (E_i E_j) - E_j\, \partial_j E_i \Big) \\ &= \epsilon_0 \Big(\partial_j (E_i E_j) - E_j\, \partial_i E_j \Big) \\ &= \epsilon_0 \partial_j \left(E_i E_j - \frac{1}{2} \delta_{ij} |E|^2 \right). \end{aligned} \tag{10.116}$$

Here, in passing from the first line to the second, we have used the fact that curl $\mathbf{E}$ is zero for static fields, and so $\partial_j E_i = \partial_i E_j$. Similarly, using $\mathbf{j} = \mathrm{curl}\,\mathbf{H}$, together with $\mathbf{B} = \mu_0 \mathbf{H}$ and $\mathrm{div}\,\mathbf{B} = 0$, we find that the force per unit volume due the magnetic field has components

$$(\mathbf{j} \times \mathbf{B})_i = \mu_0 \partial_j \left(H_i H_j - \frac{1}{2}\delta_{ij}|H|^2 \right). \tag{10.117}$$

The quantity

$$\sigma_{ij} = \epsilon_0 \left(E_i E_j - \frac{1}{2}\delta_{ij}|E|^2 \right) + \mu_0 \left(H_i H_j - \frac{1}{2}\delta_{ij}|H|^2 \right) \tag{10.118}$$

is called the *Maxwell stress tensor*. Its utility lies in in the fact that the total electromagnetic force on an isolated body is the integral of the Maxwell stress over its surface. We do not need to know the fields within the body.

Michael Faraday was the first to intuit a picture of electromagnetic stresses and attributed both a longitudinal tension and a mutual lateral repulsion to the field lines. Maxwell's tensor expresses this idea mathematically.

Exercise 10.9: Allow the fields in the preceding calculation to be time dependent. Show that Maxwell's equations

$$\begin{aligned} \mathrm{curl}\,\mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \quad \mathrm{div}\,\mathbf{B} = 0, \\ \mathrm{curl}\,\mathbf{H} &= \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}, \quad \mathrm{div}\,\mathbf{D} = \rho, \end{aligned}$$

with $\mathbf{B} = \mu_0 \mathbf{H}$, $\mathbf{D} = \epsilon_0 \mathbf{E}$, and $c = 1/\sqrt{\mu_0 \epsilon_0}$, lead to

$$(\rho \mathbf{E} + \mathbf{j} \times \mathbf{B})_i + \frac{\partial}{\partial t}\left\{ \frac{1}{c^2}(\mathbf{E} \times \mathbf{H})_i \right\} = \partial_j \sigma_{ij}.$$

The left-hand side is the time rate of change of the mechanical (first term) and electromagnetic (second term) momentum density. Observe that we can equivalently write

$$\frac{\partial}{\partial t}\left\{ \frac{1}{c^2}(\mathbf{E} \times \mathbf{H})_i \right\} + \partial_j(-\sigma_{ij}) = -(\rho \mathbf{E} + \mathbf{j} \times \mathbf{B})_i,$$

and think of this a local field-momentum conservation law. In this interpretation $-\sigma_{ij}$ is thought of as the *momentum flux* tensor, its entries being the flux in direction j of the component of field momentum in direction i. The term on the right-hand side is the rate at which momentum is being supplied to the electro-magnetic field by the charges and currents.

10.4 Further exercises and problems

Exercise 10.10: *Quotient theorem.* Suppose that you have come up with some recipe for generating an array of numbers T^{ijk} in any co-ordinate frame, and want to know whether these numbers are the components of a triply contravariant tensor. Suppose further that you know that, given the components a_{ij} of an arbitrary doubly covariant tensor, the numbers

$$T^{ijk}a_{jk} = v^i$$

transform as the components of a contravariant vector. Show that T^{ijk} does indeed transform as a triply contravariant tensor. (The natural generalization of this result to arbitrary tensor types is known as the *quotient theorem*.)

Exercise 10.11: Let $T^i{}_j$ be the 3-by-3 array of components of a tensor. Show that the quantities

$$a = T^i{}_i, \quad b = T^i{}_j T^j{}_i, \quad c = T^i{}_j T^j{}_k T^k{}_i$$

are invariant. Further show that the eigenvalues of the linear map represented by the matrix $T^i{}_j$ can be found by solving the cubic equation

$$\lambda^3 - a\lambda^2 + \frac{1}{2}(a^2 - b)\lambda - \frac{1}{6}(a^3 - 3ab + 2c) = 0.$$

Exercise 10.12: Let the covariant tensor R_{ijkl} possess the following symmetries:

i) $R_{ijkl} = -R_{jikl}$,
ii) $R_{ijkl} = -R_{ijlk}$,
iii) $R_{ijkl} + R_{iklj} + R_{iljk} = 0$.

Use the properties i),ii), iii) to show that:

a) $R_{ijkl} = R_{klij}$.
b) If $R_{ijkl}x^iy^jx^ky^l = 0$ for all vectors x^i, y^i, then $R_{ijkl} = 0$.
c) If B_{ij} is a symmetric covariant tensor and set we $A_{ijkl} = B_{ik}B_{jl} - B_{il}B_{jk}$, then A_{ijkl} has the same symmetries as R_{ijkl}.

Exercise 10.13: Write out Euler's equation for fluid motion

$$\dot{\mathbf{v}} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\nabla h$$

in Cartesian tensor notation. Transform it into

$$\dot{\mathbf{v}} - \mathbf{v} \times \boldsymbol{\omega} = -\nabla \left(\frac{1}{2}\mathbf{v}^2 + h\right),$$

where $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ is the vorticity. Deduce Bernoulli's theorem, that for steady ($\dot{\mathbf{v}} = 0$) flow the quantity $\frac{1}{2}\mathbf{v}^2 + h$ is constant along streamlines.

Exercise 10.14: The elastic properties of an infinite homogeneous and isotropic solid of density ρ are described by Lamé constants λ and μ. Show that the equation of motion for small-amplitude vibrations is

$$\rho \frac{\partial^2 \eta_i}{\partial t^2} = (\lambda + \mu) \frac{\partial^2 \eta_j}{\partial x_i \partial x_j} + \mu \frac{\partial^2 \eta_i}{\partial x_j^2}.$$

Here η_i are the cartesian components of the displacement vector $\boldsymbol{\eta}(\mathbf{x}, t)$ of the particle initially at the point $\mathbf{x}$. Seek plane wave solutions of the form

$$\boldsymbol{\eta} = \mathbf{a} \exp\{i\mathbf{k} \cdot \mathbf{x} - i\omega t\},$$

and deduce that there are two possible types of wave: longitudinal "P-waves," which have phase velocity

$$v_{\mathrm{P}} = \sqrt{\frac{\lambda + 2\mu}{\rho}},$$

and transverse "S-waves," which have phase velocity

$$v_{\mathrm{S}} = \sqrt{\frac{\mu}{\rho}}.$$

Exercise 10.15: *Symmetric integration.* Show that the n-dimensional integral

$$I_{\alpha\beta\gamma\delta} = \int \frac{d^n k}{(2\pi)^n} (k_\alpha k_\beta k_\gamma k_\delta)\, f(k^2),$$

is equal to

$$A(\delta_{\alpha\beta}\delta_{\gamma\delta} + \delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma})$$

where

$$A = \frac{1}{n(n+2)} \int \frac{d^n k}{(2\pi)^n} (k^2)^2 f(k^2).$$

Similarly evaluate

$$I_{\alpha\beta\gamma\delta\epsilon} = \int \frac{d^n k}{(2\pi)^n} (k_\alpha k_\beta k_\gamma k_\delta k_\epsilon)\, f(k^2).$$

Exercise 10.16: Write down the most general three-dimensional isotropic tensors of rank two and three.

In piezoelectric materials, the application of an electric field E_i induces a mechanical strain that is described by a rank-two symmetric tensor

$$e_{ij} = d_{ijk} E_k,$$

where d_{ijk} is a third-rank tensor that depends only on the material. Show that e_{ij} can only be non-zero in an anisotropic material.

Exercise 10.17: In three dimensions, a rank-five isotropic tensor T_{ijklm} is a linear combination of expressions of the form $\epsilon_{i_1 i_2 i_3} \delta_{i_4 i_5}$ for some assignment of the indices i, j, k, l, m to the $i_1, \ldots, i_5$. Show that, on taking into account the symmetries of the Kronecker and Levi-Civita symbols, we can construct *ten* distinct products $\epsilon_{i_1 i_2 i_3} \delta_{i_4 i_5}$. Only *six* of these are linearly independent, however. Show, for example, that

$$\epsilon_{ijk}\delta_{lm} - \epsilon_{jkl}\delta_{im} + \epsilon_{kli}\delta_{jm} - \epsilon_{lij}\delta_{km} = 0,$$

and find the three other independent relations of this sort.[7]

(*Hint*: Begin by showing that, in three dimensions,

$$\delta^{i_1 i_2 i_3 i_4}_{i_5 i_6 i_7 i_8} \stackrel{\text{def}}{=} \begin{vmatrix} \delta_{i_1 i_5} & \delta_{i_1 i_6} & \delta_{i_1 i_7} & \delta_{i_1 i_8} \\ \delta_{i_2 i_5} & \delta_{i_2 i_6} & \delta_{i_2 i_7} & \delta_{i_2 i_8} \\ \delta_{i_3 i_5} & \delta_{i_3 i_6} & \delta_{i_3 i_7} & \delta_{i_3 i_8} \\ \delta_{i_4 i_5} & \delta_{i_4 i_6} & \delta_{i_4 i_7} & \delta_{i_4 i_8} \end{vmatrix} = 0,$$

and contract with $\epsilon_{i_6 i_7 i_8}$.)

Problem 10.18: *The Plücker Relations.* This problem provides a challenging test of your understanding of linear algebra. It leads you through the task of deriving the necessary and sufficient conditions for

$$\mathbf{A} = A^{i_1 \ldots i_k}\, \mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_k} \in \bigwedge{}^k V$$

to be decomposable as

$$\mathbf{A} = \mathbf{f}_1 \wedge \mathbf{f}_2 \wedge \ldots \wedge \mathbf{f}_k.$$

The trick is to introduce two subspaces of V,

[7]Such relations are called *syzygies*. A recipe for constructing linearly independent basis sets of isotropic tensors can be found in: G. F. Smith, *Tensor*, **19** (1968) 79-88.

i) W, the smallest subspace of V such that $\mathbf{A} \in \bigwedge^k W$,
ii) $W' = \{\mathbf{v} \in V : \mathbf{v} \wedge \mathbf{A} = 0\}$,

and explore their relationship.

a) Show that if $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ constitute a basis for W', then

$$\mathbf{A} = \mathbf{w}_1 \wedge \mathbf{w}_2 \wedge \cdots \wedge \mathbf{w}_n \wedge \boldsymbol{\varphi}$$

for some $\boldsymbol{\varphi} \in \bigwedge^{k-n} V$. Conclude that that $W' \subseteq W$, and that equality holds if and only if $\mathbf{A}$ is decomposable, in which case $W = W' = \operatorname{span}\{\mathbf{f}_1 \ldots \mathbf{f}_k\}$.

b) Now show that W is the image space of $\bigwedge^{k-1} V^*$ under the map that takes

$$\boldsymbol{\Xi} = \Xi_{i_1 \ldots i_{k-1}} \mathbf{e}^{*i_1} \wedge \ldots \wedge \mathbf{e}^{*i_{k-1}} \in \bigwedge{}^{k-1} V^*$$

to

$$i(\boldsymbol{\Xi})\mathbf{A} \stackrel{\text{def}}{=} \Xi_{i_1 \ldots i_{k-1}} A^{i_1 \ldots i_{k-1} j} \mathbf{e}_j \in V$$

Deduce that the condition $W \subseteq W'$ is that

$$\Big(i(\boldsymbol{\Xi})\mathbf{A}\Big) \wedge \mathbf{A} = 0, \quad \forall \boldsymbol{\Xi} \in \bigwedge{}^{k-1} V^*.$$

c) By taking

$$\boldsymbol{\Xi} = \mathbf{e}^{*i_1} \wedge \ldots \wedge \mathbf{e}^{*i_{k-1}},$$

show that the condition in part b) can be written as

$$A^{i_1 \ldots i_{k-1} j_1} A^{j_2 j_3 \ldots j_{k+1}} \mathbf{e}_{j_1} \wedge \ldots \wedge \mathbf{e}_{j_{k+1}} = 0.$$

Deduce that the necessary and sufficient conditions for decomposibility are that

$$A^{i_1 \ldots i_{k-1} [j_1} A^{j_2 j_3 \ldots j_{k+1}]} = 0,$$

for all possible index sets $i_1, \ldots, i_{k-1}, j_1, \ldots j_{k+1}$. Here $[\ldots]$ denotes anti-symmetrization of the enclosed indices.