7   Tensors

A tensor is something that transforms like a tensor. The purpose of this section is to give a less concise but, perhaps, ultimately more helpful discussion of the properties of these objects.

7.1   What it Takes to Make a Tensor

Not any list of n numbers constitutes a vector in $\mathbb{R}^n$. For example, if you write down the heights of the first three people you met this morning, that doesn't make a vector in $\mathbb{R}^3$. Instead, a vector comes with certain responsibilities. In particular, the underlying object should be independent of the choice of basis which, as we now explain, means that the components should transform in the right way under rotations.

We consider a point $\mathbf{x} \in \mathbb{R}^n$. If we wish to attach some coordinates to this point, we first need to introduce a set of basis vectors $\{\mathbf{e}_i\}$ with $i = 1, \ldots, n$. We will take these to be orthonormal, meaning that $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. Any vector can then be expressed as

$$\mathbf{x} = x_i \mathbf{e}_i \tag{7.1}$$

Usually we conflate the coordinates $x_i = (x_1, \ldots, x_n)$ with the "vector". But, for our purposes, we should remember that these are just a useful way of representing the more abstract object $\mathbf{x}$. In particular, we're entirely at liberty to take a different set of basis vectors,

$$\mathbf{e}_i' = R_{ij} \mathbf{e}_j$$

If we ask that $\mathbf{e}_i'$ are also orthonormal, so $\mathbf{e}_i' \cdot \mathbf{e}_j' = \delta_{ij}$, then we have

$$\mathbf{e}_i' \cdot \mathbf{e}_j' = R_{ik} R_{jl} \, \mathbf{e}_k \cdot \mathbf{e}_l = R_{ik} R_{jk} = \delta_{ij}$$

or, in matrix notation,

$$R R^T = \mathbb{1}$$

Matrices of this kind are said to be *orthogonal*. We write $R \in O(n)$. Taking the determinant, we have $\det R = \pm 1$. Those matrices with $\det R = +1$ correspond to rotations and are said to be *special orthogonal*. We write $R \in SO(n)$. In $\mathbb{R}^3$, a rotation $R \in SO(3)$ takes a right-handed orthonormal basis into another right-handed orthonormal basis. Those matrices with $\det R = -1$ correspond to a rotation together with a reflection and take a right-handed basis to a left-handed basis.

Under a change of basis, the vector $\mathbf{x}$ itself doesn't change. But its components do. We have

$$\mathbf{x} = x_i \mathbf{e}_i = x_i' \mathbf{e}_i' = x_i' R_{ij} \mathbf{e}_j$$

So the components transform under the same rotation matrix R,

$$x_j = R_{ji} x_j' \quad \Rightarrow \quad x_i' = R_{ij} x_j$$

A *tensor T* is a generalisation of these ideas to an object with more indices. Just as the vector $\mathbf{x}$ has an identity independent of any choice of basis, so too does the tensor T. But when measured with respect to a chosen basis $\{\mathbf{e}_i\}$, a *tensor of rank p* has components $T_{i_1 \ldots i_p}$. When we change the basis using (7.1), the tensor transform as

$$T_{i_1 \ldots i_p}' = R_{i_1 j_1} \ldots R_{i_p j_p} T_{j_1 \ldots j_p} \tag{7.2}$$

This is known as the *tensor transformation rule*. A tensor of rank p is sometimes referred to simply as a *p-tensor*.

The simplest examples of tensors are very familiar. A tensor of rank 0 is just a number, or scalar, T. Under a rotation, it doesn't change: $T' = T$. A tensor of rank 1 is a vector, while a tensor of rank 2 is a matrix that transforms as $T_{ij}' = R_{ik} R_{jl} T_{jl}$.

7.1.1 Tensors as Maps

There is something a little strange about the definition of a tensor given above We first pick a set of coordinates, and the transformation law (7.2) then requires that the tensor transforms nicely so that, ultimately, nothing depends on these coordinates. But, if that's the case, surely there should be a definition of a tensor that doesn't rely on coordinates at all!

There is. A tensor T of rank p is a multi-linear map that takes p vectors, $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}$ and spits out a number in $\mathbb{R}$,

$$T(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}) = T_{i_1 i_2 \ldots i_p} a_{i_1} b_{i_2} \ldots c_{i_p} \tag{7.3}$$

Here "multi-linear" means that T is linear in each of the entries $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}$ individually. By evaluating T on a all possible vectors $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}$, we get the components $T_{i_1 i_2 \ldots i_p}$. The transformation rule (7.2) is simply the statement that the map T is independent of the choice of basis, and we can equally well write

$$T(\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}) = T_{i_1 i_2 \ldots i_p}' a_{i_1}' b_{i_2}' \ldots c_{i_p}'$$

Tensors as Maps Between Vectors

Rather than thinking of a tensor as a map from many vectors to $\mathbb{R}$, you can equivalently think of it as a map from some lower-rank tensor to another. For example, in (7.3), if you don't fill in the first entry, then a rank p tensor can equally well be viewed as taking $(p-1)$ vectors and spitting out a single vector

$$a_i = T_{i\,j_1\ldots j_{p-1}} = T_{ij_1\ldots j_{p-1}} b_{j_1} \ldots c_{j_{p-1}}$$

This is the way that tensors typically arise in physics or applied mathematics, where the most common example is simply a rank 2 tensor, defined as a map from one vector to another

$$\mathbf{u} = T\mathbf{v} \quad \Rightarrow \quad u_i = T_{ij}v_j$$

Until now, we've simply called T a matrix but for the equation $\mathbf{u} = T\mathbf{v}$ to make sense, M must transform as a tensor (7.2). This is inherited from the transformation rules of the vectors, $u'_i = R_{ij}u_j$ and $v'_i = R_{ij}v_j$, giving

$$u'_i = T'_{ij}v'_j \quad \text{with} \quad T'_{ij} = R_{ik}R_{jl}T_{kl}$$

Written as a matrix equation, this is $T' = RTR^T$.

Examples

Tensors of this kind arise, for example, in the description of how a current density $\mathbf{J}$ flows through a solid material due to the presence of an electric field $\mathbf{E}$,

$$J_i = \sigma_{ij}E_j$$

Here σ_{ij} is the *conductivity tensor* and this expression in known as *Ohm's law*. The underlying crystal lattice of a material can mean that electricity may preferentially flow in one direction rather than another. It is also possible that an electric field in, say, the $\hat{\mathbf{x}}$ direction can induce a current in the $\hat{\mathbf{y}}$ direction. All of this means that σ is 3×3 matrix, or tensor, rather than just a single number.

Another simple example arises in Newtonian mechanics. A *rigid body* rotating about the origin can be modelled by some number of masses m_a at positions $\mathbf{x}_a$, all moving with velocity $\dot{\mathbf{x}}_a = \boldsymbol{\omega} \times \mathbf{x}_a$. Here $\boldsymbol{\omega}$ is known as the angular velocity. The angular velocity $\boldsymbol{\omega}$ is related to the angular momentum $\mathbf{L}$ by

$$\mathbf{L} = I\boldsymbol{\omega} \tag{7.4}$$

with I the *inertia tensor*. The angular momentum does not necessarily lie parallel to the angular velocity and, correspondingly, I is in general a matrix, rather than a single number. In fact, we can easily derive an expression for the inertia tensor. The angular momentum is

$$\mathbf{L} = \sum_a m_a \mathbf{x}_a \times \dot{\mathbf{x}}_a = \sum_a m_a \mathbf{x}_a \times (\boldsymbol{\omega} \times \mathbf{x}_a) = \sum_a m_a \left(|\mathbf{x}_a|^2 \boldsymbol{\omega} - (\mathbf{x}_a \cdot \boldsymbol{\omega}) \mathbf{x}_a \right)$$

In components, $L_i = I_{ij} \omega_j$, where

$$I_{ij} = \sum_a m_a \left(|\mathbf{x}_a|^2 \delta_{ij} - (\mathbf{x}_a)_i (\mathbf{x}_a)_j \right)$$

For a continuous object with density $\rho(\mathbf{x})$, we can replace the sum with a volume integral

$$I_{ij} = \int_V \rho(\mathbf{x}) \left(|\mathbf{x}|^2 \delta_{ij} - x_i x_j \right) dV \tag{7.5}$$

So, for example, $I_{33} = \int \rho(x_1^2 + x_2^2) \, dV$ and $I_{12} = \int \rho x_1 x_2 \, dV$.

An Example of the Example

Consider a solid 3d cylinder of radius a and height $2L$, with uniform density ρ. The mass is $M = 2\pi a^2 L \rho$. We align the cylinder with the z-axis and work in cylindrical polar coordinates $x = r \cos \phi$ and $y = r \sin \phi$. The components of the inertia tensor are then

$$I_{33} = \int_V \rho(x_1^2 + x_2)^2 \, dV = \rho \int_0^{2\pi} d\phi \int_0^a dr \int_{-L}^{+L} dz \, r \, r^2 = \rho \pi L a^4$$

$$I_{11} = \int_V \rho(x_2^2 + x_3^2) \, dV = \rho \int_0^{2\pi} d\phi \int_0^a dr \int_{-L}^{+L} dz \, r(r^2 \sin^2 \phi + z^2) = \rho \pi a^2 L \left(\frac{a^2}{2} + \frac{2L^2}{3} \right)$$

By symmetry, $I_{22} = I_{11}$. For the off-diagonal elements, we have

$$I_{13} = -\int_V \rho x_1 x_3 \, dV = -\rho \int_0^a \int_0^{2\pi} \int_{-L}^{L} dz \, r^2 z \cos \phi = 0$$

where the integral vanishes due to the ϕ integration. Similarly, $I_{12} = I_{13} = 0$. We find that the inertia tensor for the cylinder is diagonal in these coordinates,

$$I = \text{diag}\left(M \left(\frac{a^2}{4} + \frac{L^2}{3} \right) , \; M \left(\frac{a^2}{4} + \frac{L^2}{3} \right) , \; \frac{1}{2} M a^2 \right) \tag{7.6}$$

Diagonalisation of a Symmetric Rank 2 Tensor

The inertia tensor I defined in (7.5) has a special property: it is symmetric

$$I_{ij} = I_{ji}$$

Any symmetric matrix I can always be diagonalised by an appropriate rotation. This means that there exists an $R \in SO(n)$ such that

$$I' = RIR^T = \text{diag}(I_1, \ldots, I_n)$$

Another way of saying this is that any symmetric rank 2 tensor has a basis of orthonormal eigenvectors $\{\mathbf{e}_i\}$, with I_i the corresponding eigenvalues.

In the case of the inertia tensor, the eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are called the *principal axes* of the solid. It means that any object, no matter how complicated, has its own preferred set of orthonormal axes embedded within it. If the object has some symmetry, then the principal axes will always be aligned with this symmetry. This, for example, was the case for the cylinder that we computed above where aligning the cylinder with the z-axis automatically gave us a diagonal inertia tensor (7.6).

In general, it will be less obvious where the principal axes lie. For example, the figure on the right shows the asteroid Toutatis, which is notable for its lumpy shape. The principal axes are shown embedded in the asteroid.
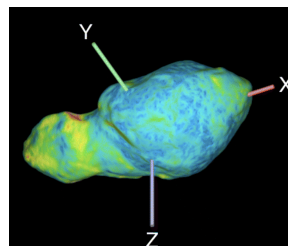
From (7.4), the angular momentum $\mathbf{L}$ is aligned with the angular velocity $\boldsymbol{\omega}$ only if a body spins about one of its principal axes. It turns out that, in this case, nice things happen and the body spins smoothly. However, if $\mathbf{L}$ and $\boldsymbol{\omega}$ are misaligned, the body exhibits more complicated tumbling, wobbling motion as it spins. You can learn all about this in the lectures on Classical Dynamics. (For what it's worth, Toutatis does not spin about a principal axes.)

7.1.2 Tensor Operations

Given a bunch of tensors, there are some manipulations that leave you with another tensor. Here we describe these operations.

- We can *add* and *subtract* tensors of the same rank, so if S and T are both tensors of rank p then so too is $S + T$. We can also multiply a tensor by a constant α and it remains a tensor.

- If S is a tensor of rank p and T a tensor of rank q, then the *tensor product* $S \otimes T$ is a tensor of rank $p + q$, defined by

$$(S \otimes T)_{i_1 \ldots i_p j_1 \ldots j_q} = S_{i_1 \ldots i_p} T_{j_1 \ldots j_q}$$

  You can check that the components of $(S \otimes T)$ do indeed satisfy the transformation rule (7.2). In particular, if we have p different vectors $\mathbf{a}$, $\mathbf{b}$, $\ldots$, $\mathbf{c}$ then we can construct a tensor

$$T = \mathbf{a} \otimes \mathbf{b} \otimes \ldots \otimes \mathbf{c} \quad \text{with} \quad T_{i_1 \ldots i_p} = a_{i_1} b_{i_2} \ldots c_{i_p}$$

- Given a tensor T of rank p, we can construct a new tensor S of rank $(p - 2)$ by *contracting* on two indices using δ_{ij},

$$S_{k_1 \ldots k_{p-2}} = \delta_{ij} T_{ijk_1 \ldots k_{p-2}}$$

  For a rank 2 tensor, the contraction is what we call the trace, $\text{Tr}\, T = T_{ii}$. It's a valid tensor operation because the end result is a scalar that does not transform under rotations

$$T'_{ii} = R_{ij} R_{ik} T_{jk} = \delta_{jl} T_{jk} = T_{ii}$$

  The same derivation shows that higher rank tensors can also be contracted, with the additional indices unaffected by the contraction.

The Quotient Rule

In practice, it's not hard to recognise a tensor when you see one. In any setting, they're usually just objects with a bunch of i and j indices, each of which clearly transforms as a vector. If in doubt, you can just check explicitly how the thing transforms. (There are cases where this check is needed. In later courses, you'll meet an object called the Levi-Civita connection Γ^i_{jk} which looks for all the world like a tensor but turns out, on closer inspection, to be something more subtle.)

There is a more formal way to say this. Let $T_{i_1 \ldots i_{p+q}}$ be a bunch of numbers that you think might comprise a tensor of rank $p + q$ in some coordinate basis. If $T_{i_1 \ldots i_{p+q}}$ are indeed the components of a tensor then you can feed it a rank q tensor $u_{j_1 \ldots j_q}$ and it will spit back a rank p tensor

$$v_{i_1 \ldots i_p} = T_{i_i \ldots i_p j_1 \ldots j_q} u_{j_1 \ldots j_q} \tag{7.7}$$

There is a converse to this statement. If for every tensor $u_{j_1 \ldots j_q}$, the output $v_{i_1 \ldots i_p}$ defined in (7.7) is a tensor, then $T_{i_i \ldots i_p j_1 \ldots j_q}$ are the components of a tensor. This is called the *quotient rule*.

It is straightforward, if a little fiddly, to prove the quotient rule. We can restrict attention to tensors u formed from the tensor product of vectors $u_{j_1...j_q} = c_{j_1} \ldots d_{j_q}$. Then, by assumption, $v_{i_1...i_p} = T_{i_1...i_p j_1...j_q} u_{j_1...j_q}$ is a tensor. If we then contract with p further vectors $\mathbf{a}, \ldots, \mathbf{b}$ then $v_{i_1...i_p} a_{i_1} \ldots b_{i_p} = T_{i_1...i_p j_1...j_q} a_{i_1} \ldots b_{i_p} c_{j_1} \ldots d_{j_q}$ is necessarily a scalar. This is then enough to ensure the correct transformation rule (7.2) for the components $T_{i_1...i_p j_1...j_q}$.

Symmetry and Anti-Symmetry

The symmetrisation properties of tensors are worthy of comment. A tensor of rank p that obeys

$$T_{ijp...q} = \pm T_{jip...q}$$

is said to be *symmetric* (for $+$) or *anti-symmetric* (for $-$) in the indices i and j. If a tensor is (anti)-symmetric in one coordinate system then it is (anti)-symmetric in any coordinate system

$$T'_{klr...s} = R_{ki} R_{lj} R_{rp} \ldots R_{sq} T_{ijp...q} = \pm R_{ki} R_{lj} R_{rp} \ldots R_{sq} T_{jip...q} = \pm T'_{lkr...s}$$

A tensor that is (anti)-symmetric in all pairs of indices is said to be *totally (anti)-symmetric*. Note that for tensors in $\mathbb{R}^n$, there are no anti-symmetric tensors of rank $p > n$ because at least one of the indices must take the same value and so the tensor necessarily vanishes. A totally anti-symmetric tensor of rank p in $\mathbb{R}^n$ has $\binom{n}{p}$ independent components.

Let's now restrict our attention to $\mathbb{R}^3$. A tensor of rank 2 is our new fancy name for a 3×3 matrix T_{ij}. In general, it has 9 independent components. We can always decompose it into the symmetric and anti-symmetric pieces

$$S_{ij} = \frac{1}{2}(T_{ij} + T_{ji}) \quad \text{and} \quad A_{ij} = \frac{1}{2}(T_{ij} - T_{ji})$$

which have 6 and 3 independent components respectively. Our discussion above shows that S and A are each, themselves, tensors. In fact, the symmetric piece can be decomposed further,

$$S_{ij} = P_{ij} + \frac{Q}{3}\delta_{ij}$$

where $Q = S_{ii}$ is the trace of S and carries a single degree of freedom, while P_{ij} is the traceless part of of S and carries 5.

In $\mathbb{R}^3$, we can also rewrite an anti-symmetric matrix in terms of a vector,

$$A_{ij} = \frac{1}{2}\epsilon_{ijk}B_k \quad \Longleftrightarrow \quad B_k = \frac{1}{2}\epsilon_{ijk}A_{ij}$$

The upshot is that in any 3×3 matrix can be decomposed as

$$T_{ij} = P_{ij} + \epsilon_{ijk}B_k + \frac{1}{3}\delta_i Q \tag{7.8}$$

where $P_{ii} = 0$.

7.1.3 Invariant Tensors

There are two important invariant tensors in $\mathbb{R}^n$. These are

- The rank 2 tensor δ_{ij}. This is invariant because

$$\delta'_{ij} = R_{ik}R_{jl}\delta_{kl} = \delta_{ij}$$

Note that this is invariant under any $R \in O(n)$. The invariance of δ_{ij} is the reason why the inner product $\mathbf{a} \cdot \mathbf{b} = \delta_{ij}a_i b_j$ between any two vectors is independent of the choice of basis.

- The rank n totally anti-symmetric tensor $\epsilon_{i_1 \dots i_n}$. This is defined by $\epsilon_{12 \dots n} = +1$. If you swap any two indices you get a minus sign. In particular, if any two indices are repeated, the epsilon symbol vanishes. This is invariant because

$$\epsilon'_{i_1 \dots i_n} = R_{i_1 j_1} \dots R_{i_n j_n}\epsilon_{j_1 \dots j_n} = \det R \, \epsilon_{i_1 \dots i_n} = \epsilon_{i_1 \dots i_n}$$

Note that the epsilon symbol is only invariant under $R \in SO(n)$. It picks up a minus sign under reflections. The invariance of ϵ_{ijk} in $\mathbb{R}^3$ is the reason why the cross-product $(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk}a_j b_k$ is itself a vector. Or, said differently, why the triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \epsilon_{ijk}a_i b_j c_k$ is independent of the choice of basis.

In general, a tensor is said to be invariant under a given rotation R if

$$T'_{i_1 \dots i_n} = R_{i_1 j_1} \dots R_{i_n j_n}T_{j_1 \dots j_n} = T_{i_1 \dots i_n}$$

A tensor that is invariant under all rotations R is said to be *isotropic*. Obviously all tensors of rank 0 are isotropic. What about higher rank tensors?

Claim: The only isotropic tensors in $\mathbb{R}^3$ of rank $p \leq 3$ are $T_{ij} = \alpha\delta_{ij}$ and $T_{ijk} = \beta\epsilon_{ijk}$ with α and β constant. In particular, there are no isotropic tensors of rank 1 (essentially because a vector always points in a preferred direction).

Proof: The idea is simply to look at how tensors transform under a bunch of specific rotations by π or $\pi/2$ about certain axes.

For example, consider a tensor of rank 1, so that

$$T_i' = R_{ij}T_j \quad \text{with} \quad R_{ij} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & +1 \end{pmatrix} \tag{7.9}$$

Requiring $T_i' = T_i$ gives $T_1 = T_2 = 0$. Clearly a similar argument, using a different R, also gives $T_3 = 0$.

For a tensor of rank 2, consider the transformation

$$T_{ij}' = \tilde{R}_{ik}\tilde{R}_{jl}T_{kl} \quad \text{with} \quad \tilde{R}_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & +1 \end{pmatrix} \tag{7.10}$$

which is a rotation by $\pi/2$ about the z-axis. The rotation gives $T_{13}' = T_{23}$ and $T_{23}' = -T_{13}$ so if $T_{ij}' = T_{ij}$, we must have $T_{13} = T_{23} = 0$. Meanwhile $T_{11}' = T_{22}$. Similar arguments tell us that all off-diagonal elements must vanish and all diagonal elements must be equal: $T_{11} = T_{22} = T_{33} = \alpha$ for some α. Hence $T_{ij} = \alpha\delta_{ij}$.

Finally, for a rank 3 tensor we have

$$T_{ijk}' = R_{il}R_{jp}R_{kq}T_{lpq}$$

If we pick R given in (7.9), then we find $T_{133}' = -T_{133}$ and $T_{111}' = -T_{111}$. Similar arguments show that an isotropic tensor must have $T_{ijk} = 0$ unless i, j and k are all distinct. Meanwhile, if we pick $R = \tilde{R}$ given in (7.10), then we get $T_{123}' = -T_{213}$. We end up with the result we wanted: T_{ijk} is isotropic if and only if $T_{ijk} = \beta\epsilon_{ijk}$ for some constant β. □

Although we won't prove it here, all other isotropic tensors can be formed from δ_{ij} and ϵ_{ijk}. For example, the only isotropic 4-tensor in $\mathbb{R}^3$ is

$$T_{ijkl} = \alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{jl} + \gamma\delta_{il}\delta_{jk}$$

with α, β and γ constants. You could try to cook up something involving ϵ_{ijk} but it doesn't give anything new. In particular, $\epsilon_{ijk}\epsilon_{ilp} = \delta_{jl}\delta_{kp} - \delta_{jp}\delta_{kl}$.

There is also an analogous result in $\mathbb{R}^n$: all isotropic tensors can be constructed from the symmetric 2-tensor δ_{ij} and the totally anti-symmetric n-tensor $\epsilon_{i_1...i_n}$.

An Example

The fact that there are only two invariant tensors can be very helpful in writing down the answers to problems. As an example, let's return to our conductivity tensor defined by

$$\mathbf{J} = \sigma \mathbf{E}$$

If we're dealing with a 3d material which we know is isotropic, with no direction different from any other, then the conductivity tensor must be proportional to the unit matrix, $\sigma_{ij} = \sigma \delta_{ij}$. Here σ is just referred to as the *conductivity*.

In contrast, suppose that we're dealing with a thin wafer of material, with both the current and electric field restricted to lie in a plane. In this case, the material is effectively two dimensional and there are two possible isotropic conductivity tensors

$$\sigma_{ij} = \sigma_{xx}\delta_{ij} + \sigma_{xy}\epsilon_{ij} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix}$$

Here σ_{xx} is called the *longitudinal conductivity* while σ_{xy} is called the *Hall conductivity*. If $\sigma_{xy} \neq 0$ then an electric field in the x-direction induces a current in the y-direction.

As an aside, it turn out that the seemingly mundane question of understanding σ_{xy} in real materials is closely tied to some of the most interesting breakthroughs in mathematics in recent decades! This is the subject of the Quantum Hall Effect.

Invariant Integrals

It is sometimes possible to use invariance properties to immediately write down the index structure of an integral, without doing the hard work of evaluating everything term by term. Suppose that we have some integral of the form

$$T_{ij...k} = \int_V f(r)x_i x_j \ldots x_k \, dV$$

with $r = |\mathbf{x}|$. Then under a rotation, we have

$$T'_{ij...k} = R_{ip}R_{jq}\ldots R_{kr}T_{ij...k} = \int_V f(r)x'_i x'_j \ldots x'_k \, dV$$

with, as usual, $x'_i = R_{ij}x_j$. But if we now change the integration variables to x', both $r = |\mathbf{x}| = |\mathbf{x}'|$ and $dV = dV'$ are invariant. (The latter because the Jacobian is $\det R = 1$). If the domain of integration is also rotationally invariant, so $V = V'$, then the final result must itself be an invariant tensor, $T'_{ij...k} = T_{ij...k}$.

As an example, consider the inertia tensor of a sphere of density $\rho(r)$ and radius R. We have

$$I_{ij} = \int_V \rho(r)(r^2 \delta_{ij} - x_i x_j) \, dV \qquad (7.11)$$

Clearly the first term gives rise to an isotropic tensor, δ_{ij}. But, by the argument above, so too does the second. This means that we must have

$$\int_V \rho(r) x_i x_j \, dV = \alpha \delta_{ij}$$

for some α. If we take the trace, we get

$$3\alpha = \int_V \rho(r) r^2 \, dV$$

Upon substitution back into (7.11), we have the

$$I_{ij} = \frac{2}{3} \delta_{ij} \int_V \rho(r) r^2 \, dV = \frac{8\pi}{3} \delta_{ij} \int_0^R dr \, \rho(r) r^4$$

For example, if $\rho(r) = \rho_0$ is constant, then $I_{ij} = \frac{8}{15} \pi \rho_0 R^5 \delta_{ij} = \frac{2}{3} MR^2 \delta_{ij}$ where M is the mass of the sphere.

7.2 Tensor Fields

A tensor field over $\mathbb{R}^3$ is the assignment of a tensor $T_{i...k}(\mathbf{x})$ to every point $\mathbf{x} \in \mathbb{R}^3$. This is the generalisation of a vector field

$$\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

to a map of the kind

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^m$$

with m the number of components of the tensor. So, for example, a map that assigns a symmetric, traceless rank 2 tensor $P_{ij}(\mathbf{x})$ to every point has $m = 5$.

The tensor field $T_{i...k}(\mathbf{x})$ is sometimes denoted as $T_{i...k}(x^l)$ which is supposed to show that the field depends on all coordinates $x^1, \ldots, x^3$. It's not great notation because the indices as subscripts are supposed to take some definite values, while the index l in the argument is supposed to denote the whole set of indices. It's especially bad notation when combined with the summation convention and we won't adopt it here.

Before we move on, it's worth pausing to mention a slightly subtle point. Not all maps $\mathbb{R}^3 \to \mathbb{R}^3$ qualify as "vector fields". The point $\mathbf{x}$ in the codomain $\mathbb{R}^3$ is a vector and so its components transform in the appropriate way under rotation. To be a vector field, the components of the map must transform under the *same* rotation. Similar comments hold for a tensor field.

To illustrate this, the electric field $\mathbf{E}(\mathbf{x})$ is an example of a vector field. If you rotate in space, and so change $\mathbf{x}$, then the direction $\mathbf{E}$ also changes: the rotation acts on both the argument $\mathbf{x}$ and the function itself $\mathbf{E}$.

In contrast, there are maps $\mathbb{R}^3 \to \mathbb{R}^3$ where, although the domain and codomain look the same, vectors in them transform under different rotations. For example, in particle physics there exists an object called a *quark field* which, for our (admittedly, slightly dumbed down) purposes, can be thought of as a map $\mathbb{R}^3 \to \mathbb{R}^3$. This is a quantum field whose ripples are particles that we call quarks, but these details can be safely ignored for the next couple of years of your life. We will write this field as $q_a(\mathbf{x})$ where the $a = 1, 2, 3$ label is the "colour" of the quark but should just be thought of as a vector in $\mathbb{R}^3$. If we rotate in space, then $\mathbf{x}$ changes but the colour of the quark does not. There is then an independent rotation that acts on the codomain and rotates the colour, but leaves the point in space unchanged. For this reason, the quark field is usually not referred to as a vector field.

Taking Derivatives

Given a suitably smooth tensor field $T(\mathbf{x})$ of rank p, we can differentiate it any number of times to get a new tensor field of rank, say, $p + q$,

$$X_{i_1 \dots i_q j_1 \dots j_p} = \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_q}} T_{j_1 \dots j_p}(\mathbf{x}) \tag{7.12}$$

To verify that this is indeed a tensor, we need to check how it changes under a rotation. In a new basis, we have $x_i' = R_{ij} x_j$, and so

$$\frac{\partial x_i'}{\partial x_j} = R_{ij} \quad \Longrightarrow \quad \frac{\partial}{\partial x_i'} = \frac{\partial x_j}{\partial x_i'} \frac{\partial}{\partial x_j} = R_{ij} \frac{\partial}{\partial x_j}$$

which is the result we need for X in (7.12) to qualify as a tensor field. Note that this is the same essential calculation that we did back in Section 1.1 when arguing that the gradient $\nabla \phi$ is a vector.

We can implement any of the ideas described above for tensor fields. For example, if we start with a vector field $\mathbf{F}(\mathbf{x})$, we can form a rank 2 tensor field

$$T_{ij}(\mathbf{x}) = \frac{\partial F_i}{\partial x_j}$$

But we saw in (7.8) that any rank 2 tensor field can be decomposed into various pieces. There is an anti-symmetric piece

$$A_{ij}(\mathbf{x}) = \epsilon_{ijk} B_k(\mathbf{x}) \quad \text{with} \quad B_k = \frac{1}{2} \epsilon_{ijk} \frac{\partial F_i}{\partial x_j} = -\frac{1}{2} (\nabla \times \mathbf{F})_k$$

and a trace piece

$$Q = \frac{\partial F_i}{\partial x_i} = \nabla \cdot \mathbf{F}$$

and, finally, a symmetric, traceless piece

$$P_{ij}(\mathbf{x}) = \frac{1}{2} \left(\frac{\partial F_i}{\partial x_j} + \frac{\partial F_j}{\partial x_i} \right) - \frac{1}{3} \nabla \cdot \mathbf{F}$$

Obviously, the first two of these are familiar tensors (in this case a scalar and vector) from earlier sections.

7.3 A Unification of Integration Theorems

The three integral theorems that we met in Section 4 are obviously closely related. To end these lectures, we show how they can be presented in a unified framework. This requires us to introduce some novel and slightly formal ideas. These go quite a bit beyond what is usually covered in an introductory course on vector calculus, but we will meet these objects again in later courses on Differential Geometry and General Relativity. View this section as a taste of things to come.

7.3.1 Integrating in Higher Dimensions

Our unified framework will give us integral theorems in any dimension $\mathbb{R}^n$. If you look back at Section 4, you'll notice that the divergence theorem already holds in any $\mathbb{R}^n$. Meanwhile, Stokes' theorem is restricted to surfaces in $\mathbb{R}^3$ for the very simple reason that the cross-product is only defined in $\mathbb{R}^3$. This suggests that before we can extend our integral theorems to higher dimensions, we should first ask a more basic question: how do we extend the cross product to higher dimensions?

The introduction of tensors gives us a way to do this. Given two vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$, the cross-product is

$$(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k$$

From this perspective, the reason that the cross product can only be employed in $\mathbb{R}^3$ is because it's only there that the ϵ_{ijk} symbol has three entries. If, in contrast, we're in $\mathbb{R}^4$ then we have ϵ_{ijkl} and so if we feed it two vectors $\mathbf{a}$ and $\mathbf{b}$, then we find ourselves with a tensor of rank 2, $T_{ij} = \epsilon_{ijkl} a_k b_l$.

The tensors that we get from an epsilon symbol are always special, in the sense that they are totally anti-symmetric. The anti-symmetry condition doesn't impose any extra constraint on a 0-tensor ϕ or a 1-tensors a_i as these are just scalar fields and vector fields respectively. It only kicks in when we get to tensors of rank 2 or higher.

With this in mind, we can revisit the cross product. We can define the cross product in any dimension $\mathbb{R}^n$: it is a map that eats two vectors $\mathbf{a}$ and $\mathbf{b}$ and spits back an anti-symmetric $(n-2)$-tensor

$$(\mathbf{a} \times \mathbf{b})_{i_1 \ldots i_{n-2}} = \epsilon_{i_1 \ldots i_n} a_{i_{n-1}} b_{i_n}$$

The only thing that's special about $\mathbb{R}^3$ is that we get back another vector, rather than a higher dimensional tensor.

There is also a slightly different role played by the epsilon symbol $\epsilon_{i_1, \ldots, i_n}$: it provides a map from anti-symmetric p-tensors to anti-symmetric $(n-p)$-tensors, simply by contracting indices,

$$\epsilon : \ T_{i_1 \ldots i_p} \ \mapsto \ \frac{1}{(n-p)!} \epsilon_{i_1 \ldots i_n} T_{i_{n-p+1} \ldots i_n} \tag{7.13}$$

This map goes by the fancy name of the *Hodge dual*. (Actually, it's an entirely trivial version of the Hodge dual. The proper Hodge dual is a generalisation of this idea to curved spaces.)

Our next step is to think about what this has to do with integration. To proceed, let's recall that earlier in these lectures we found two natural ways to integrate vector fields in $\mathbb{R}^3$. The first is along a line

$$\int_C \mathbf{F} \cdot d\mathbf{x} \tag{7.14}$$

which captures the component vector field *tangent* to the line. We can perform this procedure in any dimension $\mathbb{R}^n$. The second operation is to integrate a vector field over a surface and is restricted to $\mathbb{R}^3$,

$$\int_S \mathbf{F} \cdot d\mathbf{S} \tag{7.15}$$

This time it captures the component of the vector field *normal* to the surface. This normal, which is buried in the definition of $d\mathbf{S}$ requires us to use the cross product. For a parameterised surface $\mathbf{x}(u, v)$, the vector area element is

$$d\mathbf{S} = \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \, du \, dv$$

or, in components,

$$dS_i = \epsilon_{ijk} \frac{\partial x^j}{\partial u} \frac{\partial x^k}{\partial v} \, du \, dv$$

Now comes a mathematical sleight of hand. Rather than thinking of (7.15) as the integral of a vector field projected normal to the surface, instead think of it as the integral of an anti-symmetric 2-tensor $F_{ij} = \epsilon_{ijk} F_k$ integrated *tangent* to the surface. We then have

$$\int_S \mathbf{F} \cdot d\mathbf{S} = \int_S F_{ij} \, dS_{ij} \quad \text{with} \quad dS_{ij} = \frac{1}{2} \left(\frac{\partial x^j}{\partial u} \frac{\partial x^k}{\partial v} - \frac{\partial x^j}{\partial v} \frac{\partial x^k}{\partial u} \right) du \, dv \tag{7.16}$$

This is the same equation as before, just with the epsilon symbol viewed as part of the integrand rather than as part of the measure. Note that we've retained the anti-symmetry of the area element dS_{ij} that was inherent in our original cross product definition of $d\mathbf{S}$. Strictly speaking this isn't necessary because we're contracting with anti-symmetric indices in F_{ij}, but it turns out that it's best to think of both objects F_{ij} and dS_{ij} as individually anti-symmetric.

This new perspective suggests a way to generalise to higher dimensions. In the line integral (7.14) we're integrating a vector field over a line. In the surface integral (7.16), we're really integrating an anti-symmetric 2-tensor over a surface. The key idea is that one can integrate a totally anti-symmetric p-tensor over a p-dimensional subspace.

Specifically, given an anti-symmetric p-tensor, the generalisation of the line integral integral (7.14) is to integrate it over a p-dimensional subspace,

$$\int_M T_{i_1 \dots i_p} \, dS_{i_1 \dots i_p} \tag{7.17}$$

where $\dim(M) = p$. Here $dS_{i_1 \dots i_p}$ is a higher dimensional version of the "area element" defined in (7.16).

Alternatively, the higher dimensional version of the surface integral (7.15) involves first mapping the p-tensor to an $(n-p)$tensor using the Hodge dual. This can subsequently be integrated over an $(n-p)$-dimensional subspace

$$\int_{\tilde{M}} T_{i_1 \ldots i_p} \epsilon_{i_1 \ldots i_p j_1 \ldots j_{n-p}} \, d\tilde{S}_{j_1 \ldots j_{n-p}} \tag{7.18}$$

with $\dim(\tilde{M}) = n - p$.

In fact, we've already met an integral of the form (7.18) elsewhere in these lectures, since this is what we're implicitly doing when we integrate a scalar field over a volume. In this case the "area element" is just $dS_{i_1 \ldots i_n} = \frac{1}{n!} \epsilon_{i_1 \ldots i_n} \, dV$ and the two epsilon symbols just multiply to a constant.. When actually computing a volume integral, this extra machinery is more of a distraction than a help.. But if we want to know how to think about things more generally then it's extremely useful.

7.3.2 Differentiating Anti-Symmetric Tensors

We've now learned how to integrate anti-symmetric tensors. Our next step is to learn how to differentiate them. We've already noted in (7.12) that we can differentiate a p tensor once to get a tensor of rank $p + 1$, but in general differentiating loses the anti-symmetry property. As we now explain, there is a way to restore it so that when we differentiate a totally anti-symmetric p tensor, we end up with a totally anti-symmetric $(p+1)$-tensor.

For a scalar field, things are trivial. We can construct a vector field $\nabla \phi$ and this is automatically "anti-symmetric" because there's nothing to anti-symmetrise.

If we're given a vector field F_i, we can differentiate and then anti-symmetrise by hand. I will introduce a new symbol for "differentiation and anti-symmetrisation" and write

$$(\mathcal{D}F)_{ij} := \frac{1}{2} \left(\frac{\partial F_i}{\partial x^j} - \frac{\partial F_j}{\partial x^i} \right)$$

where the anti-symmetry is manifest on the right-hand side. I should confess that the notation $\mathcal{D}F$ is not at all standard. In subsequent courses, this object is usually viewed as something called a "differential form" and written simply as dF but the notation dF is loaded with all sorts of other connotations which are best ignored at this stage. Hence the made-up notation $\mathcal{D}F$.

In $\mathbb{R}^3$, this anti-symmetric differentiation is equivalent to the curl using the Hodge map (7.13),

$$(\nabla \times \mathbf{F})_i = \epsilon_{ijk}(\mathcal{D}F)_{jk}$$

But now we can extend this definition to any anti-symmetric p-tensor. We can always differentiate and anti-symmetrise to get a $(p+1)$-tensor defined by

$$(\mathcal{D}T)_{i_1 \ldots i_{p+1}} = \frac{1}{p+1} \left(\frac{\partial T_{i_1 \ldots i_p}}{\partial x^{i_{p+1}}} + p \text{ further terms} \right)$$

where the further terms involve replacing the derivative $\partial/\partial x^{i_{p+1}}$ with one of the other coordinates $\partial/\partial x^j$ so that the whole shebang is fully anti-symmetric.

Note that, with this definition of $\mathcal{D}$, if we differentiate twice then we take a p-tensor to a $(p+2)$-tensor. But this $(p+2)$-tensor always vanishes! This is because we'll have two derivatives contracted with an epsilon and is the higher dimensional generalisation of the statements that $\nabla \times \nabla \phi = 0$ or $\nabla \cdot (\nabla \times \mathbf{F}) = 0$.

This is actually the second time in these lectures that we've seen something vanish when you act twice. Here our new anti-symmetric derivative obeys $\mathcal{D}^2(\text{anything}) = 0$. But we previously noted that the "boundary of a boundary" is always zero. This means that if a higher dimensional space (really a manifold) M has boundary ∂M then $\partial(\partial M) = 0$. Conceptually, these two ideas are very different but one can't help but be struck by the similarity of the equations $\mathcal{D}^2(\text{anything}) = 0$ and $\partial^2(\text{anything}) = 0$. It turns out that this similarity is pointing at a deep connection between the topology of spaces and the kinds of tensors that one can put on these spaces. In fancy maths words, this is the link between homology and cohomology.

Finally, we can now state the general integration theorem. Given an anti-symmetric p-tensor T, then

$$\int_M (\mathcal{D}T)_{i_1, \ldots i_{p+1}} \, dS_{i_1 \ldots i_{p+1}} = \int_{\partial M} T_{i_1 \ldots i_p} \, dS_{i_1 \ldots i_p} \tag{7.19}$$

Here $\dim(M) = p + 1$ and, therefore the boundary has $\dim(\partial M) = p$. Note that we don't use a different letter to distinguish the integration measure over these various spaces: everything is simply dS and you have to look closer at the indices to see what kind of space you're integrating over.

The equation (7.19) contains the fundamental theorem of calculus (when $p = 0$), the divergence theorem (when $p = n - 1$) and Stokes' theorem (when $p = 1$ and $\mathbb{R}^n = \mathbb{R}^3$). Geometers refer to this generalised theorem simply as *Stokes' theorem* since that is the original result that it resembles most. The proof is simply a higher dimensional version of the proofs that we sketched previously.

There is, to put it mildly, quite a lot that I'm sweeping under the rug in the discussion above. In particular, the full Stokes' theorem does not hold only in $\mathbb{R}^n$ but in a general curved space known as a manifold. In that context, one has to be a lot more careful about what kind of tensors we're dealing with and, as I mentioned above, Stokes' theorem should be written using a kind of anti-symmetric tensor known as a *differential form*. None of this really matters when working in flat space, but the differences become crucial when thinking about curved spaces. If you want to learn more, these topics will be covered in glorious detail in later courses on Differential Geometry or, for physicists, General Relativity.