Index Gymnastics: Gauss' Theorem, Isotropic Tensors, NS Equations

The purpose of today's TA session is to mess a bit with tensors and indices, which are a necessary tool for continuum theories and in particular for Solid Mechanics. We'll see some simple examples, trying to become comfortable with these mathematical tools.

1 Gauss' integral theorem for tensors

You know from your undergrad studies that if $\vec{u}$ is a vector field in a volume $\Omega \subset \mathbb{R}^3$, then

$$\int_\Omega \operatorname{div} \vec{u}\, dV = \int_S \vec{u} \cdot d\vec{S} \tag{1}$$

where S is the surface of Ω (in mathematical notation, $S = \partial\Omega$). $d\vec{S}$ is a unit vector, perpendicular to a local surface. This is called Gauss' theorem, and it also works for tensors:

$$\int_\Omega \operatorname{div} \boldsymbol{A}\, dV = \int_{\partial\Omega} \boldsymbol{A}\, d\vec{S} \tag{2}$$

where the right-hand-side should be understood as $\boldsymbol{A}$ operating as a tensor on $d\vec{S}$, exactly like the right-hand-side of (1) represented $\vec{u}$ operating as a tensor on $d\vec{S}$, i.e. the usual dot product. Both Eqs. (1) and (2) are given here without proof.

We will now see that you already know a particular case of Eq. (2). Take Ω to be a 2-dimensional sheet in a 3D space, and a vector field $\vec{u}$ on it. The boundary of Ω is now a curve, whose tangent will by denoted by $\vec{t}$. For simplicity, we'll assume that Ω is confined to the $x-y$ plane, although this is not necessary. We'll also take $\vec{u}$ to be z-independent. We define a new tensor

$$\boldsymbol{A} \equiv \boldsymbol{\mathcal{E}}\vec{u}\ , \tag{3}$$

where $\boldsymbol{\mathcal{E}}$ is the Levi-Civita tensor. Index-wise, this means $A_{ij} = \mathcal{E}_{ijk}u_k$. We begin by calculating the left-hand-side of (2):

$$\int_\Omega \operatorname{div} \boldsymbol{A} dS = \int_\Omega \partial_j \mathcal{E}_{ijk} u_k dS = \int_\Omega \mathcal{E}_{ijk} \partial_j u_k dS = \int_\Omega \left(\vec{\nabla} \times \vec{u}\right) dS\ . \tag{4}$$

The right-hand-side gives

$$\int_{\partial\Omega} \boldsymbol{A} d\vec{n} = \int_{\partial\Omega} \mathcal{E}_{ijk} u_k dn_j = \int_{\partial\Omega} \left(\vec{u} \times d\vec{n}\right)\ . \tag{5}$$

Now, $\vec{u} \times d\vec{n}$ is a vector that is perpendicular to both $\vec{u}$ and $d\vec{n}$, that is, it is directed in the $\hat{z}$ direction. Its magnitude is $|\vec{u}||\vec{n}|\sin\theta$ where θ is the angle between $\vec{u}$ and $\vec{n}$. But the angle between $\vec{u}$ and $\vec{t}$ is $\alpha \equiv 90° - \theta$, so we can write

$$|\vec{u} \times d\vec{n}| = |\vec{u}|\, |\vec{n}| \sin\theta = |\vec{u}| \left|\vec{t}\right| \cos\alpha = \left|\vec{u} \cdot \vec{t}\right|\ . \tag{6}$$

We conclude that

$$\vec{u} \times d\vec{n} = \left(\vec{u} \cdot \vec{t}\right) \hat{z} \tag{7}$$

The theorem (2) says that (4) and (5) are equal, so we conclude that

$$\int_{\Omega} \left(\vec{\nabla} \times \vec{u}\right) dS = \oint_{\partial\Omega} \vec{u} \cdot d\vec{l} \tag{8}$$

which you know well from your happy undergrad days, under the name of Stokes' Theorem (or Green's Theorem, sometimes).

2 Isotropic tensors

A tensor is called *isotropic* if its coordinate representation is independent under coordinate rotation. Let's look at all the possible forms of isotropic tensors of low ranks.

2.0 0-rank tensors

A 0-rank tensor, a.k.a a scalar, does not change under rotations, therefore all scalars are isotropic (surprise!)

2.1 1-rank tensors

A vector $\vec{v}$ is isotropic if for every rotation matrix R_{ij} we have

$$R_{ij}\, v_j = v_i \ , \tag{9}$$

You can easily show that this condition is satisfied for arbitrary $\boldsymbol{R}$ only if $\vec{v} = 0$. So the zero vector is the only isotropic vector (surprise #2!!).

2.2 2-rank tensors

Let's hope we're gonna get something a bit more interesting. A matrix $\boldsymbol{A}$ is isotropic if for every rotation matrix $\boldsymbol{R}$ we have $A_{ij} = R_{ik} R_{jl} A_{kl}$, or in matrix notation:

$$\boldsymbol{R}\boldsymbol{A}\boldsymbol{R}^T = \boldsymbol{A} \ . \tag{10}$$

Let's choose a specific rotation matrix, say a rotation of angle α around $\hat{z}$,

$$\boldsymbol{R}^z(\alpha) \equiv \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{11}$$

The invariance equation now takes the form

$$\boldsymbol{A}(\alpha) \equiv \boldsymbol{R}^z(\alpha)\, \boldsymbol{A}\, \boldsymbol{R}^z(\alpha)^T = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{A} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{A}(0) \ . \tag{12}$$

This is a complicated equation, with cosines and sines all mixed up in a very unpleasant manner. Luckily, we can find an equivalent condition, which is significantly simpler. Differentiating with respect to α and plugging $\alpha = 0$ gives

$$\left.\frac{\partial \boldsymbol{A}(\alpha)}{\partial \alpha}\right|_{\alpha=0} = \left.\frac{\partial \boldsymbol{R}^z(\alpha)}{\partial \alpha}\right|_{\alpha=0} \boldsymbol{A}\, \boldsymbol{R}^z(0) + \boldsymbol{R}^z(0)\, \boldsymbol{A}\, \left.\frac{\partial \boldsymbol{R}^z(\alpha)^T}{\partial \alpha}\right|_{\alpha=0} , \tag{13}$$

but since $\boldsymbol{R}^z(0)$ is the identity matrix, this reduces to the simple equation

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\equiv \boldsymbol{L}^z} \boldsymbol{A} + \boldsymbol{A} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0 . \tag{14}$$

$\boldsymbol{L}^z = \partial_\alpha \boldsymbol{R}|_{\alpha=0}$ is sometimes called "the generator of rotations around the z axis", because $\boldsymbol{R}^z(\alpha) = e^{\alpha \boldsymbol{L}^z}$. We see that equation $\boldsymbol{A}(0) = \boldsymbol{A}(\alpha)$ is equivalent to the much easier equation

$$\boldsymbol{A}(0) = \boldsymbol{A}(\alpha) \quad \Longleftrightarrow \quad [\boldsymbol{A}, \boldsymbol{L}^z] = 0 . \tag{15}$$

Explicitly calculating $[\boldsymbol{A}, \boldsymbol{L}^z]$ gives

$$[\boldsymbol{A}, \boldsymbol{L}^z] = \begin{pmatrix} -A_{1,2} - A_{2,1} & A_{1,1} - A_{2,2} & -A_{2,3} \\ A_{1,1} - A_{2,2} & A_{1,2} + A_{2,1} & A_{1,3} \\ -A_{3,2} & A_{3,1} & 0 \end{pmatrix} \tag{16}$$

We see that commutation with $\boldsymbol{L}^z$ requires (a) $A_{13} = A_{31} = A_{23} = A_{32} = 0$ and (b) $A_{11} = A_{22}$. Obviously, the choice of $\hat{z}$ is arbitrary and isotropy means that $\boldsymbol{A}$ should also commute with $\boldsymbol{L}^x$ and $\boldsymbol{L}^y$. If we repeat the above procedure for the other $\boldsymbol{L}$'s, the analog of (a) will be that all off diagonal elements must vanish, and the analog of (b) will be that all diagonal elements must be equal. That is,

$$A_{ij} \propto \delta_{ij} . \tag{17}$$

I stress that this is true only in dimensions ≥ 3. In the HW you'll see that in 2D there are isotropic tensors that are not proportional to the identity (can you already see how the above argument fails in 2D?).

2.3 3-rank tensors

Here we can use the same trick. A 3rd rank tensor $\boldsymbol{A}$ is isotropic iff for every rotation matrix R_{ij} we have

$$R_{i\alpha}\, R_{j\beta}\, R_{k\gamma} A_{\alpha\beta\gamma} = A_{ijk} . \tag{18}$$

You can imagine the mess that comes out of this if you plug in a real rotation matrix with sines and cosines and whatnot, and then start using trig identities. Phew, no thanks!

So like before, we choose $\boldsymbol{R} = \boldsymbol{R}^z(\alpha)$, differentiate, and set $\alpha = 0$. This gives

$$\begin{aligned} 0 &= \left(L^z_{i\alpha}\, \delta_{j\beta}\, \delta_{k\gamma} + \delta_{i\alpha}\, L^z_{j\beta}\, \delta_{k\gamma} + \delta_{i\alpha}\, \delta_{j\beta}\, L^z_{k\gamma} \right) A_{\alpha\beta\gamma} && (19) \\ &= L^z_{i\alpha}\, A_{\alpha jk} + L^z_{j\beta}\, A_{i\beta k} + L^z_{k\gamma}\, A_{ij\gamma} . && (20) \end{aligned}$$

To see what kind of equation we got, let's choose $i = 1,\ j = 3,\ k = 3$. Since the only non-zero elements of $\boldsymbol{L}^z$ are L^z_{12} and L^z_{21}, we get

$$0 = L^z_{1\alpha}\, A_{\alpha 33} + L^z_{3\beta}\, A_{1\beta 3} + L^z_{3\gamma}\, A_{13\gamma} = A_{233} \tag{21}$$

Similarly, by choosing different combinations of i, j, k and/or different $\boldsymbol{L}$'s, you get that $A_{ijk} = 0$ whenever i, j, k are not all different, that is, if (ijk) is not a permutation of (123).

Using this knowledge, we can choose now $i = 1,\ j = 1,\ k = 3$, we get

$$A_{113} = 0 = L^z_{1\alpha}\, A_{\alpha 13} + L^z_{1\beta}\, A_{1\beta 3} + L^z_{3\gamma}\, A_{11\gamma} = A_{213} + A_{123}$$

Or put differently, $A_{213} = -A_{123}$. Similarly, we can show that every time we flip two indices we get a minus sign. Therefore, we conclude that the only isotropic 3rd rank tensor is equal, up to a multiplicative constant, to $\boldsymbol{\mathcal{E}}$,

$$\mathcal{E}_{ijk} = \begin{cases} 0 & (ijk) \text{ is not a permutation of (123)} \\ \text{sign of permutation} & \text{otherwise} \end{cases} \tag{22}$$

As you probably know, $\boldsymbol{\mathcal{E}}$ is called the Levi-Civita completely anti-symmetric tensor.

2.4 4-rank tensors

Seriously? No. We're not going to redo the algebra. But can we guess the form of some isotropic 4-rank tensors? We can easily build them from lower rank isotropic tensors. Here are a few examples that come to mind:

$$A_{ijkl} = \delta_{ij}\, \delta_{kl} \tag{23}$$
$$A_{ijkl} = \delta_{il}\, \delta_{jk} \tag{24}$$
$$A_{ijkl} = \delta_{ik}\, \delta_{jl} \tag{25}$$
$$A_{ijkl} = \mathcal{E}_{ij\alpha}\, \mathcal{E}_{\alpha kl} \tag{26}$$

We did a really good job there, because it turns out that these are the only options. In fact, this list is even redundant, because each of the lines can be written as a linear combination of the other 3 (can you find it?). You may want to prove at home that there really are no other options - it's a nice exercise that can be easily automatized on `Mathematica`, and we're going to use this result in the course.

3 Navier-Stokes equation

We are now going to use the heavy arsenal developed above, and derive the Navier-Stokes equation solely from symmetry considerations. We want to find a dynamical equation for $\partial_t \vec{v}$ as a function of $\vec{v}$ and its spatial derivatives. We take a perturbative approach, and expand $\partial_t v$ to second order in $\vec{v}$ and in its gradients:

$$\partial_t v_i = A_{ij}\, v_j + B_{ijk}\, \partial_j v_k + D_{ijkl}\, v_j \partial_k v_l + E_{ijkl}\, \partial_j \partial_k v_l + F_{ijk}\, v_j v_k \tag{27}$$

Since $\vec{v}$ is a physical quantity (specifically, a 1-rank tensor) the dynamical equation for $\partial_t \vec{v}$ should be covariant under symmetries of the physical system in question. We'll see what these symmetries impose on the form of the various tensors $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{F}$.

We begin with a Galilean transformation:

$$y_i = x_i + c_i t \tag{28}$$
$$\tau = t \tag{29}$$

Under this transformation, the velocity field now takes the form $\vec{w} = \vec{v} - \vec{c}$. Also, by the chain rule:

$$\partial_{x_i} = \frac{\partial y_j}{\partial x_i}\partial_{y_j} + \frac{\partial \tau}{\partial x_i}\partial_\tau = \partial_{y_i} \tag{30}$$
$$\partial_t = \frac{\partial y_j}{\partial t}\partial_{y_j} + \frac{\partial \tau}{\partial t}\partial_\tau = c_j \partial_{y_j} + \partial_\tau \tag{31}$$

Applying this to Eq. (27) gives

$$\begin{aligned} \partial_t w_i + c_j \partial_j w_i = A_{ij}(w_j - c_j) + B_{ijk}\partial_j w_k + D_{ijkl}(w_j - c_j)\partial_k w_l \\ + E_{ijkl}\partial_j \partial_k w_l + F_{ijk}(w_j - c_j)(w_k - c_k) \end{aligned} \tag{32}$$

If we want the NS equation to be covariant, we need to impose that (32) will be equal, term by term, to (27), i.e.

$$-A_{ij} c_j = 0 \ , \tag{33}$$
$$c_j \partial_j w_i = -D_{ijkl}\, c_j \partial_k w_l, \tag{34}$$
$$F_{ijk}(w_j - c_j)(w_k - c_k) = F_{ijk} w_j w_k \tag{35}$$

All these should hold for arbitrary $\vec{c}, \vec{w}$. The first constraint clearly means $\boldsymbol{A} = 0$. For the third one, choose for example $\vec{w} = \vec{c}$, and get that $F_{ijk} w_j w_k = 0$ for arbitrary $\vec{w}$. Note that this is exactly the last term in Eq. (27), so it we see that is vanishes identically. The constraint (34) may be written as

$$\delta_{il}\delta_{jk}\ c_j \partial_k w_l = -D_{ijkl}\ c_j \partial_k w_l$$

Since $\vec{c}, \vec{w}$ are arbitrary, this means $D_{ijkl} = -\delta_{il}\delta_{kj}$ and (27) can be written as

$$\partial_t v_i + v_j \partial_j v_i = B_{ijk}\partial_j v_k + E_{ijkl}\partial_j \partial_k v_l \tag{36}$$

You have to admit that this is a very big improvement...

Now let's look at rotations, $y_j = R_{ij} x_j$. Demanding Eq. (36) to be invariant means that the tensors $\mathbf{B}, \mathbf{E}$ are *isotropic*.

We've just seen that the only 3^{rd} rank isotropic tensor is the Levi-Civita tensor, so the $\boldsymbol{B}$ term is proportional to $\vec{\nabla} \times \vec{v}$ and thus is forbidden by reflection symmetry. It's too bad that we know already that the $\boldsymbol{A}$ and $\boldsymbol{F}$ terms are gone, because they would also be forbidden by rotational symmetry. For example, the $\boldsymbol{F}$ term must be proportional to $\vec{v} \times \vec{v}$ and therefore vanishes identically (note that we didn't show that $\boldsymbol{F} = 0$, but only that it gives zero when it acts on the same vector in its two slots).

As for $\mathbf{E}$, we know that we have exactly three choices, given in Eqs. (23),(24),(25). These give, respectively,

$$\delta_{ij}\,\delta_{kl}\,\partial_j\partial_k\,v_l = \partial_i\partial_j\,v_j = \vec{\nabla}\,(\nabla\cdot\vec{v}) = \mathrm{grad}\,(\mathrm{div}\,\vec{v}) \tag{37}$$

$$\delta_{il}\,\delta_{jk}\,\partial_j\partial_k\,v_l = \partial_j\partial_j\,v_i = \nabla^2\vec{v} = \mathrm{div}\,(\mathrm{grad}\,\vec{v}) \tag{38}$$

$$\delta_{ik}\,\delta_{jl}\,\partial_j\partial_k\,v_l = \partial_i\partial_j\,v_j = \text{same as (37)} \tag{39}$$

So the third option is redundant. Note that if we wanted to use Eq. (26) we'd get

$$\mathcal{E}_{ij\alpha}\,\mathcal{E}_{\alpha kl}\partial_j\partial_k\,v_l = \mathcal{E}_{ij\alpha}\,\partial_j\left(\vec{\nabla}\times\vec{v}\right)_\alpha = \vec{\nabla}\times\left(\vec{\nabla}\times\vec{v}\right)\ ,$$

which is also redundant because of the vector calculus identity which you all know by heart: $\nabla\times(\nabla\times\mathbf{A}) = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}$.

To sum up, we see that the only form of $\partial_t\vec{v}$ which is covariant under rotations, reflections and Galilean symmetries is

$$\left(\partial_t + \vec{v}\cdot\vec{\nabla}\right)\vec{v} = \eta\nabla^2\vec{v} + \mu\vec{\nabla}\,(\nabla\cdot\vec{v}) \tag{40}$$

where η and μ are two scalars. In incompressible flows there's only one η, as the μ term vanishes.

Lastly, note that there's another term that clearly does not violate any symmetries: $\vec{\nabla}P$ where P is some scalar function.

3.1 An historical note about the power of symmetries in continuum theories

Euler's equation $(\partial_t + v_j\partial_j)v_i = \partial_i P$, regarding inviscid flows, was derived sometime around 1750. It took the scientific community almost eighty years (!!) to understand how to incorporate viscosity into the business. Mind you, some of the greatest minds of the time were devoted to the problem, including Cauchy, Poisson, d'Alembert, Bernoulli and of course Navier and Stokes. Not exactly Elitzur Ra'anana, if you see what I mean. So what took them so long?

The answer, very very roughly, is that they tried to model viscosity on a molecular level: to understand the dissipation mechanisms, stress-transfer mechanisms and whatnot. One of the great strengths of continuum theory is that measly insignificant mortals like us were able to do here in 45 minutes a derivation that the primordial gods needed 80 years to do. Moreover, we did that *without caring even the slightest bit about the underlying physics.*

In fact, this is the crux of the matter – the use of symmetries allows us to say very powerful statements about the functional form of the viscosity term, without having to deal with the microscopic mechanisms. It allows us to develop a predictive theory, where all the "microscopics" are lumped unto a small number of parameters (in our case - μ and η), which of course must be determined experimentally.

The down side is that we can not say anything quantitative about the parameters. From our theory we can not give even an order-of-magnitude estimation of η or μ, let alone their dependence on the fluid's properties (although thermodynamics easily tells us that they are positive).