L#1 - Symmetric tensors in Mathematical Physics

Let $\Omega \subset \mathbb{R}^n$ be an open set. We are interested in symmetric tensors

$$S : \Omega \to \mathbf{Sym}_n.$$

If the entries are distributions, we define the row-wise divergence $\operatorname{Div} S : \Omega \to \mathbb{R}^n$ by

$$(\operatorname{Div} S)_i = \sum_j \partial_j s_{ij}.$$

The positivity of the tensor plays an important role. We say that S is *positive semi-definite* if for every $\mu \in \mathbb{R}^n$, the distribution $\mu^T S \mu$ is ≥ 0. We recall that non-negative distributions are locally finite measures. Hence $S \geq 0_n$ implies that every entry s_{ij} is a locally finite measure.

The following result shows how well positivity fits with divergence-freeness:

Proposition 3 *Let S be symmetric $\geq 0_n$ over $\mathbb{R}^n$, and divergence-free. Then $S \equiv 0_n$.*

Proof

The Fourier transform $\xi \mapsto \mathcal{F}S(\xi)$ is continuous. The divergence-free condition translates as $\mathcal{F}S(\xi)\xi = 0$. For v a unit vector, take $\xi = \varepsilon v$. Passing to the limit into $\mathcal{F}S(\varepsilon v)v = 0$, we obtain $\mathcal{F}S(0)v = 0$. Therefore

$$\int_{\mathbb{R}^n} S(x)\,dx = \mathcal{F}S(0) = 0_n.$$

In other words, the non-negative finite measure $\mu^T S\mu$ has total mass 0, and therefore vanishes identically.

∎

Continuum mechanics

The physical space is $\mathbb{R}^d$ for some $d \geq 1$, the space variable being denoted $y = (y_1, \ldots, y_d)$. We set $n = 1 + d$ and $x = (t, y)$ where t is the time variable.

A material is described by its mass density $\rho(t, y) \geq 0$ and velocity field $v(t, y)$. The equation of continuity (conservation of mass) writes

$$\partial_t \rho + \mathrm{div}_y(\rho v) = 0, \tag{3}$$

where ρv is the *linear momentum*.

Newton's law of acceleration writes

$$\partial_t(\rho v) + \mathrm{Div}_y(\rho v \otimes v) = \mathrm{Div}_y \Sigma, \tag{4}$$

where $\Sigma(t,y)$ is the *Cauchy's stress tensor*. An important fact, which is equivalent to the conservation of angular momentum, is that Σ is *symmetric*. The system (3,4) can therefore be recast as

$$\mathrm{Div}_x S = 0, \qquad S := \begin{pmatrix} \rho & \rho v^T \\ \rho v & \rho v \otimes v - \Sigma \end{pmatrix}$$

where the tensor S is symmetric. Thanks to Proposition 1 (take $p = 1$), it is positive definite if and only if $\rho > 0$ and $-\Sigma \in \mathbf{SPD}_d$. More generally, it is positive semi-definite if and only if $\Sigma \leq 0_d$.

- For an inviscid gas, $\Sigma = -pI_d$ where $p \geq 0$ is the pressure. Thus $S \geq 0_n$.

- For a viscous gas, $\Sigma = -pI_d + \lambda(\nabla v + \nabla v^T) + \mu(\mathrm{div} v)I_d$ and one cannot conclude.

Determinant

Suppose that a physical model involves a divergence-free tensor S. The independent variables x_j have physical dimensions ℓ_j, while those of the entries s_{ij} are d_{ij}. Because each equation $(\mathrm{Div}\, S)_i = 0$ must involve quantities of the same dimension, say m_i, we have $d_{ij} = m_i \ell_j$. Therefore the monomials (σ a permutation)

$$\prod_{i=1}^{n} s_{i\sigma(i)}$$

have a common dimension

$$\prod_i m_i \cdot \prod_j \ell_j$$

and it makes sense to form a linear combination. For instance, $\det S$ is well-defined from a physical point of view.

Another reason to consider the determinant is that a model must be invariant under the action of some group G of linear transformations: the orthogonal, Galilean or Lorentz group, depending on the context. If $R \in G$, the action is defined by

$$(R \cdot S)(x) := RS(R^{-1}x)R^T.$$

The determinant is equivariant under this action.

In continuum mechanics, the Schur complement formula yields $\det S = \rho \det(-\Sigma)$. For an inviscid gas, this gives

$$\det S = \rho p^d.$$

Relativistic gas (special relativity)

Let $c > 0$ be the speed of light. Recall that $|v| < c$.

Because of the equivalence principle between mass an energy, the continuity equation is replaced by the conservation of energy:

$$\partial_t \left(\frac{\rho c^2 + p}{c^2 - |v|^2} - \frac{p}{c^2} \right) + \mathrm{div}_y \left(\frac{\rho c^2 + p}{c^2 - |v|^2} v \right) = 0. \tag{5}$$

The law of acceleration becomes

$$\partial_t \left(\frac{\rho c^2 + p}{c^2 - |v|^2} v \right) + \mathrm{Div}_y \left(\frac{\rho c^2 + p}{c^2 - |v|^2} v \otimes v \right) + \nabla_y p = 0. \tag{6}$$

This is recast as $\mathrm{Div}_x S = 0$ where the symmetric tensor

$$S = \begin{pmatrix} \frac{\rho c^2+p}{c^2-|v|^2} - \frac{p}{c^2} & \frac{\rho c^2+p}{c^2-|v|^2} v \\ \frac{\rho c^2+p}{c^2-|v|^2} v & \frac{\rho c^2+p}{c^2-|v|^2} v \otimes v + pI_d \end{pmatrix}$$

is still positive semi-definite (homework: prove this using Schur complement). Amazingly enough one has again

$$\det S = \rho p^d.$$

This may be seen by using the Lorentz invariance, and considering the frame in which the velocity vanishes.

Kinetic models

We still take $n = 1 + d$ and $x = (t, y)$. But now the velocity $v \in \mathbb{R}^d$ is an independent variable and the particle density is a function $f(t, y, v) \geq 0$. The local mass density and the linear momentum are

$$\rho(t, y) = \int_{\mathbb{R}^d} f(t, y, v)\, dv, \qquad m = \int_{\mathbb{R}^d} f(t, y, v) v\, dv,$$

from which we can define a *mean velocity* $u := \frac{m}{\rho}$.

The motion is governed by a kinetic equation

$$\underbrace{\partial_t f + v \cdot \nabla_y f}_{\text{transport}} = Q[f]. \tag{7}$$

The right-hand side accounts for particle interactions, for instance collisions. In most models, Q is non-linear (often quadratic), non-local in the variable v, but local in x : $Q[f](t, y, v)$ depends only upon $f(t, y, \cdot)$.

All the models, among which there is that of Boltzmann, share the following properties, in which $g = g(v) \geq 0$ is an arbitrary function in $L^1((1+|v|^2)dv)$.

Minimum principle. If $g(w) = 0$, then $Q[g](w) \geq 0$.

Conservation of mass. $\int_{\mathbb{R}^d} Q[g](v)\,dv = 0$.

Conservation of momentum. $\int_{\mathbb{R}^d} Q[g](v)v\,dv = 0$.

Conservation of energy. $\int_{\mathbb{R}^d} Q[g](v)|v|^2\,dv = 0$.

The first one ensures that f stays ≥ 0. The next two are responsible for the macroscopic conservation laws of mass and momentum:

$$\begin{aligned} \partial_t \rho + \mathrm{div}_y m &= 0, \\ \partial_t m + \mathrm{Div}_y T &= 0, \end{aligned}$$

where $T := \int_{\mathbb{R}^d} Q[g](v) v \otimes v \, dv$. This is recast as $\mathrm{Div}_x S = 0$ with

$$S := \begin{pmatrix} \rho & m^T \\ m & T \end{pmatrix} = \int_{\mathbb{R}^d} f(t, y, v) V \otimes V \, dv$$

where $V = \binom{1}{v}$. The latter formula shows that S is symmetric, positive semi-definite.

Wave equation

A scalar function $u(t,y)$ obeys to

$$u_{tt} = c^2 \Delta_y u. \tag{8}$$

The conservation of energy writes

$$\partial_t \frac{1}{2}(u_t^2 + c^2|\nabla u|^2) = \mathrm{div}_y(c^2 u_t \nabla u)$$

where ∇ is the spatial gradient. Meanwhile the energy flux $u_t \nabla u$ obeys a conservation law too:

$$\partial_t(u_t \partial_j u) = \partial_j \frac{1}{2}(u_t^2 - c^2|\nabla u|^2) + c^2 \mathrm{div}(\partial_j u \nabla u).$$

All this can be recast as $\mathrm{Div}_x S = 0$ where

$$S := \begin{pmatrix} \frac{1}{2}(u_t^2 + c^2|\nabla u|^2) & -c^2 u_t \nabla u^T \\ -c^2 u_t \nabla u & \frac{1}{2}(u_t^2 - c^2|\nabla u|^2) I_d + c^2 \nabla u \otimes \nabla u \end{pmatrix}.$$

When $d = 1$, S is positive semi-definite. Otherwise the positiveness occurs only when $|u_t| \geq c|\nabla u|$. Actually we have

$$\det S = \frac{c^{2d}}{2^n}(u_t^2 - c^2|\nabla u|^2)^n.$$

Maxwell's equations

Electro-magnetism[3] in vacuum is described by a field (E,B) with values in $\mathbb{R}^6$. The standard models writes

$$\partial_t B + \operatorname{curl} E = 0, \qquad \operatorname{div} B = 0, \tag{9}$$

$$\varepsilon_0 \mu_0 \partial_t E - \operatorname{curl} B = 0, \qquad \operatorname{div} E = 0. \tag{10}$$

In presence of charges, the equations (10) incorporate the charge density ρ and the current j. One has $\varepsilon_0 \mu_0 c^2 = 1$ with c the light speed. The energy density is conserved[4]:

$$\partial_t \frac{1}{2}(|B|^2 + |E|^2) + \operatorname{div}(E \times B) = 0.$$

[3]Here $d = 3$ and $n = 4$.

[4]Here we chose units in which $c = 1$.

The energy momentum satisfies a conservation law as well:

$$\begin{aligned}\partial_t(E\times B) &= (\operatorname{curl} B)\times B+(\operatorname{curl} E)\times E\\ &= \operatorname{Div}(B\otimes B+E\otimes E)-\nabla\frac{1}{2}(|B|^2+|E|^2).\end{aligned}$$

Whence a divergence-free symmetric tensor

$$S:=\begin{pmatrix}\frac{1}{2}(|B|^2+|E|^2) & E\times B\\ E\times B & -B\otimes B-E\otimes E+\frac{1}{2}(|B|^2+|E|^2)I_3\end{pmatrix}.$$

This does not look impressive ... But let us mention that the Maxwell system might not be perfectly linear. Linearity has been criticized because a steady point charge generates a field $B\equiv 0$ and $E=q|y|^{-3}y$, for which the energy density $q^2/2|y|^4$ is not integrable at the origin. Therefore the amount of energy in a neighbourhood of the charge is infinite ! Physicists looked for alternate, non-linear models to resolve this paradox ; the most famous one is that of M. Born & L. Infeld in 1934.

Non-linear models

One postulates that the unknown is a differential form ω of degree 2, which is closed: $d\omega = 0$. In coordinates, one writes

$$\omega = (E \cdot dy) \wedge dt + B \cdot (dy \wedge dy) = E_1 dy_1 \wedge dt + \cdots + B_1 dy_2 \wedge dy_3 + \cdots ,$$

which defines the fields (E, B). Mind that these depend on the choice of coordinates. The condition $d\omega = 0$ translates as the Maxwell–Faraday equations (9).

The remaining equations are assumed to derive from a variational principle $\delta\mathcal{L} = 0$ where $\mathcal{L}[\omega] = \int\int L(E,B)\, dy\, dt$. Because the ambiant space is that of closed 2-forms, which are locally exact, we content ourselves to writing

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{L}[\omega + \varepsilon d\alpha] = 0, \tag{11}$$

for every 1-form α.

Coordinate-wise, $\alpha = \phi dt + A \cdot dy$ and

$$d\alpha = (\nabla\phi - \partial_t A) \cdot dy \wedge dt + \mathrm{curl} A \cdot (dy \wedge dy).$$

Thus (11) gives

$$\int\int \left(\frac{\partial L}{\partial B} \cdot \mathrm{curlA} + \frac{\partial L}{\partial E} \cdot (\nabla\phi - \partial_t A \right) dy\, dt = 0$$

for every function ϕ and vector field A. This yields

$$\partial_t D - \mathrm{curl}\, H = 0, \qquad \mathrm{div}\, D = 0, \tag{12}$$

where

$$D := \frac{\partial L}{\partial E}, \qquad H := -\frac{\partial L}{\partial B}.$$

This replaces the Maxwell–Gauss equations (10)

The system (9,10) is a special case where $L = \frac{1}{2}(|E|^2 - |B|^2)$.

The general model admits an energy density W, which is the Legendre transform of L with respect to E (B is kept as a parameter):

$$W(B,D) = D\cdot E - L(E,B).$$

By reciprocity $W_D = E$. The chain rule gives (B-derivatives are taken at D constant)

$$W_B = D\cdot E_B - L_E\cdot E_B - L_B = -L_B = H.$$

Thus (9.1,12.1) can be recast as

$$\partial_t B + \mathrm{curl}\, W_D = 0, \qquad \partial_t D - \mathrm{curl}\, W_B = 0.$$

We derive essentially the same conservation law (Poynting identity) as in the linear model

$$\partial_t W + \mathrm{div}\,(E\times H) = 0. \tag{13}$$

A controversy

It is tempting to look for a conservation law of the form $\partial_t(E\times H)+\mathrm{Div}\,(\cdots)=0$, as we had one in the standard linear model, but this does not seem available in general. Instead, we have

$$\partial_(D\times B)=\mathrm{Div}(W_B\otimes B+W_D\otimes D)+\nabla(W-B\cdot W_B-D\cdot W_D). \tag{14}$$

We may therefore form a divergence-free tensor

$$S=\begin{pmatrix} W & E\times H \\ D\times B & -W_B\otimes B-W_D\otimes D+(B\cdot W_B+D\cdot W_D-W)I_3 \end{pmatrix}.$$

It is however unclear at this stage whether S is symmetric or not. On the one hand the lower-right block is not clearly symmetric. On the other hand, why should $D\times B$ equal $E\times H$? This was the origin of a controversy about the so-called Poynting vector. Some people leant to the Minkowski's form $D\times B$, while the others inclined towards Abraham's form $E\times H$.

Involving Lorentz invariance

The controversy is resolved when one remarks that the model must be frame-independent, that is invariant under Lorentz transformations. In other words, the density $L(E,B)$ does not really depend upon E and B, but only on the 2-form $\omega(x)$. This means that if (E',B') represents the same form in a different admissible coordinate system, then $L(E',B') = L(E,B)$.

One proves that such a density is actually a function of two scalar quantities

$$L = \ell(\sigma,\pi), \qquad \sigma := \frac{1}{2}(|E|^2 - |B|^2), \quad \pi = E\cdot B.$$

A remarkable fact is

Theorem 1 *The tensor S is symmetric if, and only if L is invariant under the action of the Lorentz group, that is $L = \ell(\sigma,\pi)$ for some function ℓ.*

Proof

($\Longleftarrow$) If L is Lorentz-invariant, then $D = \ell_\sigma E + \ell_\pi B$ and $H = \ell_\sigma B - \ell_\pi E$. Therefore

$$D \times B = \ell_\sigma E \times B = E \times H.$$

On the other hand, the lower-right block of S is symmetric because of

$$\begin{aligned} W_B \otimes B + W_D \otimes D &= H \otimes B + E \otimes D = -L_B \otimes B + E \otimes L_E \\ &= \ell_\sigma (B \otimes B + E \otimes E). \end{aligned}$$

($\Longrightarrow$) Conversely, the condition $D \times B = E \times H$ writes $L_E \times B + E \times L_B =$. This is a set of three first-order linear PDEs $R_j \cdot \nabla L = 0$. For instance $R_1 \cdot \nabla = B_3 \partial_{E_2} - B_2 \partial_{E_3} + E_2 \partial_{B_3} - E_3 \partial_{B_2}$. They imply $R \cdot \nabla L = 0$ for every vector field R in the Lie algebra $\mathcal{A}$ spanned by $\{R_1, R_2, R_3\}$.

One verifies that

$$\mathcal{A} = \langle R_1, R_2, R_3, [R_1, R_2], [R_2, R_3], [R_3, R_1] \rangle$$

has dimension 6, and that for each point (E,B), $\dim \mathcal{A}(E,B) = 4$. This means that the PDEs admit $6-4=2$ independent solutions, from which all solutions are functionally dependent. Obviously σ and π are two such independent solutions.

$\blacksquare$

Not surprisingly (Maxwell contains the wave equation), S is not positive. Actually

$$\det S = -Z^2 \leq 0$$

where

$$Z := \ell_\sigma^2(\sigma^2 + \pi^2) - (\ell - \sigma \ell_\sigma - \pi \ell_\pi)^2.$$

Conclusion

The various examples taken form Mathematical Physics teach us the following lessons.

- There often exists a divergence-free tensor of size $n \times n$ (the divergence being taken in space and time variables). The equality $\mathrm{Div}\, S = 0$ expresses the conservation of either mass or energy, together with that of the corresponding momentum. We shall speak of the mass-momentum (energy-momentum) tensor of the model.

- The symmetry of S is a consequence of (or even is equivalent to) the invariance of the model with respect to a group of transformation (Galilean, Lorentzian).

- The tensor is not always positive (we shall elaborate later on).

Comments

* We warn the audience that in continuum mechanics, the tensor is symmetric only when the equations of motion are written in the Eulerian frame (when x stands for space and time). It is lost when the equations of the motion are written in the Lagrangian variables (material variables).

* Although the tensor S for the wave equation is not always positive, it defines a quadratic form Q such that $Q(V) \geq 0$ for every time-like vector. This expresses the property that the positivity of the energy density does not depend upon the admissible frame.

* Likewise, a natural assumption for Maxwell's models is that the energy density is positive, for every admissible frame. This means that not only $W \geq 0$, but actually $V^T SV \geq 0$ for every time-like vector,

$$V = \begin{pmatrix} 1 \\ v \end{pmatrix}, \qquad |v| \leq 1.$$

Under the natural assumption that $\ell_\sigma \geq 0$, this amounts to the differential inequality

$$\ell \leq \pi \ell_\pi + \left(\sigma + \sqrt{\sigma^2 + \pi^2}\right) \ell_\sigma. \tag{15}$$

* The above examples raise the natural (still open) question of understanding the properties of divergence-free symmetric tensor that take values in the cone defined by $V^T SV \geq 0$ for every time-like vector V. Since this is a weaker condition than positive semi-definiteness, we expect weaker results.