Maxwell's Stress Tensor and Electromagnetic Momentum

Dr. Priti Mishra

Momentum conservation is rescued in electrodynamics by the realization that *the fields themselves carry momentum.* This is not so surprising when you consider that we have already attributed energy to the fields. In the case of the two point charges in Fig. 1, whatever momentum is lost to the particles is gained by the fields. Only when the field momentum is added to the mechanical momentum of the charges is momentum conservation restored. Let us see how this works out quantitatively in the following sections.

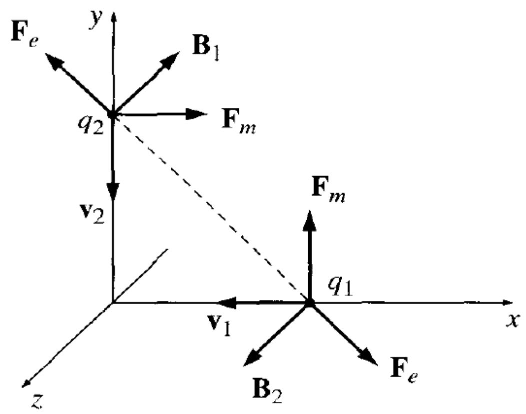

Figure 1: A point charge q_1 moving along x-axis encounters an identical one, q_2 proceeding in at the same speed along the y-axis.

1 Maxwell's Stress Tensor

Let us calculate the total electromagnetic force on the charges in volume V: As we know the force acting on a point charge q moving with velocity $\mathbf{v}$ in an electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ is given by Lorentz force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{1}$$

Similarly, the force per unit volume acting on a charge distribution ρ in a volume V:

$$\mathbf{f} = \rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{2}$$

As we know $\rho\mathbf{v} = \mathbf{J}$ where $\mathbf{J}$ is the current density.

$$\Rightarrow \mathbf{f} = \rho\mathbf{E} + \mathbf{J} \times \mathbf{B} \tag{3}$$

Let us write the above equation in terms of fields alone, eliminating ρ and $\mathbf{J}$ by using Maxwell's equations (i) and (iv):

$$\begin{aligned}
\boldsymbol{\nabla}.\mathbf{E} &= \frac{\rho}{\epsilon_0} \\
\Rightarrow \rho &= \epsilon_0 \boldsymbol{\nabla}.\mathbf{E} && (4) \\
\boldsymbol{\nabla} \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \\
\Rightarrow \mathbf{J} &= \frac{1}{\mu_0} \boldsymbol{\nabla} \times \mathbf{B} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} && (5)
\end{aligned}$$

Substituting ρ and $\mathbf{J}$ from equations (4) and (5) into equation (3) we get

$$\mathbf{f} = \epsilon_0 (\boldsymbol{\nabla}.\mathbf{E})\mathbf{E} + \left(\frac{1}{\mu_0} \boldsymbol{\nabla} \times \mathbf{B} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \times \mathbf{B} \tag{6}$$

Now

$$\frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) = \frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} + \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t} \tag{7}$$

or

$$\frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} = \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) - \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t} \tag{8}$$

and from Maxwell's third equation (Faraday's law)

$$\frac{\partial \mathbf{B}}{\partial t} = -\boldsymbol{\nabla} \times \mathbf{E} \tag{9}$$

Substituting above in equation (8) we get

$$\frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} = \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) + \mathbf{E} \times (\boldsymbol{\nabla} \times \mathbf{E}) \tag{10}$$

Substituting equation (10) in equation (6) we obtain

$$\mathbf{f} = \epsilon_0 (\boldsymbol{\nabla}.\mathbf{E})\mathbf{E} + \left(\frac{1}{\mu_0} (\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B} \right) - \epsilon_0 \left[\frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) + \mathbf{E} \times (\boldsymbol{\nabla} \times \mathbf{E}) \right] \tag{11}$$

or

$$\mathbf{f} = \epsilon_0 \left[(\boldsymbol{\nabla}.\mathbf{E})\mathbf{E} - \mathbf{E} \times (\boldsymbol{\nabla} \times \mathbf{E}) \right] + \left(\frac{1}{\mu_0} (\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B} \right) - \epsilon_0 \left[\frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) \right] \tag{12}$$

Just to make things look more symmetrical, let's add a term $(\boldsymbol{\nabla}.\mathbf{B})\mathbf{B}$; since $\boldsymbol{\nabla}.\mathbf{B} = 0$, this costs us nothing.

$$\mathbf{f} = \epsilon_0 \left[(\boldsymbol{\nabla}.\mathbf{E})\mathbf{E} - \mathbf{E} \times (\boldsymbol{\nabla} \times \mathbf{E})\right] + \frac{1}{\mu_0}\left[(\boldsymbol{\nabla}.\mathbf{B})\mathbf{B} - \mathbf{B} \times \boldsymbol{\nabla} \times \mathbf{B}\right] - \epsilon_0 \left[\frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B})\right] \tag{13}$$

Meanwhile, from the property of gradient we know that

$$\boldsymbol{\nabla}(\mathbf{A}.\mathbf{B}) = \mathbf{A} \times \boldsymbol{\nabla} \times \mathbf{B} + \mathbf{B} \times \boldsymbol{\nabla} \times \mathbf{A} + (\mathbf{A}.\boldsymbol{\nabla})\mathbf{B} + (\mathbf{B}.\boldsymbol{\nabla})\mathbf{A} \tag{14}$$

$$\begin{aligned} \Rightarrow \boldsymbol{\nabla}(E^2) = \boldsymbol{\nabla}(\mathbf{E}.\mathbf{E}) &= \mathbf{E} \times \boldsymbol{\nabla} \times \mathbf{E} + \mathbf{E} \times \boldsymbol{\nabla} \times \mathbf{E} + (\mathbf{E}.\boldsymbol{\nabla})\mathbf{E} + (\mathbf{E}.\boldsymbol{\nabla})\mathbf{E} \\ &= 2\mathbf{E} \times \boldsymbol{\nabla} \times \mathbf{E} + 2(\mathbf{E}.\boldsymbol{\nabla})\mathbf{E} \end{aligned} \tag{15}$$

So

$$\mathbf{E} \times \boldsymbol{\nabla} \times \mathbf{E} = \frac{1}{2}\boldsymbol{\nabla}(E^2) - (\mathbf{E}.\boldsymbol{\nabla})\mathbf{E} \tag{16}$$

Similarly

$$\mathbf{B} \times \boldsymbol{\nabla} \times \mathbf{B} = \frac{1}{2}\boldsymbol{\nabla}(B^2) - (\mathbf{B}.\boldsymbol{\nabla})\mathbf{B} \tag{17}$$

Substituting equations (16) and (17) in equation (13)

$$\begin{aligned} \mathbf{f} = \quad & \epsilon_0 \left[(\boldsymbol{\nabla}.\mathbf{E})\mathbf{E} - \frac{1}{2}\boldsymbol{\nabla}(E^2) + (\mathbf{E}.\boldsymbol{\nabla})\mathbf{E}\right] + \frac{1}{\mu_0}\left[(\boldsymbol{\nabla}.\mathbf{B})\mathbf{B} - \frac{1}{2}\boldsymbol{\nabla}(B^2) + (\mathbf{B}.\boldsymbol{\nabla})\mathbf{B}\right] \\ & -\epsilon_0 \left[\frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B})\right] \end{aligned} \tag{18}$$

Or

$$\begin{aligned} \mathbf{f} = \quad & \epsilon_0 \left[(\boldsymbol{\nabla}.\mathbf{E})\mathbf{E} + (\mathbf{E}.\boldsymbol{\nabla})\mathbf{E}\right] + \frac{1}{\mu_0}\left[(\boldsymbol{\nabla}.\mathbf{B})\mathbf{B} + (\mathbf{B}.\boldsymbol{\nabla})\mathbf{B}\right] \\ & - \left[\frac{1}{2}\epsilon_0 \boldsymbol{\nabla}(E^2) + \frac{1}{2\mu_0}\boldsymbol{\nabla}(B^2)\right] - \epsilon_0 \left[\frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B})\right] \end{aligned} \tag{19}$$

Above equation can be simplified by introducing the **Maxwell stress tensor**,

$$T_{ij} = \epsilon_0 \left(E_i E_j - \frac{1}{2}\delta_{ij}E^2\right) + \frac{1}{\mu_0}\left(B_i B_j - \frac{1}{2}\delta_{ij}B^2\right) \tag{20}$$

The indices i and j refer to the coordinates x, y, and z, so the stress tensor has a total of nine components $(T_{xx}, T_{xy}, T_{xz}, T_{yx}, T_{yy}, T_{yz}, T_{zx}, T_{zy}, T_{zz})$.

The **Kronecker delta**, δ_{ij}, is 1 if the indices are the same ($\delta_{xx} = \delta_{yy} = \delta_{zz} = 1$) and zero otherwise ($\delta_{xy} = \delta_{yz} = \delta_{zx} = 0$). Thus

$$T_{xx} = \frac{1}{2}\epsilon_0 \left(E_x^2 - E_y^2 - E_z^2\right) + \frac{1}{2\mu_0}\left(B_x^2 - B_y^2 - B_z^2\right) \tag{21}$$

$$T_{yy} = \frac{1}{2}\epsilon_0 \left(E_y^2 - E_x^2 - E_z^2\right) + \frac{1}{2\mu_0}\left(B_y^2 - B_x^2 - B_z^2\right) \tag{22}$$

$$T_{zz} = \frac{1}{2}\epsilon_0 \left(E_z^2 - E_x^2 - E_y^2\right) + \frac{1}{2\mu_0}\left(B_z^2 - B_y^2 - B_x^2\right) \tag{23}$$

and

$$T_{xy} = T_{yx} = \epsilon_0 E_x E_y + \frac{1}{\mu_0} B_x B_y \tag{24}$$

$$T_{yz} = T_{zy} = \epsilon_0 E_y E_z + \frac{1}{\mu_0} B_y B_z \tag{25}$$

$$T_{zx} = T_{xz} = \epsilon_0 E_z E_x + \frac{1}{\mu_0} B_z B_x. \tag{26}$$

Because it carries two indices, where a vector has only one, T_{ij} is sometimes written with a double arrow: $\overleftrightarrow{\mathbf{T}}$. $\overleftrightarrow{\mathbf{T}}$ is a rank-2 tensor, it is represented by a 2-dimensional, 3×3 matrix:

$$\overleftrightarrow{\mathbf{T}} = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix}$$

where $T_{xx}, T_{xy}, T_{xz}, T_{yy}, T_{zz}, T_{yz}$ etc are given in equations (21)-(26).

One can form the dot product of $\overleftrightarrow{\mathbf{T}}$ with a vector $\mathbf{a}$:

$$(\mathbf{a}.\overleftrightarrow{\mathbf{T}})_j = \Sigma_{x,y,z}(a_i T_{ij}) \tag{27}$$

the resulting object, which has one remaining index, is itself a vector. In particular, the divergence of $\overleftrightarrow{\mathbf{T}}$ has as its j^{th} component

$$(\boldsymbol{\nabla}.\overleftrightarrow{\mathbf{T}})_j = \epsilon_0 \left[(\boldsymbol{\nabla}.\mathbf{E})E_j + (\mathbf{E}.\boldsymbol{\nabla})E_j\right] + \frac{1}{\mu_0}\left[(\boldsymbol{\nabla}.\mathbf{B})B_j + (\mathbf{B}.\boldsymbol{\nabla})B_j\right] - \left[\frac{1}{2}\epsilon_0 \boldsymbol{\nabla}_j(E^2) + \frac{1}{2\mu_0}\boldsymbol{\nabla}_j(B^2)\right] \tag{28}$$

Thus the force per unit volume in equation (19) can be written in the much simpler form

$$f = (\boldsymbol{\nabla}.\overleftrightarrow{T}) - \epsilon_0 \mu_0 \frac{\partial \mathbf{S}}{\partial t} \tag{29}$$

where $\mathbf{S} = \mathbf{E} \times \mathbf{H} = \mathbf{E} \times \frac{\mathbf{B}}{\mu_0}$ is the **Poynting vector**.
The total force on the charges in volume V is evidently

$$\mathbf{F} = \int_V \mathbf{f} dV = \int_V (\boldsymbol{\nabla} . \overleftrightarrow{\mathbf{T}}) dV - \epsilon_0 \mu_0 \int_V \frac{\partial \mathbf{S}}{\partial t} dV \tag{30}$$

Or,

$$\boxed{\mathbf{F} = \int_V \mathbf{f} dV = \oint_S \overleftrightarrow{\mathbf{T}} . d\mathbf{a} - \epsilon_0 \mu_0 \int_V \frac{\partial \mathbf{S}}{\partial t} dV \quad (31)}$$

(I used the divergence theorem to convert the first term to a surface integral.) In the static case (or, more generally, whenever $\int_V \mathbf{S} dV$ is independent of time), the second term drops out, and the electromagnetic force on the charge configuration can be expressed entirely in terms of the stress tensor at the boundary.

Physically, $\overleftrightarrow{\mathbf{T}}$ is the force per unit area (or stress) acting on the surface. More precisely,
T_{ij} is the force (per unit area) in the ith direction acting on an element of surface oriented in the jth direction.
The "diagonal" elements (T_{xx}, T_{yy}, T_{zz}) represent pressures, and "off-diagonal" elements (T_{xy}, T_{yz}, T_{zx} , etc.) are shears.

2 Conservation of momentum

According to Newton's second law, the force on an object is equal to the rate of change of its momentum:

$$\mathbf{F} = \frac{\mathbf{p}_{\text{mech}}}{dt} \tag{32}$$

Equation (31) can therefore be written in the form

$$\mathbf{F} = \frac{d\mathbf{p}_{\text{mech}}}{dt} = -\epsilon_0\mu_0 \int_V \frac{\partial \mathbf{S}}{\partial t} dV + \oint_S \overleftrightarrow{\mathbf{T}}.d\mathbf{a} \tag{33}$$

where $\mathbf{p}_{\text{mech}}$ is the total (mechanical) momentum of the particles contained in the volume V. This expression is similar in structure to *Poynting's theorem*, and it invites an analogous interpretation: The first integral represents *momentum stored in the electromagnetic fields themselves*:

$$\mathbf{p}_{\text{em}} = \epsilon_0\mu_0 \int_V \mathbf{S} dV \tag{34}$$

while *the second integral is the momentum per unit time flowing in through the surface*. Equation (33) is the general statement of conservation of momentum in electrodynamics:
Any increase in the total momentum (mechanical plus electromagnetic) is equal to the momentum brought in by the fields.(If V is all of space, then no momentum flows in or out, and $\mathbf{p}_{\text{mech}} + \mathbf{p}_{\text{em}}$ is constant.)

As in the case of conservation of charge and conservation of energy, conservation of momentum can be given a differential formulation. Let $\wp_{\text{mech}}$ be the density of mechanical momentum, and $\wp_{\text{em}}$ the density of momentum in the fields:

$$\wp_{\text{em}} = \epsilon_0\mu_0\mathbf{S} \tag{35}$$

Then Equation (33), in differential form, says

$$\frac{\partial}{\partial t}(\wp_{\text{mech}} + \wp_{\text{em}}) = \boldsymbol{\nabla}.\overleftrightarrow{\mathbf{T}} \tag{36}$$

Evidently -$\overleftrightarrow{\mathbf{T}}$ is the **momentum flux density**, playing the role of $J\mathbf{J}$ (current density) in the continuity equation, or $\mathbf{S}$ (energy flux density) in Poynting's theorem. Specifically, -T_{ij} is the momentum in the i direction crossing a surface oriented in the j direction, per unit area, per unit time. Notice that the Poynting vector has appeared in two quite different roles: $\mathbf{S}$ itself is the energy per unit area, per unit time, transported by the electromagnetic fields, while $\epsilon_0\mu_0\mathbf{S}$ is the momentum per unit volume stored in those fields. Similarly, $\overleftrightarrow{\mathbf{T}}$ plays a dual role: It itself is the electromagnetic stress (force per unit area) acting on a surface, and $\overleftrightarrow{\mathbf{T}}$ describes the flow of momentum (the momentum current density) transported by the fields.