Momentum and Energy Conservation Laws in Electromagnetism

One of the reasons for introducing a Lagrangian formulation of a physical theory is to seek the conservation laws that are hidden in it. This means to study the Lagrangian density with an eye to finding transformations of, in general, both the independent variables(for EM, these are $(\vec{r}, t) \to (\vec{r}'(\vec{r}, t), t'(\vec{r}, t))$) and the dependent variables (for EM, they are $\left(\vec{A}(\vec{r}, t), \Phi(\vec{r}, t), \vec{x}_i(t)\right) \to \left(\vec{A}'(\vec{r}', t'), \Phi'(\vec{r}', t'), \vec{x}_i'(t')\right)$ where the $\vec{x}_i(t)$ is the position of the i^{th} charged particle) that leave the field equations unchanged. Thus under the transformation, the *mathematical* form of the Lagrangian density in the new variables is the same, to within an additive 4-divergence, as the old. Then, if the transformation depends upon a parameter, Nöther's theorem yields a conservation law related to that parameter. This would be the proper approach in a course emphasizing the Hamiltonian Action approach to the fundamentals of electromagnetism, but sadly there is not enough time in this course to show how Nöther's theorem works. Consequently, I will not use the Lagrangian approach to conservation laws to show those of momentum and energy. These four conservation laws arise from the symmetry of the Lagrangian under suitably defined translations in three-space (momentum conservation) and in time (energy conservation). In addition, the Lagrangian density is invariant under rotations in three-space, leading to angular momentum conservation and under relativistic boosts, leading to conservation of a quantity we have no name for, but which is the generalization of the three independent components of the antisymmetric 3×3 angular momentum tensor in three space to the six components of a certain antisymmetric 4×4 tensor in four space. I will not delve into the angular momentum or its generalization, restricting the discussion to an elementary approach to energy-momentum conservation.

Momentum Conservation in Electrodynamics

In general the charges and currents that produce the electromagnetic fields are due to massive charged particles, which carry momentum when moving to produce currents. Let the amount of momentum carried by the particles in volume $d^3 r$ be $\vec{\mathcal{P}}\, d^3 r$. Thus the quantity $\vec{\mathcal{P}}$ is the vectorial momentum density of the same matter generating the charge density ρ and current density $\vec{j}$. As we have seen earlier, there is a force density acting on the charge and current densities which is $\rho \vec{E} + \vec{j}\times\vec{B}$. Since momentum is defined as whatever it is whose time rate of change is the force (this is true both non-relativistically and relativistically) we can write the fundamental equation

$$\frac{\partial \vec{\mathcal{P}}}{\partial t} = \rho \vec{E} + \vec{j}\times\vec{B}.$$

At this point a natural question arises:

Should the charge and current densities be those of the free charges and currents or of the total charges and currents, i.e., including those which are bound???

The answer is that it can be either, or if you wish, whatever portions you wish to include. The only issue is what momentum and force do you wish to deal with. The two most common cases are to deal with the momentum and forces on the free charges and currents (those which an engineer has some direct control over) or you may wish to deal with the momentum and forces of *all* the matter in a problem, i.e., both the free and the bound. Thus we have the two most commonly used forms of the above equation:

$$\frac{\partial \vec{\mathcal{P}}^{\text{free}}}{\partial t} = \rho^{\text{free}} \vec{E} + \vec{j}^{\text{free}} \times \vec{B} \qquad \text{and} \qquad \frac{\partial \vec{\mathcal{P}}^{\text{all}}}{\partial t} = (\rho^{\text{free}} + \rho^{\text{bound}}) \vec{E} + \left(\vec{j}^{\text{free}} + \vec{j}^{\text{bound}}\right) \times \vec{B}\,.$$

The second equation is appropriate if you are interested in the forces acting on a piece of a permanent magnet, for example, or the forces acting on a dielectric in a static electric field, or more generally, a time varying electromagnetic field. What is interesting is that in both the above two cases, we can eliminate the charge and current densities in favor of the fields only.

■ Forces on and Momentum of All the Matter

It is easiest to start with the second equation and work out $\frac{\partial \vec{\mathcal{P}}^{\text{all}}}{\partial t}$. We have from Maxwell's equation

$$\rho^{\text{free}} + \rho^{\text{bound}} = \varepsilon_0 \vec{\nabla} \cdot \vec{E} \qquad \text{and} \qquad \vec{j}^{\text{free}} + \vec{j}^{\text{bound}} = \vec{\nabla} \times \frac{\vec{B}}{\mu_0} - \varepsilon_0 \frac{\partial \vec{E}}{\partial t}.$$

It is easier to work with Cartesian components to avoid having to remember various vector calculus identities (all the ones we need are contained in $\epsilon_{k\ell m}\, \epsilon_{k\ell' m'} = \delta_{\ell\ell'}\, \delta_{mm'} - \delta_{\ell m'}\, \delta_{\ell' m}$ – notice that the second term just has the relevant indices interchanged relative to the first term), and so we have

$$\frac{\partial \mathcal{P}_i^{\text{all}}}{\partial t} = \varepsilon_0 (\partial_j E_j)\, E_i + \epsilon_{ijk} \left(\vec{\nabla} \times \frac{\vec{B}}{\mu_0} - \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right)_j B_k$$

$$= \varepsilon_0 (\partial_j E_j)\, E_i + \epsilon_{ijk} \left(\frac{1}{\mu_0}\, \epsilon_{j\ell m} (\partial_\ell B_m) - \varepsilon_0 \frac{\partial E_j}{\partial t} \right) B_k$$

$$= \varepsilon_0 (\partial_j E_j)\, E_i + \frac{1}{\mu_0} (\delta_{k\ell}\, \delta_{im} - \delta_{km}\, \delta_{\ell i}) (\partial_\ell B_m)\, B_k - \varepsilon_0\, \epsilon_{ijk} \left(\frac{\partial E_j}{\partial t} \right) B_k$$

$$= \varepsilon_0 (\partial_j E_j)\, E_i + \frac{1}{\mu_0} ((\partial_k B_i)\, B_k - (\partial_i B_k)\, B_k) - \varepsilon_0\, \epsilon_{ijk} \left(\frac{\partial}{\partial t} (E_j B_k) - \left(\frac{\partial B_k}{\partial t} \right) E_j \right)$$

Now use the two source-free Maxwell equations, $\partial_j B_j = 0$ and $\epsilon_{k\ell m}\, \partial_\ell E_m = -\frac{\partial B_k}{\partial t}$, and $(\partial_i B_k)\, B_k = \frac{1}{2}\, \partial_i (B_k B_k) = \frac{1}{2}\, \partial_i (B_\ell B_\ell)$ to get, after rearranging a little,

$$\frac{\partial}{\partial t} (\mathcal{P}_i^{\text{all}} + \varepsilon_0\, \epsilon_{ijk}\, E_j B_k) = \varepsilon_0 (\partial_j E_j)\, E_i + \frac{1}{\mu_0} \left(\partial_k (B_i B_k) - \frac{1}{2}\, \partial_i (B_\ell B_\ell) \right) - \varepsilon_0\, \epsilon_{ijk}\, \epsilon_{k\ell m} (\partial_\ell E_m)\, E_j$$

$$= \varepsilon_0 (\partial_j E_j)\, E_i + \partial_k \frac{1}{\mu_0} \left(B_i B_k - \frac{1}{2}\, \delta_{ik}\, B_\ell B_\ell \right) - \varepsilon_0 (\delta_{i\ell}\, \delta_{jm} - \delta_{im}\, \delta_{j\ell}) (\partial_\ell E_m)\, E_j$$

$$= \varepsilon_0 (\partial_j E_j)\, E_i + \partial_k \frac{1}{\mu_0} \left(B_i B_k - \frac{1}{2}\, \delta_{ik}\, B_\ell B_\ell \right) - \varepsilon_0 (\partial_i E_j)\, E_j + \varepsilon_0 (\partial_j E_i)\, E_j$$

$$= \partial_j\, \varepsilon_0 \left(E_j E_i - \frac{1}{2}\, \delta_{ij}\, E_\ell E_\ell \right) + \partial_k \frac{1}{\mu_0} \left(B_i B_k - \frac{1}{2}\, \delta_{ik}\, B_\ell B_\ell \right)$$

$$= \partial_j \left[\varepsilon_0 \left(E_j E_i - \frac{1}{2}\, \delta_{ij}\, E_\ell E_\ell \right) + \frac{1}{\mu_0} \left(B_i B_j - \frac{1}{2}\, \delta_{ij}\, B_\ell B_\ell \right) \right]$$

$$= \partial_j S_{ji}^{\text{all}}$$

where

$$S_{ji}^{\text{all}} = S_{ij}^{\text{all}} = \varepsilon_0 \left(E_j E_i - \frac{1}{2} \delta_{ij} E_\ell E_\ell \right) + \frac{1}{\mu_0} \left(B_i B_j - \frac{1}{2} \delta_{ij} B_\ell B_\ell \right).$$

In vector notation then we get

$$\frac{\partial}{\partial t} \left(\vec{\mathcal{P}}_{\text{matter}}^{\text{all}} + \varepsilon_0 \vec{E} \times \vec{B} \right) = \frac{\partial}{\partial t} \left(\vec{\mathcal{P}}_{\text{matter}}^{\text{all}} + \vec{\mathcal{P}}_{\text{EM fields}}^{\text{all}} \right) = \partial_j \vec{S}_j.$$

This is in the form of a conservation law and it can be interpreted as follows: integrate over all of space and assume that the fields go to zero as the radius of integration goes to infinity, to get

$$\frac{d \vec{P}^{\text{all}}}{d t} = 0 \qquad \text{where} \qquad \vec{P}^{\text{all}} = \vec{P}_{\text{matter}}^{\text{all}} + \vec{P}_{\text{EM field}}^{\text{all}},$$

with the definitions

$$\vec{P}_{\text{matter}}^{\text{all}}(t) = \int \vec{\mathcal{P}}_{\text{matter}}^{\text{all}}(\vec{r}, t)\, d^3 r \qquad \text{and} \qquad \vec{P}_{\text{EM field}}^{\text{all}}(t) = \int \vec{\mathcal{P}}_{\text{EM field}}^{\text{all}}(\vec{r}, t)\, d^3 r$$

We interpret $\vec{P}$ as the total momentum in the system, $\vec{P}_{\text{matter}}$ as the momentum in matter, and $\vec{P}_{\text{EM field}}$ as that in the electromagnetic fields. Thus we see that the momentum density in an electromagnetic field is

$$\vec{\mathcal{P}}_{\text{EM field}}^{\text{all}}(\vec{r}, t) = \varepsilon_0 \vec{E} \times \vec{B} = \vec{D}_{\text{free space}} \times \vec{B}.$$

What is the interpretation of the three vectors $\vec{S}_j$? Just integrate the conservation law over a finite volume V to get

$$\frac{d}{d t} \int_V \left(\vec{\mathcal{P}}_{\text{matter}}^{\text{all}} + \varepsilon_0 \vec{E} \times \vec{B} \right) d^3 r = \int_{\partial V} \vec{S}_j^{\text{all}}\, d A_j.$$

The left hand side is the total momentum in the volume V and so the right hand side must be the vectorial force acting on the volume.

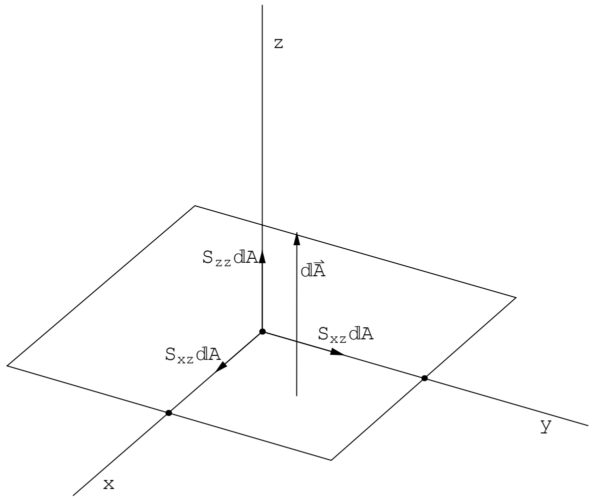
■ Interpretation of the Maxwell's Stress Tensor

To interpret the S_{ij}^{all}, usually called Maxwell's *stress tensor*, let the coordinate system at a point $\vec{r} \in \partial V$ have its z axis in the direction of the normal to the surface there and x and y axes in the surface, as indicated in the sketch (I drop the superscript "all" in the sketch).

■ Begin graphics

■ End graphics

```
Show[stressten]
```



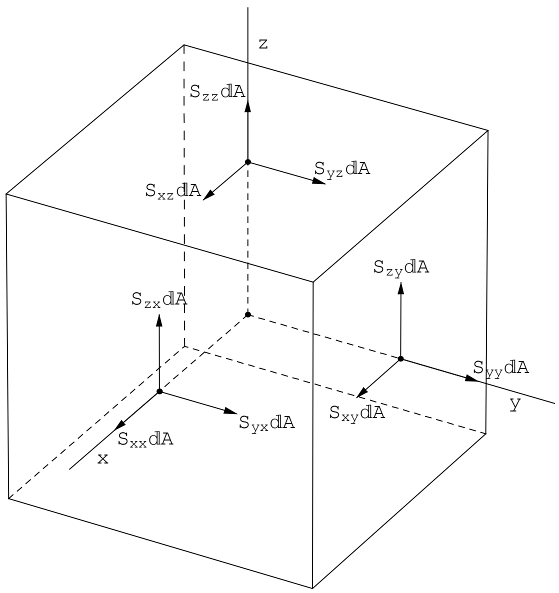
- Graphics -

From the sketch, you can see that, with the outward direction of the differential area vector along the z axis, $S_{zz}\, dA$ is the force action normal to the surface. Thus S_{zz} is a pressure, either positive or negative, and it tends to compress or expand the volume V. The two forces $S_{xz}\, dA$ and $S_{yz}\, dA$ are acting tangentially in the surface; S_{xz} and S_{yz} are called tangential stresses they tend to shear the surface of the volume. It is interesting and instructive to look at the stress forces on an infinitesimal cube, each of its walls having area dA.

■ Begin graphics

■ End graphics

```
Show[stressedcube]
```



- Graphics -

Only the forces on three of the faces is shown – you can easily imagine them on the "back" faces. The stress tensor is to be evaluated at the center of the cube; a more precise calculation would include corrections for the fact that the stresses are acting at points differentially removed from the cube's center. I will ignore these corrections. What is important to notice is the consequence of the symmetry of the stress tensor, $S_{ij} = S_{ji}$. It is easy to see that the torque tending to rotate the cube about the z axis is due to the two forces, $S_{yx}\, dA$ tending to rotate in the positive sense around the z axis, and $S_{xy}\, dA$ in a negative sense. By the symmetry, the torques due to two terms cancel; similarly the torques due to forces on the unshown faces also cancel. Similarly, the torque due to the force $S_{xz}\, dA$ acting in the top surface is cancelled by that due to $S_{zx}\, dA$ acting in the front face. And finally, the net force acting along each of the three axes vanishes to this order ($S_{zz}\, dA$ on the top surface cancelling the similar term on the bottom, not shown). These considerations shown that the net forces of order dA acting on a mass of order $d^3 r$ vanish as they should; there is a net force of order $dA\sqrt{dA}$ and evaluating it just undoes Gauss's theorem. Further the net torque vanishes to order $dA\sqrt{dA}$, the order of size of the mass in the cube. This shows that the stress tensor does not produce rotation. To work out angular momentum issues takes a more refined analysis.

■ Forces on and Momentum of the Free Matter Only

It is formally easy to redo the first part of the above analysis if want the corresponding quantities for the free charges and currents only. Here is the corresponding analysis.

For $\frac{\partial \vec{\mathcal{P}}^{\text{free}}}{\partial t} = \rho^{\text{free}}\vec{E} + \vec{j}^{\text{free}} \times \vec{B}$, we need Maxwell's equations again, but now in the form

$$\rho^{\text{free}} = \vec{\nabla}\cdot\vec{D} \qquad \text{and} \qquad \vec{j}^{\text{free}} = \vec{\nabla}\times\vec{H} - \frac{\partial \vec{D}}{\partial t}.$$

As above, then

$$\frac{\partial \mathcal{P}_i^{\text{free}}}{\partial t} = \rho^{\text{free}} E_i + \epsilon_{ijk}\, j_j^{\text{free}} B_k$$
$$= (\partial_j D_j)\, E_i + \epsilon_{ijk}\left(\epsilon_{j\ell m}(\partial_\ell H_m) - \frac{\partial D_j}{\partial t}\right) B_k$$
$$= (\partial_j D_j)\, E_i + (\partial_k H_i)\, B_k - (\partial_i H_k)\, B_k - \frac{\partial}{\partial t}\,(\epsilon_{ijk}\, D_j\, B_k) + \epsilon_{ijk}\left(\frac{\partial B_k}{\partial t}\right) D_j$$

Again, we now use the other two Maxwell equations and rearrange a bit to get

$$\frac{\partial}{\partial t}\left(\mathcal{P}_i^{\text{free}} + \left(\vec{D}\times\vec{B}\right)_i\right) = (\partial_j D_j)\, E_i + \partial_k (H_i B_k) - (\partial_i H_k)\, B_k - \epsilon_{ijk}\,\epsilon_{k\ell m}(\partial_\ell E_m)\, D_j$$
$$= (\partial_j D_j)\, E_i + \partial_k (H_i B_k) - (\partial_i H_k)\, B_k - (\partial_i E_m)\, D_m + (\partial_\ell E_i)\, D_\ell$$
$$= (\partial_k (D_k E_i) - (\partial_i E_m)\, D_m) + \partial_k (H_i B_k) - (\partial_i H_m)\, B_m.$$

In general we cannot carry the analysis further than this because in the time domain, it is not generally true that the displacement field is proportional to the electric field (and similarly in magnetism). Furthermore, we may not have a homogeneous material. In particular we have

$$\vec{D}(\vec{r},\, t) = \varepsilon_0\left(\vec{E}(\vec{r},\, t) + \int_{-\infty}^{\infty} \chi(\vec{r},\, t - t')\, \vec{E}(\vec{r},\, t')\, dt'\right) \qquad (\chi(\vec{r},\, t) = 0 \text{ for } t < 0)$$

and so it is not generally the case that $(\partial_i E_m)\, D_m = \frac{1}{2}\, \partial_i (E_m D_m)$ *since the polarization at time t is not proportional to the electric field at time t* and so we do not get a conservation law in general.

■ Low Frequency, including Statics, and Homogeneous Material

At low frequencies, it is often a good approximation to take $\vec{P}(\vec{r},\, t) \simeq \varepsilon(\vec{r},\, \omega = 0)\, \vec{E}(\vec{r},\, t)$. Of course tis is exactly true in statics, and for many materials, monatomic gasses for example, it is a good approximation at all frequencies well below optical. The reason is that the lowest energy excitation above the ground state of the atom corresponds to optical (or higher) frequencies. For diatomic or more complicated molecules, there are relatively low energy transitions associated with atomic motion (vibration and rotation bands), making $\varepsilon(\vec{r},\, \omega = 0)$ a good approximation for frequencies corresponding to wavelengths in the millimeter to centimeter range. *So at low frequencies, including statics, we can make the approximation that* ε and μ *are time independent and set*

$$\vec{D}(\vec{r},\, t) = \varepsilon(\vec{r})\, \vec{E}(\vec{r},\, t) \qquad \text{and} \qquad \vec{B}(\vec{r},\, t) = \mu(\vec{r})\, \vec{H}(\vec{r},\, t).$$

However, this still does not let us convert the $(\partial_k (D_k E_i) - (\partial_i E_m)\, D_m) + \partial_k (H_i B_k) - (\partial_i H_m)\, B_m$ into a divergence because $(\partial_i E_m)\, D_m = (\partial_i E_m)\, \varepsilon(\vec{r})\, E_m = \frac{1}{2}\, (\partial_i E_m\, E_m)\, \varepsilon(\vec{r}) \neq \frac{1}{2}\, \partial_i (E_m\, E_m\, \varepsilon(\vec{r}))$ unless $\partial_i \varepsilon = 0$.In other words, only when we are treating homogeneous material at low frequencies can we write $(\partial_i E_m)\, D_m = \frac{1}{2}\, \partial_i (E_m D_m) = \delta_{ik}\, \partial_k \left(\frac{1}{2}\, \vec{D}\cdot\vec{E}\right)$, and similarly $(\partial_i H_m)\, B_m = \frac{1}{2}\, \delta_{ik}\, \partial_k \left(\vec{B}\cdot\vec{H}\right)$. Thus at low frequency in an homogeneous material so that $\vec{D} = \varepsilon\, \vec{E}$ and $\vec{B} = \mu\, \vec{H}$we get the following conservation law

$$\frac{\partial}{\partial t}\left(\vec{\mathcal{P}}^{\text{free}} + \vec{D}\times\vec{B}\right) = \partial_j \vec{S}_j^{\text{free}}$$

where $(\vec{S}_j^{\text{free}})_i = S_{ij}^{\text{free}} = S_{ji}^{\text{free}} = \left(D_i E_j - \frac{1}{2}\delta_{ij}\vec{D}\cdot\vec{E}\right) + \left(B_i H_j - \frac{1}{2}\delta_{ij}\vec{B}\cdot\vec{H}\right)$.

So in these special cases we get the field's contribution to the momentum density associated with the free charges and currents to be $\vec{D}\times\vec{B}$.

Energy Conservation in Electrodynamics

The relevant quantity we now need is the rate at which a force $\vec{F}$ does work on a particle moving with velocity $\vec{v}$. It is simply $\vec{F}\cdot\vec{v}$, and this is the rate at which the particle's kinetic energy is increasing. If we have many charged particles q_i at position $\vec{x}_i(t)$ then the rate at which energy is being given to the i^{th} particle by an electromagnetic field is $q_{(i)}\left(\vec{E}(\vec{x}_i(t), t) + \vec{v}_i(t)\times\vec{E}(\vec{x}(t), t)\right)\cdot\vec{v}_i(t) = q_{(i)}\,\vec{v}_i(t)\cdot\vec{E}(\vec{x}_i(t), t)$ (no sum on i). Summing over all of the particles in a differential volume $d^3 r$ at $\vec{r}$ we get

$$\sum_{\substack{i \\ \vec{x}_i\in d^3 r}} q_{(i)}\,\vec{v}_i(t)\cdot\vec{E}(\vec{x}_i(t), t) = \vec{j}(\vec{r}, t)\cdot\vec{E}(\vec{r}, t)\,d^3 r$$

If we let $W(\vec{r}, t)$ be the matter's energy density at $(\vec{r}, t)$ then we have

$$\frac{\partial W}{\partial t} = \vec{j}\cdot\vec{E}$$

As with the momentum, we have two cases to consider: the kinetic energy of *all* the matter or just the matter associated with free charges and currents. First choose the "all" case.

■ Energetics for All the Matter

$$\begin{aligned}
\frac{\partial W^{\text{all}}}{\partial t} &= \vec{j}^{\text{all}}\cdot\vec{E} = E_i\left(\epsilon_{ijk}\,\partial_j\,\frac{B_k}{\mu_0} - \varepsilon_0\,\frac{\partial E_i}{\partial t}\right) \\
&= \frac{1}{\mu_0}\,\epsilon_{ijk}\,\partial_j(E_i B_k) - \frac{1}{\mu_0}\,B_k\,\epsilon_{ijk}\,\partial_j E_i - \varepsilon_0\,\frac{1}{2}\,\frac{\partial(E_i E_i)}{\partial t} \\
&= \frac{1}{\mu_0}\,\epsilon_{ijk}\,\partial_j(E_i B_k) + \frac{1}{\mu_0}\,B_k\,\epsilon_{kji}\,\partial_j E_i - \varepsilon_0\,\frac{1}{2}\,\frac{\partial(E_i E_i)}{\partial t} \\
&= \frac{1}{\mu_0}\,\epsilon_{ijk}\,\partial_j(E_i B_k) + \frac{1}{\mu_0}\,B_k\,\epsilon_{kji}\,\partial_j E_i - \varepsilon_0\,\frac{1}{2}\,\frac{\partial(E_i E_i)}{\partial t} \\
&= -\frac{1}{\mu_0}\,\partial_j(\epsilon_{jik}\,E_i B_k) - \frac{1}{\mu_0}\,B_k\,\frac{\partial B_k}{\partial t} - \varepsilon_0\,\frac{1}{2}\,\frac{\partial(E_i E_i)}{\partial t} \\
&= -\frac{1}{\mu_0}\,\partial_j(\epsilon_{jik}\,E_i B_k) - \frac{1}{\mu_0}\,\frac{1}{2}\,\frac{\partial(B_k B_k)}{\partial t} - \varepsilon_0\,\frac{1}{2}\,\frac{\partial(E_i E_i)}{\partial t}
\end{aligned}$$

So we get the conservation law for energy:

$$\frac{\partial}{\partial t}\left(W^{\text{all}} + \frac{1}{2}\,\frac{1}{\mu_0}\,\vec{B}\cdot\vec{B} + \frac{1}{2}\,\varepsilon_0\,\vec{E}\cdot\vec{E}\right) = -\vec{\nabla}\cdot\left(\frac{1}{\mu_0}\,\vec{E}\times\vec{B}\right).$$

Integrate this over a volume V to get the interpretation.

$$\frac{d}{dt}\int\left(W^{\text{all}}+\frac{1}{2}\frac{1}{\mu_0}\vec{B}\cdot\vec{B}+\frac{1}{2}\varepsilon_0\vec{E}\cdot\vec{E}\right)d^3r=-\int\left(\frac{1}{\mu_0}\vec{E}\times\vec{B}\right)\cdot d\vec{A}$$

The interpretation is that $\frac{1}{2}\frac{1}{\mu_0}\vec{B}\cdot\vec{B}+\frac{1}{2}\varepsilon_0\vec{E}\cdot\vec{E}$ is the density of stored energy in the electromagnetic field which when added to the energy density stored in the matter gives the total stored energy density in the matter. The interpretation of $-\frac{1}{\mu_0}\vec{E}\times\vec{B}$ is that it is the energy per unit area per unit time flowing into the volume supporting the increase in energy stored in V. Consequently $+\frac{1}{\mu_0}\vec{E}\times\vec{B}$ is the flow of energy per unit area per unit time flowing out of the volume.