Chapter 7
Energy and momentum

June 21, 2021

Expressions are derived for the energy and momentum density and flux, proving the conservation of total energy and momentum. The latter provides the consistency condition, in the sense of Newton's third law, between the Lorentz force law and the Maxwell equations.

Contents

1 Introduction **1**
1.1 Need for consistency 1
1.2 The framework for analysis 2

2 Energy **2**
2.1 Work done 2
2.2 Energy density and flux 2
2.3 Application to plane wave 3
2.4 Charging a capacitor 3

3 Momentum **4**
3.1 Conservation law 4
3.2 Derivation 4
3.3 Some applications 5

4 The stress tensor **5**
4.1 Tensor character and symmetry . . . 5
4.2 Stress in continuum mechanics . . . 5
4.3 Model for point particles 6

5 Uniqueness **6**
5.1 Possible alternatives? 6
5.2 It does not matter 7
5.3 Time evolution 7
5.4 Gravity 7

A Supplement: Force law from Maxwell's equations **7**

B Supplement: Magnetars **8**

C Supplement: Alfven waves **9**

D Supplement: Pinning of magnetic flux lines **10**
D.1 The phenomenon 10
D.2 Basic equations in magnetohydrodynamics 10
D.3 Derivation of pinning 10

1 Introduction

1.1 Need for consistency

The Lorentz force law (L) specifies how EM fields act on charged particles, and thereby deliver energy and momentum to matter. Maxwell's equations (M) specify how the charged particles (including the currents due to their motion) produce EM fields, and thereby deliver energy and momentum to the fields. The reciprocal relationship is illustrated schematically in **Figure 1**. If total energy and momentum are to be conserved, then the two processes have to be equal and opposite — in the case of momentum, equivalent to a generalized notion of Newton's third law. Thus there are consistency conditions between the two sets of physical laws L and M.

This Chapter analyzes this issue, and derives expressions for energy and momentum density and flux that satisfy the required consistency conditions.

Actually many of the results in this Chapter can be obtained more neatly using advanced concepts.

- There is a standard method to get the Hamiltonian, i.e., the energy, from the Lagrangian. It is easy to generalize this idea to obtain the energy density from the Lagrangian density.
- Energy and momentum can be treated together as a 4-vector.
- Density and flux can be treated as a 4-vector. Thus the (density and flux) of (energy and momentum) carries two 4-vector indices — it is a tensor. Knowing how to obtain one component (the energy density) makes it fairly obvious how to obtain the others.

Nevertheless we shall not adopt such an approach here, so that students can learn more physically about energy and momentum without having to first learn these advanced concepts.

1.2 The framework for analysis

The framework for analysis (**Figure 2**) will be explained for energy, and exactly the same idea will apply for (each component of) momentum.

Consider a volume V and conceptually divide it into two subsystems: the matter and the field. Let the work done by the field on matter, per unit volume per unit time, be denoted as[1] $\mathcal{W}$.

First look at the matter in this volume. It receives energy per unit time

$$\int_V \mathcal{W}\, d^3r$$

In addition, there is some outflow of energy of matter per unit area per unit time,[2] due to the flux $\mathbf{S}^m$, through the surface ∂V bounding V:

$$\oint_{\partial V} \mathbf{S}^m \cdot d^2r \quad = \quad \int_V \boldsymbol{\nabla} \cdot \mathbf{S}^m \; d^3r$$

(In this Chapter, except in Section 4, we shall reserve the symbol S and $\mathbf{S}$ for the energy flux; therefore elements of area will be denoted as d^2r and not dS or $d\mathbf{S}$.) As a result of this, the total energy changes by

$$\frac{d}{dt}\int_V \mathcal{U}^m \, d^3r$$

where $\mathcal{U}^m$ is the energy density of matter. These three must balance, leading to the condition

$$\boxed{\mathcal{W} = \frac{\partial \mathcal{U}^m}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{S}^m} \tag{1}$$

In exactly the same way, but recognizing that the work now enters with the opposite sign, the fields in this volume must satisfy

$$\boxed{-\mathcal{W} = \frac{\partial \mathcal{U}}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{S}} \tag{2}$$

where $\mathcal{U}$ and $\mathbf{S}$ are the energy density and the energy flux of the field.

The two equations together, linked by the same $\mathcal{W}$, then express the conservation of energy, indeed the *local* conservation of energy. The analysis of (1), purely within (fluid) mechanics, will be omitted. We focus on the following questions in relation to (2):

- What is the rate of work done per unit volume, $\mathcal{W}$?
- Can it be expressed as the sum of a time derivative and a divergence, both expressed in terms of the fields? If so, we would have identified the energy density and the energy flux, and proved conservation.

All these considerations can be repeated for (each component of) momentum.

2 Energy

2.1 Work done

The work done per unit time on a point charge is

$$W \quad = \quad \mathbf{F} \cdot \mathbf{v} = q\mathbf{E} \cdot \mathbf{v} = (q\mathbf{v}) \cdot \mathbf{E}$$

where we have put in the Lorentz force law for $\mathbf{F}$ and recognized that the magnetic field does no work. When this is generalized to a unit volume, obviously we find, since the sum of $q\mathbf{v}$ over the unit volume is just $\mathbf{J}$:

$$\boxed{\mathcal{W} = \mathbf{J} \cdot \mathbf{E}} \tag{3}$$

2.2 Energy density and flux

The task required is to rewrite $-\mathcal{W}$ in the form shown on the RHS of (2), which should only involve the fields.

Step 1

First use Maxwell's equations to express $\mathbf{J}$ in terms of the fields, namely, from Ampere's law,

$$\mathbf{J} \quad = \quad \frac{1}{\mu_0}\boldsymbol{\nabla} \times \mathbf{B} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

giving

$$\begin{aligned} & -\mathcal{W} \\ = \quad & -\left(\frac{1}{\mu_0}\boldsymbol{\nabla} \times \mathbf{B} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right) \cdot \mathbf{E} \\ = \quad & -\frac{1}{\mu_0}(\boldsymbol{\nabla} \times \mathbf{B}) \cdot \mathbf{E} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \cdot \mathbf{E} \end{aligned} \tag{4}$$

Step 2

The next step is to manipulate (4) so that it becomes a time derivative plus a divergence.

The first term in (4) is, as usual adopting the summation convention over repeated indices,

$$\begin{aligned} & -\frac{1}{\mu_0}(\epsilon_{ijk}\partial_j B_k)\, E_i \\ = \quad & -\frac{1}{\mu_0}\epsilon_{ijk}\left[\partial_j(B_k E_i) - B_k(\partial_j E_i)\right] \end{aligned}$$

[1] Densities will be denoted by script letters.

[2] This comes about through the flow of particles through the boundary ∂V, carrying energy with them.

$$
\begin{aligned}
&= -\frac{1}{\mu_0}\partial_j(\epsilon_{ijk}B_k E_i) + \frac{1}{\mu_0}B_k(\epsilon_{ijk}\partial_j E_i) \\
&= \frac{1}{\mu_0}\boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{B}) - \frac{1}{\mu_0}\mathbf{B}\cdot(\boldsymbol{\nabla}\times\mathbf{E}) \\
&= \frac{1}{\mu_0}\boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{B}) + \frac{1}{\mu_0}\mathbf{B}\cdot\frac{\partial\mathbf{B}}{\partial t} \\
&= \frac{1}{\mu_0}\boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{B}) + \frac{\partial}{\partial t}\left(\frac{1}{2\mu_0}B^2\right) \qquad (5)
\end{aligned}
$$

The second term in (4) is

$$
\epsilon_0\frac{\partial\mathbf{E}}{\partial t}\cdot\mathbf{E} = \frac{\partial}{\partial t}\left(\frac{\epsilon_0}{2}E^2\right) \qquad (6)
$$

These together then allow us to identify the energy density and the energy flux as

$$
\boxed{\mathcal{U} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2}
$$

$$
\boxed{\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}} \qquad (7)
$$

- The expression for $\mathbf{S}$ is called the *Poynting vector.*
- Note that $\mathcal{U}$ is positive definite. (It is interesting to ask what happens if Faraday's law has the opposite sign.)
- The energy density can also be written as

  $$\mathcal{U} = \frac{1}{2\mu_0}\left[(E/c)^2 + B^2\right]$$

  showing that $\mathbf{B}$ and $\mathbf{E}/c$ should be regarded as parallel.
- Note that there is no interaction term. In fact, the expression for $\mathcal{U}$ already incorporates what we normally call the interaction term, as can be seen from the following problem.

Problem 1
Consider the case of electrostatics. The energy of a charge q at position $\mathbf{r}$ due to an external potential is $q\Phi(\mathbf{r})$. Generalizing to a charge distribution we get

$$\mathcal{U} = \int \rho(\mathbf{r})\Phi(\mathbf{r})\, d^3r$$

Show that this expression agrees with the electric term in (7). There is a subtle factor of 2. §

2.3 Application to plane wave

We shall later analyze plane wave solutions to Maxwell's equation in vacuum, but the results should be well known already. We here use the example of a plane wave to illustrate some properties of the energy density and the energy flux.

Plane wave solution
Consider a plane wave propagating in the $+z$ direction, say

$$\mathbf{E} = E_0\,\hat{\mathbf{e}}_x\,\cos(kz-\omega t)$$

where $\omega/k = c$. The magnetic field must have a similar dependence, namely, only on $kz-\omega t$, so from

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

applied to the y-component, we get

$$
\begin{aligned}
-\partial_t B_y &= \partial_z E_x = -E_0 k\sin(kz-\omega t) \\
B_y &= E_0(k/\omega)\cos(kz-\omega t)
\end{aligned}
$$

Obviously this is the only component, so

$$\mathbf{B} = (E_0/c)\,\hat{\mathbf{e}}_y\,\cos(kz-\omega t)$$

- $\mathbf{E}$ and $\mathbf{B}$ are in phase.
- The amplitudes are in the ratio $|\mathbf{B}| = |\mathbf{E}|/c$.

Properties of energy and density and flux
We now evaluate $\mathcal{U}$ and $\boldsymbol{S}$ for this solution.

Problem 2
Applied to the above solution, show that
(a) the electric energy density and the magnetic energy density are equal;
(b) the Poynting vector is in the direction of propagation; and
(c) the magnitude of the Poynting vector is $S = c\mathcal{U}$. §

2.4 Charging a capacitor

Next consider the charging of a capacitor. During the charging process, there is an energy flux $\mathbf{S}$ flowing into the capacitor gap. The point is to check that the total flow of energy during the charging process just equals the final energy stored.

Problem 3
A capacitor is formed by two circular disks of radius R, separated by a distance h, with $h \ll R$ (**Figure 3**). The capacitor is charged linearly with time: $V(t) = V_0(t/\tau)$.
(a) Show that during the process of charging, there is energy flow into the gap.
(b) Show that the total energy that flows into the gap up to $t=\tau$ is exactly equal to the energy stored in the gap.
(c) Generalize to the case where $V(t) = V_0 f(t)$, where the final value is $f(\tau) = 1$. §

3 Momentum

3.1 Conservation law

We now repeat the same argument for each component of the momentum. Evidently the condition should be

$$\boxed{-\mathcal{F}_i = \frac{\partial \Pi_i}{\partial t} + \partial_j \mathcal{T}_{ij}} \tag{8}$$

where $\mathcal{F}_i$ is the i-component of the force density, i.e., the rate at which the i-component of momentum is delivered to matter, per unit mass; Π_i is the i-component of the momentum density of the field; and $\mathcal{T}_{ij}$ is the flux of the i-component of momentum in the j direction. In other words, the amount of the i-component of momentum flowing across a surface $\hat{\mathbf{n}}\, d^2r$ is

$$\mathcal{T}_{ij} n_j\, d^2r$$

The last term in (8) is a generalized divergence.

3.2 Derivation

The EM force on a unit volume is

$$\mathcal{F}_i \quad = \quad \rho E_i + \epsilon_{ijk} J_j B_k \tag{9}$$

Step 1
The first step is to use Maxwell's equations to eliminate ρ and $\mathbf{J}$. Putting in

$$\begin{aligned} \rho &= \epsilon_0 \partial_j E_j \\ J_j &= \frac{1}{\mu_0} \epsilon_{jmn}(\partial_m B_n) - \epsilon_0(\partial_t E_j) \end{aligned}$$

we find

$$\begin{aligned} & -\mathcal{F}_i = -\rho E_i - \epsilon_{ijk} J_j B_k \\ = & -\epsilon_0(\partial_j E_j)E_i - \frac{1}{\mu_0}(\epsilon_{ijk}\epsilon_{jmn})(\partial_m B_n)B_k \\ & +\epsilon_0 \epsilon_{ijk}(\partial_t E_j)B_k \end{aligned} \tag{10}$$

Step 2
The next step is again to manipulate this expression into the form of a time derivative plus a generalized divergence.

Denote the three terms in (10) as Q_1, Q_2, Q_3. The third term is

$$\begin{aligned} & Q_3 = \epsilon_0 \epsilon_{ijk}\left[(\partial_t E_j)B_k\right] \\ = & \ \epsilon_0 \epsilon_{ijk}\left[\partial_t(E_j B_k) - E_j(\partial_t B_k)\right] \\ = & \ \partial_t(\epsilon_0 \epsilon_{ijk} E_j B_k) \\ & + \epsilon_0 \epsilon_{ijk} E_j [\epsilon_{kmn}(\partial_m E_n)] \end{aligned} \tag{11}$$

where in the last step we have used Faraday's law to eliminate the time derivative of $\boldsymbol{B}$.

The first term in (11), which we shall denote as Q_{3a}, can be written as

$$\begin{aligned} Q_{3a} &= \partial_t \Pi_i \\ \mathbf{\Pi} &= \epsilon_0 \mathbf{E} \times \mathbf{B} = \mathbf{S}/c^2 \end{aligned} \tag{12}$$

The remaining terms involving $\mathbf{E}$ are

$$\begin{aligned} & Q_1 + Q_{3b} \\ = & \ \epsilon_0\left[-(\partial_j E_j)E_i + (\epsilon_{kij}\epsilon_{kmn})E_j(\partial_m E_n)\right] \\ = & \ \epsilon_0\left[-(\partial_j E_j)E_i + (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})E_j(\partial_m E_n)\right] \\ = & \ \epsilon_0\left[-(\partial_j E_j)E_i + E_j(\partial_i E_j) - E_j(\partial_j E_i)\right] \\ = & \ \epsilon_0\left[\frac{1}{2}\partial_i(E_j E_j) - \partial_j(E_i E_j)\right] \\ = & \ \partial_j\left\{\epsilon_0\left(\frac{1}{2}E^2\delta_{ij} - E_i E_j\right)\right\} \end{aligned}$$

This can be regarded as a contribution to $\partial_j \mathcal{T}_{ij}$, denoted with a superscript ^{E}, with

$$\boxed{\mathcal{T}_{ij}^E = \epsilon_0\left(\frac{1}{2}E^2\delta_{ij} - E_i E_j\right)} \tag{13}$$

Finally, the terms involving $\mathbf{B}$ are

$$\begin{aligned} & Q_2 = -\frac{1}{\mu_0}(\epsilon_{ijk}\epsilon_{jmn})(\partial_m B_n)B_k \\ = & -\frac{1}{\mu_0}(\epsilon_{jki}\epsilon_{jmn})(\partial_m B_n)B_k \\ = & -\frac{1}{\mu_0}(\delta_{km}\delta_{in} - \delta_{kn}\delta_{im})(\partial_m B_n)B_k \\ = & -\frac{1}{\mu_0}\left[(\partial_k B_i)B_k - (\partial_i B_k)B_k\right] \\ = & -\frac{1}{\mu_0}\left[\partial_k(B_i B_k) - \frac{1}{2}\partial_i(B_k B_k)\right] \\ = & \ \partial_j\left\{\frac{1}{\mu_0}\left(\frac{1}{2}B^2\delta_{ij} - B_i B_j\right)\right\} \end{aligned}$$

In reaching the second last line, in the first term we have in effect added a term proportional to $\partial_k B_k = 0$. The above expression can be regarded as another contribution to $\partial_j \mathcal{T}_{ij}$, denoted with a superscript ^{M}, with

$$\boxed{\mathcal{T}_{ij}^M = \frac{1}{\mu_0}\left(\frac{1}{2}B^2\delta_{ij} - B_i B_j\right)} \tag{14}$$

Summary
Thus we have derived

$$\begin{aligned} \mathbf{\Pi} &= \epsilon_0 \mathbf{E} \times \mathbf{B} = \mathbf{S}/c^2 \\ \mathcal{T}_{ij} &= \mathcal{T}_{ij}^E + \mathcal{T}_{ij}^M \end{aligned} \tag{15}$$

with the two terms in the last equation given above. In terms of these quantities, energy and momentum are conserved.

This laborious calculation verifies the consistency between the Lorentz force law and the Maxwell equations. Of course, you might wonder whether there is a more natural way to guarantee the consistency.

3.3 Some applications

One plane wave
Consider a plane wave propagating along $+z$, with $\mathbf{E}$ along x and $\mathbf{B}$ along y. We have already seen that the electric and magnetic energy densities are equal, i.e.,

$$\frac{\epsilon_0}{2}E^2 \quad = \quad \frac{1}{2\mu_0}B^2$$

So

$$\mathcal{U} \quad = \quad \epsilon_0 E^2$$

(Usually we would take time average and E^2 should be thus understood.)

Problem 4
Show that for a plane wave as described above, $\mathcal{T}_{ij} = 0$ if $i \neq j$, and $\mathcal{T}_{zz} = \mathcal{U}$, $\mathcal{T}_{xx} = \mathcal{T}_{yy} = 0$. §

Random plane waves
Consider a region where there are random plane waves in all directions, so that there is isotropy; this would be the case in a hot cavity containing blackbody radiation. Isotropy requires

$$E_i E_j \quad \mapsto \quad \frac{1}{3}E^2\,\delta_{ij}$$

and likewise for $B_i B_j$.

Problem 5
With the above replacement, show that

$$\boxed{\mathcal{T}_{ij} = \frac{1}{3}\mathcal{U}\,\delta_{ij}} \tag{16}$$

which is the equation of state for a gas of photons. §

4 The stress tensor

This section provides a physical interpretation of $\mathcal{T}_{ij}$.

4.1 Tensor character and symmetry

First, the quantities $\mathcal{T}_{ij}$ transform like a 3-tensor — roughly speaking, each of its two indices transforms like a 3-vector. This means that quantities such as $\partial_j \mathcal{T}_{ij}$ transform like the i component of a 3-vector. That ensures that all the equations we have written above remain valid upon rotation to other coordinate frames.

Second, it is obvious that $\mathcal{T}_{ij}$ is symmetric in its indices, so there are only 6 independent components.

Third, it is easy to show (without assuming isotropy and thus going beyond Problem 5) that the trace is just the energy density

$$\mathcal{T}_{ii} = \mathcal{U} \tag{17}$$

consistent with (16).

4.2 Stress in continuum mechanics

Definition
We now recall some concepts in continuum mechanics: either solid elasticity or fluid mechanics. **Figure 4a** shows an element of the material. To discuss the force on each surface (due to the neighboring element), we have to specify two directions: the direction j which characterizes the normal to the surface, and the direction i of the force; the relevant *force per unit area* is called the ij component of *stress*, or more precisely the *stress tensor*, usually denoted as σ_{ij}. To be more precise, the force $d\mathbf{F}$ acting on a surface element $d\mathbf{S}$ is

$$dF_i \quad = \quad \sigma_{ij}\,dS_j \tag{18}$$

Symmetry
For a small element, here shown in cross-section in the x–y plane (**Figure 4b**), the surface forces are as shown. The forces on opposite faces act in opposite directions because $d\mathbf{S}$ for the two surfaces are opposite. The torque about the center caused by the horizontal forces is

$$\begin{aligned} \tau_1 &= (\sigma_{xz}\Delta x\Delta y)\cdot\Delta z \\ &= \sigma_{xz}\Delta V \end{aligned}$$

In the first line, Δx, Δy, Δz are the dimensions of the sides, the first bracket is the force, and the last factor is the moment arm,[3] while in the last line, ΔV is the volume. In exactly the same way, the torque caused by the pair of vertical forces is

$$\tau_2 \quad = \quad -\sigma_{zx}\Delta V$$

[3] If moments are taken about the center, then the moment arm is half this amount, but there is an equal contribution from the force on the opposite face.

and the net torque is

$$\begin{aligned}\tau &= \tau_1 + \tau_2 \\ &= (\sigma_{xz} - \sigma_{zx})\,\Delta V \qquad (19)\end{aligned}$$

This goes as ϵ^3 if the linear dimension is $O(\epsilon)$.

But the moment of inertia must be of order ϵ^5, leading to an angular acceleration of order ϵ^{-2} — which is not possible for $\epsilon \to 0$. Thus we conclude that the bracket in (19) must vanish: σ_{ij} is symmetric.

Pressure
An element of stress such as σ_{xx} relates to a force in the x direction, acting on a surface whose normal is also in the x direction; thus it is a *normal* force. Such a normal force per unit area is usually called the *pressure*.[4]

EM stress tensor
Recall that $\mathcal{T}_{ij}$ in EM is defined as the *flux of momentum per unit time* (momentum component i) per unit area (surface normal along j). Here we have σ_{ij} as the *force* (force component i) per unit area (surface normal along j). But momentum per unit time is just force.

Therefore the EM tensor $\mathcal{T}_{ij}$ has exactly the same interpretation as the stress tensor in continuum mechanics, and is therefore called the EM *stress tensor*. Henceforth we shall also use $\mathcal{T}_{ij}$ for continuum mechanics.

4.3 Model for point particles

To gain a more vivid picture, suppose the continuum contains an ensemble of particles in motion. Let there be n particles per unit volume, with (individual) momentum $\mathbf{p} = \gamma m \mathbf{v}$. where $\gamma = (1 - v^2/c^2)^{-1/2}$ allows for the relativistic case.

Consider the surface in **Figure 5** perpendicular to the x direction. In a time Δt, all the particles in a volume $v_x \Delta t \times \Delta y \Delta z$ would have flowed past it, so the y component of momentum carried across the surface is

$$(n v_x \Delta t \Delta y \Delta z)\, p_y$$

The bracket is the number of particles involved. Factoring out the time interval and the area gives the stress; hence

$$\mathcal{T}_{yx} = n p_y v_x \qquad (20)$$

[4]If we only require the force to be normal in one frame, that would allow three independent diagonal entries. But unless they are equal, upon rotation to a different frame off-diagonal components will appear. Thus the stress tensor for a pure pressure is not only diagonal, but proportional to the identity.

The same argument applies to all components. We also note that $\mathbf{p}$ can be expressed in terms of $\mathbf{v}$, and since velocities of the particles are all different, the RHS should be understood as an average. Thus

$$\boxed{\mathcal{T}_{ij} = nm\,\langle \gamma v_i v_j \rangle} \qquad (21)$$

which shows its symmetry.

Problem 6
Derive the familiar expression for the pressure of a non-relativistic ideal gas. §

For a gas of photons, it is more appropriate to write

$$p_i = (E/c^2) v_i$$

so that (20) leads to

$$\mathcal{T}_{ij} = n(E/c^2)\langle v_i v_j \rangle$$

But the velocities all have magnitude c and are in random directions, so

$$\langle v_i v_j \rangle = \frac{c^2}{3}\delta_{ij}$$

Problem 7
Hence recover (16). §

5 Uniqueness

5.1 Possible alternatives?

We have derived *one* set of formulas for the energy density, energy flux, momentum density and momentum flux of the EM field, and shown that these are conserved together with the energy and momentum of matter. How do we know these are *the* correct expressions for the energy etc. of the EM field? In other words, might there not be other possible expressions, which when put into the formulas such as (2) would still be correct?[5]

So what are the possible alternatives? Let us, for simplicity, discuss just the energy. Since relativistic invariance has to be maintained, if any changes are made to the energy, then corresponding changes must be made to the momentum.[6] The question can be posed as the possibility of adding some quantities $\mathcal{U}'$ and $\mathbf{S}'$ made out of the fields, such that their integrals over all space are zero, and moreover

[5]I am grateful to Pui-tak Leung for drawing my attention to this question.

[6]In fact, this whole discussion is best carried out in a covariant formalism, but there is no need for the greater mathematical complexity to obscure the ideas.

$$\frac{\partial \mathcal{U}'}{\partial t} + \nabla \cdot \mathbf{S}' \quad = \quad 0 \qquad (22)$$

As a purely mathematical theorem (which we shall not prove), it can be shown that (22) has no solution if $\mathcal{U}'$ and $\mathbf{S}'$ are to be made up of the components of $\mathbf{E}$ and $\mathbf{B}$, but there are many nontrivial solutions if derivatives of the fields are allowed.

5.2 It does not matter

The narrow answer is that it does not matter. After all, the whole idea of energy is only an *accounting tool*; it is some conserved quantity which helps us keep track of certain aspects of the physical process. For example, if a plane wave is (partially) absorbed, then the energy transferred to matter (which can be measured through the increase in kinetic energy of the particles, for example) is just $-\Delta\mathcal{U}$, where $\mathcal{U}$ is *any* expression that satisfies the conservation laws. Certainly the one that emerged in this Chapter is the simplest one.

In fact, going back to elementary mechanics, you could have asked the same question. Expressions such as $K = (1/2)mv^2$ satisfy conservation laws and K is called the kinetic energy. Could there be other expressions that work just as well? The answer is that we do not care. All we need is that *this* expression works.

However, the question of uniqueness can also be considered from other perspectives.

5.3 Time evolution

Energy is not just a conserved quantity; it is related to the Hamiltonian which governs time evolution, whether in classical physics or quantum mechanics. Thus, in general, if we change the expression for energy, we would change the equations of motion — Maxwell's equations, which are well verified by experiments.

The objection is not only experimental in the narrow sense. If we add some extra terms (which would involve higher derivatives), the resulting equations of motion would be *qualitatively* different — they would likewise involve higher derivatives, taking us to a totally different *type* of theory difficult to reconcile with experiments.

However this line of reasoning begs a question: We have not yet shown that *this* choice of the expression for the energy actually leads, in the Hamiltonian formulation, to the Maxwell equations. This gap will be filled later in this course, at least for some simple contexts.

5.4 Gravity

Energy has yet another role. Energy is (apart from a trivial factor of c^2) the same as mass, and therefore should participate in gravity. So how do EM fields participate in gravity?

At the theoretical level, we just put whatever expression for the energy and momentum (more precisely the energy-momentum tensor) into Einstein's theory. Different proposals would have different experimental consequences, which could in principle be distinguished.

There are two types of gravitational phenomena with respect to EM fields. First, EM fields should *produce* gravity, just as mass produces gravity. It will be a very long time before any such effect can be measured in nontrivial circumstances. Second, EM fields *respond* to gravity, and this effect was observed soon after general relativity was formulated — through the bending of light rays passing through a gravitational field. But these experiments are also not sensitive to the present issue, since the path of the photon is determined by a geodesic condition, which would not be altered even if there is a different expression for the energy of a photon.

In summary, one could say that these arguments do not, logically speaking, rule out other expressions for the energy. All it means is that any nontrivial proposals must be considered together with all the consequences — and when all these are considered, nobody has found another possibility. And indeed, in classical EM, it does not matter.

A Supplement: Force law from Maxwell's equations

Each supplement will consist of one or more problems, usually more challenging and of an open nature, to stretch the better students.

It has been repeatedly stressed that the force law and Maxwell's equations form an action-reaction pair. In some sense it is possible to derive one from the other. In this Supplement, we sketch how, knowing the energy-momentum tensor of the fields, it is possible to derive the force on various sources.

Point charge

Suppose that in the region $r \geq R$ for some small R, we observe a field

$$\mathbf{E} \quad = \quad \frac{q}{4\pi\epsilon_0 r^2}\,\hat{\mathbf{e}}_r + E_0\,\hat{\mathbf{e}}_z \qquad (23)$$

We would normally interpret this as a point charge q at the origin, placed in an external field represented by the second term. But pretend we never probe inside the sphere of radius R and do not know what is inside.

Now, using the energy momentum tensor, evaluate the net momentum flowing into the sphere of radius R per unit time; this would be interpreted as the force on the point charge — we would have derived this force $qE_0\,\mathbf{e}_z$ acting *on* the charge q, starting from the $1/r^2$ field produced *by* the point charge; the latter is equivalent to Gauss' law.

A few remarks to guide the evaluation.

- The energy-momentum tensor is quadratic in the field, so there will be three types of terms, respectively going as q^2, qE_0 and E_0^2. It should be obvious that only the cross term matters.
- In the cross term, by symmetry only the z component will survive when integrated over the surface of the sphere.
- The cross term will have a magnitude $\propto R^{-2}$ on the surface, and the surface integral will contribute a factor R^2, so the answer is independent of R.
- The energy-momentum expression carries one factor of ϵ_0, which cancels the factor $1/\epsilon_0$ in the first term of (23), so the final answer in independent of ϵ_0.

Electric dipole
Generalizing the same idea, Suppose that in the region $r \geq R$ for some small R, we observe a field

$$\begin{aligned}\mathbf{E} &= \frac{1}{4\pi\epsilon_0}\left(-\boldsymbol{\nabla}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}\right) \\ &\qquad + C_i\,\hat{\mathbf{e}}_i + D_{ij}x_i\,\hat{\mathbf{e}}_j \qquad (24)\end{aligned}$$

The first term represents a dipole inside the sphere, $C_i = E_i(0)$, i.e., the components of the external field at the origin, and $D_{ij} = \partial_i E_j(0)$, i.e., the field gradients of the external field at the origin.

Carry out a similar analysis, at least for some configurations, and evaluate the force on a dipole. Apart from verifying consistency, this exercise shows that anything inside an unseen region, provided it *produces* fields that *look like* a dipole, must necessarily *respond to* an external field also exactly like a dipole.

Long wire
Consider a long wire along the z axis, carrying a current I, together with an external magnetic field $B_0\,\hat{\mathbf{e}}_x$. Write out the total field in a region $r \geq R$; the part due to I will circulate around the wire — this result follows from Ampere's law. Then go through the same type of calculation, and find the force on a unit length of the wire. This would give the magnetic part of the Lorentz force law.

Magnetic dipole
Carry out a similar analysis where the field outside the sphere of radius R consists of a piece that looks like a magnetic dipole $\boldsymbol{\mu}$ at the origin and a spatially varying external magnetic field. Again evaluate the force involved.

This derivation is important in that we do not need to ask whether it is a current loop or an intrinsic magnetic moment. So long as it generates that kind of field, it must respond in that particular way to an external field gradient. In other words, we can freely port the equations for a loop-dipole to apply to intrinsic dipoles.

B Supplement: Magnetars

This Supplement sketches a topic of current interest, in which the energy density due to magnetic fields is a dominant feature.

The term *magnetar* [1] refers to a kind of neutron star with very strong magnetic fields. Typical parameters are

$$\begin{aligned}\text{Mass } M &= 1.4 \text{ solar masses} \\ \text{Radius } R &= 10 \text{ km} \\ \text{Rotation period } T &= 5 \text{ s}\end{aligned}$$

In particular, the rotation period is much longer than most "normal" neutron stars (typically 0.5 s).

Such a star would have a rotational energy of about

$$\begin{aligned}E_{\text{rot}} &\sim \frac{1}{2}I\omega^2 \sim \frac{1}{2}\left(\frac{2}{5}MR^2\right)\left(\frac{2\pi}{T}\right)^2 \\ &= \frac{4\pi^2}{5}\,\frac{MR^2}{T^2}\end{aligned}$$

On the other hand, if there is a typical magnetic field B in the star, then there is a magnetic energy

$$E_{\text{mag}} \sim \frac{1}{2\mu_0}B^2\,V \sim \frac{2\pi}{3\mu_0}B^2R^3$$

where we have put in $V \sim (4\pi/3)R^3$. This is an underestimate, since the magnetic field must extend outside the star.

For sufficiently strong B, the magnetic energy would exceed the rotational energy, putting these stars in a qualitatively different category. Give an algebraic expression for the critical value of B that divides the domain of magnetars from "normal" neutron stars, and a numerical estimate based on the given parameters.

C Supplement: Alfven waves

This Supplement sketches the main elements of Alfven waves, in which the EM stress tensor, in particular the tension associated with magnetic field lines, plays an important part.

A bundle of flux lines
Imagine a bundle of magnetic flux lines say along the x direction, with cross section area A. It is easy to see that

$$T_{xx} = \frac{1}{\mu_0}\left(\frac{1}{2}B^2 - B^2\right) = -\frac{1}{2\mu_0}B^2 \quad (25)$$

The minus sign implies a tension (rather than a pressure). Thus this bundle of field lines can be thought of as a string, with tension $\tau = |T_{xx}|\, A$.

Now suppose there are charged particles with total mass density ρ_m circulating these flux lines. For reasons we do not go into here, the flux lines and the fluid of charged particles (plasma) are pinned to each other. This string has a mass per unit length $\sigma = \rho_\mathrm{m}A$. Thus, based on elementary notions of the transverse vibrations of a string, there should be waves with speed v, where

$$v^2 = \frac{\tau}{\sigma} = \frac{B^2}{2\mu_0\rho_\mathrm{m}} \quad (26)$$

This is the idea of *Alfven waves.*

A factor of two
Actually the above formula for v^2 is wrong by a factor of two. This error is implicit in any statement that ascribes the dynamics only to tension; for example, the Wikipedia entry on Alfven waves [2] says

> In plasma physics, an Alfven wave, named after Hannes Alfven, is a type of magnetohydrodynamic wave in which ions oscillate in response to a restoring force provided by an effective tension on the magnetic field lines.

To understand this, imagine a segment of this bundle now slanted by a small angle θ towards the y direction, and consider the vertical (i.e., y) component of the force due to the end of this segment, a plane surface of area A, acting on the adjoining segment. Because the bundle is tilted with respect to the x-axis, there is a y component due to the tension force, of magnitude

$$\begin{aligned} F_{1y} &= \tau \sin\theta \approx |T_{xx}|A\,\theta \\ &\approx \frac{1}{2\mu_0}B^2 A\,\theta \end{aligned}$$

This is the usual effect encountered in elementary discussions of the transverse vibrations of a string.

However, a bundle of field lines is not exactly a string; there is an additional effect — a *shear force* acting on the surface A, also along the y direction, given by

$$\begin{aligned} F_{2y} &= T_{yx}A = \frac{1}{2\mu_0}B_yB_xA \\ &\approx \frac{1}{2\mu_0}(B\sin\theta)(B\cos\theta)A \\ &\approx \frac{1}{2\mu_0}B^2A\,\theta \end{aligned}$$

to first order in θ; the angles appear in the second line because $\mathbf{B}$ is assumed to be in a direction making an angle θ with the x-axis.

Now it is obvious that incorporation of F_{2y} doubles the total transverse force. Following through with the usual arguments, we find, instead of (26),

$$\boxed{v^2 = \frac{\tau}{\sigma} = \frac{B^2}{\mu_0\rho_\mathrm{m}}} \quad (27)$$

A paradox?
The result (27) seems to contain a paradox: What if the magnetic field is so strong, and/or the density of the plasma so small that the RHS turns out to be larger than c^2?

There are two contributions to the inertia of the bundle. First, there is the inertia due to matter comoving with the bundle, described by mass density ρ_m. But in addition, the fields themselves contain energy, or in effect mass, with a density

$$\mathcal{U}/c^2 \sim \frac{1}{2\mu_0c^2}B^2 \quad (28)$$

The derivation above (in implicitly using $F = ma$ for the transverse motion of a segment of the bundle) has neglected (28) in comparison to ρ_m, which is valid only in circumstances where

$$\frac{1}{2\mu_0c^2}B^2 \ll \rho_\mathrm{m}$$

In these circumstances, the RHS of (27) is guaranteed to be much smaller than c^2. Consult for example Jackson [3] for the general case where both contributions to the inertia are included.

An example
In the solar corona, $\rho \sim 10^{-4}$ kg m^{-3}, $B \sim 10^{-4}$ T, leading to rather low waves velocities of a few m s^{-1}. Such Alfven waves are thought to be important in heating the corona to temperatures much higher than that on the solar surface.

D Supplement: Pinning of magnetic flux lines

D.1 The phenomenon

It turns out that, under certain circumstances (to be detailed below), the magnetic flux lines in a plasma (i.e., a fluid of ionized particles) are *pinned* to the fluid. In other words, the fluid and the flux lines move together. This phenomenon is important in many situations, for example in magnetic confinement of a hot plasma for fusion, and in many astrophysical situations where the magnetic field plays a role. We have already encountered two examples.

- When a star collapses to a neutron star, any flux lines penetrating the star would be carried along in the collapse. The same flux gets concentrated in a much smaller cross-section area, leading to a very large magnetic field. This is one (but not the only) reason for the large field in magnetars.

- In the discussion of Alfven waves, it was assumed that as the flux lines vibrate, the charged particles are carried along, which is the reason why the mass density of the ions appear in determining the wave velocity.

First, let us understand a trivial special case: static field lines. In this case, the charged particles would (a) execute circular motion around the field lines, and (b) move freely along the field lines. In that sense, each particle is always confined to the neighborhood of a flux line — a trivial case of pinning. The purpose here is to generalize this idea to the case where the flux lines move.

In a sense, this topic does not have a direct relationship with the concept of energy and momentum. However, pinning is related to both magnetars and Alfven waves, so the opportunity is taken to give a simple account.

D.2 Basic equations in magnetohydrodynamics

Generalized Ohm's law

Normal Ohm's law is, in obvious notation

$$\mathbf{J} \;=\; \sigma\,\mathbf{E} \tag{29}$$

where σ is the conductivity. But now, if the medium is moving with velocity $\mathbf{v}$, the charges are carried along with this velocity, and also experience a magnetic force. Thus the generalized Ohm's law should read

$$\mathbf{J} \;=\; \sigma\,(\mathbf{E} + \mathbf{v}\times\mathbf{B}) \tag{30}$$

Perfect conductor

We are taught in freshman physics that (on reasonable long time scales, i.e., when equilibrium is established) the electric field inside a conductor is zero. This is obtained from (29) by taking $\sigma \to \infty$. Now the corresponding statement for a charged fluid in motion becomes the bracket on the RHS of (30) being zero, or

$$\mathbf{E} \;=\; -\mathbf{v}\times\mathbf{B} \tag{31}$$

Equations of motion

To solve the system, we have to specify how the fluid velocity $\mathbf{v}$ and the fields $\mathbf{E}$ and $\mathbf{B}$ change with time. (The evolution of the fluid density is understood to be specified by the continuity equation, which we omit.) But since $\mathbf{E}$ is determined by (31), we only need to specify

$$\begin{aligned}\frac{\partial\mathbf{v}}{\partial t} &= ???\\ \frac{\partial\mathbf{B}}{\partial t} &= ???\end{aligned}$$

In other words, the system is described by the magnetic field together with the fluid motion — the subject is therefore called *magnetohydrodynamics* or MHD.

The first is just Newton's Second Law, which we skip. The second is simply given by Faraday's Law

$$\begin{aligned}\frac{\partial\mathbf{B}}{\partial t} &= -\boldsymbol{\nabla}\times\mathbf{E}\\ &= \boldsymbol{\nabla}\times(\mathbf{v}\times\mathbf{B})\end{aligned} \tag{32}$$

where in the second line we have made use of (31).

This key equation in MHD tells us how $\mathbf{B}$ evolves if we know how the fluid moves. It is claimed that this equation implies the pinning of flux lines.

D.3 Derivation of pinning

Equation for transverse motion

First, note that in (32) we can replace $\mathbf{v} \mapsto \mathbf{u}$, where the latter is the part of $\mathbf{v}$ transverse to the field lines. Now taking the i component of (32), we get

$$\begin{aligned}&\frac{\partial B_i}{\partial t}\\ &= \epsilon_{ijk}\partial_j\left(\epsilon_{kmn}u_m B_n\right)\\ &= \left(\epsilon_{ijk}\epsilon_{kmn}\right)\partial_j(u_m B_n)\\ &= \left(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}\right)\,\partial_j(u_m B_n)\\ &= \partial_j(u_i B_j) - \partial_j(u_j B_i)\end{aligned} \tag{33}$$

Now take the direction i to be along the field lines; thus $B_i = B$, where the latter is the magnitude of the field. On the RHS, the first term is zero, because $\mathbf{u}$ has no component along this direction, i.e., $u_i = 0$.

Therefore we get

$$\frac{\partial B}{\partial t} + \boldsymbol{\nabla} \cdot (B\,\mathbf{u}) \quad = \quad 0 \qquad (34)$$

This key equation describes how the magnitude of $\mathbf{B}$ changes as the flux lines execute transverse motion. It is convenient to regard this as an equation in 2D space, with B being a scalar (i.e., invariant with respects to rotations in the 2D space).

Interpretation

Recall the continuity equation for a fluid (here written for two dimensions)

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho\,\mathbf{u}) \quad = \quad 0$$

where ρ is say the density of total particles. This equation expresses the fact that the particle number $N = \int \rho\, d^2r$ over a given area remains unchanged as the area follows the fluid. In just the same way, $\varphi_B = \int B\, d^2r$, which is the magnetic flux, remains unchanged as the area follows the fluid — the flux lines are pinned to the fluid.

References

[1] See for example the Shaw Prize in Astronomy, 2021.
`https://www.shawprize.org/prizes-and-laureates/astronomy/2021/press-release`

[2] Wikipedia, “Alfven waves”.

[3] JD Jackson, *Classical Electrodynamics*, 3rd Edition, Wiley (1998).
ISBN: 978-0-471-30932-1.
An online edition (not sure whether it breaches copyright) is available at
`www.fisica.unlp.edu.ar/materias/electromagnetismo-licenciatura-en-fisica-medica/electromagnetismo-material-adicional/`
or from
`https://www.scribd.com/doc/48520397/Jackson-Classical-Electrodynamics-3rd-edition`
The first edition (1962) can be downloaded from
`https://archive.org/details/ClassicalElectrodynamics`

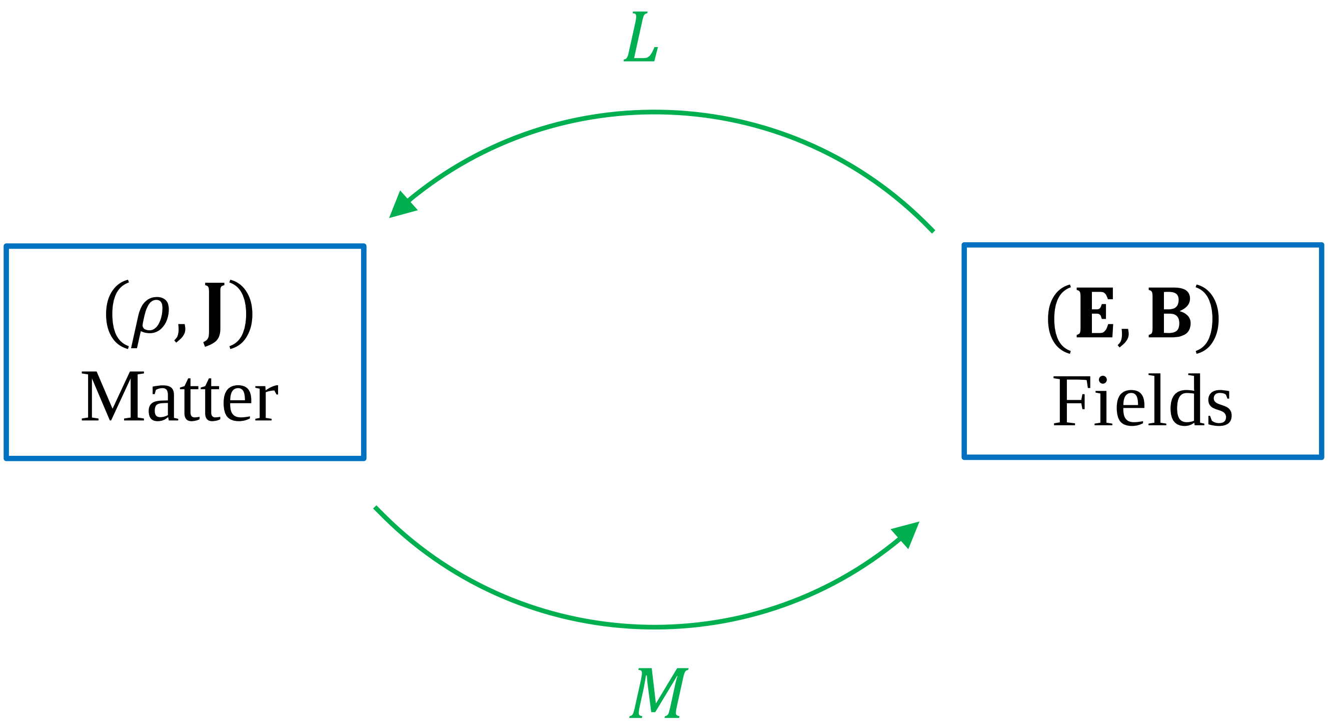


Figure 1 Matter $(\rho, \mathbf{J})$ produces fields $(\mathbf{E}, \mathbf{B})$ according to Maxwell's equation (M). The fields $(\mathbf{E}, \mathbf{B})$ act on matter $(\rho, \mathbf{J})$ according to Lorentz force law (L).

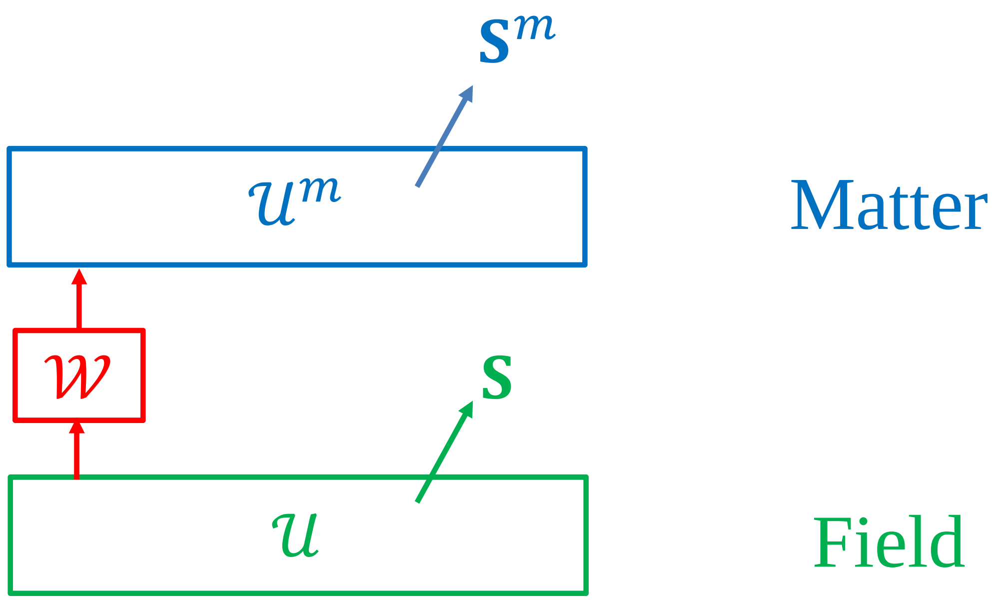


Figure 2 Schematic showing the matter and the fields in a small volume. The fields do work on the matter.

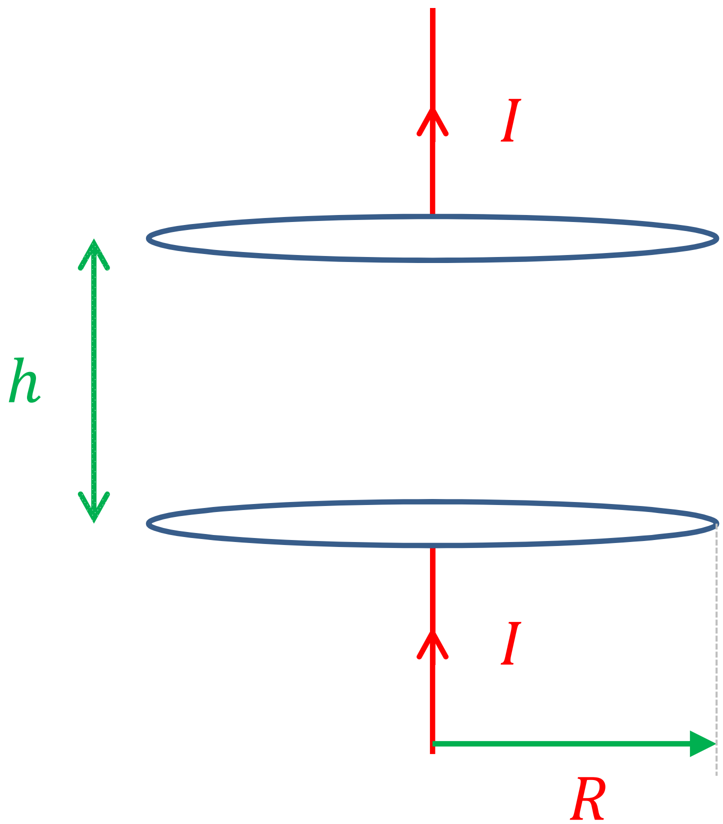
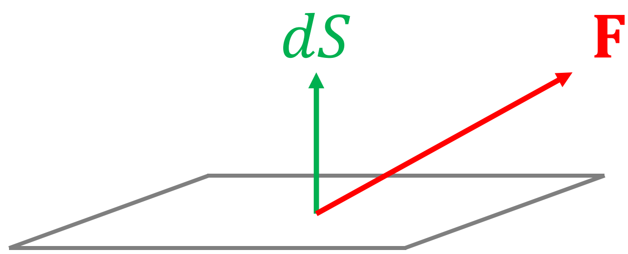


Figure 3 A capacitor is being charged. How much energy flows into the volume of the capacitor?

(a)



(b)

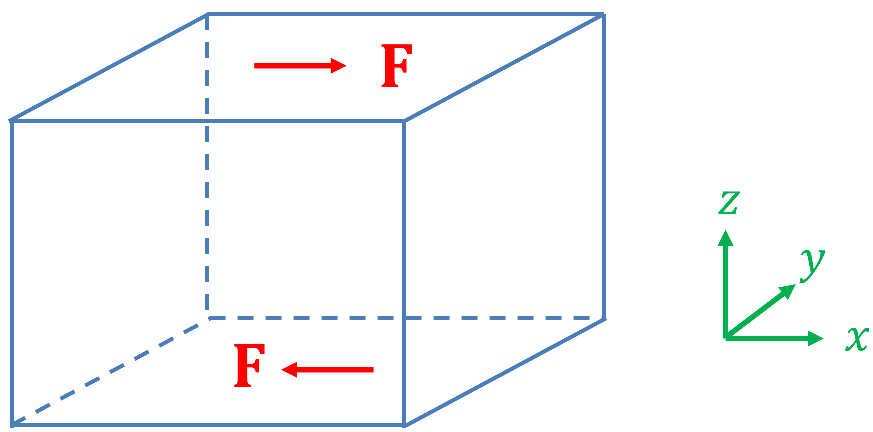


Figure 4 (a) A force **F** acting on a surface dS
(b) A pair of forces causing a torque τ_1

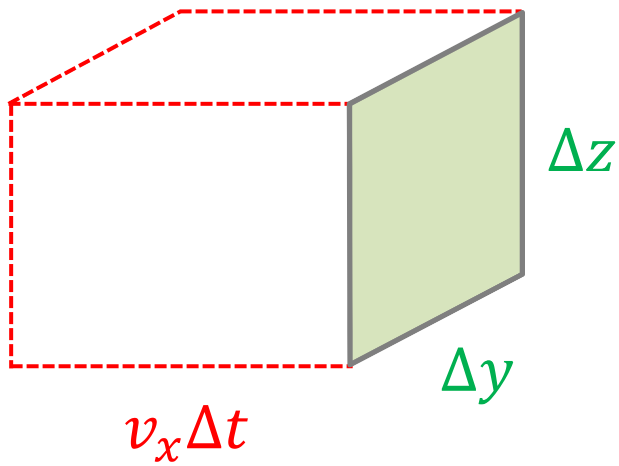


Figure 5 Momentum carried across the shaded surface per unit time