Generalizing the Lorentz force to a region containing charge

Hendrik Lorentz (Dutch, 1853-1928) was one of the physicists who contributed to the great advance in electromagnetics in the late 19th century. It was Lorentz who first formulated the expression which bears his name:

$$\vec{F}_{Lorentz} = q\vec{E} + q\vec{v} \times \vec{B} \quad (1)$$

A point particle with charge q which is located in an electric field $\vec{E}$ and magnetic field $\vec{B}$ experiences a force which is the sum of: (i) an "electric" component $q\vec{E}$ and (ii) a "magnetic" component $q\vec{v} \times \vec{B}$, where:

- the electric component of the force acts in the direction of the electric field $\vec{E}$ at the point's location;
- the magnetic component of the force is perpendicular to both the magnetic field $\vec{B}$ at the point's location and to the point's velocity $\vec{v}$;
- $\vec{v} \times \vec{B}$ is a vector cross-product, whose magnitude is equal to the product of v, B and the sine of the angle between $\vec{v}$ and $\vec{B}$ and whose direction is determined by the right-hand rule – thumb parallel to $\vec{v}$, extended fingers parallel to $\vec{B}$ and palm in the direction of the cross-product and
- the Lorentz force is valid only if the charged particle is moving at a speed well less than the speed of light. If its speed is an appreciable fraction of the speed of light, then certain delays need to be introduced to account for the fact that events at one location are not felt instantaneously elsewhere.

Lorentz's equation is ideal for use in certain circumstances, where the particle is confined to move within a wire, for example, or where the motion of only a single particle needs to be determined. Lorentz's equation is clumsy in other circumstances, where a number of charged particles are free to move within a three-dimensional region. The situation quickly gets out of hand when it is necessary to also account for the electric fields and magnetic fields generated by the charged particles themselves.

For a lump of material, in which the charges are free to move about in three dimensions, Lorentz's force is handier if written in the form:

$$\vec{f} = \rho\vec{E} + \vec{J} \times \vec{B} \quad (2)$$

Here, $\vec{f}$ is the force acting per unit volume at the point of interest, ρ is the number of charges per unit volume at the point of interest and $\vec{J}$ is the current density at the point of interest. The current density $\vec{J}$ is the number of charges moving through a unit area each second. The direction of the current density is the direction normal to the unit area through which the charges are moving.

$\vec{f}$, ρ and $\vec{J}$ are macroscopic quantities. When we consider these quantities at a particular point, we must imagine a small element of volume. The element of volume needs to be small enough that these quantities, as well as the electric field and magnetic field, are essentially constant throughout the volume. On the other hand, the element of volume must be large enough to hold enough of the charged particles for statistical averaging to take place. In reality, much jostling will occur in the element of volume. The charged particles will bump into each other, into uncharged particles which might happen to be in the way, into ions in the material if it is a metal, or into some boundary. At any given instant, a particular charged particle might be travelling in response to forces quite different from Equation (2). On average, though, or for the "average" particle, Equation (2) explains the force on the collective.

In addition, the force $\vec{f}$ in Equation (2) should be considered to be a time-averaged quantity, relevant on a time-scale much longer than atomic movements

Equation (2) is said to be a "differential" form of Lorentz's equation. This is because it applies within small elements of volume, which need to be added up, or integrated, to find the net average force on a macroscopic region.

In order to work with Equation (2), we need to refer to the four differential equations which are now called Maxwell's Equations:

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad (3)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \qquad (4)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad (5)$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} \qquad (6)$$

(3) is Gauss's Law, which states that the divergence of an electric field outwards from any small volume of space is proportional to the number of charges it contains.

(4) is Gauss's Law for magnetism, which states that a magnetic field cannot diverge (because there are no magnetic monopoles).

(5) is Faraday's Law of induction, which can be used, among other things, to integrate a changing magnetic field with respect to time in order to determine the voltage induced in a loop of wire.

(6) is Ampere's Circuital Law, which can be used, among other things, to determine the magnetic field set up around one or more wires carrying current. Equation (6) includes the correction added by Maxwell for the effect of an electric field which varies with time.

Units are defined in the SI system so that the permittivity ϵ_0 and permeability μ_0 of free space, respectively, have a product $\epsilon_0\mu_0$ equal to the reciprocal of the square of the speed of light. The permeability μ_0 of free space is equal to $4\pi \times 10^{-7}$ H/m [Henries per meter]. Since the Henry is a derived unit, the unit of permeability can also be expressed as N/A^2 [Newtons per Ampere squared] and in other forms as well.

Now, our objective in this paper is to express the Lorentz force in terms of the electric and magnetic fields only. This requires that we use Maxwell's equations to "solve" Equation (2) in such a way as to remove the explicit dependence on the charged particles' position (their charge density ρ) and speed (their current density $\vec{J}$).

Step 1: Substitute Gauss's Law and Ampere's Circuital Law to remove the charge density and the current density, respectively:

$$\begin{aligned}\vec{f} &= \epsilon_0(\vec{\nabla}\cdot\vec{E})\vec{E} + \frac{1}{\mu_0}\left(\vec{\nabla}\times\vec{B} - \epsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}\right)\times\vec{B} \\ &= \epsilon_0(\vec{\nabla}\cdot\vec{E})\vec{E} + \frac{1}{\mu_0}(\vec{\nabla}\times\vec{B})\times\vec{B} - \epsilon_0\frac{\partial\vec{E}}{\partial t}\times\vec{B} \quad (7)\end{aligned}$$

Note that this substitution, which removed all references to the charge density and current density, has also removed any reference to the sign of the distributed charges.

Step 2: Apply the product rule to the last term in Equation (7). The product rule is:

$$\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) = \left(\frac{\partial\vec{E}}{\partial t}\times\vec{B}\right) + \left(\vec{E}\times\frac{\partial\vec{B}}{\partial t}\right)$$

so Equation (7) can be written as:

$$\vec{f} = \epsilon_0(\vec{\nabla}\cdot\vec{E})\vec{E} + \frac{1}{\mu_0}(\vec{\nabla}\times\vec{B})\times\vec{B} - \epsilon_0\left\{\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) - \left(\vec{E}\times\frac{\partial\vec{B}}{\partial t}\right)\right\} \quad (8)$$

Step 3: Substitute Faraday's Law of induction to get:

$$\begin{aligned}\vec{f} &= \epsilon_0(\vec{\nabla}\cdot\vec{E})\vec{E} + \frac{1}{\mu_0}(\vec{\nabla}\times\vec{B})\times\vec{B} - \epsilon_0\left\{\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) + [\vec{E}\times(\vec{\nabla}\times\vec{E})]\right\} \\ &= \epsilon_0(\vec{\nabla}\cdot\vec{E})\vec{E} + \frac{1}{\mu_0}(\vec{\nabla}\times\vec{B})\times\vec{B} - \epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) - \epsilon_0[\vec{E}\times(\vec{\nabla}\times\vec{E})] \quad (9)\end{aligned}$$

Step 4: Use the identity, that $\vec{a}\times\vec{b} = -\vec{b}\times\vec{a}$ for any two vectors, to reverse the second term:

$$\vec{f} = \epsilon_0(\vec{\nabla}\cdot\vec{E})\vec{E} - \frac{1}{\mu_0}\vec{B}\times(\vec{\nabla}\times\vec{B}) - \epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) - \epsilon_0[\vec{E}\times(\vec{\nabla}\times\vec{E})] \quad (10)$$

Step 5: Collect like terms to get:

$$\vec{f} = \epsilon_0\{(\vec{\nabla}\cdot\vec{E})\vec{E} - \vec{E}\times(\vec{\nabla}\times\vec{E})\} + \frac{1}{\mu_0}\{\qquad\qquad -\vec{B}\times(\vec{\nabla}\times\vec{B})\} - \epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) \quad (11)$$

Step 6: The gap in Equation (11) highlights a certain absence of symmetry. The two terms in curly brackets would be identical in $\vec{E}$ and $\vec{B}$ if there happened to be a term $(\vec{\nabla}\cdot\vec{B})\vec{B}$, but there is not. However, Gauss's Law for magnetism tells us that the missing term is identically zero. We can add the term to Equation (11) to get:

$$\vec{f} = \epsilon_0\{(\vec{\nabla}\cdot\vec{E})\vec{E} - \vec{E}\times(\vec{\nabla}\times\vec{E})\} + \frac{1}{\mu_0}\{(\vec{\nabla}\cdot\vec{B})\vec{B} - \vec{B}\times(\vec{\nabla}\times\vec{B})\} - \epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) \quad (12)$$

Step 7: Curls are difficult to deal with. Fortunately, they can be expressed in alternative form. In Lemma #1 attached, the following identity is proved, for any vector $\vec{S}$:

$$\vec{S} \times (\vec{\nabla} \times \vec{S}) = \frac{1}{2}\vec{\nabla}S^2 - (\vec{S} \cdot \vec{\nabla})\vec{S} \quad (13)$$

Step 8: Using this identity in Equation (12) gives:

$$\vec{f} = \epsilon_0\left\{(\vec{\nabla} \cdot \vec{E})\vec{E} - \left[\frac{1}{2}\vec{\nabla}E^2 - (\vec{E} \cdot \vec{\nabla})\vec{E}\right]\right\} + \frac{1}{\mu_0}\left\{(\vec{\nabla} \cdot \vec{B})\vec{B} - \left[\frac{1}{2}\vec{\nabla}B^2 - (\vec{B} \cdot \vec{\nabla})\vec{B}\right]\right\} - \epsilon_0\frac{\partial}{\partial t}(\vec{E} \times \vec{B}) \quad (14)$$

Step 9: Re-arranging terms gives:

The force per unit volume on the distributed charges in a region is

$$\vec{f} = \pm\left\{\epsilon_0[(\vec{\nabla} \cdot \vec{E})\vec{E} + (\vec{E} \cdot \vec{\nabla})\vec{E}] + \frac{1}{\mu_0}[(\vec{\nabla} \cdot \vec{B})\vec{B} + (\vec{B} \cdot \vec{\nabla})\vec{B}] - \frac{1}{2}\vec{\nabla}\left[\epsilon_0 E^2 + \frac{1}{\mu_0}B^2\right] - \epsilon_0\frac{\partial}{\partial t}(\vec{E} \times \vec{B})\right\}$$

This is a very general result. It accounts for all aspects of sub-warp electromagnetism. It can handle time-varying and non-uniform electric and magnetic fields. It is independent of co-ordinate system. It does not include any of those nasty curls.

Jim Hawley
September 2011

An e-mail pointing out errors and omissions would be appreciated.

Lemma #1 – Prove that

$$\frac{1}{2}\vec{\nabla}S^2 = \vec{S}\times(\vec{\nabla}\times\vec{S}) + (\vec{S}\cdot\vec{\nabla})\vec{S}$$

It is convenient to conduct the proof in a traditional Euclidian $\hat{x}$-$\hat{y}$-$\hat{z}$ co-ordinate frame. If the expression holds true in any inertial co-ordinate frame, then it will hold true in any other inertial co-ordinate frame. Let us consider any arbitrary vector $\vec{S}$, which can be written in terms of its components as:

$$\vec{S} = S_x\hat{x} + S_y\hat{y} + S_z\hat{z} \quad (L1)$$

The curl of $\vec{S}$ is defined as the following determinant:

$$\begin{aligned}\vec{\nabla}\times\vec{S} &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ S_x & S_y & S_z \end{vmatrix} \\ &= \left(\frac{\partial S_z}{\partial y} - \frac{\partial S_y}{\partial z}\right)\hat{x} + \left(\frac{\partial S_x}{\partial z} - \frac{\partial S_z}{\partial x}\right)\hat{y} + \left(\frac{\partial S_y}{\partial x} - \frac{\partial S_x}{\partial y}\right)\hat{z} \quad (L2)\end{aligned}$$

The cross-product on the right-hand side of the subject expression can be written as the following determinant, and Equation $(L2)$ substituted therein, to give:

$$\begin{aligned}\vec{S}\times(\vec{\nabla}\times\vec{S}) &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ S_x & S_y & S_z \\ \left(\dfrac{\partial S_z}{\partial y} - \dfrac{\partial S_y}{\partial z}\right) & \left(\dfrac{\partial S_x}{\partial z} - \dfrac{\partial S_z}{\partial x}\right) & \left(\dfrac{\partial S_y}{\partial x} - \dfrac{\partial S_x}{\partial y}\right) \end{vmatrix} \\ &= \left\{\begin{array}{c} \left[S_y\left(\dfrac{\partial S_y}{\partial x} - \dfrac{\partial S_x}{\partial y}\right) - S_z\left(\dfrac{\partial S_x}{\partial z} - \dfrac{\partial S_z}{\partial x}\right)\right]\hat{x} + \cdots \\ \left[S_z\left(\dfrac{\partial S_z}{\partial y} - \dfrac{\partial S_y}{\partial z}\right) - S_x\left(\dfrac{\partial S_y}{\partial x} - \dfrac{\partial S_x}{\partial y}\right)\right]\hat{y} + \cdots \\ \left[S_x\left(\dfrac{\partial S_x}{\partial z} - \dfrac{\partial S_z}{\partial x}\right) - S_y\left(\dfrac{\partial S_z}{\partial y} - \dfrac{\partial S_y}{\partial z}\right)\right]\hat{z} \end{array}\right\} \quad (L3)\end{aligned}$$

Now, the vector dot-product $\vec{S}\cdot\vec{\nabla}$ on the right-hand side of the subject expression can be expanded as:

$$\vec{S}\cdot\vec{\nabla} = S_x\frac{\partial}{\partial x} + S_y\frac{\partial}{\partial y} + S_z\frac{\partial}{\partial z} \quad (L4)$$

Applying this dot-product operator to vector $\vec{S}$ gives:

$$(\vec{S}\cdot\vec{\nabla})\vec{S} = \left\{\begin{array}{c} \left[S_x\dfrac{\partial S_x}{\partial x} + S_y\dfrac{\partial S_x}{\partial y} + S_z\dfrac{\partial S_x}{\partial z}\right]\hat{x} + \cdots \\ \left[S_x\dfrac{\partial S_y}{\partial x} + S_y\dfrac{\partial S_y}{\partial y} + S_z\dfrac{\partial S_y}{\partial z}\right]\hat{y} + \cdots \\ \left[S_x\dfrac{\partial S_z}{\partial x} + S_y\dfrac{\partial S_z}{\partial y} + S_z\dfrac{\partial S_z}{\partial z}\right]\hat{z} \end{array}\right\} \quad (L5)$$

Equations $(L3)$ and $(L5)$ are the two terms on the right-hand side of the subject expression. We can add them together to get:

$$RHS = \left\{ \begin{array}{c} \left[S_y \left(\frac{\partial S_y}{\partial x} - \frac{\partial S_x}{\partial y} \right) - S_z \left(\frac{\partial S_x}{\partial z} - \frac{\partial S_z}{\partial x} \right) + S_x \frac{\partial S_x}{\partial x} + S_y \frac{\partial S_x}{\partial y} + S_z \frac{\partial S_x}{\partial z} \right] \hat{x} + \cdots \\ \left[S_z \left(\frac{\partial S_z}{\partial y} - \frac{\partial S_y}{\partial z} \right) - S_x \left(\frac{\partial S_y}{\partial x} - \frac{\partial S_x}{\partial y} \right) + S_x \frac{\partial S_y}{\partial x} + S_y \frac{\partial S_y}{\partial y} + S_z \frac{\partial S_y}{\partial z} \right] \hat{y} + \cdots \\ \left[S_x \left(\frac{\partial S_x}{\partial z} - \frac{\partial S_z}{\partial x} \right) - S_y \left(\frac{\partial S_z}{\partial y} - \frac{\partial S_y}{\partial z} \right) + S_x \frac{\partial S_z}{\partial x} + S_y \frac{\partial S_z}{\partial y} + S_z \frac{\partial S_z}{\partial z} \right] \hat{z} \end{array} \right\} \quad (L6)$$

There are a number of terms which cancel each other out. Cancelling like terms gives:

$$RHS = \left\{ \begin{array}{c} \left[S_y \frac{\partial S_y}{\partial x} + S_z \frac{\partial S_z}{\partial x} + S_x \frac{\partial S_x}{\partial x} \right] \hat{x} + \cdots \\ \left[S_z \frac{\partial S_z}{\partial y} + S_x \frac{\partial S_x}{\partial y} + S_y \frac{\partial S_y}{\partial y} \right] \hat{y} + \cdots \\ \left[S_x \frac{\partial S_x}{\partial z} + S_y \frac{\partial S_y}{\partial z} + S_z \frac{\partial S_z}{\partial z} \right] \hat{z} \end{array} \right\} \quad (L7)$$

Each remaining term is the partial derivative of a square, which enables us to write:

$$RHS = \frac{1}{2} \left\{ \left[\frac{\partial S_y^2}{\partial x} + \frac{\partial S_z^2}{\partial x} + \frac{\partial S_x^2}{\partial x} \right] \hat{x} + \left[\frac{\partial S_z^2}{\partial y} + \frac{\partial S_x^2}{\partial y} + \frac{\partial S_y^2}{\partial y} \right] \hat{y} + \left[\frac{\partial S_x^2}{\partial z} + \frac{\partial S_y^2}{\partial z} + \frac{\partial S_z^2}{\partial z} \right] \hat{z} \right\} \quad (L8)$$

Collecting terms in S_x^2, S_y^2 and S_z^2 gives:

$$RHS = \frac{1}{2} \left\{ \left[\frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} \right] S_x^2 + \left[\frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} \right] S_y^2 + \left[\frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} \right] S_z^2 \right\} \quad (L9)$$

The common operator can be removed from the curly brackets to give:

$$RHS = \frac{1}{2} \left[\frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} \right] \left\{ S_x^2 + S_y^2 + S_z^2 \right\} \quad (L10)$$

The expression in the curly brackets is simply the square of the length of vector $\vec{S}$, so that:

$$RHS = \frac{1}{2} \left[\frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} \right] S^2 \quad (L11)$$

and the operator in the square brackets is simply the gradient, so that:

$$RHS = \frac{1}{2} \vec{\nabla} S^2 \quad (L12)$$

This is the left-hand side.

Q.E.D.