There Must Be a Magnetic Field

© William O. Straub
Pasadena, California
April 14, 1990

I was introduced to special relativity in third-semester undergraduate physics back in 1969, and I remember that I had absolutely no idea what the professor was talking about. Late in the term he tried to explain how a stream of equal but opposing currents could induce a force on a moving charged particle that looked just like an electrostatic force. The idea, of course, was to show that electric and magnetic fields are one and the same as a consequence of special relativity. Well, the professor tripped all over himself trying to explain it, and only succeeded in thoroughly confusing everyone in the class. Many years later, I tried to explain it to a group of second-year physics students myself. While I think I got farther than my old professor, I made an equal fool of myself trying to convince the students that I really knew what I was talking about.

Many beginning physics students have difficulty understanding the problem and its solution. And while many textbooks have treated the problem, none to my knowledge have succeeded in being very helpful. This a pity, because the problem neatly encapsulates the intimate relationship between relativity and electrodynamics.

Here I'll try to make the problem transparent. I will follow Griffiths' approach (which is probably the best available), but without the typos.

1. Electrostatics and Magnetostatics in Brief

Consider a wire of negligible diameter carrying a static, uniform charge of positive density λ coulombs/meter. Using $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho/\epsilon_0$ (where the total enclosed charge per unit of length l is $q = \lambda l$) with Gauss' Law, we easily calculate the electric field E at a distance R outside the wire to be

$$E = \frac{\lambda}{2\pi\epsilon_0 R}$$

It we now place a particle with positive charge Q near the wire, it will feel a force directed *away* from the wire given by

$$F = QE$$

Conversely, a magnetic field can be set up by making the charges flow through the wire. It can be calculated using Ampere's Law $\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J}$ which, for a uniform current I, gives

$$B = \frac{\mu_0 I}{2\pi R}$$

If we now give the charge Q a velocity u parallel to the current, it feels a sideways force directed *toward* the wire in accordance with the Lorentz force law

$$\begin{aligned} F &= QuB \qquad (1.1) \\ &= \frac{\mu_0 Q u I}{2\pi R} \end{aligned}$$

If we now imagine the charge Q to be moving with the same velocity and direction as the current, we might expect the charge to "see" a static electric field in the vicinity of the moving charges in the wire, which now appear stationary to Q. The charge would then feel the resulting electrostatic field (although in this case it would be repelled, not attracted). We might even guess that the magnetic field could be "transformed away" to some extent by this moving point of view. Special relativity, which deals with moving frames of reference, confirms this guess. In the following, we will examine the details of how and why this comes about.

2. Relativity in Brief

You should already be very familiar with the Lorentz transformation equations of special relativity, which relates one inertial system in relative motion to another. Imagine an observer in a system designated as S looking at events taking place in a system S' that is moving in the positive x-direction with constant relative velocity v. At the same time, the observer in S' is also busy looking at events in S. The Lorentz transformation describes what one sees in terms of another via

$$\mathbf{x}' = \gamma\left[\mathbf{x} - \boldsymbol{\beta} x^0\right] \tag{2.1}$$

$$x^{0\prime} = \gamma\left[x^0 - \boldsymbol{\beta}\cdot\mathbf{x}\right] \tag{2.2}$$

where c is the speed of light, $x^0 = ct$, $\boldsymbol{\beta} = \mathbf{v}/c$, and

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{2.3}$$

(When the motion is in the x-direction, we obviously have $y' = y$ and $z' = z$.) The apparent x, x^0 symmetry in (2.1) and (2.2) is sometimes mistakenly taken as proof that space and time are the same thing in relativity, but that is wrong. The equations for an arbitrary direction of the relative velocity parameter v, which are quite complicated, show no such symmetry. *Relativity requires four spacetime dimensions, but time is fundamentally different from space.*

For convenience and familiarity, we now consider the Lorentz transformation in terms of the relative velocity v (taken in the positive x-direction) and the space and time *differentials* dx and dt, as we will be dealing shortly with velocities.

If you are in system S, you can determine what is going on in S' using the "inverse" Lorentz transformation

$$dx = \gamma\left[dx' + v\,dt'\right] \tag{2.4}$$

$$dt = \gamma\left[dt' + \frac{v}{c^2}\,dx'\right] \tag{2.5}$$

The parameter γ is the same, but the primes are exchanged and the sign on v is changed.

Let's say there's a clock located at a specific point ($dx' = 0$) in the moving system S'. What does S see? Using

$$dt = \gamma\, dt'$$

from this we see that S measures a *longer* period of time than the elapsed time in S' because the quantity γ is always greater or equal to unity. This is the *time dilation* effect of special relativity. You get the same result if you are cruising along in the S' system checking on a fixed clock in S. *For inertial systems, time appears to run slow from the vantage point of the "other" observer.* Why do we fix the clock in the system being observed? Because otherwise the clock would be have to be moved in that system with some additional velocity, giving a multitude of possible results for the observer. This is not how one observes a moving clock.

Now imagine that there's a rod of length dx' in the moving system. The observer in S must measure this length at a fixed time $dt = 0$ in his/her own system, and so, from (2.4) and (2.5), we have

$$dx = \gamma\left[dx' + v\,dt'\right]$$

$$0 = \gamma\left[dt' + \frac{v}{c^2}\,dx'\right]$$

Eliminating dt' from these two expressions, the observer gets

$$dx = \frac{dx'}{\gamma}$$

That is, the rod appears to be shorter than its length in S'. This is the *length contraction* effect of special relativity. The same result would be obtained for an observer in S'. Why does the measurement have to be taken at a fixed time in the observer's system? Because otherwise the observed length would appear stretched out as the observer's clock ticked away, giving a multitude of possible measurements. This is not how we measure a physical distance.

It is possible to derive expressions in special relativity that relate observer *velocities* in the two inertial systems other than the relative velocity v. For example, from (2.4) and (2.5) we can calculate

$$\frac{dx'}{dt'} = u' = \frac{u - v}{1 - uv/c^2} \tag{2.6}$$

Similarly,

$$\frac{dx}{dt} = u = \frac{u' + v}{1 + u'v/c^2} \tag{2.7}$$

These are the famous *Einstein velocity-addition* formulas of special relativity. Basically, (2.6) says that, if an observer in S observes a velocity u in his/her own system, then S' will observe that velocity to be u'. This may seem straightforward, but it is essential that you understand what is really going on.

If the observer in S associates an object with the velocity u in his system (but stays fixed himself), then the relative velocity between S and S' remains v. However, if the observer in S decides to jump on board the object, then he jumps into a new system as far as S' is concerned. In this new system, $u = 0$ as far as S is concerned, *and the relative velocity between S and S' is no longer v.*

Note that, prior to being given any movement, an observer in S sees events in system S' moving to the right with relative velocity v. By the same token, S'sees events in S moving in the opposite direction with relative velocity $-v$. Now, if system S is suddenly moved to the right with velocity u, an observer in S' will see that the velocity of S will have changed from $-v$ to something proportional to $-v + u$, or $u - v$. Similarly, if an event in $S'$is given some additional velocity u', the observer in S will perceive it to be something proportional to $u' + v$. The relativistic effect comes from the fact that the proportionality constants, which are the denominators in (2.6) and (2.7), are not unity (as they are in classical, or Galilean, relativity).

Thus, the most paradoxical aspect of this velocity addition business (and one that is often overlooked or misinterpreted) is this: As we have shown, if system S is given a velocity u, system S' will observe it to have the velocity

$$u' = \frac{u - v}{1 - uv/c^2}$$

But now, according to S, the relative velocity of the two systems changes from v to $-u'$. To see this, note that $u' = -v$ when $u = 0$ and that $u = v$ when $u' = 0$. If S is now given a velocity u, it will seem to an observer in that system to be "catching up" with S'; that is, the relative velocity between the systems will have been *reduced*. The new relative velocity according to an observer in S is given by

$$-u' = \frac{v - u}{1 - uv/c^2} \tag{2.8}$$

As a consequence, the expression for γ becomes

$$\gamma = \frac{1}{\sqrt{1 - (u'^2)/c^2}} \tag{2.9}$$

To see a classic textbook example of the velocity-addition formula, consider a particle in S' that is given a velocity $u' = 0.7c$ when $v = 0.5c$. Classical physics says that the combined velocity according to S is just $(0.7 + 0.5)c = 1.2c$. But of course this is impossible, because nothing can travel faster than light. The relativistic formula (2.7), on the other hand, gives

$$u = \frac{0.7 + 0.5}{1 + (0.7)(0.5)}\, c = 0.89c$$

which is the correct answer. It is easy to verify that if $u' = c$, then $u = c$ as well. This proves one of the central tenets of relativity: *the speed of light is the same for all observers, regardless of their relative motion.*

3. The Charge Four-Vector of Electrodynamics

The quantities (dx^0, dx, dy, dz) constitute what is known as a *four-vector* in relativistic parlance. This four-vector is designated by dx^μ, where the index μ takes on the values $0, 1, 2, 3$. The equation

$$dx^{\mu\prime} = \Lambda^\mu{}_\nu \, dx^\nu$$

is shorthand notation for the Lorentz transformation, where the matrix of coefficients $\Lambda^\mu{}_\nu$ involves the quantities γ and β. In the warm-up given previously, we took $dy' = dy$, $dz' = dz$ because all the motion took place in the x-direction (and of course in the time "direction"!)

Four-vectors occur everywhere in physics, and in the flat spacetime of special relativity they all transform via the Lorentz transformation. In addition to the spacetime four-vector, we have the four-momentum

$$p^\mu = mc \, \frac{dx^\mu}{ds} = (E/c, p_x, p_y, p_z) \tag{3.1}$$

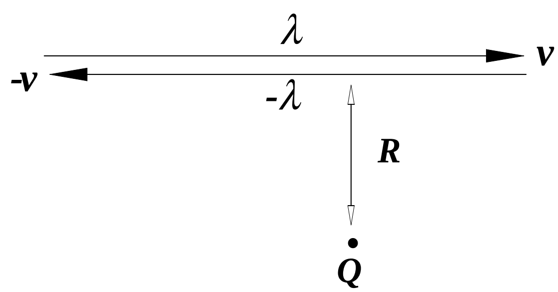
where ds is the invariant line element and E is energy. Similarly, the four-vector of electric charge density is

$$j^\mu = c\rho \, \frac{dx^\mu}{ds} = (c\rho, v_x\rho, v_y\rho, v_z\rho) \tag{3.2}$$

where ρ is the rest density of electric charge (sometimes it's written as ρ_0 to indicate that it's a proper density) and v_x, v_y, v_z are the velocities that the charge density can have in the indicated directions. (Important note: These velocities have nothing to do with the velocity v that denotes relative motion!)

4. There Must Be a Magnetic Field

Let's place a particle of charge $+Q$ sitting motionless a distance R from a line of continuous charge of positive density $\lambda^+ = \lambda$ moving at velocity v to the right. Coincident with this flow is a line of density $\lambda^- = -\lambda$ moving to the left at velocity $-v$ (see figure below). What does the charge Q see?



To start, let's assume that the charge Q is in an unmoving system that we label as S, and that the line of positive charges is in the moving system S'. For an observer moving with S' there is no current ($v'_x = 0$), only positive and negative charge densities that we will call $\lambda^{\pm\prime}$. Special relativity then says that, from (3.2),

$$\begin{aligned} c\lambda^+ &= \gamma \left[c\lambda' + \frac{v}{c} \, v'_x \lambda' \right] \\ j^+ &= \gamma \left[v'_x \lambda' + \frac{v}{c} \, c\lambda' \right] \end{aligned}$$

Similarly, for the line of negative charges ($v \to -v$), S sees

$$\begin{aligned} c\lambda^- &= \gamma \left[-c\lambda' + \frac{(-v)}{c} \, v'_x(-\lambda') \right] \\ j^- &= \gamma \left[-v'_x \lambda' + \frac{(-v)}{c} \, c(-\lambda') \right] \end{aligned}$$

But $v'_x = 0$, because the observer in S' is moving with the charge. We therefore have

$$\lambda_{\text{total}} = \lambda' - \lambda' = 0$$

so the total electrostatic charge that Q sees is zero. On the other hand, S sees a positive total current equal to

$$\begin{aligned} j_{\text{total}} &= j^+ + j^- = \gamma\left[v\lambda' + v\lambda'\right] \\ &= 2v\lambda \end{aligned}$$

where we have written $\lambda = \gamma\lambda'$, consistent with the fact that a moving charge density appears increased by a factor of γ due to length contraction. However, the charge Q feels no force from this current.

We now give the charge Q a velocity (according to system S) equal to u. Does this change anything?

In accordance with the velocity-addition theorem, the velocity of the charge according to the positive line of charge in S' is

$$u' = \frac{u - v}{1 - uv/c^2}$$

But, as explained previously, $-u'$ is the apparent relative velocity as seen by Q; that is, Q sees that it is "catching up" to the line of charge moving in the positive x-direction. In accordance with (2.8), we will call the negative of this quantity $v^+ = -u'$, which is the new relative velocity. By the same logic, Q sees itself "moving away" from the negative line of charge with a velocity v^- as given by

$$u' = v^- = -\frac{u + v}{1 + uv/c^2}$$

We now repeat the above procedure using

$$\begin{aligned} c\lambda &= \gamma^+\left[c\lambda' + \frac{v^+}{c}\, v'_x\lambda'\right] \\ j^+ &= \gamma^+\left[v'_x\lambda' + \frac{v^+}{c}\, c\lambda'\right] \end{aligned}$$

and

$$\begin{aligned} c\lambda^- &= \gamma^-\left[-c\lambda' + \frac{-(v^-)}{c}\, v'_x\lambda'\right] \\ j^- &= \gamma^-\left[-v'_x\lambda' + \frac{(-v^-)}{c}\, c\lambda'\right] \end{aligned}$$

where

$$\gamma^{\pm} = \frac{1}{\sqrt{1 - v^{\pm\,2}/c^c}}$$

in accordance with (2.9). Again, $v'_x = 0$ in the S' system and so, after some simple algebra, we get

$$\begin{aligned} \lambda_{\text{total}} &= \gamma^+\lambda' - \gamma^-\lambda' \\ &= -\frac{2uv\lambda'}{c^2\sqrt{1 - v^2/c^2}\sqrt{1 - u^2/c^2}} \\ &= -\frac{2uv\lambda}{c^2\sqrt{1 - u^2/c^2}} \end{aligned}$$

Thus, when Q is given a velocity it feels a *net negative electrostatic charge* from the line of opposing moving currents! Being positively charged itself, Q now feels an *attractive* force F_y toward S'. To determine this force, we again use Gauss' Law, which gives us

$$E = \frac{2uv\lambda}{2\epsilon_0 c^2\pi RL\sqrt{1 - u^2/c^2}} = \frac{uv\lambda}{\epsilon_0 c^2\pi R\sqrt{1 - u^2/c^2}}$$

where E is the electric field in system S at distance R from the line of currents (I've dropped the minus sign). Using $F = QE$, we then have

$$F_y = \frac{uv\lambda Q}{\epsilon_0 c^2\pi R\sqrt{1 - u^2/c^2}}$$

directed in the positive y direction.

Note that we are still in system S, which is moving to the right with velocity u in this system, while the charge Q is fixed in this system. What we would like to do is obtain an expression for the force in the stationary lab system, which is at rest with respect to both S and S'. Since

$$F_y = \frac{dp_y}{dt}$$

we can Lorentz-transform to the lab system using

$$\begin{aligned} dp_y &= dp_{y,\mathrm{lab}} \\ dt &= \gamma \left[dt_{\mathrm{lab}} - \frac{u}{c^2} dx_{\mathrm{lab}} \right] \end{aligned}$$

(the momenta in the y-direction are the same because there's no motion taking place in that direction, at least not yet). Thus, we have

$$F_y = \frac{dp_{y,\mathrm{lab}}}{\gamma \left[dt_{\mathrm{lab}} - u/c^2 \, dx_{\mathrm{lab}} \right]}$$

where

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}}$$

so that

$$F_y = \frac{[dp/dt]_{\mathrm{lab}}}{\gamma \left[1 - u^2/c^2 \right]}$$

This reduces to

$$F_y = \gamma \, F_{y,\mathrm{lab}}$$

or

$$\begin{aligned} F_{y,\mathrm{lab}} &= \frac{uv\lambda Q}{\epsilon_0 c^2 \pi R} \\ &= \frac{\mu_0 Q u I}{2\pi R} \end{aligned}$$

where we have used $c^2 = 1/\mu_0 \epsilon_0$ and $I = 2v\lambda$ is the total current. This expression for the force is identical to (1.1), which was obtained by assuming that there is a magnetic field associated with the total current $2v\lambda$ acting on the moving charge Q in the laboratory frame.

5. Conclusions

Special relativity indicates that what we perceive as a magnetic field is, in fact, only a consequence of the relativity of moving charges. So is there really a "magnetic field," or is it only an artifact of the way we observe things? Perhaps the only real thing is *electric charge* which, when set into motion, sets up an electric current and, consequently, what is mistakenly perceived as a magnetic field. But then how can one explain the fact that non-moving objects (magnets) can set up a field around them that we unambiguously see as a static magnetic field.?

Microscopically, magnets are known to be created by the aligned spins of negatively-charged electrons in ferromagnetic materials such as iron. If "spin" can be viewed as a kind of motion, then the magnetic field can indeed be explained as only an apparent manifestation of electric charge. In special relativity, one often asks the question "Are the contraction of object lengths and the dilation of time real phenomena or not?" And the answer is always given as "It depends on your point of view." Admittedly, this isn't much help.

At the same time, we have to recognize that magnetic fields can do no work and that there is no such thing as a "magnetic charge" (unless you believe in magnetic monopoles, which have never been found). In view of this, it is tempting to write off the magnetic field altogether and say that the electric field is the only real field.

In Faraday's time, electric and magnetic fields were understood to be two separate phenomena. Then Maxwell came along in the 1860s to prove that they are one and the same, and we called the unified phenomenon "electromagnetism." Can it be simplified any further? It depends on your point of view!

It is well known that the electric and magnetic fields can be expressed in terms of a scalar potential Φ and a vector potential A that satisfy

$$\begin{aligned}\mathbf{E} &= -\boldsymbol{\nabla}\Phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B} &= \boldsymbol{\nabla} \times \mathbf{A}\end{aligned}$$

(The combination $(\Phi, \mathbf{A})$ defines yet another four-vector in physics, one that Hermann Weyl tried to derive from his non-Riemannian geometry of 1918.) The latter expression appears to be almost an afterthought, given that $\mathbf{A}$ can be in principle be expressed in terms of $\mathbf{E}$ and Φ. But the most basic mathematical object in the general-relativistic theory of electromagnetic fields is the antisymmetric Maxwell tensor $F_{\mu\nu}$, which can be expressed in Cartesian matrix format as

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ +E_x & 0 & +B_z & -B_y \\ +E_y & -B_z & 0 & +B_x \\ +E_z & +B_y & -B_x & 0 \end{bmatrix}$$

where the Es and Bs are respectively the electric and magnetic components. All of Maxwell's equations can be expressed in terms of derivatives of this tensor and its upper-component form. In view of this, perhaps it is best to retain the magnetic field as a separate phenomenon, despite the fact that it can always be viewed as a moving current of electric charge. You be the judge.

References

David J. Griffiths, *Introduction to Electrodynamics*. Prentice-Hall, 3rd edition, 1999. A good introductory undergraduate text.

J.D. Jackson, *Classical Electrodynamics*. Wiley, 2nd edition, 1962. A long, dense, impenetrable graduate-level text, this book has discouraged more students from really learning the subject than I care to write about. I had to use the book, and I hated it. The mathematics is daunting, and it seems like every one of the text's problems takes hours to solve (but then I'm an idiot). The third edition (1998) is a vast improvement in terms of style, clarity and approach, but it appears to have been dumbed down somewhat, probably because of criticisms from people like me. Consider yourself warned.

Melvin Schwartz, *Principles of Electrodynamics*. Dover Publications, 1972. A great book from the great Nobel Laureate, it logically treats the electric and magnetic fields with the same units. I stole the title of this paper from Section 3.4 of Schwartz' book. May the Schwartz be with you!