Thomas precession and the BMT equation

April 12, 2016

Assigning the electron a spin, $\mathbf{s}$, and consequent magnetic moment $\boldsymbol{\mu}$, related by

$$\boldsymbol{\mu} = \frac{ge}{2mc}\mathbf{s}$$

introduced by Uhlenbeck and Goudsmit in 1926, explains the anomalous Zeeman effect if $g = 2$ and produces the correct multiplet splitting of spectral lines when an atom is in a magnetic field if $g = 1$. But it does not solve both problems at once. The next year, Thomas showed that the conflict is resolved by a relativistic effect.

1 Incorrect classical treatment

First, we work the problem incorrectly, as was done before Thomas' insight. The problem resides in the use of an incorrect equation of motion. Let an electron of magnetic moment $\boldsymbol{\mu}$ move in a magnetic field $\mathbf{B}$. Since the electron is moving, it sees a magnetic field $\mathbf{B}'$, so the rate of change of its angular momentum $\mathbf{s}$ is (incorrectly) given by the Newtonian formula

$$\left(\frac{d\mathbf{s}}{dt}\right)_{electron\,frame} = \boldsymbol{\mu} \times \mathbf{B}'$$

corresponding to an interaction energy of

$$U = -\boldsymbol{\mu} \cdot \mathbf{B}'$$

We show below that the magnetic field in the moving frame is given by

$$\begin{aligned} \mathbf{B}' &= \gamma\left(\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E}\right) - \frac{\gamma^2}{1+\gamma}\boldsymbol{\beta}\left(\boldsymbol{\beta} \cdot \mathbf{B}\right) \\ &\approx \mathbf{B} - \boldsymbol{\beta} \times \mathbf{E} \end{aligned}$$

where in the second line we take a nonrelativistic approximation, neglecting terms of order β^2 and higher. Dropping the β^2 terms is not the problem, however.

Now, the electron is in the electric field of the nucleus, which for single outer electron atoms may be written as the gradient of a spherically symmetric potential,

$$\mathbf{E} = -\frac{dV(r)}{dr}\frac{\mathbf{r}}{r}$$

With the orbital angular momentum of the electron given by

$$\mathbf{L} = \mathbf{r} \times m\mathbf{v}$$

we can write the energy as

$$\begin{aligned}
U &= -\boldsymbol{\mu}\cdot\mathbf{B}+\boldsymbol{\mu}\cdot(\boldsymbol{\beta}\times\mathbf{E}) \\
&= -\frac{ge}{2mc}\mathbf{s}\cdot\mathbf{B}+\frac{ge}{2mc^2}\mathbf{s}\cdot(\mathbf{v}\times\mathbf{E}) \\
&= -\frac{ge}{2mc}\mathbf{s}\cdot\mathbf{B}-\frac{ge}{2mc^2}\mathbf{s}\cdot(\mathbf{v}\times\mathbf{r})\,\frac{1}{r}\frac{dV(r)}{dr} \\
&= -\frac{ge}{2mc}\mathbf{s}\cdot\mathbf{B}+\frac{ge}{2m^2c^2}\mathbf{s}\cdot\mathbf{L}\frac{1}{r}\frac{dV(r)}{dr}
\end{aligned}$$

so that the second term becomes a spin-orbit interaction. With $g=2$, the first term is correct but the second term gives a spin-orbit interaction that is double the actual value.

2 Relativistic treatment

The problem with this calculation is that the equation of motion above holds only in an inertial frame of reference. However, the true inertial frame of the electron is a rotating one. Let the rest frame of the nucleus be the laboratory inertial frame, $\mathcal{O}$. Then consider two frames of reference for the electron, one at time t when the electron has velocity $\boldsymbol{\beta}$, and one a moment later at time $t+\delta t$, when the electron velocity is $\boldsymbol{\beta}+\delta\boldsymbol{\beta}$. We can write a boost that takes us to the electron rest frame at either of these times by using the appropriate velocity:

$$\begin{aligned}
(x')^\alpha &= [A_{boost}(\boldsymbol{\beta})]^\alpha{}_\beta\, x^\beta \\
(x'')^\alpha &= [A_{boost}(\boldsymbol{\beta}+\delta\boldsymbol{\beta})]^\alpha{}_\beta\, x^\beta
\end{aligned}$$

We can relate $\mathcal{O}''$ to $\mathcal{O}'$ by combining these

$$\begin{aligned}
(x'')^\alpha &= [A_{boost}(\boldsymbol{\beta}+\boldsymbol{\delta\beta})]^\alpha{}_\beta\, x^\beta \\
&= [A_{boost}(\boldsymbol{\beta}+\boldsymbol{\delta\beta})]^\alpha{}_\beta \left[A^{-1}_{boost}(\boldsymbol{\beta})\right]^\beta{}_\mu (x')^\mu \\
&= [A_{boost}(\boldsymbol{\beta}+\boldsymbol{\delta\beta})]^\alpha{}_\beta \left[A_{boost}(-\boldsymbol{\beta})\right]^\beta{}_\mu (x')^\mu
\end{aligned}$$

The essential observation here is that, since the two velocity vectors, $\boldsymbol{\beta},\boldsymbol{\beta}+\delta\boldsymbol{\beta}$ are in different directions, the transformation from one to the other involves a rotation as well as a boost, but the force does not apply any torque, so we need to remove the rotational part of the transformation.

Let the initial instantaneous rest frame of the electron, $\mathcal{O}'$, move with speed β in the x-direction, so that the boost back to the lab frame is

$$[A_{boost}(-\boldsymbol{\beta})]^\beta{}_\mu = \begin{pmatrix} \gamma & \beta\gamma & & \\ \beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Now let the velocity, $\boldsymbol{\beta}+\delta\boldsymbol{\beta}$, defining the frame $\mathcal{O}''$, lie in the xy-plane. Starting from our general formula for a boost,

$$\begin{aligned}
t' &= \gamma(ct-\boldsymbol{\beta}\cdot\mathbf{x}) \\
\mathbf{x}' &= \mathbf{x}+\frac{\gamma-1}{\beta^2}(\boldsymbol{\beta}\cdot\mathbf{x})\,\boldsymbol{\beta}-\gamma\boldsymbol{\beta}ct
\end{aligned}$$

we substitute $\boldsymbol{\beta}+\delta\boldsymbol{\beta}$ for $\boldsymbol{\beta}$ and expand to keep only terms linear in $\delta\boldsymbol{\beta}=\delta\beta_1\mathbf{i}+\delta\beta_2\mathbf{j}$. We need

$$\gamma'' = \frac{1}{\sqrt{1-(\boldsymbol{\beta}+\delta\boldsymbol{\beta})^2}}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{1-\beta^2-2\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}}} \\
&= \frac{1}{\sqrt{1-\beta^2}\sqrt{1-\frac{2\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}}{1-\beta^2}}} \\
&= \gamma\frac{1}{\sqrt{1-2\gamma^2\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}}} \\
&\approx \gamma\left(1+\gamma^2\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}\right) \\
&= \gamma+\gamma^3\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}
\end{aligned}$$

Then for the time component,

$$\begin{aligned}
t'' &= \gamma''\left(ct-(\boldsymbol{\beta}+\delta\boldsymbol{\beta})\cdot\mathbf{x}\right) \\
&= \left(\gamma+\gamma^3\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}\right)\left(ct-\beta x-\delta\boldsymbol{\beta}\cdot\mathbf{x}\right) \\
&= \gamma\left(ct-\beta x-\delta\beta_1 x-\delta\beta_2 y\right)+\gamma^3\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}\left(ct-\beta x\right) \\
&= \gamma\left(ct-\beta x-\delta\beta_1 x-\delta\beta_2 y\right)+\gamma^3\beta\delta\beta_1\left(ct-\beta x\right) \\
&= \gamma\left(1+\gamma^2\beta\delta\beta_1\right)ct-\left(\beta+\delta\beta_1+\gamma^2\beta\delta\beta_1\beta\right)\gamma x-\left(\delta\beta_2\gamma\right)y \\
&= \left(\gamma+\gamma^3\beta\delta\beta_1\right)ct-\left(\gamma\beta+\gamma^3\delta\beta_1\right)x-\left(\delta\beta_2\gamma\right)y
\end{aligned}$$

while for the spatial components,

$$\begin{aligned}
\mathbf{x}'' &= \mathbf{x}+\frac{\gamma''-1}{(\boldsymbol{\beta}+\delta\boldsymbol{\beta})^2}\left(\left((\boldsymbol{\beta}+\delta\boldsymbol{\beta})\cdot\mathbf{x}\right)(\boldsymbol{\beta}+\delta\boldsymbol{\beta})-\gamma''(\boldsymbol{\beta}+\delta\boldsymbol{\beta})ct\right) \\
&= \mathbf{x}+\left(\frac{\gamma''-1}{(\boldsymbol{\beta}+\delta\boldsymbol{\beta})^2}\left(\beta x+\delta\beta_1 x+\delta\beta_2 y\right)-\gamma'' ct\right)\boldsymbol{\beta} \\
&\quad+\left(\frac{\gamma''-1}{(\boldsymbol{\beta}+\delta\boldsymbol{\beta})^2}\left(\beta x+\delta\beta_1 x+\delta\beta_2 y\right)-\gamma'' ct\right)\delta\boldsymbol{\beta} \\
&= \mathbf{x}+\left(\frac{\gamma+\gamma^3\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}-1}{\beta^2\left(1+\frac{2}{\beta}\delta\beta_1\right)}\left(\beta x+\delta\beta_1 x+\delta\beta_2 y\right)-\left(\gamma+\gamma^3\beta\delta\beta_1\right)ct\right)\boldsymbol{\beta} \\
&\quad+\left(\frac{\gamma-1}{\beta}x-\gamma ct\right)\delta\boldsymbol{\beta} \\
&= \mathbf{x}+\left(\frac{\gamma-1}{\beta}x-\gamma ct\right)\delta\boldsymbol{\beta} \\
&\quad+\left(\frac{1}{\beta^2}\left(\gamma+\gamma^3\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}-1\right)\left(1-\frac{2}{\beta}\delta\beta_1\right)\left(\beta x+\delta\beta_1 x+\delta\beta_2 y\right)-\left(\gamma+\gamma^3\beta\delta\beta_1\right)ct\right)\boldsymbol{\beta} \\
&= \mathbf{x}+\left(\frac{\gamma-1}{\beta}x-\gamma ct\right)\delta\boldsymbol{\beta}+\left(\frac{1}{\beta^2}\left(\gamma-1+\gamma^3\boldsymbol{\beta}\cdot\delta\boldsymbol{\beta}\right)\left(\beta x-\delta\beta_1 x+\delta\beta_2 y\right)-\left(\gamma+\gamma^3\beta\delta\beta_1\right)ct\right)\boldsymbol{\beta} \\
&= \mathbf{x}+\left(\frac{\gamma-1}{\beta}x-\gamma ct\right)\delta\boldsymbol{\beta}+\left(\frac{\gamma-1}{\beta^2}\left(\beta x-\delta\beta_1 x+\delta\beta_2 y\right)+\gamma^3\delta\beta_1 x-\left(\gamma+\gamma^3\beta\delta\beta_1\right)ct\right)\boldsymbol{\beta}
\end{aligned}$$

In components,

$$\begin{aligned}
x'' &= x+\left(\frac{\gamma-1}{\beta}x-\gamma ct\right)\delta\beta_1+\left(\frac{\gamma-1}{\beta^2}\left(\beta x-\delta\beta_1 x+\delta\beta_2 y\right)+\gamma^3\delta\beta_1 x-\left(\gamma+\gamma^3\beta\delta\beta_1\right)ct\right)\beta \\
&= x+\frac{\gamma-1}{\beta}\delta\beta_1 x-\gamma\delta\beta_1 ct+\frac{\gamma-1}{\beta}\beta x-\frac{\gamma-1}{\beta}\delta\beta_1 x+\frac{\gamma-1}{\beta}\delta\beta_2 y+\gamma^3\beta\delta\beta_1 x-\gamma\beta ct-\gamma^3\beta\delta\beta_1\beta ct
\end{aligned}$$

$$
\begin{aligned}
&= \left(1+\frac{\gamma-1}{\beta}\delta\beta_1+\gamma-1-\frac{\gamma-1}{\beta}\delta\beta_1+\gamma^3\beta\delta\beta_1\right)x+\frac{\gamma-1}{\beta}\delta\beta_2 y-\gamma ct\left(\beta+\delta\beta_1\left(1+\gamma^2\beta^2\right)\right)\\
&= \left(\gamma+\gamma^3\beta\delta\beta_1\right)x+\left(\frac{\gamma-1}{\beta}\delta\beta_2\right)y-ct\left(\gamma\beta+\gamma^3\delta\beta_1\right)\\
y'' &= y+\frac{\gamma-1}{\beta}\delta\beta_2 x-\gamma\delta\beta_2 ct\\
z'' &= z
\end{aligned}
$$

Now write this as a matrix equation,

$$
\begin{pmatrix} ct'' \\ x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} \gamma+\gamma^3\beta\delta\beta_1 & -\gamma\beta-\gamma^3\delta\beta_1 & -\gamma\delta\beta_2 & 0 \\ -\beta\gamma-\gamma^3\delta\beta_1 & \gamma+\gamma^3\beta\delta\beta_1 & \frac{\gamma-1}{\beta}\delta\beta_2 & 0 \\ -\gamma\delta\beta_2 & \frac{\gamma-1}{\beta}\delta\beta_2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}
$$

so that we have

$$
\begin{aligned}
x^\alpha(t+\delta t) &= \left[A_{boost}(\boldsymbol{\beta}+\boldsymbol{\delta\beta})\right]^\alpha{}_\beta \left[A_{boost}(-\boldsymbol{\beta})\right]^\beta{}_\mu x^\mu(t)\\
&= \begin{pmatrix} \gamma+\gamma^3\beta\delta\beta_1 & -\gamma\beta-\gamma^3\delta\beta_1 & -\gamma\delta\beta_2 & 0 \\ -\beta\gamma-\gamma^3\delta\beta_1 & \gamma+\gamma^3\beta\delta\beta_1 & \frac{\gamma-1}{\beta}\delta\beta_2 & 0 \\ -\gamma\delta\beta_2 & \frac{\gamma-1}{\beta}\delta\beta_2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \beta\gamma & & \\ \beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}\\
&= \begin{pmatrix} \gamma\left(\gamma+\gamma^3\beta\delta\beta_1\right)-\gamma\beta\left(\gamma\beta+\gamma^3\delta\beta_1\right) & \gamma\beta\left(\gamma+\gamma^3\beta\delta\beta_1\right)-\gamma\left(\gamma\beta+\gamma^3\delta\beta_1\right) & -\gamma\delta\beta_2 & 0 \\ \gamma\left(-\beta\gamma-\gamma^3\delta\beta_1\right)+\gamma\beta\left(\gamma+\gamma^3\beta\delta\beta_1\right) & \gamma\beta\left(-\beta\gamma-\gamma^3\delta\beta_1\right)+\gamma\left(\gamma+\gamma^3\beta\delta\beta_1\right) & \frac{\gamma-1}{\beta}\delta\beta_2 & 0 \\ -\gamma^2\delta\beta_2+\gamma\beta\left(\frac{\gamma-1}{\beta}\delta\beta_2\right) & -\gamma^2\beta\delta\beta_2+\gamma\left(\frac{\gamma-1}{\beta}\delta\beta_2\right) & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} x^\mu(t)
\end{aligned}
$$

Simplifying,

$$
\begin{aligned}
\gamma\left(\gamma+\gamma^3\beta\delta\beta_1\right)-\gamma\beta\left(\gamma\beta+\gamma^3\delta\beta_1\right) &= \gamma^2+\gamma^4\beta\delta\beta_1-\gamma^2\beta^2-\gamma^4\beta\delta\beta_1\\
&= \gamma^2-\gamma^2\beta^2\\
&= 1\\
\gamma\beta\left(\gamma+\gamma^3\beta\delta\beta_1\right)-\gamma\left(\gamma\beta+\gamma^3\delta\beta_1\right) &= \gamma^2\beta+\gamma^4\beta^2\delta\beta_1-\gamma^2\beta-\gamma^4\delta\beta_1\\
&= \gamma^4\left(\beta^2-1\right)\delta\beta_1\\
&= -\gamma^2\delta\beta_1\\
\beta\gamma\left(-\beta\gamma-\gamma^3\delta\beta_1\right)+\gamma\left(\gamma+\gamma^3\beta\delta\beta_1\right) &= -\beta^2\gamma^2-\beta\gamma^4\delta\beta_1+\gamma^2+\gamma^4\beta\delta\beta_1\\
&= \gamma^2-\beta^2\gamma^2\\
&= 1\\
-\gamma^2\delta\beta_2+\gamma\beta\left(\frac{\gamma-1}{\beta}\delta\beta_2\right) &= -\gamma^2\delta\beta_2+\left(\gamma^2-\gamma\right)\delta\beta_2\\
&= -\gamma\delta\beta_2\\
-\gamma^2\beta\delta\beta_2+\gamma\left(\frac{\gamma-1}{\beta}\delta\beta_2\right) &= \frac{1}{\beta}\left(-\gamma^2\beta^2+\left(\gamma^2-\gamma\right)\right)\delta\beta_2\\
&= \frac{1}{\beta}\left(\gamma^2\left(1-\beta^2\right)-\gamma\right)\delta\beta_2\\
&= \frac{1-\gamma}{\beta}\delta\beta_2
\end{aligned}
$$

Therefore,

$$x^{\alpha}(t+\delta t) = \begin{pmatrix} 1 & -\gamma^2\delta\beta_1 & -\gamma\delta\beta_2 & 0 \\ -\gamma^2\delta\beta_1 & 1 & \frac{\gamma-1}{\beta}\delta\beta_2 & 0 \\ -\gamma\delta\beta_2 & \frac{1-\gamma}{\beta}\delta\beta_2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{\alpha}_{\ \ \mu} x^{\mu}(t)$$

The most useful way to think about this matrix is as the product of an infinitesimal boost and an infinitesimal rotation,

$$\begin{pmatrix} 1 & -\gamma^2\delta\beta_1 & -\gamma\delta\beta_2 & 0 \\ -\gamma^2\delta\beta_1 & 1 & \frac{\gamma-1}{\beta}\delta\beta_2 & 0 \\ -\gamma\delta\beta_2 & \frac{1-\gamma}{\beta}\delta\beta_2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \left(1 + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \left(\frac{\gamma-1}{\beta}\right)\delta\beta_2 & 0 \\ 0 & -\left(\frac{\gamma-1}{\beta}\right)\delta\beta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right) \times \left(1 + \begin{pmatrix} 0 & -\gamma^2\delta\beta_1 & -\gamma\delta\beta_2 & 0 \\ -\gamma^2\delta\beta_1 & 0 & 0 & 0 \\ -\gamma\delta\beta_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right)$$

Recall that for a rotation in the xy plane,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & \\ \sin\theta & \cos\theta & \\ & & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

so if θ is infinitesimal, $\delta\theta$, the transformation matrix is

$$\begin{pmatrix} 1 & -\delta\theta & \\ \delta\theta & 1 & \\ & & 1 \end{pmatrix} = 1 + \begin{pmatrix} 0 & -\delta\theta & 0 \\ \delta\theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and we may identify the first factor as a rotation by an angle

$$\delta\theta = -\left(\frac{\gamma-1}{\beta}\right)\delta\beta_2$$

around the z axis. We can get both angle and axis by writing this as

$$\begin{aligned} \delta\boldsymbol{\theta} &= -\left(\frac{\gamma-1}{\beta^2}\right)\boldsymbol{\beta}\times\delta\boldsymbol{\beta} \\ &= -\left(\frac{\gamma-1}{\left(\frac{\gamma^2-1}{\gamma^2}\right)}\right)\boldsymbol{\beta}\times\delta\boldsymbol{\beta} \\ &= -\left(\frac{\gamma^2(\gamma-1)}{\gamma^2-1}\right)\boldsymbol{\beta}\times\delta\boldsymbol{\beta} \\ &= -\left(\frac{\gamma^2}{\gamma+1}\right)\boldsymbol{\beta}\times\delta\boldsymbol{\beta} \end{aligned}$$

The second factor in $x^{\alpha}(t+\delta t)$ is a pure infinitesimal boost.

The rotation transformation results in an angular velocity, the Thomas precession, given by

$$\boldsymbol{\omega}_T = -\frac{d\boldsymbol{\theta}}{dt}$$

$$
\begin{aligned}
&= -\frac{1}{\gamma}\frac{d\boldsymbol{\theta}}{d\tau} \\
&= \frac{\gamma}{\gamma+1}\boldsymbol{\beta}\times\frac{d\boldsymbol{\beta}}{d\tau} \\
&= -\frac{\gamma}{\gamma+1}\frac{\mathbf{a}\times\mathbf{v}}{c^2}
\end{aligned}
$$

In the case of our atom, the (non-relativistic) acceleration is produced by the potential $V(r)$,

$$
\begin{aligned}
\mathbf{a} &= \frac{e\mathbf{E}}{m} \\
&= -\frac{e}{m}\frac{dV(r)}{dr}\frac{\mathbf{r}}{r}
\end{aligned}
$$

so that with $\gamma \approx 1$,

$$
\begin{aligned}
\boldsymbol{\omega}_T &= -\frac{\gamma}{\gamma+1}\mathbf{a}\times\frac{\mathbf{v}}{c^2} \\
&= -\frac{1}{2}\left(-\frac{e}{m}\frac{dV(r)}{dr}\frac{\mathbf{r}}{r}\right)\times\frac{\mathbf{v}}{c^2} \\
&= \frac{e}{2mc^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{r}\times\mathbf{v} \\
&= \frac{e}{2m^2c^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{L}
\end{aligned}
$$

Now return to the equation of motion for the electron. Since we have a rotating frame of reference, we must replace

$$
\frac{d}{dt} \rightarrow \frac{d}{dt} + \boldsymbol{\omega}_T \times
$$

This means that instead of the equation of motion

$$
\left(\frac{d\mathbf{s}}{dt}\right)_{electron\,frame} = \boldsymbol{\mu}\times(\mathbf{B}-\boldsymbol{\beta}\times\mathbf{E})
$$

we have

$$
\begin{aligned}
\left(\frac{d}{dt}+\boldsymbol{\omega}_T\times\right)\mathbf{s} &= \boldsymbol{\mu}\times(\mathbf{B}-\boldsymbol{\beta}\times\mathbf{E}) \\
\frac{d\mathbf{s}}{dt} &= \boldsymbol{\mu}\times(\mathbf{B}-\boldsymbol{\beta}\times\mathbf{E}) - \boldsymbol{\omega}_T\times\mathbf{s} \\
&= \boldsymbol{\mu}\times\left(\mathbf{B}-\boldsymbol{\beta}\times\mathbf{E}+\frac{2mc}{ge}\boldsymbol{\omega}_T\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{ge}{2mc}\mathbf{s}\times\mathbf{B} - \frac{ge}{2mc}\mathbf{s}\times(\boldsymbol{\beta}\times\mathbf{E}) + \mathbf{s}\times\boldsymbol{\omega}_T \\
&= \frac{ge}{2mc}\mathbf{s}\times\mathbf{B} - \mathbf{s}\times\left(\frac{ge}{2mc}(\boldsymbol{\beta}\times\mathbf{E}) - \boldsymbol{\omega}_T\right)
\end{aligned}
$$

where we use $\boldsymbol{\mu} = \frac{ge}{2mc}\mathbf{s}$. The energy is therefore modified to

$$
U = -\boldsymbol{\mu}\cdot\left(\mathbf{B}-\boldsymbol{\beta}\times\mathbf{E}+\frac{2mc}{ge}\boldsymbol{\omega}_T\right)
$$

so the energy is changed from our incorrect classical result by

$$
\begin{aligned}
-\boldsymbol{\mu}\cdot\frac{2mc}{ge}\boldsymbol{\omega}_T &= -\frac{ge}{2mc}\mathbf{s}\cdot\frac{2mc}{ge}\boldsymbol{\omega}_T \\
&= -\frac{ge}{2mc}\mathbf{s}\cdot\frac{2mc}{ge}\left(\frac{e}{2m^2c^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{L}\right) \\
&= -\frac{e}{2m^2c^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{s}\cdot\mathbf{L}
\end{aligned}
$$

so the correct energy is

$$
\begin{aligned}
U &= \left(-\frac{ge}{2mc}\mathbf{s}\cdot\mathbf{B}+\frac{ge}{2m^2c^2}\mathbf{s}\cdot\mathbf{L}\frac{1}{r}\frac{dV(r)}{dr}\right)-\frac{e}{2m^2c^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{s}\cdot\mathbf{L} \\
&= -\frac{ge}{2mc}\mathbf{s}\cdot\mathbf{B}+\frac{(g-1)e}{2m^2c^2}\mathbf{s}\cdot\mathbf{L}\frac{1}{r}\frac{dV(r)}{dr}
\end{aligned}
$$

The coefficient of the spin-orbit interaction energy is changed from $g=2$ to $g-1=1$, giving the required factor of 2.

3 The BMT equation

First consider orbital angular momentum,

$$\mathbf{L}=\mathbf{r}\times\mathbf{p}$$

We may immediately generalize this to spacetime by defining

$$L^{\alpha\beta}\equiv x^\alpha p^\beta - x^\beta p^\alpha$$

where x^α is the position 4-vector of a particle and p^α is its momentum. Consider the dual,

$$
\begin{aligned}
\mathcal{L}_{\mu\nu} &\equiv \frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}L^{\alpha\beta} \\
&= \frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}x^\alpha p^\beta
\end{aligned}
$$

and contract with $\frac{1}{c}$ times the *observer's* 4-velocity,

$$
\begin{aligned}
L_\nu &\equiv \frac{1}{c}u^\mu\mathcal{L}_{\mu\nu} \\
&= \frac{1}{c}u^\mu\varepsilon_{\mu\nu\alpha\beta}x^\alpha p^\beta
\end{aligned}
$$

Then in the rest frame of the observer, where $u^\mu=(c,\mathbf{0})$, L_ν becomes

$$
\begin{aligned}
L_\nu &\equiv \frac{1}{c}u^\mu\mathcal{L}_{\mu\nu} \\
&= \varepsilon_{0\nu\alpha\beta}x^\alpha p^\beta \\
&= \left(0,\varepsilon_{ijk}x^j p^k\right) \\
&= (0,\mathbf{L})
\end{aligned}
$$

This shows that we expect angular momentum to be described by a spacelike vector, and we'll treat the spin vector in the same way. This property, of reducing to a purely spacelike vector in the rest frame, may be characterized by orthogonality to the 4-velocity, $u^\alpha L_\alpha=0$.

We are now in a position to consider the fully relativistic evolution of spin angular momentum. We know that the equation of motion of the 3-dimensional spin vector, $\mathbf{s}$, in the rest frame $\tilde{\mathcal{O}}$ of the particle is

$$\frac{d\mathbf{s}}{d\tilde{t}} = \frac{ge}{2mc}\mathbf{s} \times \tilde{\mathbf{B}}$$

and want to generalize this to a covariant expression. We begin by generalizing the spin 3-vector to a spin 4-vector, which we assume reduces to

$$\tilde{s}^\alpha = (0, \mathbf{s})$$

in the rest frame $\tilde{\mathcal{O}}$, or equivalently,

$$\tilde{u}_\alpha \tilde{s}^\alpha = 0$$

and this relation is invariant.

To generalize the rest frame equation, we will clearly want the proper time derivative of the 4-vector spin

$$\frac{ds^\alpha}{d\tau}$$

expressed in terms of other 4-vectors. Rather than trying to transform the known expression, we try to write the most general covariant expression we can that reduces to the known non-relativistic expression. The only tensors relevant to the problem are

$$F^{\alpha\beta}, u^\alpha, s^\alpha, \frac{du^\alpha}{d\tau}$$

Notice that the 4-acceleration, $a^\alpha = \frac{du^\alpha}{d\tau}$, has a part expressible in terms of the first two but may also depend on non-electromagnetic forces, so we need to consider it separately. Our goal is to write $\frac{ds^\alpha}{d\tau}$ in terms of these. Looking at the classical limit in the rest frame, we expect that the correct expression is at most linear in the fields $F^{\alpha\beta}$, and each term is linear in the spin vector. Then there are only three possibile 4-vectors we can construct:

$$\begin{array}{c} F^{\alpha\beta}s_\beta \\ \left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha \\ \left(s_\beta \dfrac{du^\beta}{d\tau}\right)u^\alpha \end{array}$$

(Notice that $\left(u^\beta s_\beta\right)u^\alpha = 0$, $F^{\alpha\beta}u_\beta$ is independent of s^α, and $\left(F^{\mu\nu}u_\mu u_\nu\right)u^\alpha = 0$. Since the acceleration is already linear in $F^{\alpha\beta}$ we do not consider terms involving $F^{\alpha\beta}a_\beta$).

We begin with an arbitrary linear combination of these, and ask what combination gives the right behavior in the particle rest frame,

$$\frac{ds^\alpha}{d\tau} = AF^{\alpha\beta}s_\beta + B\left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha + C\left(s_\beta \frac{du^\beta}{d\tau}\right)u^\alpha \tag{1}$$

First, notice that since $u^\alpha s_\alpha = 0$, it follows that

$$\begin{aligned} 0 &= \frac{d}{d\tau}\left(u^\alpha s_\alpha\right) \\ &= \frac{du^\alpha}{d\tau}s_\alpha + u^\alpha \frac{ds_\alpha}{d\tau} \end{aligned}$$

so that

$$u^\alpha \frac{ds_\alpha}{d\tau} = -s_\alpha \frac{du^\alpha}{d\tau} \tag{2}$$

Rewrite this using the Lorentz force law, combined with other non-electromagnetic forces, F^α

$$m\frac{du^\alpha}{d\tau} = F^\alpha + \frac{e}{c}u_\beta F^{\alpha\beta}$$

Then the contraction of the rate of change of spin with the 4-velocity is

$$\begin{aligned} u^\alpha \frac{ds_\alpha}{d\tau} &= -s_\alpha \frac{du^\alpha}{d\tau} \\ &= -\frac{1}{m}s_\alpha\left(F^\alpha + \frac{e}{c}F^{\alpha\beta}u_\beta\right) \end{aligned}$$

Take this same contraction with our ansatz, eq.(1),

$$\begin{aligned} \frac{ds^\alpha}{d\tau} &= AF^{\alpha\beta}s_\beta + B\left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha + C\left(s_\beta\frac{du^\beta}{d\tau}\right)u^\alpha \\ u_\alpha\frac{ds^\alpha}{d\tau} &= Au_\alpha F^{\alpha\beta}s_\beta + B\left(s_\mu F^{\mu\nu}u_\nu\right)u_\alpha u^\alpha + C\left(s_\beta\frac{du^\beta}{d\tau}\right)u_\alpha u^\alpha \\ &= Au_\alpha F^{\alpha\beta}s_\beta - c^2B\left(s_\mu F^{\mu\nu}u_\nu\right) - c^2C\left(s_\beta\frac{du^\beta}{d\tau}\right) \end{aligned}$$

and substituting on the left and in the final term,

$$-\frac{1}{m}s_\alpha\left(F^\alpha + \frac{e}{mc}u_\beta F^{\beta\alpha}\right) = Au_\alpha F^{\alpha\beta}s_\beta - c^2B\left(s_\mu F^{\mu\nu}u_\nu\right) - \frac{c^2}{m}C\left(s_\beta\left(F^\beta + \frac{e}{mc}u_\alpha F^{\alpha\beta}\right)\right)$$

Since the non-electromagnetic force F^α is independent of the electromagnetic force, we may split this into two independent equations,

$$\begin{aligned} -\frac{1}{m}s_\alpha F^\alpha &= -\frac{c^2}{m}Cs_\beta F^\beta \\ -\frac{e}{m^2c}u_\beta F^{\beta\alpha}s_\alpha &= Au_\alpha F^{\alpha\beta}s_\beta - c^2B\left(s_\mu F^{\mu\nu}u_\nu\right) - \frac{c^2}{m}C\left(\frac{e}{mc}u_\alpha F^{\alpha\beta}s_\beta\right) \end{aligned}$$

The first of these immediately gives $C = \frac{1}{c^2}$, while the second becomes

$$\begin{aligned} 0 &= \left(A + c^2B - \frac{c^2}{m}C\frac{e}{mc} + \frac{e}{m^2c}\right)u_\beta F^{\beta\alpha}s_\alpha \\ &= \left(A + c^2B\right)u_\beta F^{\beta\alpha}s_\alpha \end{aligned}$$

With these restrictions, the equation of motion reduces to

$$\frac{ds^\alpha}{d\tau} = AF^{\alpha\beta}s_\beta - \frac{A}{c^2}\left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha + \frac{1}{c^2}\left(s_\beta\frac{du^\beta}{d\tau}\right)u^\alpha$$

Now consider the limit to the rest frame, in the case of a pure magnetic field, so $\mathbf{E} = \mathbf{0}$ in $F^{\alpha\beta}$, and the 4-velocity is given by

$$u^\alpha = (c, \mathbf{0})$$

Then $s_\mu F^{\mu\nu}u_\nu = -cs_iF^{i0} = cs_iE^i = 0$ so the middle term drops out. Then the time component simply reproduces eq.(2), while the spatial components give

$$\frac{ds^i}{dt} = AF^{ij}s_j$$

$$
\begin{aligned}
&= A \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & B^3 & -B^2 \\ 0 & -B^3 & 0 & B^1 \\ 0 & B^2 & -B^1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ s^1 \\ s^2 \\ s^3 \end{pmatrix} \\
&= A \begin{pmatrix} 0 \\ B^3 s^2 - B^2 s^3 \\ B^1 s^3 - B^3 s^1 \\ B^2 s^1 - B^1 s^2 \end{pmatrix}
\end{aligned}
$$

or simply

$$\frac{d\mathbf{s}}{dt} = A\left(\mathbf{s} \times \mathbf{B}\right)$$

and comparing this to

$$\frac{d\mathbf{s}}{dt} = \frac{ge}{2mc}\mathbf{s} \times \mathbf{B}$$

we see that

$$A = \frac{ge}{2mc}$$

and using the Lorentz force law $\frac{du^\alpha}{d\tau} = \frac{q}{mc}F^{\alpha\beta}u_\beta$ to substitute the electromagnetic force for the acceleration, the full equation of motion for $\mathbf{s}$ becomes

$$
\begin{aligned}
\frac{ds^\alpha}{d\tau} &= \frac{ge}{2mc}F^{\alpha\beta}s_\beta - \frac{ge}{2mc^3}\left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha + \frac{1}{c^2}\left(s_\beta \frac{du^\beta}{d\tau}\right)u^\alpha \\
&= \frac{ge}{2mc}F^{\alpha\beta}s_\beta - \frac{ge}{2mc^3}\left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha + \frac{1}{c^2}\left(s_\beta\left(\frac{e}{mc}F^{\alpha\beta}u_\beta\right)\right)u^\alpha \\
&= \frac{ge}{2mc}F^{\alpha\beta}s_\beta - \frac{(g-2)\,e}{2mc^3}\left(s_\mu F^{\mu\nu}u_\nu\right)u^\alpha
\end{aligned}
$$

This is the BMT equation. The signs differ from Jackson because we use the opposite convention for the sign of the metric, $\eta_{\alpha\beta}$.