

## 8.2: Angular Momentum

### Learning Objectives

- Explain angular momentum

Nonrelativistically, the angular momentum of a particle with momentum  $p$ , at a position  $r$  relative to some arbitrarily fixed point, is  $L = r \times p$ . When we generalize this equation to relativity, we run into a number of issues. Issues due to special relativity:

1. The vector cross product only makes sense in three dimensions, so it is not well defined in special relativity (section 7.6)
2. Assuming we get around issue number 1, how do we know that this quantity is conserved?

And from general relativity:

3. In general relativity, only infinitesimally small spatial or spacetime displacements  $dr$  can be treated as vectors. Larger ones cannot. This is because spacetime can be curved, and vectors can't be used to define displacements on a curved space (e.g., the surface of the earth).
4. If space has a nontrivial topology, then we may not be able to define an orientation (section 7.6)

For points 3 and 4, we refer to Hawking and Ellis in section 3.5. Number 2 is addressed in section 9.3. For number 2 we will need the stress-energy tensor, which will be described in chapter 9. Lest you feel totally cheated, we will resolve issue number 1 in this section itself, but before we do that, let's consider an interesting example that can be handled with simpler math.

### The Relativistic Bohr model

If we want to see an interesting real-world example of relativistic angular momentum, we need something that rotates at relativistic velocities. At large scales we have astrophysical examples such as neutron stars and the accretion disks of black holes, but these involve gravity and would therefore require general relativity. At microscopic scales we have systems such as hadrons, nuclei, atoms, and molecules. These are quantum-mechanical, and relativistic quantum mechanics is a difficult topic that is beyond the scope of this book, but we can sidestep that issue by using the Bohr model of the atom. In the Bohr model of hydrogen, we assume that the electron has a circular orbit governed by Newton's laws, does not radiate, and has its angular momentum quantized in units of  $\hbar$ . Let's generalize the Bohr model by applying relativity.

It will be convenient to define the constant  $\alpha = \frac{ke^2}{\hbar}$ , known as the fine structure constant, where  $k$  is the Coulomb constant and  $e$  is the fundamental charge. The fine structure constant is unitless and is approximately  $1/137$ . It is essentially a measure of the strength of the electromagnetic interaction, and in the Bohr model it also turns out to be the velocity of the electron (in units of  $c$ ) in the ground state of hydrogen. Because this velocity is small compared to 1, we expect relativistic corrections in hydrogen to be small — of relative size  $\alpha^2$ . But we have an interesting opportunity to get at some additional and more exciting physics if we consider a hydrogenlike atom, i.e., an ion with  $Z$  protons in the nucleus and only one electron. Raising  $Z$  cranks up the energy scale and therefore increases the velocity as well.

Combining the Coulomb force law with the result of Example 4.5.1, for uniform circular motion, we have

$$\frac{kZe^2}{r} = m\gamma v^2 \quad (8.2.1)$$

where the factor of  $\gamma$  is the relativistic correction. The electron's momentum is perpendicular to the radius vector, so we assume for the moment that (as turns out to be true), the angular momentum is given by  $L = rp = m\gamma vr$ , where again a relativistic correction factor of  $\gamma$  appears. This is quantized, so let  $L = l\hbar$ , where  $l$  is an integer. Solving these equations gives

$$v = \frac{Z\alpha}{l} \quad (8.2.2)$$

$$r = \frac{l^2 \hbar}{mZ\alpha\gamma} \quad (8.2.3)$$

These differ from the nonrelativistic versions only by the factor of  $\gamma$  in the second equation. The electrical energy is  $U = -kZe^2/r$ , and the kinetic energy  $K = m(\gamma - 1)$  (with  $c = 1$ ). We will find it convenient to work with the (positive) binding energy in units of the mass of the electron. Call this quantity  $\varepsilon$ . After some algebra, the result is

$$\varepsilon = 1 - \sqrt{1 - v^2} \quad (8.2.4)$$

Surprisingly, this is also the exact result given by relativistic quantum mechanics if we solve the Dirac equation for the ground state, or if we take a high-energy (nearly unbound) state with the maximum value of  $l$ , as is appropriate for a semiclassical circular orbit. So we can see that even though our quantum mechanics was crude, our relativity makes some sense and gives reasonable results. For small  $Z$ , a Taylor series approximation gives

$$E = v^2/2 + v^4/8 + \dots \quad (8.2.5)$$

where the fourth-order term represents the relativistic correction.

So far so good, but now what if we crank up the value of  $Z$  to make the relativistic effects strong? A very disturbing thing happens when we make  $Z \gtrsim 137 \approx 1/\alpha$ . In the ground state we get  $v > 1$  and a complex number for  $\varepsilon$ . Clearly something has broken down, and our results no longer make sense. We might be inclined to dismiss this as a consequence of our crude model, but remember, our calculations happened to give the same result as the Dirac equation, which has real relativistic quantum mechanics baked in. We should take this breakdown as evidence of a real physical breakdown. The interpretation is as follows.

According to quantum mechanics, the vacuum isn't really a vacuum. Particle-antiparticle pairs are continually popping into existence in empty space and then reannihilating one another. Their temporary creation is a violation of the conservation of mass-energy, but only a temporary violation, and this is allowed by the time-energy form of the Heisenberg uncertainty principle,  $\Delta E \Delta t \gtrsim \hbar$ , as long as  $\Delta t$  is short. It's as though we steal some money, but the police don't catch us as long as we put it back before anyone can notice. Because these particles are only temporarily in our universe, we call them virtual particles, as opposed to real particles that have a potentially permanent existence and can be detected as blips on a Geiger counter.

But when the vacuum contains an electric field that is beyond a certain critical strength, it becomes possible to create an electron-antielectron pair, let the opposite charges separate and release energy, and pay off the energy debt without having to reannihilate the particles. This is known as "sparking the vacuum." As of this writing, only nuclei with  $Z$  up to about 118 have been discovered, and in any case the critical  $Z$  value of  $1/\alpha \approx 137$  was only a rough estimate. But by colliding heavy nuclei such as lead, one can at least temporarily form an unstable compound system with a high  $Z$ , and attempts are being made to search for the predicted effect in the laboratory.

## The Angular Momentum Tensor

As mentioned previously, there is no such thing as a vector cross product in four dimensions, so the nonrelativistic definition of angular momentum as  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$  needs to be modified to be usable in relativity.

Given a position vector  $r^a$  and a momentum vector  $p^b$ , we expect based both on units and the correspondence principle that a relativistic definition of angular momentum must be some kind of a product of the vectors. Based on the rules of index notation, we don't have much leeway here. The only products we can form are  $r^a p^b$ , which is a rank-2 tensor, or  $r^a p_a$ , a scalar. Since nonrelativistic angular momentum is a three-vector, the correspondence principle tells us that its relativistic incarnation can't be a scalar — there simply wouldn't be enough information in a scalar to tell us the things that the nonrelativistic angular momentum vector tells us: what axis the rotation is about, and which direction the rotation is.

The tensor  $r^a p^b$  also has a problem, but one that can be fixed. Suppose that in a certain frame of reference a particle of mass  $m \neq 0$  is at rest at the origin. Then its position four-vector at time  $t$  is  $(t, 0, 0, 0)$ , and its energy-momentum vector is  $(m, 0, 0, 0)$ . These vectors are parallel. The tensor  $r^a p^b$  is nonzero and nonconserved as time flows, but clearly we want the angular momentum of an isolated particle to be conserved. Another example would be if, at a certain moment in time, we had  $r = (0, x, 0, 0)$  and  $p = (E, p, 0, 0)$ , with both  $x$  and  $p$  positive. This particle's motion is directly away from the origin, so its angular momentum should be zero by symmetry, but  $r^a p^b$  is again nonzero.

The way to fix the problem is to force the product of the position and momentum vectors to be an antisymmetric tensor:

$$L^{ab} = r^a p^b - r^b p^a \quad (8.2.6)$$

Antisymmetric means that  $L^{ab} = -L^{ba}$ , so that elements on opposite sides of the main diagonal are the same except for opposite signs. A quick check shows that this gives the expected zero result in both of the above examples. A component such as  $L^{yz}$  measures the amount of rotation in the  $y - z$  plane. In a nonrelativistic context, we would have described this as an  $x$  component  $L_x$  of the angular momentum three-vector, because a rotation of the  $y - z$  plane about the origin is a rotation about the  $x$  axis — such a rotation keeps the  $x$  axis fixed. But in four-dimensional spacetime, a rotation in the  $y - z$  plane keeps the entire  $t - x$  plane

fixed, so the notion of rotation “about an axis” breaks down. (Notice the pattern: in two dimensions we rotate about a point, in three dimensions rotation is about a line, and in four dimensions we rotate about a fixed plane.) In section 9.3, we show that  $L^{ab}$  is conserved.

If we lay the angular momentum tensor out in matrix format, it looks like this:

$$\begin{pmatrix} 0 & L^{tx} & L^{ty} & L^{tz} \\ & 0 & L^{xy} & L^{xz} \\ & & 0 & L^{yz} \\ & & & 0 \end{pmatrix} \quad (8.2.7)$$

The zeroes on the main diagonal are due to the antisymmetrization in the definition. I’ve left blanks below the main diagonal because although those components can be nonzero, they only contain a (negated) copy of the information given by the ones above the diagonal. We can see that there are really only 6 pieces of information in this  $4 \times 4$  matrix, and we’ve already physically interpreted the triangular cluster of three space-space components on the bottom right.

Why do we have the row on the top, consisting of the time-space components, and what do they mean physically? A highbrow answer would be that this is something very deep having to do with the fact that, as described in Section 8.3, rotation and linear motion are not as cleanly separated in relativity as they are in nonrelativistic physics. A more straightforward answer is that in most situations these components are actually not very interesting. Consider a cloud of particles labeled  $i = 1$  through  $n$ . Then for a representative component from the top row we have the total value

$$L^{tx} = \sum t_i p_i^x - \sum x_i E_i \quad (8.2.8)$$

Now suppose that we fix a certain surface of simultaneity at time  $t$ . The sum becomes

$$L^{tx} = t \sum p_i^x - \sum x_i E_i \quad (8.2.9)$$

There is information here, but it’s not exciting information about angular momentum, it’s boring information about the position and motion of the system’s center of mass. If we fix a frame of reference in which the total momentum is zero, i.e., the center of mass frame, then we have  $\sum p_i^x = 0$ . Let’s also define the position of the center of mass as the average position weighted by mass-energy, rather than the mass-weighted average, as we would do in Newtonian mechanics. Then the  $\sum x_i E_i$  is a constant relating to the position of the center of mass, and if we like we can make it equal zero by choosing the origin of our spatial coordinates to coincide with the center of mass.

With these choices we have a much simpler angular momentum tensor:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ & 0 & L^{xy} & L^{xz} \\ & & 0 & L^{yz} \\ & & & 0 \end{pmatrix} \quad (8.2.10)$$

If we wish, we can sprinkle some notational sugar on top of all of this using the Levi-Civita tensor described in optional Section 7.6. Let’s define a new tensor  ${}^*L$  according to

$${}^*L_{ij} = \frac{1}{2} \epsilon_{ijkl} L^{kl} \quad (8.2.11)$$

Then for an observer with velocity vector  $o^\mu$ , the quantity  $o_\mu {}^*L^{\mu\nu}$  has the form  $((0, L^{\wedge\{yz\}}, L^{\wedge\{zx\}}, L^{\wedge\{xy\}})$ . That is, its spatial components are exactly the quantities we would have expected for the nonrelativistic angular momentum three-vector (using the correct relativistic momentum).

---

This page titled [8.2: Angular Momentum](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Benjamin Crowell](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.