Chapter 9

Angular momentum, Part II (General $\hat{\mathbf{L}}$)

In Chapter 8, we discussed situations where the direction of the vector **L** remained constant, and only its magnitude changed. In this chapter, we will look at more general situations where the direction of **L** is allowed to change. The vector nature of **L** will prove to be vital here, and we will arrive at all sorts of strange results for spinning tops and such things. This chapter is rather long, alas, but the general outline is that the first three sections cover general theory, then Section 9.4 introduces some actual physical setups, and then Section 9.6 begins the discussion of tops.

9.1 Preliminaries concerning rotations

9.1.1 The form of general motion

Before getting started, we should make sure we're all on the same page concerning a few important things about rotations. Because rotations generally involve three dimensions, they can often be hard to visualize. A rough drawing on a piece of paper might not do the trick. For this reason, this chapter is one of the more difficult ones in this book. But to ease into it, the next few pages consist of some definitions and helpful theorems. This first theorem describes the general form of any motion. You might consider it obvious, but it's a little tricky to prove.

Theorem 9.1 *(Chasles' theorem) Consider a rigid body undergoing arbitrary motion. Pick any point P in the body. Then at any instant (see Fig. 9.1), the motion of the body can be written as the sum of the translational motion of P, plus a rotation around some axis (which may change with time) through P.*[1]

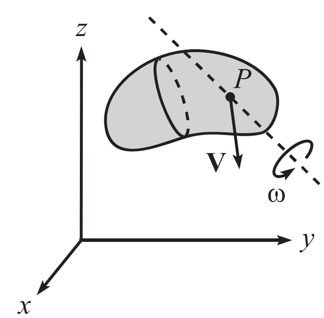

Fig. 9.1

Proof: The motion of the body can be written as the sum of the translational motion of P, plus some other motion relative to P (this is true because relative coordinates are additive quantities). We must show that this latter motion is a rotation. This seems quite plausible, and it holds because the body is rigid; that is,

[1] In other words, a person at rest with respect to a frame whose origin is P, and whose axes are parallel to the fixed-frame axes, sees the body undergoing a rotation around some axis through P.

all points keep the same distances relative to each other. If the body weren't rigid, then this theorem wouldn't be true.

To be rigorous, consider a spherical shell fixed in the body, centered at P. The motion of the body is completely determined by the motion of the points on this sphere, so we need only examine what happens to the sphere. Because distances are preserved in the rigid body, the points on the sphere must always remain the same radial distance from P. And because we are looking at motion relative to P, we have therefore reduced the problem to the following: In what manner can a rigid sphere transform into itself? We claim that any such transformation has the property that there exist two points that end up where they started.[2] These two points must then be diametrically opposite points (assuming that the whole sphere doesn't end up back where it started, in which case every point ends up where it started), because distances are preserved; given one point that ends up where it started, the diametrically opposite point must also end up where it started, to maintain the distance of a diameter.

If this claim is true, then we are done, because for an infinitesimal transformation, a given point moves in only one direction, because there is no time to do any turning. So a point that ends up where it started must have remained fixed for the whole (infinitesimal) time. Therefore, all the points on the diameter joining the two fixed points must also have remained fixed the whole time, because distances are preserved. So we are left with a rotation around this axis.

This "two points ending up where they started" claim is quite believable, but nevertheless tricky to prove. Claims with these properties are always fun to think about, so I've left this one as a problem (Problem 9.2). Try to solve it on your own. ■

We will invoke this theorem repeatedly in this chapter (often without bothering to say so). Note that we are assuming that P is a point in the body, because we used the fact that P keeps the same distances from other points in the body.

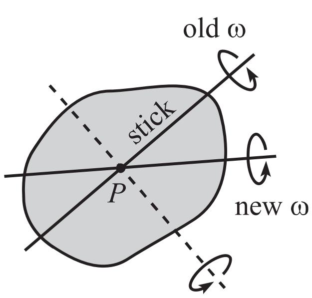

Fig. 9.2

REMARK: A situation where this theorem isn't so obvious is the following (this setup contains only rotation, with no translation of the point P). Consider an object rotating around a fixed axis, the stick shown in Fig. 9.2. But now imagine grabbing the stick and rotating it around some other axis (the dotted line shown). It isn't immediately obvious that the resulting motion is (instantaneously) a rotation around some new axis through the point P (which remains fixed). But indeed it is. We'll be quantitative about this in the "Rotating sphere" example later in this section. ♣

9.1.2 The angular velocity vector

It is extremely useful to introduce the angular velocity vector, $\boldsymbol{\omega}$, which is defined as the vector that points along the axis of rotation, and whose magnitude equals

[2] This claim is actually true for *any* transformation of a rigid sphere into itself, but for the present purposes we are concerned only with infinitesimal transformations, because we are looking only at what happens at a given instant in time.

the angular speed. The choice of the two possible directions along the axis is given by the right-hand rule: if you curl your right-hand fingers in the direction of the spin, then your thumb points in the direction of $\boldsymbol{\omega}$. For example, a spinning record has $\boldsymbol{\omega}$ perpendicular to the record, through the center (as shown in Fig. 9.3),[3] with its magnitude equal to the angular speed, ω. The points on the axis of rotation are the ones that (instantaneously) do not move. Of course, the direction of $\boldsymbol{\omega}$ may change over time, so the points that were formerly on the axis may now be moving.

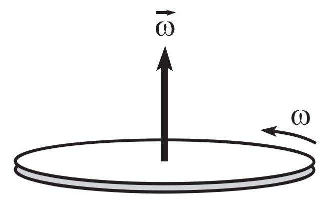

Fig. 9.3

REMARKS:

1. If you want, you can break the mold and use the left-hand rule to determine $\boldsymbol{\omega}$, as long as you use it consistently. The direction of $\boldsymbol{\omega}$ will be the opposite, but that doesn't matter, because $\boldsymbol{\omega}$ isn't really physical. Any physical result (for example, the velocity of a particle, given below in Theorem 9.2) will come out the same, independent of which hand you (consistently) use.

   When studying vectors in school,
   You'll use your right hand as a tool.
   But look in a mirror,
   And then you'll see clearer,
   It's just like the left-handed rule.

2. The fact that we can specify a rotation by specifying a vector $\boldsymbol{\omega}$ is a peculiarity to three dimensions. If we lived in one dimension, then there would be no such thing as a rotation. If we lived in two dimensions, then all rotations would take place in that plane, so we could label a rotation by simply giving its speed, ω. In three dimensions, rotations take place in $\binom{3}{2} = 3$ independent planes. And we choose to label these, for convenience, by the directions orthogonal to these planes, and by the angular speed in each plane. If we lived in four dimensions, then rotations could take place in $\binom{4}{2} = 6$ planes, so we would have to label a rotation by giving 6 planes and 6 angular speeds. Note that a vector, which has four components in four dimensions, would not do the trick. ♣

In addition to specifying the points that are instantaneously motionless, $\boldsymbol{\omega}$ also easily produces the velocity of any point in the rotating object. Consider the situation where the axis of rotation passes through the origin, which we'll generally assume to be the case in this chapter, unless otherwise stated. Then we have the following theorem.

Theorem 9.2 *Given an object rotating with angular velocity* $\boldsymbol{\omega}$*, the velocity of a point at position* $\mathbf{r}$ *is given by*

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}. \tag{9.1}$$

Proof: Drop a perpendicular from the point in question (call it P) to the axis $\boldsymbol{\omega}$. Let Q be the foot of the perpendicular, and let $\mathbf{r}'$ be the vector from Q to P

[3] It's actually meaningless to say that $\boldsymbol{\omega}$ passes through the center of the record, because you can draw the vector anywhere, and it's still the same vector, as long as it has the correct magnitude and direction. Nevertheless, it's customary to draw $\boldsymbol{\omega}$ along the axis of rotation and to say things like, "An object rotates around $\boldsymbol{\omega}$ . . ."

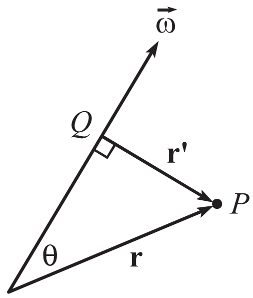

Fig. 9.4

(see Fig. 9.4). From the properties of the cross product (see Appendix B), $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$ is orthogonal to $\boldsymbol{\omega}$, $\mathbf{r}$, and also $\mathbf{r}'$ because $\mathbf{r}'$ is a linear combination of $\boldsymbol{\omega}$ and $\mathbf{r}$. Therefore, the direction of $\mathbf{v}$ is correct; it is always orthogonal to $\boldsymbol{\omega}$ and $\mathbf{r}'$, so it describes circular motion around the axis $\boldsymbol{\omega}$. Also, by the right-hand rule in the cross product (or the left-hand rule, if you had chosen to be different and defined $\boldsymbol{\omega}$ that way), $\mathbf{v}$ has the proper orientation around $\boldsymbol{\omega}$, namely into the page at the instant shown. And since

$$|\mathbf{v}| = |\boldsymbol{\omega}||\mathbf{r}| \sin\theta = \omega r', \tag{9.2}$$

we see that $\mathbf{v}$ has the correct magnitude, because $\omega r'$ is the speed of the circular motion around $\boldsymbol{\omega}$. So $\mathbf{v}$ is indeed the correct velocity vector. (If we have the special case where P lies along $\boldsymbol{\omega}$, then $\mathbf{r}$ is parallel to $\boldsymbol{\omega}$, so the cross product gives a zero result for $\mathbf{v}$, as it should.) ■

We'll make good use of Eq. (9.1) and apply it repeatedly throughout this chapter. Even if it's hard to visualize what's going on in a given rotation, all you have to do to find the speed of any point is calculate the cross product $\boldsymbol{\omega} \times \mathbf{r}$. Conversely, if the speed of every point in a body is given by $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, then the body must be undergoing a rotation with angular velocity $\boldsymbol{\omega}$, because all points on the axis $\boldsymbol{\omega}$ are motionless, and all other points move with the proper speed for this rotation.

A very nice thing about angular velocities is that they simply add. Stated more precisely:

Theorem 9.3 *Let coordinate systems S_1, S_2, and S_3 have a common origin. Let S_1 rotate with angular velocity $\boldsymbol{\omega}_{1,2}$ with respect to S_2, and let S_2 rotate with angular velocity $\boldsymbol{\omega}_{2,3}$ with respect to S_3. Then S_1 rotates (instantaneously) with angular velocity*

$$\boldsymbol{\omega}_{1,3} = \boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3} \tag{9.3}$$

with respect to S_3.

Proof: If $\boldsymbol{\omega}_{1,2}$ and $\boldsymbol{\omega}_{2,3}$ point in the same direction, then the theorem is clear; the angular speeds just add. If, however, they don't point in the same direction, then things are a bit harder to visualize. But we can prove the theorem by making abundant use of the definition of $\boldsymbol{\omega}$.

Pick a point P_1 at rest in S_1. Let $\mathbf{r}$ be the vector from the origin to P_1. The velocity of P_1 (relative to a very close point P_2 at rest in S_2) due to the rotation of S_1 around $\boldsymbol{\omega}_{1,2}$ is $\mathbf{V}_{P_1P_2} = \boldsymbol{\omega}_{1,2} \times \mathbf{r}$. The velocity of P_2 (relative to a very close point P_3 at rest in S_3) due to the rotation of S_2 around $\boldsymbol{\omega}_{2,3}$ is $\mathbf{V}_{P_2P_3} = \boldsymbol{\omega}_{2,3} \times \mathbf{r}$, because P_2 is also located essentially at position $\mathbf{r}$. Therefore, the velocity of P_1 relative to P_3 is $\mathbf{V}_{P_1P_2} + \mathbf{V}_{P_2P_3} = (\boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3}) \times \mathbf{r}$. This holds for any point P_1 at rest in S_1, so the frame S_1 rotates with angular velocity $(\boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3})$ with respect to S_3. We see that the proof basically comes down to the facts that (1) the linear velocities just add, as usual, and (2) the angular velocities differ from the linear velocities by a cross product with $\mathbf{r}$. ■

If $\boldsymbol{\omega}_{1,2}$ is constant in S_2, then the vector $\boldsymbol{\omega}_{1,3} = \boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3}$ will change with respect to S_3 as time goes by, because $\boldsymbol{\omega}_{1,2}$, which is fixed in S_2, is changing with respect to S_3 (assuming that $\boldsymbol{\omega}_{1,2}$ and $\boldsymbol{\omega}_{2,3}$ aren't parallel). But at any instant, $\boldsymbol{\omega}_{1,3}$ may be obtained by adding the present values of $\boldsymbol{\omega}_{1,2}$ and $\boldsymbol{\omega}_{2,3}$. Consider the following example.

Example (Rotating sphere): A sphere rotates with angular speed ω_3 around a stick that initially points in the $\hat{\mathbf{z}}$ direction. You grab the stick and rotate it around the $\hat{\mathbf{y}}$ axis with angular speed ω_2. What is the angular velocity of the sphere, with respect to the lab frame, as time goes by?

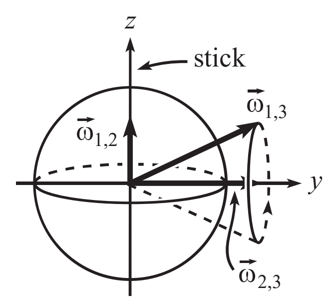

Fig. 9.5

Solution: In the language of Theorem 9.3, the sphere defines the S_1 frame; the stick and the $\hat{\mathbf{y}}$ axis define the S_2 frame; and the lab frame is the S_3 frame. The instant after you grab the stick, we are given that $\boldsymbol{\omega}_{1,2} = \omega_3\hat{\mathbf{z}}$, and $\boldsymbol{\omega}_{2,3} = \omega_2\hat{\mathbf{y}}$. Therefore, the angular velocity of the sphere with respect to the lab frame is $\boldsymbol{\omega}_{1,3} = \boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3} = \omega_3\hat{\mathbf{z}} + \omega_2\hat{\mathbf{y}}$, as shown in Fig. 9.5. Convince yourself that the combination of these two rotations yields zero motion for the points along the line of $\boldsymbol{\omega}_{1,3}$. As time goes by, the stick (and hence $\boldsymbol{\omega}_{1,2}$) rotates around the $\mathbf{y}$ axis, so $\boldsymbol{\omega}_{1,3} = \boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3}$ traces out a cone around the $\mathbf{y}$ axis, as shown.

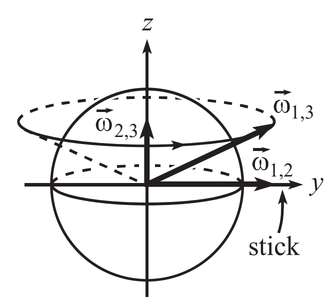

Fig. 9.6

REMARK: Note the different behavior of $\boldsymbol{\omega}_{1,3}$ for a slightly different statement of the problem: Let the sphere initially rotate with angular velocity $\omega_2\hat{\mathbf{y}}$ around a stick, and then grab the stick and rotate it with angular velocity $\omega_3\hat{\mathbf{z}}$. For this situation, $\boldsymbol{\omega}_{1,3}$ initially points in the same direction as in the original statement of the problem (it initially equals $\omega_2\hat{\mathbf{y}} + \omega_3\hat{\mathbf{z}}$). But as time goes by, it is now the horizontal component (defined by the stick) of $\boldsymbol{\omega}_{1,3}$ that changes, so $\boldsymbol{\omega}_{1,3} = \boldsymbol{\omega}_{1,2} + \boldsymbol{\omega}_{2,3}$ traces out a cone around the $\mathbf{z}$ axis, as shown in Fig. 9.6. ♣

An important point concerning rotations is that they are defined with respect to a *coordinate system*. It makes no sense to ask how fast an object is rotating with respect to a certain point, or even a certain axis. Consider, for example, an object rotating with angular velocity $\boldsymbol{\omega} = \omega_3\hat{\mathbf{z}}$ with respect to the lab frame. Saying only, "The object has angular velocity $\boldsymbol{\omega} = \omega_3\hat{\mathbf{z}}$," is not sufficient, because someone standing in the frame of the object would measure $\boldsymbol{\omega} = 0$, and would therefore be very confused by your statement. Throughout this chapter, we'll try to remember to state the coordinate system with respect to which $\boldsymbol{\omega}$ is measured. But if we forget, the default frame is the lab frame.

This section was definitely a bit abstract, so don't worry too much about it at the moment. The best strategy is perhaps to read on, and then come back for a second pass after digesting a few more sections. At any rate, we'll be discussing many other aspects (probably more than you'd ever want to know) of $\boldsymbol{\omega}$ in Section 9.7.2, so you're assured of getting a lot more practice with it. For now, if you want to strain some brain cells thinking about $\boldsymbol{\omega}$ vectors, you are encouraged to solve Problem 9.3, and also to look at the three given solutions.

9.2 The inertia tensor

Given an object undergoing general motion, the *inertia tensor* is what relates the angular momentum, $\mathbf{L}$, to the angular velocity, $\boldsymbol{\omega}$. This tensor (which is just a fancy name for "matrix" in this context) depends on the geometry of the object, as we'll see. In finding the $\mathbf{L}$ due to general motion, we'll follow the strategy of Section 8.1. We'll first look at the special case of rotation around an axis through the origin, then we'll look at the most general possible motion.

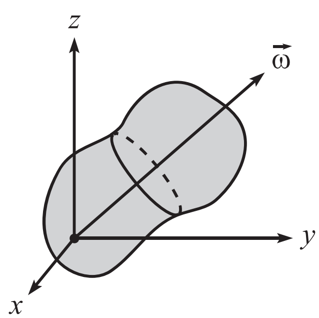

Fig. 9.7

9.2.1 Rotation around an axis through the origin

The three-dimensional object in Fig. 9.7 rotates with angular velocity $\boldsymbol{\omega}$. Consider a little piece of the body, with mass dm and position $\mathbf{r}$. The velocity of this piece is $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, so its angular momentum (relative to the origin) is $\mathbf{r} \times \mathbf{p} = (dm)\mathbf{r} \times \mathbf{v} = (dm)\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})$. The angular momentum of the entire body is therefore

$$\mathbf{L} = \int \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})\, dm, \tag{9.4}$$

where the integration runs over the volume of the body. In the case where the rigid body is made up of a collection of point masses m_i, the angular momentum is

$$\mathbf{L} = \sum_i m_i \mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i). \tag{9.5}$$

The double cross product in Eqs. (9.4) and (9.5) looks a bit intimidating, but it's actually not so bad. First, we have

$$\begin{aligned}\boldsymbol{\omega} \times \mathbf{r} &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \omega_1 & \omega_2 & \omega_3 \\ x & y & z \end{vmatrix} \\ &= (\omega_2 z - \omega_3 y)\hat{\mathbf{x}} + (\omega_3 x - \omega_1 z)\hat{\mathbf{y}} + (\omega_1 y - \omega_2 x)\hat{\mathbf{z}}. \end{aligned} \tag{9.6}$$

We're using the notation ω_1 instead of ω_x, etc., because there are already enough x, y, z's floating around here. The double cross product is then

$$\begin{aligned}\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ x & y & z \\ (\omega_2 z - \omega_3 y) & (\omega_3 x - \omega_1 z) & (\omega_1 y - \omega_2 x) \end{vmatrix} \\ &= \Big(\omega_1 (y^2 + z^2) - \omega_2 xy - \omega_3 zx\Big)\hat{\mathbf{x}} \end{aligned}$$

$$+\Big(\omega_2(z^2+x^2)-\omega_3 yz-\omega_1 xy\Big)\hat{\mathbf{y}}$$
$$+\Big(\omega_3(x^2+y^2)-\omega_1 zx-\omega_2 yz\Big)\hat{\mathbf{z}}. \tag{9.7}$$

The angular momentum in Eq. (9.4) may therefore be written in the concise matrix form,

$$\begin{aligned}\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} &= \begin{pmatrix} \int(y^2+z^2) & -\int xy & -\int zx \\ -\int xy & \int(z^2+x^2) & -\int yz \\ -\int zx & -\int yz & \int(x^2+y^2) \end{pmatrix}\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\ &\equiv \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\ &\equiv \mathbf{I}\boldsymbol{\omega}. \end{aligned} \tag{9.8}$$

For the sake of clarity, we have not bothered to write the *dm* part of each integral (and we'll continue to drop it for most of the remainder of this section). The matrix **I** is called the *inertia tensor*. If the word "tensor" scares you, just ignore it. **I** is simply a matrix. It acts on a vector (the angular velocity) and produces another vector (the angular momentum).

Example (Cube with origin at corner): Calculate the inertia tensor for a solid cube of mass M and side length L, with the coordinate axes parallel to the edges of the cube, and the origin at a corner (see Fig. 9.8).

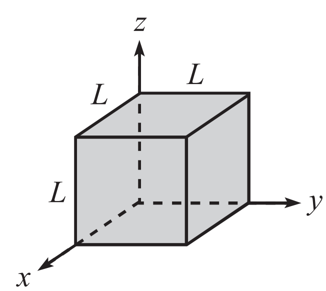

Fig. 9.8

Solution: Due to the symmetry of the cube, there are only two integrals we need to calculate in Eq. (9.8). The diagonal entries are all equal to $\int(y^2+z^2)\,dm$, and the off-diagonal entries are all equal to $-\int xy\,dm$. With $dm = \rho\,dx\,dy\,dz$, and $\rho = M/L^3$, these two integrals are

$$\int_0^L\int_0^L\int_0^L (y^2+z^2)\rho\,dx\,dy\,dz = \rho L^2\int_0^L y^2\,dy + \rho L^2\int_0^L z^2\,dz = \frac{2}{3}ML^2,$$
$$-\int_0^L\int_0^L\int_0^L xy\rho\,dx\,dy\,dz = -\rho L\int_0^L x\,dx\int_0^L y\,dy = -\frac{ML^2}{4}. \tag{9.9}$$

Therefore,

$$\mathbf{I} = ML^2\begin{pmatrix} 2/3 & -1/4 & -1/4 \\ -1/4 & 2/3 & -1/4 \\ -1/4 & -1/4 & 2/3 \end{pmatrix}. \tag{9.10}$$

Having found **I**, we can calculate the angular momentum associated with any given angular velocity. If, for example, the cube is rotating around the z axis with angular speed ω, then we can apply the matrix **I** to the vector $(0, 0, \omega)$ to find that the angular momentum is $\mathbf{L} = ML^2\omega(-1/4, -1/4, 2/3)$. Note the somewhat odd fact that L_x and L_y are nonzero, even though the rotation is only around the z axis. We'll discuss this issue after the following remarks.

Remarks:

1. The inertia tensor in Eq. (9.8) is a rather formidable-looking object. You will therefore be very pleased to hear that you rarely have to use it. It's nice to know that it's there if you need it, but the concept of *principal axes* (discussed in Section 9.3) provides a way to avoid using the inertia tensor (or more precisely, to greatly simplify it) and is therefore much more useful in solving problems.
2. $\mathbf{I}$ is a symmetric matrix, which is a fact that will be important in Section 9.3. There are therefore only six independent entries, instead of nine.
3. In the case where the rigid body is made up of a collection of point masses m_i, the entries in the matrix are just sums. For example, the upper left entry is $\sum m_i(y_i^2 + z_i^2)$.
4. $\mathbf{I}$ depends only on the geometry of the object, and not on $\boldsymbol{\omega}$.
5. To construct an $\mathbf{I}$, you not only need to specify the origin, you also need to specify the x, y, z axes of your coordinate system. And the basis vectors must be orthogonal, because the cross product calculation above is valid only for an orthonormal basis. If someone else comes along and chooses a different orthonormal basis (but the same origin), then her $\mathbf{I}$ will have different *entries*, as will her $\boldsymbol{\omega}$, as will her $\mathbf{L}$. But her $\boldsymbol{\omega}$ and $\mathbf{L}$ will be exactly the same *vectors* as your $\boldsymbol{\omega}$ and $\mathbf{L}$. They will appear different only because they are written in a different coordinate system. A vector is what it is, independent of how you choose to look at it. If you each point your arm in the direction of what you calculate $\mathbf{L}$ to be, then you will both be pointing in the same direction.
6. For the case of a pancake object rotating in the x-y plane, we have $z = 0$ for all points in the object. And $\boldsymbol{\omega} = \omega_3\hat{\mathbf{z}}$, so $\omega_1 = \omega_2 = 0$. The only nonzero term in the $\mathbf{L}$ in Eq. (9.8) is therefore $L_3 = \int (x^2 + y^2)\, dm\, \omega_3$, which is simply the $L_z = I_z\omega$ result we found in Eq. (8.5). ♣

This is all perfectly fine. Given any rigid body, we can calculate $\mathbf{I}$ (relative to a given origin, using a given set of axes). And given $\boldsymbol{\omega}$, we can then apply $\mathbf{I}$ to it to find $\mathbf{L}$. But what do these entries in $\mathbf{I}$ really mean? How do we interpret them? Note, for example, that ω_3 appears not only in L_3 in Eq. (9.8), but also in L_1 and L_2. But ω_3 is relevant to rotations around the z axis, so what in the world is it doing in L_1 and L_2? Consider the following examples.

Example 1 (Point mass in the *x*-*y* plane): Consider a point mass m traveling in a circle of radius r (centered at the origin) in the x-y plane, with frequency ω_3, as shown in Fig. 9.9. Using $\boldsymbol{\omega} = (0, 0, \omega_3)$, $x^2 + y^2 = r^2$, and $z = 0$ in Eq. (9.8) (with a discrete sum of only one object, instead of the integrals), the angular momentum with respect to the origin is

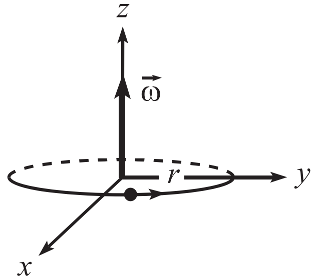

Fig. 9.9

$$\mathbf{L} = (0, 0, mr^2\omega_3). \tag{9.11}$$

The z component is $mr(r\omega_3) = mrv$, as it should be. And the x and y components are zero, as they should be. This case where $\omega_1 = \omega_2 = 0$ and $z = 0$ is simply the case we studied in Chapter 8, as mentioned in Remark 6 above.

Example 2 (Point mass in space): Consider a point mass m traveling in a circle of radius r, with frequency ω_3. But now let the circle be centered at the point $(0, 0, z_0)$, with the plane of the circle parallel to the x-y plane, as shown in Fig. 9.10. Using $\boldsymbol{\omega} = (0, 0, \omega_3)$, $x^2 + y^2 = r^2$, and $z = z_0$ in Eq. (9.8), the angular momentum with respect to the origin is

$$\mathbf{L} = m\omega_3(-xz_0,\ -yz_0,\ r^2). \tag{9.12}$$

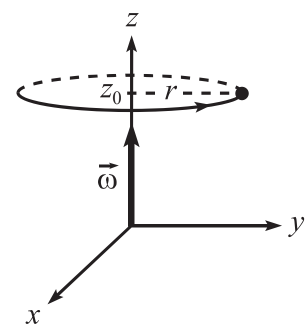

Fig. 9.10

The z component is mrv, as it should be. But surprisingly, we have nonzero L_1 and L_2, even though the mass is just rotating around the z axis. $\mathbf{L}$ does *not* point along $\boldsymbol{\omega}$ here. What's going on?

Consider an instant when the mass is in the y-z plane, as shown in Fig. 9.10. The velocity of the mass is then in the $-\hat{\mathbf{x}}$ direction. Therefore, the particle most certainly has angular momentum around the y axis, as well as the z axis. Someone looking at a split-second movie of the mass at this point can't tell whether it's rotating around the y axis, the z axis, or undergoing some complicated motion. But the past and future motion is irrelevant; at any instant in time, as far as the angular momentum goes, we are concerned only with what is happening at this instant.

At this instant, the angular momentum around the y axis is $L_2 = -mz_0v$, because z_0 is the distance from the y axis, and the minus sign comes from the right-hand rule. Using $v = \omega_3 r = \omega_3 y$, we have $L_2 = -mz_0\omega_3 y$, in agreement with Eq. (9.12). Also, at this instant, L_1 is zero, because the velocity is parallel to the x axis. This agrees with Eq. (9.12), since $x = 0$. As an exercise, you can check that Eq. (9.12) is also correct when the mass is at a general point (x, y, z_0).

We see that, for example, the $I_{yz} \equiv -\int yz$ entry in $\mathbf{I}$ tells us how much the ω_3 component of the angular velocity contributes to the L_2 component of the angular momentum. And due to the symmetry of $\mathbf{I}$, the $I_{yz} = I_{zy}$ entry in $\mathbf{I}$ also tells us how much the ω_2 component of the angular velocity contributes to the L_3 component of the angular momentum. In the former case, if we group the product of the various quantities as $-\int(\omega_3 y)z$, we see that this is simply the appropriate component of the velocity times the distance from the y axis. In the latter case with $-\int(\omega_2 z)y$, it is the opposite grouping. But in both cases there is one factor of y and one factor of z, hence the symmetry in $\mathbf{I}$.

Remark: For a point mass, $\mathbf{L}$ is actually more easily obtained by just calculating $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. The result for the instant shown in Fig. 9.10 is drawn in Fig. 9.11, where it is clear that $\mathbf{L}$ has both y and z components, and thus also clear that $\mathbf{L}$ doesn't point along $\boldsymbol{\omega}$. For a more complicated object, the tensor $\mathbf{I}$ is generally used, because it is necessary to perform the integral of the $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ contributions over the entire object, and the tensor has this integral built into it. At any rate, whatever method you use, you will find that except in special circumstances (see Section 9.3), $\mathbf{L}$ doesn't point along $\boldsymbol{\omega}$.

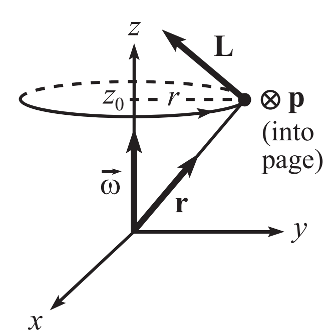

Fig. 9.11

Consider the vector of $\mathbf{L}$,
And that of $\boldsymbol{\omega}$ as well.
The erroneous claim
That they must aim the same
Is a view that you've got to dispel! ♣

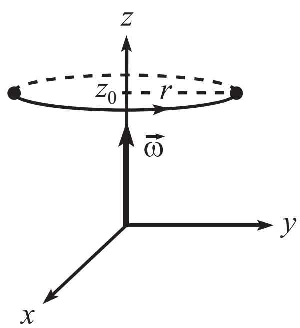

Fig. 9.12

Example 3 (Two point masses): Let's now add another point mass m to the previous example. Let it travel in the same circle, at the diametrically opposite point, as shown in Fig. 9.12. Using $\omega = (0, 0, \omega_3)$, $x^2 + y^2 = r^2$, and $z = z_0$ in Eq. (9.8), you can show that the angular momentum with respect to the origin is

$$\mathbf{L} = 2m\omega_3(0, 0, r^2). \tag{9.13}$$

Since $v = \omega_3 r$, the z component is $2mrv$, as it should be. And L_1 and L_2 are zero, unlike in the previous example, because these components of the $\mathbf{L}$'s of the two particles cancel. This occurs because of the symmetry of the masses around the z axis, which causes the I_{zx} and I_{zy} entries in the inertia tensor to vanish; they are each the sum of two terms, with opposite x values, or opposite y values. Alternatively, you can just note that adding on the mirror-image $\mathbf{L}$ vector in Fig. 9.10 produces canceling x and y components.

Let's now look at the kinetic energy of our object, which is rotating around an axis passing through the origin. To find this, we must add up the kinetic energies of all the little pieces. A little piece has energy $(dm)\, v^2/2 = dm\, |\boldsymbol{\omega} \times \mathbf{r}|^2/2$. Therefore, using Eq. (9.6), the total kinetic energy is

$$T = \frac{1}{2}\int \Big((\omega_2 z - \omega_3 y)^2 + (\omega_3 x - \omega_1 z)^2 + (\omega_1 y - \omega_2 x)^2 \Big)\, dm. \tag{9.14}$$

Multiplying this out, we see (after a little work) that we can write T as

$$\begin{aligned} T &= \frac{1}{2}(\omega_1, \omega_2, \omega_3) \cdot \begin{pmatrix} \int (y^2 + z^2) & -\int xy & -\int zx \\ -\int xy & \int (z^2 + x^2) & -\int yz \\ -\int zx & -\int yz & \int (x^2 + y^2) \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\ &= \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{I}\boldsymbol{\omega} = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{L}. \end{aligned} \tag{9.15}$$

If $\boldsymbol{\omega} = \omega_3 \hat{\mathbf{z}}$, then this reduces to $T = I_{zz}\omega_3^2/2$, which agrees with the result in Eq. (8.8), with a slight change in notation.

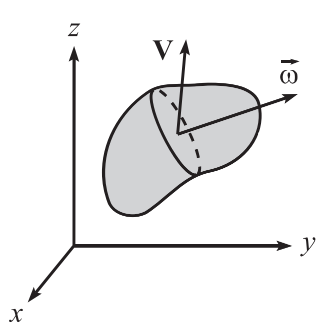

Fig. 9.13

9.2.2 General motion

How do we deal with general motion in space? That is, what if an object is both translating and rotating? For the motion in Fig. 9.13, the various pieces of mass aren't traveling in circles around the origin, so we can't write $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, as we did prior to Eq. (9.4).

To determine $\mathbf{L}$ (relative to the origin), and also the kinetic energy T, we will use Theorem 9.1 to write the motion as the sum of a translation plus a rotation. In applying the theorem, we may choose any point in the body to be the point P in the theorem. However, only in the case where P is the object's CM can we extract anything useful, as we'll see. The theorem then says that the motion of the body is the sum of the motion of the CM plus a rotation around the CM. So let the CM move with velocity $\mathbf{V}$, and let the body instantaneously rotate with angular velocity $\boldsymbol{\omega}'$ around the CM (that is, with respect to the frame whose origin is the CM, and whose axes are parallel to the fixed-frame axes).[4]

Let the position of the CM relative to the origin be $\mathbf{R} = (X, Y, Z)$, and let the position of a given piece of mass relative to the CM be $\mathbf{r}' = (x', y', z')$. Then $\mathbf{r} = \mathbf{R} + \mathbf{r}'$ is the position of a piece of mass relative to the origin (see Fig. 9.14). Let the velocity of a piece of mass relative to the CM be $\mathbf{v}'$ (so $\mathbf{v}' = \boldsymbol{\omega}' \times \mathbf{r}'$). Then $\mathbf{v} = \mathbf{V} + \mathbf{v}'$ is the velocity relative to the origin.

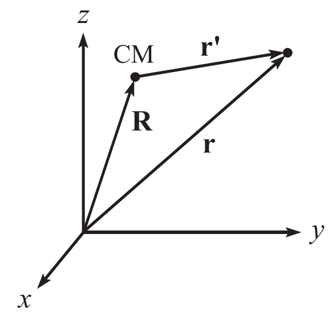

Fig. 9.14

Let's look at $\mathbf{L}$ first. The angular momentum is

$$\begin{aligned} \mathbf{L} &= \int \mathbf{r} \times \mathbf{v}\, dm = \int (\mathbf{R} + \mathbf{r}') \times \Big(\mathbf{V} + (\boldsymbol{\omega}' \times \mathbf{r}')\Big)\, dm \\ &= \int (\mathbf{R} \times \mathbf{V})\, dm + \int \mathbf{r}' \times (\boldsymbol{\omega}' \times \mathbf{r}')\, dm \\ &= M(\mathbf{R} \times \mathbf{V}) + \mathbf{L}_{\text{CM}}, \end{aligned} \tag{9.16}$$

where the cross terms vanish because the integrands are linear in $\mathbf{r}'$. More precisely, the integrals involve $\int \mathbf{r}'\, dm$, which is zero by definition of the CM (because $\int \mathbf{r}'\, dm/M$ is the position of the CM relative to the CM, which is zero). $\mathbf{L}_{\text{CM}}$ is the angular momentum relative to the CM.[5]

We see that as in the pancake case in Section 8.1.2, the angular momentum (relative to the origin) of a body can be found by treating the body like a point mass located at the CM and finding the angular momentum of this point mass relative to the origin, and by then adding on the angular momentum of the body relative to the CM. Note that these two parts of the angular momentum need not point in the same direction, as they did in the case of the pancake moving in the x-y plane.

Now let's look at T. The kinetic energy is

$$\begin{aligned} T &= \int \frac{1}{2} v^2\, dm = \int \frac{1}{2} |\mathbf{V} + \mathbf{v}'|^2\, dm \\ &= \int \frac{1}{2} V^2\, dm + \int \frac{1}{2} v'^2\, dm \\ &= \frac{1}{2} M V^2 + \int \frac{1}{2} |\boldsymbol{\omega}' \times \mathbf{r}'|^2\, dm \\ &\equiv \frac{1}{2} M V^2 + \frac{1}{2} \boldsymbol{\omega}' \cdot \mathbf{L}_{\text{CM}}, \end{aligned} \tag{9.17}$$

[4] It's not necessary to put the prime on the $\boldsymbol{\omega}$ here, because the angular velocity vector in the CM frame is the same as in the lab frame. But we'll use the prime just because we'll have primes on the other CM quantities below.

[5] By this, we mean the angular momentum as measured in the coordinate system whose origin is the CM, and whose axes are parallel to the fixed-frame axes.

where the last line follows from the steps leading to Eq. (9.15). The cross term $\int \mathbf{V} \cdot \mathbf{v}'\, dm = \int \mathbf{V} \cdot (\boldsymbol{\omega}' \times \mathbf{r}')\, dm$ vanishes because the integrand is linear in $\mathbf{r}'$ and thus yields a zero integral, by definition of the CM. As in the pancake case in Section 8.1.2, the kinetic energy of a body can be found by treating the body like a point mass located at the CM, and by then adding on the kinetic energy of the body due to the rotation around the CM.

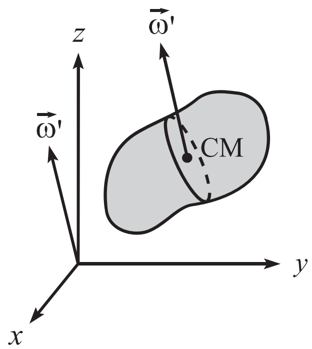

Fig. 9.15

9.2.3 The parallel-axis theorem

Consider the special case where the CM rotates around the origin with the same angular velocity at which the body rotates around the CM (see Fig. 9.15), that is, $\mathbf{V} = \boldsymbol{\omega}' \times \mathbf{R}$. This can be achieved, for example, by piercing the body with the base of a rigid "T" and then rotating the T and the body around the (fixed) line of the "upper" part of the T (the origin must pass through this line). We then have the nice situation where all points in the body travel in fixed circles around the axis of rotation. Mathematically, this follows from $\mathbf{v} = \mathbf{V} + \mathbf{v}' = \boldsymbol{\omega}' \times \mathbf{R} + \boldsymbol{\omega}' \times \mathbf{r}' = \boldsymbol{\omega}' \times \mathbf{r}$. Dropping the prime on $\boldsymbol{\omega}$, Eq. (9.16) becomes

$$\mathbf{L} = M\mathbf{R} \times (\boldsymbol{\omega} \times \mathbf{R}) + \int \mathbf{r}' \times (\boldsymbol{\omega} \times \mathbf{r}')\, dm \tag{9.18}$$

Expanding the double cross products as in the steps leading to Eq. (9.8), we can write this as

$$\begin{aligned} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} &= M \begin{pmatrix} Y^2 + Z^2 & -XY & -ZX \\ -XY & Z^2 + X^2 & -YZ \\ -ZX & -YZ & X^2 + Y^2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\ &+ \begin{pmatrix} \int (y'^2 + z'^2) & -\int x'y' & -\int z'x' \\ -\int x'y' & \int (z'^2 + x'^2) & -\int y'z' \\ -\int z'x' & -\int y'z' & \int (x'^2 + y'^2) \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\ &\equiv (\mathbf{I}_{\mathrm{R}} + \mathbf{I}_{\mathrm{CM}})\boldsymbol{\omega}. \end{aligned} \tag{9.19}$$

This is the generalized parallel-axis theorem. It says that once you've calculated $\mathbf{I}_{\mathrm{CM}}$ relative to the CM, then if you want to calculate $\mathbf{I}$ relative to another point, you simply have to add on the $\mathbf{I}_{\mathrm{R}}$ matrix, obtained by treating the object like a point mass at the CM. So you have to compute six extra numbers (there are six, instead of nine, because the $\mathbf{I}_{\mathrm{R}}$ matrix is symmetric) instead of just the one MR^2 in the parallel-axis theorem in Chapter 8, given in Eq. (8.13). Problem 9.4 gives another derivation of the parallel-axis theorem, without mentioning the angular velocity.

Remark: The name "parallel-axis" theorem is actually a misnomer here. The inertia tensor isn't associated with one particular axis, as the moment of inertia in Chapter 8 was. The moment of inertia is just one of the diagonal entries (associated with a given axis) in the inertia tensor. The inertia tensor depends on the entire coordinate system. So in that sense we should call this the "parallel-*axes*" theorem, because the coordinate axes in the CM frame are assumed to be parallel to the ones in the fixed frame. At any rate, the point is that the parallel-axis theorem in Chapter 8 dealt with shifting the axis, whereas the present theorem deals with shifting the origin (and hence all three axes in general). ♣

As far as the kinetic energy goes, if $\boldsymbol{\omega}$ and $\boldsymbol{\omega}'$ are equal, so that $\mathbf{V} = \boldsymbol{\omega}' \times \mathbf{R}$, then Eq. (9.17) gives (dropping the prime on $\boldsymbol{\omega}$)

$$T = \frac{1}{2}M|\boldsymbol{\omega} \times \mathbf{R}|^2 + \int \frac{1}{2}|\boldsymbol{\omega} \times \mathbf{r}'|^2\, dm. \tag{9.20}$$

Performing the steps leading to Eq. (9.15), this becomes

$$T = \frac{1}{2}\boldsymbol{\omega} \cdot (\mathbf{I}_{\mathrm{R}} + \mathbf{I}_{\mathrm{CM}})\boldsymbol{\omega} = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{L}. \tag{9.21}$$

9.3 Principal axes

The cumbersome expressions in the previous section may seem a bit unsettling, but it turns out that usually we can get by without them. The strategy for avoiding all of the above mess is to use the *principal axes* of a body, which we will define below.

In general, the inertia tensor $\mathbf{I}$ in Eq. (9.8) has nine nonzero entries, of which six are independent due to the symmetry of $\mathbf{I}$. In addition to depending on the origin chosen, the inertia tensor depends on the set of orthonormal basis vectors chosen for the coordinate system; the x, y, z variables in the integrals in $\mathbf{I}$ depend, of course, on the coordinate system they're measured with respect to. Given a blob of material, and given an arbitrary origin,[6] any orthonormal set of basis vectors is usable, but there is one special set that makes all our calculations very nice. These special basis vectors are called the *principal axes*. They can be defined in various equivalent ways:

- The principal axes are the orthonormal basis vectors for which $\mathbf{I}$ is diagonal, that is, for which[7]

$$\mathbf{I} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}. \tag{9.22}$$

  I_1, I_2, and I_3 are called the *principal moments*. For many objects, it is quite obvious what the principal axes are. For example, consider a uniform rectangle in the x-y plane. Pick the origin to be the CM, and let the x and y axes be parallel to the sides. Then the

[6] The CM is often chosen to be the origin, but it need not be. There are principal axes associated with any origin.

[7] Technically, we should be writing I_{11} or I_{xx} instead of I_1, etc., in this matrix, because the one-index object I_1 looks like the component of a vector, not a matrix. But the two-index notation gets cumbersome, so we'll be sloppy and just use I_1, etc.

principal axes are clearly the x, y, and z axes, because all the off-diagonal elements in the inertia tensor in Eq. (9.8) vanish, by symmetry. For example, $I_{xy} \equiv -\int xy\,dm$ equals zero, because for every point (x, y) in the rectangle, there is a corresponding point $(-x,\ y)$, so the contributions to $\int xy\,dm$ cancel in pairs. Also, the integrals involving z are identically zero, because $z = 0$.

- A principal axis is an axis $\hat{\boldsymbol{\omega}}$ for which $\mathbf{I}\hat{\boldsymbol{\omega}} = I\hat{\boldsymbol{\omega}}$. That is, a principal axis is a special direction with the property that if $\boldsymbol{\omega}$ points along it, then so does $\mathbf{L}$. The principal axes of an object are then the orthonormal set of three vectors $\hat{\boldsymbol{\omega}}_1$, $\hat{\boldsymbol{\omega}}_2$, $\hat{\boldsymbol{\omega}}_3$ with the property that

$$\mathbf{I}\hat{\boldsymbol{\omega}}_1 = I_1\hat{\boldsymbol{\omega}}_1, \qquad \mathbf{I}\hat{\boldsymbol{\omega}}_2 = I_2\hat{\boldsymbol{\omega}}_2, \qquad \mathbf{I}\hat{\boldsymbol{\omega}}_3 = I_3\hat{\boldsymbol{\omega}}_3. \tag{9.23}$$

The three statements in Eq. (9.23) are equivalent to Eq. (9.22), because the vectors $\hat{\boldsymbol{\omega}}_1$, $\hat{\boldsymbol{\omega}}_2$, and $\hat{\boldsymbol{\omega}}_3$ are simply $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ in the frame in which they are the basis vectors.

- Consider an object rotating around a fixed axis with constant angular speed. Then this axis is a principal axis if there is no need for any torque. So in some sense, the object is "happy" to spin around a principal axis. A set of three orthonormal axes, each of which has this property, is by definition what we call a set of principal axes.

This definition of a principal axis is equivalent to the previous definition for the following reason. Assume that the object rotates around a fixed axis $\hat{\boldsymbol{\omega}}_1$ for which $\mathbf{L} = \mathbf{I}\hat{\boldsymbol{\omega}}_1 = I_1\hat{\boldsymbol{\omega}}_1$, as in Eq. (9.23). Then since $\hat{\boldsymbol{\omega}}_1$ is assumed to be fixed, we see that $\mathbf{L}$ is also fixed. Therefore, $\boldsymbol{\tau} = d\mathbf{L}/dt = \mathbf{0}$.

Conversely, if the object is rotating around a fixed axis $\hat{\boldsymbol{\omega}}_1$, and if $\boldsymbol{\tau} = d\mathbf{L}/dt = \mathbf{0}$, then we claim that $\mathbf{L}$ points along $\hat{\boldsymbol{\omega}}_1$ (that is, $\mathbf{L} = I_1\hat{\boldsymbol{\omega}}_1$). This is true because if $\mathbf{L}$ does *not* point along $\hat{\boldsymbol{\omega}}_1$, then imagine painting a dot on the object somewhere along the line of $\mathbf{L}$. A little while later, the dot will have rotated around the fixed vector $\hat{\boldsymbol{\omega}}_1$. But the line of $\mathbf{L}$ must always pass through the dot, because we could have rotated our axes around $\hat{\boldsymbol{\omega}}_1$ and started the process at a slightly later time (this argument relies on $\hat{\boldsymbol{\omega}}_1$ being fixed). Therefore, we see that $\mathbf{L}$ has changed, in contradiction to the assumption that $d\mathbf{L}/dt = \mathbf{0}$. Hence, $\mathbf{L}$ must in fact point along $\hat{\boldsymbol{\omega}}_1$.

For a rotation around a principal axis $\hat{\boldsymbol{\omega}}$, the lack of need for any torque means that if the object is pivoted at the origin, and if the origin is the only place where any force is applied (which implies that there is zero torque around it), then the object can undergo rotation with constant angular velocity $\boldsymbol{\omega}$. If you try to set up this scenario with a nonprincipal axis, it won't work.

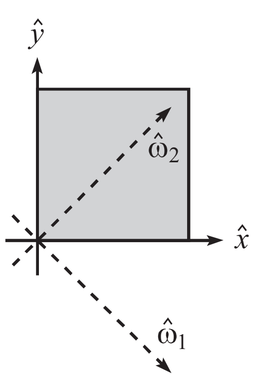

Fig. 9.16

Example (Square with origin at corner): Consider the uniform square in Fig. 9.16. In Appendix E, we show that the principal axes are the dotted lines drawn (and also the z axis perpendicular to the page). But there is no need to use the techniques in the appendix to see this, because in this new basis it is clear by symmetry that the integral $\int x_1x_2$ is zero; for every x_1 in the integral, there is a $-x_1$. And $x_3 \equiv z$

is identically zero, which makes all the other off-diagonal terms in **I** also equal to zero. Therefore, since **I** is diagonal in this new basis, these basis vectors are the principal axes.

Furthermore, it is intuitively clear that the square will be happy to rotate around any one of these axes indefinitely. During such a rotation, the pivot will certainly be applying a *force* (if the axis is $\hat{\boldsymbol{\omega}}_1$ or $\hat{\mathbf{z}}$, but not if it is $\hat{\boldsymbol{\omega}}_2$), to produce the centripetal acceleration of the CM in its circular motion. But it won't be applying a *torque* relative to the origin (because the **r** in $\mathbf{r} \times \mathbf{F}$ is **0**). This is good, because for a rotation around one of these principal axes, $d\mathbf{L}/dt = \mathbf{0}$, so there is no need for any torque.

In contrast with the off-diagonal zeros in the new basis, the integral $\int xy$ in the old basis is *not* zero, because every point gives a positive contribution. So the inertia tensor is not diagonal in the old basis, which means that $\hat{x}$ and $\hat{y}$ are not principal axes. Consistent with this, it is reasonably clear that it is impossible to make the square rotate around, say, the x axis, assuming that its only contact with the outside world is through a pivot (for example, a ball and socket) at the origin. The square simply doesn't want to remain in this circular motion. Mathematically, **L** (relative to the origin) doesn't point along the x axis, so it therefore precesses around the x axis along with the square, tracing out the surface of a cone. This means that **L** is changing. But there is no torque available (relative to the origin) to provide for this change in **L**. Hence, such a rotation cannot exist.

At the moment, it is not at all obvious that an orthonormal set of principal axes exists for an arbitrary object. But this is indeed the case, as stated in Theorem 9.4 below. Assuming for now that principal axes do exist, then in this basis the **L** and T in Eqs. (9.8) and (9.15) take on the particularly nice forms,

$$\mathbf{L} = (I_1\omega_1, I_2\omega_2, I_3\omega_3),$$
$$T = \frac{1}{2}\left(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2\right). \tag{9.24}$$

The quantities ω_1, ω_2, and ω_3 here are the components of a general vector $\boldsymbol{\omega}$ written in the principal-axis basis; that is, $\boldsymbol{\omega} = \omega_1\hat{\boldsymbol{\omega}}_1 + \omega_2\hat{\boldsymbol{\omega}}_2 + \omega_3\hat{\boldsymbol{\omega}}_3$. Equation (9.24) is a vast simplification over the general formulas in Eqs. (9.8) and (9.15). We will therefore invariably work with principal axes in the remainder of this chapter.

Note that the directions of the principal axes (relative to the body) depend only on the geometry of the body. They may therefore be considered to be painted on (or in) it. Hence, they will generally move around in space as the body rotates. For example, if the object is rotating around a principal axis, then that axis stays fixed while the other two principal axes rotate around it. In relations like $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$ and $\mathbf{L} = (I_1\omega_1, I_2\omega_2, I_3\omega_3)$, the components ω_i and $I_i\omega_i$ are measured along the *instantaneous* principal axes $\hat{\boldsymbol{\omega}}_i$. Since these axes change with time, it is quite possible that the components ω_i and $I_i\omega_i$ change with time, as we'll see in Section 9.5 (and onward).

Let's now state the theorem that implies that a set of principal axes does indeed exist for any body and any origin. The proof of this theorem involves a useful but rather slick technique, but it's slightly off the main line of thought, so we'll relegate it to Appendix D. Take a look at the proof if you wish, but if you want to just accept the fact that principal axes exist, that's fine.

Theorem 9.4 *Given a real symmetric* 3×3 *matrix,* $\mathbf{I}$*, there exist three orthonormal real vectors,* $\hat{\boldsymbol{\omega}}_k$*, and three real numbers,* I_k*, with the property that*

$$\mathbf{I}\hat{\boldsymbol{\omega}}_k = I_k \hat{\boldsymbol{\omega}}_k. \tag{9.25}$$

Proof: See Appendix D. ■

Since the inertia tensor in Eq. (9.8) is indeed symmetric for any body and any origin, this theorem says that we can always find three orthogonal basis vectors that satisfy Eq. (9.23). Or equivalently, we can always find three orthogonal basis vectors for which $\mathbf{I}$ is a diagonal matrix, as in Eq. (9.22). In other words, principal axes always exist. Problem 9.7 gives another way to demonstrate the existence of principal axes in the special case of a pancake object.

Invariably, it is best to work in a coordinate system that has principal axes as its basis, due to the simplicity of Eq. (9.24). And as mentioned in Footnote 6, the origin is generally chosen to be the CM, because from Section 8.4.3 the CM is one of the origins for which $\boldsymbol{\tau} = d\mathbf{L}/dt$ is a valid statement. But this choice is not necessary; there are principal axes associated with any origin.

For an object with a fair amount of symmetry, the principal axes are usually the obvious choices and can be written down by simply looking at the object (examples are given below). If, however, you are given an unsymmetrical body, then the only way to determine the principal axes is to pick an arbitrary basis, then find $\mathbf{I}$ in this basis, and then go through a diagonalization procedure. This diagonalization procedure basically consists of the steps at the beginning of the proof of Theorem 9.4 (given in Appendix D), with the addition of one more step to get the actual vectors, so we'll relegate it to Appendix E. There's no need to worry much about this method. Virtually every system you encounter will involve an object with sufficient symmetry to enable you to just write down the principal axes.

Let's now prove two very useful (and very similar) theorems.

Theorem 9.5 *If two principal moments are equal* ($I_1 = I_2 \equiv I$)*, then any axis (through the chosen origin) in the plane of the corresponding principal axes is a principal axis, and its moment is also* I*. Similarly, if all three principal moments are equal* ($I_1 = I_2 = I_3 \equiv I$)*, then any axis (through the chosen origin) in space is a principal axis, and its moment is also* I*.*

Proof: The first part was already proved at the end of the proof in Appendix D, but we'll do it again here. Since $I_1 = I_2 \equiv I$, we have $\mathbf{I}\mathbf{u}_1 = I\mathbf{u}_1$, and $\mathbf{I}\mathbf{u}_2 = I\mathbf{u}_2$, where the $\mathbf{u}$'s are the principal axes. Hence, $\mathbf{I}(a\mathbf{u}_1 + b\mathbf{u}_2) = I(a\mathbf{u}_1 + b\mathbf{u}_2)$, for any a and b. Therefore, any linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$ (that is, any vector in the plane spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$) is a solution to $\mathbf{I}\mathbf{u} = I\mathbf{u}$ and is thus a principal axis, by definition.

The proof of the second part proceeds in a similar manner. Since $I_1 = I_2 = I_3 \equiv I$, we have $\mathbf{I}\mathbf{u}_1 = I\mathbf{u}_1$, $\mathbf{I}\mathbf{u}_2 = I\mathbf{u}_2$, and $\mathbf{I}\mathbf{u}_3 = I\mathbf{u}_3$. Hence, $\mathbf{I}(a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3) = I(a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3)$. Therefore, any linear combination of $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ (that is, any vector in space) is a solution to $\mathbf{I}\mathbf{u} = I\mathbf{u}$ and is thus a principal axis, by definition.

In short, if $I_1 = I_2 \equiv I$, then $\mathbf{I}$ is the identity matrix (up to a multiple) in the space spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. And if $I_1 = I_2 = I_3 \equiv I$, then $\mathbf{I}$ is the identity matrix (up to a multiple) in the entire space. Note that it isn't required that the various $\mathbf{u}_i$ vectors be orthogonal. All we need is that they span the relevant space. ■

If two or three moments are equal, so that there is freedom in choosing the principal axes, then it is possible to pick a nonorthogonal group of them. We will, however, always choose ones that are orthogonal. So when we say "a set of principal axes," we mean an orthonormal set.

Theorem 9.6 *If a pancake object is symmetric under a rotation through an angle $\theta \neq 180°$ in the x-y plane (such as a hexagon), then every axis in the x-y plane (with the origin chosen to be the center of the symmetry rotation) is a principal axis with the same moment.*

Proof: Let $\hat{\boldsymbol{\omega}}_0$ be a principal axis in the plane, and let $\hat{\boldsymbol{\omega}}_\theta$ be the axis obtained by rotating $\hat{\boldsymbol{\omega}}_0$ through the angle θ. Then $\hat{\boldsymbol{\omega}}_\theta$ is also a principal axis with the same principal moment, due to the symmetry of the object. Therefore, $\mathbf{I}\hat{\boldsymbol{\omega}}_0 = I\hat{\boldsymbol{\omega}}_0$, and $\mathbf{I}\hat{\boldsymbol{\omega}}_\theta = I\hat{\boldsymbol{\omega}}_\theta$.

Now, any vector $\boldsymbol{\omega}$ in the x-y plane can be written as a linear combination of $\hat{\boldsymbol{\omega}}_0$ and $\hat{\boldsymbol{\omega}}_\theta$, provided that $\theta \neq 180°$ (or zero, of course). That is, $\hat{\boldsymbol{\omega}}_0$ and $\hat{\boldsymbol{\omega}}_\theta$ span the x-y plane. Therefore, any vector $\boldsymbol{\omega}$ can be written as $\boldsymbol{\omega} = a\hat{\boldsymbol{\omega}}_0 + b\hat{\boldsymbol{\omega}}_\theta$, and so

$$\mathbf{I}\boldsymbol{\omega} = \mathbf{I}(a\hat{\boldsymbol{\omega}}_0 + b\hat{\boldsymbol{\omega}}_\theta) = aI\hat{\boldsymbol{\omega}}_0 + bI\hat{\boldsymbol{\omega}}_\theta = I\boldsymbol{\omega}. \tag{9.26}$$

Hence, $\boldsymbol{\omega}$ is also a principal axis. Problem 9.8 gives another proof of this theorem. ■

The theorem actually holds even without the "pancake" restriction. That is, it holds for any object with a rotational symmetry around the z axis (excluding $\theta \neq 180°$). This can be seen as follows. The z axis is a principal axis, because if $\boldsymbol{\omega}$ points along $\hat{z}$, then $\mathbf{L}$ must also point along $\hat{z}$, by symmetry. There are therefore (at least) two principal axes in the x-y plane. Label one of these as $\hat{\boldsymbol{\omega}}_0$ and proceed as above.

Let's now do some quick examples. We'll state the principal axes for the objects listed below (relative to the origin). Your task is to show that they are correct. Usually, a quick symmetry argument shows that

$$\mathbf{I} \equiv \begin{pmatrix} \int (y^2+z^2) & -\int xy & -\int zx \\ -\int xy & \int (z^2+x^2) & -\int yz \\ -\int zx & -\int yz & \int (x^2+y^2) \end{pmatrix} \qquad (9.27)$$

is diagonal. In all of these examples (see Fig. 9.17), the origin for the principal axes is understood to be the origin of the given coordinate system (which is not necessarily the CM). In describing the axes, they therefore all pass through the origin, in addition to having the other properties stated.

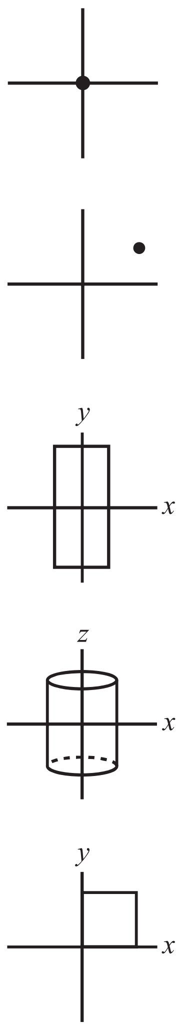

Fig. 9.17

Example 1: Point mass at the origin.
principal axes: any axes.

Example 2: Point mass at the point (x_0, y_0, z_0).
principal axes: axis through point, any axes perpendicular to this.

Example 3: Rectangle centered at the origin, as shown.
principal axes: the x, y, and z axes.

Example 4: Cylinder with axis as z axis.
principal axes: z axis, any axes in x-y plane.

Example 5: Square with one corner at origin, as shown.
principal axes: z axis, axis through CM, axis perpendicular to this.

9.4 Two basic types of problems

The previous three sections introduced a variety of abstract concepts. We will now finally look at some actual physical systems. The concept of principal axes gives us the ability to solve many kinds of problems. Two kinds, however, come up again and again. There are variations on these, of course, but they may be generally stated as follows.

- Strike a rigid object with an impulsive (that is, quick) blow. What is the motion of the object immediately after the blow?
- An object rotates around a fixed axis. A given torque is applied. What is the frequency of the rotation? Or conversely, given the frequency, what is the required torque?

We'll work through an example for each of these problems. In both cases, the solution involves a few standard steps, so we'll write them out explicitly.

9.4.1 Motion after an impulsive blow

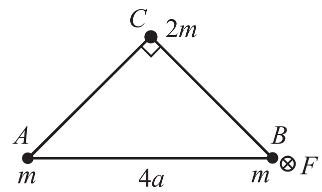

Fig. 9.18

Problem: Consider the rigid object in Fig. 9.18. Three masses are connected by three massless rods, in the shape of an isosceles right triangle with hypotenuse length $4a$. The mass at the right angle is $2m$, and the other two masses are m. Label them A, B, C, as shown. Assume that the object is floating freely in outer space. Mass B is struck with a quick blow, directed into the page. Let the imparted impulse have magnitude $\int F\,dt = P$. (See Section 8.6 for a discussion of impulse and angular impulse.) What are the velocities of the three masses immediately after the blow?

Solution: Our strategy will be to find the angular momentum of the system (relative to the CM) using the angular impulse, and then calculate the principal moments and find the angular velocity vector (which will give the velocities relative to the CM), and then finally add on the CM motion.

The altitude from the right angle to the hypotenuse has length $2a$, and the CM is easily seen to be located at its midpoint (see Fig. 9.19). Picking the CM as our origin, and letting the plane of the paper be the x-y plane, the positions of the three masses are $\mathbf{r}_A = (-2a, -a, 0)$, $\mathbf{r}_B = (2a, -a, 0)$, and $\mathbf{r}_C = (0, a, 0)$. There are now five standard steps that we must perform.

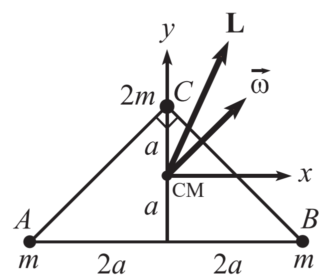

Fig. 9.19

- **Find L:** The positive z axis is directed out of the page, so the impulse vector is $\mathbf{P} \equiv \int \mathbf{F}\,dt = (0, 0, -P)$. Therefore, the angular momentum of the system (relative to the CM) is

$$\begin{aligned} \mathbf{L} = \int \boldsymbol{\tau}\,dt = \int (\mathbf{r}_B \times \mathbf{F})\,dt = \mathbf{r}_B \times \int \mathbf{F}\,dt \\ = (2a, -a, 0) \times (0, 0, -P) = aP(1, 2, 0), \end{aligned} \tag{9.28}$$

as shown in Fig. 9.19. We have used the fact that $\mathbf{r}_B$ is essentially constant during the blow (because the blow is assumed to happen very quickly) in taking $\mathbf{r}_B$ outside the integral.

- **Calculate the principal moments:** The principal axes are the x, y, and z axes, because the symmetry of the triangle makes $\mathbf{I}$ diagonal in this basis, as you can quickly check. The moments (relative to the CM) are

$$\begin{aligned} I_x &= ma^2 + ma^2 + (2m)a^2 = 4ma^2, \\ I_y &= m(2a)^2 + m(2a)^2 + (2m)0^2 = 8ma^2, \\ I_z &= I_x + I_y = 12ma^2. \end{aligned} \tag{9.29}$$

We have used the perpendicular-axis theorem to obtain I_z, although it won't be needed to solve the problem.

- **Find $\boldsymbol{\omega}$:** We now have two expressions for the angular momentum of the system. One expression is in terms of the given impulse, Eq. (9.28). The other is in terms of the moments and the angular velocity components, Eq. (9.24). Equating these gives

$$\begin{aligned}
(I_x\omega_x, I_y\omega_y, I_z\omega_z) &= aP(1,2,0) \\
\Longrightarrow \quad (4ma^2\omega_x, 8ma^2\omega_y, 12ma^2\omega_z) &= aP(1,2,0) \\
\Longrightarrow \quad (\omega_x, \omega_y, \omega_z) &= \frac{P}{4ma}(1,1,0),
\end{aligned} \tag{9.30}$$

as shown in Fig. 9.19.

- **Calculate the velocities relative to the CM:** Right after the blow, the object rotates around the CM with the angular velocity found in Eq. (9.30). The velocities relative to the CM are then $\mathbf{u}_i = \boldsymbol{\omega} \times \mathbf{r}_i$. Thus,

$$\begin{aligned}
\mathbf{u}_A &= \boldsymbol{\omega} \times \mathbf{r}_A = \frac{P}{4ma}(1,1,0) \times (-2a, -a, 0) = (0,0,P/4m), \\
\mathbf{u}_B &= \boldsymbol{\omega} \times \mathbf{r}_B = \frac{P}{4ma}(1,1,0) \times (2a, -a, 0) = (0,0,-3P/4m), \\
\mathbf{u}_C &= \boldsymbol{\omega} \times \mathbf{r}_C = \frac{P}{4ma}(1,1,0) \times (0, a, 0) = (0,0,P/4m).
\end{aligned} \tag{9.31}$$

  As a check, it makes sense that $\mathbf{u}_B$ is three times as large as $\mathbf{u}_A$ and $\mathbf{u}_C$, because B is three times as far from the axis of rotation as A and C are, as you can verify by doing a little geometry in Fig. 9.19.

- **Add on the velocity of the CM:** The impulse (that is, the change in linear momentum) supplied to the whole system is $\mathbf{P} = (0,0,-P)$. The total mass of the system is $M = 4m$. Therefore, the velocity of the CM is

$$V_{\text{CM}} = \frac{\mathbf{P}}{M} = (0,0,-P/4m). \tag{9.32}$$

  The total velocities of the masses are therefore

$$\begin{aligned}
\mathbf{v}_A &= \mathbf{u}_A + V_{\text{CM}} = (0,0,0), \\
\mathbf{v}_B &= \mathbf{u}_B + V_{\text{CM}} = (0,0,-P/m), \\
\mathbf{v}_C &= \mathbf{u}_C + V_{\text{CM}} = (0,0,0).
\end{aligned} \tag{9.33}$$

REMARKS:

1. We see that masses A and C are instantaneously at rest immediately after the blow, and mass B acquires all of the imparted impulse. In retrospect, this is clear. Basically, it is possible for both A and C to remain at rest while B moves a tiny bit, so this is what happens. If B moves into the page by a small distance ϵ, then A and C won't know that B has moved, because their distances to B will change (assuming hypothetically that they don't move) by a distance of order only ϵ^2. If we changed the problem and added a mass D at, say, the midpoint of the hypotenuse, then it would *not* be possible for A, C, and D to remain at rest while B moved a tiny bit. So there would have to be some other motion in addition to B's. This setup is the topic of Exercise 9.38.
2. As time goes on, the system undergoes a rather complicated motion. What happens is that the CM moves with constant velocity while the masses rotate around it in

a messy manner. Since there are no torques acting on the system (after the initial blow), we know that **L** forever remains constant. It turns out that $\boldsymbol{\omega}$ moves around **L** while the masses rotate around this changing $\boldsymbol{\omega}$. These matters are the subject of Section 9.6, although in that discussion we restrict ourselves to symmetric tops, that is, ones with two equal moments. But these issues aside, it's good to know that we can, without too much difficulty, determine what's going on immediately after the blow.

3. The object in this problem was assumed to be floating freely in space. If we instead have an object that is pivoted at a given fixed point, then we should use this pivot as our origin. There is then no need to perform the last step of adding on the velocity of the origin (which was the CM, above), because this velocity is now zero. Equivalently, just consider the pivot to be an infinite mass, which is therefore the location of the (motionless) CM. ♣

9.4.2 Frequency of motion due to a torque

Problem: Consider a stick of length ℓ, mass m, and uniform mass density. The stick is pivoted at its top end and swings around the vertical axis. Assume that conditions have been set up so that the stick always makes an angle θ with the vertical, as shown in Fig. 9.20. What is the frequency, ω, of this motion?

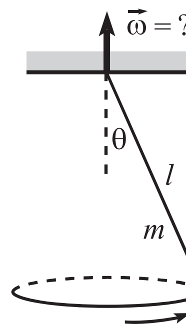

Fig. 9.20

Solution: Our strategy will be to find the principal moments and then the angular momentum of the system (in terms of ω), and then find the rate of change of **L**, and then calculate the torque and equate it with $d\mathbf{L}/dt$. We will choose the pivot to be the origin.[8] Again, there are five standard steps that we must perform.

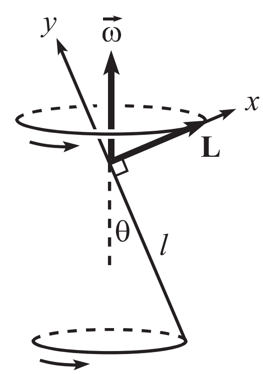

Fig. 9.21

- **Calculate the principal moments:** The principal axes are the axis along the stick, along with any two orthogonal axes perpendicular to the stick. So let the x and y axes be as shown in Fig. 9.21. The positive z axis then points out of the page. The moments (relative to the pivot) are $I_x = m\ell^2/3$, $I_y = 0$, and $I_z = m\ell^2/3$ (which won't be needed).
- **Find L:** The angular velocity vector points vertically (however, see the third remark following this solution), so in the basis of the principal axes, the angular velocity vector is $\boldsymbol{\omega} = (\omega \sin\theta, \omega\cos\theta, 0)$, where ω is yet to be determined. The angular momentum of the system (relative to the pivot) is therefore

$$\mathbf{L} = (I_x\omega_x,\ I_y\omega_y,\ I_z\omega_z) = \big((1/3)m\ell^2\omega\sin\theta, 0, 0\big). \tag{9.34}$$

- **Find $d\mathbf{L}/dt$:** The vector **L** in Eq. (9.34) points up to the right, along the x axis (at the instant shown in Fig. 9.21), with magnitude $L = (1/3)m\ell^2\omega\sin\theta$. As the stick rotates around the vertical axis, **L** traces out the surface of a cone. That is, the tip of **L** traces out a horizontal circle. The radius of this circle is the horizontal component of **L**, which is $L\cos\theta$. The speed of the tip (which is the magnitude of $d\mathbf{L}/dt$) is therefore

[8] This is a better choice than the CM because this way we won't have to worry about any messy forces acting at the pivot when computing the torque. The task of Exercise 9.41 is to work through the more complicated solution which has the CM as the origin.

$(L\cos\theta)\omega$, because $\mathbf{L}$ rotates around the vertical axis with the same frequency as the stick. So $d\mathbf{L}/dt$ has magnitude

$$\left|\frac{d\mathbf{L}}{dt}\right| = (L\cos\theta)\omega = \frac{1}{3}m\ell^2\omega^2\sin\theta\cos\theta, \tag{9.35}$$

and it points into the page.

REMARK: With more complicated objects where $I_y \neq 0$, $\mathbf{L}$ won't point nicely along a principal axis, so the length of its horizontal component (the radius of the circle that $\mathbf{L}$ traces out) won't immediately be obvious. In this case, you can either explicitly calculate the horizontal component (see the spinning-top example in Section 9.7.5), or you can just do things the formal way by finding the rate of change of $\mathbf{L}$ via the expression $d\mathbf{L}/dt = \boldsymbol{\omega}\times\mathbf{L}$, which holds for all the same reasons that $\mathbf{v} \equiv d\mathbf{r}/dt = \boldsymbol{\omega}\times\mathbf{r}$ holds. In the present problem, we obtain

$$\begin{aligned} d\mathbf{L}/dt &= (\omega\sin\theta, \omega\cos\theta, 0)\times\big((1/3)m\ell^2\omega\sin\theta, 0, 0\big) \\ &= \big(0, 0, -(1/3)m\ell^2\omega^2\sin\theta\cos\theta\big), \end{aligned} \tag{9.36}$$

in agreement with Eq. (9.35). And the direction is correct, because the negative z axis points into the page. Note that we calculated this cross product in the principal-axis basis. Although these axes are changing in time, they present a perfectly good set of basis vectors at any instant. ♣

- **Calculate the torque:** The torque (relative to the pivot) is due to gravity, which effectively acts on the CM of the stick. So $\boldsymbol{\tau} = \mathbf{r}\times\mathbf{F}$ has magnitude

$$\tau = rF\sin\theta = (\ell/2)(mg)\sin\theta, \tag{9.37}$$

and it points into the page.

- **Equate $\boldsymbol{\tau}$ with $d\mathbf{L}/dt$:** The vectors $d\mathbf{L}/dt$ and $\boldsymbol{\tau}$ both point into the page, which is good, because they had better point in the same direction. Equating their magnitudes gives

$$\frac{m\ell^2\omega^2\sin\theta\cos\theta}{3} = \frac{mg\ell\sin\theta}{2} \quad\Longrightarrow\quad \omega = \sqrt{\frac{3g}{2\ell\cos\theta}}. \tag{9.38}$$

REMARKS:

1. This frequency is slightly larger than the frequency that would arise if we instead had a mass on the end of a massless stick of length ℓ. From Problem 9.12, the frequency in that case is $\sqrt{g/\ell\cos\theta}$. So, in some sense, a uniform stick of length ℓ behaves like a mass on the end of a massless stick of length $2\ell/3$, as far as these rotations are concerned.
2. As $\theta \to \pi/2$, the frequency goes to ∞, which makes sense. And as $\theta \to 0$, it approaches $\sqrt{3g/2\ell}$, which isn't so obvious.
3. As explained in Problem 9.1, the instantaneous $\boldsymbol{\omega}$ is not uniquely defined in some situations. At the instant shown in Fig. 9.20, the stick is moving directly into the page. What if someone else wants to think of the stick as (instantaneously) rotating around the $\boldsymbol{\omega}'$ axis perpendicular to the stick (the x axis, in the above notation), instead of the vertical axis, as shown in Fig. 9.22. What is the angular speed ω'?

   Well, if ω is the angular speed of the stick around the vertical axis, then we may view the tip of the stick as instantaneously moving in a circle of radius $\ell\sin\theta$ around the

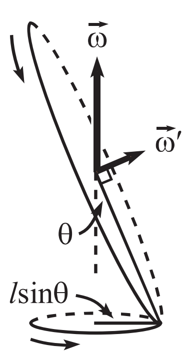

Fig. 9.22

vertical axis $\boldsymbol{\omega}$. So $\omega(\ell \sin\theta)$ is the speed of the tip of the stick. But we may also view the tip of the stick as instantaneously moving in a circle of radius ℓ around $\boldsymbol{\omega}'$, as shown. The speed of the tip is still $\omega(\ell \sin\theta)$, so the angular speed around this axis is given by $\omega'\ell = \omega(\ell \sin\theta)$. Hence $\omega' = \omega \sin\theta$, which is simply the x component of $\boldsymbol{\omega}$ that we found above, right before Eq. (9.34). The moment of inertia around $\boldsymbol{\omega}'$ is $m\ell^2/3$, so the angular momentum has magnitude $(m\ell^2/3)(\omega \sin\theta)$, in agreement with Eq. (9.34). And the direction is along the x axis, as it should be.

Note that although $\boldsymbol{\omega}$ is not uniquely defined at any instant, $\mathbf{L} \equiv \int (\mathbf{r} \times \mathbf{p})\, dm$ certainly is.[9] Choosing $\boldsymbol{\omega}$ to point vertically, as we did in the above solution, is in some sense the natural choice, because this $\boldsymbol{\omega}$ doesn't change with time. ♣

9.5 Euler's equations

Consider a rigid body instantaneously rotating around an axis $\boldsymbol{\omega}$. This $\boldsymbol{\omega}$ may change as time goes on, but all we care about for now is what it is at a given instant. The angular momentum is given by Eq. (9.8) as $\mathbf{L} = \mathbf{I}\boldsymbol{\omega}$, where $\mathbf{I}$ is the inertia tensor, calculated with respect to a given origin and a given set of axes (and $\boldsymbol{\omega}$ is written in the same basis, of course).

As usual, things are much nicer if we use the principal axes (relative to the chosen origin) as the basis vectors of our coordinate system. Since these axes are fixed with respect to the rotating object, they will rotate with respect to the fixed reference frame. In this basis, $\mathbf{L}$ takes the nice form,

$$\mathbf{L} = (I_1\omega_1, I_2\omega_2, I_3\omega_3), \tag{9.39}$$

where ω_1, ω_2, and ω_3 are the components of $\boldsymbol{\omega}$ along the principal axes. In other words, if you take the vector $\mathbf{L}$ in space and project it onto the instantaneous principal axes, then you get the components in Eq. (9.39).

On one hand, writing $\mathbf{L}$ in terms of the rotating principal axes allows us to write it in the nice form of Eq. (9.39). But on the other hand, writing $\mathbf{L}$ in this way makes it nontrivial to determine how it changes in time, because the principal axes themselves are changing. However, it turns out that the benefits outweigh the detriments, so we will invariably use the principal axes as our basis vectors.

The goal of this section is to find an expression for $d\mathbf{L}/dt$, and to then equate this with the torque. The result will be Euler's equations in Eq. (9.45).

Derivation of Euler's equations

If we write $\mathbf{L}$ in terms of the body frame, which we'll choose to be described by the principal axes painted on the body, then $\mathbf{L}$ can change (relative to the lab frame) due to two effects. It can change because its coordinates in the body frame change, and it can also change because of the rotation of the body frame. To be precise, let $\mathbf{L}_0$ be the vector $\mathbf{L}$ at a given instant. At this instant, imagine painting the vector $\mathbf{L}_0$ onto the body frame, so that $\mathbf{L}_0$ then rotates with the body. The rate

[9] The nonuniqueness of $\boldsymbol{\omega}$ arises from the fact that $I_y = 0$ here. If all the moments are nonzero, then $(L_x, L_y, L_z) = (I_x\omega_x, I_y\omega_y, I_z\omega_z)$ uniquely determines $\boldsymbol{\omega}$, given $\mathbf{L}$.

of change of $\mathbf{L}$ with respect to the lab frame may be written in the (identically true) way,

$$\frac{d\mathbf{L}}{dt} = \frac{d(\mathbf{L} - \mathbf{L}_0)}{dt} + \frac{d\mathbf{L}_0}{dt}. \tag{9.40}$$

The second term here is simply the rate of change of a body-fixed vector, which we know is $\boldsymbol{\omega} \times \mathbf{L}_0$, which equals $\boldsymbol{\omega} \times \mathbf{L}$ at this instant. The first term is the rate of change of $\mathbf{L}$ with respect to the body frame, which we'll denote by $\delta\mathbf{L}/\delta t$. This is what someone standing fixed on the body measures. So we end up with

$$\frac{d\mathbf{L}}{dt} = \frac{\delta\mathbf{L}}{\delta t} + \boldsymbol{\omega} \times \mathbf{L}. \tag{9.41}$$

This is actually a general statement, true for any vector in any rotating frame (we'll derive it in another more mathematical way in Chapter 10). There was nothing particular about $\mathbf{L}$ that we used in the above derivation. Also, there was no need to restrict ourselves to principal axes. In words, what we've shown is that the total change equals the change relative to the rotating frame, plus the change of the rotating frame relative to the fixed frame. This is just the usual way of adding velocities when one frame moves with respect to another.

Let us now make use of our choice of the principal axes as the body axes. This will put Eq. (9.41) in a usable form. Using Eq. (9.39), we can rewrite Eq. (9.41) as

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(I_1\omega_1, I_2\omega_2, I_3\omega_3) + (\omega_1, \omega_2, \omega_3) \times (I_1\omega_1, I_2\omega_2, I_3\omega_3). \tag{9.42}$$

The $\delta\mathbf{L}/\delta t$ term does indeed equal $(d/dt)(I_1\omega_1, I_2\omega_2, I_3\omega_3)$, because someone in the body frame measures the components of $\mathbf{L}$ with respect to the principal axes to be $(I_1\omega_1, I_2\omega_2, I_3\omega_3)$. And $\delta\mathbf{L}/\delta t$ is by definition the rate at which these components change.

Equation (9.42) equates two vectors. As is true for any vector, these (equal) vectors have an existence that is independent of the coordinate system we choose to describe them with (Eq. (9.41) makes no reference to a coordinate system). But since we've chosen an explicit frame on the right-hand side of Eq. (9.42), we should choose the same frame for the left-hand side. We can then equate the components on the left with the components on the right. Projecting $d\mathbf{L}/dt$ onto the instantaneous principal axes, we have

$$\left(\left(\frac{d\mathbf{L}}{dt}\right)_1, \left(\frac{d\mathbf{L}}{dt}\right)_2, \left(\frac{d\mathbf{L}}{dt}\right)_3\right) = \frac{d}{dt}(I_1\omega_1, I_2\omega_2, I_3\omega_3) + (\omega_1, \omega_2, \omega_3) \times (I_1\omega_1, I_2\omega_2, I_3\omega_3). \tag{9.43}$$

REMARK: The left-hand side looks nastier than it really is. The reason we've written it in this cumbersome way is the following (this is a remark that has to be read very slowly). We could have written the left-hand side as $(d/dt)(L_1, L_2, L_3)$, but this might cause confusion as to whether the L_i refer to the components with respect to the rotating axes, or the components with respect to the fixed set of axes that coincide with the rotating principal axes at this instant. That is, do we project **L** onto the principal axes to obtain components, and then take the derivative of these components? Or do we take the derivative of **L** and then project onto the principal axes to obtain components? The latter is what we mean in Eq. (9.43).[10] The way we've written the left-hand side of Eq. (9.43), it's clear that we're taking the derivative first. We are, after all, simply projecting Eq. (9.41) onto the principal axes. ♣

The time derivatives on the right-hand side of Eq. (9.43) are $d(I_1\omega_1)/dt = I_1\dot{\omega}_1$, etc., because the I's are constant. Performing the cross product and equating the corresponding components on each side yields the three equations,

$$\begin{aligned}
\left(\frac{d\mathbf{L}}{dt}\right)_1 &= I_1\dot{\omega}_1 + (I_3 - I_2)\omega_3\omega_2, \\
\left(\frac{d\mathbf{L}}{dt}\right)_2 &= I_2\dot{\omega}_2 + (I_1 - I_3)\omega_1\omega_3, \\
\left(\frac{d\mathbf{L}}{dt}\right)_3 &= I_3\dot{\omega}_3 + (I_2 - I_1)\omega_2\omega_1.
\end{aligned} \tag{9.44}$$

We will now invoke the results of Section 8.4.3 to say that if we have chosen the origin of our rotating frame to be either a fixed point or the CM (as we always do), then we can equate $d\mathbf{L}/dt$ with the torque, $\boldsymbol{\tau}$. We therefore have

$$\begin{aligned}
\tau_1 &= I_1\dot{\omega}_1 + (I_3 - I_2)\omega_3\omega_2, \\
\tau_2 &= I_2\dot{\omega}_2 + (I_1 - I_3)\omega_1\omega_3, \\
\tau_3 &= I_3\dot{\omega}_3 + (I_2 - I_1)\omega_2\omega_1.
\end{aligned} \tag{9.45}$$

These are *Euler's equations*. You need to remember only one of them, because the other two can be obtained by cyclic permutation of the indices.

REMARKS:

1. We repeat that the left- and right-hand sides of Eqs. (9.45) are components that are measured with respect to the instantaneous principal axes. Let's say we do a problem, for example, where τ_3 has a constant nonzero value, and τ_1 and τ_2 are always zero (as in the example in Section 9.4.2). This doesn't mean that $\boldsymbol{\tau}$ is a constant vector. On the contrary, $\boldsymbol{\tau}$ always points along the $\hat{\mathbf{x}}_3$ vector in the rotating frame, but this vector is changing in the fixed frame (unless $\hat{\mathbf{x}}_3$ points along $\boldsymbol{\omega}$).
2. The two types of terms on the right-hand sides of Eqs. (9.44) are the two types of changes that **L** can undergo. **L** can change because its components with respect to the rotating frame change, and **L** can also change because the body is rotating around $\boldsymbol{\omega}$.

[10] The former is $\delta\mathbf{L}/\delta t$, by definition. The two interpretations certainly give different results. For example, if instead of **L** we consider a vector fixed in the body (such as the $\mathbf{L}_0$ above), then the first interpretation gives a zero result, whereas the second interpretation gives a nonzero result. Considering what we mean by, say, the vector $(\omega_1, \omega_2, \omega_3)$, I think that the more logical interpretation of $(d/dt)(L_1, L_2, L_3)$ is the first one, so it should definitely be avoided.