# Operators and Matrices

You've been using operators for years even if you've never heard the term. Differentiation falls into this category; so does rotation; so does wheel-alignment. In the subject of quantum mechanics, familiar ideas such as energy and momentum will be represented by operators. You probably think that pressure is simply a scalar, but no. It's an operator.

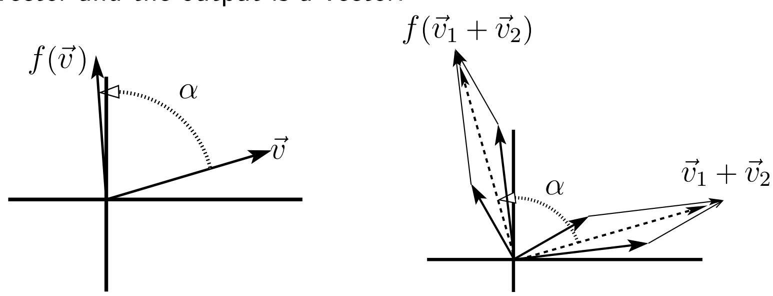
## 7.1 The Idea of an Operator

You can understand the subject of matrices as a set of rules that govern certain square or rectangular arrays of numbers — how to add them, how to multiply them. Approached this way the subject is remarkably opaque. Who made up these rules and why? What's the point? If you look at it as simply a way to write simultaneous linear equations in a compact way, it's perhaps convenient but certainly not the big deal that people make of it. It is a big deal.

There's a better way to understand the subject, one that relates the matrices to more fundamental ideas and that even provides some geometric insight into the subject. The technique of similarity transformations may even make a little sense. This approach is precisely parallel to one of the basic ideas in the use of vectors. You can draw pictures of vectors and manipulate the pictures of vectors and that's the right way to look at certain problems. You quickly find however that this can be cumbersome. A general method that you use to make computations tractable is to write vectors in terms of their components, then the methods for manipulating the components follow a few straight-forward rules, adding the components, multiplying them by scalars, even doing dot and cross products.

Just as you have components of vectors, which are a set of numbers that depend on your choice of basis, matrices are a set of numbers that are components of — not vectors, but functions (also called operators or transformations or tensors). I'll start with a couple of examples before going into the precise definitions.

The first example of the type of function that I'll be interested in will be a function defined on the two-dimensional vector space, arrows drawn in the plane with their starting points at the origin. The function that I'll use will rotate each vector by an angle $\alpha$ counterclockwise. This *is* a function, where the input is a vector and the output is a vector.




What happens if you change the argument of this function, multiplying it by a scalar? You know $f(\vec{v})$, what is $f(c\vec{v})$? Just from the picture, this is $c$ times the vector that you got by rotating $\vec{v}$. What happens when you add two vectors and then rotate the result? The whole parallelogram defining the addition will rotate through the same angle $\alpha$, so whether you apply the function before or after adding the vectors you get the same result.

This leads to the definition of the word *linearity:*

$$\boxed{f(c\vec{v}) = cf(\vec{v}), \qquad \text{and} \qquad f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2)} \tag{7.1}$$

*Keep your eye on this pair of equations!* They're central to the whole subject.

Another example of the type of function that I'll examine is from physics instead of mathematics. A rotating rigid body has some angular momentum. The greater the rotation rate, the greater the angular momentum will be. Now how do I compute the angular momentum assuming that I know the shape and the distribution of masses in the body and that I know the body's angular velocity? The body is made of a lot of point masses (atoms), but you don't need to go down to that level to make sense of the subject. As with any other integral, you start by dividing the object in to a lot of small pieces.

What is the angular momentum of a single point mass? It starts from basic Newtonian mechanics, and the equation $\vec F = d\vec p/dt$. (It's better in this context to work with this form than with the more common expressions $\vec F = m\vec a$.) Take the cross product with $\vec r$, the displacement vector from the origin.

$$\vec r \times \vec F = \vec r \times d\vec p/dt$$

Add and subtract the same thing on the right side of the equation (add zero) to get

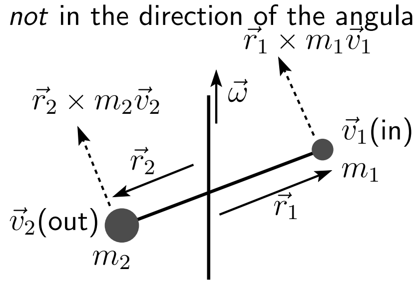
$$\begin{aligned}\vec r \times \vec F &= \vec r \times \frac{d\vec p}{dt} + \frac{d\vec r}{dt} \times \vec p - \frac{d\vec r}{dt} \times \vec p \\ &= \frac{d}{dt}(\vec r \times \vec p) - \frac{d\vec r}{dt} \times \vec p\end{aligned}$$

Now recall that $\vec p$ is $m\vec v$, and $\vec v = d\vec r/dt$, so the last term in the preceding equation is zero because you are taking the cross product of a vector with itself. This means that when adding and subtracting a term from the right side above, I was really adding and subtracting zero.

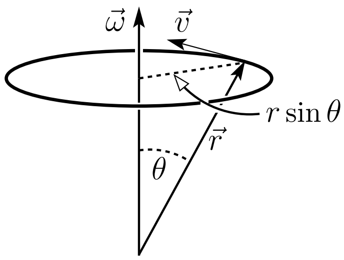
$\vec r \times \vec F$ is the torque applied to the point mass $m$ and $\vec r \times \vec p$ is the mass's angular momentum about the origin. Now if there are many masses and many forces, simply put an index on this torque equation and add the resulting equations over all the masses in the rigid body. The sums on the left and the right provide the definitions of torque and of angular momentum.

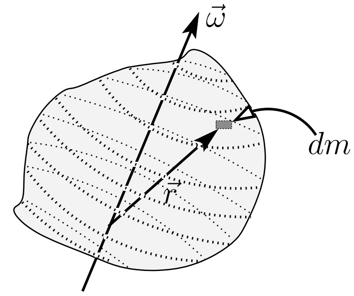
$$\vec\tau_{\text{total}} = \sum_k \vec r_k \times \vec F_k = \frac{d}{dt}\sum_k \left(\vec r_k \times \vec p_k\right) = \frac{d\vec L}{dt}$$

For a specific example, attach two masses to the ends of a light rod and attach that rod to a second, vertical one as sketched — at an angle. Now spin the vertical rod and figure out what the angular velocity and angular momentum vectors are. Since the spin is along the vertical rod, that specifies the direction of the angular velocity vector $\vec\omega$ to be upwards in the picture. (Viewed from above everything is rotating counter-clockwise.) The angular momentum of one point mass is $\vec r \times \vec p = \vec r \times m\vec v$. The mass on the right has a velocity pointing *into* the page and the mass on the left has it pointing *out.* Take the origin to be where the supporting rod is attached to the axis, then $\vec r \times \vec p$ for the mass on the right is pointing up and to the left. For the other mass both $\vec r$ and $\vec p$ are reversed, so the cross product is in exactly the same direction as for the first mass. The total angular momentum the sum of these two parallel vectors, and it is *not* in the direction of the angular velocity.



Now make this quantitative and apply it to a general rigid body. There are two basic pieces to the problem: the angular momentum of a point mass and the velocity of a point mass in terms of its angular velocity. The position of one point mass is described by its displacement vector from the origin, $\vec r$. Its angular momentum is then $\vec r\times\vec p$, where $\vec p = m\vec v$. If the rigid body has an angular velocity vector $\vec\omega$, the linear velocity of a mass at coordinate $\vec r$ is $\vec\omega\times\vec r$.




The total angular momentum of a rotating set of masses $m_k$ at respective coordinates $\vec r_k$ is the sum of all the individual pieces of angular momentum

$$\begin{aligned}\vec L &= \sum_k \vec r_k\times m_k\vec v_k, \qquad \text{and since} \qquad \vec v_k = \vec\omega\times\vec r_k, \\ \vec L &= \sum_k \vec r_k\times m_k(\vec\omega\times\vec r_k)\end{aligned} \tag{7.2}$$

If you have a continuous distribution of mass then using an integral makes more sense. For a given distribution of mass, this integral (or sum) depends on the vector $\vec\omega$. It defines a function having a vector as input and a vector $\vec L$ as output. Denote the function by $I$, so $\vec L = I(\vec\omega\,)$.

$$\vec L = \int dm\,\vec r\times\big(\vec\omega\times\vec r\big) = I(\vec\omega\,) \tag{7.3}$$

This function satisfies the same linearity equations as Eq. (7.1). When you multiply $\vec\omega$ by a constant, the output, $\vec L$ is multiplied by the same constant. When you add two $\vec\omega$'s together as the argument, the properties of the cross product and of the integral guarantee that the corresponding $\vec L$'s are added.

$$I(c\vec\omega\,) = cI(\vec\omega\,), \qquad \text{and} \qquad I(\vec\omega_1+\vec\omega_2) = I(\vec\omega_1)+I(\vec\omega_2)$$

This function $I$ is called the "inertia operator" or more commonly the "inertia tensor." It's not simply multiplication by a scalar, so the rule that appears in an introductory course in mechanics ($\vec L = I\vec\omega$) is valid only in special cases, for example those with enough symmetry.

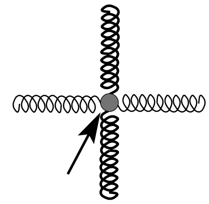
Note: $I$ is not a vector and $\vec L$ is not a function. $\vec L$ is the output of the function $I$ when you feed it the argument $\vec\omega$. This is the same sort of observation appearing in section 6.3 under "Function Spaces."

If an electromagnetic wave passes through a crystal, the electric field will push the electrons around, and the bigger the electric field, the greater the distance that the electrons will be pushed. They may not be pushed in the same direction as the electric field however, as the nature of the crystal can make it easier to push the electrons in one direction than in another. The relation between the applied field and the average electron displacement is a function that (for moderate size fields) obeys the same linearity relation that the two previous functions do.

$$\vec P = \alpha(\vec E\,)$$

$\vec{P}$ is the electric dipole moment density and $\vec{E}$ is the applied electric field. The function $\alpha$ is called the polarizability.

If you have a mass attached to six springs that are in turn attached to six walls, the mass will come to equilibrium somewhere. Now push on this mass with another (not too large) force. The mass will move, but will it move in the direction that you push it? If the six springs are all the same it will, but if they're not then the displacement will be more in the direction of the weaker springs. The displacement, $\vec{d}$, will still however depend linearly on the applied force, $\vec{F}$.



### 7.2 Definition of an Operator

An operator, also called a linear transformation, is a particular type of function. It is first of all, a vector valued function of a vector variable. Second, it is linear; that is, if $A$ is such a function then $A(\vec{v})$ is a vector, and

$$A(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha A(\vec{v}_1) + \beta A(\vec{v}_2). \tag{7.4}$$

The *domain* is the set of variables on which the operator is defined. The *range* is the set of all values put out by the function. Are there *nonlinear* operators? Yes, but not here.

### 7.3 Examples of Operators

The four cases that I started with, rotation in the plane, angular momentum of a rotating rigid body, polarization of a crystal by an electric field, and the mass attached to some springs all fit this definition. Other examples:

5. The simplest example of all is just multiplication by a scalar: $A(\vec{v}) \equiv c\vec{v}$ for all $\vec{v}$. This applies to any vector space and its domain is the entire space.
6. On the vector space of all real valued functions on a given interval, multiply any function $f$ by $1+x^2$: $(Af)(x) = (1+x^2)f(x)$. The domain of $A$ is the entire space of functions of $x$. This is an infinite dimensional vector space, but no matter. There's nothing special about $1+x^2$, and any other function will do to define an operator.
7. On the vector space of square integrable functions $\left[\int dx\,|f(x)|^2 < \infty\right]$ on $a < x < b$, define the operator as multiplication by $x$. The only distinction to make here is that if the interval is infinite, then $xf(x)$ may not itself be square integrable. The domain of this operator in this case is therefore *not* the entire space, but just those functions such that $xf(x)$ is also square-integrable. On the same vector space, differentiation is a linear operator: $(Af)(x) = f'(x)$. This too has a restriction on the domain: It is necessary that $f'$ also exist and be square integrable.
8. On the vector space of infinitely differentiable functions, the operation of differentiation, $d/dx$, is itself a linear operator. It's certainly linear, and it takes a differentiable function into a differentiable function.

So where are the matrices? This chapter started by saying that I'm going to show you the inside scoop on matrices and so far I've failed to produce even one.

When you describe vectors you can use a basis as a computational tool and manipulate the vectors using their components. In the common case of three-dimensional vectors we usually denote the basis in one of several ways

$$\hat{\imath},\quad \hat{\jmath},\quad \hat{k}, \qquad \text{or} \qquad \hat{x},\quad \hat{y},\quad \hat{z}, \qquad \text{or} \qquad \vec{e}_1,\quad \vec{e}_2,\quad \vec{e}_3$$

and they all mean the same thing. The first form is what you see in the introductory physics texts. The second form is one that you encounter in more advanced books, and the third one is more suitable when you want to have a compact index notation. It's that third one that I'll use here; it has the advantage that it doesn't bias you to believe that you must be working in three spatial dimensions. The index could go beyond 3, and the vectors that you're dealing with may not be the usual geometric arrows. (And why does it have to start with one? Maybe I want the indices 0, 1, 2 instead.) These need not be perpendicular to each other or even to be unit vectors.

The way to write a vector $\vec{v}$ in components is

$$\vec{v} = v_x\,\hat{x} + v_y\,\hat{y} + v_z\,\hat{z}, \qquad \text{or} \qquad v_1\,\vec{e}_1 + v_2\,\vec{e}_2 + v_3\,\vec{e}_3 = \sum_k v_k\,\vec{e}_k \tag{7.5}$$

Once you've chosen a basis, you can find the three numbers that form the components of that vector. In a similar way, define the components of an operator, only that will take *nine* numbers to do it (in three dimensions). If you evaluate the effect of an operator on any one of the basis vectors, the output is a vector. That's part of the definition of the word operator. This output vector can itself be written in terms of this same basis. The defining equation for the components of an operator $f$ is

$$\boxed{f(\vec{e}_i) = \sum_{k=1}^{3} f_{ki}\,\vec{e}_k} \tag{7.6}$$

For each input vector you have the three components of the output vector. Pay careful attention to this equation! It is the defining equation for the entire subject of matrix theory, and everything in that subject comes from this one innocuous looking equation. (And yes if you're wondering, I wrote the indices in the correct order.)

Why?

Take an arbitrary input vector for $f$: $\vec{u} = f(\vec{v}\,)$. Both $\vec{u}$ and $\vec{v}$ are vectors, so write them in terms of the basis chosen.

$$\vec{u} = \sum_k u_k\,\vec{e}_k = f(\vec{v}\,) = f\Big(\sum_i v_i\,\vec{e}_i\Big) = \sum_i v_i f(\vec{e}_i) \tag{7.7}$$

The last equation is the result of the linearity property, Eq. (7.1), already assumed for $f$. Now pull the sum and the numerical factors $v_i$ out in front of the function, and write it out. It is then clear:

$$f(v_1\,\vec{e}_1 + v_2\,\vec{e}_2) = f(v_1\,\vec{e}_1) + f(v_2\,\vec{e}_2) = v_1 f(\vec{e}_1) + v_2 f(\vec{e}_2)$$

Now you see where the defining equation for operator components comes in. Eq. (7.7) is

$$\sum_k u_k\,\vec{e}_k = \sum_i v_i \sum_k f_{ki}\,\vec{e}_k$$

For two vectors to be equal, the corresponding coefficients of $\vec{e}_1$, $\vec{e}_2$, *etc.* must match; their respective components must be equal, and this is

$$u_k = \sum_i v_i f_{ki}, \qquad \text{usually written} \qquad u_k = \sum_i f_{ki} v_i \tag{7.8}$$

so that in the latter form it starts to resemble what you may think of as matrix manipulation. $f_{\text{row,column}}$ is the conventional way to write the indices, and multiplication is defined so that the following product *means* Eq. (7.8).

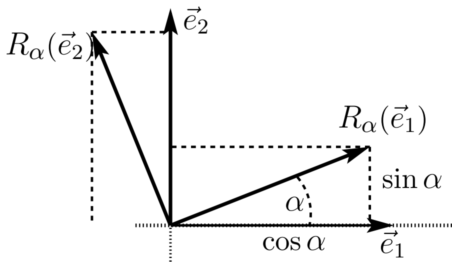
$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \tag{7.9}$$

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \qquad \text{is} \qquad u_1 = f_{11}v_1 + f_{12}v_2 + f_{13}v_3 \qquad \textit{etc.}$$

And this is the reason behind the definition of how to multiply a matrix and a column matrix. The order in which the indices appear is the conventional one, and the indices appear in the matrix as they do because I chose the order of the indices in a (seemingly) backwards way in Eq. (7.6).

**Components of Rotations**

Apply this to the first example, rotate all vectors in the plane through the angle $\alpha$. I don't want to keep using the same symbol $f$ for *every* function, so I'll call this function $R$ instead, or better yet $R_\alpha$. $R_\alpha(\vec v\,)$ is the rotated vector. Pick two perpendicular unit vectors for a basis. You may call them $\hat x$ and $\hat y$, but again I'll call them $\vec e_1$ and $\vec e_2$. Use the definition of components to get




$$R_\alpha(\vec e_1) = \sum_k R_{k1}\vec e_k$$
$$R_\alpha(\vec e_2) = \sum_k R_{k2}\vec e_k \tag{7.10}$$

The rotated $\vec e_1$ has two components, so

$$R_\alpha(\vec e_1) = \vec e_1 \cos\alpha + \vec e_2 \sin\alpha = R_{11}\vec e_1 + R_{21}\vec e_2 \tag{7.11}$$

This determines the first column of the matrix of components,

$$R_{11} = \cos\alpha, \qquad \text{and} \qquad R_{21} = \sin\alpha$$

Similarly the effect on the other basis vector determines the second column:

$$R_\alpha(\vec e_2) = \vec e_2 \cos\alpha - \vec e_1 \sin\alpha = R_{12}\vec e_1 + R_{22}\vec e_2 \tag{7.12}$$

Check: $R_\alpha(\vec e_1)\cdot R_\alpha(\vec e_2) = 0$.

$$R_{12} = -\sin\alpha, \qquad \text{and} \qquad R_{22} = \cos\alpha$$

The component matrix is then

$$(R_\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{7.13}$$

**Components of Inertia**

The definition, Eq. (7.3), and the figure preceding it specify the inertia tensor as the function that relates the angular momentum of a rigid body to its angular velocity.

$$\vec L = \int dm\, \vec r \times (\vec\omega \times \vec r\,) = I(\vec\omega\,) \tag{7.14}$$

Use the vector identity,

$$\vec A \times (\vec B \times \vec C\,) = \vec B(\vec A \cdot \vec C\,) - \vec C(\vec A \cdot \vec B\,) \tag{7.15}$$

then the integral is

$$\vec L = \int dm\, \big[\vec\omega(\vec r \cdot \vec r\,) - \vec r(\vec\omega \cdot \vec r\,)\big] = I(\vec\omega\,) \tag{7.16}$$

Pick the common rectangular, orthogonal basis and evaluate the components of this function. Equation (7.6) says $\vec r = x\,\vec e_1 + y\,\vec e_2 + z\,\vec e_3$ so

$$I(\vec e_i) = \sum_k I_{ki}\,\vec e_k$$
$$I(\vec e_1) = \int dm\,\big[\vec e_1(x^2+y^2+z^2) - (x\,\vec e_1 + y\,\vec e_2 + z\,\vec e_3)(x)\big]$$
$$= I_{11}\,\vec e_1 + I_{21}\,\vec e_2 + I_{31}\,\vec e_3$$

from which $\qquad I_{11} = \int dm\,(y^2+z^2), \qquad I_{21} = -\int dm\,yx, \qquad I_{31} = -\int dm\,zx$

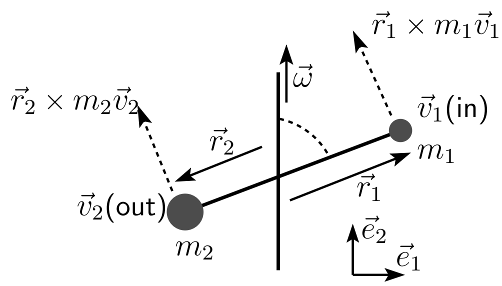
This provides the first column of the components, and you get the rest of the components the same way. The whole matrix is

$$\int dm \begin{pmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{pmatrix} \tag{7.17}$$

These are the components of the tensor of inertia. The diagonal elements of the matrix may be familiar; they are the moments of inertia. $x^2+y^2$ is the perpendicular distance-squared to the $z$-axis, so the element $I_{33}$ ($\equiv I_{zz}$) is the moment of inertia about that axis, $\int dm\, r_\perp^2$. The other components are less familiar and are called the products of inertia. This particular matrix is symmetric: $I_{ij} = I_{ji}$. That's a special property of the inertia tensor.

**Components of Dumbbell**

Look again at the specific case of two masses rotating about an axis. Do it quantitatively.




The integrals in Eq. (7.17) are simply sums this time, and the sums have just two terms. I'm making the approximation that these are point masses. Make the coordinate system match the indicated basis, with $x$ right and $y$ up, then $z$ is zero for all terms in the sum, and the rest are

$$\int dm\,(y^2+z^2) = m_1r_1^2\cos^2\alpha + m_2r_2^2\cos^2\alpha$$
$$-\int dm\,xy = -m_1r_1^2\cos\alpha\sin\alpha - m_2r_2^2\cos\alpha\sin\alpha$$
$$\int dm\,(x^2+z^2) = m_1r_1^2\sin^2\alpha + m_2r_2^2\sin^2\alpha$$
$$\int dm\,(x^2+y^2) = m_1r_1^2 + m_2r_2^2$$

The matrix is then

$$(I) = \big(m_1r_1^2 + m_2r_2^2\big)\begin{pmatrix} \cos^2\alpha & -\cos\alpha\sin\alpha & 0 \\ -\cos\alpha\sin\alpha & \sin^2\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7.18}$$

Don't count on all such results factoring so nicely.

In this basis, the angular velocity $\vec\omega$ has just one component, so what is $\vec L$?

$$(m_1r_1^2+m_2r_2^2)\begin{pmatrix}\cos^2\alpha & -\cos\alpha\sin\alpha & 0\\ -\cos\alpha\sin\alpha & \sin^2\alpha & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0\\ \omega\\ 0\end{pmatrix}= \\ (m_1r_1^2+m_2r_2^2)\begin{pmatrix}-\omega\cos\alpha\sin\alpha\\ \omega\sin^2\alpha\\ 0\end{pmatrix}$$

Translate this into vector form:

$$\vec L=(m_1r_1^2+m_2r_2^2)\omega\sin\alpha\big(-\vec e_1\cos\alpha+\vec e_2\sin\alpha\big) \tag{7.19}$$

When $\alpha=90^\circ$, then $\cos\alpha=0$ and the angular momentum points along the $y$-axis. This is the symmetric special case where everything lines up along one axis. Notice that if $\alpha=0$ then everything vanishes, but then the masses are both *on* the axis, and they have no angular momentum. In the general case as drawn, the vector $\vec L$ points to the upper left, perpendicular to the line between the masses.

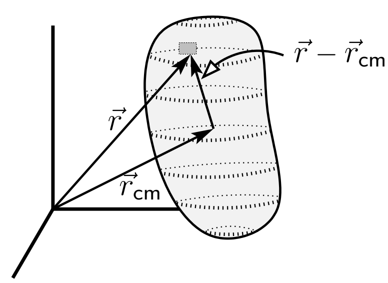
**Parallel Axis Theorem**

When you know the tensor of inertia about one origin, you can relate the result to the tensor about a different origin.

The center of mass of an object is

$$\vec r_{\text{cm}}=\frac{1}{M}\int\vec r\,dm \tag{7.20}$$

where $M$ is the total mass. Compare the operator $I$ using an origin at the center of mass to $I$ about another origin.



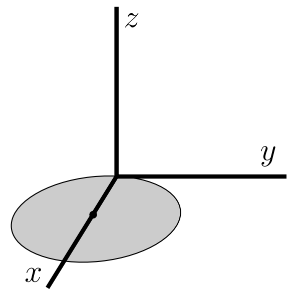

$$\begin{aligned}I(\vec\omega\,)&=\int dm\,\vec r\times(\vec\omega\times\vec r\,)=\int dm\,[\vec r-\vec r_{\text{cm}}+\vec r_{\text{cm}}]\times\big(\vec\omega\times[\vec r-\vec r_{\text{cm}}+\vec r_{\text{cm}}]\,\big)\\ &=\int dm\,[\vec r-\vec r_{\text{cm}}]\times\big(\vec\omega\times[\vec r-\vec r_{\text{cm}}]\big)+\int dm\,\vec r_{\text{cm}}\times\big(\vec\omega\times\vec r_{\text{cm}}\big)+\text{two cross terms}\end{aligned} \tag{7.21}$$

The two cross terms vanish, problem 7.17. What's left is

$$\begin{aligned}I(\vec\omega\,)&=\int dm\,[\vec r-\vec r_{\text{cm}}]\times\big(\vec\omega\times[\vec r-\vec r_{\text{cm}}]\big)+M\,r_{\text{cm}}\times\big(\vec\omega\times\vec r_{\text{cm}}\big)\\ &=I_{\text{cm}}(\vec\omega\,)+M\,\vec r_{\text{cm}}\times\big(\vec\omega\times\vec r_{\text{cm}}\big)\end{aligned} \tag{7.22}$$

Put this in words and it says that the tensor of inertia about any point is equal to the tensor of inertia about the center of mass plus the tensor of inertia of a point mass $M$ placed *at* the center of mass.

As an example, place a disk of mass $M$ and radius $R$ and uniform mass density so that its center is at $(x,y,z)=(R,0,0)$ and it is lying in the $x$-$y$ plane. Compute the components of the inertia tensor. First get the components about the center of mass, using Eq. (7.17).




The integrals such as

$$-\int dm\, xy, \qquad -\int dm\, yz$$

are zero. For fixed $y$ each positive value of $x$ has a corresponding negative value to make the integral add to zero. It is odd in $x$ (or $y$); remember that this is about the *center* of the disk. Next do the $I_{33}$ integral.

$$\int dm\,(x^2+y^2)=\int dm\, r^2=\int \frac{M}{\pi R^2}dA\, r^2$$

For the element of area, use $dA=2\pi r\,dr$ and you have

$$I_{33}=\frac{M}{\pi R^2}\int_0^R dr\, 2\pi r^3=\frac{M}{\pi R^2}2\pi\frac{R^4}{4}=\frac{1}{2}MR^2$$

For the next two diagonal elements,

$$I_{11}=\int dm\,(y^2+z^2)=\int dm\, y^2 \qquad \text{and} \qquad I_{22}=\int dm\,(x^2+z^2)=\int dm\, x^2$$

Because of the symmetry of the disk, these two are equal, also you see that the sum is

$$I_{11}+I_{22}=\int dm\, y^2+\int dm\, x^2=I_{33}=\frac{1}{2}MR^2 \tag{7.23}$$

This saves integration. $I_{11}=I_{22}=MR^2/4$.

For the other term in the sum (7.22), you have a point mass at the distance $R$ along the $x$-axis, $(x,y,z)=(R,0,0)$. Substitute this point mass into Eq. (7.17) and you have

$$M\begin{pmatrix}0 & 0 & 0\\ 0 & R^2 & 0\\ 0 & 0 & R^2\end{pmatrix}$$

The total about the origin is the sum of these two calculations.

$$MR^2\begin{pmatrix}1/4 & 0 & 0\\ 0 & 5/4 & 0\\ 0 & 0 & 3/2\end{pmatrix}$$

Why is this called the parallel axis theorem when you're translating a point (the origin) and not an axis? Probably because this was originally stated for the moment of inertia alone and not for the whole tensor. In that case you have only an axis to deal with.

### Components of the Derivative

The set of all polynomials in $x$ having degree $\le 2$ forms a vector space. There are three independent vectors that I can choose to be $1$, $x$, and $x^2$. Differentiation is a linear operator on this space because the derivative of a sum is the sum of the derivatives and the derivative of a constant times a function is the constant times the derivative of the function. With this basis I'll compute the components of $d/dx$. Start the indexing for the basis from zero instead of one because it will cause less confusion between powers and subscripts.

$$\vec e_0 = 1, \qquad \vec e_1 = x, \qquad \vec e_2 = x^2$$

By the definition of the components of an operator — I'll call this one $D$,

$$D(\vec e_0) = \frac{d}{dx}1 = 0, \qquad D(\vec e_1) = \frac{d}{dx}x = 1 = \vec e_0, \qquad D(\vec e_2) = \frac{d}{dx}x^2 = 2x = 2\vec e_1$$

These define the three columns of the matrix.

$$(D) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{check: } \frac{dx^2}{dx} = 2x \text{ is} \qquad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

There's nothing here about the basis being orthonormal. It isn't.

### 7.4 Matrix Multiplication

How do you multiply two matrices? There's a rule for doing it, but where does it come from?

The composition of two functions means you first apply one function then the other, so

$$h = f \circ g \qquad \text{means} \qquad h(\vec v\,) = f\big(g(\vec v\,)\big) \tag{7.24}$$

I'm assuming that these are vector-valued functions of a vector variable, but this is the general definition of composition anyway. If $f$ and $g$ are linear, does it follow the $h$ is? Yes, just check:

$$\begin{gathered} h(c\vec v\,) = f\big(g(c\vec v\,)\big) = f\big(c\,g(\vec v\,)\big) = c\,f\big(g(\vec v\,)\big), \qquad \text{and} \\ h(\vec v_1 + \vec v_2) = f\big(g(\vec v_1 + \vec v_2)\big) = f\big(g(\vec v_1) + g(\vec v_2)\big) = f\big(g(\vec v_1)\big) + f\big(g(\vec v_2)\big) \end{gathered}$$

What are the components of $h$? Again, use the definition and plug in.

$$h(\vec e_i) = \sum_k h_{ki}\,\vec e_k = f\big(g(\vec e_i)\big) = f\big(\sum_j g_{ji}\,\vec e_j\big) = \sum_j g_{ji} f\big(\vec e_j\big) = \sum_j g_{ji} \sum_k f_{kj}\,\vec e_k \tag{7.25}$$

and now all there is to do is to equate the corresponding coefficients of $\vec e_k$.

$$h_{ki} = \sum_j g_{ji} f_{kj} \qquad \text{or more conventionally} \qquad h_{ki} = \sum_j f_{kj} g_{ji} \tag{7.26}$$

This is in the standard form for matrix multiplication, recalling the subscripts are ordered as $f_{\mathrm{rc}}$ for row-column.

$$\begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{32} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{32} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{32} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} \tag{7.27}$$

The computation of $h_{12}$ from Eq. (7.26) is

$$\begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix}$$
$$\longrightarrow \quad h_{12} = f_{11}g_{12} + f_{12}g_{22} + f_{13}g_{32}$$

Matrix multiplication is just the component representation of the composition of two functions, Eq. (7.26), and there's nothing here that restricts this to three dimensions. In Eq. (7.25) I may have made it look too easy. If you try to reproduce this without looking, the odds are that you will not get the indices to match up as nicely as you see there. Remember: When an index is summed it is a *dummy*, and you are free to relabel it as anything you want. You can use this fact to make the indices come out neatly.

**Composition of Rotations**

In the first example, rotating vectors in the plane, the operator that rotates every vector by the angle $\alpha$ has components

$$(R_\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{7.28}$$

What happens if you do two such transformations, one by $\alpha$ and one by $\beta$? The result better be a total rotation by $\alpha+\beta$. One function, $R_\beta$ is followed by the second function $R_\alpha$ and the composition is

$$R_{\alpha+\beta} = R_\alpha R_\beta$$

This is mirrored in the components of these operators, so the matrices must obey the same equation.

$$\begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$

Multiply the matrices on the right to get

$$\begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{pmatrix} \tag{7.29}$$

The respective components must agree, so this gives an immediate derivation of the formulas for the sine and cosine of the sum of two angles. Cf. Eq. (3.8)

**7.5 Inverses**

The simplest operator is the one that does nothing. $f(\vec v) = \vec v$ for all values of the vector $\vec v$. This implies that $f(\vec e_1) = \vec e_1$ and similarly for all the other elements of the basis, so the matrix of its components is diagonal. The $2\times 2$ matrix is explicitly the identity matrix

$$(I) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{or in index notation} \qquad \delta_{ij} = \begin{cases} 1 & (\text{if } i = j) \\ 0 & (\text{if } i \neq j) \end{cases} \tag{7.30}$$

and the index notation is completely general, not depending on whether you're dealing with two dimensions or many more. Unfortunately the words "inertia" and "identity" both start with the letter "I" and this symbol is used for both operators. Live with it. The $\delta$ symbol in this equation is the Kronecker delta — very handy.

The *inverse* of an operator is defined in terms of Eq. (7.24), the composition of functions. If the composition of two functions takes you to the identity operator, one function is said to be the inverse of the other. This is no different from the way you look at ordinary real valued functions. The exponential and the logarithm are inverse to each other because*

$$\ln(e^x) = x \qquad \text{for all } x.$$

For the rotation operator, Eq. (7.10), the inverse is obviously going to be rotation by the same angle in the opposite direction.

$$R_\alpha R_{-\alpha} = I$$

Because the matrix components of these operators mirror the original operators, this equation must also hold for the corresponding components, as in Eqs. (7.27) and (7.29). Set $\beta = -\alpha$ in (7.29) and you get the identity matrix.

In an equation such as Eq. (7.7), or its component version Eqs. (7.8) or (7.9), if you want to solve for the vector $\vec u$, you are asking for the inverse of the function $f$.

$$\vec u = f(\vec v) \qquad \text{implies} \qquad \vec v = f^{-1}(\vec u)$$

The translation of these equations into components is Eq. (7.9)

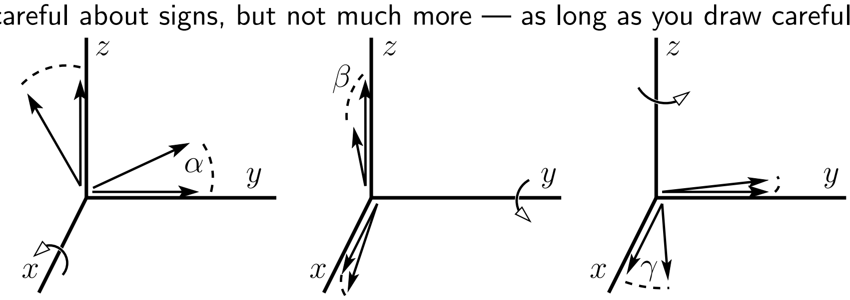
$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$\text{which implies} \qquad \frac{1}{f_{11}f_{22} - f_{12}f_{21}} \begin{pmatrix} f_{22} & -f_{12} \\ -f_{21} & f_{11} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{7.31}$$

The verification that these are the components of the inverse is no more than simply multiplying the two matrices and seeing that you get the identity matrix.

### 7.6 Rotations, 3-d

In three dimensions there are of course more basis vectors to rotate. Start by rotating vectors about the axes and it is nothing more than the two-dimensional problem of Eq. (7.10) done three times. You do have to be careful about signs, but not much more — as long as you draw careful pictures!




The basis vectors are drawn in the three pictures: $\vec e_1 = \hat x$, $\vec e_2 = \hat y$, $\vec e_3 = \hat z$.

In the first sketch, rotate vectors by the angle $\alpha$ about the $x$-axis. In the second case, rotate by the angle $\beta$ about the $y$-axis, and in the third case, rotate by the angle $\gamma$ about the $z$-axis. In the first case, the $\vec e_1$ is left alone. The $\vec e_2$ picks up a little positive $\vec e_3$, and the $\vec e_3$ picks up a little negative $\vec e_2$.

$$R_{\alpha\vec e_1}(\vec e_1) = \vec e_1, \quad R_{\alpha\vec e_1}(\vec e_2) = \vec e_2\cos\alpha + \vec e_3\sin\alpha, \quad R_{\alpha\vec e_1}(\vec e_3) = \vec e_3\cos\alpha - \vec e_2\sin\alpha \tag{7.32}$$

* The reverse, $e^{\ln x}$ works just for positive $x$, unless you recall that the logarithm of a negative number is complex. Then it works there too. This sort of question doesn't occur with finite dimensional matrices.

Here the notation $R_{\vec{\theta}}$ represents the function prescribing a rotation by $\theta$ about the axis pointing along $\hat{\theta}$. These equations are the same as Eqs. (7.11) and (7.12).

The corresponding equations for the other two rotations are now easy to write down:

$$R_{\beta\vec{e}_2}(\vec{e}_1) = \vec{e}_1\cos\beta - \vec{e}_3\sin\beta, \quad R_{\beta\vec{e}_2}(\vec{e}_2) = \vec{e}_2, \quad R_{\beta\vec{e}_2}(\vec{e}_3) = \vec{e}_1\sin\beta + \vec{e}_3\cos\beta \tag{7.33}$$

$$R_{\gamma\vec{e}_3}(\vec{e}_1) = \vec{e}_1\cos\gamma + \vec{e}_2\sin\gamma, \quad R_{\gamma\vec{e}_3}(\vec{e}_2) = -\vec{e}_1\sin\gamma + \vec{e}_2\cos\gamma, \quad R_{\gamma\vec{e}_3}(\vec{e}_3) = \vec{e}_3 \tag{7.34}$$

From these vector equations you immediate read the columns of the matrices of the components of the operators as in Eq. (7.6).

$$\begin{array}{ccc} (R_{\alpha\vec{e}_1}) & (R_{\beta\vec{e}_2}) & (R_{\gamma\vec{e}_3}) \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, & \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}, & \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \tag{7.35}$$

As a check on the algebra, did you see if the rotated basis vectors from any of the three sets of equations (7.32)-(7.34) are still orthogonal sets?

Do these rotation operations commute? No. Try the case of two $90^\circ$ rotations to see. Rotate by this angle about the $x$-axis then by the same angle about the $y$-axis.

$$(R_{\vec{e}_2\pi/2})(R_{\vec{e}_1\pi/2}) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix} \tag{7.36}$$

In the reverse order, for which the rotation about the $y$-axis is done first, these are

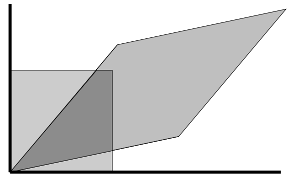
$$(R_{\vec{e}_1\pi/2})(R_{\vec{e}_2\pi/2}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{7.37}$$

Translate these operations into the movement of a physical object. Take the same $x$-$y$-$z$ coordinate system as in this section, with $x$ pointing toward you, $y$ to your right and $z$ up. Pick up a book with the cover toward you so that you can read it. Now do the operation $R_{\vec{e}_1\pi/2}$ on it so that the cover still faces you but the top is to your left. Next do $R_{\vec{e}_2\pi/2}$ and the book is face down with the top still to your left. See problem 7.57 for and algebraic version of this.

Start over with the cover toward you as before and do $R_{\vec{e}_2\pi/2}$ so that the top is toward you and the face is down. Now do the other operation $R_{\vec{e}_1\pi/2}$ and the top is toward you with the cover facing right — a different result. Do these physical results agree with the matrix products of the last two equations? For example, what happens to the vector sticking out of the cover, initially the column matrix $(\,1 \quad 0 \quad 0\,)$? This is something that you cannot simply *read*. You have to do the experiment for yourself.

## 7.7 Areas, Volumes, Determinants

In the two-dimensional example of arrows in the plane, look what happens to areas when an operator acts. The unit square with corners at the origin and $(0,1)$, $(1,1)$, $1,0)$ gets distorted into a parallelogram. The arrows from the origin to every point in the square become arrows that fill out the parallelogram.



What is the area of this parallelogram?

I'll ask a more general question. (It isn't really, but it looks like it.) Start with *any* region in the plane, and say it has area $A_1$. The operator takes all the vectors ending in this area into some new area of a size $A_2$, probably different from the original. What is the ratio of the new area to the old one? $A_2/A_1$. How much does this transformation stretch or squeeze the area? What isn't instantly obvious is that this ratio of areas depends on the operator *alone,* and not on how you chose the initial region to be transformed. If you accept this for the moment, then you see that the question in the previous paragraph, which started with the unit square and asked for the area into which it transformed, is the same question as finding the ratio of the two more general areas. (Or the ratio of two volumes in three dimensions.) See the end of the next section for a proof.

This ratio is called the determinant of the operator.

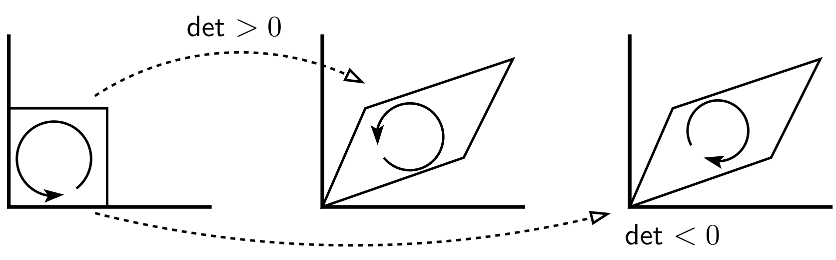
The first example is the simplest. Rotations in the plane, $R_\alpha$. Because rotations leave area unchanged, this determinant is one. For almost any other example you have to do some work. Use the component form to do the computation. The basis vector $\vec e_1$ is transformed into the vector $f_{11}\,\vec e_1 + f_{21}\,\vec e_2$ with a similar expression for the image of $\vec e_2$. You can use the cross product to compute the area of the parallelogram that these define. For another way, see problem 7.3. This is

$$(f_{11}\,\vec e_1 + f_{21}\,\vec e_2) \times (f_{12}\,\vec e_1 + f_{22}\,\vec e_2) = (f_{11}f_{22} - f_{21}f_{12})\vec e_3 \tag{7.38}$$

The product in parentheses is the determinant of the transformation.

$$\det(f) = f_{11}f_{22} - f_{21}f_{12} \tag{7.39}$$

What if I had picked a different basis, maybe even one that isn't orthonormal? From the definition of the determinant it is a property of the operator and not of the particular basis and components you use to describe it, so you must get the same answer. But will the answer be the same simple formula (7.39) if I pick a different basis? Now *that's* a legitimate question. The answer is yes, and that fact will come out of the general computation of the determinant in a moment. [What is the determinant of Eq. (7.13)?]




The determinant can be either positive or negative. That tells you more than simply how the transformation alters the area; it tells you whether it changes the *orientation* of the area. If you place a counterclockwise loop in the original area, does it remain counterclockwise in the image or is it reversed? In three dimensions, the corresponding plus or minus sign for the determinant says that you're changing from a right-handed set of vectors to a left-handed one. What does that mean? Make an $x$-$y$-$z$ coordinate system out of the thumb, index finger, and middle finger of your right hand. Now do it with your left hand. You cannot move one of these and put it on top of the other (unless you have *very* unusual joints). One is a mirror image of the other.

The equation (7.39) is a special case of a rule that you've probably encountered elsewhere. You compute the determinant of a square array of numbers by some means such as expansion in minors or Gauss reduction. Here I've defined the determinant geometrically, and it has no obvious relation the traditional numeric definition. They are the same, and the reason for that comes by looking at how the area (or volume) of a parallelogram depends on the vectors that make up its sides. The derivation is slightly involved, but no one step in it is hard. Along the way you will encounter a new and important function: $\Lambda$.

Start with the basis $\vec e_1$, $\vec e_2$ and call the output of the transformation $\vec v_1 = f(\vec e_1)$ and $\vec v_2 = f(\vec e_2)$. The final area is a function of these last two vectors, call it $\Lambda\big(\vec v_1, \vec v_2\big)$, and this function has two key properties:

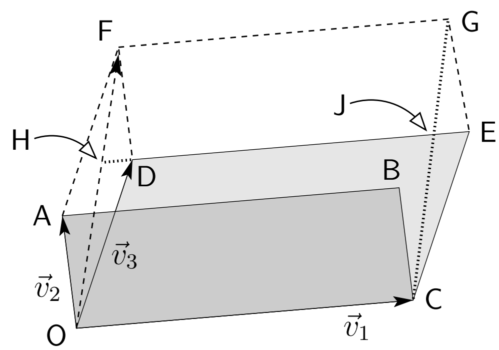
$$\Lambda\big(\vec v, \vec v\big) = 0, \qquad \text{and} \qquad \Lambda\big(\vec v_1, \alpha\vec v_2 + \beta\vec v_3\big) = \alpha\Lambda\big(\vec v_1, \vec v_2\big) + \beta\Lambda\big(\vec v_1, \vec v_3\big) \tag{7.40}$$

That the area vanishes if the two sides are the same is obvious. That the area is a linear function of the vectors forming the two sides is not so obvious. (It is linear in both arguments.) Part of the proof of linearity is easy:

$$\Lambda\big(\vec v_1, \alpha\vec v_2\big) = \alpha\Lambda\big(\vec v_1, \vec v_2\big)$$

simply says that if one side of the parallelogram remains fixed and the other changes by some factor, then the area changes by that same factor. For the other part, $\Lambda\big(\vec v_1, \vec v_2 + \vec v_3\big)$, start with a picture and see if the area that this function represents is the same as the sum of the two areas made from the vectors $\vec v_1 \& \vec v_2$ and $\vec v_1 \& \vec v_3$.

$\vec v_1 \& \vec v_2$ form the area OCBA. $\vec v_1 \& \vec v_3$ form the area OCED.




HDF $\cong$ JEG
HDO $\cong$ JEC
so
area HJGF = area DEGF = area OCBA
area OCJH = area OCED
add these equations:
area OCGF = area OCBA + area OCED

The last line is the statement that sum of the areas of the two parallelograms is the area of the parallelogram formed using the sum of the two vectors:

$$\Lambda\big(\vec v_1, \vec v_2 + \vec v_3\big) = \Lambda\big(\vec v_1, \vec v_2\big) + \Lambda\big(\vec v_1, \vec v_3\big)$$

This sort of function $\Lambda$, generalized to three dimensions, is characterized by

$$\begin{aligned} &(1) \qquad \Lambda\big(\alpha\vec v_1 + \beta\vec v_1{}', \vec v_2, \vec v_3\big) = \alpha\Lambda\big(\vec v_1, \vec v_2, \vec v_3\big) + \beta\Lambda\big(\vec v_1{}', \vec v_2, \vec v_3\big) \\ &(2) \qquad \Lambda\big(\vec v_1, \vec v_1, \vec v_3\big) = 0 \end{aligned} \tag{7.41}$$

It is linear in each variable, and it vanishes if any two arguments are equal. I've written it for three dimensions, but in $N$ dimensions you have the same equations with $N$ arguments, and these properties hold for all of them.

Theorem: Up to an overall constant factor, this function is unique.

An important result is that these assumptions imply the function is antisymmetric in any two arguments. Proof:

$$\Lambda\big(\vec v_1+\vec v_2,\vec v_1+\vec v_2,\vec v_3\big)=0=\Lambda\big(\vec v_1,\vec v_1,\vec v_3\big)+\Lambda\big(\vec v_1,\vec v_2,\vec v_3\big)+\Lambda\big(\vec v_2,\vec v_1,\vec v_3\big)+\Lambda\big(\vec v_2,\vec v_2,\vec v_3\big)$$

This is just the linearity property. Now the left side, and the $1^{\text{st}}$ and $4^{\text{th}}$ terms on the right, are zero because two arguments are equal. The rest is

$$\Lambda\big(\vec v_1,\vec v_2,\vec v_3\big)+\Lambda\big(\vec v_2,\vec v_1,\vec v_3\big)=0 \tag{7.42}$$

and this says that interchanging two arguments of $\Lambda$ changes the sign. (The reverse is true also. Assume antisymmetry and deduce that it vanishes if two arguments are equal.)

I said that this function is unique up to a factor. Suppose that there are two of them: $\Lambda$ and $\Lambda'$. Now show for some constant $\alpha$, that $\Lambda-\alpha\Lambda'$ is identically zero. To do this, take three independent vectors and evaluate the number $\Lambda'\big(\vec v_{\mathsf a},\vec v_{\mathsf b},\vec v_{\mathsf c}\big)$ There is some set of $\vec v$'s for which this is non-zero, otherwise $\Lambda'$ is identically zero and that's not much fun. Now consider

$$\alpha=\frac{\Lambda\big(\vec v_{\mathsf a},\vec v_{\mathsf b},\vec v_{\mathsf c}\big)}{\Lambda'\big(\vec v_{\mathsf a},\vec v_{\mathsf b},\vec v_{\mathsf c}\big)}\qquad\text{and define}\qquad\Lambda_0=\Lambda-\alpha\Lambda'$$

This function $\Lambda_0$ is zero for the special argument: $(\vec v_{\mathsf a},\vec v_{\mathsf b},\vec v_{\mathsf c})$, and now I'll show why it is zero for *all* arguments. That means that it is the zero function, and says that the two functions $\Lambda$ and $\Lambda'$ are proportional.

The vectors $(\vec v_{\mathsf a},\vec v_{\mathsf b},\vec v_{\mathsf c})$ are independent and there are three of them (in three dimensions). They are a basis. You can write any vector as a linear combination of these. *E.g.*

$$\vec v_1=A\vec v_{\mathsf a}+B\vec v_{\mathsf b}\qquad\text{and}\qquad\vec v_2=C\vec v_{\mathsf a}+D\vec v_{\mathsf b}\qquad\text{and}$$

Put these (and let's say $\vec v_{\mathsf c}$) into $\Lambda_0$.

$$\Lambda_0\big(\vec v_1,\vec v_2,v_{\mathsf c}\big)=AC\Lambda_0\big(\vec v_{\mathsf a},\vec v_{\mathsf a},\vec v_{\mathsf c}\big)+AD\Lambda_0\big(\vec v_{\mathsf a},\vec v_{\mathsf b},\vec v_{\mathsf c}\big)+BC\Lambda_0\big(\vec v_{\mathsf b},\vec v_{\mathsf a},\vec v_{\mathsf c}\big)+BD\Lambda_0\big(\vec v_{\mathsf b},\vec v_{\mathsf b},\vec v_{\mathsf c}\big)$$

*All these terms are zero.* Any argument that you put into $\Lambda_0$ is a linear combination of $\vec v_{\mathsf a}$, $\vec v_{\mathsf b}$, and $\vec v_{\mathsf c}$, and that means that this demonstration extends to any set of vectors, which in turn means that $\Lambda_0$ vanishes for any arguments. It is identically zero and that implies $\Lambda$ and $\Lambda'$ are, up to a constant overall factor, the same.

In $N$ dimensions, a scalar-valued function of $N$ vector variables, linear in each argument and antisymmetric under interchanging any pairs of arguments, is unique up to a factor.

I've characterized this volume function $\Lambda$ by two simple properties, and surprisingly enough this is all you need to *compute* it in terms of the components of the operator! With just this much information you can compute the determinant of a transformation.

Recall: $\vec v_1$ has for its components the first column of the matrix for the components of $f$, and $\vec v_2$ forms the second column. Adding any multiple of one vector to another leaves the volume alone. This is

$$\Lambda\big(\vec v_1,\vec v_2+\alpha\vec v_1,\vec v_3\big)=\Lambda\big(\vec v_1,\vec v_2,\vec v_3\big)+\alpha\Lambda\big(\vec v_1,\vec v_1,\vec v_3\big) \tag{7.43}$$

and the last term is zero. Translate this into components. Use the common notation for a determinant, a square array with vertical bars, but *forget that you know how to compute this symbol!* I'm going to use it simply as a notation by keep track of vector manipulations. The numerical value will come out at the end as the computed value of a volume. $\vec v_i = f\big(\vec e_i\big) = \sum_j f_{ji}\vec e_j$, then $\Lambda\big(\vec v_1, \vec v_2, \vec v_3\big) = \Lambda\big(\vec v_1, \vec v_2 + \alpha\vec v_1, \vec v_3\big) =$

$$\begin{vmatrix} f_{11} & f_{12}+\alpha f_{11} & f_{13} \\ f_{21} & f_{22}+\alpha f_{21} & f_{23} \\ f_{31} & f_{32}+\alpha f_{31} & f_{33} \end{vmatrix} = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} + \alpha \begin{vmatrix} f_{11} & f_{11} & f_{13} \\ f_{21} & f_{21} & f_{23} \\ f_{31} & f_{31} & f_{33} \end{vmatrix} = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}$$

To evaluate this object, simply choose $\alpha$ to make the element $f_{12} + \alpha f_{11} = 0$. Then repeat the operation, adding a multiple of the first column to the third, making the element $f_{13} + \beta f_{11} = 0$. This operation doesn't change the original value of $\Lambda\big(\vec v_1, \vec v_2, \vec v_3\big)$.

$$\Lambda\big(\vec v_1, \vec v_2 + \alpha\,\vec v_1, \vec v_3 + \beta\,\vec v_1\big) = \begin{vmatrix} f_{11} & 0 & 0 \\ f_{21} & f_{22}+\alpha f_{21} & f_{23}+\beta f_{21} \\ f_{31} & f_{32}+\alpha f_{31} & f_{33}+\beta f_{31} \end{vmatrix} = \begin{vmatrix} f_{11} & 0 & 0 \\ f_{21} & f'_{22} & f'_{23} \\ f_{31} & f'_{32} & f'_{33} \end{vmatrix}$$

Repeat the process to eliminate $f'_{23}$, adding $\gamma\,\vec v\,'_2$ to the third argument, where $\gamma = -f'_{23}/f'_{22}$.

$$= \begin{vmatrix} f_{11} & 0 & 0 \\ f_{21} & f'_{22} & f'_{23} \\ f_{31} & f'_{32} & f'_{33} \end{vmatrix} = \begin{vmatrix} f_{11} & 0 & 0 \\ f_{21} & f'_{22} & f'_{23}+\gamma f'_{22} \\ f_{31} & f'_{32} & f'_{33}+\gamma f'_{32} \end{vmatrix} = \begin{vmatrix} f_{11} & 0 & 0 \\ f_{21} & f'_{22} & 0 \\ f_{31} & f'_{32} & f''_{33} \end{vmatrix} \qquad (7.44)$$

Written in the last form, as a triangular array, the final result for the determinant *does not depend* on the elements $f_{21}$, $f_{31}$, $f'_{32}$. They may as well be zero. Why? Just do the same sort of column operations, but working toward the left. Eliminate $f_{31}$ and $f'_{32}$ by adding a constant times the third column to the first and second columns. Then eliminate $f_{21}$ by using the second column. You don't actually have to do this, you just have to recognize that it can be done so that you can ignore the lower triangular part of the array.

Translate this back to the original vectors and $\Lambda$ is unchanged:

$$\Lambda\big(\vec v_1, \vec v_2, \vec v_3\big) = \Lambda\big(f_{11}\,\vec e_1, f'_{22}\,\vec e_2, f''_{33}\,\vec e_3\big) = f_{11}f'_{22}f''_{33}\,\Lambda\big(\vec e_1, \vec e_2, \vec e_3\big)$$

The volume of the original box is $\Lambda\big(\vec e_1, \vec e_2, \vec e_3\big)$, so the quotient of the new volume to the old one is

$$\det = f_{11}f'_{22}f''_{33} \qquad (7.45)$$

The fact that $\Lambda$ is unique up to a constant factor doesn't matter. Do you want to measure volume in cubic feet, cubic centimeters, or cubic light-years? This algorithm is called Gauss elimination. It's development started with the geometry and used vector manipulations to recover what you may recognize from elsewhere as the traditional computed value of the determinant.

Did I leave anything out in this computation of the determinant? Yes, one point. What if in Eq. (7.44) the number $f'_{22} = 0$? You can't divide by it then. You can however interchange any two arguments of $\Lambda$, causing simply a sign change. If this contingency occurs then you need only interchange the two columns to get a component of zero where you want it. Just keep count of such switches whenever they occur.

### Trace

There's a property closely related to the determinant of an operator. It's called the trace. If you have an operator $f$, then consider the determinant of $M = I + \epsilon f$, where $I$ is the identity. This combination is very close to the identity if $\epsilon$ is small enough, so its determinant is very close to one. How close? The first order in $\epsilon$ is called the trace of $f$, or more formally

$$\operatorname{Tr}(f) = \frac{d}{d\epsilon} \det\bigl(I + \epsilon f\bigr)\bigg|_{\epsilon=0} \tag{7.46}$$

Express this in components for a two dimensional case, and

$$(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\Rightarrow\; \det\bigl(I + \epsilon f\bigr) = \det\begin{pmatrix} 1+\epsilon a & \epsilon b \\ \epsilon c & 1+\epsilon d \end{pmatrix} = (1+\epsilon a)(1+\epsilon d) - \epsilon^2 bc \tag{7.47}$$

The first order coefficient of $\epsilon$ is $a + d$, the sum of the diagonal elements of the matrix. This is the form of the result in any dimension, and the proof involves carefully looking at the method of Gauss elimination for the determinant, remembering at every step that you're looking for *only* the first order term in $\epsilon$. See problem 7.53.

### 7.8 Matrices as Operators

There's an important example of a vector space that I've avoided mentioning up to now. Example 5 in section 6.3 is the set of n-tuples of numbers: $(a_1, a_2, \ldots, a_n)$. I can turn this on its side, call it a column matrix, and it forms a perfectly good vector space. The functions (operators) on this vector space are the matrices themselves.

When you have a system of linear equations, you can translate this into the language of vectors.

$$ax + by = e \qquad \text{and} \qquad cx + dy = f \qquad \longrightarrow \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}$$

Solving for $x$ and $y$ is inverting a matrix.

There's an aspect of this that may strike you as odd. This matrix is an operator on the vector space of column matrices. What are the components of this operator? What? Isn't the matrix a set of components already? That depends on your choice of basis. Take an example

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{with basis} \qquad \vec e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \vec e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Compute the components as usual.

$$M\vec e_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 1\,\vec e_1 + 3\,\vec e_2$$

This says that the first column of the components of $M$ in this basis are $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$. What else would you expect? Now select a different basis.

$$\vec e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \vec e_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Again compute the component.

$$\begin{aligned} M\vec e_1 &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 5\,\vec e_1 - 2\,\vec e_2 \\ M\vec e_2 &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -1\,\vec e_1 \end{aligned}$$

$$\text{The components of } M \text{ in this basis are} \quad \begin{pmatrix} 5 & -1 \\ -2 & 0 \end{pmatrix}$$

It doesn't look at all the same, but it represents the same operator. Does this matrix have the same determinant, using Eq. (7.39)?

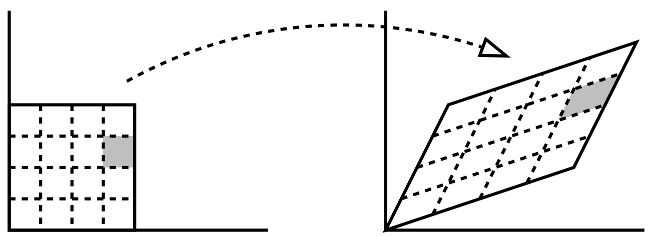
**Determinant of Composition**

If you do one linear transformation followed by another one, that is the composition of the two functions, each operator will then have its own determinant. What is the determinant of the composition? Let the operators be $F$ and $G$. One of them changes areas by a scale factor $\det(F)$ and the other ratio of areas is $\det(G)$. If you use the composition of the two functions, $FG$ or $GF$, the overall ratio of areas from the start to the finish will be the same:

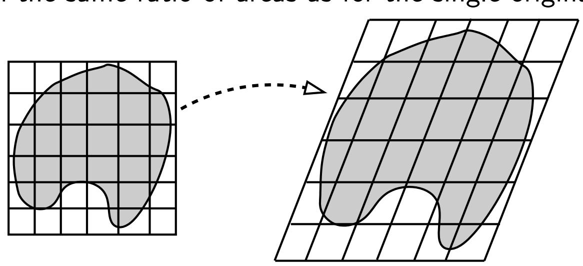
$$\det(FG) = \det(F)\cdot\det(G) = \det(G)\cdot\det(F) = \det(GF) \tag{7.48}$$

Recall that the determinant measures the ratio of areas for any input area, not just a square; it can be a parallelogram. The overall ratio of the product of the individual ratios, $\det(F)\det(G)$. The product of these two numbers is the total ratio of a new area to the original area and it is independent of the order of $F$ and $G$, so the determinant of the composition of the functions is also independent of order.

Now what about the statement that the definition of the determinant doesn't depend on the original area that you start with. To show this takes a couple of steps. First, start with a square that's not at the origin. You can always picture it as a piece of a square that *is* at the origin. The shaded square that is $1/16$ the area of the big square goes over to a parallelogram that's $1/16$ the area of the big parallelogram. Same ratio.



An arbitrary shape can be divided into a lot of squares. That's how you do an integral. The image of the whole area is distorted, but it retains the fact that a square that was inside the original area will become a parallelogram that is inside the new area. In the limit as the number of squares goes to infinity you still maintain the same ratio of areas as for the single original square.



## 7.9 Eigenvalues and Eigenvectors

There is a particularly important basis for an operator, the basis in which the components form a diagonal matrix. Such a basis almost always* exists, and it's easy to see *from the definition* as usual just what this basis must be.

$$f(\vec{e}_i) = \sum_{k=1}^{N} f_{ki}\vec{e}_k$$

* See section 7.12.

To be diagonal simply means that $f_{ki} = 0$ for all $i \neq k$, and that in turn means that all but one term in the sum disappears. This defining equation reduces to

$$f(\vec{e}_i) = f_{ii}\,\vec{e}_i \qquad \text{(with no sum this time)} \tag{7.49}$$

This is called an *eigenvalue* equation. It says that for any one of these special vectors, the operator $f$ on it returns a scalar multiple of that same vector. These multiples are called the eigenvalues, and the corresponding vectors are called the eigenvectors. The eigenvalues are then the diagonal elements of the matrix in this basis.

The inertia tensor is the function that relates the angular momentum of a rigid body to its angular velocity. The axis of rotation is defined by those points in the rotating body that aren't moving, and the vector $\vec{\omega}$ lies along that line. The angular momentum is computed from Eq. (7.3) and when you've done all those vector products and integrals you can't really expect the angular momentum to line up with $\vec{\omega}$ unless there is some exceptional reason for it. As the body rotates around the $\vec{\omega}$ axis, $\vec{L}$ will be carried with it, making $\vec{L}$ rotate about the direction of $\vec{\omega}$. The vector $\vec{L}$ is time-dependent and that implies there will be a torque necessary to keep it going, $\vec{\tau} = d\vec{L}/dt$. Because $\vec{L}$ is rotating with frequency $\omega$, this rotating torque will be felt as a vibration at this rotation frequency. If however the angular momentum happens to be parallel to the angular velocity, the angular momentum will not be changing; $d\vec{L}/dt = 0$ and the torque $\vec{\tau} = d\vec{L}/dt$ will be zero, implying that the vibrations will be absent. Have you ever taken your car in for servicing and asked the mechanic to make the angular momentum and the angular velocity vectors of the wheels parallel? It's called wheel-alignment.

How do you compute these eigenvectors? Just move everything to the left side of the preceding equation.

$$f(\vec{e}_i) - f_{ii}\,\vec{e}_i = 0, \qquad \text{or} \qquad (f - f_{ii}I)\vec{e}_i = 0$$

$I$ is the identity operator, output equals input. This notation is cumbersome. I'll change it.

$$f(\vec{v}) = \lambda\vec{v} \;\leftrightarrow\; (f - \lambda I)\vec{v} = 0 \tag{7.50}$$

$\lambda$ is the eigenvalue and $\vec{v}$ is the eigenvector. This operator $(f - \lambda I)$ takes some non-zero vector into the zero vector. In two dimensions then it will squeeze an area down to a line or a point. In three dimensions it will squeeze a volume down to an area (or a line or a point). In any case the ratio of the final area (or volume) to the initial area (or volume) is zero. That says the determinant is zero, and that's the key to computing the eigenvectors. Figure out which $\lambda$'s will make this determinant vanish.

Look back at section 4.9 and you'll see that the analysis there closely parallels what I'm doing here. In that case I didn't use the language of matrices or operators, but was asking about the possible solutions of two simultaneous linear equations.

$$ax + by = 0 \qquad \text{and} \qquad cx + dy = 0, \qquad \text{or} \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The explicit algebra there led to the conclusion that there can be a non-zero solution $(x, y)$ to the two equations only if the determinant of the coefficients vanishes, $ad - bc = 0$, and that's the same thing that I'm looking for here: a non-zero vector solution to Eq. (7.50).

Write the problem in terms of components, and of course you aren't yet in the basis where the matrix is diagonal. If you were, you're already done. The defining equation is $f(\vec{v}) = \lambda\vec{v}$, and in components this reads

$$\sum_i f_{ki} v_i = \lambda v_k, \qquad \text{or} \qquad \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \lambda\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Here I arbitrarily wrote the equation for three dimensions. That will change with the problem. Put everything on the left side and insert the components of the identity, the unit matrix.

$$\left[\begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{7.51}$$

The one way that this has a non-zero solution for the vector $\vec{v}$ is for the determinant of the whole matrix on the left-hand side to be zero. This equation is called the *characteristic equation* of the matrix, and in the example here that's a cubic equation in $\lambda$. If it has all distinct roots, no double roots, then you're guaranteed that this procedure will work, and you will be able to find a basis in which the components form a diagonal matrix. If this equation has a multiple root then there is no guarantee. It may work, but it may not; you have to look closer. See section 7.12. If the operator has certain symmetry properties then it's guaranteed to work. For example, the symmetry property found in problem 7.16 is enough to insure that you can find a basis in which the matrix for the inertia tensor is diagonal. It is even an orthogonal basis in that case.

**Example of Eigenvectors**

To keep the algebra to a minimum, I'll work in two dimensions and will specify an arbitrary but simple example:

$$f(\vec{e}_1) = 2\vec{e}_1 + \vec{e}_2, \quad f(\vec{e}_2) = 2\vec{e}_2 + \vec{e}_1 \qquad \text{with components} \qquad M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \tag{7.52}$$

The eigenvalue equation is, in component form

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad \text{or} \qquad \left[\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0 \tag{7.53}$$

The condition that there be a non-zero solution to this is

$$\det\left[\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = 0 = (2-\lambda)^2 - 1$$

The solutions to this quadratic are $\lambda = 1,\ 3$. For these values then, the apparently two equation for the two unknowns $v_1$ and $v_2$ are really one equation. The other is not independent. Solve this single equation in each case. Take the first of the two linear equations for $v_1$ and $v_2$ as defined by Eq. (7.53).

$$2v_1 + v_2 = \lambda v_1$$
$$\lambda = 1 \text{ implies} \qquad v_2 = -v_1, \qquad \lambda = 3 \text{ implies} \qquad v_2 = v_1$$

The two new basis vectors are then

$$\vec{e}\,'_1 = (\vec{e}_1 - \vec{e}_2) \qquad \text{and} \qquad \vec{e}\,'_2 = (\vec{e}_1 + \vec{e}_2) \tag{7.54}$$

and in this basis the matrix of components is the diagonal matrix of eigenvalues.

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

If you like to keep your basis vectors normalized, you may prefer to say that the new basis is $(\vec{e}_1 - \vec{e}_2)/\sqrt{2}$ and $(\vec{e}_1 + \vec{e}_2)/\sqrt{2}$. The eigenvalues are the same, so the new matrix is the same.

**Example: Coupled Oscillators**
Another example drawn from physics: Two masses are connected to a set of springs and fastened between two rigid walls. This is a problem that appeared in chapter 4, Eq. (4.45).

$$m_1 d^2x_1/dt^2 = -k_1x_1 - k_3(x_1 - x_2), \qquad \text{and} \qquad m_2 d^2x_2/dt^2 = -k_2x_2 - k_3(x_2 - x_1)$$

The exponential form of the solution was

$$x_1(t) = Ae^{i\omega t}, \qquad x_2(t) = Be^{i\omega t}$$

The algebraic equations that you get by substituting these into the differential equations are a pair of linear equations for $A$ and $B$, Eq. (4.47). In matrix form these equations are, after rearranging some minus signs,

$$\begin{pmatrix} k_1 + k_3 & -k_3 \\ -k_3 & k_2 + k_3 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \omega^2 \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}$$

You can make it look more like the previous example with some further arrangement

$$\left[ \begin{pmatrix} k_1 + k_3 & -k_3 \\ -k_3 & k_2 + k_3 \end{pmatrix} - \omega^2 \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \right] \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The matrix on the left side maps the column matrix to zero. That can happen only if the matrix has zero determinant (or the column matrix is zero). If you write out the determinant of this $2 \times 2$ matrix you have a quadratic equation in $\omega^2$. It's simple but messy, so rather than looking first at the general case, look at a special case with more symmetry. Take $m_1 = m_2 = m$ and $k_1 = k_2$.

$$\det \left[ \begin{pmatrix} k_1 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{pmatrix} - \omega^2 m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = 0 = \left(k_1 + k_3 - m\omega^2\right)^2 - k_3^2$$

This is now so simple that you don't even need the quadratic formula; it factors directly.

$$\left(k_1 + k_3 - m\omega^2 - k_3\right)\left(k_1 + k_3 - m\omega^2 + k_3\right) = 0$$

The only way that the product of two numbers is zero is if one of the numbers is zero, so either

$$k_1 - m\omega^2 = 0 \qquad \text{or} \qquad k_1 + 2k_3 - m\omega^2 = 0$$

This determines two possible frequencies of oscillation.

$$\omega_1 = \sqrt{\frac{k_1}{m}} \qquad \text{and} \qquad \omega_2 = \sqrt{\frac{k_1 + 2k_3}{m}}$$

You're not done yet; these are just the eigenvalues. You still have to find the eigenvectors and *then* go back to apply them to the original problem. This is $\vec F = m\vec a$ after all. Look back to section 4.10 for the development of the solutions.

### 7.10 Change of Basis

In many problems in physics and mathematics, the correct choice of basis can enormously simplify a problem. Sometimes the obvious choice of a basis turns out in the end not to be the best choice, and you then face the question: Do you start over with a new basis, or can you use the work that you've already done to transform everything into the new basis?

For linear transformations, this becomes the problem of computing the components of an operator in a new basis in terms of its components in the old basis.

First: Review how to do this for vector components, something that ought to be easy to do. The equation (7.5) defines the components with respect to a basis, *any* basis. If I have a second proposed basis, then by the definition of the word basis, every vector in that second basis can be written as a linear combination of the vectors in the first basis. I'll call the vectors in the first basis, $\vec e_i$ and those in the second basis $\vec e_i^{\,\prime}$, for example in the plane you could have

$$\vec e_1 = \hat x, \quad \vec e_2 = \hat y, \qquad \text{and} \qquad \vec e_1^{\,\prime} = 2\hat x + 0.5\hat y, \quad \vec e_2^{\,\prime} = 0.5\hat x + 2\hat y \tag{7.55}$$

Each vector $\vec e_i^{\,\prime}$ is a linear combination* of the original basis vectors:

$$\vec e_i^{\,\prime} = S(\vec e_i) = \sum_j S_{ji}\vec e_j \tag{7.56}$$

This follows the standard notation of Eq. (7.6); you have to put the indices in this order in order to make the notation come out right in the end. One vector expressed in two different bases is still one vector, so

$$\vec v = \sum_i v_i'\vec e_i^{\,\prime} = \sum_i v_i\vec e_i$$

and I'm using the fairly standard notation of $v_i'$ for the $i^{\text{th}}$ component of the vector $\vec v$ with respect to the second basis. Now insert the relation between the bases from the preceding equation (7.56).

$$\vec v = \sum_i v_i'\sum_j S_{ji}\vec e_j = \sum_j v_j\vec e_j$$

and this used the standard trick of changing the last dummy label of summation from $i$ to $j$ so that it is easy to compare the components.

$$\sum_i S_{ji}v_i' = v_j \qquad \text{or in matrix notation} \qquad (S)(v') = (v), \implies (v') = (S)^{-1}(v)$$

#### Similarity Transformations

Now use the definition of the components of an operator to get the components in the new basis.

$$f\big(\vec e_i^{\,\prime}\big) = \qquad\qquad\qquad\qquad = \sum_j f_{ji}'\vec e_j^{\,\prime}$$

$$f\Big(\sum_j S_{ji}\vec e_j\Big) = \sum_j S_{ji}f\big(\vec e_j\big) = \sum_j S_{ji}\sum_k f_{kj}\vec e_k = \sum_j f_{ji}'\sum_k S_{kj}\vec e_k$$

* There are two possible conventions here. You can write $\vec e_i^{\,\prime}$ in terms of the $\vec e_i$, calling the coefficients $S_{ji}$, or you can do the reverse and call *those* components $S_{ji}$. [$\vec e_i = S(\vec e_i^{\,\prime})$] Naturally, both conventions are in common use. The reverse convention will interchange the roles of the matrices $S$ and $S^{-1}$ in what follows.

The final equation comes from the preceding line. The coefficients of $\vec{e}_k$ must agree on the two sides of the equation.

$$\sum_j S_{ji} f_{kj} = \sum_j f'_{ji} S_{kj}$$

Now rearrange this in order to place the indices in their conventional $_{\text{row,column}}$ order.

$$\begin{gathered}\sum_j S_{kj} f'_{ji} = \sum_j f_{kj} S_{ji} \\ \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} f'_{11} & f'_{12} \\ f'_{21} & f'_{22} \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\end{gathered} \tag{7.57}$$

In turn, this matrix equation is usually written in terms of the inverse matrix of $S$,

$$(S)(f') = (f)(S) \qquad \text{is} \qquad (f') = (S)^{-1}(f)(S) \tag{7.58}$$

and this is called a similarity transformation. For the example Eq. (7.55) this is

$$\vec{e}\,'_1 = 2\hat{x} + 0.5\hat{y} = S_{11}\vec{e}_1 + S_{21}\vec{e}_2$$

which determines the first column of the matrix $(S)$, then $\vec{e}\,'_2$ determines the second column.

$$(S) = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 2 \end{pmatrix} \qquad \text{then} \qquad (S)^{-1} = \frac{1}{3.75}\begin{pmatrix} 2 & -0.5 \\ -0.5 & 2 \end{pmatrix}$$

**Eigenvectors**

In defining eigenvalues and eigenvectors I pointed out the utility of having a basis in which the components of an operator form a diagonal matrix. Finding the non-zero solutions to Eq. (7.50) is then the way to find the basis in which this holds. Now I've spent time showing that you can find a matrix in a new basis by using a similarity transformation. Is there a relationship between these two subjects? Another way to ask the question: I've solved the problem to find all the eigenvectors and eigenvalues, so what is the similarity transformation that accomplishes the change of basis (and why is it necessary to know it if I already know that the transformed, diagonal matrix is just the set of eigenvalues, and I already know them)?

For the last question, the simplest answer is that you *don't* need to know the explicit transformation once you already know the answer. It is however useful to know that it exists and how to construct it. *If* it exists — I'll come back to that presently. Certain manipulations are more easily done in terms of similarity transformations, so you ought to know how they are constructed, especially because almost all the work in constructing them is done when you've found the eigenvectors.

The equation (7.57) tells you the answer. Suppose that you want the transformed matrix to be diagonal. That means that $f'_{12} = 0$ and $f'_{21} = 0$. Write out the first column of the product on the right.

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \longrightarrow \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} S_{11} \\ S_{21} \end{pmatrix}$$

This equals the first column on the left of the same equation

$$f'_{11} \begin{pmatrix} S_{11} \\ S_{21} \end{pmatrix}$$

This is the eigenvector equation that you've supposedly already solved. The first column of the component matrix of the similarity transformation is simply the set of components of the first eigenvector. When you write out the second column of Eq. (7.57) you'll see that it's the defining equation for the second eigenvector. You already know these, so you can immediately write down the matrix for the similarity transformation.

For the example Eq. (7.52) the eigenvectors are given in Eq. (7.54). In components these are

$$\vec e_1' \to \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad \text{and} \qquad \vec e_2' \to \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \text{implying} \qquad S = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

The inverse to this matrix is

$$S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

You should verify that $S^{-1}MS$ is diagonal.

### 7.11 Summation Convention

In all the manipulation of components of vectors and components of operators you have to do a lot of sums. There are so many sums over indices that a convention* was invented (by Einstein) to simplify the notation.

A repeated index in a term is summed.

Eq. (7.6) becomes $f(\vec e_i) = f_{ki}\vec e_k$.
Eq. (7.8) becomes $u_k = f_{ki}v_i$.
Eq. (7.26) becomes $h_{ki} = f_{kj}g_{ji}$.
$IM = M$ becomes $\delta_{ij}M_{jk} = M_{ik}$.

What if there are *three* identical indices in the same term? Then you made a mistake; that can't happen. What about Eq. (7.49)? That has three indices. Yes, and there I explicitly said that there is no sum. This sort of rare case you have to handle as an exception.

### 7.12 Can you Diagonalize a Matrix?

At the beginning of section 7.9 I said that the basis in which the components of an operator form a diagonal matrix "almost always exists." There's a technical sense in which this is precisely true, but that's not what you need to know in order to manipulate matrices; the theorem that you need to have is that every matrix is the limit of a sequence of diagonalizable matrices. If you encounter a matrix that cannot be diagonalized, then you can approximate it as closely as you want by a matrix that can be diagonalized, do your calculations, and finally take a limit. You already did this if you did problem 4.11, but in that chapter it didn't look anything like a problem involving matrices, much less diagonalization of matrices. Yet it is the same.

Take the matrix

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

You can't diagonalize this. If you try the standard procedure, here is what happens:

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad \text{then} \qquad \det\begin{pmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{pmatrix} = 0 = (1-\lambda)^2$$

The resulting equations you get for $\lambda = 1$ are

$$0v_1 + 2v_2 = 0 \qquad \text{and} \qquad 0 = 0$$

---

* There is a modification of this convention that appears in chapter 12, section 12.5

This provides only one eigenvector, a multiple of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. You need two for a basis.

Change this matrix in any convenient way to make the two roots of the characteristic equation different from each other. For example,

$$M_\epsilon = \begin{pmatrix} 1+\epsilon & 2 \\ 0 & 1 \end{pmatrix}$$

The eigenvalue equation is now

$$(1+\epsilon-\lambda)(1-\lambda) = 0$$

and the resulting equations for the eigenvectors are

$$\lambda = 1: \ \epsilon v_1 + 2v_2 = 0, \quad 0 = 0 \qquad \lambda = 1+\epsilon: \ 0v_1 + 2v_2 = 0, \quad \epsilon v_2 = 0$$

Now you have two distinct eigenvectors,

$$\lambda = 1: \ \begin{pmatrix} 1 \\ -\epsilon/2 \end{pmatrix}, \qquad \text{and} \qquad \lambda = 1+\epsilon: \ \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

You see what happens to these vectors as $\epsilon \to 0$.

**Differential Equations at Critical**

Problem 4.11 was to solve the damped harmonic oscillator for the critical case that $b^2 - 4km = 0$.

$$m\frac{d^2x}{dt^2} = -kx - b\frac{dx}{dt} \tag{7.59}$$

Write this as a pair of equations, using the velocity as an independent variable.

$$\frac{dx}{dt} = v_x \qquad \text{and} \qquad \frac{dv_x}{dt} = -\frac{k}{m}x - \frac{b}{m}v_x$$

In matrix form, this is a matrix differential equation.

$$\frac{d}{dt}\begin{pmatrix} x \\ v_x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -k/m & -b/m \end{pmatrix}\begin{pmatrix} x \\ v_x \end{pmatrix}$$

This is a linear, constant-coefficient differential equation, but now the constant coefficients are matrices. Don't let that slow you down. The reason that an exponential form of solution works is that the derivative of an exponential is an exponential. Assume such a solution here.

$$\begin{pmatrix} x \\ v_x \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} e^{\alpha t}, \qquad \text{giving} \qquad \alpha\begin{pmatrix} A \\ B \end{pmatrix} e^{\alpha t} = \begin{pmatrix} 0 & 1 \\ -k/m & -b/m \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} e^{\alpha t} \tag{7.60}$$

When you divide the equation by $e^{\alpha t}$, you're left with an eigenvector equation where the eigenvalue is $\alpha$. As usual, to get a non-zero solution set the determinant of the coefficients to zero and the characteristic equation is

$$\det\begin{pmatrix} 0-\alpha & 1 \\ -k/m & -b/m-\alpha \end{pmatrix} = \alpha(\alpha + b/m) + k/m = 0$$

with familiar roots

$$\alpha = \big(-b \pm \sqrt{b^2 - 4km}\big)/2m$$

If the two roots are equal you *may* not have distinct eigenvectors, and in this case you *do* not. No matter, you can solve any such problem for the case that $b^2 - 4km \neq 0$ and then take the limit as this approaches zero.

The eigenvectors come from the either one of the two equations represented by Eq. (7.60). Pick the simpler one, $\alpha A = B$. The column matrix $\begin{pmatrix} A \\ B \end{pmatrix}$ is then $A \begin{pmatrix} 1 \\ \alpha \end{pmatrix}$.

$$\begin{pmatrix} x \\ v_x \end{pmatrix}(t) = A_+ \begin{pmatrix} 1 \\ \alpha_+ \end{pmatrix} e^{\alpha_+ t} + A_- \begin{pmatrix} 1 \\ \alpha_- \end{pmatrix} e^{\alpha_- t}$$

Pick the initial conditions that $x(0) = 0$ and $v_x(0) = v_0$. You must choose *some* initial conditions in order to apply this technique. In matrix terminology this is

$$\begin{pmatrix} 0 \\ v_0 \end{pmatrix} = A_+ \begin{pmatrix} 1 \\ \alpha_+ \end{pmatrix} + A_- \begin{pmatrix} 1 \\ \alpha_- \end{pmatrix}$$

These are two equations for the two unknowns

$$A_+ + A_- = 0, \qquad \alpha_+ A_+ + \alpha_- A_- = v_0, \qquad \text{so} \qquad A_+ = \frac{v_0}{\alpha_+ - \alpha_-}, \quad A_- = -A_+$$

$$\begin{pmatrix} x \\ v_x \end{pmatrix}(t) = \frac{v_0}{\alpha_+ - \alpha_-}\left[\begin{pmatrix} 1 \\ \alpha_+ \end{pmatrix} e^{\alpha_+ t} - \begin{pmatrix} 1 \\ \alpha_- \end{pmatrix} e^{\alpha_- t}\right]$$

If you now take the limit as $b^2 \to 4km$, or equivalently as $\alpha_- \to \alpha_+$, this expression is just the definition of a derivative.

$$\begin{pmatrix} x \\ v_x \end{pmatrix}(t) \longrightarrow v_0 \frac{d}{d\alpha}\begin{pmatrix} 1 \\ \alpha \end{pmatrix} e^{\alpha t} = v_0 \begin{pmatrix} te^{\alpha t} \\ (1+\alpha t)e^{\alpha t} \end{pmatrix} \qquad \alpha = -\frac{b}{2m} \tag{7.61}$$

### 7.13 Eigenvalues and Google

The motivating idea behind the search engine Google is that you want the first items returned by a search to be the most important items. How do you do this? How do you program a computer to decide which web sites are the most important?

A simple idea is to count the number of sites that contain a link to a given site, and the site that is linked to the most is then the most important site. This has the drawback that all links are treated as equal. If your site is referenced from the home page of Al Einstein, it counts no more than if it's referenced by Joe Blow. This shouldn't be.

A better idea is to assign each web page a numerical importance rating. If your site, #1, is linked from sites #11, #59, and #182, then your rating, $x_1$, is determined by adding those ratings (and multiplying by a suitable scaling constant).

$$x_1 = C\big(x_{11} + x_{59} + x_{182}\big)$$

Similarly the second site's rating is determined by what links to it, as

$$x_2 = C\big(x_{137} + x_{157983} + x_1 + x_{876}\big)$$

But this assumes that you already know the ratings of the sites, and that's what you're trying to find!

Write this in matrix language. Each site is an element in a huge column matrix $\{x_i\}$.

$$x_i = C\sum_{j=1}^{N} \alpha_{ij}x_j \qquad \text{or} \qquad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix} = C\begin{pmatrix} 0 & 0 & 1 & 0 & 1 & \dots \\ 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 1 & 1 & \dots \\ \dots \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix}$$

An entry of 1 indicates a link and a 0 is no link. This is an eigenvector problem with the eigenvalue $\lambda = 1/C$, and though there are many eigenvectors, there is a constraint that lets you pick the right one. All the $x_i$s must be non-negative, and there's a theorem (Perron-Frobenius) guaranteeing that you can find such an eigenvector. This algorithm is a key idea behind Google's ranking methods. They have gone well beyond this basic technique of course, but the spirit of the method remains.

See www-db.stanford.edu/~backrub/google.html for more on this.

### 7.14 Special Operators

Symmetric
Antisymmetric
Hermitian
Antihermitian
Orthogonal
Unitary
Idempotent
Nilpotent
Self-adjoint

In no particular order of importance, these are names for special classes of operators. It is often the case that an operator defined in terms of a physical problem will be in some way special, and it's then worth knowing the consequent simplifications. The first ones involve a scalar product.

**Symmetric**: $\big\langle \vec u, S(\vec v\,)\big\rangle = \big\langle S(\vec u\,), \vec v\,\big\rangle$
**Antisymmetric**: $\big\langle \vec u, A(\vec v\,)\big\rangle = -\big\langle A(\vec u\,), \vec v\,\big\rangle$

The inertia operator of Eq. (7.3) is symmetric.

$$I(\vec\omega\,) = \int dm\,\vec r\times\big(\vec\omega\times\vec r\,\big) \quad \text{satisfies} \quad \big\langle\vec\omega_1, I(\vec\omega_2)\big\rangle = \vec\omega_1\cdot I(\vec\omega_2) = \big\langle I(\vec\omega_1), \vec\omega_2\,\big\rangle = I(\vec\omega_1)\cdot\vec\omega_2$$

Proof: Plug in.

$$\begin{aligned}\vec\omega_1\cdot I(\vec\omega_2) &= \vec\omega_1\cdot\int dm\,\vec r\times\big(\vec\omega_2\times\vec r\,\big) = \vec\omega_1\cdot\int dm\,\big[\vec\omega_2\,r^2 - \vec r(\vec\omega_2\cdot\vec r\,)\big] \\ &= \int dm\,\big[\vec\omega_1\cdot\vec\omega_2\,r^2 - (\vec\omega_1\cdot\vec r\,)(\vec\omega_2\cdot\vec r\,)\big] = I(\vec\omega_1)\cdot\vec\omega_2\end{aligned}$$

What good does this do? You will be guaranteed that all eigenvalues are real, all eigenvectors are orthogonal, and the eigenvectors form an orthogonal basis. In this example, the eigenvalues are moments of inertia about the axes defined by the eigenvectors, so these moments better be real. The magnetic field operator (problem 7.28) is antisymmetric.

**Hermitian** operators obey the same identity as symmetric: $\big\langle\vec u, H(\vec v\,)\big\rangle = \big\langle H(\vec u\,), \vec v\,\big\rangle$. The difference is that in this case you allow the scalars to be complex numbers. That means that the scalar product has a complex conjugation implied in the first factor. You saw this sort of operator in the chapter on Fourier series, section 5.3, but it didn't appear under this name. You will become familiar with this class of operators when you hit quantum mechanics. Then they are ubiquitous. The same theorem as for symmetric operators applies here, that the eigenvalues are real and that the eigenvectors are orthogonal.

**Orthogonal** operators satisfy $\big\langle O(\vec{u}\,), O(\vec{v}\,)\big\rangle = \big\langle \vec{u}, \vec{v}\,\big\rangle$. The most familiar example is rotation. When you rotate two vectors, their magnitudes and the angle between them do not change. That's all that this equation says — scalar products are preserved by the transformation.

**Unitary** operators are the complex analog of orthogonal ones: $\big\langle U(\vec{u}), U(\vec{v})\big\rangle = \big\langle \vec{u}, \vec{v}\,\big\rangle$, but all the scalars are complex and the scalar product is modified accordingly.

The next couple you don't see as often. **Idempotent** means that if you take the square of the operator, it equals the original operator.

**Nilpotent** means that if you take successive powers of the operator you eventually reach the zero operator.

**Self-adjoint** in a finite dimensional vector space is exactly the same thing as Hermitian. In infinite dimensions it is not, and in quantum mechanics the important operators are the self-adjoint ones. The issues involved are a bit technical. As an aside, in infinite dimensions you need one extra hypothesis for unitary and orthogonal: that they are invertible.

## Problems

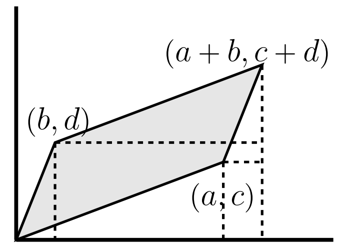
**7.1** Draw a picture of the effect of these linear transformations on the unit square with vertices at $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$. The matrices representing the operators are

$$\text{(a)} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad \text{(b)} \begin{pmatrix} 1 & -2 \\ 2 & -4 \end{pmatrix}, \qquad \text{(c)} \begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix}$$

Is the orientation preserved or not in each case? See the figure at the end of section 7.7

**7.2** Using the same matrices as the preceding question, what is the picture resulting from doing (a) followed by (c)? What is the picture resulting from doing (c) followed by (a)? The results of section 7.4 may prove helpful.

**7.3** Look again at the parallelogram that is the image of the unit square in the calculation of the determinant. In Eq. (7.39) I used the cross product to get its area, but sometimes a brute-force method is more persuasive. If the transformation has components $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ The corners of the parallelogram that is the image of the unit square are at $(0,0)$, $(a,c)$, $(a+b,c+d)$, $(b,d)$. You can compute its area as sums and differences of rectangles and triangles. Do so; it should give the same result as the method that used a cross product.




**7.4** In three dimensions, there is an analogy to the geometric interpretation of the cross product as the area of a parallelogram. The triple scalar product $\vec{A}\cdot\vec{B}\times\vec{C}$ is the volume of the parallelepiped having these three vectors as edges. Prove both of these statements starting from the geometric definitions of the two products. That is, from the $AB\cos\theta$ and $AB\sin\theta$ definitions of the dot product and the magnitude of the cross product (and its direction).

**7.5** Derive the relation $\vec{v} = \vec{\omega}\times\vec{r}$ for a point mass rotating about an axis. Refer to the figure before Eq. (7.2).

**7.6** You have a mass attached to four springs in a plane and that are in turn attached to four walls as on page 3; the mass is at equilibrium. Two opposing spring have spring constant $k_1$ and the other two are $k_2$. Push on the mass with a (small) force $\vec{F}$ and the resulting displacement of $m$ is $\vec{d} = f(\vec{F})$, defining a linear operator. Compute the components of $f$ in an obvious basis and check a couple of special cases to see if the displacement is in a plausible direction, especially if the two $k$'s are quite different.

**7.7** On the vector space of quadratic polynomials, degree $\le 2$, the operator $d/dx$ is defined: the derivative of such a polynomial is a polynomial. (**a**) Use the basis $\vec{e}_0 = 1$, $\vec{e}_1 = x$, and $\vec{e}_2 = x^2$ and compute the components of this operator. (**b**) Compute the components of the operator $d^2/dx^2$. (**c**) Compute the square of the first matrix and compare it to the result for (**b**). Ans: $(\mathbf{a})^2 = (\mathbf{b})$

**7.8** Repeat the preceding problem, but look at the case of cubic polynomials, a four-dimensional space.

**7.9** In the preceding problem the basis $1$, $x$, $x^2$, $x^3$ is too obvious. Take another basis, the Legendre polynomials:

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, \qquad P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$$

and repeat the problem, finding components of the first and second derivative operators. Verify an example explicitly to check that your matrix reproduces the effect of differentiation on a polynomial of your choice. Pick one that will let you test your results.

**7.10** What is the determinant of the inverse of an operator, explaining why?
Ans: $1/\det(\text{original operator})$

**7.11** Eight identical point masses $m$ are placed at the corners of a cube that has one corner at the origin of the coordinates and has its sides along the axes. The side of the cube is length $= a$. In the basis that is placed along the axes as usual, compute the components of the inertia tensor. Ans: $I_{11} = 8ma^2$

**7.12** For the dumbbell rotating about the off-axis axis in Eq. (7.19), what is the time-derivative of $\vec{L}$? In very short time $dt$, what new direction does $\vec{L}$ take and what then is $d\vec{L}$? That will tell you $d\vec{L}/dt$. Prove that this is $\vec{\omega} \times \vec{L}$.

**7.13** A cube of uniform volume mass density, mass $m$, and side $a$ has one corner at the origin of the coordinate system and the adjacent edges are placed along the coordinate axes. Compute the components of the tensor of inertia. Do it (**a**) directly and (**b**) by using the parallel axis theorem to check your result.
Ans: $ma^2 \begin{pmatrix} 2/3 & -1/4 & -1/4 \\ -1/4 & 2/3 & -1/4 \\ -1/4 & -1/4 & 2/3 \end{pmatrix}$

**7.14** Compute the cube of Eq. (7.13) to find the trigonometric identities for the cosine and sine of triple angles in terms of single angle sines and cosines. Compare the results of problem 3.9.

**7.15** On the vectors of column matrices, the operators are matrices. For the two dimensional case take $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and find its components in the basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
What is the determinant of the resulting matrix? Ans: $M_{11} = (a+b+c+d)/2$, and the determinant is *still* $ad - bc$.

**7.16** Show that the tensor of inertia, Eq. (7.3), satisfies $\vec{\omega}_1 \cdot I(\vec{\omega}_2) = I(\vec{\omega}_1) \cdot \vec{\omega}_2$. What does this identity tell you about the components of the operator when you use the ordinary orthonormal basis? First determine in such a basis what $\vec{e}_1 \cdot I(\vec{e}_2)$ is.

**7.17** Use the definition of the center of mass to show that the two cross terms in Eq. (7.21) are zero.

**7.18** Prove the Perpendicular Axis Theorem. This says that for a mass that lies flat in a plane, the moment of inertia about an axis perpendicular to the plane equals the sum of the two moments of inertia about the two perpendicular axes that lie in the plane and that intersect the third axis.

**7.19** Verify in the conventional, non-matrix way that Eq. (7.61) really does provide a solution to the original second order differential equation (7.59).

**7.20** The Pauli spin matrices are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Show that $\sigma_x\sigma_y = i\sigma_z$ and the same for cyclic permutations of the indices $x$, $y$, $z$. Compare the products $\sigma_x\sigma_y$ and $\sigma_y\sigma_x$ and the other pairings of these matrices.

**7.21** Interpret $\vec{\sigma}\cdot\vec{A}$ as $\sigma_x A_x + \sigma_y A_y + \sigma_z A_z$ and prove that

$$\vec{\sigma}\cdot\vec{A}\,\vec{\sigma}\cdot\vec{B} = \vec{A}\cdot\vec{B} + i\vec{\sigma}\cdot\vec{A}\times\vec{B}$$

where the first term on the right has to include the identity matrix for this to make sense.

**7.22** Evaluate the matrix

$$\frac{I}{I - \vec{\sigma}\cdot\vec{A}} = \big(I - \vec{\sigma}\cdot\vec{A}\big)^{-1}$$

Evaluate this by two methods: (**a**) You may assume that $\vec{A}$ is in some sense small enough for you to manipulate by infinite series methods. This then becomes a geometric series that you can sum. Use the results of the preceding problem.
(**b**) You can manipulate the algebra directly without series. I suggest that you recall the sort of manipulation that allows you to write the complex number $1/(1-i)$ without any $i$'s in the denominator. I suppose you could do it a third way, writing out the $2\times 2$ matrix and explicitly inverting it, but I definitely don't recommend this.

**7.23** Evaluate the sum of the infinite series defined by $e^{-i\sigma_y\theta}$. Where have you seen this result before? The first term in the series must be interpreted as the identity matrix. Ans: $I\cos\theta - i\sigma_y\sin\theta$

**7.24** For the moment of inertia about an axis, the integral is $\int r_\perp^2\,dm$. State precisely what this $m$ function must be for this to make sense as a Riemann-Stieltjes integral, Eq. (1.28). For the case that you have eight masses, all $m_0$ at the 8 corners of a cube of side $a$, write explicitly what this function is and evaluate the moment of inertia about an axis along one edge of the cube.

**7.25** The summation convention allows you to write some compact formulas. Evaluate these, assuming that you're dealing with three dimensions. Note Eq. (7.30). Define the alternating symbol $\epsilon_{ijk}$ to be
1: It is totally anti-symmetric. That is, interchange any two indices and you change the sign of the value.
2: $\epsilon_{123} = 1$. [*E.g.* $\epsilon_{132} = -1$, $\epsilon_{312} = +1$]

$$\delta_{ii}, \qquad \epsilon_{ijk}A_jB_k, \qquad \delta_{ij}\epsilon_{ijk}, \qquad \delta_{mn}A_mB_n, \qquad S_{mn}u_mv_n, \qquad u_nv_n,$$
$$\epsilon_{ijk}\epsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}$$

Multiply the last identity by $A_jB_mC_n$ and interpret.

**7.26** The set of Hermite polynomials starts out as

$$H_0 = 1, \qquad H_1 = 2x, \qquad H_2 = 4x^2 - 2, \qquad H_3 = 8x^3 - 12x, \qquad H_4 = 16x^4 - 48x^2 + 12,$$

(**a**) For the vector space of cubic polynomials in $x$, choose a basis of Hermite polynomials and compute the matrix of components of the differentiation operator, $d/dx$.
(**b**) Compute the components of the operator $d^2/dx^2$ and show the relation between this matrix and the preceding one.

**7.27** On the vector space of functions of $x$, define the translation operator

$$T_af = g \qquad \text{means} \qquad g(x) = f(x-a)$$

This picks up a function and moves it by $a$ to the right.
(**a**) Pick a simple example function $f$ and test this definition graphically to verify that it does what I said.
(**b**) On the space of cubic polynomials and using a basis of your choice, find the components of this operator.
(**c**) Square the resulting matrix and verify that the result is as it should be.
(**d**) What is the inverse of the matrix? (You should be able to guess the answer and then verify it. Or you can work out the inverse the traditional way.)
(**e**) What if the parameter $a$ is *huge*? Interpret some of the components of this first matrix and show why they are clearly correct. (If they are.)
(**f**) What is the determinant of this operator?
(**g**) What are the eigenvectors and eigenvalues of this operator?

**7.28** The force by a magnetic field on a moving charge is $\vec{F} = q\vec{v} \times \vec{B}$. The operation $\vec{v} \times \vec{B}$ defines a linear operator on $\vec{v}$, stated as $f(\vec{v}\,) = \vec{v} \times \vec{B}$. What are the components of this operator expressed in terms of the three components of the vector $\vec{B}$? What are the eigenvectors and eigenvalues of this operator? For this last part, pick the basis in which you want to do the computations. If you're not careful about this choice, you are asking for a lot of algebra. Ans: eigenvalues: $0,\ \pm iB$

**7.29** In section 7.8 you have an operator $M$ expressed in two different bases. What is its determinant computed in each basis?

**7.30** In a given basis, an operator has the values

$$A(\vec{e}_1) = \vec{e}_1 + 3\vec{e}_2 \qquad \text{and} \qquad A(\vec{e}_2) = 2\vec{e}_1 + 4\vec{e}_4$$

(**a**) Draw a picture of what this does. (**b**) Find the eigenvalues and eigenvectors and determinant of $A$ and see how this corresponds to the picture you just drew.

**7.31** The characteristic polynomial of a matrix $M$ is $\det(M - \lambda I)$. $I$ is the identity matrix and $\lambda$ is the variable in the polynomial. Write the characteristic polynomial for the general $2\times 2$ matrix. Then in place of $\lambda$ in this polynomial, put the matrix $M$ itself. The constant term will have to include the factor $I$ as usual. For this $2\times 2$ case verify the Cayley-Hamilton Theorem, that the matrix satisfies its own characteristic equation, making this polynomial in $M$ the zero matrix.

**7.32** (**a**) For the magnetic field operator defined in problem 7.28, place $\hat{z} = \vec{e}_3$ along the direction of $\vec{B}$. Then take $\vec{e}_1 = (\hat{x} - i\,\hat{y})/\sqrt{2}$, $\vec{e}_2 = (\hat{x} + i\,\hat{y})/\sqrt{2}$ and find the components of the linear operator representing the magnetic field. (**b**) A charged particle is placed in this field and the equations of motion are $m\vec{a} = \vec{F} = q\vec{v} \times \vec{B}$. Translate this into the operator language with a matrix like that of problem 7.28, and write $\vec{F} = m\vec{a}$ in this language and this basis. Ans: (part) $m\ddot{r}_1 = -iqBr_1$, where $r_1 = (x + iy)/\sqrt{2}$. $\quad m\ddot{r}_2 = +iqBr_2$, where $r_2 = (x - iy)/\sqrt{2}$.

**7.33** For the operator in problem 7.27 part (**b**), what are the eigenvectors and eigenvalues?

**7.34** A *nilpotent* operator was defined in section 7.14. For the operator defined in problem 7.8, show that it is nilpotent. How does this translate into the successive powers of its matrix components?

**7.35** A cube of uniform mass density has side $a$ and mass $m$. Evaluate its moment of inertia about an axis along a longest diagonal of the cube. Note: If you find yourself entangled in a calculation having multiple integrals with hopeless limits of integration, toss it out and start over. You may even find problem 7.18 useful. Ans: $ma^2/6$

**7.36** Show that the set of all $2 \times 2$ matrices forms a vector space. Produce a basis for it, and so what is its dimension?

**7.37** In the vector space of the preceding problem, the following transformation defines an operator. $f(M) = SMS^{-1}$. For $S$, use the rotation matrix of Eq. (7.13) and compute the components of this operator $f$. The obvious choice of basis would be matrices with a single non-zero element 1. Instead, try the basis $I$, $\sigma_x$, $\sigma_y$, $\sigma_z$. Ans: A rotation by $2\alpha$ about the $y$-axis, *e.g.* $f(\vec{e}_1\,) = \vec{e}_1 \cos 2\alpha - \vec{e}_3 \sin 2\alpha$.

**7.38** What are the eigenvectors and eigenvalues of the operator in the preceding problem? Now you'll be happy I suggested the basis that I did.

**7.39** (**a**) The commutator of two matrices is defined to be $[A, B\,] = AB - BA$. Show that this commutator satisfies the Jacobi identity.

$$[A, [B, C\,]] + [B, [C, A\,]] + [C, [A, B\,]] = 0$$

(**b**) The anti-commutator of two matrices is $\{A, B\,\} = AB + BA$. Show that there is an identity like the Jacobi identity, but with one of the two commutators (the inner one or the outer one) replaced by an anti-commutator. I'll leave it to you to figure out which.

**7.40** Diagonalize each of the Pauli spin matrices of problem 7.20. That is, find their eigenvalues and specify the respective eigenvectors as the basis in which they are diagonal.

**7.41** What are the eigenvalues and eigenvectors of the rotation matrix Eq. (7.13)? Translate the answer back into a statement about rotating vectors, not just their components.

**7.42** Same as the preceding problem, but replace the circular trigonometric functions in Eq. (7.13) with hyperbolic ones. Also change the sole minus sign in the matrix to a plus sign. Draw pictures of what this matrix does to the basis vectors. What is its determinant?

**7.43** Compute the eigenvalues and eigenvectors of the matrix Eq. (7.18). Interpret each.

**7.44** Look again at the vector space of problem 6.36 and use the basis $f_1$, $f_2$, $f_3$ that you constructed there. (**a**) In this basis, what are the components of the two operators described in that problem?
(**b**) What is the product of these two matrices? Do it in the order so that it represents the composition of the first rotation followed by the second rotation.
(**c**) Find the eigenvectors of this product and from the result show that the combination of the two rotations is a third rotation about an axis that you can now specify. Can you anticipate before solving it, what one of the eigenvalues will be?
(**d**) Does a sketch of this rotation axis agree with what you should get by doing the two original rotations in order?

**7.45** Verify that the Gauss elimination method of Eq. (7.44) agrees with (7.38).

**7.46** What is the determinant of a nilpotent operator? See problem 7.34.

**7.47** (**a**) Recall (or look up) the method for evaluating a determinant using cofactors (or minors). For 2×2, 3×3, and in fact, for $N \times N$ arrays, how many multiplication operations are required for this. Ignore the time to do any additions and assume that a computer can do a product in $10^{-10}$ seconds. How much time does it take by this method to do the determinant for 10×10, 20×20, and 30×30

arrays? Express the times in convenient units.
(**b**) Repeat this for the Gauss elimination algorithm at Eq. (7.44). How much time for the above three matrices and for 100×100 and 1000×1000? Count division as taking the same time as multiplication.
Ans: For the first method, 30×30 requires 10 000×age of universe. For Gauss it is 3 $\mu$s.

**7.48** On the vector space of functions on $0 < x < L$, (**a**) use the basis of complex exponentials, Eq. (5.20), and compute the matrix components of $d/dx$.
(**b**) Use the basis of Eq. (5.17) to do the same thing.

**7.49** Repeat the preceding problem, but for $d^2/dx^2$. Compare the result here to the squares of the matrices from that problem.

**7.50** Repeat problem 7.27 but using a different vector space of functions with basis
(**a**) $e^{n\pi i x/L}$, $(n = 0, \pm 1, \pm 2, \ldots)$
(**b**) $\cos(n\pi x/L)$, and $\sin(m\pi x/L)$.
These functions will be a basis in the set of periodic functions of $x$, and these will be *very* big matrices.

**7.51** (**a**) What is the determinant of the translation operator of problem 7.27?
(**b**) What is the determinant of $d/dx$ on the vector space of problem 7.26?

**7.52** (**a**) Write out the construction of the trace in the case of a three dimensional operator, analogous to Eq. (7.47). What are the coefficients of $\epsilon^2$ and $\epsilon^3$? (**b**) Back in the two dimensional case, draw a picture of what $(I + \epsilon f)$ does to the basis vectors to first order in $\epsilon$.

**7.53** Evaluate the trace for arbitrary dimension. Use the procedure of Gauss elimination to compute the determinant, and note at every step that you are keeping terms only through $\epsilon^0$ and $\epsilon^1$. Any higher orders can be dropped as soon as they appear. Ans: $\sum_{i=1}^{N} f_{ii}$

**7.54** The set of all operators on a given vector space forms a vector space.* (Show this.) Consider whether you can or should restrict yourself to real numbers or if you ought to be dealing with complex scalars.
Now what about the list of operators in section 7.14. Which of them form vector spaces?
Ans: Yes(real), Yes(real), No, No, No, No, No, No, No

**7.55** In the vector space of cubic polynomials, choose the basis

$$\vec e_0 = 1, \qquad \vec e_1 = 1 + x, \qquad \vec e_2 = 1 + x + x^2, \qquad \vec e_3 = 1 + x + x^2 + x^3.$$

In this basis, compute the matrix of components of the operator $P$, where this is the parity operator, defined as the operator that takes the variable $x$ and changes it to $-x$. For example $P(\vec e_1) = 1 - x$. Compute the square of the resulting matrix. What is the determinant of $P$? If you had only the quadratic polynomials with basis $\vec e_0$, $\vec e_1$, $\vec e_2$, what is the determinant? What about linear polynomials, with basis $\vec e_0$, $\vec e_1$? Maybe even constant polynomials?

**7.56** On the space of quadratic polynomials define an operator that permutes the coefficients: $f(x) = ax^2 + bx + c$, then $Of = g$ has $g(x) = bx^2 + cx + a$. Find the eigenvalues and eigenvectors of this operator.

---

* If you're knowledgeable enough to recognize the difficulty caused by the question of domains, you'll recognize that this is false in infinite dimensions. But if you know that much then you don't need to be reading this chapter.

**7.57** The results in Eq. (7.36) is a rotation about *some* axis. Where is it? Notice that a rotation about an axis leaves the axis itself alone, so this is an eigenvector problem. If it leaves the vector alone, you even know what the eigenvalue is, so you can easily find the vector. Repeat for the other rotation, found in Eq. (7.37) Ans: $\vec{e}_1 + \vec{e}_2 - \vec{e}_3$

**7.58** Find the eigenvectors and eigenvalues of the matrices in problem 7.1.