Chapter 7

Non-inertial Reference Frames

> Absolute space, in its own nature, without relation to anything external, remains always similar and immovable. Absolute, true and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.
>
> -Isaac Newton

7.1 Overview

In freshman physics we are warned always to work in an inertial reference frame. Only then can Newton's 2nd Law be reliably applied. But there are certainly times when a non-inertial frame, particularly a rotating one, looks like a very good thing to use. One obvious case is the earth itself. Since it rotates on its axis, a reference frame in which the earth is fixed is definitely not inertial. For a lot of things we can ignore that fact because the rotation is not so fast. For atmospheric phenomena, however, we must either come to grips with what the Laws of motion in a rotating frame are, or use an inertial frame, which is very inconvenient. This example will be examined later in this chapter, in section 7.5. The use of a non-inertial frame also suggests itself in trying to describe the motion of a rotating rigid body. This application will be taken up in the next chapter.

In this chapter, we first have a critical look at the entire concept of frames of reference in sectionref frames. We'll see how a reference frame is a way to relate the spatial organization of the world at different instants of time and that some reference frames are "better" than others. Having satisfied ourselves that we know what we're trying to do, the quantitative relation between descriptions of motion in two mutually moving and rotating frames is worked out in section 7.4. Since this involves some new linear algebra, those tools are developed in section 7.3. As we shall see, motion described in a rotating frame is as if we were operating under Newton's 2nd Law, but in the presence of some additional fictitious forces. This is all put to work in section 7.5, where we look at some of the manifestations of these fictitious forces. Finally, the chapter winds up with another critical look at some of the assumptions about the structure of space and time which underlie the entire construction of a reference frame.

7.2 Reference Frames

Since this chapter is about describing motion in non-inertial reference frames, it seems appropriate to have a fresh look at the fundamental notion of reference frame of whatever sort. Figure 7.1 depicts what we would call a rotating reference frame.

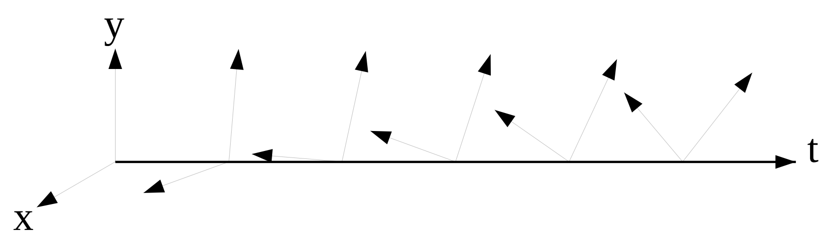

Figure 7.1: A rotating refernence frame. (Or not?) The z-axis is suppressed for obvious reasons.

It's an x, y, z coordinate system at each instant of time in which the axes don't point in the 'same direction' at different times. But the classification of those axes as rotating requires comparison to something else. How do we know it's not that standard which is actually rotating? This question, and related ones, is not trivial. It's answer involves fundamental principles of mechanics as well as suppositions (postulates) about the structure of space and time.

Newtonian mechanics is founded upon the postulate that time is something completely distinct from anything else (such as space). So the question of the time interval between two events, and most importantly the question of whether two events are simultaneous, has an unambiguous answer. At each instant of time, space is presumed to have a three dimensional Euclidean structure.

reference frame

A **reference frame** is a procedure for determining whether something is not moving, and which is consistent with this Euclidean structure. What I mean by "consistent" is that two objects which are judged to be stationary must remain forever at the same distance from one another. This requires also that angles between stationary lines are constant since a triangle is completely determined by the lengths of its sides. This requirement on a reference frame stems from the assumption that the Euclidean structure of space is not changing with time.

From an operational perspective, the simplest way to specify a reference frame is to choose one point to act as origin, attach a Cartesian coordinate system to it and keep the axes and origin stationary. In this way, questions about the relation of one reference frame to another are reduced to questions about the relative positions of their origins and orientations of their axes.

This is not the end of the story. Not all reference frames are created equal. The requirement that a reference frame be consistent with the Euclidean structure of space can be characterized as a kinematical condition. A dynamical condition is used to winnow the reference frames further into good and not-so-good. The Principle of Inertia, otherwise known as Newton's First Law of Motion asserts that there is a special class of reference frames, known as **inertial reference frames**. These are the ones in which a particle free from external forces (a free particle) moves at constant velocity. There is no trick to finding a reference frame in which one free particle moves at constant velocity. That it is even *possible* to do it for all free particles says something new about the nature of motion.

inertial reference frame

Newton's Second Law is true only in inertial reference frames. That is the reason we generally like to work in inertial reference frames. However, it is not absolutely necessary. Sometimes a non-inertial reference frame is much more convenient for one reason or another. If we can correctly relate coordinates in this non-inertial frame to those in an inertial frame, we can determine how to modify Newton's 2nd Law for use with the non-inertial system of coordinates. That is our basic task of this chapter.

The definition of reference frame presented here is very strong. Often the term is used in a somewhat weaker sense. A pair of frames, the origins of which maintain a constant displacement from one another, and whose axes maintain a constant relative orientation are not really much different. Particularly when speaking of one or another inertial frame, one really is referring to an entire class of this sort.

I have belabored some very foundational things in this section. We will continue to take them for granted. But how did Newton know that they are correct? He did not. Absolute time and the Euclidean structure of space seemed natural to him, and he had no evidence to suggest that these suppositions were wrong. The situation remained more-or-less thus for a couple of hundred years. Then Einstein threw them both overboard. Later (I hope) we will look into that in more detail.

7.3 Further Adventures in Linear Algebra

7.3.1 operator inverses and changing basis

If the linear operator B is nonsingular (does not map any nonzero vector to zero) then it is invertible. There exists a linear operator B^{-1} such that $\mathsf{B}^{-1}\mathsf{B} = \mathsf{B}\mathsf{B}^{-1} = 1$ (1 is the identity operator). According to the rules for matrix multiplication,

matrix inverse

$$[AB]_{ij} = \sum_k A_{ik}B_{kj},$$

the matrices for B and its inverse satisfy

$$\sum_k [B^{-1}]_{ik}B_{kj} = \delta_{ij}.$$

Knowing that an inverse exists is not at all the same as having an efficient algorithm for computing them. There is a formula, called **Cramer's rule**, for inverting a matrix. The very explicit formula for a 2×2 matrix is given below. You really don't want to use it for anything much larger than that. There are other, more efficient, algorithms which computers are very good at following.

But here is the rule. If you remove the i-th row from a matrix A, close up the gap to make a nonsquare array, then remove the j-th column and close up again into a square array, compute the determinant of that and multiply by $(-1)^{i+j}$, you have what is known as the i, j cofactor (abbreviated 'cof') of A:

$$[\mathrm{cof}A]_{ij} = (-1)^{i+j}(\text{det of } A \text{ with i-th row, j-th column removed}). \tag{3.1}$$

The infamous formula for the entries in the inverse matrix is then

$$[A^{-1}]_{ij} = \frac{[\mathrm{cf}A]_{ji}}{\det A}. \tag{3.2}$$

(Note the reversal of the indices!)

This is obviously something you should avoid if you can. In the case of a 2×2 matrix, though, things are very simple. You can check by direct multiplication that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{3.3}$$

Exercise 7-1 check equation (3.3).

Sometimes one wants to change from one basis to another. For instance, we may want to use the eigenvectors of some symmetric operator as a basis and then we need to know how to re-express our vectors in this basis. It is not always desirable to use the general method developed here. Particularly in two dimensions, it is often simpler to use more haphazard techniques. For general abstract arguments, however, there is no substitute.

Let's consider the old basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ and a new basis $\mathbf{e}'_1, \dots, \mathbf{e}'_n$, such that the new basis is expressed in terms of the old by

$$\mathbf{e}'_i = \sum_j \mathbf{e}_j S_{ji}. \tag{3.4}$$

At the moment, the S_{ij} are just a collection of numbers, not an operator. So it is possibly a little confusing that we're about to use matrix techniques to handle them. But it is in fact the matrix for a recognizable operator. Interpreting it as such,

$$\mathsf{S}\mathbf{e}_i = \sum_j \mathbf{e}_j S_{ji} = \mathbf{e}'_i \tag{3.5}$$

so S_{ij} is the matrix, *relative to the old basis* of the operator which sends $\mathbf{e}_i$ to $\mathbf{e}'_i$. One thing which this tells us is that the operator (hence the matrix) is invertible, since any vector annihilated by it would have to have all components zero.

So, the formula can be inverted to yield

$$\mathbf{e}_i = \sum_j \mathbf{e}'_j [S^{-1}]_{ji}. \tag{3.6}$$

This shows how to determine the components of any vector with respect to the new basis, since

$$\mathbf{v} = \sum_i \mathbf{e}_i v_i = \sum_{i,j} \mathbf{e}'_j [S^{-1}]_{ji} v_i. \tag{3.7}$$

The new and old components are therefore related by

$$v'_j = \sum_i [S^{-1}]_{ji} v_i \quad \Leftrightarrow \quad v_j = \sum_i S_{ji} v'_i. \tag{3.8}$$

The rule for the matrix of an operator in the new basis is now almost as easy. Computing the action of A on an arbitrary vector $\mathbf{v}$,

$$\mathsf{A}\mathbf{v} = \sum_{i,j} \mathbf{e}_i A_{ij} v_j = \sum_{i,j,k,l} \mathbf{e}'_k [S^{-1}]_{ki} A_{ij} S_{jl} v'_l = \sum_{k,l} \mathbf{e}'_k \left(\sum_{i,j} [S^{-1}]_{ki} A_{ij} S_{jl} \right) v'_l$$

This gives us the following formula:

$$A'_{ij} = \sum_{k,l} [S^{-1}]_{ik} A_{kl} S_{lj}. \tag{3.9}$$

Example
As an example of the use of these formulas, let's consider an operator whose matrix is

$$B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

with respect to the standard basis $\mathbf{e}_1$, $\mathbf{e}_2$ in the plane. Now we want to change to a new basis given by

$$\mathbf{e}'_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The matrix which effects this as in equation (3.4) is

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

and, according to equation (3.3), the inverse is

$$S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Then, the matrix for B in the new basis is found by a matrix sandwich

$$\begin{aligned} B' \equiv S^{-1}BS &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix}. \end{aligned}$$

To check this and also to show that other methods are often just as good, let's calculate the action of B on the new basis vectors by expressing them in the old basis and using the original matrix. We get

$$\begin{aligned} \mathbf{e}'_1 = \mathbf{e}_1 &\mapsto \mathbf{e}_1 + \mathbf{e}_2 = \mathbf{e}'_2 \\ \mathbf{e}'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} &\mapsto \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \end{pmatrix} = -\mathbf{e}'_1 + 3\mathbf{e}'_2. \end{aligned}$$

This also illustrates two slightly different ways of organizing things. Putting those together immediately gives a matrix in agreement with B'.

7.3.2 orthogonal transformations

On a real vector space $\mathcal{V}$ which is equipped with an inner (dot) product, there is a special class of operators which are very important. In discussing symmetric operators, I pointed out that any basis could be used to *define* an inner product. Here, we suppose there is some special fixed once-for-all inner product. Vectors in real, physical three-dimensional space have such an inner product defined on them and that is the one which will most concern us, but for the moment we can be a little more general.

orthogonal operator

An **orthogonal transformation** is one which preserves all lengths:

$$|\mathsf{O}\mathbf{u}| = |\mathbf{u}|. \tag{3.10}$$

(I will use the letter O a lot to denote an unspecified orthogonal operator.) This has as a consequence that all angles are also preserved, by which we mean

$$(\mathsf{O}\mathbf{u}) \cdot (\mathsf{O}\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}, \tag{3.11}$$

for any pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{V}$. That is maybe a little surprising, but not hard to see. After all,

$$|\mathsf{O}(\mathbf{u}+\mathbf{v})|^2 = |\mathbf{u}+\mathbf{v}|^2,$$

by the original definition. Expanding those,

$$|\mathsf{O}\mathbf{u}|^2 + 2(\mathsf{O}\mathbf{u}) \cdot (\mathsf{O}\mathbf{v}) + |\mathsf{O}\mathbf{v}|^2 = |\mathbf{u}|^2 + 2\mathbf{u} \cdot \mathbf{v} + |\mathbf{v}|^2.$$

Now cancelling the parts which are equal by definition of orthogonal, equation (3.11) emerges.

In ordinary three-dimensional space, rotations are orthogonal transformations. (More about that later)

orthogonal basis

An **orthonormal basis** is one made of unit vectors which are orthogonal to one another. The standard basis $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$ and $\hat{\mathbf{e}}_z$ for ordinary three-dimensional vectors is an orthonormal basis. If we use an orthonormal basis, then

$$\mathbf{u} \cdot (\mathsf{O}\mathbf{v}) = \left(\sum_i u_i \hat{\mathbf{e}}_i\right) \cdot \left(\sum_{jk} \hat{\mathbf{e}}_j A_{jk} v_k\right) = \sum_{ik} u_i A_{ik} v_k.$$

transpose

The **transpose** A^T of a matrix A is defined by

$$A^T_{ij} = A_{ji}, \tag{3.12}$$

i.e., the entries are reflected across the diagonal. A symmetric matrix is equal to its transpose. The transpose of a product of matrices is the product of the transposes *in the reversed order*:

$$[AB]^T = B^T A^T.$$

Taking a transpose is like inverting in this respect.

The transpose of a *linear operator* O is a notion which only makes sense relative to an inner product. It is defined by

$$\mathbf{u} \cdot \mathsf{A}\mathbf{v} = (\mathsf{A}^T \mathbf{u}) \cdot \mathbf{v}, \tag{3.13}$$

for all $\mathbf{u}$ and $\mathbf{v}$. Since

$$(\mathsf{A}^T \mathbf{u}) \cdot \mathbf{v} = \mathbf{v} \cdot (\mathsf{A}^T \mathbf{u}) = (\mathsf{A}^{TT} \mathbf{v}) \cdot \mathbf{u},$$

transposing twice is the same as doing nothing: $\mathsf{A}^{TT} = \mathsf{A}$.

The matrix of the transpose of an operator is not necessarily the transpose of its matrix (see example below). However, it is true if the basis is orthogonal.

By the definition of the transpose of an operator, we have the relation

$$(\mathsf{O}\mathbf{u}) \cdot (\mathsf{O}\mathbf{v}) = \mathbf{u} \cdot (\mathsf{O}^T \mathsf{O}\mathbf{v})$$

So, if O is othogonal,

$$\mathbf{u} \cdot (\mathsf{O}^T \mathsf{O}\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}.$$

As this must hold for all vectors $\mathbf{u}$ and $\mathbf{v}$, the inescapable conclusion is

$$\mathsf{O} \text{ orthogonal} \quad \Leftrightarrow \quad \mathsf{O}^T \mathsf{O} = 1 \quad \Leftrightarrow \quad \mathsf{O}^{-1} = \mathsf{O}^T \tag{3.14}$$

I'll bring up one more property of orthogonal transformations here. They form a group. What this means is that the product of two orthogonal transformations is another, the identity is an orthogonal transformation and every orthogonal transformation has an inverse (its transpose) which is also orthogonal. The only part which requires any comment is the closure under taking products. Suppose A and B are orthogonal. Then

$$|(\mathsf{AB})\mathbf{v}| = |\mathsf{A}(\mathsf{B}\mathbf{v})| = |\mathsf{B}\mathbf{v}| = |\mathbf{v}|,$$

the first equality resulting from what we mean by a product of operators, the second from the fact that A is orthogonal and the last from orthogonality of B.

Example: matrix of transpose generally not the transpose of matrix.

Consider the operator O defined on the vectors in the plane by

$$\mathsf{O}\hat{\mathbf{e}}_x = \hat{\mathbf{e}}_y, \quad \mathsf{O}\hat{\mathbf{e}}_y = -\hat{\mathbf{e}}_x \tag{3.15}$$

Since this is a rotation by 90° counterclockwise, it is clearly orthogonal. The transpose is its inverse, which is rotation by 90° clockwise:

$$\mathsf{O}^T\hat{\mathbf{e}}_x = -\hat{\mathbf{e}}_y, \quad \mathsf{O}^T\hat{\mathbf{e}}_y = \hat{\mathbf{e}}_x. \tag{3.16}$$

Relative to this orthonormal basis, these operators have matrices

$$[\mathsf{O}] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad [\mathsf{O}]^T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.17}$$

So the matrix of O^T is the transpose of the matrix of O. But let's try it with the basis

$$\mathbf{e}'_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{3.18}$$

Since under O,

$$\mathbf{e}'_1 \mapsto \mathbf{e}_y = \mathbf{e}'_2 - \mathbf{e}'_1, \quad \mathbf{e}'_2 \mapsto \mathbf{e}_y - \mathbf{e}_x = \mathbf{e}'_2 - 2\mathbf{e}'_1,$$

the matrix of O with respect to the new basis is

$$[\mathsf{O}]' = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}. \tag{3.19}$$

If this is multiplied by the matrix for the transpose of O, the result should be the identity matrix. However, the product with the transpose of the matrix $[\mathsf{O}]'$ is

$$\begin{pmatrix} 0 & 3 \\ 3 & 5 \end{pmatrix},$$

which is not the identity at all! So, relative to this basis, the matrix of O^T must be something else.

Exercise 7-2 Carry out that multiplication of a matrix by its transpose to see if I lied.

Fortunately, we will be using orthonormal bases almost exclusively in this chapter and the next, because it is the natural thing to do. As a result, you can probably forget the distinctions we've drawn between operators and matrices without danger of making mistakes in the near future. But don't forget it on purpose.

7.3.3 Orthogonal Transformations on Euclidean 3-Space

The orthogonal transformations on Euclidean 3-space are certainly especially interesting. These transformations collectively constitute a group labelled $O(3)$. It contains ordinary simple rotations, but it also contains some other things. In particular the **inversion** $\mathbf{-1}$, which maps every vector to its inverse. One way this operation can be achieved is to reflect first through the y-z plane, then through the x-z plane and then through the x-y plane, since these operations successively change the signs of the components of a vector. The ordinary rotations make up the part of $O(3)$ called $SO(3)$. These are sometimes called **proper rotations** to distinguish them from other elements of $O(3)$. Every other orthogonal transformation can be realized as a rotation followed by inversion. The proper rotations form a group by themselves, and it is almost exclusively with those that we will be concerned. The fact that every orthogonal transformation does take one of these two forms is proven in section 7.3.4. But you might not feel any need to see that.

The operation of taking a cross product with $\mathbf{u}$,

$$\mathbf{u} \times \mathbf{r} = \sum \hat{\mathbf{e}}_i \epsilon_{ijk} u_j r_k, \tag{3.20}$$

is not an orthogonal operator, but it is closely related. It is clearly linear, and a little manipulation will reveal its matrix. Calling the matrix A,

$$A_{ij} = \sum_k \epsilon_{ikj} u_k = -\sum_k \epsilon_{ijk} u_k. \tag{3.21}$$

matrix for cross product

Writing that out in the matrix form,

$$A_{ij} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}. \tag{3.22}$$

This matrix is **antisymmetric**, which means that taking the transpose is the same as changing the sign: $A^T = -A$. In fact, every antisymmetric matrix is the matrix for a cross product operation because the three entries above the diagonal of an antisymmetric matrix determine the whole thing, and those can be made whatever we please by choosing the right $\mathbf{u}$.

There is a simple way to manufacture a linear operator on an inner product space from a pair of vectors $\mathbf{u}$ and $\mathbf{v}$ which will be useful for writing the inertia tensor when we study rigid body motion, so I will describe it now. It is defined by

$$(\mathbf{u} \otimes \mathbf{v})\mathbf{w} \stackrel{def}{=} \mathbf{u}[\mathbf{v} \cdot \mathbf{w}]. \tag{3.23}$$

Such an operator is sometimes referred to as the **outer product** of $\mathbf{u}$ and $\mathbf{v}$. Described in words, its effect is to project a vector onto $\mathbf{v}$ and then swap that projection for the same multiple of $\mathbf{u}$. If we use $\mathbf{u}$ in both slots, we get an **orthogonal projection**,

$$\Pi_{\hat{\mathbf{u}}} \stackrel{def}{=} \frac{\mathbf{u} \otimes \mathbf{u}}{|\mathbf{u}|^2}. \tag{3.24}$$

Applied to a vector $\mathbf{w}$, this simply throws away the part which is orthogonal to $\mathbf{u}$.

We can combine this projection operator with the cross product operator discussed just before to write down a representation of an arbitrary rotation about $\mathbf{u}$. It's not a useful as you might think, but interesting nonetheless.

Rotation about a unit vector $\hat{\mathbf{u}}$ is a linear operation. So we can easily reconstruct it from its effect on vectors parallel $\hat{\mathbf{u}}$ (it does nothing to them), and its effect on vectors in the plane orthogonal to $\hat{\mathbf{u}}$. On the latter, the rotation through an angle θ is

$$\mathsf{R}_\theta(\hat{\mathbf{u}})\mathbf{v} = (\cos\theta)\mathbf{v} + (\sin\theta)\hat{\mathbf{u}} \times \mathbf{v}.$$

Adding in the 'do nothing' on vectors parallel to $\hat{\mathbf{u}}$,

$$\mathsf{R}_\theta(\hat{\mathbf{u}})\mathbf{v} = (\hat{\mathbf{u}} \otimes \hat{\mathbf{u}})\mathbf{v} + (1 - \hat{\mathbf{u}} \otimes \hat{\mathbf{u}})\left[(\cos\theta)\mathbf{v} + (\sin\theta)\hat{\mathbf{u}} \times \mathbf{v}\right]. \tag{3.25}$$

7.3.4 $O(3)$ = rotations + inversion

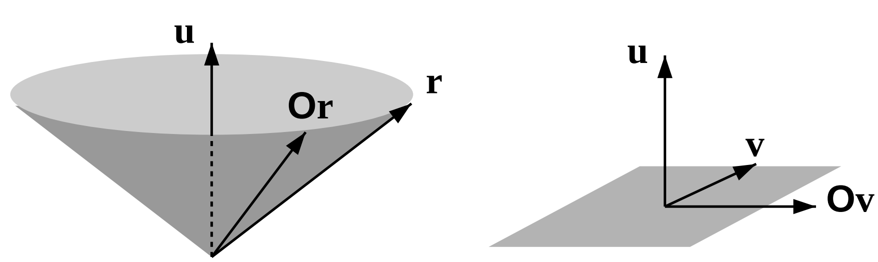

Figure 7.2: A rotation about $\mathbf{u}$ moves $\mathbf{r}$ into another vector on a cone centered on $\mathbf{u}$, as at left. The component of $\mathbf{r}$ along $\mathbf{u}$ is unaffected by the rotation. We subtract that part out to form $\mathbf{v} = \mathsf{O}\mathbf{r} - \mathbf{r}$, which then lies in the plane perpendicular to $\mathbf{u}$ as at right. $\mathbf{v}$ and its image under O can be used to find $\mathbf{u}$ because the span the plane orthogonal to it.

In this subsection, we prove the assertion that all orthogonal transformations on Euclidean 3-space is either a simple rotation about some axis, or a rotation followed by an inversion.

The first step is to show that the orthogonal transformation O has a real eigenvector. This can be done by appealing to a general result which asserts that for each distinct root of the characteristic equation (eigenvalue), there is an eigenvector. But, in our case it can be done quite directly, so let's see that. We have used complex eigenvectors without much comment before, so I want to emphasize that I'm talking about an eigenvector whose components are real. This means that the eigenvalue must be real, too, since O is real.

Assume (contrary to what we intend to show) that O has no real eigenvectors. Then an arbitrarily chosen vector $\mathbf{r}$ is not mapped into a multiple of itself by O. And therefore

$$\mathbf{v} = \mathsf{O}\mathbf{r} - \mathbf{r}$$

is a nonzero vector. Since, by assumption, $\mathbf{v}$ isn't an eigenvector either,

$$\mathbf{u} = \mathsf{O}\mathbf{v} \times \mathbf{v} = [\mathsf{O}(\mathsf{O}\mathbf{r} - \mathbf{r})] \times (\mathsf{O}\mathbf{r} - \mathbf{r})$$

is also nonzero. We will show that it is an eigenvector of O. First,

$$\mathbf{u} \perp \mathbf{v} \Rightarrow \mathsf{O}\mathbf{u} \perp \mathsf{O}\mathbf{v}.$$

But $\mathsf{O}\mathbf{u}$ is also perpendicular to $\mathbf{v}$ since

$$\begin{aligned}(\mathsf{O}\mathbf{u}) \cdot \mathbf{v} &= (\mathsf{O}\mathbf{u}) \cdot (\mathsf{O}\mathbf{r}) - (\mathsf{O}\mathbf{u}) \cdot \mathbf{r} \\ &= \mathbf{u} \cdot \mathbf{r} - \mathbf{u} \cdot (\mathsf{O}\mathbf{r}) \\ &= -\mathbf{u} \cdot (\mathsf{O}\mathbf{r} - \mathbf{r}) = 0.\end{aligned}$$

Therefore $\mathsf{O}\mathbf{u}$ is perpendicular to both $\mathbf{v}$ and $\mathsf{O}\mathbf{v}$, so is proportional to $\mathbf{u}$.

That settles that part. O has at least one eigenvector, which we'll call $\mathbf{u}$ for the duration. The associated eigenvalue must be either 1 or -1. After all, $\mathsf{O}\mathbf{u}$ has the same length as $\mathbf{u}$ and $\pm\mathbf{u}$ are the only vectors proportional to $\mathbf{u}$ with that length.

Then vectors in the plane perpendicular to $\mathbf{u}$ must be kept in that plane by O since their images must stll be orthogonal to $\mathbf{u}$. (see figure 7.2 (a)) If O has an eigenvector $\mathbf{v}$ in this plane then the vector orthogonal to both must also be an eigenvector, since its image under O must still be orthogonal to both of those. Put aside for the moment the possibility that there are any eigenvectors in the plane perpendicular to $\mathbf{u}$, and pick any vector in that plane. Its image, $\mathsf{O}\mathbf{v}$, is the same length and still in this plane, so must just be $\mathbf{v}$ rotated around $\mathbf{u}$ by some angle. Since all angles are preserved by O, it must rotate all the vectors in that plane by the same angle.

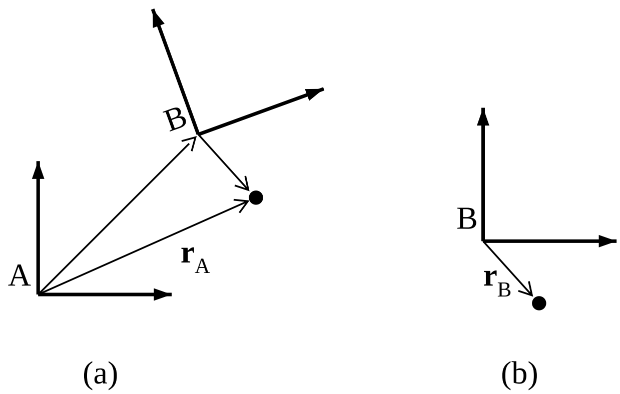

Figure 7.3: A pair of reference frames.

So, in this case, O is a rotation in the plane perpendicular to $\mathbf{u}$. If $\mathsf{O}\mathbf{u} = \mathbf{u}$, then it is a simple rotation. A rotation by θ in that plane, followed by a reflection through it, to change the sign of $\mathbf{u}$ is the same thing as rotating by $\theta + \pi$ and then performing an inversion. (picture it!) If O has just one eigenvector then, it is either a simple rotation or a rotation followed by inversion.

Now we have to clean up the case put aside earlier: O has three mutually orthogonal eigenvectors. If all three eigenvalues are $+1$, it is the identity, which is a rotation by zero (about any axis you choose!). If one eigenvalue is -1, O is a 180° rotation about the corresponding eigenvector followed by inversion. If two eigenvalues are -1, it is a 180° rotation about the leftover eigenvector. And, finally, if all three eigenvalues are -1, it is the inversion itself.

Exercise 7-3 Picture those operations in your head to verify the assertions.

Well, that takes care of that.

7.4 Description of Motion in Differing Reference Frames

Now we get down to business and look into the problem of relating the description of motion in two different reference frames. In general, their origins may not coincide and may be in relative motion. As well, the coordinate axes of the two frames may not be oriented in the same way. The second problem is the more difficult and unfamiliar. We will put off adding the distraction of non-coincident origins until section 7.4.2.

Figure 7.3 depicts a pair of reference frames, and the specification of an object's position by vectors relative to each of them. The position vector of an object is determined by the orientation of that frame's axes. Hence the position $\mathbf{r}_{\mathrm{B}}$ has a negative y-component because the vector is below the x-axis in frame B.

7.4.1 relative rotations

Until further notice, we'll assume that we are dealing with two reference frames, $\mathcal{A}$ and $\mathcal{A}$, whose origins coincide at all times. They are therefore related by a pure rotation. Each point in space is located by a vector from the origin of a reference frame to that point, and the coordinate axes are used to identify that vector. A point one meter from the origin along the x' axis is referred to by the vector $\hat{\mathbf{e}}_x$ in the frame $\mathcal{A}'$. In the $\mathcal{A}$ frame, it corresponds to a different vector.

It can help to avoid confusion if we try to think about position vectors in different reference frames as belonging to completely separate vector spaces.

We have just seen that the same point in space is referenced by different vectors in the two reference frames $\mathcal{A}$ and $\mathcal{A}'$. They are related by

$$\mathbf{r}(t) = \mathsf{O}(t)\mathbf{r}'(t). \tag{4.26}$$

Inverting this to get the relation the other direction,

$$\mathbf{r}'(t) = \mathsf{O}(t)^{-1}\mathbf{r}(t) = \mathsf{O}(t)^T\mathbf{r}(t). \tag{4.27}$$

Now, to relate velocities measured relative to the two reference frames, we need to take a time derivative, to get

$$\frac{d\mathbf{r}}{dt} = \frac{d}{dt}\left(\mathsf{O}\mathbf{r}'\right) = \frac{d\mathsf{O}}{dt}\mathbf{r}' + \mathsf{O}\frac{d\mathbf{r}'}{dt}. \tag{4.28}$$

The first term of the final expression here can be rewritten in terms of $\mathbf{r}$ instead of $\mathbf{r}'$. Inserting equation (4.27), it becomes

$$\dot{\mathsf{O}}\mathsf{O}^T\mathbf{r}.$$

Since O is orthogonal, $\mathsf{O}\mathsf{O}^T = \mathbf{1}$ is time independent. Differentiating, you discover that

$$\dot{\mathsf{O}}\mathsf{O}^T + \mathsf{O}\dot{\mathsf{O}}^T = 0.$$

But,

$$\left[\dot{\mathsf{O}}\mathsf{O}^T\right]^T = \mathsf{O}\dot{\mathsf{O}}^T,$$

which together with the previous equation implies that $\dot{\mathsf{O}}\mathsf{O}^T$ is an antisymmetric operator. According to section 7.3.3, that means that it can be written as a cross-product operation:

$$\mathsf{B} \equiv \dot{\mathsf{O}}\mathsf{O}^T = \boldsymbol{\omega} \times (\quad). \tag{4.29}$$

In a moment, we'll see what the meaning of $\boldsymbol{\omega}$ is, but first put it back into equation (4.28) to get

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r} + \mathsf{O}\frac{d\mathbf{r}'}{dt}. \tag{4.30}$$

From this you can see that any object stationary with respect to $\mathcal{A}'$ (so $\dot{\mathbf{r}}' = 0$) has an angular velocity $\boldsymbol{\omega}$ as seen from $\mathcal{A}$. Since the coordinate axes of $\mathcal{A}'$ can themselves be considered objects stationary in frame $\mathcal{A}'$, we refer to $\boldsymbol{\omega}$ as the angular velocity of $\mathcal{A}'$ relative to $\mathcal{A}$ (as measured in $\mathcal{A}$).

relative angular velocity

The angular velocity of $\mathcal{A}'$ relative to $\mathcal{A}$, *as measured in* $\mathcal{A}$, on the other hand, is $\mathsf{O}^T\boldsymbol{\omega}$. We need the O^T because the coordinate axes are generally not pointing in the same direction. Suppose $\mathcal{A}'$ had its z-axis along the x-axis of $\mathcal{A}$ and was spinning about that axis. According to $\mathcal{A}$, the relative angular velocity is in the x direction, but according to $\mathcal{A}'$, it is in the z direction.

Question 7-1 In one of the steps leading to equation (4.29), there was an $\dot{\mathsf{O}}^T$. Since $\mathsf{O}^T = \mathsf{O}^{-1}$, we might have written that as $\dot{\mathsf{O}}^{-1}$. I chose not to do so because the latter notation is ambiguous, whereas the former is not. How? (Hint: Is the time derivative of an inverse the same as the inverse of the time derivative?)

accelerations

Since Newton's Second Law involves accelerations, we really want to see how the second time derivatives of $\mathbf{r}$ and $\mathbf{r}'$ are related. Following our noses, we could differentiate again and try to identify $\boldsymbol{\omega}$'s and so forth in the debris. That's not so hard, but let's take a slightly different tack.

The key is the identity

$$\mathsf{O}\,\frac{d}{dt}\,\mathsf{O}^T = \frac{d}{dt} - \mathsf{B}. \tag{4.31}$$

What this means is that if anything (for instance $\mathbf{r}(t)$) is put to the right of these expressions, and the time derivatives allowed to act on everything to their right, the results will be equal. To see that this is true only requires computing the time derivative of O^T, which we do now. From

$$0 = \frac{d}{dt}(\mathsf{O}\mathsf{O}^T) = \dot{\mathsf{O}}\mathsf{O}^T + \mathsf{O}\dot{\mathsf{O}}^T,$$

you can immediately deduce

$$\dot{\mathsf{O}}^T = -\mathsf{O}^T \dot{\mathsf{O}} \mathsf{O}^T. \tag{4.32}$$

With this result in hand,

$$\mathsf{O}\frac{d}{dt}\mathsf{O}^T = \mathsf{O}\left[-\mathsf{O}^T\dot{\mathsf{O}}\mathsf{O}^T + \mathsf{O}^T\frac{d}{dt}\right].$$

Cleaning up a bit and substituting the definition of B results in equation (4.31).

Exercise 7-4 I said you could deduce eq. (4.32). Demonstrate that I did not lie. Also make that last step to recover eq. (4.31).

Perhaps you are wondering what the point of such a strange-looking identity is. Let's see. Apply it to $\mathbf{r}$. From the left-hand side, you get

$$\mathsf{O}\,\frac{d}{dt}\,\mathsf{O}^T\mathbf{r} = \mathsf{O}\,\frac{d}{dt}\,\mathbf{r}' = \mathsf{O}\,\dot{\mathbf{r}}',$$

and from the right-hand side

$$\left(\frac{d}{dt} - \mathsf{B}\right)\mathbf{r}.$$

This relates the time derivative of $\mathbf{r}'$ to that of $\mathbf{r}$. Repeating the operation,

$$\mathsf{O}\,\frac{d}{dt}\,\mathsf{O}^T\mathsf{O}\,\frac{d}{dt}\,\mathsf{O}^T = \left(\frac{d}{dt} - \mathsf{B}\right)\left(\frac{d}{dt} - \mathsf{B}\right).$$

The O and O^T in the left-hand side pair up to give the identity, so that this is equivalent to

$$\mathsf{O}\,\frac{d^2}{dt^2}\,\mathsf{O}^T = \left(\frac{d}{dt} - \mathsf{B}\right)^2. \tag{4.33}$$

The right-hand side expands to

$$\left(\frac{d}{dt} - \mathsf{B}\right)\left(\frac{d}{dt} - \mathsf{B}\right) = \frac{d^2}{dt^2} - 2\mathsf{B}\frac{d}{dt} - \dot{\mathsf{B}} + \mathsf{B}\mathsf{B}.$$

Applying this to $\mathbf{r}$ now,

$$\ddot{\mathbf{r}} = \mathsf{O}\ddot{\mathbf{r}}' + (2\boldsymbol{\omega} \times \dot{\mathbf{r}}) + (\dot{\boldsymbol{\omega}} \times \mathbf{r}) - (\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r})\,. \tag{4.34}$$

We want to rewrite this with the rôles of the two reference frames reversed. The transformation of position vectors from $\mathcal{A}$ to $\mathcal{A}'$ is achieved not by O, but by its inverse O^T. Since nothing was priviledged about one reference frame or the other, we could just go through, replacing O by O^T, and primed things by unprimed and vice-versa. $\boldsymbol{\omega}'$ would be the angular velocity of $\mathcal{A}$ relative to $\mathcal{A}'$ in frame $\mathcal{A}'$. The angular velocity of $\mathcal{A}'$ relative to $\mathcal{A}$ in $\mathcal{A}'$ is minus this, and is henceforth denoted by $\boldsymbol{\Omega}$. In fact,

$$\boldsymbol{\Omega} = \mathsf{O}^T\boldsymbol{\omega} = -\boldsymbol{\omega}.$$

With those replacements,

$$\ddot{\mathbf{r}}' = \mathsf{O}^T\ddot{\mathbf{r}} - \left(2\boldsymbol{\Omega} \times \dot{\mathbf{r}}'\right) - \left(\dot{\boldsymbol{\Omega}} \times \mathbf{r}'\right) - \left(\boldsymbol{\Omega} \times \boldsymbol{\Omega} \times \mathbf{r}'\right). \tag{4.35}$$

The parentheses in this equation only serve to help the eye sort things out. The last term means to perform the cross product operations one-by-one, working from right to left. That is, it should be read as $\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r})$. It makes a difference – $\boldsymbol{\Omega} \times \boldsymbol{\Omega} = 0$! The "extra" terms on the right hand side, apart from the first, are a little mysterious. They be interpreted in section 7.5.

Exercise 7-5 Work out the analogue of equation (4.33) for higher derivatives and see that all the internal O's and O's pair up and disappear.

7.4.2 relative translation

With just a little more work, we can extend what we've done to the case of reference frames whose origins are not coincident and possibly moving with respect to each other (this situation is shown in figure 7.3). Again, the two frames we're interested in are labelled $\mathcal{A}$ and $\mathcal{A}'$. To get us over the hump, invent a new one, $\mathcal{A}''$. It has its origin in the same place as does $\mathcal{A}'$, but its axes are aligned with those of $\mathcal{A}$. The location of this origin in frame $\mathcal{A}$ is denoted by $\mathbf{R}_0(t)$. Now we transfer position vectors from $\mathcal{A}'$ to $\mathcal{A}$ in two stages. The angular velocity of $\mathcal{A}'$ is the same with respect to $\mathcal{A}''$ and $\mathcal{A}$, namely $\boldsymbol{\omega}$. Then,

$$\begin{aligned} \mathbf{r}'' &= \mathsf{O}\mathbf{r}' \\ \mathbf{r} &= \mathbf{R}_0 + \mathbf{r}''. \end{aligned} \tag{4.36}$$

When we take time derivatives now, the only extra thing is derivatives of $\mathbf{R}_0$. Here they are:

$$\begin{aligned} \dot{\mathbf{r}} &= \dot{\mathbf{R}}_0 + \mathsf{O}\dot{\mathbf{r}'} + \boldsymbol{\omega} \times \mathbf{r} \\ \ddot{\mathbf{r}} &= \ddot{\mathbf{R}}_0 + \mathsf{O}\ddot{\mathbf{r}}' + (2\boldsymbol{\omega} \times \dot{\mathbf{r}}) + (\dot{\boldsymbol{\omega}} \times \mathbf{r}) - (\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r}). \end{aligned} \tag{4.37}$$

7.5 Centrifugal and Coriolis "Forces"

That's all the hard work for now. The situation in which we are mostly interested in applying this is that in which $\mathcal{A}'$ is an inertial reference frame and we need to work out the equation of motion for a particle's position with respect to $\mathcal{A}$. If it's mass is m, equation (4.35) immediately gives us

$$\ddot{\mathbf{r}}' = \mathbf{F}'/m - \mathsf{O}^T\ddot{\mathbf{R}}_0 + -\dot{\mathbf{\Omega}} \times \mathbf{r}' - 2\mathbf{\Omega} \times \dot{\mathbf{r}}' - \mathbf{\Omega} \times \mathbf{\Omega} \times \mathbf{r}'. \tag{5.38}$$

$\mathbf{F}' = \mathsf{O}^T\mathbf{F}$ is the force as measured in frame $\mathcal{A}'$, and $\mathbf{R}_0' = \mathsf{O}^T\mathbf{R}_0$ is the position of the origin of $\mathcal{A}'$ relative to that of $\mathcal{A}$ as measured in $\mathcal{A}'$. The O^T is needed to correct for the differing alignments of the axes in the two reference frames. $\mathcal{A}$ is not an inertial frame if it is rotating with respect to $\mathcal{A}'$, so that $\mathbf{\Omega} \neq 0$. Some extra terms appear in the equation of motion by comparison to what is there for an inertial reference frame.

fictitious forces

If the other terms were interpreted as extra forces, we could pretend that we were working in an inertial reference frame. The advantage of this is not clear to me, but there are traditional names attached to these things to reflect that point of view. The term $-m\mathbf{\Omega} \times \mathbf{\Omega} \times \mathbf{r}'$ is called the **centrifugal force**, and is no doubt somewhat familiar, though this form probably isn't. Without the m, as it actually appears in the equation, it is the centrifugal acceleration. Using the identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$ (the "BAC − CAB" rule),

$$-\mathbf{\Omega} \times \mathbf{\Omega} \times \mathbf{r}' = -(\mathbf{\Omega} \cdot \mathbf{r}')\,\mathbf{\Omega} + |\mathbf{\Omega}|^2\,\mathbf{r}'. \tag{5.39}$$

If $\mathbf{r}$ is perpendicular to $\mathbf{\Omega}$, this reduces to the centrifugal acceleration you've seen before.

The term $2m\mathbf{\Omega} \times \dot{\mathbf{r}}'$ is called the **Coriolis force**. The remaining term, $m\dot{\mathbf{\Omega}} \times \mathbf{r}'$ seems to usually go without a name. In most applications it disappears anyway because $\dot{\mathbf{\Omega}} = 0$ unless the rotation of $\mathcal{A}$ and $\mathcal{A}'$ is nonuniform.

From now on, I'm going to mostly omit the primes on position vectors etc. measured in the moving reference frame. This should cause no confusion because we will do all the calculations in that frame.

7.5.1 Life on a Turntable

Actually, none of the extra terms in the noninertial-frame equation of motion are terribly mysterious. Let's look at a simple situation to understand that. Imagine yourself situated at the axis of a giant turntable, the surface of which is frictionless. Your reference frame has axes fixed in the structure of the turntable. We suppose it to be rotating with angular velocity $\mathbf{\Omega}$ with relative to an inertial reference frame. To be more precise, say Ω counterclockwise, so that $\mathbf{\Omega} = \Omega\,\hat{\mathbf{e}}_z$ is directed upward.

Your dog is radially outward from you, lying on the ground at the end of his leash. (don't ask how he got there, but he gave up trying to walk on the frictionless surface). From your perspective, his acceleration is zero. Yet, you must pull on the leash in order to keep Rover from flying off. This is required in order to counter the centrifugal force

$$m\Omega^2 \mathbf{r}$$

which is directed outward. From an inertial observers point of view, your pulling on the leash provides the only force acting on Rover. It is simply the force required to accelerate him into a circular trajectory. Rover

The leash snaps. There are no real forces acting on Rover now. In the inertial reference frame $\mathcal{A}$, he now moves in a straight line at constant speed. In your frame, he initially has no tangential velocity, but immediately begins to accelerate outward, away from you. After he has slid a short distance, his tangential velocity will no longer be as large as that of the turntable material above which he is situated. Thus, in the turntable frame, Rover appears to *veer off clockwise*. You can interpret this as the Coriolis force. Let's check the magnitude and, most importantly, the sign:

$$\begin{aligned} \text{"Coriolis force"} &= -2m\boldsymbol{\Omega} \times \dot{\mathbf{r}} \\ &= -2m\Omega\, \hat{\mathbf{e}}_z \times \dot{\mathbf{r}} \\ &= -2m\Omega\, \hat{\mathbf{e}}_z \times (\dot{r}\, \hat{\mathbf{e}}_r + r\dot{\theta}\, \hat{\mathbf{e}}_\theta) \\ &= 2m\Omega(-\dot{r}\, \hat{\mathbf{e}}_\theta + r\dot{\theta}\, \hat{\mathbf{e}}_r). \end{aligned}$$

So, a positive $\dot{r}$ gives rise to a fictitious force in the $-\hat{\mathbf{e}}_\theta$ direction, which is clockwise. That's good.

Bye, Rover.

7.5.2 Rocks, Rivers and Drains

In this section, we look into some effects of the rotation of the earth, working in a frame fixed with respect to the earth, and which is therefore non-inertial. The magnitude of the earth's angular velocity is $\Omega = 2\pi/(1 \text{ day}) = 7.27 \times 10^{-5}\,\text{s}^{-1}$. This gives rise to a centrifugal acceleration which is directed away from the earth's axis. That direction is vertical at the equator, and becomes horizontal at the poles. Since the radius of the earth is 6380 km, $\Omega^2 R_E = 0.0337$ m/s$^2 \approx 3 \times 10^{-3} g$. At the equator, the centrifugal force is 3/1000 times the weight. In general

$$a_{\text{cent}} = \Omega^2 R_E \sin\lambda = 3.4 \times 10^{-3} g \cos\lambda,$$

where λ is the latitude. This acceleration is not directed vertically but is totally independent of the state of motion of objects, so is indistiguishable from a small

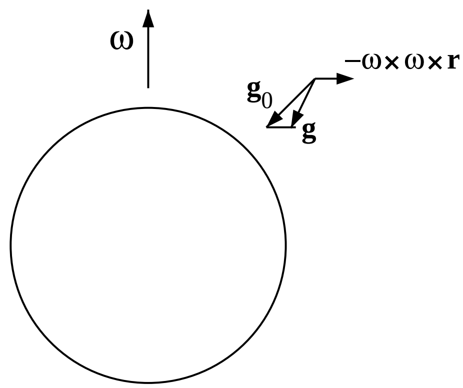

Figure 7.4: The centrifugal force modifies the local acceleration of gravity from $\mathbf{g}_0$ which it would be if the earth did not rotate, to $\mathbf{g}$. The change is extremely exaggerated in this figure; it is never greater than a degree.

change of the local acceleration of gravity, so that $\mathbf{g}$ does not point quite toward the center of the earth, as shown in figure 7.4.

A moving object will also experience a Coriolis acceleration $-2\mathbf{\Omega} \times \mathbf{v}$. If we adopt a local coordinate system in which the z axis is along the vertical and the positive x direction points down a line of longitude toward the equator,

$$\mathbf{\Omega} = \Omega[\hat{\mathbf{e}}_z \sin\lambda - \hat{\mathbf{e}}_x \cos\lambda], \tag{5.40}$$

where λ is the latitude again. The acceleration of an object moving under gravity is then

$$\mathbf{a} = \mathbf{g} - 2\mathbf{\Omega} \times \mathbf{v}. \tag{5.41}$$

Example.
A rock is dropped down a mine shaft 250 m deep at the latitude of St. Petersburg (60° north). How far will it deviate from a straight drop along $\mathbf{g}$?

If it were not for the Coriolis acceleration, the velocity of the rock would be $\mathbf{v} = \mathbf{g}t = -gt\,\hat{\mathbf{e}}_z$. As the rock falls, the Coriolis acceleration will cause it to develop a horizontal velocity (see equation 5.40). Putting this into the Coriolis term of equation (5.41) results in

$$\dot{\mathbf{v}} = -g\,\hat{\mathbf{e}}_z - 2\Omega\cos\lambda v_z \hat{\mathbf{e}}_y = -g\,\hat{\mathbf{e}}_z - 2\Omega gt\,\cos\lambda\hat{\mathbf{e}}_y.$$

We have not put the entire velocity into the coriolis term. In principle the y component of $\mathbf{v}$ should also go into that, but it is very small, so can be neglected. This is easily integrated twice to give the position:

$$\mathbf{r}(t) = -\frac{gt^2}{2}\hat{\mathbf{e}}_z - \frac{\Omega g t^3}{3}\hat{\mathbf{e}}_y.$$

The time to fall a distance D is $\sqrt{2g/D}$, so the net deflection is

$$\Delta y = g\Omega\cos\lambda\frac{1}{3}\left(\frac{2D}{g}\right)^{3/2} = \frac{\Omega\cos\lambda(2D)^{3/2}}{3g^{1/2}}.$$

Inserting numerical values,

$$\Delta y = 4.3\,\mathrm{cm}.$$

One lesson of this example is that the effects of the Coriolis acceleration are pretty slight over modest distances like this. Over longer distances, such as occur in hurricanes, it can amount to a lot.

In the northern hemisphere, the vertical component of $\boldsymbol{\Omega}$ is positive and in the southern hemisphere negative. This means that the horizontal motion of an object is deflected to the right in the northern hemisphere. It would close into a clockwise circle (seen from above), given enough distance. In the southern hemishpere, the direction is reversed. This is definitely significant for weather patterns. Possibly also rivers. The Volga tends to undermine its right bank in the middle of its course where the curvature is slight.[1] Sometimes people make the claim that water going down a drain will swirl clockwise in the northern hemisphere on account of the Coriolis acceleration. This is certainly false as even the slightest current in the water before the drain is opened or asymmetry in the drain itself would surely swamp the effect.

7.5.3 The Foucault Pendulum

An interesting manifestation of the Coriolis force is provided by the Foucault pendulum. A Foucault pendulum is really nothing other than a pendulum which is free to swing in two different directions, so is different from a plane pendulum whose motion is restricted to a plane by swinging on a fixed axle. A weight at the end of a wire will do.

Ignoring the Coriolis acceleration, we are dealing with a familiar situation. Orient the coordinate system so that z is vertical and x points down a line of longitude, as for the rock in the mine shaft. For a pendulum of length ℓ, the Coriolis-free

[1] So says V. I. Arnold, I presume the Mississippi does the same!

equations of motion are simply

$$\begin{aligned}\ddot{x} &= -\alpha x \\ \ddot{y} &= -\alpha y,\end{aligned}$$

where

$$\alpha = \frac{g}{\ell}$$

is the frequency of small vibrations of the pendulum.

Decompose the angular velocity of the earth as in equation (5.40),

$$\mathbf{\Omega} = \Omega[\hat{\mathbf{e}}_z \sin\lambda - \hat{\mathbf{e}}_x \cos\lambda],$$

and add the Coriolis acceleration $-2\mathbf{\Omega} \times \mathbf{v}$ to the equations of motion. Only the z component of $\mathbf{\Omega}$ makes any difference to the horizontal motion, so

$$\begin{aligned}\ddot{x} &= -\alpha x + 2\Omega_z \dot{y} \\ \ddot{y} &= -\alpha y - 2\Omega_z \dot{x}.\end{aligned} \tag{5.42}$$

We have seen ways to solve coupled equations like these. But let's use a clever trick. Combine x and y, which are real (Duh!) into a complex variable,

$$w = x + iy.$$

Then the two equations of motion become a single equation for this new complex coordinate:

$$\ddot{w} + 2i\Omega_z \dot{w} + \alpha^2 w = 0. \tag{5.43}$$

This is precisely the equation of motion for a damped harmonic oscillator, studied in chapter 3, except that the damping is imaginary. Just substitute α for Ω_0 and $2i\Omega_z$ for $\omega_0/Q = \Gamma$ in equation (3.4.9) We can immediately write down the solution from equation (3.4.11):

$$w(t) = \left(Ae^{i\omega' t} + Be^{-i\omega' t}\right) e^{-i\Omega_z t}, \tag{5.44}$$

with

$$\omega' = \sqrt{\alpha^2 + \Omega_z^2} \approx \alpha.$$

The difference from the oscillator is that this time we really do want a complex-valued solution. The real part of w is x and the imaginary part is y. Now the x-y plane is identified with the complex w-plane. If we arrange initial conditions so that the factor in parentheses in the solution (5.44) is real, near time $t = 0$ the pendulum is oscillating along the x-direction only. Slowly, the phase factor $e^{-i\Omega_z t}$ rotates it in the x-y plane. The line along which the pendulum swings rotates around with a period

$$\tau = \frac{2\pi}{\Omega_z} = \frac{2\pi}{\Omega_{\text{earth}} \sin\lambda} = \frac{1\,\text{day}}{\sin\lambda}, \tag{5.45}$$

where λ is the latitude.

7.6 Relativity and "Obvious" Things

The discussion of reference frames at the very beginning of this chapter was at the brink of both special and general relativity. In this section, we'll have another quick look at those issues. An absolutely crucial assumption in making everything work was that there is no ambiguity in the idea of the time interval between a pair of events, and particularly in the notion of simultaneity. It is now beyond doubt that this is not rigorously true. Events which are judged simultaneous in one frame are not necessarily so in a reference frame moving with respect to it. However, the discrepancy is very slight unless the relative speed of the two frames is an appreciable fraction of the speed of light.

This does not mean that one observer is right and the other wrong. They are both right. In each reference frame, space is still Euclidean at each instant of time. But the specification of a particular instant of time is no longer something which is the same from one reference frame to another. As a result, time gets tangled up with the transformation from one reference frame to another even when they are not rotating relative to one another.

You will notice that the fictitious forces which appear in a rotating frame, the centrifugal and Coriolis forces, are really accelerations because the force on an object is proportional to its mass. Indeed I called them accelerations. This is a clue to the fact that they are not real forces, but are kinematic properties of the reference frame itself. The gravitational force on a body shares this property. Perhaps it is not really a force either? This idea is one opening through which general relativity can be developed. It immediately points up a problem in the use of our definition of an inertial reference frame. That definition utilized the concept of a free particle – one with no forces acting on it. This seems innocent enough, but to test a reference frame for "inertial-ness" we should get a free particle. There's no trouble getting a particle which is not subject to electromagnetic forces — just find a neutral one. Also, we can keep other bodies from coming into contact with our test particle and pushing on it. But there's no way to isolate it from the effects of gravity. The only way, within the Newtonian framework, to find an inertial reference frame in the presence of gravity is to explicitly calculate the gravitational forces and compensate for them.

Maybe a particle moving only under the influence of gravity really is free and a *freely-falling* reference frame is an inertial reference frame. This is the point of view of general relativity. Gravity becomes no longer a force, but a property of the space (more properly spacetime) through which the particle moves. This does require us to give up the global Euclidean structure of space, however. Near the earth, you can set up a freely-falling Cartesian coordinate system over a distance small compared to the size of the earth, but if you try to extend it to the other side of the planet, particles stationary with respect to that frame are not freely falling. In fact, they're accelerating upward! Something has to give, and that something is the geometry of space.