Contents

II Part II: Action, Symmetry, and Conservation

III Part III: Space and Time Unified

IV Part IV: Electromagnetism and Gravity

TWO Book Two: From the Happiest Thought to the Universe

V Part V: Equivalence Principle and Curved Spacetime

VI Part VI: Einstein's Field Equation Derived and Put to Work

VII Part VII: Black Holes

VIII Part VIII: Introduction to Our Universe

THREE Book Three: Gravity at Work and at Play

IX Part IX: Aspects of Gravity

I.1 Newton's Laws

The foundational equation of our subject

> For in those days I was in the prime of my age for invention and minded Mathematicks & Philosophy more than at any time since.
>
> —Newton describing his youth in his memoirs

Let us start with one of Newton's laws, which curiously enough is spoken as $F = ma$ but written as $ma = F$. For a point particle moving in D-dimensional space with position given by $\vec{x}(t) = (x^1(t), x^2(t), \cdots, x^D(t))$, Mr. Newton taught us that

$$m\frac{d^2x^i}{dt^2} = F^i \tag{1}$$

with the index* $i = 1, \cdots, D$. For $D \leq 3$ the coordinates have traditional "names": for example, for $D = 3$, x^1, x^2, x^3 are often called, with some affection, x, y, z, respectively.

Bad notation alert! In teaching physics, I sometimes feel, with only slight exaggeration, that students are confused by bad notation almost as much as by the concepts. I am using the standard notation of x and t here, but the letter x does double duty, as the position of the particle, which more strictly should be denoted by $x^i(t)$ or $\vec{x}(t)$, and as the space coordinates x^i, which are variables ranging from $-\infty$ to ∞ and which certainly are independent of t.

The different status between x and t in say (1) is particularly glaring if $N > 1$ particles are involved, in which case we write $m\frac{d^2x^i{}_a}{dt^2} = F^i{}_a$ or $m\frac{d^2\vec{x}_a}{dt^2} = \vec{F}_a$ with $x^i{}_a(t)$ for $a = 1, 2, \cdots, N$. But certainly t_a is a meaningless concept in Newtonian physics. In the Newtonian universe, t is the time ticked off by a universal clock, while $\vec{x}_a(t)$ is each particle's private business. We will have plenty more to say about this point. Here $x^i{}_a(t)$ are $3N$ functions of t, but there are still only 3 x^i.

* See appendix 2.

Some readers may feel that I am overly pedantic here, but in fact this fundamental inequality of status between x and t will come to a head when we get to the special theory of relativity. (I now drop the arrow on $\vec{x}$.) Perhaps Einstein as a student was bothered by this bad notation. One way to remedy the situation is to use q (or q_a) to denote the position of particles, as in more advanced treatments. But here I bow to tradition and continue to use x.

Have differential equation, will solve

After Newton's great insight, we "merely" have to solve some second order differential equations.

To understand Newton's fabulous equation, it's best to work through a few examples. (I need hardly say that if you do not already know Newtonian mechanics, you are unlikely to be able to learn it here.)

A priori, the force F^i could depend on any number of things, but from experience we know that in many simple cases, it depends only on x and not on t or $\frac{dx}{dt}$. As physicists unravel the mysteries of Nature, it becomes increasingly clear that fundamental forces are derived from an underlying quantum field theory and that they have simple forms. Complicated forces often merely result from some approximations we make in particular situations.

Example A

A particle in 1-dimensional space tied to a spring oscillates back and forth.

The force F is a function of space. Newton's equation

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

is easily solved in terms of two integration constants: $x(t) = a\cos\omega t + b\sin\omega t$, with $\omega = \sqrt{\frac{k}{m}}$. The two constants a and b are determined by the initial position and initial velocity, or alternatively* by the initial position at $t = 0$ and by the final position at some time $t = T$. Energy, but not momentum, is conserved.

Example B

We kick a particle in 1-dimensional space at $t = 0$.

The force F is a function of time. This example allows me to introduce the highly useful Dirac[1] delta function, or simply delta function.[2] By the word "kick" we mean that the time scale τ during which the force acts is much less than the other time scales we are

* See part II.

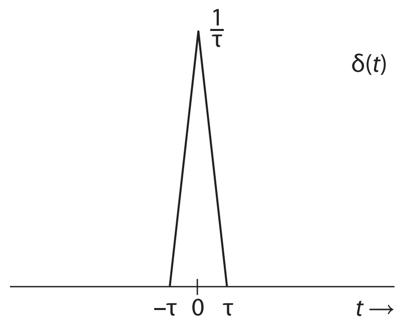


Figure 1 The delta function, which could be thought of as an infinitely sharp spike, is strictly speaking not a function, but the limit of a sequence of functions.

interested in. Thus, take $F(t) = w\delta(t)$, where the function $\delta(t)$ rises sharply just before $t = 0$, rapidly reaches its maximum at $t = 0$, and then sharply drops to 0. Because we included a multiplicative constant w, we could always normalize $\delta(t)$ by

$$\int dt\ \delta(t) = 1 \tag{3}$$

As we will see presently, the precise form of $\delta(t)$ does not matter. For example, we could take $\delta(t)$ to rise linearly from 0 at $t = -\tau$, reach a peak value of $1/\tau$ at $t = 0$, and then fall linearly to 0 at $t = \tau$. For $t < -\tau$ and for $t > \tau$, the function $\delta(t)$ is defined to be zero. Take the limit $\tau \to 0$, in which this function is known as the delta function. In other words the delta function is an infinitely sharp spike. See figure 1.

The δ function is somehow treated as an advanced topic in mathematical physics, but in fact, as you will see, it is an extremely useful function that I will use extensively in this book, for example in chapters II.1 and III.6. More properties of the δ function will be introduced as needed.

Integrating

$$\frac{d^2x}{dt^2} = \frac{w}{m}\delta(t) \tag{4}$$

from some time $t_- < 0$ to some time $t_+ > 0$, we obtain the change in velocity $v \equiv \frac{dx}{dt}$:

$$v(t_+) - v(t_-) = \frac{w}{m} \tag{5}$$

Note that in this example, neither energy nor momentum is conserved. The lack of conservation is easy to understand: (4) does not include the agent administering the kick. In general, a time-dependent force indicates that the description is not dynamically complete.

Example C

A planet approximately described as a point particle of mass m goes around its sun of mass $M \gg m$.

This is of course the celebrated problem Newton solved to unify celestial and terrestrial mechanics, previously thought to be two different areas of physics. His equation now reads

$$m\frac{d^2\vec{r}}{dt^2} = -GMm\frac{\vec{r}}{r^3} \tag{6}$$

where we use the notation $\vec{r} = (x, y, z)$ and $r = \sqrt{\vec{x}\cdot\vec{x}} = \sqrt{x^2 + y^2 + z^2}$.

John Wheeler has emphasized the interesting point that while Newton's law (1) tells us how a particle moves in space as a function of time, we tend to think of the trajectory of a particle as a curve fixed in space. For example, when we think of the motion of a planet around the sun, we think of an ellipse rather than a spiral around the time axis. Even in Newtonian mechanics, it is often illuminating to think in terms of a spacetime picture rather than a picture in space.[3]

Newton and his two distinct masses

> By thinking on it continually.
> —Newton (reply given when asked how he discovered the law of gravity)

Conceptually, in (6), m represents two distinct physical notions of mass. On the left hand side, the inertial mass measures the reluctance of the object to move. On the right hand side, the gravitational mass measures how strongly the object responds to a gravitational field. The equality of the inertial and the gravitational mass was what Galileo tried to verify in his famous apocryphal experiment dropping different objects from the Leaning Tower of Pisa. Newton himself experimented with a pendulum consisting of a hollow wooden box, which he proceeded to fill with different substances, such as sand and water. In our own times, this equality has been experimentally verified[4,5] to incredible accuracy.

That the same m appears on both sides of the equation turns out to be one of the greatest mysteries in physics before Einstein came along. His great insight was that this unexplained fact provided the clue to a deeper understanding of gravity. At this point, all we care about this mysterious equality is that m cancels out of (6), so that $\ddot{\vec{r}} = -\kappa\frac{\vec{r}}{r^3}$, with $\kappa \equiv GM$.

Celestial mechanics solved

Since the force is "central," namely it points in the direction of $\vec{r}$, a simple symmetry argument shows that the motion is confined to a plane, which we take to be the $(x\text{-}y)$ plane. Set $z = 0$ and we are left with

$$\ddot{x} = -\kappa x/r^3 \quad \text{and} \quad \ddot{y} = -\kappa y/r^3 \tag{7}$$

I have already, without warning, switched from Leibniz's notation to Newton's dot notation

$$\dot{x} \equiv \frac{dx}{dt} \quad \text{and} \quad \ddot{x} \equiv \frac{d^2x}{dt^2} \tag{8}$$

Since this is one of the most beautiful problems[6] in theoretical physics, I cannot resist solving it here in all its glory. Think of this as a warm-up before we do the heavy lifting of learning Einstein gravity. Also, later, we can compare the solution here with Einstein's solution.

Evidently, we should change from Cartesian coordinates (x, y) to polar coordinates (r, θ). We will do it by brute force to show, in contrast, the elegance of the formalism we will develop later. Differentiate

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta \tag{9}$$

twice to obtain first

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta\,\dot{\theta} \quad \text{and} \quad \dot{y} = \dot{r}\sin\theta + r\cos\theta\,\dot{\theta} \tag{10}$$

and then

$$\begin{aligned} &\ddot{x} = \ddot{r}\cos\theta - 2\dot{r}\sin\theta\,\dot{\theta} - r\cos\theta\,\dot{\theta}^2 - r\sin\theta\,\ddot{\theta} \\ \text{and} \quad &\ddot{y} = \ddot{r}\sin\theta + 2\dot{r}\cos\theta\,\dot{\theta} - r\sin\theta\,\dot{\theta}^2 + r\cos\theta\,\ddot{\theta} \end{aligned} \tag{11}$$

(Note that in each pair of these equations, the second could be obtained from the first by the substitution $\theta \to \theta - \frac{\pi}{2}$, so that $\cos\theta \to \sin\theta$, and $\sin\theta \to -\cos\theta$.)

Multiplying the first equation in (7) by $\cos\theta$ and the second by $\sin\theta$ and adding, we obtain, using (11),

$$\ddot{r} - r\dot{\theta}^2 = -\frac{\kappa}{r^2} \tag{12}$$

On the other hand, multiplying the first equation in (7) by $\sin\theta$ and the second by $\cos\theta$ and subtracting, we have

$$2\dot{r}\dot{\theta} + r\ddot{\theta} = 0 \tag{13}$$

I remind the reader again that we are doing all this in a clumsy brute force way to show the power of the formalism we are going to develop later.

After staring at (13) we recognize that it is equivalent to

$$\frac{d}{dt}(r^2\dot{\theta}) = 0 \tag{14}$$

which implies that

$$\dot{\theta} = \frac{l}{r^2} \tag{15}$$

for some constant l. Inserting this into (12), we have

$$\ddot{r} = \frac{l^2}{r^3} - \frac{\kappa}{r^2} = -\frac{dv(r)}{dr} \tag{16}$$

where we have defined

$$v(r) = \frac{l^2}{2r^2} - \frac{\kappa}{r} \tag{17}$$

Multiplying (16) by $\dot{r}$ and integrating over t, we have

$$\int dt\, \frac{1}{2}\frac{d}{dt}\dot{r}^2 = \int dt\, \dot{r}\ddot{r} = -\int dt \frac{dr}{dt}\frac{dv(r)}{dr} = -\int dr \frac{dv(r)}{dr}$$

so that finally

$$\frac{1}{2}\dot{r}^2 + v(r) = \epsilon \tag{18}$$

with ϵ an integration constant.

This describes a unit mass particle moving in the potential $v(r)$ with energy ϵ. Plot $v(r)$. Clearly, if ϵ is equal to the minimum of the potential $v_{min} = -\frac{\kappa^2}{2l^2}$, then $\dot{r} = 0$ and r stays constant. The planet follows a circular orbit of radius l^2/κ. If $\epsilon > v_{min}$ the orbit is elliptical, with r varying between r_{min} (perihelion) and r_{max} (aphelion) defined by the solutions to $\epsilon = v(r)$. For $\epsilon > 0$ the planet is not bound and should not even be called a planet.

We have stumbled across two conserved quantities, the angular momentum l and the energy ϵ per unit mass, seemingly by accident. They emerged as integration constants, but surely there should be a more fundamental and satisfying way of understanding conservation laws. We will see in chapter II.4 that there is.

Orbit closes

One fascinating apparent mystery is that the orbit closes. In other words, as the particle goes from r_{min} to r_{max} and then back to r_{min}, θ changes by precisely 2π. To verify that this is so, solve (18) for $\dot{r}$ and divide by (15) to obtain $\frac{dr}{d\theta} = \pm(r^2/l)\sqrt{2(\epsilon - v(r))}$. Changing variable from r to $u = 1/r$, we see, using (17), that $2(\epsilon - v(r))$ becomes the quadratic polynomial $2\epsilon - l^2u^2 + 2\kappa u$, which we can write in terms of its two roots as $l^2(u_{max} - u)(u - u_{min})$. Since u varies between u_{min} and u_{max}, we are led to make another change of variable from $u = u_{min} + (u_{max} - u_{min}) \sin^2 \zeta$ to ζ, so that ζ ranges from 0 to $\frac{\pi}{2}$. Thus, as the particle completes one round trip excursion in r, the polar angle changes by (note that $u_{min} = 1/r_{max}$ and $u_{max} = 1/r_{min}$)

$$\begin{aligned}\Delta\theta &= 2\int_{r_{min}}^{r_{max}} \frac{l\,dr}{r^2\sqrt{2(\epsilon - v(r))}} = 2\int_{u_{min}}^{u_{max}} \frac{l\,du}{\sqrt{2\epsilon - l^2u^2 + 2\kappa u}} \\ &= 2\int_{u_{min}}^{u_{max}} \frac{du}{\sqrt{(u_{max} - u)(u - u_{min})}} = 4\int_0^{\frac{\pi}{2}} d\zeta = 2\pi \end{aligned} \tag{19}$$

That this integral turns out to be exactly 2π is at this stage nothing less than an apparent miracle. Surely, there is something deeper going on, which we will reveal in chapter I.4. Note also that the inverse square law is crucial here. Incidentally, the change of variable here indicates how the Newtonian orbit* (and also the Einsteinian orbit, as we will see in part VI) could be determined. See exercise 2.

Bad notation alert! In (1), the force on the right hand side should be written as $F^i(x(t))$ (in many cases). In C, the gravitational force exists everywhere, namely $F(x)$ exists as a function, and what appears in Newton's equation is just $F(x)$ evaluated at the position of the particle $x(t)$. In contrast, in A, with a mass pulled by a spring, $F(x)$ does not make sense, only $F(x(t))$ does. The force exerted by the spring does not pervade all of space, and hence is defined only at the position of the particle $x(t)$, not at any old x. I can practically hear the reader chuckling, wondering what kind of person I could be addressing here, but believe me, I have encountered plenty of students who confuse these two basic concepts: spatial coordinates and the location of particles. I may sound awfully pedantic, but when we get to curved spacetime, it is often important to be clear that certain quantities are defined only on so-called geodesic curves, while others are defined everywhere in spacetime.

A historical digression on the so-called Newton's constant

> Wouldn't we be better off with the two eyes we now have plus a third that would tell us what is sneaking up behind? . . . With six eyes, we could have precise stereoscopic vision in all directions at once, including straight up. A six-eyed Newton might have dodged that apple and bequeathed us some levity rather than gravity.
>
> —George C. Williams[7]

Physics textbooks by necessity cannot do justice to physics history. As you probably know, in the *Principia,* Newton (1642–1727) converted his calculus-based calculations to geometric arguments,[8] which most modern readers find rather difficult to follow. Here I want to mention another curious point: Newton never did specifically define what we call his constant G. What he did with $ma = GMm/r^2$ was to compare the moon's acceleration with the apple's acceleration: $a_{\text{moon}} R^2_{\text{lunar orbit}} = GM_{\text{earth}} = a_{\text{apple}} R^2_{\text{radius of earth}}$. But to write $GM_{\text{earth}} = a_{\text{apple}} R^2_{\text{radius of earth}}$, he had to prove what is sometimes referred to as the first of Newton's two "superb theorems," namely that with the inverse square law the gravitational force exerted by a spherical mass distribution acts as if the entire mass were concentrated in a point at the center of the distribution. (See exercise 4.) Even with his abilities, Newton had to struggle for almost 20 years, the length of which contributed to the bitter priority fight he had with Hooke on the inverse square law, with Newton claiming that he had the law a long time before publication. You should be able to do it faster by a factor of $\sim 10^4$ as an exercise.

* On the old one pound note, a portrait of Newton together with his orbits appears on the back. Amusingly, the artist felt compelled to put the sun at the center, rather than one of the foci, of the ellipse.

Knowing the moon's period and $R_{\text{lunar orbit}}$, Newton could calculate a_{moon}. Since $R_{\text{radius of earth}}$ had been known since antiquity, he was thus able to calculate a_{apple} and obtained agreement* with Galileo's measurement of a_{apple}. This of course represents one of the most magnificent advances in physics history, with Newton unifying[9] the previously disparate subjects of celestial and terrestrial mechanics in one stroke. I don't have space to dwell on this here, but I do want to call your attention to the fact that Newton did not need to know G and M_{earth} to perform his feat.

Indeed, G was not measured until 1798 by Henry Cavendish (1731–1810) using equipment built and designed by his friend John Michell (1724–1793), now of black hole fame, who died before he could carry out the experiment.

Needless to say, what I have presented here should only be regarded as a comic book version of history.

Appendix 1: Where is hell?

You will find it in this appendix, sort of.

Curiously, contrary to what some textbooks and popular books stated, Cavendish's goal was not to measure G, but M_{earth} and hence the earth's density. Why this was of more interest to physicists of the time than G is in itself another interesting tidbit in physics history.

I mentioned that Newton had two superb theorems and that the first triggered his feud with Hooke. His second superb theorem states that there is no gravitational force inside a spherical shell.[10] Are you curious why Newton would even attack such a problem? An erroneous calculation had convinced him that the earth was much less dense than the moon, which led his friend Edmond Halley (1656–1742), who by the way published the *Principia* at his expense, to propose the hollow earth theory.[11] Witness the popularity of the idea in science fiction, notably Jules Verne's *Journey to the Center of the Earth* (1864). The idea may seem absurd to us, but at that time, a location for hell had to be found, and leading physicists gave serious thought to this pressing problem. Every epoch in physics has its own top ten problems.

So now we understand Cavendish's interest in M_{earth} and hence in the density of the earth rather than in G. Some textbooks give the impression that people easily obtained M_{earth} by multiplying the density of rock and the volume of the earth. Not so easy if you think that the earth might be hollow! We learn from Newton's second theorem that there is no gravitational force in hell, so the usual portrayal of the leaping flames can't be right!

Appendix 2: Fear of indices

Occasionally, a student or two would profess, unaccountably, a "fear of indices." In fact, there is nothing to fear.[12] At this stage, just stand back and admire how clever the invention of indices is. Instead of giving names to each coordinate axis, such as x, y, and z, we could pass fluidly between different dimensions by writing x^i, with $i = 1, 2, \cdots, D$. The length of the alphabet we use does not limit us, and we could easily go beyond 26 dimensions.

When we get to Einstein's theory, there will be a flood of indices, and we will have to distinguish between upper and lower indices. In Newtonian mechanics, there is no significance to whether we write the index as a superscript or a subscript. Have no fear: we will discuss each of these features of indices when the need arises. At this point, we merely note that a quantity can carry more than one index. In the text, we wrote $x^i{}_a$, with $i = 1, 2, \cdots, D$ labeling the different spatial directions, and $a = 1, 2, \cdots, N$ labeling the different particles. We will encounter more examples as we go along.

* Newton's first try did not lead to excellent agreement, because the value for the earth's equatorial radius was off. Just a reminder that physics never progresses as smoothly as textbooks say.

With only slight exaggeration, we could say that the invention of indices represents one of the really clever ideas[13] in the history of physics and mathematics, almost a "magic trick" that enables us to deal with as many particles in as many spatial dimensions as we like with the mere addition of some subscripts and superscripts.

Exercises

1 Show that for some suitably smooth function $f(x)$, the integral $\int_{-\infty}^{\infty} dx \delta(x) f(x) = f(0)$. Then show that $\delta(ax) = \delta(x)/|a|$ by evaluating the integral $\int_{-\infty}^{\infty} dx \delta(ax) f(x)$ for some smooth function $f(x)$.

2 Determine the orbit $r(\theta)$ by changing variable from r to $u = 1/r$. We will need the result of this exercise later.

3 Newton thought that light consists of "corpuscles." Calculate the deflection of light by the sun, applying what you learned in the text to the case $\epsilon > 0$. Note that the mass of these minute "particles of light" drops out in Newtonian theory anyway. We will need this result to compare with Einstein's theory later in chapter VI.3.

4 Prove Newton's first superb theorem: the gravitational force exerted by a spherical mass distribution acts as if the entire mass were concentrated in a point at the center of the distribution.

5 Prove Newton's second superb theorem.

6 Suppose engineers can build a straight tunnel connecting two cities on earth. Then we could have a free unpowered "gravity express"[14] by simply dropping a railroad car into the tunnel, allowing it to fall from one city to the other. Use Newton's two superb theorems to calculate the transit time.

Notes

1. Also introduced by Cauchy, Poisson, Hermite, Kirchoff, Kelvin, Helmholtz, and Heaviside. See J. D. Jackson, *Am. J. Phys.* 76 (2008), pp. 707–709.
2. Rigorous mathematicians go berserk at physicists' use of the word "function" here; they prefer to call it a distribution, defined as the limit of a function. But working physicists do not give a flying barf about such niceties. In any case, I do not personally know a theoretical physicist suffering any harm by calling $\delta(t)$ a function.
3. Consider a game of tennis. Compare a hard drive down the line and a soft lob high over the net. In both cases, we are to solve Newton's law $\frac{d^2x}{dt^2} = 0$, $\frac{d^2y}{dt^2} = -g$, with the boundary conditions $x(0) = 0$, $x(T) = L$, and $y(0) = y(T) = 0$. (The problem is so elementary that we won't bother to explain the notation, that y denotes the vertical direction, that $y = 0$ is the ground, that T is the time of flight before the ball hits the ground, that L is the length of the tennis court, and so on and so forth. You might want to draw your own figure.) The solution is $x = Lt/T$, $y = \frac{1}{2}g(T - t)t$. Note that the two types of tennis shots are governed by the same equation and the same L. Hence we obtain the same solution, but keep in mind that T is small in the case of the hard drive and that T is large in the case of the soft lob. Now eliminate t to obtain y as a function of x, namely $y(x) = \frac{1}{2}gT^2(1 - \frac{x}{L})\frac{x}{L}$, a parabola in both cases (of course). But compare the curvature of the two parabolas: we have $\frac{d^2y}{dx^2} = -g(T/L)^2$, very small in the case of the hard drive (small T) and very large in the case of the lob (large T). The hard drive down the line barely skimming over the net, and the soft lob climbing lazily high up into the sky, look and feel totally different pictured in space. In contrast, consider y as a function of t. We also have two parabolas (of course), namely $y(t) = \frac{1}{2}g(T - t)t$, as given earlier. Now compare the curvature of the two parabolas: we have $\frac{d^2y}{dt^2} = -g$, the same in both cases. The curvature of the ball's trajectory in spacetime is universal (universal gravity, get it?). But we tend to see in our mind's eye the two parabolas $y(x)$ in space, one for the hard drive and one for the lob, which look quite different, rather than the parabolas $y(t)$ in spacetime, which have the same curvature. I learned this long ago from John Wheeler.
4. Currently to one part in 10^{13}. The modern round of experiments started with Loránd Eötvös in 1885 and continues with the Eöt-Wash experiment led by E. Adelberger in our days.

5. The equality of the gravitational and inertial mass of the neutron has also been verified to good accuracy using neutron interferometry.
6. For Newton's letter to Halley about Hooke on the inverse square, see P. J. Nahin, *Mrs. Perkins's Electric Quilt,* Princeton University Press, 2009.
7. G. C. Williams, *The Pony Fish's Glow,* Basic Books, 1997, p. 128.
8. S. Chandrasekhar, *Newton's Principia for the Common Reader,* Oxford University Press, 2003.
9. *Fearful,* pp. 74–75.
10. For a popular account, see *Toy/Universe*.
11. N. Kollerstrom, "The Hollow World of Edmond Halley," *J. Hist. Astronomy* 23 (1992), p. 185.
12. Surely most readers are familiar with indices. My son the biologist informs me that even biologists use indices routinely; for example, on p. 20 of *Genetics and Analysis of Quantitative Traits* by M. Lynch and B. Walsh, indices appear without explanation or apology.
13. A colleague told me to mention that indices are crucial in computer programming, something that many readers can relate to.
14. *Toy/Universe,* p. xxix.

I.2 Conservation Is Good

An integrability condition

Conservation has been important to physics from day one.[1] In this chapter, we discuss the origin of various conservation laws in Newtonian mechanics.

The most important case is when the force F^i depends only on x and can be written in the form

$$F^i(x) = -\frac{\partial V(x)}{\partial x^i} \tag{1}$$

for $i = 1, 2, \cdots, D$. As we all learned, $V(x)$ is called the potential.

Suppose such a function $V(x)$ exists; then a clever person might have the insight to multiply each of Newton's equations

$$m\frac{d^2x^i}{dt^2} = F^i = -\frac{\partial V(x)}{\partial x^i} \tag{2}$$

by $\frac{dx^i}{dt}$ to obtain the D equations

$$m\frac{d^2x^i}{dt^2}\frac{dx^i}{dt} = -\frac{\partial V(x)}{\partial x^i}\frac{dx^i}{dt}, \qquad \text{with } i = 1, \cdots, D \tag{3}$$

He or she would then recognize that the sum of these D equations could be written as

$$\frac{d}{dt}\left[\frac{1}{2}m\sum_i\left(\frac{dx^i}{dt}\right)^2 + V(x)\right] = 0 \tag{4}$$

which we could verify by explicit differentiation. Lo and behold, the total energy, defined by

$$E = \frac{1}{2}m\sum_i\left(\frac{dx^i}{dt}\right)^2 + V(x) \tag{5}$$

is conserved. It does not change in time.

For $D = 1$, (1) holds automatically: $V(x)$ is simply given by $-\int^x dx' F(x')$. For $D > 1$, the D equations in (1), namely $F^i(x) = -\frac{\partial V(x)}{\partial x^i}$, imply the consistency or integrability condition

$$\frac{\partial F^i(x)}{\partial x^j} = \frac{\partial F^j(x)}{\partial x^i} \tag{6}$$

(Since derivatives commute, both sides of (6) are equal to $-\frac{\partial^2 V(x)}{\partial x^i \partial x^j}$.) Thus, given $F^i(x)$, we merely have to check to see whether (6) holds. If not, then V does not exist. If yes, then we could integrate $F^i(x) = -\frac{\partial V(x)}{\partial x^i}$ for each i to determine V.

Apples do not fall down

Suppose $V(r)$ depends only on $r \equiv \left(\sum_{i=1}^D (x^i)^2\right)^{\frac{1}{2}}$. In other words, the potential does not pick out any preferred direction. We take this for granted nowadays, but it represents one of the most astonishing insights of physics.[2] Newton realized that the apple did not fall down, but toward the center of the earth.

Differentiating $r^2 = \sum_{i=1}^D (x^i)^2$, we obtain $r dr = \sum_i x^i dx^i$ (an "identity," which we will use again and again in this text) or $\frac{\partial r}{\partial x^j} = \frac{x^j}{r}$, so that

$$F^i = -\frac{x^i}{r} V'(r) \quad \text{and} \quad \frac{\partial F^i(x)}{\partial x^j} = -\frac{1}{r}[\delta^{ij} V'(r) + \frac{x^i x^j}{r^2}(-V'(r) + r V''(r))]$$

which is manifestly symmetric under $i \leftrightarrow j$.

Here we have introduced the Kronecker delta δ^{ij}, defined by

$$\delta^{kj} = 1 \text{ if } k = j, \qquad \delta^{kj} = 0 \text{ if } k \neq j \tag{7}$$

(which we can think of as an ancestor of the Dirac delta function[3] introduced in chapter I.1).

The important point is not the somewhat involved expression for $\frac{\partial F^i(x)}{\partial x^j}$, but that it is a linear combination of δ^{ij} and $x^i x^j$. We haven't talked about tensors yet (see chapter I.4), but this result could have been anticipated by a "what else can it be?" type of argument. Not having any preferred direction, we could only construct an object with indices i and j out of δ^{ij} and $x^i x^j$. We could have seen immediately that the integrability condition (6) holds.

Note that this discussion holds for any value of D.

Conservation of angular momentum

Suppose the force in (2) points toward the center, so that it has the form $F^i = f(r)x^i$ (with $f(r) = -V'(r)/r$, as we just saw). Then we obtain angular momentum conservation immediately. To see this, multiply Newton's equation (2)

$$m\frac{d^2 x^i}{dt^2} = f(r)x^i \tag{8}$$

by x^j, so that $m\frac{d^2x^i}{dt^2}x^j = f(r)x^i x^j$. Subtract from this the same equation but with i and j interchanged. Regardless of the function $f(r)$, we find

$$x^j \frac{d^2x^i}{dt^2} - x^i \frac{d^2x^j}{dt^2} = 0 \tag{9}$$

But this is the same as

$$\frac{d}{dt}\left(x^j \frac{dx^i}{dt} - x^i \frac{dx^j}{dt}\right) = 0 \tag{10}$$

Clever, eh? I am constantly amazed by how brilliant early physicists were.

The quantity $l^{ij} \equiv \left(x^j \frac{dx^i}{dt} - x^i \frac{dx^j}{dt}\right)$, the angular momentum per unit mass, is conserved. Recall that in the preceding chapter, this fact seemingly fell out when we changed to polar coordinates. Note also that the argument given here holds for any $D \geq 2$.

Exercise

1 Let N particles interact according to

$$m_a \frac{d^2x_a^i}{dt^2} = -\frac{\partial V(x)}{\partial x_a^i} \tag{11}$$

with $a = 1, \cdots, N$. Suppose $V(x_1, \cdots, x_N)$ depends only on the differences $x_a^i - x_b^i$, with $a, b = 1, \cdots, N$. Show that the total momentum $\sum_a m_a \frac{dx_a^i}{dt}$ is conserved.

Notes

1. *Fearful.*
2. I once explained this point to humanists using Einstein's terminology by saying that "The words up and down have no place in the Mind of the Creator." See A. Zee, *New Lit. Hist.* 23 (1992), pp. 815–838. See also web.physics.ucsb.edu/jatila/supplements/zee lecture.pdf.
3. In the sense that $\delta(x - y)$ is zero for $x \neq y$.

I.3 Rotation: Invariance and Infinitesimal Transformation

Rotation in the plane

My pedagogical strategy for this chapter is to take something you know extremely* well, namely rotations in the plane, present it in a way possibly unfamiliar to you, and go through it slowly in great detail, "beating the subject to death," so to speak.

I have already mentioned that Monsieur Descartes had the clever idea of reducing geometry to algebra. Put down Cartesian coordinate axes so that a point P is labeled by two real numbers (x, y). Suppose another observer (call him Mr. Prime) puts down coordinate axes rotated by angle θ with respect to the axes put down by the first observer (call her Ms. Unprime) but sharing the same origin O. Elementary trigonometry tells us that the coordinates (x, y) and (x', y') assigned by the two observers to the same point P are related by† (see figure 1)

$$x' = \cos\theta\, x + \sin\theta\, y, \qquad y' = -\sin\theta\, x + \cos\theta\, y \tag{1}$$

The distance from P to the origin O of course has to be the same for the two observers. According to Pythagoras, this requires $\sqrt{x'^2 + y'^2} = \sqrt{x^2 + y^2}$, which you can check using (1).

Introduce the column vectors $\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $\vec{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ and the rotation matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{2}$$

so that we can write (1) more compactly as $\vec{r}' = R(\theta)\vec{r}$.

* If you don't know rotations in the plane extremely well, then perhaps you are not ready for this book. A nodding familiarity with matrices and linear algebra is among the prerequisites.

† For example, by comparing similar triangles in the figure, we obtain $x' = (x/\cos\theta) + (y - x\tan\theta)\sin\theta$.

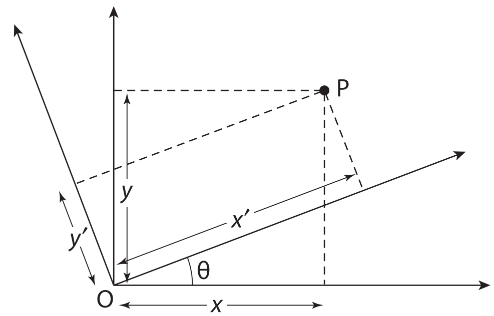


Figure 1 The same point P is labeled by (x, y) and (x', y'), depending on the observer's frame of reference.

As you recall from a course on mechanics, we can either envisage rotating the physical body we are studying or rotating the observer. We will consistently rotate the observer.

We have already used the word "vector." A vector is a physical quantity (for example the velocity of a particle in the plane) consisting of two real numbers, so that if Ms. Unprime represents it by $\vec{p} = \begin{pmatrix} p^1 \\ p^2 \end{pmatrix}$, then Mr. Prime will represent it by $\vec{p}' = R(\theta)\vec{p}$. In short, a vector is something that transforms like the coordinates $\begin{pmatrix} x \\ y \end{pmatrix}$ under rotation.

Given two vectors $\vec{p} = \begin{pmatrix} p^1 \\ p^2 \end{pmatrix}$ and $\vec{q} = \begin{pmatrix} q^1 \\ q^2 \end{pmatrix}$, the scalar or dot product is defined by $\vec{p}^T \cdot \vec{q} = p^1 q^1 + p^2 q^2$. Here T stands for transpose and $\vec{p}^T$ the row vector (p^1, p^2). By definition, rotations leave $\vec{p}^2 \equiv \vec{p}^T \cdot \vec{p} = (p^1)^2 + (p^2)^2$ invariant. In other words, if $\vec{p}' = R(\theta)\vec{p}$, then $\vec{p}'^2 = \vec{p}^2$. Since this works for any vector $\vec{p}$, including the case in which $\vec{p}$ happens to be the sum of two vectors $\vec{p} = \vec{u} + \vec{v}$, and since $\vec{p}^2 = (\vec{u} + \vec{v})^2 = \vec{u}^2 + \vec{v}^2 + 2\vec{u}^T \cdot \vec{v}$, rotation also leaves the dot product between two arbitrary vectors invariant: the invariance of $\vec{p}^2$ implies that $\vec{u}'^T \cdot \vec{v}' = \vec{u}^T \cdot \vec{v}$.

Since $\vec{u}' = R\vec{u}$ (to unclutter things, we often suppress the θ dependence in $R(\theta)$) and so $\vec{u}'^T = \vec{u}^T R^T$, we now have $\vec{u}^T \cdot \vec{v} = \vec{u}'^T \cdot \vec{v}' = (\vec{u}^T R^T) \cdot (R\vec{v}) = \vec{u}^T \cdot (R^T R)\vec{v}$. (The transpose M^T of a matrix M is of course obtained by interchanging the rows and columns of M.) As this holds for any two vectors $\vec{u}$ and $\vec{v}$, we must have the matrix equation

$$R^T R = I \tag{3}$$

where, as usual, I denotes the identity or unit matrix: $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Indeed, we could verify (3) explicitly:

$$R(\theta)^T R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4}$$

Matrices that satisfy (3) are called orthogonal.

Taking the determinant of (3), we obtain $(\det R)^2 = 1$, that is, $\det R = \pm 1$. The determinant of an orthogonal matrix may be -1 as well as $+1$. In other words, orthogonal matrices also include reflection matrices, such as $\mathcal{P} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, a reflection in the y-axis.

To focus on rotations, let us exclude reflections by imposing the condition (since $\det \mathcal{P} = -1$)

$$\det R = 1 \tag{5}$$

Matrices with unit determinant are called special.

We define a rotation as a matrix that is both orthogonal and special, that is, a matrix that satisfies both (3) and (5). Thus, the rotation group of the plane consists of the set of all special orthogonal 2 by 2 matrices and is known as $SO(2)$.

Note that matrices of the form $\mathcal{P}R$ for any rotation R are also excluded by (5), since $\det(\mathcal{P}R) = \det \mathcal{P} \det R = (-1)(+1) = -1$. In particular, a reflection in the x-axis $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, which is the product of $\mathcal{P}$ and a rotation through 90°, is also excluded.

Act a little bit at a time

The Norwegian physicist Marius Sophus Lie (1842–1899) had the almost childishly obvious but brilliant idea that to rotate through, say, 29°, you could just as well rotate through a zillionth of a degree and repeat the process 29 zillion times. To study rotations, it suffices to study rotation through infinitesimal angles. Shades of Newton and Leibniz! A rotation through a finite angle could always be obtained by performing infinitesimal rotations repeatedly. As is typical with many profound statements in physics and mathematics, Lie's idea is astonishingly simple. Replace the proverb "Never put off until tomorrow what you have to do today" by "Do what you have to do a little bit at a time."

When the angle is small enough, the rotation is almost the identity, that is, no rotation at all. Thus, we can write

$$R(\theta) \simeq I + A \tag{6}$$

where A denotes some infinitesimal matrix.

Now suppose we have never seen (2). Indeed, suppose we have never even heard of sine and cosine. Instead, let us define rotations as the set of linear transformations on 2-component objects $\vec{u}' = R\vec{u}$ and $\vec{v}' = R\vec{v}$ that leave $\vec{u}^T \cdot \vec{v}$ invariant. Following Lie, we solve this condition on R, namely (3) $R^T R = I$, by considering an infinitesimal transformation $R(\theta) \simeq I + A$. Since by assumption, A^2 can be neglected relative to A, $R^T R \simeq (I + A^T)(I + A) \simeq (I + A^T + A) = I$. We thus obtain $A^T = -A$, namely that A must be antisymmetric. But there is basically only one 2-by-2 antisymmetric matrix:

$$\mathcal{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{7}$$

In other words, the solution of $A^T = -A$ is $A = \theta \mathcal{J}$ for some real number θ. Thus, rotations close to the identity have the form $R = I + \theta \mathcal{J} + O(\theta^2) = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix} + O(\theta^2)$. The antisymmetric matrix $\mathcal{J}$ is known as the generator of the rotation group.

An equivalent way of saying this is that for infinitesimal θ, the transformation $x' \simeq x + \theta y$ and $y' \simeq y - \theta x$ (you could verify that (1) indeed reduces to this to leading order in

θ) obviously satisfies the Pythagorean condition $x'^2 + y'^2 = x^2 + y^2$ to first order in θ. Or, write $x' = x + \delta x$, $y' = y + \delta y$ and solve $x\delta x + y\delta y = 0$.

Alternatively, simply draw figure 1 for θ infinitesimal. Since we know the transformation is linear, we could determine the matrix R in (6) by looking at the figure to see what happens to the two points ($x = 1$, $y = 0$) and ($x = 0$, $y = 1$) under an infinitesimal rotation.

Now recall the identity $e^x = \lim_{N\to\infty}(1 + \frac{x}{N})^N$ (which you can easily prove by differentiating both sides). Then, for a finite (that is, not infinitesimal) angle θ, we have

$$R(\theta) = \lim_{N\to\infty} R\left(\frac{\theta}{N}\right)^N = \lim_{N\to\infty}\left(1 + \frac{\theta\mathcal{J}}{N}\right)^N = e^{\theta\mathcal{J}} \tag{8}$$

The first equality represents Lie's profound idea. For the last equality, we use the identity just mentioned, which amounts to the definition of the exponential.

Some readers may not be familiar with the exponential of a matrix. Given a well-behaved function f with a power series expansion, we can define $f(M)$ for an arbitrary matrix M using that power series. For example, define $e^M \equiv \sum_{n=0}^{\infty} M^n/n!$; since we know how to multiply and add matrices, this series makes perfect sense. (Whether or not any given series converges is of course another issue.) We must be careful, however, in using various identities that may or may not generalize. For example, the identity $e^a e^a = e^{2a}$ for a a real number, which we could prove by applying the binomial theorem to the product of two series (square of a series in this case) generalizes immediately. Thus, $e^M e^M = e^{2M}$. But for two matrices M_1 and M_2 that do not commute with each other, $e^{M_1}e^{M_2} \neq e^{M_1+M_2}$.

This provides an alternative but of course equivalent path to our result. To leading order, we have every right to write $R\left(\frac{\theta}{N}\right) = 1 + \frac{\theta\mathcal{J}}{N} \simeq e^{\frac{\theta\mathcal{J}}{N}}$ and thus $R(\theta) = R\left(\frac{\theta}{N}\right)^N = e^{\theta\mathcal{J}}$.

Finally, we easily check that the formula $R(\theta) = e^{\theta\mathcal{J}}$ reproduces (2) for any value of θ. We simply note that $\mathcal{J}^2 = -I$ and separate the exponential series into even and odd terms. Thus

$$\begin{aligned} e^{\theta\mathcal{J}} &= \sum_{n=0}^{\infty}\theta^n\mathcal{J}^n/n! = \left(\sum_{k=0}^{\infty}(-1)^k\theta^{2k}/(2k)!\right)I + \left(\sum_{k=0}^{\infty}(-1)^k\theta^{2k+1}/(2k+1)!\right)\mathcal{J} \\ &= \cos\theta\, I + \sin\theta\,\mathcal{J} = \cos\theta\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \sin\theta\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix} \end{aligned} \tag{9}$$

which is precisely $R(\theta)$ as given in (2). Note this works because $\mathcal{J}$ plays the same role as i in the identity $e^{i\theta} = \cos\theta + i\sin\theta$.

Poor Lie, he never made it into the 20th century.

Two approaches to rotation

Notice that I actually gave you two different approaches to rotation. Let us summarize the two approaches. In the first approach, applying trigonometry to figure 1, we write down (1) and hence (2). In the second approach, we specify what is to be left invariant by rotations and hence define rotations by the condition (3) that rotations must satisfy. Lie then tells us that it suffices to solve (3) for infinitesimal rotations. We could then build up rotations through finite angles by multiplying infinitesimal rotations together, thus also arriving at (2).

It might seem that the first approach is much more direct. One writes down (2) and that is that. The second approach appears more roundabout and involves some "fancy math." It might even provoke an adherent of the first "more macho" approach to wisecrack, "Why, with a bit of higher education, sine and cosine are not good enough for you any more? You have to go around doing fancy math!" The point is that the second approach generalizes to higher dimensional spaces (and to other situations) much more readily than the first approach does, as we will see presently. Dear reader, in going through life, you would be well advised to always separate fancy but useful math from fancy but useless math.

Before we go on, let us take care of one technical detail. We assumed that Mr. Prime and Ms. Unprime set up their coordinate systems to share the same origin O. We now show that this condition is unnecessary if we consider two points P and Q (rather than one point, as in our discussion above) and study how the vector connecting P to Q transforms.

Let Ms. Unprime assign the coordinates $\vec{r}_P = (x, y)$ and $\vec{r}_Q = (\tilde{x}, \tilde{y})$ to P and Q, respectively. Then Mr. Prime's coordinates $\vec{r}'_P = (x', y')$ for P and $\vec{r}'_Q = (\tilde{x}', \tilde{y}')$ for Q are then given by $\vec{r}'_P = R(\theta)\vec{r}_P$ and $\vec{r}'_Q = R(\theta)\vec{r}_Q$. Subtracting the first equation from the second and defining $\Delta x = \tilde{x} - x$, $\Delta y = \tilde{y} - y$, and the corresponding primed quantities, we obtain

$$\begin{pmatrix} \Delta x' \\ \Delta y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \tag{10}$$

Rotations leave the distance between the points P and Q unchanged: $(\Delta x')^2 + (\Delta y')^2 = (\Delta x)^2 + (\Delta y)^2$. You recognize of course that this is a lot of tedious verbiage stating the perfectly obvious, but I want to be precise here. Of course, the distance between any two points is left unchanged by rotations. (This also means that the distance between P and the origin is left unchanged by rotations; ditto for the distance between Q and the origin.)

Invariance and geometry

There is no royal road to geometry.
—Euclid's advice to a prince

Let no one unversed in geometry enter here.
—Plato's motto, carved over the entrance to his academy

Let us take the two points P and Q to be infinitesimally close to each other and replace the differences $\Delta x'$, Δx, and so forth by differentials dx', dx, and so forth. Indeed, 2-dimensional Euclidean space is defined by the distance squared between two nearby points:

$$ds^2 = dx^2 + dy^2 \tag{11}$$

Rotations are defined as linear transformations* $(x, y) \to (x', y')$, such that

$$dx^2 + dy^2 = dx'^2 + dy'^2 \tag{12}$$

The whole point is that this now makes no reference to the origin O (and whether Mr. Prime and Ms. Unprime even share the same origin).

The column $d\vec{x} = \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix} \equiv \begin{pmatrix} dx \\ dy \end{pmatrix}$ is defined as the basic or ur-vector, the template for all other vectors. To repeat, a vector is defined as something that transforms like $d\vec{x}$ under rotations.

So, a vector is defined by how it transforms. An array of two numbers $\vec{p} = \begin{pmatrix} p^1 \\ p^2 \end{pmatrix}$ is a vector if it transforms according to $\vec{p}' = R(\theta)\vec{p}$.

Sometimes it is very helpful, in order to understand what something is, to be given an example of something that is not. As a simple example, given a $\vec{p}$, then $\begin{pmatrix} ap^1 \\ bp^2 \end{pmatrix}$ is definitely not a vector if $a \neq b$. (You could easily write down more outrageous examples, such as $\begin{pmatrix} (p^1)^2 p^2 \\ (p^1)^3 + (p^2)^3 \end{pmatrix}$. That ain't no vector!) You will work out further examples in exercise 1. An array of numbers is not a vector unless it transforms in the right way.[1]

Oh, about the advice Euclid gave to the prince who wanted to know a quick way of mastering geometry. Mr. E is also telling you that, to master the material covered in this book, there is no way other than to cogitate over the material until you get it and to work through as many exercises as possible.

From the plane to higher dimensional space

The reader who has wrestled with Euler angles in a mechanics course knows that the analog of (2) for 3-dimensional space is already quite a mess. In contrast, Lie's approach allows us, as mentioned above, to immediately jump to D-dimensional Euclidean space, defined by specifying the distance squared between two nearby points (compare this with (11)), as given by the obvious generalization of Pythagoras' theorem:

$$ds^2 = \sum_{i=1}^{D} \left(dx^i\right)^2 = \left(dx^1\right)^2 + \left(dx^2\right)^2 + \cdots + \left(dx^D\right)^2 \tag{13}$$

This is as good a place as any to say a word about indices. As I said in chapter I.1, in my experience teaching, there are always a couple of students confounded by indices. Dear reader, if you are not, you could simply laugh and skip to the next paragraph. Indices provide a marvelous notational device to save us from having to give names to individual elements belonging to a set. (For example, consider all humans h^i now alive, with $i = 1, 2, \cdots, P$ where P denotes the population size.) Take a look at the 19th century physics literature, before the use of indices became widespread. I am always amazed by

* Indeed, most, but not all, of the readers[2] of this book are constantly rotating between two coordinate systems.

the fact that, for example, Maxwell could see through the morass of the electromagnetic equations written out component by component.

Rotations are defined as linear transformations $d\vec{x}' = Rd\vec{x}$ that leave ds unchanged. The preceding discussion allows us to write this condition as $R^T R = I$. As before, we want to focus on rotations by imposing the additional condition $\det R = 1$. The set of D-by-D matrices R that satisfy these two conditions forms the simple orthogonal group $SO(D)$, which is just a fancy way of saying the rotation group in D-dimensional space.

Lie in higher dimensions

The power of Lie now shines through when we want to work out rotations in higher dimensional spaces. All we have to do is satisfy the two conditions $R^T R = I$ and $\det R = 1$.

So let us follow Lie and write $R \simeq I + A$. Then $R^T R = I$ is solved by requiring $A = -A^T$, namely that A must be antisymmetric. But it is very easy to write down all possible antisymmetric D-by-D matrices! For $D = 2$, there is basically only one: the $\mathcal{J}$ introduced earlier. For $D = 3$, there are basically three of them:

$$\mathcal{J}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathcal{J}_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{14}$$

Any 3-by-3 antisymmetric matrix can be written as $A = \theta_x \mathcal{J}_x + \theta_y \mathcal{J}_y + \theta_z \mathcal{J}_z$, with three real numbers θ_x, θ_y, and θ_z. At this point, you can verify that $R \simeq I + A$, with A as given here, satisfies the condition $\det R = 1$.

The three matrices $\mathcal{J}_x$, $\mathcal{J}_y$, $\mathcal{J}_z$ are known as the generators of the 3-dimensional rotation group $SO(3)$. They generate rotations, but are of course not to be confused with rotations, which are by definition 3-by-3 orthogonal matrices with determinant equal to 1.

The upshot of this whole discussion is that any 3-dimensional rotation (not necessarily infinitesimal) can be written as $R(\theta) = e^A$ and is thus characterized by three real numbers. As I said, those readers who have suffered through the rotation of a rigid body in a course on mechanics must appreciate the simplicity of studying the generators of infinitesimal rotations and then simply exponentiating them.

Index notation and rotations

Some readers will find this obvious, but others might find it helpful if we derive the condition $R^T R = I$ explicitly once again using the index notation. I prefer to go slow here, since we will need some of the same formalism later when we get to special relativity. Once the reader feels sure-footed, we could then dispense with indices.

Let me start by reminding the reader that a D-by-D matrix M carries two indices and has entries M^{ij}, with the standard convention that the first index labels the rows, the second the column (for $i, j = 1, 2, \cdots, D$). For example, for $D = 2$, $M = \left(\begin{smallmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{smallmatrix}\right)$, and M^{12} is the entry in the first row and the second column, whereas M^{21} is the entry in the second row and the first column. Note that the transpose of a matrix M is given by $(M^T)^{ji} \equiv M^{ij}$. Thus, if $\vec{v}$ is a column vector with entries v^j, then the entries of the column vector $\vec{u} = M\vec{v}$ are given by $u^i = \sum_j M^{ij} v^j$. For A and B two D-by-D matrices, the product AB is defined as the matrix with the entries $(AB)^{ij} = \sum_k A^{ik} B^{kj}$. (If everything here is news to you, see the first footnote in this chapter.)

Under a rotation,

$$dx'^i = \sum_j R^{ij} dx^j = R^{i1} dx^1 + R^{i2} dx^2 + \cdots + R^{iD} dx^D \tag{15}$$

(I have written the sum out explicitly for the benefit of the rare reader afflicted by fear of indices.) Also, as was mentioned in chapter I.1, at this stage it is completely arbitrary whether we write upper or lower indices.

Let us pause and recall the Kronecker delta symbol δ^{ij} introduced in (I.2.7), defined to be equal to $+1$ if $i = j$ and 0 otherwise, and which we can also think of as a D-by-D unit matrix. We will be encountering the highly useful Kronecker delta often in this book. For example, $\sum_j A^j B^j = \sum_k \sum_j \delta^{kj} A^k B^j$. Since δ^{kj} vanishes unless k is equal to j, the double sum on the right hand side collapses to the single sum on the left hand side. In other words, the Kronecker delta allows us to write a single sum as a double sum. It seems like a really silly thing to do, but as we will see presently, it is an extremely useful trick that we use quite often in this book.

We now determine how the matrix R must be restricted for it to be a rotation. The statement that $ds^2 = \sum_{i=1}^{D} (dx^i)^2$ as defined in (13) is left unchanged by the rotation implies that (with all indices running over $1, \cdots, D$)

$$\sum_i (dx'^i)^2 = \sum_i \sum_k \sum_j R^{ik} dx^k R^{ij} dx^j = \sum_j (dx^j)^2 = \sum_k \sum_j \delta^{kj} dx^k dx^j \tag{16}$$

In the last step, we used what we just learned.

Since the infinitesimals dx^i can take on arbitrary values, to have the second term equal to the last term in (16), we must equate the coefficients of $dx^k dx^j$ and demand that

$$\sum_i R^{ik} R^{ij} = \delta^{kj} = \sum_i (R^T)^{ki} R^{ij} = (R^T R)^{kj} \tag{17}$$

Indeed, we obtain $R^T R = I$ just as in (3), but now in D-dimensional space for any D.

We end this section with a trivial remark. So far in this chapter, we have written the column vectors as columns. But columns take up so much space, and so for typographical convenience (editors must be placated!) we will henceforth write the entries of a column vector as $d\vec{x} = (dx^1, dx^2, \cdots, dx^D)$, a practice we will indulge in throughout this book. (If we want to be insufferably pedantic, we could put in a T for transpose: the column ur-vector $d\vec{x} = (dx^1, dx^2, \cdots, dx^D)^T$.)

Einstein's repeated index summation

Observe that in all those sums in (16) the indices to be summed over always appear twice, that is, they are repeated. For example, in the second term in (16), $\sum_i \sum_k \sum_j R^{ik} dx^k R^{ij} dx^j$, the indices i, k, and j all appear repeated. Thus, we could adopt the so-called repeated index summation convention proposed by Albert Einstein himself: omit the pesky summation symbol and agree that if an index is repeated, then it is to be summed over. For example, $dx'^i = \sum_j R^{ij} dx^j$ can now be written as $dx'^i = R^{ij} dx^j$: in the expression on the right hand side, the index j appears twice and is thus to be summed over.* In contrast, i is a "free" index and does not appear twice in the same expression. Notice that free indices must match on opposite sides of any equation. It is rightly said that one of Einstein's greatest contribution to physics is the repeated index summation convention.† When we get to Einstein gravity, we will meet lots of indices to be summed over, and it would be silly to keep on writing the summation symbol.

Vector fields

The vectors we encounter may well vary in space. For example, the flow velocity in a fluid in general would depend on where we are. We are then dealing with a vector field $\vec{V}(\vec{x})$. Again, consider two observers studying the same vector field. Mr. Prime would see

$$\vec{V}'(\vec{x}') = R\vec{V}(\vec{x}) \tag{18}$$

with $\vec{x}' = R\vec{x}$ of course. In other words, the two observers are studying the same vector field at the same point P. See figure 2. As another example, the familiar electric $\vec{E}(\vec{x})$ and magnetic fields $\vec{B}(\vec{x})$ are both vector fields.

Physics should not depend on the observer

Let me stress again why physicists constantly talk about vectors. The laws of physics often involve the statement that one vector is equal to another, for example, Newton's law states $m\vec{a} = \vec{F}$. Applying a rotation matrix $R(\theta)$, we obtain $mR(\theta)\vec{a} = R(\theta)\vec{F}$. If $\vec{F}$ transforms like a vector, then $m\vec{a}' = \vec{F}'$. Ms. Unprime and Mr. Prime see the same Newton's law, and more generally, the same laws of physics!

This statement, while self-evident, is profound, and in some sense, it is what makes physics possible. Physics should not depend on the physicist. Ms. Unprime and Mr. Prime

* When a pair of repeated indices, such as j here, is summed over, they are often said to be contracted with each other. In a tiny abuse of terminology, people also say that R^{ij} is contracted with dx^j.

† It appeared only in his later work. In 1905, Einstein did not even use vector notation! In one system, the coordinates were denoted by x, y, z, in the other, by ξ, η, ζ; the components of the force acting on the electron were called X, Y, Z. To modern eyes, his notation was a horrific mess.

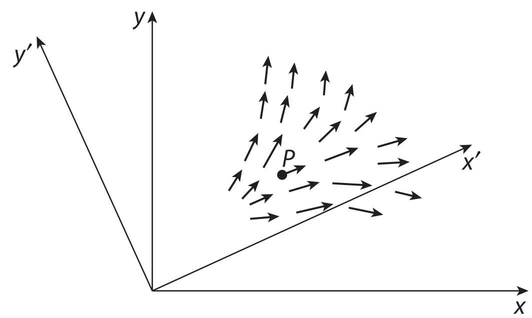


Figure 2 Two observers studying the same vector field.

see different accelerations $\vec{a}$ and $\vec{a}'$, and different forces $\vec{F}$ and $\vec{F}'$, but the same Newton's law. We say that Newton's law is invariant—that is, it does not change—under rotation.*

We should also remind ourselves that mass is an example of a scalar: a physical quantity that does not change under rotation. If it does change, Newton's law would not be invariant under rotation and one observer would be preferred over another, which is unacceptable. Physics rests on the democratic ideal.

Let me remind you that the gravitational force in the planetary problem studied in chapter I.1 is derived from what is sometimes called a central potential, namely one without a preferred direction: $F^i(x) = -\frac{\partial}{\partial x^i}V(r) = -\frac{x^i}{r}V'(r)$. Hence, $\vec{F}$ is proportional to $\vec{x}$ and so a fortiori transforms like a vector.

At this point, it may be worthwhile to be a bit more pedantic and professorial. Some authors give long-winded speeches about covariance versus invariance, and take great pain to distinguish the two. We should too. The equation $m\vec{a} = \vec{F}$ is covariant, that is, the two sides transform the same way under rotations. The physics expressed by Newton's second law is, however, invariant, that is, independent of observers related by a rotation. If physics depends on how you tilt your head, we are in trouble. Physics does not, but the way physics is expressed, in terms of equations, does.

Here is the profound and trivial statement. Under a certain set of transformations, a purportedly fundamental equation is said to be covariant if the two sides of the equation transform in the same way. If so, then that transformation is known as a symmetry of physics.[3] Physics is said to be invariant under that transformation. As we will see, both sides of Einstein's field equation transform in the same way, as tensors, under what are known as general coordinate transformations. I will explain what a tensor is in the next chapter. I will allow myself the luxury of using the words invariance and covariance interchangeably and simply trust you to be discerning.

Since we can always move the quantity on the right hand side of an equation to the left hand side, we can rewrite a physical law of the form $\vec{u} = \vec{v}$ in the form $\vec{w} \equiv \vec{u} - \vec{v} = 0$. Physics students sometimes joke that they could already write down the ultimate

* The reader who has already been exposed to the special theory of relativity knows that this notion of invariance represents the essence of Einstein's insight. We will of course have a great deal more to say about that!

equation of physics, namely $\mathcal{X} = 0$, whatever $\mathcal{X}$ is. Thus, the statement of invariance merely expresses the mathematically obvious fact that if $\vec{w} = 0$, then $R(\theta)\vec{w} = 0$. (Strictly speaking, the 0 on the right hand side should be written as $\vec{0}$, but we don't want to be that pedantic!)

Descartes versus Euclid

I remember how excited I was when I learned about analytic geometry. Surely you were excited too. What a genius, that Descartes! Henceforth, we could prove geometric theorems by doing algebra. After Descartes,[4] physics can no longer live without the concept of coordinates,* but he also managed to obscure what was once obvious to Euclid. We now must also insist on invariance. Indeed, the notion of invariance is at the heart of what we mean by geometry.

For example, suppose somebody hands you a formula for the area of a triangle with vertices at (a_1, b_1), (a_2, b_2), (a_3, b_3). You better insist that the formula is invariant under rotation. In fact, this requirement, plus the requirement that the area should scale as the square of the separation between the three vertices, suffices to determine the formula. This simple example rings in the central motif of this book.

Appendix 1: Differential operators rather than matrices

Here I have to divide readers into the haves and the have-nots, but only temporarily. What I will say may sound difficult, but really, it amounts to not much more than a notational triviality.

If you have studied quantum mechanics, you would know that the generators $\mathcal{J}$ of rotation studied here are related to angular momentum operators. You would also know that in quantum mechanics, observables are represented by hermitean operators. However, in our discussion, the $\mathcal{J}$s come out naturally as antisymmetric matrices and are thus antihermitean. To make them hermitean, we multiply them by some multiples of i.

If you have not studied quantum mechanics, then the preceding would sound like gibberish to you, but do not worry. Simply take the attitude that, hey, it is a free country, and we can always invite ourselves to define a new set of physical quantities by multiplying an existing set of physical quantities by some constant. Heck, we could multiply by $\sqrt{17i}$ if we want.

Even though here we are nowhere near quantum mechanics, we will bow to customary usage and define $J_x \equiv -i\mathcal{J}_x$ and so forth. From (14) we see that, for example, J_z acting on the column vector (x, y, z) gives $i(y, -x, 0)$. Thus, instead of using matrices, we could also represent J_z by $i(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y})$, since $J_z x = i(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y})x = iy$, $J_z y = i(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y})y = -ix$, and $J_z y = i(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y})z = 0$. Note that J_z is precisely the z-component of the angular momentum operators in quantum mechanics. We can naturally pass back and forth between matrices and differential operators. We will not make use of this differential representation until a later chapter.

* Regarding the argument (which I mentioned in a footnote in the preface) between those who live with coordinates and those who live coordinate free, I would say that the proof of angular momentum conservation, which I already gave, not once, but twice in the two preceding chapters using coordinates, provides an example in favor of the latter group: $\frac{d}{dt}\vec{l} = \frac{d}{dt}(\vec{r} \times \vec{p}) = m\frac{d}{dt}(\vec{r} \times \frac{d\vec{r}}{dt}) = m\frac{d\vec{r}}{dt} \times \frac{d\vec{r}}{dt} + m\vec{r} \times \frac{d^2\vec{r}}{dt^2} = 0$ for rotationally symmetric potentials. While this indeed looks simpler than the two previous discussions, the former group could also say that this requires learning "considerable formal math," such as the cross product and its various properties.

Appendix 2: Rotations in higher dimensional space

Here we discuss rotations in D-dimensional Euclidean space. As you have no doubt heard, Einstein combined space and time into a 4-dimensional spacetime. Thus, what you will learn here about $SO(4)$ will be put to good use.* If you prefer, you could skip this discussion and come back to it later.

Start with a D-by-D matrix with 0 everywhere. Generalize (14). Stick a 1 into the mth row and nth column, and a (-1) into the nth row and mth column. Call this matrix $J_{(mn)}$. We put the subscripts (mn) in parentheses to emphasize that (mn) labels the matrix. They are not indices to tell us which element of the matrix we are talking about. As explained before, we define $J_{(mn)} = -i\mathcal{J}_{(mn)}$ so that explicitly

$$J^{ij}_{(mn)} = -i(\delta^{mi}\delta^{nj} - \delta^{mj}\delta^{ni}) \tag{19}$$

To repeat, in the symbol $J^{ij}_{(mn)}$, the indices i and j indicate respectively the row and column of the entry $J^{ij}_{(mn)}$ of the matrix $J_{(mn)}$, while the indices m and n, which I put in parentheses for pedagogical clarity, indicate which matrix we are talking about. The first index m on $J_{(mn)}$ can take on D values, and then the second index n can take on only $(D-1)$ values since, obviously, $J_{(mm)} = 0$. Also, since $J_{(nm)} = -J_{(mn)}$, we require $m > n$ to avoid double counting. Thus, there are only $\frac{1}{2}D(D-1)$ real antisymmetric D-by-D matrices $J_{(mn)}$, and A could be written as a linear combination of them: $A = i\sum_{m>n}\theta_{mn}J_{(mn)}$, where θ_{mn} denote $\frac{1}{2}D(D-1)$ real numbers. (As a check, for $D=2$ and 3, $\frac{1}{2}D(D-1)$ equals 1 and 3, respectively.) The matrices $J_{(mn)}$ are known as the generators of the group $SO(D)$.

Notice a notational peculiarity: for $SO(3)$, the Js could be labeled with one index rather than two indices. The reason is simple. In this case, the indices m, n take on 3 values, and so we could write $J_x = J_{23}$, $J_y = J_{31}$, and $J_z = J_{12}$. We will, as we do here, often pass freely between the index sets (123) and (xyz). In general, rotations are labeled by the plane they occur in, say the (m-n) plane spanned by the mth and nth axes. In 3-dimensional space, and only in 3-dimensional space, a plane is uniquely specified by the vector perpendicular to it. Thus, a rotation commonly spoken of as a rotation around the z-axis is better thought of as a rotation in the (1-2) plane, that is, the (x-y) plane. (In this connection, note that the $\mathcal{J}$ in (7) appears as the upper left 2-by-2 block in $\mathcal{J}_z$ in (14).) In contrast, for $SO(4)$ it makes no sense to speak of a rotation around, say, the third axis.

The reader who has studied some group theory knows that the essence of the group is captured by the extent to which the multiplication of two group elements does not commute. For rotations, everyday observations show that $R(\theta)R(\theta')$ is in general quite different from $R(\theta')R(\theta)$. See figure 3.

Following Lie, we could try to capture this essence by focusing on infinitesimal rotations. Let $R_1 \simeq I + A$ and $R_2 \simeq I + B$. Then $R_1R_2 \simeq (I+A)(I+B) \simeq I + A + B + AB + O(A^2, B^2)$ (where rather pedantically we have indicated that to the desired order if we keep AB, we should also keep terms of order $O(A^2, B^2)$, but we will see immediately that they are irrelevant). If we multiply in the other order, we simply interchange A and B, thus $R_2R_1 \simeq (I+A)(I+B) \simeq I + B + A + BA + O(A^2, B^2)$. Hence, R_1R_2 and R_2R_1 differ by the amount $[A, B] \equiv AB - BA$, a quantity known as the commutator between A and B.

More formally, given two matrices X and Y, to measure how they differ from each other, we could ask how $X^{-1}Y$ differs from the identity. If $X = Y$, then this product is equal to the identity. Now, the inverse of a matrix $I + A$ infinitesimally close to the identity is easy to determine: it is just $I - A$, since $(I-A)(I+A) = I + O(A^2)$. Thus, let us calculate $(R_2R_1)^{-1}R_1R_2$:

$$\begin{aligned}(R_2R_1)^{-1}R_1R_2 &= [I - (B + A + BA + O(A^2, B^2))][I + A + B + AB + O(A^2, B^2)] \\ &= I + [A, B] + \cdots\end{aligned} \tag{20}$$

For $SO(3)$, for example, A is a linear combination of the J_is, known as the generators of the Lie algebra. Thus, we could write $A = i\sum_i \theta_i J_i$ and similarly $B = i\sum_j \theta'_j J_j$. Hence $[A, B] = i^2\sum_{ij}\theta_i\theta'_j[J_i, J_j]$, and so it suffices to calculate the commutators $[J_i, J_j]$.

Recall that for two matrices M_1 and M_2, $(M_1M_2)^T = M_2^T M_1^T$. Transpose reverses the order. Thus $([J_i, J_j])^T = -[J_i, J_j]$. In other words, the commutator $[J_i, J_j]$ is itself an antisymmetric 3-by-3 matrix and thus could be written as a linear combination of the J_ks:

$$[J_i, J_j] = ic_{ijk}J_k \tag{21}$$

* Higher dimensional rotation groups often pop up in the most unlikely places in theoretical physics. For example, $SO(4)$ is relevant for a deeper understanding of the spectrum of the hydrogen atom.[5]

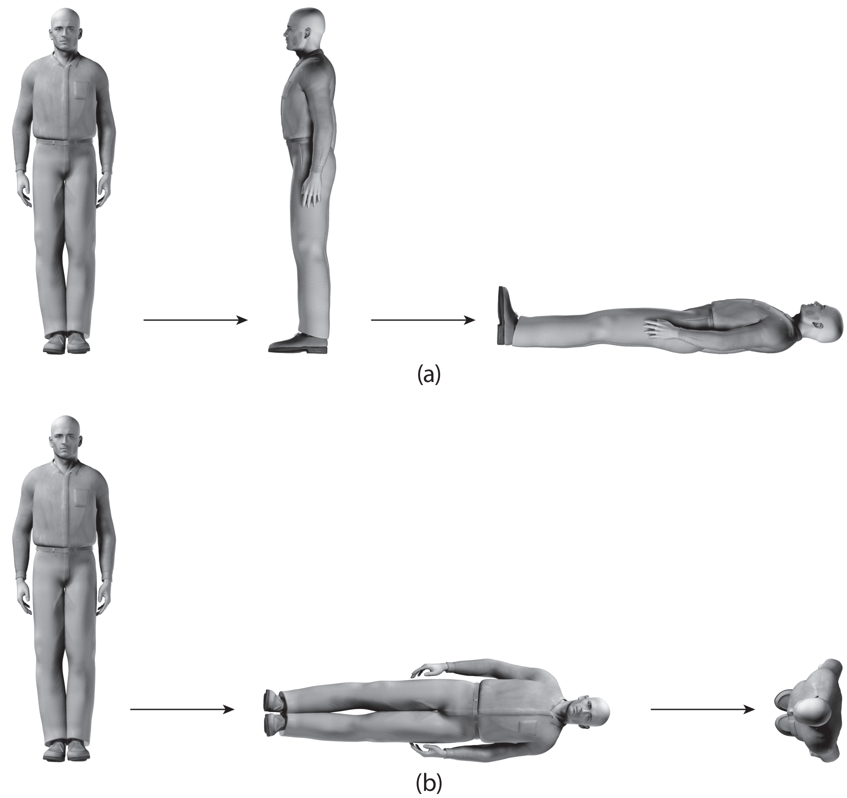


Figure 3 A marine recruit in a boot camp is standing and facing north. When the drill sergeant shouts, "Rotate by 90° eastward around the vertical axis" our recruit turns to face east. Suppose the sergeant next shouts, "Rotate by 90° westward around the north-south axis." Our recruit ends up lying down on his back with his head pointing west, his feet pointing east. But what would happen if the sergeant reverses his two commands? You could easily verify that our recruit now ends up lying down on his left elbow, with his head pointing north. The order matters. For this reason, the study of rotations has been a *bête noire* for generations of physics students.

for a set of real (convince yourself of this!) numbers c_{ijk}. The summation over k is implied by the repeated index summation convention.

By explicit computation using (14), we find

$$[J_x, J_y] = i J_z \tag{22}$$

You should work out the other commutators or argue by cyclic substitution $x \to y \to z \to x$. The three commutation relations may be summarized by

$$[J_i, J_j] = i\epsilon_{ijk} J_k \tag{23}$$

We define the totally antisymmetric symbol ϵ_{ijk} by saying that it changes sign upon the interchange of any pair of indices (and hence it vanishes when any two indices are equal) and by specifying that $\epsilon_{123} = 1$. In other words, we found that $c_{ijk} = \epsilon_{ijk}$.

Lie's great insight is that the preceding discussion holds for any group whose elements are labeled by a set of continuous parameters (such as θ_i, $i = 1, 2, 3$ in the case of $SO(3)$), groups now known as Lie groups. Expanding the group elements around the origin, we arrive at (20) and hence the structure (21) for any continuous group. The set of all commutation relations of the form (21) is said to define a Lie algebra, with c_{ijk} referred to as the structure constants of the algebra. The matrices J_i are called the generators of the Lie algebra. The idea is that by studying the Lie algebra, we go a long way toward understanding the group.

You should now work out (exercise 4), starting from (19), the Lie algebra for $SO(D)$:

$$[J_{(mn)}, J_{(pq)}] = i(\delta_{mp} J_{(nq)} + \delta_{nq} J_{(mp)} - \delta_{np} J_{(mq)} - \delta_{mq} J_{(np)}) \tag{24}$$

This may look rather involved to the uninitiated, but in fact it is quite simple. First, the right hand side, a linear combination of the Js, as required by the general argument above, is completely fixed by the first term by noting that the left hand side is antisymmetric under three separate interchanges: $m \leftrightarrow n$, $p \leftrightarrow q$, and $(mn) \leftrightarrow (pq)$. Next, all those Kronecker deltas just say that if the two sets (mn) and (pq) have no integer in common, then the commutator vanishes. If they do have an integer in common, you simply "cross off" that integer. This is best explained by using $SO(4)$ as an example. We have $[J_{(12)}, J_{(34)}] = 0$, $[J_{(12)}, J_{(14)}] = iJ_{(24)}$, $[J_{(23)}, J_{(31)}] = -iJ_{(21)} = iJ_{(12)}$, and so forth. The first of these relations says that rotations in the (1-2) plane and in the (3-4) plane commute, as you might expect. Do write down a few more and you will get it.

Exercises

1 Suppose we are given two vectors $\vec{p}$ and $\vec{q}$ in ordinary 3-dimensional space. Consider this array of three numbers: $\begin{pmatrix} p^2q^3 \\ p^3q^1 \\ p^1q^2 \end{pmatrix}$. Prove that it is not a vector, even though it looks like a vector. (Check how it transforms under rotation!) In contrast, $\begin{pmatrix} p^2q^3 - p^3q^2 \\ p^3q^1 - p^1q^3 \\ p^1q^2 - p^2q^1 \end{pmatrix}$ does transform like a vector. It is in fact the vector cross product $\vec{p} \times \vec{q}$.

2 Show that the product of two delta functions $\delta(x)\delta(y)$ is invariant under rotation around the origin.

3 Using (14) show that a rotation around the x-axis through angle θ_x is given by

$$R_x(\theta_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x \\ 0 & -\sin\theta_x & \cos\theta_x \end{pmatrix}$$

Write down $R_y(\theta_y)$. Show explicitly that $R_x(\theta_x)R_y(\theta_y) \neq R_y(\theta_y)R_x(\theta_x)$.

4 Calculate $[J_{(mn)}, J_{(pq)}]$.

5 Given a 3-vector $\vec{p}$, show that the quantity $\vec{p}^i \vec{p}^j$ when averaged over the direction of $\vec{p}$ is given by $\frac{1}{4\pi}\int d\theta d\varphi \cos\theta\, \vec{p}^i \vec{p}^j = \frac{1}{3}\vec{p}^2\delta^{ij}$.

Notes

1. Outside of physics, people often erroneously call any array of numbers a vector. Of course, people are free to call anything anything, so let's not quibble about the word "erroneously."
2. I say "most, but not all," because it is conceivable that you are a native speaker of Guugu Yimithirr. See G. Deutscher, *Through the Language Glass,* H. Holt and Co., 2010, p. 161.
3. The intellectual precision of our definition of symmetry is necessary lest we make the same mistake as the ancient Greeks. See *Fearful,* pp. 11–12 and figure 2.2.
4. According to one story, take it or leave it, Descartes was lying in bed when he noticed a fly buzzing around the room. He then realized that he could fix the fly's position given how far the fly was from two intersecting walls and the ceiling.
5. For example, J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics,* pp. 265–268.

I.4 Who Is Afraid of Tensors?

A tensor is something that transforms like a tensor

Long ago, an undergrad who later became a distinguished condensed matter physicist came to me after a class on group theory and asked me, "What exactly is a tensor?" I told him that a tensor is something that transforms like a tensor. When I ran into him many years later, he regaled me with the following story. At his graduation, his father, perhaps still smarting from the hefty sum he had paid to the prestigious private university his son attended, asked him what was the most memorable piece of knowledge he acquired during his four years in college. He replied, "A tensor is something that transforms like a tensor."

But this should not perplex us. A duck is something that quacks like a duck. Mathematical objects could also be defined by their behavior. We already saw in the preceding chapter that a vector is defined by how it transforms: $V'^i = R^{ij}V^j$. Consider a collection of "mathematical entities" T^{ij} with $i, j = 1, 2, \cdots, D$ in D-dimensional space. If they transform under rotations according to

$$T^{ij} \to T'^{ij} = R^{ik}R^{jl}T^{kl} \tag{1}$$

then we say that T transforms like a tensor, and hence is a tensor. (Here we are using the Einstein summation convention introduced in the previous chapter: The right hand side actually means $\sum_{k=1}^{D}\sum_{l=1}^{D} R^{ik}R^{jl}T^{kl}$ and is a sum of D^2 terms.) Indeed, we see that we are just generalizing the transformation law of a vector.

Fear of tensors

In my experience teaching, a couple of students are invariably confused by the notion of tensors. The very word "tensor" apparently make them tense. Dear reader, if you are not one of these unfortunates, so much the better for you! You could zip through this chapter. But to allay the nameless fear of the tensorphobe, I will go slow and be specific.

Think of the tensor T^{ij} as a collection of D^2 mathematical entities that transform into linear combinations of one another. To help the reader focus, I will often specialize to $D = 3$. Compounded and intertwined with their fear of tensors, the unfortunates mentioned above are also unaccountably afraid of indices, as mentioned in chapter I.1. For them, let us list T^{ij} explicitly for $D = 3$. There are $3^2 = 9$ of them: T^{11}, T^{12}, T^{13}, T^{21}, T^{22}, T^{23}, T^{31}, T^{32}, T^{33}. That's it, 9 objects that transform into linear combinations of one another. For example, (1) says that $T'^{21} = R^{2k}R^{1l}T^{kl} = R^{21}R^{11}T^{11} + R^{21}R^{12}T^{12} + R^{21}R^{13}T^{13} + R^{22}R^{11}T^{21} + R^{22}R^{12}T^{22} + R^{22}R^{13}T^{23} + R^{23}R^{11}T^{31} + R^{23}R^{12}T^{32} + R^{23}R^{13}T^{33}$. This shows explicitly, as if there were any doubt to begin with, that T'^{21} is given by a particular linear combination of the 9 objects. That's all: the tensor T^{ij} consists of 9 objects that transform into linear combinations of themselves under rotations.

We could generalize further and define* 3-indexed tensors, 4-indexed tensors, and so forth by such transformation laws as $W'^{ijn} = R^{ik}R^{jl}R^{nm}W^{klm}$. Here we will focus on 2-indexed tensors, and if we say tensor without any qualifier, we often, but not always, mean a 2-indexed tensor. With this definition, we might say that a vector is a 1-indexed tensor and a scalar is a 0-indexed tensor, but this usage is not common. A scalar transforms as a tensor with no index at all, namely $S' = S$; in other words, a scalar does not transform.

Tensor field

In the preceding chapter, we introduced the notion of a vector field $V^i(\vec{x})$, nothing more or less than a vector function of position. That it is a vector means that it transforms according to $V'^i(\vec{x}') = R^{ij}V^j(\vec{x})$. Now consider the derivative of this vector field $\frac{\partial V^j(\vec{x})}{\partial x^k}$, which we will call $W^{kj}(\vec{x})$.

Use the fact that $\vec{x}' = R\vec{x}$ implies $\vec{x} = R^{-1}\vec{x}' = R^T\vec{x}'$ and thus $\frac{\partial x^k}{\partial x'^h} = (R^T)^{kh} = R^{hk}$. (The O in the rotation group $SO(D)$ is crucial: the inverse of a rotation is its transpose.) Then

$$\frac{\partial}{\partial x'^h} = \frac{\partial x^k}{\partial x'^h}\frac{\partial}{\partial x^k} = R^{hk}\frac{\partial}{\partial x^k} \tag{2}$$

Thus

$$W'^{hi}(\vec{x}') \equiv \frac{\partial V'^i(\vec{x}')}{\partial x'^h} = R^{hk}\frac{\partial}{\partial x^k}(R^{ij}V^j(\vec{x})) = R^{hk}R^{ij}\frac{\partial V^j(\vec{x})}{\partial x^k} = R^{hk}R^{ij}W^{kj}(\vec{x}) \tag{3}$$

Comparing with (1) we see that $W^{kj}(\vec{x})$ transforms like a tensor and, hence, is a tensor. Indeed, it is a tensor field.

Notice that a tensor T^{ij} transforms as if it were composed of two vectors v^iw^j, that is, T^{ij} and v^iw^j transform in the same way. (Compare $v^iw^j \to v'^iw'^j = R^{ik}v^kR^{jl}w^l = R^{ik}R^{jl}v^kw^l$ with (1).) It is important to recognize that only in exceptional cases does a tensor T^{ij} happen to be equal to v^iw^j for some v and w. In general, a tensor cannot be

* Our friend the Jargon Guy tells us that the number of indices carried by a tensor is known as its rank. (The Jargon Guy is a new friend of the author; he did not appear in *QFT Nut*.)

written in the form $v^i w^j$. Our tensor field $W^{kj}(\vec{x})$ offers a ready example: in general, it is not equal to some vector U^k multiplied by $V^j(\vec{x})$.

Also, note in our example that the differential operator $\frac{\partial}{\partial x^k}$ transforms (2) like a vector. For example, if $\phi'(x') = \phi(x)$ transforms like a scalar, then $\frac{\partial \phi}{\partial x^k}$ transforms like a vector. Indeed, that's why you have encountered the notation $\vec{\nabla}$ for the gradient in an elementary physics course. This remark will be important later when we revisit Newton's inverse square law in chapter II.3. Do exercise 1 now.

Representation theory

Go back to the 9 objects T^{ij} that form a tensor. Mentally arrange them in a column

$$\begin{pmatrix} T^{11} \\ T^{12} \\ \vdots \\ T^{33} \end{pmatrix}$$

The linear transformation on the 9 objects can then be represented by a 9-by-9 matrix $\mathcal{D}(R)$ acting on this column. (Here we are going painfully slowly because of common confusion on this point. Some authors refer to this column as a 9-component "vector," which is a horrible abuse of terminology. We reserve the word "vector" for something that transforms like a vector $V'^i = R^{ij} V^j$. It is not true that any old collection of stuff arranged in a column is a vector. Don't call anything with feathers a duck!)

For every rotation, specified by a 3-by-3 matrix R, we could thus associate a 9-by-9 matrix $\mathcal{D}(R)$ transforming the 9 objects T^{ij} linearly among themselves. We say that the 9-by-9 matrix $\mathcal{D}(R)$ represents the rotation matrix R in the sense that

$$\mathcal{D}(R_1)\mathcal{D}(R_2) = \mathcal{D}(R_1 R_2) \tag{4}$$

Multiplication of $\mathcal{D}(R_1)$ and $\mathcal{D}(R_2)$ mirrors the multiplication of R_1 and R_2, as it were. The tensor T is said to furnish a 9-dimensional representation of the rotation group $SO(3)$. The 9-by-9 matrices $\mathcal{D}(R)$ represent R. Notice that with this jargon, the vector furnishes a 3-dimensional representation of the rotation group, known as the defining or fundamental representation.

Reducible versus irreducible

Let us now pose the central question of representation theory. Given these 9 entities T^{ij} that transform into each other, consider the 9 independent linear combinations that we can form out of them. Is there a subset among them that only transform into each other? A secret in-club, as it were.

A moment's thought reveals that there is indeed an in-club. Consider $A^{ij} \equiv T^{ij} - T^{ji}$. Under a rotation,

$$A^{ij} \to A'^{ij} = T'^{ij} - T'^{ji} = R^{ik}R^{jl}T^{kl} - R^{jk}R^{il}T^{kl}$$
$$= R^{ik}R^{jl}T^{kl} - R^{jl}R^{ik}T^{lk} = R^{ik}R^{jl}(T^{kl} - T^{lk}) = R^{ik}R^{jl}A^{kl} \quad (5)$$

I have again gone painfully slow here, but it is obvious, isn't it? We just verified in (5) that A^{ij} transforms like a tensor and is thus a tensor. Furthermore, this tensor changes sign upon interchange of its two indices ($A^{ij} = -A^{ji}$) and is said to be antisymmetric. The transformation law (1) treats the two indices democratically, without favoring one over the other, and thus preserves the antisymmetric character of a tensor: if $A^{ij} = -A^{ji}$, then $A'^{ij} = -A'^{ji}$ also.

Let us count. The index i in A^{ij} could take on D values; for each of these values, the index j could take on only $D-1$ values (since the D diagonal elements $A^{ii} = 0$ for $i = 1, 2, \cdots, D$, no Einstein repeated index summation here); but to avoid double counting (since $A^{ij} = -A^{ji}$) we should divide by 2. Hence, the number of independent components in A is equal to $\frac{1}{2}D(D-1)$. For example, for $D = 3$, we have the 3 objects: A^{12}, A^{23}, and A^{31}. The attentive reader would recall that we did the same counting in the previous chapter.

Obviously, the same goes for the symmetric combination $S^{ij} \equiv T^{ij} + T^{ji}$. You could verify as a trivial exercise that $S'^{ij} = R^{ik}R^{jl}S^{kl}$. A tensor S^{ij} that does not change sign upon interchange of its two indices ($S^{ij} = S^{ji}$) is said to be symmetric. Evidently, the symmetric tensor S has more components than the antisymmetric tensor A. In addition to the components S^{ij} with $i \neq j$, S also has D diagonal components, namely $S^{11}, S^{22}, \cdots, S^{DD}$. Thus, the number of independent components in S is equal to $\frac{1}{2}D(D-1) + D = \frac{1}{2}D(D+1)$.

For $D = 3$, the number of components in A and S are $\frac{1}{2} \cdot 3 \cdot 2 = 3$ and $\frac{1}{2} \cdot 3 \cdot 4 = 6$, respectively. (For $D = 4$, the number of components in A and S are 6 and 10, respectively.) Thus, in a suitable basis, the 9-by-9 matrix referred to above actually breaks up into a 3-by-3 block and a 6-by-6 block. We say that the 9-dimensional representation is reducible: it could be reduced to smaller representations.

But we are not done yet. The 6-dimensional representation is also reducible. To see this, note

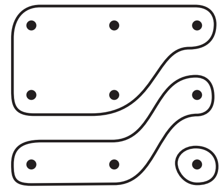
$$S'^{ii} = R^{ik}R^{il}S^{kl} = (R^T)^{ki}R^{il}S^{kl} = (R^{-1})^{ki}R^{il}S^{kl} = \delta^{kl}S^{kl} = S^{kk} \quad (6)$$

where we have used the O in $SO(D)$. (Here we are using repeated index summation: the indices i and k are both summed over.) In other words, the linear combination $S^{11} + S^{22} + \cdots + S^{DD}$, the trace of S, transforms into itself, that is, does not transform at all. It is a loner forming an in-club of one. The 6-by-6 matrix describing the linear transformation of the 6 objects S^{ij} breaks up into a 1-by-1 block and a 5-by-5 block. See figure 1.

Again, for the sake of the beginning student, let us work out explicitly the 5 objects that furnish the representation 5 of $SO(3)$. First define a traceless symmetric tensor $\tilde{S}$ by

$$\tilde{S}^{ij} = S^{ij} - \delta^{ij}(S^{kk}/D) \quad (7)$$

(The repeated index k is summed over.) Explicitly, $\tilde{S}^{ii} = S^{ii} - D(S^{kk}/D) = 0$, and $\tilde{S}$ is traceless. Specialize to $D = 3$. Now we have only 5 objects, namely $\tilde{S}^{11}$, $\tilde{S}^{22}$, $\tilde{S}^{12}$, $\tilde{S}^{13}$, $\tilde{S}^{23}$.



$9 \to 5 + 3 + 1$

Figure 1 How the collection of 9 objects T^{ij} splits up. The figure is meant to be schematic: the dots do not represent the original 9 objects, but linear combinations of them, and the positions of the dots are not meaningful.

We do not count $\tilde{S}^{33}$ separately, since it is equal to $-(\tilde{S}^{11}+\tilde{S}^{22})$. Under an $SO(3)$ rotation, these 5 objects transform into linear combinations of one another, as we just explained.

Let us be specific: the object $\tilde{S}^{13}$, for example, transforms into $\tilde{S}'^{13} = R^{1k}R^{3l}\tilde{S}^{kl} = R^{11}R^{31}\tilde{S}^{11} + R^{11}R^{32}\tilde{S}^{12} + R^{11}R^{33}\tilde{S}^{13} + R^{12}R^{31}\tilde{S}^{21} + R^{12}R^{32}\tilde{S}^{22} + R^{12}R^{33}\tilde{S}^{23} + R^{13}R^{31}\tilde{S}^{31} + R^{13}R^{32}\tilde{S}^{32} + R^{13}R^{33}\tilde{S}^{33} = (R^{11}R^{31} - R^{13}R^{33})\tilde{S}^{11} + (R^{11}R^{32} + R^{12}R^{31})\tilde{S}^{12} + (R^{11}R^{33} + R^{13}R^{31})\tilde{S}^{13} + (R^{12}R^{32} - R^{13}R^{33})\tilde{S}^{22} + (R^{12}R^{33} + R^{13}R^{32})\tilde{S}^{23}$, where in the last equality, we used $\tilde{S}^{ij} = \tilde{S}^{ji}$ and $\tilde{S}^{33} = -(\tilde{S}^{11}+\tilde{S}^{22})$. Indeed, $\tilde{S}^{13}$ transforms into a linear combination of $\tilde{S}^{11}$, $\tilde{S}^{22}$, $\tilde{S}^{12}$, $\tilde{S}^{13}$, $\tilde{S}^{23}$.

To summarize, what we found is that if, instead of the basis consisting of the 9 entities T^{ij}, we use the basis consisting of the 3 entities A^{ij}, the single entity S^{kk} (remember repeated index summation!), and the 5 entities $\tilde{S}^{ij}$, the 9-by-9 matrix $\mathcal{D}(R)$ (that represents rotation in the sense of (4)) breaks up into a 3-by-3 matrix, a 1-by-1 matrix, and a 5-by-5 matrix "stacked on top of each other." This is represented schematically as

$$\mathcal{D}(R) = (\text{9-by-9 matrix}) \to \left[\begin{array}{c|c|c} \text{(3-by-3 block)} & 0 & 0 \\ \hline 0 & \text{(1-by-1 block)} & 0 \\ \hline 0 & 0 & \text{(5-by-5 block)} \end{array}\right] \qquad (8)$$

Note that once we chose the new basis, this decomposition holds true for all rotations. (For the readers who know their linear algebra, the technical statement is that there exists a similarity transformation that block-diagonalizes $\mathcal{D}(R)$ for all R. Incidentally, we will encounter plenty of similarity transformations later.)

More generally, the D^2 representation furnished by a general 2-indexed tensor decomposes into a $\frac{1}{2}D(D-1)$-dimensional representation, a $(\frac{1}{2}D(D+1)-1)$-dimensional representation, and a 1-dimensional representation. We say that in $SO(3)$, $9 = 5 + 3 + 1$. (In $SO(4)$, $16 = 9 + 6 + 1$.)

You might have noticed that in this entire discussion we never had to write out R explicitly in terms of the 3 rotation angles and how the 5 objects $\tilde{S}^{11}, \cdots, \tilde{S}^{23}$ transform into one another in terms of these angles. It is only the counting that matters. You might regard that as the difference between mathematics and arithmetic.

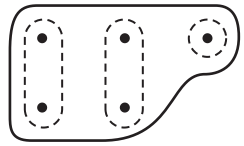


Figure 2 Under $SO(3)$, the 5 objects inside the solid line transform into linear combinations of each other, but under the smaller group of transformations $SO(2)$, the objects inside each of the 3 dashed lines transform into linear combinations of each other. The 5 breaks up as $5 \rightarrow 2+2+1$. As in figure 1, this figure is meant to be schematic.

Restriction to a subgroup

You definitely do not have to master group theory[1] to read this book, but it would be useful for you to learn a few basic concepts and to be able to count. For instance, the notion of a subgroup. Consider the group $SO(2)$ that we studied to exhaustion, consisting of rotations around the z-axis, say. Evidently, $SO(2)$ is a subgroup of $SO(3)$ in that its elements are all elements of $SO(3)$ and form a group all by themselves. The components of the 3-vector V^i could be split into two sets: (V^1, V^2) and V^3. Under a rotation around the z-axis, (V^1, V^2) transform as a 2-vector and V^3 as a scalar. We say that upon restriction to the subgroup $SO(2)$, the irreducible representation 3 breaks up into the representations 2 and 1 of the subgroup, a decomposition we write as $3 \rightarrow 2+1$. All the group theoretic results we need in this book could be obtained by explicit listing and simple counting.

Look at the 5 objects, S^{11}, S^{22}, S^{12}, S^{13}, S^{23}, that furnish the representation 5 of $SO(3)$. Now consider a restriction to the subgroup $SO(2)$. In other words, we restrict ourselves to rotations around the z-axis, that is, rotations under which $V^3 \rightarrow V'^3 = V^3$, namely rotations with $R^{33} = 1$ and R^{13}, R^{23}, R^{31}, R^{32} all vanishing. Since $SO(2)$ does not touch the index 3, we conclude immediately that the combination $S^{11} + S^{22} = -S^{33}$ does not transform, or in other words, it transforms as a singlet under $SO(2)$. Similarly, the pair (S^{13}, S^{23}) transforms as a doublet, since the index 3 is "invisible" to $SO(2)$: the group transforms the indices 1 and 2 into each other, while leaving the index 3 alone. Indeed, we see that our earlier expression for S'^{13} collapses to $S'^{13} = R^{11}S^{13} + R^{12}S^{23}$, as expected. Finally, you can verify that the remaining combinations $(S^{12}, S^{11} - S^{22})$ transform like a doublet. These results could be summarized by saying that, upon restriction to the subgroup $SO(2)$, the irreducible representation 5 of the group $SO(3)$ breaks up as $5 \rightarrow 2+2+1$. See figure 2.

Tensors in Newtonian mechanics

Let us give another example, particularly apt for a book on gravity, of a Newtonian tensor. Consider two nearby particles moving in a potential. Denote their trajectories by $\vec{x}(t)$

and $\vec{y}(t)$, respectively, determined by $\frac{d^2x^i}{dt^2} = -\partial^i V(\vec{x})$ and $\frac{d^2y^i}{dt^2} = -\partial^i V(\vec{y})$. (I am also testing whether there are any readers who do not understand thoroughly the concept of notational freedom.) We want to know how the separation vector $\vec{s} \equiv \vec{y} - \vec{x}$ changes with time, keeping terms to leading order in $\vec{s}$:

$$\frac{d^2s^i}{dt^2} = \frac{d^2y^i}{dt^2} - \frac{d^2x^i}{dt^2} = -\partial^i[V(\vec{y}) - V(\vec{x})] = -\partial^i[V(\vec{x}+\vec{s}) - V(\vec{x})] \simeq -\partial^i\partial^j V(\vec{x})s^j$$

The object $\mathcal{R}^{ij}(\vec{x}) \equiv \partial^i\partial^j V(\vec{x})$ is manifestly a tensor if $V(\vec{x})$ is a scalar. For example, verify that $\mathcal{R}^{ij} = GM(\delta^{ij}r^2 - 3x^ix^j)/r^5$ for the gravitational potential $V(\vec{x}) = -GM/r$. Note that $\mathcal{R}^{ij}$ is a symmetric traceless tensor. Since $\mathcal{R}^{ii} = \partial^i\partial^i V(\vec{x}) = \vec{\nabla}^2 V$, the tracelessness merely reaffirms the fact that the $1/r$ potential satisfies Laplace's equation $\vec{\nabla}^2 V = 0$. Also, $\mathcal{R}^{ij}$ is manifestly not the product of two vectors, but it transforms as if it were.

Let us see how rotational covariance works in the equation

$$\frac{d^2s^i}{dt^2} = -\mathcal{R}^{ij}s^j \tag{9}$$

The right hand side has to be linear in the vector $\vec{s}$. Since the left hand side transforms like a vector, the right hand side must also: indeed, it is given by a tensor $\mathcal{R}$ contracted* with a vector $\vec{s}$. A tensor is needed on the right hand side.

Imagine yourself falling toward a spherical planet or star. With no loss of generality, let your location at some instant be $(0, 0, r)$ along the z-axis. The tensor $\mathcal{R}$ written out as a matrix is then diagonal and is given by (for example, $\mathcal{R}^{33} = GM(\delta^{33}r^2 - 3x^3x^3)/r^5 = GM(1-3)/r^3$)

$$\mathcal{R} = \frac{GM}{r^3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{10}$$

Thus, the sign of $\frac{d^2\vec{s}}{dt^2}$ depends on the orientation of $\vec{s}$.

To see why this is so and to understand what tensors are all about, imagine surrounding yourself with a circular arrangement of balls lying in the $(x$-$z)$ plane (see figure 3a) and initially at rest in your frame. Using (9) and (10), we can now write down how the separation between two balls along different directions changes.

Since we are going to specify the direction, we will denote the separation simply by s. Along the z-axis, s grows according to (see (9)) $\frac{d^2s}{dt^2} = -\mathcal{R}^{33}s = +2\frac{GM}{r^3}s$. The plus sign indicates that the two balls move away from each other. In contrast, along the x-axis, s decreases according to $\frac{d^2s}{dt^2} = -\mathcal{R}^{11}s = -\frac{GM}{r^3}s$. The two balls approach each other. (Similarly for two balls aligned along the y-axis.) (Note that acting on $\vec{s}$ on the right hand side of (9) by a tensor makes it possible for $\frac{d^2s}{dt^2}$ to change sign depending on the orientation of $\vec{s}$.)

Inspecting figure 3a, you see why. Look at it as an observer on the planet. In the first case, one of the two balls, being closer to the planet, is falling faster than the other. Thus, they

* When a pair of repeated indices, such as j in (9), is summed over, they are often said to be contracted with each other (as mentioned in a footnote in the preceding chapter) in the sense that this index no longer appears in the result, as shown by the left hand side of (9).

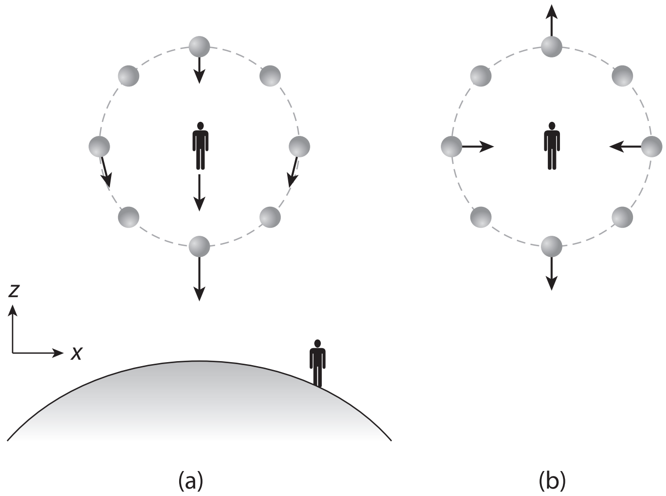


Figure 3 A falling ring of balls as seen by an observer on the planet (a), and as seen by an observer falling with the balls (b).

are moving away from each other. In the second case, the two balls are coming closer due to spherical symmetry: they are both heading toward the center of the planet. As Newton pointed out, objects do not fall down to earth, but toward the center of the earth.

In your rest frame (figure 3b) as you fall along with the balls, however, you see a tidal force acting on the circular ring (or a spherical shell if you prefer) of balls. The force appears to stretch the ring in the z-direction and to squeeze it in the orthogonal direction. When we come to Einstein's prediction of gravitational waves in chapter IX.4, we will see that gravitational waves act on the detector according to equations analogous to (9) and (10). Note also for future reference that the tidal force $\mathcal{R}^{ij}(\vec{x}) \equiv \partial^i \partial^j V(\vec{x})$ involves two derivatives acting on the gravitational potential $V(\vec{x})$.

Invariant tensors

In D-dimensional space, define the antisymmetric symbol $\varepsilon^{ijk\cdots n}$ carrying D indices to have the following properties:

$$\varepsilon^{\cdots l \cdots m \cdots} = -\varepsilon^{\cdots m \cdots l \cdots} \quad \text{and} \quad \varepsilon^{12\cdots D} = 1 \tag{11}$$

In other words, the antisymmetric symbol ε flips sign upon the interchange of any pair of indices. It follows that ε vanishes when two indices are equal. (Note that the second property listed is just normalization.) Since each index can take on only values $1, 2, \cdots, D$, the antisymmetric symbol for D-dimensional space must carry D indices as already noted. For example, for $D = 2$, $\varepsilon^{12} = -\varepsilon^{21} = 1$, with all other components vanishing. For $D = 3$, $\varepsilon^{123} = \varepsilon^{231} = \varepsilon^{312} = -\varepsilon^{213} = -\varepsilon^{132} = -\varepsilon^{321} = 1$, with all other components vanishing (as was already noted in the preceding chapter).

Using the Kronecker delta and the antisymmetric symbol, we can write the defining properties of rotations $R^T R = I$ and $\det R = 1$ as

$$\delta^{ij} R^{ik} R^{jl} = \delta^{kl} \tag{12}$$

and

$$\varepsilon^{ijk\cdots n} R^{ip} R^{jq} R^{kr} \cdots R^{ns} = \varepsilon^{pqr\cdots s} \det R = \varepsilon^{pqr\cdots s} \tag{13}$$

respectively. In (13) we used the definition of det R. (Verify this for $D = 2$ and 3.)

Referring to (1), we see that we can describe δ^{ij} and $\varepsilon^{ijk\cdots n}$ as invariant tensors: they transform into themselves. For the rest of this text, we will often use, implicitly or explicitly, the notion of invariant tensors.

For example, for $SO(3)$, using (13) you can show that $\varepsilon^{ijk} A^i B^j \equiv C^k$ defines a vector $\vec{C} = \vec{A} \times \vec{B}$, the familiar cross product. Various identities follow. Consider, for example,

$$\varepsilon^{ijk} \varepsilon^{lnk} = \delta^{il} \delta^{jn} - \delta^{in} \delta^{jl} \tag{14}$$

To prove this, simply note that both sides transform as invariant tensors with four indices, and the symmetry properties (such as under $i \leftrightarrow j$) of the two sides match. Contracting with A^j, B^l, and C^n, we obtain an identity you might recognize: $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$.

Closing of Newtonian orbits once again

We can now go back to the apparent mystery in chapter I.1, that the Newtonian orbits in a $1/r$ potential close. Out of the conserved angular momentum vector $\vec{l} = \vec{r} \times \vec{p} = \vec{r} \times \dot{\vec{r}}$ (we are using the notation of chapter I.1; we have effectively set the mass to unity and hence the second equality) we can form the Laplace-Runge-Lenz vector $\vec{\mathcal{L}} \equiv \vec{l} \times \dot{\vec{r}} + \frac{\kappa}{r}\vec{r}$. Computing the time derivative $\dot{\vec{\mathcal{L}}}$, you can verify (see exercise 4) that $\vec{\mathcal{L}}$ is conserved for an inverse square central force. When $\dot{\vec{r}}$ is perpendicular to $\vec{r}$, which occurs at perihelion and aphelion, the vector $\vec{\mathcal{L}}$ points in the direction of $\vec{r}$. We could take the constant vector $\vec{\mathcal{L}}$ to point toward the perihelion, and thus the position of the perihelion does not change. Hence the orbit closes.

This result does not hold in Einstein gravity. The precession of the perihelion of Mercury, which we will discuss in chapter VI.3, is of course one of the classic tests of general relativity.

Appendix: Two lemmas for future use

There is a lot more we could say about tensors, but let me mention two simple lemmas that we will happen to need later.

Let S^{ij} and A^{ij} be two arbitrary and unrelated tensors, symmetric and antisymmetric, respectively. Then $S^{ij} A^{ij} = 0$. (See exercise 5.)

Tensors can have all kinds of symmetry properties, which you can explore on your own and in the exercises. For example, a totally antisymmetric 3-indexed tensor T^{ijk} is such that T flips sign under the interchange of any pair of indices (for example, $T^{ijk} = -T^{jik} = +T^{jki}$). A multi-indexed tensor can also have symmetry properties under the interchange of a specific pair, or may have no symmetry at all. Consider, for example, a tensor G^{kij} symmetric under the interchange of the first pair of indices only, that is, $G^{kij} = G^{ikj}$. To be pedantic and absolutely clear, sometimes I like to put a space or a dot between the indices, thus $G^{ki\ j}$ or $G^{ki \cdot j}$ to separate the "special" pair from the other indices. For example, our tensor could happen to be $G^{ki\cdot j} = \partial^k \partial^i W^j(\vec{x})$ for some vector field W^j.

Given $G^{ki\cdot j}$, define $H^{k\cdot ij} \equiv G^{ki\cdot j} + G^{kj\cdot i}$. (Note that $H^{k\cdot ij} = H^{k\cdot ji}$ by definition, but $H^{i\cdot kj}$ is in general not equal to $H^{k\cdot ij}$.) Then we can solve for G in terms of H:

$$G^{ki\cdot j} = \frac{1}{2}(H^{k\cdot ij} + H^{i\cdot jk} - H^{j\cdot ki}) \tag{15}$$

(See exercise 8.)

Exercises

1 Define $\vec{\nabla} \equiv \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \cdots, \frac{\partial}{\partial x^D}\right)$. Show that if ϕ is a scalar, then $(\vec{\nabla}\phi)^2 = \vec{\nabla}\phi \cdot \vec{\nabla}\phi = \sum_k \left(\frac{\partial \phi}{\partial x^k}\right)^2$ and $\nabla^2 \phi$ transform like a scalar. The Laplacian is defined by

$$\nabla^2 = \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial (x^1)^2} + \frac{\partial^2}{\partial (x^2)^2} + \cdots + \frac{\partial^2}{\partial (x^D)^2}$$

2 Show that the symmetric tensor S^{ij} is indeed a tensor.

3 Show that the infinitesimal volume element d^3x is a scalar.

4 Show that the Laplace-Runge-Lenz vector is conserved.

5 Show that $S^{ij}A^{ij} = 0$ if S^{ij} is a symmetric tensor and A^{ij} an antisymmetric tensor.

6 Let T^{ijk} be a totally antisymmetric 3-indexed tensor. Show that T has $\frac{1}{3!}D(D-1)(D-2)$ components. Identify the one component for $D = 3$.

7 Consider for $SO(3)$ the tensor T^{ijk} from exercise 6. Show that it transforms as a scalar.

8 Prove the lemma in (15).

9 Verify (13) for $D = 2$ and 3.

Note

1. **For a concise introduction to some of the group theory needed in theoretical physics, see *QFT Nut,* appendix B.**

(***CHAPTER 1 CONTINUES...***)

Index

Page numbers followed by letters e, f, and n refer to exercises, figures, and notes, respectively.