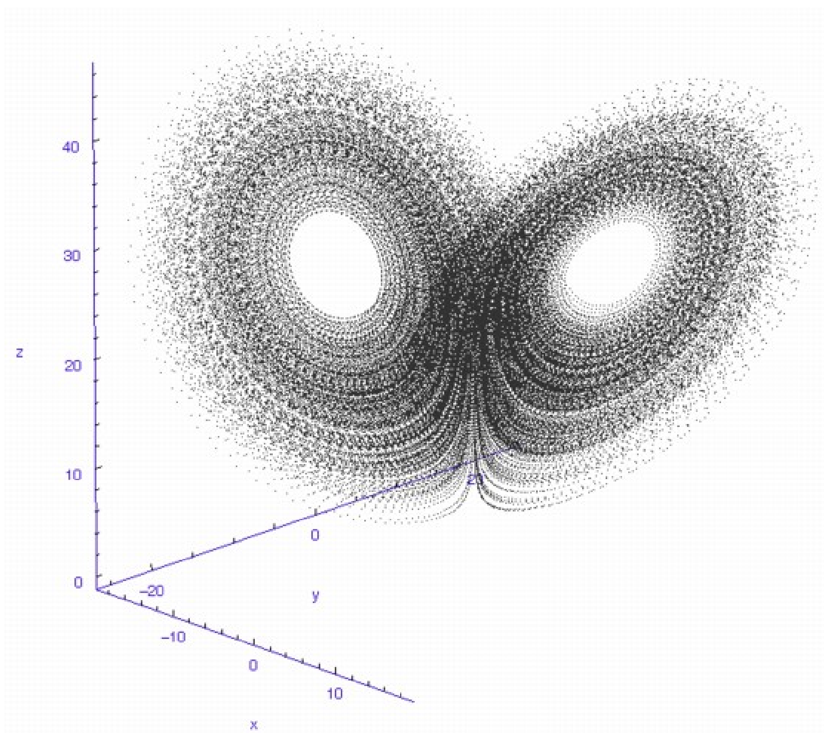
PHYS 363

Intermediate Classical Mechanics

Dr. Kevin Resch



Typeset by Aiden Diaz

Credit to Dr. Kevin Resch for His Lectures on PHYS 363

Preface

A majority of the information presented in this note package is transcribed directly from the *PHYS 363: Intermediate Classical Mechanics* lectures during the Spring 2025 term, taught by Dr. Kevin Resch.

These notes are by no means exhaustive nor are they guaranteed to be error-free, and I do not make any claims about the reliability or accuracy of these set of notes. They are intended primarily for personal use and for those who may benefit from an alternative presentation of the material.

Please use these notes at your own discretion.

Contents

List of Figures

1 Calculus of Variations

There is a specific application of calculus which pertains to the Lagrangian and Hamiltonian formalisms discussed in this course. This application is called the **calculus of variations**, and hinges on the idea of finding the extrema of **functionals**.

A functional is of the form

$$J = \int_{x_1}^{x_2} f(y, y', x)\, \mathrm{d}x\,, \quad \text{where } y = y(x), \;\; y' = y'(x) \equiv \mathrm{d}y(x)\, x. \tag{1}$$

Here, the arguments of (1) contain a dependent variable, $y(x)$, and a independent variable, x. Let us now consider a motivating example,

$$y(\alpha, x) = x + \alpha \sin(x)$$

for the free-varying parameter α, and independent variable, x. Notice, that this general function has all its unique counterparts, for $\alpha \in \mathbb{Z}$, converge at the values of $x = n\pi$.

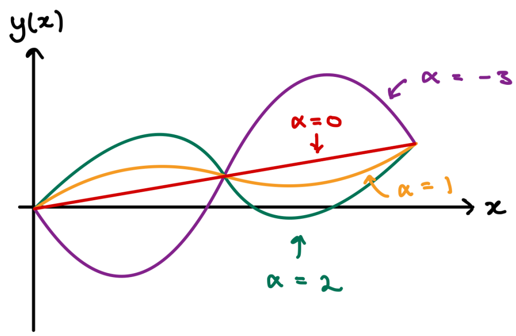


Figure 1: Function $y(\alpha, x)$ for the parameters $\alpha = -3, 0, 1, 2$, over a range $x \in [0, 2\pi]$.

Let us consider only the domain $x \in [0, 2\pi]$ (arbitrary), and determine the value of α for which its functional, $J(\alpha)$, is minimized. For this particular example, we will assume the form of $J(\alpha)$ to be

$$J(\alpha) = \int_0^{2\pi} (y')^2\, \mathrm{d}x\,.$$

Taking the derivative of $y(\alpha, x)$ with respect to x,

$$\frac{\partial y(\alpha, x)}{\partial x} = 1 + \alpha \cos(x)$$

so the functional becomes

$$J(\alpha) = \int_0^{2\pi} [1 + \alpha\cos(x)]^2\, \mathrm{d}x = \int_0^{2\pi} \left[1 + \cancel{2\alpha\cos(x)} + \alpha^2\cos^2(x)\right] \mathrm{d}x = \pi(2 + \alpha^2).$$

Since $J(\alpha)$ increases for any value, α, its minimum must be at the point $\alpha = 0$. Thus, $J(\alpha) = 2\pi$.

Here is an important thing to note. This motivating example only discusses a **particular** family of functions. However, the methods of Lagrangian and Hamiltonian mechanics are interested in the minima of the set of **all possible** functions passing through the endpoints of a boundary. For this reason, we must develop a set of functions which satisfy the demand that $J(\alpha)$ is minimized.

1.1 Euler's Equation

Here is Euler. For now, let us begin developing this method by supposing the general form of our functional, (1). In the most general case, we have that the endpoints are $y(x_1) = y_1$ and $y(x_2) = y_2$. Now, let us construct an arbitrary family of functions, such that

$$\begin{aligned} y(\alpha, x) &= y(0, x) + \alpha\eta(x) \\ \frac{\partial y(\alpha, x)}{\partial x} &= \frac{\partial y(0, x)}{\partial x} + \alpha\frac{\partial \eta(x)}{\partial x} \end{aligned} \tag{2}$$

where $y(0, x)$ is the **optimal path** and $\eta(x)$ is the **deviation**. Let us condense the partial derivatives such that $\partial y(\alpha, x)/\partial x = y'$ and $\partial\eta(x)/x = \eta'(x)$. We also must determine the boundary conditions of our problem. In particular, let us assume that the endpoints are fixed and that every function must pass through these endpoints, such that the deviation $\eta(x_1) = \eta(x_2) = 0$.

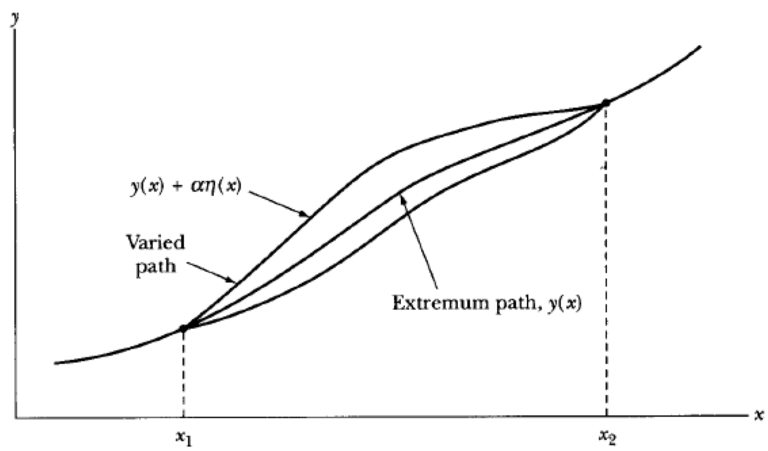


Figure 2: Varying paths of $y(\alpha, x)$. The point x_2 is associated with y_2, and x_1 associated with y_1.

Reconstructing our functional, we find that

$$J(\alpha) = \int_{x_1}^{x_2} f\big(y(\alpha, x), y'(\alpha, x), x\big)\,\mathrm{d}x\,.$$

Since we are interested in the minima of $J(\alpha)$, we take the partial in α, and by chain rule,

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\right]\mathrm{d}x = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial y'}\eta'(x)\right]\mathrm{d}x\,.$$

However, integration by parts[1] shows that

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial y'}\eta'(x)\,\mathrm{d}x = \cancel{\frac{\partial f}{\partial y'}\eta(x)\bigg|_{x_1}^{x_2}} - \int_{x_1}^{x_2} \left[\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right]\eta(x)\,\mathrm{d}x\,.$$

[1] Integration by parts states that

$$\int_b^a f\frac{\mathrm{d}g}{\mathrm{d}x}\,\mathrm{d}x = fg\bigg|_b^a - \int_b^a g\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x\,.$$

The first term is zero by the previously assumed boundary conditions, and since $\eta'(x)$ is simply $\mathrm{d}\eta(x)/\mathrm{d}x$, the functional becomes

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right]\eta(x)\,\mathrm{d}x\,.$$

Furthermore, we are interested in the extrema of $J(\alpha)$ at the point where $\alpha = 0$ for the path of minimum deviation. Then,

$$\left.\frac{\partial J(\alpha)}{\partial \alpha}\right|_{\alpha=0} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right]\eta(x)\,\mathrm{d}x = 0.$$

Since this is true only if the argument of the integral is zero, we arrive at the following condition

$$\boxed{\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} = 0} \tag{3}$$

which is known as **Euler's Equation**, as if we did not already have enough of him. Notice that we can also throw away the $\eta(x)$ term, since we initially assumed that it was an arbitrary deviation. Let us now see how we use this relation to find minimized functions.

Example 1.1 (Planar Surface)**.** Suppose we have a planar surface of arbitrary length and width. If the path is considered as $y(x)$, determine the form of $y(x)$ which minimizes the path length.

We begin by representing the length of the path as a functional,

$$J = \int_1^2 \mathrm{d}s = \int_1^2 \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2} = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = \int_{x_1}^{x_2} f(y, y', x)\,\mathrm{d}x$$

where we have implied that the argument inside J is $f(y, y', x) = \sqrt{1 + (\mathrm{d}y/\mathrm{d}x)^2}$.

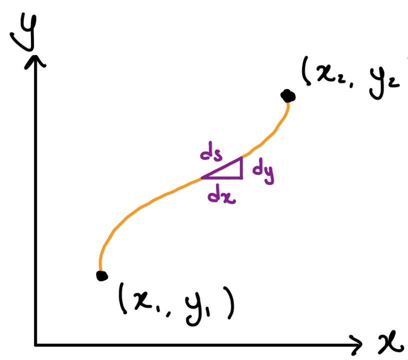


Figure 3: Arbitrary path, $y(x)$ (orange curve), with the differentials superimposed on a small segment of the curve.

Computing the partials, we see that

$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial y'} = \frac{1}{\cancel{2}}\frac{\cancel{2}y'}{\sqrt{1 + (y')^2}} = \frac{y'}{\sqrt{1 + (y')^2}}.$$

So, Euler's Equation says that

$$\frac{\partial f}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} \implies 0 = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{y'}{\sqrt{1+(y')^2}}\right] \implies \frac{y'}{\sqrt{1+(y')^2}} = \alpha.$$

Rearranging this result to be in terms of y', we have

$$(y')^2 = c^2\left[1+(y')^2\right] \implies y' = \sqrt{\frac{\alpha}{1-\alpha^2}} = \beta$$

where we have redefined $\sqrt{\alpha/(1-\alpha^2)} = \beta$ as a new constant. Thus, the function which minimizes the path on a planar surface is

$$y(x) = \beta x + \gamma.$$

Imposing boundary conditions such that the endpoints are fixed, we have that $y(x_2) = y_2$ and $y(x_1) = y_1$. Then,

$$\begin{aligned} y_1 &= \beta x_1 + \gamma \\ y_2 &= \beta x_2 + \gamma \end{aligned} \implies \beta = \frac{y_2 - y_1}{x_2 - x_1}.$$

The fully determined minimized path is then of the form

$$y(x) = \left[\frac{y_2-y_1}{x_2-x_1}\right]x + y_1 - \left[\frac{y_2-y_1}{x_2-x_1}\right]x_1 = \left[\frac{y_2-y_1}{x_2-x_1}\right](x - x_1) + y_1.$$

Notice that $y(x)$ is of the form $y = \beta(x-x_1) + y(x_1)$. As you would expect, the opposite is also true: $y(x) = \beta(x - x_2) + y_2$.

Typically, the general form of the minimizing path is all that we need. In fact, we will always find that the function which minimizes the path is its general form *plus* some arbitrary boundary conditions which distinguish it from other possible endpoints. For this reason, applying the boundary conditions is rather repetitive, and does not serve a useful purpose unless otherwise necessary. From now on, we will ignore these boundary conditions and keep in mind that *usually* we must apply them, instead giving only the general solution form.

Example 1.2 (Spherical Surface)**.** Suppose we have a path, $s = s(x,y,z)$, constrained to lie on the surface of a sphere of radius, R, such that two points in θ and ϕ travel $(\theta_1,\phi_1) \to (\theta_2,\phi_2)$. Determine the form of $s(x,y,z)$ which minimizes the path length.

We begin by writing the differential path length, $\mathrm{d}s$, in spherical coordinates as

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2} = R\sqrt{(\mathrm{d}\theta)^2 + \sin^2(\phi)(\mathrm{d}\phi)^2}.$$

Then, the functional, J, is of the form

$$J = \int_1^2 R\sqrt{(\mathrm{d}\theta)^2 + \sin^2(\phi)(\mathrm{d}\phi)^2}.$$

Notice we can manipulate this integral to be in terms of ϕ or in terms of θ. This is an important choice. For now, we will state what choice we have made and explain after. We choose the argument within J such that it is independent of θ.

$$R\sqrt{(\mathrm{d}\theta)^2+\sin^2(\phi)(\mathrm{d}\phi)^2} = R\sqrt{1+\sin^2(\theta)\left(\frac{\mathrm{d}\phi}{\mathrm{d}\theta}\right)^2} = R\sqrt{1+\sin^2(\theta)(\phi')^2} = f.$$

Then, taking the partial derivatives for Euler's Equation, we get

$$\frac{\partial f}{\partial \phi} = 0, \qquad \frac{\partial f}{\partial \phi'} = \frac{R}{\cancel{2}}\frac{\cancel{2}\sin^2(\theta)\phi'}{\sqrt{1+\sin^2(\theta)(\phi')^2}} = \frac{R\sin^2(\theta)}{\sqrt{1+\sin^2(\theta)(\phi')^2}}\phi'.$$

Euler's Equation then shows

$$\frac{\partial f}{\partial \phi} = \frac{\mathrm{d}}{\mathrm{d}\theta}\frac{\partial f}{\partial \phi'} \implies 0 = \frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{R\sin^2(\theta)}{\sqrt{1+\sin^2(\theta)(\phi')^2}}\phi'\right] \implies \frac{\sin^2(\theta)}{\sqrt{1+\sin^2(\theta)(\phi')^2}}\phi' = \alpha$$

for some constant, α, and where we have divided through by R, since it is also a constant. We can rearrange this result for $\phi' = \mathrm{d}\phi/\mathrm{d}\theta$ to find ϕ,

$$\frac{\mathrm{d}\phi}{\mathrm{d}\theta} = \pm\frac{\alpha}{\sin^2(\theta)\sqrt{1-\alpha^2/\sin^2(\theta)}} \implies \phi = \pm\int \frac{\alpha}{\sin^2(\theta)\sqrt{1-\alpha^2/\sin^2(\theta)}}\,\mathrm{d}\theta$$

which is an incomplete elliptic integral of the third kind. Specifically, ϕ has the solution

$$\phi = \pm\arcsin\left[\frac{\cot(\theta)}{\beta}\right] + \gamma$$

where β and γ are newly defined constants, and $\alpha = -1/(1+\beta^2)$. With further manipulations, we can rewrite ϕ as $\beta\sin(\phi-\gamma) = \pm\cot(\theta)$. With the angle-difference trigonometric identity, we finally arrive at

$$[-\beta\sin(\gamma)]R\sin(\theta)\cos(\phi) + [\beta\cos(\gamma)]R\sin(\theta)\sin(\phi) \pm R\cos(\theta) = 0.$$

Notice, $x = R\sin(\theta)\cos(\phi)$, $y = R\sin(\theta)\sin(\phi)$, and $z = \pm R\cos(\theta)$, this new representation is effectively of the form

$$Ax + By + Cz = 0$$

for some constants, A, B, and C, which is the equation of a plane passing through the origin. Physically, the **geodesics** (path of shortest distance) follow the intersection of the plane with the sphere of radius, R.

In Ex. 1.2, we saw that we could manipulate the integral in J to be either in terms of ϕ or in terms of θ. This distinction is important, because it allows us to simplify calculations significantly. In fact, let us compare the two. We define f_1 as the argument independent of ϕ, while f_2 as the argument independent of θ. Then,

$$f_1 = R\sqrt{\left(\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\right)^2 + \sin^2(\theta)}$$

$$f_2 = R\sqrt{1+\sin^2(\theta)\left(\frac{\mathrm{d}\phi}{\mathrm{d}\theta}\right)^2}.$$

Comparing the two, we can see that f_1 depends on both θ and θ', so Euler's Equation would have given us a result that isn't simply zero as we had before. f_2, on the other hand, depends only on ϕ', and simplifies nicely as we saw. The form of the solution doesn't change, as you would expect, but one is exceedingly more tedious than the other.

A slightly important note is that Beltrami's identity, which we will see soon, simplifies to the f_1 case; in fact, we should be mindful about how we are simplifying J, since extra work may involve extra mistakes.

1.1.1 Beltrami's Identity

Until now, we have solved the minimizing functions analytically, and although this works 100% of the time, it can become exceedingly difficult for more complicated paths. A trick which simplifies this slightly makes use of **Beltrami's identity**. Let us derive the identity itself.

Suppose we have the typical functional, J, with an argument $f = f(y, y', x)$. We will consider a special case, where f is independent of x, such that $f(y, y', x) = f(y, y')$. Furthermore, $y' \equiv \mathrm{d}y/\mathrm{d}x$, and so our functional becomes

$$J = \int_{x_1}^{x_2} f(y, y')\,\mathrm{d}x\,.$$

Let us now consider a quantity of the form

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f - y'\frac{\partial f}{\partial y'}\right]$$

and see if we can simplify this result to something we can work with. By chain rule (three times),

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[f - \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right] &= \frac{\mathrm{d}f}{\mathrm{d}x} - \left[\frac{\mathrm{d}y'}{\mathrm{d}x}\right]\frac{\partial f}{\partial y'} - y'\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} \\
&= \left[\frac{\partial f}{\partial y}y' + \cancel{\frac{\partial f}{\partial y'}y''} + \cancel{\frac{\partial f}{\partial x}}\right] - \cancel{y''\frac{\partial f}{\partial y'}} - y'\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} \\
&= y'\left[\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right]
\end{aligned}$$

where $\partial f/\partial x = 0$ since f is independent of x, and two of the terms cancel. Notice that the simplified result is in terms of Euler's Equation, and since we have imposed that it must be zero, we find that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f - \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right] = 0 \implies \boxed{f - \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\partial f}{\partial y'} = C} \tag{4}$$

for some constant, C, if $f = f(y, y')$.

This new representation of $f(y, y')$ is useful in certain calculus of variation problems. However, much of its significance comes later in the course, which involve conservation laws and Hamiltonian mechanics.

Example 1.3 (The Brachistochrone)**.** The Brachistochrone curve, derived from Ancient Greek's *brákhistos khrónos*, is the unique curve which pertains to the path of shortest time. It predates Euler's Equation, and we will derive its form here.

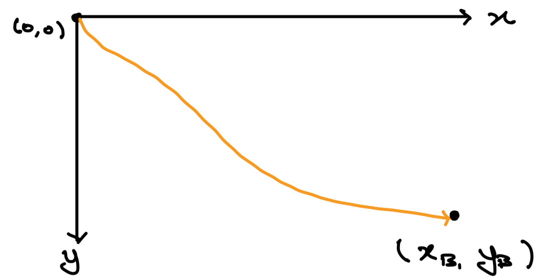


Figure 4: Arbitrary path, $y(x)$, that starts at the origin, $(x, y) = (0, 0)$, and ends at some point $(x, y) = (x_B, y_B)$.

Suppose a particle can slide frictionlessly on a track under the influence of gravity—that is, $|\mathbf{F}| = mg$—starting from rest. Determine the path, $y(x)$, which minimizes the time taken to travel from point $A \to B$. We begin by writing the time taken during travel as an integral functional with bounds $s \in [A, B]$.

$$t = \int_A^B \frac{\mathrm{d}s}{v}.$$

Since the track is frictionless, conservation of energy shows that

$$E = T + V = \frac{1}{2}mv^2 - mgy = 0 \implies v = \sqrt{\frac{2\cancel{m}gy}{\cancel{m}}} = \sqrt{2gy} \implies v = \sqrt{2gy}$$

where $y = -y$ since the particle travels downward. Since the differential path length, $\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}$, we can rearrange for $\mathrm{d}x$ (or $\mathrm{d}y$), and find that t is of the form

$$t = \int_A^B \frac{\sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}}{\sqrt{2gy}} = \int_0^{x_B} \sqrt{\frac{1 + (y')^2}{2gy}}\,\mathrm{d}x$$

where $y \equiv \mathrm{d}y \,/\, \mathrm{d}x$ as before. Notice that the argument of t is of the form $f(y, y')$, so Beltrami's identity says that

$$f - y'\frac{\partial f}{\partial y'} = \alpha \implies \sqrt{\frac{1 + (y')^2}{2gy}} - y'\left[\frac{1}{\sqrt{2gy}}\right]\left[\frac{1}{\cancel{2}}\frac{\cancel{2}y'}{\sqrt{1 + (y')^2}}\right] = \alpha$$

$$\frac{1 + \cancel{(y')^2 - (y')^2}}{\sqrt{2gy[1 + (y')^2]}} = \alpha \implies 1 + (y')^2 = \frac{1}{\alpha^2(2gy)}.$$

Now, let us introduce a new parameter, k, such that $k^2 = 1/(2g\alpha^2)$, which shows that

$$1 + (y')^2 = \frac{k^2}{y}, \quad k \equiv \frac{1}{2g\alpha^2}.$$

This differential equation is not the easiest to solve since it is nonlinear, so we will state the solution, which is typically expressed in parametric form.

$$x = \frac{k^2}{2}[\theta - \sin(\theta)] + C$$
$$y = \frac{k^2}{2}[1 - \cos(\theta)]$$

where θ is a free-varying parameter, C is an arbitrary constant due to integration, and k is the substitution we applied before. The form of this solution is known as the **cycloid** function, and is expressed as a parametric equation since there is no elementary inverse for x. Since we have just given the solution, let us at least confirm that it is correct.

First, we have $y' \equiv \mathrm{d}y \,/\, \mathrm{d}x$, and we can re-express this as

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y \,/\, \mathrm{d}\theta}{\mathrm{d}x \,/\, \mathrm{d}\theta} = \frac{\cancel{k^2} \sin(\theta)/\cancel{2}}{\cancel{k^2}[1-\cos(\theta)]/\cancel{2}} = \frac{\sin(\theta)}{1-\cos(\theta)}$$

then,

$$1 + (y')^2 = 1 + \frac{\sin^2(\theta)}{[1-\cos(\theta)]^2} = \frac{[1-\cos(\theta)]^2 + \sin^2(\theta)}{[1-\cos(\theta)]^2} = \frac{2}{1-\cos(\theta)}.$$

We can rearrange the parametric equation in y to give

$$\frac{k^2}{y} = \frac{\cancel{k^2}}{\cancel{k^2}[1-\cos(\theta)]/2} = \frac{2}{1-\cos(\theta)}$$

which agree, showing that the solution is indeed correct.

There is actually a lot more significance to the Brachistochrone that is not in Ex. (1.3) that will we discuss here. The Brachistochrone curve is actually a consequence of a more general function, called the cycloid function. In fact, the curve is generated by the path that a circle of radius $r = k^2/2$ traces for the angle revolution $\theta \in [0, \pi]$.

We can calculate the time to each point, but as it is slightly unnecessary, the result is just given and is

$$t = \frac{k}{\sqrt{2g}}\theta_B \implies \theta_B = \frac{\sqrt{2g}}{k}t = \omega t$$

where θ_B is the upper bound of a change of variable applied to the original time integral functional. We can see that a particle lies on the trajectory of the cycloid only if it rotates in the special case $\omega = \sqrt{2g}/k$, where ω is the angular velocity. At the lowest point of the curve, $\theta = \pi$, the time taken is

$$t = \frac{\pi k}{\sqrt{2g}}$$

which is as important as the Brachistochrone itself, since it corresponds to its sister curve, the Tautochrone.

Example 1.4 (The Tautochrone)**.** Consider the same setup as in Ex. (1.3). Determine the path, $y(x)$, such that the amount of time the slide to a fixed point is constant no matter the initial position on the curve.

As before, our integral functional is

$$t = \int_{x_A}^{x_B} \frac{\mathrm{d}s}{v} = \int_{x_A}^{x_B} \sqrt{\frac{1+(y')^2}{2g(y-y_A)}}\,\mathrm{d}x$$

where we have defined a point y_A to be where the particle initially sits. We can parametrize this curve (this is the method to find the the time at each point in the Brachistochrone), such that

$$t = \frac{1}{\sqrt{2g}}\frac{k}{\sqrt{2}}\int_{\theta_A}^{\pi}\sqrt{\frac{1+[\cot^2(\theta/2)]}{1-\cos-[1-\cos(\theta_A)]}}[1-\cos(\theta)]\,\mathrm{d}\theta = \frac{\pi k}{\sqrt{2g}}$$

As before, this integral is not easy to solve, so we have stated the solution. However, notice that this result matches exactly with that of the Brachistochrone. In fact, the Tautochrone is just a special case of the Brachistrochrone for the angle $\theta \in [0,\pi]$.

1.2 Multiple Degrees of Freedom

Let us now consider the case where we want to determine the form of multiple functional maximizing functions. In this case, let us suppose that we have the functions $y_1(x), y_2(x), \cdots, y_n(x)$, such that our integral functional is

$$J(\alpha) = \int_{x_1}^{x_2} f(y_1, y_1', y_2, y_2', \cdots, y_n, y_n', x)\,\mathrm{d}x\,.$$

As before, our family of functions become

$$\begin{aligned} y_1(\alpha,x) &= y_1(0,x) + \alpha\eta_1(x) \\ y_2(\alpha,x) &= y_2(0,x) + \alpha\eta_2(x) \\ &\vdots \\ y_n(\alpha,x) &= y_n(0,x) + \alpha\eta_n(x) \end{aligned}$$

for the free-varying parameter, α, and deviations $\eta_1(x), \eta_2(x), \cdots, \eta_n(x)$. We impose the same boundary conditions at the endpoints, such that $\eta_i(x_1) = \eta_i(x_2) = 0$, for $i \in \mathbb{Z}$. Then, taking the derivative of J with respect to α, we have

$$\begin{aligned} \frac{\partial J}{\partial \alpha} &= \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y_1}\frac{\partial y_1}{\partial \alpha} + \frac{\partial f}{\partial y_1'}\frac{\partial y_1'}{\partial \alpha} + \frac{\partial f}{\partial y_2}\frac{\partial y_2}{\partial \alpha} + \frac{\partial f}{\partial y_2'}\frac{\partial y_2'}{\partial \alpha} + \cdots\right]\mathrm{d}x \\ &= \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y_1}\eta_1(x) + \frac{\partial f}{\partial y_1'}\eta_1'(x) + \frac{\partial f}{\partial y_2}\eta_2(x) + \frac{\partial f}{\partial y_2'}\eta_2'(x) + \cdots\right]\mathrm{d}x \end{aligned}$$

where $\partial y_i/\partial\alpha = \eta_i(x)$ and $\partial y_i'(x)/\partial\alpha = \eta_i'(x)$. Applying integration by parts on every term, we see that

$$\cancel{\frac{\partial f}{\partial y_1}\eta_1(x)\bigg|_{x_1}^{x_2}} + \cancel{\frac{\partial f}{\partial y_2}\eta_2(x)\bigg|_{x_1}^{x_2}} + \cdots + \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y_1} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y_1'}\right]\eta_1(x)\,\mathrm{d}x + \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y_2} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y_2'}\right]\eta_2(x)\,\mathrm{d}x + \cdots$$

and so evaluating $\partial J/\partial\alpha$ at $\alpha = 0$ and demanding it be zero, we find

$$\boxed{\frac{\partial f}{\partial y_1} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y_1'} = 0, \qquad \frac{\partial f}{\partial y_2} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y_2'} = 0, \qquad \cdots} \tag{5}$$

where we have canceled $\eta_i(x)$ since they are already arbitrary. Notice that this result gives an Euler Equation for each dependent variable, as could have been expected. Furthermore, it is important to note that this derivation assumes that each $y_i(x)$ was free to vary independently. This, however, may not always be the case if there are **constraints** imposed on the paths.

1.3 Constraints

In some cases, we are interested in extremizing an integral functional in the presence of constraints. One example of this is Ex. (1.2), which was a geodesic on a sphere. In particular, the particle was forced to lie on its surface, which we can express as

$$g(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0$$

where R is fixed the radius of the sphere. This constraint is an example of a **holonomic** constraint, which are generally written in the form independent of its argument's derivatives,

$$g(y_1, y_2, \cdots, y_n, x) = 0.$$

There are two methods of determining the form of our functions in the presence of constraints. Let us discuss these.

1.3.1 Holonomic Constraints via Coordinate System

The first way, and often the best way, to enforce a holonomic constraint is to build it in directly into the choice of our coordinates. In the geodesic on a sphere example, we can switch our coordinate system to be

$$x = R\sin(\theta)\cos(\phi)$$
$$y = R\sin(\theta)\sin(\phi)$$
$$z = R\cos(\theta)$$

where we have the parameters (θ, ϕ), and the constant R to reflect the radius of the sphere. This automatically satisfies our constraint. Alternatively, we my also wish to eliminate a variable. In which case,

$$z^2 = R^2 - x^2 - y^2$$

to which further analysis can be done. Some other examples are *rolling without slipping*, in which case the constraint is

$$g(x, \theta) = x - r\theta = 0$$

or a mass *sliding along a plane*, such that

$$g(x, y) = y - x\tan(\theta) = 0$$

or a *particle constrained to a rotating wire*,

$$g(\theta) = \theta - \omega t = 0.$$

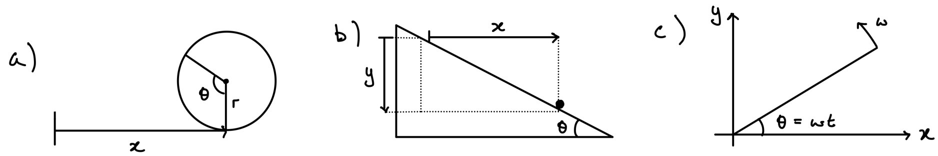


Figure 5: (a) Disk rolling without slipping. (b) Particle sliding along a plane. (c) Particle constrained to a rotating wire.

Notice that each holonomic constraint reduces the number of degrees of freedom[2] (DoF) by

[2] A degree of freedom of a system is an independent way a system can move.

one. In the case of a sphere, a constraint on the radius effectively takes the 3D system into a 2D surface, governed by the coordinates (θ, ϕ).

1.3.2 Holonomic Constraints via Lagrange Undetermined Multipliers

Let us consider a simpler case for multiple degrees of freedom. The integral functional we wish to extremize is

$$J(\alpha) = \int_{x_1}^{x_2} f(y, y', z, z', x)\,\mathrm{d}x$$

subject to the holonomic constraint

$$g(y, z, x) = 0.$$

As before, we consider the family of functions

$$\begin{aligned} y(\alpha, x) &= y(0, x) + \alpha\eta_1(x) \\ z(\alpha, x) &= z(0, x) + \alpha\eta_2(x) \end{aligned}$$

and apply the same process in Chapter 1.2. We should arrive at

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2} \left[\left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right)\eta_1(x) + \left(\frac{\partial f}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial z'}\right)\eta_2(x)\right]\mathrm{d}x\,.$$

At this point, we cannot assume that the variations are independent. So, to ensure that the variations do not violate our constraint, we calculate the differential $\mathrm{d}g$ with respect to α.

$$\mathrm{d}g = \left(\frac{\partial g}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial \alpha} + \frac{\partial g}{\partial x}\cancel{\frac{\partial x}{\partial \alpha}}\right)\mathrm{d}\alpha = \left(\frac{\partial g}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial \alpha}\right)\mathrm{d}\alpha\,.$$

where $\partial x/\partial\alpha = 0$. For the constraint to remain satisfied, we must have that the differential is zero, since any change in $\mathrm{d}g$ would imply that g would also change. Thus,

$$\mathrm{d}g = \left(\frac{\partial g}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial \alpha}\right)\mathrm{d}\alpha = 0$$

which is only satisfied if the argument is zero, so

$$\frac{\partial g}{\partial y}\eta_1(x) = -\frac{\partial g}{\partial z}\eta_2(x), \quad \text{where } \frac{\partial y}{\partial \alpha} = \eta_1(x), \;\; \frac{\partial z}{\partial \alpha} = \eta_2(x).$$

Isolating for $\eta_2(x)$ so that we may rewrite $\partial J/\partial\alpha$ in a simpler form,

$$\eta_2(x) = -\frac{\partial g/\partial y}{\partial g/\partial z}\eta_1(x).$$

Thus, evaluating $\partial J/\partial\alpha$ at $\alpha = 0$ and imposing it to be zero,

$$\begin{aligned} \frac{\partial J}{\partial \alpha} &= \int_{x_1}^{x_2} \left[\left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right)\eta_1(x) + \left(\frac{\partial f}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial z'}\right)\eta_2(x)\right]\mathrm{d}x \\ &= \int_{x_1}^{x_2} \left[\left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right) - \left(\frac{\partial f}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial z'}\right)\frac{\partial g/\partial y}{\partial g/\partial z}\right]\eta_1(x)\,\mathrm{d}x = 0. \end{aligned}$$

Then,

$$0 = \left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right) - \left(\frac{\partial f}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial z'}\right)\frac{\partial g/\partial y}{\partial g/\partial z}$$
$$\left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right)\frac{1}{\partial g/\partial y} = \left(\frac{\partial f}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial z'}\right)\frac{1}{\partial g/\partial z}.$$

Notice that both sides of the equation are functions of x. So, we denote this function as $-\lambda(x)$ and substitute it back into $\partial J/\partial\alpha$, giving the system of equations

$$\begin{aligned} \frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} + \lambda(x)\frac{\partial g}{\partial y} &= 0 \\ \frac{\partial f}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial z'} + \lambda(x)\frac{\partial g}{\partial z} &= 0, \end{aligned} \qquad g(y,z,x) = 0$$

which is a system with three unknowns, $y(x)$, $z(x)$, and $\lambda(x)$. For several independent variables, we can extend this system equations similarly. Let us denote $y_i(x)$ for $i = 1, 2, \cdots, n$, to be our independent variables, and our holonomic constraints to be $g_j(\{y_j\}, z)$ for $j = 1, 2, \cdots m$. The system of equations is represented by

$$\boxed{\frac{\partial f}{\partial y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y_i'} + \sum_{j=1}^{m}\lambda_j(x)\frac{\partial g_j}{\partial y_i} = 0, \quad g_j(\{y_j\}, x) = 0 + \text{Boundary Conditions.}} \tag{6}$$

Here, we distinguish the difference between the constraint functions $g_j(\{y_i\}, x)$ and the actual functions $y_i(x)$ with the indices j and i. Furthermore, notice that there are n equations for each $y_i(x)$. This observation is related to the difference between Lagrange's equations and Hamilton's equations, which we will see eventually.

There is also an alternate way to get to the same expression, where we suppose we have an expanded function, $h(x)$, such that

$$h(x) = f(x) + \sum_{j=1}^{m}\lambda_j(x)g_j.$$

We will not go through this derivation since we arrive at the same result, but in principle all you would need to do is apply Euler's Equation, (3), and cancel terms that do not depend on each other to arrive back at (6).

Of course, there are more other general types of constraints; however, these are largely beyond the scope of the course. For example, we may have a particle's path which may be forbidden from a region $y(x) \geq 0$, or may have a maximum slope.

Example 1.5 (Geodesic on a Sphere)**.** Let us revisit Ex. (1.2) and set up the solution using Lagrange undetermined multipliers (LUM). Recall, the function are we interested in is

$$J = \int_1^2 \mathrm{d}s = \int_{x_1}^{x_2} \sqrt{1 + (y')^2 + (z')^2}\,\mathrm{d}x$$

with the constraint

$$g(y,z,x) = x^2 + y^2 + z^2 - R^2 = 0.$$

By LUM, we have two Euler equations of the form

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} + \lambda(x)\frac{\partial g}{\partial y} = 0$$
$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial y}{\partial z'} + \lambda(x)\frac{\partial g}{\partial z} = 0.$$

Computing the partials, we find that

$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2 + (z')^2}}, \quad \frac{\partial g}{\partial y} = 2y, \quad \frac{\partial g}{\partial z} = 2z$$

and so the Euler equations are

$$-\frac{\mathrm{d}}{\mathrm{d}x}\frac{y'}{\sqrt{1 + (y')^2 + (z')^2}} + \lambda(x)2y = 0$$
$$-\frac{\mathrm{d}}{\mathrm{d}x}\frac{y'}{\sqrt{1 + (y')^2 + (z')^2}} + \lambda(x)2z = 0.$$

The method of LUM is rather lengthy, but boils down to a simple algorithm:

(1) Compute the derivatives of the constraint.

(2) Compute the partials of the functional within the integral.

(3) Obtain the relationship between the coordinates.

1.3.3 Integral Constraints

Another type of constraint that we may encounter is one that is expressed as an integral. To see how this affects our system, let us consider the functional $J = J(y, y', x)$, such that

$$J = \int_{x_1}^{x_2} f(y, y', x)\,\mathrm{d}x$$

subject to the constraint

$$I = \int_{x_1}^{x_2} g(y, y', x)\,\mathrm{d}x - L = 0$$

where L is an arbitrary constant that depends on the system. There is, of course, boundary conditions that are involved in our systems, and this together with the calculus of variation problem is called an **isoperimetric problem**.[3]

Here is an important issue: if we attempt to vary y according to the family of functions

$$y(\alpha, x) = y(0, x) + \alpha\eta(x)$$

we may violate our integral constraint. One method of avoiding this is by considering two variations instead, where $\eta(x)$ is the usual variation in the path, but $\xi(x)$ is a variation which keeps the path within the constraint. We express this as

$$y(\alpha, \beta, x) = y(0, x) + \alpha\eta(x) + \beta\xi(x).$$

[3] A simple example of such a problem, is attempting to maximize the area enclosed by a fence of length, L.

Now, the integral functional can be considered a function of two variables, $J(\alpha, \beta)$ and $I(\alpha, \beta)$, and is maximized exactly when their gradients are parallel.

$$\vec{\nabla}_{\alpha,\beta} J(\alpha, \beta) = -\lambda \vec{\nabla}_{\alpha,\beta} I.$$

We should be careful here. λ is **not** the same as $\lambda(x)$, despite that they are both called Lagrange multipliers. In fact, we refer to λ as a "Lagrange multiplier" while $\lambda(x)$ is our "Lagrange undetermined multiplier".

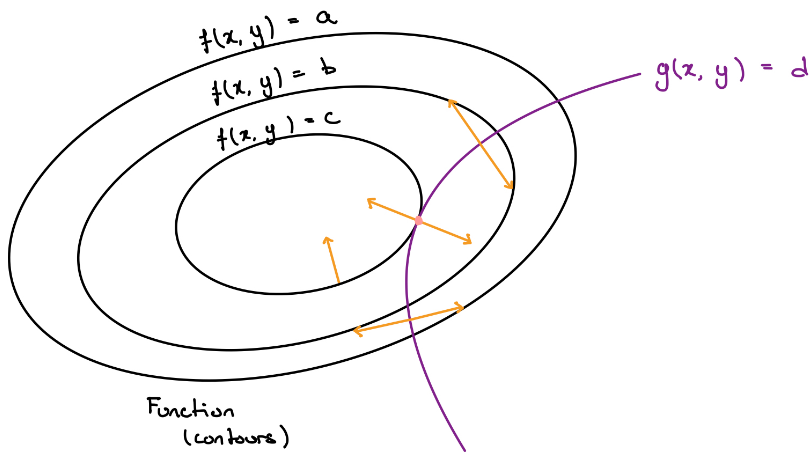


Figure 6: Contours functions (black), $f(x, y) = a, b, c$, with the constraint function (purple), $g(x, y) = d$, superimposed along c, with gradient lines (orange) placed along g. The pink point represents the maximum value of f subject to the constraint.

Taking the partials with respect to α and β, we have

$$\frac{\partial J}{\partial \alpha} = -\lambda \frac{\partial I}{\partial \alpha}$$
$$\frac{\partial J}{\partial \beta} = -\lambda \frac{\partial I}{\partial \beta}.$$

Now, to ensure that the gradient condition is satisfied, we must construct a *new*, expanded functional. We denote this as K, and is expressed as

$$\boxed{K(\alpha) = J(\alpha) + \lambda I(\alpha)} \tag{7}$$

As before, we extremize K at the point $\alpha = 0$, which follows the same process as that for Euler's equation, to which we should arrive at the form

$$\boxed{\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} + \lambda\left(\frac{\partial g}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial g}{\partial y'}\right) = 0, \quad I = \int_{x_1}^{x_2} g(y, y', x)\,\mathrm{d}x - L = 0} \tag{8}$$

with some associated boundary conditions due to the system of interest.

Example 1.6 (The Catenary)**.** Let us suppose we want to find the shape of a hanging chain of a fixed length. In particular, we want to minimize the energy supposing that the chain has a linear mass density and a small length mass element $\mathrm{d}m = \mu\,\mathrm{d}s$.

Our functional is of the form

$$J = \int_1^2 gy\,\mathrm{d}m = \int_1^2 \mu g y\sqrt{1+(y')^2}\,\mathrm{d}x\,.$$

Since μ and g are arbitrary constants, we can drop them and determine the form simply from

$$J = \int_1^2 y\sqrt{1+(y')^2}\,\mathrm{d}x = \int_1^2 f(y,y')\,\mathrm{d}x\,.$$

The integral constraint related to the problem is of the form

$$I = \int_{x_1}^{x_2} \mathrm{d}s - L = \int_{x_1}^{x_2} \sqrt{1+(y')^2}\,\mathrm{d}x - L = \int_{x_1}^{x_2} g(y')\,\mathrm{d}x - L = 0.$$

Now, defining the expanded functional, we have

$$K = J + \lambda I = \int_{x_1}^{x_2} \left[y\sqrt{1+(y')^2} + \lambda\sqrt{1+(y')^2}\right]\mathrm{d}x\,.$$

But, notice that K is also independent of x, so by Beltrami's identity, we have

$$\begin{aligned}
(f+\lambda g) - y'\frac{\partial(f+\lambda g)}{\partial y'} &= \alpha \\
(y+\lambda)\sqrt{1+(y')^2} - y'(y+\lambda)\frac{\cancel{2}y'}{\cancel{2}\sqrt{1+(y')^2}} &= \alpha \\
(y+\lambda)\left[\frac{1+\cancel{(y')^2-(y')^2}}{\sqrt{1+(y')^2}}\right] &= \alpha \\
\frac{y+\lambda}{\sqrt{1+(y')^2}} &= \alpha
\end{aligned}$$

for some constant α. Rearranging this, we find the differential equation

$$\alpha\frac{\mathrm{d}y}{\mathrm{d}x} = \pm\sqrt{(y+\lambda)^2-\alpha^2}$$

which has the solution

$$y = \alpha\cosh\left(\frac{x-\beta}{\alpha}\right) - \lambda.$$

We can determine the constants α, β, and λ from the initial conditions of the problem and the constraint. In particular, the constraint demands that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left[\alpha\cosh\left(\frac{x-\beta}{\alpha}\right) - \lambda\right] = \sinh\left(\frac{x-\beta}{\alpha}\right)$$

and so

$$1+(y')^2 = 1+\sinh^2\left(\frac{x-\beta}{x}\right) = \cosh^2\left(\frac{x-\beta}{\alpha}\right).$$

Thus, the length of the chain is

$$L = \int_{x_1}^{x_2} = \sqrt{1+(y')^2}\,\mathrm{d}x = \int_{x_1}^{x_2} \cosh\left(\frac{x-\beta}{\alpha}\right)\mathrm{d}x = \alpha\sinh\left(\frac{x-\beta}{\alpha}\right)\bigg|_{x_1}^{x_2}$$
$$= \alpha\left[\sinh\left(\frac{x_2-\beta}{\alpha}\right) - \sinh\left(\frac{x_1-\beta}{\alpha}\right)\right].$$

Notice that this constraint imposes that α, β, and λ satisfy L, so our original problem effectively becomes a boundary value problem (BVP) with the constraint

$$L = \alpha\left[\sinh\left(\frac{x_2-\beta}{\alpha}\right) - \sinh\left(\frac{x_1-\beta}{\alpha}\right)\right]$$

and boundary conditions

$$y_1 = \alpha\cosh\left(\frac{x_1-\beta}{\alpha}\right) - \lambda$$
$$y_2 = \alpha\cosh\left(\frac{x_2-\beta}{\alpha}\right) - \lambda.$$

Obviously, integral constraints are not too common in the methods of Lagrangian and Hamiltonian mechanics (at least from what's covered in this course), but understanding the principles behind them are important. They translate nicely into the general type of constraint problems that we wish to solve, and furthermore act as the basis for other topics such as D'Alembert's Principle.

2 Lagrangian Formalism

2.1 Minimal Principles in Physics

We begin our discussion on the formalism of Lagrangians and Hamiltonians by first looking at how Newton's Laws differ. In particular, Newton's laws suggest a view of particles as a sequence of tiny steps. That is, we can describe the motion of a particle by the following idea:

$$\begin{array}{ccc} \mathbf{r}(t_0) & & \mathbf{r}(t_0+\mathrm{d}t) = \mathbf{r}(t_0) + \mathbf{v}(t_0)\,\mathrm{d}t \\ \mathbf{v}(t_0) & \xrightarrow{\mathrm{d}t} & \mathbf{v}(t_0+\mathrm{d}t) = \mathbf{v}(t_0) + \mathbf{a}(t_0)\,\mathrm{d}t \\ \mathbf{F}(t_0) \to \mathbf{a}(t_0) & & \mathbf{F}(t_0+\mathrm{d}t) \to \mathbf{a}(t_0+\mathrm{d}t). \end{array}$$

Minimal principles, however, do not use this approach. Instead, they attempt to identify some physical property of a path which is minimized. For example, **Fermat's principle of least time** identifies that light takes the path of shortest time within a medium. This, in turn, predicts Snell's law and other refractive-based optical components.

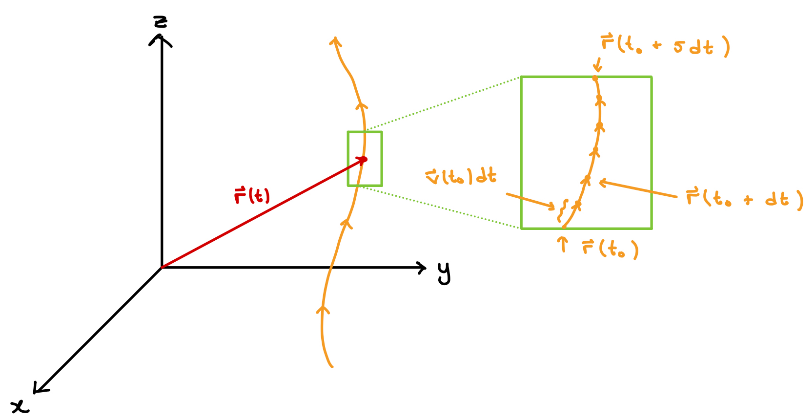


Figure 7: Newtonian view of classical mechanics.

In 1747, Maupertuis developed his **principle of least action**, which, although outdated now, predates the definition of action that is used as the foundation of Lagrangian and Hamiltonian mechanics today. Later, in 1829, Friedrich Gauss added additional constraints on this principle, and in 1896, Heinrich Hertz used it as a basis to develop the principle of least curvature.

2.2 Hamilton's Principle and the Euler-Lagrange Equation

We now direct our attention to formalizing Lagrangian mechanics. In particular, it is built on Hamilton's Principle—the **principle of least action**—which predicts the motion of a classical system where all forces aside from constraints arise from a **generalized scalar potential** of the form

$$V(q, \dot{q}, t)$$

where q is a generalized coordinate. The force which naturally follows is of the form

$$F_i = -\frac{\partial V}{\partial q_i} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V}{\partial \dot{q}_i}$$

where the index i corresponds to the DoF of the system of interest. Systems characterized by this approach are considered **monogenic**, and in the special case where $V = V(q_i)$, that is, V is independent of $\dot{q}_i$, we say that the system is **conservative**.

2.2.1 Hamilton's Principle

Hamilton's Principle, in essence, says that the actual path which a particle follows between two points $1 \to 2$ in a given time interval, $t_1 \to t_2$, is such that the action integral

$$\boxed{S = \int_{t_1}^{t_2} \mathcal{L}(q_i, \dot{q}_i, t)\, \mathrm{d}t} \tag{9}$$

is stationary along the path. By stationary, we mean that the functional, S, is extremized according to the dependent variables $q_i(t)$ and independent variable, t.

In (9), the quantity $\mathcal{L}(q_i, \dot{q}_i, t)$ is called the **Lagrangian**, and its arguments are those of generalized coordinates, q_i, and generalized velocities, $\dot{q}_i$, where $i = 1, 2 \cdots, n$. In our case of monogenic systems, we define the Lagrangian as

$$\boxed{\mathcal{L} = T - V} \tag{10}$$

and for S to be extremized, the solution follows the **Euler-Lagrange Equations** (ELE)

$$\boxed{\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = 0, \quad i = 1, 2, \cdots, n.} \tag{11}$$

The index, i, implies that there is one ELE per generalized coordinate, and together define the *equations of motion* (EoM) for the system. Furthermore, notice that (10) is in the form of Euler's Equation, where in particular we have that $f \to \mathcal{L}$ and $x \to t$.

Now, we must ask how many DoF do we need for a given system. It follows that we need a single generalized coordinate, q_i, for each independent direction a particle in the system can move. For example, a particle moving along a 1D line needs one q_i, whereas a particle moving in 3D needs three q_i's.

In the case of rigid bodies in 3D, we find that there are 6 DoF, where there are three DoF's to specify the Center of Mass (CoM) of the system, and three DoF's to specify the orientation of the system.

DoF's are also additive, in that for every number of particle in the system, there are always $\sum_{i=1}^{n} d_i$ amount of DoF, where d_i is the DoF for a single particle, and n are the number of particles in the system. In the case of a two-particle system in 3D, we have $d_i = 3$ DoF's for both particles. Since there are two particles, $n = 2$, and so the total number of DoF is $\sum_{i=1}^{2} d_i = 3 + 3 = 6$ DoF's.

Of course, we may recall that constraints reduce the amount of DoF's we have on a system, so we can express this most generally as

$$\mathrm{DoF} = \sum_{i=1}^{N} d_i - C$$

where d_i are the DoF of a single particle in the system, and C are the number of independent constraints acting on the system. Let us now discuss how to interpret this new Lagrangian formalism of classical mechanics.

Example 2.1. Consider a 1D conservative system governed by the potential $V = V(x)$. Determine the EoM of the system in terms of the potential.

In this case, we may write the energies as

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{x}^2, \quad V = V(x).$$

By (10), we find that

$$\mathcal{L} = T - V = \frac{1}{2}m\dot{x}^2 - V(x)$$

and we can compute the partials to find

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial V}{\partial x}, \quad \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\ddot{x}.$$

Then, the ELE says that

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 0 \implies m\ddot{x} = -\frac{\partial V}{\partial x}$$

which is exactly equivalent to Newton's 2nd law, in that

$$\mathbf{F} = -\vec{\nabla}V = -\frac{\partial V}{\partial x}$$

where the force, $\mathbf{F}$, is $m\ddot{x}$. This result is easily extendable to 3D, where instead of the Lagrangian only depending on the quantity, x, it would depend on (x, y, z) with a potential of the same dependence.

Ex. (2.1) served to act as an introduction to the methods of Lagrangian mechanics. Now, we shall consider a 2D Lagrangian and see how this gives rise to the different types of quantities acting on a system.

Suppose we have a Lagrangian governed by the energies

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2), \quad V = V(x, y).$$

As before, we have

$$\mathcal{L} = T - V = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y).$$

We will now apply the polar-coordinate transformation, such that $\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2\dot{\phi}^2$, and thus

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - V(r, \phi).$$

Now, our partials in r are

$$\frac{\partial \mathcal{L}}{\partial r} = mr\dot{\phi}^2 - \frac{\partial V}{\partial r}, \quad \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r} \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{r}} = m\ddot{r}$$

and in ϕ,

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{\partial V}{\partial \phi}, \quad \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = mr^2\dot{\phi} \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{r}} = mr^2\ddot{\phi} + 2mr\dot{r}\dot{\phi}$$

where we have used the chain rule on the time derivative of $\partial\mathcal{L}/\partial\dot{\phi}$. We arrive at the ELE's,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial\mathcal{L}} = \frac{\partial\mathcal{L}}{\partial r} &\implies m\ddot{r} = mr\dot{\phi}^2 - \frac{\partial V}{\partial r} \\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \frac{\partial\mathcal{L}}{\partial\phi} &\implies mr^2\ddot{\phi} + 2mr\dot{r}\dot{\phi} = -\frac{\partial V}{\partial\phi}.
\end{aligned}$$

Notice, in this case, we end up with a few unfamiliar terms. We note them momentarily here:

1. $mr^2\ddot{\phi}$ corresponds to the moment of inertia of a particle of mass, m.
2. $mr\dot{\phi}^2$ corresponds to the centripetal acceleration of the particle.
3. $2mr\dot{r}\dot{\phi}$ is a Coriolis term, arising from rotational motion.
4. $-\partial V/\partial r$ corresponds to the radial force.

We have neglected to reference $-\partial V/\partial\phi$. In fact, a quick dimensional analysis would show that $[\partial V/\partial\phi] = [J]$, which is **not** the units of force. If this is the case, we must reconcile what this quantity *actually* is, and what it does to our system. Luckily, with some algebraic manipulations, we can see the following,

$$-\frac{\partial V}{\partial\phi} = -\frac{\partial V}{\partial x}\frac{\partial x}{\partial\phi} - \frac{\partial V}{\partial y}\frac{\partial y}{\partial\phi} = F_x[-r\sin(\phi)] + F_y[r\cos(\phi)].$$

However, $-r\sin(\phi)$ is simply $-y$, and $r\cos(\phi)$ is x, so we find that

$$-\frac{\partial V}{\partial\phi} = -yF_x + xF_y = (\mathbf{r}\times\mathbf{F})_z = \mathbf{N}_z$$

which shows that this new term is actually the z-component of the torque exerted on the particle. Notice that this derivation assumed that r varied. We can, however, describe circular motion as a special case of the dynamics of this system.

Let us assume $r =$ constant, such that $\dot{r} = \ddot{r} = 0$. From this, we can see that

$$-\frac{\partial V}{\partial\phi} = mr^2\ddot{\phi} + \cancel{2mr\dot{r}\dot{\phi}} = mr^2\ddot{\phi}.$$

We already know that $mr^2\ddot{\phi}$ is the moment of inertia of the particle, $I\alpha$, and $-\partial V/\partial\phi = \mathbf{N}_z$. Then, the ELE in ϕ becomes equivalent to the Newtonian relation for torque.

$$\mathbf{N}_z = I\alpha_z.$$

We must notice now that we have switched from purely Cartesian coordinates to a mix of both Cartesian and angular coordinates. However, the EoM that we arrive at could have been obtained through Newtonian mechanics all the same. Thus, the Lagrangian and Newtonian formalisms of classical mechanics are equivalent.

2.3 Generalized Coordinates

We will now discuss the importance of the term, **generalized coordinates**, something we have thrown out loosely without defining. There is a specific property of these coordinates, which is that they are **invariant**.

To see this, let us suppose we have two sets of coordinates, x_j and q_i, where j and i are the indices corresponding to each coordinate within a single system. We say that a coordinate is **invertible** if they satisfy the property

$$\boxed{x_j = x_j(\{q_i\}, t) \Longleftrightarrow q_i = q_i(\{x_j\}, t)} \tag{12}$$

that is, x_j can be expressed as a function of the coordinates q_i, or vice versa. If the transformation $x_j \Leftrightarrow q_i$ is constant in time, then we call them **fixed** or **scleronomic**. On the other hand, if the transformation is time-dependent, then they are called **rheonomic**.

Furthermore, we define two types of transformations: **passive** and **active**. A passive transformation is one which describes the same point. That is, $x_j = q_i$; an active transformation moves the particle in space-time. That is, $x_j \neq q_i$.

Now, suppose that we did not know about a transformation $x_j \Leftrightarrow q_i$ and are interested in how other coordinates interact with x_j. We approach this by deriving some new relations which depend on both coordinates. Let us first assume the ELE in x is valid.

$$\frac{\partial \mathcal{L}}{\partial x_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_j} = 0.$$

We also assume that there exists a coordinate, q_i, such that we can write $x_j(\{q_i\})$, and x_j is time independent.[4]

Relation 1. Let us take the time derivative of x_i, such that

$$\dot{x}_j = \frac{\mathrm{d}x_j}{\mathrm{d}t} = \sum_{m=1}^{n} \frac{\partial x_j}{\partial q_m}\frac{\partial q_m}{\partial t} + \frac{\partial x_j}{\partial t} = \sum_{m=1}^{n} \frac{\partial x_j}{\partial q_m}\dot{q}_m + \cancel{\frac{\partial x_j}{\partial t}} = \sum_{m=1}^{n} \frac{\partial x_j}{\partial q_m}\dot{q}_m$$

Then, dropping the sum since we are only interested in a single coordinate, we have

$$\boxed{\frac{\partial \dot{x}_j}{\partial \dot{q}_m} = \frac{\partial x_j}{\partial q_m}.} \tag{13}$$

Relation 2. Let us now take the time derivative of $\partial x_j / \partial q_m$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_j}{\partial q_m} = \sum_{k=1}^{n-1}\left[\frac{\partial}{\partial q_k}\frac{\partial x_j}{\partial q_m}\right]\frac{\partial q_k}{\partial t} + \frac{\partial}{\partial t}\frac{\partial x_j}{\partial q_m}$$

where we have introduced the new index, k, to represent the derivative of q after using the chain rule. Since partial derivatives commute, we can rewrite this as

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_j}{\partial q_m} = \frac{\partial}{\partial q_m}\left[\sum_{k=1}^{n-1}\frac{\partial x_j}{\partial q_k}\frac{\partial q_k}{\partial t} + \frac{\partial x_j}{\partial t}\right] = \frac{\partial}{\partial q_m}\frac{\mathrm{d}x_j}{\mathrm{d}t}.$$

[4]Subtly, this new approach of writing the each of the generalized coordinates q_i, as a set, $\{q_i\}$, introduces the idea of a *configuration* space.

Thus,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_j}{\partial q_m} = \frac{\partial}{\partial q_m}\frac{\mathrm{d}x_j}{\mathrm{d}t}.} \tag{14}$$

Now, using these results, let us consider the time derivative of the Lagrangian with respect to its coordinates. We have,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_m} = \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial x_j}\cancel{\frac{\partial x_j}{\partial \dot{q}_m}} + \sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial \dot{x}_j}\frac{\partial \dot{x}_j}{\partial \dot{q}_m}\right].$$

However, we have assumed that x_j has no $\dot{q}_i$ dependence, so this simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_m} = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial \dot{q}_m}.$$

By Relation 1,

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial \dot{q}_m} = \sum_{i=1}^{n}\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right]\frac{\partial x_i}{\partial q_m} + \sum_{i=1}^{n}\frac{\partial}{\partial \mathcal{L}}\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_i}{\partial q_m}\right].$$

The first term is simply the ELE expressed only by the time derivative component, while the second term invokves Relation 2, so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial \dot{q}_m} = \sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial x_i}\frac{\partial x_i}{\partial q_m} + \sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial q_m} = \sum_{i=1}^{n}\frac{\partial \mathcal{L}}{\partial q_m}.$$

Thus, dropping the sums since they are no longer necessary, we finally arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_m} = \frac{\partial \mathcal{L}}{\partial q_m}$$

which shows that the form of the ELE are unchanged. The implications of this result is such that, if one ELE gives the correct EoM in one coordinate system, we can always transform them into another coordinate system and expect the result to hold true as well.

2.3.1 Choice of Generalized Coordinates

That's great and all, but now that we know this, we must make a distinction between the different choices we have for our coordinate systems. In fact, with the result we found previously, we are now open to choosing coordinates that are non-inertial, or coordinates that do not explicitly depend on units of length.

Let us refer back to $\{q_i\}$. We refer to these set of coordinates as **complete**, **generalized coordinates**. That is, they are capable of locating all components of a system at all times.

In particular, we say that if the number of generalized coordinates, n, are equal to the DoF of the system, then $\{q_i\}$ is referred to as a **proper set** of generalized coordinates. Typically, this may arise if there are no holonomic constraints posed in the system, or if we have already built all constraints into the coordinate system itself. Let's consider an example.

The Compound Pendulum. We can represent the compound pendulum either in Cartesian or angular coordinates.

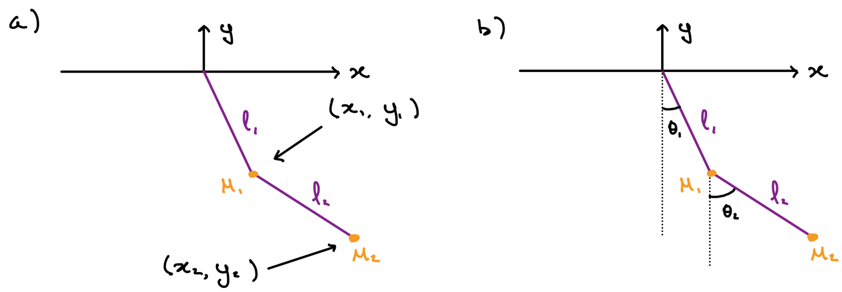


Figure 8: (a) Compound pendulum represented in Cartesian coordinates. (b) Compound pendulum represented in angular coordinates.

In Cartesian, the compound pendulum imposes two constraints of the forms

$$g_1(x_1, y_1) = x_1^2 + x_1^2 - l_1^2 = 0$$
$$g_2(x_1, x_2, y_1, y_2) = (x_2 - x_1)^2 + (y_2 - y_1)^2 - l_2^2 = 0.$$

However, these constraints make the problem nasty to solve, even with Lagrangian mechanics. Consider instead angular coordinates, which has no constraints. Why is this the case? Notice, in Fig. 8b, we can see that the angles θ_1 and θ_2 are free to vary for all θ. There is nothing obstructing each component of the pendulum from making a full 360° rotation, as opposed to Cartesian which inherently limits the rod length to the coordinates of the masses.

In this way, we can see that $\{\theta_1, \theta_2\}$ are a proper set of generalized coordinates. In essence, the best choice for generalized coordinates is always choosing the ones which minimize the constraints that the system pose. In most cases, this is the proper set of coordinates. In the rest, it's the one with the least constraints.

2.3.2 Generalized Momenta and Generalized Forces

Let us now consider a Lagrangian of the form, $\mathcal{L} = \mathcal{L}(x, \dot{x})$ and investigate its properties to see how generalized coordinates give rise to generalized momenta and forces. For a 1D system, we can write the Lagrangian subjected to a conservative potential $V = V(x)$ as

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x).$$

Our generalized momentum and force, in turn, is

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}, \qquad \frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial V}{\partial x}.$$

Here, $p = m\dot{x}$ is the linear momentum associated in the x-direction, and $-\partial V/\partial x$ is the force arising from our conservative potential. Furthermore, if we take the time derivative of p, we find that

$$\dot{p} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\ddot{x} = -\frac{\partial V}{\partial x}$$

which is the generalized force we find before. In particular, we can see that if a force was applied to the system as it moved, we can define the differential work as

$$\mathrm{d}W = F\,\mathrm{d}x$$

with units $[\mathrm{d}W] = [F][\mathrm{d}x] = [N][m] = [J]$ of energy. This example only discusses a particular case in 1D, but we can easily see how this extends to a generalized set of coordinates $\{q_i\}$. In this case, we have $\mathcal{L} = \mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t)$. Our generalized momenta, then, is

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \tag{15}$$

and the set of all generalized momenta is given by $\{p_i\}$ after taking the partial derivative of the Lagrangian with respect to the set of all generalized coordinates. In this way, we say that the momenta, $\{p_i\}$, is **canonically conjugate** to the generalized coordinates, $\{q_i\}$.[5] Now, notice that because we describe the set of momenta $\{p_i\}$ in this way, it naturally follows that our generalized momenta do not need to have units of momentum, that is $[p] = [kg][m][s^{-2}]$.

We can give rise to our generalized forces in a similar manner. In fact,

$$F_i = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i} \tag{16}$$

which also does not need to have units of force. However, the differential work is still units of energy, where

$$\mathrm{d}W = F_i\,\mathrm{d}q_i\,. \tag{17}$$

In the case of $[q_i] = [\mathrm{rad}]$, then $[\mathrm{d}W] = [F_i][\mathrm{d}q_i] = [N][m][\mathrm{rad}] = [J]$, where F_i is a torque, giving rise to energy.

2.3.3 Dissipation and Forces of Constraint

Notice, our discussion of generalized forces and momenta do not give rise to certain types of forces. In particular, we cannot derive constraint or dissipative forces from a scalar potential. These are inherently external, and arise from conditions on the system itself.

In this case, we can manually include them into our ELE, such that

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i} + Q_i} \tag{18}$$

where Q_i denotes any additional force imposed on the system. We must be careful here. Note that if we add generalized forces to the ELE, they must also have the same units as the generalized forces derived from $\mathcal{L}$ and $\partial\mathcal{L}/\partial q_i$. For example, suppose we want to include a drag force, Q_i. We would have to account for the unit conversions with additional constants, such that

$$Q_i = -b\dot{q}_i$$

[5]This is slightly loosely defined. By "canonically conjugate", we sort of imply that the variables $\{p_i\}$ and $\{q_i\}$ play a reciprocal role. This is a lot more evident when we begin discussing Poisson brackets, but for now, take this as the explanation.

where b would contain any units necessary to have units of drag. Now, let us recall the discussion in Chapter 1.3. We saw that Euler's Equations were modified to contain a constraint term, $-\lambda(x)$. This is exactly the same as in our Lagrangian approach, where

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i} + \sum_{j=1}^{m} \lambda_j(t)\frac{\partial g_j}{\partial q_j}, \quad g_j(\{q_i\}, t) = 0 + \text{Boundary Conditions.}} \tag{19}$$

Here, our generalized forces are the set, $\{Q_{ij}\}$, given by $\sum_{j=1}^{m} \lambda_j(t)\partial g_j/\partial q_j$, and constraint is $g_j(\{q_i\}, t) = 0$. We will now look at an example which incorporates these qualities.

Example 2.2 (Particle on a Dome)**.** Consider a particle of mass, m, starting from rest at the top of a hemispherical dome of radius, a. Under the influence of gravity, determine when the particle loses contact with the dome.

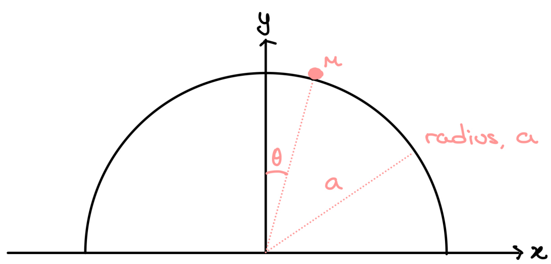


Figure 9: Hemispherical dome constrained to have a radius, a, with a particle, m, tracing an angle, θ, as it travels.

We begin by choosing polar coordinates to take advantage of the geometry of our system. The radius of the dome, a, creates a holonomic constraint on the particle, which we can model as

$$g(r, \theta) = r - a = 0$$

which enforces that the particle remains on the surface. Now, the generalized coordinates in Cartesian are

$$\begin{matrix} x = r\sin(\theta) \\ y = r\cos(\theta) \end{matrix} \implies \begin{matrix} \dot{x} = \dot{r}\sin(\theta) + r\cos(\theta)\dot{\theta} \\ \dot{y} = \dot{r}\cos(\theta) - r\sin(\theta)\dot{\theta}. \end{matrix}$$

Computing $\dot{x}^2 + \dot{y}^2$, we find

$$\begin{aligned} \dot{x}^2 + \dot{y}^2 &= \left[\dot{r}\sin(\theta) + r\cos(\theta)\dot{\theta}\right]^2 + \left[\dot{r}\cos(\theta) - r\sin(\theta)\dot{\theta}\right]^2 \\ &= \dot{r}^2\sin^2(\theta) + \cancel{2r\dot{r}\sin(\theta)\cos(\theta)\dot{\theta}} + r^2\cos[2]\dot{\theta}^2 + \dot{r}^2\cos^2(\theta) - \cancel{2r\dot{r}\cos\theta\sin\theta\dot{\theta}} + r^2\sin^2(\theta)\dot{\theta}^2 \\ &= \dot{r}^2 + r^2\dot{\theta}^2 \end{aligned}$$

which shows that the kinetic energy is

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)$$

and the potential,

$$V = mgy = mgr\cos(\theta).$$

Our Lagrangian follows and is of the form

$$\mathcal{L} = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgr\cos(\theta).$$

Now, the problem becomes a matter of applying the ELE for each generalized coordinate. Computing the partials in r, we get

$$\frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r} \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{r}} = m\ddot{r}, \qquad \frac{\partial \mathcal{L}}{\partial r} = mr\dot{\theta}^2 - mg\cos(\theta), \qquad \frac{\partial g}{\partial r} = 1$$

and in θ,

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta} \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta}, \qquad \frac{\partial \mathcal{L}}{\partial \theta} = mgr\sin(\theta), \qquad \frac{\partial g}{\partial \theta} = 0.$$

Thus, we obtain the ELE's,

$$\begin{aligned} m\ddot{r} &= mr\dot{\theta}^2 - mg\cos(\theta) + \lambda(t) \\ mgr\sin(\theta) &= mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} \end{aligned}$$

and with addition of the constraint, is a system of three equations with three unknowns. Solving this, we begin with the constraint, which shows that

$$r = a \implies \dot{r} = \ddot{r} = 0$$

and using this to simplify, we arrive at

$$\begin{aligned} \dot{\theta}^2 &= \frac{g}{a}\cos(\theta) - \frac{\lambda(t)}{ma} \\ \ddot{\theta} &= \frac{g}{a}\sin(\theta). \end{aligned}$$

To determine when the particle loses contact with the dome, we use chain rule,

$$\ddot{\theta} = \frac{\mathrm{d}\dot{\theta}}{\mathrm{d}t} = \frac{\mathrm{d}\dot{\theta}}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \dot{\theta}\frac{\mathrm{d}\dot{\theta}}{\mathrm{d}\theta}$$

so the second equation simplifies to

$$\begin{aligned} \dot{\theta}\frac{\mathrm{d}\dot{\theta}}{\mathrm{d}\theta} &= \frac{g}{a}\sin(\theta) \\ \int \dot{\theta}\,\mathrm{d}\dot{\theta} &= \frac{g}{a}\int \sin(\theta)\,\mathrm{d}\theta \\ \frac{1}{2}\dot{\theta}^2 &= \frac{g}{a}[C - \cos(\theta)]. \end{aligned}$$

The initial conditions of our system assume that $\dot{\theta}(0) = \theta(0) = 0$,[6] and so we find that

$$0 = \frac{g}{a}[C-1] \implies C = 1 \implies \dot{\theta}^2 = \frac{2g}{a}[1 - \cos(\theta)].$$

[6] Technically, the initial condition on $\theta(0)$ should be $\theta(0) \approx \varepsilon$, for some small perturbation, ε. However, for our purposes we can approximate this as zero.

Recalling our first ELE, we now have

$$\begin{aligned}\frac{2g}{\cancel{d}}[1-\cos(\theta)] - \frac{g}{\cancel{d}}\cos(\theta) + \frac{\lambda(t)}{m\cancel{d}} &= 0\\ 2g - 3g\cos(\theta) + \frac{\lambda(t)}{m} &= 0\\ 3mg\cos(\theta) - 2mg = \lambda(t) &= \lambda(\theta).\end{aligned}$$

Now, by our implicit definition of Q_{ij}, we can see that the force of constraint in r is

$$Q_r = \lambda(\theta)\frac{\partial g}{\partial r} = 3mg\cos(\theta) - 2mg$$

which is, in fact, the normal force exerted on the particle for any angle θ. We can determine the critical angle of the particle by taking note of when Q_r changes sign. That is, $Q_r < 0$, which is nonphysical. The critical point is then at the point $Q_r = 0$, and so

$$3mg\cos(\theta_c) = 2mg \implies \theta_c = \arccos\left(\frac{2}{3}\right) \approx 122^\circ.$$

There are a few important observations to make in regards to Ex. (2.2). The first is that we *could* have solved it by using θ as a proper set of generalized coordinates. However, this approach would not have given us the normal force exerted on the particle, and by extension not have given us the opportunity to determine the critical angle of the particle.

In a sense, embedding constraints into our system "loses" information that we could have known about our constraint force, since it becomes encoded into the system itself.

Another note is the coordinate transformation from Cartesian to polar. In general, it's easier to write your energy terms in Cartesian coordinates first, then transform them into generalized coordinates. However, it is possible to skip straight to the known velocity results. In this case, it would have been $\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2\dot{\theta}^2$.

2.4 Gauge Invariance

If you are observant, you may notice that when computing the EoMs, constant terms within the Lagrangian do not contribute. In fact, this is a more general consequence of a property called **gauge invariance**. A Lagrangian transformation which encapsulates these ideas is of the form

$$\boxed{\mathcal{L}' = \alpha\mathcal{L} + \beta + f(t).} \tag{20}$$

where, α and β are arbitrary constants, and $f(t)$ is some arbitrary function of time only. We must make a note on the constant α. In situations with constraining forces, we can only retain their proper interpretations if the constant α also ensures that the units of the generalized force remains consistent. Otherwise, terms like $f(t)$ or β do not impact the EoM, and can thus be discarded if needed.

There is also another less obvious transformation to consider,

$$\boxed{\mathcal{L}(q_i, \dot{q}_i, t) \to \mathcal{L}(q_i, \dot{q}_i, t) + \frac{\mathrm{d}\Lambda(q_i, t)}{\mathrm{d}t}.} \tag{21}$$

where $\Lambda = \Lambda(\{q_i\}, t)$, an arbitrary scalar function, and is independent of the generalized velocities. In this case, we find that the ELE pick up additional terms. In particular (11) says that,

$$\frac{\partial}{\partial q_i}\frac{\mathrm{d}\Lambda(q_i,t)}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial q_i}\frac{\mathrm{d}\Lambda(q_i,t)}{\mathrm{d}t}\right] = \frac{\partial}{\partial q_i}\frac{\mathrm{d}\Lambda(q_i,t)}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial q_i}\left(\sum_{j=1}^{n}\frac{\partial\Lambda(q_i,t)}{\partial q_j}q_j\right)\right] - \cancel{\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial q_i}\frac{\partial\Lambda(q_i,t)}{\partial t}\right]}$$

where the last term is zero since $\Lambda(q_i, t)$ is independent of $\dot{q}_i$. Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial q_i}\left(\sum_{j=1}^{n}\frac{\partial\Lambda(q_i,t)}{\partial q_j}q_j\right)\right] = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\Lambda(q_i,t)}{\partial q_i}$$

to give the result

$$\frac{\partial}{\partial q_i}\frac{\mathrm{d}\Lambda(q_i,t)}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial q_i}\frac{\mathrm{d}\Lambda(q_i,t)}{\mathrm{d}t}\right] = \frac{\partial}{\partial q_i}\frac{\mathrm{d}\Lambda(q_i,t)}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\Lambda(q_i,t)}{\partial q_i} = 0.$$

Notice, we have invoked Relation 2 on the final step, and that this derivation assumes that the functions are well behaved, and as well as contain well defined derivatives. We can see, then, that adding a full time derivative of a function with arguments of the generalized coordinates and time does not impact the EoM.

2.4.1 Gauge Invariance in Electricity and Magnetism

A topic where gauge invariance becomes rather important is in electrodynamics. To see this, consider the following example.

Suppose we have a particle of charge, q, moving with a velocity, $\mathbf{v}$, through an electric field, $\mathbf{E}$, and subjected to a magnetic field, $\mathbf{B}$. The particle, then, experiences the Lorentz force,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

We can rewrite the electric and magnetic fields in terms of a scalar potential, $\phi(\mathbf{x}, t)$ and a vector potential, $\mathbf{A}(\mathbf{x}, t)$, such that

$$\begin{aligned}\mathbf{E}(\mathbf{x},t) &= -\vec{\nabla}\phi(\mathbf{x},t) - \frac{\partial\mathbf{A}(\mathbf{x},t)}{\partial t}\\ \mathbf{B}(\mathbf{x},t) &= \vec{\nabla}\times\mathbf{A}(\mathbf{x},t).\end{aligned}$$

This rewriting of $\mathbf{E}$ and $\mathbf{B}$ is not an issue, since they automatically satisfy Maxwell's equations[7]

$$\begin{aligned}\vec{\nabla}\cdot\mathbf{B} &= 0\\ \vec{\nabla}\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t}.\end{aligned}$$

Then, we can write the potential of the particle in the form

$$V = -q[\mathbf{v}\cdot\mathbf{A}(\mathbf{x},t) - \phi(\mathbf{x},t)]$$

[7] To see this explicitly, you could use the vector calculus identities $\vec{\nabla}\times\vec{\nabla}f = 0$ and $\vec{\nabla}\cdot(\vec{\nabla}\times\mathbf{F}) = 0$, where f is a scalar fuction and $\mathbf{F}$ is a vector field.

where $\mathbf{v}$ is the velocity of the particle. Now, we shall justify this potential by writing the Lagrangian

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m|\mathbf{v}|^2 + q\mathbf{v}\cdot\mathbf{A}(\mathbf{x},t) - q\phi \\ &= \frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2) + q[\dot{x}A_x(\mathbf{x},t)+\dot{y}A_y(\mathbf{x},t)+\dot{z}A_z(\mathbf{x},t)] - q\phi(\mathbf{x},t).\end{aligned}$$

Computing the partials of our Lagrangian, we find

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial x} &= q\left[\dot{x}\frac{\partial A_x(\mathbf{x},t)}{\partial x} + \dot{y}\frac{\partial A_y(\mathbf{x},t)}{\partial x} + \dot{z}\frac{\partial A_z(\mathbf{x},t)}{\partial x} - \frac{\partial\phi(\mathbf{x},t)}{\partial x}\right] \\ \frac{\partial\mathcal{L}}{\partial\dot{x}} &= m\dot{x} + qA_x \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}} &= m\ddot{x} + q\left[\frac{\partial A_x(\mathbf{x},t)}{\partial x}\dot{x} + \frac{\partial A_x(\mathbf{x},t)}{\partial y}\dot{y} + \frac{\partial A_x(\mathbf{x},t)}{\partial z}\dot{z} + \frac{\partial A_x(\mathbf{x},t)}{\partial t}\right].\end{aligned}$$

Notice that the quantities $q(\partial A_x/\partial x)\dot{x}$ cancel in the ELE, and after some simplifications, we find

$$\begin{aligned}m\ddot{x} &= q\left(-\frac{\partial\phi(\mathbf{x},t)}{\partial x} - \frac{\partial A_x(\mathbf{x},t)}{\partial t}\right) + q\left[\dot{y}\left(\frac{\partial A_y(\mathbf{x},t)}{\partial x} - \frac{\partial A_x(\mathbf{x},t)}{\partial y}\right) + \dot{z}\left(\frac{\partial A_z(\mathbf{x},t)}{\partial x} - \frac{\partial A_x(\mathbf{x},t)}{\partial z}\right)\right] \\ &= qE_x + q\dot{y}B_z + q\dot{z}(-B_y) \\ m\ddot{x} &= q[E_x + (\dot{\mathbf{x}}\times\mathbf{B})_x]\end{aligned}$$

which is the EoM in x. Repeating this process, we find the EoM in y and z, which are

$$\begin{aligned}m\ddot{y} &= q[E_y + (\dot{\mathbf{x}}\times\mathbf{B})_y] \\ m\ddot{z} &= q[E_z + (\dot{\mathbf{x}}\times\mathbf{B})_z].\end{aligned}$$

which justifies our potential term. Now, we shall add a full time derivative to the Lagrangian and see how this interacts with the system.

$$\mathcal{L} = \frac{1}{2}m|\mathbf{v}|^2 + q\mathbf{v}\cdot\mathbf{A} - q\phi + q\frac{\mathrm{d}\Lambda}{\mathrm{d}t}$$

where $\Lambda = \Lambda(\mathbf{x},t)$. We will drop the arguments of the functions, since they are redundant, and we rewrite the time derivative as

$$q\left(\frac{\partial\Lambda}{\partial x}\dot{x} + \frac{\partial\Lambda}{\partial y}\dot{y} + \frac{\partial\Lambda}{\partial z}\dot{z} + \frac{\partial\Lambda}{\partial t}\right) = q\left(\mathbf{v}\cdot\vec{\nabla}\Lambda + \frac{\partial\Lambda}{\partial t}\right)$$

to give

$$\mathcal{L} = \frac{1}{2}m|\mathbf{v}|^2 + q\mathbf{v}\cdot\left(\mathbf{A}+\vec{\nabla}\Lambda\right) - q\left(\phi - \frac{\partial\Lambda}{\partial t}\right).$$

Let us apply the transformations

$$\mathbf{A}' = \mathbf{A} + \vec{\nabla}\Lambda, \quad \phi' = \phi - \frac{\partial\Lambda}{\partial t}$$

to give our transformed Lagrangian

$$\mathcal{L}' = \frac{1}{2}m|\mathbf{v}|^2 + q\mathbf{v}\cdot\mathbf{A}' - q\phi'$$

which is actually of the same form as the Lagrangian we had before adding the time derivative. Naturally, our EoM are the same, and to see this more explicitly, you would only need to apply the same process as before. The transformations, in turn, are **gauge transformations**.

2.5 Setting up Lagrangians

One of the most difficult parts of Lagrangian and Hamiltonian mechanics is actually setting up the problems themselves. However, doing so really boils down to a methodical step-by-step process which never fails.

1. Draw a clear starting diagram.
2. Determine the DoF in the system.
3. Choose generalized coordinates and label them on the diagram.
4. Determine any constraints within the system.
5. Label all distances on the diagram.
6. Draw an inertial, right-handed, Cartesian coordinate system with the origin at a reasonable position.
7. Write the position of each mass in the Cartesian coordinate system.
8. Incorporate holonomic constraints, or utilize LUM.
9. Write the expression for the energies.
10. Write out the Lagrangian.
11. Identify any conserved quantities.
12. Calculate derivatives for the ELE's.
13. Write out the EoM.
14. Confirm that your answer is reasonable.

In point (3.), we should be careful. In particular, it's important to draw the axis as a **directional** arrow, and not a two-sided one. The direction matters severely, as it impacts the EoM.

In point (6.), reasonability means an origin which would make the calculations the easiest to deal with, or quantities the clearest.

In point (11.), conserved quantities can be identified through **cyclic coordinates** or by **Noether's Theorem**. We discuss this in more detail later.

In point (12.), it is recommended to calculate each derivative component separately, then combine them later.

In point (14.), reasonability refers to checking units and the basic behaviour of the EoM. In particular, an EoM should align with the expected motion of the system.

These are a lot of steps, but familiarity makes them simple. To see how we could employ this method, let us discuss a particularly inconspicuous example.

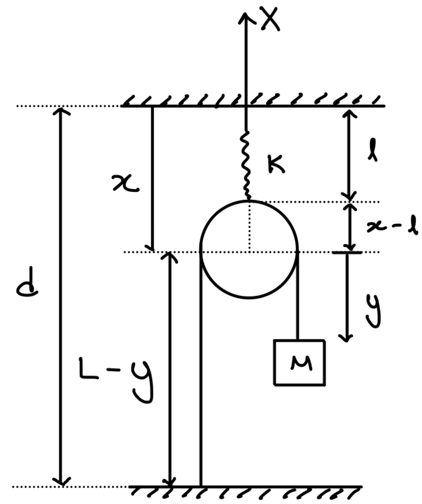


Figure 10: The "funny scale" problem, with all labeled quantities as per the 11 steps.

Consider the "funny scale" in Fig. 10. We are interested in how much the spring stretches compared to a regular scale with with a mass, m.

1. To begin, we draw a starting diagram of the system, as seen in 10. Notice, this diagram also includes steps 3, 5, and 6, so we will at most just discuss them here.

2. In this scenario, we can see that both the mass and the pulley are free to move. However, the relative motion of the system is constrained *if* the string remains in contact with the pulley. We then have

$$\text{DoF} = \sum_{i=1}^{2} d_i - C = (1+1) - 1 = 1$$

which is a single DoF. Thus, we expect a single generalized coordinate *provided* that we incorporate the constraint into our Lagrangian. In the case of LUM, we also expect a single DoF, since we would have a two generalized coordinates affected by a holonomic constraint.

3. Included in Fig. 10.

4. The constraint in this system is the fixed string length. Informally, we will say that by *"conservation of string"*, the system is constrained by a function of the form

$$g(x, y) = x + (L - y) = d$$

where d is the fixed distance from the ceiling to the floor.

5. Included in Fig. 10. Conservation of string implies that the distance from the pulley to the floor must be $(L - y)$. Furthermore, the equilibrium distance of the spring we denote as l, which implies that the extension of the spring is $x - l$.

6. Included in Fig. 10. We need a single inertial coordinate, which we define as X. Let us consider the origin to be at the (fixed) point where the spring meets the ceiling. Other options may work too; for example, one could instead choose the positive X direction to be down. However, a choice such as placing the origin at the pulley itself would not be wise—the pulley can accelerate, so X would not be an inertial coordinate in this case.

7. The position of the mass in this system is given by

$$X = -x - y.$$

Note the sign, which says that if the scalar mass position increases (downward), then the distance from the mass to the ceiling would also increase. That is, X becomes increasingly negative as x becomes increasingly positive.

8. We must choose between incorporating the constraint into the Lagrangian, or using LUM. In this case, we will build the constraint into the Lagrangian, to which we use the constraint function to find

$$y = -d + L + x.$$

Thus, the X position of the mass is

$$X = -x - (-d + L + x) = -2x + d - L \implies \dot{X} = -2\dot{x}.$$

Should we have used LUM, we would have instead defined the Lagrangian with both x and y, and used two LUM with partials in g of x and y.

9. We now write the expression for the energies. In an inertial Cartesian coordinate system, we have the kinetic energy term to be

$$T = \frac{1}{2}m(\dot{X}^2 + \cancel{\dot{Y}^2 + \dot{Z}^2}) = \frac{1}{2}m\dot{X}^2 = \frac{1}{2}m(-2\dot{x})^2 = 2m\dot{x}^2.$$

The potential term has a contribution from both gravity and the spring, so we naturally expect the potential to be of the form

$$V = V_{\text{grav.}} + V_{\text{spring}} = mgX + \frac{1}{2}k(x-l)^2 = mg(2x + \cancel{d - L}) + \frac{1}{2}kx^2 - klx + \cancel{\frac{1}{2}kl^2} = -2mgx + \frac{1}{2}kx(x-2l)$$

where we have thrown out any terms that did not depend on the generalized coordinates since they do not impact the EoM.

10. We now write the Lagrangian, which is of the form

$$\mathcal{L} = T - V = 2m\dot{x}^2 + 2mgx - \frac{1}{2}kx(x - 2l).$$

11. As we have not yet discussed conservation laws, this step is slightly out of place. However, for posterity's sake, we note that the Lagrangian is time independent and the generalized coordinates are related to the Cartesian coordinates by a time-independent relation. We thus expect the Hamiltonian to be conserved and equal to the total energy, $\mathcal{H} = T + V$.

12. We compute the partial derivatives in the Lagrangian separately. Here, we have

$$\frac{\partial \mathcal{L}}{\partial x} = 2mg - k(x-l), \quad \frac{\partial \mathcal{L}}{\partial \dot{x}} = 4m\dot{x} \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 4m\ddot{x}.$$

13. The next step is to write the EoM. In particular, the ELE shows that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} &= \frac{\partial \mathcal{L}}{\partial x} \\
4m\ddot{x} &= 2mg - k(x-l) \\
\ddot{x} &= -\frac{k}{4m}x + \frac{2mg+kl}{4m}.
\end{aligned}$$

14. Let us confirm the units and behaviour of our EoM. Here, we must have that the right hand side be units of acceleration to match the units of the left. So,

$$\begin{aligned}
\left[\frac{k}{m}x + \frac{mg+kl}{m}\right] &= \frac{[N]\cancel{[m^{-1}]}\cancel{[m]}}{[kg]} + \frac{\cancel{[kg]}[m][s^{-2}]}{\cancel{[kg]}} + \frac{[N]\cancel{[m^{-1}]}\cancel{[m]}}{[kg]} \\
&= \frac{\cancel{[kg]}[m][s^{-2}]}{\cancel{[kg]}} + \frac{[m]}{[s^2]} + \frac{\cancel{[kg]}[m][s^{-2}]}{\cancel{[kg]}} \\
&= 3\frac{[m]}{[s^{-2}]} \\
&\sim \frac{[m]}{[s^{-2}]}
\end{aligned}$$

which agrees with what we expect. Furthermore, notice that the EoM is one for a harmonic oscillator with a constant term. We thus expect an oscillation with an offset, which agrees with the setup of the equilibrium stretch of our loaded spring.

We have thus properly solved a majority of the problem. Now, all that is left to do is determine what the ratio is between the frequency of this "funny scale" and that of a regular scale. In this problem, we see that

$$\omega_0 = \frac{1}{2}\sqrt{\frac{k}{m}}.$$

A regular scale has frequency $\omega_0' = \sqrt{k/m}$, so we find that the ratio between them is

$$\frac{\omega_0'}{\omega_0} = 2\frac{\cancel{\sqrt{k/m}}}{\cancel{\sqrt{k/m}}} = 2$$

which shows that the "funny scale" should stretch half as much as a regular scale. We could be getting ripped off!

2.6 Theorem Concerning Kinetic Energy

We will now deviate slightly from our discussion of the Lagrangian formalism to discuss a general theorem of kinetic energy. Suppose that we have a system of masses in $\mathbb{R}^3$. We can represent the total kinetic energy of the system as the sum of the kinetic energies of the masses, such that

$$T = \frac{1}{2}\sum_{\alpha=1}^{n}\sum_{i=1}^{3} m_\alpha \dot{x}_{\alpha,i}^2$$

where the sum in i accounts for the three spatial directions. Let us now consider a set of generalized coordinates, $\{q_i\}$, such that we can express the Cartesian coordinates as

$$x_{\alpha,i} = x_{\alpha,i}(\{q_j\}, t).$$

We note that $\{q_j\}$ has a different subscript to distinguish it from the subscript, i, in x. Using this, we can write the time derivative of x as

$$\dot{x}_{\alpha,i} = \frac{\mathrm{d}x_{\alpha,i}}{\mathrm{d}t} = \sum_{j=1}^{n} \frac{\partial x_{\alpha,i}}{\partial q_j}\dot{q}_j + \frac{\partial x_{\alpha,i}}{\partial t}$$

and so

$$\dot{x}^2_{\alpha,i} = \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial x_{\alpha,i}}{\partial q_j}\dot{q}_j \frac{\partial x_{\alpha,i}}{\partial q_k}\dot{q}_k + 2\sum_{j=1}^{n} \frac{\partial x_{\alpha,i}}{\partial q_j}\dot{q}_j \frac{\partial x_{\alpha,i}}{\partial t} + \left[\frac{\partial x_{\alpha,i}}{\partial t}\right]^2.$$

Thus, we can write the kinetic energy as

$$\begin{aligned} T &= \sum_{\alpha=1}^{n}\sum_{i,j,k=1}^{n} \frac{1}{2}m_\alpha \frac{\partial x_{\alpha,i}}{\partial q_j}\dot{q}_j \frac{\partial x_{\alpha,i}}{\partial q_k}\dot{q}_k + \sum_{\alpha=1}^{n}\sum_{i,j=1}^{n} m_\alpha \frac{\partial x_{\alpha,i}}{\partial q_j}\dot{q}_j \frac{\partial x_{\alpha,i}}{\partial t} + \sum_{\alpha=1}^{n}\sum_{i=1}^{n} \frac{1}{2}m_\alpha \left[\frac{\partial x_{\alpha,i}}{\partial t}\right]^2 \\ &= \sum_{j,k=1}^{n} a_{jk}\dot{q}_j\dot{q}_k + \sum_{j=1}^{n} b_j\dot{q}_j + c \end{aligned}$$

where we let

$$a_{jk} = \sum_{\alpha,j=1}^{n} \frac{1}{2}m_\alpha \frac{\partial x_{\alpha,i}}{\partial q_j}\frac{\partial x_{\alpha,i}}{\partial q_k}, \quad b_j = \sum_{\alpha,i=1}^{n} m_\alpha \frac{\partial x_{\alpha,i}}{\partial q_j}\frac{\partial x_{\alpha,i}}{\partial t}, \quad c = \sum_{\alpha,i=1}^{n} \frac{1}{2}m_\alpha \left[\frac{\partial x_{\alpha,i}}{\partial t}\right]^2.$$

This substitution for a_{jk}, b_j, and c is a more general consequence for any system of particles. In fact, it is not always true that a_{jk}, b_j, and c are equal to the above.

We are particularly interested in the case where the transformation $x_{\alpha,i} = x_{\alpha,i}(\{q_j\})$. That is, the transformation is *scleronomic* and thus independent of time. In this case, the terms in b_j and c go to zero, since $\partial x_{\alpha,i}/\partial t = 0$, to give

$$T = \sum_{j,k}^{n} a_{jk}\dot{q}_j\dot{q}_k.$$

Computing the partial of T with respect to some generalized velocity $\dot{q}_l$, we find

$$\frac{\partial T}{\partial \dot{q}_l} = \sum_{k=1}^{n} a_{lk}\dot{q}_k + \sum_{j=1}^{n} a_{jl}\dot{q}_j$$

and we multiply through by $\dot{q}_l$ and sum over l to obtain

$$\sum_{l=1}^{n} \dot{q}_l \frac{\partial T}{\partial \dot{q}_l} = \sum_{k,l=1}^{n} a_{lk}\dot{q}_k\dot{q}_l + \sum_{j,l=1}^{n} a_{jl}\dot{q}_j\dot{q}_l = 2\sum_{k,l=1}^{n} a_{lk}\dot{q}_k\dot{q}_l = 2T$$

where we have used the fact that the indices were dummy indices. We thus arrive at the general relation,

$$\boxed{\sum_{l=1}^{n} \dot{q}_l \frac{\partial T}{\partial \dot{q}_l} = 2\sum_{k,l=1}^{n} a_{lk}\dot{q}_k\dot{q}_l = 2T.} \tag{22}$$

Although it seems rather sudden and hard to see, this theorem becomes particularly useful in our upcoming discussion on conservation laws.

2.7 Conservation Laws

As we would expect in Newtonian mechanics, there are also dynamical quantities that are conserved under the Lagrangian formalism. The most intuitive examples are conservation of momentum, or conservation of energy. We can identify these quantities by taking note of some Lagrangian-inherent properties.

2.7.1 Cyclic Coordinates

We begin with the ELE,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i}.$$

Notice, if the Lagrangian is independent of a coordinate, q_i. That is, $\mathcal{L} = \mathcal{L}(\dot{q}_i, t)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = 0 \implies \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = p_i$$

for some constant p_i. In this way, we say that the momentum, p_i, which is canonically conjugate to the coordinate q_i, is conserved. Generalized coordinates of this kind are **cyclic**.

Let us illustrate this with a free particle in 1D example subjected to zero potential. Here, our Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2$$

which shows that x is cyclic in $\mathcal{L}$. Thus, we expect the canonical momentum, p_x, to be conserved

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$$

which turns out to be the linear momentum in x. A slightly more complicated example may be a particle subjected to a central potential, $V(r, \theta) = V(r)$ in 2D. Here, we have

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r)$$

where the coordinate, θ, is cyclic instead. Thus, we expect p_θ to be conserved, such that

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta} = r(mr\dot{\theta}) = L_z$$

where L turns out to be the angular momentum along the z-direction.

2.7.2 Time-independent Lagrangians

Let us recall that Beltrami's identity applies to functionals of the form $f = f(y, y')$. The corresponding equivalence in Lagrangian mechanics are Lagrangians of the form $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i)$. We thus expected that

$$\dot{q}_i\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} = \alpha$$

for some constant α. You may notice that Beltrami's identity is flipped here. Particularly, we pulled out a negative sign to flip the order of the terms in this conserved quantity.

Now, we must ask: what is this quantity? We already know that if a time-independent transformation takes us from $x_i \to q_i$, then (22) holds. Furthermore, if our potential term is only dependent on the generalized coordinate, q_i. That is, $V = V(q_i)$, then we have that

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_l} = \frac{\partial T}{\partial \dot{q}_l} - \cancel{\frac{\partial V(q_i)}{\partial \dot{q}_l}} \implies \frac{\partial \mathcal{L}}{\partial \dot{q}_l} = \frac{\partial T}{\partial \dot{q}_l}$$

where we have switched from $i \to l$ to to keep consistency with (22). If both of these conditions are met, we can show that

$$\dot{q}_l \frac{\partial \mathcal{L}}{\partial \dot{q}_l} - \mathcal{L} = \dot{q}_l \frac{\partial T}{\partial \dot{q}_l} - (T - V).$$

Invoking (22),

$$\dot{q}_l \frac{\partial T}{\partial \dot{q}_l} - (T - V) = 2T - T + V = T + V = E$$

which is the total energy contained within the system. There are, in fact, three conditions which *always* lead to this result. They are:

1. The Lagrangian is time-independent. That is, $\mathcal{L} = \mathcal{L}(\{q_i\}, \{\dot{q}_i\})$.
2. The set of generalized coordinates, $\{q_i\}$, are related to the Cartesian coordinates by a time-independent (rheonomic) transformation.
3. The potential energy is a function of only the generalized coordinates. That is, $V = V(\{q_i\})$.

We switch from a single generalized coordinate to the configuration space interpretation to consider *all* possible generalized coordinates. Plus it's clearer.

We should make an important observation here. If point (1.) holds, then it is true that the quantity

$$\dot{q}_l \frac{\partial \mathcal{L}}{\partial \dot{q}_l} - \mathcal{L}$$

is conserved. However, if the coordinate transformation $x_l \Leftrightarrow q_l$ is *no longer* rheonomic, then it is **not** true that it is the total energy:

$$\dot{q}_l \frac{\partial \mathcal{L}}{\partial \dot{q}_l} - \mathcal{L} \neq T + V = E.$$

In particular, it is a consequence of the fact that

$$\dot{q}_l \frac{\partial T}{\partial \dot{q}_l} \neq 2T.$$

2.7.3 Multiple Degrees of Freedom

Let us now suppose that we have a Lagrangian which is time-independent, but has multiple degrees of freedom. This Lagrangian would be of the form

$$\mathcal{L} = \mathcal{L}(\{q_i\}, \{\dot{q}_i\}).$$

If we take the time derivative of the Lagrangian,

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial q_i} \dot{q}_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \ddot{q}_i + \cancel{\frac{\partial \mathcal{L}}{\partial t}}.$$

The first term is simply the ELE, and the last term cancels since the Lagrangian is time-independent. Thus, we find that

$$\begin{aligned}\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\dot{q}_i\right] + \sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{\mathrm{d}\dot{q}_i}{\mathrm{d}t}\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\dot{q}_i\right]\end{aligned}$$

where we have reversed the chain rule. Rearranging this result, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\dot{q}_i - \mathcal{L}\right] &= 0\\ \sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\dot{q}_i - \mathcal{L} &= \mathcal{H}\end{aligned}$$

for some constant $\mathcal{H}$. This quantity, if you are observant, is so important that it is called the **Hamiltonian**; and, conventionally, is written in terms of the conjugate momenta, which is $p_i = \partial\mathcal{L}/\partial q_i$.

$$\mathcal{H} = \sum_{i=1}^{n} p_i\dot{q}_i - \mathcal{L}.$$

We do not box this result since it is discussed later. Notice the relation of this quantity to that of the conditions discussed in Chapter 2.7.2.

1. If the Lagrangian $\mathcal{L} \propto t$, then $\mathcal{H}$ is no longer conserved.
2. If the potential, $V = V(\dot{q}_i)$ or the transformation is non-rheonomic. That is, $x_i \Leftrightarrow q_i$, then $\mathcal{H} \neq T + V = E$.

2.8 Noether's Theorem

This next topic should really be in the previous subsection, but since this theorem is so important, we will differentiate it from the others.

The previous conservation laws, in fact, were special cases of a theorem called **Noether's Theorem**, which states that *"for any continuous symmetry in a physical system, there is a conserved quantity."* Here, symmetries refer to invariance of a system under certain transformations. We can split these symmetries into two types.

Continuous symmetries correspond to transformations which can be implemented by a sequence of infinitesimal steps. For example, a space translation transformation or a rotation transformation.

On the other hand, **discrete symmetries** cannot be implemented by a sequence of infinitesimally small steps. For example, reflection symmetries.

Let us now consider a Lagrangian of the form $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i, t)$ to see how Noether's theorem gives rise to a conserved quantity. We furthermore assume that we have already found a solution to the ELE as a classical trajectory, given by $q_i = q_i(t)$.

Now, let us consider small perturbations about this classical trajectory. In particular we can write these perturbations in the form

$$
\begin{aligned}
q_i' &= q_i + \varepsilon K_i(q_i, \dot{q}_i, t) \\
t' &= t + \varepsilon\theta
\end{aligned}
$$

where ε is an infinitesimal constant, θ is some arbitrary constant, and $K_i(q_i, \dot{q}_i, t)$ is some arbitrary function that depends on the generalized coordinates, velocities and time.

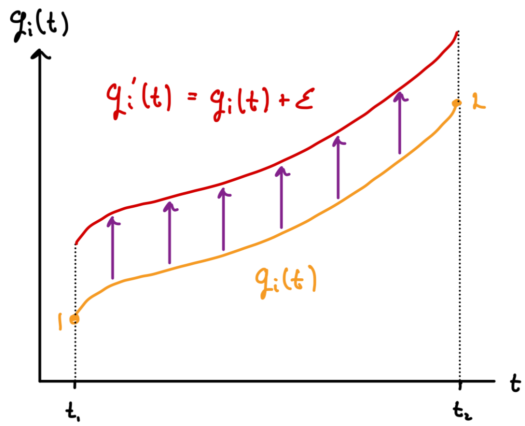


Figure 11: Trajectory, $q_i(t)$, subject to the small perturbation, ε.

For the transformations to be considered a continuous symmetry, we must have that the original Lagrangian obeys a Taylor expansion about the primed quantities, with the expansion order governed by the constant ε.

$$
\mathcal{L}(q_i, \dot{q}_i, t) = \mathcal{L}(q_i', \dot{q}_i', t') + \varepsilon \frac{\mathrm{d}f(q_i', t')}{\mathrm{d}t'} + \mathcal{O}\big(\varepsilon^2\big).
$$

Notice, $f = f(q_i', t')$ is an arbitrary function which does not involve the generalized, primed velocities, and is subject to a full time derivative with respect to t'. Let us evaluate the first two terms. In particular, we have that

$$
\begin{aligned}
\mathcal{L}(q_i', \dot{q}_i', t') + \varepsilon \frac{\mathrm{d}f(q_i', t')}{\mathrm{d}t'} &= \mathcal{L}(q_i + \varepsilon K_i, \dot{q}_i + \varepsilon \dot{K}_i, t + \varepsilon\theta) + \varepsilon \frac{\mathrm{d}}{\mathrm{d}t}[f(q_i + \varepsilon K_i, t + \varepsilon\theta)] \cancel{\frac{\mathrm{d}t}{\mathrm{d}t'}} \\
&= \mathcal{L}(q_i, \dot{q}_i, t) + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial q_i} \varepsilon K_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \varepsilon \dot{K}_i + \frac{\partial \mathcal{L}}{\partial t} \varepsilon\theta + \varepsilon \frac{\mathrm{d}f(q_i, t)}{\mathrm{d}t} + \mathcal{O}\big(\varepsilon^2\big) \\
&= \mathcal{L}(q_i, \dot{q}_i, t) + S_1 + S_2 + S_3 + \varepsilon \frac{\mathrm{d}f(q_i, t)}{\mathrm{d}t} + \mathcal{O}\big(\varepsilon^2\big)
\end{aligned}
$$

where we have denoted the second, third, and fourth terms by S_1, S_2, and S_3 respectively. We shall approach S_1 and S_2 together by considering that S_2 is a series expansion, and by letting $p_i = \partial \mathcal{L}/\partial \dot{q}_i$. Then,

$$
S_1 + S_2 = \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial q_i} \varepsilon K_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \varepsilon \dot{K}_i = \sum_{i=1}^{n} \left[\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right] \varepsilon K_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \varepsilon \dot{K}_i.
$$

The first term can be substituted for the ELE, and by the reverse chain rule,

$$S_1+S_2=\sum_{i=1}^{n}\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right]\varepsilon K_i+\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\varepsilon\dot{K}_i=\sum_{i=1}^{n}\left[\frac{\mathrm{d}p_i}{\mathrm{d}t}\right]\varepsilon K_i+\sum_{i=1}^{n}p_i\varepsilon\frac{\mathrm{d}K_i}{\mathrm{d}t}=\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}p_iK_i\right].$$

We approach S_3 by manually computing the derivative, which is

$$\begin{aligned}S_3=\frac{\partial\mathcal{L}}{\partial t}&=\left[\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t}-\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial q_i}\dot{q}_i-\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\ddot{q}_i\right]\theta\\&=\left[\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t}-\sum_{i=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right)\dot{q}_i-\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{\mathrm{d}\dot{q}_i}{\mathrm{d}t}\right]\theta\\&=\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathcal{L}-\sum_{i=1}^{n}p_i\dot{q}_i\right]\theta\end{aligned}$$

where we have, again, used the reverse chain rule. Now, returning back to our original derivation, we find that

$$\begin{aligned}\mathcal{L}(q_i',\dot{q}_i',t')+\varepsilon\frac{\mathrm{d}f(q_i',t')}{\mathrm{d}t'}&=\mathcal{L}(q_i,\dot{q}_i,t)+\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}p_iK_i\right]+\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathcal{L}-\sum_{i=1}^{n}p_i\dot{q}_i\right]\theta+\varepsilon\frac{\mathrm{d}f(q_i,t)}{\mathrm{d}t}+\mathcal{O}\big(\varepsilon^2\big)\\&=\mathcal{L}(q_i,\dot{q}_i,t)+\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}p_iK_i+\left(\mathcal{L}-\sum_{i=1}^{n}p_i\dot{q}_i\right)\theta+f(q_i,t)\right]+\mathcal{O}\big(\varepsilon^2\big).\end{aligned}$$

We can ignore the second order corrections in ε, so we arrive at

$$\mathcal{L}(q_i,\dot{q}_i,t)=\mathcal{L}(q_i,\dot{q}_i,t)+\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n}p_iK_i+\left(\mathcal{L}-\sum_{i=1}^{n}p_i\dot{q}_i\right)\theta+f(q_i,t)\right].$$

For this to be true, we then find that the quantity within the derivative must be constant. Thus, we have that $f(q,t)=f(q)$, and we expect a conserved quantity of the form

$$\boxed{P=\sum_{i=1}^{n}p_iK_i+\left(\mathcal{L}-\sum_{i=1}^{n}p_i\dot{q}_i\right)\theta+f(q_i)}\tag{23}$$

which must be time-independent, and therefore conserved. Let us quickly discuss some connections to our conservation laws we had before.

Previously, we saw that a cyclic coordinate required that the Lagrangian did not depend on that coordinate. That is, $\mathcal{L}=\mathcal{L}(\dot{q}_i,t)$. In this sense, the coordinate, q_i, is unchanged under a transformation of the form

$$q_i'=q_i+\varepsilon K$$

for any constant K. Thus, p_iK is conserved, and is a constant multiple of the corresponding generalized momentum.

For a time-independent Lagrangian, we have that the time is invariant under a transformation of the form

$$t'=t+\varepsilon\theta$$

under any path. Correspondingly, we find that

$$\mathcal{L} - \sum_{i=1}^{n} p_i \dot{q}_i$$

is conserved, and is the Hamiltonian as before.

In a more complicated, 2D example, we may have a Lagrangian of the form

$$\mathcal{L} = \frac{1}{2} m |\dot{\mathbf{x}}_1|^2 + \frac{1}{2} m |\dot{\mathbf{x}}_2|^2 - V(|\mathbf{x}_1 - \mathbf{x}_2|).$$

Notice that this Lagrangian is, in general, **not** invariant under a single space translation. However, it **is** invariant if both positions are displaced by the same amount. That is, the generalized coordinates are transformed by

$$\begin{aligned} x_1' &= x_1 + \varepsilon K \\ x_2' &= x_2 + \varepsilon K. \end{aligned}$$

Now, we expect that the quantity

$$\sum_{i=1}^{2} p_i K_i = (p_{x,1} + p_{x,2}) K$$

is conserved, for some constant K. In particular, the quantity within the parentheses is the total momentum along the direction x, and we can find that the total momentum along y and z are also conserved and of the same form.

3 Examples in Lagrangian Mechanics

We now begin our discussion on certain types of mechanics problems that can be solved with Lagrangians. This chapter is dedicated solely on solving examples and techniques to investigate properties of mechanical systems. We will also box the solutions to the examples, but note that for this chapter specifically, they are not generally important formulas which you *need* to know. They are just derivations that you should end up getting, should you do the problem yourself.

We begin with the simple Atwood machine.

3.1 The Atwood Machine

3.1.1 The Inertial Pulley

Consider an Atwood machine characterized by a fixed pulley and an ideal string. The pulley is attached to a ceiling, and two masses, m_1 and m_2, are attached by the string to the pulley itself. Both masses are free to move in the vertical direction. Determine the EoM of the system and any forces of constraint.

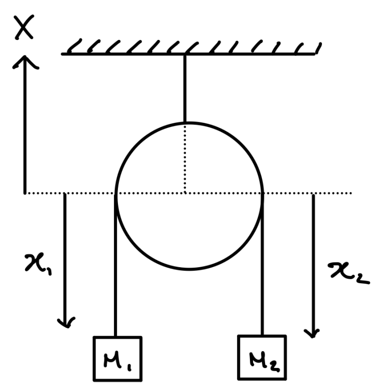


Figure 12: Atwood machine of two masses.

We note that there are two methods of solving this problem. We could (a) embed the constraint into the system, by which we find the EoM, or (b), use the method of LUM to determine constraining forces and the EoM.

Because the problem asks for the forces of constraints, the method of LUM would be "better" in this scenario, but we will illustrate both methods to become familiar with the techniques.

Embedding the Constraint. Starting with method (a), we write down the energies of the system, which are

$$T = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2$$
$$V = -m_1gx_1 - m_2gx_2$$

Then, the Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2 + m_1gx_1 + m_2gx_2.$$

The constraint on the system is such that conservation of string is equal to the distance both masses can travel. Thus,

$$g(x_1, x_2) = x_1 + x_2 - L = 0.$$

We then find that

$$x_2 = L - x_1 \implies \dot{x}_2 = -\dot{x}_1$$

to which we can embed into the Lagrangian to find

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\dot{x}_1^2 + m_1gx_1 + m_2g(L - x_1) = \frac{1}{2}(m_1 + m_2)\dot{x}_1^2 + (m_1 - m_2)gx_1$$

which is a Lagrangian which depends only on the coordinate, x_1. Thus, we can find the EoM by computing the partials in $\mathcal{L}$, which are

$$\frac{\partial\mathcal{L}}{\partial\dot{x}_1} = (m_1 + m_2)\dot{x}_1 \implies \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} = (m_1 + m_2)\ddot{x}_1$$
$$\frac{\partial\mathcal{L}}{\partial x_1} = (m_1 - m_2)g$$

and by the ELE, we have

$$(m_1 + m_2)\ddot{x}_1 = (m_1 - m_2)g \implies \ddot{x}_1 = \frac{m_1 - m_2}{m_1 + m_2}g.$$

Notice that the constraint also imposes that $\ddot{x}_1 = -\ddot{x}_2$, so we have the full relation

$$\boxed{\ddot{x}_1 = \frac{m_1 - m_2}{m_1 + m_2}g = -\ddot{x}_2.}$$

Notice that this method never actually *tells* us what the force of constraint is. That is, the tension force. Because we embed the constraint into the Lagrangian, we effectively ”lose” this information, as we implicitly incorporate it into the EoM.

Method of LUM. Recall the Lagrangian we found previously (the setup is the same) as well as the holonomic constraint.

$$\mathcal{L} = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2 + m_1gx_1 + m_2gx_2$$
$$g(x_1, x_2) = x_1 + x_2 - L = 0.$$

Together, we have a system of two equations which we can solve by using the method of LUM we found before. Here, the constrained ELE in x_1 is

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} = \frac{\partial\mathcal{L}}{\partial x_1} + \lambda(t)\frac{\partial g}{\partial x_1} \implies m\ddot{x}_1 = m_1g + \lambda(t)$$

and the constrained ELE in x_2 is

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_2} = \frac{\partial\mathcal{L}}{\partial x_2} + \lambda(t)\frac{\partial g}{\partial x_2} \implies m\ddot{x}_2 = m_2g + \lambda(t).$$

The constraint, if you may recall, shows that $\ddot{x}_1 = -\ddot{x}_2$, so solving the system of three equations, we find that

$$\begin{aligned} m_1\ddot{x}_1 - m_1 g &= m_2\ddot{x}_2 - m_2 g \\ m_1\ddot{x}_1 - m_1 g &= -m_2\ddot{x}_2 - m_2 g \\ (m_1 + m_2)\ddot{x}_1 &= (m_1 - m_2)g \\ \ddot{x}_1 &= \frac{m_1 - m_2}{m_1 + m_2} g = -\ddot{x}_2 \end{aligned}$$

which is what we found before. However, we can now define function $\lambda(t)$, which is

$$\lambda(t) = m_1\ddot{x}_1 - m_1 g = m_1\left[\frac{m_1 - m_2}{m_1 + m_2} g\right] - m_1 g = \frac{\cancel{m_1^2} - m_1 m_2 - \cancel{m_1^2} - m_1 m_2}{m_1 + m_2} g = -2\frac{m_1 m_2}{m_1 + m_2} g.$$

Therefore, the force of constraint is

$$\boxed{\lambda(t)\frac{\partial g}{\partial x_1} = \frac{-2m_1 m_2}{m_1 + m_2} g}$$

since $\partial g/\partial x_1 = 1$. Here, notice that the force is in the negative direction, in other words, down from the pulley. This is because the force that the mass must exert on the string itself is directed downwards, which, in Newtonian mechanics, would be the reactive force of the tension force exerted by the string. This holds for the other mass as well.

3.1.2 The Accelerating Pulley

We will now consider a more complicated version of the Atwood machine. Suppose we have the same setup as before, but with the added condition that the Atwood machine is within an elevator, accelerating down with an acceleration condition $\alpha < g$.

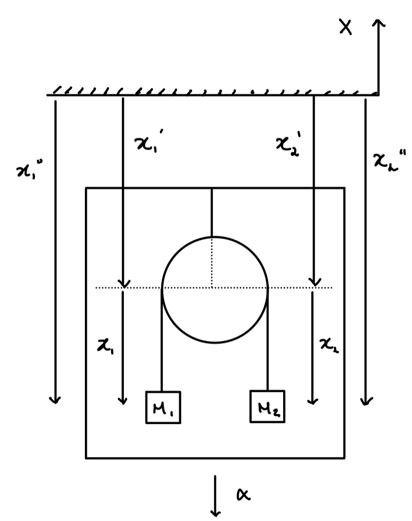


Figure 13: Atwood machine within an accelerating frame of acceleration $\alpha < g$.

In this case, we have that the generalized coordinates, x_1 and x_2 are non-inertial. It is still possible to use them, however, we must be careful with how we define our coordinates. In particular, we must define our energies **relative** to an inertial reference frame.

In this case, if x_1 and x_2 are the non-inertial vertical distances the mass can travel, we must define the distances from the ceiling to x_1 and x_2. This, effectively, adjusts for the non-inertial component of our system. Let us define x_1' and x_2' to be the distances from the ceiling to x_1 and x_2, and x_1'' and x_2'' to be the total vertical distance, such that

$$\begin{aligned} x_1'' &= x_1' + x_1 = x_0 + x_1 + \frac{1}{2}\alpha t^2 \\ x_2'' &= x_2' + x_2 = x_0 + x_2 + \frac{1}{2}\alpha t^2 \end{aligned} \implies \begin{aligned} (\dot{x}_1'') &= \alpha t + \dot{x}_1 \\ (\dot{x}_2'') &= \alpha t + \dot{x}_2 \end{aligned}$$

Notice we have used the fact that $x_1' = x_0 + \alpha t^2/2$, where x_0 is some arbitrary initial position (constant). The same goes for x_2'. Thus, the energies in the system are

$$T = \frac{1}{2}m_1(\dot{x}_1'')^2 + \frac{1}{2}m_2(\dot{x}_2'')^2 = \frac{1}{2}m_1(\dot{x}_1^2 + 2\alpha t\dot{x}_1 + \alpha^2t^2) + \frac{1}{2}m_2(\dot{x}_2^2 + 2\alpha t\dot{x}_2 + \alpha^2t^2)$$

$$V = -m_1 g x_1'' - m_2 g x_2'' = -m_1 g(x_0 + x_1 + \frac{1}{2}\alpha t^2) - m_2 g(x_0 + x_2 + \frac{1}{2}\alpha t^2)$$

to give the Lagrangian

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}m_1\left(\dot{x}_1^2 + 2\alpha t\dot{x}_1 + \cancel{\alpha^2t^2}\right) + \frac{1}{2}m_2\left(\dot{x}_2^2 + 2\alpha t\dot{x}_2 + \cancel{\alpha^2t^2}\right) \\ &\qquad + m_1 g\left(\cancel{x_0} + x_1 + \frac{1}{2}\alpha t^2\right) + m_2 g\left(\cancel{x_0} + x_2 + \frac{1}{2}\alpha t^2\right) \\ &= \frac{1}{2}m_1\left(\dot{x}_1^2 + 2\alpha t\dot{x}_1\right) + \frac{1}{2}m_2\left(\dot{x}_2^2 + 2\alpha t\dot{x}_2\right) + m_1 g x_1 + m_2 g x_2. \end{aligned}$$

where we have thrown out non-impacting terms in the Lagrangian. We proceed with incorporating the constraint, so find that $x_2 = L - x_1 \implies \dot{x}_2 = -\dot{x}_1 \implies \ddot{x}_2 = -\ddot{x}_1$ as before, so

$$\mathcal{L} = \frac{1}{2}m_1\left(\dot{x}_1^2 + 2\alpha t\dot{x}_1\right) + \frac{1}{2}m_2\left(\dot{x}_1^2 - 2\alpha t\dot{x}_1\right) + m_1 g x_1 - m_2 g x_1.$$

Let us make a note here: the Lagrangian has explicit time dependence. That is, $\mathcal{L} \propto t$ in this scenario. We expect, then, that the total energy within the system is not conserved. Regardless, computing the partials in x_1, we find

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{x}_1} &= m_1\alpha t + m_1\dot{x}_1 - m_2\alpha t + m_2\dot{x}_1 \\ &= (m_1 - m_2)\alpha t + (m_1 + m_2)\dot{x}_1 \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_1} &= (m_1 - m_2)\alpha + (m_1 + m_2)\ddot{x}_1 \\ \frac{\partial \mathcal{L}}{\partial x_1} &= (m_1 - m_2)g. \end{aligned}$$

Thus, the ELE reads

$$(m_1 + m_2)\ddot{x}_1 + (m_1 - m_2)\alpha = (m_1 - m_2)g \tag{24}$$

and simplifying, the final result is

$$\boxed{\ddot{x}_1 = \frac{m_1 - m_2}{m_1 + m_2}(g - \alpha) = -\ddot{x}_2.}$$

This result is rather surprising. Contrary to what one would expect—that the acceleration would remain unchanged—the accelerations in the masses are *reduced* by the factor of α. We call this the **"effective g"**. In fact, this result is a consequence of the *translational force* that is present within non-inertial reference frames. It is a fictitious force which affects the force observed in an inertial reference frame.

3.2 Relative Plane Motion

3.2.1 Ball Rolling Without Slipping on an Inclined Plane

Let us now consider a ball, situated on an inclined plane with mass, m, and moment of inertia, I, rolling without slipping down an inclined plane making an angle, θ, with the horizontal. Determine the EoM of the ball.

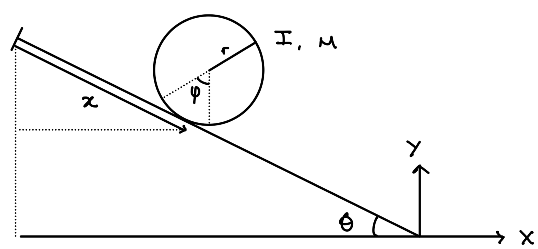


Figure 14: Ball rolling without slipping on an inclined plane making an angle, θ.

We begin by choosing x and ϕ to be our generalized coordinates in the system, where x is the distance travelled along the incline, and ϕ is the angular rotation of the ball. In terms of the inertial Cartesian coordinates, our generalized coordinates can be expressed as

$$\begin{aligned} X &= x_0 + x\cos(\theta) \\ Y &= y_0 - x\sin(\theta) \end{aligned} \implies \begin{aligned} \dot{X} &= \dot{x}\cos(\theta) \\ \dot{Y} &= -\dot{x}\sin(\theta) \end{aligned} \implies \dot{X}^2 + \dot{Y}^2 = \dot{x}^2.$$

Then, the energies in the system are

$$T = T_{\text{trans.}} + T_{\text{rot.}} = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\phi}^2$$
$$V = -mg\sin(\theta)x.$$

We make a note here: the rotational energy term, $T_{\text{rot.}}$ is *only* valid when the rotation is about a principal axis. In this case, the principal axis is $-\hat{k}$, which points into the page. By the right hand rule (RHR), the ball's rotation is directed in the $-\hat{k}$ direction, satisfying our condition on $T_{\text{rot.}}$.

Our Lagrangian is then

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\phi}^2 + mg\sin(\theta)x.$$

From the setup of our problem, we also know that there is a constraint in the system, which imposes that the ball must always make contact with the inclined plane. The constraint function is of the form

$$g(x, \phi) = x - r\phi = 0$$

where r is the radius of the ball, and furthermore implies that $\dot{x} = r\dot{\phi}$. Using this relation, we incorporate the constraint into the Lagrangian, giving

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\frac{\dot{x}^2}{r^2} + mg\sin(\theta)x.$$

We only need to compute the partials in $\mathcal{L}$ for x, which are

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot{x}} &= \left(m + \frac{I}{r^2}\right)\dot{x} \\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} &= \left(m + \frac{I}{r^2}\right)\ddot{x} \\
\frac{\partial \mathcal{L}}{\partial x} &= mg\sin(\theta)
\end{aligned}$$

and thus the ELE is

$$\left(m + \frac{I}{r^2}\right)\ddot{x} = mg\sin(\theta).$$

Rearranging this result,

$$\boxed{\ddot{x} = \frac{mgr^2\sin(\theta)}{mr^2 + I}.}$$

Again, notice that by incorporating the constraint, we are not able to determine the frictional force that the plane exerts on the ball. We could have found this by the method of LUM, but we will not go through it here.

3.2.2 Block Rolling Without Slipping on a Moving Inclined Plane

Consider now a similar setup as before, but the inclined plane of mass, M, is free to move horizontally along the Cartesian, inertial X-direction.

This next part is important. You may be tempted to write down the energies of the individual bodies as before, where

$$\begin{aligned}
T_{\text{plane}} &= \frac{1}{2}M\dot{x}_1^2 & \quad\text{and}\quad T_{\text{block}} &= \frac{1}{2}m\dot{x}_2^2 \\
V_{\text{plane}} &= 0 & V_{\text{block}} &= -mg\sin(\theta)x_2
\end{aligned}$$

and the energies in the block are analogous to the stationary inclined plane situation. In this way, we have the Lagrangian,

$$\mathcal{L} = \frac{1}{2}M\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 + mg\sin(\theta)x_2.$$

This approach, however, is **wrong**. To see why, we must look at what our Lagrangian is really stating about our system.

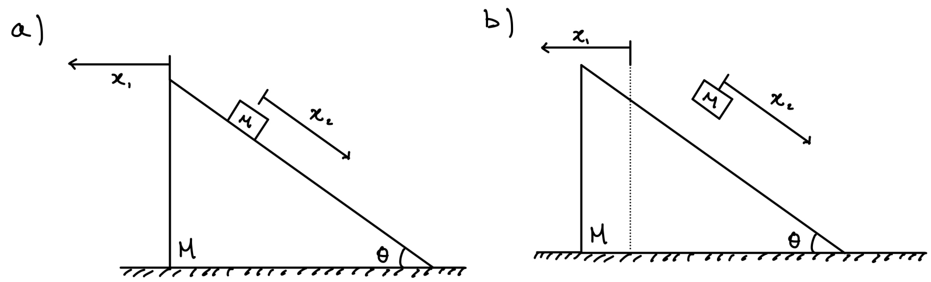


Figure 15: Isolated dynamics of the inclined plane problem using non-coupled energy terms.

Let's focus first on the block. It's true that if we were to release the block from rest, there is a constraint which enforces that the block must stay attached to the surface of the plane. **However**, let us now imagine we move the inclined plane instead and release the block from rest. This is where the issue arises. Our constraint, which would be of the form

$$g(x_2, y_2) = y_2 - x_2 \tan(\theta) = 0$$

assuming that y_2 is directed downwards, does **not** encode that the inclined plane is moving along the inertial Cartesian X-direction. This, in turn, says that when the inclined plane is moving, the block travels downward **as if** the inclined plane were stuck at its initial position.

This is obviously wrong. How, then, do we fix this issue? There are two methods, which we will call *absolute coordinates* and *relative coordinates*. The first attempts to solve the problem by setting **absolute coordinates**, which the bodies in the system move relative to, while the second attempts to solve the problem by setting **relative coordinates** with respect to each body. We will show both methods, though the second is preferred.

Absolute Coordinates. Notice that if the plane is fixed, and we move the block by some horizontal distance, x_2, it must drop by some vertical distance, y_2. These two coordinates are related by our original constraint, which is

$$y_2 = x_2 \tan(\theta).$$

Now, suppose that leave the position of the block to itself and instead move the plane. If we were to impose that the block remains on the plane, we should expect that the block must drop by some vertical distance

$$y_2 = x_1 \tan(\theta)$$

which actually should arise from a constraint of the form $g(x_1, y_2) = y_2 - x_1 \tan(\theta) = 0$. Now, for the system to have the proper dynamics, these constraints must both be active, and we can define an *effective* constraint which includes both possibilities. This new, effective constraint would be of the form

$$g(x_1, x_2, y_2) = y_2 - (x_1 + x_2) \tan(\theta) = 0.$$

From here, the process is the same as before. We define the energy terms, which are

$$\begin{aligned} T_{\text{plane}} &= \frac{1}{2} M \dot{x}_1^2 \\ V_{\text{plane}} &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} T_{\text{block}} &= \frac{1}{2} m(\dot{x}_2^2 + \dot{y}_2^2) = \frac{1}{2} m\left[\dot{x}_2^2 + (\dot{x}_1 + \dot{x}_2)^2 \tan^2(\theta)\right] \\ V_{\text{block}} &= -mgy_2 = -mg(x_1 + x_2) \tan(\theta) \end{aligned}$$

and find that the Lagrangian is

$$\mathcal{L} = \frac{1}{2}M\dot{x}_1^2 + \frac{1}{2}m\big[\dot{x}_2^2 + (\dot{x}_1 + \dot{x}_2)^2\tan^2(\theta)\big] + mg(x_1 + x_2)\tan(\theta).$$

Simplifying this further, we find that

$$(\dot{x}_1 + \dot{x}_2)^2 = \dot{x}_1^2 + 2\dot{x}_1\dot{x}_2 + \dot{x}_2^2$$

such that

$$\mathcal{L} = \frac{1}{2}\big[M + m\tan^2(\theta)\big]\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2\sec^2(\theta) + m\dot{x}_1\dot{x}_2\tan^2(\theta) + mg(x_1 + x_2)\tan(\theta).$$

Notice now that this constraint enforces a *coupling factor* of $2\dot{x}_1\dot{x}_2$. This encodes that the inclined plane and the mass are acting together and not separately. It now remains to compute the partials in both directions. We have in x_1,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} &= \big[M + m\tan^2(\theta)\big]\ddot{x}_1 + m\ddot{x}_2\tan^2(\theta)\\ \frac{\partial\mathcal{L}}{\partial x_1} &= mg\tan(\theta)\end{aligned}$$

and in x_2,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_2} &= m\ddot{x}_2\sec^2(\theta) + m\ddot{x}_1\tan^2(\theta)\\ \frac{\partial\mathcal{L}}{\partial x_2} &= mg\tan(\theta).\end{aligned}$$

Then, our ELE's are

$$\begin{aligned}\big[M + m\tan^2(\theta)\big]\ddot{x}_1 + m\ddot{x}_2\tan^2(\theta) &= mg\tan(\theta)\\ m\ddot{x}_2\sec^2(\theta) + m\ddot{x}_1\tan^2(\theta) &= mg\tan(\theta)\end{aligned}$$

which is a system of two equations. We can isolate for the acceleration, $\ddot{x}_2$ to find that

$$M\ddot{x}_1 - m\ddot{x}_2 = 0 \implies \ddot{x}_2 = \frac{M}{m}\ddot{x}_1$$

and so we simplify the first ELE to

$$\begin{aligned}\big[M + m\tan^2(\theta)\big]\ddot{x}_1 + \cancel{m}\left[\frac{M}{\cancel{m}}\right]\ddot{x}_1\tan^2(\theta) &= mg\tan(\theta)\\ \ddot{x}_1\big[M + (M + m)\tan^2(\theta)\big] &= mg\tan(\theta)\\ \ddot{x}_1 &= \frac{mg\tan(\theta)}{M + (M + m)\tan^2(\theta)}\\ &= \frac{mg\sin(\theta)\cos(\theta)}{M + m\sin^2(\theta)}.\end{aligned}$$

By our previous constraint, the full relation thus becomes

$$\boxed{\ddot{x}_2 = \frac{Mg\sin(\theta)\cos(\theta)}{M + m\sin^2(\theta)} = \frac{M}{m}\ddot{x}_1, \quad \ddot{x}_1 = \frac{mg\sin(\theta)\cos(\theta)}{M + m\sin^2(\theta)}.}$$

Relative Coordinates. Rather than having to determine two *individual* constraints and stitching them together, we can utilize our original generalized coordinates and be more mindful about the dynamics of our system. Here, let us redefine our generalized coordinates to be

$$\begin{aligned} X_{\text{plane}} &= -x_1 \\ Y_{\text{plane}} &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} X_{\text{block}} &= x_0 - x_1 + x_2\cos(\theta) \\ Y_{\text{block}} &= y_0 - x_2\sin(\theta) \end{aligned}$$

where x_0 and y_0 are arbitrary constants. We note that our Cartesian coordinate system is defined at the initial position of the plane's corner, and the block positions depend on both x_1 and x_2. Why does this work?

If we move the plane some horizontal distance, d, we expect that the block remains in contact with the inclined surface. The subscripts 1 and 2 designate the plane and the block respectively, so we expect the block to travel some horizontal distance, $x_2\cos(\theta)$, which is mitigated by $-x_1$ if the plane moves left (here, it does). The same applies for the vertical distance for the block. If the plane is situated at some arbitrary height, y_0, then we expect the block, as the plane moves, to drop by $-x_2\sin(\theta)$. Thus, we contain both dynamics in the system with each generalized coordinate.

We can now redefine our energies, which are

$$\begin{aligned} T_{\text{plane}} &= \frac{1}{2}M\dot{x}_1^2 \\ V_{\text{plane}} &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} T_{\text{block}} &= \frac{1}{2}m\left(\dot{X}_{\text{block}}^2 + \dot{Y}_{\text{block}}^2\right) = \frac{1}{2}m[-\dot{x}_1 + \dot{x}_2\cos(\theta)]^2 + \frac{1}{2}m[-\dot{x}_2\sin(\theta)]^2 \\ V_{\text{block}} &= -mgx_2\sin(\theta) \end{aligned}$$

where we have thrown out non-impacting terms. Simplifying the terms, we see that

$$[-\dot{x}_1 + \dot{x}_2\cos(\theta)]^2 = \dot{x}_1^2 - 2\dot{x}_1\dot{x}_2\cos(\theta) + \dot{x}_2^2\cos^2(\theta)$$

and so our Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}(M+m)\dot{x}_1^2 - m\dot{x}_1\dot{x}_2\cos(\theta) + \frac{1}{2}m\dot{x}_2^2 + mgx_2\sin(\theta).$$

Before we begin computing the partials, we make an important note. x_1 is a **cyclic** coordinate. Thus, we expect the conjugate momentum in x_1 to be conserved. This does **not** mean that the linear momentum, $m\dot{x}_1$, is conserved! In fact, the conjugate momentum is of the form

$$\frac{\partial\mathcal{L}}{\partial\dot{x}_1} = (M+m)\dot{x}_1 - m\dot{x}_2\cos(\theta)$$

which, in this case, is the total momentum in the Cartesian X-direction. This behaviour is common in coupled systems like these, in that the conserved quantities depend on at the very least two bodies within the system.

Regardless, computing the partials, we find that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} &= (M+m)\ddot{x}_1 - m\ddot{x}_2\cos(\theta), \quad \frac{\partial\mathcal{L}}{\partial x_1} = 0 \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_2} &= m\ddot{x}_2 - m\ddot{x}_1\cos(\theta), \quad \frac{\partial\mathcal{L}}{\partial x_2} = mg\sin(\theta) \end{aligned}$$

and so our ELE's read,

$$\begin{aligned} (M+m)\ddot{x}_1 - m\ddot{x}_2\cos(\theta) &= 0 \\ \ddot{x}_2 - \ddot{x}_1\cos(\theta) &= g\sin(\theta). \end{aligned}$$

We isolate $\ddot{x}_1$ in the first ELE to simplify the second, and find that

$$\ddot{x}_1 = \frac{m\cos(\theta)}{M+m}\ddot{x}_2.$$

Thus,

$$\ddot{x}_2 - \left[\frac{m\cos(\theta)}{M+m}\right]\ddot{x}_2\cos(\theta) = g\sin(\theta)$$

and so our acceleration in the mass is

$$\boxed{\ddot{x}_2 = \frac{(M+m)g\sin(\theta)}{M+m\sin^2(\theta)}.}$$

Notice, this does not look the same as our previous result, but if we substitute this into the relation for $\ddot{x}_1$, we get

$$\ddot{x}_1 = \frac{mg\cos(\theta)\sin(\theta)}{M+m\sin^2(\theta)}$$

which *does* agree with our previous result! Since x_1 is the same in both examples, it follows that this "new" result for $\ddot{x}_2$ is still technically valid and correct. What, then, happened here? Our new result for $\ddot{x}_2$ is *actually* the acceleration along the inclined plane. We can get our original result back if we account for the fact that the block is along an incline. In particular,

$$\ddot{X}_{\text{block}} = \ddot{x}_2\cos(\theta) - \ddot{x}_1 = \left[\frac{(M+m)g\sin(\theta)}{M+m\sin^2(\theta)}\right]\cos(\theta) - \frac{mg\cos(\theta)\sin(\theta)}{M+m\sin^2(\theta)} = \frac{Mg\cos(\theta)\sin(\theta)}{M+m\sin^2(\theta)}$$

as before.

Let us now discuss the limiting cases of our result. It is easier to see how the problem reduces to that of the inertial plane if we look at the purely Cartesian coordinates. For the case $M \gg m$, we can see that

$$\ddot{x}_2 = g\sin(\theta)\cos(\theta)$$

which is the $\hat{X}$ component of $g\sin(\theta)$ *along* the plane, and furthermore $\ddot{x}_1 \to 0$, as which is the same as our inertial scenario.

In the case $M \ll m$, we find that $\ddot{x}_1 \to g\tan(\theta)$, which is what we would expect if the mass accelerated *straight down* with an acceleration of g, and the plane was effectively null.

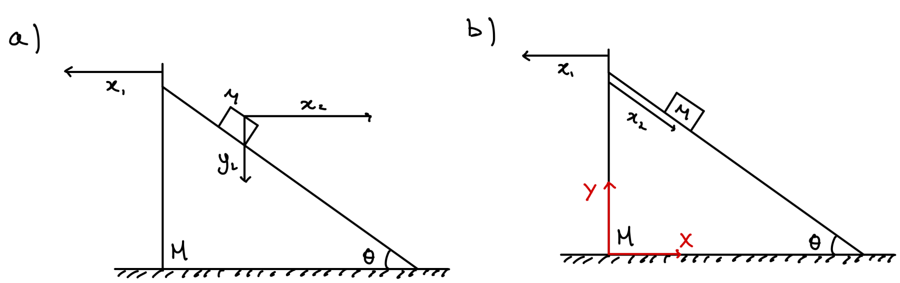


Figure 16: Two methods to solve the relative motion of the inclined plane. (a) Using absolute coordinates, where the constraint is imposed by the motion of each individual moving body. (b) Using relative coordinates, where the constraint is imposed by accounting for the relative motion between the bodies.

3.3 Orbiting Mass on a Table

Consider a system of two particles of masses m_1 and m_2, joined by a string of fixed length, L, through a hole in a frictionless table. Determine the EoM and determine when the motion of the system is stable.

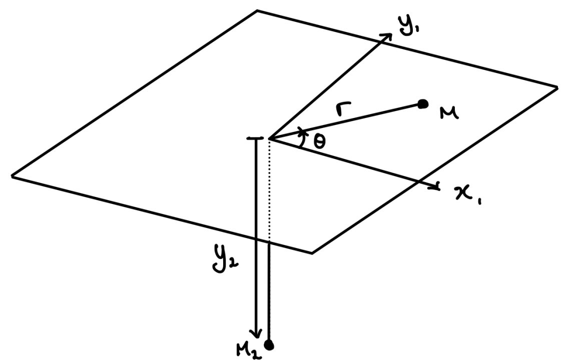


Figure 17: System of two masses, m_1 and m_2, where m_1 orbits around a table.

We begin by determining the DoF of our system. In particular, we notice that m_2 is free to move vertically, m_1 is free to move along the plane x_1 and y_1, and is free to rotate about the hole with an angle, θ. It follows that our three generalized coordinates are r, θ, and y_2, where $x_1^2 + y_1^2 = r^2$. The constraint on the system is given by the conservation of string length, and is of the form

$$g(r, y_2) = r + y_2 - L = 0.$$

Rather than computing the Cartesian energies, we will use the known result that $\dot{X}^2 + \dot{Y}^2 = \dot{R}^2 + r^2\dot{\Theta}^2$, where R and Θ are arbitrary radial and angles of the particle of interest. We find that the energies in the system are

$$T = \frac{1}{2}m_1(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{1}{2}m_2\dot{y}_2^2$$

$$V = -mgy_2$$

and so the Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}m_1(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{1}{2}m_2\dot{y}_2^2 + mgy_2.$$

We embed the constraint into the Lagrangian, such that $y_2 = L - r$ and $\dot{y}_2 = -\dot{r}$, so that

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\dot{r}^2 + \frac{1}{2}m_1 r^2\dot{\theta}^2 - m_2 gr.$$

Notice that θ is a cyclic coordinate, so we expect its corresponding generalized momentum to be conserved, such that

$$p_\theta = \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = m_1 r\dot{\theta} = l \implies \boxed{\dot{\theta} = \frac{l}{m_1 r}.}$$

It helps to notice that this quantity, $m_1 r\dot{\theta}$, is the angular momentum of the particle of mass, m_2. Then, taking the partials in r, we find that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{dt}}\frac{\partial \mathcal{L}}{\partial \dot{r}} &= m_1\ddot{r} + m_2\ddot{r}\\ \frac{\partial \mathcal{L}}{\partial r} &= m_1 r\dot{\theta}^2 - m_2 g\end{aligned}$$

and so the ELE reads

$$(m_1 + m_2)\ddot{1} = m_1 r\dot{\theta}^2 - m_2 g.$$

Using the equation for $\dot{\theta}$ to simplify this result, we arrive at

$$\boxed{\ddot{r} = \frac{l^2}{m_1(m_1+m_2)r^3} - \frac{m_2 g}{m_1+m_2}.}$$

Now, we are interested in when the motion of this system is *stable*. Stable motion occurs as a result of equilibrium points, which impose that the **second derivative of the generalized coordinates are zero**. In this case, we notice that for stable motion, $\ddot{r} = 0$, which is a consequence of circular motion, and so we find that

$$0 = \frac{l^2}{m_1 \cancel{(m_1+m_2)} r_c^3} - \frac{m_2 g}{\cancel{m_1+m_2}} \implies r_c^3 = \frac{l^2}{m_1 m_2 g}$$

where we have denoted r_c to be the radius of stable equilibrium. This condition occurs at radial equilibrium. However, we are now stuck. All we've gotten is the radius of stable equilibrium, and if we return back to our original EoM in r, we cannot proceed with further analysis since it is nonlinear. How do we go about this? The answer is **Linearization**.

3.3.1 The Method of Linearization

Linearization attempts to simplify our nonlinear EoM in to one that is, well, linear. We begin by supposing that we perturb our circular orbit. Now, we ask two questions: (1) *is the system stable or unstable?* (2) *what is the natural frequency and period?* Let us suppose that we have a perturbation of the form

$$\begin{aligned}r &= r_c + \varepsilon\\ \ddot{r} &= 0 + \ddot{\varepsilon}\end{aligned}$$

where ε is a small parameter which reflects the perturbative motion. Notice, we have implicitly imposed the condition that $\ddot{r}_c = 0$ to reflect that circular motion occurs about a stable equilibrium. Returning back to our radial equation and assuming circular motion, we find

$$\ddot{\varepsilon} = \frac{l^2}{m_1(m_1+m_2)(r_c+\varepsilon)^3} - \frac{m_2 g}{m_1+m_2}.$$

Now, here is the trick, we *Taylor expand* our nonlinear term. That is, we find that

$$\frac{1}{r_c^3(1+\varepsilon/r_c)^3} = \frac{1}{r_c^3}\left[1 - \frac{3\varepsilon}{r_c} + \mathcal{O}(\varepsilon^2)\right].$$

Ignoring second order terms since the parameter is small, we arrive at the **binomial approximation**, and find that

$$\ddot{\varepsilon} = \frac{l^2}{m_1(m_1+m_2)}\left[\frac{1}{r_c^3}\left(1-\frac{3\varepsilon}{r_c}\right)\right] - \frac{m_2 g}{m_1+m_2} = \cancel{\frac{l^2}{m_1(m_1+m_2)}\frac{1}{r_c^3} - \frac{m_2 g}{m_1+m_2}} - \frac{l^2}{m_1(m_1+m_2)}\frac{3\varepsilon}{r_c^4}$$

where we have cancelled the first two terms since we impose that it is zero for circular motion. We thus arrive at the following EoM in ε after substituting l,

$$\boxed{\ddot{\varepsilon} = -\frac{3m_2 g}{(m_1+m_2)r_c}\varepsilon.}$$

Notice that this equation is in the form of simple harmonic motion, such that we can say $\ddot{\varepsilon} \approx -\omega_0^2\varepsilon$. We can determine the natural frequency of these oscillations, which is simply

$$\boxed{\omega_0 = \sqrt{\frac{3m_2 g}{(m_1+m_2)r_c}}.}$$

Notice, that by assuming small perturbations, we have determined the stability of our equilibrium. In particular, if the EoM was $\ddot{\varepsilon} \approx \omega_0^2\varepsilon$, where the negative is missing, we would have found that the motion is unstable, and our particle would eventually tend further from our equilibrium point.

3.4 The Spring Pendulum and Autoparametric Resonance

Consider a pendulum whose rod is replaced by a spring with an equilibrium length of l, spring constant, k, and holds a mass, m, at the bob. Determine the motion of the pendulum and its properties.

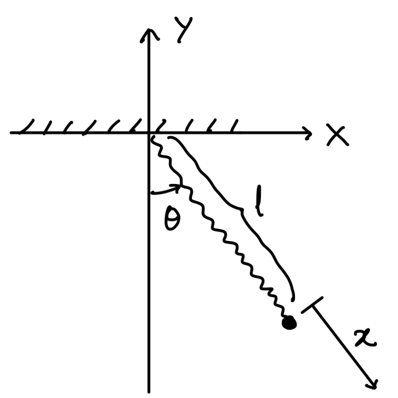


Figure 18: The spring pendulum of mass, m.

We begin by writing the positions of the mass, which are

$$\begin{aligned} X &= (l+x)\sin(\theta) \\ Y &= -(l+x)\cos(\theta) \end{aligned} \implies \begin{aligned} \dot{X} &= (l+x)\cos(\theta)\dot{\theta} + \dot{x}\sin(\theta) \\ \dot{Y} &= (l+x)\sin(\theta)\dot{\theta} - \dot{x}\cos(\theta). \end{aligned}$$

The total in the system has a contribution from the gravitational and spring energies, and thus

$$T = \frac{1}{2}m(\dot{X}^2 + \dot{Y}^2) = \frac{1}{2}\left[\dot{x}^2 + (l+x)^2\dot{\theta}^2\right]$$
$$V = mgY + \frac{1}{2}kx^2 = -mg(l+x)\cos(\theta) + \frac{1}{2}kx^2.$$

Then, the Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}\left[\dot{x}^2 + (l+x)^2\dot{\theta}^2\right] + mg(l+x)\cos(\theta) - \frac{1}{2}kx^2.$$

Then, computing the partials in x, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\ddot{x}$$
$$\frac{\partial \mathcal{L}}{\partial x} = m(l+x)\dot{\theta}^2 + mg\cos(\theta) - kx$$

and in θ,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = m(l+x)^2\ddot{\theta} + 2m(l+x)\dot{x}\dot{\theta}$$
$$\frac{\partial \mathcal{L}}{\partial \theta} = -mg(l+x)\sin(\theta).$$

Thus, the ELE gives

$$m\ddot{x} = m(l+x)\dot{\theta}^2 + mg\cos(\theta) - kx$$
$$m(l+x)^2\ddot{\theta} + 2m(l+x)\dot{x}\dot{\theta} = -mg(l+x)\sin(\theta)$$

which is a system of two equations in x and θ. We could identify the terms at this point, but this is easier to do with the Newtonian approach and trying to keep track of all forces and torques. Thus, we will instead skip to solving these equations numerically, and analyze the results.

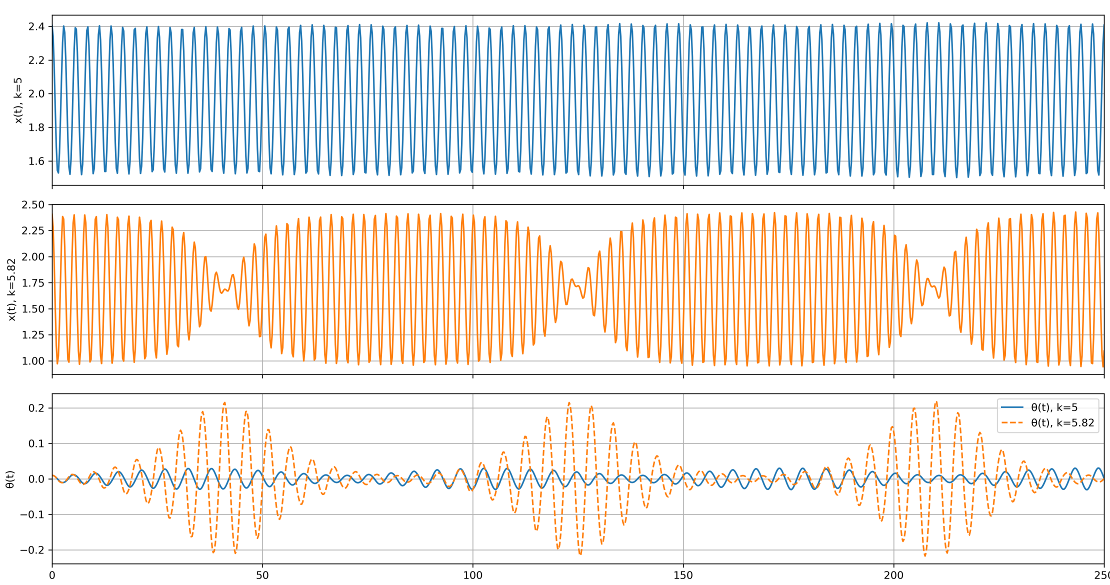


Figure 19: Plot of the spring pendulum over the range $t \in [0, 250\text{s}]$. The orange curve is the motion with $k = 5.82$, and the blue curve is the motion with $k = 5$.

Assuming that the initial conditions are $\dot{x}(0) = \dot{\theta}(0) = 0$, $x(0) = 2.405\text{m}$, $\theta(0) = 0.01$ rad, with the constant parameter $l = 5$, we can see two distinctive types of motion in Fig. 19.

Qualitatively, we can see that the blue curve has (roughly) constant motion. That is, it remains in (roughly) simple harmonic motion for all $t \in [0, 250\text{s}]$. On the other hand, we can see that the orange curve has a slightly different kind of motion. It is still harmonic motion, but creates **beats** with itself at certain times.

Notice that this is still present with the blue curve, however, the simple difference in the k value, $\Delta k = 5.82 - 5 = 0.82$ is enough to show what we call **autoparametric resonance**. In fact, the motion in the orange curve predicts that the pendulum will swing, then at some point (near the origin), begin swinging vertically, then swing regularly again, and repeat this process for all t.

This motion is a consequence of a more general type of oscillator called the **parametric oscillator**. In particular, we can explain this type of motion by acknowledging that there is some type of *driving force* which enforces motion onto the system. Here, the slight change in the spring constant shows that the behaviour varies far differently, which suggests that the system is far more sensitive to our parameters and when our motion diverges.

However, now we might ask *what is the difference between these two curves*? From our EoM, we may notice that the natural frequency of our pendulum are

$$\omega_0 = \sqrt{\frac{g}{l + x}}.$$

If we were to adjust the spring constant, a larger k effectively *changes* the average length of the pendulum. Why? For larger k, we expect the spring to be stiffer. So, if k is large, then at some point l becomes more defined. That is, it does not stretch enough to cause large changes in a typical pendulum. This is important because now we can define two different types of frequencies, that of our spring itself, and that of the pendulum.

Graphically, we see that the natural frequency of our pendulum are

$$\omega_0^B \approx \sqrt{\frac{9.8}{7}} \approx 1.18\text{s}^{-1}, \quad \omega_0^O \approx \sqrt{\frac{9.8}{6.7}} \approx 1.21\text{s}^{-1}$$

where we have denoted B to be the blue curve, and O to be the orange curve. Comparing this to frequency of our spring,

$$\omega_{\text{spring}}^B = \sqrt{\frac{k}{m}} = \sqrt{5} \approx 2.23\text{s}^{-1} \quad \omega_{\text{spring}}^O = \sqrt{\frac{k}{m}} = \sqrt{5.82} \approx 2.41\text{s}^{-1}.$$

At first glance, this doesn't seem too different. *However*, we can see that for the orange curve, the ratio between the frequencies is $(\omega_{\text{spring}}/\omega_0)_O = 2.41/1.21 = 1.99$, while while for the blue curve, $(\omega_{\text{spring}}/\omega_0)_B = 2.23/1.18 = 1.89$. The orange curve has a spring frequency essentially *twice* that of the pendulum. This, in fact, leads to parametric resonance, and since there is no external driving factor, we call it autoparametric resonance as before. The small difference in 0.10 between the two ratios is enough to show that autoparametric resonance is highly sensitive to the rate at which the motion is driven.

3.5 Coupled Oscillators

These next few examples are *extremely* important. In fact, coupled, dynamical systems appear in a large variety of areas. We see them in electric circuits, condensed matter, acoustics, quantum computing, and much more. Thus, we will dedicate a lot of detail to investigating the motion of these types of systems.

3.5.1 Two-System Coupled Pendula

Consider a system of two identical pendula of massless rods of length, l, with a bob mass, m. Their pivots are separated by some arbitrary distance, l_{eq}, and the bobs are connected by a spring with a spring constant, k and the same equilibrium length, l_{eq}. Determine the motion of the system and its significance.

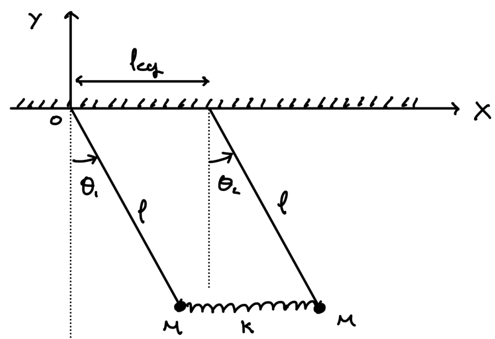


Figure 20: System of two pendula coupled by a spring of constant, k, and separated by a distance, l_{eq}.

We begin by determining the position of the masses, to which we use subscripts 1 and 2 to designate the left and right masses respectively.

$$\begin{aligned} X_1 &= l\sin(\theta_1) \\ Y_1 &= -l\cos(\theta_1) \end{aligned} \quad \text{and} \quad \begin{aligned} X_2 &= l\sin(\theta_2) + l_{\text{eq}} \\ Y_2 &= -l\cos(\theta_2) \end{aligned}$$

so that

$$\dot{X}_1^2 + \dot{Y}_1^2 = l^2\dot{\theta}_1^2 \quad \text{and} \quad \dot{X}_2^2 + \dot{Y}_2^2 = l^2\dot{\theta}_2^2.$$

Due to the coupled nature of the system, we expect the spring potential energy to be more complicated, which is of the form

$$\begin{aligned} V &= \frac{1}{2}k\Delta x^2 \\ &= \frac{1}{2}k\Big[\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} - l_{\text{eq}}\Big]^2. \end{aligned}$$

Simplifying the square root term,

$$\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} = [l\sin(\theta_2) + l_{\text{eq}} - l\sin(\theta_1)]^2 + [-l\cos(\theta_2) + l\cos(\theta_1)]^2.$$

Here, we proceed by assuming the small angle approximation. *This is important.* We will see later that there are two ways to incorporate approximate solutions into our systems. For now we will assume it now, then proceed with analysis later.

$$\begin{aligned} [l\sin(\theta_2) + l_{\text{eq}} - l\sin(\theta_1)]^2 + [-l\cos(\theta_2) + l\cos(\theta_1)]^2 &\approx (l\theta_2 + l_{\text{eq}} - l\theta_1)^2 + (1-1)^2 \\ &\approx \cancel{l^2(\theta_2 - \theta_1)^2} + 2l(\theta_2 - \theta_1)l_{\text{eq}} + l_{\text{eq}}^2. \end{aligned}$$

Notice that we have thrown out the term $(\theta_2-\theta_1)^2$ since it is quadratic in θ. Keeping it would have been a problem, since we would *not* be able to proceed with analysis on our system. Generally, if we want to find approximate solutions to our system, we must throw out terms that are higher than quadratic order in the variable of interest, *should* you proceed with this method. Regardless, or spring potential is

$$V = \frac{1}{2}k\left[\sqrt{l_{\text{eq}} + 2l(\theta_2-\theta_1)l_{\text{eq}}} - l_{\text{eq}}\right]^2 = \frac{1}{2}k\left[l_{\text{eq}}\sqrt{1+\frac{2l(\theta_2-\theta_1)}{l_{\text{eq}}}} - l_{\text{eq}}\right].$$

By the binomial approximation, we further find that

$$\Delta x \approx l_{\text{eq}}\left[1+\frac{1}{\cancel{2}}\frac{\cancel{2}l(\theta_2-\theta_1)}{l_{\text{eq}}}\right] - l_{\text{eq}} = \cancel{l_{\text{eq}}} + l(\theta_2-\theta_1) - \cancel{l_{\text{eq}}} = l(\theta_2-\theta_1)$$

so that

$$V = \frac{1}{2}kl^2(\theta_2-\theta_1)^2.$$

Be careful! We do **not** throw out the quadratic term here. The reason why is, not only have we already thrown it out before, but also because this quadratic term in this case would turn into a linear term after computing the partials in the Lagrangian. In fact, if we kept our previous quadratic term before, our potential energy term would be proportional to θ^4, which would *still* be non-linear after computing the partials in the Lagrangian! Regardless, the full energy terms are

$$T = T_1 + T_2 = \frac{1}{2}ml^2\dot{\theta}_1^2 + \frac{1}{2}ml^2\dot{\theta}_2^2$$

$$V = mgY_1 + mgY_2 + V_{\text{spring}} = -mgl\cos(\theta_1) - mgl\cos(\theta_2) + \frac{1}{2}kl^2(\theta_2-\theta_1)^2$$

to give the Lagrangian

$$\mathcal{L} = \frac{1}{2}ml^2\dot{\theta}_1^2 + \frac{1}{2}ml^2\dot{\theta}_2^2 + mgl\cos(\theta_1) + mgl\cos(\theta_2) - \frac{1}{2}kl^2(\theta_2-\theta_1)^2.$$

Here, our term $(\theta_2-\theta_1)^2$ is a coupling term, which connects the two individual pendula together. Then, computing the partials in θ_1, we find

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\theta}_1} &= ml^2\ddot{\theta}_1\\ \frac{\partial\mathcal{L}}{\partial\theta_1} &= -mgl\sin(\theta_1) - \frac{1}{\cancel{2}}kl^2\cancel{2}(\theta_2-\theta_1)(-1) \approx -mgl\theta_1 - kl^2(\theta_1-\theta_2).\end{aligned}$$

Here, we have again used the small angle approximation for $\sin(\theta_1)$ (lots of approximations here). It's not a problem to do this again, but you should be mindful of where and when you do this. The partials in θ_2 are similar,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\theta}_2} &= ml^2\ddot{\theta}_2\\ \frac{\partial\mathcal{L}}{\partial\theta_2} &= -mgl\sin(\theta_1) - \frac{1}{\cancel{2}}kl^2\cancel{2}(\theta_2-\theta_1)(1) \approx -mgl\theta_1 - kl^2(\theta_2-\theta_1).\end{aligned}$$

Thus, our ELE's are

$$\boxed{\begin{aligned}\ddot{\theta}_1 &= -\frac{g}{l}\theta_1 - \frac{k}{m}(\theta_1-\theta_2)\\ \ddot{\theta}_2 &= -\frac{g}{l}\theta_2 - \frac{k}{m}(\theta_2-\theta_1)\end{aligned}}$$

which is a coupled system of linear second order ODE's. We are now interested in determining the significance of this system. In particular, we want to know how the system evolves in time. There are two methods to determining this: **algebraically** and **with vectors**. We will show the algebraic method first.

Algebraic Method. Let us assume that we are interested in the sum and difference of the equations. We will see why soon. This gives us two more equations,

$$\ddot{\theta}_1 + \ddot{\theta}_2 = -\frac{g}{l}(\theta_1 + \theta_2) = -\frac{g}{l}\theta_S$$
$$\ddot{\theta}_1 - \ddot{\theta}_2 = -\left[\frac{g}{l} + \frac{2k}{m}\right](\theta_1 - \theta_2) = -\left[\frac{g}{l}+\frac{2k}{m}\right]\theta_A$$

which effectively **decouples** our system and puts them in the form of SHM for two new variables, θ_S and θ_A. However, what *is* this motion?

If you are observant, you may notice that the first equation says that if pendula 1 and 2 were released from rest, their *effective* acceleration is is greater than its components. That is, $\ddot{\theta}_S > \ddot{\theta}_1, \ddot{\theta}_2$. The second equation says that if we release both pendula from rest, their *effective* acceleration may not necessarily greater than its components. That is, $\ddot{\theta}_1 \leq \ddot{\theta}_A \leq \ddot{\theta}_2$ or $\ddot{\theta}_2 \leq \ddot{\theta}_A \leq \ddot{\theta}_1$. These two types of motion are what we refer to as **symmetric** and **antisymmetric** motion.

Physically, we can see that $\ddot{\theta}_S$ corresponds to the motion of both pendula in the *same* way. If pendulum 1 moves left, then pendulum 2 would also move left, and vice versa. On the other hand, $\ddot{\theta}_A$ corresponds to the motion of both pendula in the *opposite* way. If pendulum 1 moves left, then pendulum 2 moves right, and vice versa. We can determine the frequency of these dynamics, which are

$$\boxed{\omega_S = \sqrt{\frac{g}{l}} \quad \text{and} \quad \omega_A = \sqrt{\frac{g}{l} + \frac{2k}{m}}.}$$

The first frequency is the same as that of a regular pendulum. This is to be expected. The same motion can be converted into a single, effective pendulum, with frequency ω_S. The second frequency is larger than the first, $\omega_A > \omega_S$, and is also dependent on the spring constant. This, also, is to be expected. If the spring constant is larger, then the *coupled* dynamics of the system are more defined, thus making the frequency larger since both pendula swing faster.

To be more general, this type of motion is what we call a **normal mode**. It is the motion in our system which is *purely* sinusoidal and depends on a single frequency. In fact, the linear combination of our generalized coordinates are what we call **normal coordinates**, which is simply a change of basis to convert to our normal mode representation.

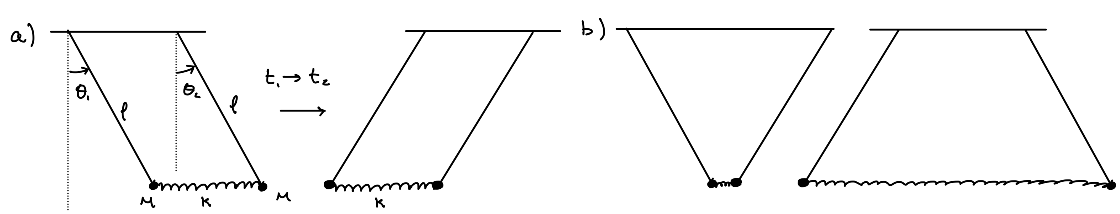


Figure 21: (a) Symmetric mode of the 2 coupled pendula system. (b) Asymmetric mode of the 2 coupled pendula system.

We can, in fact, determine the motion of the individual pendula as well. However, we must impose some initial conditions. For this case, (arbitrary), let us suppose that $\theta_2(0) = \dot{\theta}_1(0) = \dot{\theta}_2(0) = 0$, and $\theta_1(0) = \varepsilon$ for some small displacement ε. We now assume solutions in θ_S and θ_A of the forms

$$\theta_S(t) = A\cos(\omega_S t) + B\sin(\omega_S t)$$
$$\theta_A(t) = C\cos(\omega_A t) + D\sin(\omega_A t).$$

Applying the conditions, we find that

$$\theta_1(0) + \theta_2(0) = \theta_S(0) = \varepsilon$$
$$\theta_1(0) - \theta_2(0) = \theta_A(0) = \varepsilon$$

and furthermore

$$A = C = \varepsilon \quad \text{and} \quad B = D = 0.$$

Thus, we arrive at

$$\theta_S(t) = \varepsilon\cos(\omega_S t) = \theta_1(t) + \theta_2(t)$$
$$\theta_A(t) = \varepsilon\cos(\omega_A t) = \theta_1(t) - \theta_2(t).$$

Changing basis back to our individual pendulums and with the use of trigonometric identities, we finally end up with

$$\theta_1(t) = \frac{\varepsilon}{2}[\cos(\omega_S t) + \cos(\omega_A t)] = \varepsilon\cos\left(\frac{\omega_S + \omega_A}{2}t\right)\cos\left(\frac{\omega_S - \omega_A}{2}t\right)$$
$$\theta_2(t) = \frac{\varepsilon}{2}[\cos(\omega_S t) - \cos(\omega_A t)] = -\varepsilon\sin\left(\frac{\omega_S + \omega_A}{2}t\right)\sin\left(\frac{\omega_S - \omega_A}{2}t\right)$$

which is **not** simple sinusoidal nor co-sinusoidal motion. In fact, plotting these equations for the numerical values $\varepsilon = 1$, $\omega_S = 5\text{s}^{-1}$, and $\omega_A = 4.8\text{s}^{-1}$, we can see that the motion of the individual pendula create beats.

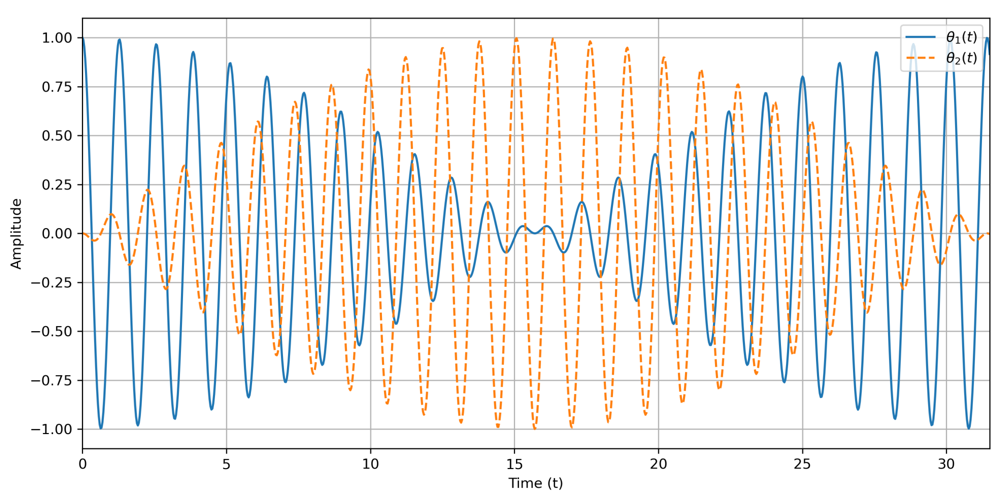


Figure 22: Motion of the individual pendula. The blue curve is the pendulum of angle $\theta_1(t)$, and the orange curve is the pendulum of angle $\theta_2(t)$.

However, what is the significance of these beats? We can see that as pendulum 1 evolves in time, it oscillates normally with a modulated amplitude until it momentarily stops. At this point, all the amplitude of oscillation is in the second pendulum, and this motion repeats for all t. This is actually a consequence of the fact that energy transfers continuously through the system. As the first pendulum moves, it provides energy to the second pendulum, which begins oscillating with an increasing amplitude, until all the energy from the first pendulum is gone. Quantitatively, we can argue that

$$E_1 \propto A_1 \quad \text{and} \quad E_2 \propto A_2$$

and since A_1 decreases for $t \to T$ and A_2 increases for $t \to T$, by conservation of energy, it must be true that $E_1 \to E_2$ from $t \to T$.

Vector Method. The second method is to convert our system of ODEs into a vector equation. Generally, this is preferred, since the previous method becomes rather complicated for more intricate systems. Converting our EoM into vectors, we find that

$$\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} = \begin{bmatrix} -\frac{g}{l} - \frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & -\frac{g}{l} - \frac{k}{m} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}.$$

Now, as we would for regular vector DEs, we make the exponential ansatz $\vec{\theta} = \mathbf{c}e^{-i\lambda t}$ where $\mathbf{c}$ is a column matrix of same dimension as the matrix $[\theta_1\ \theta_2]$.[8] Substituting this back into our equation, we have

$$(-\lambda^2\mathbf{I})\cancel{e^{-\lambda t}} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -\frac{g}{l} - \frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & -\frac{g}{l} - \frac{k}{m} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \cancel{e^{-\lambda t}}.$$

Rearranging this to have zero on the LHS, we get

$$\mathbf{0} = \begin{bmatrix} \lambda^2 - \frac{g}{l} - \frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & \lambda^2 - \frac{g}{l} - \frac{k}{m} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Obviously, trivial solutions where $\mathbf{c} = 0$ imply that there is no motion in the system. So, for nontrivial solutions, we require that the determinant of the coefficient matrix is zero. That is, $\det(\mathbf{K}) = 0$, where $\mathbf{K}$ is our matrix in λ^2. We find that

$$\begin{aligned} 0 = \begin{vmatrix} \lambda^2 - \frac{g}{l} - \frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & \lambda^2 - \frac{g}{l} - \frac{k}{m} \end{vmatrix} &= \left(\lambda^2 - \frac{g}{l} - \frac{k}{m}\right)^2 - \left(\frac{k}{m}\right)^2 \\ &= \lambda^4 - 2\lambda^2\left(\frac{g}{l} + \frac{k}{m}\right) + \frac{g^2}{l^2} + \frac{2gk}{ml} + \cancel{\frac{k^2}{m^2}} - \cancel{\frac{k^2}{m^2}} \end{aligned}$$

and furthermore applying the quadratic formula, we find that our eigenvalues are

$$\lambda^2 = \frac{g}{l} + \frac{k}{m} \pm \frac{k}{m}.$$

We thus have two solutions, which are

$$\boxed{\lambda_1 = \pm\sqrt{\frac{g}{l} + \frac{2k}{m}} \quad \text{and} \quad \lambda_2 = \pm\sqrt{\frac{g}{l}}.}$$

[8]You could alternative use the real ansatz $\vec{\theta} = \mathbf{c}e^{\lambda t}$ as well. Adding the imaginary term simply encodes the sinusoidal behaviour of our equations.

There are some important observations to make here. Our ansatz, $\exp(-i\lambda t)$ implies that real valued λ corresponds to oscillatory solutions, which is to be expected. Our second observation is that our frequencies, which are λ, come in pairs. In principle, we can combine our complex, exponential solutions to get real-valued sine and co-sinusoidal solutions.

Naturally, we are also interested in finding the eigenvectors of our system. If we recall our ansatz'd equation, we can substitute our squared eigenvalues back in to determine the corresponding eigenvectors. Here's a neat trick: we can keep the general expression for our eigenvalues (with the $\pm$ sign), and determine our vectors by row reducing $\mathbf{K}$ that way.

$$\begin{bmatrix} \cancel{\frac{g}{l}+\frac{k}{m}} \pm \frac{k}{m} - \cancel{\frac{g}{l}-\frac{k}{m}} & \frac{k}{m} \\ \frac{k}{m} & \cancel{\frac{g}{l}+\frac{k}{m}} \pm \frac{k}{m} - \cancel{\frac{g}{l}-\frac{k}{m}} \end{bmatrix} = \begin{bmatrix} \pm\frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & \pm\frac{k}{m} \end{bmatrix}.$$

Since we know that the original vector DE is zero, we can cancel out the k/m term (if you're unconvinced, leave it in then you'll find it cancels later). Thus, we will end up having to row reduce

$$\begin{bmatrix} \pm1 & 1 \\ 1 & \pm1 \end{bmatrix} \xrightarrow{R_2\leftrightarrow R_1} \begin{bmatrix} 1 & \pm1 \\ \pm1 & 1 \end{bmatrix} \xrightarrow{R_2/(\pm1)} \begin{bmatrix} 1 & \pm1 \\ 1 & \pm1 \end{bmatrix} \xrightarrow{R_2-R_1} \begin{bmatrix} 1 & \pm1 \\ 0 & 0 \end{bmatrix}.$$

Now, letting $c_2 = t$, we find that

$$\begin{aligned} c_1 \pm c_2 &= 0 \\ c_2 &= t \end{aligned} \implies \begin{aligned} c_1 &= \mp t \\ c_2 &= t \end{aligned} \implies \mathbf{v}_{1,2} = \begin{bmatrix} \mp1 \\ 1 \end{bmatrix}.$$

We thus arrive at our two eigenvectors, which are

$$\boxed{\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.}$$

Notice that eigenvector $\mathbf{v}_1$ corresponds to the symmetric motion we found before in $\theta_S(t)$, while the eigenvector $\mathbf{v}_2$ corresponds to the asymmetric motion in $\theta_A(t)$. However, by the method of vector DEs, we also obtain more information on how these normal modes evolve. In fact, since the arguments of the vectors are $1:1$, we can infer that each pendula swings an equal distance compared to the other.

Normal Coordinate Lagrangian. There is another method to simplify calculations that we neglected to show. This method encodes the normal coordinates into the Lagrangian, where the Lagrangian itself keeps terms only up to the *quadratic* order. Let us recall our Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}ml^2\dot{\theta}_1^2 + \frac{1}{2}ml^2\dot{\theta}_2^2 + mgl\cos(\theta_1) + mgl\cos(\theta_2) - \frac{1}{2}kl^2(\theta_2-\theta_1)^2.$$

We can determine our EoM by simplifying this to an *approximate* Lagrangian. Specifically, we Taylor expand $\sin(\theta_{1,2})$ and $\cos(\theta_{1,2})$ to the second order, such that

$$\sin(\theta) \approx \theta, \quad \cos(\theta) \approx 1 - \frac{\theta^2}{2}.$$

Doing so, we find that

$$\mathcal{L}_{\text{approx}} \approx \frac{1}{2}ml^2(\dot{\theta}_1^2+\dot{\theta}_2^2) - \frac{1}{2}mgl(\theta_1^2+\theta_2^2) - \frac{1}{2}kl(\theta_1-\theta_2)^2.$$

We can rewrite this new Lagrangian in terms of our *normalized* normal coordinates, where $\mathbf{v}^T\mathbf{v} = 1$, such that

$$\mathbf{v}_S = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \quad \mathbf{v}_A = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix}$$

and so we find that

$$\theta_S(t) = \frac{1}{\sqrt{2}}[\theta_1(t) + \theta_2(t)], \quad \theta_A(t) = \frac{1}{\sqrt{2}}[\theta_1(t) - \theta_2(t)]$$
$$\dot\theta_S(t) = \frac{1}{\sqrt{2}}\left[\dot\theta_1(t) + \dot\theta_2(t)\right], \quad \dot\theta_A(t) = \frac{1}{\sqrt{2}}\left[\dot\theta_1(t) - \dot\theta_2(t)\right]$$

and so rewriting our approximate Lagrangian once more, we arrive at

$$\mathcal{L}'_{\text{approx}} = \frac{1}{2}ml^2\left(\dot\theta_S^2 + \dot\theta_A^2\right) - \frac{1}{2}mgl\theta_S^2 - \frac{1}{2}(mgl + 2kl^2)\theta_A^2.$$

Notice that this new Lagrangian is the sum of the Lagrangians for our *decoupled* harmonic oscillators with natural frequencies rewritten in the basis of our eigenfrequencies. However, this approach *requires* that we know the normalized eigenvectors, which does scale for more complex systems as well.

3.5.2 The Compound Pendulum

Consider now a special kind of coupled oscillator consisting of two pendula, each of length, l, with a massless rod. A mass, m_1, is attached to the end of the first rod, and a mass, m_2, is attached to the end of the second rod, such that the second rod swings in conjunction with the first with m_1 as a pivot. Using θ_1 and θ_2 as a proper set of generalized coordinates, determine the EoM of the system.

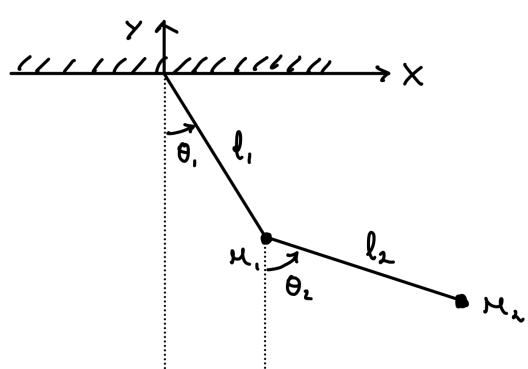


Figure 23: Double pendulum with masses m_1 and m_2.

As always, we set up the generalized coordinates for our system letting the subscripts 1 and 2 designate the respective pendula.

$$X_1 = l_1\sin(\theta_1) \quad \text{and} \quad X_2 = l_1\sin(\theta_1) + l_2\sin(\theta_2)$$
$$Y_1 = -l_1\cos(\theta_1) \qquad\qquad Y_2 = -l_1\cos(\theta_1) - l_2\cos(\theta_2).$$

Computing the energies, we find

$$\begin{aligned} T_1 &= \frac{1}{2}m_1 l_1^2\dot{\theta}_1^2 \\ T_2 &= \frac{1}{2}m_2\Big\{\Big[l_1\cos(\theta_1)\dot{\theta}_1 + l_2\cos(\theta_2)\dot{\theta}_2)^2 + (l_1\sin(\theta_1)\dot{\theta}_1 + l_2\sin(\theta_2)\dot{\theta}_2\Big]^2\Big\} \\ &= \frac{1}{2}m_2\Big\{l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1l_2[\cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)]\dot{\theta}_1\dot{\theta}_2\Big\} \end{aligned}$$

$$V = -m_1gl_1\cos(\theta_1) - m_2gl_1\cos(\theta_1) - m_2gl_2\cos(\theta_2).$$

We thus arrive at the Lagrangian,

$$\begin{aligned} \mathcal{L} = \frac{1}{2}m_1l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2\Big\{l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1l_2[\cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)]\dot{\theta}_1\dot{\theta}_2\Big\} \\ + m_1gl_1\cos(\theta_1) + m_2gl_1\cos(\theta_1) + m_2gl_2\cos(\theta_2). \end{aligned}$$

This Lagrangian is very messy, and so we will state the result of the EoM. You can compute them yourself to be sure.

$$\boxed{\begin{aligned} &(m_1+m_2)l_1^2\ddot{\theta}_1 + m_2l_1l_2[\cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)]\ddot{\theta}_2 \\ &\quad + m_2l_1l_2[-\cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)]\ddot{\theta}_2^2 = -(m_1+m_2)gl_1\sin(\theta_1) \end{aligned}}$$

$$\boxed{\begin{aligned} &m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2[\cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)]\ddot{\theta}_2 \\ &\qquad + m_2l_1l_2[-\sin(\theta_1)\cos(\theta_2) + \cos(\theta_1)\sin(\theta_2)]\ddot{\theta}_1^2 = -m_2gl_2\sin(\theta_2). \end{aligned}}$$

These EoM are extremely complex, so analyzing them in this form is too much of a headache. Generally, we are interested in the approximate behaviour of our systems, so we want to investigate the small angle behaviour of our equations. There are two ways of accomplishing this. We could use the EoM itself, or we can, as before, determine an approximate Lagrangian. This process of simplifying our equations is *also* called **Linearization**.

Using the EoM. At the EoM state, we make the small angle approximations $\sin(\theta) \approx \theta$ and $\cos(\theta) \approx 1$. Furthermore, we drop **any** nonlinear turms, such as

$$\theta^2,\ \theta_a\theta_b,\ \dot{\theta}^2,\ \theta_a\dot{\theta}_b,\ \theta_a^2\theta_b,\ \theta_a\ddot{\theta}_b,\ \text{etc.}$$

For the first EoM, notice that $\sin(\theta_1)\sin(\theta_2) \approx \theta_1\theta_2$ which is nonlinear, so we throw it out. Similarly, $\cos(\theta_1)\cos(\theta_2) \approx 1$. For the second EoM, we apply the same process, and so we find

$$\begin{aligned} (m_1+m_2)l_1^2\ddot{\theta}_1 + m_2l_1l_2\ddot{\theta}_2 = -(m_1+m_2)gl_1\theta_1 \\ m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2\ddot{\theta}_1 = -m_2gl_2\theta_2. \end{aligned}$$

At this point, we can write this system of two equations as a vector DE, such that

$$\begin{bmatrix} (m_1+m_2)l_1^2 & m_2l_1l_2 \\ m_2l_1l_2 & m_2l_2^2 \end{bmatrix}\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} = \begin{bmatrix} -(m_1+m_2)gl_1 & 0 \\ 0 & -m_2gl_2 \end{bmatrix}\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}.$$

Assuming a solution of the form $\vec{\theta} = \mathbf{c}e^{-i\lambda t}$, we thus find that

$$\mathbf{0} = \begin{bmatrix} \lambda^2(m_1+m_2)l_1^2 - (m_1+m_2)gl_1 & \lambda^2m_2l_1l_2 \\ \lambda^2m_2l_1l_2 & \lambda^2m_2l_2^2 - m_2gl_2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

As before, we require the coefficient matrix to be zero for nontrivial solutions, so we arrive at

$$\begin{aligned}\left[\lambda^2(m_1+m_2)l_1^2-(m_1+m_2)gl_1\right](\lambda^2 m_2 l_2^2-m_2 g l_2)-\lambda^4 m_2^2 l_1^2 l_2^2=0\\ \lambda^4 m_1 l_1 l_2-\lambda^2 g(l_1+l_2)(m_1+m_2)+g^2(m_1+m_2)=0\end{aligned}$$

which is the **characteristic equation** for the system. Thus, by the quadratic formula, we find the eigenvalues

$$\boxed{\lambda^2=\frac{g}{2}\frac{(l_1+l_2)(m_1+m_2)}{l_1 l_2 m_1}\pm\frac{g}{2}\frac{\sqrt{(m_1+m_2)[(l_1-l_2)^2 m_1+(l_1+l_2)^2 m_2]}}{l_1 l_2 m_1}.}$$

which is horribly long. Notice that $\omega=\sqrt{\lambda^2}$ are the frequencies of our normal modes of the system. We *could* find the eigenvectors and normal coordinates, but this is complicated, so we will not do so.

Approximate Lagrangian. As seen before, we could also solve this by creating an approximate Lagrangian, which in most cases is significantly easier. Rather than simply dropping nonlinear terms at the EoM stage, at the Lagrangian stage, we instead drop any terms *higher* than quadratic order in the coordinates, and keep those lower. Notice that this changes our small angle approximation, which is now

$$\sin(\theta)\approx\theta,\quad \cos(\theta)\approx 1-\frac{\theta_2}{2}.$$

Recalling the original Lagrangian, we find that $\cos(\theta_1)\cos(\theta_2)\approx 1$ (we drop the quadratic term in θ in this case), $\sin(\theta_1)\sin(\theta_2)=0$, and proceed with the approximation for just $\cos(\theta)$. If the process is applied correctly, we should arrive at

$$\mathcal{L}_{\text{approx}}=\frac{1}{2}m_1 l_1^2\ddot{\theta}_1^2+\frac{1}{2}m_2(l_1^2\dot{\theta}_1^2+l_2^2\dot{\theta}_2^2+2l_1 l_2\dot{\theta}_1\dot{\theta}_2)-\frac{1}{2}(m_1+m_2)gl_1\theta_1^2-\frac{1}{2}m_2 g l_2\theta_2^2.$$

Then, from here, we calculate the EoM as before. We will not show this here, since we have already derived the result, but you may want to go through this as an exercise. Now, let us investigate the properties of our equations of motion. Assuming our initial conditions are $\dot{\theta}_1(0)=\dot{\theta}_2(0)=0$, $\theta_1(0)=0.1$, and $\theta_2(0)=0.14$ for the symmetric mode, and $\dot{\theta}_1(0)=\dot{\theta}_2(0)=0$, $\theta_1(0)=0.1$, and $\theta_2(0)=-0.14$ for the antisymmetric mode, we can see that there are two distinct types of motion.

As expected, the symmetric mode has curves which are roughly in phase, the antisymmetric mode has curves are almost exactly out of phase. Furthermore, notice that the motion is not exactly simple harmonic motion. In fact, there is some amplitude modulation in the pendula, which is a result of the complicate setup of our system.

We can as well determine the frequencies, and assuming that $l_1=l_2=1\text{m}$ and $m_1=m_2=1\text{kg}$, we have that

$$\lambda^2=\frac{g}{2}\left(4\pm\sqrt{8}\right)$$

to show that

$$\omega_S=\sqrt{\frac{g}{2}(4-\sqrt{8}}\approx 2.4\text{s}^{-1},\quad \omega_A=\sqrt{\frac{g}{2}(4+\sqrt{8}}\approx 5.8\text{s}^{-1}.$$

Graphically, we can also see that for the symmetric mode, we have that there are 5 oscillations in about 13s, and for the antisymmetric there are about 10 oscillations in about 10.8s.

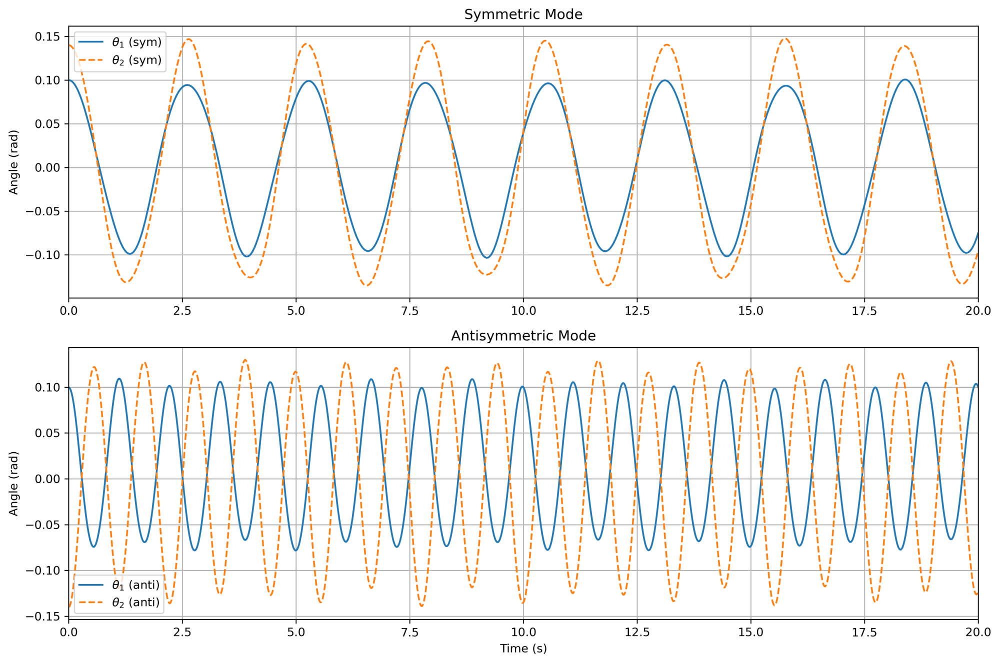


Figure 24: Symmetric and antisymmetric modes of the double pendulum. The blue curve is the pendulum with coordinate, $\theta_1(t)$, and the orange curve is the pendulum with coordinate, $\theta_2(t)$.

We can thus find our graphical frequencies, which are

$$\omega_S' = \frac{5.2\pi}{13} \approx 2.4\text{s}^{-1}, \quad \omega_A' = \frac{10.2\pi}{10.8} \approx 5.8\text{s}^{-1}$$

which does agree with our previous calculation. Now, how do our relative displacements and normal coordinates evolve in time? We can, in fact, find this from our matrix linearized equation.

$$\begin{bmatrix} (m_1+m_2)l_1^2 & m_2 l_1 l_2 \\ m_2 l_1 l_2 & m_2 l_2^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -(m_1+m_2)gl_1 & 0 \\ 0 & -m_2 g l_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Substituting our values for l and m, we can reduce this to

$$-\lambda^2 \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -2g & 0 \\ 0 & -g \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

and solving for the eigenvalue in this case, we have

$$\lambda^2 = \frac{g}{2}\left(4-\sqrt{8}\right) - \frac{g}{2}\left(4-\sqrt{8}\right)(2c_1+c_2) = -2gc_1.$$

Setting $c_1 = s$ as a parameter, we thus arrive at

$$2s + c_2 = \frac{4s}{4-\sqrt{8}} = \frac{(4+\sqrt{8})s}{2} \implies c_2 = \sqrt{2}s$$

to give our eignevector

$$\mathbf{v} = \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}.$$

This quantity is actually rather important. It represents the ratio of the amplitudes of the coordinates, which if we assume the same setup as before, shows that

$$\frac{\text{amplitude}(\theta_1)}{\text{amplitude}(\theta_2)} = \frac{0.01}{0.014} \approx \frac{1}{\sqrt{2}}$$

as before. We will stop with our analysis here, but it is worth going through the rest of the eigenvectors and normal coordinates yourself.

As an aside, another unique property of the double pendulum is that it is *extremely* sensitive to initial conditions. In principle, if we were to change the initial conditions by even 0.01 radian, our trajectory would change wildly.

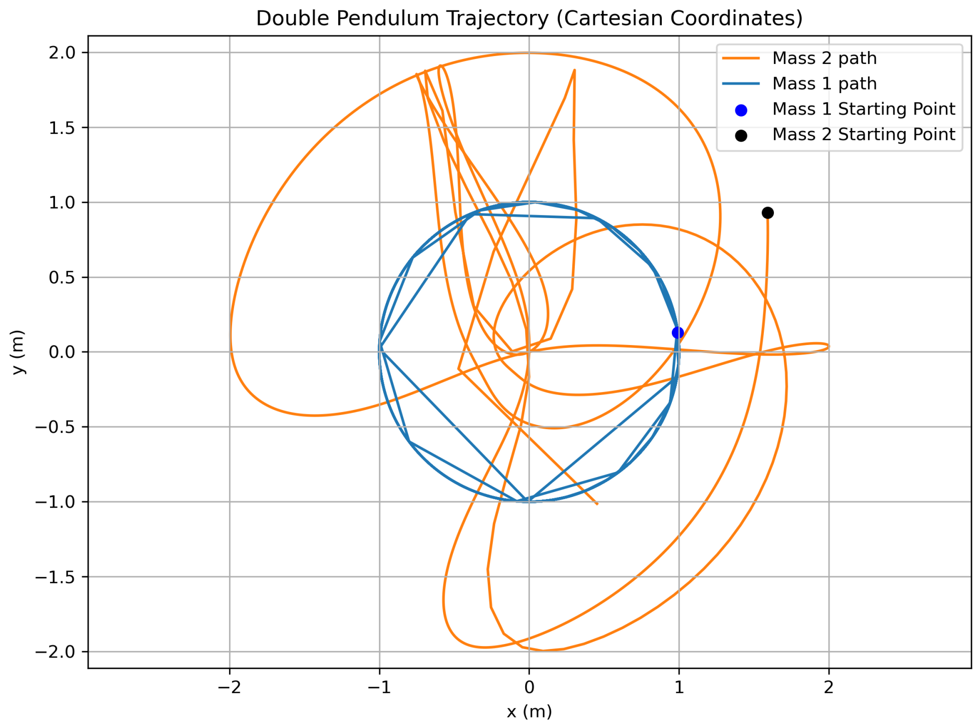


Figure 25: Cartesian evolution of the double pendulum for initial conditions $\theta_1(0) = 1.7$ and $\theta_2(0) = 2.502$.

3.5.3 Three-Pendula Degenerate System

Consider a system of three identical pendula with massless rod length, l, and a mass, m, attached at the end. The pendula are suspended from a platform of mass, M, which is free to slide frictionlessly along a surface. Determine the equation of motion of the system and its significance.

Here, we will approach the problem by the method of approximate Lagrangian. Our generalized coordinates for the system are

$$\begin{aligned} X_M &= x_M \\ Y_M &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} X_{m,i} &= x_M + d_i + l\sin(\theta_i) \approx x_m + d_i l\theta_i \\ Y_{m,i} &= -l\cos(\theta_i) \approx -l\left(1 - \frac{\theta_i^2}{2}\right) \end{aligned}$$

where $d_i = a, b$ or c depending on the subscript i, and $i = 1, 2, 3$ to represent each pendula.

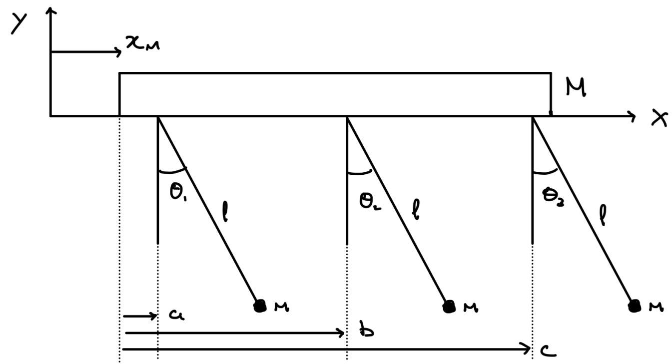


Figure 26: Three-pendula system free to move, attached to a platform of mass, M.

Then, we find our approximate Lagrangian, which is of the form

$$\begin{aligned}\mathcal{L}_{\text{approx}} =& \frac{1}{2}m\Big(l^2\dot{\theta}_1^2 + 2l\dot{\theta}_1\dot{x}_M + \dot{x}_M^2\Big) + \frac{1}{2}m\Big(l^2\dot{\theta}_2^2 + 2l\dot{\theta}_2\dot{x}_M + \dot{x}_M^2\Big) + \frac{1}{2}m\Big(l^2\dot{\theta}_3^2 + 2l\dot{\theta}_3\dot{x}_M + \dot{x}_M^2\Big)\\ &+ \frac{1}{2}M\dot{x}_M^2 - \frac{1}{2}mgl(\theta_1^2 + \theta_2^2 + \theta_3^2).\end{aligned}$$

Notice, x_M is a cyclic coordinate, so we do expect its conjugate momentum to be conserved,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_M} = 3m\ddot{x}_M + M\ddot{x}_M + ml(\ddot{\theta}_1 + \ddot{\theta}_2 + \ddot{\theta}_3) \implies \ddot{x}_M = -\frac{ml}{3m+M}(\ddot{\theta}_1 + \ddot{\theta}_2 + \ddot{\theta}_3).$$

Computing the rest of the partials, we find that they follow the form

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\theta}_i} &= ml^2\ddot{\theta}_i + ml\ddot{x}_M\\ \frac{\partial \mathcal{L}}{\partial \theta_i} &= -mgl\theta_i\end{aligned}$$

and so we arrive at the EoM

$$\begin{aligned}\ddot{\theta}_1 &= -\frac{g}{l}\theta_1 - \frac{\ddot{x}_M}{l}\\ \ddot{\theta}_2 &= -\frac{g}{l}\theta_2 - \frac{\ddot{x}_M}{l}\\ \ddot{\theta}_3 &= -\frac{g}{l}\theta_3 - \frac{\ddot{x}_M}{l}.\end{aligned}$$

Using our relation for $\ddot{x}_M$, we can eliminate $\ddot{x}_M$ and find the coupled ODEs,

$$\boxed{\begin{aligned}\ddot{\theta}_1 &= -\frac{g}{l}\theta_1 + \frac{m}{3m+M}(\ddot{\theta}_1 + \ddot{\theta}_2 + \ddot{\theta}_3)\\ \ddot{\theta}_2 &= -\frac{g}{l}\theta_2 + \frac{m}{3m+M}(\ddot{\theta}_1 + \ddot{\theta}_2 + \ddot{\theta}_3)\\ \ddot{\theta}_3 &= -\frac{g}{l}\theta_3 + \frac{m}{3m+M}(\ddot{\theta}_1 + \ddot{\theta}_2 + \ddot{\theta}_3)/\end{aligned}}$$

We make an observation on our system here. The EoMs have a *symmetric coupling*, in that the EoMs are unchanged if we transform from $\theta_1 \Leftrightarrow \theta_2$ or from $\theta_1 \Leftrightarrow \theta_3$ or from $\theta_2 \Leftrightarrow \theta_3$. Regardless, to make our equations simpler, we define $\alpha = g/l$ and $\beta = m/(3m+M)$, then make the exponential ansatz $\vec{\theta} = \mathbf{c}e^{\lambda t}$. We find, then, that

$$\lambda^2 \mathbf{I} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -\alpha & 0 & 0 \\ 0 & -\alpha & 0 \\ 0 & 0 & -\alpha \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} + \lambda^2 \begin{bmatrix} \beta & \beta & \beta \\ \beta & \beta & \beta \\ \beta & \beta & \beta \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

where we have cancelled out $e^{\lambda t}$. To simplify further, we let $\lambda^2 = \Lambda$, and so we find that our coefficient determinant is

$$0 = \begin{vmatrix} -\Lambda - \alpha + \Lambda\beta & \Lambda\beta & \Lambda\beta \\ \Lambda\beta & -\Lambda - \alpha + \Lambda\beta & \Lambda\beta \\ \Lambda\beta & \Lambda\beta & -\Lambda - \alpha + \Lambda\beta \end{vmatrix} = -(\alpha + \Lambda)^2(\alpha + \Lambda - 3\beta\Lambda).$$

We have not gone through the calculations explicitly, but you should confirm this is true. We thus arrive at our two roots,

$$\Lambda_1 = -\alpha = -\frac{g}{l}, \quad \Lambda_2 = -\frac{\alpha}{1-3\beta} = -\frac{g}{l}\frac{3m+M}{M}.$$

Let us state something here. We say that Λ_1 is a **degenerate** eigenvalue. In particular, it is **two-fold degenerate**, suggesting that it corresponds to two independent parameters in our system after computing our eigenvectors. We will discuss this later. Another useful observation is that in the limit that $M \to \infty$, we see that $\Lambda \to -g/l$ which is that of a regular pendulum.

For non-generate Λ, we find that our matrix equation becomes

$$0 = \frac{\alpha\beta}{1-3\beta} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

This is a matrix with a *single* redundant equation. Stating the result, we will set $c_3 = t$ as the parameter and instead solve the system

$$\begin{aligned} 2c_1 - c_2 - t &= 0 \\ -c_1 + 2c_2 - t &= 0 \end{aligned}$$

which has the solutions $c_1 = c_2 = t$. Thus, we find the corresponding eigenvalue to be

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

For degenerate Λ, we find that our matrix equation becomes

$$0 = \alpha\beta \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

which has *two* redundant equations. Setting $c_3 = s$ and $c_2 = t$ as our parametrs, we immediately find that $c_1 = -s - t$, and so our eigenvector is of the form

$$\mathbf{v}_2 = \begin{bmatrix} -s - t \\ t \\ s \end{bmatrix}.$$

This, however, does not span the subspace of our system. To do so, we must find *two* orthogonal vectors stemming from $\mathbf{v}_2$ so that we can properly describe the motion of our system. We may choose the parameters (arbitrary) $t = 1$ and $s = 1$, so that $\mathbf{v}_2 = [-2, 1, 1]$. To find our second eigenvector, we take the dot product and impose that it is zero, such that

$$\mathbf{v}_2 \cdot \mathbf{v}_2' = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -s - t \\ t \\ s \end{bmatrix} = 0$$

to which we find $t = -s$. Our corresponding eigenvector after substituting is then

$$\mathbf{v}_2' = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Now, we must ask, what is the motion of our system in these two cases? Our first eigenvector, $\mathbf{v}_1$, describes motion which is completely symmetric. That is, each pendulum moves in the same direction at the same time. On the other hand, our second eigenvector describes motion that is asymmetric in at least *one* pendulum. In fact, our second eigenvector is *arbitrary.* It is simply one possible way to describe our system, which just so happens that $\mathbf{v}_2$ corresponds to one pendulum moving opposite to the other two, with twice the speed, and $\mathbf{v}_2'$ corresponds to having one pendulum stationary, with the other two pendula moving asymmetrically.

Generally, any *linear combination* of our modes is still also a mode with the same frequency, so in principle we could have multiple possible types of motion in our system.

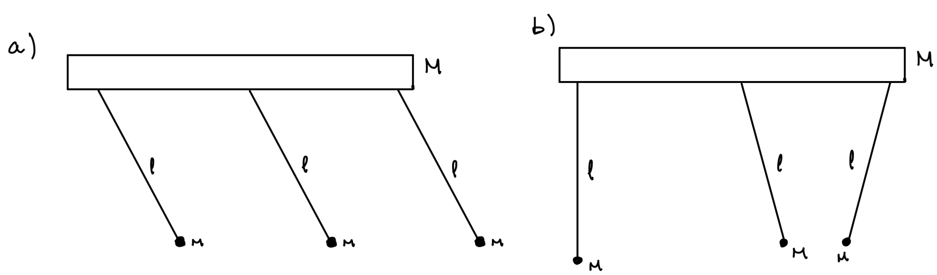


Figure 27: Motion in eigenvector $\mathbf{v}_1$ and $\mathbf{v}_2'$. (a) Symmetric mode in all three pendula. (b) Antisymmetric mode in pendula 2 and 3, and stationary mode in pendula 1.

Furthermore, we should notice that both modes oscillate with a frequency $\omega_0 = \sqrt{-\Lambda} = \sqrt{g/l}$. With this motion, the platform does *not* move to maintain the same frequency as our uncoupled pendula.

Another important observation, which is general, is *that for any n-fold degeneracy, we will have n-parameters.* Here, we have two-fold degeneracy, so we should have expected two parameters in our system.

3.5.4 Many Coupled Oscillators

Consider a string with fixed endpoints in $n + 1$ sections and a fixed distance $\Delta x = d$ for each section, such that the total length in the string is $l = (n + 1)d$. Assume that the string remains under tension, τ, at equilibrium, and a mass, m, is placed at each section. Furthermore, assume that the i-th mass can be displaced a distance y_i transversely such that its x-position is fixed, and ignore gravity. Determine the equations of motion of the system and discuss the generalized case.

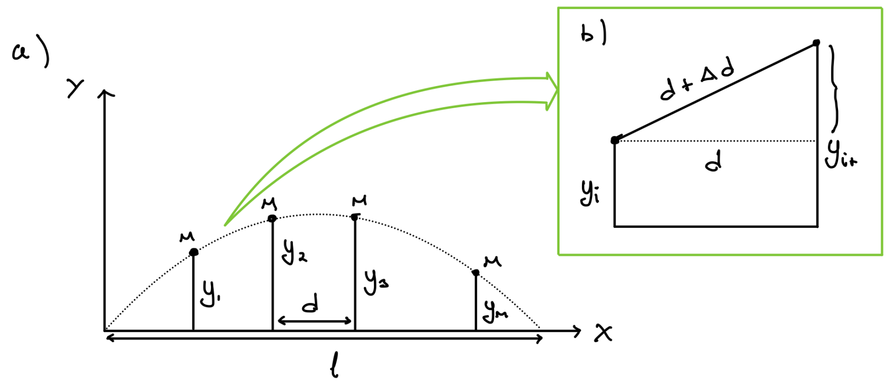


Figure 28: Many oscillator system coupled by a string in $n+1$ sections. (a) Macroscopic setup of the system. (b) Microscopic view of the i-th and i-th plus one mass.

The total kinetic energy in the system is the sum of kinetic energies of each mass, such that

$$T = \frac{1}{2}m\dot{y}_1^2 + \frac{1}{2}m\dot{y}_2^2 + \frac{1}{2}m\dot{y}_3^2 + \cdots + \frac{1}{2}m\dot{y}_n^2$$

and the total potential energy is the work done against the string tension for a given displacement. Assuming that the displacements are small, we have that

$$\begin{aligned} d + \Delta d = \sqrt{d^2 + (y_{i+1} - y_i)^2} &= d\sqrt{1 + \frac{(y_{i+1} - y_i)^2}{d^2}} \\ &\approx d\left[1 + \frac{1}{2}\frac{(y_{i+1} - y_i)^2}{d^2}\right] \\ \implies \Delta d &= \frac{1}{2}\frac{(y_{i+1} - y_i)^2}{d} \end{aligned}$$

by the binomial approximation. Then, we can define the work done against the tension assuming constant τ, such that

$$V_{i,i+1} = \tau\Delta d = \frac{1}{2}\frac{\tau}{d}(y_{i+1} - y_i)^2.$$

Then, the Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}m\dot{y}_1^2 + \frac{1}{2}m\dot{y}_2^2 + \cdots + \frac{1}{2}m\dot{y}_n^2 - \frac{1}{2}\frac{\tau}{d}\left[y_1^2 + (y_2 - y_1)^2 + \cdots + (y_n - y_{n-1})^2 + y_n^2\right].$$

Notice that because the endpoints are fixed, the Lagrangian has terms in the potential that are fixed, that is y_1 and y_n. Computing the partials in y_1, we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{y}_1} &= m\ddot{y}_1 \\ \frac{\partial\mathcal{L}}{\partial y_1} &= -\frac{1}{2}\frac{\tau}{d}\frac{\partial}{\partial y_1}\left[y_1^2 + y_2^2 - 2y_1y_2 + y_1^2 + \cdots\right] = -\frac{\tau}{d}(2y_1 - y_2) \end{aligned}$$

and in y_2, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{y}_2} &= m\ddot{y}_2 \\ \frac{\partial \mathcal{L}}{\partial y_2} &= -\frac{1}{2}\frac{\tau}{d}\frac{\partial}{\partial y_2}\left[\cdots + (y_2 - y_1)^2 + (y_3 - y_2)^2 + \cdots\right] = -\frac{\tau}{d}(-y_1 + 2y_2 - y_3).\end{aligned}$$

Then, the EoMs are

$$\boxed{\begin{aligned}\ddot{y}_1 &= -\frac{\tau}{md}(2y_1 - y_2) \\ \ddot{y}_2 &= -\frac{\tau}{md}(-y_1 + 2y_2 - y_3) \\ \ddot{y}_3 &= -\frac{\tau}{md}(-y_2 + 2y_3 - y_4) \\ &\vdots \\ \ddot{y}_n &= -\frac{\tau}{md}(-y_{n-1} + 2y_n).\end{aligned}}$$

Take note of the edge effects of our EoMs, which are $2y_1$ and $2y_n$. Now, defining $\omega_0^2 \equiv \tau/md$ and making the exponential ansatz, $\mathbf{y} = \mathbf{c}e^{i\omega t}$, we can write this set of ODEs as a vector DE,

$$\omega^2 \mathbf{I}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 2\omega_0^2 & -\omega_0^2 & 0 & \cdots & 0 \\ -\omega_0^2 & 2\omega_0^2 & -\omega_0^2 & \cdots & 0 \\ 0 & -\omega_0^2 & 2\omega_0^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2\omega_0^2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix}$$

where the negative in $-\omega^2$ cancels with the diagonals of the coefficient matrix. In this case, our coefficient matrix is called a **tridiagonal matrix**, reflecting the "nearest neighbour" coupling effects in our system. All off-diagonals are zero, *except* for those closest to each entry in the diagonal. Let us now examine cases of increasing n. For $n = 1$, we have simply a harmonic oscillator of frequency $\omega^2 = 2\omega_0^2$. The $n = 2$ case becomes slightly more complicated. We find,

$$\omega^2 \mathbf{I}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2\omega_0^2 & -\omega_0^2 \\ -\omega_0^2 & 2\omega_0^2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

and solving this, we find our eigenvalues to be $\omega^2 = 2\omega_0^2 \pm \omega_0^2$. We can find our eigenvectors, which we will leave as exercise, which are

$$\mathbf{v}_{S/A} = \begin{bmatrix} \pm 1 \\ 1 \end{bmatrix}.$$

For the $n = 3$ case, we have a matrix

$$\omega^2 \mathbf{I}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 2\omega_0^2 & -\omega_0^2 & 0 \\ -\omega_0^2 & 2\omega_0^2 & -\omega_0^2 \\ 0 & -\omega_0^2 & 2\omega_0^2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

and solving we find the eigenvalues $\omega^2 = 2\omega_0^2$ and $\omega^2 = \left(2 \pm \sqrt{2}\right)\omega_0^2$. Furthermore, our eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_{2,3} = \begin{bmatrix} 1 \\ \pm\sqrt{2} \\ 1 \end{bmatrix}.$$

The process is the same for higher cases of n. In particular, notice that for larger n, there is always at least $n-1$ node.

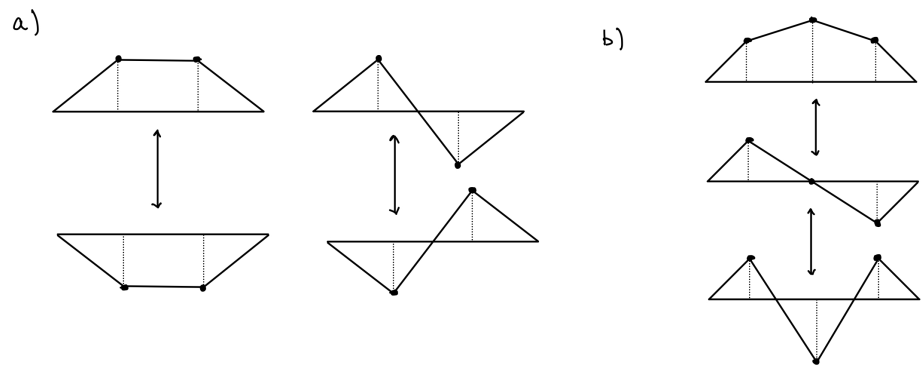


Figure 29: $n = 2$ and $n = 3$ modes for the many coupled oscillator system. (a) Symmetric and antisymmetric modes for $n = 2$. (b) Symmetric, antisymmetric, and semi-symmetric modes for $n = 3$.

Furthermore, notice that when we increase to $n > 2$, the amplitude of the furthest oscillation is $\sqrt{n-1}$ larger than its nearest neighbours (conjecture). For the case $n = 3$, we have that the maximal amplitude is $\sqrt{(3-1)} = \sqrt{2}$ larger than the nearest neighbours.

Now, suppose we want to generalize this setup to arbitrarily many oscillators. It turns out, in fact, that this can be easily summarized in our EoMs, which would be of the form

$$\ddot{y}_j = -\frac{\tau}{md}(-y_{i-1} + 2y_j - y_{j+1})$$

where j is an index that ranges from $j = 1, 2, \cdots, n$. We have excluded the edge effects, but including them is simply a matter of setting $y_0 = 0$ and $y_{n+1} = 0$. Then, we make the ansatz $\mathbf{y} = \mathbf{c}e^{i\omega t} = c\sin(j\gamma - \delta)e^{i\omega t}$. Notice, we have switched from a vector representation to a scalar one in terms of the index, j. Since we expect the motion to be oscillatory, we can assume the solution is of the form $\sin(j\gamma - \delta)$, which is a spatial ansatz with phase, δ and scaling factor, γ.

Substituting this into our equation, we find

$$-\omega^2 \sin(j\gamma - \delta) = -\frac{\tau}{md}\{-\sin[(j-1)\gamma - \delta] + 2\sin(j\gamma - \delta) - \sin[(j+1)\gamma - \delta]\}$$

where we have cancelled out the constant, c, and the exponential term. Isolating for $-\omega^2$, we find

$$-\omega^2 = -\frac{\tau}{md}\left\{2 - \frac{\sin[(j-1)\gamma - \delta] + \sin[(j+1)\gamma - \delta]}{\sin(j\gamma - \delta)}\right\}.$$

Using the trigonometric identities $\sin(a+b) + \sin(a-b) = 2\sin(a)\cos(b)$ for $a = (j\gamma - \delta)$ and $b = -\gamma$, we finally find

$$\omega^2 = \frac{2\tau}{md}[1 - \cos(\gamma)] = \frac{4\tau}{md}\sin^2\left(\frac{\gamma}{2}\right).$$

If our system has n masses, our characteristic equation will also be of order n, with n solutions for ω^2. We can index these solutions with an integer, say r, where $r = 1, 2, \cdots, n$, to re-represent ω as

$$\boxed{\omega_r = \pm 2\sqrt{\frac{\tau}{md}}\sin\left(\frac{\gamma_r}{2}\right).}$$

Notice that ω_r are our normal mode frequencies, and that the presence of negative frequencies are also possible solutions. Furthermore, we require negative frequencies to get real-valued solutions from our original, complex exponential ansatz.

Then, using this, we can return back to our equation for y_j, to find that

$$y_{j,r} = c_r \sin(j\gamma_r - \delta_r)e^{i\omega_r t}.$$

Our boundary conditions on the problem demands that $y_{0,r} = y_{n+1,r} = 0$, such hat

$$c_r \sin(-\delta_r)e^{i\omega_r t} = 0$$
$$c_r \sin\left[(n+1)\gamma_r\right]e^{i\omega_r t} = n+1.$$

In the first boundary equation, δ_r is nonzero since we cannot expect the beginning of the oscillation to begin with zero phase. In the second boundary equation, γ_r is nonzero since j is also nonzero. We thus find that

$$\delta_r = k\pi, \quad \gamma_r = \frac{s\pi}{n+1}$$

for parameters $k, s = 1, 2, \cdots, n$. Let us choose $k = 0$ for simplicity. It is important to observe that the order of ω_r determines the number of spatial frequencies by our original equation for ω_r. Since we have $\mathcal{O}(\omega_r) = n$, we also expect $\text{num}(\gamma_r) = n$. Thus, we can equate our indices, $s = r$, and find that

$$\omega_r = 2\sqrt{\frac{\tau}{md}}\sin\left(\frac{\gamma_r}{2}\right) = 2\sqrt{\frac{\tau}{md}}\sin\left[\frac{r\pi}{2(n+1)}\right]$$

and so

$$\boxed{y_{j,r} = c_r \sin\left(\frac{jr\pi}{n+1}\right)e^{i\omega_r t}.}$$

Now, how does this relate to our original setup? For simplicity, we will consider the case $n = 3$. If the order in ω_r is n, we expect that there are three temporal frequencies, $\text{num}(\gamma_r) = n$. We know that our function is determined by the index, r, however r is related to n in the same way. Thus, we will expect that there are three nodes as in Fig. 29b, for $r = 1, 2, 3$. Furthermore, for $r = 4$, we will find a **null node**, where there is *zero* motion in the system.

If we were to continue past $r = 4$, however, we will find that the motion in our system repeats. That is, $r = 5$ has 2 nodes, $r = 6$ has 1 node, $r = 7$ has zero nodes, until $r = 8$ again has three nodes.

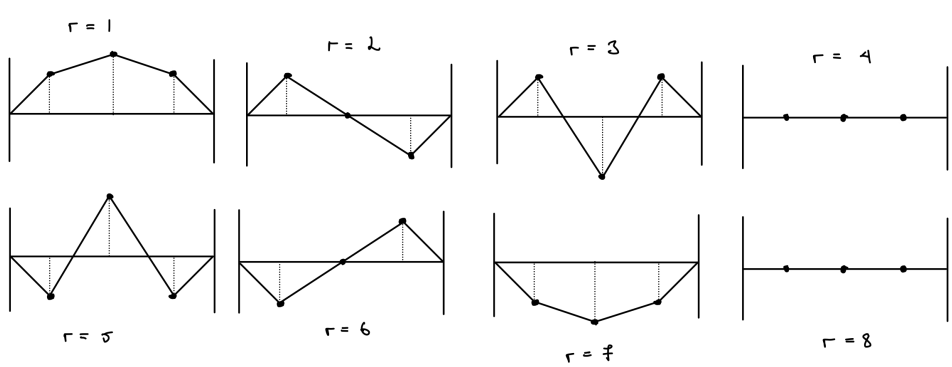


Figure 30: Many coupled oscillator modes for the case $n = 3$ and values $r = 1, 2, \cdots, 8$.

The Continuum Limit. With this derivation, we can now discuss the continuum limit of our results. Consider now the same string, this time with length, l, and a total mass, M, under the same, constant tension, τ. We arrive at the continuum limit when $n \to \infty$ as the spacing in the masses $d \to 0$, such that the total length is still $l = (n+1)d$ stays fixed and finite.

Furthermore, assume that each individual mass, $m \to 0$ such that the linear mass density is $\mu = m/d$ is constant, and $\mu l = m(n+1)d/d = M$. In the continuum limit, we find that

$$\omega_r = 2\sqrt{\frac{\tau}{(\mu d)d}}\sin\left(\frac{r\pi}{2l/d}\right) = 2\sqrt{\frac{\tau}{\mu d^2}}\sin\left(\frac{r\pi d}{2l}\right)$$

and making the small angle approximation,

$$\omega_r \approx \cancel{2}\sqrt{\frac{\tau}{m\cancel{d^2}}}\frac{r\pi\cancel{d}}{\cancel{2}l} \implies \boxed{\omega_r = \frac{r\pi}{l}\sqrt{\frac{\tau}{\mu}}.}$$

For the mode functions, we replace $j/(n+1) \to x/l$ for the continuum limit, where we implied that x ranges from 0 to l, and so we arrive at

$$\boxed{y_{j,r} = c_r \sin\left(\frac{r\pi x}{l}\right)e^{i\omega_r t}.}$$

Notice the following: if the argument within sin is kx, we have

$$kx = \frac{2\pi x}{\lambda} = \frac{r\pi x}{l} \implies \frac{\lambda}{2} = \frac{l}{r} \implies \lambda = \frac{2l}{r}.$$

In particular, the spatial mode functions are sinusoidal with a half-integral number of wavelengths fitting on our string. We will not consider initial conditions, however for you may want to consider the orthogonality of the components $y_{j,r}$ to find a Fourier series, then propagate the state forward in time using the frequencies of the normal modes.

3.5.5 Lagrangian Density and the Wave Equation

There is an alternate approach to modelling the transverse waves of a stretched spring in the continuum limit by the method of Lagrangian densities. Let us assume the string is of equilibrium length, l, with a linear mass density, μ, and fixed endpoints all under a constant tension, τ.

Rather than consider the position of each mass, y_i, we can describe the shape of the string as a function $y(x,t)$.

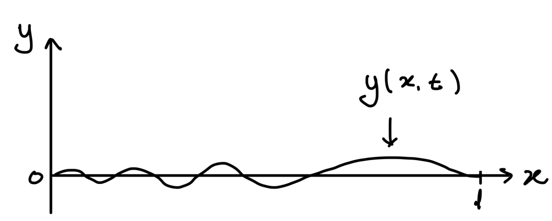


Figure 31: Stretched spring modelled as a function, $y(x,t)$.

Here, we have the differential mass unit, $\mathrm{d}m = \mu\,\mathrm{d}x$. Furthermore, we require the space and time derivatives, which are

$$\dot{y} = \frac{\partial y(x,t)}{\partial t}, \quad y' = \frac{\partial y(x,t)}{\partial x}$$

where the dot is the derivative with respect to time, and the prime is the derivative with respect to space. The kinetic energy of a differential mass element is

$$\mathrm{d}T = \frac{1}{2}\dot{y}^2\,\mathrm{d}m = \frac{1}{2}\mu\dot{y}^2\,\mathrm{d}x$$

and thus the total kinetic energy is

$$T = \int_0^l \frac{1}{2}\mu\dot{y}^2\,\mathrm{d}x\,.$$

When the spring is displaced transversely according to $y(x,t)$, the length of the string increases by

$$l + \Delta l = \int_1^2 \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2} = \int_0^l \sqrt{1 + (y')^2}\,\mathrm{d}x\,.$$

Furthermore, if we assume the transverse displacements are small, we thus have

$$l + \Delta l \approx \int_0^l \left[1 + \frac{1}{2}(y')^2\right]\mathrm{d}x \implies \Delta l = \int_0^l \frac{1}{2}(y')^2\,\mathrm{d}x\,.$$

We find the potential energy by the work done to the stretch of the string against the tension,

$$V = \tau\Delta l = \int_0^l \frac{1}{2}\tau(y')^2\,\mathrm{d}x\,.$$

Thus, our Lagrangian is of the form

$$\mathcal{L} = \int_0^l \left[\frac{1}{2}\mu\dot{y}^2 - \frac{1}{2}\tau(y')^2\right]\mathrm{d}x\,.$$

Let us make note of the integral here. Typically, our Lagrangians are not in integral form, and furthermore, it is unusual that we are integrating over space, unlike the action which is an integration over time. We thus define a new term for this type of Lagrangian, called the **Lagrangian density**. Here, it is of the form

$$\mathcal{L} = \mathcal{L}(y, \dot{y}, y', x) = \frac{1}{2}\mu\dot{y}^2 - \frac{1}{2}\tau(y')^2.$$

Redefining this in terms of the action, we have

$$S = \int_{t_1}^{t_2} \int_0^l \mathcal{L}(y, \dot{y}, y', x)\,\mathrm{d}x\,\mathrm{d}t\,.$$

Similarly, we consider small variations of the path, which we call δS, in both space and time and furthermore impose that it is zero.

$$\delta S = \int_{t_1}^{t_2} \int_0^l \left[\frac{\partial \mathcal{L}}{\partial y}\delta y + \frac{\partial \mathcal{L}}{\partial \dot{y}}\delta\dot{y} + \frac{\partial \mathcal{L}}{\partial y'}\delta y'\right]\mathrm{d}x\,\mathrm{d}t = 0$$

and applying IBP twice, we have

$$\frac{\partial \mathcal{L}}{\partial y} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{y}} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial y'} = 0$$

where we have assumed that the boundary terms are zero. For our specific Lagrangian density, we take our partials which are

$$\begin{aligned}
\frac{\partial}{\partial \mathcal{L}} &= 0 \\
\frac{\partial \mathcal{L}}{\partial \dot{y}} &= \mu \dot{y} \\
\frac{\partial \mathcal{L}}{\partial y'} &= -\tau y'
\end{aligned}$$

and so our EL-type equation is

$$\cancel{\frac{\partial \mathcal{L}}{\partial y}} - \mu \ddot{y} +'' = 0 \implies \boxed{\ddot{y} = \frac{\tau}{\mu} y''.}$$

This quantity, in fact, is so important that we call it the **wave equation** for a constants $c^2 = \tau/\mu$, which is the velocity of the wave. We can connect this back to our previous result by considering the string as an infinite number of small masses. Then, using the phase velocity, we have

$$c = \frac{\omega}{k} = \lambda f = \frac{\omega_r}{2\pi}\lambda_r = \frac{1}{\cancel{2}\pi}\frac{r\cancel{\pi}}{l}\sqrt{\frac{\tau}{\mu}\frac{2\cancel{l}}{\cancel{r}}} = \sqrt{\frac{\tau}{\mu}}$$

which is what we find earlier.

3.5.6 Coupled Metronome Model

Consider a platform of mass, M, and length, l, free to oscillate due to its attachment to a spring of constant, k, at the end of a wall. Suppose that there are two pendula attached at the bottom of the platform, free to oscillate with angles, θ_1 and θ_2; pendula 1 is situated a distance, a, away from the left edge of the platform, and pendula 2 is situated a distance, b, away from the left edge of the platform. Determine the equations of motion of the system and its significance.

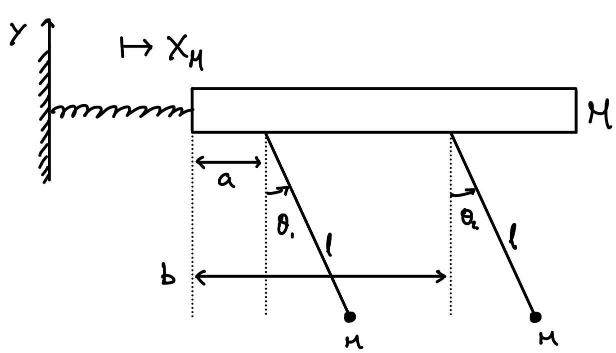


Figure 32: Coupled metronome model of two pendula free to oscillate due to a spring of constant, k.

The generalized coordinates for the first mass are

$$\begin{aligned} x_1 &= a + x_M + l\sin(\theta_1) \\ y_1 &= -l\cos(\theta_1) \end{aligned} \implies \begin{aligned} \dot{x}_1 &= \dot{x}_M + l\cos(\theta_1)\dot{\theta}_1 \\ \dot{y}_1 &= l\sin(\theta_1)\dot{\theta}_1. \end{aligned}$$

Then, the contribution to the Lagrangian for the first mass is

$$\frac{1}{2}m\left[l^2\dot{\theta}_1^2 + 2l\cos(\theta_1)\dot{\theta}_1\dot{x}_M + \dot{x}_M^2\right] + mgl\cos(\theta_1).$$

The contribution from mass 2 assumes a similar setup. In fact, we will let θ_1 become θ_2 and you can confirm this yourself. The contribution from the platform is

$$\frac{1}{2}M\dot{x}_M^2 - \frac{1}{2}kx_M^2$$

and so the Lagrangian is

$$\begin{aligned} \mathcal{L} = \frac{1}{2}m\left[l^2\dot{\theta}_1^2 + 2l\cos(\theta_1)\dot{\theta}_1\dot{x}_M + \dot{x}_M^2\right] + mgl\cos(\theta_1) \\ + \frac{1}{2}m\left[l^2\dot{\theta}_2^2 + 2l\cos(\theta_2)\dot{\theta}_2\dot{x}_M + \dot{x}_M^2\right] + mgl\cos(\theta_2) \\ + \frac{1}{2}M\dot{x}_M^2 - \frac{1}{2}kx_M^2. \end{aligned}$$

Then, computing that partials in θ_1, we find

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\dot{\theta}_1} &= ml^2\dot{\theta}_1 + ml\cos(\theta_1)\dot{x}_M \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\theta}_1} &= ml^2\ddot{\theta}_1 - ml\sin(\theta_1)\dot{\theta}_1\dot{x}_M + ml\cos(\theta_1)\ddot{x}_M \\ \frac{\partial\mathcal{L}}{\partial\theta_1} &= -ml\sin(\theta_1)\dot{\theta}_1\dot{x}_M - mgl\sin(\theta_1). \end{aligned}$$

Similarly, the partials in θ_2 can be found by letting $\theta_1 \to \theta_2$. In x_M, we have

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\dot{x}_M} &= M\dot{x}_M + ml\cos(\theta_1)\dot{\theta}_1 + m\dot{x}_M + ml\cos(\theta_2)\dot{\theta}_2 + m\dot{x}_M \\ &= (2m+M)\dot{x}_M + ml\cos(\theta_1)\dot{\theta}_1 + ml\cos(\theta_2)\dot{\theta}_2 \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_M} &= (2m+M)\ddot{x}_M - ml\sin(\theta_1)\dot{\theta}_1^2 + ml\cos(\theta_1)\ddot{\theta}_1 - ml\sin(\theta_2)\dot{\theta}_2^2 + ml\cos(\theta_2)\ddot{\theta}_2. \end{aligned}$$

Thus, our ELE's yield

$$\boxed{\begin{aligned} ml^2\ddot{\theta}_1 + ml\cos(\theta_1)\ddot{x}_M + mgl\sin(\theta_1) = 0 \\ ml^2\ddot{\theta}_2 + ml\cos(\theta_2)\ddot{x}_M + mgl\sin(\theta_2) = 0 \\ (2m+M)\ddot{x}_M - ml\sin(\theta_1)\dot{\theta}_1^2 + ml\cos(\theta_1)\ddot{\theta}_1 - ml\sin(\theta_2)\dot{\theta}_2^2 + ml\cos(\theta_2)\ddot{\theta}_2 + kx_M = 0 \end{aligned}}$$

which is a system of there equations. Now, there is a unique type of movement dedicated to the system that we are dealing with. In fact, if you were to set this system up in real life, you would find that releasing the pendula from rest, at some time, t, the pendula will synchronize and swing with the same frequency. This happens *regardless* of the frequency you release each pendula, as well as the amplitude.

However, do our system of equations explain this spontaneous synchronization? The answer is no.

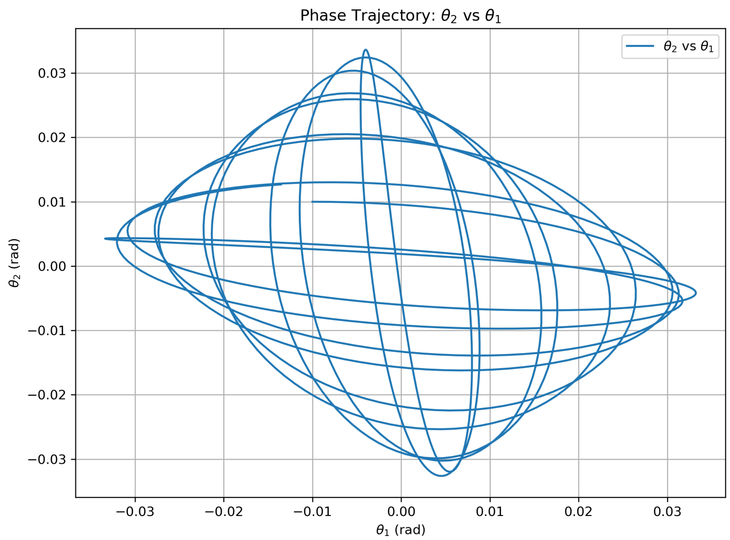


Figure 33: Plot of $\theta_2(t)$ vs. $\theta_1(t)$ over the time interval $t \in [0, 20\text{s}]$.

So, obvious our coupled equations are not enough to explain this motion. In fact, we know that coupled oscillations *transfer* energy between bodies in our system, but this does *not* tell us anything about the spontaneous synchronization of our system.

We will not go into detail on how to show that spontaneous synchronization can be derived, but we will state that the result is related to **Louville's Theorem**, which says that *phase space density* is conserved under a Hamiltonian evolution. We will look at this more closely later.

There is, however, another kind of dynamic which causes a similar result to spontaneous synchronization. This type of motion is **damped motion**, which bring any system to the same final state *regardless* of the starting state. However, we are also interested in the continuous evolution of our system, so we can't have *only* damping. How do we circumvent this? We aren't interested in adding a global drive, as in shaking the platform itself, since that makes the result simple. Instead, we will consider an oscillator which has both types of dynamics contained into one.

The Van der Pol Oscillator. One oscillator which does this is the Van der Pol oscillator (VDPO). In particular, the VDPO winds a spring to keep a metronome ticking, which in turn causes a driving force to keep our system in motion despite the presence of a damping force.

We can easily convert a standard pendulum into a VDPO by adding a generalized force to its EoM. In fact,

$$ml^2\ddot{\theta}_1 + mgl\sin(\theta_1) = Q_1$$

where Q_1 is also related to the velocity by

$$Q_1 = \varepsilon\left[1 - \left(\frac{\theta_1}{\theta_0}\right)^2\right]\dot{\theta}_1.$$

Here, Q_1 is a velocity dependent generalized force, which also depends on the ratio between the varying angle θ_1 and some constant θ_0. The 1 in front of the squared ratio *adds* energy to the system, and we have a small parameter $\varepsilon > 0$ which allows Q_1 to be positive.

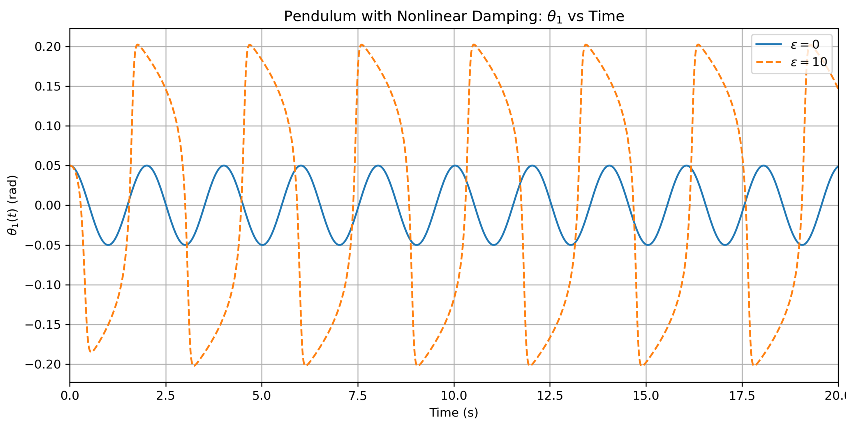


Figure 34: Van der Pol oscillator for parameters $\varepsilon = 0$ (blue curve) and $\varepsilon = 10$ (dashed, orange curve).

Numerically plotting this for parameters $m = 1$kg, $l = 1$m, $g = 9.8$m/s^2, $\theta_0 = 0.1$, and initial conditions $\theta_1(0) = 0.05$ and $\dot{\theta}_1(0) = 0$, we can see that the value of ε determines the type of damped motion we have. Furthermore, notice that we did *not* start the pendulum at the highest amplitude point. In fact, no matter where we start the pendulum, we will always end up with an amplitude of 0.20 rad, assuming the same parameters here.

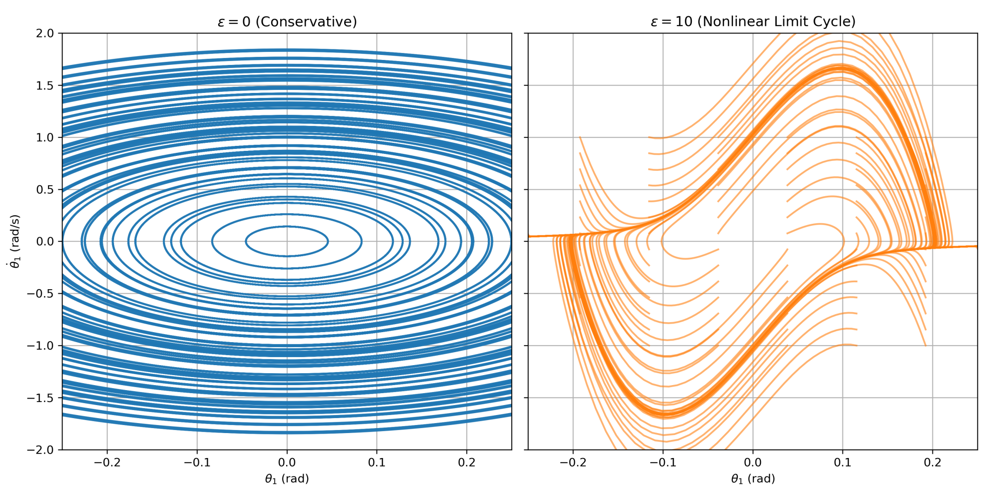


Figure 35: Phase portrait of the Van der Pol oscillator over the range $t \in [0, 20\text{s}]$.

This motion, in particular, is an example of an **attractor** called a **limit cycle**. Here, an attractor refers to a state or set of states in which the system tends to evolve. In the orange curve in Fig. 35, we can see that the motion for $\varepsilon = 10$ tends to follow a sort of sponge. This is the *limiting cycle*. Notice, also, that there is one unstable equilibrium at the origin.

We can now revisit our original coupled metronome model. However, with the knowledge just gained, we can add generalized forces to see how our VDP-type oscillators evolve to synchronize in time. Assuming the same generalized force as before, we can write our equations now as

$$ml^2\ddot{\theta}_1 + ml\cos(\theta_1)\ddot{x}_M + mgl\sin(\theta_1) = Q_1$$
$$ml^2\ddot{\theta}_2 + ml\cos(\theta_2)\ddot{x}_M + mgl\sin(\theta_2) = Q_2$$
$$(2m+M)\ddot{x}_M - ml\sin(\theta_1)\dot{\theta}_1^2 + ml\cos(\theta_1)\ddot{\theta}_1 - ml\sin(\theta_2)\dot{\theta}_2^2 + ml\cos(\theta_2)\ddot{\theta}_2 + kx_M = 0$$

where

$$Q_i = \varepsilon\left[1 - \left(\frac{\theta_i}{\theta_0}\right)^2\right]\dot{\theta}_i.$$

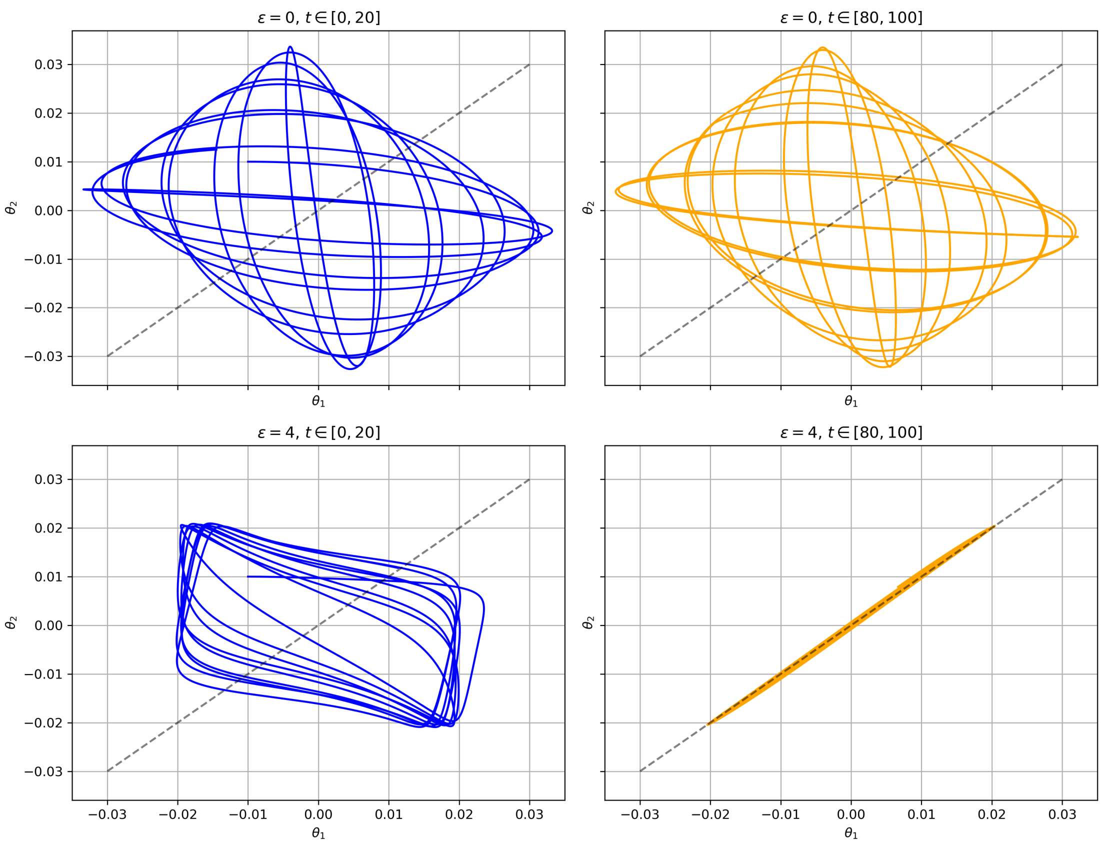


Figure 36: Plot of the Van der Poll-type coupled metronome oscillator for parameters $\varepsilon = 0$ and $\varepsilon = 4$ over time intervals $t \in [0, 20\text{s}]$ and $t \in [80, 100\text{s}]$.

Numerically simulating these equations, we find that the VDPO now synchronize over a later time. In fact, we can see that for $\varepsilon = 0$, the VDP-type oscillators do not synchronize as there is no driving force, but for $\varepsilon = 4$, they synchronize and therefore satisfy Louville's Theorem by at least $t = 80\text{s}$ into the motion.

4 Non-inertial Dynamics

Much of our discussion of classical mechanics has been in the context of non-inertial reference frames. However, non-inertial frames are generally not too common. Now, we will direct our attention to the dynamics of non-inertial systems.

4.1 Vectors in Rotating Frames

Consider a vector, **A**, that is **fixed** relative to a non-inertial frame, rotating with an angular velocity, ω. Since we are dealing with a non-inertial system, we will have to be clear about the

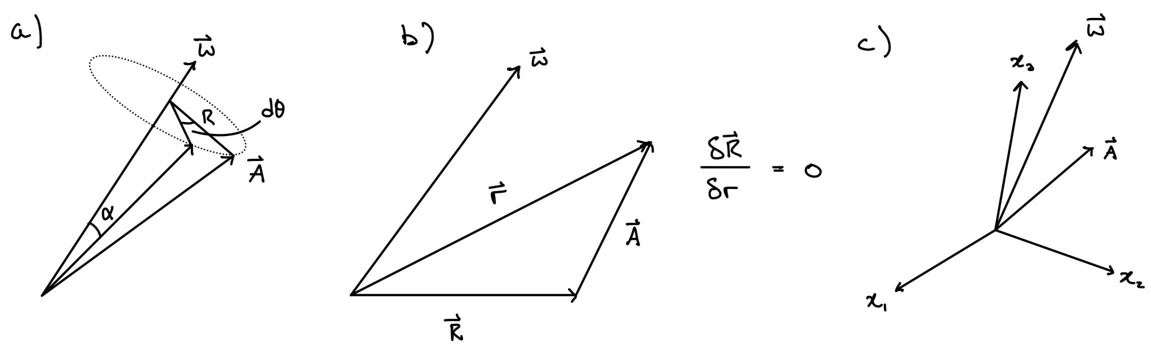


Figure 37: Three different ways of looking at a vector in a non-inertial frame. (a) Vector fixed in a frame rotating with angular velocity, ω. (b) Vector **A** in the general case that it does not lie on the axis of rotation. (c) Vector ω starting from the origin of a rotating frame, and the associated vector **A** also starting from the origin.

types of derivatives we are computing. Right now, we will state that the derivative of a vector, **J**, *within* a rotating frame is denoted by a δ (i.e., the time derivative of **J** in a rotating frame is $\delta\mathbf{J}/\delta t$). In our rotating frame, we have that the time derivative of the vector, **A**, in the rotating frame is

$$\frac{\delta \mathbf{A}}{\delta t} = 0.$$

We can compute the change in **A** in terms of its differential,

$$|\,\mathrm{d}\mathbf{A}\,| = R|\,\mathrm{d}\vec{\theta}\,| = |\mathbf{A}||\,\mathrm{d}\vec{\theta}\,|\sin(\alpha)$$

where we should take note that that change is perpendicular to both **A** and $\vec{\omega}$. Furthermore, the time derivative of **A** relative to a fixed frame is

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \vec{\omega} \times \mathbf{A} = \left[\frac{\mathrm{d}\vec{\theta}}{\mathrm{d}t}\right] \times \mathbf{A}.$$

Let's unpack this first so we know what we are doing. We assumed that **A** is *fixed* within the non-inertial frame, so it makes sense that its time derivative *within* that frame is zero. However, we **cannot** say this for the inertial frame. In fact, as time passes, the inertial frame would observe **A** *rotating* about the axis, $\vec{\omega}$, with angular velocity $\vec{\omega}$, hence why its time derivative is non-zero.

Here, we also made the assumption that the tail of $\mathbf{A}$ lied on the axis of rotation. We can still derive this result in a more general form (Fig. 37b). If we have a vector, $\mathbf{R}$, starting from the origin of the non-inertial frame and travelling to the tail of $\mathbf{A}$, and we assume that $\mathbf{A}$ is still fixed, we find that

$$\frac{\delta \mathbf{R}}{\delta t} = 0$$

Then, making the transformation $\mathbf{r} = \mathbf{R} + \mathbf{A}$, we can take this time derivative to find

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} - \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} = \vec{\omega} \times \mathbf{r} - \omega \times \mathbf{R} = \vec{\omega} \times (\mathbf{r} - \mathbf{R}) = \vec{\omega} \times \mathbf{A}$$

which is the same relation as before. This was a simple case. Now we will suppose that $\mathbf{A}$ is no longer fixed in the rotating frame. That is,

$$\frac{\delta \mathbf{A}}{\delta t} \neq 0.$$

We expect, now, that the differential $\mathrm{d}\mathbf{A}$ to have a contribution from both the inertial frame and the non-inertial frame. In fact, we should find that

$$\mathrm{d}\mathbf{A} = \delta \mathbf{A} + \mathrm{d}\vec{\theta} \times \mathbf{A}$$

where $\delta \mathbf{A}$ is the contribution from the rotating frame. Dividing this equation by $\mathrm{d}t = \delta t$, we arrive at the relation

$$\boxed{\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \frac{\delta \mathbf{A}}{\delta t} + \vec{\omega} \times \mathbf{A}} \tag{24}$$

which is the velocity of $\mathbf{A}$ as seen in the inertial frame. In result, in fact, has two consequences. Notice that $\mathrm{d}\vec{\omega} = \delta \vec{\omega}$, so we expect that

$$\frac{\mathrm{d}\vec{\omega}}{\mathrm{d}t} = \frac{\delta \vec{\omega}}{\delta t}.$$

Furthermore, the basis vectors of our reference frames also follow this result, and in fact,

$$\frac{\mathrm{d}\hat{x}_i}{\mathrm{d}t} = \vec{\omega} \times \hat{x}_i.$$

We can also derive this result using a different method. Consider the configuration in Fig. 37c. Here, we can write the vector $\mathbf{A}$ in term of its components, then take the time derivative

$$\mathbf{A} = \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \implies \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \begin{bmatrix} \dot{A}_x \\ \dot{A}_y \\ \dot{A}_z \end{bmatrix} + \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \cdot \begin{bmatrix} \frac{\mathrm{d}\hat{x}}{\mathrm{d}t} & \frac{\mathrm{d}\hat{y}}{\mathrm{d}t} & \frac{\mathrm{d}\hat{z}}{\mathrm{d}t} \end{bmatrix}$$

where $\dot{A}_x = \delta A_i / \delta x_i$. Furthermore, noting that $\mathrm{d}\hat{x}_i/\mathrm{d}t = \vec{\omega} \times \hat{x}_i$ before, we can simplify this to

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \begin{bmatrix} \dot{A}_x \\ \dot{A}_y \\ \dot{A}_z \end{bmatrix} + \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \cdot \begin{bmatrix} \vec{\omega} \times \hat{x} \\ \vec{\omega} \times \hat{y} \\ \vec{\omega} \times \hat{z} \end{bmatrix} = \frac{\delta \mathbf{A}}{\delta t} + \vec{\omega} \times \mathbf{A}$$

as found before.

4.2 Fictitious Forces

Now that we have a proper description for the mechanics of vectors in non-inertial frames, we can begin discussing how unique forces arise in these scenarios.

Let us consider a point, P, situated within a rotating reference frame with origin O, possessing a vector, $\mathbf{r}$, travelling from O to P. Now consider an inertial reference frame with origin O', away from the non-inertial one, with a vector, $\mathbf{R}$, connecting O' to O. We furthermore denote $\mathbf{r}'$ to be the vector connecting the O to P, and we differentiate between the two frames by letting x_i be the basis vectors of the non-inertial frame, while x_i' are the basis vectors for the inertial frame.

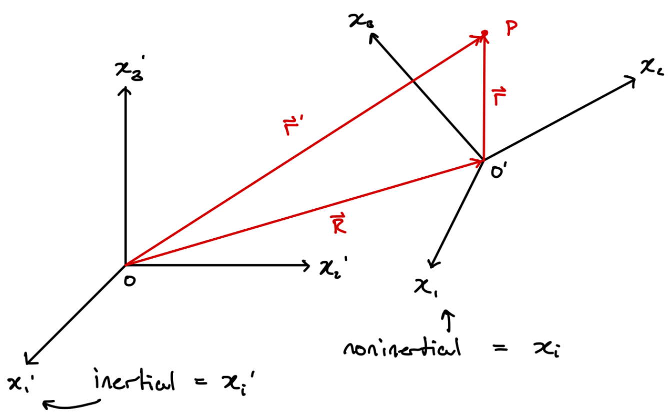


Figure 38: Two reference frames of origins O and O' with a point of interest, P, located within the reference frame of O'.

The relationship between the vectors $\mathbf{r}'$ and $\mathbf{r}$ is such that

$$\boxed{\mathbf{r}' = \mathbf{R} + \mathbf{r}} \tag{25}$$

and decomposing each component of this equation into their basis representation, we have

$$\begin{aligned}
\mathbf{r}' &= r_1'\hat{x}_1' + r_2'\hat{x}_2' + r_3'\hat{x}_3' \\
\mathbf{R} &= R_1\hat{x}_1' + R_2\hat{x}_2' + R_3\hat{x}_3' \\
\mathbf{r} &= r_1\hat{x}_1 + r_2\hat{x}_2 + r_3\hat{x}_3.
\end{aligned}$$

Notice that $\mathbf{r}'$ and $\mathbf{R}$ are written in terms of the inertial basis, however $\mathbf{r}$ is written in the non-inertial basis. We are interested in the velocity and accelerations of the point, P, measured in each frame, so we can simply take the the derivative of (25) to find

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t} &= \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} + \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \\
\mathbf{v}' &= \mathbf{V} + \frac{\delta\mathbf{r}}{\delta t} + \vec{\omega}\times\mathbf{r}.
\end{aligned}$$

Condensing further, we arrive at

$$\boxed{\mathbf{v}' = \mathbf{V} + \dot{\mathbf{r}} + \vec{\omega}\times\mathbf{r}} \tag{26}$$

where $\mathbf{V} = \mathrm{d}\mathbf{R}/\mathrm{d}t$ and $\dot{\mathbf{r}} = \delta\mathbf{r}/\delta t$. Here, $\mathbf{v}'$ is the velocity measured in the inertial basis, while $\dot{\mathbf{r}}$ is the velocity measured in the non-inertial basis. The acceleration follows the same process, to which we find

$$\frac{\mathrm{d}\mathbf{v}'}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\mathbf{r}}{\delta t} + \frac{\mathrm{d}}{\mathrm{d}t}[\vec{\omega}\times\mathbf{r}].$$

This expression is more complicated, so we will evaluate each component separately. For the second term, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\mathbf{r}}{\delta t} = \frac{\delta^2\mathbf{r}}{\delta t^2} + \vec{\omega}\times\frac{\delta\mathbf{r}}{\delta t} = \ddot{\mathbf{r}} + \vec{\omega}\times\dot{\mathbf{r}}$$

by the chain rule, and for the third term we find that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\vec{\omega}\times\mathbf{r}) &= \frac{\mathrm{d}\vec{\omega}}{\mathrm{d}t}\times\mathbf{r} + \vec{\omega}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \\ &= \dot{\vec{\omega}}\times\mathbf{r} + \vec{\omega}\times\left[\frac{\delta\mathbf{r}}{\delta t} + \vec{\omega}\times\mathbf{r}\right] \\ &= \dot{\vec{\omega}}\times\mathbf{r} + \vec{\ }\times\dot{\mathbf{r}} + \vec{\omega}\times[\vec{\omega}\times\mathbf{r}].\end{aligned}$$

Thus, our acceleration is

$$\boxed{\mathbf{a}' = \mathbf{A} + \ddot{\mathbf{r}} + 2\vec{\omega}\times\dot{\mathbf{r}} + \dot{\vec{\omega}}\times\mathbf{r} + \vec{\omega}\times(\vec{\omega}\times\mathbf{r})} \tag{27}$$

where, like before, $\mathbf{A} = \mathrm{d}\mathbf{V}/\mathrm{d}t$ and $\ddot{\mathbf{r}} = \delta^2\mathbf{r}/\delta t^2$. Here, $\mathbf{a}'$ is the acceleration measured in the inertial basis, and $\ddot{\mathbf{r}}$ is the acceleration in the non-inertial basis.

Now, let us suppose that we have an external force on a mass, m, present within our inertial frame, such that

$$\mathbf{F} = m\mathbf{a}'.$$

We will, in fact, find that the acceleration in the non-inertial frame is simply

$$\boxed{m\ddot{\mathbf{r}} = \mathbf{F} - m\mathbf{A} - m\vec{\omega}\times(\vec{\omega}\times\mathbf{r}) - 2m\vec{\omega}\times\dot{\mathbf{r}} - m\dot{\vec{\omega}}\times\mathbf{r}} \tag{28}$$

which is considerably different from a system which is completely inertial. In fact, we have *four* new quantities, which we must name. Here, $-m\mathbf{A}$ is the translational force of our non-inertial frame, $-m\vec{\omega}\times(\vec{\omega}\times\mathbf{r})$ is the centrifugal force, $-2m\vec{\omega}\times\dot{\mathbf{r}}$ is the Coriolis force, and $-m\vec{\omega}\times\mathbf{r}$ is the azimuthal or, if you're annoying, Euler force. In fact, we can rewrite the force equation more simply as

$$m\ddot{\mathbf{r}} = \mathbf{F} + \mathbf{F}_{\text{translational}} + \mathbf{F}_{\text{centrifugal}} + \mathbf{F}_{\text{Coriolis}} + \mathbf{F}_{\text{Euler}}.$$

Let us now discuss the significance of these forces.

The Translational Force. This force, $\mathbf{F}_{\text{trans.}} = -m\mathbf{A}$, arises from the linear acceleration of the origin of our non-inertial reference frame. If you recall the accelerating pulley, we found an "effective" g of the value $(g-\alpha)$, where α is the acceleration of our elevator. Here, g is reduced by α since $\alpha < g$. This, in turn, causes the feeling of "weightlessness" that you may be familiar with in an elevator.

The Centrifugal Force. The force, $\mathbf{F}_{\text{centr.}} = -m\vec{\omega}\times(\vec{\omega}\times\mathbf{r})$ arises from the rotation of our non-inertial reference frame. The most intuitive example is what you feel when you're rounding a corner in a car. There's a *fictitious* force pulling you toward the door of the car as you round a turn, since you, the origin, is turning around an inertial point, the corner.

The Coriolis Force. This force, $\mathbf{F}_{\text{cor.}} = -2m\vec{\omega} \times \dot{\mathbf{r}}$, is rather counterintuitive. There is no "everyday" occurrence that this can be likened to. However, if you've found yourself spinning (that is, you are a rotating reference frame), about some point, and attempted to throw a ball, you will find that the ball tends to veer to either the left or right of you.

See this Youtube short, called the Coriolis Carousel: Short.

A slightly cool connection of the Coriolis force to a real-world occurrence (though rare) is the weather. Hurricanes actually arise when air rushes into a low pressure region. However, depending on whether the air rushes within the northern or southern hemisphere, the air will tend to deflect to either its left or right. This, in turn, creates a counter-clockwise swirl, which gradually builds in what we call a hurricane, if the air is deflected right. If the air is deflected left, we call this a cyclone.

The Euler Force. The Euler force, $\mathbf{F}_{\text{Euler}} = -m\dot{\vec{\omega}} \times \mathbf{r}$, is relatively intuitive when we consider an angular velocity which changes magnitude but not direction. Here, the Euler force arises from the linear acceleration of the origin *perpendicular* to $\mathbf{r}$.

We can also have this effect when the direction of rotation of the non-inertial frame changes for direction *and* magnitude. This, however, is not very intuitive.

4.3 Non-inertial Lagrangian Mechanics

Let us now consider a particle of mass, m, in a potential which can be written relative to a non-inertial frame of reference, $U(\mathbf{r})$.

We can construct our Lagrangian by writing our kinetic energies *relative* to the inertial frame by our velocity relation (26). Here, we have that

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m|\mathbf{v}'|^2 - U(\mathbf{r})\\ &= \frac{1}{2}m|\mathbf{V} + \dot{\mathbf{r}} + \vec{\omega}\times\mathbf{r}|^2 - U(\mathbf{r})\\ &= \frac{1}{2}m\Big[\cancel{|\mathbf{V}|^2} + |\dot{\mathbf{r}}|^2 + |\vec{\omega}\times\mathbf{r}|^2 + 2\mathbf{V}\cdot\dot{\mathbf{r}} + 2(\mathbf{V}+\dot{\mathbf{r}})\cdot(\omega\times\mathbf{r})\Big] - U(\mathbf{r})\end{aligned}$$

where we can cancel the first term since $\mathbf{V}$ does not depend on the coordinates x_i. Now, we will simplify this Lagrangian with the identities

$$\begin{gathered}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{d})\\ \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a})\end{gathered}$$

to arrive that

$$\mathcal{L} = \frac{1}{2}m\big[|\dot{\mathbf{r}}|^2 + |\vec{\omega}|^2|\mathbf{r}|^2 - |\mathbf{r}\cdot\vec{\omega}|^2 + 2\mathbf{V}\cdot\dot{\mathbf{r}} + 2(\mathbf{V}+\dot{\mathbf{r}})\cdot(\vec{\omega}\times\mathbf{r})\big] - U(\mathbf{r}).$$

Assuming that the inertial and non-inertial coordinates are Cartesian, we can compute our partials for the ELE,

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\dot{x}} &= m\dot{x} + m\mathbf{V}\cdot\hat{x} + m(\vec{\omega}\times\mathbf{r})_x\\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}} &= m\ddot{x} + m\Big[\dot{\mathbf{V}}\cdot\hat{x} + \dot{\mathbf{V}}\cdot(\vec{\omega}\times\hat{x})\Big] + m\Big[(\dot{\vec{\omega}}\times\mathbf{r})\cdot\hat{x} + (\vec{\omega}\times\dot{\mathbf{r}})\cdot\hat{x}\Big]\\ \frac{\partial\mathcal{L}}{\partial x} &= m|\vec{\omega}|^2x - m\omega_x(\mathbf{r}\cdot\vec{\omega}) + m(\mathbf{V}\times\vec{\omega})_x + m(\dot{\mathbf{r}}\times\vec{\omega})_x - \frac{\partial U(\mathbf{r})}{\partial x}.\end{aligned}$$

We will skip simplifying the ELE since it is lengthy, and instead state the known result. However, if you wish to do this yourself, you should make use of the cyclic permutations of the dot and cross product to cancel some terms. We thus find

$$m\ddot{x} = F_x - mA_x - m[\vec{\omega} \times (\vec{\omega} \times \mathbf{r})]_x - 2m(\vec{\omega} \times \mathbf{r})_x - m(\vec{\omega} \times \mathbf{r})_x$$

and by the symmetry of our problem, we should find similar expression for y and z. Then, we can find that our vector representation of the EOM agrees with our previous result, which is

$$m\ddot{\mathbf{r}} = \mathbf{F} - m\mathbf{A} - m\vec{\omega} \times (\vec{\omega} \times \mathbf{r}) - 2m\vec{\omega} \times \dot{\mathbf{r}} - m\dot{\vec{\omega}} \times \mathbf{r}.$$

This, as we saw just now, should be enough to convince you that Lagrangian mechanics and handle non-inertial frames very well, as long as we write the kinetic energy related to an inertial point.

4.3.1 Constantly Accelerating Pendulum

Consider a pendulum with a massless rod of length, l, and a bob of mass, m, attached at its end. The pendulum is attached to a ceiling which accelerates with a constant acceleration, $\mathbf{a}$. Determine the equation of motion for the system and discuss its significance.

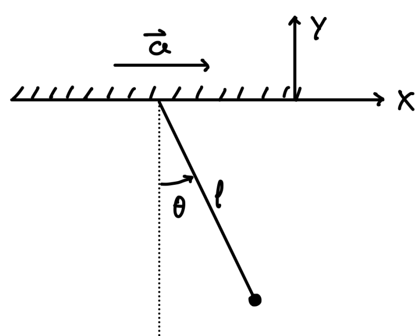


Figure 39: Pendulum subject to constant horizontal acceleration, $\mathbf{a}$.

The generalized coordinates for our system are

$$\begin{aligned} X &= l\sin(\theta) + \frac{1}{2}at^2 \\ Y &= -l\cos(\theta) \end{aligned} \implies \begin{aligned} \dot{X} &= l\cos(\theta)\dot{\theta} + at \\ \dot{Y} &= l\sin(\theta)\dot{\theta} \end{aligned}$$

and so our energies are

$$T = \frac{1}{2}m\Big[l^2\dot{\theta}^2 + 2atl\cos(\theta)\dot{\theta} + a^2t^2\Big]$$
$$V = -mgl\cos(\theta).$$

This gives a Lagrangian of the form

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}m\Big[l^2\dot{\theta}^2 + 2atl\cos(\theta)\dot{\theta} + \cancel{a^2t^2}\Big] + mgl\cos(\theta) \\ &= \frac{1}{2}m\Big[l^2\dot{\theta}^2 + 2atl\cos(\theta)\dot{\theta}\Big] + mgl\cos(\theta) \end{aligned}$$

where we have cancelled out the single, non-impacting term. Computing our partials, we find

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot{\theta}} &= ml^2\dot{\theta} + matl\cos(\theta) \\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\theta}} &= ml^2\ddot{\theta} + mal\cos(\theta) - matl\sin(\theta)\dot{\theta} \\
\frac{\partial \mathcal{L}}{\partial \theta} &= -matl\sin(\theta)\dot{\theta} - mgl\sin(\theta)
\end{aligned}$$

and so the ELE reads

$$\boxed{\ddot{\theta} = -\frac{1}{l}[g\sin(\theta) + a\cos(\theta)].}$$

At equilibrium, we have $\ddot{\theta} = 0$, and so we find that

$$0 = -\frac{1}{l}[g\sin(\theta) + a\cos(\theta)] \implies \tan(\theta_{\text{eq}}) = -\frac{a}{g}.$$

Notice that at this result implies that the equilbrium position for constant acceleration, $\mathbf{a}$, is **not** straight down. In fact, it is at a negative angle relative to the $\theta = 0$ point, such that the translational force, $\mathbf{F}_{\text{trans.}}$ balances the gravitational force, $\mathbf{F}_g$.

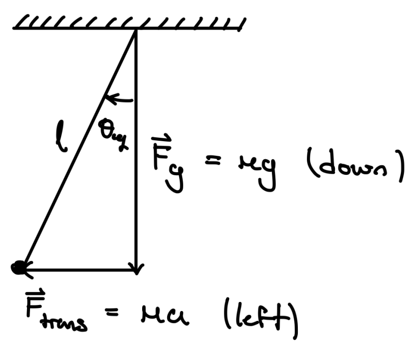


Figure 40: The equilibrium case for a pendulum under constant horizontal acceleration, $\mathbf{a}$.

We can, now, define a new angular variable which is relative to the equilibrium angle if we let $\tilde{\theta} = \theta - \theta_{\text{eq}}$, such that $\ddot{\tilde{\theta}} = \ddot{\theta}$ and

$$\ddot{\tilde{\theta}} = -\frac{1}{l}\left[g\sin\left(\tilde{\theta} + \theta_{\text{eq}}\right) + a\cos\left(\tilde{\theta} + \theta_{\text{eq}}\right)\right].$$

Using some trigonometric identities, we can thus show that

$$\boxed{\ddot{\tilde{\theta}} = -\frac{1}{l}\sqrt{a^2 + g^2}\sin\left(\tilde{\theta}\right).}$$

What does this tell us? We know that this equation is in the form of SHM with a natural frequency,

$$\omega_0 = \frac{(a^2 + g^2)^{1/4}}{l^{1/2}}$$

and in fact tells us what the ratio is between the gravitational and translational force of our system.

4.3.2 Coriolis Ballistics

A relatively important concept in non-inertial mechanics is that of Coriolis ballistics. In general, the trajectories of our particles become more complicated by the rotation of the Earth. We will see how this comes into contact with our systems and projectile motion.

Consider a projectile, launched from the origin of a non-inertial frame of reference, with an initial velocity of

$$\mathbf{v}(0) = v_x\hat{x} + v_y\hat{y} + v_z\hat{z}$$

on the surface of the Earth. Let $\vec{\omega}$ denote the angular velocity of the earth (constant), and determine the equations of motion for the projectile and its significance.

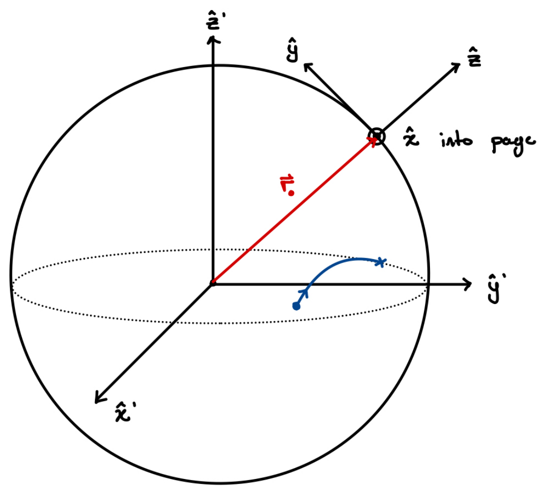


Figure 41: Coriolis ballistics of a projectile being launched relative to the origin of Earth.

Here, the Earth will be our non-inertial frame of reference, and the projectile is the non-inertial frame, free to rotate and translate. We will ignore air resistance and ignore the centrifugal force on our system, that is, we calculate the trajectory to **first order** in $|\vec{\omega}|$.

We must be careful here. In this scenario, it is fine to define the velocity in the non-inertial frame relative to the origin *at* the surface of the Earth. In general, we may not be able to do this, but we can see that the relation

$$\mathbf{r} = \mathbf{r}_0 + \begin{bmatrix} x \\ y \\ z \end{bmatrix} \implies \dot{\mathbf{r}} = \mathbf{0} + \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}$$

makes our calculations easier.

We can assume that $\vec{\omega}$ is a small quantity in comparison to everything else, so to first order in $\vec{\omega}$, we can approximate $\mathbf{F}_g = -mg\hat{z}$ to arrive at

$$m\ddot{\mathbf{r}} = \mathbf{F}_g - 2m\vec{\omega} \times \dot{\mathbf{r}}$$

where $\vec{\omega} \times \dot{\mathbf{r}}$ is our Coriolis force. Expanding this into vector form, we can see that

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} - 2 \begin{bmatrix} 0 \\ \omega \sin(\theta) \\ \omega \cos(\theta) \end{bmatrix} \times \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}.$$

Evaluating the cross product, we arrive at

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 2\omega \cos(\theta)\dot{y} - 2\omega \sin(\theta)\dot{z} \\ -2\omega \cos(\theta)\dot{x} \\ -g + 2\omega \sin(\theta)\dot{x} \end{bmatrix}.$$

This is rather hard to solve. However, we can find an approximate solution to these coupled equations by a **perturbative** approach. We will assume a solution can be written in the form

$$x_i = x_{i,0} + \varepsilon x_{i,1} + \mathcal{O}(\varepsilon^2).$$

Here, ε is a small parameter to which we can keep track of the order of $\vec{\omega}$. Furthermore, we can ignore orders higher than second order, so substituting this relation into our vector DE,

$$\begin{bmatrix} \ddot{x}_0 + \varepsilon\ddot{x}_1 \\ \ddot{y}_0 + \varepsilon\ddot{y}_1 \\ \ddot{z}_0 + \varepsilon\ddot{z}_1 \end{bmatrix} = \begin{bmatrix} 2\omega \cos(\theta)(\dot{y}_0 + \varepsilon\dot{y}_1) - 2\omega \sin(\theta)(\dot{z}_0 + \varepsilon\dot{z}_1) \\ -2\omega \cos(\theta)(\dot{x}_0 + \varepsilon\dot{x}_1) \\ -g + 2\omega \sin(\theta)(\dot{x}_0 + \varepsilon\dot{x}_1) \end{bmatrix}.$$

By our perturbative approach, we know that $\mathbf{r}_1 \approx \mathbf{r}_0 + \varepsilon\mathbf{r}_1$, so we have

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{y}_1 \\ \ddot{z}_1 \end{bmatrix} \approx \begin{bmatrix} 2\omega \cos(\theta)(\dot{y}_0 + \varepsilon\dot{y}_1) - 2\omega \sin(\theta)(\dot{z}_0 + \varepsilon\dot{z}_1) \\ -2\omega \cos(\theta)(\dot{x}_0 + \varepsilon\dot{x}_1) \\ -g + 2\omega \sin(\theta)(\dot{x}_0 + \varepsilon\dot{x}_1) \end{bmatrix}.$$

Since we are using ε to keep track of the order in our system, we can notice that $\varepsilon\dot{x}_i$ is of second order. Ignoring zeroth order portions in our first order as well, we find that the first order correction term is

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{y}_1 \\ \ddot{z}_1 \end{bmatrix} = \begin{bmatrix} 2\omega \cos(\theta)\dot{y}_0 - 2\omega \sin(\theta)\dot{z}_0 \\ -2\omega \cos(\theta)\dot{x}_0 \\ 2\omega \sin(\theta)\dot{x}_0 \end{bmatrix}.$$

We already know our zeroth order equation with no dependence on ε or ω, which is

$$\begin{bmatrix} \ddot{x}_0 \\ \ddot{y}_0 \\ \ddot{z}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} \implies \begin{bmatrix} \dot{x}_0 \\ \dot{y}_0 \\ \dot{z}_0 \end{bmatrix} = \begin{bmatrix} v_{x_0} \\ v_{y_0} \\ v_{z_0} - gt \end{bmatrix} \implies \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} v_{x_0}t \\ v_{y_0}t \\ v_{z_0} - \frac{1}{2}gt^2 \end{bmatrix}$$

and substituting this into our first order correction, we find

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} \omega \cos(\theta)v_{y_0}t^2 - 2\omega \sin(\theta)\left(\frac{1}{2}v_{z_0}gt^2 - \frac{1}{6}gt^3\right) \\ -\omega \cos(\theta)v_{x_0}t^2 \\ \omega \sin(\theta)v_{x_0}t^2 \end{bmatrix}.$$

Notice that we have used our initial conditions for the first order terms to arrive at this result. Thus, putting this together, we can arrive at the solution

$$\boxed{\begin{bmatrix} x \\ y \\ z \end{bmatrix} \approx \mathbf{r}_0 + \mathbf{r}_1 = \begin{bmatrix} v_{x_0}t + \omega \cos(\theta)v_{y_0}t^2 - \omega \sin(\theta)v_{z_0}t^2 + \frac{1}{3}\omega \sin(\theta)gt^3 \\ v_{y_0}t - \omega \cos(\theta)v_{x_0}t^2 \\ v_{z_0}t - \frac{1}{2}gt^2 + \omega \sin(\theta)v_{x_0}t^2 \end{bmatrix}}$$

which is our solution to first order in $\vec{\omega}$.

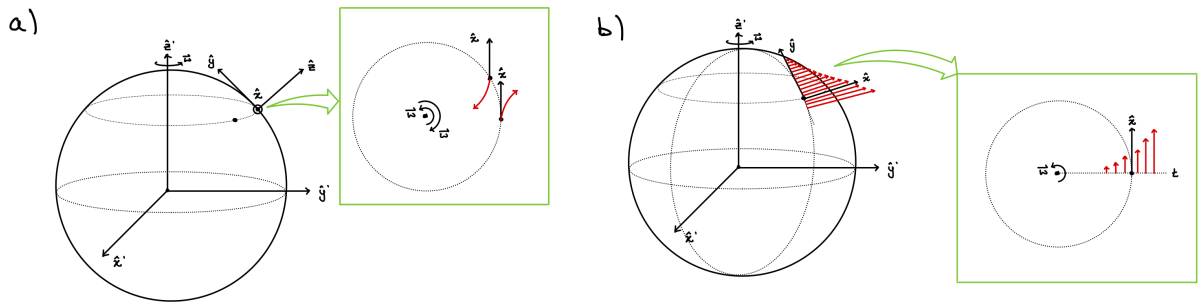


Figure 42: Visualization of the Coriolis effects on a projectile launched from the surface of the Earth.

What are the consequences of this result? We can, in fact, see that each term in our solution contributions to a different effect, which, when added together, cause the total Coriolis force exerted on the particle.

Let's look first at y_1 and z_1 terms of our first order correction. These effects are analogous to a target on a rotating turntable. Specifically, it is the portion of the force which causes the particle to rotate about the axis, $\vec{\omega}$.

The x_1 term in our first order correction has two contributions. The first is the term in v_{y0}. This term acts as a *velocity gradient* from the North to South pole of Earth. In particular, if you were to release a particle at some angle, θ, on the surface of the Earth, its N-S deflection is governed by this term.

The second contribution is in the v_{z_0} term in x_1. This is our *vertical velocity gradient*. At some initial time, t_0, if you were to increase the angle at which the particle was released, we would reach a *maximum* velocity oriented tangential to the surface of the Earth.

These three terms, when combined, create the entire Coriolis force, which is responsible for the deflection of our particle within the non-inertial frame.

4.3.3 The Rotating Saddle Potential

Consider a particle of mass, m, constrained to move in 2D under a saddle function potential of the form

$$V(x, y) = \frac{1}{2}m\omega^2(x^2 - y^2).$$

Here, we will state that the potential has an unstable equilibrium at the origin, and that particles will be bound only if $y = 0$; any small perturbation in y will diverge and cause the particle to escape.

We will skip the setup and instead state the Lagrangian assuming that the potential is static (time-independent).

$$\mathcal{L} = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - \frac{1}{2}m\omega^2(x^2 - y^2).$$

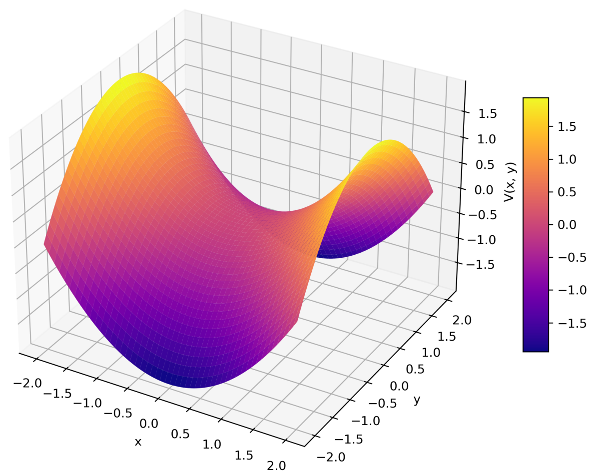


Figure 43: Saddle potential $V(x, y)$.

Computing our partials, we find that

$$\frac{\partial \mathcal{L}}{\partial x} = -m\omega^2 x, \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\ddot{x}$$
$$\frac{\partial \mathcal{L}}{\partial y} = m\omega^2 y, \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{y}} = m\ddot{y}$$

and so our EOMs are

$$\boxed{\begin{aligned} \ddot{x} &= -\omega^2 x \\ \ddot{y} &= \omega^2 y. \end{aligned}}$$

Here, notice that our EOMs display two types of motion. The EOM in x is oscillatory with SHM, but the EOM in y is exponential and unstable. Now, let us consider rotating the potential with a constant frequency, Ω, about the z-axis. In this case, we are interested in the EOM and how the motion evolves in time.

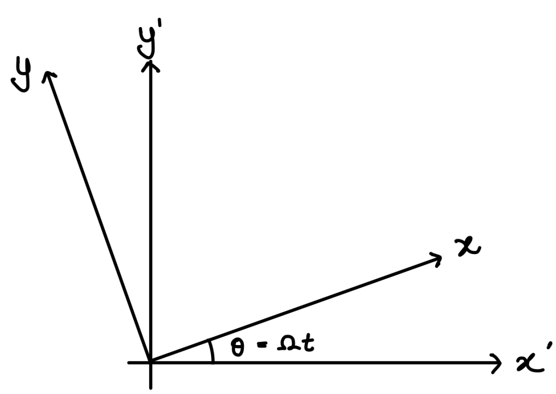


Figure 44: Rotating reference frame, x_i, superimposed on top of the inertial reference frame, x_i', with time-varying angle, $\theta = \Omega t$.

To proceed with our analysis in the rotating frame, we can apply a passive rotation transformation from the inertial to fixed, and fixed to inertial

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\Omega t) & \sin(\Omega t) \\ -\sin(\Omega t) & \cos(\Omega t) \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}, \qquad \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\Omega t) & -\sin(\Omega t) \\ \sin(\Omega t) & \cos(\Omega t) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

To see which frame is easier to work with, we can write the potential in both systems, which are

$$\begin{aligned} V(x,y) &= \frac{1}{2}m\omega^2(x^2-y^2) \\ V(x',y') &= \frac{1}{2}m\omega^2\big[(x'^2-y'^2)\cos(2\Omega t)+2x'y'\sin(2\Omega t)\big]. \end{aligned}$$

Clearly, we can see that V is simpler in the rotating frame. On the other hand,

$$\begin{aligned} T(x,y) &= \frac{1}{2}m\big[\dot{x}^2+\dot{y}^2+2(x\dot{y}-y\dot{x})\Omega+(x^2+y^2)\Omega^2\big] \\ T(x',y') &= \frac{1}{2}m(\dot{x}'^2+\dot{y}'^2) \end{aligned}$$

where T is simpler in the fixed frame. We will proceed with our analysis in the rotating frame. However, we will show that there are two methods to calculating T in the rotating frame.

Method 1. We could use the rotation matrix from the non-inertial to the inertial frame. Here, we find that

$$\begin{aligned} \dot{x}' &= \cos(\Omega t)\dot{x}-\Omega\sin(\Omega t)x-\sin(\Omega t)\dot{y}-\Omega\cos(\Omega t)y = \cos(\Omega t)(\dot{x}-\Omega y)-\sin(\Omega t)(\dot{y}+\Omega x) \\ \dot{y}' &= \sin(\Omega t)\dot{x}+\Omega\cos(\Omega t)x+\cos(\Omega t)\dot{y}-\Omega\sin(\Omega t)y = \sin(\Omega t)(\dot{x}-\Omega y)+\cos(\Omega t)(\dot{y}+\Omega x). \end{aligned}$$

Thus, the kinetic energy is

$$\dot{x}'^2+\dot{y}'^2 = (\dot{x}-\Omega y)^2+(\dot{y}+\Omega x)^2.$$

Method 2. We could, on the other hand, start with our general relation for velocity and compute the cross product. Here $\mathbf{V}=0$, since the origin does not translate, and so we have

$$\begin{aligned} |\mathbf{v}|^2 = |\dot{\mathbf{r}}+\vec{\omega}\times\mathbf{r}|^2 = |[\dot{x},\dot{y},0]+[0,0,\Omega]\times[x,y,z]|^2 &= |[\dot{x},\dot{y},0]+[-\Omega y,\Omega x,0]|^2 \\ &= (\dot{x}-\Omega y)^2+(\dot{y}+\Omega x)^2 \end{aligned}$$

as we have arrived at before. Computing the Lagrangian, we find

$$\mathcal{L} = \frac{1}{2}m\big[\dot{x}^2+\dot{y}^2+2(x\dot{y}-y\dot{x})\Omega+(x^2+y^2)\Omega^2\big]-\frac{m}{2}\omega^2\big(x^2-y^2\big)$$

and, evaluating the partials, we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}} &= m\ddot{x}-m\Omega\dot{y} \\ \frac{\partial\mathcal{L}}{\partial x} &= m\Omega\dot{y}+m\Omega^2x-m\omega^2x \end{aligned}$$

and in y,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{y}} &= m\ddot{y}+m\Omega\dot{x} \\ \frac{\partial\mathcal{L}}{\partial y} &= -m\Omega\dot{x}+m\Omega^2y+\omega^2y. \end{aligned}$$

Thus, our ELEs read

$$\boxed{\begin{aligned}\ddot{x} &= \Omega^2 x + 2\Omega\dot{y} - \omega^2 x \\ \ddot{y} &= \Omega^2 y - 2\Omega\dot{x} + \omega^2 y.\end{aligned}}$$

Notice that we have two non-inertial terms here. $\Omega^2 x$ is due to the centrifugal force, while $2\Omega\dot{y}$ is a result of the Coriolis force. We can now make the exponential ansatz, $\mathbf{r} = \mathbf{c}e^{-i\lambda t}$, to which we find

$$-\lambda^2\mathbf{I}\begin{bmatrix}c_1\\c_2\end{bmatrix} = \begin{bmatrix}\Omega^2-\omega^2\\\Omega^2+\omega^2\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix} + i\lambda\mathbf{I}\begin{bmatrix}-2\Omega\\2\Omega\end{bmatrix}\begin{bmatrix}c_2\\c_1\end{bmatrix}.$$

Thus,

$$0 = \begin{vmatrix}\lambda^2+\Omega^2-\omega^2 & -i2\Omega\lambda\\ i2\Omega\lambda & \lambda^2+\Omega^2+\omega^2\end{vmatrix} = \lambda^4 - 2\Omega^2\lambda^2 + (\Omega^4-\omega^4)$$

which has the solutions

$$\lambda^2 = \Omega^2 \pm \omega^2.$$

Notice that the solutions do *not* grow exponentially, so require oscillatory solutions, and thus all values of λ must be real. However, this *only* occurs in the case that

$$\Omega \geq \omega.$$

That is, the potential function must rotate faster than the natural frequency of our solutions. This process, although long, is an example of **linear stability analysis**. We could solve this analytically, but we will instead look at a numerical solution. For that, we must define the Lagrangian in the inertial frame, which we simply state is

$$\mathcal{L}' = \frac{1}{2}m(\dot{x}'^2+\dot{y}'^2) - \frac{1}{2}m\omega^2\big[(x'^2-y'^2)\cos(2\Omega t) + 2x'y'\sin(2\Omega t)\big]$$

to give the EOMs

$$\boxed{\begin{aligned}\ddot{x}' &= -\omega^2\big[x'\cos(2\Omega t) + y'\sin(2\Omega t)\big]\\ \ddot{y}' &= -\omega^2\big[-y'\cos(2\Omega t) + x'\sin(2\Omega t)\big].\end{aligned}}$$

Using the initial conditions $x(0) = y(0) = \dot{y}(0) = 0$, $\dot{x}(0) = 0.01$, and parameters $\Omega = 0.99$, $\Omega = 2.2$, and $\omega = 1$ we can see that the motion of a particle within the saddle potential varies largely for both cases. $\Omega = 0.99 < \omega = 1$ has the particle, as a function of $y(x)$, spiralling away from the origin. Furthermore, a plot of its displacement over time shows that near the end of the time frame, the particle travels further from 0.

In the $\Omega = 2.2 > \omega = 1$ case, the particle instead creates a complex "micro-motion" in the $y(x)$ plane. In fact, it creates circular patterns, in which it rotates about the origin.

This type of motion is actually related to an important application in physics called **the Paul Trap**. Wolfgang Paul, who won the Novel Prize in Physics, circumvented Earnshaw's Theorem—which forbids trapping charged particles in 3D with static fields—by instead considering an oscillating electric field to trap the charged particles. Here, the potential is of a different form, but is non-inertial and rotates to trap the particle under a similar condition.

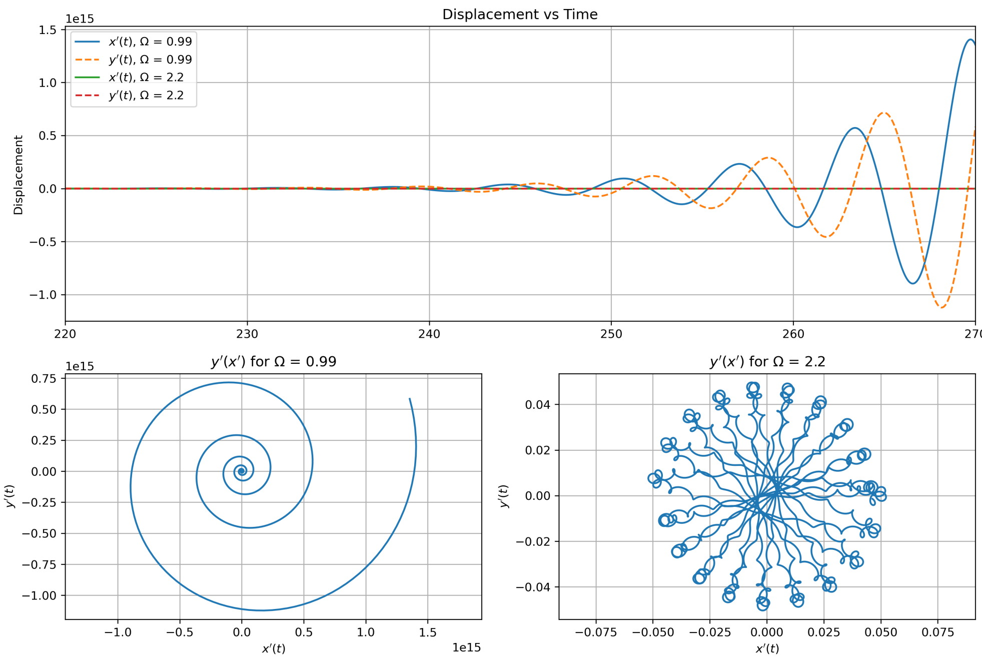


Figure 45: Three plots of a particle subject to $V(x, y)$ for parameters $\Omega = 0.99$ and $\Omega = 2.2$.

4.3.4 The Foucault Pendulum

Consider a pendulum, free to swing in 2D, but is initially set to swing in a plane. The pendulum is of length, l, with a mass, m, hanging somewhere on Earth at a colatitude, θ. We will investigate the properties of this pendulum, called a **Foucault pendulum**.

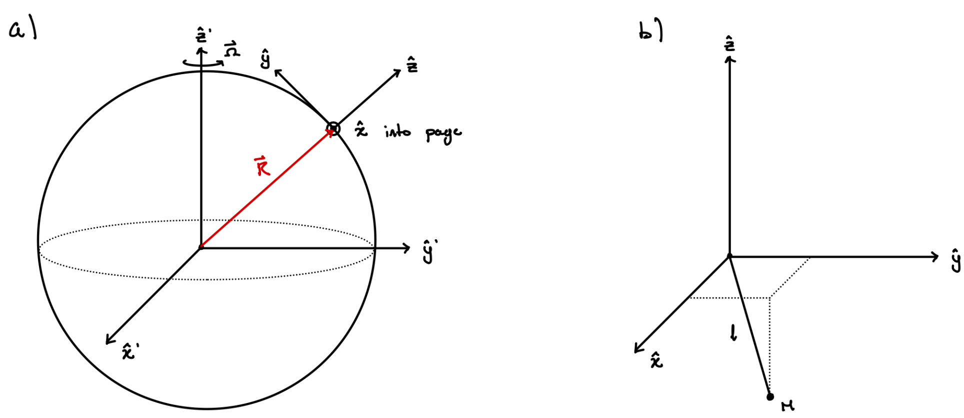


Figure 46: (a) Reference frame relative to Earth. (b) Foucault pendulum as seen by an observed on Earth.

Let us assume that the Earth is only rotating, that is, $\mathbf{V} = 0$, and that we are only interested in the first order in $\vec{\Omega}$. Here, our Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}\Big[|\dot{\mathbf{r}}|^2 + 2m\dot{\mathbf{r}} \cdot (\vec{\Omega} \times \mathbf{r})\Big] - U(\mathbf{r})$$

for some non-inertial potential, $U(\mathbf{r})$. Assuming the constraint of a pendulum, such that $g(r) = r - l = 0$, we find that

$$z = -l\sqrt{1 - \frac{x^2}{l^2} - \frac{y^2}{l^2}}$$

where we have chosen the negative sign since the pendulum lies below the z-axis. Furthermore, we will assume that x and y are small, such that by the binomial approximation,

$$z \approx -l\left(1 - \frac{x^2}{2l^2} - \frac{y^2}{2l^2}\right) \implies \dot{z} = \frac{\dot{x}x}{l} + \frac{\dot{y}y}{l}.$$

We can write out the individual terms in the Lagrangian and find that

$$\frac{1}{2}m|\dot{\mathbf{r}}|^2 = \frac{1}{2}m\Big[\dot{x}^2 + \dot{y}^2 + \Big(\frac{x}{l}\dot{x} + \frac{y}{l}\dot{y}^2\Big)\Big]$$

and keeping only leading order terms in x/l and y/l, we have

$$\frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \approx \frac{1}{2}m(\dot{x}^2 + \dot{y}^2).$$

Take note that this is fine, since we are using a perturbative approach. In fact, we can find the first order correction term, which is

$$m\dot{\mathbf{r}} \cdot (\vec{\Omega} \times \mathbf{r}) = m\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{z}\end{bmatrix} \cdot \begin{bmatrix}0\\ \Omega\sin(\theta)\\ \Omega\cos(\theta)\end{bmatrix} \times \begin{bmatrix}x\\ y\\ z\end{bmatrix} = m\Omega[\sin(\theta)z\dot{x} - \cos(\theta)y\dot{x} + \cos(\theta)x\dot{y} - \sin(\theta)x\dot{z}].$$

We can furthermore substitute our constraint relation in z to find that the quantities dependent on $z\dot{x}$ and $x\dot{z}$ are *not* leading order in x/l and y/l, and since $l\sin(\theta)\dot{x}$ is a full time derivative, we can drop it to simplify this term to

$$m\dot{\mathbf{r}} \cdot (\vec{\Omega} \times \mathbf{r}) \approx m\Omega[-\cos(\theta)y\dot{x} + \cos(\theta)x\dot{y}].$$

Evaluating our potential, we have

$$U(\mathbf{r}) = mgz = -mgl\left(1 - \frac{x^2}{2l^2} - \frac{y^2}{2l^2}\right) = \frac{mg}{2l}(x^2 + y^2).$$

Thus, we can approximate our Lagrangian to be of the simplified form,

$$\mathcal{L} \approx \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + m\Omega[-\cos(\theta)y\dot{x} + \cos(\theta)x\dot{y}] - \frac{mg}{2l}\big(x^2 + y^2\big).$$

Notice that this Lagrangian describes a 2D harmonic oscillator with a Coriolis correction term. So, we expect our pendulum motion to be slightly unnatural. Regardless, our partials in x are

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial\mathcal{L}} &= m\ddot{x} - m\Omega\cos(\theta)\dot{y}\\ \frac{\partial\mathcal{L}}{\partial x} &= m\Omega\cos(\theta)\dot{y} - \frac{mgx}{l}\end{aligned}$$

and in y,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{y}} &= m\ddot{y} + m\Omega\cos(\theta)\dot{x} \\ \frac{\partial\mathcal{L}}{\partial y} &= -m\Omega\cos(\theta)\dot{x} - \frac{mgy}{l}.\end{aligned}$$

Thus, we arrive at the EOMs

$$\boxed{\begin{aligned}\ddot{x} &= 2\Omega\cos(\theta)\dot{y} - \omega^2 x \\ \ddot{y} &= -2\Omega\cos(\theta)\dot{x} - \omega^2 y.\end{aligned}}$$

Here is a new problem. We have already used our perturbative approach, so we've already made several assumptions on our system. Ideally, we would like to solve these EOMs with a different method. To do this, we will assume that we can write an ansatz

$$\eta = x + jy \implies \dot{\eta} = \dot{x} + j\dot{y} \implies \ddot{\eta} = \ddot{x} + j\ddot{y}$$

for some constant, j. Thus, we have

$$\ddot{\eta} = \ddot{x} + j\ddot{y} = 2\Omega\cos(\theta)\dot{y} - 2j\Omega\cos(\theta)\dot{x} - \omega^2 x - j\omega^2 y = -2j\Omega\cos(\theta)(\dot{x} + j\dot{y}) - \omega^2(x + j\dot{y})$$

and so we find

$$\ddot{\eta} = -2j\Omega\cos(\theta)\dot{\eta} - \omega^2\eta.$$

Here, we have ignored the portions in x since we are not interested in them.[9] Now, we will make the ansatz $\vec{\eta} = \mathbf{c}e^{-i\lambda t}$, such that our characteristic equation is

$$\lambda^2 - 2\Omega\cos(\theta)\lambda - \omega^2 = 0 \implies \lambda = \left[\Omega\cos(\theta) \pm \sqrt{\Omega^2\cos^2(\theta) + \omega^2}\right].$$

We will now assume that $\omega \gg \Omega$ (lots of assumptions here), such that

$$\lambda \approx \Omega\cos(\theta) \pm \omega$$

and so we arrive at the approximate solution,

$$\eta = c_1 e^{-i[\Omega\cos(\theta)+\omega]t} + c_2 e^{-i[\Omega\cos(\theta)-\omega]t}.$$

Assuming the initial conditions $x(0) = \eta(0) = \varepsilon$, $y(0) = \dot{x}(0) = \dot{y}(0) = \dot{\eta}(0) = 0$, we find the system of equations

$$\begin{aligned}\eta(0) &= c_1 + c_2 = \varepsilon \\ \dot{\eta}(0) &= c_1[-i\Omega\cos(\theta) + \omega] + c_2[-i\Omega\cos(\theta) - \omega] = 0.\end{aligned}$$

However, we made the previous assumption that $\omega \gg \Omega$, so the second condition shows that

$$\dot{\eta}(0) \approx c_1\omega - c_2\omega = 0 \implies c_1 = c_2$$

and so

$$c_1 = c_2 = \frac{\varepsilon}{2}.$$

[9] Gotta check this later.

Thus, we find that

$$\eta = \frac{\varepsilon}{2} e^{-i\Omega\cos(\theta)t} e^{-i\omega t} + \frac{\varepsilon}{2} e^{-i\Omega\cos(\theta)t} e^{i\omega t} = \varepsilon e^{-i\Omega\cos(\theta)t} \cos(\omega t)$$

to show that our solution is of the form

$$\boxed{\eta = x + jy = \varepsilon\cos\left[\Omega\cos(\theta)t\right]\cos(\omega t) - j\sin\left[\Omega\cos(\theta)t\right]\cos(\omega t).}$$

Here, we can see that $x(t) = \varepsilon\cos\left[\Omega\cos(\theta)t\right]\cos(\omega t)$, and $y(t) = -\sin\left[\Omega\cos(\theta)t\right]\cos(\omega t)$, which suggests that our pendulum has a relatively fast oscillation compared to a normal pendulum with a *slower precession* at a modulated frequency, $\Omega\cos(\theta)$.

If we consider the Coriolis effect on the different hemispheres on Earth, we will find that the pendulum will precess in the opposite direction to that of the other hemisphere. We can analyze this further by taking the limits of our result. In fact, if we sit at the equation with our Foucault pendulum, $\Omega\cos(\theta) \to \infty$, which suggests that the period of our pendulum is infinite. The minimum case is at the poles of Earth, where we have $\Omega\cos(\theta) \to \Omega$, where our period would be a single day.

5 Hamiltonian Formalism

Let us now begin our discussion on Hamiltonian mechanics. First, we should recall our expression for the generalized momentum, which is

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}.$$

This might not look like much, but this quantity is *central* in Hamiltonian mechanics. However, to discuss Hamiltonian mechanics, we must start with Lagrangian mechanics. Here, we have $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i, t)$ as before, and we can construct, if you remember, the **Hamiltonian** which is

$$\boxed{\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L}.} \tag{29}$$

The difference, however, between Lagrangian and Hamiltonian mechanics is that it is **essential** to write our Hamiltonian in the form

$$\mathcal{H} = \mathcal{H}(q_i, p_i, t)$$

where p_i is our conjugate momentum. In particular, we must eliminate $\dot{q}_i$, by writing our generalized velocities in the form

$$\dot{q}_i = \dot{q}_i(\{q_i\}, \{p_i\}, t)$$

to transform our Lagrangian to the Hamiltonian, such that $\mathcal{L} \to \mathcal{H}$. Notice that we have switched to configuration space to stress that we must write our *all* our generalized velocities in terms of our conjugate momenta.

You may recall that if the Lagrangian is time-independent, then the Hamiltonian is conserved. We can expand this further by stating that *if our generalized coordinates are related to the Cartesian coordinates by a time-independent transformation, then the Hamiltonian corresponds to the total energy within our system.*

Thus, we expect that if the Lagrangian is time-independent, then so is our Hamiltonian. What if this is not the case? Then we can simply take the time derivative of the Hamiltonian, and find that

$$\frac{\partial \mathcal{H}}{\partial t} = \sum_{i=1}^{n} p_i \frac{\partial \dot{q}_i}{\partial t} - \frac{\partial \mathcal{L}}{\partial t} - \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial t}.$$

However, we know that $p_i = \partial \mathcal{L} / \partial \dot{q}_i$, so we find

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}$$

which is non-zero, and, in fact, depends on the Lagrangian. Again, we stress that we have written the Hamiltonian in terms of our conjugate momenta as well, so we really started with

$$\mathcal{H} = \mathcal{H}(q_i, p_i, t) = \sum_{i=1}^{n} p_i [\dot{q}_i(q_i, p_i, t)] - \mathcal{L}[q_i, \dot{q}_i(q_i, p_i, t), t]$$

to arrive at a consistent result. With this new perspective, we can find a third method to solve classical mechanics problems in addition to Newtonian mechanics and Lagrangian mechanics.

5.1 Constructing a Hamiltonian

To see how to work with Hamiltonians, let us consider a simple example. Suppose we have a Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$$

such that

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} \implies \dot{x} = \frac{p}{m}.$$

It only remains that we construct our Hamiltonian, which is simply

$$\mathcal{H} = p\dot{x} - \mathcal{L} = p\Big[\frac{p}{m}\Big] - \Big[\frac{1}{2}m\dot{x}^2 - V(x)\Big] = \frac{p^2}{m} - \frac{1}{2}\frac{p^2}{m} + V(x) = \frac{p^2}{2m} + V(x).$$

Immediately, you should notice that $p^2/2m = T$, which is in terms of linear momentum; and, in fact, $\mathcal{H} = T+V = E$, where E is the total energy in our system. This result arises because the Lagrangian that we considered was time-independent, and we have our generalized coordinates related to the Cartesian coordinates. Furthermore, our potential is time-independent, so our Hamiltonian has zero explicit time-independence.

5.1.1 The Electromagnetic Hamiltonian

Let us revisit our electromagnetic setup of the Lagrangian, where

$$\mathcal{L} = \frac{1}{2}m|\mathbf{v}|^2 + q\mathbf{v}\cdot\mathbf{A} - q\phi$$

where $\mathbf{A}$ is some vector potential and ϕ is some scalar potential. We can expand our Lagrangian in terms of the Cartesian coordinates, to find that

$$\mathcal{L} = \frac{1}{2}m\big[\dot{x}^2 + \dot{y}^2 + \dot{z}^2\big] + q[\dot{x}A_x(\mathbf{x},t) + \dot{y}A_y(\mathbf{x},t) + \dot{z}A_z(\mathbf{x},t)] - q\phi(x,t).$$

To construct our Hamiltonian with this Lagrangian, we now need to include the other generalized coordinates. In particular, we find that in x,

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} + qA_x \implies \dot{x} = \frac{1}{m}(p_x - qA_x).$$

Similarly, in y and z, we should find that

$$\dot{y} = \frac{1}{m}(p_y - qA_y), \quad \dot{z} = \frac{1}{m}(p_z - qA_z).$$

Thus, our Hamiltonian is

$$\begin{aligned}
\mathcal{H} &= \sum_{i=1}^{3} p_i\dot{q}_i - \mathcal{L} \\
&= \frac{p_x}{m}(p_x - qA_x) + \frac{p_y}{m}(p_y - qA_y) + \frac{p_z}{m}(p_z - qA_z) \\
&\qquad - \frac{1}{2}m\left[\frac{1}{m^2}(p_x - qA_x)^2) + \frac{1}{m^2}(p_y - qA_y)^2) + \frac{1}{m^2}(p_z - qA_z)^2)\right] + q\phi \\
&= \frac{|\mathbf{p} - q\mathbf{A}|^2}{2m} + q\phi(x,t).
\end{aligned}$$

This result, if we had attempted to guess it from the Lagrangian, is not obvious. It stands to reason that although the Hamiltonian is still of the form $\mathcal{H} = T + V = E$, when dealing with multiple coordinates, we must be careful in determining the form of the Hamiltonian.

5.2 Hamilton's Equations of Motion

Hamiltonian mechanics also derives its own set of equations of motion. To derive them, we will begin our Hamiltonian for some number of generalized coordinates,

$$\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}_i(q_i, p_i, t) - \mathcal{L}[q_i, \dot{q}_i(q_i, p_i, t), t].$$

To approach this, we will expand the sum in terms of its components. Furthermore, we will also drop the arguments, but keep note that we are doing this derivation in terms of the conjugate momenta. We have

$$\mathcal{H} = p_1\dot{q}_1 + p_2\dot{q}_i + \cdots + p_n\dot{q}_n - \mathcal{L}[q_i, \dot{q}_i(q_i, p_i, t), t].$$

For the first EOM, we will take the derivative of $\mathcal{H}$ with respect to q_1[10] to find that

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial q_1} &= p_1\frac{\partial \dot{q}_1}{\partial q_1} + p_2\frac{\partial \dot{q}_2}{\partial q_1} + \cdots + p_n\frac{\partial \dot{q}_n}{\partial q_1} - \left[\frac{\partial \mathcal{L}}{\partial q_1} + \frac{\partial \mathcal{L}}{\partial \dot{q}_1}\frac{\partial \dot{q}_1}{\partial_1} + \frac{\partial \mathcal{L}}{\partial \dot{q}_2}\frac{\partial \dot{q}_2}{\partial q_1} + \cdots + \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\frac{\partial \dot{q}_n}{\partial q_1}\right] \\
&= \cancel{p_1\frac{\partial \dot{q}_1}{\partial q_1}} + \cancel{p_2\frac{\partial \dot{q}_2}{\partial q_1}} + \cdots + \cancel{p_n\frac{\partial \dot{q}_n}{\partial q_1}} - \left[\frac{\partial \mathcal{L}}{\partial q_1} + \cancel{p_1\frac{\partial \dot{q}_1}{\partial q_1}} + \cancel{p_2\frac{\partial \dot{q}_2}{\partial q_1}} + \cdots \cancel{p_n\frac{\partial \dot{q}_n}{\partial q_1}}\right] \\
&= -\frac{\partial \mathcal{L}}{\partial q_1}.
\end{aligned}$$

However, by the ELE, this simply becomes

$$\frac{\partial \mathcal{H}}{\partial q_1} = -\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_1} = -\dot{p}_1.$$

We can repeat this process for other coordinates, $q_2, q_3, \cdots, q_n$, and find that in the most general case,

$$\boxed{\frac{\partial \mathcal{H}}{\partial q_i} = -\dot{p}_i.} \tag{30}$$

Applying the same process when taking the derivative of $\mathcal{H}$ with respect to p_1,

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial p_1} &= \dot{q}_1 + p_1\frac{\partial \dot{q}_1}{\partial p_1} + p_2\frac{\partial \dot{q}_2}{\partial p_1} + \cdots + p_n\frac{\partial \dot{q}_n}{\partial p_1} - \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_1}\frac{\partial \dot{q}_1}{\partial p_1} + \frac{\partial \mathcal{L}}{\partial \dot{q}_2}\frac{\partial \dot{q}_2}{\partial p_1} + \cdots + \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\frac{\partial \dot{q}_n}{\partial p_1}\right] \\
&= \dot{q}_1 + \cancel{p_1\frac{\partial \dot{q}_1}{\partial p_1}} + \cancel{p_2\frac{\partial \dot{q}_2}{\partial p_1}} + \cdots + \cancel{p_n\frac{\partial \dot{q}_n}{\partial p_1}} - \left[\cancel{p_1\frac{\partial \dot{q}_1}{\partial p_1}} + \cancel{p_2\frac{\partial \dot{q}_2}{\partial p_1}} + \cdots + \cancel{p_n\frac{\partial \dot{q}_n}{\partial p_1}}\right] \\
&\dot{q}_1.
\end{aligned}$$

As before, in the most general case, we have

$$\boxed{\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i.} \tag{31}$$

[10] Thornton and Marion approach this slightly differently. Here, our derivation is more informal, but it gets the job done.

Together, (30) and (31) comprise Hamilton's equations of motion.

You may notice that comparing this to the the Euler-Lagrange equation, that it takes Hamiltonian mechanics an extra equation for the entire system. In fact, we can find that Lagrangian mechanics gives n second order DEs, describing the trajectories, $q_i(t)$, of a given particle within a system. On the other hand, Hamiltonian mechanics gives $2n$ first order DEs, which describe a new quantity, which is called **phase space**, and depends on both the trajectory of a particle and its momentum, $p_i(t)$.

5.3 Liouville's Theorem

Previously, we briefly introduced the concept of phase space.[11] In Lagrangian mechanics, we implicitly define an n-dimensional space, called a *configuration space* where every generalized coordinate, q_i, represents a state of a particle at some time, t. Similarly, we also implicitly define another n-dimensional *momentum space* for every generalized momentum, p_i, which describes every possible motion of a particle at some time, t. Furthermore, the value of n is determined by the number of generalized coordinates within the system.

We can similarly define a space for Hamiltonian mechanics, which utilize both coordinates. In particular, we say that the **phase space** of a particle is a $2n$-dimensional space consisting of both the generalized coordinates, q_i, and generalized momenta, p_i. In principle, if we knew at some given time, the position ane momenta of a particle within a system, then together with the IC's we know the motion of the entire system.

This can become increasingly taxing when consider an ensemble of particles, so we will aim to determine another quantity of a particle within phase space. Beginning with some arbitrary particle in phase space, we say that it has the coordinates

$$\mathbf{z} = [q_i, p_i].$$

Following this, we define the velocity of these coordinates as

$$\dot{\mathbf{z}} = [\dot{q}_i, \dot{p}_i] = \left[\frac{\partial \mathcal{H}}{\partial p_i}, -\frac{\partial \mathcal{H}}{\partial q_i}\right]$$

which is given by Hamilton's equations of motion. Let us now consider a collection of closely (densely packed) space points in phase space which can be enclosed by a surface.[12] Enclosed by this surface is the volume of the points within phase space. Generally, we expect that the surface of the points in phase space to vary as the system evolves in time. However, we want to see if the volume remains changes as well.

[11] This concept of *space* (in my opinion) is rather hard to rationalize intuitively. To see more on the formalism of this, consider reading Appendix A.

[12] In 2D, we refer to a surface as a *closed path*, and the volume within the surface is the corresponding *area* within the path. This, in fact, is linked to Stokes' Theorem.

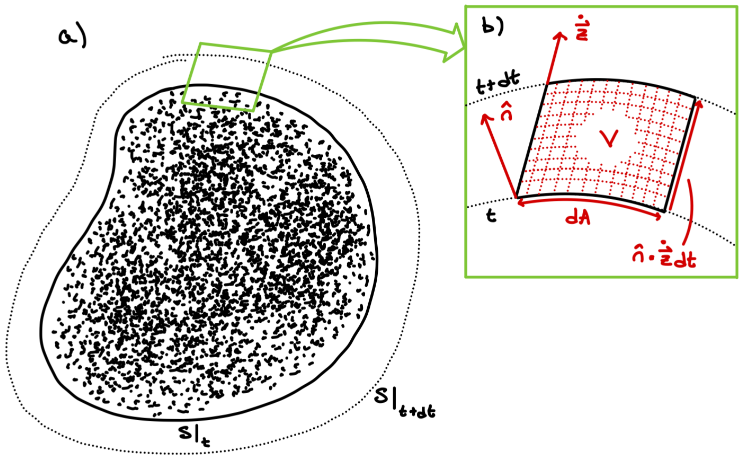


Figure 47: (a) Arbitrary surface, S, enclosing an collection of points. (b) Microscopic view of the surface at time $t + \mathrm{d}t$.

We can define the change in volume as

$$\mathrm{d}V = \iint_S (\mathbf{n} \cdot \dot{\mathbf{z}}\,\mathrm{d}t)\,\mathrm{d}A$$

and, dividing by the time differential,

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \iint_S (\mathbf{n} \cdot \dot{\mathbf{z}})\,\mathrm{d}A\,.$$

However, by the Divergence Theorem, we can rewrite the double integral in S as

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \iint_S (\mathbf{n} \cdot \dot{\mathbf{z}})\,\mathrm{d}A = \iiint_V \left(\vec{\nabla} \cdot \dot{\mathbf{z}}\right)\mathrm{d}V\,.$$

For phase space, we can define the divergence of $\dot{\mathbf{z}}$ as[13]

$$\vec{\nabla} \cdot \dot{\mathbf{z}} = \sum_{i=1}^{n} \left[\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i}\right]$$

so that we can turn our time derivative of the volume into

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \iiint_V \sum_{i=1}^{n} \left[\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i}\right]\mathrm{d}V\,.$$

However, invoking Hamilton's EOMs, we can find that this term simplifies to

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \iiint_V \sum_{i=1}^{n} \left[\frac{\partial}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} + \frac{\partial}{\partial p_i}\left(-\frac{\partial \mathcal{H}}{\partial q_i}\right)\right]\mathrm{d}V = \iiint_V \sum_{i=1}^{n} [0]\,\mathrm{d}V = 0$$

[13] Be careful here. We are working in phase space, so the gradient $\vec{\nabla}$ is really

$$\vec{\nabla} = \frac{\partial}{\partial q_i} + \frac{\partial}{\partial p_i}$$

which is completely analogous to the Cartesian version.

which is zero. Thus, we expect that the volume of the surface enclosing a set of closely spaced points, in phase space, is conserved under a Hamiltonian evolution. That is, phase space density is conserved, such that

$$\boxed{\frac{\mathrm{d}V}{\mathrm{d}t} = 0} \tag{32}$$

which is precisely **Liouville's Theorem**.

5.4 Poisson Brackets

A relatively important method of determining conserved quantities is by the use of **Poisson Brackets**. Not only are they useful in classical mechanics, they are analogous in almost every fashion to quantum commutators.

To see this, we will consider two functions of dynamical variables, such that

$$\begin{aligned} f &= f(\{q_i\}, \{p_i\}, t) \\ g &= g(\{q_i\}, \{p_i\}, t). \end{aligned}$$

The Poisson Bracket of the two functions, f and g is given by

$$\boxed{\{f, g\} = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right].} \tag{33}$$

There are, in fact, useful properties of the Poisson Bracket. In particular,

1. Anti-commutativity property:
$$\{f, g\} = -\{g, f\}.$$

2. Bilinearity property:
$$\begin{aligned} \{af + bg, h\} &= a\{f, h\} + b\{g, h\} \\ \{h, af + bg\} &= a\{h, f\} + b\{h, g\} \end{aligned}$$
for some constants, $a, b \in \mathbb{R}$.

3. Leibniz rule:
$$\begin{aligned} \{gf, h\} &= f\{g, h\} + \{f, h\}g. \\ \{f, gh\} &= g\{f, h\} + \{f, g\}h. \end{aligned}$$

4. Jacobi identity:
$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0.$$

Most of these proofs are pretty annoying, so we will only show a case of the bilinearity property to convince you that the others work. Here,

$$\begin{aligned}
\{af+bg,h\} &= \sum_{i=1}^{n}\left[\frac{\partial(af+bg)}{\partial q_i}\frac{\partial h}{\partial p_i}-\frac{\partial(af+bg)}{\partial p_i}\frac{\partial h}{\partial q_i}\right]\\
&= \sum_{i=1}^{n}\left[\frac{\partial(af)}{\partial q_i}+\frac{\partial(bg)}{\partial q_i}\right]\frac{\partial h}{\partial p_i}-\sum_{i=1}^{n}\left[\frac{\partial(af)}{\partial p_i}+\frac{\partial(bg)}{\partial p_i}\right]\frac{\partial h}{\partial q_i}\\
&= \sum_{i=1}^{n}a\left[\frac{\partial f}{\partial q_i}\frac{\partial h}{\partial p_i}-\frac{\partial g}{\partial p_i}\frac{\partial h}{\partial q_i}\right]+\sum_{i=1}^{n}b\left[\frac{\partial g}{\partial q_i}\frac{\partial h}{\partial p_i}-\frac{\partial g}{\partial p_i}\frac{\partial h}{\partial q_i}\right]\\
&= a\{f,h\}+b\{g,h\}
\end{aligned}$$

as before. We hinted at this earlier, but if you are observant, you may notice that these identities are exactly analogous to the commutator identities in quantum mechanics. In quantum mechanics, the identities are of the form

$$\begin{aligned}
\left[\hat{A},\hat{B}\right] = -\left[\hat{B},\hat{A}\right] &\sim \{f,g\} = -\{g,f\}\\
\left[\hat{A}\hat{B},\hat{C}\right] = \hat{A}\left[\hat{B},\hat{C}\right]+\left[\hat{A},\hat{C}\right]\hat{B} &\sim \{gf,h\} = f\{g,h\}+\{f,h\}g.
\end{aligned}$$

The only caveat is that in quantum mechanics, the arguments of the commutator relations do not permute. With this new formulation, we can further find a few important Poisson bracket relations. In particular, we find

1. Two generalized coordinates,
$$\{q_i,q_j\}=0.$$
2. Two generalized momenta,
$$\{p_i,p_j\}=0.$$
3. A generalized coordinate and a generalized momenta,
$$\{q_i,p_j\}=\delta_{ij}.$$
4. A generalized coordinate and a momenta to the power, n,
$$\left\{q_i,p_j^n\right\}=np_j^{n-1}\delta_{ij}.$$
5. An arbitrary function of dynamical variables,
$$\{f,f\}=0.$$

Notice that the presence of the Kronecker delta in relation 3. In quantum mechanics, the analogous identity is multiplied by an $i\hbar$, which suggests the uncertain nature of the particles.

5.4.1 Time-evolution of the Properties Within a System

Suppose now that we have a system described by its corresponding Hamiltonian, $\mathcal{H}$. If we are interested in the time evolution of some arbitrary function within the system, $f = f(\{q_i\}, \{p_i\}, t)$, we can explicitly take its time derivative, which is

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial \dot{q}_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i \right] + \frac{\partial f}{\partial t} = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right] + \frac{\partial f}{\partial t}$$

where we have used Hamilton's equations of motion. Notice that the first term is the Poisson bracket of the function with its Hamiltonian, and so we find

$$\boxed{\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t}.} \tag{34}$$

Immediately, we can see the most important consequence of this result. If f is time independent, that is, $f = f(\{q_i\}, \{p_i\})$, *and* the Poisson bracket of f commutes with the Hamiltonian, we find that

$$\boxed{\{f, \mathcal{H}\} = 0 \implies \frac{\mathrm{d}f}{\mathrm{d}t} = 0}$$

which is simply the case that f is a conserved quantity. An extremely interesting connection of this time-evolution result is Heisenberg's equation of motion, which describes the time evolution of observables within the quantum system. There, his equation says that

$$\frac{\mathrm{d}\hat{A}}{\mathrm{d}t} = \frac{1}{i\hbar} \left[\hat{A}, \mathcal{H} \right] + \frac{\partial \hat{A}}{\partial t} \sim \frac{\mathrm{d}f}{\mathrm{d}t} = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t}$$

which is almost the exact same form if we consider the uncertain nature of the particles.

Hamilton's Equations from Poisson Brackets. We can backtrack to Hamilton's equations by the time-evolution of our generalized coordinates. In particular, for the first equation we see

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \{q_i, \mathcal{H}\} + \cancel{\frac{\partial q_i}{\partial t}} = \sum_{j=1}^{n} \left[\frac{\partial q_i}{\partial q_j} \frac{\partial \mathcal{H}}{\partial p_j} - \cancel{\frac{\partial q_i}{\partial p_j} \frac{\partial \mathcal{H}}{\partial q_j}} \right].$$

In the first term, we assume that the coordinates are time-independent, and for the second term we can see that $\partial q_i / \partial q_j = \delta_{ij}$. The second term cancels since q_j is independent of p_j, and so dropping the sum since we are only interested in the case $i = j$, we find

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}$$

as before. We can repeat the same process for the second equation of motion, such that

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \{p_i, \mathcal{H}\} + \cancel{\frac{\partial p_i}{\partial t}} = \sum_{j=1}^{n} \left[\cancel{\frac{\partial p_i}{\partial q_j} \frac{\partial \mathcal{H}}{\partial p_j}} - \frac{\partial p_i}{\partial p_j} \frac{\partial \mathcal{H}}{\partial q_j} \right]$$

where we have cancelled terms for the same reason as before and substituted $\partial p_i / \partial p_j = \delta_{ij}$. Thus, for the case $i = j$, we find

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}$$

as before.

5.5 Generalized Forces and Constraints in Hamiltonian Mechanics

Generally, we handle generalized and constraint forces using Lagrangian mechanics, instead building our constraints into the Hamiltonian. However, it is possible to handle them using the Hamiltonian formulation. The ELE's are modified by incorporating the generalized forces as

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = Q_i.$$

Recalling Hamilton's equations, we have

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}.$$

If we notice that the differential equation in p yields the result $\mathbf{F} = m\mathbf{a}$ when our conjugate momentum is linear and the Hamiltonian is the total system energy, we can immediately account for our generalized forces by adjusting this term specifically. In fact,

$$\boxed{\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} + Q_i.} \tag{35}$$

This, however, is ruins the symmetry of Hamilton's equations. Furthermore, since we now have a generalized momentum which is generalized-force dependent, Liouville's theorem can no longer apply. The particles at any point in phase space are affected by the generalized force, and thus do not grow as a constant.

We can confirm this adjustment gives the correct EOM by recalling that

$$\frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i}.$$

Furthermore, we also know that

$$\dot{p}_i = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}$$

so we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i} + Q_i$$

from Hamilton's second equation of motion. This result is, in fact, consistent with what we found before, so our guess at adding the addition Q_i term into the second EOM is valid.

However, we note that the generalized forces *must* be written in terms of the phase space coordinates for this to work. Otherwise the result would be inconsistent.

6 Examples in Hamiltonian Mechanics

Let us now revisit some of the examples we saw in Lagrangian mechanics and see how they translate to the Hamiltonian formulation.

6.1 Mass on a Spring

We will consider the familiar mass on a spring. Here, our Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2.$$

This Lagrangian is in terms of $(q_i, \dot{q}_i)$, so we need to find the conjugate momentum of the system to transform it into the Hamiltonian formulation. Here,

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} \implies \dot{x} = \frac{p}{m}.$$

Then, constructing our Hamiltonian, we have

$$\mathcal{H} = p\dot{x} - \mathcal{L} = p\Big[\frac{p}{m}\Big] - \Big[\frac{1}{2}m\Big(\frac{p}{m}\Big)^2 - \frac{1}{2}kx^2\Big] = \frac{p^2}{2m} + \frac{1}{2}kx^2.$$

Again, notice that the Hamiltonian is the total energy in the system: $\mathcal{H} = T + V = E$. We can now construct Hamilton's EOMs, which are

$$\begin{aligned} \dot{x} &= \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{m} \\ \dot{p} &= -\frac{\partial \mathcal{H}}{\partial x} = -kx \end{aligned}$$

so

$$\boxed{\begin{aligned} \dot{x} &= \frac{p}{m} \\ \dot{p} &= -kx. \end{aligned}}$$

You may notice that one of the equations repeats a previous definition. In fact, it will always be true that $\dot{q}_i = \partial\mathcal{H}/\partial p_i$ will repeat the definition $p_i = \partial\mathcal{L}/\partial\dot{q}_i$. While it is redundant, it does serve as a confirmation that you have not done wrong in constructing the Hamiltonian.

As our EOMs are a coupled set of first order DEs, we can make the exponential ansatz, $x = c_1 e^{i\lambda t}$ and $p = c_2 e^{-\lambda t}$. It's worth mentioning here that we did not simply throw this into a single vector. We *could*, and if we did, it would be of the form $\mathbf{r} = \mathbf{c}e^{i\lambda t}$, where $\mathbf{r}$ is just some arbitrary vector ansatz dependent on the system you're dealing with. However, we did not do this since we expect the constants, c_1 and c_2 to have different units.

The ansatz, in fact, *are* different, since they correspond to different quantities within the system.

Regardless, we find that

$$\begin{aligned} \begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} &= \begin{bmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} \\ i\lambda\mathbf{I} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ \mathbf{0} &= \begin{bmatrix} -i\lambda & \frac{1}{m} \\ -k & -i\lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \end{aligned}$$

As before, we demand that the coefficient matrix in λ is zero, so

$$0 = \begin{vmatrix} -i\lambda & \frac{1}{m} \\ -k & -i\lambda \end{vmatrix} = (-i\lambda)^2 + \frac{k}{m} = -\lambda^2 + \frac{k}{m} \implies \boxed{\lambda = \pm\sqrt{\frac{k}{m}}}$$

which are our eigenvalues for the system. As always, we want to find the eigenvectors, so we row-reduce the coefficient matrix after substituting in λ such that

$$\begin{bmatrix} \mp i\sqrt{\frac{k}{m}} & \frac{1}{m} \\ -k & \mp i\sqrt{\frac{k}{m}} \end{bmatrix} \xrightarrow{R_1/R_{11}} \begin{bmatrix} 1 & \pm\frac{i}{\sqrt{km}} \\ -k & \mp i\sqrt{\frac{k}{m}} \end{bmatrix} \xrightarrow{-R_2/k} \begin{bmatrix} 1 & \mp\frac{1}{i\sqrt{km}} \\ 1 & \mp\frac{1}{i\sqrt{km}} \end{bmatrix} \xrightarrow{R_1-R_2} \begin{bmatrix} 0 & 0 \\ 1 & \mp\frac{1}{i\sqrt{km}} \end{bmatrix}.$$

Notice that we have intentially done $R_1 - R_2$ here, since we know that the constants have different units. In particular, we expect that quantity $\sqrt{km}$ to have units of momentum, and 1 to have units of distance. Setting the free parameter, c_1, to t, we thus find the eigenvectors

$$\boxed{\mathbf{v}_{1,2} = \begin{bmatrix} 1 \\ \pm\frac{1}{i\sqrt{km}} \end{bmatrix}.}$$

Using this, we can construct our general solution as a vector equation, which is

$$\begin{bmatrix} x(t) \\ p(t) \end{bmatrix} = a\begin{bmatrix} 1 \\ \frac{1}{i\sqrt{km}} \end{bmatrix} e^{i\sqrt{\frac{k}{m}}t} + b\begin{bmatrix} 1 \\ -\frac{1}{i\sqrt{km}} \end{bmatrix} e^{-i\sqrt{\frac{k}{m}}t} = a\begin{bmatrix} 1 \\ \frac{1}{i\sqrt{km}} \end{bmatrix} e^{i\omega t} + b\begin{bmatrix} 1 \\ -\frac{1}{i\sqrt{km}} \end{bmatrix} e^{-i\omega t}$$

where we have substituted $\sqrt{k/m} = \omega$. Let us now suppose we are given initial initial conditions on the system, such that $x(0) = x_0$ and $p(0) = 0$. Here, we find that

$$\begin{aligned} x(0) &= a + b = x_0 \\ p(0) &= i\sqrt{km}(a - b) = 0. \end{aligned}$$

We thus arrive at $a = b = x_0/2$, and

$$\begin{aligned} x(t) &= \frac{x_0}{2}\left(e^{i\omega t} + e^{-i\omega t}\right) = x_0\cos(\omega t) \\ p(t) &= \frac{x_0}{2}\left(i\sqrt{km}\right)e^{i\omega t} - \frac{x_0}{2}\left(i\sqrt{km}\right)e^{-i\omega t} = \frac{ix_0\sqrt{km}}{2}\left(e^{i\omega t} - e^{-i\omega t}\right) = -x_0\sqrt{km}\sin(\omega t). \end{aligned}$$

To confirm this, we will do a quick unit check. Here,

$$\begin{aligned} [x_0] &= [m] \\ \left[x_0\sqrt{km}\right] &= \frac{[m][N^{1/2}][kg^{1/2}]}{[m^{-1/2}]} = \frac{[m][kg^{1/2}]\cancel{[m^{1/2}]}[kg^{1/2}]}{\cancel{[m^{-1/2}]}[s^{-2}]} = [m]\frac{[kg]}{[s]} \end{aligned}$$

which are units of length and momentum respectively, as we expected.

6.2 Sliding Pendulum on a Bead

Consider a pendulum of length, l, with a bob of mass, m_2, hanging from a bead of mass, m_1, free to slide along a horizontal surface without friction. Determine the equations of motion of the system using Hamilton's equations.

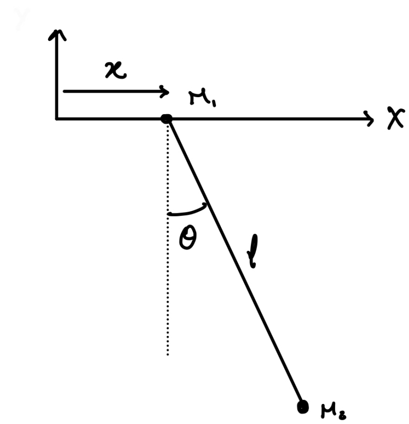


Figure 48: Pendulum of length, l, oscillating in conjunction with a sliding bead.

Here, our generalized coordinates for the system are

$$\begin{array}{c} X_1 = x \\ Y_1 = 0 \end{array} \implies \begin{array}{c} \dot{X}_1 = \dot{x} \\ \dot{Y}_1 = 0 \end{array}$$

and

$$\begin{array}{c} X_2 = x + l\sin(\theta) \\ Y_2 = -l\cos(\theta) \end{array} \implies \begin{array}{c} \dot{X}_2 = \dot{x} + l\cos(\theta)\dot{\theta} \\ \dot{Y}_2 = l\sin(\theta)\dot{\theta}. \end{array}$$

Our energies are

$$T = \frac{1}{2}(m_1 + m_2)\dot{x}^2 + \frac{1}{2}m_2 l^2\dot{\theta}^2 + m_2 l\cos(\theta)\dot{x}\dot{\theta}$$
$$V = -m_2 gl\cos(\theta)$$

to give a Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\dot{x}^2 + \frac{1}{2}m_2 l^2\dot{\theta}^2 + m_2 l\cos(\theta)\dot{x}\dot{\theta} + m_2 gl\cos(\theta).$$

Here, the respective momenta are

$$p_x = \frac{\partial\mathcal{L}}{\partial\dot{x}} = (m_1 + m_2)\dot{x} + m_2 l\cos(\theta)\dot{\theta}$$
$$p_\theta = \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = m_2 l^2\dot{\theta} + m_2 l\cos(\theta)\dot{x}.$$

We will note that the total linear momentum along the x-direction is conserved as x is a cyclic coordinate. Then, our Hamiltonian is of the form

$$\begin{aligned}
\mathcal{H} &= \sum_i p_i\dot{q}_i - \mathcal{L} \\
&= p_x\dot{x} + p_\theta\dot{\theta} - \mathcal{L} \\
&= \Big[(m_1 + m_2)\dot{x} + m_2 l\cos(\theta)\dot{\theta}\Big]\dot{x} + \Big[m_2 l^2\dot{\theta} + m_2 l\cos(\theta)\dot{x}\Big]\dot{\theta} \\
&\qquad - \left[\frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\dot{x}^2 + \frac{1}{2}m_2 l^2\dot{\theta}^2 + m_2 l\cos(\theta)\dot{x}\dot{\theta}\right] \\
&= \frac{1}{2}(m_1 + m_2)\dot{x}^2 + \frac{1}{2}m_2 l^2\dot{\theta}^2 + m_2 l\cos(\theta)\dot{x}\dot{\theta} - mgl\cos(\theta).
\end{aligned}$$

Again, we expect that the Hamiltonian is the total energy in the system. Notice, however, that because our potential is negative, we have that $\mathcal{H} = T + V = T + (-V) = E$. This, in fact, suggests that the Hamiltonian has *stable* or *bounded* dynamics. In particular, we have that the particles within the system exhibit bound motion. Regardless, we must transform the Hamiltonian to phase space, to which we substitute

$$\begin{aligned}\dot{x} &= \frac{lp_x - l_\theta \cos(\theta)}{l[m_1 + m_2 - m_2 \cos^2(\theta)]} \\ \dot{\theta} &= \frac{(m_1 + m_2)p_\theta - lm_2 p_x \cos(\theta)}{l^2 m_2 [m_1 + m_2 - m_2 \cos^2(\theta)]}\end{aligned}$$

to eliminate $\dot{x}$ and $\dot{\theta}$ and find

$$\boxed{\mathcal{H} = \frac{p_x^2 + (m_1 + m_2)p_\theta^2/(l^2 m_2) - 2\cos(\theta)p_x p_\theta/l}{2m_1 + m_2[1 - \cos(2\theta)]} - mgl\cos(\theta).}$$

Here, p_x and p_θ have different units so that the Hamiltonian is in units of energy.

6.3 The Accelerating Pulley Revisited

Recall Fig. 13, the pulley accelerating with an acceleration, $\alpha < g$. Here, our Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}m_1\big(\dot{x}_1^2 + 2\alpha t\dot{x}_1\big) + \frac{1}{2}m_2\big(\dot{x}_1^2 - 2\alpha t\dot{x}_1\big) + m_1 g x_1 - m_2 g x_1.$$

To construct our Hamiltonian, we find that the momentum is

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}_1} = (m_1 + m_2)\dot{x} + (m_1 - m_2)\alpha t.$$

With our understanding of non-inertial frames, we will show why the momentum appears in this form. We know that

$$\begin{aligned}\dot{x}_1'' &= \dot{x}_1 + \alpha t \\ \dot{x}_2'' &= \dot{x}_2 + \alpha t = -\dot{x}_1 + \alpha t = -(\dot{x}_1 - \alpha t).\end{aligned}$$

The linear momentum in each component is correspondingly

$$\begin{aligned}p_{x_1} &= m_1\dot{x}_1'' = m_1(\dot{x}_1 + \alpha t) \\ p_{x_2} &= m_2\dot{x}_2'' = -m_2(\dot{x}_1 - \alpha t).\end{aligned}$$

Now, since we know that the frame with acceleration α will generally move slower than the acceleration due to gravity itself, the total momentum is the difference in the Cartesian X direction (which, if you remember, is defined downward),

$$m_1(\dot{x}_1 + \alpha t) - [-m_2(\dot{x}_1 - \alpha t)] = m_1(\dot{x}_1 + \alpha t) + m_2(\dot{x}_1 - \alpha t) = (m_1 + m_2)\dot{x}_1 + (m_1 - m_2)\alpha t$$

as we found before. We now return to constructing the Hamiltonian, which is

$$\begin{aligned}\mathcal{H} &= p\dot{x}_1 - \mathcal{L} \\ &= (m_1 + m_2)\dot{x}_1^2 + \cancel{(m_1 - m_2)\alpha t\dot{x}_1} - \left[\frac{1}{2}(m_1 + m_2)\dot{x}_1^2 + \cancel{(m_1 - m_2)\alpha t\dot{x}_1} + (m_1 - m_2)g x_1\right] \\ &= \frac{1}{2}(m_1 + m_2)\dot{x}_1^2 - (m_1 - m_2)g x_1.\end{aligned}$$

Notice that because of the time dependence in our Lagrangian, the Hamiltonian is not simply $\mathcal{H} = T + V = E$. Removing the $\dot{x}$ dependence in the Hamiltonian, we find

$$\dot{x}_1 = \frac{p - (m_1 - m_2)\alpha t}{m_1 + m_2}$$

such that

$$\begin{aligned}\mathcal{H} &= \frac{1}{2}(m_1 + m_2)\left[\frac{p - (m_1 - m_2)\alpha t}{m_1 + m_2}\right]^2 - (m_1 - m_2)gx_1 \\ &= \frac{1}{2}\left[\frac{p^2 - 2(m_1 - m_2)\alpha tp + (m_1 - m_2)^2\alpha^2 t^2}{m_1 + m_2}\right] - (m_1 - m_2)gx_1\end{aligned}$$

and simplifying further, we find

$$\boxed{\mathcal{H} = \frac{p^2}{2(m_1 + m_2)} - \left[\frac{m_1 - m_2}{m_1 + m_2}\right]\alpha tp + \frac{[(m_1 - m_2)\alpha t]^2}{2(m_1 + m_2)} - (m_1 - m_2)gx_1.}$$

Take a careful note at this point. You may notice that the third term is only a function of time. That is, $f = f(t)$. We *could* have cancelled this term in the Lagrangian stage, but since the Lagrangian itself impacts both the Hamiltonian and the conjugate momenta, we would have run into problems. We *need* to have the same Lagrangian, so be careful not to arbitrarily throw out terms unless there is good reason to.

6.4 Central Potentials

We will point out that this example is relatively important. Many systems in real life are subject to central potentials, and we will discuss the consequences of this here.

Consider a Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z).$$

Applying a transformation on the Cartesian coordinates such that $r = \sqrt{x^2 + y^2 + z^2}$, we can write our Hamiltonian in the form

$$\mathcal{H} = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + V(r).$$

Now, defining the angular momentum vector, we have

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \times \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} yp_z - p_y z \\ zp_x - p_z x \\ xp_y - p_x y \end{bmatrix} = \begin{bmatrix} L_x \\ L_y \\ L_z \end{bmatrix}.$$

Here, we are mixing Cartesian and polar coordinates. It's worth being careful, although this is not a problem. Suppose now we are interested in how the components of the angular momentum interact with the Poisson bracket of $\mathcal{H}$. Considering, the z-component first, we find

$$\{L_z, \mathcal{H}\} = \{xp_y, \mathcal{H}\} - \{p_x y, \mathcal{H}\} = x\{p_y, \mathcal{H}\} + \{x, \mathcal{H}\}p_y - y\{p_x, \mathcal{H}\} - \{y, \mathcal{H}\}p_x$$

where we have used our Poisson bracket identities. We will first consider a Poisson bracket of the form $\{p_i, \mathcal{H}\}$ where we have switched to the index i, j, z to keep things general. We find,

$$\{p_i, \mathcal{H}\} = \left[\frac{\cancel{\partial p_i}}{\cancel{\partial} x_i}\frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial p_i}{\partial p_i}\frac{\partial \mathcal{H}}{\partial x_i}\right] + \left[\frac{\cancel{\partial p_i}}{\cancel{\partial} x_j}\frac{\partial \mathcal{H}}{\partial p_j} - \frac{\cancel{\partial p_i}}{\cancel{\partial} p_j}\frac{\partial \mathcal{H}}{\partial x_j}\right] + \left[\frac{\cancel{\partial p_i}}{\cancel{\partial} x_k}\frac{\partial \mathcal{H}}{\partial p_k} - \frac{\cancel{\partial p_i}}{\cancel{\partial} p_k}\frac{\partial \mathcal{H}}{\partial x_k}\right]$$

where we have cancelled some terms since $\partial p_i/\partial x_{j,k} = 0$. Furthermore, $\partial p_i/\partial p_i = 1$, so we find

$$\{p_i, \mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial x_i} = -\frac{\partial V(r)}{\partial x_i} = -\frac{\partial V(r)}{\partial r}\frac{\partial r}{\partial x_i}.$$

In general, we have that

$$\frac{\partial r}{\partial x_i} = \frac{\partial}{\partial x_i}\left[\sqrt{x_i^2 + x_j^2 + x_k^2}\right] = \frac{x_i}{\sqrt{x_i^2 + x_j^2 + x_k^2}} = \frac{x_i}{r}$$

which follows for any cyclic permutation of $i \to j \to k$. Thus,

$$\{p_i, \mathcal{H}\} = -\frac{\partial V(r)}{\partial r}\frac{x_i}{r}$$

under cyclic permutations as well. We now consider the other general case, which is

$$\begin{aligned}
\{x_i, \mathcal{H}\} &= \left[\frac{\partial x_i}{\partial x_i}\frac{\partial \mathcal{H}}{\partial p_i} - \cancel{\frac{\partial x_i}{\partial p_i}}\frac{\partial \mathcal{H}}{\partial x_i}\right] + \left[\cancel{\frac{\partial x_i}{\partial x_j}}\frac{\partial \mathcal{H}}{\partial p_j} - \cancel{\frac{\partial x_i}{\partial p_j}}\frac{\partial \mathcal{H}}{\partial x_j}\right] + \left[\cancel{\frac{\partial x_i}{\partial x_k}}\frac{\partial \mathcal{H}}{\partial p_k} - \cancel{\frac{\partial x_i}{\partial p_k}}\frac{\partial \mathcal{H}}{\partial x_k}\right] \\
&= \frac{\partial \mathcal{H}}{\partial p_i} \\
&= \frac{p_i}{m}
\end{aligned}$$

where we have cancelled by similar arguments. This relation is also valid under a cyclic permutation. Thus, we find that our total Poisson bracket in L_z and $\mathcal{H}$ is

$$\begin{aligned}
\{L_z, \mathcal{H}\} &= x\left[-\frac{\partial V(r)}{\partial r}\frac{y}{r}\right] + \left[\frac{p_x}{m}\right]p_y - y\left[-\frac{\partial V(r)}{\partial r}\frac{x}{r}\right] - \left[\frac{p_y}{m}\right]p_x \\
&= \frac{\partial V(r)}{\partial r}\left[-\frac{xy}{r} + \frac{xy}{r}\right] - \frac{1}{m}[p_x p_y - p_y p_x] \\
&= 0.
\end{aligned}$$

Since our previous relations hold under a cyclic permutation, it only remains to construct the Poisson bracket of L_x and L_y with $\mathcal{H}$, which are of similar form. In fact,

$$\begin{aligned}
\{L_x, \mathcal{H}\} &= y\{p_z, \mathcal{H}\} + \{y, \mathcal{H}\}p_z - z\{p_y, \mathcal{H}\} - \{z, \mathcal{H}\}p_y = 0 \\
\{L_y, \mathcal{H}\} &= z\{p_x, \mathcal{H}\} + \{z, \mathcal{H}\}p_x - x\{p_z, \mathcal{H}\} - \{x, \mathcal{H}\}p_z = 0.
\end{aligned}$$

If it is the case that all three Poisson brackets of L_x, L_y, and L_z are zero with $\mathcal{H}$, it then immediately follows that

$$\boxed{\{\mathbf{L}, \mathcal{H}\} = 0.}$$

In other words, angular momentum is **always** conserved with systems under central potentials. This follows from the fact that we have furthermore implied

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = \{\mathbf{L}, \mathcal{H}\} + \cancel{\frac{\partial \mathbf{L}}{\partial t}} = \{\mathbf{L}, \mathcal{H}\} = 0.$$

6.5 Bead on a Parabolic Wire

Consider a bead of mass, m, constrained to a vertically oriented, parabolic wire governed by $y = ax^2/2$ under the influence of gravity. Determine the equations of motion of the system by Hamilton's equations and discuss the result. Note that the constant, a, is one with units of inverse length for the question to be consistent.

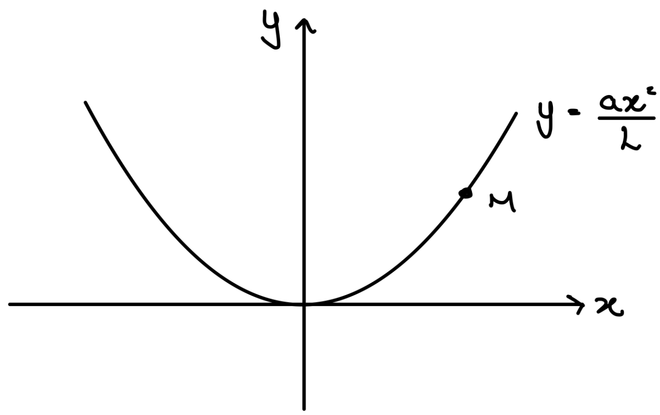


Figure 49: Bead of mass, m, constrained to a parabolic wire $y = ax^2/2$.

Here, the wire imposes a constraint between the coordinates x and y, such that

$$y = \frac{ax^2}{2} \implies \dot{y} = ax\dot{x}.$$

Assuming that the energies are purely kinetic and gravitational, the Lagrangian is of the form

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy = \frac{1}{2}m(\dot{x}^2 + a^2x^2\dot{x}^2) - mga\frac{x^2}{2}.$$

The generalized momentum in the system is simply

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} + ma^2x^2\dot{x} \implies \dot{x} = \frac{p}{m(1+a^2x^2)}$$

and so we can construct the Hamiltonian

$$\begin{aligned}\mathcal{H} &= p\dot{x} - \mathcal{L} \\ &= p\left[\frac{p}{m(1+a^2x^2)}\right] - \frac{1}{2}m(1+a^2x^2)\left[\frac{p^2}{m^2(1+a^2x^2)^2}\right] + mga\frac{x^2}{2} \\ &= \frac{1}{2m}\frac{p^2}{(1+a^2x^2)} + \frac{mgax^2}{2}.\end{aligned}$$

Using Hamilton's equations, we find that

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial p} \implies \boxed{\dot{x} = \frac{p}{m(1+a^2x^2)}}$$

which is the redundant relation, and

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial x} = \frac{p^2}{2m}\left[-\frac{2a^2x}{(1+a^2x^2)^2}\right] + mgax \implies \boxed{\dot{p} = \frac{p^2a^2x}{2m(1+a^2x^2)^2} - mgax.}$$

Plotting this numerically for the values $m = 1$, $a = 1$, and $g = 9.8$, as well as with the initial conditions $x(0) = 0$ and $p(0) \in [-20, 20]$, we can see that the movement of the particle is slightly unusual.

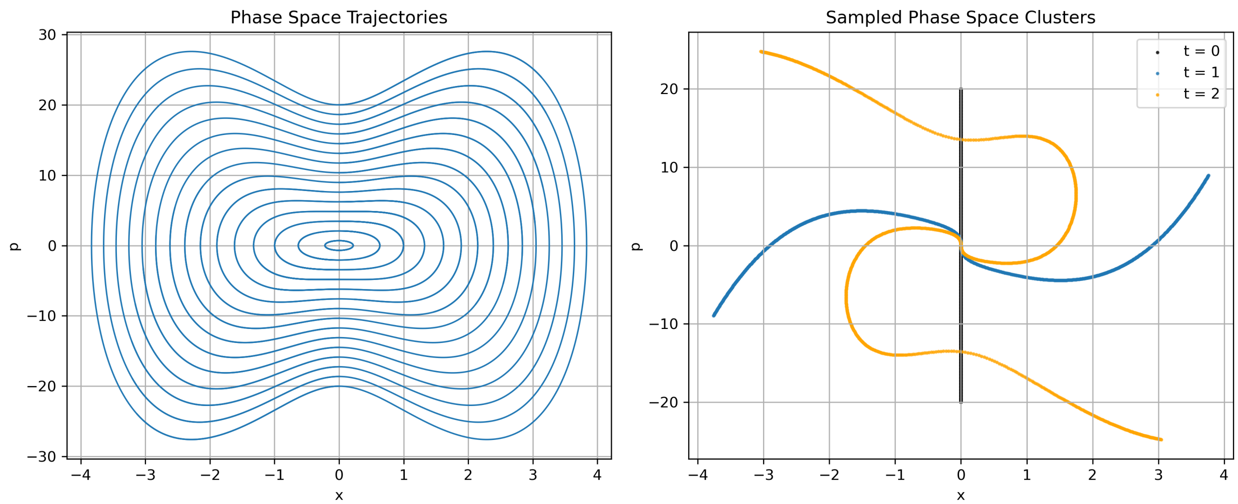


Figure 50: Numerical results for a bead on a parabolic wire. The left side are a range of trajectories and the right side are a cluster of trajectories sampled at $t = 0$ (black), $t = 1$ (blue), and $t = 2$ (red).

However, if we recall that the canonical momentum here is *not* simply the linear momentum, the deviations as the trajectories grow larger become dominated by the $1/(1+a^2x^2)$ factor. In fact, if we take the ratio of the linear momentum, Π_x (conventionally used), to the canonical momentum, p, we have

$$\frac{\Pi_x}{p} = \frac{m\dot{x}}{p} = \frac{\cancel{m}}{\cancel{p}}\frac{\cancel{p}}{\cancel{m}(1+a^2x^2)} = \frac{1}{1+a^2x^2}$$

which shows that for large, p, $1 + x^2$ must also grow large and mitigate the linear momentum correspondingly.

Furthermore, for a large cluster of trajectories, we can see that Liouville's theorem suggests that the area enclosed by each cluster is constant. That is, we if were to measure the area of the blue, orange, and black cluster, their areas would all be the same.

6.6 The Van der Pol Oscillator Revisited

Recall the VDPO, with a Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos(\theta)$$

with the generalized force

$$Q = \varepsilon\left[1 - \left(\frac{\theta}{\theta_0}\right)^2\right]\dot{\theta}.$$

We can construct the Hamiltonian now, which is

$$\mathcal{H} = \dot{\theta}_\theta - \mathcal{L} = \frac{p_\theta^2}{2ml^2} - mgl\cos(\theta), \quad p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = ml^2\dot{\theta}.$$

Hamilton's equations show that

$$\dot{\theta} = \frac{\partial \mathcal{H}}{\partial p_\theta} \implies \boxed{\dot{\theta} = \frac{p_\theta}{ml^2}}$$

$$\dot{p}_\theta = -\frac{\partial \mathcal{H}}{\partial \theta} + Q \implies \boxed{\dot{p}_\theta = -mgl\sin(\theta) + \varepsilon\left[1 - \left(\frac{\theta}{\theta_0}\right)^2\right]\frac{p_\theta}{ml^2}.}$$

We can revisit Fig. 35 now to understand what is happening. If we take the ratio of the second equation to the first. That is, the derivative of the angular momentum to the angular velocity, we can see how the trajectories tend to act for large θ. Here,

$$\frac{\dot{p}_\theta}{\dot{\theta}} = \frac{ml^2}{p_\theta}\left\{-mgl\sin(\theta) + \varepsilon\left[1 - \left(\frac{\theta}{\theta_0}\right)^2\right]\frac{p_\theta}{ml^2}\right\} = \varepsilon\left[1 - \left(\frac{\theta}{\theta_0}\right)^2\right] - \frac{m^2gl^3}{p_\theta}\sin(\theta).$$

Then, for $\theta \gg 1$, we find that

$$\left(\frac{\theta}{\theta_0}\right)^2 \gg 1, \quad \frac{\sin(\theta)}{p_\theta} \to 0$$

and so

$$\frac{\dot{p}_\theta}{\dot{\theta}} \approx -\varepsilon\left(\frac{\theta}{\theta_0}\right)^2.$$

We can employ the trick that the derivatives are related to their anti-derivatives, such that

$$\frac{\dot{p}_\theta}{\dot{\theta}} = \frac{\mathrm{d}p_\theta/\mathrm{d}t}{\mathrm{d}\theta/\mathrm{d}t} = \frac{\mathrm{d}p_\theta}{\mathrm{d}\theta} \approx -\varepsilon\frac{\theta^2}{\theta_0^2}$$

and so

$$p_\theta(\theta) = \int \frac{\mathrm{d}p_\theta}{\mathrm{d}\theta}\,\mathrm{d}\theta \approx -\frac{\varepsilon}{\theta_0^2}\int \theta^2\,\mathrm{d}\theta = C - \frac{\varepsilon}{3\theta_0^2}\theta^3.$$

Notice, with p_θ as a function of θ, we can see that as θ increases, that is $\theta \gg 1$, we have the proportionality

$$p_\theta(\theta) \propto -\theta^3$$

which suggests that momentum decays cubically with the angular position. It is furthermore *not* periodic, with a nonlinear damping that grows larger with respect to θ. We can, if you are interested, do further analysis on determining the critical angle of the limit cycle for the VDPO, but this will suffice. However, to do this yourself, all you would need to do is take that average time derivative of the energy over a single period; in principle, that will give you the expected angle at the limit cycle.

6.7 Particle on a Cone

Consider a particle of mass, m, constrained to move the surface of a cone governed by $z = cr$. That is, the cone grows linearly with the radius. Determine the equations of motion of the system and discuss the result.

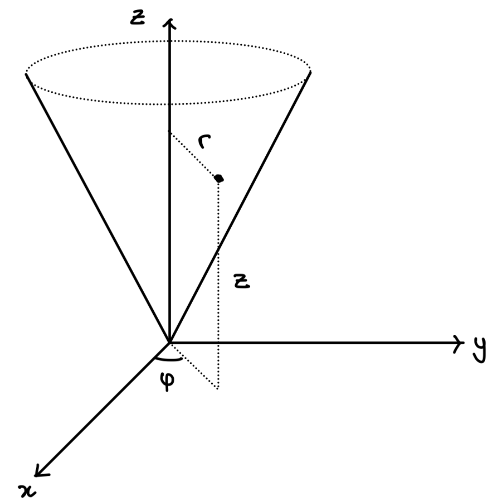


Figure 51: Particle of mass, m, constrained to move on a cone governed by $z = cr$.

Similar to the particle on a parabolic wire case, the cone constrains the particle by the relationship

$$\frac{z}{r} = c$$

where c is some constant, positive slope. We can write out the Lagrangian, which is

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + z]^2) - mgz$$

and, incorporating the constraint such that $z = cr \implies \dot{z} = c\dot{r}$, we have

$$\mathcal{L} = \frac{1}{2}m\Big[(1+c^2)\dot{r}^2 + r^2\dot{\theta}^2\Big] - mgcr$$

where we have also switched to cylindrical coordinates. Here, there are two generalized momenta, such that

$$p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m(1+c^2)\dot{r} \implies \dot{r} = \frac{p_r}{m(1+c^2)}$$
$$p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta} \implies \dot{\theta} = \frac{p_\theta}{mr^2}.$$

Our Hamiltonian is of the form

$$\mathcal{H} = p_r\dot{r} + p_\theta\dot{\theta} - \mathcal{L} = \frac{p_r^2}{2m(1+c^2)} + \frac{p_\theta^2}{2mr^2} + mgcr.$$

Hamilton's equations in r are then

$$\dot{r} = \frac{\partial \mathcal{H}}{\partial p_r} \implies \boxed{\dot{r} = \frac{p_r}{m(1+c^2)}}$$
$$\dot{p}_r = -\frac{\partial \mathcal{H}}{\partial r} \implies \boxed{\dot{p}_r = \frac{p_\theta^2}{mr^3} - mgc}$$

and in θ,

$$\dot{\theta} = \frac{\partial \mathcal{H}}{\partial p_\theta} \implies \boxed{\dot{\theta} = \frac{p_\theta}{mr^2}}$$
$$\dot{p}_\theta = -\frac{\partial \mathcal{H}}{\partial \theta} \implies \boxed{\dot{p}_\theta = 0.}$$

We could have found that p_θ is conserved quantity just by realizing it is a cyclic coordinate, or you could also take the Poisson bracket of p_θ with $\mathcal{H}$. Let us now consider a specific case of circular motion. Here, the condition is

$$\dot{r} = 0 \implies \dot{p}_r = 0$$

and so Hamilton's second EOM in r, shows that

$$\frac{p_\theta^2}{mr_c^3} = mgc \implies p_\theta = m\sqrt{gr_c^3}.$$

To determine the stability of this specific type of motion, we take the derivative of Hamilton's first EOM in r, to find that

$$\ddot{r} = \frac{\dot{p}_r}{m(1+c^2)} = \frac{1}{m(1+c^2)}\left[\frac{p_\theta^2}{mr^3 - mgc}\right].$$

Now, we will assume a small perturbation in our circular orbits, such that $r = r_c + \delta r$, to find

$$\ddot{r} = \cancel{\ddot{r}_c} + \delta\ddot{r} = \delta\ddot{r}$$

and so

$$\delta\ddot{r} = \frac{1}{m(1+c^2)}\left[\frac{p_\theta^2}{m(r_c+\delta r)^3} - mgc\right].$$

By the binomial expansion, the fraction in r_c becomes

$$\frac{p_\theta^2}{m(r_c+\delta r)^3} = \frac{p_\theta^2}{mr_c^3(1+3\delta r/r_c)} \approx \frac{p_\theta^2}{mr_c^3}\left[1 - \frac{3\delta r}{r_c}\right]$$

to find that

$$\delta\ddot{r} = \frac{1}{m(1+c^2)}\left[\frac{p_\theta^2}{mr_c^3}\left(1 - \frac{3\delta r}{r_c}\right) - mgc\right] = -\frac{3}{m(1+c^2)}\frac{p_\theta^2}{mr_c^4}\delta r$$

which is the form of SHM. Thus, we expect our orbits to be stable. This type of stability, in fact, is a general consequence of a type of central potential that is attractive. It turns out that *any* attractive central potential has circular, closed orbits. However, there are only *two* potentials where a perturbation to its circular orbit is bound and closed. These potentials turn out to be of the form

$$V = -\frac{k}{r}, \quad \text{and} \quad V = kr^2$$

which is a consequence of **Bertrand's Theorem**.

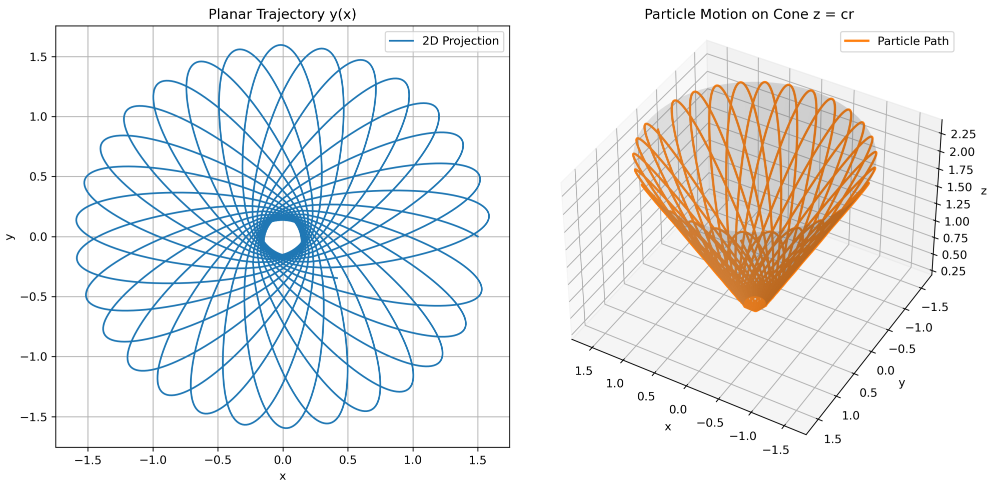


Figure 52: Two plots of the particle constrained to move along a cone

6.8 Particle Constrained to a Rotating Wire

Consider a similar setup to that of the previous example. However, rather than a cone, the particle is constrained to move on a rotating wire governed by the linear relationship $z = cr$. Furthermore, the wire rotates with a constant angular frequency, $\vec{\Omega} = \Omega\hat{z}$.

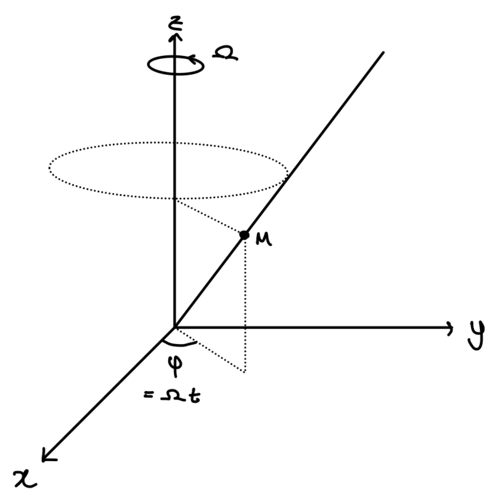


Figure 53: Particle constrained to move on a rotating wire of frequency $\vec{\Omega} = \Omega\hat{z}$.

Here, we will use cylindrical coordinates for the generalized coordinates, such that

$$\begin{aligned} x &= r\cos(\Omega t) \\ y &= r\sin(\Omega t) \\ z &= cr. \end{aligned}$$

The Lagrangian of our system is thus

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\Omega^2 + c^2\dot{r}^2) - mgcr.$$

The generalized momentum is only in the radial direction, and is of the form

$$p = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m(1+c^2)\dot{r}$$

and so the Hamiltonian is of the form

$$\mathcal{H} = p\dot{r} - \mathcal{L} = \frac{p^2}{2m(1+c^2)} - \frac{1}{2}mr^2\Omega^2 + mgcr.$$

Hamilton's equations of motion further state that

$$\dot{r} = \frac{\partial \mathcal{H}}{\partial p} \implies \boxed{\dot{r} = \frac{p}{m(1+c^2)}}$$

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial r} \implies \boxed{\dot{p} = mr\Omega^2 - mgc.}$$

We can use these EOMs to determine the second derivative in r, which is

$$\ddot{r} = \frac{\dot{p}}{m(1+c^2)} = \frac{r\Omega^2 - gc}{(1+c^2)}.$$

As before, we are interested in the case of circular orbits, so $\ddot{r} = 0$ and

$$r_c\Omega^2 = gc \implies r_c = \frac{gc}{\Omega^2}.$$

For small perturbations, we set $r = r_c + \delta r$ such that $\ddot{r} = \delta\ddot{r}$, and so

$$\delta\ddot{r} = \frac{(r_c + \delta r)\Omega^2 - gc}{1+c^2} = \frac{\Omega^2}{1+c^2}\delta r$$

which is unstable motion. However, these two examples, qualitatively, are similar. The rotating wire creates a cone structure as it rotates, which is analogous to the actual cone case, so what is the difference?

In the cone problem, you may recall that $\dot{p}_\theta = 0$. Thus, the angular momentum is conserved and furthermore energy is conserved. In this case, however, we have that $\dot{\phi} = \Omega$ is constant. If we integrate this expression, we find that the angle $\phi = \Omega t + C$ for some constant, C. We can drop the constant and furthermore find that

$$\dot{p} = mr\left[\frac{\phi}{t}\right]^2 - mgc$$

which says that the momentum is a time-dependent quantity. Energy, in this case, is not conserved, which is a more general consequence of the normal force being imposed from the wire. In fact, if we substitute the relation $\Omega = \phi/t$ into the Lagrangian, we find that it *is* time-dependent under two coordinates ϕ and r, to which we can use LUM to determine the normal force in the system.

7 Rigid Body Mechanics

Let us now extend our discussion of mechanics to rigid bodies. A rigid body, in general, is a collection of point masses with some spatial extend. Furthermore, we say that a rigid body is a *set* of n point masses where the *relative* positions of all the point masses, $\mathbf{r}_{ij}$, are fixed. That is,

$$\mathbf{r}_{ij} = \text{constant}.$$

As a result of this condition, our rigid body shape is fixed for all dynamics of the system. In reality, this is not the case; rigid body mechanics is an *approximation* of certain types of bodies that can be analyzed.

You may recall that a point particle in 3D has 3 independent DoF that define its position, $\mathbf{r}(t)$. We extended this to a rigid body with 6 DoF, where 3 DoF describe the COM of the system, and another 3 defines the DoF of any individual mass, m_α. Here, the COM specifies the position of the body, while the individual masses specify the orientation.

You may also notice that is, in fact, a consequence of another result. In reality, we should have $3n$ DoF for the position of every mass, m_α. However, since we impose that $\mathbf{r}_{ij}$ is constant, we reduce this by a factor of n entirely. That is

$$\text{DoF} = \frac{1}{n}\sum_{i=1}^{n} d_i - \varnothing = \frac{3n}{n} + d_{\text{CM}} = 3 + 3 = 6.$$

Let us now formalize specific quantities about our system.

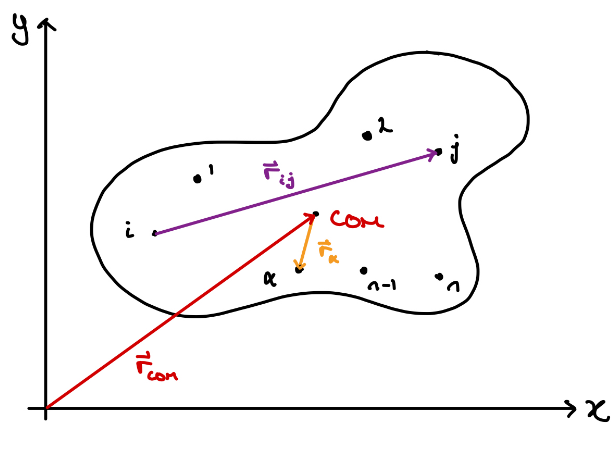


Figure 54: Rigid body with the vector $\mathbf{r}_{\text{COM}}$, $\mathbf{r}_\alpha$, and $\mathbf{r}_{ij}$.

We can define the centre of mass vector, $\mathbf{r}_{\text{CM}}$ as the *weighted sum* of the moment of inertia's due to each mass, with respect to the total mass. That is,

$$\boxed{\mathbf{r}_{\text{CM}} = \frac{1}{M}\sum_{\alpha=1}^{n} m_\alpha \mathbf{r}_\alpha.} \tag{36}$$

For a continuous mass distribution, this can be extended to its integral form such that

$$\boxed{\mathbf{r}_{\text{CM}} = \frac{1}{M}\left[\lim_{n\to\infty}\sum_{\alpha=1}^{n}\mathbf{r}_\alpha \Delta m_\alpha\right] = \frac{1}{M}\int \mathbf{r}\,\mathrm{d}m} \tag{37}$$

For most cases, we are interested in the origin being situated *at* the CM, such that we have the condition

$$\boxed{\sum_{\alpha=1}^{n} m_\alpha \mathbf{r}_\alpha = \sum_{\alpha=1}^{n}\dot{\mathbf{r}}_\alpha = 0.} \tag{38}$$

We typically define the orientation of the rigid body in terms of the CM, and the rotation of the body with the particles *about* the CM.

7.1 Dynamical Quantities of a Rigid Body

7.1.1 The Velocity of the Body

Suppose now we have a rigid body rotating (co-moving) with a non-inertial frame within an inertial frame. We can describe the position of some point, P, inside the rigid body in the inertial frame, $\hat{x}_i$, or the non-inertial frame, $\hat{x}'_i$.

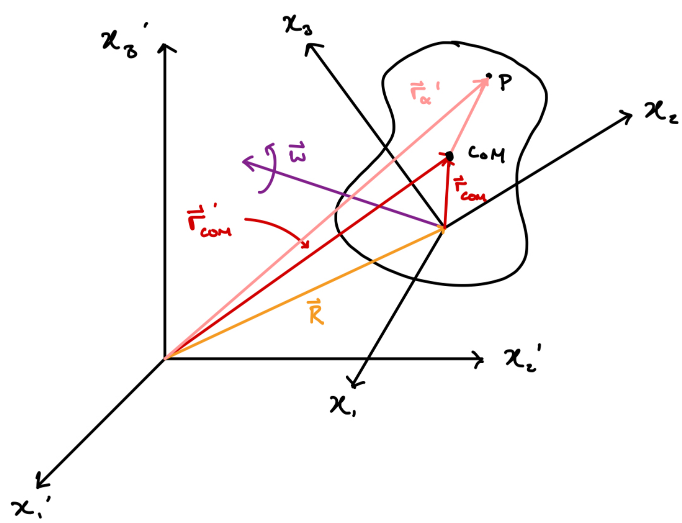


Figure 55: Rotating rigid body about the vector, $\vec{\omega}$.

In particular, we can write the position of the CM as

$$\mathbf{r}'_{\text{CM}} = \mathbf{R} + \mathbf{r}_{\text{CM}}$$

where $\mathbf{R}$ is the vector from $O' \to O$, and $\mathbf{r}_{\text{CM}}$ is the vector from $O \to$ CM, and O' is the origin in the inertial frame, and O is the origin of the non-inertial frame. With our previous definitions, the velocity of the CM is

$$\dot{\mathbf{r}}'_{\text{CM}} = \dot{\mathbf{R}} + \cancel{\frac{\delta \mathbf{r}_{\text{CM}}}{\delta t}} + \vec{\omega} \times \mathbf{r}_{\text{CM}} = \dot{\mathbf{R}} + \vec{\omega} \times \mathbf{r}_{\text{CM}}$$

where the derivative of $\mathbf{r}_{\text{CM}}$ is zero since we assumed that the body is co-moving with the non-inertial frame. We can thus define the position of any mass within the body, m_α, in the non-inertial frame with

$$\mathbf{r}'_\alpha = \mathbf{R} + \mathbf{r}_{\text{CM}} + \mathbf{r}_\alpha = \mathbf{r}'_{\text{CM}} + \mathbf{r}_\alpha, \quad \text{where } \mathbf{r}'_{\text{CM}} = \mathbf{R} + \mathbf{r}_{\text{CM}}.$$

Thus, the velocity of any particle within the inertial frame is given by

$$\dot{\mathbf{r}}'_\alpha = \dot{\mathbf{R}} + \vec{\omega} \times (\mathbf{r}_{\text{CM}} + \mathbf{r}_\alpha) = (\dot{\mathbf{R}} + \vec{\omega} \times \mathbf{r}_{\text{CM}}) + \vec{\omega} \times \mathbf{r}_\alpha = \dot{\mathbf{r}}'_{\text{CM}} + \vec{\omega} \times \mathbf{r}_\alpha$$

where we have distributed the cross product term. Furthermore, we note that $\dot{\mathbf{r}}_{\text{CM}}$ is the motion *of* the CM, while $\vec{\omega} \times \mathbf{r}_\alpha$ is the motion *about* the CM. For more clarity, we can rewrite this term in the form

$$\boxed{\mathbf{v}'_\alpha = \mathbf{V} + \vec{\omega} \times \mathbf{r}_\alpha.} \tag{39}$$

Notice that the form of this expression is what we would expect if the origin of the co-moving non-inertial frame is at the CM. In this way, we are inclined to (and will) use this representation of the rigid body as it is the most intuitive and simplifies expressions.

7.1.2 The Kinetic Energy of the Body

It naturally follows that we should determine the kinetic energy from an inertial reference frame. If we recall our expression for the rigid body, we find that the kinetic energy is

$$\begin{aligned} T &= \frac{1}{2} \sum_{\alpha=1}^{n} m_\alpha |\mathbf{v}'_\alpha|^2 \\ &= \frac{1}{2} \sum_{\alpha=1}^{n} m_\alpha |\mathbf{V} + \vec{\omega} \times \mathbf{r}_\alpha|^2 \\ &= \frac{1}{2} \sum_{\alpha=1}^{n} m_\alpha |\mathbf{V}|^2 + \sum_{\alpha=1}^{n} m_\alpha \mathbf{V} \cdot (\vec{\omega} \times \mathbf{r}_\alpha) + \frac{1}{2} \sum_{\alpha=1}^{n} m_\alpha |\vec{\omega} \times \mathbf{r}_\alpha|^2. \end{aligned}$$

Starting with the first term, we can see that $|\mathbf{V}|^2$ is not dependent on the mass, so we simply have $M|\mathbf{V}|^2/2$. On the other hand, we can use the cyclic permutations for our second term to find

$$\sum_{\alpha=1}^{n} m_\alpha \mathbf{V} \cdot (\vec{\omega} \times \mathbf{r}_\alpha) = \sum_{\alpha=1}^{n} m_\alpha \mathbf{r}_\alpha \cdot (\mathbf{V} \times \vec{\omega}) = 0$$

since we assumed the origin is now situated at the CM. The third term we can use a vector calculus identity to find that

$$|\vec{\omega} \times \mathbf{r}_\alpha|^2 = |\vec{\omega}|^2 |\mathbf{r}_\alpha|^2 - |\vec{\omega} \cdot \mathbf{r}_\alpha|^2$$

and so the kinetic energy term becomes

$$T = \frac{1}{2} M |\mathbf{V}|^2 + \frac{1}{2} \sum_{\alpha=1}^{n} m_\alpha \big[|\vec{\omega}|^2 |\mathbf{r}_\alpha|^2 - |\vec{\omega} \cdot \mathbf{r}_\alpha|^2 \big].$$

Notice that this kinetic energy is split into two components. In fact, we can see that

$$T = T_{\text{trans.}} + T_{\text{rot.}}$$

where $T_{\text{trans.}}$ is the kinetic energy as if *all* mass were concentrated at the CM moving with a velocity, $\mathbf{V}$, and $T_{\text{rot.}}$ is the kinetic energy of the particles from rotations *about* the CM.

7.1.3 The Inertia Tensor

Beginning with the expression for the kinetic energy we found earlier, we can expand $\vec{\omega}$ and $\mathbf{r}_\alpha$ in terms of their components,

$$\begin{aligned}\vec{\omega} &= \omega_1\hat{x}_1 + \omega_2\hat{x}_2 + \omega_3\hat{x}_3 \\ \mathbf{r}_\alpha &= x_{\alpha,1}\hat{x}_1 + x_{\alpha,2}\hat{x}_2 + x_{\alpha,3}\hat{x}_3.\end{aligned}$$

Furthermore, we also have that

$$|\omega|^2 = \sum_{i=1}^{3}\omega_i^2 \quad \text{and} \quad \vec{\omega}\cdot\mathbf{r}_\alpha = \sum_{i=1}^{3}\omega_i x_{\alpha,i}.$$

Using this, we can rewrite our rotational kinetic energy expression to find

$$T_{\text{rot.}} = \frac{1}{2}\sum_{\alpha=1}^{n} m_\alpha\left[\left(\sum_{i=1}^{3}\omega_i^2\right)\left(\sum_{k=1}^{3}x_{\alpha,k}^2\right) - \left(\sum_{i=1}^{3}\omega_i x_{\alpha,i}\right)\left(\sum_{j=1}^{3}\omega_j x_{\alpha,j}\right)\right].$$

Furthermore, we can expand our ω_i^2 term into two sums, such that

$$\sum_{i=1}^{3}\omega_i^2 = \sum_{i=1}^{3}\sum_{j=1}^{3}\omega_i\omega_j\delta_{ij} = \sum_{i,j=1}^{3}\omega_i\omega_j\delta_{ij}.$$

Substituting this back into the kinetic energy expression, we find that

$$T_{\text{rot.}} = \frac{1}{2}\sum_{\alpha=1}^{n} m_\alpha\left[\left(\sum_{i,j=1}^{3}\omega_i\omega_j\delta_{ij}\right)\left(\sum_{k=1}^{3}x_{\alpha,k}^2\right) - \left(\sum_{i=1}^{3}\omega_i x_{\alpha,i}\right)\left(\sum_{j=1}^{3}\omega_j x_{\alpha,j}\right)\right].$$

Pulling out the sums in i and j, we can thus simplify T to be

$$\boxed{T_{\text{rot.}} = \frac{1}{2}\sum_{i,j=1}^{3}\omega_i\omega_j\sum_{\alpha=1}^{n}m_\alpha\left[\delta_{ij}\sum_{k=1}^{3}x_{\alpha,k}^2 - x_{\alpha,i}x_{\alpha,j}\right].} \tag{40}$$

Here, notice that the term before the sum in m_α depends on the angular velocity, $\vec{\omega}$, while the term involving the sum depends on the mass distribution. In fact, we assign the latter term a new quantity, which we call the **Inertia Tensor**, I_{ij}, for the indices i and j.

$$\boxed{I_{ij} = \sum_{\alpha=1}^{n}\left[\delta_{ij}\sum_{k=1}^{3}x_{\alpha,k}^2 - x_{\alpha,i}x_{\alpha,j}\right].} \tag{41}$$

There are several conventional notations to represent tensors. Thornton and Marion, for one, use the notation $\{\mathbf{I}\}$ to represent the full inertia tensor, but there are others in use such as $\mathbf{I}$ and $\vec{\vec{I}}$. Personally, we will use $\{\mathbf{I}\}$ for the expanded tensor form, that is its 3×3 matrix form. We can explicitly compute each of the components for the inertia tensor, but we will see shortly that it is somewhat symmetric; thus, we opt to show only the case I_{11} and I_{13}.

$$I_{11} = \sum_{\alpha=1}^{n} m_\alpha \big[\delta_{11}\big(x_{\alpha,1}^2 + x_{\alpha,2}^2 + x_{\alpha,3}^2\big) - x_{\alpha,1}x_{\alpha,1}\big] = \sum_{\alpha=1}^{n} m_\alpha (x_{\alpha,2}^2 + x_{\alpha,3}^2)$$
$$I_{13} = \sum_{\alpha=1}^{n} m_\alpha \big[\delta_{13}(x_{\alpha,1}^2 + x_{\alpha,2}^2 + x_{\alpha,3}^2) - x_{\alpha,1}x_{\alpha,3}\big] = -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,1}x_{\alpha,3}.$$

Notice, in the I_{13} case, the Kronecker delta is zero, and so the I_{13} component of the inertia tensor is negative. If we explicitly compute each argument of the tensor, we eventually arrive at

$$\boxed{\{\mathbf{I}\} = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} = \begin{bmatrix} \sum_{\alpha=1}^{n} m_\alpha\big(x_{\alpha,2}^2 + x_{\alpha,3}^2\big) & -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,1}x_{\alpha,2} & -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,1}x_{\alpha,3} \\ -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,1}x_{\alpha,2} & \sum_{\alpha=1}^{n} m_\alpha\big(x_{\alpha,1}^2 + x_{\alpha,3}^2\big) & -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,2}x_{\alpha,3} \\ -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,1}x_{\alpha,3} & -\sum_{\alpha=1}^{n} m_\alpha x_{\alpha,2}x_{\alpha,3} & \sum_{\alpha=1}^{n} m_\alpha\big(x_{\alpha,1}^2 + x_{\alpha,2}^2\big) \end{bmatrix}.}$$

Here, the diagonal elements of the inertia tensor are called the **moments of inertia**, while the off-diagonal elements are called the **products of inertia**. Furthermore, we say that the inertia tensor is a **second order tensor**. It adheres to specific transformation rules, which transform to

$$\boxed{\{\mathbf{I}'\} = \mathbf{R}\{\mathbf{I}\}\mathbf{R}^T} \tag{42}$$

where R is some transformation matrix, typically a rotation matrix. In general, a tensor is unique from a matrix in the sense that its transformation is not simply a single rotation. Without getting into details, we can say that a tensor is transformed as a multilinear mapping (essentially, a generalization to a bilinear[14] mapping) $\mathbb{I} : V^* \times \cdots \times V \times \cdots \times V \to \mathbb{R}$. If this is slightly confusing, then don't stress over it; it is added for clarity to show that a tensor and a matrix are **not** the same.

Back to the topic at hand, we can thus rewrite the rotational kinetic energy as

$$T_{\text{rot.}} = \frac{1}{2}\sum_{i,j=1}^{3} \omega_i \omega_j I_{ij}$$

and for a more familiar notation as adopted by Thornton and Marion, is rewritten as

$$\boxed{T_{\text{rot.}} = \frac{1}{2}\vec{\omega}^T\{\mathbf{I}\}\vec{\omega} = \frac{1}{2}\begin{bmatrix}\omega_1 \\ \omega_2 \\ \omega_3\end{bmatrix}^T \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \begin{bmatrix}\omega_1 \\ \omega_2 \\ \omega_3\end{bmatrix}.} \tag{43}$$

[14]A bilinear map is essentially a function which takes elements of two vector spaces and spits out a third one, which is linear in each of its arguments. That is,

$$f(cu, v) = cf(u, v), \quad f(cu, v) = f(u, cv)$$

for some scalar quantity, c.

Here, $\{\mathbf{I}\}\vec{\omega}$ is a column vector, such that $\vec{\omega}^T\{\mathbf{I}\}\vec{\omega}$ returns a scalar quantity. That is, $\vec{\omega}^T\{\mathbf{I}\}\vec{\omega}$ is a dot product of $\vec{\omega}^T$ on $\{\mathbf{I}\}\vec{\omega}$.

We can easily define the inertia tensor for continuous mass distributions as well if we consider the limit of a differential mass element. Here, the argument I_{ij} simply becomes

$$\boxed{I_{ij} = \int_V \rho(\mathbf{r})\left[\delta_{ij}\sum_{k=1}^{3} x_{\alpha,k}^2 - x_{\alpha,i}x_{\alpha,j}\right]\mathrm{d}V\,.} \tag{44}$$

7.1.4 Linear and Angular Momentum of the Body

We can naturally write the linear momentum of a single particle, which is

$$\mathbf{p}'_\alpha = m_\alpha \mathbf{v}'_\alpha = m_\alpha(\mathbf{V} + \vec{\omega}\times\mathbf{r}_\alpha)$$

in the inertial reference frame. The total linear momentum is then the sum of each α-th particle, so

$$\mathbf{P}' = \sum_{\alpha=1}^{n}\mathbf{p}'_\alpha = \sum_{\alpha=1}^{n} m_\alpha(\mathbf{V} + \vec{\omega}\times\mathbf{r}_\alpha) = \sum_{\alpha=1}^{n} m_\alpha\mathbf{V} + \vec{\omega}\times\cancel{\left(\sum_{\alpha=1}^{n} m_\alpha\mathbf{r}_\alpha\right)} = M\mathbf{V}.$$

Thus, $\mathbf{P} = M\mathbf{V}$ is the total linear momentum, which is the momentum of the body as if all the mass were concentrated at the CM with velocity, $\mathbf{V}$. We can furthermore find the angular momentum of the body *about* the origin of the inertial frame, which is given by

$$\mathbf{L}'_\alpha = \mathbf{r}'_\alpha \times m_\alpha\mathbf{v}'_\alpha = \left(\mathbf{r}'_{\mathrm{CM}} + \mathbf{r}_\alpha\right)\times m_\alpha(\mathbf{V} + \vec{\omega}\times\mathbf{r}_\alpha).$$

The total angular momentum is simply the sum in α, such that

$$\begin{aligned}\mathbf{L}' &= \sum_{\alpha=1}^{n}\left[(\mathbf{r}_{\mathrm{CM}} + \mathbf{r}_\alpha)\times m_\alpha(\mathbf{V} + \vec{\omega}\times\mathbf{r}_\alpha)\right] \\ &= \sum_{\alpha=1}^{n} m_\alpha\mathbf{r}'_{\mathrm{CM}}\times\mathbf{V} + \cancel{\sum_{\alpha=1}^{n} m_\alpha(\mathbf{r}_{\mathrm{CM}}\times\vec{\omega}\times\mathbf{r}_\alpha)} + \cancel{\sum_{\alpha=1}^{n} m_\alpha\mathbf{r}_\alpha}\times\mathbf{V} + \sum_{\alpha=1}^{n} m_\alpha(\mathbf{r}_\alpha\times\vec{\omega}\times\mathbf{r}_\alpha).\end{aligned}$$

The second and third term cancels since we situate the origin at the CM, and we can rearrange the first term to find

$$M\mathbf{r}'_{\mathrm{CM}}\times\mathbf{V} = \mathbf{r}'_{\mathrm{CM}}\times M\mathbf{V} = \mathbf{r}'_{\mathrm{CM}}\times\mathbf{P}'.$$

Using vector calculus identities, we can also simplify the second term to be

$$\mathbf{r}_\alpha\times\vec{\omega}\times\mathbf{r}_\alpha = \vec{\omega}|\mathbf{r}_\alpha|^2 - \mathbf{r}_\alpha|\mathbf{r}_\alpha\cdot\vec{\omega}|$$

Thus, we find that

$$\mathbf{L}' = \mathbf{r}'_{\mathrm{CM}}\times\mathbf{P}' + \sum_{\alpha=1}^{n} m_\alpha\left[\vec{\omega}|\mathbf{r}_\alpha|^2 - \mathbf{r}_\alpha|\mathbf{r}_\alpha\cdot\vec{\omega}|\right].$$

Like the rotational kinetic energy, this angular momentum within the inertial frame is split into two components.

$$\mathbf{L}' = \mathbf{L}'_{\mathrm{trans.}} + \mathbf{L}_{\mathrm{rot.}}$$

We should be **very** careful here. If you are observant, you will notice that the angular momentum component in the inertial frame is the sum of the translational component of the angular momentum as seen *within* the inertial frame, *and* the angular momentum about the CM in the non-inertial frame.

It turns out, in fact, that the basis in which we calculate the rotational angular momentum component can be derived all the same. In this derivation, we have chosen that the basis to write the angular velocity vector and the vectors to each mass were done in the inertial frame. We *know* that there should be some component of this angular momentum that is dependent on the non-inertial frame, and that turns out to be the rotational component.

There is nothing large that changes if we redo this derivation in the non-inertial frame. It would just be more complicated.

However, it is true that for a rotating body, the inertia tensor becomes time-dependent in the inertial frame. This can be avoided by working within the CM frame, however the angular velocity compensates by becoming time-dependent. In any case, the matter is that these derivations (and henceforth solutions to our problems if we proceed further) are *only valid* at the instant of interest.

Returning back to our derivation, we can, in fact, simplify the rotational angular momentum if we look at the components of $\mathbf{L}_{\text{rot.}}$,

$$L_{\text{rot.},i} = \sum_{\alpha=1}^{n} m_\alpha \left[\omega_i \sum_{k=1}^{3} x_{\alpha,k}^2 - x_{\alpha,i}\left(\sum_{j=1}^{3} x_{\alpha,j}\omega_j\right)\right].$$

We can write the component ω_i as

$$\omega_i = \sum_{j=1}^{3} \omega_j \delta_{ij}$$

and so this term simplifies to

$$\begin{aligned}
L_{\text{rot.},i} &= \sum_{\alpha=1}^{n} m_\alpha \left[\left(\sum_{j=1}^{3} \omega_j \delta_{ij}\right) \sum_{k=1}^{3} x_{\alpha,k}^2 - x_{\alpha,i}\left(\sum_{j=1}^{3} x_{\alpha,j}\omega_j\right)\right] \\
&= \sum_{j=1}^{3} \omega_j \sum_{\alpha=1}^{n} m_\alpha \left[\delta_{ij} \sum_{k=1}^{3} x_{\alpha,k}^2 - x_{\alpha,i}x_{\alpha,j}\right] \\
&= \sum_{j=1}^{3} \omega_j I_{ij}.
\end{aligned}$$

Since this is just the i-th component of $\mathbf{L}_{\text{rot.}}$, we can write the full relation as

$$\mathbf{L}_{\text{rot.}} = \begin{bmatrix} L_{\text{rot.},1} \\ L_{\text{rot.},2} \\ L_{\text{rot.},3} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{3} \omega_j I_{1j} \\ \sum_{j=1}^{3} \omega_j I_{2j} \\ \sum_{j=1}^{3} \omega_j I_{3j} \end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}. \tag{45}$$

Thus, condensing this further, we arrive at

$$\boxed{\mathbf{L}_{\text{rot.}} = \{\mathbf{I}\}\vec{\omega}.} \tag{46}$$

We can see that the inertia tensor appears once more. In a way, you may be inclined to assume that because this is the case, the angular momentum and angular velocity vectors are parallel. However, this is not true in general due to the off-diagonal presence of the inertia tensor.

Notice, however, that the rotational angular momentum is very closely related to the rotational kinetic energy. In fact, we can rewrite our rotational kinetic energy,

$$T_{\text{rot.}} = \frac{1}{2}\sum_{i,j=1}^{3}\omega_i\omega_j I_{ij} = \frac{1}{2}\sum_{i=1}^{3}\omega_i\left[\sum_{j=1}^{3}\omega_j I_{ij}\right] = \frac{1}{2}\sum_{i=1}^{3}\omega_i L_{\text{rot.},i} = \frac{1}{2}\vec{\omega}^T\cdot\mathbf{L}_{\text{rot.}}$$

which is now in terms of the rotational angular momentum.

7.1.5 Forces and Torques of a Body

Assuming that there is an external force imposed on our system, that is, it acts on the rigid body in question, we can write the total force on the α-th mass as

$$\mathbf{F}_\alpha = \mathbf{F}_\alpha^{(e)} + \mathbf{f}_\alpha$$

where the superscript (e) denotes the external force imposed on the body. The vector, $\mathbf{f}_\alpha$ is the sum of the forces from all the other masses, which is

$$\mathbf{f}_\alpha = \sum_{\beta=1}^{n}\mathbf{f}_{\alpha\beta}$$

which is the force on the α-th mass due to the β-th mass. We can use Newton's 2nd law to determine the impulse of the α-th mass as

$$\dot{\mathbf{p}}'_\alpha = m_\alpha\ddot{\mathbf{r}}'_\alpha = \mathbf{F}_\alpha + \sum_{\beta=1}^{n}\mathbf{f}_{\alpha\beta}.$$

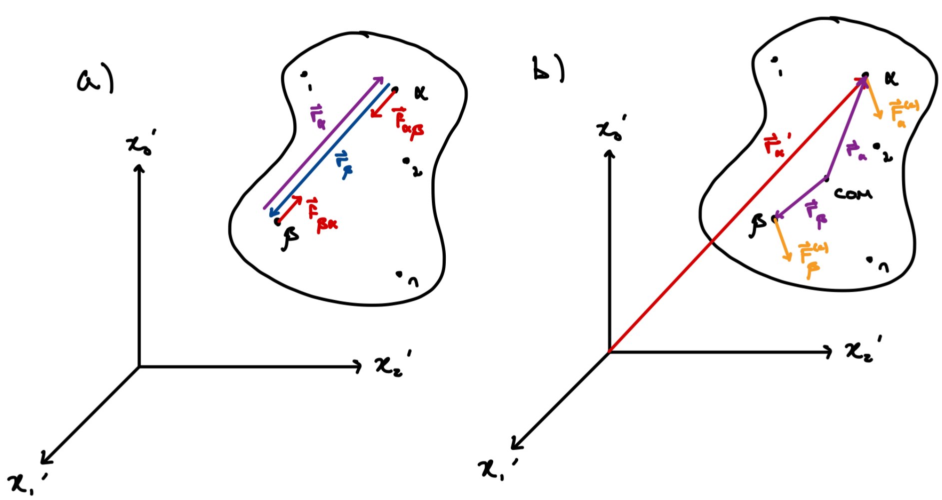


Figure 56: (a) Vectors $\mathbf{r}_\alpha$, $\mathbf{r}_\beta$, and their corresponding forces in the non-inertial frame. (b) Vectors $\mathbf{r}'_\alpha$ and its non-inertial counterparts.

For the entire rigid body, we thus have then total impulse

$$\dot{\mathbf{P}}' = \sum_{\alpha=1}^{n} \dot{\mathbf{p}}'_\alpha = \sum_{\alpha=1}^{n} \mathbf{F}_\alpha + \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \mathbf{f}_{\alpha\beta}\Bigg|_{\beta\neq\alpha}$$

where we have added the condition that we cannot have $\alpha = \beta$. Why? If we were to include that $\alpha = \beta$-th mass, we would be double-counting the force between each mass, which would affect the entire derivation.

Continuing, we know that Newton's 3rd law states

$$\mathbf{f}_{\alpha\beta} = -\mathbf{f}_{\beta\alpha}.$$

For the second term, we can rewrite the double sum expression by realizing that it is the sum of twice the force of an α-th mass to the β-th mass. However, twice the sum of this is the same as taking the sum of both the α-th and β-th mass on each other, with the bounds adjusted for double counting correction. In fact, we can see how they are equal by considering the simpler case $n = 3$. Here, we have

$$\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3} \mathbf{f}_{\alpha\beta} = [\mathbf{f}_{12} + \mathbf{f}_{21}] + [\mathbf{f}_{13} + \mathbf{f}_{31}] + [\mathbf{f}_{23} + \mathbf{f}_{32}].$$

If we shift the upper bound for the sum in β to $(\alpha - 1)$, we can further see that

$$\sum_{\alpha=1}^{3}\sum_{\beta=1}^{\alpha-1} = \left[0\right|_{\alpha=1} + \left[\mathbf{f}_{12} + \mathbf{f}_{21}\right|_{\alpha=2} + \left[\mathbf{f}_{13} + \mathbf{f}_{31} + \mathbf{f}_{23} + \mathbf{f}_{32}\right|_{\alpha=3}$$

and so it must be true that these sums are equal. Thus, for arbitrary n, it follows that

$$\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \mathbf{f}_{\alpha\beta} = \sum_{\alpha=1}^{n}\sum_{\beta=1}^{\alpha-1} [\mathbf{f}_{\alpha\beta} + \mathbf{f}_{\beta\alpha}].$$

Combining these two results, we can return to our equation for the total impulse, such that

$$\dot{\mathbf{P}}' = \sum_{\alpha=1}^{n} \mathbf{F}_\alpha + \cancel{\sum_{\alpha=1}^{n}\sum_{\beta=1}^{\alpha-1} [\mathbf{f}_{\alpha\beta} - \mathbf{f}_{\alpha\beta}]} = \sum_{\alpha=1}^{n} \mathbf{F}_\alpha = \mathbf{F}.$$

Furthermore, since we know that $\mathbf{P}' = M\mathbf{V}$, we can take the derivative of this quantity and equate it to the previous result, thus arriving at

$$\boxed{\dot{\mathbf{P}}' = M\dot{\mathbf{V}} \implies \mathbf{F} = M\mathbf{A}.} \tag{47}$$

Notice that this result shows that the action of the external forces on a rigid body is *equivalent* to summing them together, *and* acceleration the CM of the body. Thus, the total force exerted on the body itself acts as if all the mass were concentrated at the CM.

We can similarly discuss the rate of time derivative of angular momentum of any α-th mass, about the origin of the inertial frame. Here, we have

$$\dot{\mathbf{L}}'_\alpha = \mathbf{r}'_\alpha \times \dot{\mathbf{p}}_\alpha = \mathbf{r}'_\alpha \times \left[\mathbf{F}_\alpha^{(e)} + \sum_{\beta=1}^{n} \mathbf{f}_{\alpha\beta}\right]$$

and taking the sum over all α's,

$$\dot{\mathbf{L}}' = \sum_{\alpha=1}^{n} \dot{\mathbf{L}}'_\alpha = \sum_{\alpha=1}^{n} \left[\mathbf{r}'_\alpha \times \mathbf{F}_\alpha^{(e)}\right] + \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \left[\mathbf{r}'_\alpha \times \mathbf{f}_{\alpha\beta}\right].$$

We can look at the second term by again splitting it into its constituent parts. In fact,

$$\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \left[\mathbf{r}'_\alpha \times \mathbf{f}_{\alpha\beta}\right] = \sum_{\alpha=1}^{n}\sum_{\beta=1}^{\alpha-1} [\mathbf{r}_\alpha \times \mathbf{f}_{\alpha\beta} + \mathbf{r}_\beta \times \mathbf{f}_{\beta\alpha}] = \sum_{\alpha=1}^{n}\sum_{\beta=1}^{\alpha-1} [(\mathbf{r}_\alpha - \mathbf{r}_\beta) \times \mathbf{f}_{\alpha\beta}]$$

where we have invoked Newton's 3rd law. We neglected to state a specific property of Newton's 3rd law. The previous condition that $\mathbf{f}_{\alpha\beta} = -\mathbf{f}_{\beta\alpha}$ is known as the *weak* form of Newton's 3rd law. It states that the vectors are opposite in magnitude. However, there is also the *strong* form of Newton's 3rd law, which states that

$$\mathbf{f}_{\alpha\beta} = -\mathbf{f}_{\beta\alpha}, \quad (\mathbf{r}_\alpha \parallel \mathbf{r}_\beta).$$

That is, the forces are directed along the vector joining them. In this way, we will also invoke the strong from, and find that

$$\sum_{\alpha=1}^{n}\sum_{\beta=1}^{\alpha-1} [(\mathbf{r}_\alpha - \mathbf{r}_\beta) \times \mathbf{f}_{\alpha\beta}] = 0.$$

Therefore, we find that the time derivative of the angular momentum in the inertial frame is

$$\dot{\mathbf{L}}' = \sum_{\alpha=1}^{n} \mathbf{r}'_\alpha \times \mathbf{F}_\alpha^{(e)} = \sum_{\alpha=1}^{n} \mathbf{N}_\alpha = \mathbf{N}$$

where $\mathbf{N}_\alpha$ is the torque on the α-th particle. That is, we have

$$\boxed{\dot{\mathbf{L}}' = \mathbf{N}.} \tag{48}$$

which is the total external torque on the body.

7.2 Properties of the Inertia Tensor

Let us now revisit our discussion on the inertia tensor and define some important relations from it.

7.2.1 Parallel Axis Theorem

Consider the case of two non-inertial reference frames co-moving with a rigid body. The first non-inertial reference frame moves with the rigid body centred on the CM of the system, while the second non-inertial reference frame moves relative to the first by the vector $\mathbf{r}_{\text{CM}}$.

We shall denote each reference frame by the notation $\{x_i\}$, and using this the angular momentum about the origin of $\{\tilde{x}_i\}$ is

$$\tilde{\mathbf{L}} = \sum_{\alpha=1}^{n} \{[\mathbf{r}_{\text{CM}} + \mathbf{r}_\alpha] \times m_\alpha[\vec{\omega} \times (\mathbf{r}_{\text{CM}} + \mathbf{r}_\alpha)]\}.$$

Expanding this quantity out, we can find that

$$\tilde{\mathbf{L}} = \sum_{\alpha=1}^{n} [m_\alpha \mathbf{r}_{\text{CM}} \times \vec{\omega} \times \mathbf{r}_{\text{CM}}] + \cancel{\sum_{\alpha=1}^{n} [m_\alpha \mathbf{r}_{\text{CM}} \times \vec{\omega} \times \mathbf{r}_\alpha]} + \cancel{\sum_{\alpha=1}^{n} [m_\alpha \mathbf{r}_\alpha \times \vec{\omega} \times \mathbf{r}_{\text{CM}}]} + \sum_{\alpha=1}^{n} [m_\alpha \mathbf{r}_\alpha \times \vec{\omega} \times \mathbf{r}_\alpha].$$

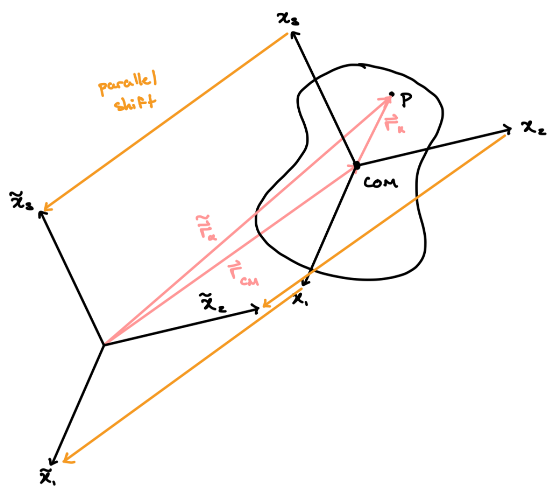


Figure 57: Two reference frames co-moving with a rigid body. Reference frame one, denoted with x_i is centred on the CM, while reference frame two, denoted with $\tilde{x}_i$, moves relative to reference frame one.

The second and third terms cancel since we are working in the CM frame, and so

$$\tilde{\mathbf{L}} = \sum_{\alpha=1}^{n} [m_\alpha \mathbf{r}_{\text{CM}} \times \vec{\omega} \times \mathbf{r}_{\text{CM}}] + \sum_{\alpha=1}^{n} [m_\alpha \mathbf{r}_\alpha \times \vec{\omega} \times \mathbf{r}_\alpha].$$

We now proceed by the same approach as before. With our vector calculus identites, we find

$$\mathbf{r}_{\text{CM}} \times \vec{\omega} \times \mathbf{r}_{\text{CM}} = \vec{\omega}|\mathbf{r}_{\text{CM}}|^2 - \mathbf{r}_{\text{CM}}|\mathbf{r}_{\text{CM}} \cdot \vec{\omega}|$$

and we already know the identity in $\mathbf{r}_\alpha$, and thus

$$\tilde{\mathbf{L}} = \sum_{\alpha=1}^{n} m_\alpha \big[\vec{\omega}|\mathbf{r}_{\text{CM}}|^2 - \mathbf{r}_{\text{CM}}|\mathbf{r}_{\text{CM}} \cdot \vec{\omega}|\big] + \sum_{\alpha=1}^{n} m_\alpha \big[\vec{\omega}|\mathbf{r}_\alpha|^2 - \mathbf{r}_\alpha|\mathbf{r}_\alpha \cdot \vec{\omega}|\big].$$

We furthermore rewrite the components of $\vec{\omega}$ and $\mathbf{r}_\alpha$ in terms of the Kronecker delta to show that the components of the angular momentum are given by

$$\tilde{L}_i = \sum_{j=1}^{3} \omega_j M \left[\delta_{ij} \sum_{k=1}^{3} x_{\text{CM},k}^2 - x_{\text{CM},i} x_{\text{CM},j}\right] + \sum_{j=1}^{3} \omega_j \sum_{\alpha=1}^{n} m_\alpha \left[\delta_{ij} \sum_{k=1}^{3} x_{\alpha,k}^2 - x_{\alpha,i} x_{\alpha,j}\right].$$

Notice that the first term looks exceedingly similar to that of the inertia tensor. In fact, it *is* the inertia tensor with the caveat that it is of a point mass, M, located at the position $\mathbf{r}_{\text{CM}}$.

Furthermore, since we are already working in the coordinate system $\{\tilde{x}_i\}$, we can write this simply as

$$\tilde{L}_i = \sum_{j=1}^{3} \omega_j \left[M\left(\delta_{ij} \sum_{k=1}^{3} x_{\mathrm{CM},k}^2 - x_{\mathrm{CM},i} x_{\mathrm{CM},j} \right) + \sum_{\alpha=1}^{n} m_\alpha \left(\delta_{ij} \sum_{k=1}^{3} x_{\alpha,k}^2 - x_{\alpha,i} x_{\alpha,j} \right) \right].$$

Assigning a new variable, J_{ij}, to the term in the square brackets, we thus arrive at

$$\tilde{L}_i = \sum_{j=1}^{3} \omega_j J_{ij}, \quad J_{ij} = M\left(\delta_{ij} \sum_{k=1}^{3} x_{\mathrm{CM},k}^2 - x_{\mathrm{CM},i} x_{\mathrm{CM},j} \right) + I_{ij}. \tag{49}$$

This is still ugly, so we can assign the inertia tensor about the CM to be $P_{ij,\mathbf{x}}$.

$$P_{ij} = M\left(\delta_{ij} \sum_{k=1}^{3} x_{\mathrm{CM},k}^2 - x_{\mathrm{CM},i} x_{\mathrm{CM},j} \right).$$

If we add a subscript, $\mathbf{x}$, to P_{ij} to imply that we are computing the terms in the reference frame of the CM system, we finally arrive at

$$\tilde{L}_i = \sum_{j=1}^{3} \omega_j J_{ij}, \quad J_{ij} = P_{ij,\mathbf{x}} + I_{ij}.$$

This, however, is quite annoying to see without expanding in its full tensor form, and so the full relation is

$$\boxed{\tilde{\mathbf{L}} = \{\mathbf{J}\}\vec{\omega}.} \tag{50}$$

The tensor, $\{\mathbf{J}\}$, is in fact a more general relation known as the **parallel axis theorem**, and its full form is

$$\boxed{\{\mathbf{J}\} = \{\mathbf{P}\}_{\mathbf{x}} + \{\mathbf{I}\}.} \tag{51}$$

It is useful to expand the additional contribution, $\{\mathbf{P}\}_{\mathbf{x}}$ in its arguments, to see that

$$\{\mathbf{P}\}_{\mathbf{x}} = M \begin{bmatrix} x_2^2 + x_3^2 & -x_1 x_2 & -x_1 x_3 \\ -x_1 x_2 & x_1^2 + x_3^2 & -x_2 x_3 \\ -x_1 x_3 & -x_2 x_3 & x_1^2 + x_2^2 \end{bmatrix}$$

which shows that $\{\mathbf{P}\}_{\mathbf{x}}$ is the inertia tensor of a single point about the CM. Immediately, we can see a very useful consequence of this relation. It is possible to convert the actual inertia tensor between two parallel, co-moving frames, neither of which may be centred at the CM. In this way, we use the CM frame as an intermediary, to find that

$$\begin{aligned} \{\mathbf{J}\} &= \{\mathbf{P}\}_{\mathbf{x}} + \{\mathbf{I}\} \\ \{\mathbf{K}\} &= \{\mathbf{P}\}_{\mathbf{y}} + \{\mathbf{I}\} \end{aligned}$$

where $\{\mathbf{K}\}$ is our intermediary tensor displaced from the CM by some vector $\mathbf{y}$ (note that $\mathbf{x}$ and $\mathbf{y}$ do not generally act in the direction of x_i). We can subtract these two equations, and find

$$\{\mathbf{J}\} = \{\mathbf{K}\} + \{\mathbf{P}\}_{\mathbf{x}} - \{\mathbf{P}\}_{\mathbf{y}}.$$

7.2.2 Principal Axes of Inertia

We already know that the angular frequency and angular momentum are not general. However, every rigid body that we deal with has a special set of axes of rotation where they *are* parallel. These axis are called the **principal axes**, and can be immensely useful for computation. To determine them, we need only to construct an eigenvalue problem of the form

$$\mathbf{L} = \{\mathbf{I}\}\vec{\omega} = \lambda\vec{\omega}$$

where the inertia tensor is written in its eigenbasis. That is, it is a diagonal matrix with the eigenvalues called the **principal moments of inertia**. In every case, we find that there are *always* 3 principal moments, and in certain cases they have different terms. We say that if

$$I_1 = I_2 = I_3$$

then the rigid body is a **spherical top**. If, instead, we have

$$I_1 = I_2 \neq I_3$$

then our body is considered a **symmetric top**. Naturally,

$$I_1 \neq I_2 \neq I_3$$

is an **asymmetric top**, and the case

$$I_1 = 0, \quad I_2 = I_3$$

is called a **rotor**. For these four cases, we find that the angular momentum and kinetic energies are simplified. In particular, we have

$$L_i = \sum_{j=1}^{3} \omega_j I_{ij}\delta_{ij} \implies L_i = I_i\omega_i$$

since $I_{ij}\delta_{ij}$ is simply a diagonal tensor; and,

$$T = \frac{1}{2}\sum_{i,j=1}^{3} I_{ij}\delta_{ij}\omega_i\omega_j \implies T = \frac{1}{2}\sum_{i=1}^{3} I_i\omega_i^2.$$

It should be obvious to you that these special cases are consequences of Newtonian energy and angular momentum relations. In fact, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ reduces to its arguments being $L_i = I_i\omega_i$ for each axes, and $T = mv^2/2$ reduces to $T = I_i\omega_i^2/2$.

7.3 Examples with the Inertia Tensor

7.3.1 The Dumbbell

Consider an object consisting of two point masses of mass, m, displaced from the origin by a distance, r. The masses are rotated from the $\hat{x}_1$-axis by an angle, θ, within the $\hat{x}_1$-$\hat{x}_2$ plane. Determine the inertia tensor by the parallel axis theorem.

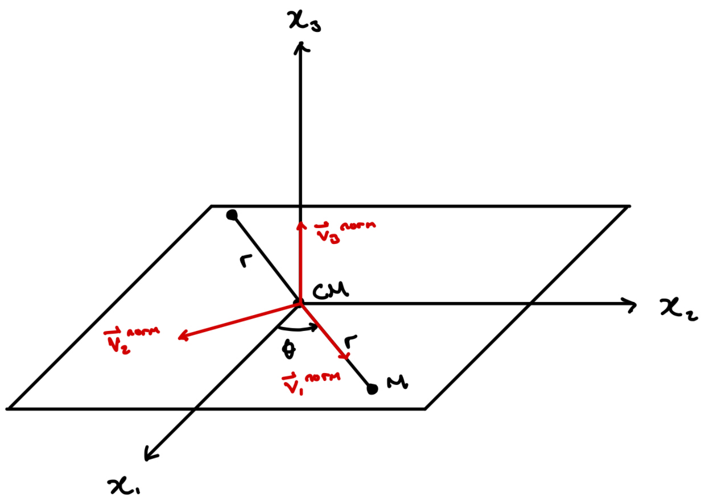


Figure 58: Two point masses of mass, m, rotating about the axis, x_3.

Here, we note that a point mass has zero inertia tensor about its own CM, so the contribution to the inertia tensor for each mass must be a result of parallel displacements, such that

$$\{\mathbf{I}\} = \{\mathbf{P}\}_{\mathbf{r}} + \{\mathbf{P}\}_{-\mathbf{r}}$$

where $\mathbf{r}$ denote the vector in making an angle θ with the $\hat{x}_1$-direction and furthermore

$$\mathbf{r} = [r\cos(\theta), r\sin(\theta), 0].$$

From our general expression of $\{\mathbf{P}\}_{\mathbf{x}}$, we find that

$$\{\mathbf{P}\}_{\mathbf{r}} = mr^2 \begin{bmatrix} \sin^2(\theta) & -\cos(\theta)\sin(\theta) & 0 \\ -\cos(\theta)\sin(\theta) & \cos^2(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Furthermore, notice that the two terms in $\mathbf{r}$ and $-\mathbf{r}$ are identical, and since the direction of our tensor is already encoded in its matrix form, we find

$$\{\mathbf{I}\} = 2\{\mathbf{P}\}_{\mathbf{r}} = 2mr^2 \begin{bmatrix} \sin^2(\theta) & -\cos(\theta)\sin(\theta) & 0 \\ -\cos(\theta)\sin(\theta) & \cos^2(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let us suppose now that the body is rotating about $\hat{x}_3$ with a constant angular velocity, $\vec{\omega} = [0, 0, \omega]$. We can compute the angular momentum of the body, which is

$$\mathbf{L} = \{\mathbf{I}\}\vec{\omega} = 2mr^2 \begin{bmatrix} \sin^2(\theta) & -\cos(\theta)\sin(\theta) & 0 \\ -\cos(\theta)\sin(\theta) & \cos^2(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \omega \end{bmatrix} = 2mr^2\omega \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Notice here that $\mathbf{L} \parallel \vec{\omega}$, which suggests that $\hat{x}_3$ is a principal axis. This, of course, is at the instant in the diagram shown, and we note that in general this is not the case, *unless* we have special cases where $\sin(\theta) = 0$ or $\sin(\theta) = 1$.

Given this, we are able to compute the principle axes and principal moments of inertia. By our inertia eigenvalue problem, we have

$$\mathbf{L} = \{\mathbf{I}\}\vec{\omega} = \lambda\vec{\omega}$$

and so we should arrive at

$$0 = \begin{vmatrix} \lambda - 2mr^2\sin^2(\theta) & 2mr^2\cos(\theta)\sin(\theta) & 0 \\ 2mr^2\cos(\theta)\sin(\theta) & \lambda - 2mr^2\cos^2(\theta) & 0 \\ 0 & 0 & \lambda - 2mr^2 \end{vmatrix}.$$

Solving this determinant, we find

$$-\lambda(\lambda - 2mr^2)^2 = 0$$

and so our eigenvalues are

$$\lambda_1 = 0 \quad \lambda_2 = 2mr^2$$

which has the eigenvalue λ_2 as a two-fold degenerate eigenvalue. Furthermore, since we already know one of the principal axes, $\hat{x}_3$, we can look to the $\hat{x}_1$-$\hat{x}_2$ plane to determine the direction of the other two. In this case, we can determine them by looking at our eigenvectors. Here, we set the parameter $c_3 = 0$ as we have already solved that case for the third row, giving us the new system

$$\begin{bmatrix} \lambda - 2mr^2\sin^2(\theta) & 2mr^2\cos(\theta)\sin(\theta) \\ 2mr^2\cos(\theta)\sin(\theta) & \lambda - 2mr^2\cos^2(\theta) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

We look at the first eigenvalue, $\lambda_1 = 0$, which suggests that

$$\begin{aligned} \cancel{2mr^2}\begin{bmatrix} -\sin^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & -\cos^2(\theta) \end{bmatrix} &\xrightarrow{R_1/R_{11}} \begin{bmatrix} 1 & -\cot(\theta) \\ \cos(\theta)\sin(\theta) & -\cos^2(\theta) \end{bmatrix} \\ &\xrightarrow{R_2/R_{21}} \begin{bmatrix} 1 & -\cot(\theta) \\ 1 & -\cot(\theta) \end{bmatrix} \\ &\xrightarrow{R_2-R_1} \begin{bmatrix} 1 & -\cot(\theta) \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Setting the free parameter to be $v_{12} = 0$, we find

$$\begin{aligned} v_{11} - v_{12}\cot(\theta) &= 0 \\ v_{12} &= t \end{aligned} \implies \mathbf{v}_1 = \begin{bmatrix} \cot(\theta) \\ 1 \end{bmatrix} \implies \mathbf{v}_1^{\text{norm.}} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

where we have normalized the vector also. Of course, since the system is really in 3D, the normalized vector should also contain a third argument in $\hat{x}_3$, which is zero. We can repeat this same process for the eigenvalue $\lambda_2 = 2mr^2$, and so we arrive at the two eigenvectors

$$\mathbf{v}_1^{\text{norm.}} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{bmatrix} \quad \mathbf{v}_2^{\text{norm.}} = \begin{bmatrix} \sin(\theta) \\ -\cos(\theta) \\ 0 \end{bmatrix}.$$

Now, what does this result tell us about our system? If we recall that the principal axes are given by the eigenvalues we found before, we can construct the principal axes inertia tensor, which is

$$\boxed{\{\mathbf{I}\}_{\text{PA}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2mr^2 & 0 \\ 0 & 0 & 2mr^2 \end{bmatrix}}.$$

Furthermore, this result arises from the direction of the eigenvectors. $\mathbf{v}_1^{\text{norm.}}$ suggests that the one of the principal moments of inertia lies at the position of the mass. Here, this is I_1, and naturally, suggests that the moment of inertia of I_1 is zero as we have stated that I_{ij} are zero at the masses. It follows that the other two principal moments are oriented orthogonal to $\mathbf{v}_1^{\text{norm.}}$, and we find that $\mathbf{v}_2^{\text{norm.}}$ is oriented perpendicular to the two masses. The final principal moment, of course, is located at $\hat{x}_3$.

7.3.2 The Rectangular Prism

Consider a rectangular prism with a uniform mass density, ρ, and total mass, M. Suppose that the rectangular prism is situated such that its corner lies at the origin of the reference frame $\{\hat{x}, \hat{y}, \hat{z}\}$. Determine from this the inertia tensor *about* the origin.

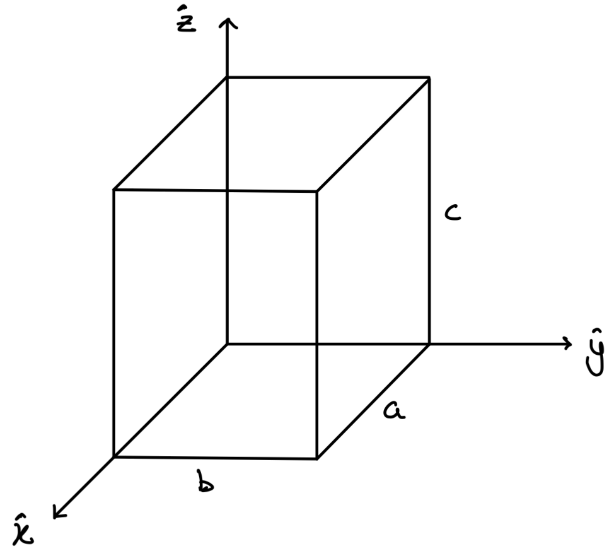


Figure 59: Rectangular prism of side lengths a, b, and c situated with its corner at the origin.

The arguments of the inertia tensor, for the prism corner resting at the origin, is given by

$$J_{ij} = \int_V \rho(\mathbf{r}) \left[\delta_{ij} \sum_{k=1}^{3} x_k^2 - x_i x_j \right] \mathrm{d}V \,.$$

For the case J_{11}, we have that

$$\begin{aligned} J_{11} &= \rho \int_{x=0}^{a} \int_{y=0}^{b} \int_{c=0}^{c} \left[\delta_{11} \left(x^2 + y^2 + z^2 \right) - x^2 \right] \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\ &= \rho \left[\int_0^a \mathrm{d}x \int_0^b y^2 \, \mathrm{d}y \int_0^c \mathrm{d}z + \int_0^a \mathrm{d}x \int_0^b \mathrm{d}y \int_0^c z^2 \, \mathrm{d}z \right] \\ &= \rho \left[a \frac{b^3}{3} c + ab \frac{c^3}{3} \right] = \frac{\rho}{3} (abc)(b^2 + c^2) \end{aligned}$$

but we know $\rho V = \rho(abc) = M$, so

$$J_{11} = \frac{M}{3}(b^2 + c^2).$$

We can use the argument that the diagonals are valid under a cyclic permutation of the coordinates, so we expect

$$J_{22} = \frac{M}{3}(a^2 + c^2) \quad \text{and} \quad J_{33} = \frac{M}{3}(a^2 + b^2).$$

Now we will consider the case J_{12}. Here, we have

$$\begin{aligned} J_{12} &= \rho \int_{x=0}^{a} \int_{y=0}^{b} \int_{z=0}^{c} \left[\delta_{12}(x^2 + y^2 + z^2) - xy\right] \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\ &= -\rho \int_0^a x \, \mathrm{d}x \int_0^b y \, \mathrm{d}y \int_0^c \mathrm{d}z \\ &= \rho \frac{a^2}{2} \frac{b^2}{2} c \\ &= -\frac{1}{4} \rho(abc) \\ &= -\frac{M}{4} ab = J_{21}. \end{aligned}$$

Again, we argue that this result is valid under a cyclic permutation of the coordinates, so

$$J_{13} = -\frac{M}{4} ab = J_{31} \quad \text{and} \quad J_{23} = -\frac{M}{4} bc = J_{32}.$$

Thus the expression of the inertia tensor about the origin is given by

$$\boxed{\{\mathbf{J}\} = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix} = \frac{M}{4} \begin{bmatrix} \frac{4}{3}(b^2 + c^2) & -ab & -ac \\ -ab & \frac{4}{3}(a^2 + c^2) & -bc \\ -ac & -bc & \frac{4}{3}(a^2 + b^2) \end{bmatrix}.}$$

A very useful fact about this derivation is that because the shape of the rectangular prism is general, we can take various limits of the expression to obtain the inertia tensor of other, related objects. For a cube, we can take the limit $(a = b = c)$, for a rectangular plane, $(a, b, c = 0)$, a square plane, $(a = b, c = 0)$, or a stick for $(a = b = 0, c)$.

7.3.3 Principal Moments of a Cube (Origin at Corner)

Suppose we have taken the limit of the rectangular prism to arrive at the expression for the inertia tensor of a cube *about* its corner. We wish to derive the principal moments of inertia and principal axes of the cube.

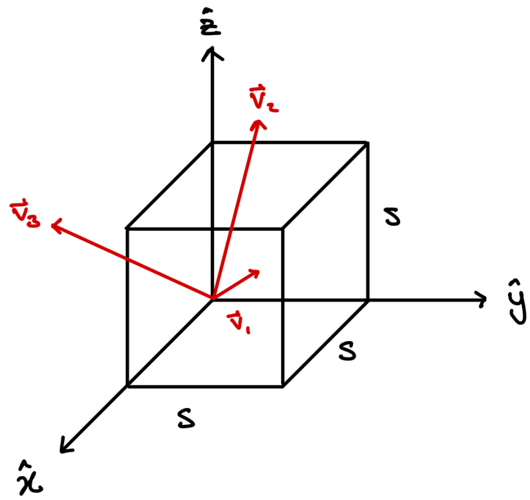


Figure 60: Cube with side lengths s, placed with its corner at the origin and eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ superimposed.

Here, our inertia tensor is of the form

$$\{\mathbf{J}\} = \frac{M}{4}s^2 \begin{bmatrix} \frac{8}{3} & -1 & -1 \\ -1 & \frac{8}{3} & -1 \\ -1 & -1 & \frac{8}{3} \end{bmatrix}$$

where we have factored out some quantities to make the result look nicer (make sure to not forget them when determining the principal axes). If you wish, you may want to confirm this yourself as an exercise.

To determine the principal axes, we use the inertia tensor eigenvalue problem,

$$\{\mathbf{J}\}\vec{\omega} = \lambda\vec{\omega}.$$

As the computation is rather formulaic, we will simply state the characteristic equation, which is of the form

$$M\lambda^3 + 2Ms^2\lambda^2 - \frac{55}{48}Ms^4 + \frac{121}{864}Ms^6 = 0.$$

From numerical calculations, the eigenvalues are

$$\lambda_1 = \frac{M}{6}s^2, \quad \lambda_2 = \frac{11M}{12}s^2$$

which is, again, two-fold degenerate in λ_2. The eigenvectors naturally follow and are of the form

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

however, in doing this calculation you will find that $\mathbf{v}_2$ and $\mathbf{v}_3$ are *not* orthogonal. In fact, the non-orthogonality of the eigenvectors arise from the two-fold degeneracy of our eigenvalues. This is a problem, since we cannot span the $\mathbb{R}^3$ space with two non-orthogonal vectors.

The solution to this issue is by constructing another eigenvector which is a linear combination of the two non-orthogonal ones. It turns out that any linear combination of $\mathbf{v}_2$ and $\mathbf{v}_3$ (and in general) is *also* an eigenvector with the same corresponding eigenvalue.

In this case, we will choose the eigenvector $\mathbf{v}_2$ to construct a new eigenvector, which is of the form

$$\mathbf{v}_3 = a \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

and now we impose the condition that

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 2a + b = 0. \tag{52}$$

In this way, we require another *two* parameters. Here, if we choose $a = 1$ and $b = -2$, then we find that our third eigenvector is

$$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Thus, we construct the principal moments about the corner of the cube.

7.3.4 Centre of Mass Inertia Tensor of the Rectangular Prism about

Suppose now we are interested in the centre of mass inertia tensor of the rectangular prism. There are two ways of finding this. We could calculate the inertia tensor explicitly while adjusting the bounds to account for the centre of mass. That is, integrate from $-r/2$ to $r/2$, for some length r. Or, we would use the parallel axis theorem. Here, we will use the latter.

We first define the centre of mass vector, which is

$$\mathbf{r} = \frac{1}{2} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Then, by the parallel axis theorem, we have

$$\{\mathbf{I}\} = \{\mathbf{J}\} - \{\mathbf{P}\}_{\mathbf{r}}.$$

The tensor, $\mathbf{P}_{\mathbf{r}}$ turns out to be

$$\{\mathbf{P}\}_{\mathbf{r}} = \frac{M}{4} \begin{bmatrix} b^2 + c^2 & -ab & -ac \\ -ab & a^2 + c^2 & -bc \\ -ac & -bc & a^2 + b^2 \end{bmatrix}$$

after evaluation, and the tensor $\{\mathbf{J}\}$ is

$$\{\mathbf{J}\} = \frac{M}{4} \begin{bmatrix} \frac{4}{3}(b^2 + c^2) & -ab & -ac \\ -ab & \frac{4}{3}(a^2 + c^2) & -bc \\ -ac & -bc & \frac{4}{3}(a^2 + b^2) \end{bmatrix}$$

as before. We can calculate certain components of this to determine the entire tensor, $\{\mathbf{I}\}$. Here, we have

$$I_{11} = \frac{M}{3}(b^2 + c^2) - \frac{M}{4}(b^2 + c^2) = \frac{M}{12}(b^2 + c^2)$$

which holds for all diagonal entries under a cyclic permutation of the axes lengths, and

$$I_{12} = -\frac{1}{4}ab + \frac{1}{4}ab = 0$$

which turns out to be the same for all off diagonal terms. Thus, the inertia tensor about the CM is

$$\boxed{\{\mathbf{I}\} = \frac{M}{12}\begin{bmatrix} b^2 + c^2 & 0 & 0 \\ 0 & a^2 + c^2 & 0 \\ 0 & 0 & a^2 + b^2 \end{bmatrix}}$$

which is a diagonal tensor, indicating that the axes $\hat{x}_1$, $\hat{x}_2$, and $\hat{x}_3$ are also the principal axes. Furthermore, the diagonal terms themselves are the principal moments of inertia, and as before, we can take appropriate limits to determine the inertia tensor of other, similar objects.

7.4 Euler Angles and Equations

In much of our problems, we have worked with two reference frames: lab (inertial) frame and the non-inertial frame rotating with the rigid body. We discuss the difference here.

The Lab (Inertial) Frame. The lab frame, which is our initial, inertial frame, is a frame which satisfies Newton's laws. The lab frame is different (we will see soon) from another, specialized inertial frame which handles the discussion of torques of a rigid body. In most cases, we are interested in the lab frame because it is the frame from which we observe the physical dynamics of our system.

The Body (Non-inertial) Frame. The body frame is the frame which typically co-rotates with the rigid body. In this frame, the inertia tensor, angular momentum, and kinetic energy terms are very simple when written in terms of the principal axes. Conventionally, we do calculations in terms of the principal axes in the body frame.

A large issue is connecting these two frames. Since our two frames share a common origin, we can transform from one to the other by a series of three rotations, called **Euler angles**. We denote the rotation types by the angles ϕ, θ, and ψ, which can describe the orientation of any rigid body in space.

We will, in particular, adopt the conventional sequence used in Thornton and Marion. That is, rotations about the axis $\hat{x}_3'$, then $\hat{x}_1''$, then $\hat{x}_3'''$. Here, the single prime refers to a inertial frame, but it may not be true that the frame for basis of more than a single prime are inertial.

Note also that other conventions are in use. For example, aerospace or robotics use different conventions, so it is important to exercise caution when referencing other literature; otherwise, you would arrive at an inconsistency.

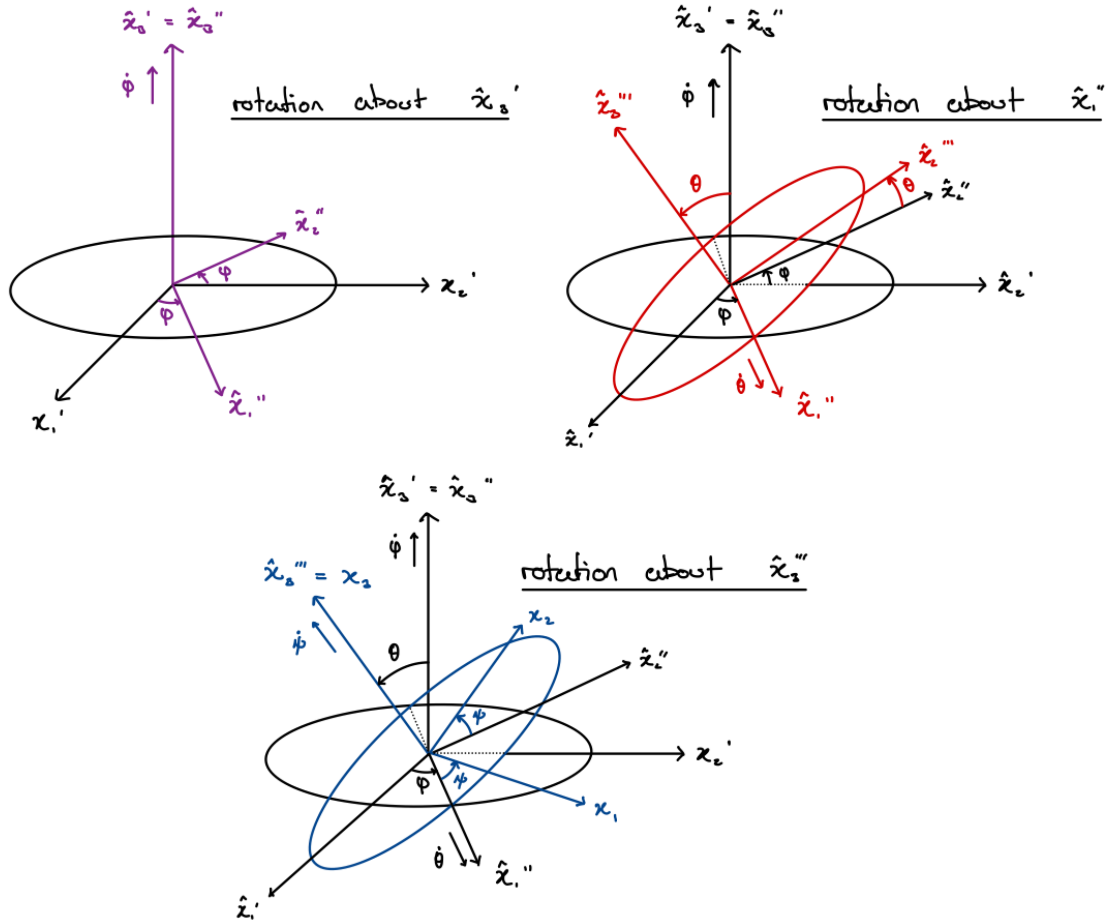


Figure 61: All three Euler angle rotations of an arbitrary rigid body. (Left) Rotation about $\hat{x}_3' = \hat{x}_3''$ for angle ϕ. (Middle) Rotation about $\hat{x}_1'' = \hat{x}_1'''$ for angle θ. (Right) Rotation about $\hat{x}_3''' = \hat{x}_3$ for angle ψ.

We, in particular, say that the Euler angles and thereby rotations are a set of *passive* transformations. As a whole, their net effect is to transform any vector in our lab frame, $\{\hat{x}_i'\}$, to a vector in the body frame. This transformation is always of the form

$$\boxed{\mathbf{A} = \mathbf{R}\mathbf{A}'} \tag{53}$$

for some vector, $\mathbf{A}'$ originally in the lab frame denoted by the prime. Fig. 60 shows the three steps.

Step 1 (Rotation about ϕ). We apply a counter-clockwise rotation about the $\hat{x}_3'$ axis according to the angle, ϕ, given by

$$\mathbf{R}_\phi^{(3)} = \begin{bmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We give $\mathbf{R}$ the subscript ϕ to denote that it is a rotation about the angle ϕ, and additionally add the superscript (3) to remind you that it is about the $\hat{x}_3'$ axis. In turn, we arrive at the vector

$$\mathbf{A}'' = \mathbf{R}_\phi^{(3)}\mathbf{A}'.$$

Step 2 (Rotation about θ). We next apply another counter-clockwise rotation, this time about the $\hat{x}_1''$ axis according to the angle θ, given by

$$\mathbf{R}_\theta^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

In turn, our vector becomes

$$\mathbf{A}''' = \mathbf{R}_\theta^{(1)}\mathbf{A}'' = \mathbf{R}_\theta^{(1)}\mathbf{R}_\phi^{(3)}\mathbf{A}'.$$

We will stress that the order in which these vectors are multiplied matters. Rotations **do not commute**, and so for consistency we must first compute the rotation $\mathbf{R}_\phi^{(3)}\mathbf{A}'$.

Step 3 (Rotation about ψ). We lastly apply the counter-clockwise rotation about $\hat{x}_3'''$ according to the angle ψ, given by

$$\mathbf{R}_\psi^{(3)} = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In turn, we arrive at the final transformed vector,

$$\mathbf{A} = \mathbf{R}_\psi^{(3)}\mathbf{A}''' = \mathbf{R}_\psi^{(3)}\mathbf{R}_\theta^{(1)}\mathbf{R}_\phi^{(3)}\mathbf{A}'.$$

Notice, in this way we find the full form of our rotation matrix. For the three Euler angles, we can compute the full, single rotation

$$\mathbf{R} = \mathbf{R}_\psi^{(3)}\mathbf{R}_\theta^{(1)}\mathbf{R}_\phi^{(3)}.$$

Expanding this out, the rotation matrix is thus

$$\mathbf{R} = \begin{bmatrix} \cos(\phi)\cos(\psi) - \cos(\theta)\sin(\phi)\sin(\psi) & \sin(\phi)\cos(\psi) + \cos(\theta)\cos(\phi)\sin(\psi) & \sin(\theta)\sin(\psi) \\ -\sin(\phi)\cos(\theta)\cos(\psi) - \cos(\phi)\sin(\psi) & \cos(\phi)\cos(\theta)\cos(\psi) - \sin(\phi)\sin(\psi) & \sin(\theta)\cos(\psi) \\ \sin(\phi)\sin(\theta) & -\cos(\phi)\sin(\theta) & \cos(\theta) \end{bmatrix}$$

and of course, for compactness, this can be given by

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{32} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

where the arguments are given by the transformations above. Naturally, we can take the inverse of this matrix to find the body $\to$ lab frame counterpart, but manually computing the inverse is abysmal so we exploit the orthogonality of the transformations to find

$$\mathbf{A}' = \mathbf{R}^T\mathbf{A}$$

where $\mathbf{R}^T$ denotes the transpose of $\mathbf{R}$.

We can now revisit our discussion on kinetic energies and rotational angular momentum. To calculate these quantities, we will, in general, need to determine the angular velocities in both frames. Our Euler angles that we have just defined allow us to do just this. In accordance with each rotation, we can assign its corresponding angular velocity vector, such that

$$\vec{\omega}_\phi = \dot{\phi}\hat{x}_3', \quad \vec{\omega}_\theta = \dot{\theta}\hat{x}_1'', \quad \vec{\omega}_\psi = \dot{\psi}\hat{x}_3'''.$$

To write our angular velocities in the body frame, we require the transformation of these vectors by the Euler angles. In general, we can write the transformed, full angular velocity vector as

$$\vec{\omega} = \mathbf{R}_\psi^{(3)}\mathbf{R}_\theta^{(1)}\mathbf{R}_\phi^{(3)}\vec{\omega}_\phi + \mathbf{R}_\psi^{(3)}\mathbf{R}_\theta^{(1)}\vec{\omega}_\theta + \mathbf{R}_\psi^{(3)}\vec{\omega}_\psi$$

and, in doing so, we find that our angular velocity is given by

$$\boxed{\vec{\omega} = \begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\end{bmatrix} = \begin{bmatrix}\dot{\phi}\sin(\theta)\sin(\psi) + \dot{\theta}\cos(\psi)\\ -\dot{\theta}\sin(\psi) + \dot{\phi}\sin(\theta)\cos(\psi)\\ \dot{\phi}\cos(\theta) + \dot{\psi}\end{bmatrix}.} \tag{54}$$

Furthermore, we can invert this expression to obtain the coupled DE's for the Euler angles, which turns out to be the system

$$\begin{aligned}\dot{\phi} &= \frac{1}{\sin(\theta)}[\omega_1\sin(\psi) + \omega_2\cos(\psi)]\\ \dot{\theta} &= \omega_1\cos(\psi) - \omega_2\sin(\psi)\\ \dot{\psi} &= \omega_3 - \frac{\cos(\theta)}{\sin(\theta)}[\omega_1\sin(\psi) + \omega_2\cos(\psi)].\end{aligned}$$

Analogously, in the lab frame, our angular velocity is given by

$$\vec{\omega}' = \mathbf{R}_\phi^{-1(3)}\mathbf{R}_\theta^{-1(1)}\mathbf{R}_\psi^{-1(3)}\vec{\omega}_\psi + \mathbf{R}_\phi^{-1(3)}\mathbf{R}_\theta^{-1(1)}\vec{\omega}_\theta + \mathbf{R}_\phi^{-1(3)}\vec{\omega}_\phi.$$

This, in turn, becomes

$$\boxed{\vec{\omega}' = \begin{bmatrix}\omega_1'\\ \omega_2'\\ \omega_3'\end{bmatrix} = \begin{bmatrix}\dot{\psi}\sin(\phi)\sin(\theta) + \dot{\theta}\cos(\phi)\\ -\dot{\psi}\cos(\phi)\sin(\theta) + \dot{\theta}\sin(\phi)\\ \dot{\psi}\cos(\theta) + \dot{\phi}\end{bmatrix}.} \tag{55}$$

If we now recall that the angular velocities within the body frame lie on a principal axis, that is, they satisfy the moment of inertia eigenvalue problem

$$\mathbf{L} = \{\mathbf{I}\}\vec{\omega} = \lambda\vec{\omega}$$

then we can define the angular momentum vector in the form

$$\mathbf{L} = \begin{bmatrix}I_1\omega_1\\ I_3\omega_3\\ I_3\omega_3\end{bmatrix}$$

where I_i, $i = 1, 2, 3$ refer to the principal moments of inertia. In the body frame, however, we know that this quantity is non-inertial, and so we can take the time derivative of the angular momentum and find the torque in an inertial frame, such that

$$\mathbf{N}' = \dot{\mathbf{L}} + \vec{\omega}\times\mathbf{L} = \frac{\delta\mathbf{L}}{\delta t} + \begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\end{bmatrix} \times \begin{bmatrix}I_1\omega_1\\ I_2\omega_2\\ I_3\omega_3\end{bmatrix}.$$

Explicitly calculating this quantity, we thus arrive at the system of equations

$$\mathbf{N}' = \begin{bmatrix} I_1\dot{\omega}_1 - \omega_2\omega_3(I_2 - I_3) \\ I_2\dot{\omega}_2 - \omega_3\omega_1(I_3 - I_1) \\ I_3\dot{\omega}_3 - \omega_1\omega_2(I_1 - I_2) \end{bmatrix} = \mathbf{N}. \tag{56}$$

which are known as **Euler's equations**, which relate the dynamics of the angular frequency in the body frame to the external torque imposed on the system. These equations, however, are challenging to solve even under the condition of constant torque. The components of these equations change as the body rotates, and so these equations are, in general, time-dependent.

Furthermore, to solve these equations to must express everything in the system in terms of a common set of coordinates; typically, we do this by writing the components of $\mathbf{N}$ in the body frame.

7.5 A Note on Basis and Euler's Equations

We momentarily stop here to discuss about the basis of our system when finding angular momentum and torque in our frames. If you are observant, you may have noticed that at some point, we switched basis to determine Euler's equations. This may strike you as odd, as we know that in general, the torque in the body frame is not the same as an inertial frame. Let us consider two specific cases.

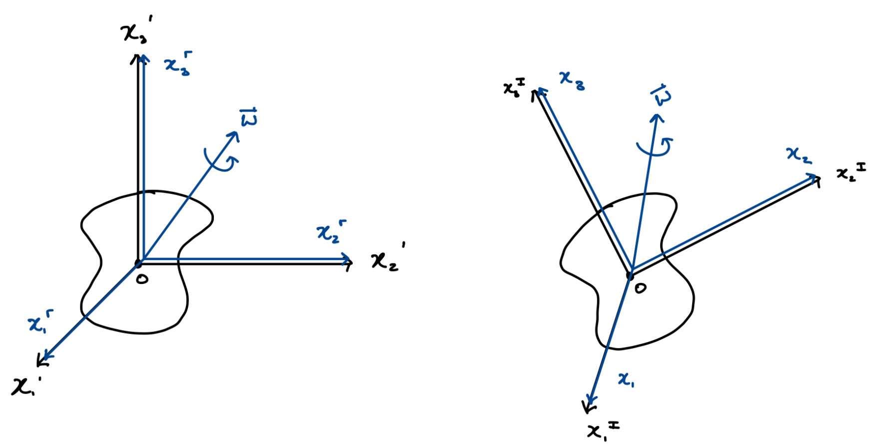


Figure 62: Two cases of rotating and non-inertial frame. (Left) Lab frame and an arbitrary rotating frame. (Right) Body frame and some arbitrary inertial frame.

First, we will discuss the case of an inertial, lab frame and some arbitrary rotating frame which coincides with the lab frame at *some* instant. Here, we denote the lab frame as $\{\hat{x}_i'\}$, and the arbitrary rotating frame as $\{\hat{x}_i^r\}$. We will stress that this arbitrary frame is *not* the body frame.

We can determine the torque in the lab frame by taking the time-derivative of the angular momentum in its corresponding frame. That is, we find

$$\frac{\mathrm{d}\mathbf{L}'}{\mathrm{d}t} = \mathbf{N}' \implies \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} L_1' \\ L_2' \\ L_3' \end{bmatrix} = \begin{bmatrix} N_1' \\ N_2' \\ N_3' \end{bmatrix}.$$

At the same time, we also know that the time derivative of the angular momentum is given by

$$\frac{\mathrm{d}\mathbf{L}'}{\mathrm{d}t} = \frac{\delta \mathbf{L}^r}{\delta t} + \vec{\omega} \times \mathbf{L}^r = \begin{bmatrix} \dot{L}_1^r \\ \dot{L}_2^r \\ \dot{L}_3^r \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \times \begin{bmatrix} L_1^r \\ L_2^r \\ L_3^r \end{bmatrix} = \begin{bmatrix} N_1' \\ N_2' \\ N_3' \end{bmatrix} = \begin{bmatrix} N_1^r \\ N_2^r \\ N_3^r \end{bmatrix}.$$

At the instant where the frame coincides, the torque in the lab frame and the arbitrary rotating frame *agree*. However, in general, this is *not* the case. If we consider the full time rotation, the angular momentum in the rotating frame is given by

$$\mathbf{L}^r = \{\mathbf{I}\}^r \vec{\omega} \implies \dot{\mathbf{L}}^r = \{\dot{\mathbf{I}}\}^r \vec{\omega} + \{\mathbf{I}\} \dot{\vec{\omega}}$$

by product rule. Furthermore, $\{\mathbf{I}\}^r$ will, in general, be a function of time as the frame is not necessarily co-rotating with the body *and* $\{\mathbf{I}\}^r$ is not, in general, diagonal.

On the other hand, let us now consider the case of the body frame and some arbitrary inertial frame that coincides with the body frame at an instant in time. Again, we stress that the inertial frame is *not* the lab frame, and we denote each frame as $\{\hat{x}_i\}$ and $\{\hat{x}_i^I\}$ respectively.

In the inertial frame, the torque is given by

$$\mathbf{N}^I = \frac{\mathrm{d}\mathbf{L}^I}{\mathrm{d}t}$$

while in the body frame we also have

$$\mathbf{N}^I = \frac{\delta \mathbf{L}}{\delta t} + \vec{\omega} \times \mathbf{L} = \mathbf{N}$$

at the instant where the frames coincide. In this case, we know that $\dot{\mathbf{L}}$ is rather simple, since $\{\mathbf{I}\}$ is a *constant and diagonal* in the body frame. This simpler version, in fact, is related to the two frames which we define the Euler equations. We choose the simpler basis, where the body frame is co-rotating with the body itself, and the non-inertial frame in which we derive the Euler equations is arbitrary, valid *only* in the instant in time where the frames coincide.

7.6 Torque-Free Tops and Conservation Laws

Let us now consider the special case of zero external torque. Rigid bodies, in this case, are governed by Euler's equations on the condition that $\mathbf{N} = 0$. Here, we have the system of equations

$$\begin{aligned} I_1\dot{\omega}_1 - \omega_2\omega_3(I_2 - I_3) &= 0 \\ I_2\dot{\omega}_2 - \omega_3\omega_1(I_3 - I_1) &= 0 \\ I_3\dot{\omega}_3 - \omega_1\omega_2(I_1 - I_2) &= 0. \end{aligned}$$

Note that, if we consider the component $\omega_i\omega_j$, we can use the relation that $L_i = I_i\omega_i$ to find

$$\omega_i\omega_j = \frac{L_iL_j}{I_iI_j}$$

and furthermore $I_i\dot{\omega}_i = L_i$. Thus we can rewrite our Euler equations in the form

$$\begin{aligned} \dot{L}_1 &= \frac{I_2 - I_3}{I_2I_3}L_2L_3 \\ \dot{L}_2 &= \frac{I_3 - I_1}{I_3I_1}L_3L_1 \\ \dot{L}_3 &= \frac{I_1 - I_2}{I_1I_2}L_1L_2. \end{aligned}$$

Multiplying each equation through by L_i we arrive at the new system of equations

$$L_1\dot{L}_1 = \frac{I_2 - I_3}{I_2 I_3} L_2 L_3 L_1$$
$$L_2\dot{L}_2 = \frac{I_3 - I_1}{I_3 I_1} L_3 L_1 L_2$$
$$L_3\dot{L}_3 = \frac{I_1 - I_2}{I_1 I_2} L_1 L_2 L_3.$$

Notice that if we sum these expressions, we find

$$L_1\dot{L}_1 + L_2\dot{L}_2 + L_3\dot{L}_3 = \left[\frac{I_2 - I_3}{I_2 I_3} + \frac{I_3 - I_1}{I_3 I_1} + \frac{I_1 - I_2}{I_1 I_2}\right] L_1 L_2 L_3.$$

The left hand side is the derivative of the squared components up to a constant and the right side equates to zero. We can see this explicitly, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L_1^2 + L_2^2 + L_3^2\right] = 2\left[L_1\dot{L}_1 + L_2\dot{L}_2 + L_3\dot{L}_3\right] = \frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{L}\cdot\mathbf{L}]$$

and

$$\frac{I_2 - I_3}{I_2 I_3} + \frac{I_3 - I_1}{I_3 I_1} + \frac{I_1 - I_2}{I_1 I_2} = \frac{I_1(I_2 - I_3) + I_2(I_3 - I_1) + I_3(I_1 - I_2)}{I_1 I_2 I_3} = 0$$

and so we find

$$\frac{\mathrm{d}(\mathbf{L}\cdot\mathbf{L})}{\mathrm{d}t} = 0.$$

This, in fact, tells us that

$$L^2 = \text{conserved}$$

where L^2 is the magnitude of the angular momentum, and furthermore suggests a spherical geometry for our system. Correspondingly, we find that

$$\mathbf{L}' = \text{conserved}$$

as observed in the fixed frame. On the other hand, we can switch to the body frame to find that

$$T = \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2$$

and taking the time derivative,

$$\dot{T} = \omega_1\dot{\omega}_1 I_1 + \omega_2\dot{\omega}_2 I_2 + \omega_3\dot{\omega}_3 I_3.$$

However, we know that by Euler's equations, we can rewrite this as

$$\dot{T} = \omega_1\omega_2\omega_3[(I_2 - I_3) + (I_3 - I_1) + (I_1 - I_2)] = 0$$

which suggests that

$$T = \text{conserved}$$

as well. Thus, in the case of a torque-free top, the kinetic energy is conserved within the body frame. This quantity is scalar, so we naturally expect it to be conserved in the lab frame as well.

On the other hand, we find that L^2, the magnitude of angular momentum, to be conserved in the body frame. This is also scalar, so in the lab frame it remains conserved. *Furthermore*, we find that $\mathbf{L}'$ is conserved since we also know that $\mathbf{N} = 0$; this, however, does *not* hold in the body frame. That is, $\mathbf{L}$ is *not* conserved in the body frame.

7.7 Torque-Free Symmetric Top

Consider the case of a rigid body where $I_1 = I_2$, but $I_1 \neq I_3$ or equivalently $I_2 \neq I_3$. We say that the rigid body adhering to these conditions is a **symmetric top**; and, furthermore, if

$$\begin{cases} I_1 < I_3, & \text{rigid body is } \textbf{oblate} \\ I_1 > I_3, & \text{rigid body is } \textbf{prolate}. \end{cases}$$

An oblate rigid body is an example of a coin; the first moment is less than the third moment, and so the body is skewed to have most of its mass compressed about $\hat{x}_1$ and $\hat{x}_2$. A prolate rigid body is an example of a pen; the first moment is larger than the third moment, and so the body is skewed along $\hat{x}_3$.

We begin the discussion of the symmetric top in the body frame. Here, the Euler equations show that

$$\begin{aligned} I_1\dot{\omega}_1 - \omega_2\omega_3(I_1 - I_3) &= 0 \\ I_1\dot{\omega}_2 - \omega_3\omega_1(I_3 - I_1) &= 0 \\ I_3\dot{\omega}_3 &= 0 \end{aligned}$$

where we have substituted I_2 for I_1. Notice that the third equation, assuming that I_1, I_2, and I_3 are constant, reduces to

$$\dot{\omega}_3 = 0 \implies \omega_0 = \omega_{3,0}$$

for some constant $\omega_{3,0}$. In this we, we can rewrite the first Euler equation to be

$$I_1\dot{\omega}_1 = \omega_2\omega_{3,0}(I_1 - I_3) \implies \dot{\omega}_1 = \left[\omega_{3,0}\frac{I_1 - I_3}{I_1}\right]\omega_2.$$

If we make the substitution

$$\Omega = \left[\omega_{3,0}\frac{I_1 - I_3}{I_1}\right]$$

we can reduce this expression to

$$\dot{\omega}_1 = \Omega\omega_2.$$

Similarly, we can find that

$$\dot{\omega_2} = \left[\omega_{3,0}\frac{I_3 - I_1}{I_1}\right]\omega_2 = -\Omega\omega_1$$

and so we have the two DE's

$$\begin{aligned} \dot{\omega}_1 &= \Omega\omega_2 \\ \dot{\omega}_2 &= -\Omega\omega_1. \end{aligned}$$

The form of these equations should be immediately familiar to you. We have seen them in the example of a ball on a rotating platform, or a particle constrained to a rotating potential.

Solving this system, we can find immediately that

$$\begin{aligned} \ddot{\omega}_1 &= \Omega\dot{\omega}_2 = -\Omega^2\omega_1 \\ \ddot{\omega}_2 &= -\Omega\dot{\omega}_1 = -\Omega^2\omega_2 \end{aligned}$$

which are both in the form of SHM. Furthermore, their solutions are governed by

$$\omega_1 = A\cos(\Omega t) + B\sin(\Omega t)$$
$$\omega_2 = -A\sin(\Omega t) + B\cos(\Omega t).$$

We can determine the coefficients of the solutions by the initial conditions. Let us consider the case

$$\mathbf{L}'(0) = L\hat{x}_3'.$$

That is, the angular momentum is initially directed upwards. Since there are no torques present on the system, we know that $\mathbf{L}'$ will be a conserved quantity in the fixed frame. However, we cannot work with this vector in the fixed frame, so applying the transformation matrix $\mathbf{R}$, we have

$$\mathbf{L} = \mathbf{R}\mathbf{L}' = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{32} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ L \end{bmatrix} = \begin{bmatrix} R_{13}L \\ R_{23}L \\ R_{33}L \end{bmatrix} = L \begin{bmatrix} \sin(\theta_0)\sin(\psi_0) \\ \sin(\theta_0)\cos(\psi_0) \\ \cos(\theta_0) \end{bmatrix}$$

where we denote the initial angles accordingly. Notice that θ_0, ψ_0, and ϕ_0 determine the initial orientation of the body. In fact, it is not enough to simply specify $\mathbf{L}'(0)$, since that does not tell us how the body is able to evolve given its initial orientation.

In the body frame, we also know that the angular momentum vector follows the moment of inertia eigenvalue problem. In this case, we can write

$$\begin{bmatrix} \omega_1(0) \\ \omega_2(0) \\ \omega_3(0) \end{bmatrix} = \begin{bmatrix} \sin(\theta_0)\sin(\psi_0)\frac{L}{I_1} \\ \sin(\theta_0)\cos(\psi_0)\frac{L}{I_1} \\ \cos(\theta_0)\frac{L}{I_3} \end{bmatrix} = \begin{bmatrix} A \\ B \\ \omega_{3,0} \end{bmatrix}.$$

With this, we are able to solve for the time evolution of the angular velocity. Beginning with the first component,

$$\begin{aligned} \omega_1 = A\cos(\Omega t) + B\sin(\Omega t) &= \frac{L}{I_1}[\sin(\theta_0)\sin(\psi_0)\cos(\Omega t) + \sin(\theta_0)\cos(\psi_0)\sin(\Omega t)] \\ &= \frac{L}{I_1}\sin(\theta_0)\sin(\Omega t + \psi_0) \end{aligned}$$

by the sum angle identity. Similarly, we have in ω_2,

$$\omega_2 = \frac{L}{I_1}\sin(\theta_0)\cos(\Omega t + \psi_0)$$

and in ω_3,

$$\omega_3 = \frac{L}{I_3}\cos(\theta_0).$$

Thus, our angular velocity is given by

$$\boxed{\vec{\omega}(t) = L \begin{bmatrix} \sin(\theta_0)\sin(\Omega t + \psi_0)/I_1 \\ \sin(\theta_0)\cos(\Omega t + \psi_0)/I_1 \\ \cos(\theta_0)/I_3 \end{bmatrix}}.$$

If we recall, furthermore, for a torque-free top that K^2 is conserved in both frame, we can simply take the squared magnitude of $\mathbf{L}$ to find that the trajectories of $\mathbf{L}$ are paths on a sphere. Specifically,

$$|\vec{\omega}(t)|^2 = L^2\left[\frac{\sin^2(\theta_0)}{I_1^2} + \frac{\cos^2(\theta_0)}{I_3^2}\right]$$

and since the magnitude of $\vec{\omega}(t)$ is independent of t and constant, it thus always is the same distance away from the origin of our non-inertial reference frame. The radius, in particular, is given by

$$|\vec{\omega}(t)| = L\sqrt{\frac{\sin^2(\theta_0)}{I_1^2} + \frac{\cos^2(\theta_0)}{I_3^2}}$$

where the trajectory of the rigid body follow lines of latitude with constant L_3.

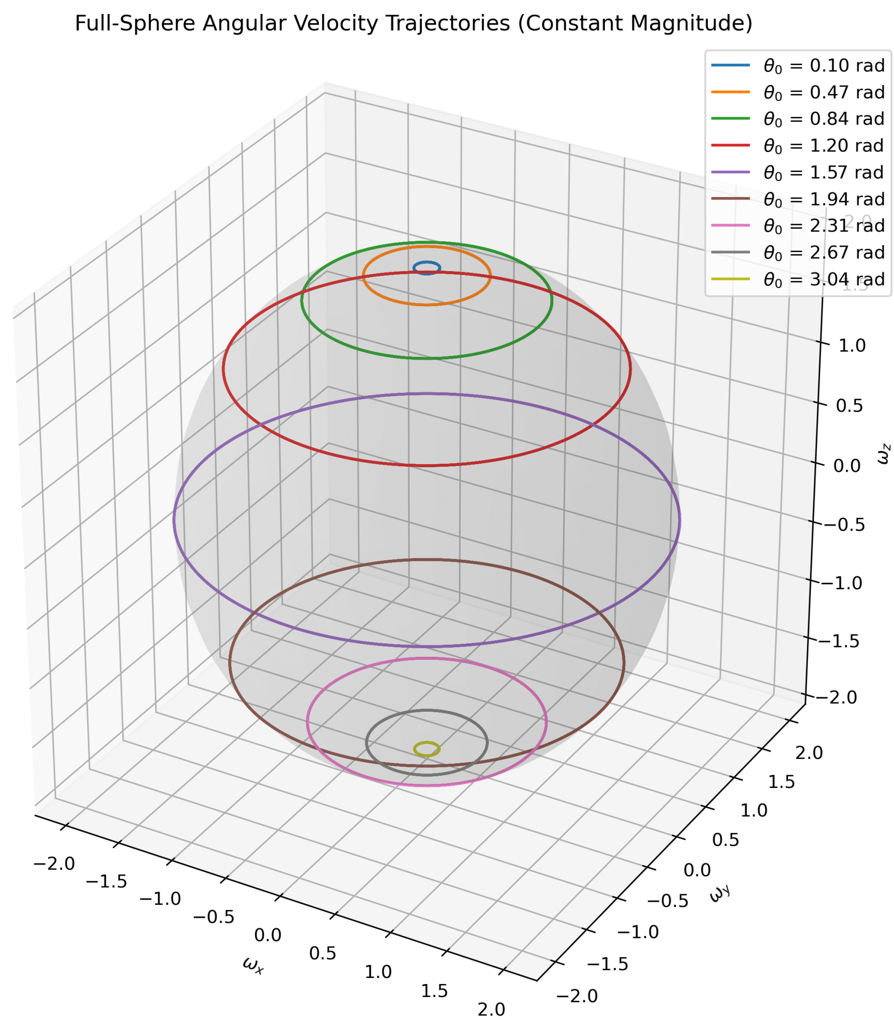


Figure 63: Several trajectories of a rigid body making a trajectory of constant radius, R, around a sphere.

There are other quantities of the system that we can analyze. The precession frequency, we found was given by

$$\Omega = \omega_{3,0}\frac{I_1 - I_3}{I_1}.$$

If we have $\Omega, \omega_{3,0} > 0$, then we have the case $I_1 > I_3$ for a prolate top. On the other hand, if $\Omega < 0$ and $\omega_{3,0} > 0$ or $\Omega > 0$ and $\omega_{3,0} < 0$, then we have the case $I_1 < I_3$ for a oblate top. In the special case, $\Omega = \omega_{3,0} = 0$, then we find that there is no trajectory, which lies exactly at the equator of our sphere. It may also not surprise you that there are no trajectories on the pole of the sphere, but rather constant angular motion. In fact, for the case $\sin(\theta) = 0, \pi$, we have

$$\vec{\omega}(t) = \begin{bmatrix} 0 \\ 0 \\ L/I_3 \end{bmatrix} \quad \text{or} \quad \vec{\omega}(t) = \begin{bmatrix} 0 \\ 0 \\ -L/I_3 \end{bmatrix}$$

which is time-independent. Physically, these cases correspond to rotating the top about its principal axes.

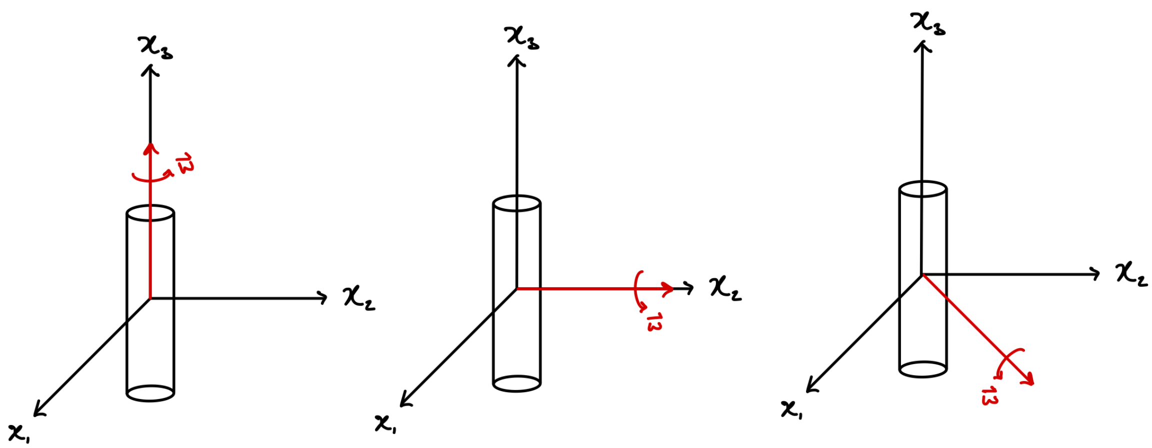


Figure 64: Prolate top rotating about its principal axes.

These quantities which we've just discussed, however, do *not* tell us about the orientation of the body as a function of time. To determine how the rigid body evolves in time, we must return back to the fixed frame and utilize the Euler angles. In particular, our angular velocity suggests

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = \begin{bmatrix} L_1/I_1 \\ L_2/I_2 \\ L_3/I_3 \end{bmatrix} = L \begin{bmatrix} \sin(\theta_0)\sin(\Omega t + \psi_0)/I_1 \\ \sin(\theta_0)\cos(\Omega t + \psi_0)/I_1 \\ \cos(\theta_0)/I_3 \end{bmatrix} = \mathbf{L} \cdot \begin{bmatrix} 1/I_1 \\ 1/I_1 \\ 1/I_3 \end{bmatrix}.$$

With this, we can work backwards by using the angular velocity as a function of the time-derivatives of the Euler angles. For ϕ, we have

$$\dot{\phi} = \frac{1}{\sin(\theta)}\left[\sin(\theta)\sin^2(\psi)\frac{L}{I_1} + \sin(\theta)\cos^2(\psi)\frac{L}{I_1}\right] = \frac{L}{I_1}.$$

You can work through the other angles to eventually find that

$$\boxed{\dot{\phi} = \frac{L}{I_1}, \quad \dot{\theta} = 0, \quad \dot{\psi} = L\cos(\theta)\left[\frac{1}{I_3} - \frac{1}{I_1}\right].}$$

As the are all constant terms, the angles as a function of time is simply

$$\boxed{\theta(t) = \theta_0, \quad \phi(t) = \frac{L}{I_1}t + \phi_0, \quad \psi(t) = L\cos(\theta_0)\left[\frac{1}{I_3} - \frac{1}{I_1}\right]t + \psi_0.}$$

With this, we are now able to determine the angular velocity in the fixed frame, which is given by

$$\vec{\omega}' = \begin{bmatrix}\omega_1' \\ \omega_2' \\ \omega_3'\end{bmatrix} = \begin{bmatrix}\dot{\psi}\sin(\phi)\sin(\theta) + \dot{\theta}\cos(\phi) \\ -\dot{\psi}\cos(\phi)\sin(\theta) + \dot{\theta}\sin(\phi) \\ \dot{\psi}\cos(\theta) + \dot{\phi}\end{bmatrix}.$$

Applying the previous solutions, we thus arrive at

$$\vec{\omega}'(t) = L\begin{bmatrix}\cos(\theta_0)\sin(\theta_0)[1/I_3 - 1/I_1]\sin(Lt/I_1 + \phi_0) \\ -\cos(\theta_0)\sin(\theta_0)[1/I_3 - 1/I_1]\cos(Lt/I_1 + \phi_0) \\ \sin^2(\theta_0)/I_1 + \cos^2(\theta_0)/I_3\end{bmatrix}.$$

Here, $L/I_1 = \Omega'$ becomes the precession frequency of the angular velocity in the fixed frame about the $\hat{x}_3$ axis. Furthermore, you may notice that ω_3' is constant, as was the case in the non-inertial frame.

In fact, this result reveals that the rigid body *must* spin its body around the $\hat{x}_3$ axis if it isn't rotating about a principal axis.

7.8 Torque-Free Asymmetric Top

Consider now the case of a rigid body where $I_1 > I_2 > I_3$. The rigid body of this condition is referred to as an **asymmetric top**, and we begin our discussion of this body by analyzing its linear stability. Let us assume that the start the system rotating about an axis close to the direction $\hat{x}_1$ in the body frame. We can approximate the angular velocity in this case to be

$$\vec{\omega} \approx \begin{bmatrix}\omega_1 \\ \lambda \\ \mu\end{bmatrix}$$

where we introduce the small parameters, λ and μ, such that $\lambda, \mu \ll 1$. Here, we can write Euler's equations in the form

$$\begin{aligned}\dot{\omega}_1 &= \lambda\mu\frac{I_2 - I_3}{I_1} \approx 0 \\ \dot{\omega}_2 &= \dot{\lambda} = \omega_1\mu\frac{I_3 - I_1}{I_2} \\ \dot{\omega}_3 &= \dot{\mu} = \omega_1\lambda\frac{I_1 - I_2}{I_3}\end{aligned}$$

where we can approximate the first equation to be zero since $\lambda\mu$ are both small, and thus its product is sufficiently close enough to zero. We can combine, on the other hand, the last two equations by taking the derivative of, say, the first equation,

$$\ddot{\lambda} = \omega_1\frac{I_3 - I_1}{I_2}\dot{\mu} = \omega_1^2\left[\frac{I_3 - I_1}{I_2}\right]\left[\frac{I_1 - I_2}{I_3}\right]\lambda.$$

We make the important observation about what we assumed on our system initially. We know that $I_1 > I_3$, so we naturally expect that quantity

$$\frac{I_3 - I_1}{I_2} < 0.$$

On the other hand, we know that $I_1 > I_2$, so we also expect that

$$\frac{I_1 - I_2}{I_3} > 0.$$

In this way, we can see that the quantity $\ddot{\lambda}$ depends on a constant term which multiplies a negative and a positive. That is,

$$\left[\frac{I_3 - I_1}{I_2}\right]\left[\frac{I_1 - I_2}{I_3}\right] < 0 \quad \text{since} \;\; (I_3 - I_1) < 0 \wedge (I_1 - I_2) > 0.$$

If we refine this constant to be κ, we can thus simplify our second order equation in λ to be

$$\ddot{\lambda} - \kappa\lambda$$

which is, in fact, in the form of SHM. That is, if we were to start our system near that $\hat{x}_1$ axis, our rigid body would exhibit stable rotations. Notice, that we do *not* know if the object precesses or similar. We have simply determined if the object tends to evolve in a stable manner.

Now let us suppose that we begin the system rotating near that axis $\hat{x}_2$. In this case, we can approximate our angular velocity to be

$$\vec{\omega} = \begin{bmatrix} \lambda \\ \omega_2 \\ \mu \end{bmatrix}$$

to give Euler's equations

$$\begin{aligned} \dot{\lambda} &= \omega_2 \frac{I_2 - I_3}{I_1} \mu \\ \dot{\omega}_2 &= \lambda\mu \frac{I_3 - I_1}{2} \approx 0 \\ \dot{\mu} &= \omega_2 \frac{I_1 - I_2}{I_3} \lambda. \end{aligned}$$

Applying the same steps as before, we thus find that

$$\ddot{\lambda} = \omega_2 \frac{I_2 - I_3}{I_1} \dot{\mu} = \omega_2^2 \left[\frac{I_2 - I_3}{I_1}\right]\left[\frac{I_1 - I_2}{I_3}\right] \lambda.$$

Notice here, that the quantities $(I_2 - I_3) > 0$ and $(I_1 - I_2) > 0$. We thus except

$$\left[\frac{I_2 - I_3}{I_1}\right]\left[\frac{I_1 - I_2}{I_3}\right] > 0$$

and so our solution takes on the form

$$\ddot{\lambda} = \kappa\lambda$$

which is *not* SHM. In this case, if we were to start our system near that $\hat{x}_2$ axis (which we call the **intermediate axis**), then we expect small deviations about our motion to grow exponentially. The motion is thus unstable.

We can determine the trajectories according to the angular momentum by Euler's equations by assuming a range of initial conditions. However, if we were to solve Euler's equation explicitly, we will find that they don't have a nice closed form; that is, $I_1 > I_2 > I_3$ make the analytic calculations difficult. For this reason, we instead opt for numerical calculations and assume the parameters

$$I_1 = 1, \quad I_2 = 0.7, \quad I_3 = 0.5.$$

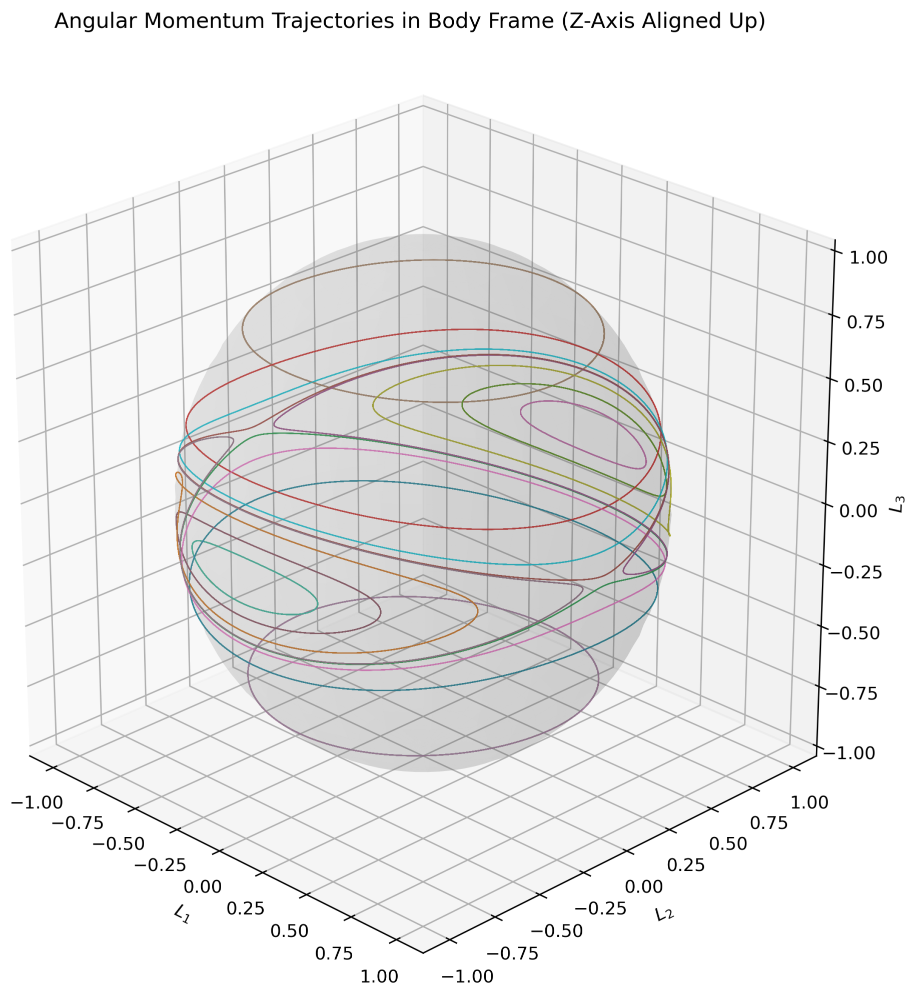


Figure 65: Varying trajectories of the asymmetric top given by the components of $\mathbf{L}$ in the body frame.

Notice that we no longer see simple lines of latitude coming from the trajectories of our sphere. In fact, if we look at the trajectories of $\mathbf{L}$ where $\vec{\omega}$ is near a principal axis, we can see that for $\hat{x}_2$, that is, the axis of L_2, the motion tends to be stable. On the other hand, we have unstable motion if we start the rigid body away at $\hat{x}_1$, where it tends to move as far as $-\hat{x}_1$ and back. Similarly, motion near $\hat{x}_3$ tends to be stable, and is in fact stable for a large range of initial conditions.

Let us now assume that the initial angular momentum is given by $\mathbf{L}' = L\hat{x}_3'$. In the body frame, we can determine this to be

$$\mathbf{L} = \mathbf{R}\mathbf{L}' = \begin{bmatrix} \sin(\theta)\cos(\psi)L \\ \sin(\theta)\cos(\psi)L \\ \cos(\theta)L \end{bmatrix}$$

and so we arrive back at

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = \begin{bmatrix} \sin(\theta)\sin(\psi)L/I_1 \\ \sin(\theta)\cos(\psi)L/I_2 \\ \cos(\theta)L/I_3 \end{bmatrix}.$$

If we relate the angular velocity back to the Euler angles, we find that

$$\begin{aligned} \dot{\phi} &= L\left[\frac{cos[2](\psi)}{I_2} + \frac{\sin^2(\psi)}{I_1}\right] \\ \dot{\theta} &= -L\left[\frac{1}{I_2} - \frac{1}{I_1}\right]\sin(\theta)\cos(\psi)\sin(\psi) \\ \dot{\psi} &= L\cos(\theta)\left[\frac{1}{I_3} - \frac{\cos^2(\psi)}{I_2} - \frac{\sin^2(\psi)}{I_1}\right] \end{aligned}$$

which does not turn out to be as nice as we had in the symmetric top for ψ.

The solutions to these equations are, as we would expect, consistent with our initial conditions. Furthermore, we can find the orientation of the rigid body by rotating the coordinates from the body frame to the fixed frame using the inverse (transposed) rotation matrix, $\mathbf{R}^T$. However, this can be handled numerically, and will not be discussed here.

8 Concluding Remarks

We have discussed a large quantity of topics in the formulation of Lagrangian, Hamiltonian, and non-inertial dynamics. In a way, we have nicely wrapped up a large majority of the classical mechanics exposed in current literature.

To that end, I would like to thank Dr. Kevin Resch for his lectures during the Spring 2025 term. The class was very enjoyable. We're done now, probably. Have fun.

Kevin please hire me.

References

[1] Resch, K. (2025). *PHYS 363: Intermediate Classical Mechanics*. University of Waterloo.

[2] Marion, J., Thornton, S. (1965). *Classical Dynamics of Particles and Systems: Fifth Edition*. Thomson.