6 Variational Principles

So far, we have discussed a variety of clever ways to solve differential equations, but have given less attention to where these differential equations come from. In this chapter we will look at a very powerful general approach to finding governing equations for a broad class of systems: variational principles. These replace local rules with global constraints which can be much easier to understand, and which can then be used to derive the local equations. The foundations of many disciplines can be written either in local or global forms; historically this has been the subject of intense religious debates (figuratively and literally).

6.1 VARIATIONAL CALCULUS

6.1.1 Euler's Equation

A variational principle is one that states a problem in terms of an unknown function that makes an integral take on an extremum (a maximum or a minimum; frequently a problem is constrained so that it is not necessary to distinguish between these). For example, let's say that we seek the function $y(x)$ that minimizes the distance between two points in a plane (not a very ambitious problem). If ds is an element of the path length along this curve, the total length of the curve is

$$\mathcal{I} = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{dx^2 + dy^2} = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \tag{6.1}$$

(the integral is broken into segments if the curve is not single-valued). More generally (in 2D), an integral defining a variational constraint may be written as

$$\mathcal{I} = \int_{x_1}^{x_2} f[y(x), \dot{y}(x), x]\, dx \tag{6.2}$$

($\dot{y} = dy/dx$). We are writing y and $\dot{y}$ independently because the function may involve either or both; they will soon become related through a differential equation. To find the solution to this problem, let's start as we often do by assuming that we've already solved it. Let $y(x)$ be that solution, and let

$$y(x, \alpha) = y(x) + \alpha\eta(x) \tag{6.3}$$

represent the solution with an arbitrary curve η added that has the property $\eta(x_1) =$

$\eta(x_2) = 0$ (Figure 6.1). Therefore, $y(x, 0) = y(x)$ and by definition I is extremal for $\alpha = 0$:

$$\frac{d\mathcal{I}}{d\alpha} = 0 \quad . \tag{6.4}$$

This is sometimes written $\delta I = 0$. To solve this equation, first differentiate under the integral and apply the chain rule:

$$\frac{d\mathcal{I}}{d\alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha} \right] dx = 0 \quad . \tag{6.5}$$

The second term can be integrated by parts:

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha} = \int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{d}{dx}\frac{\partial y}{\partial \alpha}\, dx = \underbrace{\frac{\partial f}{\partial \dot{y}}\frac{\partial y}{\partial \alpha}\bigg|_{x_1}^{x_2}}_{0} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) \underbrace{\frac{\partial y}{\partial \alpha}}_{\eta(x)}\, dx \quad . \tag{6.6}$$

Therefore

$$\frac{d\mathcal{I}}{d\alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) \right] \eta(x)\, dx = 0 \quad . \tag{6.7}$$

Since this must hold for all choices of η, the expression in the square brackets must vanish:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0 \quad . \tag{6.8}$$

This is *Euler's equation* for a variational extremum. Repeating this derivation for a function that depends on D variables $f(y_1(x), y_2(x), \ldots, y_D(x),\ \dot{y}_1(x), \dot{y}_2(x), \ldots, \dot{y}_D(x), x)$, Euler's equation becomes a set of D equations

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}_i}\right) = 0 \quad . \tag{6.9}$$

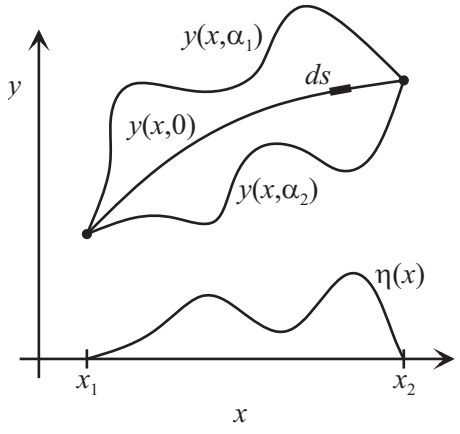

Figure 6.1. Variational paths.

This is a common, correct, and confusing notation. The partial derivatives are taken with respect to the symbols y and $\dot{y}$ as if they are unrelated, then their usual relationship

is restored to solve the resulting differential equation. This procedure keeps track of the chain rule for how they appear in the function, and is straightforward to perform in practice. But the notation is strained because it is really expressing an algorithm rather than a functional relationship. An intriguing alternative made possible by computer symbolic manipulation is to give up on conventional mathematical notation entirely and do variational problems entirely in an algorithmic language [Abelson *et al*., 1997].

Returning to equation (6.1), we can now see that

$$f = \sqrt{1+\dot{y}^2} \Rightarrow \frac{\partial f}{\partial y} = 0, \ \frac{\partial f}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{1+\dot{y}^2}} \tag{6.10}$$

and so Euler's equation is

$$\frac{d}{dx}\left(\frac{\dot{y}}{\sqrt{1+\dot{y}^2}}\right) = 0 \quad . \tag{6.11}$$

Integrating,

$$\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = C \Rightarrow \dot{y}^2 = \frac{C^2}{1-C^2} \tag{6.12}$$

for any C. The slope is constant; we have just made the significant discovery that the shortest distance between two points is a straight line.

6.1.2 Integrals and Missing Variables

If some of the variables are missing in the function f, first integrals of Euler's equation exist and can be used to help find solutions. This is an example of the deep connection between symmetry and conserved quantities, to be developed in Section 6.2.3.

If f does not depend on y, then

$$\frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0 \ \Rightarrow \ \frac{\partial f}{\partial \dot{y}} = C \tag{6.13}$$

where C is an integration constant determined by the boundary conditions. Similarly, if f does not depend on x, Euler's equations becomes a first-order ordinary differential equation

$$f - \dot{y}\frac{\partial f}{\partial \dot{y}} = C \quad . \tag{6.14}$$

To see this, differentiate with respect to x:

$$\begin{aligned}\frac{d}{dx}\left(f - \dot{y}\frac{\partial f}{\partial \dot{y}}\right) &= \frac{\partial f}{\partial y}\dot{y} + \frac{\partial f}{\partial \dot{y}}\ddot{y} - \frac{\partial f}{\partial \dot{y}}\ddot{y} - \frac{\partial^2 f}{\partial y \partial \dot{y}}\dot{y}^2 - \frac{\partial^2 f}{\partial \dot{y}^2}\dot{y}\ddot{y} \\ &= \dot{y}\left(\frac{\partial f}{\partial y} - \frac{\partial^2 f}{\partial y \partial \dot{y}}\dot{y} - \frac{\partial^2 f}{\partial \dot{y}^2}\ddot{y}\right) \quad .\end{aligned} \tag{6.15}$$

Equation (6.14) then follows from this because the right hand side vanishes, since for the case $\partial f/\partial x = 0$ Euler's equation becomes

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) = \frac{\partial f}{\partial y} - \frac{\partial^2 f}{\partial y \partial \dot{y}}\dot{y} - \frac{\partial^2 f}{\partial \dot{y}^2}\ddot{y} - \underbrace{\frac{\partial^2 f}{\partial x \partial y}}_{0} = 0 \quad . \tag{6.16}$$

Finally, if f does not depend on $\dot{y}$ then Euler's equation reduces to the algebraic equation

$$f(x, y) = C \quad . \tag{6.17}$$

The integrals frequently involve a path length $\sqrt{1 + \dot{y}^2}$ in the argument, which can lead to *hyperbolic functions* in the solution. These are defined by analogy with trigonometric functions

$$\begin{aligned} \cos(x) &= \frac{e^{ix} + e^{-ix}}{2} & \sin(x) &= \frac{e^{ix} - e^{-ix}}{2i} \\ \cosh(x) &\equiv \frac{e^{x} + e^{-x}}{2} & \sinh(x) &\equiv \frac{e^{x} - e^{-x}}{2} \end{aligned} \tag{6.18}$$

(pronounced "cosh" and "sinch"). They arise because the derivatives of their inverses have the simple form

$$\frac{d}{dx} \cosh^{-1}\left(\frac{x}{a}\right) = \frac{a}{\sqrt{x^2 - a^2}} \qquad \frac{d}{dx} \sinh^{-1}\left(\frac{x}{a}\right) = \frac{a}{\sqrt{x^2 + a^2}} \quad .$$

6.1.3 Constraints and Lagrange Multipliers

Often, a variational problem $\delta \int_1^2 f \; dx = 0$ comes with an ancillary integral constraint

$$\int_{x_1}^{x_2} g[y(x), \dot{y}(x), x] \; dx = C \tag{6.19}$$

for some constant C. For example, the problem may be to find a minimal energy curve or surface with a given length or area (Problem 4.1). This is handled by recognizing that if f solves the variational problem and g satisfies the constraint equation, then $h = f + \lambda g$ will also satisfy Euler's equation for any λ. This is because

$$\int_{x_1}^{x_2} h \; dx = \int_{x_1}^{x_2} (f + \lambda g) \; dx = \int_{x_1}^{x_2} f \; dx + \lambda \int_{x_1}^{x_2} g \; dx = \int_{x_1}^{x_2} f \; dx + \lambda C \tag{6.20}$$

and so if f is extremal then h will also be (the other term is a constant). Solving Euler's equation for $f + \lambda g$ introduces the new variable, λ, called a *Lagrange multiplier*, into the solution. It parameterizes a family of solutions all of which are extremal with a value of the constraint integral that depends on the choice of λ, which is then determined by the boundary conditions on the problem. If there is more than one constraint,

$$\begin{aligned} \int_{x_1}^{x_2} g_1[y(x), \dot{y}(x), x] \; dx &= C_1 \\ \int_{x_1}^{x_2} g_2[y(x), \dot{y}(x), x] \; dx &= C_2 \quad , \end{aligned} \tag{6.21}$$

then there will be a Lagrange multiplier for each constraint equation $h = f + \lambda_1 g_1 + \lambda_2 g_2$.

The logic of using a Lagrange multiplier is hard to follow when first encountered, but straightforward to implement and is an extremely important trick. For example, statistical mechanics is derived by finding the population distribution that maximizes the entropy of a system with a set of constraints; the Lagrange multiplier associated with a fixed average energy gives the temperature of the system, and the one associated with a fixed average number of particles gives the chemical potential of a particle.

6.2 VARIATIONAL PROBLEMS

6.2.1 Optics: Fermat's Principle

Many physical laws can be derived from variational principles; depending on one's taste this may be viewed as simply convenient or deeply significant. An important example is provided by optics. The velocity v of a photon in a medium is related to its velocity in a vacuum c by the *index of refraction* $n = c/v$. *Fermat's principle* states that a light ray will choose the path that minimizes the total transmit time. This means that the following variation vanishes between the start and the end of the path:

$$\delta \int_1^2 n \, ds = \delta c \int_1^2 \frac{ds}{v} = \delta c \int_1^2 dt = 0 \quad , \tag{6.22}$$

and so $f = n \, ds$ satisfies Euler's equation. In 2D this leads to the equation

$$\begin{aligned} 0 &= \frac{\partial}{\partial y}\left(n(y)\sqrt{1+\dot{y}^2}\right) - \frac{d}{dx}\left[\frac{\partial}{\partial \dot{y}}\left(n(y)\sqrt{1+\dot{y}^2}\right)\right] \\ &= \frac{\partial n(y)}{\partial y}\sqrt{1+\dot{y}^2} - \frac{d}{dx}\frac{n(y)\,\dot{y}}{\sqrt{1+\dot{y}^2}} \quad . \end{aligned} \tag{6.23}$$

Problem 4.2 solves this at the interface between two mediums.

6.2.2 Analytical Mechanics: Hamilton's Principle

For a classical conservative system (one in which all the forces can be derived from potentials, which means that there is no dissipation), the *Lagrangian* is equal to the difference between the kinetic energy U and the potential energy V

$$\mathcal{L}(q_1(t), \dot{q}_1(t), \ldots, t) = U - V \quad . \tag{6.24}$$

The *generalized coordinates* $q_i(t)$ are degrees of freedom that can be varied independently and that are chosen by their convenience for specifying the potential and kinetic energy. The great virtue of a Lagrangian formulation of a problem is that by using generalized coordinates it is possible to avoid explicitly writing the constraint forces that appear in a conventional coordinate system.

The integral over a path of the Lagrangian is called the *action*

$$\mathcal{I} = \int_1^2 \mathcal{L}(q_1(t), \dot{q}_1(t), \ldots, t) \, dt \quad . \tag{6.25}$$

According to Hamilton's variational principle, the system's trajectory in its configuration space will be the one that makes the action extremal ($\delta I = 0$). Therefore, the Lagrangian satisfies Euler's equations:

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = 0 \quad . \tag{6.26}$$

In this context these are called the Euler–Lagrange equations. They allow the governing equations to be found directly from knowledge of the energy in a system. As a simple

example, take a particle in a quadratic potential so that the Lagrangian is

$$\mathcal{L} = \frac{1}{2}m\dot{y}^2 - \frac{1}{2}ky^2$$

$$\frac{\partial \mathcal{L}}{\partial y} = -ky$$

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} = \frac{d}{dt}m\dot{y} = m\ddot{y}$$

$$m\ddot{y} + ky = 0 \quad . \tag{6.27}$$

This of course is just a simple harmonic oscillator; Problem 4.3 shows the use of a Lagrangian for a nontrivial example.

6.2.3 Symmetry: Noether's Theorem

Symmetries of systems are coordinate transformations that leave the governing equations unchanged; in this section we will see that for each symmetry there is a conserved quantity (an *integral of the motion*) that is invariant along the system's trajectory in its configuration space. These invariants are so valuable because each one decreases by 1 the effective dimensionality of the system of differential equations that must be solved.

Let f_s be a coordinate transformation operator that is parameterized by s (for example, a shift f_s that replaces q with $q + s$) and that does not change the Lagrangian, with $f_{s=0}$ equal to the identity transformation that makes no change to the system. If $\vec{q}(t) = (q_1(t), q_2(t), \ldots)$ is a solution of the Euler–Lagrange equations then by assumption $\vec{q}(s,t) = f_s[\vec{q}(t)]$ is also a solution:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}[\vec{q}(s,t), \dot{\vec{q}}(s,t)] = \frac{\partial \mathcal{L}}{\partial q_i}[\vec{q}(s,t), \dot{\vec{q}}(s,t)] \quad . \tag{6.28}$$

Also, since the Lagrangian is invariant under the transformation

$$0 = \frac{d}{ds}\mathcal{L}[\vec{q}(s,t), \dot{\vec{q}}(s,t)] = \sum_i \left[\frac{\partial \mathcal{L}}{\partial q_i}\frac{dq_i(s,t)}{ds} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\frac{d\dot{q}_i(s,t)}{ds}\right] \quad . \tag{6.29}$$

Combining these two equations gives

$$\begin{aligned} 0 &= \sum_i \left[\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)\frac{dq_i}{ds} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\frac{d}{dt}\left(\frac{dq_i}{ds}\right)\right] \\ &\equiv \frac{d\mathcal{I}}{dt} \quad , \end{aligned} \tag{6.30}$$

where

$$\mathcal{I} = \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\frac{d}{ds}f_s(q_i) \quad . \tag{6.31}$$

This is called *Noether's theorem* (after the mathematician Emmy Noether). It is true for all s, but in particular it is true at $s = 0$. Therefore, the integral invariants can be found by considering an infinitesimal neighborhood of the identity transformation.

The simplest example is translational invariance $f_s(q_i) = q_i + s$. The corresponding integral is

$$\mathcal{I} = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d}{ds}(q_i + s) = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad . \tag{6.32}$$

If $\dot{q}_i$ enters into $\mathcal{L}$ as $m\dot{q}_i^2/2$, then this integral is just the momentum $m\dot{q}_i$ (which could be linear or angular depending on the generalized coordinate q_i). Conservation of momentum is really a statement of the translational invariance of a system.

6.3 RIGID BODY MOTION

An important application of Lagrange's equations is to the motion of *rigid bodies*, such as a spinning top. Developing this connection will require a small detour to describe the configuration of a rigid body.

Let the distribution of mass be $\rho(\vec{x})$, and let the origin for $\vec{x}$ be taken to be the center of mass so that

$$\int \rho(\vec{x})\vec{x}\, d\vec{x} = 0 \quad . \tag{6.33}$$

This discussion is for a continuous mass distribution, but it also applies to a discrete distribution if the integrals are replaced with sums.

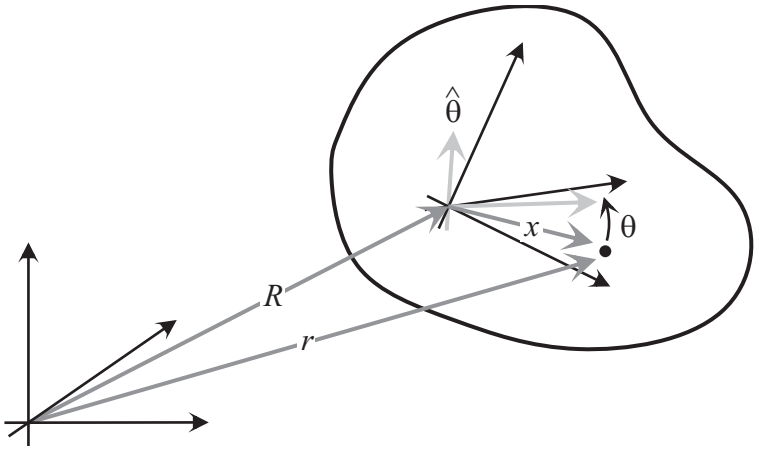

Figure 6.2. Axes for the motion of a rigid body.

Consider the motion of a point $\vec{x}$ in the rigid body. Its position can also be measured from a fixed external coordinate system ($\vec{r}$), and in this frame let the position of the center of mass be $\vec{R}$. In addition to translation the body can also rotate; let the rotation be by an angle θ around an axis $\hat{\theta}$ (Figure 6.2). If both the position of the center of mass and the angle change, the position of the point as seen in the fixed frame changes by

$$d\vec{r} = d\vec{R} + d\hat{\theta} \times \vec{x} \quad . \tag{6.34}$$

Dividing both sides by dt shows that the velocity of the point is the sum of the velocity

of the center of mass of the rigid body and a term that depends on the rotation:

$$\vec{v} = \vec{V} + \vec{\omega} \times \vec{x} \tag{6.35}$$

($\vec{\omega} \equiv d\hat{\theta}/dt$). Therefore in a rotating frame the time derivative operator gains an extra cross-product term

$$\left(\frac{d\vec{f}}{dt}\right)_{\text{rotating}} \mapsto \left(\frac{d\vec{f}}{dt}\right)_{\text{fixed}} + \omega \times \vec{f} \quad . \tag{6.36}$$

Now let's calculate the kinetic energy of the body so that we can find the Lagrangian:

$$\begin{aligned} U &= \frac{1}{2}\int \rho(\vec{x})|\vec{v}(\vec{x})|^2 \, d\vec{x} \\ &= \frac{1}{2}\int \rho(\vec{x})|\vec{V} + \vec{\omega} \times \vec{x}|^2 \, d\vec{x} \\ &= \frac{1}{2}|\vec{V}|^2 \int \rho(\vec{x}) \, d\vec{x} + \vec{V} \cdot \int \rho(\vec{x}) \, \vec{\omega} \times \vec{x} \, d\vec{x} + \frac{1}{2}\int \rho(\vec{x})|\vec{\omega} \times \vec{x}|^2 \, d\vec{x} \\ &= \frac{1}{2}M|\vec{V}|^2 + (\vec{V} \times \vec{\omega}) \cdot \underbrace{\int \rho(\vec{x})\vec{x} \, d\vec{x}}_{0} + \frac{1}{2}\int \rho(\vec{x})|\vec{\omega} \times \vec{x}|^2 \, d\vec{x} \end{aligned} \tag{6.37}$$

(the last line uses the identity $\vec{A} \cdot (\vec{B} \times \vec{C}) = (\vec{A} \times \vec{B}) \cdot \vec{C}$). The kinetic energy is the sum of two terms. The first one just depends on the translational motion of the center of mass, and the second depends only on the internal rotation. The second term can be simplified by writing it in terms of a dot product. Letting the angle between $\vec{\omega}$ and $\vec{x}$ be α,

$$\begin{aligned} |\vec{\omega} \times \vec{x}|^2 &= |\vec{\omega}|^2|\vec{x}|^2 \sin^2(\alpha) \\ &= |\vec{\omega}|^2|\vec{x}^2|(1 - \cos^2(\alpha)) \\ &= |\vec{\omega}^2||\vec{x}|^2 - (\vec{\omega} \cdot \vec{x})^2 \\ &= \sum_i \sum_j \omega_i(|\vec{x}|^2\delta_{ij} - x_i x_j)\omega_j \quad . \end{aligned} \tag{6.38}$$

Plugging this back into equation (6.37),

$$\begin{aligned} U &= \frac{1}{2}M|\vec{V}|^2 + \frac{1}{2}\sum_i \sum_j \omega_i \underbrace{\int \rho(\vec{x})[|\vec{x}|^2\delta_{ij} - x_i x_j] \, d\vec{x}}_{\equiv \mathbf{J}_{ij}} \omega_j \\ &= \frac{1}{2}M|\vec{V}|^2 + \frac{1}{2}\vec{\omega}^T \cdot \mathbf{J} \cdot \vec{\omega} \end{aligned} \tag{6.39}$$

(where $\vec{\omega}^T$ is the transpose of the vector $\vec{\omega}$). $\mathbf{J}$ is called the *inertia tensor*. If the coordinate system in the rigid body is transformed to diagonalize the inertia tensor (as we did for normal modes), the eigenvalues (I_1, I_2, I_3) are called the *moments of inertia* around the eigenvectors called the *principal axes*, and the kinetic energy reduces to

$$U = \frac{1}{2}M|\vec{V}|^2 + \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2 \quad . \tag{6.40}$$

The inertia tensor is diagonal with respect to the principal axes, but it is not diagonal

when viewed from a fixed external frame. Therefore, to calculate the kinetic energy it will be convenient to transform from the fixed frame into the moving principal axes frame. This transformation can be parameterized by taking for generalized coordinates the *Euler angles* which can be varied independently (Figure 6.3). In the standard definition used for rigid body motion, there is a rotation of φ around the third axis, a rotation of θ around the new position of the first axis, and then a rotation of ψ around the new position of the third axis. As the body moves, these angles are measured between the principal axes and a coordinate system that moves with the center of mass but remains aligned with the external frame.

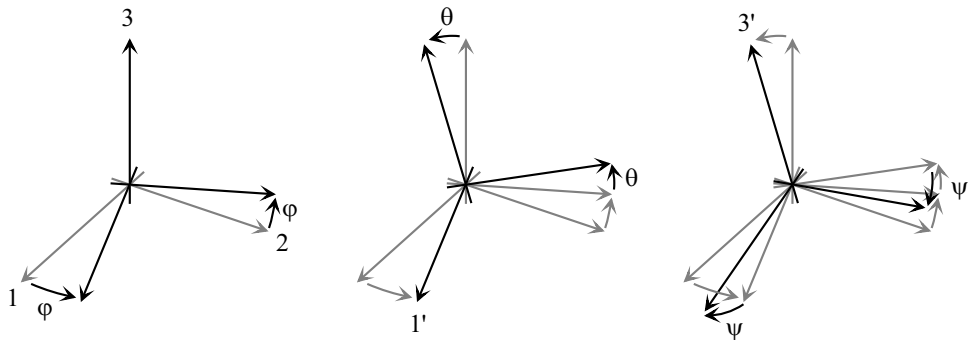

Figure 6.3. Definition of Euler angles.

The angular velocities around the body's principal axes can be related to the derivatives of the Euler angles through simple trigonometry in Figure 6.3:

$$\begin{aligned}\omega_1 &= \dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi \\ \omega_2 &= -\dot{\theta}\sin\psi + \dot{\varphi}\sin\theta\cos\psi \\ \omega_3 &= \dot{\varphi}\cos\theta + \dot{\psi} \quad .\end{aligned} \tag{6.41}$$

Therefore, the rotational kinetic energy is

$$\begin{aligned}U_{rot} &= \frac{1}{2}\sum I_i\omega_i^2 \\ &= \frac{1}{2}I_1(\dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi)^2 + \frac{1}{2}I_2(-\dot{\theta}\sin\psi + \dot{\varphi}\sin\theta\cos\psi)^2 \\ &\quad + \frac{1}{2}I_3(\dot{\varphi}\cos\theta + \dot{\psi})^2 \quad .\end{aligned} \tag{6.42}$$

Applying Lagrange's equation to this gives the *Euler equations* for rigid body motion. Ignoring the translational energy, and assuming there are no external forces, these take a particularly simple form:

$$\begin{aligned}I_1\dot{\omega}_1 &= (I_2 - I_3)\omega_2\omega_3 \\ I_2\dot{\omega}_2 &= (I_3 - I_1)\omega_3\omega_1 \\ I_3\dot{\omega}_3 &= (I_1 - I_2)\omega_1\omega_2 \quad .\end{aligned} \tag{6.43}$$

If the principal axes are ordered so that $I_3 > I_2 > I_1$, the middle equation has a positive coefficient while the first and third have negative coefficients. Small perturbations about

the middle axis will grow exponentially. To see this experimentally, take an object such as an eraser or a book and try to throw it up while spinning it around each of the principal axes in turn.

6.4 SELECTED REFERENCES

[Goldstein, 1980] Goldstein, Herbert (1980). *Classical Mechanics*. 2nd edn. Reading, MA: Addison-Wesley.

[Scheck, 1990] Scheck, Florian (1990). *Mechanics : From Newton's Laws to Deterministic Chaos*. New York, NY: Springer-Verlag.

The development of variational methods is closely tied to classical mechanics. Goldstein's book is a classic text, and Scheck provides a nice treatment of modern techniques.

6.5 PROBLEMS

(4.1) Consider a chain of length L and density ρ that is hanging between two posts. Find the general form for the shape that minimizes the potential energy. Remember that the potential energy of a segment ds is $\rho\ ds\ g\ y$.

(4.2) Consider a light ray travelling in a medium that has index of refraction n_1 for $x < 0$ and n_2 for $x > 0$. As it crosses the line $x = 0$ its angle with respect to the x axis will change. Solve Euler's equation to find the simple relationship between the two angles.

(4.3) Consider a pendulum with a mass m connected by a massless rod of length l to a moveable platform at height z (Figure 6.4). Write down the Lagrangian for this system, and find a differential equation for the angle θ as z is varied. What is its form for small oscillations?

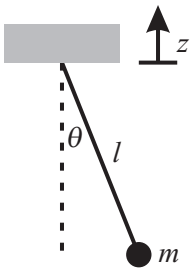

Figure 6.4. A pendulum attached to a moving platform.