PHYS3001
Classical Mechanics

Robert L. Dewar
Department of Theoretical Physics
Research School of Physical Sciences & Engineering
The Australian National University
Canberra ACT 0200
Australia

robert.dewar@anu.edu.au

Version 1.51

Contents

1 Generalized Kinematics **1**

1.1 Introduction 1
1.2 Generalized coordinates 2
1.3 Example: The ideal fluid 5
1.4 Variational Calculus 6
1.4.1 Example: Geodesics 9
1.4.2 Trial function method 10
1.5 Constrained variation: Lagrange multipliers 10
1.6 Problems 13
1.6.1 Rigid rod 13
1.6.2 Ecliptic 13
1.6.3 Curvature of geodesics 13

2 Lagrangian Mechanics **15**

2.1 Introduction 15
2.2 Generalized Newton's 2nd Law 16
2.2.1 Generalized force 16
2.2.2 Generalized equation of motion 20
2.2.3 Example: Motion in Cartesian coordinates 21
2.3 Lagrange's equations (scalar potential case) 21
2.3.1 Hamilton's Principle 23
2.4 Lagrangians for some Physical Systems 24
2.4.1 Example 1: 1-D motion—the pendulum 24
2.4.2 Example 2: 2-D motion in a central potential 25
2.4.3 Example 3: 2-D motion with time-varying constraint . 26
2.4.4 Example 4: Atwood's machine 27
2.4.5 Example 5: Particle in e.m. field 28
2.4.6 Example 6: Particle in ideal fluid 29
2.5 Averaged Lagrangian 30
2.5.1 Example: Harmonic oscillator 31
2.6 Transformations of the Lagrangian 32

2.6.1 Point transformations 32
2.6.2 Gauge transformations 34
2.7 Symmetries and Noether's theorem 35
2.7.1 Time symmetry 37
2.8 Problems 37
2.8.1 Coriolis force and cyclotron motion 37
2.8.2 Anharmonic oscillator 38

3 Hamiltonian Mechanics **41**
3.1 Introduction: Dynamical systems 41
3.2 Mechanics as a dynamical system 41
3.2.1 Lagrangian method 41
3.2.2 Hamiltonian method 43
3.2.3 Example 1: Scalar potential 45
3.2.4 Example 2: Physical pendulum 47
3.2.5 Example 3: Motion in e.m. potentials 48
3.2.6 Example 4: The generalized N-body system 48
3.3 Time-Dependent and Autonomous Hamiltonian systems . . . 50
3.4 Hamilton's Principle in phase space 50
3.5 Problems 53
3.5.1 Constraints and moving coordinates 53
3.5.2 Anharmonic oscillator phase space 53
3.5.3 2-D motion in a magnetic field 53

4 Canonical transformations **55**
4.1 Introduction 55
4.2 Generating functions 56
4.2.1 Example 1: Adiabatic Oscillator 60
4.2.2 Example 2: Point transformations 62
4.3 Infinitesimal canonical transformations 63
4.3.1 Time evolution 64
4.4 Poisson brackets 65
4.4.1 Symmetries and integrals of motion 66
4.4.2 Perturbation theory 66
4.5 Action-Angle Variables 67
4.6 Properties of canonical transformations 69
4.6.1 Preservation of phase-space volume 69
4.6.2 Transformation of Poisson brackets 74
4.7 Problems 74
4.7.1 Coriolis yet again 74
4.7.2 Difference approximations 75

5 Answers to Problems **77**
5.1 Chapter 1 Problems 77
5.2 Chapter 2 Problems 80
5.3 Chapter 3 Problems 86
5.4 Chapter 4 Problems 93

6 References and Index **99**

Chapter 1

Generalized coordinates and variational principles

1.1 Introduction

In elementary physics courses you were introduced to the basic ideas of Newtonian mechanics via concrete examples, such as motion of a particle in a gravitational potential, the simple harmonic oscillator etc. In this course we will develop a more abstract viewpoint in which one thinks of the dynamics of a system described by an arbitrary number of *generalized coordinates*, but in which the dynamics can be nonetheless encapsulated in a single scalar function: the *Lagrangian*, named after the French mathematician Joseph Louis Lagrange (1736–1813), or the *Hamiltonian*, named after the Irish mathematician Sir William Rowan Hamilton (1805–1865).

This abstract viewpoint is enormously powerful and underpins quantum mechanics and modern nonlinear dynamics. It may or may not be more efficient than elementary approaches for solving simple problems, but in order to feel comfortable with the formalism it is very instructive to do some elementary problems using abstract methods. Thus we will be revisiting such examples as the harmonic oscillator and the pendulum, but when examples are set in this course please remember that you are expected to use the approaches covered in the course rather than fall back on the methods you learnt in First Year.

In the following notes the convention will be used of *italicizing* the first use or definition of a concept. The index can be used to locate these definitions and the subsequent occurrences of these words.

The present chapter is essentially geometric. It is concerned with the description of possible motions of general systems rather than how to calculate

physical motions from knowledge of forces. Thus we call the topic *generalized kinematics*.

1.2 Generalized coordinates

Suppose we have a system of N particles each moving in 3-space and interacting through arbitrary (finite) forces, then the dynamics of the total system is described by the motion of a point $\boldsymbol{q} \equiv \{q_i | i = 1, 2, \ldots, 3N\} = \{x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_N, y_N, z_N\}$ in a $3N$-dimensional generalized *configuration space*. The number $n = 3N$ of *generalized coordinates* q_i is called the number of *degrees of freedom*. No particular metric is assumed—e.g. we could equally as well use spherical polar coordinates (see Fig. 1.1), $q_i = \{r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2, \ldots, r_N, \theta_N, \phi_N\}$, or a more general curvilinear coordinate system.

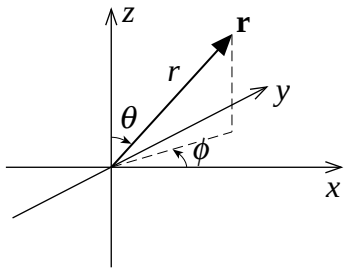

Figure 1.1: Position vector $\mathbf{r}$ in Cartesian and Spherical Polar coordinates.

In other systems the generalized "coordinates" need not even be spatial coordinates—e.g. they could be the charges flowing in an electrical circuit. Thus the convenient vector-like notation $\boldsymbol{q}$ for the array of generalized coordinates should not be confused with the notation $\mathbf{r}$ for the position vector in 3-space. Often the set of generalized coordinates is simply denoted q, but in these notes we use a bold font, distinguishing generalized coordinate arrays from 3-vectors by using a bold *slanted* font for the former and a bold *upright* font for the latter.

Vectors are entities independent of which coordinates are used to represent them, whereas the set of generalized coordinates changes if we change variables. For instance, consider the position vector of a particle in Cartesian coordinates x, y, z and in spherical polar coordinates r, θ, ϕ in Fig. 1.1. The vector $\mathbf{r}$ represents a point in physical 3-space and thus does not change when we change coordinates,

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z = r\mathbf{e}_r(\theta, \phi) \; . \tag{1.1}$$

However its *representation* changes, because of the change in the unit basis vectors from $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ to $\mathbf{e}_r$, $\mathbf{e}_\phi$ and $\mathbf{e}_z$. On the other hand, the sets of

generalized coordinates $\boldsymbol{r}_{\rm C} \equiv \{x, y, z\}$ and $\boldsymbol{r}_{\rm sph} \equiv \{r, \phi, z\}$ are distinct entities: they are points in two different (though related) configuration spaces describing the particle.

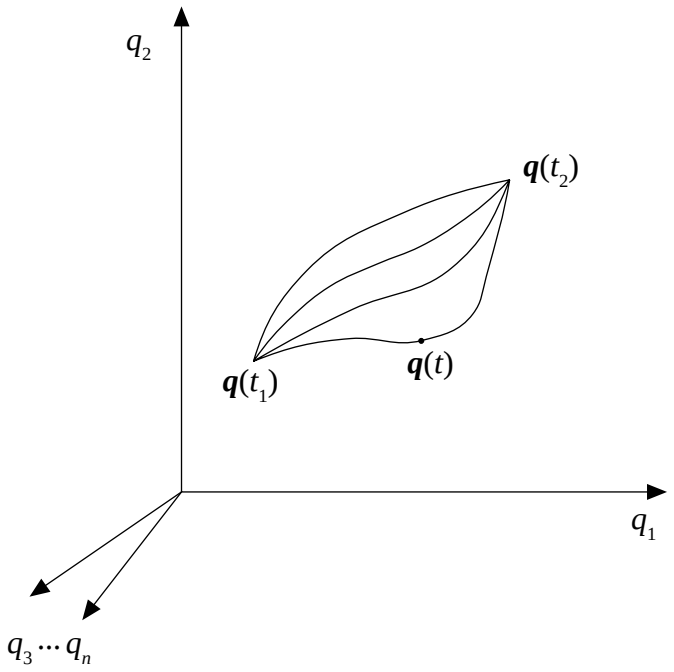

Figure 1.2: Some possible paths in configuration space, each parametrized by the time, t.

Sometimes the motion is constrained to lie within a submanifold of the full configuration space.[1] For instance, we may be interested in the motion of billiard balls constrained to move within a plane, or particles connected by rigid rods. In such cases, where there exists a set of (functionally independent) constraint equations, or *auxiliary conditions*

$$f_j(\boldsymbol{q}) = 0 \ , \quad j = 1, 2, \cdots, m < n \ , \tag{1.2}$$

the constraints are said to be *holonomic*.

Each holonomic constraint reduces the number of degrees of freedom by one, since it allows us to express one of the original generalized coordinates as

[1] A *manifold* is a mathematical space which can everywhere be described locally by a Cartesian coordinate system, even though extension to a global nonsingular curvilinear coordinate system may not be possible (as, e.g. on a sphere). A manifold can always be regarded as a surface embedded in a higher dimensional space.

a function of the others and delete it from the set. For instance, for a particle constrained to move in a horizontal plane (an idealized billiard table), the vertical position $z = \text{const}$ is a trivial function of the horizontal coordinates x and y and the configuration space becomes two dimensional, $\boldsymbol{q} = \{x, y\}$.

Consider the set of all conceivable *paths* through configuration space (see Fig. 1.2). Each one may be parametrized by the time, t: $\boldsymbol{q} = \boldsymbol{q}(t)$. By differentiating eq. (1.2) with respect to time, we find a set of constraints on the generalized velocities, $\dot{q}_i \equiv \mathrm{d}q_i/\,\mathrm{d}t$, which we write in differential notation as

$$\sum_{i=1}^{n} \frac{\partial f_j(\boldsymbol{q})}{\partial q_i}\,\mathrm{d}q_i \equiv \frac{\partial f_j(\boldsymbol{q})}{\partial \boldsymbol{q}}\cdot \mathrm{d}\boldsymbol{q} = 0\ , \tag{1.3}$$

where $\partial f/\partial \boldsymbol{q} \equiv \{\partial f/\partial q_1, \ldots, \partial f/\partial q_n\}$ and we use the shorthand dot-product notation

$$\boldsymbol{a}\cdot\frac{\partial f}{\partial \boldsymbol{b}} \equiv \sum_{i=1}^{n} a_i \frac{\partial f}{\partial b_i}\ , \tag{1.4}$$

where a_i and b_i are arbitrary configuration space variables.

The condition for functional independence of the m constraints is that there be m nontrivial solutions of eq. (1.3), i.e. that the rank of the matrix $\partial f_j(\boldsymbol{q})/\partial q_i$ be its maximal possible value, m.

Note that not *all* such differential constraints lead to holonomic constraints. If we are given constraints as a general set of differential forms

$$\sum_{i=1}^{n} \omega_i^{(j)}(\boldsymbol{q})\,\mathrm{d}q_i = 0\ , \tag{1.5}$$

then we may or may not be able to integrate the constraint equations to the form eq. (1.2). When we can, the forms are said to be *complete differentials*. When we cannot, the constraints are said to be *nonholonomic*.

The latter case, where we cannot reduce the number of degrees of freedom by the number of constraints, will not be considered explicitly in these notes. Furthermore, we shall normally assume that any holonomic constraints have been used to reduce $\{q_i | i = 1, \ldots, n\}$ to a minimal, unconstrained set. However, we present in Sec. 1.5 an elegant alternative that may be used when this reduction is not convenient, or is impossible due to the existence of nonholonomic constraints.

There are situations where there is an *infinite* number of generalized coordinates. For instance, consider a scalar *field* (such as the instaneous amplitude of a wave), $\psi(\mathbf{r}, t)$. Here ψ is a generalized coordinate of the system and the position vector $\mathbf{r}$ replaces the index i. Since $\mathbf{r}$ is a continuous variable it ranges over an infinite number of values.

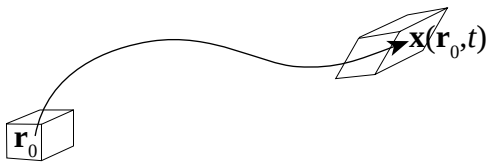

Figure 1.3: A fluid element advected from point $\mathbf{r} = \mathbf{r}_0$ at time $t = 0$ to $\mathbf{r} = \mathbf{x}(\mathbf{r}_0, t)$ at time t.

1.3 Example: The ideal fluid

As an example of a system with both an infinite number of degrees of freedom and holonomic constraints, consider a fluid with density field $\rho(\mathbf{r}, t)$, pressure field $p(\mathbf{r}, t)$ and velocity field $\mathbf{v}(\mathbf{r}, t)$.

Here we are using the *Eulerian description*, where the fluid quantities ρ, p and $\mathbf{v}$ are indexed by the actual position, $\mathbf{r}$, at which they take on their physical values at each point in time.

However, we can also index these fields by the *initial* position, $\mathbf{r}_0$, of the fluid particle passing through the point $\mathbf{r} = \mathbf{x}(\mathbf{r}_0, t)$ at time t (see Fig. 1.3). This is known as the *Lagrangian description*. (cf. the Schrödinger and Heisenberg pictures in quantum mechanics.) We shall denote fields in the Lagrangian description by use of a subscript L: $\rho_{\rm L}(\mathbf{r}_0, t)$, $p_{\rm L}(\mathbf{r}_0, t)$ and $\mathbf{v}_{\rm L}(\mathbf{r}_0, t) = \partial_t \mathbf{x}(\mathbf{r}_0, t)$.

The field $\mathbf{x}(\mathbf{r}_0, t)$ may be regarded as an infinite set of generalized coordinates, the specification of which gives the state of the fluid at time t. The *Jacobian* $J(\mathbf{r}_0, t)$ of the change of coordinates $\mathbf{r} = \mathbf{x}(\mathbf{r}_0, t)$ is defined by

$$\begin{aligned} J &\equiv \frac{\partial \mathbf{x}}{\partial x_0} \cdot \frac{\partial \mathbf{x}}{\partial y_0} \times \frac{\partial \mathbf{x}}{\partial z_0} \\ &= \begin{vmatrix} \frac{\partial x}{\partial x_0} & \frac{\partial y}{\partial x_0} & \frac{\partial z}{\partial x_0} \\ \frac{\partial x}{\partial y_0} & \frac{\partial y}{\partial y_0} & \frac{\partial z}{\partial y_0} \\ \frac{\partial x}{\partial z_0} & \frac{\partial y}{\partial z_0} & \frac{\partial z}{\partial z_0} \end{vmatrix} , \end{aligned} \tag{1.6}$$

where x_0, y_0 and z_0 are Cartesian components of $\mathbf{r}_0$ and x, y and z are the corresponding components of $\mathbf{x}(\mathbf{r}_0, t)$. This gives the change of volume of a fluid element with initial volume $\mathrm{d}V_0$ and final volume (at time t) $\mathrm{d}V$ through

$$\mathrm{d}V = J(\mathbf{r}_0, t)\, \mathrm{d}V_0 \, . \tag{1.7}$$

To see this, consider $\mathrm{d}V_0 = \mathrm{d}x_0\, \mathrm{d}y_0\, \mathrm{d}z_0$ to be an infinitesimal rectangular box, as indicated in Fig. 1.3, with sides of length $\mathrm{d}x_0$, $\mathrm{d}y_0$, $\mathrm{d}z_0$. This fluid

element is transformed by the effect of compression and shear to an infinitesimal parallelipiped with sides given by $\mathrm{d}\mathbf{l}_x \equiv \mathrm{d}x_0 \partial\mathbf{x}/\partial x_0$, $\mathrm{d}\mathbf{l}_y \equiv \mathrm{d}y_0 \partial\mathbf{x}/\partial y_0$, $\mathrm{d}\mathbf{l}_z \equiv \mathrm{d}z_0 \partial\mathbf{x}/\partial z_0$. The volume of such a parallepiped is $\mathrm{d}\mathbf{l}_x \cdot \mathrm{d}\mathbf{l}_y \times \mathrm{d}\mathbf{l}_z = J\,\mathrm{d}x_0\,\mathrm{d}y_0\,\mathrm{d}z_0$ □.

Are the fields $\rho(\mathbf{r}_0, t)$ and $p(\mathbf{r}_0, t)$ additional generalized coordinates which need to be specified at each point in time? In an ideal fluid (i.e. one with no dissipation, also called an Euler fluid) the answer is *no*, because mass conservation, $\rho\,\mathrm{d}V = \rho_0\,\mathrm{d}V_0$, allows us to write

$$\rho_{\mathrm{L}}(\mathbf{r}_0, t) = \rho_0(\mathbf{r}_0)/J(\mathbf{r}_0, t)\ , \tag{1.8}$$

while the ideal equation of state $p(\,\mathrm{d}V)^\gamma = p_0(\,\mathrm{d}V_0)^\gamma$, where γ is the ratio of specific heats, gives

$$p_{\mathrm{L}}(\mathbf{r}_0, t) = p_0(\mathbf{r}_0)/J^\gamma(\mathbf{r}_0, t)\ , \tag{1.9}$$

where ρ_0 and p_0 are the initial density and pressure fields, respectively. These are, by definition, fixed in time, so the only time dependence occurs through the Jacobian J, which we showed in eq. (1.6) to be completely determined by the Lagrangian displacement field $\mathbf{x}(\mathbf{r}_0, t)$. Thus eqs. (1.8) and (1.9) have allowed us to reduce the number of generalized coordinate fields from 5 to 3 (the three components of $\mathbf{x}$)—mass conservation and the equation of state have acted as *holonomic constraints*.

Note: Mass conservation is valid even for *nonideal* fluids (provided they are not reacting and thus changing from one state to another). However, in a fluid with finite dissipation, heat will be generated by the motion and entropy will be increased in each fluid element, thus invalidating the use of the adiabatic equation of state. Further, the entropy increase depends on the complete path of the fluid through its state space, not just on its instantaneous state. Thus the pressure cannot be holonomically constrained in a nonideal fluid.

Remark 1.1 *A useful model for a hot plasma is the magnetohydrodynamic (MHD) fluid—an ideal fluid with the additional property of being a perfect electrical conductor. This leads to the magnetic field* $\mathbf{B}(\mathbf{r}, t)$ *being "frozen in" to the plasma, so that* $\mathbf{B}_{\mathrm{L}}$ *also obeys a holonomic constraint in the Lagrangian representation, but as it is a vector constraint it is a little too complicated to give here.*

1.4 Variational Calculus

Consider an *objective functional* $I[\boldsymbol{q}]$, defined on the space of all differentiable paths between two points in configuration space, $\boldsymbol{q}(t_1)$ and $\boldsymbol{q}(t_2)$, as depicted

in Fig. 1.2

$$I[\boldsymbol{q}] \equiv \int_{t_1}^{t_2} \mathrm{d}t\, f(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t) \; . \tag{1.10}$$

(As we shall wish to integrate by parts later, we in fact assume the paths to be slightly smoother than simply differentiable, so that $\ddot{\boldsymbol{q}}$ is also defined.)

We suppose our task is to find a path that makes I a maximum or minimum (or at least stationary) with respect to neighbouring paths. Thus we *vary* the path by an amount $\delta\boldsymbol{q}(t)$: $\boldsymbol{q}(t) \mapsto \boldsymbol{q}(t) + \delta\boldsymbol{q}(t)$. Then the *first variation*, δI, is defined to be the change in I as estimated by linearizing in $\delta\boldsymbol{q}$:

$$\delta I[\boldsymbol{q}] \equiv \int_{t_1}^{t_2} \mathrm{d}t\, \left[\delta\boldsymbol{q}(t) \cdot \frac{\partial f}{\partial \boldsymbol{q}} + \delta\dot{\boldsymbol{q}}(t) \cdot \frac{\partial f}{\partial \dot{\boldsymbol{q}}} \right] \; . \tag{1.11}$$

Our first task is to evaluate eq. (1.11) in terms of $\delta\boldsymbol{q}(t)$. The crucial step here is the lemma "*delta and dot commute*". That is

$$\delta\dot{\boldsymbol{q}} \equiv \frac{\mathrm{d}\delta\boldsymbol{q}}{\mathrm{d}t} \; . \tag{1.12}$$

To prove this, simply go back to definitions: $\delta\dot{\boldsymbol{q}} \equiv \mathrm{d}(\boldsymbol{q} + \delta\boldsymbol{q})/\,\mathrm{d}t - \mathrm{d}\boldsymbol{q}/\,\mathrm{d}t = \mathrm{d}\delta\boldsymbol{q}/\,\mathrm{d}t$ □.

We can now integrate by parts to put δI in the form

$$\delta I[\boldsymbol{q}] = \left[\delta\boldsymbol{q} \cdot \frac{\partial f}{\partial \dot{\boldsymbol{q}}} \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \mathrm{d}t\, \delta\boldsymbol{q}(t) \cdot \frac{\delta f}{\delta \boldsymbol{q}} \; . \tag{1.13}$$

This consists of an endpoint contribution and an integral of the *variational derivative* $\delta f/\delta\boldsymbol{q}$, defined by

$$\frac{\delta f}{\delta \boldsymbol{q}} \equiv \frac{\partial f}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial f}{\partial \dot{\boldsymbol{q}}} \; . \tag{1.14}$$

Remark 1.2 *The right-hand side of eq. (1.14) is also sometimes called the* functional derivative *or* Fréchet derivative *of $I[\boldsymbol{q}]$. When using this terminology the notation $\delta I/\delta\boldsymbol{q}$ is used instead of $\delta f/\delta\boldsymbol{q}$ so that we can write eq. (1.13), for variations δq_i which vanish in the neighbourhood of the endpoints, as*

$$\begin{aligned} \delta I[\boldsymbol{q}] &\equiv \int_{t_1}^{t_2} \mathrm{d}t \sum_{i=1}^{n} \delta q_i(t) \frac{\delta I}{\delta q_i}(t) \; , \\ &\equiv \left(\delta\boldsymbol{q}, \frac{\delta I}{\delta \boldsymbol{q}} \right) \; , \end{aligned} \tag{1.15}$$

which may be taken as the most general defining equation for $\delta I/\delta \boldsymbol{q}$*. The* inner product *notation* $(\cdot,\cdot)$ *used above is a kind of infinite-dimensional dot product where we not only sum over the index* i*, but integrate over the "index"* t*. If we recall that the change in the value of a field defined on 3-space, e.g.* $\varphi(\mathbf{r})$*, due to an arbitrary infinitesimal change* $\delta\mathbf{r}$ *is* $\delta\varphi = \delta\mathbf{r}\cdot\boldsymbol{\nabla}\varphi$*, which may be regarded as the* definition *of the gradient* $\boldsymbol{\nabla}\varphi$*, we see that the functional derivative* $\delta I/\delta \boldsymbol{q}$ *may be thought of, by analogy, as an infinite-dimensional gradient defined on the function space of paths.*

A typical variational problem is to make I *extremal* or *stationary* under arbitrary variations $\delta\boldsymbol{q}(t)$ holding the endpoints fixed. That is, we require

$$\delta I = 0 \quad \forall \text{ functions } \delta\boldsymbol{q}(t) \text{ such that } \delta\boldsymbol{q}(t_1) = \delta\boldsymbol{q}(t_2) = 0 \; . \tag{1.16}$$

Note that this condition does not necessarily require I to be a minimum or maximum—it can be a kind of saddle point in function space, with some ascending and some descending "directions". To determine the nature of a stationary point we would need to expand I to second order in $\delta\boldsymbol{q}$—the *second variation.*

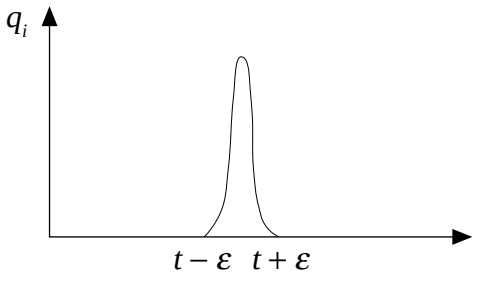

Figure 1.4: A time-localized variation in generalized coordinate q_i with support in the range $t-\epsilon$ to $t+\epsilon$.

In the class of variations in eq. (1.16), the endpoint contribution in eq. (1.13) vanishes, leaving only the contribution of the integral over t. Since $\delta\boldsymbol{q}(t)$ is arbitrary, we can, in particular, consider functions with arbitrarily localized support in t, as indicated in Fig. 1.4. (The *support* of a function is just the range over which it is nonzero.) As $\epsilon \to 0$, $\delta f/\delta\boldsymbol{q}(t)$ becomes essentially constant over the support of $\boldsymbol{q}$ in eq. (1.13) and we can move it outside the integral. Clearly then, I can only be stationary for all such variations *if and only if the variational derivative vanishes for each value of* t *and each index* i

$$\frac{\delta f}{\delta \boldsymbol{q}} = 0 \; . \tag{1.17}$$

These n equations are known as the *Euler–Lagrange equations*. Sometimes we encounter variational problems where we wish to extremize I under

variations of the endpoints as well, $\delta\boldsymbol{q}(t_1) \neq \delta\boldsymbol{q}(t_2) \neq 0$. In such cases we see from eq. (1.13) that, in addition to eq. (1.17), stationarity implies the *natural boundary conditions*

$$\frac{\partial f}{\partial \dot{\boldsymbol{q}}} = 0 \tag{1.18}$$

at t_1 and t_2.

1.4.1 Example: Geodesics

In the above development we have used the symbol t to denote the independent variable because, in applications in dynamics, paths in configuration space are naturally parametrized by the time. However, in purely geometric applications t is simply an arbitrary label for the position along a path, and we shall in this section denote it by τ to avoid confusion.

The distance along a path is given by integrating the lengths $\mathrm{d}l$ of infinitesimal line elements, given a *metric tensor* $g_{i,j}$ such that

$$(\,\mathrm{d}l)^2 = \sum_{i,j=1}^{n} \mathrm{d}q_i \, g_{i,j} \, \mathrm{d}q_j \ . \tag{1.19}$$

In terms of our parameter τ, we thus have the length l as a functional of the form discussed above

$$l = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \ \left[\sum_{i,j=1}^{n} \dot{q}_i g_{i,j}(\boldsymbol{q},\tau) \dot{q}_j \right]^{1/2} \ . \tag{1.20}$$

A *geodesic* is a curve between two points whose length (calculated using the given metric) is stationary against infinitesimal variations about that path. Thus the task of finding geodesics fits within the class of variational problems we have discussed, and we can use the Euler–Lagrange equations to find them. Perhaps the best known result on geodesics is the fact that the shortest path between two points in a Euclidean space (one where $g_{i,j} = 0$ for $i \neq j$ and $g_{i,j} = 1$ for $i = j$) is a straight line. Another well-known result is that the shortest path between two points on the surface of a sphere is a great circle (see Problem 1.6.3 for a general theorem on geodesics on a curved surface).

Geodesics are not necessarily purely geometrical objects, but can have physical interpretations. For instance, suppose we want to find the shape of an elastic string stretched over a slippery surface. The string will adjust its shape to minimize its elastic energy. Since the elastic potential energy is a

monotonically increasing function of the length of the string, the string will settle onto a geodesic on the surface.

Geodesics also play an important role in General Relativity, because the world line of a photon is a geodesic in 4-dimensional space time, with the metric tensor obeying Einstein's equations. If the metric is sufficiently distorted, it can happen that there is not one, but *several* geodesics between two points, a fact which explains the phenomenon of *gravitational lensing* (multiple images of a distant galaxy behind a closer massive object).

1.4.2 Trial function method

One advantage of the variational formulation of a problem is that we can use *trial function* methods to find approximate solutions. That is, we can make a clever guess, $\boldsymbol{q}(t) = \boldsymbol{q}_K(t, a_1, a_2, \ldots, a_K)$ as to the general form of the solution, using some specific function $\boldsymbol{q}_K$ (the trial function) involving a finite number of parameters a_k, $k = 1, \ldots, K$. Then we evaluate the integral in eq. (1.11) (analytically or numerically) and seek a stationary point of the resulting function $I(a_1, a_2, \ldots, a_K)$ in the K-dimensional space of the parameters a_k. Varying the a_k the variation in I is

$$\delta I = \sum_{k=1}^{K} \frac{\partial I}{\partial a_k} \delta a_k \ . \tag{1.21}$$

The condition for a stationary point is thus

$$\frac{\partial I}{\partial a_k} = 0, \quad k = 1, \ldots, K \ , \tag{1.22}$$

that is, that the K-dimensional gradient of I vanish.

Since the true solution makes the objective functional I stationary with respect to small variations, if our guessed trial function solution is close to the true solution the error in I will be small. (Of course, because $\boldsymbol{q}_K$ may not be a reasonable guess for the solution in all ranges of the parameters, there may be spurious stationary points that must be rejected because they cannot possibly be close to a true solution—see the answer to Problem 2.8.2 in Sec. 5.2.)

1.5 Constrained variation: Lagrange multipliers

As mentioned in Sec. 1.2 we normally assume that the holonomic constraints have been used to reduce the dimensionality of the configuration space so

that all variations are allowed. However, it may not be possible to do an analytic elimination explicitly. Or it may be that some variables appear in a symmetric fashion, making it inelegant to eliminate one in favour of the others.

Thus, even in the holonomic case, it is worth seeking a method of handling *constrained variations*: when there are one or more differential auxiliary conditions of the form $\delta f^{(j)} = 0$. In the nonholonomic case it is mandatory to consider such variations because the auxiliary conditions cannot be integrated.

We denote the dimension of the configuration space by n. Following eq. (1.5) we suppose there are $m < n$ auxiliary conditions of the form

$$\delta f^{(j)} \equiv \boldsymbol{\omega}^{(j)}(\boldsymbol{q}, t) \cdot \delta \boldsymbol{q} = 0 \ . \tag{1.23}$$

The vectors $\boldsymbol{\omega}^{(j)}$, $j = 1, \ldots, m$ may be assumed linearly independent (else some of the auxiliary conditions would be redundant) and thus span an m-dimensional subspace, $V_m(t)$, of the full n-dimensional linear vector space V_n occupied by the unconstrained variations. Thus the equations eq. (1.23) constrain the variations $\delta \boldsymbol{q}$ to lie within an $(n-m)$-dimensional subspace, $V_{n-m}(\boldsymbol{q}, t)$, complementary to V_n.

The variational problem we seek to solve is to find the conditions (the generalizations of the Euler–Lagrange equations) under which the objective functional $I[\boldsymbol{q}]$ is stationary with respect to all variations $\delta \boldsymbol{q}$ in $V_m(\boldsymbol{q}, t)$. Apart from this restriction on the variations, the problem is the same as that described by eq. (1.16). The generalization of eq. (1.17) is

$$\frac{\delta f}{\delta \boldsymbol{q}} \cdot \delta \boldsymbol{q} = 0 \quad \forall \ \delta \boldsymbol{q} \in V_{n-m}(\boldsymbol{q}, t) \ . \tag{1.24}$$

If there are no constraints, so that $m = 0$, then $\delta f / \delta \boldsymbol{q}$ is orthogonal to all vectors in V_n and the only solution is that $\delta f / \delta \boldsymbol{q} \equiv 0$. Thus eq. (1.24) and eq. (1.17) are equivalent in this case. However, if $m < n$, then $\delta f / \delta \boldsymbol{q}$ can have a *nonvanishing* component in the subspace V_m and eq. (1.17) is no longer valid.

An elegant solution to the problem of generalizing eq. (1.17) was found by Lagrange. Expressed in our linear vector space language, his idea was that eq. (1.24) can be regarded as the statement that the *projection*, $(\delta f / \delta \boldsymbol{q})_{n-m}$, of $\delta f / \delta \boldsymbol{q}$ into $V_{n-m}(\boldsymbol{q}, t)$ is required to vanish.

However, we can write $(\delta f / \delta \boldsymbol{q})_{n-m}$ as $\delta f / \delta \boldsymbol{q} - (\delta f / \delta \boldsymbol{q})_m$, where $(\delta f / \delta \boldsymbol{q})_m$ is the projection of $\delta f / \delta \boldsymbol{q}$ into V_m. Now observe that we can write any vector in V_m as a linear superposition of the $\boldsymbol{\omega}^{(j)}$ since they form a basis spanning

this space. Thus we write $(\delta f/\delta \boldsymbol{q})_m = -\sum \lambda_j \boldsymbol{\omega}^{(j)}$, or, equivalently,

$$\left(\frac{\delta f}{\delta \boldsymbol{q}}\right)_m + \sum_{j=1}^{m} \lambda_j \boldsymbol{\omega}^{(j)} = 0 \,, \tag{1.25}$$

where the $\lambda_j(\boldsymbol{q}, t)$ coefficients, as yet to be determined, are known as the *Lagrange multipliers*. They can be determined by dotting eq. (1.25) with each of the m basis vectors $\boldsymbol{\omega}^{(j)}$, thus providing m equations for the m unknowns. Alternatively, we can express this variationally as

$$\left[\left(\frac{\delta f}{\delta \boldsymbol{q}}\right)_m + \sum_{j=1}^{m} \lambda_j \boldsymbol{\omega}^{(j)}\right] \cdot \delta \boldsymbol{q} = 0 \quad \forall\ \delta \boldsymbol{q} \in V_m(\boldsymbol{q}, t) \,. \tag{1.26}$$

Since $-\sum \lambda_j \boldsymbol{\omega}^{(j)}$ is the projection into V_m of $\delta f/\delta \boldsymbol{q}$, the projection into the complementary subspace V_{n-m} is found by subtracting $(-\sum \lambda_j \boldsymbol{\omega}^{(j)})$ from $\delta f/\delta \boldsymbol{q}$. That is,

$$\begin{aligned} \left(\frac{\delta f}{\delta \boldsymbol{q}}\right)_{n-m} &= \frac{\delta f}{\delta \boldsymbol{q}} + \sum_{j=1}^{m} \lambda_j \boldsymbol{\omega}^{(j)} \\ &= 0 \,, \end{aligned} \tag{1.27}$$

where the V_{n-m} component of the second equality follows by eq. (1.24) and the V_n component from eq. (1.25). Thus we have n generalized Euler–Lagrange equations, but they incorporate the m equations for the, so far arbitrary, $\lambda^{(j)}$ implicit in eq. (1.25). Thus we really only gain $(n-m)$ equations from the variational principle, which is at it should be because we also get m kinematic equations from the constraint conditions—if we got more from the variational principle the problem would be overdetermined.

The variational formulation of the second equality in eq. (1.27) is

$$\left[\frac{\delta f}{\delta \boldsymbol{q}} + \sum_{j=1}^{m} \lambda_j \boldsymbol{\omega}^{(j)}\right] \cdot \delta \boldsymbol{q} = 0 \quad \forall\ \delta \boldsymbol{q} \in V_n(\boldsymbol{q}, t) \,. \tag{1.28}$$

That is, *by using the Lagrange multipliers we have turned the constrained variational problem into an unconstrained one.*

In the holonomic case, when the auxiliary conditions are of the form in eq. (1.2), we may derive eq. (1.28) by unconstrained variation of the modified objective functional

$$I_*[\boldsymbol{q}] \equiv \int_{t_1}^{t_2} \mathrm{d}t \, (f + \sum_{j=1}^{m} \lambda_j f_j) \,. \tag{1.29}$$

The auxiliary conditions also follow from this functional if we require that it be stationary under variation of the λ_j.

1.6 Problems

1.6.1 Rigid rod

Two particles are connected by a rigid rod so they are constrained to move a fixed distance apart. Write down a constraint equation of the form eq. (1.2) and find suitable generalized coordinates for the system incorporating this holonomic constraint.

1.6.2 Ecliptic

Suppose we know that the angular momentum vectors $\mathbf{r}_k \times m_k \dot{\mathbf{r}}_k$ of a system of particles are all nonzero and parallel to the z-axis in a particular Cartesian coordinate system. Write down the differential constraints implied by this fact, and show that they lead to a set of holonomic constraints. Hence write down suitable generalized coordinates for the system.

1.6.3 Curvature of geodesics

Show that any geodesic $\mathbf{r} = \mathbf{x}(\tau)$ on a two-dimensional manifold S : $\mathbf{r} = \mathbf{X}(\theta, \zeta)$ embedded in ordinary Euclidean 3-space, where θ and ζ are arbitrary curvilinear coordinates on S, is such that the curvature vector $\boldsymbol{\kappa}(\tau)$ is everywhere **normal** to S (or zero).

The curvature vector is defined by $\boldsymbol{\kappa} \equiv \mathrm{d}\mathbf{e}_\parallel / \mathrm{d}l$, where $\mathbf{e}_\parallel(\tau) \equiv \mathrm{d}\mathbf{x} / \mathrm{d}l$ is the unit tangent vector at each point along the path $\mathbf{r} = \mathbf{x}(\tau)$.

Hint: First find $f(\theta, \zeta, \dot{\theta}, \dot{\zeta}) = \dot{l}$, the integrand of the length functional, $l = \int f \, \mathrm{d}\tau$ (which involves finding the metric tensor in θ, ζ space in terms of $\partial\mathbf{X}/\partial\theta$ and $\partial\mathbf{X}/\partial\zeta$). Then show that, for any path on S,

$$\frac{\partial f}{\partial \dot{\theta}} = \mathbf{e}_\parallel \cdot \frac{\partial \mathbf{X}}{\partial \theta}$$

(and similarly for the $\dot{\zeta}$ derivative) and

$$\frac{\partial f}{\partial \theta} = \mathbf{e}_\parallel \cdot \frac{\mathrm{d}}{\mathrm{d}\tau} \frac{\partial \mathbf{X}}{\partial \theta} \ ,$$

and again similarly for the ζ derivative.

Chapter 2

Lagrangian Mechanics

2.1 Introduction

The previous chapter dealt with generalized *kinematics*—the description of given motions in time and space. In this chapter we deal with one formulation (due to Lagrange) of generalized *dynamics*—the derivation of differential equations (*equations of motion*) for the time evolution of the generalized coordinates. Given appropriate initial conditions, these (in general, nonlinear) equations of motion specify the motion uniquely. Thus, in a sense, the most important task of the physicist is over when the equations of motion have been derived—the rest is just mathematics or numerical analysis (important though these are). The goal of generalized dynamics is to find universal forms of the equations of motion.

From elementary mechanics we are all familiar with Newton's Second Law, $\mathbf{F} = m\mathbf{a}$ for a particle of mass m subjected to a force $\mathbf{F}$ and undergoing an acceleration $\mathbf{a} \equiv \ddot{\mathbf{r}}$. If we know the Cartesian components $F_i(\mathbf{r}, \dot{\mathbf{r}}, t)$, $i = 1, 2, 3$, of the force in terms of the Cartesian coordinates $x_1 = x$, $x_2 = y$, $x_3 = z$ and their first time derivatives then the equations of motion are the set of three second-order differential equations $m\ddot{x}_i - F_i = 0$.

To give a physical framework for developing our generalized dynamical formalism we consider a set of N Newtonian point masses, which may be connected by holonomic constraints so the number n of generalized coordinates may be less than $3N$. Indeed, in the case of a rigid body N is essentially infinite, but the number of generalized coordinates is finite. For example, the generalized coordinates for a rigid body could be the three Cartesian coordinates of the centre of mass and three angles to specify its orientation (known as the Euler angles), so $n = 6$ for a rigid body allowed to move freely in space.

Whether the point masses are real particles like electrons, composite particles like nuclei or atoms, or mathematical idealizations like the infinitesimal volume elements in a continuum description, we shall refer to them generically as "particles".

Having found a very general form of the equations of motion (Lagrange's equations), we then find a variational principle (Hamilton's Principle) that gives these equations as its Euler–Lagrange equations in the case of no frictional dissipation. This variational principle forms a basis for generalizing even beyond Newtonian mechanics (e.g. to dynamics in Special Relativity).

2.2 Generalized Newton's 2nd Law

2.2.1 Generalized force

Let the (constrained) position of each of the N particles making up the system be given as a function of the n generalized coordinates $\boldsymbol{q}$ by $\mathbf{r}_k = \mathbf{x}_k(\boldsymbol{q}, t)$, $k = 1, \ldots, N$. If there are holonomic constraints acting on the particles, the number of generalized coordinates satisfies the inequality $n \leq 3N$. Thus, the system may be divided into two *subsystems*—an "exterior" subsystem described by the n generalized coordinates and an "interior" constraint subsystem whose $(3N - n)$ coordinates are rigidly related to the $\boldsymbol{q}$ by the geometric constraints.

In a naive Newtonian approach we would have to specify the forces acting on each particle (taking into account Newton's Third Law, "action and reaction are equal and opposite"), derive the $3N$ equations of motion for each particle and then eliminate all the interior subsystem coordinates to find the equations of motion of the generalized coordinates only. This is clearly very inefficient, and we already know from elementary physics that it is unnecessary—we do not really need an infinite number of equations to describe the motion of a rigid body. What we seek is a formulation in which only the generalized coordinates, and generalized forces conjugate to them, appear explicitly. All the interior coordinates and the forces required to maintain their constrained relationships to each other (the "forces of constraint") should be implicit only.

To achieve this it turns out to be fruitful to adopt the viewpoint that the total mechanical *energy* (or, rather, its change due to the performance of external *work*, W, on the system) is the primitive concept, rather than the vector quantity *force*. The basic reason is that the energy, a scalar quantity, needs only specification of the coordinates for its full description, whereas the representation of the force, a vector, depends also on defining a basis

set on which to resolve it. The choice of basis set is not obvious when we are using generalized coordinates. [Historically, force came to be understood earlier, but energy also has a long history, see e.g. Ernst Mach "The Science of Mechanics" (Open Court Publishing, La Salle, Illinois, 1960) pp. 309–312, QA802.M14 Hancock. With the development of Lagrangian and Hamiltonian methods, and thermodynamics, energy-based approaches can now be said to be dominant in physics.]

To illustrate the relation between force and work, first consider just one particle. Recall that the work δW done on the particle by a force $\mathbf{F}$ as the particle suffers an infinesimal displacement $\delta\mathbf{r}$ is $\delta W = \mathbf{F}\cdot\delta\mathbf{r} \equiv F_x\delta x + F_y\delta y + F_z\delta z$. A single displacement $\delta\mathbf{r}$ does not give enough information to determine the three components of $\mathbf{F}$, but if we imagine the thought experiment of displacing the particle in the three independent directions, $\delta\mathbf{r} = \delta x\,\mathbf{e}_x$, $\delta y\,\mathbf{e}_y$ and $\delta z\,\mathbf{e}_z$, determining the work, δW_x, δW_y, δW_z, done in each case, then we will have enough equations to deduce the three components of the force vector, $F_x = \delta W_x/\delta x$, $F_y = \delta W_y/\delta y$, $F_z = \delta W_z/\delta z$. [If δW can be integrated to give a function $W(\mathbf{r})$ (which is not always possible), then we may use standard partial derivative notation: $F_x = \partial W/\partial x$, $F_y = \partial W/\partial y$, $F_z = \partial W/\partial z$.]

The displacements $\delta x\,\mathbf{e}_x$, $\delta y\,\mathbf{e}_y$ and $\delta z\,\mathbf{e}_z$ are historically called *virtual displacements*. They are really simply the same displacements as used in the mathematical definition of partial derivatives. Note that the virtual displacements are done at a fixed instant in time as if by some "invisible hands", which perform the work δW: this is a thought experiment—the displacements are fictitious, not dynamical.

Suppose we now transform from the Cartesian coordinates x, y, z to an arbitrary curvilinear coordinate system q^1, q^2, q^3 (the superscript notation being conventional in tensor calculus). Then an arbitrary virtual displacement is given by $\delta\mathbf{r} = \sum_i \delta q^i\,\mathbf{e}_i$, where the basis vectors $\mathbf{e}_i$ are in general *not* orthonormal. The corresponding "virtual work" is given by $\delta W = \sum_i \delta q^i\,F_i$, where $F_i \equiv \mathbf{e}_i\cdot\mathbf{F}$. As in the Cartesian case, this can be used to determine the *generalized forces* F_i by determining the virtual work done in three independent virtual displacements.

In tensor calculus the set $\{F_i\}$ is known as the *covariant representation* of $\mathbf{F}$, and is in general distinct from an alternative resolution, the *contravariant* representation $\{F^i\}$. (You will meet this terminology again in relativity theory.) The energy approach shows that the covariant, rather than the contravariant, components of the force form the natural generalized forces conjugate to the generalized coordinates q^i.

Turning now to the N-particle system as a whole, the example above suggests we *define* the set of n generalized forces, $\{Q_i\}$, conjugate to each

of the n degrees of freedom q_i, to be such that the virtual work done on the system in displacing it by an arbitrary infinitesimal amount $\delta\boldsymbol{q}$ at fixed time t is given by

$$\delta W \equiv \sum_{i=1}^{n} Q_i \delta q_i \quad \forall\, \delta\boldsymbol{q}\ . \tag{2.1}$$

We now calculate the virtual work in terms of the displacements of the N particles assumed to make up the system and the forces $\mathbf{F}_k$ acting on them. The virtual work is

$$\begin{aligned} \delta W &= \sum_{k=1}^{N} \mathbf{F}_k \cdot \delta\mathbf{r}_k \\ &= \sum_{i=1}^{n} \delta q_i \left(\sum_{k=1}^{N} \mathbf{F}_k \cdot \frac{\partial \mathbf{x}_k}{\partial q_i} \right)\ . \end{aligned} \tag{2.2}$$

Comparing eq. (2.1) and eq. (2.2), and noting that they hold for *any* $\delta\boldsymbol{q}$, we can in particular take all but one of the δq_i to be zero to pick out the ith component, giving the generalized force as

$$Q_i = \sum_{k=1}^{N} \mathbf{F}_k \cdot \frac{\partial \mathbf{x}_k}{\partial q_i}\ . \tag{2.3}$$

When there are holonomic constraints on the system we decompose the forces acting on the particles into what we shall call *explicit forces* and *forces of constraint*. (The latter terminology is standard, but the usage "explicit forces" seems new—often they are called "applied forces", but this is confusing because they need not originate externally to the system, but also from interactions between the particles.)

By forces of constraint, $\mathbf{F}_k^{\mathrm{cst}}$, we mean those imposed on the particles by the rigid rods, joints, sliding planes etc. that make up the holonomic constraints on the system. These forces simply adjust themselves to whatever values are required to maintain the geometric constraint equations and can be regarded as "private" forces that, for most purposes, we do not need to know[1]. Furthermore, we may not be able to tell what these forces need to be until we have solved the equations of motion, so they cannot be assumed known *a priori*.

The explicit force on each particle, $\mathbf{F}_k^{\mathrm{xpl}}$, is the vector sum of any *externally imposed* forces, such as those due to an external gravitational or electric field, plus any *interaction* forces between particles such as those due to elastic

[1] Of course, in practical engineering design contexts one should at some stage check that the constraint mechanism is capable of supplying the required force without deforming or breaking!

springs coupling point masses, or to electrostatic attractions between charged particles. If there is friction acting on the particle, including that due to constraint mechanisms, then that must be included in $\mathbf{F}_k^{\rm xpl}$ as well. These are the "public" forces that determine the dynamical evolution of the degrees of freedom of the system and are determined by the configuration $\boldsymbol{q}$ of the system at each instant of time, and perhaps by the generalized velocity $\dot{\boldsymbol{q}}$ in the case of velocity-dependent forces such as those due to friction and those acting on a charged particle moving in a magnetic field.

Figure 2.1 shows a simple system with a holonomic constraint—a particle sliding on a plane inclined at angle α. It is subject to the force of gravity, $m\mathbf{g}$, the normal force $\mathbf{N}$, and a friction force F in the directions shown. The force of constraint is $\mathbf{N}$. It does no work because it is orthogonal to the direction of motion, and its magnitude is that required to null out the normal component of the gravitational force, $|\mathbf{N}| = mg\cos\alpha$, so as to give no acceleration in the normal direction and thus maintain the constraint.

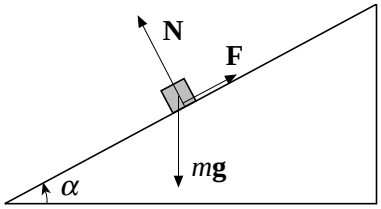

Figure 2.1: A body on an inclined plane as described in the text.

We now make the crucial observation that, because the constraints are assumed to be provided by rigid, undeformable mechanisms, no work can be done on the interior constraint subsystem by the virtual displacements. That is, *no (net) virtual work is done against the nonfriction forces imposed by the particles on the constraint mechanisms* in performing the variations $\delta\boldsymbol{q}$. By Newton's Third Law, the nonfriction forces acting on the constraint mechanisms are equal and opposite to the forces of constraint, $\mathbf{F}_k^{\rm cst}$. Thus the *sum over* $\mathbf{F}_k^{\rm cst}\cdot\delta\mathbf{x}_k$ *vanishes* and we can replace the total force $\mathbf{F}_k$ with $\mathbf{F}_k^{\rm xpl}$ in eqs. (2.2) and (2.3). **Note:** If there are friction forces associated with the constraints there is work done against these, but this fact does not negate the above argument because we have included the friction forces in the *explicit* forces—any work done against friction forces goes into heat which is dissipated into the external world, not into mechanical energy within the constraint subsystem.

That is,

$$Q_i = \sum_{k=1}^{N} \mathbf{F}_k^{\rm xpl}\cdot\frac{\partial\mathbf{x}_k}{\partial q_i}\ . \tag{2.4}$$

For the purpose of calculating the generalized forces, this is a much more practical expression than eq. (2.3) because the $\mathbf{F}_k^{\rm xpl}$ are known in terms of

the instantaneous positions and velocities. Thus $Q_i = Q_i(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$.

2.2.2 Generalized equation of motion

We now suppose that eq. (2.4) has been used to determine the $\boldsymbol{Q}$ as functions of the $\boldsymbol{q}$ (and possible $\dot{\boldsymbol{q}}$ and t if we have velocity-dependent forces and time-dependent constraints, respectively). Then we rewrite eq. (2.3) in the form

$$\sum_{k=1}^{N} \mathbf{F}_k \cdot \frac{\partial \mathbf{x}_k}{\partial q_i} = Q_i(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \ . \tag{2.5}$$

To derive an equation of motion we use Newton's second law to replace $\mathbf{F}_k$ on the left-hand side of eq. (2.5) with $m_k \ddot{\mathbf{r}}_k$,

$$\begin{aligned} &\sum_{k=1}^{N} m_k \ddot{\mathbf{r}}_k \cdot \frac{\partial \mathbf{x}_k}{\partial q_i} \\ &\equiv \sum_{k=1}^{N} m_k \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\dot{\mathbf{r}}_k \cdot \frac{\partial \mathbf{x}_k}{\partial q_i} \right) - \dot{\mathbf{r}}_k \cdot \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathbf{x}_k}{\partial q_i} \right] = Q_i \ . \end{aligned} \tag{2.6}$$

We now differentiate $\mathbf{r}_k = \mathbf{x}_k(\boldsymbol{q}(t), t)$ with respect to t to find the function $\mathbf{v}_k$ such that $\dot{\mathbf{r}}_k = \mathbf{v}_k(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$:

$$\dot{\mathbf{r}}_k = \frac{\partial \mathbf{x}_k}{\partial t} + \sum_{j=1}^{n} \dot{q}_j \frac{\partial \mathbf{x}_k}{\partial q_j} \equiv \mathbf{v}_k(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \ . \tag{2.7}$$

Differentiating $\mathbf{v}_k$ wrt $\dot{q}_i$ (treating $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ as independent variables in partial derivatives) we immediately have the lemma

$$\frac{\partial \mathbf{x}_k}{\partial q_i} = \frac{\partial \mathbf{v}_k}{\partial \dot{q}_i} \ . \tag{2.8}$$

The second lemma about partial derivatives of $\mathbf{v}_k$ that will be needed is

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathbf{x}_k}{\partial q_i} = \frac{\partial \mathbf{v}_k}{\partial q_i} \ , \tag{2.9}$$

which follows because $\mathrm{d}/\,\mathrm{d}t$ can be replaced by $\partial/\partial t + \dot{\boldsymbol{q}} \cdot \partial/\partial \boldsymbol{q}$, which commutes with $\partial/\partial q_i$ (cf. the "interchange of delta and dot" lemma in Sec. 1.4).

Applying these two lemmas in eq. (2.6) we find

$$\begin{aligned} &\sum_{k=1}^{N} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(m_k \mathbf{v}_k \cdot \frac{\partial \mathbf{v}_k}{\partial \dot{q}_i} \right) - m_k \mathbf{v}_k \cdot \frac{\partial \mathbf{v}_k}{\partial q_i} \right] \\ &= \sum_{k=1}^{N} \left[\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{q}_i} \left(\tfrac{1}{2} m_k \mathbf{v}_k^2 \right) - \frac{\partial}{\partial q_i} \left(\tfrac{1}{2} m_k \mathbf{v}_k^2 \right) \right] = Q_i \ , \end{aligned} \tag{2.10}$$

In terms of the total kinetic energy of the system, T

$$T \equiv \sum_{k=1}^{N} \tfrac{1}{2} m_k \mathbf{v}_k^2 \ , \tag{2.11}$$

we write eq. (2.10) compactly as the *generalized Newton's second law*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i \ . \tag{2.12}$$

These n equations are sometimes called Lagrange's equations of motion, but we shall reserve this term for a later form [eq. (2.24)] arising when we assume a special (though very general) form for the Q_i. They are also sometimes called (e.g. Scheck p. 83) d'Alembert's equations, but this may be historically inaccurate so is best avoided.

2.2.3 Example: Motion in Cartesian coordinates

Let us check that we can recover Newton's equations of motion as a special case when $\boldsymbol{q} = \{x, y, z\}$. In this case

$$T = \tfrac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \tag{2.13}$$

so

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0 \tag{2.14}$$

and

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x} \ , \quad \frac{\partial T}{\partial \dot{y}} = m\dot{y} \ , \quad \frac{\partial T}{\partial \dot{z}} = m\dot{z} \ . \tag{2.15}$$

Also, from eq. (2.4) we see that $Q_i \equiv F_i$. Substituting in eq. (2.12) we immediately recover Newton's 2nd Law in Cartesian form

$$m\ddot{x} = F_x \quad m\ddot{y} = F_y \quad m\ddot{z} = F_y \tag{2.16}$$

as expected.

2.3 Lagrange's equations (scalar potential case)

In many problems in physics the forces $\mathbf{F}_k$ are derivable from a *potential*, $V(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$. For instance, in the classical N-body problem the particles are assumed to interact pairwise via a two-body interaction potential

$V_{k,l}(\mathbf{r}_k, \mathbf{r}_l) \equiv U_{k,l}(|\mathbf{r}_k - \mathbf{r}_l|)$ such that the force on particle k due to particle l is given by

$$\begin{aligned} \mathbf{F}_{k,l} &= -\boldsymbol{\nabla}_k V_{k,l} \\ &= -\frac{(\mathbf{r}_k - \mathbf{r}_l)}{|\mathbf{r}_k - \mathbf{r}_l|} U'_{k,l}(r_{k,l}) \ , \end{aligned} \tag{2.17}$$

where the prime on U denotes the derivative with respect to its argument, the interparticle distance $r_{k,l} \equiv |\mathbf{r}_k - \mathbf{r}_l|$. Then the total force on particle k is found by summing the forces on it due to all the other particles

$$\begin{aligned} \mathbf{F}_k &= -\sum_{l \neq k} \boldsymbol{\nabla}_k V_{k,l} \\ &= -\boldsymbol{\nabla}_k V \ , \end{aligned} \tag{2.18}$$

where the N-body potential V is the sum of all *distinct* two-body interactions

$$\begin{aligned} V &\equiv \sum_{k=1}^{N} \sum_{l \leq k} V_{k,l} \\ &= \tfrac{1}{2} \sum_{k,l=1}^{N}{}' V_{k,l} \ . \end{aligned} \tag{2.19}$$

In the first line we counted the interactions once and only once: noting that $V_{k,l} = V_{l,k}$, so that the matrix of interactions is symmetric we have kept only those entries below the diagonal to avoid double counting. In the second, more symmetric, form we have summed *all* the off-diagonal entries of the matrix but have compensated for the double counting by dividing by 2. The exclusion of the diagonal "self-interaction" potentials is indicated by putting a prime on the $\sum$.

Physical examples of such an N-body system with binary interactions are:

- An unmagnetized plasma, where $V_{k,l}$ is the Coulomb interaction

$$U_{k,l}(r) = \frac{e_k e_l}{\epsilon_0 r} \ , \tag{2.20}$$

  where the e_k are the charges on the particles and ϵ_0 is the permittivity of free space. We could also allow for the effect of gravity by adding the potential $\sum_k m_k g z_k$ to V, where z_k is the height of the kth particle with respect to a horizontal reference plane and m_k is its mass.

- A globular cluster of stars, where $V_{k,l}$ is the gravitational interaction

$$U_{k,l}(r) = \frac{G m_k m_l}{r} \ , \tag{2.21}$$

  where G is the gravitational constant.

- A dilute monatomic gas, where $V_{k,l}$ is the Van der Waal's interaction. However, if the gas is too dense (or becomes a liquid) we would have to include 3-body or higher interactions as the wave functions of more than two atoms could overlap simultaneously.

Even when the system is subjected to external forces, such as gravity, and/or holonomic constraints, we can often still assume that the "explicit forces" are derivable from a potential

$$\mathbf{F}_k^{\rm xpl} = -\boldsymbol{\nabla}_k V \ . \tag{2.22}$$

Taking into account the constraints, we see that the potential $V(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$ becomes a function, $V(\boldsymbol{q}, t)$, in the reduced configuration space. Then, from eq. (2.4) we have

$$\begin{aligned} Q_i &= -\sum_{k=1}^{N} \frac{\partial \mathbf{x}_k}{\partial q_i} \boldsymbol{\cdot} \boldsymbol{\nabla}_k V \\ &= -\frac{\partial}{\partial q_i} V(\boldsymbol{q}, t) \ . \end{aligned} \tag{2.23}$$

Substituting this form for Q_i in eq. (2.12) we see that the generalized Newton's equations of motion can be encapsulated in the very compact form (Lagrange's equations of motion)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \ , \tag{2.24}$$

where the function $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, called the *Lagrangian*, is defined as

$$L \equiv T - V \ . \tag{2.25}$$

2.3.1 Hamilton's Principle

Comparing eq. (2.24) with eq. (1.17) we see that Lagrange's equations of motion have exactly the same form as the Euler–Lagrange equations for the variational principle $\delta S = 0$, where the functional $S[\boldsymbol{q}]$, defined by

$$S \equiv \int_{t_1}^{t_2} \mathrm{d}t \ L(\dot{\boldsymbol{q}}, \boldsymbol{q}, t) \ , \tag{2.26}$$

is known as the *action integral*. Since the natural boundary conditions eq. (1.18) are not physical, the variational principle is one in which the endpoints are to be kept fixed.

We can now state *Hamilton's Principle*: *Physical paths in configuration space are those for which the action integral is stationary against all infinitesimal variations that keep the endpoints fixed.*

By *physical paths* we mean those paths, out of all those that are consistent with the constraints, that actually obey the equations of motion with the given Lagrangian.

To go beyond the original Newtonian dynamics with a scalar potential that we used to motivate Lagrange's equations, we can instead take Hamilton's Principle, being such a simple and geometrically appealing result, as a more fundamental and natural starting point for Lagrangian dynamics.

2.4 Lagrangians for some Physical Systems

2.4.1 Example 1: 1-D motion—the pendulum

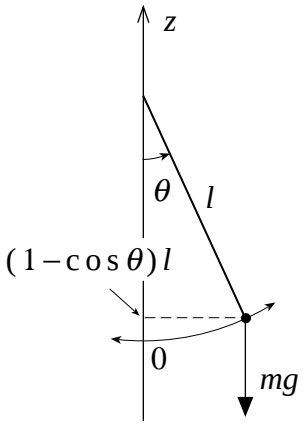

Figure 2.2: Physical pendulum.

One of the simplest nonlinear systems is the one-dimensional *physical pendulum* (so called to distinguish it from the linearized harmonic oscillator approximation). As depicted in Fig. 2.2, the pendulum consists of a light rigid rod of length l, making an angle θ with the vertical, swinging from a fixed pivot at one end and with a bob of mass m attached at the other.

The constraint $l = \text{const}$ and the assumption of plane motion reduces the system to one degree of freedom, described by the generalized coordinate θ. (This system is also called the *simple pendulum* to distinguish it from the spherical pendulum and compound pendula, which have more than one degree of freedom.) The potential energy with respect to the equilibrium position $\theta = 0$ is $V(\theta) = mgl(1 - \cos\theta)$, where g is the acceleration due to gravity, and the velocity of the bob is $v_\theta = l\dot{\theta}$, so that the kinetic energy $T = \frac{1}{2}mv_\theta^2 = \frac{1}{2}ml^2\dot{\theta}^2$. The Lagrangian, $T - V$, is thus

$$L(\theta, \dot{\theta}) = \tfrac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos\theta) \; . \tag{2.27}$$

This is also essentially the Lagrangian for a particle moving in a sinusoidal spatial potential, so the physical pendulum provides a paradigm for problems such as the motion of an electron in a crystal lattice or of an ion or electron in a plasma wave.

From eq. (2.27) $\partial L/\partial\dot{\theta} = ml^2\dot{\theta}$ and $\partial L/\partial\theta = -mgl\sin\theta$. Thus, the Lagrangian equation of motion is

$$ml^2\ddot{\theta} = -mgl\sin\theta \ . \tag{2.28}$$

Expanding the cosine up to quadratic order in θ gives the harmonic oscillator oscillator approximation (see also Sec. 2.6.2)

$$L \approx L_{\rm lin} \equiv \tfrac{1}{2}ml^2\dot{\theta}^2 - \tfrac{1}{2}mgl\theta^2 \ , \tag{2.29}$$

for which the equation of motion is, dividing through by ml^2, $\ddot{\theta} + \omega_0^2\theta = 0$, with $\omega_0 \equiv \sqrt{g/l}$.

2.4.2 Example 2: 2-D motion in a central potential

Let us work in plane polar coordinates, $\boldsymbol{q} = \{r, \theta\}$, such that

$$x = r\cos\theta \ , \quad y = r\sin\theta \ , \tag{2.30}$$

so that

$$\begin{aligned} \dot{x} &= \dot{r}\cos\theta - r\dot{\theta}\sin\theta \ , \\ \dot{y} &= \dot{r}\sin\theta + r\dot{\theta}\cos\theta \ , \end{aligned} \tag{2.31}$$

whence the kinetic energy $T \equiv \frac{1}{2}(\dot{x}^2 + \dot{y}^2)$ is found to be

$$T = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) \ . \tag{2.32}$$

An alternative derivation of eq. (2.32) may be found by resolving $\mathbf{v}$ into the components $\dot{r}\mathbf{e}_r$ and $r\dot{\theta}\mathbf{e}_\theta$, where $\mathbf{e}_r$ is the unit vector in the radial direction and $\mathbf{e}_\theta$ is the unit vector in the azimuthal direction.

We now consider the restricted two body problem—one light particle orbiting about a massive particle which may be taken to be fixed at $r = 0$ (e.g. an electron orbiting about a proton in the Bohr model of the hydrogen atom, or a planet orbiting about the sun). Then the potential $V = V(r)$ (given by eq. (2.20) or eq. (2.21)) is a function only of the radial distance from the central body and not of the angle.

Then, from eq. (2.25) the Lagrangian is

$$L = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r)\ . \tag{2.33}$$

First we observe that L is independent of θ (in which case θ is said to be *ignorable*). Then $\partial L/\partial\theta \equiv 0$ and the θ component of Lagrange's equations, eq. (2.24) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{\theta}} = 0\ , \tag{2.34}$$

which we may immediately integrate once to get an *integral of the motion*, i.e. a dynamical quantity that is constant along the trajectory

$$\frac{\partial L}{\partial\dot{\theta}} = \text{const}\ . \tag{2.35}$$

From eq. (2.33) we see that $\partial L/\partial\dot{\theta} = mr^2\dot{\theta}$, which is the *angular momentum*. Thus eq. (2.33) expresses conservation of angular momentum.

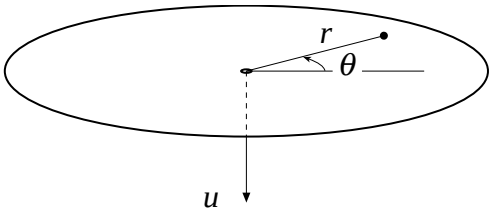

Figure 2.3: Planar motion with a time-varying centripetal constraint as described in the text.

Turning now to the r-component of Lagrange's equations, we see from eq. (2.33)

$$\frac{\partial L}{\partial r} = mr\dot{\theta}^2 - V'(r)\ ,$$
$$\frac{\partial L}{\partial\dot{r}} = m\dot{r}\ , \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{r}} = m\ddot{r}\ . \tag{2.36}$$

From eq. (2.24) we find the radial equation of motion to be

$$m\ddot{r} - mr\dot{\theta}^2 = -V'(r)\ . \tag{2.37}$$

2.4.3 Example 3: 2-D motion with time-varying constraint

Instead of free motion in a central potential, consider instead a weight rotating about the origin on a frictionless horizontal surface (see Fig. 2.3) and constrained by a thread, initially of length a, that is being pulled steadily downward at speed u through a hole at the origin so that the radius $r = a - ut$.

Then the Lagrangian is, substituting for r in eq. (2.32),

$$L = T = \tfrac{1}{2}m\left[u^2 + (a-ut)^2\dot{\theta}^2\right]\ . \tag{2.38}$$

Now only θ is an unconstrained generalized coordinate. As before, it is ignorable, and so we again have conservation of angular momentum

$$m(a-ut)^2\dot{\theta} \equiv l = \text{const} \ , \tag{2.39}$$

which equation can be integrated to give θ as a function of t, $\theta = \theta_0 + (l/mu)[1/(a-ut) - 1/a] = \theta_0 + lt/[ma(a-ut)]$.
Clearly angular momentum *is* conserved, because the purely radial string cannot exert any torque on the weight. Thus Lagrange's equation of motion gives the correct answer. However, the string is obviously doing work on the system because $T = \frac{1}{2}[mu^2 + (l^2/m)/(a-ut)^2]$ is *not* conserved. Have we not therefore violated the postulate in Sec. 2.2.2 of no work being done by the constraints? The answer is "no" because what we assumed in Sec. 2.2.2 was that no *virtual* work was done by the constraints. The fact that the constraint is time-dependent is irrelevant to this postulate, because virtual displacements are done instantaneously at any given time.

2.4.4 Example 4: Atwood's machine

Consider two weights of mass m_1 and m_2 suspended from a frictionless, inertialess pully of radius a by a rope of fixed length, as depicted in Fig. 2.4. The height of weight 1 is x with respect to the chosen origin and the holonomic constraint provided by the rope allows us to express the height of weight 2 as $-x$, so that there is only one degree of freedom for this system.

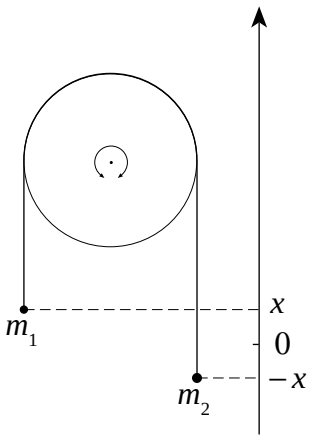

Figure 2.4: Atwood's machine.

The kinetic and potential energy are $T = \frac{1}{2}(m_1+m_2)\dot{x}^2$ and $V = m_1gx - m_2gx$. Thus $L = T - V$ is given by

$$L = \tfrac{1}{2}(m_1+m_2)\dot{x}^2 - (m_1-m_2)gx \tag{2.40}$$

and its derivatives are $\partial L/\partial x = -(m_1-m_2)g$
and $\partial L/\partial\dot{x} = (m_1+m_2)\dot{x}$, so that the equation of motion $\mathrm{d}(\partial L/\partial\dot{x}) = \partial L/\partial x$ becomes

$$\ddot{x} = -\frac{m_1-m_2}{m_1+m_2}g \ . \tag{2.41}$$

2.4.5 Example 5: Particle in e.m. field

The fact that Lagrange's equations are the Euler–Lagrange equations for the extraordinarily simple and general Hamilton's Principle (see Sec. 2.3.1) suggests that Lagrange's equations of motion may have a wider range of validity than simply problems where the force is derivable from a scalar potential. Thus we *do not* define L as $T-V$, but rather postulate the universal validity of Lagrange's equations of motion (or, equivalently, Hamilton's Principle), for describing non-dissipative classical dynamics and accept any Lagrangian as valid that gives the physical equation of motion.

In particular, it is obviously of great physical importance to find a Lagrangian for which Lagrange's equations of motion eq. (2.24) reproduce the equation of motion of a charged particle in an electromagnetic field, under the influence of the Lorentz force,

$$m\ddot{\mathbf{r}} = e\mathbf{E}(\mathbf{r},t) + e\dot{\mathbf{r}}\times\mathbf{B}(\mathbf{r},t)\ , \tag{2.42}$$

where e is the charge on the particle of mass m.

We assume the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, respectively, to be given in terms of the scalar potential Φ and vector potential $\mathbf{A}$ by the standard relations

$$\begin{aligned} \mathbf{E} &= -\boldsymbol{\nabla}\Phi - \partial_t\mathbf{A}\ , \\ \mathbf{B} &= \boldsymbol{\nabla}\times\mathbf{A}\ . \end{aligned} \tag{2.43}$$

The electrostatic potential energy is $e\Phi$, so we expect part of the Lagrangian to be $\frac{1}{2}m\dot{\mathbf{r}}^2 - e\Phi$, but how do we include the vector potential? Clearly we need to form a scalar since L is a scalar, so we need to dot $\mathbf{A}$ with one of the naturally occurring vectors in the problem to create a scalar. The three vectors available are $\mathbf{A}$ itself, $\mathbf{r}$ and $\dot{\mathbf{r}}$. However we do not wish to use $\mathbf{A}$, since $\mathbf{A}\cdot\mathbf{A}$ in the Lagrangian would give an equation of motion that is nonlinear in the electromagnetic field, contrary to eq. (2.42). Thus we can only use $\mathbf{r}$ and $\dot{\mathbf{r}}$. Comparing eqs. (2.42) and (2.43) we see that $\dot{\mathbf{r}}\cdot\mathbf{A}$ has the same dimensions as Φ, so let us try adding that to form the total Lagrangian

$$L = \tfrac{1}{2}m\dot{\mathbf{r}}^2 - e\Phi + e\dot{\mathbf{r}}\cdot\mathbf{A}\ . \tag{2.44}$$

Taking $\boldsymbol{q} \equiv \{q_1, q_2, q_3\} = \{x, y, z\}$ we have

$$\frac{\partial L}{\partial q_i} = -e\frac{\partial\Phi}{\partial q_i} + e\sum_{j=1}^{3}\dot{q}_j\frac{\partial A_j}{\partial q_i}\ , \tag{2.45}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = m\ddot{q}_i + e\frac{\mathrm{d}A_i}{\mathrm{d}t} = m\ddot{q}_i + e\left[\frac{\partial A_i}{\partial t} + \sum_{j=1}^{3}\dot{q}_j\frac{\partial A_i}{\partial q_j}\right] . \tag{2.46}$$

Substituting eqs. (2.45) and (2.46) in eq. (2.24) we find

$$m\ddot{q}_i = e\left[-\frac{\partial\Phi}{\partial q_i} - \frac{\partial A_i}{\partial t}\right] + e\sum_{j=1}^{3}\dot{q}_j\left[\frac{\partial A_j}{\partial q_i} - \frac{\partial A_i}{\partial q_j}\right] . \tag{2.47}$$

This is simply eq. (2.42) in Cartesian component form, so our guessed Lagrangian is indeed correct.

2.4.6 Example 6: Particle in ideal fluid

In Sec. 1.3 we presented a fluid as a system with an infinite number of degrees of freedom. However, if we concentrate only on the motion of a single fluid element (a *test particle*), taking the pressure p and mass density ρ as known, prescribed functions of $\mathbf{r}$ and t the problem becomes only three-dimensional.

Dividing by ρ we write the equation of motion of a fluid element as

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\boldsymbol{\nabla}p}{\rho} - \boldsymbol{\nabla}V , \tag{2.48}$$

where $V(\mathbf{r},t)$ is the potential energy (usually gravitational) per unit mass.

To find a Lagrangian for this motion we need to be able to combine the pressure gradient and density into an effective potential. In an ideal compressible fluid we have the equation of state $p(\,\mathrm{d}V)^\gamma = p_0(\,\mathrm{d}V_0)^\gamma$, where γ is the ratio of specific heats, which we can write as

$$\frac{p}{\rho^\gamma} = \text{const} , \tag{2.49}$$

where the right-hand side is a constant of the motion for the given test particle. If we further assume that it is the same constant for *all* fluid elements in the neighbourhood of the test particle, then we can take the gradient of the log of eq. (2.49) to get $(\boldsymbol{\nabla}p)/p = \gamma(\boldsymbol{\nabla}\rho)/\rho$. Thus

$$\begin{aligned}\frac{\boldsymbol{\nabla}p}{\rho} &= \boldsymbol{\nabla}\left(\frac{p}{\rho}\right) - \frac{p}{\rho^2}\boldsymbol{\nabla}\rho \\ &= \boldsymbol{\nabla}\left(\frac{p}{\rho}\right) - \frac{\boldsymbol{\nabla}p}{\gamma\rho} .\end{aligned}$$

Solving for $\boldsymbol{\nabla} p/\rho$ we get

$$\frac{\boldsymbol{\nabla} p}{\rho} = \boldsymbol{\nabla} h \tag{2.50}$$

where the *enthalpy* (per unit mass) is defined by

$$h \equiv \frac{\gamma}{\gamma - 1}\frac{p}{\rho} \, . \tag{2.51}$$

Using eq. (2.50) in eq. (2.48) we recognize it as the equation of motion for a particle of unit mass with total potential energy $h + V$. Thus the Lagrangian is

$$L = \tfrac{1}{2}\dot{\mathbf{r}}^2 - h - V \, . \tag{2.52}$$

2.5 Averaged Lagrangian

From Sec. 2.3.1 we know that Lagrangian dynamics has a variational formulation, and so we expect that trial function methods (see Sec. 1.4.2) may be useful as a way of generating approximate solutions of the Lagrangian equations of motion. In particular, suppose we know that the solutions are oscillatory functions of t with a frequency much higher than the inverse of any characteristic time for slow changes in the parameters of the system (the changes being then said to occur *adiabatically*). Then we may use a trial function of the form

$$\boldsymbol{q}(t) = \widetilde{\boldsymbol{q}}(\phi(t), A_1(t), A_2(t), \ldots) \tag{2.53}$$

where $\widetilde{\boldsymbol{q}}$ is a 2π-periodic function of ϕ, the *phase* of the rapid oscillations, and the A_k are a set of slowly varying amplitudes characterizing the waveform (e.g. see Problem 2.8.2). Thus,

$$\dot{\boldsymbol{q}} = \omega(t)\frac{\partial \widetilde{\boldsymbol{q}}}{\partial \phi} + \dot{A}_1 \frac{\partial \widetilde{\boldsymbol{q}}}{\partial A_1} + \dot{A}_2 \frac{\partial \widetilde{\boldsymbol{q}}}{\partial A_2} + \cdots \, , \tag{2.54}$$

where the *instantaneous frequency* is defined by

$$\omega(t) \equiv \dot{\phi}(t) \, . \tag{2.55}$$

Since $\omega \gg \mathrm{d} \ln A_k / \mathrm{d}t$, to a first approximation we may keep only the first term in eq. (2.54). Thus our approximate L is a function of ω, but *not* of $\dot{A}_1$, $\dot{A}_2$ etc.

Now take the time integration in the action integral to be over a time long compared with the period of oscillation, but short compared with the

timescale for changes in the system parameters. Thus only the *phase-average* of L,

$$\bar{L} \equiv \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi}\, L \tag{2.56}$$

contributes to the action,

$$S \approx \int_{t_1}^{t_2} \mathrm{d}t\, \bar{L}(\omega, A_1(t), A_2(t), \ldots) \tag{2.57}$$

for the class of oscillatory physical solutions we seek. Note that the averaging in eq. (2.56) removes all direct dependence on ϕ, so $\bar{L}$ depends only on its time derivative ω.

Thus, we have a new, approximate form of Hamilton's Principle in which the averaged Lagrangian replaces the exact Lagrangian, and in which the set $\{\phi, A_1, A_2, \ldots\}$ replaces the set $q_1, q_2, \ldots$ as the generalized variables. Requiring S to be stationary within the class of quasiperiodic trial functions then gives the new adiabatic Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \bar{L}}{\partial \dot{\phi}} = 0\ , \tag{2.58}$$

and

$$\frac{\partial \bar{L}}{\partial A_k} = 0\ , \quad k = 1, 2, \ldots\ . \tag{2.59}$$

The first equation expresses the conservation of the *adiabatic invariant*, $J = \text{const}$, where J is known as the oscillator *action*

$$J \equiv \frac{\partial \bar{L}}{\partial \omega}\ . \tag{2.60}$$

The second set of equations gives relations between ω and the A_ks that give the waveform and the instantaneous frequency.

2.5.1 Example: Harmonic oscillator

Consider a weight of mass m at the end of a light spring with spring constant $k = m\omega_0^2$. Then the kinetic energy is $T = \frac{1}{2}m\dot{x}^2$ and the potential energy is $V = \frac{1}{2}kx^2 = \frac{1}{2}m\omega_0^2 x^2$. Thus the *natural form* of the Lagrangian, $T - V$, is

$$L = \tfrac{1}{2}m(\dot{x}^2 - \omega_0^2 x^2)\ . \tag{2.61}$$

Now consider that the length of the spring, or perhaps the spring constant, changes slowly with time, so that $\omega_0 = \omega_0(t)$, with $\mathrm{d}\ln\omega_0/\,\mathrm{d}t \ll \omega_0$. Because

the system is changing with time, the amplitude of oscillation is not constant. We now use the average Lagrangian method to find the adiabatic invariant, which allows us to predict how the amplitude changes with time. (For an alternative derivation of adiabatic invariance in this case, see Sec. 4.2.1.)

Since the problem is linear, we use a sinusoidal trial function with no harmonics

$$x = A \sin \phi(t) \; . \tag{2.62}$$

In response to the slow changes in ω_0 we expect A and $\omega \equiv \dot{\phi}$ also to be slowly varying functions of time. Calculating the averaged Lagrangian as in eq. (2.56) we have

$$\bar{L} = \frac{mA^2}{4} \left(\omega^2 - \omega_0^2 \right) \; . \tag{2.63}$$

The Euler–Lagrange equation eq. (2.59) from varying A gives

$$\frac{\partial \bar{L}}{\partial A} = \frac{mA}{2} \left(\omega^2 - \omega_0^2 \right) = 0 \; . \tag{2.64}$$

Since we assume that the amplitude is not zero, this implies that the instaneous frequency tracks the changing natural frequency, $\omega = \pm\omega_0(t)$ (which also implies that $\bar{L} = 0$).

The adiabatically conserved action J, eq. (2.60), is

$$J = \frac{m\omega A^2}{2} \; . \tag{2.65}$$

From the adiabatic invariance of J we see that A varies as the inverse of $\omega_0^{1/2}$.

Remark 2.1 *Our derivation of adiabatic invariance has been carried out only to first order in ϵ, the small parameter expressing the ratio between the short oscillatory timescale and the long slow timescale. However, it can be shown that an averaged Lagrangian asymptotically independent of ϕ to all orders in a power series in ϵ can be found. Thus an adiabatic invariant can be defined to arbitrary order in ϵ. (This does not however mean that adiabatic invariance is exact, since asymptotically small changes, e.g. going as $\exp(-1/|\epsilon|)$, are not ruled out.)*

2.6 Transformations of the Lagrangian

2.6.1 Point transformations

Given an arbitrary Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ in one generalized coordinate system, $\boldsymbol{q} \equiv \{q_i | i = 1, n\}$ (e.g. a Cartesian frame), we often want to know

the Lagrangian $L'(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t)$ in another generalized coordinate system, $\boldsymbol{Q} \equiv \{Q_i | i = 1, n\}$ (e.g. polar coordinates as in the examples in Sec. 2.4.2 and Sec. 2.4.3). Thus, suppose there exists a set $\boldsymbol{g} \equiv \{g_i | i = 1, n\}$ of twice-differentiable functions g_i such that

$$q_i = g_i(\boldsymbol{Q}, t) \ , \quad i = 1, \ldots, n \ . \tag{2.66}$$

We require the inverse function of $\boldsymbol{g}$ also to be twice differentiable, in which case $\boldsymbol{g} : \boldsymbol{Q} \mapsto \boldsymbol{q}$ is said to be a C^2 *diffeomorphism*. Note that we have allowed the transformation to be time dependent, so transformations to a moving frame are allowed.

The transformation $\boldsymbol{g}$ maps a path in $\boldsymbol{Q}$-space to a path in $\boldsymbol{q}$-space. However, it is physically the same path—all we have changed is its *representation*. What we need in order to discuss how the Lagrangian transforms is a *coordinate-free* formulation of Lagrangian dynamics. This is another virtue of Hamilton's Principle— $\delta S = 0$ on a physical path for all variations with fixed endpoints—since the action integral S given by eq. (2.26) is an integral over time only. Thus, if we can define L' so that

$$L'(\boldsymbol{Q}(t), \dot{\boldsymbol{Q}}(t), t) = L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t), t) \tag{2.67}$$

for *any* path, then S is automatically invariant under the coordinate change and will be stationary for the same physical paths, irrespective of what coordinates they are represented in.

We can guarantee this trivially, simply by choosing the new Lagrangian to be the old one in the new coordinates:

$$L'(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t) = L(\boldsymbol{g}(\boldsymbol{Q}, t), \dot{\boldsymbol{g}}(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t), t) \ , \tag{2.68}$$

where

$$\dot{g}_i \equiv \frac{\partial g_i}{\partial t} + \sum_{j=1}^{n} \dot{Q}_j \frac{\partial g_i}{\partial Q_j} \ , \quad i = 1, \ldots, n \ . \tag{2.69}$$

One can prove that eq. (2.68) gives the correct dynamics by calculating the transformation of the Euler–Lagrange equations explicitly (see e.g. Scheck, 1990) and showing that eq. (2.24) implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial \dot{Q}_i}\right) - \frac{\partial L'}{\partial Q_i} = 0 \ , \tag{2.70}$$

but clearly Hamilton's Principle provides a much simpler and more elegant way of arriving at the same result, since eqs. 2.70 are simply the Euler–Lagrange equations for S to be stationary in the new variables.

We have in fact already run into a special case of this transformation rule, in Sec. 2.4.2 and Sec. 2.4.3, where we transformed the kinetic and potential energies from Cartesians to polars in the obvious way.

2.6.2 Gauge transformations

We saw in Sec. 2.4.5 that we cannot always expect that the Lagrangian is of the form $T - V$, but nevertheless found a function for which Lagrange's equations gave the correct dynamical equations of motion. Thus it is the requirement that the equations of motion be in the form of Lagrange's equations (or, equivalently, that Hamilton's Principle apply) that is fundamental, rather than the specific form of L.

That naturally raises the question: for a given system is there only one Lagrangian giving the correct equations of motion, or are there many?

Clearly there is a trivial way to generate multiple physically equivalent Lagrangians, and that is to multiply L by a constant factor. However, we usually normalize L in a natural way, e.g. by requiring that the part linear in the mass be equal to the kinetic energy, so this freedom is not encountered much in practice.

However, there is a more important source of nonuniqueness, known as a *gauge transformation* of the Lagrangian in which L is replaced by L', defined by

$$L'(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \frac{\partial M}{\partial t} + \dot{\boldsymbol{q}} \cdot \frac{\partial M}{\partial \boldsymbol{q}} \tag{2.71}$$

[using the shorthand notation defined in eq. (1.4)], where $M(\boldsymbol{q}, t)$ is any C^3 function. We recognize the term in M as the total time derivative, $\dot{M}$, and thus, from eq. (2.26) and eq. (2.71), the new action integral becomes

$$\begin{aligned} S' &\equiv \int_{t_1}^{t_2} \mathrm{d}t \; L'(\dot{\boldsymbol{q}}, \boldsymbol{q}, t) \\ &= \int_{t_1}^{t_2} \mathrm{d}t \; \left[L(\dot{\boldsymbol{q}}, \boldsymbol{q}, t) + \frac{\mathrm{d}}{\mathrm{d}t} M(\boldsymbol{q}, t) \right] \\ &= S + M(\boldsymbol{q}(t_2), t_2) - M(\boldsymbol{q}(t_1), t_1) \; . \end{aligned} \tag{2.72}$$

However, since the endpoints are held fixed in Hamilton's Principle, the endpoint terms $M(\boldsymbol{q}(t_{1,2}), t_{1,2})$ do not affect the allowed variations of S (i.e. $\delta S' \equiv \delta S$ for all admissible variations) and therefore S' is stationary if and only if S is. Thus both Lagrangians give identical physical paths and L and L' are completely equivalent dynamically.

Example 1: Harmonic oscillator

Consider the harmonic oscillator Lagrangian, eq. (2.61), for a particle of unit mass

$$L = \tfrac{1}{2}(\dot{x}^2 - \omega_0^2 x^2) \; , \tag{2.73}$$

where we take ω_0 to be constant. Since $\partial L/\partial \dot{x} = \dot{x}$ and $\partial L/\partial x = \omega_0^2 x$, we immediately verify that Lagrange's equation, eq. (2.24), gives the *harmonic oscillator equation*

$$\ddot{x} = -\omega_0^2 x \ . \tag{2.74}$$

Now add a gauge term, taking $M = \frac{1}{2}\omega_0 x^2$. Then, from eq. (2.71), the new Lagrangian becomes

$$L' = \tfrac{1}{2}(\dot{x}^2 + 2\omega_0 x\dot{x} - \omega_0^2 x^2) \ . \tag{2.75}$$

Calculating $\partial L'/\partial \dot{x} = \dot{x} + \omega_0 x$, and $\partial L'/\partial x = \omega_0^2 x + \omega_0 \dot{x}$ and substituting into the Lagrangian equation of motion, we see that the gauge contributions cancel. Thus we do indeed recover the harmonic oscillator equation, eq. (2.74), and the new Lagrangian is a perfectly valid one despite the fact that it is no longer in the natural form, $T - V$.

Example 2: Electromagnetic gauge transformation

It is well known that the scalar and vector potentials in eqs. (2.43) are not unique, since the electric and magnetic fields are left unchanged by the gauge transformation

$$\begin{aligned} \mathbf{A}(\mathbf{r},t) &\to \mathbf{A}'(\mathbf{r},t) \equiv \mathbf{A}(\mathbf{r},t) + \boldsymbol{\nabla}\chi(\mathbf{r},t) \\ \Phi(\mathbf{r},t) &\to \Phi'(\mathbf{r},t) \equiv \Phi(\mathbf{r},t) - \partial_t \chi(\mathbf{r},t) \ , \end{aligned} \tag{2.76}$$

where χ is an arbitrary scalar function (a *gauge potential*).

Using the gauge-transformed potentials Φ' and $\mathbf{A}'$ to define a new Lagrangian L' in the same way as L was defined by eq. (2.44) we have

$$L' = \tfrac{1}{2}m\dot{\mathbf{r}}^2 - e\Phi' + e\dot{\mathbf{r}}\cdot\mathbf{A}' \ . \tag{2.77}$$

Substituting eqs. (2.76) in eq. (2.77) we find

$$L' = L + \frac{\partial(e\chi)}{\partial t} + \dot{\mathbf{r}}\cdot\boldsymbol{\nabla}(e\chi) \ , \tag{2.78}$$

which is exactly of the form eq. (2.71) with $M = e\chi$. Thus electromagnetic and Lagrangian gauge transformations are closely related.

2.7 Symmetries and Noether's theorem

A Lagrangian L is said to have a *continuous symmetry* if it is *invariant* under the transformation $\boldsymbol{q} \mapsto \boldsymbol{h}_s(\boldsymbol{q})$, where s is a real continuous parameter and

$\boldsymbol{h}_{s=0}$ is the identity transformation. That is, given any path (not necessarily physical) $\boldsymbol{q}(t)$, L has the same value on all members of the continuous *family* of paths $\boldsymbol{q}(t,s) \equiv \boldsymbol{h}_s(\boldsymbol{q}(t))$. For example, in the case of motion in a central potential, the continuous symmetry is invariance of L under rotations: $\theta \mapsto \theta + s$.

Comparing two infinitesimally close paths we see that δL also possesses this symmetry, so that if $\delta S \equiv \int \delta L \, \mathrm{d}t = 0$ on the path $\boldsymbol{q}(t)$ then it also vanishes for all members of the family $\boldsymbol{q}(t,s)$. Thus, from Hamilton's Principle, *if* $\boldsymbol{q}(t,0) \equiv \boldsymbol{q}(t)$ *is a physical path, then so are all the paths* $\boldsymbol{q}(t,s)$ *generated by the symmetry operation* $\boldsymbol{h}_s$. Thus, given one solution of the equations of motion we can use a symmetry to generate a continuous family of other solutions.

We saw in Sec. 2.4.2 that a system with an ignorable coordinate always has a conservation equation, giving rise to an integral of motion. This is a special case of a general theorem proved by Emmy Noether in 1918[2] that *any time-independent continuous symmetry of the Lagrangian,* $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ *generates an integral of motion.* This ability of the Lagrangian formalism to generate conservation relations is one of its most important features.

Consider a family $\boldsymbol{q}(t,s)$ obeying the Lagrange equation of motion eq. (2.24), which we write in the form

$$\frac{\partial L}{\partial \boldsymbol{q}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) , \tag{2.79}$$

where $\mathrm{d}/\,\mathrm{d}t$ denotes the total derivative along the path $s = \text{const}$.

Since L is constant over this family of paths, its derivative with respect to s at each point in time vanishes

$$\frac{\partial L}{\partial s} = \frac{\partial L}{\partial \boldsymbol{q}} \cdot \frac{\partial \boldsymbol{q}}{\partial s} + \frac{\partial L}{\partial \dot{\boldsymbol{q}}} \cdot \frac{\partial \dot{\boldsymbol{q}}}{\partial s} . \tag{2.80}$$

Using eq. (2.79) to eliminate $\partial L/\partial \boldsymbol{q}$ in eq. (2.80) and using the time-independence of $\boldsymbol{h}$ to allow interchange of the order of differention with respect to s and t we derive the conservation equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) \cdot \frac{\partial \boldsymbol{q}}{\partial s} + \frac{\partial L}{\partial \dot{\boldsymbol{q}}} \cdot \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \boldsymbol{q}}{\partial s}\right) \equiv \frac{\mathrm{d}I}{\mathrm{d}t} = 0 , \tag{2.81}$$

where

$$I(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \equiv \frac{\partial L}{\partial \dot{\boldsymbol{q}}} \cdot \frac{\partial \boldsymbol{q}}{\partial s} \tag{2.82}$$

is the integral of motion I generated by the symmetry $\boldsymbol{h}$. □

[2]English translation in *Transport Theory and Statistical Physics*, **1**, 186 (1971).

For example, expressing L in polar coordinates we see that the integral of motion generated by rotational symmetry is the angular momentum, eq. (2.35), but eq. (2.82) applies in *any* coordinate system. (Similarly, invariance under translations in the x-direction generates the conservation of linear momentum in the x-direction, and so on.)

2.7.1 Time symmetry

For an *autonomous system* (i.e. one not subject to external forcing), L is independent of *time*. That is, L possesses the symmetry of invariance with respect to translations in t. At first sight this seems very different from the spatial symmetries just discussed, but in fact it is possible to generate an integral of motion (the *energy integral*), using the trick of regarding t as a generalized coordinate by making it a function of a new "time" parameter τ: $t = t(\tau)$, $\boldsymbol{q} = \boldsymbol{q}(\tau)$. (E.g. τ could be the "proper time" in relativity theory.)

In order for the new action integral $\int \bar{L}\,\mathrm{d}\tau$ to be the same as the old one, the new Lagrangian $\bar{L}$ is defined to be

$$\bar{L}\left(\boldsymbol{q}, \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}\tau}, \frac{\mathrm{d}t}{\mathrm{d}\tau}\right) \equiv L\left(\boldsymbol{q}, \frac{1}{\mathrm{d}t/\,\mathrm{d}\tau}\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}\tau}\right)\frac{\mathrm{d}t}{\mathrm{d}\tau}\ . \tag{2.83}$$

Then, using invariance under $t \mapsto t + s$, $\boldsymbol{q} \mapsto \boldsymbol{q}$, eq. (2.82) gives the integral of motion

$$I = L - \dot{\boldsymbol{q}}\cdot\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\ , \tag{2.84}$$

where we have set $t = \tau$ after doing the partial differentiations, as this applies on a physical path. For Lagrangians of the form $T - V$ is easily seen that $I = -T - V$, so I can be identified as the negative of the total energy (cf. Sec. 3.3).

2.8 Problems

2.8.1 Coriolis force and cyclotron motion

As a model of the motion of a fluid element or dust particle in a planetary (e.g. Earth's) atmosphere, consider the motion of particle of unit mass constrained to move on the surface of a perfectly smooth sphere of radius R rotating with angular velocity Ω about the z-axis. Suppose the force on the particle is given by an effective potential $V(\theta, \phi)$ (see Sec. 2.4.6), where θ and ϕ are the latitude and longitude respectively.

(a) Write down the Lagrangian in a frame rotating with the planet, taking the generalized coordinates to be the latitude and longitude so that $z = R\sin\theta$, $x = R\cos\theta\,\cos(\phi+\Omega t)$, $y = R\cos\theta\,\sin(\phi+\Omega t)$, where x, y, z is a non-rotating Cartesian frame. (Draw a diagram illustrating these coordinates, noting that the origin of the θ coordinate is at the equator, not at the pole as in normal spherical polar coordinates.)

Write down the equations of motion, and find a first integral (i.e. constant of the motion) in the case where V is independent of longitude. Among this class of potentials find the special case $V = V_0(\theta)$ required to make equilibrium possible (i.e. so that the equations of motion admit the solution $\dot\theta = \dot\phi = 0$ at each latitude).

Now assume that the *general* form of the potential is $V(\theta,\phi) = V_0(\theta) + h(\theta,\phi)$, where h is the enthalpy (proportional to $p^{(\gamma-1)/\gamma}$) defined in Sec. 2.4.6. Assuming the winds are slow, so that second order time derivatives of θ and ϕ (and products of first order time derivatives) can be neglected, show that the velocity of the particle in the rotating frame is at right angles to the pressure gradient $-\boldsymbol{\nabla}p$. (**Hint:** The pressure gradient is parallel to the enthalpy gradient. You don't have to work out the velocity and gradient vectors in the rotating angular coordinates. It is much simpler to work with the scalar quantity $\mathbf{w}\cdot\boldsymbol{\nabla}p$, where $\mathbf{w}$ is the "wind velocity", i.e. the particle velocity with respect to the rotating frame.)

Draw a sketch of a typical weather-map "low" (i.e. a localized depression in p) showing the pressure contours, the direction of the force on a fluid element and the direction of motion. Hence show that motion in a depression or low is cyclonic, where "cyclonic" in geophysical fluid dynamics means "in the direction of the planet's rotation" (i.e. clockwise in the Southern Hemisphere, counter-clockwise in the Northern Hemisphere for $\Omega > 0$).

(c) Consider a charged particle constrained to move on a *non-rotating* smooth insulating sphere, immersed in a uniform magnetic field $\mathbf{B} = B\mathbf{e}_z$, on which the electrostatic potential is a function of latitude and longitude. Write down the Lagrangian in the same generalized coordinates as above and show it is the same as that for the particle on the rotating planet with appropriate identifications of Ω and V.

2.8.2 Anharmonic oscillator

Consider the following potential V, corresponding to a particle of mass m oscillating along the x-axis under the influence of a nonideal spring (i.e. one

with a nonlinear restoring force),

$$V(x) = \frac{m\omega_0^2}{2}\left(x^2 + \sigma\frac{x^4}{l_0^2}\right) ,$$

where the constant ω_0 is the angular frequency of oscillations having amplitude small compared with the characteristic length l_0, and $\sigma = \pm 1$ depends on whether the spring is "soft" ($\sigma = -1$) or "hard" ($\sigma = +1$).

Consider the trial function

$$x = l_0\left[A\cos\omega t + B\cos 3\omega t + C\sin 3\omega t\right] ,$$

where A, B, C are the nondimensionalized amplitudes of the fundamental and third harmonic, respectively, and ω is the nonlinearly shifted frequency. By using this trial function in the time-averaged Hamilton's Principle, find implicit relations giving approximate expressions for ω, B and C as functions of A. Show that $C \equiv 0$. The trial function is strictly appropriate only to the case $A \ll 1$, but plot ω/ω_0 and B *vs.* A from 0 to 1 in the case of both a hard and a soft spring. (You are encouraged to use Maple or Mathematica and/or MatLab in this problem.)

Chapter 3

Hamiltonian Mechanics

3.1 Introduction: Dynamical systems

Mathematically, a *continuous-time dynamical system* is defined to be a system of first order differential equations

$$\dot{\boldsymbol{z}} = \boldsymbol{f}(\boldsymbol{z}, t) \;, \quad t \in \mathbb{R} \;, \tag{3.1}$$

where $\boldsymbol{f}$ is known as the *vector field* and $\mathbb{R}$ is the set of real numbers. The space in which the set of time-dependent variables $\boldsymbol{z}$ is defined is called *phase space*.

Sometimes we also talk about a *discrete-time dynamical system*. This is a recursion relation, difference equation or *iterated map*

$$\boldsymbol{z}_{t+1} = \boldsymbol{f}(\boldsymbol{z}, t) \;, \quad t \in \mathbb{Z} \;, \tag{3.2}$$

where $\boldsymbol{f}$ is known as the map (or mapping) and $\mathbb{Z}$ is the set of all integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

Typically, such systems exhibit long-time phenomena like attracting and repelling fixed points and limit cycles, or more complex structures such as strange attractors. In this chapter we show how to reformulate nondissipative Lagrangian mechanics as a dynamical system, but shall find that it is a very special case where the above-mentioned phenomena cannot occur.

3.2 Mechanics as a dynamical system

3.2.1 Lagrangian method

Lagrange's equations do not form a dynamical system, because they implicitly contain second-order derivatives, $\ddot{\boldsymbol{q}}$. However, there is a standard way to

obtain a system of first-order equations from a second-order system, which is to double the size of the space of time-dependent variables by treating the generalized velocities $\boldsymbol{u}$ as independent of the generalized coordinates, so that the dynamical system is $\dot{\boldsymbol{q}} = \boldsymbol{u}$, $\dot{\boldsymbol{u}} = \ddot{\boldsymbol{q}}(\boldsymbol{q}, \boldsymbol{u}, t)$. Then the phase space is of dimension $2n$. This trick is used very frequently in numerical problems, because the standard numerical integrators require the problem to be posed in terms of systems of first-order differential equations.

In the particular case of Lagrangian mechanics, we have from eq. (2.24), expanding out the total derivative using the chain rule and moving all but the highest-order time derivatives to the right-hand side,

$$\sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \ddot{q}_j = \frac{\partial L}{\partial q_i} - \frac{\partial^2 L}{\partial \dot{q}_i \partial t} - \sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j} \dot{q}_j \ . \tag{3.3}$$

The matrix H acting on $\ddot{\boldsymbol{q}}$, whose elements are given by

$$H_{i,j} \equiv \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \ , \tag{3.4}$$

is called the *Hessian matrix*. It is a kind of generalized mass tensor (see Sec. 3.2.6), and for our method to work we require it to be *nonsingular*, so that its inverse, H^{-1}, exists and we can find $\ddot{\boldsymbol{q}}$. Then our dynamical system becomes

$$\begin{aligned} \dot{\boldsymbol{q}} &= \boldsymbol{u} \ , \\ \dot{\boldsymbol{u}} &= \mathsf{H}^{-1} \boldsymbol{\cdot} \left[\frac{\partial L}{\partial \boldsymbol{q}} - \frac{\partial^2 L}{\partial \dot{\boldsymbol{q}} \partial t} - \frac{\partial^2 L}{\partial \dot{\boldsymbol{q}} \partial \boldsymbol{q}} \boldsymbol{\cdot} \dot{\boldsymbol{q}} \right] \ . \end{aligned} \tag{3.5}$$

Remark 3.1 *The Lagrangian method* per se *does not break down if the Hessian is singular, only our attempt to force it into the standard dynamical system framework. We can still formally solve the dynamics in the following manner: Suppose* H *is singular, with rank* $n - m$*. Then, within the* n*-dimensional linear vector space* $\mathcal{N}$ *on which* H *acts, there is an* m*-dimensional subspace* $\mathcal{M}$ *(the* nullspace*) such that* $\mathsf{H} \boldsymbol{\cdot} \ddot{\boldsymbol{q}} \equiv 0$ *for all* $\ddot{\boldsymbol{q}} \in \mathcal{M}$*. We can solve eq. (3.3) for the component of* $\ddot{\boldsymbol{q}}$ *lying in the complementary subspace* $\mathcal{N} \setminus \mathcal{M}$ *provided the right-hand side satisfies the* solubility condition *that it have no component in* $\mathcal{M}$*. The component of* $\ddot{\boldsymbol{q}}$ *lying in* $\mathcal{M}$ *cannot be found directly, but the solubility condition provides* m constraints *that complete the determination of the dynamics.*

As a simple example, suppose L does not depend on one of the generalized velocities, $\dot{q}_i$. Then $\partial L / \partial \dot{q}_i \equiv 0$ and the ith component of the Lagrange equations of motion eq. (2.24) reduces to the constraint $\partial L / \partial q_i = 0$. □

3.2.2 Hamiltonian method

We can achieve our aim of finding $2n$ first-order differential equations by using many choices of auxiliary variables other than $\boldsymbol{u}$. These will be more complicated functions of the generalized velocities, but the extra freedom of choice may also bring advantages.

In particular, Hamilton realised that it is very natural to use as the new auxiliary variables the set $\boldsymbol{p} = \{p_i | i = 1, \ldots, n\}$ defined by

$$p_i \equiv \frac{\partial}{\partial \dot{q}_i} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \ , \tag{3.6}$$

where p_i is called the *generalized momentum canonically conjugate to* q_i.

We shall always assume that eq. (3.6) can be solved to give $\dot{\boldsymbol{q}}$ as a function of $\boldsymbol{q}$ and $\boldsymbol{p}$

$$\dot{\boldsymbol{q}} = \boldsymbol{u}(\boldsymbol{q}, \boldsymbol{p}, t) \ . \tag{3.7}$$

Remark 3.2 *We have in effect changed variables from* $\boldsymbol{u}$ *to* $\boldsymbol{p}$*, and such a change of variables can only be performed if the Jacobian matrix* $\partial p_i / \partial u_j = \partial^2 L / \partial \dot{q}_i \partial \dot{q}_j$ *is nonsingular. From eq. (3.4) we recognize this matrix as being the Hessian we encountered in the Lagrangian approach to constructing a dynamical system. Thus in either approach we require the Hessian to be nonsingular (i.e. for its determinant to be nonzero). This condition is usually trivially satisfied, but there are physical problems (e.g. if the Lagrangian does not depend on one of the generalized velocities) when this is not the case. However, in the standard Hamiltonian theory covered in this course it is always assumed to hold.*

The reason for defining $\boldsymbol{p}$ as in eq. (3.6) is that $\partial L / \partial \dot{\boldsymbol{q}}$ occurs *explicitly* in Lagrange's equations, eq. (2.24), so we immediately get an equation of motion for $\boldsymbol{p}$

$$\dot{\boldsymbol{p}} = \left. \frac{\partial L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)}{\partial \boldsymbol{q}} \right|_{\dot{q} = u(q,p,t)} \ . \tag{3.8}$$

Equations (3.7) and (3.8) do indeed form a dynamical system, but so far it looks rather unsatisfactory: now $\boldsymbol{u}$ is defined only implicitly as a function of the phase-space variables $\boldsymbol{q}$ and $\boldsymbol{p}$, yet the right-hand side of eq. (3.8) involves a partial derivative in which the $\boldsymbol{q}$-dependence of $\boldsymbol{u}$ is ignored!

We can fix the latter problem by holding $\boldsymbol{p}$ fixed in partial derivatives with respect to $\boldsymbol{q}$ (because it is an independent phase-space variable) but then subtracting a correction term to cancel the contribution coming from

the $\boldsymbol{q}$-dependence of $\boldsymbol{u}$. Applying the chain rule and then using eqs. (3.6) and (3.7) we get

$$\begin{aligned}
\dot{\boldsymbol{p}} &= \frac{\partial L(\boldsymbol{q},\boldsymbol{u},t)}{\partial \boldsymbol{q}} - \sum_{i=1}^{n} \frac{\partial L}{\partial u_i}\frac{\partial u_i}{\partial \boldsymbol{q}} \\
&= \frac{\partial L(\boldsymbol{q},\boldsymbol{u},t)}{\partial \boldsymbol{q}} - \sum_{i=1}^{n} p_i \frac{\partial u_i}{\partial \boldsymbol{q}} \\
&= \frac{\partial}{\partial \boldsymbol{q}}[L(\boldsymbol{q},\boldsymbol{u},t) - \boldsymbol{p}\cdot\boldsymbol{u}] \\
&\equiv -\frac{\partial H}{\partial \boldsymbol{q}} \ , \qquad (3.9)
\end{aligned}$$

where we have defined a new function to replace the Lagrangian, namely the *Hamiltonian*

$$H(\boldsymbol{q},\boldsymbol{p},t) \equiv \boldsymbol{p}\cdot\boldsymbol{u} - L(\boldsymbol{q},\boldsymbol{u},t) \ . \qquad (3.10)$$

Given the importance of $\partial H/\partial \boldsymbol{q}$ it is natural to investigate whether $\partial H/\partial \boldsymbol{p}$ plays a significant role as well. Differentiating eq. (3.10) we get

$$\begin{aligned}
\frac{\partial H}{\partial \boldsymbol{p}} &= \boldsymbol{u}(\boldsymbol{q},\boldsymbol{p},t) + \sum_{i=1}^{n}\left[p_i - \frac{\partial}{\partial u_i}L(\boldsymbol{q},\boldsymbol{u},t)\right]\frac{\partial u_i}{\partial \boldsymbol{p}} \\
&= \dot{\boldsymbol{q}} \ , \qquad (3.11)
\end{aligned}$$

where the vanishing of the expression in the square bracket and the identification of $\boldsymbol{u}$ with $\dot{\boldsymbol{q}}$ follows from eqs. (3.6) and (3.7).

Summarizing eqs. (3.9) and (3.11),

$$\begin{aligned}
\dot{\boldsymbol{q}} &= \frac{\partial H}{\partial \boldsymbol{p}} \\
\dot{\boldsymbol{p}} &= -\frac{\partial H}{\partial \boldsymbol{q}} \ , \qquad (3.12)
\end{aligned}$$

These equations are known as *Hamilton's equations of motion*. As with Lagrange's equations they express the dynamics of a system with an arbitrary number of degrees of freedom in terms of a single scalar function! Unlike the Lagrangian dynamical system, the phase-space variables are treated on a completely even footing—in Hamiltonian mechanics both the configuration-space variables $\boldsymbol{q}$ and momentum-space variables $\boldsymbol{p}$ are generalized coordinates. We define *canonical coordinates* as phase-space coordinates such that the equations of motion can be expressed in the form of eq. (3.12) and a *canonical system* as one for which canonical coordinates exist.

Remark 3.3 *The transition from the Lagrangian to the Hamiltonian in order to handle the changed meaning of partial derivatives after a change of variable is a special case of a technique known as a* Legendre transformation. *It is encountered quite often in physics and physical chemistry, especially in thermodynamics.*

3.2.3 Example 1: Scalar potential

Consider the Lagrangian for a particle in Cartesian coordinates, so $\boldsymbol{q} = \{x, y, z\}$ may be replaced by $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$. Also assume that it moves under the influence of a scalar potential $V(\mathbf{r}, t)$ so that the natural form of the Lagrangian is

$$L = T - V = \tfrac{1}{2}m|\dot{\mathbf{r}}|^2 - V(\mathbf{r}, t) \; . \tag{3.13}$$

Then, from eq. (3.6)

$$\boldsymbol{p} \to \mathbf{p} \equiv \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} \; , \tag{3.14}$$

so that in this case the canonical momentum is the same as the ordinary kinematic momentum. Equation (3.14) is solved trivially to give $\dot{\boldsymbol{q}} = \mathbf{u}(\mathbf{p})$ where $\mathbf{u}(\mathbf{p}) = \mathbf{p}/m$. Thus, from eq. (3.10) we have

$$\begin{aligned} H &= \frac{|\mathbf{p}|^2}{m} - \left(\frac{|\mathbf{p}|^2}{2m} - V(\mathbf{r}, t) \right) \\ &= \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{r}, t) \\ &= T + V \; . \end{aligned} \tag{3.15}$$

That is, the Hamiltonian is equal to the *total energy* of the system, kinetic plus potential. The fact that the Hamiltonian is an important physical quantity, whereas the physical meaning of the Lagrangian is more obscure, is one of the appealing features of the Hamiltonian approach. Both the Lagrangian and Hamiltonian have the dimensions of energy, and both approaches can be called *energy methods*. They are characterized by the use of scalar quantities rather than the vectors encountered in the direct use of Newton's second law. This has both the theoretical advantage of leading to very general formulations of mechanics and the practical benefit of avoiding some vector manipulations when changing between coordinate systems (in fact, Lagrangian and Hamiltonian methods were developed before modern vector notation was invented).

Harmonic Oscillator

An example is the harmonic oscillator Hamiltonian corresponding to the Lagrangian, eq. (2.61)

$$H = \frac{p^2}{2m} + \frac{m\omega_0 x^2}{2} \ . \tag{3.16}$$

From eq. (3.12) the Hamiltonian equations of motion are

$$\begin{aligned} \dot{x} &= \frac{p}{m} \\ \dot{p} &= -m\omega_0 x \ . \end{aligned} \tag{3.17}$$

Gauge-transformed Harmonic Oscillator

Now consider the gauge-transformed harmonic oscillator Lagrangian eq. (2.75)

$$L' = \tfrac{1}{2}m(\dot{x}^2 + 2\omega_0 x\dot{x} - \omega_0^2 x^2) \ .$$

The canonical momentum is thus

$$p = \partial L'/\partial\dot{x} = m(\dot{x} + \omega_0 x) \tag{3.18}$$

and we see that the gauge transformation has effected a transformation of the canonical momentum, even though the generalized coordinate x remains the same. This is an example of a *canonical transformation*, about which we will have more to say later.

Solving eq. (3.18) for $\dot{x}$ we find $u(p) = (p - m\omega_0 x)/m$. Hence

$$\begin{aligned} H &= p\frac{(p - m\omega_0 x)}{m} - \left(\frac{(p - m\omega_0 x)^2}{2m} + \omega_0 x(p - m\omega_0 x) - \tfrac{1}{2}m\omega_0^2 x^2\right) \\ &= \frac{(p - m\omega_0 x)^2}{2m} + \tfrac{1}{2}m\omega_0^2 x^2 \\ &= T + V \ . \end{aligned} \tag{3.19}$$

Thus, even though L was not of the natural form $T - V$ in this case, the Hamiltonian remains equal to the total energy, thus confirming that it is a quantity with a more direct physical signifance than the Lagrangian. (The *functional form* of the Hamiltonian changes under the gauge transformation however, because the meaning of p changes.)

3.2.4 Example 2: Physical pendulum

A nonlinear one-dimensional case is provided by the physical pendulum, introduced in Sec. 2.4.1. Using eq. (2.27) in eq. (3.6) we get $p_\theta = ml^2\dot{\theta}$. Thus $\dot{\theta} = p_\theta/ml^2$ and so the Hamiltonian, $H = p_\theta\dot{\theta} - L$, becomes

$$H(\theta, p_\theta) = \frac{p_\theta^2}{2ml^2} + mgl(1 - \cos\theta) \ , \tag{3.20}$$

which again is of the form $T + V$.

By conservation of energy (see also Sec. 3.3), the Hamiltonian $H = T + V$ is a constant of the motion so the nature of the orbits in phase space can be found simply by plotting the contours of H as in Fig. 3.1 (which is in units such that $m = l = g = 1$). We see that the structure of the phase space is rather more complicated than in the case of the harmonic oscillator since there are two topologically distinct classes of orbit. One class is the *rotating* orbits for which the pendulum has enough energy, $H > 2mgl$, to swing entirely over the top so θ increases or decreases secularly (though the physical position $x = l\sin\theta$ does not).

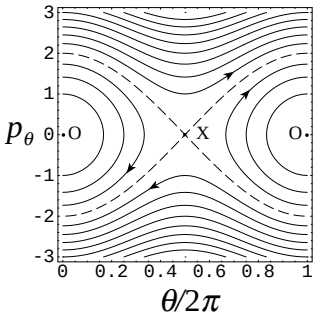

Figure 3.1: Phase space of the physical pendulum.

The other class is the *librating* orbits for $H < 2mgl$ implying that the pendulum is trapped in the gravitational potential well and oscillates like a pair of scales (hence the name). For $|\theta| \ll 1$ we may expand the cosine so $V \approx mgl\theta^2/2$ and the system is approximately a harmonic oscillator. The equilibrium point $\theta = 0$ or 2π, labelled O in Fig. 3.1, is a *fixed point*. The orbits in its neighbourhood, like those of the harmonic oscillator, remain in the neighbourhood for all time (i.e. the fixed point is *stable* or *elliptic*).

The dividing line $H = 2mgl$ between the two topological classes of orbit is called the *separatrix*, and on the separatrix lies another fixed point, labelled X in Fig. 3.1. This corresponds to the case where the pendulum is just balanced "upside down". Almost all orbits in the neighbourhood of an X point eventually are repelled far away from it, and thus it is referred to as an *unstable* or *hyperbolic* fixed point.

3.2.5 Example 3: Motion in e.m. potentials

Now consider the case of a charged particle in an electromagnetic field with magnetic vector potential $\mathbf{A}$ and electrostatic potential Φ. The Lagrangian is given by eq. (2.44) and thus, from eq. (3.6),

$$\mathbf{p} \equiv \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} + e\mathbf{A}(\mathbf{r}, t) \; . \tag{3.21}$$

Thus $\mathbf{u}(\mathbf{p}) = \mathbf{p}/m - e\mathbf{A}/m$ and, from eq. (3.10) we have

$$\begin{aligned} H &= \frac{(\mathbf{p} - e\mathbf{A})\cdot\mathbf{p}}{m} - \left(\frac{|\mathbf{p} - e\mathbf{A}|^2}{2m} + \frac{e(\mathbf{p} - e\mathbf{A})\cdot\mathbf{A}}{m} - e\Phi(\mathbf{r}, t) \right) \\ &= \frac{|\mathbf{p} - e\mathbf{A}(\mathbf{r}, t)|^2}{2m} + e\Phi(\mathbf{r}, t) \\ &= T + V \; . \end{aligned} \tag{3.22}$$

Thus we find again that, although the Lagrangian cannot be put into the natural form $T - V(\mathbf{r})$, the Hamiltonian is still the total energy, kinetic plus electrostatic potential energy.

3.2.6 Example 4: The generalized N-body system

Let us now revisit the general case which led to the original definition of the Lagrangian in Sec. 2.3, the case of N particles interacting via a scalar N-body potential, possibly with constraints (which we here assume to be independent of time), so that the number of generalized coordinates is $n \leq 3N$. Then, by eq. (2.25) and eq. (2.7) (assuming the function $\mathbf{x}_k(\boldsymbol{q})$ to be independent of time) the natural form of the Lagrangian is

$$\begin{aligned} L &= T - V \\ &= \tfrac{1}{2} \sum_{i,j=1}^{n} \sum_{k=1}^{N} m_k \dot{q}_i \frac{\partial \mathbf{x}_k}{\partial q_i} \cdot \frac{\partial \mathbf{x}_k}{\partial q_j} \dot{q}_j - V(\boldsymbol{q}) \\ &= \tfrac{1}{2} \sum_{i,j=1}^{n} \dot{q}_i \mu_{i,j} \dot{q}_j - V(\boldsymbol{q}) \; , \end{aligned} \tag{3.23}$$

where the symmetric *mass matrix*

$$\mu_{i,j}(\boldsymbol{q}) \equiv \sum_{k=1}^{N} m_k \frac{\partial \mathbf{x}_k}{\partial q_i} \cdot \frac{\partial \mathbf{x}_k}{\partial q_j} \tag{3.24}$$

is the metric tensor for a configuration-space mass-weighted "length" element $\mathrm{d}s$ defined by $(\mathrm{d}s)^2 \equiv \sum_k m_k \,\mathrm{d}\mathbf{l}_k^2$. From eq. (3.4), we see that it is the Hessian matrix for this system.

Then, using eq. (3.23) in eq. (3.6)

$$p_i = \sum_{j=1}^{n} \mu_{i,j}(\boldsymbol{q})\dot{q}_j \ . \tag{3.25}$$

Assuming none of the particles is massless, $\mu_{i,j}$ is a positive-definite matrix, so its inverse $\mu_{i,j}^{-1}$ exists and we can formally solve eq. (3.25) for $\dot{q}_i$ to give

$$u_i(\boldsymbol{q},\boldsymbol{p},t) = \sum_{j=1}^{n} \mu_{i,j}^{-1} p_j \ . \tag{3.26}$$

Then, from eq. (3.10) we have

$$\begin{aligned} H &= \sum_{i,j=1}^{n} p_i \mu_{i,j}^{-1} p_j - \left(\tfrac{1}{2} \sum_{i,j,i',j'=1}^{n} p_i \mu_{i,i'}^{-1} \mu_{i',j'} \mu_{j',j}^{-1} p_j - V(\boldsymbol{q}) \right) \\ &= \tfrac{1}{2} \sum_{i,j=1}^{n} p_i \mu_{i,j}^{-1} p_j + V(\boldsymbol{q}) \\ &= T + V \ . \end{aligned} \tag{3.27}$$

Thus, the Hamiltonian is again equal to the total energy of the system. This result does *not* hold in the case of a time-dependent representation, $\mathbf{x}_k(\boldsymbol{q},t)$. [See Problems 3.5.1(c) and 4.7.1.]

Particle in a central potential

As a simple, two-dimensional example of a problem in non-Cartesion coordinates we return to the problem of motion in a central potential, expressed in plane polar coordinates in Sec. 2.4.2. Recapitulating eq. (2.33),

$$L = \tfrac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right) - V(r) \ .$$

Comparing with eq. (3.23) we see that the mass matrix is diagonal

$$\boldsymbol{\mu} = \begin{bmatrix} m & 0 \\ 0 & mr^2 \end{bmatrix} \ , \tag{3.28}$$

and thus can be inverted simply taking the reciprocal of the diagonal elements. Hence, from eq. (3.27)

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r) \ . \tag{3.29}$$

3.3 Time-Dependent and Autonomous Hamiltonian systems

An *autonomous* dynamical system is one in which the vector field $\boldsymbol{f}$ [see eq. (3.1)] depends only on the phase-space coordinates and has no explicit dependence on t. (The reason for the name is that explicit time dependence would come from external forcing of the system, whereas "autonomous" means independent or self-governing.) In the Hamiltonian case this means that H has no explicit time dependence. Conversely, if there *is* an external time-varying perturbation, then $H = H(\boldsymbol{p}, \boldsymbol{q}, t)$.

Consider the time rate of change of the Hamiltonian, $\dot{H} \equiv \mathrm{d}H/\,\mathrm{d}t$, following the phase-space trajectory. Using the Hamiltonian equations of motion, eq. (3.12),

$$\begin{aligned} \frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\partial H}{\partial t} + \dot{\boldsymbol{q}}\cdot\frac{\partial H}{\partial \boldsymbol{q}} + \dot{\boldsymbol{p}}\cdot\frac{\partial H}{\partial \boldsymbol{p}} \\ &= \frac{\partial H}{\partial t} + \frac{\partial H}{\partial \boldsymbol{p}}\cdot\frac{\partial H}{\partial \boldsymbol{q}} - \frac{\partial H}{\partial \boldsymbol{q}}\cdot\frac{\partial H}{\partial \boldsymbol{p}} \\ &\equiv \frac{\partial H}{\partial t}\ . \end{aligned} \tag{3.30}$$

Thus, if H does not depend explicitly on time, so its partial time derivative vanishes, then its total time derivative also vanishes. That is, *in an autonomous Hamiltonian system the Hamiltonian is an integral of motion.*

We have seen above that H can often be identified as the total energy. When this is the case we may interpret the completely general result above as a statement of *conservation of energy.* More generally, comparing eq. (3.10) and eq. (2.84) we recognize H as the negative of the energy integral I for autonomous systems as predicted by Noether's theorem.

3.4 Hamilton's Principle in phase space

Let us express the action integral S in terms of $\boldsymbol{q}$ and $\boldsymbol{p}$ and show that we may derive Hamilton's equations of motion by making S stationary under variations of the trajectory in phase space, rather than in configuration space.

To find a suitable definition for the action integral, we first rearrange eq. (3.10) to get L on the left-hand side. Then we replace $\boldsymbol{u}(\boldsymbol{q}, \boldsymbol{p}, t)$ by $\dot{\boldsymbol{q}}$ in the term $\boldsymbol{p}\cdot\boldsymbol{u}$ and thus define the *phase-space Lagrangian*

$$L_{\mathrm{ph}}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{p}, t) \equiv \boldsymbol{p}\cdot\dot{\boldsymbol{q}} - H(\boldsymbol{q}, \boldsymbol{p}, t)\ , \tag{3.31}$$

If $\dot{\boldsymbol{q}} = \boldsymbol{u}(\boldsymbol{q}, \boldsymbol{p}, t)$ were identically satisfied, even on arbitrarily varied phase-space paths, then $L_{\rm ph}$ would simply be L expressed in phase-space coordinates. However, one can easily construct a counter example to show that this is not the case: consider a variation of the path in which we can vary the direction of its tangent vector, at some point $(\boldsymbol{q}, \boldsymbol{p})$, while keeping this point fixed (see Sec. 4.2 for an illustration of this). Then $\dot{\boldsymbol{q}}$ changes, but $\boldsymbol{u}$ remains the same. Thus $L_{\rm ph}$ and L are the same value only on the subset of paths (which includes the physical paths) for which $\boldsymbol{p}\cdot\dot{\boldsymbol{q}} = \boldsymbol{p}\cdot\boldsymbol{u}(\boldsymbol{q}, \boldsymbol{p}, t)$.

Replacing L by $L_{\rm ph}$ in eq. (2.26) we define the *phase-space action integral*

$$S_{\rm ph}[\boldsymbol{q}, \boldsymbol{p}] = \int_{t_1}^{t_2} \mathrm{d}t \; L_{\rm ph}(\boldsymbol{q}, \boldsymbol{p}, \dot{\boldsymbol{q}}, t) \; . \tag{3.32}$$

We shall now show that requiring $\delta S_{\rm ph} = 0$ for arbitrary variations about a given path implies that that path is such that Hamilton's equations of motion are satisfied at all points along it. Since we concluded above that $L_{\rm ph}$ and L were not the same on arbitrary nonphysical paths, this variational principle is subtly different from the Lagrangian version of Hamilton's principle, and is sometimes called the *modifed Hamilton's Principle*.

We know from variational calculus that $S_{\rm ph}$ is stationary under arbitrary variations of the phase-space path (with endpoints fixed) if and only if the Euler–Lagrange equations $\delta L_{\rm ph}/\delta \boldsymbol{q} = \delta L_{\rm ph}/\delta \boldsymbol{p} = 0$ are satisfied at all points on the path, where the variational derivatives are as defined in eq. (1.14), with $\boldsymbol{p}$ now being regarded as an independent coordinate

$$\begin{aligned} \frac{\delta L_{\rm ph}}{\delta \boldsymbol{q}} = \frac{\partial L_{\rm ph}}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_{\rm ph}}{\partial \dot{\boldsymbol{q}}} &= -\frac{\partial H}{\partial \boldsymbol{q}} - \dot{\boldsymbol{p}} = 0 \; , \\ \frac{\delta L_{\rm ph}}{\delta \boldsymbol{p}} = \frac{\partial L_{\rm ph}}{\partial \boldsymbol{p}} &= \dot{\boldsymbol{q}} - \frac{\partial H}{\partial \boldsymbol{p}} = 0 \; . \end{aligned} \tag{3.33}$$

These are identical to Hamilton's equations of motion, eq. (3.12), as claimed.

Remark 3.4 *Note that, although we fix the endpoints in the modified Hamilton's Principle, we will not be able to satisfy Hamilton's equations of motion at* all *points on the path unless the endpoints are both chosen to lie on a physical path. This is because the fixing of the initial phase-space point specifies* $2n$ *initial conditions, which completely determines the physical path (i.e. the one that obeys the equations of motion). This is in contrast with the the Lagrangian version of Hamilton's Principle, where we fixed only the* n *configuration-space coordinates* $\boldsymbol{q}$*. The problem of finding the physical path between two given configuration-space points thus involves specifying only* n *initial and* n *final data, which is generally exactly the right number to allow*

us to find one (or more) solutions of the equations of motion passing through arbitrarily chosen initial and final points.

3.5 Problems

3.5.1 Constraints and moving coordinates

(a) This revisits Problem 1.6.1. Write down a Lagrangian for the problem of two particles of mass m_1 and m_2 connected by a light rigid rod of length l in a gravitational field g. Take the generalized coordinates of the system to be $\boldsymbol{q} = \{x, y, z, \theta, \phi\}$, with the coordinates of the two particles being given by

$$\begin{aligned} x_1 &= x - \alpha_1 l \sin\theta \cos\phi \\ y_1 &= y - \alpha_1 l \sin\theta \sin\phi \\ z_1 &= z - \alpha_1 l \cos\theta \\ x_2 &= x + \alpha_2 l \sin\theta \cos\phi \\ y_2 &= y + \alpha_2 l \sin\theta \sin\phi \\ z_2 &= z + \alpha_2 l \cos\theta \ , \end{aligned}$$

where $\alpha_1 \equiv m_2/(m_1 + m_2)$ and $\alpha_2 \equiv m_1/(m_1 + m_2)$ [so that (x, y, z) is the centre of mass].

(b) Use this Lagrangian to construct the Hamiltonian for the system. Is it of the form $T + V$?

(c) Find the Hamiltonian corresponding to the Coriolis Lagrangian in Problem 2.8.1,

$$L = \tfrac{1}{2} R^2 \left[\dot{\theta}^2 + \cos^2\theta \, (\Omega + \dot{\phi})^2\right] - V(\theta, \phi) \ .$$

Is it of the form $T + V$?

3.5.2 Anharmonic oscillator phase space

Find the Hamiltonian corresponding to the Lagrangian of the anharmonic oscillator in Problem 2.8.2. Sketch contours of $H(x, p)$ in the soft potential case for (a) $H < m\omega_0^2 l_0^2/8$ and (b) $H > m\omega_0^2 l_0^2/8$. Indicate representative trajectories of the system in phase-space, showing the direction of travel. Try to include representatives of each qualitatively distinct orbit.

3.5.3 2-D motion in a magnetic field

Consider the motion of a particle of charge e and mass m in a straight, infinitely long magnetic confinement system with vector potential $\mathbf{A} = \psi(x, y)\mathbf{e}_z$.

(a) Show that $B_x = \partial\psi/\partial y$, $B_y = -\partial\psi/\partial x$, $B_z = 0$ and thus that contours of $\psi(x, y)$ define magnetic field lines.

(b) Show that the Hamiltonian is

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{(p_z - e\psi)^2}{2m} \ ,$$

write down the Hamiltonian equations of motion and give two integrals of the motion.

(c) It can be shown that in any region with no electric current, $\psi(x, y)$ can be represented as the real part of an analytic function of the complex variable $\zeta \equiv x + \mathrm{i}y$, $\psi(x, y) = \mathrm{Re}\,\Psi(\zeta)$. Show that the equations of motion for ζ and $p_\zeta \equiv p_x + \mathrm{i}p_y$ are

$$\begin{aligned} \dot{\zeta} &= \frac{1}{m} p_\zeta \ , \\ \dot{p}_\zeta &= \frac{e}{m} \left[p_z - e\mathrm{Re}\,\Psi(\zeta) \right] \left[\Psi'(\zeta) \right]^* \ , \end{aligned}$$

where the prime on Ψ means derivative with respect to its argument, and $*$ means complex conjugate.

Chapter 4

Canonical transformations

4.1 Introduction

A *canonical transformation* is a transformation of the phase-space coordinates $\boldsymbol{q}$ and $\boldsymbol{p}$ to new coordinates $\boldsymbol{Q}$ and $\boldsymbol{P}$ that obey Hamilton's equations of motion, eq. (3.12), with a new Hamiltonian $K(\boldsymbol{Q}, \boldsymbol{P}, t)$.

The discussion in Sec. 3.2.6 shows the general form of $\boldsymbol{p}$ for an N-particle system in an arbitrary, stationary generalized coordinate system. In this case, if we change from one configuration-space coordinate system, $\boldsymbol{q}$, to another, $\boldsymbol{Q}$, then the transformation of $\boldsymbol{p}$ to $\boldsymbol{P}$ follows from comparing eq. (3.25) in the two coordinate systems. For instance, if we go from coordinates x, y in a two-dimensional system to polar coordinates r, θ, the canonical transformation is

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ p_x &= m\dot{x} = m(\dot{r}\cos\theta - r\dot{\theta}\sin\theta) = p_r\cos\theta - \frac{p_\theta \sin\theta}{r} \\ p_y &= m\dot{y} = m(\dot{r}\sin\theta + r\dot{\theta}\cos\theta) = p_r\sin\theta + \frac{p_\theta \cos\theta}{r} \\ K(r,\theta,p_r,p_\theta) &= H(x,y,p_x,p_y)\ . \end{aligned} \tag{4.1}$$

Thus one obvious class of canonical transformation is produced by a change of configuration-space coordinates, or point transformation, $\boldsymbol{q} = \boldsymbol{g}(\boldsymbol{Q}, t)$ as discussed in the context of the general Lagrangian formalism in Sec. 2.6. Applying eq. (2.68) and using eq. (2.69) we find the general relationship

$$P_i = \frac{\partial}{\partial \dot{Q}_i} L(\boldsymbol{g}(\boldsymbol{Q}, t), \dot{\boldsymbol{g}}(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t), t)$$

$$= \boldsymbol{p}\cdot\frac{\partial\dot{\boldsymbol{g}}}{\partial\dot{Q}_i}$$
$$= \sum_{j=1}^{n}\frac{\partial g_j}{\partial Q_i}p_j \ . \tag{4.2}$$

Thus by inverting the matrix $T_{i,j} \equiv \partial g_j/\partial Q_i$, we complete the construction of the canonical transformation by finding $\boldsymbol{p} = \mathsf{T}^{-1}\cdot\boldsymbol{P}$. We can also show by direct calculation that $K(\boldsymbol{Q},\boldsymbol{P}) = H(\boldsymbol{q},\boldsymbol{p})$ for time-independent transformations as in eq. (4.1). The discussion of the time dependent case will be left to the next section.

We have also encountered another special case of a canonical transformation (for which $\boldsymbol{p}$ was changed but $\boldsymbol{q}$ remained the same) in Sec. 3.2.3 where we compared the Hamiltonian for the harmonic oscillator derived from the natural form for the Lagrangian with that from a gauge-transformed version. Likewise, in the case of the particle in an electromagnetic field treated in Sec. 3.2.5 an electromagnetic gauge transformation effects a change in the momenta but leaves the configuration-space coordinates unaffected.

This suggests that gauge transformations may hold the clue to constructing a general theory of canonical transformations and we show in the next section that this is indeed the case.

4.2 Generating functions

To discuss gauge transformations in the Hamiltonian context we use the modified Hamilton's Principle based on the phase-space action integral eq. (3.32),

$$S_{\rm ph}[\boldsymbol{q},\boldsymbol{p}] = \int_{t_1}^{t_2} \mathrm{d}t\ L_{\rm ph}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{p},\dot{\boldsymbol{p}},t)\ , \tag{4.3}$$

which we showed to be stationary on physical paths with the form of phase-space Lagrangian given in eq. (3.31),

$$L^{(1)}_{\rm ph} \equiv \boldsymbol{p}\cdot\dot{\boldsymbol{q}} - H(\boldsymbol{q},\boldsymbol{p},t)\ . \tag{4.4}$$

However, as in Sec. 2.6.2, we can find a physically equivalent Lagrangian by adding to $L^{(1)}_{\rm ph}$ the total time derivative of an arbitrary function, $M(\boldsymbol{q},\boldsymbol{p},t)$ say. Although it clearly has no effect on Hamilton's equations of motion, and thus does not generate a canonical transformation, we shall have use of the particular choice $M = -\boldsymbol{q}\cdot\boldsymbol{p}$. This gives the alternative, gauge-transformed phase-space Lagrangian

$$L^{(2)}_{\rm ph} \equiv -\boldsymbol{q}\cdot\dot{\boldsymbol{p}} - H(\boldsymbol{q},\boldsymbol{p},t)\ . \tag{4.5}$$

Now assume we have a phase-space transformation $(\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{Q}, \boldsymbol{P})$ where the mapping between the two coordinates is invertible and sufficiently smooth that we can evaluate the required derivatives. Since we are seeking a *canonical* transformation, there must also exist an action integral $S'_{\rm ph}$ of the form

$$S'_{\rm ph}[\boldsymbol{Q}, \boldsymbol{P}] = \int_{t_1}^{t_2} \mathrm{d}t \; L'_{\rm ph}(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, \boldsymbol{P}, \dot{\boldsymbol{P}}, t) \tag{4.6}$$

that is stationary under variations vanishing at the endpoints for the same paths (represented in the new coordinates) as $S_{\rm ph}$ is, where $L'_{\rm ph}$ is a phase-space Lagrangian in either of the forms given in eqs. (4.4) or (4.5) with $\boldsymbol{q}, \boldsymbol{p}$ replaced by $\boldsymbol{Q}, \boldsymbol{P}$ and H replaced by K. For example,

$$L'^{(1)}_{\rm ph}(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, \boldsymbol{P}, t) \equiv \boldsymbol{P}\cdot\dot{\boldsymbol{Q}} - K(\boldsymbol{Q}, \boldsymbol{P}, t) \; . \tag{4.7}$$

As for the gauge transformation discussed above, we ensure simultaneous stationarity of the two functionals by requiring that the old and new phase-space Lagrangians differ only by the total time derivative of some function, which we denote by F,

$$L_{\rm ph}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{p}, \dot{\boldsymbol{p}}, t) = L'_{\rm ph}(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, \boldsymbol{P}, \dot{\boldsymbol{P}}, t) + \frac{\mathrm{d}F}{\mathrm{d}t} \; . \tag{4.8}$$

This equality is to hold on *all* phase-space paths, not only the physical ones.

Do we take F to be a function of the old or of the new variables? The traditional answer is to take F to be a *mixed* function of n of the old and n of the new (since there are only $2n$ independent variables), giving four possibilities

$$F_1(\boldsymbol{q}, \boldsymbol{Q}, t), \quad F_2(\boldsymbol{q}, \boldsymbol{P}, t), \quad F_3(\boldsymbol{p}, \boldsymbol{Q}, t), \quad F_4(\boldsymbol{p}, \boldsymbol{P}, t) \; . \tag{4.9}$$

These functions are known as the *generating functions* of types 1 to 4. We require the two sets of coordinates used to be functionally independent, so that F really is a function of $2n$ independent phase-space coordinates. This condition rules out, for example, the use of the type 1 generating function for point transformations, including the identity transformation. Otherwise the choice of generating function is somewhat arbitrary and we choose whichever is convenient for the problem.

The total time derivatives in the four cases are given by

$$\frac{\mathrm{d}F_1}{\mathrm{d}t} = \dot{\boldsymbol{q}}\cdot\frac{\partial F_1}{\partial \boldsymbol{q}} + \dot{\boldsymbol{Q}}\cdot\frac{\partial F_1}{\partial \boldsymbol{Q}} + \frac{\partial F_1}{\partial t} \tag{4.10}$$

$$\frac{\mathrm{d}F_2}{\mathrm{d}t} = \dot{\boldsymbol{q}}\cdot\frac{\partial F_2}{\partial \boldsymbol{q}} + \dot{\boldsymbol{P}}\cdot\frac{\partial F_2}{\partial \boldsymbol{P}} + \frac{\partial F_2}{\partial t} \tag{4.11}$$

$$\frac{\mathrm{d}F_3}{\mathrm{d}t} = \dot{\boldsymbol{p}}\cdot\frac{\partial F_3}{\partial \boldsymbol{p}} + \dot{\boldsymbol{Q}}\cdot\frac{\partial F_3}{\partial \boldsymbol{Q}} + \frac{\partial F_3}{\partial t} \tag{4.12}$$

$$\frac{\mathrm{d}F_4}{\mathrm{d}t} = \dot{\boldsymbol{p}}\cdot\frac{\partial F_4}{\partial \boldsymbol{p}} + \dot{\boldsymbol{P}}\cdot\frac{\partial F_4}{\partial \boldsymbol{P}} + \frac{\partial F_4}{\partial t}\ . \tag{4.13}$$

The four types of generating function correspond to choosing the two alternative forms, eqs. (4.4) and (4.5), on either side of the gauge transformation, eq. (4.8), as indicated in the following table:

$$\begin{array}{ll} L^{(1)}_{\rm ph}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{p}) = & L^{(1)}_{\rm ph}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{p}) = \\ L'^{(1)}_{\rm ph}(\boldsymbol{Q},\dot{\boldsymbol{Q}},\boldsymbol{P}) + \dot{F}_1(\boldsymbol{q},\boldsymbol{Q})\ , & L'^{(2)}_{\rm ph}(\boldsymbol{Q},\boldsymbol{P},\dot{\boldsymbol{P}}) + \dot{F}_2(\boldsymbol{q},\boldsymbol{P})\ , \\ L^{(2)}_{\rm ph}(\boldsymbol{q},\boldsymbol{p},\dot{\boldsymbol{p}}) = & L^{(2)}_{\rm ph}(\boldsymbol{q},\boldsymbol{p},\dot{\boldsymbol{p}}) = \\ L'^{(1)}_{\rm ph}(\boldsymbol{Q},\dot{\boldsymbol{Q}},\boldsymbol{P}) + \dot{F}_3(\boldsymbol{p},\boldsymbol{Q})\ , & L'^{(2)}_{\rm ph}(\boldsymbol{Q},\boldsymbol{P},\dot{\boldsymbol{P}}) + \dot{F}_4(\boldsymbol{p},\boldsymbol{P})\ , \end{array} \tag{4.14}$$

where $\dot{F}_i$ denotes $\mathrm{d}F_i/\,\mathrm{d}t$ and we have suppressed explicit time dependences for brevity.

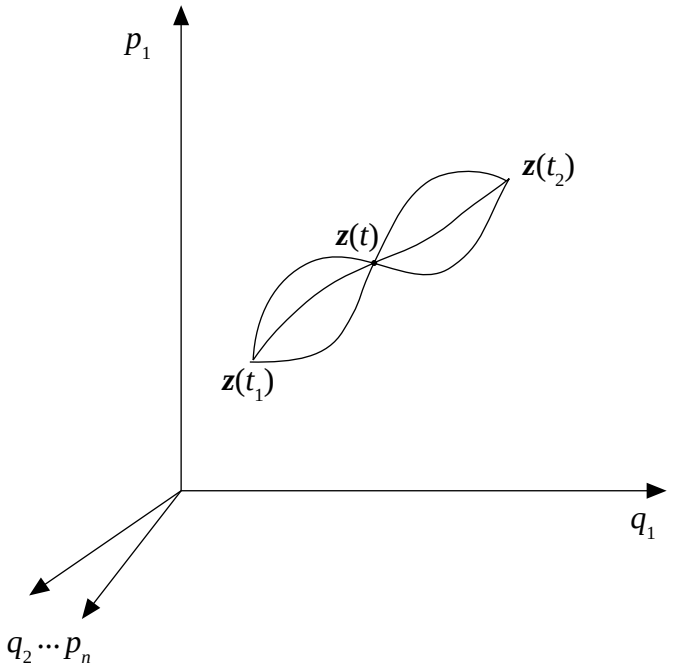

Figure 4.1: Some possible paths $\boldsymbol{z}(t) \equiv (\boldsymbol{q}(t),\boldsymbol{p}(t))$ in phase space, each passing through the same point at time t but with different $\dot{\boldsymbol{z}}$.

As indicated in eq. (4.14), **Type 1** generating functions combine most naturally with the choice $L_{\rm ph} = L^{(1)}_{\rm ph}$, $L'_{\rm ph} = L'^{(1)}_{\rm ph}$, since both the Lagrangians

and $\dot{F}_1$ involve only $\dot{\boldsymbol{q}}$ and $\dot{\boldsymbol{Q}}$, but not $\dot{\boldsymbol{p}}$ or $\dot{\boldsymbol{P}}$. Substituting eq. (4.10) in eq. (4.8) with these choices and collecting terms in $\dot{\boldsymbol{q}}$ and $\dot{\boldsymbol{Q}}$ we get

$$\begin{aligned}&\left(\boldsymbol{p}-\frac{\partial F_1}{\partial \boldsymbol{q}}\right)\cdot\dot{\boldsymbol{q}}-\left[H(\boldsymbol{q},\boldsymbol{p},t)+\frac{\partial F_1}{\partial t}\right]\\&=\left(\boldsymbol{P}+\frac{\partial F_1}{\partial \boldsymbol{Q}}\right)\cdot\dot{\boldsymbol{Q}}-K(\boldsymbol{Q},\boldsymbol{P},t)\;.\end{aligned}\tag{4.15}$$

Since eq. (4.8) is to hold on all paths, it must in particular hold at any given phase-space point $(\boldsymbol{q},\boldsymbol{p})$ for all possible paths passing through this point. As illustrated in Fig. 4.1 the possible total time derivatives $(\dot{\boldsymbol{q}},\dot{\boldsymbol{p}})$ span the full $2n$-dimensional space of tangent directions through the given point. Since we require $\boldsymbol{q}$ and $\boldsymbol{Q}$ to be functionally independent, the possible choices of $\dot{\boldsymbol{q}}$ and $\dot{\boldsymbol{Q}}$ also span a $2n$-dimensional space, and thus eq. (4.15) can hold identically only if the coefficients of $\dot{\boldsymbol{q}}$ and $\dot{\boldsymbol{Q}}$ vanish separately

$$\boldsymbol{p} = \frac{\partial F_1}{\partial \boldsymbol{q}}\tag{4.16}$$

$$\boldsymbol{P} = -\frac{\partial F_1}{\partial \boldsymbol{Q}}\;,\tag{4.17}$$

finally leaving

$$K = H+\frac{\partial F_1}{\partial t}\;.\tag{4.18}$$

These equations define, albeit implicitly, a canonical transformation.

For **type 2** generating functions we choose $L_{\rm ph}=L_{\rm ph}^{(1)}$, $L'_{\rm ph}=L_{\rm ph}^{\prime(2)}$ so that eq. (4.8) becomes

$$\begin{aligned}&\left(\boldsymbol{p}-\frac{\partial F_2}{\partial \boldsymbol{q}}\right)\cdot\dot{\boldsymbol{q}}-\left[H(\boldsymbol{q},\boldsymbol{p},t)+\frac{\partial F_2}{\partial t}\right]\\&=-\left(\boldsymbol{Q}-\frac{\partial F_2}{\partial \boldsymbol{P}}\right)\cdot\dot{\boldsymbol{P}}-K(\boldsymbol{Q},\boldsymbol{P},t)\;.\end{aligned}\tag{4.19}$$

Equating the coefficients of $\dot{\boldsymbol{q}}$ and $\dot{\boldsymbol{P}}$ to zero we get the implicit equations defining type 2 canonical transformations

$$\boldsymbol{p} = \frac{\partial F_2}{\partial \boldsymbol{q}}\tag{4.20}$$

$$\boldsymbol{Q} = \frac{\partial F_2}{\partial \boldsymbol{P}}\tag{4.21}$$

and

$$K = H+\frac{\partial F_2}{\partial t}\;.\tag{4.22}$$

For **type 3** we choose $L_{\rm ph} = L^{(2)}_{\rm ph}$, $L'_{\rm ph} = L'^{(1)}_{\rm ph}$ and find

$$\boldsymbol{q} = -\frac{\partial F_3}{\partial \boldsymbol{p}} \tag{4.23}$$

$$\boldsymbol{P} = -\frac{\partial F_3}{\partial \boldsymbol{Q}} \tag{4.24}$$

and

$$K = H + \frac{\partial F_3}{\partial t} \ . \tag{4.25}$$

For **type 4** we choose $L_{\rm ph} = L^{(2)}_{\rm ph}$, $L'_{\rm ph} = L'^{(2)}_{\rm ph}$, to find

$$\boldsymbol{q} = -\frac{\partial F_4}{\partial \boldsymbol{p}} \tag{4.26}$$

$$\boldsymbol{Q} = \frac{\partial F_4}{\partial \boldsymbol{P}} \tag{4.27}$$

and

$$K = H + \frac{\partial F_4}{\partial t} \ . \tag{4.28}$$

4.2.1 Example 1: Adiabatic Oscillator

Consider the harmonic oscillator Lagrangian eq. (2.61) and Hamiltonian eq. (3.16), taking ω_0 to be a slowly varying function of time which never vanishes, $\omega_0 = \omega_0(\epsilon t) \neq 0$, where ϵ is small parameter. (Introducing ϵ is a useful device for the formal asymptotic theory, in which we consider the scaling of error terms as $\epsilon \to 0$.)

Now consider the Hamiltonian, $H = p^2/2m + m\omega_0^2 x^2/2$ and try the type 1 generating function

$$F_1 = \tfrac{1}{2} m\omega_0 x^2 \cot\phi \ , \tag{4.29}$$

with x playing the role of the old generalized coordinate, q, and ϕ playing the role of the new generalized coordinate, Q. Using eq. (4.29) in eqs. (4.16) and (4.17) we have

$$p = m\omega_0 x \cot\phi \tag{4.30}$$

$$P = \tfrac{1}{2} m\omega_0 x^2 \,\mathrm{cosec}^2\phi \ . \tag{4.31}$$

Eliminating x between these equations we find the canonical transformation between x, p and ϕ, P

$$p = (2m\omega_0 P)^{1/2} \cos\phi \ , \tag{4.32}$$

$$x = \left(\frac{2P}{m\omega_0}\right)^{1/2} \sin\phi \ . \tag{4.33}$$

Comparing eq. (4.33) with eq. (2.62) we identify $P = m\omega_0 A^2/2$. Comparing with eq. (2.65) we see that P is in fact J, the action. Thus we should be able to show its adiabatic invariance from the Hamiltonian equations of motion. First we need the new Hamiltonian, which from eq. (4.18) is given by

$$\begin{aligned} K &= \frac{p^2}{2m} + \frac{m\omega_0^2 x^2}{2} + \frac{m\dot{\omega}_0 x^2 \cot\phi}{2} \\ &= \omega_0 P + \frac{1}{2}\frac{\mathrm{d}\ln\omega_0}{\mathrm{d}t} P \sin 2\phi \ . \end{aligned} \tag{4.34}$$

The equations of motion are thus

$$\dot{\phi} = \frac{\partial K}{\partial P} = \omega_0 + \frac{1}{2}\frac{\mathrm{d}\ln\omega_0}{\mathrm{d}t}\sin 2\phi \tag{4.35}$$

and

$$\dot{P} = -\frac{\partial K}{\partial \phi} = \frac{\mathrm{d}\ln\omega_0}{\mathrm{d}t} P\cos 2\phi \ . \tag{4.36}$$

So far the results are exact, but to proceed further (unless we solve these equations numerically) we need to make use of the expansion parameter ϵ. Observe that $\mathrm{d}\ln\omega_0/\,\mathrm{d}t = \epsilon\omega_0'(\epsilon t)/\omega_0$ is small, $O(\epsilon)$. Thus to leading order we may drop the $\sin 2\phi$ term in eq. (4.35) because it is small in comparison with ω_0, so that we are making the same approximation, $\omega \equiv \dot{\phi} \approx \omega_0$, as we were led to in the averaged Lagrangian approach. Integrating the equation to get the phase as a function of time we have

$$\phi(t) \approx \int_{t_0}^{t} \omega_0(\epsilon t) \ , \tag{4.37}$$

where t_0 is an arbitrary starting time. Thus the phase increases with time without bound. Functions with this property are known as *secular* terms.

The right-hand side of eq. (4.36) is $O(\epsilon)$ so we see that P is approximately constant over short times. However we can say more than that since we can estimate the long-term change by integrating the equation,

$$\ln P(t) - \ln P_0 \approx \int_{t_0}^{t} \mathrm{d}t \ \frac{\mathrm{d}\ln\omega_0}{\mathrm{d}t}\cos 2\phi \ . \tag{4.38}$$

Using eq. (4.37), we see that the integrand averages to zero over times long compared with $1/\omega_0$ and thus the integral will also be oscillatory, with *no secular increase*. Thus the canonical transformation method confirms the adiabatic invariance of the action.

Remark 4.1 *As implied by the remark at the end of Sec. 2.5, we could go on to construct a formal asymptotic procedure in which the canonical transformation was successively refined to remove the oscillatory terms in P at each order in ϵ to produce an adiabatic invariant that is conserved to arbitrarily high order (but not exactly, because of the possibility of asymptotically small corrections "beyond all orders").*

Note that, apart from the small correction term $\partial F_1/\partial t$, K is still approximately the total energy of the oscillator, $E \equiv T + V$, so eq. (4.34) gives $P \approx E/\omega_0$. The quantum mechanical interpretation of this was pointed out around 1911 by Einstein: If there are $N >> 1$ quanta of energy in the oscillator, then $E = N\hbar\omega_0$ (ignoring the zero-point energy $\frac{1}{2}\hbar\omega_0$). Thus the adiabatic invariance of $P = N\hbar$ can be interpreted as saying *the number of quanta is conserved during adiabatic changes in the system.*

4.2.2 Example 2: Point transformations

As an example of a type 2 generating function, take

$$F_2(\boldsymbol{q}, \boldsymbol{P}, t) = \boldsymbol{G}(\boldsymbol{q}, t)\cdot\boldsymbol{P} \ , \tag{4.39}$$

where $\boldsymbol{G}(\boldsymbol{q}, t)$ is a function such that the matrix $\partial G_j/\partial q_i$ is nonsingular. Then eqs. (4.20) and (4.21) give

$$\begin{aligned} p_i &= \sum_{j=1}^{n} \frac{\partial G_j}{\partial q_i} P_j \ , \\ \boldsymbol{Q} &= \boldsymbol{G}(\boldsymbol{q}, t) \end{aligned} \tag{4.40}$$

The second equation shows that eq. (4.39) is the generating function for a point transformation, and the first equation is the same as eq. (4.2) with the roles of the old and new variables interchanged. However, in the case of a time-dependent transformation, the generating function method now provides us with the transformation of the Hamiltonian, a useful result that we did not have easily before. That is, substituting eq. (4.39) in eq. (4.22) gives

$$K(\boldsymbol{Q}, \boldsymbol{P}, t) = H(\boldsymbol{q}, \boldsymbol{p}, t) + \frac{\partial \boldsymbol{G}(\boldsymbol{q}, t)}{\partial t}\cdot\boldsymbol{P} \ . \tag{4.41}$$

Note that included among the point transformations is the identity transformation. Thus, the class of canonical transformations that can be generated by a type 2 generating function includes the identity, whereas the type 1 class does not [though a type 1 transformation may come arbitrarily close to the identity—see Problem 4.7.2(c) and consider the limit $\Delta t \to 0$].

4.3 Infinitesimal canonical transformations

The class of canonical transformations continuously connected to the identity transformation is important for several reasons—it includes time evolution, continuous symmetry operations and canonical perturbation theory.

Since type 1 transformations do not include the identity, we consider the type 2 generating functions $F_2(\boldsymbol{q}, \boldsymbol{P}, t)$ and note from eqs. (4.20) and (4.21) that $F_2 = \boldsymbol{q}\cdot\boldsymbol{P}$ gives $\boldsymbol{p} = \partial F_2/\partial \boldsymbol{q} = \boldsymbol{P}$ and $\boldsymbol{Q} = \partial F_2/\partial \boldsymbol{P} = \boldsymbol{q}$. That is, $\boldsymbol{q}\cdot\boldsymbol{P}$ generates the identity transformation.

We now move continuously away from the identity among a family of canonical transformations parameterized by ϵ, say, such that $\epsilon = 0$ is the identity. The generating function of such a family can be written in the form

$$F_2 = \boldsymbol{q}\cdot\boldsymbol{P} + \sigma(\boldsymbol{q}, \boldsymbol{P}, t)\epsilon + O(\epsilon^2) \; , \tag{4.42}$$

where the notation $O(\epsilon^2)$ means a term scaling like ϵ^2 in the limit $\epsilon \to 0$, and thus negligibly small compared with the $O(\epsilon)$ term. The term σ is called the *infinitesimal generator* of the family of transformations.

From eqs. (4.20) and (4.21) we have

$$\begin{aligned} \boldsymbol{Q} = \frac{\partial F_2}{\partial \boldsymbol{P}} &= \boldsymbol{q} + \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{P}, t)}{\partial \boldsymbol{P}}\epsilon + O(\epsilon^2) \; , \\ \boldsymbol{p} = \frac{\partial F_2}{\partial \boldsymbol{q}} &= \boldsymbol{P} + \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{P}, t)}{\partial \boldsymbol{q}}\epsilon + O(\epsilon^2) \; . \end{aligned} \tag{4.43}$$

Due to the nature of the generating function method these equations are mixed between the perturbed and unperturbed canonical variables, but they can easily be solved by iteration to give, up to first order in ϵ,

$$\begin{aligned} \boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{p}, t, \epsilon) &= \boldsymbol{q} + \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial \boldsymbol{p}}\epsilon + O(\epsilon^2) \; , \\ \boldsymbol{P}(\boldsymbol{q}, \boldsymbol{p}, t, \epsilon) &= \boldsymbol{p} - \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial \boldsymbol{q}}\epsilon + O(\epsilon^2) \; . \end{aligned} \tag{4.44}$$

Also, inserting eq. (4.42) in eq. (4.22),

$$K(\boldsymbol{Q}, \boldsymbol{P}, t, \epsilon) = H(\boldsymbol{q}, \boldsymbol{p}, t) + \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial t}\epsilon + O(\epsilon^2) \; . \tag{4.45}$$

This is again a mixed expression in the perturbed and unperturbed variables so we use eq. (4.44) to get an explicit expression

$$\begin{aligned} K(\boldsymbol{q}, \boldsymbol{p}, t, \epsilon) = H(\boldsymbol{q}, \boldsymbol{p}, t) + \Bigg(&\frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial \boldsymbol{q}}\cdot\frac{\partial H}{\partial \boldsymbol{p}} - \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial \boldsymbol{p}}\cdot\frac{\partial H}{\partial \boldsymbol{q}} \\ &+ \frac{\partial \sigma(\boldsymbol{q}, \boldsymbol{p}, t)}{\partial t}\Bigg)\epsilon + O(\epsilon^2) \; . \end{aligned} \tag{4.46}$$

4.3.1 Time evolution

Normally ϵ is a parameter that we imagine varies at fixed time, but it is also of great interest to consider the case where the parameter is time itself. That is, we consider transformations at times $t + \delta t$ where t is fixed but the time variation comes from taking $\epsilon = \delta t$. Now consider the choice $\sigma(\boldsymbol{q}, \boldsymbol{p}, t) = H(\boldsymbol{q}, \boldsymbol{p}, t)$. Then we see by comparing eq. (4.44) with Hamilton's equations of motion, eq. (3.12), that

$$\begin{aligned} \boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{p}, t, \delta t) = \boldsymbol{q}(t) + \dot{\boldsymbol{q}}(t)\delta t + O((\delta t)^2) &= \boldsymbol{q}(t + \delta t) + O((\delta t)^2) \,, \\ \boldsymbol{P}(\boldsymbol{q}, \boldsymbol{p}, t, \delta t) = \boldsymbol{p}(t) + \dot{\boldsymbol{p}}(t)\delta t + O((\delta t)^2) &= \boldsymbol{p}(t + \delta t) + O((\delta t)^2) \end{aligned} \quad (4.47)$$

That is, we have shown that the Hamiltonian is the infinitesimal generator for dynamical evolution in time.

Since the result is true at *all* times t, we can integrate it to obtain a statement valid over finite time intervals. First we introduce some dynamical systems jargon by defining the Hamiltonian *flow*, which is the function $\boldsymbol{\varphi}_{t,s}(\boldsymbol{q}_s, \boldsymbol{p}_s)$ evolving any phase space point $(\boldsymbol{q}_s, \boldsymbol{p}_s)$ at time s to the corresponding point on the physical path at time t: $\boldsymbol{\varphi}_{t,s}(\boldsymbol{q}_s, \boldsymbol{p}_s) = (\boldsymbol{q}(t), \boldsymbol{p}(t))$, with $\dot{\boldsymbol{q}} = \partial H/\partial \boldsymbol{p}$ and $\dot{\boldsymbol{p}} = -\partial H/\partial \boldsymbol{q}$. Then we can state the integrated result in the following form: *Hamiltonian flows are canonical transformations.*

A commonly encountered example of such a flow is what we shall call a *stroboscopic map*, $\boldsymbol{\varphi}_{t_{n+1},t_n}(\boldsymbol{q}_n, \boldsymbol{p}_n)$, which evolves a phase-space point at time $t_n = n\Delta t$ to the corresponding point at time $t_{n+1} = (n+1)\Delta t$ as if the state of the system were captured periodically in the flashes of a stroboscope.

The main use of of such a stroboscopic map is the analysis of the dynamics of a nonautonomous, periodically forced system, in which one takes Δt to be the period of the forcing. In this case the map is called a *period map*. If we visualize the motion in the $(2n+1)$-dimensional space formed by adding the time dimension to phase space, and evaluate t modulo Δt (i.e. make the space topologically toroidal), we may think of the period map as giving the intersections of orbits with the *Poincaré surface of section*, $t = 0$. (This is most useful when $n = 1$, for then the surface of section is two-dimensional and easy to visualize.) For this reason systems where the Hamiltonian has an explicit periodicity in time are sometimes called $(2n+1)$-*dimensional Hamiltonian systems*.

Another example is the construction of a *numerical integrator* of the ODEs making up Hamilton's equations of motion. Of necessity we replace the infinite number of points in time of the exact problem with a finite number of points in a numerical approximation. Although we do not know the flow exactly (otherwise there would be no point in integrating the equations

numerically) it is sometimes very important to require the approximate stroboscopic map to be exactly canonical. (The full significance of the fact that a map is canonical will become apparent when we have established some more properties of canonical transformations.) In this case the numerical algorithm is called a *symplectic integrator*. An example is the discrete-time dynamical system derived in Problem 4.7.2(c).

4.4 Poisson brackets

We define the canonical Poisson bracket $\{f, g\}$ between any two phase-space functions f and g by[1]

$$\{f, g\} \equiv \frac{\partial f}{\partial \boldsymbol{q}} \cdot \frac{\partial g}{\partial \boldsymbol{p}} - \frac{\partial f}{\partial \boldsymbol{p}} \cdot \frac{\partial g}{\partial \boldsymbol{q}} \; . \tag{4.48}$$

Note that the bracket is an *antisymmetric* bilinear form in f and g: $\{f, g\} = -\{g, f\}$, so that $\{f, f\} \equiv 0$. We observe that this algebraic property is analogous to the properties of the *commutator* of two operators, A and B say: $[A, B] \equiv AB - BA$. Indeed there is a deep physical connection between the commutator and the Poisson bracket through the quantum-mechanical correspondence principle: if $a_{\rm op}$ and $b_{\rm op}$ are quantum-mechanical observables, then the classical limit may be obtained by replacing $[a_{\rm op}, b_{\rm op}]$ with $i\hbar\{a, b\}$, where a and b are the classical counterparts of the respective quantum-mechanical operators.

A property shared with the commutator is the *Jacobi identity*

$$\{u, \{v, w\}\} + \{v, \{w, u\}\} + \{w, \{u, v\}\} \equiv 0 \; . \tag{4.49}$$

The utility of the Poisson bracket notation lies in the fact that Hamilton's equations of motion, eq. (3.12), can be written in the symmetric fashion

$$\begin{aligned} \dot{\boldsymbol{q}} &= \{\boldsymbol{q}, H\} \\ \dot{\boldsymbol{p}} &= \{\boldsymbol{p}, H\} \; , \end{aligned} \tag{4.50}$$

which allows the more compact notation

$$\dot{\boldsymbol{z}} = \{\boldsymbol{z}, H\} \; , \tag{4.51}$$

where $\boldsymbol{z} \equiv (\boldsymbol{q}, \boldsymbol{p})$ is the $2n$-dimensional phase-space position defined by concatenating $\boldsymbol{q}$ and $\boldsymbol{p}$.

[1] Do not confuse these curly brackets with the similar notation we have been using for sets of variables. We follow the sign convention used by Goldstein (see Sec. 6) whereas Scheck defines the Poisson bracket with the opposite sign.

We can generalize this result to the time derivative of any function, $f(\boldsymbol{z}, t)$,

$$\frac{\mathrm{d}f}{\mathrm{d}t} \equiv \dot{\boldsymbol{z}} \cdot \frac{\partial f}{\partial \boldsymbol{z}} + \frac{\partial f}{\partial t} = \{f, H\} + \frac{\partial f}{\partial t} \ . \tag{4.52}$$

For example, if we take $f = H$ we immediately get

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} \ , \tag{4.53}$$

which agrees with eq. (3.30). Note that, in an autonomous system, f is an integral of the motion if it does not depend explicitly on time and its Poisson bracket with the Hamiltonian vanishes.

We can also write eqs. (4.44) and (4.46) in the more compact forms

$$\boldsymbol{Z}(\boldsymbol{z}, t, \epsilon) = \boldsymbol{z} + \{\boldsymbol{z}, \sigma(\boldsymbol{z}, t)\}\epsilon + O(\epsilon^2) \ , \tag{4.54}$$

where $\boldsymbol{Z} \equiv (\boldsymbol{Q}, \boldsymbol{P})$, and

$$K(\boldsymbol{z}, t, \epsilon) = H(\boldsymbol{z}, t) + \left(\{\sigma(\boldsymbol{z}, t), H(\boldsymbol{z}, t)\} + \frac{\partial \sigma(\boldsymbol{z}, t)}{\partial t} \right) \epsilon + O(\epsilon^2) \ . \tag{4.55}$$

4.4.1 Symmetries and integrals of motion

In this section we consider only autonomous systems. From eq. (4.55) the Hamiltonian is invariant under the canonical transformations generated by σ if and only if its Poisson bracket with $\sigma(\boldsymbol{z})$ vanishes. But eq. (4.52) shows that this implies that σ is an integral of the motion. That is, *the infinitesimal generator of a symmetry of the Hamiltonian is an integral of the motion.*

For example, take $\sigma = p_\phi$, the momentum canonically conjugate to $q = \phi$, the angle measured about the z-axis. Then eq. (4.44) implies that ϕ is mapped to $\phi + \epsilon$. Thus we see that p_ϕ is the infinitesimal generator for rotations about the z-axis. If the system is rotationally symmetric then the Hamiltonian is invariant under increments in ϕ. Then the above result shows that p_ϕ is an integral of the motion. This is a Hamiltonian form of Noether's theorem. (Of course the result is obvious if we use ϕ as a generalized coordinate, since the absence of ϕ from the Hamiltonian immediately gives $\dot{p}_\phi = 0$, but we now do not need to use such coordinates—all we need to do is evaluate the Poisson bracket of H with p_ϕ.)

4.4.2 Perturbation theory

Often we are given a system for which the Hamiltonian may be resolved into the form $H_0 + \epsilon H_1$, where the case $\epsilon = 0$ is exactly solvable and H_1

represents the effect of a small perturbation. (An example is the planets moving around the sun—the motion of a single planet was exactly solved by Newton, and shown to give rise to Kepler's Laws. Thus we can take H_0 to be the Hamiltonian for the planets with only the gravitational field of the sun included, and put the small gravitational interactions between the planets into H_1. This is known as a problem in *celestial mechanics*.)

One strategy for solving the perturbed problem is to successively transform the new Hamiltonian into an exactly integrable one order by order in a power series in ϵ. This is known as *canonical perturbation theory*.

The classical method of generating the canonical transformations is called *Von Zeipel's method*, and this consists of extending eq. (4.42) to higher order in ϵ. The problem with this method is the awkward mixture of old and new variables that has to be unscrambled. A more modern approach, amenable to computerized algebraic manipulation, is known as the *Lie transform*. This is an alternative to the traditional generating function method that can be motivated by noting that time evolution generates a canonical transformation continously connected to the identity. If we regard ϵ as pseudo-time, then we can use a pseudo-Hamiltonian, W say, to generate a canonical transformation that is an explicit function of only the old variables.

There are various variants of the Lie transform method, but they all make use of the idea that we can represent the Poisson bracket with an infinitesimal generator W as an operator $\mathcal{L}_W$ such that its action on any phase-space function f is given by

$$\mathcal{L}_W f \equiv \{f, W\} \; . \tag{4.56}$$

In the case that W does not depend on ϵ, the canonical transformation is effected by exponentiating $\mathcal{L}_W$

$$\exp(\epsilon \mathcal{L}_W) = 1 + \epsilon \mathcal{L}_W + \tfrac{1}{2}\epsilon^2 \mathcal{L}_W^2 + \ldots \; . \tag{4.57}$$

4.5 Action-Angle Variables

It is clear from the equations of motion that if the Hamiltonian does not depend on one of the generalized coordinates q_i then the canonical momentum conjugate to it, p_i, is conserved. (We can as before refer to q_i as ignorable, though the terminology *cyclic coordinate* is also used in this context.)

This suggests a general strategy for solving Hamiltonian dynamics problems—make a canonical transformation to new canonical coordinates $\boldsymbol{Q}, \boldsymbol{P}$ such that the new Hamiltonian is cyclic in all the new generalized coordinates

$K = K(\boldsymbol{P})$ (we consider the autonomous case only here). Then the dynamics is very simple since all the new momenta are constants of the motion and thus the generalized velocities $\partial K/\partial P_i$ are also constant, so all the new generalized coordinates evolve linearly in time

$$Q_i = Q_{i0} + \frac{\partial K}{\partial P_i} t \; . \tag{4.58}$$

A system for which this strategy can, in principle at least, be carried out is called *integrable*. Clearly, since there are n of the P_i, a necessary condition for integrability is that there must exist n independent integrals of the motion that are well behaved as functions of the old canonical coordinates for all time, including the limits $t \to \pm\infty$. (For an autonomous system in fact the integrals of motion have no explicit dependence on t at all.)

If the problem is such that the orbits remain in a bounded region of phase space, then the coordinates have to oscillate in a quasiperiodic fashion, rather than increasing secularly. In this, rather general, case the orbits are confined to topological tori labelled by the new momenta, and the new coordinates Q_i can be taken to be angles defined on the tori. The dimensions of the P_i are then those of action, and thus such a set of canonical coordinates is called *action-angle variables*. Since the Hamiltonian flow maps the $P_i = \text{const}$ tori onto themselves, they are invariant sets of the dynamical system (just as a fixed point is an invariant set consisting of just one point) and are often called *invariant tori*.

We have already come across a transformation to action-angle coordinates, eq. (4.29), for the harmonic oscillator problem (in the case when $\omega_0 = \text{const}$). Other well known integrable systems are the particle moving under the influence of an inverse-square law force, and the physical pendulum, but time does not permit the development of explicit action-angle transformations for these systems.

Although much of the classical mechanics literature is concerned with integrable problems it must be emphasized that integrability is *not* generic—given an arbitrarily chosen Hamiltonian, from the set of all possible ones, it will almost always not be integrable. Nevertheless we are very often interested in cases which are not far integrability, in the sense that the Hamiltonian can be decomposed into the form $H_0 + \epsilon H_1$ required to apply perturbation theory (see Sec. 4.4.2), where H_0 is integrable. Thus the study of integrable systems remains an important one. Of particular interest are infinite-dimensional integrable Hamiltonian systems, such as the Korteweg–de Vries equation and nonlinear Schrödinger equation that give rise to soliton phenomena.

If the problem is not integrable, perturbation theory will in general give results useful over only a finite time interval as the long-time behaviour may be chaotic and impossible to describe in detail. However by choosing the region of phase space carefully one may be able to construct an invariant torus by perturbation theory, even in a nonintegrable system, as shown by Kolmogorov, Arnol'd and Moser (the *KAM theorem*).

The essential idea in this theory is to find an invariant torus of the unperturbed system which is as far from *resonance* as possible and to keep changing parameters in such as way as ϵ is increased away from zero that the torus remains nonresonant.

Although we shall not go into this further, some idea of the meaning of resonance can be had by considering the pendulum problem. For $H \gg 2mgl$ the potential energy term is small compared with the kinetic energy and we can use $H_0 = p_\theta^2/2m$ and take g as the expansion parameter. This will give a perturbation theory involving inverse powers of p_θ, which will obviously blow up as $|p_\theta| \to 0$, i.e. as the rotation frequency approaches zero and we approach the separatrix. This blowup is an example of a resonance between the frequency of rotation and the frequency of the perturbation (in this case zero frequency).

4.6 Properties of canonical transformations

4.6.1 Preservation of phase-space volume

One of the most important properties of canonical transformations is that their Jacobian determinants $\det(\partial Z_i/\partial z_j)$ are always unity[2]. That is, *a region of phase space has the same volume whether represented in the old or new canonical coordinates.*

This is most easily seen (in the case of transformations continuously connected with the identity at least) by showing that Hamiltonian flows are *incompressible* (see Fig. 4.2). To do this we calculate the divergence of $\dot{\boldsymbol{z}} = (\partial H/\partial \boldsymbol{p}, -\partial H/\partial \boldsymbol{q})$:

$$\operatorname{div} \dot{\boldsymbol{z}} = \frac{\partial}{\partial \boldsymbol{q}} \cdot \frac{\partial H}{\partial \boldsymbol{p}} - \frac{\partial}{\partial \boldsymbol{p}} \cdot \frac{\partial H}{\partial \boldsymbol{q}} \equiv 0 \, . \tag{4.59}$$

Then the conservation of the phase-space volume follows from the application of Gauss' theorem since the surface integral of $\dot{\boldsymbol{z}}$ is the rate of change of the volume within the surface. This result obviously applies to any flows built up from infinitesimal canonical transformations, not only to Hamiltonian

[2]Unless we also include time reversal, $\mathbf{x} \mapsto \mathbf{x}$, $\mathbf{p} \mapsto -\mathbf{p}$, where the Jacobian is -1.

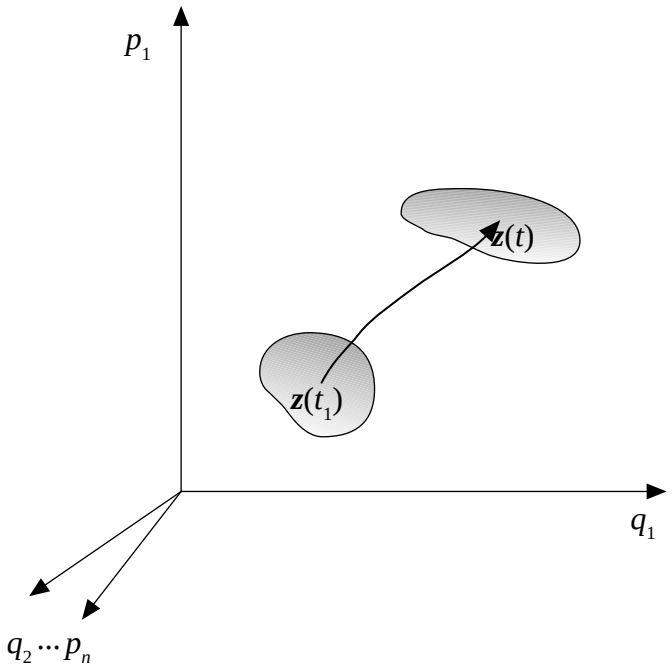

Figure 4.2: Incompressible motion of a volume in phase space.

time evolution. (In fact the conservation of phase-space volume holds for all canonical transformations.)

This result has profound implications both in statistical mechanics and in dynamical systems theory.

Liouville's Theorem

In statistical mechanics we exploit the fluid dynamical analogy implied by the above "proof" by defining a *distribution function*, $f(\boldsymbol{q}, \boldsymbol{p}, t)$, which is the phase-space density of a statistical ensemble of systems normalized so that the probability of finding a system in a small region of phase space of volume $\mathrm{d}^{2n}\boldsymbol{z}$ is $f(\boldsymbol{z}, t)\,\mathrm{d}^{2n}\boldsymbol{z}$. (Think of phase space as being filled with a dust cloud, each particle of which represents a system of the ensemble.)

Then, since the number of systems in the ensemble is conserved, f obeys the continuity equation

$$\frac{\partial f}{\partial t} + \mathrm{div}(\dot{\boldsymbol{z}} f) = 0 \,, \tag{4.60}$$

or, using the incompressibility result eq. (4.59),

$$\frac{\partial f}{\partial t} + \dot{\boldsymbol{z}} \cdot \frac{\partial f}{\partial \boldsymbol{z}} = 0 \, . \tag{4.61}$$

That is, *the distribution function is constant on a physical path in phase space*. This is Liouville's theorem of statistical mechanics.

Using $\dot{\boldsymbol{z}} = (\partial H/\partial \boldsymbol{p}, -\partial H/\partial \boldsymbol{q})$ we can write the result in terms of Poisson brackets

$$\frac{\partial f}{\partial t} + \{f, H\} = 0 \tag{4.62}$$

As a consequence we see that statistical equilibrium, $\partial f/\partial t = 0$, is obtained when $\{f, H\} = 0$, which can be achieved by taking f to be a function of the integrals of the motion. A general closed N-body system in statistical mechanics is nearly always assumed to be completely chaotic (the *ergodic hypothesis*), so that the only integral of motion in a closed system is H itself. In this case the equilibrium distribution function is a function of H only [usually the Maxwellian distribution, proportional to $\exp(-H/kT)$ where k is Boltzmann's constant and T is the temperature, though in a "collisionless" plasma it can be appropriate to consider other possibilities].

Hamiltonian mechanics as a dynamical system

The conservation of phase-space volume provides a strong constraint on the long-time behaviour of a system in that it rules out the possibility of purely attracting or repelling sets. For instance, if the basin of attraction of an attracting fixed point had the same dimensions as phase space we could draw a surface entirely enclosing the fixed point which would contract with time, thus contradicting the conservation of phase-space volume. This does not mean that fixed points that are attracting in some directions cannot exist in Hamiltonian systems, merely that for any attracting direction there must be a repelling direction to balance it.

The simplest system in which to see some of the consequences of the conservation of phase-space volume is provided by the period map of a periodically forced one-degree-of-freedom system (sometimes called a $1\frac{1}{2}$ degree of freedom Hamiltonian system). Then the phase-space is only two dimensional and the "volume" is in fact an area. Thus the period map is *area preserving*.

A standard example is the "kicked rotor"—a pendulum (see Fig. 2.2) in which gravity is varied periodically (say by oscillating the support vertically

and transforming to the accelerating frame in which the support is stationary). Furthermore we assume the gravity to be applied purely impulsively

$$g(t) = g_0 \Delta t \sum_{n=-\infty}^{\infty} \delta(t - t_n) \tag{4.63}$$

where $t_n \equiv n\Delta t$, with n an integer, and $\delta(t)$ is the Dirac δ function, so that g_0 is the time average of $g(t)$. The form of the equations of motion is unaffected by the fact that g depends on time, so the Hamiltonian is the same as in eq. (3.20)

$$H(\theta, p) = \frac{p^2}{2ml^2} + mg(t)l(1 - \cos\theta) \ . \tag{4.64}$$

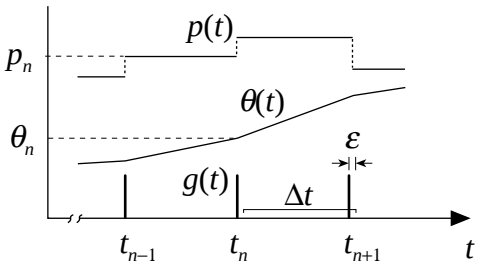

Figure 4.3: Trial function for the kicked rotor.

Between the kicks the motion is one of uniform rotation with constant angular velocity p/ml. Thus we can give a potentially exact representation of the phase-space path with θ a continuous, piecewise linear function of t and p a piecewise constant step function with only the values θ_n at, and p_n immediately before, $t = t_n$ as yet unknown. As illustrated in Fig. 4.3,

$$\begin{aligned} \theta(t) &= \frac{1}{\Delta t}\left[(t_{n+1} - t)\theta_n + (t - t_n)\theta_{n+1}\right] \ , \\ p(t) &= p_{n+1} \end{aligned} \tag{4.65}$$

for t in each range $t_n \equiv n\Delta t < t < t_{n+1} \equiv (n+1)\Delta t$. Inserting this trial function in the phase-space action integral eq. (3.32) over the interval $t_{-N} + \varepsilon \le t \le t_N + \varepsilon$, $N \ge 1$, $\varepsilon \to 0+$, we get

$$S_{\rm ph} = \sum_{n'=-N}^{N-1} \left[p_{n'+1}\left(\theta_{n'+1} - \theta_{n'}\right) - \frac{p_{n'+1}^2 \Delta t}{2ml^2} - mg_0 l \Delta t \left(1 - \cos\theta_{n'+1}\right) \right] \ . \tag{4.66}$$

The conditions for $S_{\rm ph}$ to be stationary under variations of θ_n and p_{n+1}, $-N < n < N$, (noting that θ_n occurs in the terms of the sum for both $n' = n$ and $n' = n - 1$) are

$$\frac{\partial S_{\rm ph}}{\partial p_{n+1}} = \theta_{n+1} - \theta_n - \frac{p_{n+1}\Delta t}{ml^2} = 0 \ , \tag{4.67}$$

and

$$\frac{\partial S_{\rm ph}}{\partial \theta_n} = p_n - p_{n+1} - mg_0 l \Delta t \sin\theta_n = 0 \ . \tag{4.68}$$

These equations define an area-preserving map known as the *Standard Map*. It is clearly seen that they approximate the Hamiltonian equations of motion for the physical pendulum as $\Delta t \to 0$. Commonly the map is re-expressed in terms of an angle coordinate x with period normalized to unity and a nondimensionalized momentum y, which is p expressed in units of $2\pi ml^2/\Delta t$,

$$\begin{aligned} \theta &= 2\pi x \ , \\ p &= \frac{2\pi ml^2}{\Delta t} y \ . \end{aligned} \tag{4.69}$$

Then the map becomes

$$\begin{aligned} y_{n+1} &= y_n - \frac{k}{2\pi} \sin 2\pi x_n \ , \\ x_{n+1} &= x_n + y_{n+1} \ , \end{aligned} \tag{4.70}$$

where the *chaos parameter* k is defined by

$$k \equiv \frac{g_0 (\Delta t)^2}{l} \ . \tag{4.71}$$

Figure 4.4 shows some typical iterated orbits of the map, showing both chaotic and regular regions. The remnants of the librating pendulum orbits around the O points in Fig. 3.1 are seen as large "islands" around $(x, y) = (0, 0)$ and $(0, 1)$ [also $(1, 0)$ and $(1, 1)$ because the phase space is periodic in y as well as x], but the region corresponding to the X point in Fig. 3.1 is highly chaotic. There are also new islands due to resonances not present in the physical pendulum problem (in fact an infinite number of them), each with chaotic separatrices.

The value of k used in Fig. 4.4 is $k_{\rm c} = 0.971635406\ldots$, which is the value at which the last KAM invariant curves with the topology of a rotating pendulum orbit become unstable and break up into invariant Cantor sets. Two of these KAM curves are shown, being the only curves completely crossing the figure from left boundary to right. For $k < k_{\rm c}$ the motion is bounded in y, while for k greater than this value a phase-space point can diffuse without bound in the positive or negative y-direction.

From eq. (4.71) we see that the integrable limit $\Delta t \to 0$ corresponds to $k \to 0$ [noting that eq. (4.69) shows that the width of the physical pendulum's separatrix shrinks to zero as $k \to 0$ when represented in terms of y].

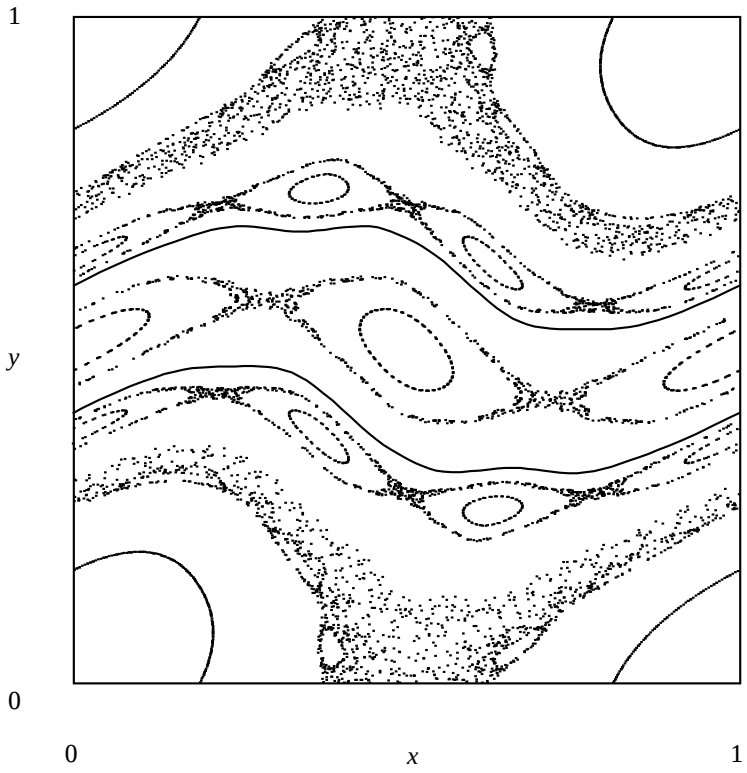

Figure 4.4: Some representative orbit segments for the Standard Map showing both chaotic and regular motion.

4.6.2 Transformation of Poisson brackets

Finally, we note the important result *the Poisson bracket is invariant under canonical transformation*

$$\{f,g\} = \frac{\partial f}{\partial \boldsymbol{q}}\cdot\frac{\partial g}{\partial \boldsymbol{p}} - \frac{\partial f}{\partial \boldsymbol{p}}\cdot\frac{\partial g}{\partial \boldsymbol{q}} = \frac{\partial f}{\partial \boldsymbol{Q}}\cdot\frac{\partial g}{\partial \boldsymbol{P}} - \frac{\partial f}{\partial \boldsymbol{P}}\cdot\frac{\partial g}{\partial \boldsymbol{Q}}\ . \tag{4.72}$$

4.7 Problems

4.7.1 Coriolis yet again

Given the Hamiltonian for a particle on a sphere in nonrotating latitude and longitude coordinates, θ and ϕ', respectively,

$$K = \frac{p_\theta^2}{2mR^2} + \frac{p_{\phi'}^2}{2mR^2\cos^2\theta} + V'(\theta,\phi',t)\ ,$$

find the transformation to the $\theta, p_\theta, \phi, p_\phi$ canonical coordinates, where ϕ is the usual latitude measured in the rotating frame defined in problem 2.7.1. That is, $\phi' = \Omega t + \phi$. Find the Hamiltonian H in the rotating coordinates.

4.7.2 Difference approximations

Using the harmonic oscillator Lagrangian $L = \frac{1}{2}m(\dot{x}^2 - \omega_0^2 x^2)$
(a) calculate an approximate action integral S using the piecewise-linear trial function

$$x(t) = \frac{1}{\Delta t}\left[(t_{n+1} - t)x_n + (t - t_n)x_{n+1}\right]$$

for t in each range $t_n \equiv n\Delta t < t < t_{n+1} \equiv (n+1)\Delta t$.
(b) Show that the approximate action integral from t_{-N} to t_N ($N \geq 1$ being an arbitrary integer) is stationary for $-N < n < N$ if x_n obeys the second-order difference equation

$$x_{n-1} - 2x_n + x_{n+1} = -\frac{\omega_0^2(\Delta t)^2}{6}\left(x_{n-1} + 4x_n + x_{n+1}\right)$$

(c) Using the approximate action integral $\bar{S}(x_n, x_{n+1})$ evaluated over the range $t_n < t < t_{n+1}$, as a type 1 generating function, $F_1(q, Q) = -\bar{S}(x_n, x_{n+1})$, with $q = x_n$ and $Q = x_{n+1}$, find the linear canonical transformation from $(x_n, p_n \equiv p)$ to $(x_{n+1}, p_{n+1} \equiv P)$. Represent this linear transformation as a matrix and show that the determinant is unity. (It would be advisable to use Maple or Mathematica for the algebra.)

This provides a discrete-time dynamical system approximation to the true continuous-time system for the harmonic oscillator.
(d) Iterate the map obtained in (c) 100 times taking the initial point as $x_0 = 1$, $p_0 = 0$ and plot the result. Use units such that $m = \omega_0 = 1$ and use three timesteps: $\Delta t = 0.1$, $\Delta t = 3.45$ and $\Delta t = 3.47$.

For a bonus mark verify one of the solutions by showing that the second-order difference equation derived in (b) above (when arranged to give x_{n+1} in terms of x_n and x_{n-1}) gives the same sequence of x-values as the discrete-time dynamical system in (c) if two successive values of x given by the dynamical system are used to start the second-order difference equation. I.e. the discrete time "Lagrangian" and "Hamiltonian" discriptions are dynamically equivalent. For another bonus mark calculate the eigenvalues for the three values of Δt given above. (Again, use Maple or Mathematica.)

Chapter 5

Answers to Problems

5.1 Chapter 1 Problems

Problem 1.6.1

Two particles are connected by a rigid rod so they are constrained to move a fixed distance apart. Write down a constraint equation of the form eq. (1.2) and find suitable generalized coordinates for the system incorporating this holonomic constraint.

Answer

Let the position of particle 1 with respect to a stationary Cartesian frame be $\{x_1, y_1, z_1\}$ and that of particle 2 be $\{x_2, y_2, z_2\}$. The rigid rod constraint equation is then

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = l^2 \ . \tag{5.1}$$

This reduces the number of degrees of freedom from 6 to 5, which we could take to be position of particle 1, $\{x_1, y_1, z_1\}$, and the spherical polar angles θ and ϕ depicted in Fig. 1.1 to specify the orientation of the rod. However, a better choice for the three position coordinates would be the coordinates of the centre of mass [see Problem 3.5.1(a)], whose position vector is

$$\mathbf{r}_{\rm CM} = \frac{m_2\mathbf{r}_1 + m_1\mathbf{r}_2}{m_1 + m_2} \ . \tag{5.2}$$

Problem 1.6.2

Suppose we know that the angular momentum vectors $\mathbf{r}_k \times m_k\dot{\mathbf{r}}_k$ of a system of particles are all nonzero and parallel to the z-axis in a particular Cartesian

coordinate system. Write down the differential constraints implied by this fact, and show that they lead to a set of holonomic constraints. Hence write down suitable generalized coordinates for the system.

Answer

The differential constraints are provided by the vanishing of the x and y components of the angular momentum vectors (multiplied by $\mathrm{d}t$)

$$\begin{aligned} y_k\,\mathrm{d}z_k - z_k\,\mathrm{d}y_k &= 0 \\ z_k\,\mathrm{d}x_k - x_k\,\mathrm{d}z_k &= 0\,. \end{aligned} \tag{5.3}$$

for $k = 1, 2, \ldots, N$.

These are satisfied by taking

$$z_k \equiv 0\,, \tag{5.4}$$

which form a set of holonomic constraints. Thus the motions are constrained to lie in the x, y plane and we can either take $\boldsymbol{q} = \{x_1, y_1, x_2, y_2, \ldots, x_N, y_N\}$ or transform to polar coordinates using $x_k = r_k \cos\theta_k$, $y_k = r_k \sin\theta_k$, so that $\boldsymbol{q} = \{r_1, \theta_1, r_2, \theta_2, \ldots, r_N, \theta_N\}$.

Remark 5.1 *This problem has been called "Ecliptic" because this is the astronomical name of the plane to which the planets are approximately "constrained". It suggests that the reason for this constraint is because the planets were all created by condensation from the same spinning disk of dust and gas and thus have parallel angular momenta.*

Note: Although eq. (5.4) clearly solves eq. (5.3) we might ask whether this is the *only* solution. Thus, suppose that z_k is *not* equal to zero. We can assume that, except for isolated instants, x_k and y_k are nonzero also. Thus we can divide the first eq. (5.3) by $z_k x_k$, the second by $y_k z_k$, and add to show (since the terms $\mathrm{d}z_k/z_k$ cancel) that $\mathrm{d}x_k/x_k - \mathrm{d}y_k/y_k = 0$. However we are told that the z components of the angular momenta are nonzero, so that $x_k\,\mathrm{d}y_k - y_k\,\mathrm{d}x_k \neq 0$. Dividing by $x_k y_k$ we clearly have a contradiction, so the assumption that $z_k \neq 0$ must be wrong. □

Problem 1.6.3

Show that any geodesic $\mathbf{r} = \mathbf{x}(\tau)$ on a two-dimensional manifold S : $\mathbf{r} = \mathbf{X}(\theta, \zeta)$ embedded in ordinary Euclidean 3-space, where θ and ζ are arbitrary curvilinear coordinates on S, is such that the curvature vector $\boldsymbol{\kappa}(\tau)$ is everywhere **normal** to S (or zero).

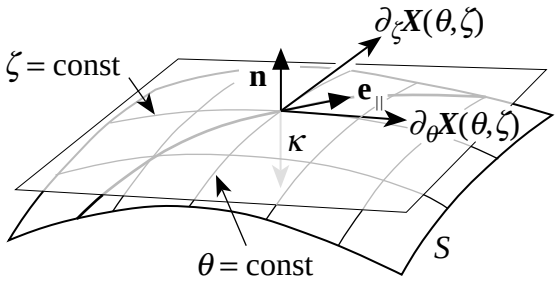

Figure 5.1: Surface S with curvilinear coordinate system, a geodesic, the three tangent-plane vectors $\mathbf{e}_\parallel$, $\partial_\theta \mathbf{X}$ and $\partial_\zeta \mathbf{X}$, and the normal vectors $\mathbf{n}$ and $\boldsymbol{\kappa}$.

The curvature vector is defined by $\boldsymbol{\kappa} \equiv \mathrm{d}\mathbf{e}_\parallel / \mathrm{d}l$, where $\mathbf{e}_\parallel(\tau) \equiv \mathrm{d}\mathbf{x}/\mathrm{d}l$ is the unit tangent vector at each point along the path $\mathbf{r} = \mathbf{x}(\tau)$.

Answer

The vector length element is $\mathbf{dl} = \mathrm{d}\theta \mathbf{X}_\theta + \mathrm{d}\zeta \mathbf{X}_\zeta$, where $\mathbf{X}_\theta \equiv \partial_\theta \mathbf{X} \equiv \partial \mathbf{X}/\partial\theta$ and $\mathbf{X}_\zeta \equiv \partial_\zeta \mathbf{X} \equiv \partial \mathbf{X}/\partial\zeta$. Thus $l = \int f(\theta, \zeta, \dot{\theta}, \dot{\zeta})\, \mathrm{d}\tau$ with

$$f \equiv \left| \frac{\mathbf{dl}}{\mathrm{d}\tau} \right| = \left(\mathbf{X}_\theta^2 \dot{\theta}^2 + 2\mathbf{X}_\theta \cdot \mathbf{X}_\zeta \dot{\theta}\dot{\zeta} + \mathbf{X}_\zeta^2 \dot{\zeta}^2 \right)^{1/2} \tag{5.5}$$

and

$$\mathbf{e}_\parallel \equiv \frac{1}{f} \frac{\mathbf{dl}}{\mathrm{d}\tau} = \frac{\dot{\theta}\mathbf{X}_\theta + \dot{\zeta}\mathbf{X}_\zeta}{f} \ . \tag{5.6}$$

From eq. (5.5), we have, for any path

$$\begin{aligned} \partial_{\dot{\theta}} f &= \frac{1}{f} \left(\mathbf{X}_\theta^2 \dot{\theta} + \mathbf{X}_\theta \cdot \mathbf{X}_\zeta \dot{\zeta} \right) \\ &= \frac{1}{f} \left(\mathbf{X}_\theta \dot{\theta} + \mathbf{X}_\zeta \dot{\zeta} \right) \cdot \mathbf{X}_\theta \\ &\equiv \mathbf{e}_\parallel \cdot \mathbf{X}_\theta \end{aligned} \tag{5.7}$$

Similarly, by interchanging θ and ζ we have

$$\partial_{\dot{\zeta}} f = \mathbf{e}_\parallel \cdot \mathbf{X}_\theta \ . \tag{5.8}$$

Also,

$$\partial_\theta f = \frac{1}{2f} \left[\dot{\theta}^2 \partial_\theta \left(\mathbf{X}_\theta^2 \right) + 2\dot{\theta}\dot{\zeta} \partial_\theta \left(\mathbf{X}_\theta \cdot \mathbf{X}_\zeta \right) + \dot{\zeta}^2 \partial_\theta \left(\mathbf{X}_\zeta^2 \right) \right]$$

$$
\begin{aligned}
&= \frac{1}{f}\left(\dot\theta^2 \mathbf{X}_\theta{\cdot}\mathbf{X}_{\theta\theta} + \dot\theta\dot\zeta \mathbf{X}_{\theta\theta}{\cdot}\mathbf{X}_\zeta + \dot\theta\dot\zeta \mathbf{X}_\theta{\cdot}\mathbf{X}_{\theta\zeta} + \dot\zeta^2 \mathbf{X}_\zeta{\cdot}\mathbf{X}_{\theta\zeta}\right) \\
&= \frac{1}{f}\left(\dot\theta \mathbf{X}_\theta + \dot\zeta \mathbf{X}_\zeta\right)\cdot\left(\dot\theta \mathbf{X}_{\theta\theta} + \dot\zeta \mathbf{X}_{\theta\zeta}\right) \\
&\equiv \mathbf{e}_\parallel{\cdot}\frac{\mathrm{d}\mathbf{X}_\theta}{\mathrm{d}\tau} \,. \qquad (5.9)
\end{aligned}
$$

Similarly, interchanging θ and ζ, we have

$$\partial_\zeta f \equiv \mathbf{e}_\parallel{\cdot}\frac{\mathrm{d}\mathbf{X}_\zeta}{\mathrm{d}\tau} \,. \qquad (5.10)$$

Now consider the Euler–Lagrange equation for a geodesic

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\partial_{\dot\theta} f - \partial_\theta f = 0 \,. \qquad (5.11)$$

Using the identities eq. (5.7) and eq. (5.9) we can write this as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\mathbf{e}_\parallel{\cdot}\mathbf{X}_\theta\right) - \mathbf{e}_\parallel{\cdot}\frac{\mathrm{d}\mathbf{X}_\theta}{\mathrm{d}\tau} = 0 \,. \qquad (5.12)$$

Or, expanding out and cancelling,

$$\mathbf{X}_\theta{\cdot}\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{e}_\parallel = 0 \,. \qquad (5.13)$$

Similarly, the Euler–Lagrange equation for ζ variations becomes

$$\mathbf{X}_\zeta{\cdot}\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{e}_\parallel = 0 \,. \qquad (5.14)$$

Recognizing that $\boldsymbol{\kappa} \equiv (\,\mathrm{d}/f\,\mathrm{d}\tau)\mathbf{e}_\parallel$ we can write eqs. (5.13) and (5.14) as

$$\mathbf{X}_\zeta{\cdot}\boldsymbol{\kappa} = \mathbf{X}_\theta{\cdot}\boldsymbol{\kappa} = 0 \,. \qquad (5.15)$$

Since the vectors $\mathbf{X}_\theta$ and $\mathbf{X}_\zeta$ span the tangent plane at the point $\mathbf{r} = \mathbf{X}(\theta,\zeta)$ (see Fig. 5.1), we have shown that the curvature vector $\boldsymbol{\kappa}$ has no component in the tangent plane. Thus either $\boldsymbol{\kappa} = 0$ or it is parallel to the unit normal, $\mathbf{n}$. □

5.2 Chapter 2 Problems

Problem 2.8.1(a)

As a model of the motion of a fluid element or dust particle in a planetary (e.g. Earth's) atmosphere, consider the motion of particle of unit mass constrained to move on the surface of a perfectly smooth sphere of radius R

rotating with angular velocity Ω about the z-axis. Suppose the force on the particle is given by an effective potential $V(\theta, \phi)$ (see Sec. 2.4.6), where θ and ϕ are the latitude and longitude respectively.

(a) Write down the Lagrangian in a frame rotating with the planet, taking the generalized coordinates to be the latitude and longitude so that $z = R\sin\theta$, $x = R\cos\theta\,\cos(\phi+\Omega t)$, $y = R\cos\theta\,\sin(\phi+\Omega t)$, where x, y, z is a non-rotating Cartesian frame.

Write down the equations of motion, and find a first integral (i.e. constant of the motion) in the case where V is independent of longitude. Among this class of potentials find the special case $V = V_0(\theta)$ required to make equilibrium possible (i.e. so that the equations of motion admit the solution $\dot\theta = \dot\phi = 0$ at each latitude).

Answer

$$\begin{aligned}
\dot x &= -R\dot\theta\sin\theta\,\cos(\phi+\Omega t) - R(\Omega+\dot\phi)\cos\theta\,\sin(\phi+\Omega t)\\
\dot y &= -R\dot\theta\sin\theta\,\sin(\phi+\Omega t) + R(\Omega+\dot\phi)\cos\theta\,\cos(\phi+\Omega t)\\
\dot z &= R\dot\theta\cos\theta
\end{aligned} \tag{5.16}$$

Thus the kinetic energy is given by

$$\begin{aligned}
T &= \tfrac{1}{2}\left(\dot x^2+\dot y^2+\dot z^2\right)\\
&= \tfrac{1}{2}R^2\Big[\dot\theta^2\sin^2\theta\left(\cos^2(\phi+\Omega t)+\sin^2(\phi+\Omega t)\right)\\
&\qquad +(\Omega+\dot\phi)^2\cos^2\theta\left(\sin^2(\phi+\Omega t)+\cos^2(\phi+\Omega t)\right)\\
&\qquad +\dot\theta^2\cos^2\theta\Big]\\
&= \tfrac{1}{2}R^2\left[\dot\theta^2+(\Omega+\dot\phi)^2\cos^2\theta\right]\ ,
\end{aligned} \tag{5.17}$$

and the Lagrangian, $L = T - V$, by

$$L = \tfrac{1}{2}R^2\left[\dot\theta^2+(\Omega+\dot\phi)^2\cos^2\theta\right] - V(\theta,\phi)\ . \tag{5.18}$$

Thus

$$\frac{\partial L}{\partial\dot\theta} = R^2\dot\theta \tag{5.19}$$

and

$$\frac{\partial L}{\partial\dot\phi} = R^2(\Omega+\dot\phi)\cos^2\theta \tag{5.20}$$

and thus the Lagrange equations of motion $(\mathrm{d}/\mathrm{d}t)\partial L/\partial\dot{\theta} = \partial L/\partial\theta$ and $(\mathrm{d}/\mathrm{d}t)\partial L/\partial\dot{\phi} = \partial L/\partial\phi$ become

$$R^2\ddot{\theta} = -R^2(\Omega+\dot{\phi})^2\sin\theta\cos\theta - \frac{\partial V}{\partial\theta} \tag{5.21}$$

and

$$R^2\ddot{\phi}\cos^2\theta - 2R^2\dot{\theta}(\Omega+\dot{\phi})\sin\theta\cos\theta = -\frac{\partial V}{\partial\phi}\ . \tag{5.22}$$

In the case where $V = V_0(\theta)$, ϕ is an ignorable coordinate and thus $\partial L/\partial\dot{\phi}$, given by eq. (5.20), is an integral of the motion.

If in this latter case we further demand that V_0 be such that $\dot{\theta} \equiv \dot{\phi} = 0$ for all θ and ϕ, and use the identity $\sin 2\theta \equiv 2\sin\theta\cos\theta$, then eq. (5.21) gives

$$\frac{\partial V_0}{\partial\theta} = -\tfrac{1}{2}R^2\Omega^2\sin 2\theta\ , \tag{5.23}$$

or, integrating,

$$V_0(\theta) = \tfrac{1}{4}R^2\Omega^2\cos 2\theta + \text{const} = \tfrac{1}{2}R^2\Omega^2\cos^2\theta\ , \tag{5.24}$$

where the second equality follows from the identity $\cos 2\theta \equiv 2\cos^2\theta - 1$ and the choice const $= \frac{1}{4}R^2$.

Problem 2.8.1(b)

Now assume that the *general* form of the potential is $V(\theta,\phi) = V_0(\theta) + h(\theta,\phi)$, where h is the enthalpy (proportional to $p^{(\gamma-1)/\gamma}$) defined in Sec. 2.4.6. Assuming the winds are slow, so that second order time derivatives of θ and ϕ (and products of first order time derivatives) can be neglected, show that the velocity of the particle in the rotating frame is at right angles to the pressure gradient $-\boldsymbol{\nabla}p$.

Draw a sketch of a typical weather-map "low" (i.e. a localized depression in p) showing the pressure contours, the direction of the force on a fluid element and the direction of motion. Hence show that motion in a depression or low is cyclonic, where "cyclonic" in geophysical fluid dynamics means "in the direction of the planet's rotation" (i.e. clockwise in the Southern Hemisphere, counter-clockwise in the Northern Hemisphere for $\Omega > 0$).

Answer

Making the slow approximation, and taking into account eq. (5.23) the equations of motion, eqs. (5.21) and (5.22), become

$$R^2\Omega\dot{\phi}\sin 2\theta \approx -\frac{\partial h}{\partial \theta} \tag{5.25}$$

and

$$R^2\Omega\dot{\theta}\sin 2\theta \approx \frac{\partial h}{\partial \phi}\ . \tag{5.26}$$

Thus the pressure gradient causes a wind proportional to spatial derivatives of h, and the slow approximation will be best satisfied when h is small.

These two equations give the wind speed in angular velocities. We now calculate the time rate of change of the enthalpy along a stream line

$$\begin{aligned}\dot{h} &\equiv \mathbf{w}\cdot\boldsymbol{\nabla} h \\ &\equiv \dot{\theta}\frac{\partial h}{\partial \theta} + \dot{\phi}\frac{\partial h}{\partial \phi} \\ &\approx \frac{1}{R^2\Omega\sin 2\theta}\left(\frac{\partial h}{\partial \phi}\frac{\partial h}{\partial \theta} - \frac{\partial h}{\partial \theta}\frac{\partial h}{\partial \phi}\right) \equiv 0\end{aligned} \tag{5.27}$$

from eqs. (5.25) and (5.26). Thus the effect of the Coriolis forces is to make the air move *along* the pressure contours, even though the forces acting on a fluid element act at right angles to the contours. It is for this reason that the vortices caused by highs and lows of pressure are so long lived—the air does not rush in to equalize the pressure differences but, rather, rotates around the maxima and minima of pressure.

In order to establish the direction of rotation it is sufficient to draw a sketch in a Mercator's projection (i.e. in the θ-ϕ plane) and use eqs. (5.25) and use (5.26) to show the wind direction at a few representative points on a typical pressure contour.

Note: It is easy to show from eq. (2.49) that $p/\rho = \text{const}p^{(\gamma-1)/\gamma}$, and so, since $\gamma > 1$, the enthalpy is an increasing function of the pressure.

Problem 2.8.1(c)

Consider a charged particle constrained to move on a *non-rotating* smooth insulating sphere, immersed in a uniform magnetic field $\mathbf{B} = B\mathbf{e}_z$, on which the electrostatic potential is a function of latitude and longitude. Write down the Lagrangian in the same generalized coordinates as above and show it is the same as that for the particle on the rotating planet with appropriate identifications of Ω and V.

Answer

The Lagrangian is $L = T + e\dot{\mathbf{r}}\cdot\mathbf{A} - e\Phi$, where T is given by eq. (5.17) with Ω set to zero. We now need a vector potential $\mathbf{A}$ such that $B_z \equiv \partial_x A_y - \partial_y A_x = B$. A suitable choice is $A_x = -\frac{1}{2}By$, $A_y = \frac{1}{2}Bx$, $A_z = 0$. Thus, from eq. (5.16),

$$\begin{aligned} e\dot{\mathbf{r}}\cdot\mathbf{A} &= \tfrac{1}{2}eB(x\dot{y} - y\dot{x}) \\ &= \tfrac{1}{2}eBR^2\cos\theta\Big[\cos\phi(-\dot{\theta}\sin\theta\,\sin\phi + \dot{\phi}\cos\theta\,\cos\phi) \\ &\qquad + \sin\phi(\dot{\theta}\sin\theta\,\cos\phi + \dot{\phi}\cos\theta\,\sin\phi)\Big] \\ &= \tfrac{1}{2}eBR^2\dot{\phi}\cos^2\theta\ . \end{aligned} \tag{5.28}$$

Thus the Lagrangian is

$$L = \tfrac{1}{2}mR^2\left(\dot{\theta}^2 + \dot{\phi}^2\cos^2\theta\right) + \tfrac{1}{2}eBR^2\dot{\phi}\cos^2\theta - e\Phi(\theta,\phi)\ . \tag{5.29}$$

Completing the square we can rewrite the Lagrangian in the form

$$L = \tfrac{1}{2}mR^2\left[\dot{\theta}^2 + \left(\dot{\phi} + \frac{eB}{2m}\right)^2\cos^2\theta\right] - \frac{e^2B^2R^2}{8m}\cos^2\theta - e\Phi(\theta,\phi)\ . \tag{5.30}$$

This is the same as eq. (5.18) with the identifications

$$\Omega = \frac{eB}{2m} \tag{5.31}$$

and

$$V(\theta,\phi) = \frac{e^2B^2R^2}{8m}\cos^2\theta + e\Phi(\theta,\phi)\ . \tag{5.32}$$

Note that eq. (5.31) can also be written $\Omega = \frac{1}{2}\omega_{\rm c}$, where $\omega_{\rm c} \equiv eB/m$ is the *cyclotron frequency*.

Problem 2.8.2

Consider the following potential V, corresponding to a particle of mass m oscillating along the x-axis under the influence of a nonideal spring (i.e. one with a nonlinear restoring force),

$$V(x) = \frac{m\omega_0^2}{2}\left(x^2 + \sigma\frac{x^4}{l_0^2}\right)\ ,$$

where the constant ω_0 is the angular frequency of oscillations having amplitude small compared with the characteristic length l_0, and $\sigma = \pm 1$ depends on whether the spring is "soft" ($\sigma = -1$) or "hard" ($\sigma = +1$).

Consider the trial function

$$x = l_0 \left[A \cos \omega t + B \cos 3\omega t + C \sin 3\omega t \right] \ ,$$

where A, B, C are the nondimensionalized amplitudes of the fundamental and third harmonic, respectively, and ω is the nonlinearly shifted frequency. By using this trial function in the time-averaged Hamilton's Principle, find implicit relations giving approximate expressions for ω, B and C as functions of A. Show that $C \equiv 0$. The trial function is strictly appropriate only to the case $A \ll 1$, but plot ω/ω_0 and B *vs.* A from 0 to 1 in the case of both a hard and a soft spring.

Answer

See the following *Maple* notebook.

5.3 Chapter 3 Problems

Problem 3.5.1(a)

Write down a Lagrangian for the problem of two particles of mass m_1 and m_2 connected by a light rigid rod of length l in a gravitational field g. Take the generalized coordinates of the system to be $\boldsymbol{q} = \{x, y, z, \theta, \phi\}$, with the coordinates of the two particles being given by

$$\begin{aligned} x_1 &= x - \alpha_1 l \sin\theta \cos\phi \\ y_1 &= y - \alpha_1 l \sin\theta \sin\phi \\ z_1 &= z - \alpha_1 l \cos\theta \\ x_2 &= x + \alpha_2 l \sin\theta \cos\phi \\ y_2 &= y + \alpha_2 l \sin\theta \sin\phi \\ z_2 &= z + \alpha_2 l \cos\theta \ , \end{aligned}$$

where $\alpha_1 \equiv m_2/(m_1 + m_2)$ and $\alpha_2 \equiv m_1/(m_1 + m_2)$ (so that (x, y, z) is the centre of mass).

Answer

$$\begin{aligned} \dot{x}_1 &= \dot{x} + \alpha_1 l \dot{\phi} \sin\theta \sin\phi - \alpha_1 l \dot{\theta} \cos\theta \cos\phi \\ \dot{x}_2 &= \dot{x} - \alpha_2 l \dot{\phi} \sin\theta \sin\phi + \alpha_2 l \dot{\theta} \cos\theta \cos\phi \end{aligned} \tag{5.33}$$

Because $m_1\alpha_1 = m_2\alpha_2$ the cross terms cancel when we expand the x-contribution to the kinetic energy,

$$\tfrac{1}{2}m_1\dot{x}_1^2 + \tfrac{1}{2}m_2\dot{x}_2^2 = \frac{m_1 + m_2}{2}\dot{x}^2 + \frac{m_1 m_2 l^2}{2(m_1 + m_2)}(\dot{\phi}\sin\theta \sin\phi - \dot{\theta}\cos\theta \cos\phi)^2 \ . \tag{5.34}$$

Similarly

$$\tfrac{1}{2}m_1\dot{y}_1^2 + \tfrac{1}{2}m_2\dot{y}_2^2 = \frac{m_1 + m_2}{2}\dot{y}^2 + \frac{m_1 m_2 l^2}{2(m_1 + m_2)}(\dot{\phi}\sin\theta \cos\phi + \dot{\theta}\cos\theta \sin\phi)^2 \ , \tag{5.35}$$

and

$$\tfrac{1}{2}m_1\dot{z}_1^2 + \tfrac{1}{2}m_2\dot{z}_2^2 = \frac{m_1 + m_2}{2}\dot{z}^2 + \frac{m_1 m_2 l^2}{2(m_1 + m_2)}\dot{\theta}^2 \sin^2\theta \ . \tag{5.36}$$

Thus, adding, the kinetic energy is

$$T = \frac{m_1 + m_2}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{m_1 m_2 l^2}{2(m_1 + m_2)}(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2)\ . \tag{5.37}$$

This is the kinetic energy of a free particle of mass (m_1+m_2) plus that of a particle with the *reduced mass* $\mu \equiv m_1 m_2/(m_1+m_2)$ constrained to move on the surface of a sphere of radius l. [Cf. eq. (5.17) of the Coriolis problem with Ω set to zero, and recalling that we are here using spherical polar coordinates where the origin of θ is at the pole, in contrast to the latitude angle whose origin was at the equator—hence $\sin\theta$ and $\cos\theta$ are interchanged.] This is a consequence of the well-known result that the kinetic energy of a system of particles is the kinetic energy of the centre of mass plus the kinetic energy in the centre-of-mass frame.

The potential energy is

$$V = (m_1 z_1 + m_2 z_2)g = (m_1 + m_2)zg\ , \tag{5.38}$$

and the Lagrangian is $L = T - V$. We see from eqs. (5.37) and (5.38) that, by using the centre of mass as the origin for our polar coordinates, we have made L ignorable in x, y and ϕ and that the centre-of-mass coordinates x, y, z form a subsystem completely independent from that of the angular coordinates.

Problem 3.5.1(b)

Use this Lagrangian to construct the Hamiltonian for the system. Is it of the form $T + V$?

Answer

Answering the second part of the question first, the Hamiltonian must be of the form $T + V$ because the constraint $|\mathbf{r}_2 - \mathbf{r}_1| = l$ does not depend explicitly on time and thus the system comes under the general case treated in Sec. 3.2.6. From eq. (5.37) the mass matrix is

$$\boldsymbol{\mu} = \begin{bmatrix} M & 0 & 0 & 0 & 0 \\ 0 & M & 0 & 0 & 0 \\ 0 & 0 & M & 0 & 0 \\ 0 & 0 & 0 & \mu l^2 & 0 \\ 0 & 0 & 0 & 0 & \mu l^2 \sin^2\theta \end{bmatrix}\ , \tag{5.39}$$

where $M \equiv m_1 + m_2$ is the total mass and μ is the reduced mass defined earlier. This is trivially inverted to give, from Sec. 3.2.6, eq. (3.27),

$$H = \frac{p_x^2}{2M} + \frac{p_y^2}{2M} + \frac{p_z^2}{2M} + \frac{p_\theta^2}{2\mu l^2} + \frac{p_\phi^2}{2\mu l^2 \sin^2\theta} + Mgz \; . \tag{5.40}$$

Note: Because x, y and ϕ are cyclic (ignorable) coordinates, p_x, p_y and p_ϕ are constants of the motion. The motion of the rod can thus be found by studying two independent, one-dimensional dynamical problems in z and θ, with Hamiltonians $p_z^2 + Mgz$ and $p_\theta^2/2\mu l^2 + p_\phi^2/2\mu l^2 \sin^2\theta$ respectively. The z-dynamics is just uniformly accelerated motion, with solution $z = z_0 + (p_{z,0}/m)t - \frac{1}{2}gt^2$, $p_z = p_{z,0} - mgt$. A qualitative understanding of the θ-dynamics can be had by considering that the "pseudo-potential" $p_\phi^2/2\mu l^2 \sin^2\theta$ forms a well about $\theta = \pi/2$ (assuming $p_\phi \neq 0$) with sides becoming infinitely high at $\theta = 0$ and π. Thus θ is an oscillatory function of time. (In fact we know from elementary physics that the motion can be integrated in closed form because the rod spins at right angles to a constant angular momentum vector with arbitrary, fixed direction.)

Problem 3.5.1(c)

Find the Hamiltonian corresponding to the Coriolis Lagrangian in Problem 2.8.1,

$$L = \tfrac{1}{2}R^2\left[\dot\theta^2 + \cos^2\theta\,(\Omega + \dot\phi)^2\right] - V(\theta,\phi) \; .$$

Is it of the form $T + V$?

Answer

From the above,

$$p_\theta \equiv \frac{\partial L}{\partial \dot\theta} = R^2\dot\theta \tag{5.41}$$

so

$$\dot\theta = \frac{p_\theta}{R^2} \; . \tag{5.42}$$

Also

$$p_\phi \equiv \frac{\partial L}{\partial \dot\phi} = R^2\cos^2\theta\,(\Omega + \dot\phi) \; . \tag{5.43}$$

so

$$\dot\phi = \frac{p_\phi}{R^2\cos^2\theta} - \Omega \; . \tag{5.44}$$

Thus

$$\begin{aligned} H &\equiv p_\theta\dot\theta + p_\phi\dot\phi - L \\ &= \frac{p_\theta^2}{R^2} + \frac{p_\phi^2}{R^2\cos^2\theta} - \Omega p_\phi - \frac{p_\theta^2}{2R^2} - \frac{p_\phi^2}{2R^2\cos^2\theta} + V(\theta,\phi) \\ &= \frac{p_\theta^2}{2R^2} + \frac{p_\phi^2}{2R^2\cos^2\theta} - \Omega p_\phi + V(\theta,\phi)\ . \end{aligned} \tag{5.45}$$

Given that

$$\begin{aligned} T &= \tfrac{1}{2}R^2\left[\dot\theta^2 + \cos^2\theta\,(\Omega+\dot\phi)^2\right] \\ &= \frac{p_\theta^2}{2R^2} + \frac{p_\phi^2}{2R^2\cos^2\theta} \end{aligned} \tag{5.46}$$

it is clear that the Hamiltonian differs from $T+V$ by the term $-\Omega p_\phi$.

Problem 3.5.2

Find the Hamiltonian corresponding to the Lagrangian of the anharmonic oscillator in Problem 2.8.2. Sketch contours of $H(x,p)$ in the soft potential case for (a) $H < m\omega_0^2 l_0^2/8$ and (b) $H > m\omega_0^2 l_0^2/8$. Indicate representative trajectories of the system in phase-space. Try to include representatives of each qualitatively distinct orbit.

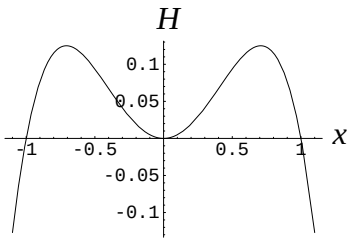

Figure 5.2: $H(x,0)$.

Answer

The Lagrangian is $L = \frac{1}{2}m\dot x^2 + V(x)$, which is in the standard form treated in Sec. 3.2.3. Thus $\dot x = p/m$ and the Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{m\omega_0^2}{2}\left(x^2 - \frac{x^4}{l_0^2}\right)\ . \tag{5.47}$$

Before sketching the phase space we first get a qualitative feel for its salient features by looking at the behaviour of H as a function of x and p and finding its stationary points, i.e. where $\partial H/\partial x = \partial H/\partial p = 0$. (From Hamilton's equations of motion we see that such points correspond to $\dot x = \dot p = 0$ and are thus equilibrium points, where the system does not move—fixed

points of the dynamical system. Points where the Hamiltonian is a minimum are stable, whereas saddle points correspond to unstable equilibria.)

For any fixed x the graph of H is a parabola, concave upwards, with its minimum at $p = 0$. Thus H approaches $+\infty$ as $|p| \to \infty$.

We now examine the behaviour of H in the limits of large and small x. For $|x| \ll l_0$ the term quartic in x is very small compared with the quadratic term so the graph of H is a parabola, concave upwards, with its minimum at $x = 0$.

On the other hand, for $|x| \gg l_0$ the negative quartic term dominates over the quadratic term and H approaches $-\infty$ as $|x| \to \infty$. The behaviour of H on the x-axis is graphed in Fig. 5.2, using units such that $l_0 = \omega_0 = m = 1$. It is seen that there are maxima of H at finite values of $|x|$. The maxima and minima of H as a function of x are the roots of

$$\frac{\partial H}{\partial x} = m\omega_0^2 x \left(1 - \frac{2x^2}{l_0^2}\right) , \tag{5.48}$$

which are at $x = 0$ (the minimum) and $x = \pm l_0/\sqrt{2}$ (the maxima).

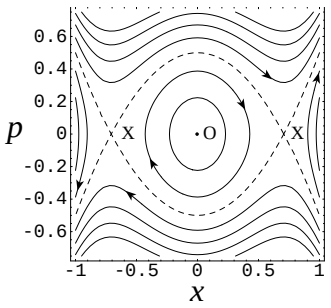

Figure 5.3: Phase space of the anharmonic oscillator.

Since the origin is the intersection of the lines of minima, $x = 0$, $p = 0$, there is a stable equilibrium point at $(x, p) = (0, 0)$. There are mimimax or saddle points at the intersections of the line of minima $p = 0$ and the lines of maxima $x = \pm l_0/\sqrt{2}$. Thus the points $(x, p) = (\pm l_0/\sqrt{2}, 0)$ are also equilibrium points, but are unstable towards small perturbations. At both these saddle points, $H = m\omega_0^2 l_0^2/8$, so that they are linked by the contour $H = \text{const} = m\omega_0^2 l_0^2/8$.

Contours of H in units such that $l_0 = \omega_0 = m = 1$ are plotted in Fig. 5.3. As in Fig. 3.1 we have marked the stable and unstable fixed points by O and X, respectively, and have indicated the contour $H = m\omega_0^2 l_0^2/8$ by a dashed line. Directions of motion (to the right in the upper half plane because $\dot{x} > 0$ and to the left in the lower half plane) are marked by arrows. It is seen that the contour $H = m\omega_0^2 l_0^2/8$ forms the separatrix between five classes of qualitatively different orbit:

- bounded oscillatory (librating) orbits contained within the separatrix, so that $H < m\omega_0^2 l_0^2/8$, corresponding to a particle trapped in the potential well;

- unbounded orbits with $H < m\omega_0^2 l_0^2/8$ which cross the x-axis ($p = 0$) at a value $x < l_0/\sqrt{2}$ or $x > l_0/\sqrt{2}$, corresponding to particles approaching from the left or right, respectively, and being reflected by a potential barrier;

- unbounded orbits with $H > m\omega_0^2 l_0^2/8$, which cross the p-axis ($x = 0$) at a value $x < m\omega_0 l_0/2$ or $x > m\omega_0 l_0/2$, corresponding to particles with sufficient energy to cross the potential barriers (from right to left, or left to right, respectively) without being reflected.

Problem 3.5.3

Consider the motion of a particle of charge e and mass m in a straight, infinitely long magnetic confinement system with vector potential $\mathbf{A} = \psi(x, y)\mathbf{e}_z$.

(a) Show that $B_x = \partial\psi/\partial y$, $B_y = -\partial\psi/\partial x$, $B_z = 0$ and thus that contours of $\psi(x, y)$ define magnetic field lines.

Answer

Given $A_x = A_y = 0$, $A_z = \psi(x, y)$ we calculate $\mathbf{B} \equiv \boldsymbol{\nabla}\times\mathbf{A}$:

$$\begin{aligned} B_x = \partial_y A_z - \partial_z A_y &= \partial_y\psi \ , \\ B_y = \partial_z A_x - \partial_x A_z &= -\partial_x\psi \ . \end{aligned} \tag{5.49}$$

(Alternatively, we can observe that $\mathbf{B} = \boldsymbol{\nabla}\psi\times\mathbf{e}_z$ and take the x and y components.) From eq. (5.49) we have $\mathbf{B}\cdot\boldsymbol{\nabla}\psi = \partial_y\psi\,\partial_x\psi - \partial_x\psi\,\partial_y\psi \equiv 0$. Thus $\psi = $ const on a magnetic field line and so plotting contours of ψ maps out the field.

(b) Show that the Hamiltonian is

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{[p_z - e\psi(x, y)]^2}{2m} \ ,$$

write down the Hamiltonian equations of motion and give two integrals of the motion.

Answer

The given expression for H follows directly by writing eq. (3.2.5) in Cartesian components and using $A_x = A_y = 0$, $A_z = \psi(x, y)$. The Hamiltonian

equations of motion are

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p_x} = \frac{p_x}{m} \,, \\ \dot{y} &= \frac{\partial H}{\partial p_y} = \frac{p_y}{m} \,, \\ \dot{z} &= \frac{\partial H}{\partial p_z} = \frac{(p_z - e\psi)}{m} \end{aligned} \tag{5.50}$$

and

$$\begin{aligned} \dot{p_x} &= -\frac{\partial H}{\partial x} = \frac{e(p_z - e\psi)}{m}\frac{\partial \psi}{\partial x} \,, \\ \dot{p_y} &= -\frac{\partial H}{\partial y} = \frac{e(p_z - e\psi)}{m}\frac{\partial \psi}{\partial y} \,, \\ \dot{p_z} &= -\frac{\partial H}{\partial z} = 0 \,. \end{aligned} \tag{5.51}$$

From the last equation above, p_z is clearly an integral of the motion. Also, since H does not depend explicitly on time, H itself constitutes a second integral of motion as shown in Sec. 3.3.

(c) It can be shown[1] that in any region with no electric current, $\psi(x, y)$ can be represented as the real part of an analytic function of the complex variable $\zeta \equiv x + \mathrm{i}y$. Thus, write $\psi(x, y) = \mathrm{Re}\,\Psi(\zeta)$ and show that the equations of motion for ζ and $p_\zeta \equiv p_x + \mathrm{i}p_y$ are

$$\begin{aligned} \dot{\zeta} &= \frac{1}{m} p_\zeta \,, \\ \dot{p_\zeta} &= \frac{e}{m}\left[p_z - e\mathrm{Re}\,\Psi(\zeta)\right]\left[\Psi'(\zeta)\right]^* \,, \end{aligned}$$

where the prime on Ψ means derivative with respect to its argument, and $*$ means complex conjugate.

Answer

Equation (5.50) implies that $\dot{\zeta} \equiv \dot{x} + \mathrm{i}\dot{y} = (p_x + \mathrm{i}p_y)/m \equiv p_\zeta/m$, as was to be shown.

[1] This is because Ampère's law $\nabla \times \mathbf{B} = \mu_0 \mathbf{j} = 0$ leads to Laplace's equation, $\nabla^2 \psi = 0$, and it is a standard result of complex analysis that the real and imaginary parts of analytic functions obey the two-dimensional Laplace's equation.

We are given that $\psi(x,y) = \mathrm{Re}\,\Psi(\zeta)$. Therefore

$$\begin{aligned} \partial_x\psi &= \mathrm{Re}\,[\partial_x\zeta\,\Psi'(\zeta)] = \mathrm{Re}\,\Psi'(\zeta)\ , \\ \partial_y\psi &= \mathrm{Re}\,[\partial_y\zeta\,\Psi'(\zeta)] = \mathrm{Re}\,[\mathrm{i}\Psi'(\zeta)] \equiv -\mathrm{Im}\,\Psi'(\zeta)\ . \end{aligned} \tag{5.52}$$

Thus, from eq. (5.51),

$$\begin{aligned} \dot{p}_\zeta &= \frac{e}{m}(p_z - e\psi)(\partial_x\psi + \mathrm{i}\partial_y\psi) \\ &= \frac{e}{m}(p_z - e\mathrm{Re}\,\Psi)[\mathrm{Re}\,\Psi'(\zeta) - \mathrm{i}\,\mathrm{Im}\,\Psi'(\zeta)] \\ &\equiv \frac{e}{m}(p_z - e\mathrm{Re}\,\Psi)[\Psi'(\zeta)]^* \end{aligned} \tag{5.53}$$

as was to be shown.

5.4 Chapter 4 Problems

Problem 4.7.1

Given the Hamiltonian for a particle of unit mass on a sphere in nonrotating latitude and longitude coordinates, θ and ϕ', respectively,

$$K = \frac{p_\theta^2}{2R^2} + \frac{p_{\phi'}^2}{2R^2\cos^2\theta} + V'(\theta,\phi',t)\ , \tag{5.54}$$

find the transformation to the $\theta, p_\theta, \phi, p_\phi$ canonical coordinates, where ϕ is the usual latitude measured in the rotating frame defined in Problem 2.8.1. That is, $\phi' = \Omega t + \phi$. Find the Hamiltonian H in the rotating coordinates.

Answer

We treat this as a "point transformation", as in Sec. 4.2.2, with $\boldsymbol{q} = \{\theta,\phi\}$, $\boldsymbol{Q} = \{\theta',\phi'\}$, $\boldsymbol{p} = \{p_\theta, p_\phi\}$ and $\boldsymbol{P} = \{p_{\theta'}, p_{\phi'}\}$.

We thus have the generating function

$$F_2(\theta,\phi,p_{\theta'},p_{\phi'},t) = \theta' p_{\theta'} + \phi' p_{\phi'}\ , \tag{5.55}$$

where $\theta' \equiv \theta$ and $\phi' \equiv \Omega t + \phi$. That is,

$$F_2(\theta,\phi,p_{\theta'},p_{\phi'},t) = \theta p_{\theta'} + (\Omega t + \phi) p_{\phi'}\ . \tag{5.56}$$

[Note that in eq. (5.55) we have committed the usual physicist's "abuse of notation" in taking θ' and ϕ' to denote *functions* of θ and ϕ rather than

independent variables. We avoided this abuse during the formal development in Sec. 4.2.2 by denoting the corresponding function by $\boldsymbol{G}$ rather than $\boldsymbol{Q}$. In practice it is convenient to take it as understood from the context that $\boldsymbol{Q}$ is a function of $\boldsymbol{q}$, rather than generate a new notation such as $\boldsymbol{G}$. This is convenient, but one needs to be aware of what one is doing.]

The transformation equations are, from Sec. 4.2,

$$\begin{aligned} p_\theta &= \frac{\partial F_2}{\partial \theta} = p_{\theta'} \\ p_\phi &= \frac{\partial F_2}{\partial \phi} = p_{\phi'} \\ \theta' &= \frac{\partial F_2}{\partial p_{\theta'}} = \theta \\ \phi' &= \frac{\partial F_2}{\partial p_{\phi'}} = \Omega t + \phi \, , \end{aligned} \tag{5.57}$$

and

$$K = H + \frac{\partial F_2}{\partial t} = H + \Omega p_{\phi'} \, . \tag{5.58}$$

That is, from eq. (5.54),

$$H = K - \Omega p_{\phi'} = \frac{p_\theta^2}{2R^2} + \frac{p_\phi^2}{2R^2 \cos^2\theta} - \Omega p_\phi + V'(\theta, \Omega t + \phi, t) \, . \tag{5.59}$$

With the identification $V'(\theta, \Omega t + \phi, t) = V(\theta, \phi)$ we recognize eq. (5.59) as the same Hamiltonian derived in Problem 3.5.1(c). This derivation makes it clear that the reason H is not of the form $T+V$ is because the transformation to the rotating frame is time dependent.

Problem 4.7.2

Using the harmonic oscillator Lagrangian $L = \frac{1}{2}m(\dot{x}^2 - \omega_0^2 x^2)$
(a) calculate an approximate action integral S by evaluating the action integral exactly using the piecewise-linear trial function

$$x(t) = \frac{1}{\Delta t}[(t_{n+1} - t)x_n + (t - t_n)x_{n+1}]$$

for t in each range $t_n \equiv n\Delta t < t < t_{n+1} \equiv (n+1)\Delta t$.

Answer

In the interval $t_n \le t \le t_{n+1}$ the velocity is constant, $\dot{x} = (x_{n+1} - x_n)/\Delta t$. Thus the contribution to the action from this interval is given by

$$\begin{aligned}
\bar{S}(x_n, x_{n+1}) &\equiv \int_{t_n}^{t_{n+1}} \mathrm{d}t\ L \\
&= \frac{m\Delta t}{2(x_{n+1}-x_n)} \int_{x_n}^{x_{n+1}} \mathrm{d}x\ \left[\frac{(x_{n+1}-x_n)^2}{(\Delta t)^2} - \omega_0^2 x^2\right] \\
&= \frac{m\Delta t}{2(x_{n+1}-x_n)} \left[\frac{(x_{n+1}-x_n)^3}{(\Delta t)^2} - \omega_0^2 \frac{(x_{n+1}^3 - x_n^3)}{3}\right] \\
&= \frac{m}{2}\left[\frac{(x_{n+1}-x_n)^2}{\Delta t} - \frac{\omega_0^2 \Delta t}{3}(x_{n+1}^2 + x_{n+1}x_n + x_n^2)\right] \quad (5.60)
\end{aligned}$$

(b) Show that the approximate action integral from t_{-N} to t_N ($N > 1$ being an arbitrary integer) is stationary for $-N < n < N$ if x_n obeys the second-order difference equation

$$x_{n-1} - 2x_n + x_{n+1} = -\frac{\omega_0^2 (\Delta t)^2}{6}\,(x_{n-1} + 4x_n + x_{n+1})$$

Answer

The total action S is a sum, $\ldots + \bar{S}(x_{n-1}, x_n) + \bar{S}(x_n, x_{n+1}) + \ldots$, in which x_n occurs in two terms. The condition for S to be stationary with respect to variations in x_n is $\partial S/\partial x_n = 0$. From eq. (5.60), after multiplying by $-\Delta t/m$, this may be written

$$\begin{aligned}
&-\frac{\Delta t}{m}\left(\frac{\partial \bar{S}(x_{n-1}, x_n)}{\partial x_n} + \frac{\partial \bar{S}(x_n, x_{n+1})}{\partial x_n}\right) \\
&= -(x_n - x_{n-1}) + \frac{(\omega_0 \Delta t)^2}{6}(2x_n + x_{n-1}) \\
&\quad + (x_{n+1} - x_n) + \frac{(\omega_0 \Delta t)^2}{6}(x_{n+1} + 2x_n) \\
&= x_{n-1} - 2x_n + x_{n+1} + \frac{(\omega_0 \Delta t)^2}{6}(x_{n-1} + 4x_n + x_{n+1}) = 0 \quad \square \quad (5.61)
\end{aligned}$$

(c) Using the approximate action integral $\bar{S}(x_n, x_{n+1})$ evaluated over the range $t_n < t < t_{n+1}$, as a type 1 generating function, $F_1(q, Q) = -\bar{S}(x_n, x_{n+1})$, with $q = x_n$ and $Q = x_{n+1}$, find the linear canonical transformation from $(x_n, p_n \equiv p)$ to $(x_{n+1}, p_{n+1} \equiv P)$. Show that the determinant of the map is unity.

This provides a discrete-time dynamical system approximation to the true continuous-time system for the harmonic oscillator.

Answer

From eq. (5.60), the generating function is

$$F_1(x_n, x_{n+1}) = -\frac{m}{2}\left[\frac{(x_{n+1}-x_n)^2}{\Delta t} - \frac{\omega_0^2 \Delta t}{3}(x_{n+1}^2 + x_{n+1}x_n + x_n^2)\right] . \quad (5.62)$$

Thus, from eqs. (4.16) and (4.17) the transformation is given by

$$\begin{aligned} p_n &= \frac{\partial F_1}{\partial x_n} \\ &= \frac{m}{2}\left[\frac{2(x_{n+1}-x_n)}{\Delta t} + \frac{\omega_0^2 \Delta t}{3}(2x_n + x_{n+1})\right] , \\ p_{n+1} &= -\frac{\partial F_1}{\partial x_{n+1}} \\ &= \frac{m}{2}\left[\frac{2(x_{n+1}-x_n)}{\Delta t} - \frac{\omega_0^2 \Delta t}{3}(2x_{n+1} + x_n)\right] . \end{aligned} \quad (5.63)$$

For further analysis see the following *Maple* printout where it is verified that the transformation matrix has unit determinant. That this must be the case follows both from the general principal that canonical transformations preserve phase-space volume (or, in this case, area) and from the fact that the determinant in this case is also the Poisson bracket between q and p, which is unity and conserved under canonical transformation. Thus we have constructed a *symplectic integrator*—a discrete-time dynamical system that preserves some of the Hamiltonian character of the continuous-time dynamical system.

(d) Iterate the map obtained in (c) 100 times taking the initial point as $x_0 = 1$, $p_0 = 0$ and plot the result. Use units such that $m = \omega_0 = 1$ and use three timesteps: $\Delta t = 0.1$, $\Delta t = 3.45$ and $\Delta t = 3.47$.

For a bonus mark verify one of the solutions by showing that the second-order difference equation derived in (b) above (when arranged to give x_{n+1} in terms of x_n and x_{n-1}) gives the same sequence of x-values as the discrete-time dynamical system in (c) if two successive values of x given by the dynamical system are used to start the second-order difference equation. I.e. the discrete time "Lagrangian" and "Hamiltonian" discriptions are dynamically equivalent. For another bonus mark calculate the eigenvalues for the three values of Δt given above.

Again, see the following *Maple* printout. For $\Delta t = 0.1$ the difference approximation gives results very close to the true continuous-time harmonic

oscillator solution, which is a circle in phase space. For large enough Δt the iteration goes unstable, which occurs somewhere between $\Delta t = 3.45$ and $\Delta t = 3.47$. This can be predicted by evaluating the eigenvalues of the transformation matrix—because the determinant is unity they must satisfy $\lambda_1 \lambda_2 = 1$, which means they are either complex-conjugate pairs on the unit circle in the complex plane (the stable case), or they are real, mutually reciprocal numbers, one of which must be greater than unity, leading to instability (exponential growth).

Chapter 6

References and Index

SOME TEXTS on classical dynamics are listed below.

1. The most referred to text is undoubtedly *Classical Mechanics* (Addison-Wesley, 1st ed. 1953; 2nd ed. 1980) by Herbert Goldstein, although it does not have much on modern developments in nonlinear Hamiltonian dynamics. The call number is QA805.G6 in the Hancock and Earth Sciences libraries.

2. Another classic, concise text is *Mechanics* (Pergamon 1st ed. 1960, 2nd ed. 1969) by L. D. Landau and E. M. Lifshitz; tr. from the Russian by J. B. Sykes and J. S. Bell. The call number is QA805.L283 in the Hancock, Stromlo, Chemistry and Earth Sciences libraries.

3. A clearly written text of similar vintage is *Principles of Mechanics* by J.L. Synge and B.A. Griffiths (McGraw-Hill, New York, 1959). The call number is QA807.S9 in the Hancock, Stromlo, Chemistry and Physical Sciences libraries.

4. A good, more modern text is *Mechanics : from Newton's Laws to Deterministic Chaos* (Springer, 1990) by Florian Scheck. The call number is TA350.S35313 in the Physical Sciences library.

5. Another interesting, though not as carefully written, text is *Dynamics* (Wiley, 1983) by S. Neil Rasband. The call number is QA845.R35 in the Hancock library.

6. A good introductory book is I. Percival and D. Richards, *Introduction to Dynamics* (Cambridge University Press, 1982). By limiting attention to systems with one degree of freedom, the authors have space to devote to the qualitative analysis of nonlinear dynamical systems, including

such modern topics as chaos. The call number is QC133.P46 in the Hancock library.

7. For a great deal of detail on nonlinear dynamics, written in a style easily accessible by physicists, see *Regular and Chaotic Dynamics* (Springer, 2nd ed. 1992; 1st ed. 1983 called *Regular and stochastic motion*) by Allan J. Lichtenberg and Michael A. Lieberman. The call number is QA1.A647 in the Hancock and Stromlo libraries.

8. *Classical Dynamics: a Contemporary Approach* Jorge V. José and Eugene J. Saletan (Cambridge U. P., New York, 1998) is a large modern text. See *Physics Today* May 1999, p. 66, for a favourable review. The call number is QA805.J73 in the Hancock library.

9. *Fundamentals of Applied Dynamics* by James H. Williams (New York, Wiley, c1996) is a large text used in the ANU Engineering Dynamics course. It includes the Lagrangian approach and Hamilton's Principle, but not Hamilton's equations. It is written in an interesting manner with many physical examples having a modern flavour, including treatment of electrical circuits using Lagrangian methods. Despite the use of the word "applied" in the title, it is actually a very scholarly work and contains an unusual amount of thoughtful historical discussion. The call number is QA845.W68 in the Hancock library.

10. A more mathematically oriented book by one of the modern grand masters in the field is *Mathematical Methods of Classical Mechanics* by V.I. Arnol'd, translated from the Russian by K. Vogtmann and A. Weinstein (Springer–Verlag, New York, 1980).

11. An older mathematically oriented classic is *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies* by E.T. Whittaker (Cambridge University Press, Cambridge UK, 1927). QA845.W62 Hancock.

Index

N-body problem, 21

abuse of notation, 93
action, 31
action integral, 24
action integral, phase space, 51
action-angle coordinates, 68
adiabatic invariance, 30, 31
angular momentum, 26, 27, 37
anharmonic oscillator, 38, 84, 89
asymptotic expansion, 32, 62
attractor, 41
Atwood's machine, 27
autonomous system, 37, 50, 66
auxiliary condition, 3, 10
averaged Lagrangian, 30

billiards, 3

canonical coordinates, 44
canonical momentum, 43, 45, 46
canonical system, 44
canonical transformation, 46, 55
  infinitesimal, 63
  time evolution, 64
celestial mechanics, 67
chaos parameter, 73
commutator, 65
configuration space, 2
conservation equation, 36
constant of motion, 38, 81
constraint, 42
  holonomic, 3, 5, 6, 13, 16, 23, 24, 48, 53, 77, 86
  nonholonomic, 4
  time-varying, 20, 26
coordinate-free formulation, 33
Coriolis force, 37, 74, 80, 82, 87, 88, 93
curvature vector, 13, 79
cycle, 41
cyclic coordinate, 67
cyclotron motion, 37, 83

d'Alembert
  equations, 21
degrees of freedom, 2
diffeomorphism, 33
difference equation, 75, 94
distribution function, 70
dynamical system, 41
  continuous time, 41
  discrete time, 41
  Hamiltonian, 43, 71
dynamics, 15

ecliptic, 77
energy integral, 37, 50
energy methods, 17, 45
enthalpy, 30
equation of motion, 15, 20, 23, 25
  Coriolis, 82
  discretization, 64, 73
  Hamiltonian, 44
  harmonic oscillator, 25, 35, 46
  particle in compressible fluid, 29
  particle in e.m. field, 28, 92

Poisson-bracket form, 65
radial, 26
ergodic hypothesis, 71
Euler angles, 15
Euler–Lagrange equation, 8, 23, 51
Eulerian description, 5
expontially small term, 32, 62
extremum condition, 8

field, 5
fixed point, 41, 47
elliptic, 47
hyperbolic, 47
stable, 47
unstable, 47
flow, Hamiltonian, 64
fluid
Euler, 6
ideal, 5, 6, 29, 37, 80
ideal equation of state, 6, 29
test-particle Lagrangian, 29
force, 16
applied, 18
constraint, 16, 18
explicit, 18, 23
friction, 19
implicit, 16
Fréchet derivative, 7
functional derivative, 7

gauge potential, 35
gauge transformation, 34, 35, 46, 56
generalized coordinates, 2
generalized force, 17
generating function, 56, 57
geodesic, 9
curvature, 13, 78
gradient
definition, 8
infinite-dimensional, 8

Hamilton's equations of motion, 44
Hamilton's Principle, 24, 28, 31, 33, 50
Hamilton's principle
modified, 51
Hamiltonian, 1, 43, 44
particle in e.m. field, 48, 91
Hamiltonian system
$(2n+1)$-dimensional, 64
harmonic oscillator, 25, 34, 46
Hessian matrix, 42, 43, 49

ignorable coordinate, 26, 27, 35, 67
infinitesimal generator, 63
inner product, 7
instantaneous frequency, 30
integrability, 68
integral of motion, 26, 36, 38, 50, 66, 81
integrator, numerical, 64
integrator, symplectic, 65
interchange of delta and dot, 7
invariant tori, 68

Jacobian, 5, 43, 69

KAM theorem, 69, 73
kicked rotor, 71
kinematic momentum, 45
kinematics, 1, 15
Korteweg–de Vries equation, 68

Lagrange multiplier, 10, 12
Lagrange's equations, 21, 23
Lagrangian, 1, 23, 41
natural form, 31, 35, 45
particle in e.m. field, 28
particle in ideal fluid, 29
Lagrangian description, 5
Legendre transformation, 45
libration, 47
Lie transform, 66, 67

Liouville's Theorem, 70
Lorentz force, 28

manifold, 3, 13, 78
map
    area preserving, 71
    iterated, 41
    period, 64, 71
    stroboscopic, 64
mass conservation, 6
mass matrix, 48
Maxwellian distribution, 71
metric tensor, 9, 13, 49
motion in a central potential, 25

natural boundary conditions, 9
natural form, 46, 48
Newton's second law, 15, 20, 21, 45
    generalized, 21
Newton's third law, 16, 19
Noether's theorem, 35, 66
nonlinear Schrödinger equation, 68
nullspace, 42

objective functional, 6, 10

particle, 16
path, 3, 4, 6, 58
    physical, 24
pendulum
    physical, 24
    simple, 24
perturbation theory, canonical, 66, 67
phase space, 41, 43, 44
phase-space Lagrangian, 50, 57
physical pendulum, 47
Poincaré section, 64
point transformations, 32, 55, 57, 62
Poisson bracket, 65
polar coordinates, 25, 49
potential
    N-body, 21
    two-body, 21

quantum mechanics, 62
quasiperiodic motion, 30

reduced mass, 87
reduction to first-order ODEs, 42
representation, 2, 33
resonance, 69
rigid rod, 3, 13, 18, 24, 53, 77, 86
rotation, 47

secular behaviour, 47, 61, 68
separatrix, 47
solubility condition, 42
spurious solution, 10
Standard Map, 73
statistical mechanics, 70
subsystem, 16, 87
support of a function, 8
symmetry, 35, 66

test particle, 29
texts on classical dynamics, 99
time reversal, 69
total energy, 37, 45, 46, 48, 50, 87
trial function method, 10, 30, 39, 85

variation
    constrained, 10
    first, 7
    second, 8
variational calculus, 6, 51
variational derivative, 7
vector field, 41
virtual displacement, 17
virtual work, 17