Notes on Analytical Mechanics

Ingemar Bengtsson (2018)

Contents

1 The Best of all Possible Worlds

> Mechanics is the paradise of the mathematical
> sciences, because with it one comes to the fruits
> of mathematics
>
> Leonardo da Vinci

Sir Isaac Newton was a Master of the Mint. He also formulated three celebrated laws of mechanics, which we can paraphrase as follows:

1. A particle not subject to any force moves on a straight line at constant speed.

2. In the presence of a force, the position of a particle obeys the equations of motion

$$m\ddot{x}_i = F_i(x, \dot{x}) \ . \tag{1.1}$$

3. The force exerted by a particle on another is equal in magnitude, but opposite in direction, to the force exerted by the other particle on the first.

A "particle" is here thought of as an entity characterized by its mass m, its location in space, and by nothing else.[1] The aim of Newton's mechanics is to predict the location at arbitrary times, given the position and velocity at some initial time. This is done by means of a solution of the differential equations above.

An overdot denotes differentiation with respect to the time parameter t (this notation, as well as Differential Calculus itself, was invented by Newton), and x_i may denote a vector in three-dimensional space. Sometimes it will be understood that we are dealing with a set of N particles, and moreover we often "suppress indices". Then the force $F_i(x, \dot{x})$ is a $3N$ component function of the $3N$ variables x_i and their $3N$ derivatives $\dot{x}_i$. Since the index notation may be a bit unfamiliar, let me note that whenever indices occur in a formula, it is understood that they can take any of a specified set of integer values.

---

[1] For further discussion see A. Jenkins, *On the title of Moriarty's 'Dynamics of an Asteroid'*, eprint arXiv:1302.5855.

If $i \in \{1, 2, ..., n\}$ then eq. (1.1) stands for n separate equations. Throughout we employ Einstein's summation convention, which means that whenever a certain index occurs twice in a particular term, a sum over all its allowed values is understood, eg.

$$x_i y_i \equiv \sum_{i=1}^{n} x_i y_i = \sum_{j=1}^{n} x_j y_j = x_j y_j \; . \tag{1.2}$$

It does not matter which letter is being used for a repeated index. To avoid confusion, the same index never occurs thrice or more in a single term. In section 8.5 we will introduce index notation in a more sophisticated "tensorial" way, but for the time being this is all there is to it. By the way index notation is not always the best choice—it does not, for instance, make use of any special properties of three dimensional space—but it has the advantage that it can be used for everything, which is why I always use it.

1.1　Analytical and Hamiltonian mechanics

What are we to think of Newton's laws? A physicist might follow Newton in using them to predict the position of the planets as they go around the sun, and will conclude that they are very meaningful. A mathematician might say that they do not say very much, only that the position of a particle is described by a set of ordinary differential equations. A philosopher might object that they say nothing at all—the first and second law together seem to state that a particle moves in a straight line, unless it does something else, in which case we say that it is subject to a force. But the philosopher Kant valued Newton's laws highly, and tried to prove that they are somehow necessary consequences of the way our minds perceive the world, and have a status similar to Euclid's axioms in geometry. Kant overestimated Newton's laws, but they do have content as they stand. It is a highly non-trivial fact that a second order differential equation is being postulated, since this means that the position and the velocity can be chosen arbitrarily at a given instant, but not the acceleration. Moreover, the use of differential equations guarantees that both the past and the future are uniquely determined by the initial values of position and velocity.

　　Analytical mechanics is at once more general and more special than Newton's theory. It is more general because it is more abstract. Its equations do not necessarily describe the positions of particles, but may be applied to much more general physical systems (such as field theories, including Einstein's general relativity theory). In the version we will study it is more special because only a restricted set of forces will be allowed in eq. (1.1). Let us see what kind of restrictions on the function F_i that are of physical interest. Newton's third law is already a restriction. It can be reformulated as the statement that the total *momentum* of a system composed of several particles is conserved:

$$\frac{dP_i}{dt} \equiv \frac{d}{dt} \sum_{particles} m\dot{x}_i = 0 \ , \tag{1.3}$$

where the sum is over all the particles in the system. The existence of such a conserved vector is clearly an interesting fact. By the way this formulation is quite superior when we deal with time dependent masses, say with rockets (see exercise 3). Now consider the function

$$E = T + V = \frac{m\dot{x}^2}{2} + V(x) \ , \tag{1.4}$$

where V is some function of x, known as the *potential energy*. The function T is called the *kinetic energy*, while E itself is the *energy* of the system. Clearly

$$\dot{E} = \dot{x}_i \left(m\ddot{x}_i + \partial_i V(x) \right) \ . \tag{1.5}$$

It follows that if the force is given by

$$F_i(x, \dot{x}) = F_i(x) = -\partial_i V(x) \ , \tag{1.6}$$

then the energy of the system is conserved. Systems for which a conserved energy function exists are called *conservative*. In our example, and indeed in many interesting cases, the energy can be divided into kinetic and potential parts, and the equation of motion is given by

$$m\ddot{x}_i = -\partial_i V(x) \ . \tag{1.7}$$

This move is typical of analytical mechanics, where vectors are usually derived from scalar functions.

Analytical mechanics devises methods to derive the differential equations describing a given system, strategies for solving them, and ways of describing the solutions if they cannot be obtained in explicit form.[2]

We will tentatively restrict ourselves to conservative systems only. If you like this is a strengthening of the third law, and it is believed that all isolated systems in Nature are of this type.

What we are trying to do is to find some properties that all the Laws of Physics, and in particular all allowed equations of motion, have in common. Now the philosopher Leibniz—who was the other of the two inventors of Differential Calculus—argued that we live in the best of all possible worlds. Is it evident from eq. (1.7) that this is so? Indeed it is, as was realized half a century after the publication of Newton's *Principia*. The inspiration came from optics, and the laws of reflection and refraction. It was known that the angle

---

[2] As a definition, this is a little vague. *Méchanique Analitique* was the title of a book written by Lagrange—"the beauty of the method so suiting the dignity of the results, as to make of his great work a kind of scientific poem", to quote Hamilton.

of reflection is equal to the angle of incidence, and it was observed by the Greeks that this implies that light always travels on the shortest path available between two points A and B, subject to the restriction that it should be reflected against the surface. If the angle of reflection were to differ from the angle of incidence, the distance covered by light in going from A to B would be greater than it has to be. For refraction, we have Snell's Law. Any medium can be assigned an index of refraction n, and the angle of refraction is related to the angle of incidence through

$$n_1 \sin \theta_i = n_2 \sin \theta_r \ . \tag{1.8}$$

Fermat noted that if

$$n = \frac{c}{v} \ , \tag{1.9}$$

where v is the velocity of light in the medium and c is a constant (independent of the medium), then Snell's law can be derived from what is now known as *Fermat's principle*, namely that the time taken for light to go from A to B is a minimum. Fermat's principle unifies the laws of refraction and reflection, since it also implies the equality between the angles of incidence and reflection.

More generally the index of refraction may be a function $n(x)$ of position, say through a dependence on temperature. This is what causes mirages. To study this mathematically we imagine that we evaluate the curve integral

$$I = c \int_\gamma dt = \int_\gamma \frac{cds}{v} = \int_\gamma n(x(s))ds \tag{1.10}$$

along an arbitrary path $\gamma(s)$ between A and B. Then the path actually taken by light in going from A to B through the medium is that specific path which results in the smallest possible value of the integral I. This path may well not be a straight line. The question is how to do the optimization. We will soon come to it.

Is there a similar principle underlying mechanics? Maupertius realized that, at least for systems obeying eq. (1.7), there is.[3] Consider two points A and B, and suppose that a particle starts out at A at time $t = t_1$, and then moves along an arbitrary path from A to B with whatever speed that is consistent with the requirement that it should arrive at B at the time $t = t_2$. In mathematical terms we are dealing with a function $x(t)$ such that

$$x(t_1) = x_A \qquad x(t_2) = x_B \ , \tag{1.11}$$

but otherwise arbitrary. For any such function $x(t)$ we can evaluate the integral

---

[3] It was claimed that Leibniz knew the result before him, but in the resulting priority fight Mauperties was strongly supported by Euler. Historians have since found out that the result was, in fact, first arrived at in an unpublished investigation by Euler—who, unlike some scientists one could mention, never cared strongly about priority as far as he himself was concerned.

$$S[x(t)] = \int_{t_1}^{t_2} dt \ (T - V) = \int_{t_1}^{t_2} dt \ \left(\frac{m\dot{x}^2}{2} - V(x) \right) \ . \qquad (1.12)$$

S is known as the *action*. It is a functional, i.e. a function of a function—the functional $S[x(t)]$ assigns a real number to any function $x(t)$. Note that $S[x(t)]$ is *not* a function of t, hence the square bracket notation. On the other hand it is a function of x_A, x_B, t_1, and t_2, but this is rarely written out explicitly.

The statement, to be verified in the next section, is that the action functional (1.12) has an extremum (not necessarily a minimum) for precisely that function $x(t)$ which obeys the differential equation (1.7). This is known as Hamilton's Principle, or—with less than perfect historical and mathematical accuracy—as the Principle of Least Action.

Hamiltonian mechanics deals with those, and only those, equations of motion which can be derived from Hamilton's Principle, for some choice of the action functional.

This is a much more general class than that given by eq. (1.7), but it does exclude some cases of physical interest. Hamiltonian mechanics forms only a part of analytical mechanics—namely that part that we will focus on.

Note once again what is going on. The original task of mechanics was to predict the trajectory of a particle, given a small set of data concerning its state at some intial time t. We claim that there exists another formulation of the problem, where we can deduce the trajectory given half as much data at each of two different times. So there seems to be a local, causal way of looking at things, and an at first sight quite different global, teleological viewpoint. The claim begins to look reasonable when we observe that the amount of "free data" in the two formulations are the same. Moreover, if the two times t_1 and t_2 approach each other infinitesimally closely, then what we are in effect specifying is the position and the velocity at time t_1, just as in the causal approach.

Why do principles like Fermat's and Hamilton's work? In both cases, we are extremizing a quantity evaluated along a path, and the path actually taken by matter in nature is the one which makes the quantity in question assume an extremal value. The point about extrema—not only minima—is that if the path is varied slightly away from the extremal path, to a path which differs to order ϵ from the extremal one, then the value of the path dependent quantity suffers a change which is of order ϵ squared. At an extremum the first derivative vanishes. In the case of optics, we know that the description of light as a bundle of rays is valid only in the approximation where the wavelength of light is much less than the distance between A and B. In the wave theory, in a way, every path between A and B is allowed. If we vary the path slightly, the time taken by light to arrive from A to B changes, and this means that it arrives out of phase with the light arriving along the first path. If the wavelength is very small, phases from light arriving by different paths will be randomly

distributed, and will cancel each other out through destructive interference. This argument fails precisely for the extremal paths: for them, neighbouring paths take approximately the same time, light from all neighbouring paths will arrive with the same phase, and constructive interference takes place. Thus, whenever the wavelength is negligibly small, it will appear that light always travels along extremal paths.

Only in the twentieth century was it realized that Hamilton's Principle works for the same reason that Fermat's Principle works. Classical mechanics is a kind of geometrical optics limit of a "wave mechanics" of matter, operating in configuration space. But that is another story.[4]

1.2 The calculus of variations

Let us now verify the claim made in the first section, namely that Newton's differential equations, for suitable choices of the dynamical system, are mathematically equivalent to the requirement that a certain functional of all possible paths of the particles should assume an extremum value at the actual trajectory. First we stare at the definition of the action functional:

$$S[x(t)] = \int_{t_1}^{t_2} dt \, \left(\frac{m\dot{x}^2}{2} - V(x) \right) \ . \tag{1.13}$$

How do we find the extrema of such a function of a function? Let us begin with some formal considerations. For a function f of an ordinary number x, it is easy enough to find the extrema. We consider how the function values $f(x)$ change as we change the number x:

$$\delta f(x) \equiv f(x + \delta x) - f(x) = \delta x \partial_x f(x) \ . \tag{1.14}$$

We assume that δx is so small that second order terms can be ignored. If the derivative is zero at the point x, the function has a minimum, or a maximum, or at least an inflection point there. For a function of several variables, the condition for an extremum (a minimum, a maximum, or a saddle point) is that

$$\delta f(x) = \sum_i \delta x_i \frac{\partial f}{\partial x_i}(x_1, \ldots, x_N) = 0 \tag{1.15}$$

for arbitrary choices of the δx_i, which means that all the N partial derivatives have to vanish at the extremal points. Now a functional of a function $x(t)$ can be regarded as a function of an infinite number of variables, say of the Fourier coefficients of the original function. You can also regard t as a label of the infinite number of variables on which the functional depends—a kind of

---

[4] Namely that of quantum mechanics, as was realized by Lanczos, by Klein and—decisively—by Schrödinger.

continuous index—and then what we have to do is to replace the sum in eq. (1.15) with an integral. Like this:

$$\delta S = S[x(t) + \delta x(t)] - S[x(t)] = \int_{t_1}^{t_2} dt \ \delta x(t) \frac{\delta S}{\delta x}(t) \ . \qquad (1.16)$$

We assume that it is possible to bring δS to this form. Then the *functional derivative* of $S[x]$ will be defined as the very expression that occurs to the right in the integrand. The equations of motion, as obtained from Hamilton's Principle, then state that the functional derivative of the action is zero, since the form of the function $\delta x(t)$ is arbitrary.

It remains to be seen if we really can bring δS to this form—if not, we would have to conclude that $S[x(t)]$ is "not differentiable". First of all, note that we are all the time evaluating the action between definite integration limits. Then the extremum, if it exists, will be given by that particular trajectory which starts at the point $x(t_1)$ at time t_1, and ends at the point $x(t_2)$ at time t_2, and for which the functional derivative vanishes. We can make this work for the action functional (1.13). Imagine that we know its value for a particular function $x(t)$, and ask how this value changes if we evaluate it for a slightly different function

$$\tilde{x}(t) = x(t) + \delta x(t) \ , \qquad \delta x(t) = \epsilon f(t) \ , \qquad (1.17)$$

where $f(t)$ is, for the time being, an arbitrary function while ϵ is an infinitesimally small constant. It is important for the following argument that $f(t)$ is arbitrary, or nearly so. That ϵ is "infinitesimally small" simply means that we will neglect terms of quadratic and higher orders in ϵ in the calculation which follows:

$$\delta S = S[\tilde{x}(t)] - S[x(t)] = \int_{t_1}^{t_2} dt \ \left(\frac{m}{2}(\dot{x} + \delta \dot{x})^2 - V(x + \delta x)\right) - S[x(t)]$$

$$= \int_{t_1}^{t_2} dt \ \left(\frac{m}{2}\dot{x}^2 + m\dot{x}\delta \dot{x} - V(x) - \delta x \partial_x V(x) + o(\epsilon^2)\right) - S[x(t)] = \quad (1.18)$$

$$= \int_{t_1}^{t_2} dt \ (m\dot{x}\delta \dot{x} - \delta x \partial_x V(x)) + o(\epsilon^2) \ .$$

The action functional has an extremum at the particular function $x(t)$ for which this expression vanishes to first order in ϵ. What we want to see is what kind of restrictions this requirement sets on the function. To see this, we perform a partial integration

$$\delta S = \int_{t_1}^{t_2} dt \ \left(-\delta x(m\ddot{x} + \partial_x V(x)) + \frac{d}{dt}(m\delta x\dot{x})\right) \ . \qquad (1.19)$$

Unfortunately this is not quite of the form (1.16), due to the presence of the

total derivative in the integrand. Therefore we impose a restriction on the so far arbitrary function $f(t)$ that went into the definition of δx, so that

$$\delta x(t_1) = \delta x(t_2) = 0 \ . \tag{1.20}$$

This is a way of saying that we are interested only in functions $x(t)$ that have certain preassigned starting and end points at specified times. With this restriction, the total derivative in eq. (1.19) goes away. The first term has to vanish for all allowed choices of the functions $\delta x(t)$. After a moment's reflection, we see that this can happen only if the factor multiplying δx in the integrand is zero! Hence we have proved that the action functional has an extremum, among all possible functions obeying

$$x(t_1) = x_A \qquad x(t_2) = x_B \ , \tag{1.21}$$

for those and only those functions which obey

$$m\ddot{x} = -\partial_x V(x) \tag{1.22}$$

at all intermediate points.

So we have proved, in this particular case at least, that Newton's equations of motion can be derived from the condition that a certain action functional shall have an extremum value. Note also that the restrictions that we had to set on the function $x(t)$, eqs. (1.21), make perfect sense. To obtain a definite trajectory it is not enough to impose the equations of motion. It is also necessary to set initial conditions. For differential equations of second order, it is natural to make a choice of $x(0)$ and $\dot{x}(0)$. From the point of view of the action, it is natural to impose the value of $x(t)$ at two different times, which is the same amount of information. It should be noted though that whatever values of $x(0)$ and $\dot{x}(0)$ we choose there is always a unique solution for some range of t, while it is perfectly possible that the equation of motion is such that there is no solution, or several solutions, for a given pair of $x(t_1)$ and $x(t_2)$.

The rest of this course is an elaboration of the contents of this section. If you have not understood everything perfectly yet there is still time!

1.3   How to solve equations

It is one thing to be able to set up equations for a physical system, and perhaps to prove theorems to the effect that a solution always exists and is unique, given suitable initial conditions. Another issue of obvious interest is how to solve these equations, or at least how to extract information from them. What precisely do we mean when we say that a differential equation is "soluble"? Consider, as an exercise, a first order differential equation for a single variable:

$$\dot{x} = f(x) \ , \tag{1.23}$$

where f is some function. This can be solved by means of separation of variables:

$$dt = \frac{dx}{f(x)} \quad \Rightarrow \quad t(x) = \int_c^x \frac{dx'}{f(x')} \ , \tag{1.24}$$

where c is a constant determined by the initial condition. If we do this integral, and then invert the resulting function $t(x)$ to obtain the function $x(t)$, we have solved the equation. We will regard eq. (1.24) as an implicit definition of $x(t)$, and eq. (1.23) is soluble in this sense. This is reasonable, since the manipulations required to extract $t(x)$ can be easily done on a computer, to any desired accuracy, even if we cannot express the integral in terms of elementary functions. But there are some limitations here: It may not be possible to invert the function $t(x)$ except for small times.

Next consider a second order equation, such as the equation of motion for a harmonic oscillator:

$$m\ddot{x} = -ax \ . \tag{1.25}$$

This is a linear equation, and we know how to express the solution in terms of trigonometric functions, but our third example—a pendulum of length l—is already somewhat worse:

$$ml^2\ddot{\theta} = -gml \sin\theta \ . \tag{1.26}$$

Let us therefore approach eq. (1.25) in a systematic fashion, which might yield results also for the pendulum. As a first step, note that any second order differential equation can be rewritten as a pair of coupled first order equations:

$$\dot{p} = -ax \qquad m\dot{x} = p \ . \tag{1.27}$$

The second equation defines the new variable p. Unfortunately coupled first order equations are difficult to solve, except in the linear case when they can be decoupled through a Fourier transformation.

The *number of degrees of freedom* of a dynamical system is defined to be one half times the number of first order differential equations needed to describe the evolution.

It will turn out that, for systems whose equations of motion are derivable from the action principle, the number of first order equations will always be even, so the number of degrees of freedom is always an integer for such systems. A system with n degrees of freedom will be described by a set of $2n$ in general coupled first order equations, and the difficulties one encounters in trying to solve them will rapidly become severe.

In the cases at hand, with one degree of freedom only, one uses the fact that

these are conservative systems, which will enable us to reduce the problem to that of solving a single first order equation. For the harmonic oscillator the conserved quantity is

$$E = \frac{m\dot{x}^2}{2} + \frac{ax^2}{2} \ . \tag{1.28}$$

The number E does not depend on t. Equivalently

$$\dot{x}^2 = \frac{2E}{m} - \frac{a}{m}x^2 \ . \tag{1.29}$$

Taking a square root we are back to the situation we know, and we proceed as before:

$$dt = dx\sqrt{\frac{m}{2E - ax^2}} \quad \Leftrightarrow \quad t(x) = \int_c^x dx'\sqrt{\frac{m}{2E - ax'^2}} \ . \tag{1.30}$$

Inverting the function defined by the integral, we find the solution $x(t)$. The answer is a trigonometric function, with two arbitrary constants E and c determining its phase and its amplitude. For our purposes the trigonometric function is defined by this procedure!

We can play the same trick with the non-linear equation for the pendulum, and we end up with

$$t(\theta) = \int_c^\theta \frac{d\theta'}{\sqrt{\frac{2}{ml^2}(E + gml\cos\theta')}} \ . \tag{1.31}$$

We integrate, and we invert. This *defines* the function $\theta(t)$. We could leave it at that, but since our example is a famous one, we manipulate the integral a bit further for the fun of it. Make the substitution

$$\sin\frac{\theta'}{2} \equiv k\sin\phi' \quad \Rightarrow \quad d\theta' = \frac{2k\cos\phi' \ d\phi'}{\sqrt{1 - k^2\sin^2\phi'}} \ . \tag{1.32}$$

The constant k is undetermined at this stage. The integral becomes

$$t(\theta) = \sqrt{\frac{l}{2g}} \int_c^{\phi(\theta)} \frac{2k\cos\phi'd\phi'}{\sqrt{1 - k^2\sin^2\phi'}\sqrt{\frac{E}{gml} + 1 - 2\sin^2\frac{\theta'}{2}}} \ . \tag{1.33}$$

The integrand simplifies if we choose the constant k such that

$$2k^2 \equiv \frac{E}{gml} + 1 \ . \tag{1.34}$$

One further substitution takes us to our desired standard form;

$$t(\theta) = \sqrt{\frac{l}{g}} \int_c^{\phi(\theta)} \frac{d\phi'}{\sqrt{1 - k^2 \sin^2 \phi'}} = \quad \Big/ \sin \phi' \equiv x' \Big/$$

$$= \sqrt{\frac{l}{g}} \int_c^{x(\theta)} \frac{dx'}{\sqrt{(1 - x'^2)(1 - k^2 x'^2)}} \ . \tag{1.35}$$

Just as eq. (1.30) can be taken as an implicit definition of a trigonometric function, this integral implicitly defines the function $\theta(t)$ as an *elliptic function*. If you compare it with the previous integral (1.30), you see that an elliptic function is a fairly natural generalization of a trigonometric, i.e. "circular", function. Since elliptic functions turn up in many contexts they have been studied in depth by mathematicians. Their work remains relevant, even if Mathematica will plot the solution $\theta(t)$ in no time.

Anyway, the above examples were some of the simplest examples of completely soluble dynamical systems. Just wait till we get to the insoluble ones!

Why did this work at all? The answer is that we had one degree of freedom, and one constant of the motion, namely E. This reduced the problem to that of solving a single uncoupled equation. This suggests a general strategy for solving the equations of motion for a system containing n degrees of freedom, i.e. solving $2n$ coupled first order equations: One must find a set of n constants of the motion with suitable properties, so that the problem reduces to that of computing n integrals. This idea forms the core of the theory of *integrable systems*. It works sometimes, but not very often. As a result the notion of what it means to "solve" a set of differential equations evolved somewhat: a solution might consist, say, of a convergent power series in t. But frequently this strategy also fails. A typical Hamiltonian system will exhibit an amount of "chaotic" behaviour, and there may not exist any effective procedure to generate the long term behaviour of the solutions on a computer. What one has to do then is to find out which questions one can reasonably ask concerning such systems.

Even in situations where one can solve the equations, things may not be altogether simple. Consider two harmonic oscillators, with the explicit solution

$$x = a \cos \left(\omega_1 t + \delta_1 \right) \qquad y = b \cos \left(\omega_2 t + \delta_2 \right) \ . \tag{1.36}$$

The trajectory in the x-y-plane is a *Lissajous figure*. Fig. 1.1 explains how to draw them; further examples are readily produced with a computer. If $\omega_1 = \omega_2$ the trajectory is an ellipse, with circles and straight lines as special cases. More generally, if there exist integers m and n such that

$$m\omega_1 = n\omega_2 \tag{1.37}$$

the trajectory is a closed curve. If there are no such integers the trajectory eventually fills a rectangle densely, and never closes on itself. Now put your-
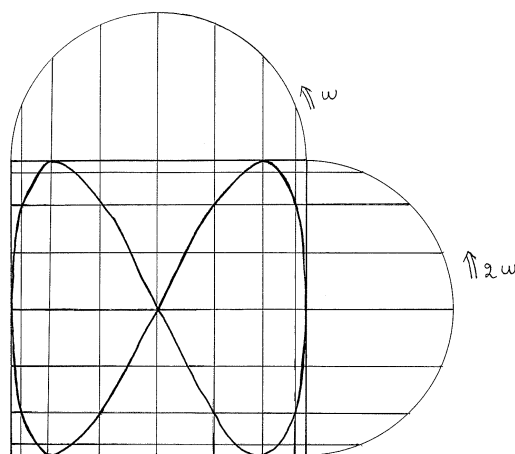
Figure 1.1. A Lissajous figure, and how to draw it; $x = \cos \omega t$, $y = \sin 2\omega t$. You may also enjoy the case $x = \cos \omega t$, $y = \cos n\omega t$, which will give you the graphs of the Chebyshev polynomials $T_n = \cos(n \arccos x)$.

self into the position of an experimentalist trying to determine by means of measurements whether the trajectory will be closed or not!

Another example of this type is a particle moving on a straight line on a plane, but confined to a quadratic box and bouncing elastically from its walls. Let us ask whether the trajectory is periodic or whether it will eventually come arbitrarily close to any point in the box. The answer depends on the initial condition for the direction of motion. If the angle between this direction and one of the walls is called α, the question is whether $\tan \alpha$ is a rational number or not. Theoretically this is fine, but for someone who wants to decide the question by means of measurements of the initial velocity it is not!

1.4 Phase space

It is worthwhile formalizing things a bit further. With the understanding that every set of ordinary differential equations can be written in first order form, we write down the general form of N coupled first order equations for N real variables z_i:

$$\dot{z}_i = f_i(z_1, \ldots, z_N; t) , \qquad 1 \le i \le N , \qquad (1.38)$$

where the N functions f_i are smooth, but otherwise arbitrary. We simplify things by assuming that there is no explicit dependence on time. We then have the equations that describe an *autonomous dynamical system*, namely

$$\dot{z}_i = f_i(z_1, \ldots, z_N) . \qquad (1.39)$$

There are theorems that guarantee the existence and uniqueness of such systems for some range of the parameter t. Thus

$$z_i = z_i(z_{01}, \ldots, z_{0N}, t) , \tag{1.40}$$

where z_{0i} are the initial values of z_i.

There is no guarantee that such solutions can be obtained in any explicit form. If we discretize the time variable a computer can easily generate approximative solutions, but it may be practically impossible to produce accurate solutions over long intervals of time.

We assume that the physical systems we are interested in—as far as we attempt to describe them—can be fully characterized by the N real numbers z_i. We imagine a space whose points are labelled in a one-to-one fashion by these numbers, and call it *phase space*.

The set of all possible states of a physical system is in one-to-one correspondence with the points of an N dimensional phase space. The time development of a system is uniquely determined by its position in phase space.

This is the first of several abstract spaces that we will encounter, and you must get used to the idea of abstract spaces.

A particle moving in space has a 6 dimensional phase space, because its position (3 numbers) and its velocity (3 numbers) at a given time determine its position at all times, given Newton's laws. Anything else can either be computed from these numbers—this is true for its acceleration—or else it can be ignored—this would be true for how it smells, if it does. The particle also has a mass, but this number is not included in phase space because it is given once and for all. Two particles moving in space have a 12 dimensional phase space, so high dimensional phase spaces are often encountered. We will have to picture them as best we may.

Now consider time evolution according to eq. (1.39). Because of the theorems I alluded to, we know that through any point z_0 there passes a unique curve $z_i(t)$, with a unique tangent vector $\dot{z}_i$. These curves never cross each other. When the system is at a definite point in phase space, it knows where it is going. The curves are called *trajectories*, and their tangent vectors define a *vector field* on phase space called the *phase space flow*. Imagine that we can see such a flow. Then there are some interesting things to be observed. We say that the flow has a *fixed point* wherever the tangent vectors vanish. If the system starts out at a fixed point at $t = 0$, it stays there forever. There is an important distinction to be made between *stable* and *unstable* fixed points. If you start out a system close to an unstable fixed point it starts to move away from it, while in the stable case it will stay close forever. The stable fixed point may be an *attractor*, in which case a system that starts out close to the fixed point will start moving towards it. The region of phase space which is close enough for this to happen is called the *basin of attraction* for the attractor.

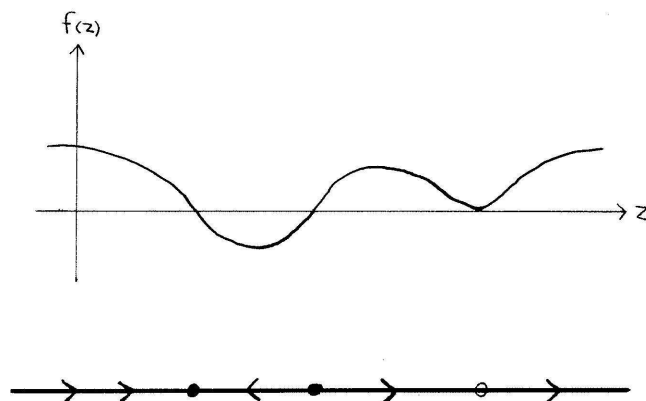Consider a one dimensional phase space, with the first order system

Figure 1.2. A one dimensional phase space, containing one stable and one unstable fixed point, as well as one fixed point which is structurally unstable.

$$\dot{z} = f(z) \; . \tag{1.41}$$

For generic choices of the function f all fixed points are either stable attractors, or unstable repellors, but for special choices of f we can have fixed points that are approached by the flow only on one side. The latter are *structurally unstable*, in the sense that the smallest change in f will either turn them into pairs of attractors and repellors, or cause them to disappear altoghether.

In two dimensions there are more possibilities. We can have *sources* and *sinks*, as well as stable *elliptic* and unstable *hyperbolic* fixed points. To see what the latter two look like, we return to the examples given in section 1.3. The phase space of the harmonic oscillator is $\mathbf{R}^2$, and it contains one elliptic fixed point. It is elliptic because it is surrounded by closed trajectories, and hence it is stable. In the case of the pendulum phase space has a non-trivial topology: since the coordinate θ is a periodic angle phase space is the surface of an infinitely long cylinder. It contains two fixed points. One of them is elliptic, and the other—the state where the pendulum is pointing upwards—is hyperbolic. What is special about the hyperbolic fixed point is that there are two trajectories leading into it, and two leading out of it. The length of the tangent vectors $\dot{\theta}$ decrease as the fixed point is approached. Taking the global structure of phase space into account we see that a trajectory leaving the fixed point is in fact identical to one of the incoming ones. Hence there are really only two special trajectories. A striking fact about them is that they divide phase space into regions with qualitatively different behaviour. One region where the trajectories go around the elliptic fixed point, and two regions where the trajectories go around the cylinder. For this reason the special trajectories are called *separatrices*, and the regions into which they divide phase space are called *invariant sets* —by definition an invariant set in phase space is a region that one cannot leave by following the phase space flow.

Figure 1.3. Fixed points in a two dimensional phase space: a source, a sink, a limit cycle, an elliptic fixed point, and a hyperbolic fixed point.

It is very important that you see how to relate this abstract discussion of the phase space of the pendulum to known facts about real pendula. Do this!

It is not by accident that the phase space of the pendulum is free of sources and sinks. The reason is, as we will see in section 8.1, that only elliptic or hyperbolic fixed points can occur in Hamiltonian mechanics. Real pendula tend to have some amount of dissipation present (because they are imperfectly isolated from the environment), and then the situation changes; see exercise 12. Speaking of Hamiltonian systems it is worthwhile to point out that the example of the two harmonic oscillators in eq. (1.36) is less frivolous than it may appear. The phase space is four dimensional, but there are two conserved quantities

$$2E_1 = p_1^2 + \omega_1^2 x_1^2 \qquad 2E_2 = p_2^2 + \omega_2^2 x_2^2 \ . \tag{1.42}$$

This means that any given trajectory will be confined to a two dimensional surface in phase space, labelled by E_1 and E_2. This surface is a *torus*, with topology $\mathbf{S}^1 \times \mathbf{S}^1$. In a sense to be made precise later, non-chaotic motion in a Hamiltonian system always takes place on a torus in phase space.

Finally we observe that we have the beginnings of a strategy to understand any given dynamical system. We begin by locating the fixed points of the phase space flow. Then we try to determine the nature of these fixed points. If the equations are linear this is straightforward. If not, we can try linearization of the equations around the fixed points. There is a theorem we can lean on here:

The Hartman-Grobman theorem: The nature of the fixed points is unchanged by linearization, as long as the fixed points are isolated and as long as no elliptic fixed points occur.

The caveat in the statement will be explained presently.

Now consider the pendulum. Its phase space is a cylinder described by the coordinates (θ, p_θ). To see if the phase space flow has any fixed points, you set

$$\dot{\theta} = \frac{1}{ml^2} p_\theta = 0 \qquad \dot{p}_\theta = -gml \sin\theta = 0 \ . \tag{1.43}$$

Hence there are fixed points at $(\theta, p_\theta) = (0,0)$ and $(\pi, 0)$. Linearizing around

them you find the former to be elliptic and the latter to be hyperbolic. If this remains true for the non-linear equations you can easily draw a qualitatively correct picture of the phase space flow. No integration is needed.

Were we justified in assuming that the fixed points are elliptic? To see what can go wrong, consider the non-linear equation

$$\ddot{x} + \epsilon x^2 \dot{x} + x = 0 \ . \tag{1.44}$$

In the linearised case ($\epsilon = 0$) there is a single elliptic fixed point. In the non-linear system the flow will actually spiral in or out from the fixed point, depending on the sign of ϵ, so this is an example where the exceptions to the Hartman-Grobman theorem are important. But in the case of the "pure" pendulum we know that the non-linear system is Hamiltonian, and therefore sources and sinks cannot appear—our analysis of the pendulum was therefore accurate.

Our tentative strategy works very well when the phase space is two dimensional, but if the dimension of phase space exceeds two things can get very complicated indeed. A famous example is the at first sight innocent looking *Lorenz equations*

$$\begin{aligned}
\dot{z}_1 &= -az_1 + az_2 \\
\dot{z}_2 &= bz_1 - z_2 - z_1 z_3 \\
\dot{z}_3 &= -cz_3 + z_1 z_2
\end{aligned} \tag{1.45}$$

They capture some aspects of thermal convection in a fluid. The non-linear terms have a dramatic effect, and the slightest change in the initial data will cause the trajectory to go to completely different regions of the three dimensional phase space. In particular Lorenz found by means of a Royal McBee LGP-30 electronic computing machine—an advanced machine at the time— that the system may behave almost periodically for some length of time, then suffer a sudden change so that some quite different periodic behaviour is approximated, followed by a sudden change back to the original quasi-periodic behaviour, and so on.[5] Lorenz was a metereologist interested in the long term accuracy of weather prediction, and used his model to argue that precise very-long-range forecasting may be impossible. The behaviour of the Lorenz equations is *chaotic* in a technical sense to be explained in chapter 10.

$\diamond$ **Problem 1.1**   Newton's Second Law says that the position and the velocity of a particle can be freely specified; then the trajectory $x(t)$, and therefore all derivatives of order higher than one, is determined by the equation of motion. Suppose instead that either

a) only the position can be freely specified, and that the equation of motion determines all the derivatives, or

---

[5] See Fig. 2 in E. N. Lorenz, *Deterministic nonperiodic flow*, J. Atmospheric Sciences **20** (1963) 130.

b) position, velocity and acceleration can be freely specified, and that the equation of motion determines all derivatives of order higher than two.
Discuss these assumptions in the light of Newton's First Law.

◇ **Problem 1.2** A particle of mass $m_1 = 1$ kg and a particle of mass $m_2 = -1$ kg (negative mass) interact with each other according to Newton's Law of Gravity. Describe in qualitative terms the behaviour of the system. Is the energy conserved?

◇ **Problem 1.3** Derive the rocket equation

$$F_i = m\dot{v}_i - \dot{m}u_i \;, \tag{1.46}$$

where F_i is an external force, v_i is the velocity of the rocket, and u_i is the exhaust velocity (relative to the rocket).

◇ **Problem 1.4** Prove Snell's Law of optics, starting from Fermat's principle. Also argue for it using properties of plane waves.

◇ **Problem 1.5** An elastic bar extends between $x = 0$ and $x = L$. It resists bending, has a load per unit length given by $\rho(x)$, and is subject to gravity. We may therefore assume that its energy is given by

$$E = \int_0^L dx \; \left(\frac{k}{2} y'' y'' - \rho(x) y \right) \;,$$

where the slash denotes differentiation with respect to x and k is a constant. The bar will minimize its energy. Analyse the variational problem to see what equation determines the equilibrium position, and what conditions one must impose on the end of the bar in order to obtain a unique solution. Archers want their bows to bend like circles. Conclude that bows must have a value of k that depends on x.

◇ **Problem 1.6** Consider the differential equation $\ddot{x} = \sqrt{x}$, with initial conditions $x(0) = \dot{x}(0) = 0$. Is the solution unique? If not, why is the example pathological?

◇ **Problem 1.7** Is it true that a once differentiable function $x(t)$ is a solution of eq. (1.25) if and only if it is a solution of eq. (1.29)? If not, find a non-trivial counterexample.

◇ **Problem 1.8** Using the general solution for the pendulum, eq. (1.35), solve for $\theta(t)$ in the special case $k = 1$. Physically, what does this solution correspond to?

◇ **Problem 1.9** Use Mathematica to compare the solutions for the pendulum to those of the harmonic oscillator, for various values of the energy (which you adjust so that $E = 0$ corresponds to the stable fixed point in both cases).

◇ **Problem 1.10** Consider a projectile that is fired straight up in a gravitational field ($V = -GM/r$), reaches a maximum height r_{max}, and falls back again. Prove that the solution has the parametric form

$$r = \frac{r_{max}}{2}(1 - \cos\theta) \;, \qquad t = \frac{r_{max}}{2}\sqrt{\frac{r_{max}}{2GM}}(\theta - \sin\theta) \;.$$

Show that the resulting curve in the t–r plane is a cycloid, the curve followed by a point on the perimeter of a circular disk rolling without slipping on the t–axis.

$\diamondsuit$ **Problem 1.11**   Using the construction sketched in Fig. 1.1, draw Lissajous figures for $(x, y) = (\cos \omega t, \cos (2\omega t + \delta))$, for $\delta = 0$, $\pi/4$, $\pi/2$. The first of these is the graph of the Chebyshev polynomial T_2, but what is it usually called?

$\diamondsuit$ **Problem 1.12**   Linearize the pendulum around its fixed points, and then draw a careful picture of its phase space. Add a friction term to the equation, $\ddot{\theta} + \gamma \dot{\theta} + \sin \theta = 0$, and see in qualitative terms what this does to the phase space flow.

$\diamondsuit$ **Problem 1.13**   Give a simple example where linearization around a fixed point gives an erroneous impression of its nature because the fixed point does not stay isolated.

$\diamondsuit$ **Problem 1.14**   Show that a non-autonomous dynamical system can always be rewritten as an autonomous dynamical system by increasing the dimension of its phase space. Do so in the simplest possible way. Supposing the original non-autonomous system has trajectories that form circles and figures-of-eight. What do they look like in the corresponding autonomous system?

2   Lagrangian mechanics

With the agreement that the action integral is an important object, we give a name also to its integrand, and call it the *Lagrangian*. In the examples that we considered so far, and in fact in most cases of interest, the Lagrangian is a function of a set of n variables q_i and their n first order derivatives $\dot{q}_i$:

$$S[q(t)] = \int_{t_1}^{t_2} dt \; L(q_i, \dot{q}_i) \; . \tag{2.1}$$

We use "q" to denote the coordinates because the Lagrangian formalism is very general, and can be applied to all sorts of systems where the interpretation of the variables may differ from the interpretation of "x" as the position of some particle. The space on which q_i are the coordinates is called the *configuration space*. Its dimension is one half that of phase space. It is an intrinsic property of the physical system we are studying, and is a very useful concept. You should try to think as much as possible in terms of the configuration space itself, and not in terms of the particular coordinates that we happen to use (the qs), since the latter can be changed by coordinate transformations. In fact one of the advantages of the Lagrangian formalism is that it is easy to perform coordinate transformations directly in the Lagrangian. We will see examples of this later on. Moreover there are situations—such as that of a particle moving on a sphere—when several coordinate systems are needed to cover the whole configuration space. The important thing is the sphere itself, not the coordinates that are being used to describe it. Which is not to say that coordinates are not useful in calculations—they definitely are!

The pair $(q, \dot{q})$ determines a tangent vector (to some curve) at the point whose coordinate is q. Taken together q and $\dot{q}$ are coordinates on the *tangent bundle* of configuration space. The dimension of the tangent bundle equals that of phase space.

2.1   The scope of Lagrangian mechanics

Among all those functions $q_i(t)$ for which $q_i(t_1)$ and $q_i(t_2)$ are equal to some arbitrarily prescribed values, the action functional has an extremum for precisely those functions $q_i(t)$ which obey the *Euler-Lagrange equations*

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \ , \tag{2.2}$$

provided such functions exist. This is straightforward to verify by means of the calculus of variations; indeed (suppressing indices)

$$\delta S = \int_{t_1}^{t_2} dt \left(\delta q \frac{\partial L}{\partial q} + \delta\dot{q}\frac{\partial L}{\partial \dot{q}}\right) = \int_{t_1}^{t_2} dt \left[\delta q \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right) + \frac{d}{dt}\left(\delta q \frac{\partial L}{\partial \dot{q}}\right)\right] . \tag{2.3}$$

The total derivative term gives rise to a boundary term that vanishes because we are only varying functions whose values at t_1 and t_2 are kept fixed, so that δq is zero at the boundary. The Euler-Lagrange equations follow as advertized. The question is to what extent the equations of motion that actually occur in physics are of this form.

There are some that cannot be brought to quite this form by any means, including some of considerable physical interest; most of them involve dissipation of energy of some sort. An example is that of a white elephant sliding down a hillside covered with flowers.[1] But then frictional forces are not fundamental forces. A complete description of the motion of the elephant would involve the motion of the atoms in the elephant and in the flowers, both being "heated" by friction. It is believed that all complete, fundamental equations are derivable from Hamilton's principle, and hence that they fall within the scope of Lagrangian mechanics—or of quantum mechanics, which is structurally similar in this regard.

Generally speaking we expect Lagrangian mechanics to be applicable whenever there is no dissipation of energy. For many simple mechanical systems the Lagrangian equals the difference between the kinetic and the potential energy,

$$L(x, \dot{x}) = T(\dot{x}) - V(x) \ . \tag{2.4}$$

Exercise 10 will tell you exactly when this holds. An example is

$$L = \frac{m\dot{x}^2}{2} - V(x) \quad \Rightarrow \quad \frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = -\frac{\partial V}{\partial x} - m\ddot{x} \ . \tag{2.5}$$

Even in some situations where there is no conservation of energy, analytical mechanics applies. The simplest examples involve Lagrangians which depend explicitly on the time t. Dissipation is not involved because we keep careful track of the way that energy is entering or leaving the system.

Now for an example where the Lagrangian formalism is useful. Suppose we wish to describe a free particle in spherical polar coordinates

$$x = r \cos\phi \sin\theta \qquad y = r \sin\phi \sin\theta \qquad z = r \cos\theta \ . \tag{2.6}$$

That is to say, we wish to derive the equations for $\ddot{r}$, $\ddot{\theta}$, and $\ddot{\phi}$. This requires

---

[1] This problem was first considered by Eddington. See, however, exercise 3.

an amount of calculation, but the amount shrinks if we perform the change of variables directly in the Lagrangian:

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right) \ . \tag{2.7}$$

Then we obtain the answer as the Euler-Lagrange equations from this Lagrangian. (Do the calculation both ways, and see!) This is often the simplest way to perform a coordinate transformation even if the Lagrangian is not known, so that one first has to spend some time in deriving it.

A famous example for which $L \neq T - V$ is that of an electrically charged particle moving in an external electromagnetic field. This example is so important that we will give it in some detail. First of all, "external" signifies that we are dealing with an approximation, in which we ignore that the presence of the electrically charged particle will affect the electromagnetic field in which it moves. In many situations, this is an excellent approximation. The equations of motion to be derived are the Lorentz equations

$$m\ddot{x}_i = e\left(E_i(x,t) + \epsilon_{ijk}\dot{x}_j B_k(x,t)\right) \ . \tag{2.8}$$

The *epsilon tensor* occurring here may be unfamiliar (but see exercise 1). For the moment let me just say that the second term on the right hand side means the cross product of the velocity and the magnetic field. With this hint you should be able to follow the argument at least in outline, so we proceed. This example is more tricky than the previous ones, since the force depends not only on the position but also on the velocity of the particle (as well as explicitly on time, but this is no big deal). It turns out that in order to derive the Lorentz equation from a Lagrangian, we need not only one but four potentials, as follows:

$$E_i(x,t) = -\partial_i\phi(x,t) - \partial_t A_i(x,t) \qquad B_i(x,t) = \epsilon_{ijk}\partial_j A_k(x,t) \ . \tag{2.9}$$

Here ϕ is known as the scalar potential and A_i as the vector potential. (They are both parts of a relativistic four vector.) It is possible to show that the following action yields the Lorentz equation when varied with respect to x:

$$S[x(t)] = \int dt \ \left(\frac{m\dot{x}^2}{2} + e\dot{x}_i A_i(x,t) - e\phi(x,t)\right) \ . \tag{2.10}$$

Please verify this!

If we consider a time independent electric field with no magnetic field present, the Lorentz equation reduces to the more familiar form

$$m\ddot{x}_i = -e\partial_i\phi(x) \ . \tag{2.11}$$

This has the same form as Newton's Law of Gravity, if the potential is specified correctly. The reason why the full Lorentz equation is much more complicated has to do with the special relativity theory. The magnetic field is a relativistic

complication. The relativistic version of Newton's law of gravity is yet more complicated, and is given by Einstein's general relativity theory.

An important difference between gravity and electricity, also in the non-relativistic case, is that particles couple to gravity through the mass, and all particles have mass while only some have electric charge. Moreover the mass serving as "charge" for gravitational forces is the *same* as the mass occurring on the left hand side of Newton's equations.

2.2 Constrained systems

A strength of the Lagrangian formalism is the way it deals with constrained systems. An example of a constrained system is the pendulum with a rigid rod, which featured already in section 1.3. Deriving eq. (1.26) using a Cartesian coordinate system and Newton's methods is not trivial, since we have to figure out the constraint force acting in the direction of the rod. Using the Lagrangian formalism we simply write down the Lagrangian as $L = T - V$, change to polar coordinates, and quickly arrive at the Euler-Lagrange equation in the form of eq. (1.26). For a more complicated system, such as the double pendulum in exercise 5, the latter method wins hands down.

Another problem in the same vein is that of a particle constrained to move on the surface of a sphere

$$x^2 + y^2 + z^2 = 1 \ . \tag{2.12}$$

The action is still given by

$$S[x, y, z] = \int dt \ \frac{m}{2} \left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2 \right) \ . \tag{2.13}$$

This action is a functional of three functions which are constrained to obey the constraint (2.12), and we are going to extremize it only with respect to variations that obey the constraint. In this particular case there is an easy way to proceed. Using spherical polar coordinates the constraint becomes $r = 1$, while the angular coordinates can be varied freely. From eq. (2.7) we see immediately that the Lagrangian describing the motion in the angular directions is

$$L = \frac{m}{2} \left(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2 \right) \ . \tag{2.14}$$

Of course this depended on the lucky accident that we have a good parametrization of the constraint surface. Correct results would also be obtained if we simply solve for

$$z = z(x, y) = \pm \sqrt{1 - x^2 - y^2} \ , \tag{2.15}$$

and insert the result back into the action that describes the free particle, ie.

$$S[x, y] = \int dt \; \frac{m}{2} \left(\dot{x}^2 + \dot{y}^2 + \dot{z}(x, y)^2 \right) \; . \tag{2.16}$$

Now we can vary x and y freely, except that they are not allowed to exceed one in absolute value. The variations in z are now

$$\delta z = \delta x \frac{\partial z}{\partial x} + \delta y \frac{\partial z}{\partial y} \; , \tag{2.17}$$

and the equations of motion can be derived at the expense of some effort.

There are some unavoidable weaknesses here. From eq. (2.16) it appears as if the configuration space were the unit disk in the plane, since x and y are not allowed to take values outside this disk. Or perhaps the configuration space is two copies of the unit disk, since there are two branches of the square root? But the true configuration space is a sphere. What we see is a reflection of the known fact that it is impossible to cover a sphere with a single coordinate system—our equations have only a "local" validity. This kind of difficulties will become more pronounced in the general problem we are heading for: Consider a Lagrangian L_0 defined on an n dimensional configuration space, with coordinates $q_1, \ldots, q_n$, and suppose that the system is confined to live in the $(n - m)$ dimensional submanifold defined by the m conditions

$$\Phi_I(q_1, \ldots, q_n) = 0 \; , \qquad 1 \leq I \leq m \; . \tag{2.18}$$

Derive equations of motion consistent with this requirement. One way to do this is to solve for m of the qs, $q_1, \ldots, q_m$ say, by means of the m conditions (2.18), and insert the result in the action. In general this will be a lot of hard work, and the difficulties we had with coordinatizing the sphere will recur with a vengeance.

The fact that the procedure avoids dealing with the constraint forces is a weakness too. If we try to design a pendulum in such a way that the approximation of a totally rigid rod holds for the kind of motion the pendulum will be subject to, we will want to know how strong the constraint force actually is. The method of *Lagrange multipliers* solves this problem, and at the same time has the advantage that the difficulties with coordinatizing the constraint surface are postponed to a later stage. The claim we will verify is this: Extremizing the action

$$S[q] = \int dt \; L_0(q, \dot{q}) \tag{2.19}$$

using only variations consistent with the constraints (2.18) is equivalent to extremizing the action

$$S[q, \lambda] = \int dt \; L_0(q, \dot{q}) + \lambda_1 \Phi_1(q) + \cdots + \lambda_m \Phi_m(q) \tag{2.20}$$

under arbitrary variations of the functions q and λ. The λs are the Lagrange multipliers, and are treated as new dynamical variables.

Indeed, when the action (2.20) is varied with respect to the λs we obtain the constraints (2.18) as equations of motion. When we vary with respect to the qs the resulting equations will contain the otherwise undetermined Lagrange multipliers, and it not obvious that these equations have anything to do with the problem we wanted to consider. But they do. Consider the analogous problem encountered in trying to find the extrema of an ordinary function $f(q)$ of the n variables q, subject to the m conditions $\Phi(q) = 0$. (Remember suppression of indices!) First suppose that we use the constraints to solve for m of the qs—it will not matter which ones—and call them y, leaving $n - m$ independent variables x. The extrema of $f(q)$ may be found through the equations

$$0 = \delta f = \delta x \partial_x f + \delta y \partial_y f \ , \tag{2.21}$$

where, however, the variations δy are not independent variations, but have to be consistent with the constraints. In fact they are linear function of the δxs, given by the conditions

$$0 = \delta \Phi = \delta x \partial_x \Phi + \delta y \partial_y \Phi \ . \tag{2.22}$$

This equation has to be solved for δy and the result inserted into eq. (2.21), which is therefore really an expression of the form $\delta x(\partial_x f + \text{something else}) = 0$. It does not imply $\partial_x f = 0$.

Since $\delta \Phi = 0$ for the variations we consider, nothing prevents us from rewriting eq. (2.21) in the form

$$0 = \delta f = \delta f + \lambda \delta \Phi = \delta x(\partial_x f + \lambda \partial_x \Phi) + \delta y(\partial_y f + \lambda \partial_y \Phi) \ , \tag{2.23}$$

where the λs are arbitrary functions. The δys are still given in terms of the δxs, so it would seem at first sight that we cannot conclude that $\partial_x f + \lambda \partial_x \Phi = 0$, but—and here comes the punch line—in fact we can, provided we choose the so far arbitrary functions λ in such a way that $\partial_y f + \lambda \partial_y \Phi = 0$. Since the division of the qs into xs and ys was arbitrary, we see that the "restricted" way of finding the extrema—making variations consistent with the constraints—is equivalent to solving the $n + m$ equations

$$\Phi(q) = 0 \qquad \partial_q f + \lambda \partial_q \Phi = 0 \tag{2.24}$$

for q and λ. But these are precisely the equations that we obtain from the Lagrange multiplier method, in which we do not care about the constraints while varying the action! In all fairness though, we have not solved the equations, we have just derived them in a convenient way.

As long as the constraints depend only on q (and not on $\dot{q}$) it is straightforward to generalize the argument from functions to functionals. From the action

$$S[q, \lambda] = S_0[q] + \int dt \; \lambda \Phi(q) \tag{2.25}$$

we rederive the constraints, together with the equations of motion

$$\frac{\delta S}{\delta q} = \frac{\delta S_0}{\delta q} + \lambda \partial_q \Phi = 0 \; . \tag{2.26}$$

This is the analogue of the second equation (2.24). Written out, if $L = L(q, \dot{q})$ and if there is only one constraint, this is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} + \lambda \frac{\partial \Phi}{\partial q_i} \tag{2.27}$$

$$\Phi(q) = 0 \; . \tag{2.28}$$

These equations have a simple interpretation. The constraint defines a surface in configuration space. We have modified the unconstrained systems by adding an extra force term $\lambda \partial_i \Phi$ to the equations. This force is directed along the gradient of the constraint function, which means that it acts in a direction orthogonal to the constraint surface. To ensure that the trajectory is confined to the surface we must choose the strength of the force (given by λ) in such a way that this is ensured. In other words, once we have solved these equations we know the strength of the constraint forces.

We have also refrained from committing us to any coordinate system adapted to the specific form of the constraint surface. That this is an advantage becomes evident when we return to the problem of the particle on a sphere, starting from the Lagrangian

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + \lambda\left(x^2 + y^2 + z^2 - 1\right) \; . \tag{2.29}$$

The equations of motion (in these inertial coordinates) are

$$m\ddot{x} = 2\lambda x \qquad m\ddot{y} = 2\lambda y \qquad m\ddot{z} = 2\lambda z \; . \tag{2.30}$$

By inspection we see that there are three constants of the motion,

$$J_x = y\dot{z} - z\dot{y} \qquad J_y = z\dot{x} - x\dot{z} \qquad J_z = x\dot{y} - y\dot{x} \; . \tag{2.31}$$

At this point we go over to polar coordinates, using eqs. (2.6) with $r = 1$. The constants of the motion become

$$J_x = -\dot{\theta}\sin\phi - \dot{\phi}\cos\phi\cos\theta\sin\theta \qquad J_y = \dot{\theta}\cos\phi - \dot{\phi}\sin\phi\cos\theta\sin\theta$$
$$\tag{2.32}$$
$$J_z = \dot{\phi}\sin^2\theta \; .$$

It is possible to check directly, using the equations of motion for θ and ϕ, that these are constants of the motion—but only J_z is "obviously" conserved. The coordinate system (θ, ϕ) somehow "hides" the others.

By the way, the kinetic energy can be expressed as

$$T = \frac{m}{2}\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) = \frac{m}{2}\left(J_x^2 + J_y^2 + J_z^2\right) \ . \tag{2.33}$$

This is the angular momentum squared.

So far we have dealt only with *holonomic* constraints, that is constraints involving the configuration space variables only. But consider a ball moving without friction across a table. The configuration space has five dimensions: the position (x, y) of the centre of mass, and three angular coordinates describing the orientation of the ball. Now suppose instead that the ball rolls without slipping. If we are given the position of the centre of mass as a function of time then the motion of the ball is fully determined, which suggests that x and y are the "true" degrees of freedom, and that the constrained configuration space is two dimensional. But the situation is more complicated than that (and cannot be described by holonomic constraints). It is impossible to solve for the angular coordinates in terms of x and y. Indeed from our experience with such things we know that the orientation of the ball at a given point depends on how it got there. Mathematically there is a constraint relating the velocity of the centre of mass to the angular velocity; the point of contact between ball and table is always momentarily at rest.

Constraints that cannot be expressed as conditions on the configuration space are called *anholonomic*. The Lagrange multiplier method can be generalized to handle some anholonomic constraints as well, but we will not do so here.

2.3   Symmetries

Let us return to Newton's Third Law. It amounts to a restriction on the kind of forces that are allowed in the second law, and implies that there exist a set of constants of the motion, namely the momenta. (The terminology is a little unfortunate, since we will soon introduce something called "canonical momenta". They are indeed identical with the conserved momenta in simple cases, but logically there need be no connection.) Constants of the motion are useful when trying to solve the equations of motion, and Emmy Noether proved a theorem explaining when and why they exist. We present the proof for a Lagrangian of the general form $L = L(q, \dot{q})$, and afterwards we discuss a simple example. Let us say at the outset that the argument is quite subtle.

Consider first an *arbitrary* variation of the action. According to eq. (2.3) the result is

$$\delta S = \int_{t_1}^{t_2} dt \, \delta q \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right) + \left[\delta q\frac{\partial L}{\partial \dot{q}}\right]_{t_1}^{t_2} \ . \tag{2.34}$$

In deriving the equations of motion the variations $\delta q(t)$ were restricted in such a way that the boundary terms vanish. This time we do something different. The variations are left unrestricted, but we assume that the function $q(t)$ that we vary around obeys the Euler-Lagrange equations. Then the only non-vanishing term is the boundary term, and

$$\delta S = \epsilon \left(Q(t_2) - Q(t_1) \right) , \qquad \epsilon Q(t) \equiv \delta q \frac{\partial L}{\partial \dot{q}} . \qquad (2.35)$$

Here and in the following ϵ is the constant occuring in $\delta q = \epsilon f$, where f is an arbitrary function of t. The point is to ensure that there is nothing infinitesimal about Q.

So far nothing has been assumed about the variations. Now suppose that, for the given Lagrangian, there exists a set of variations δq of some specified form

$$\delta q = \delta q(q, \dot{q}) , \qquad (2.36)$$

such that for these special variations

$$\delta S = 0 . \qquad (2.37)$$

It is understood that the Lagrangian is such that eq. (2.37) holds as an identity, regardless of the choice of $q(t)$, for the special variations δq. (Note that, given a Lagrangian, it is not always the case that such variations exist. But sometimes they do.)

Next comes the crux of the argument. Consider variations of the particular kind that makes eq. (2.37) hold as an identity—so that $\delta q = \epsilon f$ is a known function—and restrict attention to $q(t)$s that obey the equations of motion. With both these restrictions in force, we can combine eqs. (2.37) and (2.35) to conclude that

$$0 = \delta S = \epsilon \left(Q(t_2) - Q(t_1) \right) . \qquad (2.38)$$

The times t_1 and t_2 are arbitrary, and therefore we can conclude that $Q = Q(q, \dot{q})$ is a constant of the motion.

What this theorem does for us is to transform the problem of looking for constants of the motion to the problem of looking for variations under which the variation of the action is identically zero. Before we turn to examples we generalize the argument slightly, and state the theorem properly. Thus, suppose that there exists a special form of δq, such that

$$\delta S = \int_{t_1}^{t_2} dt \ \frac{d}{dt} \Lambda(q, \dot{q}) . \qquad (2.39)$$

Here Λ can be any function—the important and unusual thing is that the integrand is a total time derivative. Then the quantity Q, defined by

$$\epsilon Q(q, \dot{q}) = \delta q_i \frac{\partial L}{\partial \dot{q}_i} - \Lambda(q, \dot{q}) \qquad (2.40)$$

is a constant of the motion. This is easy to see along the lines we followed above.

The theorem can now be stated as follows:

Noether's theorem: To any variation for which δS takes the form (2.39), there corresponds a constant of the motion given by eq. (2.40).

We will have to investigate whether Lagrangians can be found for which such variations exist, otherwise the theorem is empty. Fortunately it is by no means empty, indeed eventually we will see that all useful constants of the motion arise in this way.

For now, one example—but one that has many symmetries—will have to suffice. Consider a free particle described by

$$L = \frac{m}{2} \dot{x}_i \dot{x}_i \ . \qquad (2.41)$$

Since only $\dot{x}$ appears in the Lagrangian, we can choose

$$\delta x_i = \epsilon_i \ , \qquad (2.42)$$

where ϵ_i is independent of time. Then the variation of the action is automatically zero, Noether's theorem applies, and we obtain a vector's worth of conserved charges

$$P_i = \frac{\partial L}{\partial \dot{x}_i} = m \dot{x}_i \ . \qquad (2.43)$$

We use the letter P rather than Q because this is the familiar conserved momentum vector whose presence is postulated in Newton's Third Law.

Another set of three conserved charges can be found easily, since

$$\delta x_i = \epsilon_{ijk} \epsilon_j x_k \quad \Rightarrow \quad \delta S = 0 \ . \qquad (2.44)$$

Here ϵ_i is again independent of t, and ϵ_{ijk} is the totally anti-symmetric epsilon tensor. Noether's theorem now implies the existence of another conserved vector, namely

$$L_i = \epsilon_{ijk} x_j \dot{x}_k \ . \qquad (2.45)$$

This is the angular momentum vector.

We know that there is at least one more conserved quantity, namely the kinetic energy. Actually there are several, but the story now becomes a bit more complicated because we have to deal with variations for which the variation of the Lagrangian is a total derivative, as in eq. (2.39), rather than zero. Thus

$$\delta x_i = \epsilon \dot{x}_i \quad \Rightarrow \quad \delta S = \int dt\, \frac{d}{dt}\left(\frac{\epsilon m}{2}\dot{x}^2\right) \; . \tag{2.46}$$

Using eq. (2.40) we obtain the constant of the motion

$$E = \frac{m}{2}\dot{x}_i\dot{x}_i \; . \tag{2.47}$$

This is the conserved energy of the particle. There is yet another conserved quantity that differs from the others in being an explicit function of time—but its total time derivative vanishes since it also depends on the time dependent dynamical variables. Thus

$$\delta x_i = -\epsilon_i t \quad \Rightarrow \quad \delta S = \int dt\, \frac{d}{dt}\left(-m\epsilon_i x_i\right) \; . \tag{2.48}$$

Eq. (2.40) gives the conserved charge

$$Q_i = mx_i - tm\dot{x}_i \; , \tag{2.49}$$

and it is easy to check that its total time derivative vanishes as a consequence of the equations of motion. Our analysis of the free particle ends here, but we will return to it in a moment, to show that the conserved quantities have a clear physical meaing.

What does it all mean? What does it mean for an action $S[q(t)]$ to admit variations $\delta q(t)$ leaving the action unaffected? To see this, select a solution $q(t)$ of the equations of motion. We know that this gives an extremum of the action. Then consider $q'(t) = q(t) + \delta q(t)$, where the variation is of the special kind that leaves the value of the action unchanged. Obviously then $S[q'(t)] = S[q(t)]$, so that the extremum is not an isolated point in the space of all qs, but rather occurs for a set of qs that can be reached from each other by means of iteration of the special variation $\delta q(t)$, In other words, given a particular solution of the equations of motion, we can get a whole set of new solutions if we apply the special variation, without going through the work of solving the equations of motion again. This leads to an important definition:

A symmetry transformation is any transformation of the space of functions $q(t)$ having the property that it maps solutions of the equations of motion to other solutions.

This is not a property of the individual solutions, but of the set of all solutions. The special variations occurring in the statement of Noether's theorem are examples of symmetry transformations. Given the converse of the statement that we proved, namely that any constant of the motion gives rise to a special variation of the kind considered by Noether, we observe that any constant of the motion arises because of the presence of a *symmetry*. (There is a converse of this statement that we will come to in section 8.3.)

Let us interpret the symmetry transformations that we found for the free particle, beginning with eq. (2.42). This is clearly a translation in space. Therefore momentum conservation is a consequence of *translation invariance*. It is immediate that we can iterate the infinitesimal translations used in Noether's theorem to obtain finite translations, and the statement is that given a solution to the equations of motion all trajectories that can be obtained by translating this solution are solutions, too. To be definite, given that $(vt, 0, 0)$ is a solution for constant v, $(a + vt, b, c)$ is a solution too, for all real values of (a, b, c). Translation invariance acquires more content when used in the fashion of Newton's third law, which we can restate as "the action for a set of particles has translation symmetry". For free particles this is automatic. When interactions between two particles are added, the law becomes a restriction on the kind of potentials that are admitted in

$$L = \frac{m_1}{2}\dot{\mathbf{x}}_1^2 + \frac{m_2}{2}\dot{\mathbf{x}}_2^2 - V(\mathbf{x}_1, \mathbf{x}_2) \ . \tag{2.50}$$

(Here I use boldface notation for the vectors because I need subscripts to label the particles. Always adapt notation to the circumstances!) Indeed invariance under (2.42) requires that

$$V(\mathbf{x}_1, \mathbf{x}_2) = V(\mathbf{x}_1 - \mathbf{x}_2) \ . \tag{2.51}$$

This is a strong restriction since the function V now depends on only three variables, as opposed to six in the general case.

Eq. (2.44) expresses the fact that the Lagrangian has rotation symmetry, while eq. (2.46) is an infinitesimal translation in time: Given a solution $x(t)$, the function

$$x'(t) = x(t + t_0) = x(t) + t_0 \dot{x}(t) + o(t_0^2) \tag{2.52}$$

is a solution too. So we can make the elegant summary that conservation of momentum, angular momentum and energy are consequences of symmetries under translations and rotations in space, together with translations in time. Eq. (2.48) expresses invariance under "boosts", since it changes all velocities by a constant amount. The free particle is exceptional because we can reach every solution by a symmetry transformation, starting from any given solution.

Having said all this, it is not true that every symmetry gives rise to a constant of the motion. Discrete symmetries like reflections, that do not arise by iterating an infinitesimal symmetry, are counterexamples, and we will see another in section 4.5.

To sum up, symmetries are important from two quite different points of view. Given the equations they facilitate the search for solutions, but they also facilitate the search for the correct equations (if we believe that they should exhibit a certain symmetry). Noether's theorem is a tool for discovering symmetries, as well as for deducing their corresponding constants of the motion.

◇   **Problem 2.1**   Prove the *epsilon-delta identity*

$$\epsilon_{ijk}\epsilon_{kmn} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm} \ .$$

What does it look like in the "cross product" notation?

◇   **Problem 2.2**   Derive Lorentz' equation from the action (2.10).

◇   **Problem 2.3**   Prove that the equation

$$m\ddot{x} + \gamma\dot{x} + kx = 0$$

cannot be derived from an autonomous Lagrangian, that is to say a Lagrangian that does not explicitly depend on time. Then derive the equation from a Lagrangian of the form $L = L(x, \dot{x}, t)$.

◇   **Problem 2.4**   Consider the Lagrangian

$$L = \frac{1}{2}M_{ij}(q)\dot{q}_i\dot{q}_j \ ,$$

where the matrix elements of M_{ij} depend on the configuration space coordinates, and the matrix is assumed to have an inverse M_{ij}^{-1}. Write down the Euler-Lagrange equations and solve for the accelerations.

◇   **Problem 2.5**   Write down the Lagrangian for a double pendulum. (The rod of the second is attached to the bob of the first. Bobs are heavy, the rigid rods not.) How many constants of the motion can you find?

◇   **Problem 2.6**   Take N positive numbers summing to one, $p_1 + p_2 + \cdots + p_N = 1$. Their geometric mean is defined as $(p_1 p_2 \cdot \cdots \cdot p_N)^{1/N}$. What is the maximum of the geometric mean?

◇   **Problem 2.7**   Check that the result from eq. (2.17) is the same as that obtained from the recipe in eq. (2.22). Beware of changes in notation!

◇   **Problem 2.8**   Consider a particle with kinetic energy $T = m(\dot{x}^2 + \dot{y}^2 - \dot{z}^2)/2$, and constrain it to the hyperboloid $x^2 + y^2 - z^2 = -1$, $z > 0$. Treat this both with coordinates adapted to the hyperboloid and with the Lagrange multiplier method. Show that the kinetic energy is positive and find three constants of the motion.

◇   **Problem 2.9**   In the brachistrone problem one considers a particle sliding along a curve in the x-z-plane (z is vertical) under the influence of gravity. Choose this curve so that the time of descent from $(x, z) = (x_0, z_0)$ to the origin is minimal.

◇   **Problem 2.10**   Given a mechanical system for which you have identified a kinetic energy $T = T(q, \dot{q})$ and a potential energy $V = V(q)$. Let Noether's theorem tell you under what conditions on the function $T(q, \dot{q})$ the equations $E = T + V$ and $L = T - V$ are consistent with each other.

◇   **Problem 2.11**   Consider the Lagrangian

$$L = \frac{1}{2}\dot{q}^2 - \lambda q^n \ ,$$

where λ is a real number and n is an integer. Determine those values of n for which the Lagrangian transforms into a total derivative under

$$\delta q = \epsilon \left(t\dot{q} - \frac{q}{2} \right) \ .$$

This is known as conformal symmetry.

$\diamond$ **Problem 2.12**   Show that the Lagrangian (2.50), under the restriction (2.51), transforms into a total derivative under the transformation

$$\delta \mathbf{x}_1 = \delta \mathbf{x}_2 = -\mathbf{v}t \ . \tag{2.53}$$

Give a physical interpretation of the corresponding Noether charge.

$\diamond$ **Problem 2.13**   For a free particle, consider the action integral

$$S = \int_{t_1}^{t_2} \frac{m}{2}\dot{x}^2 dt \ .$$

Evaluate this integral for an $x(t)$ that solves the equations of motion, and express the answer as a function of t_1, t_2, and the initial and final positions x_1 and x_2. Repeat the exercise for a harmonic oscillator.

3 Interlude: Conic sections

The theory of conic sections was one of the crowning achievements of the Greeks. After Descartes, it has become a habit to think of an ellipse as the set of points that obey

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \ . \tag{3.1}$$

However, the equation that we come across when we solve the gravitational two-body problem is

$$\frac{p}{r} = 1 + e\cos\phi \ . \tag{3.2}$$

If you do not recognize it, the following account may be helpful.

By definition a conic section is the intersection of a circular cone with a a plane. The straight lines running through the apex of the cone are called its generators—and we will consider a cone that extends in both directions from its apex. If you like, it is the set of one dimensional subspaces in a three dimensional vector space. Generically, the plane will intersect the cone in such a way that every generator crosses the plane once, or in such a way that exactly two of the generators miss the plane. Apollonius proved that the intersection is an ellipse in the first case, and a hyperbola in the second. There is a borderline case when exactly one generator is missing. Then the intersection is a parabola. We ignore the uninteresting case when the plane goes through the apex of the cone.

This is all very easy if we use the machinery of analytic geometry. For simplicity, choose a cone with circular base, symmetry axis orthogonal to the base, and opening angle 90 degrees. It consists of all points obeying

$$x^2 + y^2 - z^2 = 0 \ . \tag{3.3}$$

Without loss of generality, the plane can be described by

$$cx + z = d \ . \tag{3.4}$$

A small calculation shows that the intersection of these two surfaces is either
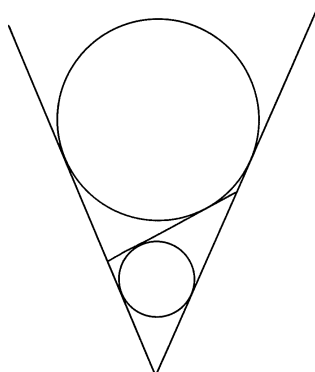
Figure 3.1. A vertical cross section through Apollonius' proof—but to under-
stand the proof you have to think in three dimensions.

an ellipse, a hyperbola, or a parabola—provided you recognize their equations,
as I assume. The section is a circle if $c = 0$ and a parabola if $c = \pm 1$.

It is an interesting exercise to prove this in the style of Apollonius. Let the
cone have arbitrary opening angle. Take the case when the plane intersects
every generator once in the upper half of the cone. Place two spheres inside
the cone, one above and one below the plane, and let them grow until each
touches the plane in a point and the cone in a circle. (See Fig. 3.1.) This clearly
defines the spheres uniquely. Denote the points by F_1 and F_2, and the circles
by C_1 and C_2. Now consider a point P in the intersection of the cone and the
plane. The generator passing through P intersects the circles C_1 and C_2 in
the points Q_1 and Q_2. Now the trick is to prove that the distance PF_1 equals
the distance PQ_1, and similarly the distance PF_2 equals the distance PQ_2.
This is true because the distances measure the lengths of two line segments
that are tangent to the sphere and end at the same point. It then follows that
the sum of the distances PF_1 and PF_2 is constant and equal to the length of
the segment of the generator between the circles C_1 and C_2, independently of
which point P on the intersection we choose. This property defines the ellipse.
This is the proof that the intersection between the cone and the plane is an
ellipse with its foci at F_1 and F_2. If you are unable to see this, consult an old
fashioned geometry book.

For our purposes it is convenient to define the ellipse somewhat differently.
An ellipse of *eccentricity* $e < 1$ can be defined as the set of points whose
distance from a given point, called a *focus*, is e times the distance to a straight
line, called a *directrix*. For the circle $e = 0$, and the directrix is at infinity.

The *latus rectum* of an ellipse is a chord through the focus parallel to the
directrix, and has length $2p$. Now place the origin of a coordinate system at
that focus, with the x-axis pointing towards the directrix. The distance of a
point on the ellipse to the focus is

$$r = e(\text{distance to the directrix}) = e\left(\frac{p}{e} - x\right) \ . \tag{3.5}$$
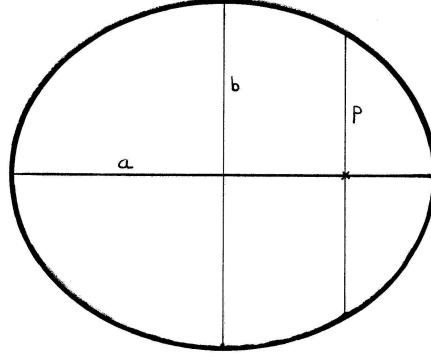
Figure 3.2. An ellipse, with eccentricity 0.58, semi-major axis a, semi-minor axis b, and latus rectum p. The directrix appears on the right. The distance between the centre and the focus is $0.58a$.

(To see this, note that the distance from the focus to the directrix is p/e.) Otherwise expressed

$$r = p - er\cos\phi \quad \Leftrightarrow \quad \frac{p}{r} = 1 + e\cos\phi \ , \tag{3.6}$$

where $\phi = 0$ gives the point closest to the directrix. For a general point on the ellipse we find

$$x^2 + y^2 = r^2 = (p - ex)^2 \quad \Leftrightarrow \quad \frac{(x + ea)^2}{a^2} + \frac{y^2}{b^2} = 1 \ , \tag{3.7}$$

where

$$a \equiv \frac{p}{1 - e^2} \qquad b^2 = pa = (1 - e^2)a^2 \ . \tag{3.8}$$

The *major axis* of the ellipse has length $2a$, and the *minor axis* has length $2b$. The eccentricity is given in terms of these by

$$e^2 = \frac{a^2 - b^2}{a^2} \ . \tag{3.9}$$

Finally the distance between the centre and the focus equals ea. To see this we set $\phi = 0$ and $\phi = \pi$ in eq. (3.6), and calculate

$$\frac{r(\pi) - r(0)}{2} = \frac{p}{2}\left(\frac{1}{1 - e} - \frac{1}{1 + e}\right) = ea \ . \tag{3.10}$$

For some further information consult exercise 1.

Similar treatments can be given for the hyperbola, for which $e > 1$, and for the parabola, for which $e = 1$.

In our study of the two-body problem we will find it interesting to relate an ellipse centred at a focus to an ellipse centred at the origin. The latter is described in Cartesian coordinates by the complex trajectory

$$w(t) = a \cos t + ib \sin t \ . \tag{3.11}$$

The parameter t must not be confused with the angle ϕ between the radius vector and the x-axis. Surprisingly, if we square this ellipse we obtain an ellipse with its focus at the origin. First we see that

$$Z(t) = w^2 = \frac{a^2 + b^2}{2} \cos 2t + iab \sin 2t + \frac{a^2 - b^2}{2} \ . \tag{3.12}$$

Using eq. (3.9) we see that the eccentricity E of the new ellipse is

$$E = \frac{a^2 - b^2}{a^2 + b^2} \ . \tag{3.13}$$

So we can rewrite the equation for the new ellipse as

$$Z(t) = A \cos 2t + iB \sin 2t + EA \ , \tag{3.14}$$

where

$$A \equiv \frac{a^2 + b^2}{2} \ , \qquad B \equiv ab \ , \qquad EA = \sqrt{A^2 - B^2} \ . \tag{3.15}$$

E is the eccentricity of an ellipse with semi-major axis A and semi-minor axis B, and consequently EA is the distance between its focus and its centre. This is again an ellipse, but shifted to be centred at one of its foci—in the sense that the angle is now seen from a focus—and traversed twice as the original ellipse is traversed once. When we put these observations to use in section 4.3, it will be important that every ellipse described on the form (3.14) can be obtained in this way from a function $w(t)$, with suitable choices of a and b.

This trick was introduced by Karl Bohlin, working at the Pulkovo observatory in Russia in 1911. His point was that the transformation $w = \sqrt{Z}$ is not analytic at the origin. This enabled him to deal with collisions between point particles, thought of as limiting cases of elliptical orbits whose eccentricity approaches 1. At the collision the particle presumably reverses its direction, but this is a rather singular occurence. In terms of the variable w it is an undramatic event.

$\diamondsuit$ **Problem 3.1**   Using the definition in terms of the directrix, prove that the sum of the distances from the two foci to a point on an ellipse is constant.

◇   **Problem 3.2**   Prove that the area of an ellipse equals πab. Set up an expression that gives the circumference of the ellipse.

◇   **Problem 3.3**   Place a lamp at one focus of an ellipse, and let the circumference of the ellipse act as a mirror. Prove that all planar light rays reconverge at the same time at the other focus. Try to do this in two different ways: by means of a calculation, and by means of an argument that makes it all obvious.

◇   **Problem 3.4**   For the parabola $y = x^2$, where is its focus? Its directrix?

◇   **Problem 3.5**   Consider a trajectory that is a straight line through the origin in the w-plane. Apply Bohlin's trick. What trajectory results in the Z–plane?

4 The central force two-body problem

Johannes Kepler spent his life pondering the observations of the solar system made by Tycho Brahe, and found that the motion of the planets around the sun follows three simple rules:

1. A planet moves along an ellipse with the sun in one of the foci.
2. The radius vector covers equal areas in equal times.
3. The square of the period of all the planets is proportional to the cube of their major axes.

To appreciate Kepler's work fully, note that there are important facts about the solar system (such as what the distances are) that do not follow simple rules. Moreover the observational data gave the planetary orbits projected on a sphere centred at a point which itself moves along an ellipse around the sun, so it was not obvious that they admitted of a simple description at all.

Kepler was clear about his aims: "My goal is to show that the heavenly machine is not a kind of divine living being but similar to a clockwork in so far as almost all the manifold motions are taken care of by one single absolutely simple magnetic bodily force, as in a clockwork all motion is taken care of by a simple weight."[1] Eventually Newton derived Kepler's laws from his own Laws, with the necessary assumption that the force between the planets and the sun is directed along the radius vector (the force is central) and is inversely proportional to the square of the distance. This remains the number one success story of physics, so we should be clear about why this is so. Naively Kepler's laws may seem simpler than Newton's, but this is not so, for at least two reasons. One is that Newton's laws unify a large body of phenomena, from the motion of planets to the falling of stones close to the Earth. The other reason is that improved observations reveal that Kepler's laws are not quite exact, and the corrections can be worked out mathematically from Newton's laws.

For Mercury (which is hard to observe) the eccentricity $e = 0.21$, for the Earth $e = 0.02$, and for Mars $e = 0.09$. Kepler's main concern was with Mars. If $e = 0$ the ellipse becomes a circle. Unbound motion through the solar system is described by hyperbolas with $e > 1$, but Kepler did not know this.

---

[1] J. Kepler, letter to Herwart von Hohenburg, February 10, 1605.

4.1   The problem and its formal solution

We want to derive Kepler's laws. As an approximation we assume that it suffices to treat the planets independently of each other. Moreover we assume that the precise shapes of the sun and the planets are unimportant and that they can be approximated as being pointlike. Later on, we can go back to these assumptions and see if we can relax them—this will give the corrections referred to above.

Actually, in the *Principia*, Newton proved a comfortable theorem: The gravitational force acting on a particle outside a spherical body of total mass m is identical to that from a particle of mass m sitting at the centre of the body. This is a consequence of the inverse square law, and shows that our second assumption is exact for spherical bodies. Newton also proved that the inverse square law implies that there is no force at all on a particle inside a spherical shell, regardless of the mass of the shell. Concerning the first assumption, a quick estimate shows that the force due to Jupiter acting on Venus is about $2 \cdot 10^{-5}$ times that due to the Sun, and the influence of the other planets is even smaller. So it makes sense to proceed.

We have decided that the configuration space of our problem has six dimensions, spanned by the positions of the sun (X) and one planet (x_P), and we try the Lagrangian

$$L = \frac{M}{2}\dot{X}_i\dot{X}_i + \frac{m_p}{2}\dot{x}_{Pi}\dot{x}_{Pi} - V(X, x_P) \ . \tag{4.1}$$

Depending on the form of the function V we may have to exclude the points $X = x_P$ from the configuration space—we insist that the function V takes finite values only as a function on configuration space. Anyway this will give six coupled second order differential equations. In general they will not admit any simple solutions. Kepler's work implies that the solutions should be simple, so we try to build some symmetries into the problem. We aim for at least six conserved quantities, one for each degree of freedom, since this should result in a soluble problem. Conservation of momentum and energy is already postulated, so we need two more. The answer is rotational symmetry. It is not an objection that an ellipse is not symmetric under rotations. All we need is that if a particular ellipse is a solution, then any ellipse which can be obtained from it by means of rotations is a solution too, even if there is no planet moving along it due to the choice of initial conditions. It might seem that rotational symmetry is overdoing it, since it will yield three conserved quantities, but—for reasons fully explained by the theory of integrable systems, see chapter 10—only two of these are really useful.

With translational and rotational symmetry in place, we find

$$V(X, x_P) = V(X - x_P) = V(|X - x_P|) \ . \tag{4.2}$$

This works for any function V of one variable. Next we introduce coordinates $x_i = x_{Pi} - X_i$ which are invariant under translations, together with coordinates

describing the centre of mass. Then the centre of mass coordinates decouple, and their equations can be solved and set aside. There remains a Lagrangian for a one-body problem, involving only three degrees of freedom:

$$L = \frac{m}{2}\dot{x}_i^2 - V(|x|) \ . \tag{4.3}$$

Here m is the *reduced mass*

$$m = \frac{Mm_p}{M + m_p} \ . \tag{4.4}$$

The reduced mass is almost equal to the mass of the planet since the sun is very heavy in comparison. The coordinate x_i vanishes at the centre of mass of the system, which is well inside the sun, and can be approximately identified with the centre of the sun. This maneouvre should be familiar from elementary mechanics, I just want to emphasize that it is translational symmetry in action.

Rotational symmetry implies the existence of a conserved vector

$$L_i = m\epsilon_{ijk}x_j\dot{x}_k \ . \tag{4.5}$$

This vector is orthogonal to both x and $\dot{x}$, which means that the motion is confined to a plane orthogonal to the angular momentum vector. We are down to a two dimensional configuration space. To take maximal advantage of spherical symmetry we introduce spherical polar coordinates, chosen so that the plane containing the orbit is at $\theta = \pi/2$. The Lagrangian simplifies to

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\phi}^2) - V(r) \ . \tag{4.6}$$

We have used the constant direction of the angular momentum vector. But its magnitude is constant too. This happens because the Lagrangian (4.6) is invariant under translations in the angle ϕ. Using Noether's theorem we find the constant of the motion

$$l = \frac{\partial L}{\partial \dot{\phi}} = mr^2\dot{\phi} \ . \tag{4.7}$$

This equation is of considerable interest in itself. It says that

$$\dot{A} = \frac{r^2\dot{\phi}}{2} = \frac{l}{2m} = \text{constant} \ , \tag{4.8}$$

where $\dot{A}$ is the area covered by the radius vector per unit time. But this is Kepler's Second Law, which therefore holds for all central forces. We are on the right track!

Together with Kepler's second law, energy conservation is enough to solve the problem. Using eq. (4.7) the conserved energy is

$$E = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\phi}^2\right) + V(r) = \frac{m\dot{r}^2}{2} + \frac{l^2}{2mr^2} + V(r) = \text{constant} . \qquad (4.9)$$

This gives the formal solution

$$dt = \frac{dr}{\sqrt{\frac{2}{m}\left(E - \frac{l^2}{2mr^2} - V(r)\right)}} . \qquad (4.10)$$

It may or may not, depending on our choice of $V(r)$, be possible to do the integral in terms of elementary functions, but anyway this equation determines the function $r(t)$, and hence solves the problem. To find $\phi(t)$ we combine eqs. (4.7) and (4.10), and get

$$d\phi = \frac{ldt}{mr^2} = \frac{ldr}{r^2\sqrt{2m\left(E - \frac{l^2}{2mr^2} - V(r)\right)}} . \qquad (4.11)$$

This equation determines $\phi(r(t))$, and the central force problem is thereby fully solved at the formal level. If we are only interested in the form of the orbits, and not the time development, eq. (4.11) is all we need—it will give us $\phi(r)$, and after inversion $r(\phi)$, which is the equation for the form of the orbit.

4.2 Existence and stability of circular orbits

There is a special kind of solution that we can look for directly, with no great expense of effort, namely circular orbits. From the expression (4.9) we see that the two body central force problem has been reduced to one dimensional motion in the *effective potential*

$$V_{eff}(r) = \frac{l^2}{2mr^2} + V(r) . \qquad (4.12)$$

A simple case is $V(r) = 0$, i.e. no force at all. In the effective one dimensional problem this corresponds to a repulsive V_{eff}—the particle comes in from infinity, reaches a minimum value of r, and then disappears to infinity again. If we want a bound orbit the potential $V(r)$ must be attractive.

A circular orbit is one for which $\dot{r} = 0$ identically, which means that the particle is sitting at the bottom of the effective potential—if it does have a bottom. The radius r of the circular orbit must obey

$$V'_{eff}(r) = 0 . \qquad (4.13)$$

If this happens at a local maximum of V_{eff} the solution is unstable, and unlikely to be realized in Nature. The orbit is stable under small perturbations if and only if

$$V''_{eff} > 0 \qquad (4.14)$$

at the value of r for which $V'_{eff} = 0$.

We look into these equations for the choice

$$V(r) = -kr^\beta \quad \Rightarrow \quad V_{eff}(r) = \frac{l^2}{2mr^2} - kr^\beta \ , \qquad (4.15)$$

with β arbitrary (except that we exclude the obviously special case $\beta = -2$). The radius of the circular orbit is found to be

$$r = \left(-\frac{l^2}{\beta k m} \right)^{\frac{1}{\beta+2}} \ . \qquad (4.16)$$

This makes sense only if $l^2 \neq 0$—a question of initial conditions—and $\beta k < 0$. Then the second derivative will be positive if and only if

$$\beta > -2 \ . \qquad (4.17)$$

Thus stability of the circular orbit requires that the force does not fall off too quickly with distance.

There is still the question whether small departures from the circular orbit will give rise to ellipses, or to something more complicated. This is really a question about the ratio between the time it takes for the planet to complete a full revolution in ϕ, to the time it takes to complete a full oscillation in r. If the orbit is an ellipse centred at a focus these times must be equal, and this is likely to happen only for a very special $V(r)$. Compare the discussion of Lissajous figures in section 1.3. For general bounded motion the amount by which the perihelion precesses during one period of the radial motion follows from eq. (4.11). It is

$$\Delta\phi = 2 \int_{r_{\min}}^{r_{\max}} \frac{l\,dr}{r^2\sqrt{2m\left(E - V_{eff}(r)\right)}} \ . \qquad (4.18)$$

For the planets, Kepler's first law requires that $\Delta\phi = 2\pi$.

A final comment: if we do choose the exponent $\beta = -1$ we have the problem that the energy is unbounded from below. We now see that this problem cannot be too serious, because

$$V_{eff}(r) = \frac{l^2}{2mr^2} - \frac{k}{r} \qquad (4.19)$$

is in fact bounded from below whenever $l \neq 0$. The case when $l = 0$ is indeed troublesome from a physical point of view: the two bodies will collide, and we do not have a prescription for what is to happen after the collision. The case of coinciding particles is not included in our configuration space.

4.3 Kepler's First Law

What force laws are consistent with Kepler's First Law? Hooke's law does give elliptical orbits. This is most easily seen by transforming the Lagrangian back to Cartesian coordinates:

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\phi}^2) - kr^2 = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - k(x^2 + y^2) \ . \qquad (4.20)$$

This is two harmonic oscillators of equal frequencies, and the corresponding Lissajous figures are indeed ellipses. They are not the right kind of ellipses however, since they are centred at the origin. Let us call them Hooke ellipses. They can be related to Kepler ellipses (centred at a focus) by means of Bohlin's trick (chapter 3). If

$$w = |w|e^{i\phi} = a\cos t + ib\sin t \qquad (4.21)$$

is a Hooke ellipse, then

$$Z = w^2 = |w|^2 e^{2i\phi} = \frac{a^2 - b^2}{2} + \frac{a^2 + b^2}{2}\cos 2t + iab\sin 2t \ . \qquad (4.22)$$

This is indeed an ellipse centred at a focus, as we saw in chapter 3. For the Hooke ellipse $w(t)$ we know that the force law is

$$\ddot{w} = -w \quad \Rightarrow \quad |\dot{w}|^2 + |w|^2 = 2\epsilon \ . \qquad (4.23)$$

Moreover Kepler's second law holds, so that

$$|w|^2\frac{d\phi}{dt} = \text{constant} \ . \qquad (4.24)$$

This relates the parameter t to the angle ϕ between the radius vector and the major axis. The idea now is to introduce a new time τ, related to t in such a way that Kepler's second law holds also for the ellipse we get when we square the Hooke ellipse. Thus, remembering that the phase of Z is 2ϕ, we require

$$2|Z|^2\frac{d\phi}{d\tau} = \text{constant} \ . \qquad (4.25)$$

A suitable choice of the two constants gives the desired relation

$$\frac{d\tau}{dt} = \frac{|Z|^2}{|w|^2} = |w|^2 \quad \Leftrightarrow \quad \frac{d}{d\tau} = \frac{1}{|w|^2}\frac{d}{dt} \ . \qquad (4.26)$$

At this point then we have obtained a function $Z(\tau)$ that describes motion in accordance with Kepler's First and Second Laws. It only remains to investigate what differential equation it obeys. But using eqs. (4.23) this is a simple calculation:

$$\frac{d^2Z}{d\tau^2} = \frac{1}{|w|^2}\frac{d}{dt}\left(\frac{1}{|w|^2}\frac{dw^2}{dt}\right) = \frac{2}{|w|^2}\frac{d}{dt}\left(\frac{\dot{w}}{\bar{w}}\right) = \cdots = -4\epsilon\frac{Z}{|Z|^3} \ , \qquad (4.27)$$

where ϵ is the constant energy of the Hooke ellipse. This is precisely Newton's force law for gravity. So we conclude that Kepler's First and Second Laws together imply the inverse square law, with the potential

$$V(r) = -\frac{k}{r} \ . \qquad (4.28)$$

We simply calculated the force law, and conclude that Kepler's First and Second Laws hold if and only if the inverse square law holds for the force. The argument is watertight because Bohlin's trick can be used to parametrize every Kepler ellipse in this convenient way.[2]

To confirm our conclusion let us go back to eq. (4.11), which gives a formal solution for the form of the orbit. We choose eq. (4.28) for $V(r)$, and we also perform the substitution

$$u = \frac{1}{r} \quad \Rightarrow \quad du = -\frac{dr}{r^2} \ . \qquad (4.29)$$

The result is

$$d\phi = -\frac{ldu}{\sqrt{2mE - l^2u^2 + 2mku}} \ . \qquad (4.30)$$

This defines u as a trigonometric function of ϕ, with the energy $E < 0$ and the phase ϕ_0 as undetermined integation constants. In fact

$$\phi = \phi_0 - \arccos\frac{\frac{l^2u}{mk} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}} \ . \qquad (4.31)$$

Inverting this, and cleaning up the answer a little, we obtain

$$\frac{l^2}{mkr} = \frac{l^2u}{mk} = 1 + \sqrt{1 + \frac{2El^2}{mk^2}}\cos(\phi - \phi_0) \ . \qquad (4.32)$$

The constant ϕ_0 is the value of the coordinate ϕ for which the planet is at its perihelion, that is when it is closest to the sun. Comparing to eq. (3.6) we read off that the eccentricity of the ellipse is

$$e = \sqrt{1 + \frac{2El^2}{mk^2}} \ . \qquad (4.33)$$

The semi-major axis of the ellipse is

---

[2] For a better description of this beautiful argument, see V. I. Arnold: *Huygens and Barrow, Newton and Hooke*, Springer 1990.

$$a = \frac{p}{1 - e^2} = -\frac{l^2}{mk}\frac{mk^2}{2El^2} = \frac{k}{2|E|} \ . \tag{4.34}$$

The solution remains valid also for $E > 0$, in which case it describes a hyperbola with $e > 1$. Physically this is an unbound trajectory, like that of a spaceship heading for the stars.

4.4 Kepler's Third Law

Kepler's Third Law awaits proof. If you recall the basic facts about ellipses the following is easy: since the areal velocity is constant the period T is simply related to the area of the ellipse. Starting with eq. (4.8) and recalling exercise 3.2 we obtain

$$\frac{l}{2m}T = A = \pi a b \ . \tag{4.35}$$

Remembering that $b^2 = ap$, and using eq. (4.32) to identify the latus rectum p, gives

$$T^2 = \frac{4\pi^2 m^2 a^3 p}{l^2} = 4\pi^2 \frac{m}{k}a^3 \ . \tag{4.36}$$

This is Kepler's Third Law, since $k = GmM$ so that the proportionality constant is independent of what planet we are looking at, to the approximation that we can take M equal to the mass of the Sun.

Here is a more involved proof, using the full force of our solution (4.10). We rewrite it using our expressions for a and e:

$$dt = \sqrt{\frac{ma}{k}}\frac{r dr}{\sqrt{-\frac{l^2}{2m|E|} + 2ar - r^2}} = \sqrt{\frac{ma}{k}}\frac{r dr}{\sqrt{a^2 e^2 - (r - a)^2}} \ . \tag{4.37}$$

We can do the integral if we can find a substitution that simplifies the integral

$$\int^r \frac{r dr}{\sqrt{1 - (r - 1)^2}} \ . \tag{4.38}$$

Such a substitution is readily found. Reinserting the constants we set

$$r = a + ae \cos \sigma \tag{4.39}$$

Now we can do the integral. With a suitable choice of the integration constant we obtain

$$t = \sqrt{\frac{ma^3}{k}}(\sigma + e \sin \sigma) \ . \tag{4.40}$$

Taken together, eqs. (4.39-4.40) provide a parametric representation of the orbit, and we can read off its period

$$T = 2\pi\sqrt{\frac{ma^3}{k}} \ . \tag{4.41}$$

This is Kepler's Third Law once again.

4.5 Self-similarity and the virial theorem

Kepler's Third Law says that, given a solution, one can simply enlarge it to get another solution—provided one also slows down the rate at which things are happening. It is really a consequence of mechanical similarity or *self-similarity*, a kind of symmetry not covered by Noether's theorem. It arises as follows. Take the Lagrangian

$$L = \frac{m}{2}\left(\frac{dq}{dt}\right)^2 - V(q) \ , \tag{4.42}$$

where we do not use the dot notation because we will soon have two different time parameters to reckon with. Assume that the potential is *homogeneous of degree* β, meaning that there exists a real number β such that for any real non-zero number λ

$$V(\lambda q) = \lambda^\beta V(q) \ . \tag{4.43}$$

There could be several variables q_i. For simplicity I write only q. Let us also change the time scale, and define a new function q' by

$$q(t) \to q'(t') = \lambda q(t) \ , \qquad t' = \lambda^{\frac{2-\beta}{2}} t \ . \tag{4.44}$$

It follows that

$$\frac{dq'}{dt'} = \frac{dt}{dt'}\frac{d}{dt}(\lambda q(t)) = \lambda^{\frac{\beta}{2}}\frac{dq}{dt} \ . \tag{4.45}$$

We can now check that our rescalings represent a symmetry because, under this transformation,

$$L\left(q, \frac{dq}{dt}\right) \to L\left(q', \frac{dq'}{dt'}\right) = \lambda^\beta L\left(q, \frac{dq}{dt}\right) \ . \tag{4.46}$$

This has the effect of changing the value of the action with a constant factor, and it follows that $q'(t')$ is an extremum of $S[q'(t')]$ if $q(t)$ is an extremum of $S[q(t)]$. In this sense rescaling is a symmetry of the action. (Compare problem 2.11. If you find the argument difficult, you can check directly that $q'(t')$ is a solution whenever $q(t)$ is.)

The harmonic oscillator has a potential $V \sim q^2$, homogeneous with $\beta = 2$. Scaling symmetry is present with $t' = t$. Given a solution $q(t)$ there is another solution that is a blown up version of this, with amplitude a factor of λ larger. Because $t = t'$ the period of the oscillations are unaffected by the scaling, and we see—without looking at any explicit solutions—that the period of the oscillations are independent of their amplitudes. Galilei first made this observation while celebrating mass in the cathedral of Pisa.

Newton's law of gravity uses a homogeneous potential with $\beta = -1$, so similarity holds with $t \to t' = \lambda^{3/2} t$. Two ellipses with the same shape (and the planetary orbits are all close to circular) will therefore have their periods and their axes related by

$$ R \to R' = \lambda R \qquad T \to T' = \lambda^{3/2} T \quad \Rightarrow \quad \frac{T'^2}{T^2} = \frac{R'^3}{R^3} \; . \tag{4.47} $$

This is Kepler's Third Law for the third time.

Another dramatic theorem can be proved for self-similar systems. It is called the *Virial Theorem*, and relates the time averages of the kinetic and potential energies to each other. If it exists, the time average of a function $f(t)$ is defined by

$$ \langle f \rangle \equiv \lim_{t \to \infty} \frac{1}{t - t_0} \int_{t_0}^{t} dt' \; f(t') \; . \tag{4.48} $$

For the argument to follow it is important that the time average of the derivative of a bounded function is zero, i.e.

$$ \left\langle \frac{df}{dt} \right\rangle = \lim_{t \to \infty} \frac{1}{t} \left(f(t) - f(t_0) \right) = 0 \tag{4.49} $$

whenever $f(t) < \infty$ for all t.

We are ready to study the time average of the kinetic energy, given the assumptions that the system obeys Newton's law

$$ m\ddot{x}_i = -\partial_i V(x) \; , \tag{4.50} $$

that the potential is homogeneous of degree β, that the motion is bounded in space, and that the velocities are everywhere finite. On the other hand we are not restricting the index i. It could run between $1 \le i \le 3N$, in which case we are actually studying an N-body problem; this could be a cluster of galaxies under the tentative assumption that the cluster is a bound system, or it could be 10^{23} atoms confined in a box. For the argument we will need Euler's theorem on homogeneous functions. Regardless of the number of variables it states that

$$ V(\lambda x) = \lambda^\beta V(x) \quad \Rightarrow \quad x_i \partial_i V(x) = \beta V(x) \; . \tag{4.51} $$

(Proof: Take the derivative with respect to λ, and then set $\lambda = 1$.)

Now the calculation is easy:

$$2\langle T \rangle = \langle m\dot{x}^2 \rangle = \left\langle \frac{d}{dt}(mx_i\dot{x}_i) - mx_i\ddot{x}_i \right\rangle = -\langle x_i m\ddot{x}_i \rangle =$$

$$= \langle x_i \partial_i V(x) \rangle = \beta \langle V(x) \rangle \ . \tag{4.52}$$

This is the conclusion we were after.

For bounded motion in homogeneous potentials

$$2\langle T \rangle = \beta \langle V \rangle \ , \tag{4.53}$$

where β is the degree of homogeneity of $V(x)$.

For the inverse square law the virial theorem implies that

$$\langle 2T + V \rangle = 0 \quad \Rightarrow \quad \langle E \rangle = \langle T + V \rangle = -\langle T \rangle \le 0 \ . \tag{4.54}$$

This is the familiar fact that motion bounded by gravity can take place only if the total energy is negative. For the harmonic oscillator we deduce that the time averages $\langle T \rangle$ and $\langle V \rangle$ are equal.

The virial theorem has been used by astronomers to estimate the masses of clusters of stars and clusters of galaxies, assuming that they are gravitationally bound. This led to the first evidence for dark matter.[3] Note also the counterintuitive fact that if energy leaves a self-gravitating cluster of particles its energy $\langle E \rangle$ becomes more negative, which means that its average kinetic energy $\langle T \rangle$ grows. In some sense the "gas" becomes hotter when energy is lost.

The calculation in eq. (4.52) is of interest even for non-potential forces, if we break it off after the first line:

$$2\langle T \rangle = -\langle x_i F_i \rangle \ . \tag{4.55}$$

If the forces are the constraint forces keeping an ideal gas contained inside a box, we can use this relation to deduce the ideal gas law. We turn the sum into an integral, recall the definition of the pressure P as force per unit area, and apply Gauss' law to the result:

$$2\langle T \rangle = P \int dA_i x_i = P \int dV \partial_i x_i = 3PV \ . \tag{4.56}$$

If we are willing to identify $\langle T \rangle$ with (a factor times) the temperature T we obtain

---

[3] The pioneering work on the Coma cluster is described, very readably, in F. Zwicky, *On the Masses of Nebulae, and Clusters of Nebulae*, Astrophys. J. **86** (1937) 217. Zwicky did overestimate the amount of dark matter, largely because he was assuming the cluster to be much closer to us than it actually is, but his methods were sound and his basic conclusion correct.

$$PV = RT \ . \tag{4.57}$$

This is Boyle's Law for ideal gases.

4.6 The three-body problem

The three-body problem—three masses interacting according to Newton's Law of Gravity—is not soluble in the sense that the two-body problem is. The number of conserved quantities is the same in both problems, and for the nine degrees of freedom in the three-body problem this is not enough. But the three-body problem is very important, and in fact motivated many of the developments that we will come to later on.

A natural first step is to look for special exact solutions, which may be used as starting points for perturbation theory, or in other ways. An interesting example was found by Lagrange. Let us begin by assuming that the motion takes place in a plane, and use complex numbers $z_i(t)$ to denote the trajectories. The equations are

$$\begin{aligned}
\ddot{z}_1 &= -m_2 \frac{z_1 - z_2}{|z_1 - z_2|^3} - m_3 \frac{z_1 - z_3}{|z_1 - z_3|^3} \\
\ddot{z}_2 &= -m_3 \frac{z_2 - z_3}{|z_2 - z_3|^3} - m_1 \frac{z_2 - z_1}{|z_2 - z_1|^3} \\
\ddot{z}_3 &= -m_1 \frac{z_3 - z_1}{|z_3 - z_1|^3} - m_2 \frac{z_3 - z_2}{|z_3 - z_2|^2} \ .
\end{aligned} \tag{4.58}$$

We assume that the centre of mass is at rest,

$$m_1 z_1 + m_2 z_2 + m_3 z_3 = 0 \ . \tag{4.59}$$

The particles form a triangle, with sides represented by

$$w_1 = z_3 - z_2 \ , \quad w_2 = z_1 - z_3 \ , \quad w_3 = z_2 - z_1 \ . \tag{4.60}$$

It turns out to be convenient to rewrite the equations of motion in terms of these variables. A small calculation shows them to take the form

$$\begin{aligned}
\ddot{w}_1 &= -m \frac{w_1}{|w_1|^3} + m_1 a \\
\ddot{w}_2 &= -m \frac{w_2}{|w_2|^3} + m_2 a \\
\ddot{w}_3 &= -m \frac{w_3}{|w_3|^3} + m_3 a \ ,
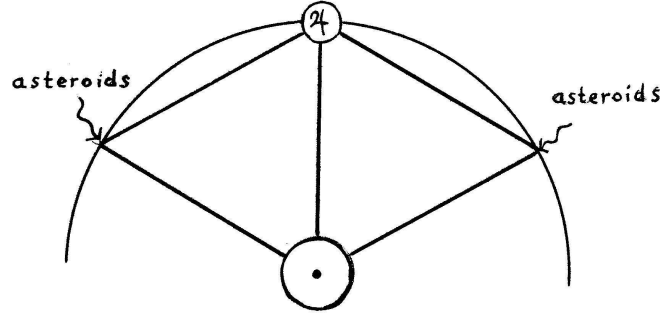\end{aligned} \tag{4.61}$$

where $m = m_1 + m_2 + m_3$ and

Figure 4.1. The two stable Lagrange points in Jupiter's orbit. The Greek and Trojan asteroids lie within roughly one astronomical unit from the Lagrange points.

$$a = \frac{w_1}{|w_1|^3} + \frac{w_2}{|w_2|^3} + \frac{w_3}{|w_3|^3} \ . \tag{4.62}$$

This time we are looking for a special solution, not at the general case. So let us assume that the triangle is an equilateral one,

$$w_2 = e^{2\pi i/3} w_1 \ , \qquad w_3 = e^{4\pi i/3} w_1 \ . \tag{4.63}$$

A glance at eqs. (4.61) shows that this property can be preserved in time. Then we have that $a = 0$, and the only equation we need to solve is

$$\ddot{w}_1 = -m\frac{w_1}{|w_1|^3} \ . \tag{4.64}$$

This we know how to do.

To interpret the solution, solve for

$$m z_1 = m_3 w_2 - m_2 w_3 \quad \Rightarrow \quad m^2|z_1|^2 = (m_2^2 + m_2 m_3 + m_3^2)|w_1|^2 \tag{4.65}$$

and so on. A small calculation then shows that

$$\ddot{z}_1 = -\frac{(m_2^2 + m_2 m_3 + m_3^2)^{\frac{3}{2}}}{m^2}\frac{z_1}{|z_1|^3} \ , \tag{4.66}$$

and similarly for the other two particles. Hence the particles are all being accelerated towards their common center of mass, with "effective masses" that take an unexpected form. Each individual particle travels on an ellipse, but the tree of them do so in unison, in such a way that they always span an equilateral triangle.

A special case of this solution is of considerable physical interest. Let one of the particles have negligibly small mass. Then the remaining pair trace out

the same orbits that they would follow in the absence of the third member. Nevertheless the three particles span an equilateral triangle. This is the origin of the two *Lagrange points* on the orbit of a planet, where small bodies may sit. To draw this conclusion we should also investigate whether the exact solution is stable under small perturbations. This turns out to be the case.

The Lagrange points we have found are called L_4 and L_5, since there is another set of three equilibria on the axis through the two bodies—although they are of less interest since they are unstable. About a thousand asteroids have been found close to the Lagrange points L_4 and L_5 on the orbit of Jupiter. They are known as the Greek and Trojan asteroids (with names taken from Homeros). It has been observed that the Earth's Lagrange points are suitable places where an alien civilisation could place a satellite surveying the Earth; however, when the STEREO spacecrafts passed through (in 2009) they found nothing of the sort.

But what can we say about the three-body problem in general? To celebrate the sixtieth anniversary of King Oscar II of Sweden and Norway a large prize was offered for a solution to the following problem: "For a system of arbitrarily many mass points that attract each other according to Newton's laws, assuming that no two points ever collide, give the coordinates of the individual points for all time as the sum of a uniformly convergent series whose terms are made up of known functions." The prize was awarded to Henri Poincaré, who did not solve the problem as stated. Instead he laid the foundations of the modern theory of dynamical systems, including chaotic behaviour. For the three-body problem a solution was in fact found by Karl F. Sundman in 1912. He did express a generic solution as a uniformly convergent power series in $t^{1/3}$. The catch is that the series converges very slowly. It is estimated that, in order to get useful information, one would have to sum the first $10^{8000000}$ terms. Hence the interest in the exact general solution dwindled from that point in.

With the advent of the computer it has become possible to follow a large number of solutions to the three body problem on the screen, with no special effort. The zoo of solutions include ones where the third body escapes from the system, leaving the remaining pair more tightly bound than before.

◇ **Problem 4.1**   Solve Lorentz' equation for a charged particle moving in a constant magnetic field.

◇ **Problem 4.2**   The conservation of angular momentum is used in the gravitational two body problem to show that the trajectory is confined to a plane. Now consider an electrically charged particle moving in the electromagnetic field of a magnetic monopole, that is to say a hypothetical particle which, if placed at the origin, gives rise to the magnetic field

$$B_i = b\frac{x_i}{r^3} \ .$$

Here b is the magnetic charge of the monopole, and the electrically charged particle

obeys Lorentz' Law. Compute the time derivative of the particle's angular momentum. Find a conserved quantity which is a modified version of the angular momentum. Use the existence of this quantity, and the result of ex. 1, to qualitatively describe the trajectory of the particle.

◇   **Problem 4.3**   Consider the Earth-Moon system. Because of the tides some dissipation of energy takes place. How does this affect the distance of the moon from the earth?

◇   **Problem 4.4**   Consider the Yukawa potential

$$V(r) = -k\frac{e^{-\mu r}}{r} \ , \qquad k > 0 \ , \ \mu > 0 \ .$$

What can you say about the existence and stability of circular orbits?

◇   **Problem 4.5**   In the theory of black holes one encounters the following equation for particles orbiting the black hole,

$$\dot{r}^2 + \left(1 - \frac{2m}{r}\right)\left(\frac{L^2}{r^2} + 1\right) = E^2 \ ,$$

where $r(t)$ is related to the distance to the event horizon at $r = 2m$, t is related to time, m is the mass of the black hole, E is the energy of the particle, and L its angular momentum. The equations make sense only if $r > 2m$. You can choose L and E freely. Compute the smallest possible value of r for which a (marginally) stable circular orbit (with $r = $ constant) exists. A warning to the uninitiated is that relativistic units are used, so that the velocity of light and Newton's constant are set to 1.

◇   **Problem 4.6**   If the Sun is flattened at its poles it will have a quadrupole moment, and the potential is

$$V(r) = -\frac{k}{r} + \frac{q}{r^3} \ .$$

The orbits will no longer be closed. Compute the angle by which the perihelion moves during one revolution, to first order in q.

◇   **Problem 4.7**   For the Newtonian potential (4.28) the two body problem admits an additional conserved vector

$$M_i = \epsilon_{ijk}\dot{x}_j L_k - \frac{k}{r}x_i \ .$$

This is known as the Runge-Lenz vector. Check that it is indeed conserved. In what direction does it point?

◇   **Problem 4.8**   The mass of Jupiter is $2 \cdot 10^{27}$ kg and its distance from the Sun is $8 \cdot 10^{11}$ m. Suppose the Sun suddenly disappears. How long would it take for the Trojan asteroids to crash onto Jupiter? (Hint: Recall exercise 1.10.)

◇   **Problem 4.9**   Consider a spherical cloud of dust, that is to say the cloud is composed of particles interacting only through gravity and its initial density ρ is

independent of the angular coordinates. Assume that initially the cloud is at rest. Now let the dust cloud collapse due to gravity. Think of the cloud as composed of spherical shells, and show that the time it takes for a shell of radius r to reach the centre is independent of r.

$\diamond$ **Problem 4.10** Consider a family of highly eccentric Kepler ellipses $Z(t)$. One particle is on the ellipse and one at its focus. Take the limit corresponding to colliding particles. Exactly what happens at the collision when it is described by the Hooke ellipse $w(t) = \sqrt{Z(t)}$? Use this transformation to solve exercise 1.10 again.

$\diamond$ **Problem 4.11** Let

$$\ddot{w} = -w|w|^{a-1} , \qquad Z = w^\alpha .$$

Choose a time parameter $\tau = \tau(t)$ so that Kepler's Second Law holds for $Z(\tau)$, and prove that

$$\frac{d^2 Z}{d\tau^2} = -cZ|Z|^{A-1} ,$$

where c is a constant and

$$\alpha = \frac{a+3}{2} , \qquad (a+3)(A+3) = 4 .$$

5 Small oscillations

An important class of equations that we can actually solve are the linear ones. Their importance stems from the fact that departures from equilibrium are described by equations that are linear to first order—near a minimum, most smooth potentials look like a collection of harmonic oscillators. When left alone they are very simple, but once we couple them to external forces many delightful things happen. Or harmful things, depending on your point of view. Either way the subject is of interest to engineers and physicists alike.

It is perhaps worth remarking that once we allow the external forces to depend on time the phase space picture is affected in a significant way— the phase space trajectories are now allowed to cross themselves, because the external conditions may have changed by the time the trajectory returns to the initial point.

5.1 Forced oscillations

We begin with a single degree of freedom obeying the equation

$$m\ddot{x} + kx = 0 \quad \Leftrightarrow \quad \ddot{x} + \omega^2 x = 0 \ , \quad \omega = \sqrt{\frac{k}{m}} \ . \tag{5.1}$$

The assumption that the string constant k is positive was slipped in when we wrote $\omega^2 = k/m$. The general solution is

$$x(t) = a_1 \cos \omega t + a_2 \sin \omega t = A \cos (\omega t + \delta) = \mathrm{Re} \left[\alpha e^{i\omega t} \right] \ , \tag{5.2}$$

where

$$\alpha = A e^{i\delta}, \qquad A = \sqrt{a_1^2 + a_2^2} \ , \qquad \tan \delta = -\frac{a_2}{a_1} \ . \tag{5.3}$$

Observe that α is a complex constant carrying information about both the amplitude A and the phase δ. It is in fact highly convenient to write the solution as the real part of a complex solution. This is a trick that works because the equation is linear: two solutions can be added, and the result is

still a solution. It does not matter if we take the real part before or after the addition—but care must be exercised if we multiply two solutions together.

The total energy of the oscillator is

$$E = \frac{m}{2}\dot{x}^2 + \frac{k}{2}x^2 = \frac{m}{2}\omega^2 A^2 \ . \tag{5.4}$$

It is manifestly independent of time. And the subject is exhausted.

To make matters more interesting we introduce an external force, and consider the *forced oscillator*

$$m\ddot{x} + kx = F(t) \quad \Leftrightarrow \quad \ddot{x} + \omega^2 x = \frac{1}{m}F(t) \ . \tag{5.5}$$

To solve this equation, with some specified function $F(t)$, we appeal to the general theory of ordinary differential equations: it is enough to find one particular solution and then add the general solution of the homogeneous equation (the one with $F = 0$).

It goes without saying that some forces are more important than others. One example is the periodically varying force

$$F(t) = f \cos{(\Omega t + \phi)} \ , \tag{5.6}$$

which we will solve under the assumption that $\Omega \neq \omega$. For a particular solution we try the Ansatz

$$x_{\text{part}}(t) = B \cos{(\Omega t + \phi)} \ . \tag{5.7}$$

Plugging this into the equation will determine B. Adding the general homogeneous solution we find the noteworthy general solution

$$x(t) = A \cos{(\omega t + \delta)} + \frac{f}{m(\omega^2 - \Omega^2)} \cos{(\Omega t + \phi)} \ . \tag{5.8}$$

The remarkable thing is that the amplitude will get very large if the system is driven by a periodic force whose frequency is close to the natural frequency of the system. This phenomenon is called *resonance*. In fact, it may well be that resonance drives the system out of the regime in which the harmonic oscillator approximation is valid. (For the special case $\Omega = \omega$, do exercise 2.)

The solution is a superposition of two harmonics with different frequencies. Close to resonance the frequencies of the two harmonics almost coincide, and we will observe the phenomenon of *beats*. Recall that

$$\cos{\omega_1 t} + \cos{\omega_2 t} = 2 \cos{\frac{(\omega_1 - \omega_2)t}{2}} \cos{\frac{(\omega_1 + \omega_2)t}{2}} \ . \tag{5.9}$$

If $\omega_1 \approx \omega_2$ this can be regarded as a vibration with frequency $\omega \approx \omega_1 \approx \omega_2$, but with an amplitude modulated by a sine-wave of very low frequency. Our case is a bit more complicated because of the differing amplitudes and phases

of the harmonics that we superpose. Let us assume that $\Omega = \omega + \epsilon$, and rewrite the general solution as the real part of the complex amplitude

$$x(t) = \alpha e^{i\omega t} + \beta e^{i(\omega + \epsilon)t} = \left(\alpha + \beta e^{i\epsilon t}\right) e^{i\omega t} \ . \tag{5.10}$$

Provided ϵ is small compared to ω it makes sense to look at this as if the system were oscillating at frequency ω, but modulated by a slowly time varying amplitude whose square is given by

$$|\alpha + \beta e^{i\epsilon t}|^2 = |\alpha|^2 + |\beta|^2 + 2|\alpha||\beta| \cos\left(\epsilon t + \text{a phase}\right) \ . \tag{5.11}$$

Thus the amplitude varies between $|\beta| - |\alpha|$ and $|\beta| + |\alpha|$. If ϵ is small we see beats.

To deal with a force of a quite arbitrary form we call on the complex numbers to do a trick. Observe first that

$$\ddot{x} + \omega^2 x = \frac{d}{dt}(\dot{x} + i\omega x) - i\omega(\dot{x} + i\omega x) \ . \tag{5.12}$$

Hence we define the complex variable

$$z = \dot{x} + i\omega x \ , \tag{5.13}$$

which—if you recall your quantum mechanics—is the "annihilation operator" in slight classical disguise. The original equation becomes

$$\dot{z} - i\omega z = \frac{1}{m} F(t) \ . \tag{5.14}$$

Since the variable is now complex this is actually a pair of real first order equations, so we are not violating any rules concerning how to count the degrees of freedom. Once we have solved for $z(t)$ we recover $x(t)$ through

$$x(t) = \frac{1}{\omega} \text{Im}\left[z(t)\right] \ . \tag{5.15}$$

Solving the first order system is straightforward. A particular solution to eq. (5.14) can be written in integral form. Adding the general solution of the homogeneous equation we obtain

$$z(t) = e^{i\omega t} \left(z_0 + \frac{1}{m} \int_{-\infty}^{t} F(t') e^{-i\omega t'} dt'\right) \ . \tag{5.16}$$

The particular solution is known as Duhamel's integral.

Let us take a look at the energy budget. Assume that $z(-\infty) = 0$, so that the system starts out at rest. The energy we end up with is then

$$E(\infty) = \frac{m}{2} |z(\infty)|^2 = \frac{1}{2m} \left|\int_{-\infty}^{\infty} F(t) e^{-i\omega t} dt\right|^2 \ . \tag{5.17}$$

There is a net transfer of energy if the external force has a Fourier component corresponding to the intrinsic frequency of the oscillator. If the force acts for a short time—as compared with the natural time scale in the problem, which is set by ω—the exponential in the integrand can be ignored, and the energy imparted to the system is the kinetic energy associated to the momentum $\int F dt$.

5.2 Damped and forced oscillations

Let us consider a particle which is losing energy to its environment, and let us assume that there are very many degrees of freedom in the latter. It often happens that the frequencies associated with the environment are very much larger than the frequencies associated with the system of interest. In many such cases we can introduce friction as an effective force. The coupling is expected to grow with velocity, so we try the equation

$$m\ddot{x} + \gamma\dot{x} + kx = 0 \quad \Leftrightarrow \quad \ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = 0 \,, \quad \omega_0^2 = \frac{k}{m} \,, \quad \lambda = \frac{\gamma}{2m} \,. \quad (5.18)$$

This is the *damped oscillator*.

Depending on circumstances, other equations may be preferable. The motion of a small body falling rapidly in air (so that turbulence appears behind the body) is described quite accurately by

$$m\ddot{z} + c\dot{z}^2 - mg = 0 \,. \quad (5.19)$$

We leave this equation aside however, and concentrate on the damped oscillator. At least, it is easily realized in electrical circuit theory. A capacitor leads to a voltage difference $V_C = q/C$, where q is the charge and C the capacitance. Across an inductor there is a voltage difference $V_L = L\dot{I}$, where $I = \dot{q}$ is the current and L the inductance. Coupling the capacitor, the inductor, and a resistor with resistance R in a series we obtain the equation

$$L\ddot{q} + R\dot{q} + \frac{1}{C}q = 0 \,. \quad (5.20)$$

Thus $\omega_0 = 1/\sqrt{LC}$. There are numerous useful analogies between electrical, acoustical, and mechanical systems—but this is by the way.

To find the solution of the damped oscillator we try the Ansatz $x = e^{i\omega t}$. Inserting this, and cancelling the exponential, leads to the equation

$$-\omega^2 + 2i\lambda\omega + \omega_0^2 = 0 \quad \Leftrightarrow \quad \omega = i\lambda \pm \sqrt{\omega_0^2 - \lambda^2} \,. \quad (5.21)$$

Hence the general solution is

$$x(t) = a_1 e^{-\lambda t + i\sqrt{\omega_0^2 - \lambda^2}\,t} + a_2 e^{-\lambda t - i\sqrt{\omega_0^2 - \lambda^2}\,t} \,. \quad (5.22)$$

There are two qualitatively distinct cases to consider. If the damping is weak, that is if $\omega_0^2 > \lambda^2$, the general solution is a damped oscillation

$$x(t) = e^{-\lambda t} \left(a_1 e^{i\omega_{\mathrm{eff}} t} + a_2 e^{-i\omega_{\mathrm{eff}} t} \right) . \tag{5.23}$$

Note that the effective frequency ω_{eff} is always smaller than the "bare" frequency ω_0, as is reasonable. If the damping is strong enough so that $\omega_0^2 < \lambda^2$ the solution decays exponentially without oscillations.

It is instructive to consider the critical case $\omega_0^2 = \lambda^2$. In the limit we obtain the solution $ae^{-\lambda t}$, but this cannot be the general solution since it contains only one integration constant. In fact the general solution is

$$x(t) = (a + bt)e^{-\lambda t} . \tag{5.24}$$

This must be the general solution because it contains two integration constants.

Finally we come to the forced and damped oscillator

$$m\ddot{x} + \gamma \dot{x} + kx = F(t) = f \cos(\Omega t) , \tag{5.25}$$

where we restrict ourselves to the important case of a periodic external force. We rewrite this—in its complex formulation—as

$$\ddot{x} + 2\lambda \dot{x} + \omega_0^2 x = \frac{f}{m} e^{i\Omega t} , \tag{5.26}$$

make the Ansatz

$$x_{\mathrm{part}}(t) = B e^{i\Omega t} , \tag{5.27}$$

insert in the equation, cancel the exponential, and solve for B. The result is

$$(-\Omega^2 + 2i\lambda\Omega + \omega_0^2)B = \frac{f}{m} \quad \Leftrightarrow \quad B = \frac{1}{m} \frac{f}{\omega_0^2 - \Omega^2 + 2i\lambda\Omega} . \tag{5.28}$$

The amplitude is the absolute value of B. In fact

$$B = b e^{i\delta} , \quad b = \frac{1}{m} \frac{f}{\sqrt{(\omega_0^2 - \Omega^2)^2 + 4\lambda^2\Omega^2}} , \quad \tan\delta = \frac{2\lambda\Omega}{\Omega^2 - \omega_0^2} . \tag{5.29}$$

The real solution is then

$$x(t) = b \cos(\Omega t + \delta) . \tag{5.30}$$

Of course we could add the general homogeneous solution, but since this decays to zero we ignore it. When the force has acted for some time the transient part will be effectively zero, and a steady state described by the above solution sets in.
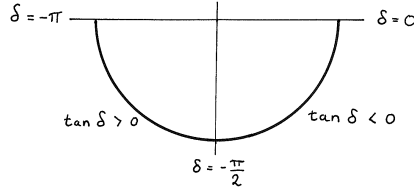
Figure 5.1. Phase shift of the damped and driven oscillator.

Because of the damping the amplitude no longer goes to infinity at resonance, but it still has a pronounced maximum. The phase δ shows that the system does not oscillate in phase with the external force. When Ω is very small so is δ, meaning that the oscillations follow the external force with no phase shift. The sign of δ is negative, but its tangent switches sign at $\Omega = \omega_0$. Indeed δ always lies between 0 and $-\pi$, meaning that the system always lags behind the force. For very large Ω the phase shift is close to $-\pi$, because the accelerations are large, and the acceleration of an oscillator is $180°$ out of phase with its displacement.

We should look at the energy budget of the damped oscillator. We begin with the damped but free oscillator ($F = 0$). Then the time derivative of the energy function must be negative. Indeed

$$\dot{E} = \frac{d}{dt}\left(\frac{m}{2}\dot{x}^2 + \frac{k}{2}x^2\right) = \dot{x}(m\ddot{x} + kx) = -\gamma\dot{x}^2 = -2m\lambda\dot{x}^2 \ . \qquad (5.31)$$

Now consider the forced and damped oscillator in steady state. It is still losing energy to the frictional forces at exactly this rate. This energy must then be supplied as work by the external force, which is something we may want to know about. Inserting the solution (5.30) we obtain the energy supplied by the external force per unit time as

$$|\dot{E}| = 2m\lambda b^2\Omega^2 \sin^2\left(\Omega t + \delta\right) \ . \qquad (5.32)$$

For most purposes it will be enough to know the time average of this quantity. This time average is a function of the frequency Ω; recalling from (5.29) how the amplitude b depends on Ω we obtain

$$I(\Omega) \equiv \langle|\dot{E}|\rangle = \frac{f^2\lambda}{m}\frac{\Omega^2}{(\omega_0^2 - \Omega^2)^2 + 4\lambda^2\Omega^2} \ . \qquad (5.33)$$

This has a maximum at $\Omega = \omega_0$. Let us suppose that $\Omega = \omega_0 + \epsilon$, and that both ϵ and the damping λ are small compared to ω_0. In this approximation we can replace $I(\Omega)$ with the function

$$I(\epsilon) = \frac{f^2\lambda}{4m}\frac{1}{\epsilon^2 + \lambda^2} \ . \qquad (5.34)$$

This is the famous Lorentzian line shape function, giving the sharp resonant response of a low-loss system. Among many other applications it explains the intrinsic width of spectral absorption (and emission) lines—although the dominating effect when astronomers observe these is the broadening due to the Doppler shift, since the atoms are in thermal motion.

We see that the width of the Lorentzian line shape function grows with the damping λ. On the other hand the total area under the curve is (almost) independent of λ, as follows from the calculation

$$\int_0^\infty I(\Omega)d\Omega = \int_{-\omega_0}^\infty I(\epsilon)d\epsilon \approx \int_{-\infty}^\infty I(\epsilon)d\epsilon = \frac{f}{4m}\int_{-\infty}^\infty \frac{1}{1+\frac{\epsilon^2}{\lambda^2}}\frac{d\epsilon}{\lambda} = \frac{f\pi}{4m} \ . \quad (5.35)$$

Which is mildly surprising.

We have only scratched the surface here. The feed-back of the environment on an oscillating system can happen in many other ways. Equations of interest include

$$\ddot{x} + \omega^2(t)x = 0 \qquad\qquad (5.36)$$

with some prescribed function $\omega^2(t)$. This leads to the theory of *parametric resonance*—a child setting a swing in motion belongs to this class. Another interesting class is given by

$$\ddot{x}(t) + \omega^2 x(t) + gx(t-c) = 0 \ . \qquad\qquad (5.37)$$

Provided that $0 < c\omega < \pi$ the force is in phase with the velocity, and any initial oscillation tends to grow. Energy is being absorbed from the environment. In the limit of small c this is a negatively damped oscillator. See exercise 6. The Tacoma Narrows bridge—a famous example in the theory of small, and not so small, oscillations—belongs here.[1]

5.3   Several degrees of freedom

Let us now consider N degrees of freedom. The variables are x_i, and the equations to be solved are

$$m_{ij}\ddot{x}_j + \gamma_{ij}\dot{x}_j + k_{ij}x_j = F_i(t) \ , \qquad 1 \le i,j \le N \ . \qquad (5.38)$$

When $\gamma_{ij} = 0$ they follow from the Lagrangian

$$L = \frac{1}{2}m_{ij}\dot{x}_i\dot{x}_j - \frac{1}{2}k_{ij}x_ix_j + x_iF_i \ , \qquad\qquad (5.39)$$

and it follows that m_{ij} and k_{ij} are symmetric when considered as matrices. It is less obvious that

[1] For a nice account, with many historical glimpses, see A. Jenkins, *Self-oscillation*, Phys. Rep. **525** (2013) 167.

$$\gamma_{ij} = \gamma_{ji} \ . \tag{5.40}$$

Nevertheless it is true, because of a subtle argument from statistical physics that we take on trust here.

To start with we now try to solve the equations for a set of harmonic oscillators that interact among themselves, but not with the outside world. That is

$$m_{ij}\ddot{x}_j + k_{ij}x_j = 0 \ . \tag{5.41}$$

We know that the general solution must contain $2N$ integration constants, or equivalently N complex integration constants. So we try the Ansatz

$$x_i = \alpha_i e^{i\omega t} \ . \tag{5.42}$$

Inserting this into the equation, and then cancelling the exponential, gives the matrix equation

$$(-\omega^2 m_{ij} + k_{ij})\alpha_j = 0 \ . \tag{5.43}$$

There is a non-zero solution for the vector α_i if and only if the determinant of the matrix vanishes, that is if and only if

$$\left|k_{ij} - \omega^2 m_{ij}\right| = 0 \ . \tag{5.44}$$

This is an eigenvalue equation for ω^2. It is also a polynomial equation of order N, which means that it will have N not necessarily distinct roots. On physical grounds all the roots must be real and positive. If not there would be exponentially growing or decaying solutions, and the system would be unstable. (Note the logic: We try to model an actually existing system. Therefore the matrices m_{ij} and k_{ij} must be special in various ways.)

The conclusion so far is that we have found N in general distinct frequencies, called *normal frequencies*. The next step is to actually solve eq. (5.43) for the eigenvector α_i. It will be determined only up to an overall complex constant. Doing so for each normal frequency we end up with N particular solutions, each coming with one undetermined complex number. Adding all these particular solutions together gives the general solution to our problem.

A simple example of a system admitting two normal modes is given by two identical pendula connected by an elastic spring. Assuming that the oscillations are small we use the Lagrangian

$$L = T - V = \frac{1}{2}\left(\dot{\theta}_1^2 + \dot{\theta}_2^2 - \theta_1^2 - \theta_2^2 - k(\theta_1 - \theta_2)^2\right) \ . \tag{5.45}$$

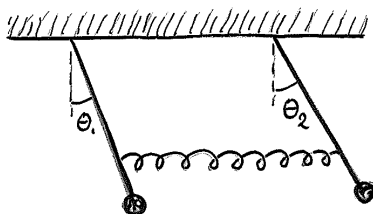To find the normal modes we must diagonalize T and V. This is achieved by

Figure 5.2. Two identical pendula connected by an elastic spring.

$$q_1 = \frac{1}{\sqrt{2}}(\theta_1 + \theta_2) , \qquad q_2 = \frac{1}{\sqrt{2}}(\theta_1 - \theta_2) . \qquad (5.46)$$

The Lagrangian becomes

$$L = \frac{1}{2}\left(\dot{q}_1^2 + \dot{q}_2^2 - q_1^2 - \omega^2 q_2^2\right) , \qquad \omega^2 = 1 + 2k . \qquad (5.47)$$

The normal modes have a simple interpretation. If $q_2 = 0$ the two pendula move in phase with the original frequency (equal to 1), if $q_1 = 0$ the pendula move in opposite phase with a frequency increased by the spring. As an afterthought we observe that, in this case, the two normal modes could have been identified from the outset using only symmetry considerations.

The general solution for θ_1, θ_2 is thus a sum of two harmonics—and if the difference in frequency between the two is small we should be able to observe the interesting phenomenon of beats, as in eq. (5.11). See exercise 7.

When we couple a set of harmonic oscillators to an external force it is the normal modes that count. As the simplest example, consider

$$\ddot{x}_1 + \omega_0^2 x_1 + k(x_1 - x_2) = f \cos \Omega t$$
$$\ddot{x}_2 + \omega_0^2 x_2 - k(x_1 - x_2) = 0 \qquad . \qquad (5.48)$$

In terms of the normal modes $q_1 = x_1 + x_2$, $q_2 = x_1 - x_2$ this becomes

$$\ddot{q}_1 + \omega_0^2 q_1 = f \cos \Omega t$$
$$\ddot{q}_2 + (\omega_0^2 + 2k)q_2 = f \cos \Omega t \qquad . \qquad (5.49)$$

Each normal mode gives rise to its own resonance frequency. The mathematical problem has been reduced to that of studying two—or in general N—independent forced oscillators.

Finally, let us consider eq. (5.41) for N degrees of freedom. We can decouple the degrees of freedom in three steps: First we diagonalise the symmetric matrix m_{ij}, using rotations. Its eigenvalues must be positive (to prevent exercise 1.2 from becoming relevant). In the next step we rescale its eigenvectors so that the diagonalized matrix becomes the identity matrix. These coordinate changes are conveniently described on the Lagrangian level,

$$L = \frac{1}{2}\sum_{i,j}(\dot{x}_i m_{ij}\dot{x}_j - x_i k_{ij}y_j) = \frac{1}{2}\sum_{i,j}(\dot{y}_i\lambda_i\delta_{ij}\dot{y}_j - y_i k'_{ij}y_j) =$$

$$\tag{5.50}$$

$$= \frac{1}{2}\sum_{i,j}\left(\dot{z}_i\delta_{ij}\dot{z}_j - z_i\frac{1}{\sqrt{\lambda_i}}k'_{ij}\frac{1}{\sqrt{\lambda_j}}z_j\right) \ .$$

For once we did not use the Einstein summation convention, since it would not work here. Note that k'_{ij} is not the same matrix as k_{ij}, since the coordinate system was rotated. In the final step we define the symmetric matrix

$$k''_{ij} = \frac{1}{\sqrt{\lambda_i}}k'_{ij}\frac{1}{\sqrt{\lambda_j}} \ , \tag{5.51}$$

we diagonalize it—by means of a rotation that leaves δ_{ij} invariant—and arrive at the Lagrangian

$$L = \frac{1}{2}\sum_i(\dot{u}_i\dot{u}_i - \omega_i^2 u_i u_i) \ . \tag{5.52}$$

The normal modes u_i are now decoupled, and the equations of motion trivial to solve.

   If you remember your linear algebra, you may be a little surprised by our success. In general the matrices m_{ij} and k_{ij} do not commute, so how could we bring both of them to diagonal form? The answer of course lies in the rescaling of the eigenvectors. We did bring both matrices to diagonal form, but not by means of rotations only.

◇   **Problem 5.1**   Verify eqs. (5.3). By the way you are supposed to verify all equations as simple to derive as those two.

◇   **Problem 5.2**   Solve eq. (5.5) when $F(t) = f\cos(\omega t + \phi)$.

◇   **Problem 5.3**   Derive eq. (5.8) by inserting the relevant force into Duhamel's integral.

◇   **Problem 5.4**   To understand Fig. 5.1 better, make a Mathematica plot of how the complex number B, in eq. (5.29), changes as you vary Ω.

◇   **Problem 5.5**   Consider a bouncing ball, obeying $\ddot{z} = -g$ but with a floor at $z = 0$. At each bounce the absolute value of its velocity decreases by a factor ϵ, $0 < \epsilon < 1$. Assume it left the floor with velocity v_0 at time $t = 0$ and that bounces are instantaneous. At what time does the motion stop? Sketch the motion in the z–t plane.

◇   **Problem 5.6**   Consider the self-oscillator described by eq. (5.37). Assume the time lag c is small. Expand the last term to first order in c and check that you get a negatively damped oscillator.

◇ **Problem 5.7** Consider the two pendula connected by the spring, and let the spring be very weak ($k << 1$). Choose initial data $\theta_1 = \theta_2 = 0$, $\dot{\theta}_1 = v$, $\dot{\theta}_2 = 0$. Show that the amplitudes of the two pendula are modulated in such a way that after some time T the first pendulum is stationary and all the energy has gone to the second.

◇ **Problem 5.8** Consider a particle in a plane, fastened to the corners of an equilateral triangle by means of identical springs (using Hooke's law). How will it move?

◇ **Problem 5.9** You can think of a metal as N independent harmonic oscillators (Einstein), or as a gas of sound waves with a certain number of allowed frequencies (Debye). Show that the number of allowed frequencies equals the number of degrees of freedom in the oscillators.

6 Rotation and rigid bodies

We have all played with spinning tops, and know that their dynamics is very rich. The subject is of considerable practical importance, say to spacecraft designers. To understand it we must understand the Lie group of rotations in three dimensional space, and this is where we begin the story.

6.1 Rotations

In the plane, the mathematics of rotations is fairly trivial. Any rotation takes place around a fixed point, and is uniquely characterized by its angle of rotation ϕ. Its action on the coordinates is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} . \tag{6.1}$$

What does this mean? Actually it can mean two quite different things: A *passive* coordinate transformation changing the coordinates used to describe a given point, or an *active* rotation moving a given point to another point described by different values of the coordinates. These are two very different operations, although the formulæ describing them are mostly identical. One must stay awake during calculations.

This attended to, we observe that rotations form a *group*. A group is a set of objects g_1, g_2, etc, together with a rule for combining them, so that $g_1 \circ g_2 = g_3$ is again a member of the set. This rule must be associative. One of the elements is the identity element e, and has the property that $e \circ g = g \circ e = g$ for any other element g, and finally every element g possesses an inverse g^{-1} such that $g \circ g^{-1} = g^{-1} \circ g = e$. In general it may or may not be true that $g_1 \circ g_2 = g_2 \circ g_1$.

For the two dimensional rotation group the elements can be written $g(\phi)$, so the number of elements are continuously infinite. The set of all group elements thus forms an abstract space, called the *group manifold*. In this case the group manifold is a circle, coordinatized by ϕ. It is also a *Lie group*. This means that the continuous parameters of a product are given by analytic functions of the parameters of its factors. Thus

$$g(\phi_1) \circ g(\phi_2) = g(\phi_3(\phi_1, \phi_2)) . \tag{6.2}$$

where ϕ_3 is an analytic function of its arguments. In our example

$$\phi_3 = \phi_1 + \phi_2 \ , \tag{6.3}$$

which is certainly analytic. "Analytic" here means that the function is fully determined by its Taylor series. The surprising thing about Lie groups is that they can be almost completely understood through a careful study of their properties in a small neighbourhood of the identity element. All rotation groups are Lie groups.

Rotations in three dimensional space are hard to understand, in the first place because rotations do not commute in general. Rotations can be represented by matrices acting on vectors,

$$x'_i = R_{ij} x_j \ . \tag{6.4}$$

By definition this is a rotation if the lengths of all vectors are preserved,

$$x'_i x'_i = x_i x_i \ . \tag{6.5}$$

Hence the matrix R_{ij} must be subject to some restrictions. Still it is immediate from the definition that rotations form a group. (Why?) To see how the matrix is restricted, we observe that scalar products must be preserved too, and then we perform a little calculation:

$$x'_i y'_i = R_{ki} x_i R_{kj} y_j = x_i R_{ki} R_{kj} y_j = x_i y_i \quad \Rightarrow \quad x_i (R_{ki} R_{kj} - \delta_{ij}) y_j = 0 \ . \tag{6.6}$$

Here δ_{ij} is the Kronecker delta. Since the vectors are arbitrary this is equivalent to

$$R_{ki} R_{kj} = \delta_{ij} \ . \tag{6.7}$$

In matrix notation this is

$$R^{\mathrm{T}} R = R R^{\mathrm{T}} = \mathbf{1} \ . \tag{6.8}$$

The group properties can be now be checked on the level of matrices. The columns, and the rows, of the matrix R must form an orthonormal triplet of vectors. Such matrices are called *orthogonal*. Matrix multiplication is associative, and moreover

$$R_1 R_2 = R_3 \quad \Rightarrow \quad R_3^{\mathrm{T}} R_3 = (R_1 R_2)^{\mathrm{T}} R_1 R_2 = R_2^{\mathrm{T}} R_1^{\mathrm{T}} R_1 R_2 = 1 \ . \tag{6.9}$$

This confirms that R_3 is a group element, as it had to be. We observe that

$$1 = \det R R^{\mathrm{T}} = \det R \det R^{\mathrm{T}} = (\det R)^2 \quad \Rightarrow \quad \det R = \pm 1 \ . \tag{6.10}$$
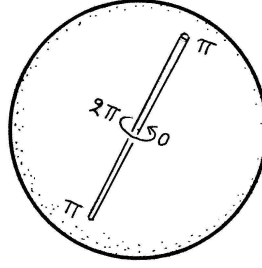
Figure 6.1. To coordinatize the set of all rotations (the group manifold of the rotation group), choose a rotation axis and a number along it. Note that the two endpoints of the axis represent the same rotation.

The group can be restricted by insisting that $\det R = 1$. This restriction can be formulated in terms of the epsilon tensor:

$$R_{im}R_{jn}R_{kp}\epsilon_{mnp} = \epsilon_{ijk} \det R = \epsilon_{ijk} \;, \tag{6.11}$$

where the definition (sic!) of the determinant was used. In N dimensions the restricted rotation group is called the *special orthogonal group*, denoted $SO(N)$. "Special" refers to the restriction that the determinant equals one. If matrices with determinant -1 are admitted the group is called $O(N)$, and includes reflections.

Euler's theorem states that any rotation in 3-space is a rotation around a fixed axis. To prove it, note that an orthogonal matrix is unitary, and the eigenvalues of a unitary matrix always take the form

$$\lambda = e^{i\phi} \tag{6.12}$$

for some angle ϕ. (Proof: Check the conditions that make a diagonal matrix unitary.) But an orthogonal matrix is a real matrix too, which means that

$$\det(R - \lambda \mathbf{1}) = 0 \quad \Rightarrow \quad (\det(R - \lambda \mathbf{1}))^* = \det(R - \lambda^* \mathbf{1}) = 0 \;. \tag{6.13}$$

Hence complex eigenvalues, if they occur, must occur in pairs because their complex conjugates are eigenvalues too. Since the number of eigenvalues of an $SO(3)$ matrix is odd, one of them must be real, and in fact it must equal one because the determinant does. The corresponding eigenvector is the fixed axis of rotation. Note that rotations in even dimensions (like two dimensions) work differently. Note also that rotations can be hard to grasp: if the rotation axes of R_1 and R_2 are known, what is the rotation axis of $R_1 R_2$?

The group manifold of $SO(3)$, that is to say the set of all rotations, is easy to visualize. The set of all rotation axes can be identified with the surface of a sphere, or more precisely with the pairs of antipodal points where the axis cuts the sphere. An arbitrary rotation is determined by its axis and an angle ϕ, hence we can think of the set of all rotations as a solid ball with the identity

matrix at its center, with ϕ as a radial coordinate, and with antipodal points on its surface identified because of the periodicity of ϕ. It sounds simple, but there are some subtleties. The topology of this group manifold is such that there are closed curves, starting and ending at the identity element, that cannot be shrunk to zero in a continuous way. There is a famous trick one can play with a pair of scissors sliding along a belt, to verify that this property has tangible consequences.

We will need a coordinate system on the group manifold, and we choose the *Euler angles* for this purpose. From our present perspective it is a little difficult to make them appear natural. We introduce them by brute force, as follows: Define

$$R_\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad R_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

$$R_\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.14}$$

Compute

$$R(\psi, \theta, \phi) = R_\psi R_\theta R_\phi = \tag{6.15}$$

$$= \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix} .$$

In the absence of the argument that makes this construction appear natural, how do we know that this is correct in the sense that an arbitrary rotation matrix can be expressed in terms of the Euler angles? Recall that every orthogonal matrix can be thought of as three orthonormal column vectors. By inspection we see that ψ and θ can be chosen so that the third column agrees with any unit vector, and further inspection of the remaining columns shows that they are restricted only to the extent needed for them to fill out a right handed orthonormal triad. All this is true provided that

$$0 \leq \phi < 2\pi , \qquad 0 \leq \psi < 2\pi , \qquad 0 \leq \theta \leq \pi . \tag{6.16}$$

We accept this, and now we have a coordinate system on the group manifold $SO(3)$ available whenever we need one.

In writing eq. (6.15) we proved that an arbitrary rotation can be effected by first rotating around the z-axis, then around the x-axis, and finally around the z-axis again. There are $3 \times 2 \times 2 = 12$ different ways of choosing the axes here, leading to 12 different ways of parametrizing an arbitrary rotation matrix.

This has the consequence that whenever Euler angles are encountered in the literature, one must check which of the 12 possible definitions that was used.

A final point: I advertized that Lie groups can be understood through a study of what they look like close to the origin. To see how for the orthogonal groups, note that an orthogonal matrix can be written in the form

$$R = e^A \equiv \mathbf{1} + A + \frac{1}{2!}A^2 + \dots \ . \tag{6.17}$$

This is an orthogonal matrix if and only if the matrix A is anti-symmetric,

$$A^{\mathrm{T}} = -A \ . \tag{6.18}$$

We can define a special class of curves in the group by

$$R(t) = e^{tA} \quad \Rightarrow \quad R(0) = \mathbf{1} \ . \tag{6.19}$$

We would like to claim that every rotation group element can be written as e^{tA}, for some choice of anti-symmetric matrix A. This is actually so for any $SO(3)$ rotation, but not for the additional reflections present in $O(3)$. To see how it works, we begin by using the definition (6.17) to conclude that

$$\exp\left[t\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right] = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \ . \tag{6.20}$$

So the statement is true for any two dimensional rotation matrix with unit determinant. But we can use Euler's theorem to reduce the three dimensional case to the two dimensional one—we simply adapt our coordinates so that one of the axes points along the eigenvector of the given but otherwise arbitrary rotation matrix.

We can now see by means of a second order Taylor expansion what the non-commutativity means in terms of what goes on close to the identity element:

$$R_1(t_1)R_2(t_2)R_1^{-1}(t_1)R_2^{-1}(t_2) \approx$$

$$(\mathbf{1} + t_1 A_1 + \frac{t_1^2}{2}A_1^2)(\mathbf{1} + t_2 A_2 + \frac{t_2^2}{2}A_2^2)(\mathbf{1} - t_1 A_1 + \frac{t_1^2}{2}A_1^2)(\mathbf{1} - t_2 A_2 + \frac{t_2^2}{2}A_2^2)$$
$$\tag{6.21}$$
$$\approx \mathbf{1} + t_1 t_2 (A_1 A_2 - A_2 A_1) \ .$$

The group elements commute only if the *commutator* $[A_1, A_2]$ vanishes. Using something known as the Baker-Campbell-Hausdorff formula one can in fact prove that e^{iA_1} and e^{iA_2} commute if and only if $[A_1, A_2] = 0$, but doing so would take us a little too far afield.

Anti-symmetric matrices can be added and multiplied with real numbers, while staying anti-symmetric. Therefore they form a vector space, known as the *tangent space* of the group manifold at the identity element. This vector

space is also known as the *Lie algebra* of the group, because the commutator $[A_1, A_2]$ provides us with a definite way to obtain a third "vector" from any given pair of "vectors". For $SO(3)$ any anti-symmetric matrix can be expressed as a linear combination of the three matrices

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \quad (6.22)$$

They obey

$$[A_1, A_2] = A_3 \qquad [A_2, A_3] = A_1 \qquad [A_3, A_1] = A_2 . \qquad (6.23)$$

They form a basis for the Lie algebra—and would perhaps look more familiar had we multiplied the Lie algebra elements A_i with an imaginary factor of i, to make them Hermitian matrices. The idea here (a grand one!) is that it works backwards too—from a Lie algebra of commutators one can reconstruct a Lie group.

In index notation an arbitrary element of the $SO(3)$ Lie algebra can be written as

$$A_{ij} = \epsilon_{ikj}\omega_k , \qquad (6.24)$$

where ω_i is a vector. If a physical system is subject to a time dependent rotation $R(t)$ such that $R(0)$ is the identity matrix and

$$\dot{R}_{ij}(0) = \epsilon_{ikj}\omega_k , \qquad (6.25)$$

then ω_i is referred to as the *angular velocity vector* at time $t = 0$. In space it is pointing along Euler's fixed axis of rotation. If we return to the picture of the group manifold as a solid ball with antipodal points on its surface identified, we recognize the angular velocity vectors as a set of tangent vectors sitting at the identity element, that is at the centre of the ball.

6.2   Rotating coordinate systems

Newton's Laws take a simple form only in *inertial* coordinate systems, that is in coordinate systems adapted to the straight lines referred to in his First Law (or to Absolute Space, as Newton himself expressed things). If the coordinate system is not inertial more work is needed—it would not have been easy to figure out the laws if all experiments had been carried out on roundabouts. Let $X_i(t)$ denote the trajectory of a particle relative to an inertial system, and let $x_i(t)$ denote the same trajectory described relative to a rotating coordinate system. Assume that the coordinates agree at $t = 0$. Then there exists a t-dependent rotation matrix R such that

$$X_i(t) = R_{ij}(t)x_j(t) = \left(\delta_{ij} + \epsilon_{ikj}\omega_k t + o(t^2)\right) x_j(t) \ . \qquad (6.26)$$

At time $t = 0$ the inertial and rotating coordinate systems agree. (There will always be such an inertial coordinate system, so this is no restriction.) We assume that the angular velocity vector ω_i is constant. The time derivatives are related by

$$\dot{X}_i = R_{ij}\dot{x}_j + \dot{R}_{ij}x_j \ . \qquad (6.27)$$

At $t = 0$ (a point in time picked by convention) our assumption gives

$$\dot{X}_i = \dot{x}_i + \epsilon_{ikj}\omega_k x_j \ . \qquad (6.28)$$

The dot-notation can get confusing at this point, so it may be helpful to write this as an operator relation

$$\dot{X}_i = D_{tij}x_j = \left(\delta_{ij}\frac{d}{dt} + \epsilon_{ikj}\omega_k\right) x_j \ , \qquad (6.29)$$

If we differentiate twice we obtain Newton's second law in the form

$$F_i = m\ddot{X}_i = mD_{tij}D_{tjk}x_k = m\ddot{x}_i + 2m\epsilon_{ijk}\omega_j\dot{x}_k + m\epsilon_{ijk}\epsilon_{kmn}\omega_j\omega_m x_n \ . \quad (6.30)$$

By the ϵ-δ-identity, or equivalently by means of the formula for repeated cross products, this is

$$m\ddot{x}_i = F_i - 2m\epsilon_{ijk}\omega_j\dot{x}_k + m\left(\omega^2\delta_{ij} - \omega_i\omega_j\right) x_j \ . \qquad (6.31)$$

This is what Newton's second law looks like on the roundabout. The "extra" terms on the right hand side are known as *inertial* or *fictitious* forces—but they feel real enough.

The third term on the right hand side of eq. (6.31) is known as the *centrifugal force*. It is responsible for the repulsive part of the effective potential (4.9) in the two body problem. The second term is the *Coriolis force*. It is perpendicular both to the velocity $\dot{x}_i$ and to the angular velocity ω_i. To see that such a term must be there, consider a free particle moving out from the centre on a frictionless rotating disk, and anchor the coordinate system to the disk. Alternatively, consider the *Foucault pendulum* somewhere on earth. Choose a coordinate system with a vertical z-axis, and let the pendulum perform small oscillations in the x-y-plane. We ignore all terms of second order in the angular velocity Ω_i of the earth. In particular we ignore the centrifugal force, but the angular velocity Ω_i enters the equations through the Coriolis term:

$$\ddot{x} = -k^2 x + 2\Omega_z\dot{y} \qquad \ddot{y} = -k^2 y - 2\Omega_z\dot{x} \ . \qquad (6.32)$$

Here $\Omega_z = |\Omega|\sin\theta$, and θ is the latitude of the pendulum (59 degrees 21
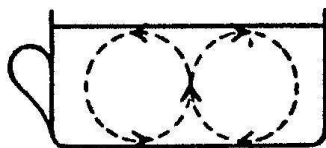
Figure 6.2. This picture illustrates Einstein's article on Baer's Law, so presumably this is what the great man's tea cup looked like.

minutes, if it is in Stockholm). To solve the equations, introduce the complex variable $w = x + iy$ and make the Ansatz $w = e^{\lambda t}$:

$$\ddot{w} + k^2 w + 2i\Omega_z \dot{w} = 0 \quad \Rightarrow \quad \lambda^2 + k^2 + 2i\Omega_z\lambda = 0 \ . \qquad (6.33)$$

The solutions are

$$\lambda = -i\Omega_z \pm i\sqrt{k^2 + \Omega_z^2} \approx -i\Omega_z \pm i\left(k + \frac{\Omega_z^2}{2k}\right) \approx -i\Omega_z \pm ik \ . \qquad (6.34)$$

Therefore

$$w = e^{-i\Omega_z t}\left(c_1 e^{ikt} + c_2 e^{-ikt}\right) \ . \qquad (6.35)$$

At the equator, where $\Omega_z = 0$, we can arrange the initial conditions so that the pendulum swings in a plane with $y = 0$ (say), but at non-zero latitudes the plane will turn at a rate given by $\sin\theta$ times the rate of rotation of the earth. At the Poles it will make one revolution in 24 hours.

It is interesting to see how the fictitious forces shape the face of the earth. They explain the erosion of river beds, as shown by Albert Einstein. He began by looking at a simpler problem: Take a cup of tea, with some tea leaves in it, and set the tea into rapid rotation by means of a spoon. You will see the tea leaves gather into a little clump at the bottom, just at the center of the cup. This happens because the flow velocity of the tea is zero at the edges of the cup, and rises gradually from zero as one moves away from the edges. Therefore the angular velocity will be greater at the tea surface, and lower close to the bottom. Therefore the outwards directed centrifugal force will be greater at the surface, and as a result a vertical velocity profile will be set up in the cup. Therefore the tea leaves—staying close to the bottom due to their specific weight—will be carried to the centre of the cup, as was to be explained.

A very similar story can be told about rivers, if there is a river bend. The vertical velocity profile will carry fast moving surface water to the outer river bank, and will cause erosion there. The latter will therefore suffer from more severe erosion than the inner bank, and the net result is that the whole river

will start to move outwards—it is meandering. The Coriolis force is also of importance to rivers, provided the rivers are large and do not bend too much. In fact it causes Baer's Law, an observation due to geographers according to which large rivers on the Northern hemisphere undermine their right banks, and large rivers on the Southern hemisphere undermine their left banks. Baer's law does not apply to small and bending rivers, where the centrifugal force dominates.

Even more dramatically, fix a coordinate system in the earth so that one of its axes point at the sun. It will make one revolution per year, and create a centrifugal force just balancing the gravitational attraction felt by the sun. This enables the sun to stand still. Now let the coordinate system follow the earth in its daily rotation. As a result a huge centrifugal force will act on the sun, completely overwhelming the gravitational force. In this coordinate system the sun appears to be in rapid motion ($\dot{x}_i \neq 0$), and the Coriolis force steps in to the rescue and prevents the sun from disappearing into outer space.[1]

6.3   The inertia tensor

A rigid body is defined as a number of particles at fixed distances from each other, or equivalently as a continuous mass distribution whose shape is fixed. In the latter case the mass points m are replaced by mass elements ρdV. The centre of mass of the body can move, and its overall orientation can change by means of a rotation. Any rotation. Hence the configuration space of a rigid body is $\mathbf{R}^3 \times SO(3)$. This is not only 6 dimensional, it has a non-trivial topology since the topology of the group manifold $SO(3)$ is non-trivial.

We introduce an internal coordinate system whose origin is fixed at some point O within the body. These coordinates are called x_i. We also introduce an inertial coordinate system so that O has coordinates X_i. The velocity of any given point x_i within the body with respect to the inertial coordinate system is denoted v_i, and the velocity of the point O is denoted V_i. The body rotates around O with angular velocity Ω_i, but the internal coordinate system (with origin O) is not rotating. Make sure to note that the velocity v_i is *not* the time derivative of x_i, and that we will *not* be bothered by fictitious forces. The various quantities we introduced are related by

$$v_i = V_i + \dot{x}_i = V_i + \epsilon_{ikj}\Omega_k x_j \ , \tag{6.36}$$

To what extent does it matter where the origin O is placed? If we translate it so that

$$x_i \rightarrow x'_i = x_i + a_i \ , \tag{6.37}$$

we see that the instantaneous velocities transform according to

---

[1] Thoughtful discussion of inertial forces can be found in two popular books by D. W. Sciama, *The Unity of the Universe*, Doubleday, New York 1959, and chapter 1 of *The Physical Foundations of General Relativity*, Doubleday, New York 1969.

$$v_i = V_i + \epsilon_{ikj}\Omega_k x_j = V_i - \epsilon_{ikj}\Omega_k a_j + \epsilon_{ikj}\Omega_k x'_j = V'_i + \epsilon_{ikj}\Omega_k x'_j \ . \qquad (6.38)$$

We observe that the angular velocity is unchanged, so we can talk of the angular velocity *of the body* without caring about the point relative to which we compute it.

The kinetic energy of a rigid body is

$$T = \sum \frac{m}{2}(V_i + \epsilon_{ikj}\Omega_k x_j)^2 =$$

$$= \sum \left(\frac{m}{2}V^2 + m\epsilon_{ikj}V_i\Omega_k x_j + \frac{m}{2}(\epsilon_{ikj}\Omega_k x_j)^2 \right) \ . \qquad (6.39)$$

(The sum runs over all the particles in the body, even though no explicit indices have been placed on them.) This can be rewritten in terms of the total mass M as

$$T = \frac{M}{2}V^2 + \frac{1}{2}\sum m(x^2\delta_{ij} - x_i x_j)\Omega_i\Omega_j + \epsilon_{ikj}V_i\Omega_k \sum m x_j \ . \qquad (6.40)$$

If O, the origin within the body, is placed at the centre of mass the last term vanishes. It also vanishes if the origin coincides with a point that is fixed in space (when $V_i = 0$). Now define the *inertia tensor* with respect to O,

$$I_{ij} = \sum m(x^2\delta_{ij} - x_i x_j) \ . \qquad (6.41)$$

Again the sum runs over all the particles in the body. We observe that the angular momentum with respect to O is

$$L_i = \sum \epsilon_{ijk}x_j m\dot{x}_k = \sum \epsilon_{ijk}\epsilon_{kmn}m x_j \Omega_m x_n = I_{ij}\Omega_j \ . \qquad (6.42)$$

Unlike the angular velocity, both the angular momentum and the inertia tensor change if we shift the position of O. Let us assume that O is placed at the centre of mass. Then

$$T = \frac{1}{2}MV^2 + \frac{1}{2}I_{ij}\Omega_i\Omega_j = \frac{1}{2}M^{-1}P_iP_i + \frac{1}{2}I_{ij}^{-1}L_iL_j \ . \qquad (6.43)$$

The inertia tensor describes the resistance of the body to changes of its rotation, just as the mass describes its resistance to changes of its translational state. But the former is a tensor, not a scalar, and hence much harder to grasp.

Written out, the inertia tensor is

$$I = \begin{pmatrix} \sum m(x_2^2 + x_3^2) & -\sum mx_1x_2 & -\sum mx_1x_3 \\ -\sum mx_2x_1 & \sum m(x_1^2 + x_3^2) & -\sum mx_2x_3 \\ -\sum mx_3x_1 & -\sum mx_3x_2 & \sum m(x_1^2 + x_2^2) \end{pmatrix} . \qquad (6.44)$$

This is a symmetric matrix, hence it can be diagonalized by rotations. This means that we can rotate the internal coordinate system so that

$$I = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} . \qquad (6.45)$$

The eigenvalues are known as *moments of inertia*. The eigenvectors are known as *principal axes* of the body, and are orthogonal to each other.

The moments of inertia are all positive since, for an arbitrary vector n_i,

$$n_i I_{ij} n_j = \sum m \left(n^2 x^2 - (n_i x_i)^2 \right) \geq 0 . \qquad (6.46)$$

Indeed the individual terms are all positive. A vanishing moment of inertia can occur only if all the particles lie on a line (because it would be necessary that $(n \cdot x)^2 = n^2 x^2$ for each individual particle—recall that the notation suppresses the sum over all particles in the body, but in this case all the individual terms are positive or zero.) Once we have adapted the coordinates so that the inertia tensor is diagonal we see immediately that

$$I_1 + I_2 = \sum m(x_1^2 + x_2^2 + 2x_3^2) \geq \sum m(x_1^2 + x_2^2) = I_3 . \qquad (6.47)$$

Hence no moment of inertia can exceed the sum of the other two. Equality happens if and only if $x_3 = 0$ for all the particles, that is for a planar body.

Some general facts about the inertia tensor can be stated:

- If the body is symmetric under reflection in a plane, two of the principal axes lie in it.

This must be so because the reflection must preserve the principal axes. The only way to arrange this is to let two of them lie in the plane. The third axis is then orthogonal to the plane and is taken into itself by the reflection. (Note that this theorem is enough to identify the principal axes of an ellipsoid.)

The body can also be symmetric under rotations around some axis through an angle which is some fixed fraction of 2π:

- If the body has a symmetry axis of order higher than two this axis must be a principal axis, and the other two moments of inertia are equal (because the corresponding eigenvectors cannot be unique).

A body with two equal moments of inertia is called a *symmetrical top*. If there are at least three higher order symmetry axes—this is true for a cube, say—it follows that all the eigenvalues are equal, so in fact every axis is a principal axis.

The shape of the body is reflected in the inertia tensor. In the gravitational N-body problem Newton proved that if all bodies are spherical, they can be regarded as mass points. Now we see that if a body is rigid it can be regarded as a homogeneous ellipsoid, since every tensor of intertia can be thus reproduced. Nothing else matters as far as the response to arbitrary forces is concerned.

The inertia tensor depends on the point with respect to which it is computed:

Steiner's theorem: For a body of total mass M, let I_{ij} be its inertia tensor relative to the centre of mass, and I'_{ij} its inertia tensor relative to a point translated from the centre of mass by the vector a_i. Then

$$I'_{ij} = I_{ij} + M(a^2\delta_{ij} - a_i a_j) . \tag{6.48}$$

Roughly speaking the theorem says that it is easiest to rotate the body around its centre of mass. The proof is easy, once we recall that

$$\sum m a_i x_i = a_i \sum m x_i = 0 , \tag{6.49}$$

where x_i is position relative to the centre of mass.

It is instructive to prove that any inertia tensor can be reproduced by placing four particles of equal mass at appropriate distances from each other. To do so, let the coordinates of the particles be

$$(x_i, y_i, z_i) , \qquad 1 \le i \le 4 . \tag{6.50}$$

This permits us to define three vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbf{R}^4$. We define a fourth auxiliary vector whose components are equal,

$$\mathbf{v}^{\mathrm{T}} = (1, 1, 1, 1) . \tag{6.51}$$

That is to say that we are looking at the matrix

$$[\mathbf{x}\ \mathbf{y}\ \mathbf{z}\ \mathbf{v}] = \begin{bmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{bmatrix} , \tag{6.52}$$

where the row index labels the four particles. Now we insist that the inertia tensor be diagonal and that the centre of mass is at the origin. This translates into the six conditions

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{z} = \mathbf{y} \cdot \mathbf{z} = \mathbf{v} \cdot \mathbf{x} = \mathbf{v} \cdot \mathbf{y} = \mathbf{v} \cdot \mathbf{z} = 0 . \tag{6.53}$$
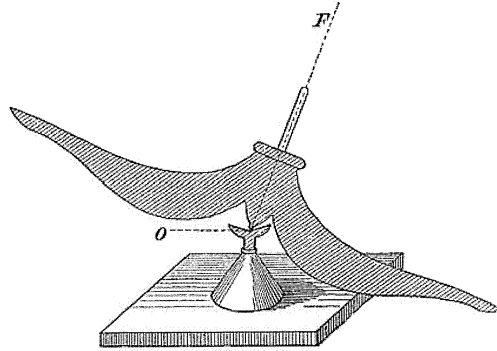
Fig. 1.

Figure 6.3. A top carefully constructed so that, with the addition of small weights, it can be given the equilibrium position shown. The illustration is from F. Klein and A. Sommerfeld: *Theorie des Kreisels I*, recently reissued by Birkhäuser in a good English translation.

In other words the four vectors must be mutually orthogonal, which is easily arranged in $\mathbf{R}^4$. It remains to check that we can reproduce the most general diagonal inertia tensor; the solution is given in exercise 4.

6.4 Euler's equations

The following picture emerges: Associated with any rigid body there are three kinds of axes in space, none of which coincide in general. Its principal axes are associated with its shape and—with a little practice—they can be literally seen. Then its angular velocity vector, which can be made visible. See exercise 7. But the axis that matters most to the motion of the body is its invisible angular momentum vector, because this is the axis that is conserved in force free motion. The visible axes will rotate around it as if by magic.

When its centre of mass is fixed, the equation of motion for a rigid body subject to a *torque* τ_i is

$$\dot{L}_i = \sum m\epsilon_{ijk}x_j F_k = \tau_i \ , \tag{6.54}$$

where both the angular momentum and the torque are computed with respect to the centre of mass. The equation is complicated by the fact that the inertia tensor is time dependent,

$$\dot{L}_i = I_{ij}\dot{\Omega}_j + \dot{I}_{ij}\Omega_j \ . \tag{6.55}$$

This is awkward, whether there is a torque or no.

For the rest of this section we confine ourselves to torque free motion, with the centre of mass kept fixed. (Even if we are in the gravitational field of the earth, it is possible to manufacture tops so that they are suspended at their

centres of mass, and realize torque free motion. See Fig. 6.3.) Theoretically, the guiding idea will be to formulate the equations of motion with respect to a coordinate system fixed within the body, because then the inertia tensor is time independent. The direction of the angular momentum vector will then change with time because the coordinate system rotates. We denote this vector by J_i. We proceed in analogy with eq. (6.28), that is to say we assume that the inertial and body fixed coordinate systems agree at $t = 0$, and get for torque free motion

$$\dot{J}_i + \epsilon_{ikj}\Omega_k J_j = I_{ij}\dot{\Omega}_j + \epsilon_{ikj}\Omega_k J_j = 0 \ . \tag{6.56}$$

Being moderately intelligent we adapt the rotating coordinate system so that its axes coincide with the principal axes, and arrive at *Euler's equations*

$$\begin{aligned} I_1\dot{\Omega}_1 + (I_3 - I_2)\Omega_2\Omega_3 &= 0 \\ I_2\dot{\Omega}_2 + (I_1 - I_3)\Omega_3\Omega_1 &= 0 \\ I_3\dot{\Omega}_3 + (I_2 - I_1)\Omega_1\Omega_2 &= 0 \end{aligned} \tag{6.57}$$

These equations can be solved exactly in terms of elliptic functions. In the absence of torque the spinning top is an integrable system.

Much of the relevant information can be easily extracted. By inspection we see that the body can execute constant rotation around any principal axis, i.e. a solution is $\Omega_1 = \omega_0 =$ constant, $\Omega_2 = \Omega_3 = 0$. Is this solution stable? Set

$$\Omega_1 = \omega_0 + \omega_1 \qquad \Omega_2 = \omega_2 \qquad \Omega_3 = \omega_3 \ , \tag{6.58}$$

where ω_1, ω_2, ω_3 are time dependent and assumed to be small. Then expand the equations to first order in the unknowns:

$$\begin{aligned} I_1\dot{\omega}_1 &= 0 & I_2\ddot{\omega}_2 + \tfrac{(I_1-I_3)(I_1-I_2)}{I_3}\omega_0^2\omega_2 &= 0 \\ I_2\dot{\omega}_2 + (I_1 - I_3)\omega_0\omega_3 &= 0 & \Rightarrow & \\ I_3\dot{\omega}_3 + (I_2 - I_1)\omega_0\omega_2 &= 0 & I_3\ddot{\omega}_3 + \tfrac{(I_2-I_1)(I_3-I_1)}{I_2}\omega_0^2\omega_3 &= 0 \end{aligned} \tag{6.59}$$

The perturbation will grow, and the solution will be unstable, unless

$$I_1 > I_2, I_3 \qquad \text{or} \qquad I_1 < I_2, I_3 \ . \tag{6.60}$$

We conclude that rotation around the largest and smallest of the principal axes is stable, rotation around the remaining principal axis is unstable.

This can be seen more elegantly. The energy surface is the ellipsoid

$$\frac{J_1^2}{I_1} + \frac{J_2^2}{I_2} + \frac{J_3^2}{I_3} = 2E \ . \tag{6.61}$$

The angular momentum vector itself is changing (in this coordinate system), but its magnitude remains constant:
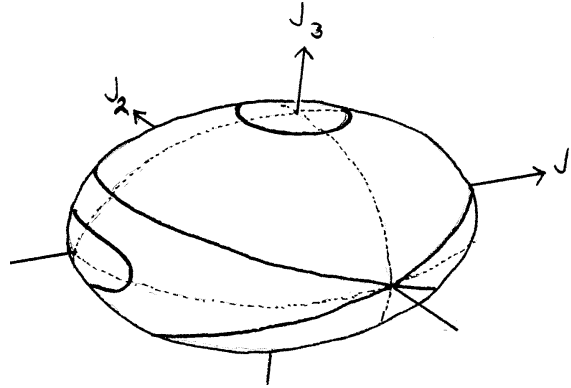
Figure 6.4. The inertia ellipsoid, and its intersections (as curves) with spheres of three different sizes.

$$J_1^2 + J_2^2 + J_3^2 = \text{constant} . \tag{6.62}$$

This is a sphere, intersecting the energy surface along one dimensional curves. The system must move along them. It is seen that there are elliptic fixed points at the major and minor axes, and hyperbolic fixed points at the middle axis. Although it says nothing about the speed of the motion, this analysis does say more than the perturbative calculation since it describes all solutions exactly.

Let us think a little bit more about this. The Euler equations define a dynamical system whose trajectories are confined to a three dimensional space, for which we can use the coordinates J_1, J_2, J_3. A solution of the Euler equations is a curve in this three dimensional space, and curves in a three dimensional space can be very unruly indeed. But the Euler equations are special because of the existence of two well behaved constants of the motion. Each of them defines a set of easy-to-describe surfaces that fill phase space, and any solution must lie where two such surfaces intersect. Hence the solutions are also easy to describe. But this is in fact a quite special property of the Euler equations. The Lorenz equations (1.45) behave quite differently: for them there are no well behaved constants of the motion, the solutions are unruly, and the system is chaotic. We will come back to this kind of issues in chapter 10.

We still lack a complete description of the motion in space. The *Poinsot construction* provides this. It describes how the angular velocity vector moves, relative to the body, and relative to absolute space.[2] The argument may be a bit hard to follow, so please begin by looking at Fig. 6.4. From the point of view of the top the angular momentum vector is moving along a curve on an ellipsoid, but from the point of view of Absolute Space the angular momentum

---

[2] The first edition of Poinsot's book *Théorie nouvelle de la rotation des corps* (1834) contained fifty-six pages of text, no figures, and no equations; he preferred "the simple and natural method of considering things in themselves". It was said in his obituary (1860) that his life "passed happily".

vector is fixed while the top—and its inertia ellipsoid—is moving. We want to turn this simple observation into a more detailed description.

Using inertial coordinates we define the *ellipsoid of inertia*

$$X_i I_{ij} X_j = 2E \ . \tag{6.63}$$

Think of it as a massless shell surrounding the body. If we compute the inertia tensor with respect to Absolute Space we will find that it moves. The numerical value $2E$ is at the moment just a convention, but since the kinetic energy E is constant the convention guarantees that the angular velocity vector Ω_i moves on this ellipsoid. There it traces out a curve known as the *polhode* (meaning "axis path"). In absolute space it traces out a curve known as the *herpolhode*. The angular velocity vector also obeys

$$L_i \Omega_i = 2E \ . \tag{6.64}$$

Since the angular momentum vector is fixed in absolute space we can assume that it points along the X_3-axis, and then Ω_3 is a constant, which means that the herpolhode is a plane curve. The plane to which it is confined is orthogonal to L_i, and is known as the *invariable plane*. Now the normal vector of the ellipsoid of inertia is

$$n_i(X) = 2I_{ij} X_j \ . \tag{6.65}$$

Evaluated at the point Ω_i this is

$$n_i(\Omega) = 2I_{ij}\Omega_j = 2L_i \ . \tag{6.66}$$

This coincides with the normal of the invariable plane. Because of eq. (6.64) the distance from the centre of the ellipsoid of inertia to the invariable plane is constant, and having placed the latter appropriately we conclude that the ellipsoid of inertia rolls on the invariable plane. It rolls without slipping because the point of tangency lies on the instantaneous rotation axis, so its velocity equals zero. Hence the polhode rolls without slipping on the herpolhode. The polhode is always a closed curve, but the herpolhode need not be—when the point of tangency has made one full revolution on the ellipsoid, the body will have turned through some angle around the X_3-axis. Hence there are two frequencies involved, and if they are not commensurable in the sense of eq. (1.37) the herpolhode never closes. This is quite reminiscent of the Lissajous figures.

Things simplify for a symmetrical top because then both the polhode and the herpolhode are circles. Moreover the symmetry axis of the top coincides with a principal axis of the ellipsoid of inertia, which means that the tip of the symmetry axis also traces out a circle around L_i. This motion is called *regular precession*, and is not to be confused with the precession of a top placed in a gravitational field, or with the precession of the equinoxes caused by the

spinning earth. Note that Euler's equations (6.57) for a symmetrical top are easily integrated in terms of trigonometric functions.

The Earth is a symmetrical top, with $I_1 \neq I_2 = I_3$ and

$$\frac{I_1 - I_3}{I_1} \approx \frac{1}{305} \ . \tag{6.67}$$

This is the ellipticity of the Earth. Moreover its rotation axis differs slightly from its geometrical symmetry axis. Judging from Euler's equations, in particular eqs. (6.59) with $\omega_0 = 2\pi/\text{day}$, we expect the rotation axis to move at the rate of one revolution per 305 days, and the polhode to be a circle surrounding the geometrical North Pole. Although small—the polhode is only about 15 meters across—such a motion is indeed observed, and is known as the *Chandler wobble*. However, the polhode is not a circle, and there is a period of close to 14 months as well as an annual periodicity. The latter is presumably caused by metereological disturbances, while the longer period is due to the effect we have discussed. The discrepancy with our prediction arises because the Earth is not perfectly rigid, and can be explained if the Earth has an elasticity approximating that of steel. Poincaré produced arguments showing that a fluid core inside a rigid shell need not invalidate the argument.

Incidentally, from this example one can see how difficult it would be to devise an experimental test of rigid body mechanics. Rigid body mechanics is a collection of theorems that are simply true. Experiments are needed to check if the theorems apply, to estimate the importance of friction, the departure from rigidity, and so on, and to guide mathematical modelling of such effects.

6.5   The Lagrangian description

Evidently the Lagrangian of the rigid body must be $L = T - V$, but going back to our expression (6.43) for the kinetic energy of the body we are momentarily confused. We recognize neither q nor $\dot{q}$. They are still there though. Recall that the configuration space is the group $SO(3)$, whose tangent vectors are Lie algebra elements. According to the discussion is section 6.1 an arbitrary rotation can be obtained by exponentiating an anti-symmetric matrix, of the general form

$$A_{ij} = \epsilon_{ikj}\Omega_k \quad \Leftrightarrow \quad A = \Omega_1 A_1 + \Omega_2 A_2 + \Omega_3 A_3 \ , \tag{6.68}$$

where we use the basis introduced in eqs. (6.22) and Ω_i is the angular velocity vector. Furthermore

$$R = e^{-tA} \quad \Rightarrow \quad A = -\dot{R}R^{-1} \ . \tag{6.69}$$

A minus sign was used because we are now considering an active rotation of the body, rather than a passive change to a rotating coordinate system. Once we have A we can read off the angular velocity vector by comparing to eq. (6.68).

And at the expense of a minor amount of work we can express A in terms of the Euler angles, introduced in section 6.1 as an example of an explicit set of coordinates on the group manifold of $SO(3)$. This is how we will express the Lagrangian of the top in terms of coordinates. The calculation goes as follows:

$$-A = \dot{R}R^{-1} = \frac{d}{dt}(R_\psi R_\theta R_\phi)(R_\psi R_\theta R_\phi)^{-1} =$$

$$= \dot{R}_\psi R_\psi^{-1} + R_\psi \dot{R}_\theta R_\theta^{-1} R_\psi^{-1} + R_\psi R_\theta \dot{R}_\phi R_\phi^{-1} R_\theta^{-1} R_\phi^{-1} = \tag{6.70}$$

$$= -\dot{\psi} A_3 - \dot{\theta}(\cos\psi A_1 + \sin\psi A_2) - \dot{\phi}(\sin\theta(\sin\psi A_1 - \cos\psi A_2) + \cos\theta A_3) \ .$$

In the final step we fell back on the explicit formulæ (6.14). Comparing to eq. (6.68) we read off that

$$\begin{aligned} \Omega_1 &= \dot{\theta}\cos\psi + \dot{\phi}\sin\theta\sin\psi \\ \Omega_2 &= \dot{\theta}\sin\psi - \dot{\phi}\sin\theta\cos\psi \\ \Omega_3 &= \dot{\psi} + \dot{\phi}\cos\theta \end{aligned} \tag{6.71}$$

This we can insert in a Lagrangian.

We will choose a Lagrangian to describe a symmetrical top with one point fixed, subject to a gravitational force. The restriction to a symmetrical top implies that two of the eigenvalues of the inertia tensor are equal, say $I_1 = I_2$; we make this restriction for the pragmatic reason that the general case is not soluble. To connect the Euler angles to the orientation of the top we place the origin at the fixed point, recall eq. (6.15), and let the identity matrix correspond to a top whose symmetry axis is vertical. When the symmetry axis is inclined the angle ϕ describes its rotation around the vertical, the angle θ its inclination, and the angle ψ the rotation of the body around its symmetry axis. Hence these angles are referred to as the angle of precession, nutation, and spin, respectively (or yaw, pitch, and roll, if you are a spacecraft designer). In these coordinates the gravitational potential energy is

$$V = V(\theta) = Mgl\cos\theta \ . \tag{6.72}$$

Now we can write the Lagrangian. We obtain

$$L = \frac{I_1}{2}\left(\Omega_1^2 + \Omega_2^2\right) + \frac{I_3}{2}\Omega_3^2 - V(\theta) =$$

$$\tag{6.73}$$

$$= \frac{I_1}{2}\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 - Mgl\cos\theta \ .$$

For the symmetrical top neither ϕ nor ψ appear in the Lagrangian, which means that—energy included—we will get three constants of the motion, enough to solve the equations of motion explicitly.

To be precise, we find the constants of motion

$$L_z = \frac{\partial L}{\partial \dot{\phi}} = I_1 \dot{\phi} \sin^2 \theta + I_3 \cos \theta (\dot{\psi} + \dot{\phi} \cos \theta) \tag{6.74}$$

$$L_3 = \frac{\partial L}{\partial \dot{\psi}} = I_3(\dot{\psi} + \dot{\phi} \cos \theta) \ . \tag{6.75}$$

They are the angular momentum along the vertical axis (in absolute space) and along the symmetry axis of the top, respectively. We can solve these equations for the velocities:

$$\dot{\psi} + \dot{\phi} \cos \theta = \frac{L_3}{I_3} \tag{6.76}$$

$$\dot{\phi} = \frac{L_z - L_3 \cos \theta}{I_1 \sin^2 \theta} \ . \tag{6.77}$$

Next we turn to the energy, which is

$$E = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi} \cos \theta)^2 + Mgl \cos \theta =$$

$$\tag{6.78}$$

$$= \frac{I_1}{2}\dot{\theta}^2 + \frac{(L_z - L_3 \cos \theta)^2}{2I_1 \sin^2 \theta} + Mgl \cos \theta + \frac{L_3^2}{2I_3} \ .$$

This leads to a first order differential equation which can be explicitly solved in terms of elliptic functions. Inserting the result successively in eqs. (6.77) and (6.76) will lead us to the complete solution for the heavy symmetrical top.

As usual one can go a long way with qualitative arguments. Eq. (6.78), which governs the *nutation* of the top, can be described as one dimensional motion in an effective potential. It can be written on the form

$$\alpha = \dot{\theta}^2 + \frac{(b - a \cos \theta)^2}{\sin^2 \theta} + \beta \cos \theta \ , \tag{6.79}$$

where

$$\alpha = \frac{2EI_3 - L_3^2}{I_1 I_3} \ , \qquad b = \frac{L_z}{I_1} \ , \qquad a = \frac{L_3}{I_1} \ , \tag{6.80}$$

$$\beta = \frac{2Mgl}{I_1} > 0 \ . \tag{6.81}$$

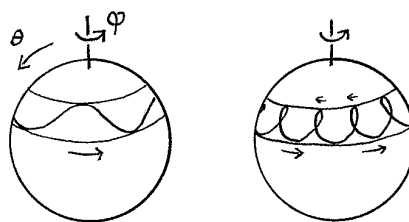If we introduce the variable $u = \cos \theta$ the effective potential becomes a cubic polynomial,

Figure 6.5. Precession might possibly turn retrograde.

$$\dot{u}^2 = \beta u^3 - (\alpha + a^2)u^2 + (2ab - \beta)u + \alpha - b^2 \ . \tag{6.82}$$

The velocity $\dot{\theta}$ can change sign only when the right hand side vanishes. The polynomial on the right hand side goes from $-\infty$ to ∞, and

$$u = \pm 1 \quad \Rightarrow \quad \dot{u}^2 = -(b \mp a)^2 \leq 0 \ . \tag{6.83}$$

On physical grounds $\dot{u}^2$ must be positive somewhere within the physically relevant interval $-1 \leq u \leq 1$. Except for the special case $b = a$ it then follows that the cubic polynomial has two roots in the physically relevant interval. This observation is enough to reveal the key qualitative features of the nutation of the top—the motion 'turns' at the roots of the cubic, which means that the axis is nutating between a maximal and a minimal value of θ.

The precise shape of the cubic, and hence the details of the motion, will depend on the choice of E, L_z, and L_3, or in other words on the initial conditions. For the *precession*, described by eq. (6.77), an interesting qualitative question is whether the derivative $\dot{\phi}$ changes its sign as θ evolves. The answer depends on the relative size of the initial data L_z and L_3; see Fig. 6.5 for the possibilities.

To see why there is precession in the first place, take a top whose point of contact with the ground is fixed but which is otherwise not subject to forces. Let it rotate around its angular momentum vector. Then grab hold of it, trying to increase the angle its axis makes with the vertical axis. If you pull hard enough in a downwards direction you will succeed to some extent, but there will be no z-component of the torque, and hence L_z must remain constant. The top manages to keep L_z constant because it begins to precess, thus adding an extra precessional contribution to L_z.

These considerations are applicable to the Earth, which is an oblate symmetrical top subject to tidal forces (primarily from the Moon) trying to decrease the angle between its axis and the normal of the ecliptic. The monotonic precession of the Earth was known to the Greeks. It has a period of 26 000 years—which incidentally means that the position of the Sun relative to the Zodiac has drifted noticeably since the Greeks determined it, with no apparent consequences to astrology. See exercise 13. The nutation of the Earth was reported by Bradley in the 18th century, who first observed it for a complete

period, 18.7 years. A lesson, perhaps, for modern astronomers. It is a small effect, but significantly greater than the Chandler wobble.

6.6   The tippe top

Turning from the large to the small we find the tippe top, a toy that amazed Bohr and Pauli (and before them William Thompson). What is remarkable about it is that, once let loose, it slides on the floor for a while, and then suddenly turns upside down, spinning stably in this position for quite some time.

Te tippe top can conveniently be approximated as a spherical top of radius R, whose centre of mass is displaced from its geometrical centre by an amount eR. It has five degrees of freedom, the Euler angles, and the position of its centre of mass given by the coordinates x and y. The z-coordinate of the centre of mass is constrained by

$$z = R(1 - e\cos\theta) \ . \tag{6.84}$$

If there were no friction, its kinetic energy would be given by

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + (me^2 R^2 \sin^2\theta + I_1)\frac{\dot{\theta}^2}{2} + \frac{I_1}{2}\dot{\phi}^2 \sin^2\theta + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 \ . \tag{6.85}$$

The angular momentum projected along the vertical axis is

$$p_\phi = I_1\dot{\phi}\sin^2\theta + I_3\cos\theta(\dot{\psi} + \dot{\phi}\cos\theta) \ , \tag{6.86}$$

and the angular momentum projected along the symmetry axis of the top is

$$p_\psi = I_3(\dot{\psi} + \dot{\phi}\cos\theta) \ . \tag{6.87}$$

The motion is a combination of gliding, rolling, and spinning. And each type of motion is associated with some friction. Note that it takes energy to turn the top upside down since its centre of mass will be raised, and this energy must come from the rotational motion, so that the angular momentum must decrease. Hence frictional torques acting on the top will play a key role.

However, in the initial stages of its dramatic life, during which the top suddenly turns upside down, gliding friction is the dominant factor. So let us ignore the friction working against rolling and spinning.

Considering gliding friction, an *Aha-Erlebnis* lies in wait.[3] Let **r** be the vector connecting the centre of mass to the point of contact. Clearly there is no frictional torque acting along this particular direction, which means that we can define a conserved quantity

[3] According to H. Leutwyler, *Why some tops tip*, Eur. J. Phys. **15** (1994) 59.
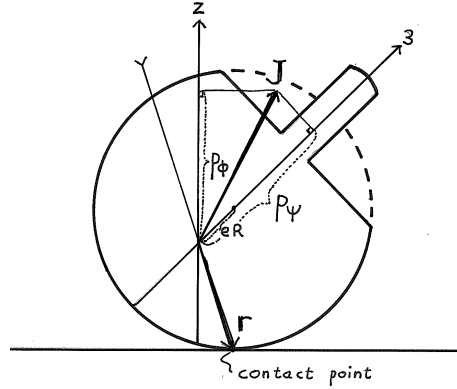
Figure 6.6. The tippe top.

$$J = \mathbf{J} \cdot \mathbf{r} = \mathbf{J} \cdot (e\,\mathbf{e}_3 - \mathbf{e}_z) = ep_\psi - ep_\phi \ . \tag{6.88}$$

Note that both the direction and the magnitude of $\mathbf{r}$ is changing with time. But so is the angular momentum $\mathbf{J}$, and this particular combination stays constant.

Of course we are going to assume that the top is trying to minimize its energy. It is reasonable, considering the initial stages of the motion, to assume that the top is spinning so fast that this means that it minimizes its kinetic energy. Clearly, at the minimum there will be no gliding ($\dot{x} = \dot{y} = 0$) and no rolling ($\dot{\theta} = 0$). Therefore, at the minimum, the kinetic energy depends only on $\dot{\phi}$ and $\dot{\psi}$ and θ.

Now it is a calculation. Setting $\dot{x} = \dot{y} = \dot{\theta} = 0$ and using the conservation law (6.88) to eliminate $\dot{\psi}$ one finds that the kinetic energy becomes

$$T = T(\dot{\phi}, \theta) = \frac{I_1}{2}\dot{\phi}^2 \sin^2\theta + \frac{1}{2I_3}\frac{(J + I_1\dot{\phi}\sin^2\theta)^2}{(\cos\theta - e)^2} \ . \tag{6.89}$$

The minimum with respect to $\dot{\phi}$ is achieved at

$$\dot{\phi} = \frac{-J}{I_3(\cos\theta - e)^2 + I_1\sin^2\theta} \ . \tag{6.90}$$

A further calculation shows that $\dot{\psi} + e\dot{\phi} = 0$ at the minimum. At this minimum the kinetic energy reads

$$T_{\min}(\theta) = \frac{J^2}{2\left(I_3(\cos\theta - e)^2 + I_1\sin^2\theta\right)} \ . \tag{6.91}$$

This is now to be minimized with respect to θ. We see immediately that $T_{\min}(0) > T_{\min}(\pi)$, which is enough to show that the top will not have the orientation it has in the static case. To make sure that the inverted position,

$\theta = \pi$, is the position it will reach, we have to check if $T_{\min}$ is a monotoneously decreasing function of θ. A simple calculation shows that this is the case provided that

$$(1 - e)I_3 < I_1 < (1 + e)I_3 \ . \tag{6.92}$$

Provided that the top is indeed constructed according to this recipe, it will turn upside down during the fast spinning phase of the experiment.

These observations suffice in order to give a qualitative explanation for the remarkable behaviour of our top. A complete description of its motion would tax our abilities to the utmost.

$\diamond$ **Problem 6.1**   You are given an $SO(3)$ matrix explicitly. You know it describes a rotation by an angle α through some axis, but you are not told what axis. What is the quickest way to compute α?

$\diamond$ **Problem 6.2**   Compute the inertia tensor with respect to the centre of mass for a sphere, a cube, a circular cylinder and a circular cone, all of them having constant density.

$\diamond$ **Problem 6.3**   Place four equal masses at the corners of a regular tetrahedron and compute the inertia tensor with respect to their centre of mass. To what extent is the result obvious?

$\diamond$ **Problem 6.4**   Complete the proof that the inertia tensor of any body can be reproduced by placing four equal masses at appropriate distances from each other.

$\diamond$ **Problem 6.5**   The mass of the Sun is $2 \cdot 10^{30}$ kg, its equatorial radius is $7 \cdot 10^8$ m, and its sidereal rotation period is 25 days. Approximate the Sun as a homogeneous sphere and compute its angular momentum. The mass of Jupiter is $2 \cdot 10^{27}$ kg, its semimajor axis is $8 \cdot 10^{11}$ m, and its orbital period is 4332 days. Approximate Jupiter as a point mass in circular orbit and compute its orbital angular momentum. Comment?

$\diamond$ **Problem 6.6**   Suspend a large key-ring in a twisted thread and let go. What happens? Analyze the situation using Euler's equations.

$\diamond$ **Problem 6.7**   Maxwell invented an ingenious device by means of which he could see the instantaneous rotation axis of a spinning top.[4] Read his paper and do the experiment.

$\diamond$ **Problem 6.8**   In the course of a pirouette a skater can increase her angular velocity by drawing her arms close to her body. Set up a model that enables you to discuss the work she must do to increase her kinetic energy.

$\diamond$ **Problem 6.9**   Stare at Fig. 6.4 until it becomes obvious that it must look this way. Put your reasoning in words.

$\diamond$ **Problem 6.10**   Solve Euler's equations for a symmetrical top, $I_2 = I_3$, and draw the analogue of Fig. 6.4.

---

[4] J. C. Maxwell, *On an instrument to illustrate Poinsôt's theory of rotation*, Trans. Royal Scottish Society of Arts, 1855.

◇   **Problem 6.11**   Repeat the derivation leading to eqs. (6.71), but use the equally valid definition

$$-A = R^{-1}\dot{R} \ .$$

What difference does it make?

◇   **Problem 6.12**   For the symmetrical top in a gravitational field there must be a solution for a "sleeping top", that is a top spinning around the vertical axis. Show that this solution is stable if and only if

$$L_z^2 > 4MglI_1 \ . \tag{6.93}$$

The top "wakes up" when friction has diminished its spin so that this bound is violated.

◇   **Problem 6.13**   Around 150 B.C. Hipparchos established the dates when the Sun is in Capricorn. Given that the period of the precession is around 26 000 years, use your understanding of the symmetrical spinning top to establish the direction in which this assignment has been drifting since then, i.e. establish in which sign the Sun actually is when it says in the astrology column that the Sun is in Capricorn.

7 Interlude: Legendre transformations

It is frequently a good idea to perform a transformation of some function that one may be working on. The Fourier transform is an example of this idea. It may turn out that the kind of manipulations that one may wish to perform on the original function are easier to do on its Fourier transform, and once this has been done one may revert to the original kind of function by means of an inverse Fourier transform.

Another example is the *Legendre transformation* of a function $f = f(x)$. We will use it to pass from the Lagrangian to the Hamiltonian formulation of the equations of motion, and later on to obtain an interesting class of "canonical" transformations of phase space. The Legendre transformation is also used to great effect in thermodynamics. In this interlude we discuss it from a purely mathematical point of view, assuming for simplicity that the function depends on one variable only.

We assume that the function to be transformed is twice differentiable, and obeys

$$\frac{d^2 f}{dx^2} > 0 \ . \tag{7.1}$$

This is not strictly necessary (see exercise 2) but convenient for our purposes. A function obeying this condition is said to be *convex*. The Legendre transform of the convex function f is a function g defined as

$$g(p) = \max_x \big(xp - f(x) \big) \ . \tag{7.2}$$

The maximum that we ask for is achieved if and only if

$$p = \frac{df}{dx} \ . \tag{7.3}$$

Precisely because we imposed condition (7.1) the derivative of f is a monotoneously increasing function of x, which means that eq. (7.3) can be solved to give the unique solution $x = x(p)$. Inserting this into the right hand side of eq. (7.2) defines the function $g(p)$ uniquely.

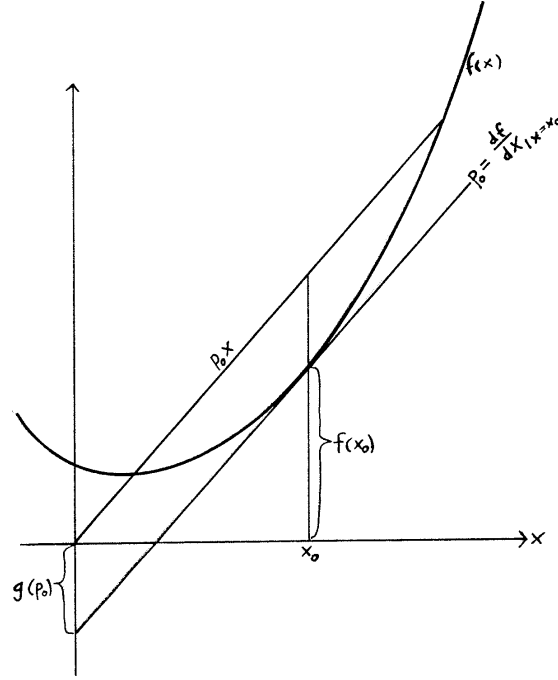A function and its Legendre transform are related by

Figure 7.1. The Legendre transformation of a function $f(x)$. The connection to eqs. (7.2) and (7.4) should be evident.

$$f(x) + g(p) = xp \ . \tag{7.4}$$

If eq. (7.3) holds this equation defines the Legendre transform $g(p)$ once the original function $f(x)$ is known. But the equation is fully symmetric between f and g, which means that if

$$x = \frac{dg}{dp} \tag{7.5}$$

then eq. (7.4) defines $f(x)$ as the Legendre transform of the function $g(p)$. To make sure that this claim is valid one must show that the function g is convex because the function f is. But this is so because

$$\frac{d^2 g}{dp^2} = \frac{dx}{dp} = \left(\frac{dp}{dx}\right)^{-1} = \left(\frac{d^2 f}{dx^2}\right)^{-1} > 0 \ . \tag{7.6}$$

Hence eq. (7.5) can be solved for $p = p(x)$. We conclude that all the information in the function f is present in its Legendre transform g, and also that—unlike the Fourier transformation, say—the Legendre transformation is its own inverse. Generalization to an arbitrary number of variables is quite straightforward.

There is an interesting geometric interpretation of the Legendre transformation, and I hope this can be deciphered from Fig. 7.1. To find out the value of the function g at p_0, start by drawing the straight line $p_0 x$, and then use the definition of the Legendre transform to find $g(p_0)$ as an intercept of a tangent of the curve with the ordinate axis. If the original function describes a curve by giving the set of points on the curve, the Legendre transform describes it by giving the set of lines that are tangent to it. This is a rather deep idea, but one that will be kept in the background of our story.

One branch of physics that uses Legendre transformations extensively is that of thermodynamics. You may recall that a thermodynamic system can be defined by specifying the energy U as a suitable function of the entropy S and the volume V (say). This function is chosen in such a way that the partial derivatives of the function $U = U(S, V)$ are the temperature and the pressure of the system,

$$T = \frac{\partial U}{\partial S} \ , \qquad P = -\frac{\partial U}{\partial V} \ . \tag{7.7}$$

(Sign conventions in thermodynamics are not chosen with mathematical convenience in mind.) This information is often summarized by writing

$$dU = TdS - PdV \ . \tag{7.8}$$

Here we use the notation of *differential forms*. For the present you can simply regard the formula as equivalent to the pair of equations that precedes it. Now the point is that physicists find the entropy a hard variable to control. We would prefer to use the temperature $T = T(S, V)$ as a coordinate instead. To do so we introduce a new function $F = F(T, V)$, called the free energy, related to the energy $U = U(S, V)$ by a Legendre transformation:

$$F = U - TS \quad \Rightarrow \quad dF = -SdT - PdV \ . \tag{7.9}$$

Again: the notation used *means* that F is a function of T and V whose partial derivative with respect to T equals $S = S(T, V)$. Provided that the function $U(S, V)$ is chosen suitably this is indeed the *same* function that you would obtain by solving the previous equation $T = T(S, V)$ for S.

Chemists do keep track of heat exchange but find the volume a hard variable to control in their bubbling concoctions. They therefore perform a different Legendre transformation, and define the enthalphy

$$H = U + PV \quad \Rightarrow \quad dH = TdS + VdP \ . \tag{7.10}$$

Sometimes one makes a Legendre transformation in both variables, and obtains the Gibbs free energy $G = G(T, P)$. The grand idea behind all this—that all the physical information about a system is coded into a single function—is peculiar to thermodynamics, but the mathematical manoeuvres will recur in section 8.4.

$\diamond$ **Problem 7.1**   Compute the Legendre transform of a parabola, i.e. the function $f(x) = c_1 + c_2 x^2$ where c_1 and c_2 are constants. Check that eq. (7.6) holds.

$\diamond$ **Problem 7.2**   Consider the function

$$f(x) = \begin{cases} 1 - x & , & x < 1 \\ 0 & , & 1 \le x \le 2 \\ x - 2 & , & 2 < x \end{cases} \quad . \tag{7.11}$$

Compute the Legendre transform using the definition (7.2).

8 The Hamiltonian formulation

Any set of ordinary differential equations can be written in first order form, provided we introduce enough extra variables. The Hamiltonian formulation is a special way of doing this to the Euler-Lagrange equations, and reveals that the latter have a very special form. The central features of mechanics—those that classical and quantum mechanics have in common—are brought out very clearly by the Hamiltonian formulation.

8.1 Hamilton's equations and Hamiltonian flows

For a Lagrangian of the form $L(q, \dot{q})$, which is general enough for our purposes, the Euler-Lagrange equations are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \ . \tag{8.1}$$

These are second order ODEs. One obvious way to turn them into first order equations is to define the *canonical momenta*

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \ . \tag{8.2}$$

The expression on the right hand side will be equal to $m\dot{q}_i$ in simple cases, and in general it will be some function of $\dot{q}$ and q. We already know that the right hand side is of some importance; it occurs when one sets boundary conditions in the variational principle, and in connection with Noether's theorem. See eq. (2.34). So the canonical momenta are the extra variables to be used in turning the Euler-Lagrange equations into first order form. The use of "q" for generalized coordinates and "p" for their momenta goes back to Jacobi, and was solidified by Whittaker in a famous textbook.

Eqs. (8.1) now take the first order form

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \ . \tag{8.3}$$

But we need equations for $\dot{q}$ too. To this end we assume that eqs. (8.2) can be inverted, that is

$$p_i = p_i(q, \dot{q}) \quad \Leftrightarrow \quad \dot{q}_i = \dot{q}_i(q, p) \; . \tag{8.4}$$

Now we do have a first order system equivalent to the original Euler-Lagrange equations. As a matter of fact it can happen that it is impossible to solve eqs. (8.2) for the velocities; exactly how to deal with every possibility was explained by Dirac when he developed his theory of constrained Hamiltonian systems.[1] Although this is interesting, from now on we assume for simplicity that the equations are invertible.

Now form the function

$$H(q, p) = \dot{q}_i p_i - L(q, \dot{q}) \; . \tag{8.5}$$

Thus H and L are Legendre transformations of each other, in the sense of chapter 7. It is a partial Legendre transformation, since only the $\dot{q}$ are involved while the coordinates q are left untouched.

It looks as if

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \; . \tag{8.6}$$

This is true, but not quite trivial. However, recalling that $\dot{q}_i = \dot{q}_i(q, p)$ and using the chain rule, we obtain

$$\frac{\partial H}{\partial p_i} = \dot{q}_i + \frac{\partial \dot{q}_j}{\partial p_i} \left(p_j - \frac{\partial L}{\partial \dot{q}_j} \right) \; . \tag{8.7}$$

The definition of the canonical momenta gives the result. Recalling the Euler-Lagrange equations we also have

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \; . \tag{8.8}$$

Taken together, eqs. (8.6-8.8) are known as *Hamilton's equations*. They are fully equivalent to the Euler-Lagrange equations. The function H is known as the *Hamiltonian*, and the variables p_i are known as the *canonical momenta*. By means of the Legendre transformation we have traded the coordinate $\dot{q}$ for the coordinate p.

In practice, the Legendre transformation is usually easy to perform. We know that

$$L = \frac{m}{2} \dot{q}^2 \quad \Leftrightarrow \quad H = \frac{1}{2m} p^2 \; . \tag{8.9}$$

An apparently more complicated case, involving many degrees of freedom, is

[1] He wrote an admirably short book about it: P. A. M. Dirac: *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, New York 1964.

$$L = \frac{1}{2}M_{ij}(q)\dot{q}_i\dot{q}_j - V(q) \; . \tag{8.10}$$

Here the matrix M depends on the configuration space variables. See exercise 2.4, or for a concrete example see the Lagrangian for a spinning top in section 6.5. Using matrix notation the Hamiltonian is immediately found to be

$$H = \frac{1}{2}p_j M_{ij}^{-1} p_j + V(q) \; . \tag{8.11}$$

We only have to check that the matrix M is invertible everywhere in configuration space. This in fact amounts to checking that the Lagrangian is a convex function of the variables $\dot{q}_i$.

Our phase space is spanned by the $2n$ variables q_i, p_i. If we compare to the general phase spaces discussed in section 1.4 we see that this is already a restriction: the phase space of a Hamiltonian system is always even dimensional. But there is another and more dramatic difference. In section 1.4 time evolution was described by the equations $\dot{z}_i = f_i(z)$, so that the general case is obtained by choosing $2n$ independent functions f_i. In the Hamiltonian case the time evolution is determined by a single function $H(q,p)$. This is indeed a very strong restriction, but it is one that Nature seems to respect for all her fundamental equations.

There are consequences. One of them is that

$$\dot{H} = \dot{q}_i\frac{\partial H}{\partial q_i} + \dot{p}_i\frac{\partial H}{\partial p_i} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \; . \tag{8.12}$$

Thanks to the minus sign in eq. (8.8) this is is zero whenever the Hamiltonian is not an explicit function of the time t. Hence the Hamiltonian is always a conserved quantity, unless it depends explicitly on time. The set of points in phase space obeying

$$H(q,p) = E \tag{8.13}$$

is called the *energy surface*. Time evolution takes place within the energy surface. In two dimensions (for one degree of freedom) the energy 'surface' coincides with the one dimensional trajectories; hence Hamiltonian systems with one degree of freedom are always soluble.

Next let us imagine time evolution as the flow of a fluid. The flow lines are defined by the little arrows in phase space. The fluid is incompressible if, for any fixed finite volume in phase space, the amount of fluid going out through its surface equals the amount that is going in. By Gauss' theorem the difference between them is equal to the integral of the divergence of the flow over the volume. Since this must vanish for every volume we conclude that the flow is that of an incompressible fluid if and only if its divergence vanishes. For Hamiltonian time evolution this is indeed so:
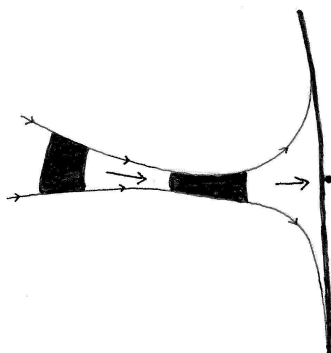
Figure 8.1. Liouville's theorem in action: a volume element in phase space—here shown at three successive moments—preserves its volume, but may become very distorted—here because it approaches a hyperbolic fixed point.

$$\mathbf{div}\dot{\mathbf{z}} = \nabla \cdot \dot{\mathbf{z}} = \partial_{q_i}\dot{q}_i(q,p) + \partial_{p_i}\dot{p}_i(q,p) = \partial_{q_i}\partial_{p_i}H + \partial_{p_i}(-\partial_{q_i}H) = 0 \; . \quad (8.14)$$

This conclusion is worth stating as a theorem.

Liouville's theorem: In Hamiltonian mechanics the phase space flow preserves volume.

Liouville's theorem is the origin of the claim—made in section 1.4—that sources and sinks do not occur in Hamiltonian systems. In two dimensions the only fixed points that occur are either elliptic or hyperbolic.

On reflection one sees that the behaviour allowed by Liouville's theorem can still be very complex. Already in two dimensions the shape of a small piece of the phase space fluid passing close to a hyperbolic fixed point will be squeezed in one direction and stretched in another. In the end the original volume element can acquire a very involved shape, and for an observer with limited resolution it may in fact seem as if it has been smeared all over the energy surface, even though on microscopic scales it does preserve its volume. Indeed Liouville's theorem rather enhances this effect.

8.2   The algebraic structure of mechanics

In Hamiltonian mechanics a single function $H = H(q,p)$ on phase space is used to specify time evolution. This is reminiscent of quantum mechanics. Let us recall the basic facts about the latter. Time evolution is specified by choosing a definite Hermitian operator $\hat{H}$. In the Heisenberg picture the time evolution of an arbitrary operator $\hat{A}$ is determined by the equation

$$\frac{d\hat{A}}{dt} = i\hbar[\hat{A}, \hat{H}] \ . \tag{8.15}$$

Here the bracket represents a commutator. It obeys the algebraic relations

$$[a_1\hat{A}_1 + a_2\hat{A}_2, \hat{B}] = a_1[\hat{A}_1, \hat{B}] + a_2[\hat{A}_2, \hat{B}] \ , \tag{8.16}$$

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \ , \tag{8.17}$$

$$[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \ , \tag{8.18}$$

$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0 \ , \tag{8.19}$$

where a_1, a_2 are arbitrary numbers. These relations also characterize the *Lie bracket* occurring in the study of Lie groups (see section 6.1); the last relation is known as the *Jacobi identity*.

In classical mechanics all functions commute, so at first sight there does not seem to be a classical analogue of the commutator. But a possible candidate was found by Poisson in the early nineteenth century. It is called the *Poisson bracket*, and it is defined, for two arbitrary phase space functions $A = A(q, p)$ and $B = B(q, p)$, as

$$\{A, B\} = \frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} \ . \tag{8.20}$$

(Yes, as always a sum over repeated indices is understood.) By inspection, we see that it enjoys all the algebraic properties of the commutator, namely

$$\{a_1 A_1 + a_2 A_2, B\} = a_1\{A_1, B\} + a_2\{A_2, B\} \ , \tag{8.21}$$

$$\{A, B\} = -\{B, A\} \ , \tag{8.22}$$

$$\{A, BC\} = B\{A, C\} + \{A, B\}C \ , \tag{8.23}$$

$$\{A, \{B, C\}\} + \{C, \{A, B\}\} + \{B, \{C, A\}\} = 0 \ . \tag{8.24}$$

Therefore phase space comes equipped with a Lie bracket. Finally, using Hamilton's equations, the time derivative of any phase space function can be written as

$$\frac{dA}{dt} = \frac{\partial A}{\partial q_i}\dot{q}_i + \frac{\partial A}{\partial p_i}\dot{p}_i = \frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} = \{A, H\} \ . \tag{8.25}$$

(Notice that the minus sign in Hamilton's equations is essential here.) The analogy to the quantum mechanics is complete.

These observations led Dirac, in the course of a Sunday walk in Cambridge, to the belief that any classical system can be "quantized" by setting up a correspondence between functions on phase space and operators on a (complex) linear space, and by replacing the Poisson brackets with commutators according to the rule

$$\{A, B\} \rightarrow \frac{1}{i\hbar}[\hat{A}, \hat{B}] \ . \tag{8.26}$$

On the whole, although various complications do arise, Dirac's idea has proved to be correct.

We can now write Hamilton's equation in two equivalent forms,

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases} \Leftrightarrow \begin{cases} \dot{q} = \{q, H\} \\ \dot{p} = \{p, H\} \ . \end{cases} \tag{8.27}$$

For the explicit evaluation of Poisson brackets we look carefully at the definition, and observe that

$$\{A, F\} = (\partial_q A \partial_p - \partial_p A \partial_q)F \ . \tag{8.28}$$

Hence we can think of $\{A, \ \}$ as a differential operator acting on phase space functions. It obeys Leibniz' rule, namely the important eq. (8.23). In this sense—which is a very important one—it acts like a derivative. Using this the evaluation of any Poisson bracket can ultimately be made starting from the *fundamental Poisson brackets*

$$\{q_i, p_j\} = \delta_{ij} \ , \qquad \{q_i, q_j\} = \{p_i, p_j\} = 0 \ . \tag{8.29}$$

We confirm that the Poisson bracket "acts like a derivative" in the sense that

$$\{q_i, A\} = \partial_{p_i} A \ , \qquad \{p_i, A\} = -\partial_{q_i} A \ , \tag{8.30}$$

where $A = A(q, p)$ is any function on phase space. Note that the quantum mechanical commutator behaves in a similar way, except that in quantum mechanics we have to care about the order in which we place things. The ordering on the right hand side of eq. (8.23) would be the correct one to use there, while ordering does not matter in classical mechanics.

Still the Poisson bracket seems to appear out of thin air. In section 8.5 we will start an excursion into the geometry of phase space, to understand where it came from.

8.3 Canonical transformations

We have yet to discuss the classical analogue of unitary transformations. Unitary transformations preserve commutators, so we define *canonical transformations* as transformations that preserve Poisson brackets. It will eventually become clear that this requirement is analogous to the definition of rotations as those transformations that leave a specific metric invariant, eq. (8.73). As always, we can regard the transformations as active, giving rise to a movement in phase space (or Hilbert space), or as passive transformations from one coordinate description to another. Both viewpoints are useful, but for the general discussion the active viewpoint is the better.

Anyway, if we are given a transformation on phase space

$$q \to q' = q'(q,p) , \qquad p \to p' = p'(q,p) , \tag{8.31}$$

then phase space functions are transformed according to

$$A'(q',p') = A(q,p) . \tag{8.32}$$

("The new function takes the same value at the new point as the old function does at the old point.") The transformation is canonical if, for all phase space functions,

$$\{A,B\} = C \quad \Rightarrow \quad \{A',B'\} = C' . \tag{8.33}$$

How do we find such transformations?

We begin by looking for infinitesimal canonical transformations. Let us express the function A' as a function of q and p, and consider

$$A'(q,p) = A(q,p) + \delta A(q,p) . \tag{8.34}$$

The function δA must be chosen so that eq. (8.33) holds, and is assumed to be small so that terms of second order in δA can be ignored when this property is checked. To obtain a candidate δA we choose a function $F = F(q,p)$ on phase space, and set

$$\delta A = \epsilon \{A,F\} . \tag{8.35}$$

Hence, to first order in ϵ,

$$A' = A + \epsilon \{A,F\} , \qquad B' = B + \epsilon \{B,F\} , \tag{8.36}$$

and so on. Computing to first order in ϵ and using the Jacobi identity we find that

$$\{A', B'\} = \{A, B\} + \epsilon \left(\{\{A, F\}, B\} + \{A, \{B, F\}\} \right) =$$

$$= C - \epsilon\{F, \{A, B\}\} = C + \epsilon\{C, F\} = C + \delta C = C' \tag{8.37}$$

(to that order). This works for *any* phase space function F. In principle it is possible to integrate eq. (8.35) to obtain a finite canonical transformation—the equation has the same form as Hamilton's equations of motion, and the only catch is that it may be difficult to do the integration explicitly.

Again, note the similarity to quantum mechanics: in quantum mechanics any Hermitian operator generates a unitary transformation, in classical mechanics any phase space function generates a canonical transformation. At the same time there is an interesting difference between canonical and unitary transformations: there are many more phase space functions than Hermitian operators, hence there are many more canonical transformations. This has to do with the fact that the unitary transformations also preserve a scalar product, and in this sense they are analogous to rotations. If the space has a finite dimension, there are indeed rather "few" rotations, but the set of canonical transformations is always infinite dimensional because the space of all functions has infinite dimensions. (The discerning reader may think that the comparison is unfair, since rotations take place in a vector space, while the kind of transformation we now allow is more general than that. However, it can be shown that the set of transformations leaving a given metric tensor invariant is at most as large as the set of translations and rotations in a vector space, and indeed often smaller than that. So the objection has no force.)

Let us revisit the discussion of Noether's theorem in section 2.3. We used it already to motivate the definition of the canonical momenta, but there is more to say about it. Let us choose spatial rotation, eq. (2.44), as an example of a transformation that we will want to make. It is easy to see that

$$L_i = \epsilon_{ijk} x_j p_k \quad \Rightarrow \quad \{x_i, \epsilon_j L_j\} = \epsilon_{ijk} \epsilon_j x_k = \delta x_i . \tag{8.38}$$

The Noether charge generates the transformation via the Poisson bracket! More is true. The ϵ-tensor obeys the identity

$$\epsilon_{ijk} \epsilon_{kmn} + \epsilon_{ink} \epsilon_{kjm} + \epsilon_{imk} \epsilon_{knj} = 0 . \tag{8.39}$$

By means of this identity it is easy to show that

$$\{L_i, L_j\} = \epsilon_{ijk} L_k . \tag{8.40}$$

The educated way of describing this result is to say that the rotational Noether charges form a Poisson bracket representation of the Lie algebra of the rotation group. In quantum mechanics, we have a commutator representation of the same Lie algebra.

These results concerning rotations can be generalized to any symmetry and any Noether charges. Let us just mention time translation invariance, eq.

(2.46). Here the Noether charge is the Hamiltonian itself. In perfect analogy to our discussion of rotations, we find

$$\delta x_i = \epsilon \dot{x}_i = \{x_i, \epsilon H\} \ . \tag{8.41}$$

Again the Noether charge generates the transformation to which it owes its existence.

8.4 General transformation theory

So far we have discussed infinitesimal canonical transformations only. Now we turn to finite ones. We will give the general transformation theory for the special case of two dimensional phase spaces only; however, every detail of the following arguments does have a natural—and in most cases immediate—generalization to the higher dimensional cases, so the loss of generality is only apparent.

First of all we adopt a practice introduced by Whittaker, who denoted the transformed coordinates by capital letters rather than primes.[2] Thus—in the active interpretation—the point labelled by (q, p) is transformed to the point labelled (Q, P),

$$Q = Q(q, p) \qquad P = P(q, p) \ . \tag{8.42}$$

This is a canonical transformation if the Poisson brackets obey

$$\{Q, P\} = \{q, p\} \ . \tag{8.43}$$

But in the two dimensional case this is recognizable as the requirement that the Jacobian of the coordinate transformation equals one,

$$\{Q, P\} = \frac{\partial(Q, P)}{\partial(q, p)} = \begin{vmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\[2mm] \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{vmatrix} = 1 \ . \tag{8.44}$$

So the requirement is that the area element stays the same,

$$dA = dq\, dp = dQ\, dP \ . \tag{8.45}$$

Canonical transformations preserve areas. Since our phase space is two dimensional, this is just a restatement of Liouville's theorem.

We can go a little deeper into this by using Stokes' theorem to convert an integral over an area to an integral along its boundary,

---

[2] Whittaker's *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies* was the standard reference at the time quantum mechanics was developed.

$$\int_A dqdp = \oint_C qdp = \oint_C (d(qp) - pdq) = -\oint_C pdq \; . \qquad (8.46)$$

This gives us a few new ways to formulate the requirement that a transformation be canonical. For instance

$$\oint_C pdq = \oint_C PdQ \; . \qquad (8.47)$$

Since this is to hold for every curve that encloses an area, however small, we can take away the integrals. In doing so we must remember that the integral is unchanged if we add a total divergence to the integrand. Hence the transformation is canonical if and only if there exists a function F on phase space such that

$$pdq = PdQ + dF \; . \qquad (8.48)$$

The expression pdq is the *symplectic one-form*, defined up to the addition of the gradient of an arbitrary function and discussed in more detail in section 8.7.

At this point it is helpful to recall your thermodynamics, or those parts of it that were discussed in chapter 7. Let us assume that the canonical transformation we are about to construct is such that q and Q together also can serve as coordinates on phase space. (See exercise 8.) Denote the function F by $F_1 = F_1(q, Q)$. Then we derive that

$$(p - \partial_q F_1)\, dq - (P + \partial_Q F_1)\, dQ = 0 \; . \qquad (8.49)$$

Since q and Q serve as coordinates—since dq and dQ are linearly independent one-forms—we can conclude from this that

$$p = \partial_q F_1(q, Q) \qquad P = -\partial_Q F_1(q, Q) \; . \qquad (8.50)$$

We assume that these equations are invertible, that is to say that they can be used to derive the explicit formulas

$$\begin{cases} Q = Q(q, p) \\ P = P(q, p) \end{cases} \Leftrightarrow \begin{cases} q = q(Q, P) \\ p = p(Q, P) \end{cases} \qquad (8.51)$$

This is the canonical transformation *generated* by the *generating function* $F_1(q, Q)$.

You may have wondered how I knew that I could use q and Q as coordinates on phase space. The answer is that I did not know, but in a concrete situation I would have noticed whether anything was wrong once I tried to invert eqs. (8.50) in order to obtain the explicit canonical transformation. If you think about it, it is clear that there cannot exist a generating function $F_1(q, Q)$

generating the simplest canonical transformation we can think of, namely the identity transformation

$$Q = q \qquad P = p \ . \tag{8.52}$$

Evidently the pair q and Q does not coordinatize phase space in this case. To circumvent this difficulty we define a new generating function by means of a Legendre transformation from the first:

$$F_2 = F_1(q, Q) + QP \ . \tag{8.53}$$

The sign conventions here are a little odd. Still the relation between the functions F_2 and F_1 is recognisably a Legendre transformation. In fact, using eqs. (8.50),

$$dF_2 = dq \frac{\partial F_1}{\partial q} + dQ \frac{\partial F_1}{\partial Q} + PdQ + QdP = pdq + QdP \ . \tag{8.54}$$

It follows that F_2 is a function of q and P only. If we assume that q and P together coordinatize phase space—which is certainly true for the identity transformation—we can use F_2 to generate canonical transformations by repeating the previous logic:

$$pdq = d(PQ) - QdP + dF_1 = -QdP + dF_2 = -QdP + \partial_q F_2 dq + \partial_P F_2 dP \ . \tag{8.55}$$

We compare the coefficients in front of the one-forms dq and dP, keeping in mind that they can be regarded as linearly independent vectors because q and P serve as coordinates. See section 8.7. We conclude that

$$p = \partial_q F_2 \qquad Q = \partial_P F_2 \ . \tag{8.56}$$

Again we assume that these equations can be inverted to yield eqs. (8.51) in explicit form. The identity transformation is generated by

$$F_2(q, P) = qP \ . \tag{8.57}$$

We can go on to define the generating functions $F_3 = F_3(p, Q)$ and $F_4 = F_4(p, P)$ by means of further Legendre transformations, in the obvious way.

We have recovered the conclusion from section 8.3, namely that every function on phase space generates a canonical transformation, but now we have a recipe for how to find such transformations in finite—not only infinitesimal—form. Hamilton's equations retain their form under all canonical transformations, with a new Hamiltonian defined as the function

$$K(Q, P) = H(q, p) = H(q(Q, P), p(Q, P)) \ . \tag{8.58}$$

In chapter 10 we will regard canonical transformations as passive coordinate

transformations, and we will try to design a canonical transformation to ensure that the new Hamiltonian takes a very simple standard form, so that the equations of motion are always soluble once we have arrived at the new coordinates Q and P.

The setup can be generalized somewhat, by allowing the generating transformations to depend on time. Hamilton's equations will again retain their form, but the numerical value of the Hamiltonian will change. In fact, at least on the level of Gedanken calculations, it is possible to choose a time dependent generating function ensuring that the transformed Hamiltonian vanishes, so that the resulting equations of motion solve themselves. This is the subject of chapter 9.

8.5   Kets and bras and all that

There is an intriguing geometrical structure behind the Hamiltonian equations of motion. To appreciate it, we begin by recalling the "kets and bras" notation in quantum mechanics. In quantum mechanics the state space is a vector space. Vectors are called *kets* and denoted by $|\psi\rangle$. They form a vector space because one can take linear combinations of kets: if $|\psi_1\rangle$ and $|\psi_2\rangle$ are kets and a_1, a_2 are numbers then $|\psi_3\rangle$ is a ket too, where

$$|\psi_3\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle \ . \tag{8.59}$$

In quantum mechanics the numbers a_1, a_2 are complex, but from now on we assume that all such numbers are real. It is important to realize that, in quantum mechanics, one never takes the scalar product of two vectors—that is to say of two kets. What one does is to introduce another vector space whose vectors are called *bras*, and denoted by $\langle\phi|$. By definition a bra is a linear function from the vector space of kets to the real numbers. Thus, given a bra $\langle\phi|$,

$$\langle\phi|\psi\rangle = a \tag{8.60}$$

is defined for all kets $|\psi\rangle$. Dirac calls this a "bracket", which explains the names "kets and bras". (It is believed that Dirac did not know that "bra" already had a meaning in English.) An added complication in quantum mechanics is that the vector spaces are often infinite dimensional function spaces, but we do not bother with this here. We do need to check that the bras form a linear space. But this is so because the map to the real numbers is linear,

$$(a_1\langle\phi_1| + a_2\langle\phi_2|)|\psi\rangle = a_1\langle\phi_1|\psi\rangle + a_2\langle\phi_2|\psi\rangle \ . \tag{8.61}$$

Hence $\langle\phi_3| = a_1\langle\phi_1| + a_2\langle\phi_2|$ is a bra.

At this point we have two vector spaces, and *a priori* they are different. The equation $\langle\psi| = |\psi\rangle$ makes no sense at all. Apples and pears are never equal.

However, there is a final twist to the story, because there does exist a one-to-one correspondence between the kets and the bras. If you have a ket written out as a column vector with respect to some basis, there is a unique bra obtained by transposing the column to a row (and taking the complex conjugates of all the numbers, but this is irrelevant here since all our components are real). This is an extra piece of structure, giving rise to a one-to-one correspondence $\langle\psi| \leftrightarrow |\psi\rangle$. It means that we can associate a unique number to any ket, namely the real number $\langle\psi|\psi\rangle$. This number is the "length squared" of the ket $|\psi\rangle$.

How can we make the distinction between bras and kets in the index notation? The answer is simple. Let us denote all the original vectors, the kets, by x^μ. We are using Greek rather than Latin indices, and will continue to do so whenever it is our intention to use the indices in a correct "tensorial" way (that is, in the way I am just going to explain). When the indices are only used to label a set of objects we continue to use Latin indices. The important thing is that we place the index on a ket "upstairs". Such a vector is called *contravariant*. A bra will be denoted by u_μ, with its index "downstairs", and is called a *covariant* vector. This should represent a linear map from the space of all kets to the real numbers, and we define this map as

$$u(x) = u_\mu x^\mu = a \ , \tag{8.62}$$

where Einstein's summation convention is understood. It is important to note that in the upstairs-downstairs notation, the expressions $x^\mu y^\mu$ and $u_\mu v_\mu$ are simply not allowed, so that we cannot define the length squared of a vector x^μ as $x^\mu x^\mu$. It is forbidden. There is no notion of length squared until we have introduced further structure. With these new rules, many of the formulæ in, say, section 6.1, are simply forbidden. We still want to use them however, so we have to modify them in a suitable way.

Let us agree that x^μ simply stands for the components of a ket relative to some basis in the ket vector space. That is to say that

$$|\psi\rangle = |e_\mu\rangle x^\mu \ , \tag{8.63}$$

where the vectors $|e_\mu\rangle$ form a basis (and must not be confused with the components u_μ of a bra vector!). A matrix operating on the ket vector space will be written as a "mixed" tensor $A^\mu{}_\nu$. Then the equations

$$y^\mu = A^\mu{}_\nu x^\nu \quad \text{and} \quad v_\mu = u_\nu B^\nu{}_\mu \tag{8.64}$$

do make sense—given that we always use the Einstein summation convention for repeated indices in "upstairs-downstairs" position. As a matter of fact these equations make two different kinds of sense. The vector

$$|\psi'\rangle = |e_\mu\rangle y^\mu = |e_\mu\rangle A^\mu{}_\nu x^\nu = A|\psi\rangle \tag{8.65}$$

is a new vector, obtained by transforming the old vector $|\psi\rangle$. This is an active transformation of the vector. On the other hand the vector

$$|\psi\rangle = |e_\mu\rangle x^\mu = |e_\sigma\rangle (A^{-1})^\sigma{}_\mu A^\mu{}_\nu x^\nu = |e'_\mu\rangle y^\mu \tag{8.66}$$

is just the original vector $|\psi\rangle$ in a new coordinate system. This is a passive transformation of the vector.

In a passive transformation the components and the basis vectors transform in opposite ways. This is an important idea. We will insist that, whatever linear transformation we do of the bra and ket vector spaces, the numbers $u_\mu x^\mu$ should remain unchanged. We can arrange this by insisting that whenever we transform the kets, we also transform the bras in the opposite way. That is to say

$$x^\mu \rightarrow A^\mu{}_\nu x^\nu \qquad \Leftrightarrow \qquad u_\mu \rightarrow u_\nu (A^{-1})^\nu{}_\mu \ . \tag{8.67}$$

This is part of the origin of the names "contravariant" and "covariant" vectors—they transform in opposite directions, to ensure that $u_\mu x^\mu$ remains unchanged.

We can go on to define tensors with more than one index in the same way. A tensor with k indices running from 1 to n is defined as a collection of kn components transforming in specific way under changes of basis. Examples include

$$T^{\mu\nu\sigma} \rightarrow A^\mu{}_\alpha A^\nu{}_\beta A^\sigma{}_\gamma T^{\alpha\beta\gamma} \ , \qquad S_{\mu\nu\sigma} \rightarrow S_{\alpha\beta\gamma}(A^{-1})^\alpha{}_\mu (A^{-1})^\beta{}_\nu (A^{-1})^\gamma{}_\sigma \ . \tag{8.68}$$

Students are usually disturbed by the fact that the components of a tensor with more than two indices cannot be displayed as a matrix, but really the definition does not require this.

Still something is missing. The whole point about chapter 6 was to discuss the special matrices that preserve the length of the vectors. To do so here we need to set up a one-to-one correspondence between the set of x^μ and the set of u_μ, and then define the analogue of $\langle\psi|\psi\rangle$.

The Kronecker delta is the key. We write it as a covariant tensor—with both indices downstairs—as $\delta_{\mu\nu}$. In fact we can be a bit more general. We introduce a *metric tensor*. By definition this is any covariant symmetric tensor with two indices downstairs,

$$g_{\mu\nu} = g_{\nu\mu} \ , \tag{8.69}$$

which has a contravariant inverse $g^{\mu\nu}$, in the sense that

$$g^{\mu\sigma} g_{\sigma\nu} = \delta^\mu{}_\nu \ . \tag{8.70}$$

Here the left hand side defines an operation which acts as the identity on both our vector spaces,

$$\delta^\mu{}_\nu x^\nu = x^\mu \quad \text{and} \quad u_\nu \delta^\nu{}_\mu = u_\mu \ . \tag{8.71}$$

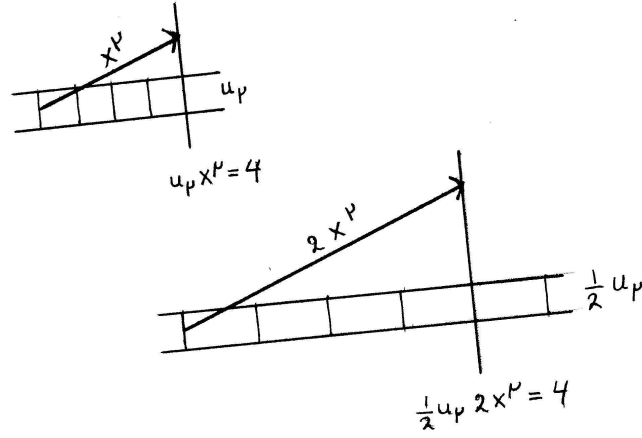Thus $\delta_{\mu\nu}$ is a special example of a metric tensor, while $\delta^\mu{}_\nu$ is an operator that

Figure 8.2. Bras and kets: kets are arrows, bras collections of parallel planes (or "measuring tapes"). Their lengths are undefined but one can multiply them with real numbers, and the quantity $u_\mu x^\mu$ has a clear meaning.

acts on kets—even though, when you write them out as matrices, they look the same.

Now we have the means to set up a one-to-one correspondence between bras and kets:

$$x^\mu \leftrightarrow x_\mu = g_{\mu\nu} x^\nu \quad \Leftrightarrow \quad u_\mu \leftrightarrow u^\mu = g^{\mu\nu} u_\nu . \tag{8.72}$$

We define the length squared of a vector x^μ as $x^\mu g_{\mu\nu} x^\nu$. If $g_{\mu\nu} = \delta_{\mu\nu}$ we recover the naive definition, used implicitly in section 6.1. Finally, let us consider eq. (6.7), which defined those special matrices that correspond to rotations, that is to say those matrices that transform the vectors in such a way that their lengths are preserved. It is not hard to rewrite this equation in such a way that they are allowed by our new conventions. The result is

$$R^\sigma_{\ \mu} R^\rho_{\ \nu} g_{\sigma\rho} = g_{\mu\nu} . \tag{8.73}$$

If the metric is the Kronecker delta, this equation says that the rotation is done by means of an orthogonal matrix. In this way we recover the full content of section 6.1.

There is a simple geometrical picture of kets and bras that may be helpful. Represent the kets as arrows pointing from the origin. If the ket x^μ is multiplied with (say) 2, the length of the arrow in the picture doubles. Represent the bras with measuring tapes through the origin, or more accurately as a set of parallel hyperplanes with constant spacing (level curves of a linearly rising function). The bras transform oppositely to the kets, so if the ket is multiplied by 2 the bra must be multiplied with $1/2$. The number $u_\mu x^\mu$ should stay unchanged under this transformation. If we interpret this number as the number of level

surfaces pierced by the arrow, we see that multiplying the ket with 2 means that the spacing between the level surfaces doubles—so the function giving rise to the level surfaces has been multiplied by 1/2. We have arrived at a consistent geometrical picture.

The true length of x^μ is not defined yet. You might think that it could be defined by choosing some special measuring tape, rotating this tape until it is aligned with the arrow, and afterwards compute $u_\mu x^\mu$. But this assumes that the measuring tape is unchanged by the rotation, and the catch is that "rotation" is defined as a transformation that preserves the lengths of all arrows, or the spacing of all measuring tapes. We are caught in a circular definition, and forced to introduce some further structure—a metric!—before we can define length.

The reason why we did not bother about these things in chapter 6 is that there was simply no need for them. Because the metric tensor was so simple—equal to the Kronecker delta—the one-to-one correspondence between bras and kets was so obvious that we simply identified them from the outset, without comment. But the more careful treatment here is useful for various generalisations. In particular it is useful to understand the Hamiltonian formulation of the equations of motion. There may even be a moral here. Forbidding certain things (in this case summing over repeated indices in any way we please) sometimes gives much more freedom.

8.6   The symplectic form

To begin with we assume that phase space is a vector space. More involved cases—such as the phase space of a rigid body—can wait. If we have a single vector x^μ, how can we assign a number to it? We know the answer. We introduce a metric tensor, and define

$$||x||^2 = x^\mu g_{\mu\nu} x^\nu \ . \tag{8.74}$$

The simple choice $g_{\mu\nu} = \delta_{\mu\nu}$ is especially interesting, but in principle any choice of $g_{\mu\nu}$ is allowed, as long as it is invertible and symmetric. Once we have a metric we can also assign a real number to any pair of vectors, namely the scalar product

$$y^\mu g_{\mu\nu} x^\nu = ||x|| \ ||y|| \cos\theta \ . \tag{8.75}$$

But geometrically it is obvious that there is another number that we can assign to the pair, namely the (oriented) area A that they span. Let us adapt our coordinates so that the two vectors have only two non-zero components each. Then A is given by a determinant

$$A = \left| \begin{array}{cc} x^1 & y^1 \\ x^2 & y^2 \end{array} \right| \ . \tag{8.76}$$
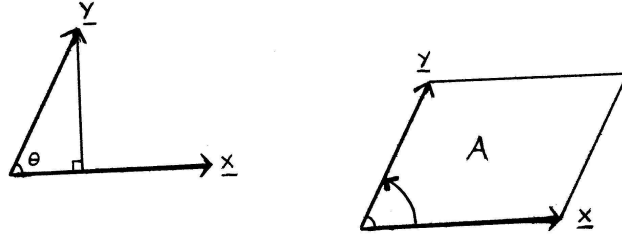
Figure 8.3. Two ways of associating a number to a pair of vectors: the angle they subtend, and the area they span.

We can write this as

$$A = y^\mu \omega_{\mu\nu} x^\nu \ , \tag{8.77}$$

where

$$\omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ . \tag{8.78}$$

Like the scalar product the area is a bilinear function of the vectors.

The next step is to "liberate" eq. (8.77) from its origins. *Any* anti-symmetric and invertible tensor can serve as a *symplectic form* on a vector space, just as any symmetric invertible tensor can serve as a metric. In other words eq. (8.77) will be taken to be a meaningful number associated to any pair of vectors, regardless of the dimension of the vector space, provided only that the symplectic form obeys

$$\omega_{\mu\nu} = -\omega_{\nu\mu} \ , \tag{8.79}$$

and provided that there exists an inverse matrix $\omega^{\mu\nu}$ so that

$$\omega^{\mu\sigma} \omega_{\sigma\nu} = \delta^\mu_{\ \nu} \ . \tag{8.80}$$

The latter equation is analogous to equation (8.70) for the metric. But it happens that an anti-symmetric $N \times N$ matrix has determinant zero if N is odd. If the determinant is zero its inverse cannot exist. Hence symplectic forms exist only on even dimensional vector spaces, while metrics exist on vector spaces of all dimensions. (This fact gave us Euler's theorem in section 6.1—the antisymmetric matrix defining an infinitesimal rotation in 3-space necessarily has a zero eigenvalue, and hence the rotation it generates has a fixed axis.)

Just as we may wish to define metrics on curved spaces, not only on vector spaces, so we may wish to define symplectic forms on more general phase

spaces. A symplectic form on an arbitrary phase space will be represented by an invertible and anti-symmetric tensor $\omega_{\mu\nu}$, but in principle its components may depend on the particular point of phase space where it sits. This cannot happen in a quite arbitrary fashion though, in fact the third and final requirement on a symplectic form is

$$\partial_\mu \omega_{\nu\sigma} + \partial_\sigma \omega_{\mu\nu} + \partial_\nu \omega_{\sigma\mu} = 0 \ . \tag{8.81}$$

We will mostly be interested in the case when the components of $\omega_{\mu\nu}$ are constant, and then this extra requirement is trivial. For now we only remark that the equation is recognisable as one of Maxwell's equations in electrodynamics. There it guarantees the existence of the vector potential. We will see that a similar object arises in Hamiltonian mechanics, once we come to section 8.4.

The most common situation is that the symplectic form is constant and "block diagonal", that is to say it has the form

$$\omega_{\mu\nu} = \begin{pmatrix} 0 & -1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & -1 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & & & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & -1 \\ 0 & 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix} \ . \tag{8.82}$$

In exercise 13 you will show that, on a vector space, one can always choose the coordinates in such a way that any symplectic form takes this form. In fact there is a stronger statement known as *Darboux' theorem*: On any symplectic space with topology $\mathbf{R}^{2n}$, coordinates can always be introduced so that the symplectic form takes this standard form. Although we do not need to go into it now, it is interesting to know that metric tensors behave in a completely different way in this respect.

Because the symplectic form has an inverse, it can be used to relate contravariant vectors with covariant ones in a unique manner, just as a metric can. The archetypical contravariant vector is the little arrow $\dot{z}^\mu$ that gives the time evolution of a system, while the archetypical covariant vector is the gradient of a function, say $\partial_\mu H$. Let us relate them:

$$\dot{z}^\mu = \omega^{\mu\nu} \partial_\nu H \ . \tag{8.83}$$

Let us further suppose that the symplectic form takes the standard form (8.82), and let us adapt the description of the coordinates according to

$$z^\mu = \begin{pmatrix} z^1 \\ z^2 \\ z^3 \\ z^4 \\ \vdots \\ z^{2n-1} \\ z^{2n} \end{pmatrix} = \begin{pmatrix} q_1 \\ p_1 \\ q_2 \\ p_2 \\ \vdots \\ q_n \\ p_n \end{pmatrix} . \tag{8.84}$$

Remembering that we agreed to be careless about the position of Latin indices we then find that

$$\dot{z}^\mu = \omega^{\mu\nu}\partial_\nu H \quad \Leftrightarrow \quad \begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} \\[2mm] \dot{p}_i = -\frac{\partial H}{\partial q_i} . \end{cases} \tag{8.85}$$

This is indeed Hamilton's equations of motion—hereby revealed as an ingredient of *symplectic geometry.*

Finally, the Poisson bracket appears very naturally. We define it as

$$\{A, B\} = \partial_\mu A \omega^{\mu\nu} \partial_\nu B . \tag{8.86}$$

It is easily seen that this agrees with the definition in section 8.2, provided we use the canonical coordinates q and p. It can also be proved, independently of any special coordinates, that the Poisson bracket so redefined obeys all the algebraic properties required of it. In fact anti-symmetry and linearity are obvious, while the Jacobi identity requires a bit of an effort. See exercise 14.

8.7 The symplectic one-form

Behind the symplectic two-form there stands the symplectic one-form. It is worth knowing, especially since the discussion will provide the rules behind the manipulations we made on section 8.4. We begin with a little discussion of vector fields. A contravariant *vector field* (with indices upstairs) is an assignment of a little arrow to each point in a space. These arrows can always be regarded as the tangent vectors of a set of curves—the flowlines of the vector field. Hence the general definition: "A dynamical system is a vector field on a manifold". This is a mathematical way of saying that we are dealing with equations of the form $\dot{z}^\mu = f^\mu(z)$, for some functions f^μ.

In vector spaces we thought of covariant vectors (with indices downstairs) as assigning a set of parallel planes, as shown in Fig. 8.2. So we would expect that a covariant vector field is a set of level surfaces of some function. But this is simply not true in general. It is true if and only if the covariant vector field is the gradient of some function, $u_\mu = \partial_\mu f$. And it can be shown that this is so if and only if

$$\partial_\mu u_\nu - \partial_\nu u_\mu = 0 \ . \tag{8.87}$$

If this condition fails the level planes that exist at each point do not fit together to form level surfaces extending over all space. Actually this phenomenon is well known from the theory of curve integrals. A covariant vector field is just what we need in the integrand of such an integral. Suppose that a curve γ is defined explicitly by the functions $z^\mu = z^\mu(t)$, where t is some parameter along the curve. Then

$$\int_\gamma dz^\mu u_\mu(z) = \int_{t_1}^{t_2} dt \frac{dz^\mu}{dt} u_\mu(z(t)) \ . \tag{8.88}$$

It is known that this integral is independent of the path and defines a function $f(z(t))$, if and only if condition (8.87) holds. But the curve integral is well defined regardless of whether this is true or not.

It makes sense: in a vector space a covariant vector is a map from the set of all arrows to the real numbers, and in general a covariant vector field is a map from the set of all flowlines (at all points) to the real numbers.

This leads to a useful piece of notation. If the coordinates on our space are denoted by z^μ and if u_μ is a covariant vector field, then we define the *differential one-form u* as

$$u = dz^\mu u_\mu \ . \tag{8.89}$$

That is, the coordinate differentials dz^μ are chosen as the basis in which we expand the covariant vector. This is the idea that lies behind the manipulations performed in section 8.4. In the bra-ket notation we would write

$$\langle u| = \langle dz^\mu|u_\mu \ . \tag{8.90}$$

The gradient of a function f defines a one-form of a special kind, namely

$$df = dx^\mu \partial_\mu f \ . \tag{8.91}$$

The notation is useful since it gives the correct behaviour of $u_\mu(x)$ under coordinate transformations. Let $x^{\mu\prime} = x^{\mu\prime}(x)$. Then

$$u = dx^\mu u_\mu = dx^{\mu\prime} \frac{\partial x^\nu}{\partial x^{\mu\prime}} u_\nu = dx^{\mu\prime} u_\mu' \quad \Rightarrow \quad u_\mu'(x') = \frac{\partial x^\nu}{\partial x^{\mu\prime}} u_\nu(x) \ . \tag{8.92}$$

Notice that everything happens automatically.

We brought this up because the *symplectic one-form* is lurking behind the symplectic *two-form* $\omega_{\mu\nu}$. To see why we come back to eq. (8.81), which was part of the definition of the latter. If the components of $\omega_{\mu\nu}$ are constants it obviously holds, but in general it is something that must be checked. The same equation occurs in electrodynamics when written in tensor form. It is this equation that implies that the vector potential exists:

$$\partial_\mu \omega_{\nu\sigma} + \partial_\sigma \omega_{\mu\nu} + \partial_\nu \omega_{\sigma\mu} = 0 \quad \Leftrightarrow \quad \omega_{\mu\nu} = \partial_\mu \theta_\nu - \partial_\nu \theta_\mu \ . \tag{8.93}$$

Here θ_μ is some covariant vector field, known in symplectic geometry as the symplectic one-form. (If you like, "one-form" is simply another name for "covariant vector field". The "if and only if" statement is true only if the topology is trivial.)

Although I have stated it without proof, this result is extremely important. When electrodynamics is written in tensor notation one finds that the electric and magnetic fields can be written as a two-form

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_i \\ E_i & \epsilon_{ijk}B_k \end{pmatrix} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \ . \tag{8.94}$$

Eq. (8.93) are part of Maxwell's equations; they guarantee the existence of the vector potential, which is a one-form. There is also a second tensor equation relating the two-form to the electric current, but this has no analogue in symplectic geometry.

What does the symplectic one-form look like if we use our standard coordinates (q_i, p_i) on phase space, so that the symplectic two-form takes the form (8.82)? There is no unique answer to this question since the previous discussion implies that if

$$\theta'_\mu = \theta_\mu + \partial_\mu \Lambda \tag{8.95}$$

then θ'_μ and θ_μ give rise to the same symplectic two-form. However, a possible choice of symplectic one-form is

$$\theta = p_i dq_i \ . \tag{8.96}$$

This serves as our standard choice. To make sure that you see how it works, consider a two dimensional phase space with

$$z^\mu = \begin{pmatrix} q \\ p \end{pmatrix} \ , \qquad \theta_\mu = \begin{pmatrix} p , & 0 \end{pmatrix} \ . \tag{8.97}$$

It is simple to verify that

$$\omega = \begin{pmatrix} 0 & \partial_q \theta_2 - \partial_p \theta_1 \\ \partial_p \theta_1 - \partial_q \theta_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \Rightarrow \quad \{q,p\} = 1 \ . \tag{8.98}$$

More generally, using eq. (8.84) we recover our standard symplectic two-form (8.82) in any dimension.

8.8   The sphere as a phase space

We end this chapter with an example of a phase space having an unusual topology, namely that of Euler's spinning top. At first sight the equations, (6.57), cannot be cast into Hamiltonian for the simple reason that phase space is three dimensional. This is a bit deceptive though. Recall that $J^2 = J_1^2 + J_2^2 + J_3^2$ is constant. If we keep J^2 fixed we are on a two dimensional sphere, and a Hamiltonian description becomes thinkable.

Let us therefore define angular coordinates θ, ϕ through

$$J_1 = J \cos\phi \sin\theta \;, \quad J_2 = J \sin\phi \sin\theta \;, \quad J_3 = J \cos\theta \;. \tag{8.99}$$

If we write out Euler's equations in terms of $J_i = I_i \Omega_i$ we find, after a calculation, that they become

$$\dot\theta = J \left(\frac{1}{I_1} - \frac{1}{I_2} \right) \cos\phi \sin\phi \sin\theta$$

$$\dot\phi = J \left(\frac{1}{I_1} - \frac{1}{I_3} \right) \cos^2\phi \cos\theta - J \left(\frac{1}{I_3} - \frac{1}{I_2} \right) \sin^2\phi \cos\theta \;. \tag{8.100}$$

Now stare at this. Then define a function on the sphere,

$$H(\theta,\phi) = \frac{J_1^2}{2I_1} + \frac{J_2^2}{2I_2} + \frac{J_3^2}{2I_3} \;. \tag{8.101}$$

We also define the symplectic form

$$\omega_{\mu\nu} = J \begin{pmatrix} 0 & \sin\theta \\ -\sin\theta & 0 \end{pmatrix} \quad \Leftrightarrow \quad \omega^{\mu\nu} = \frac{1}{J\sin\theta} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \;. \tag{8.102}$$

(The indexing here is such that $\omega_{\theta\phi} = J\sin\theta$.) Thus we have the Poisson bracket

$$\{\theta,\phi\} = -\frac{1}{J\sin\theta} \;. \tag{8.103}$$

A calculation now verifies that Euler's equations for fixed angular momentum squared take the Hamiltonian form

$$\dot\theta = \omega^{\theta\phi} \partial_\phi H \;, \qquad \dot\phi = \omega^{\phi\theta} \partial_\theta H \;. \tag{8.104}$$

Moreover it is easily checked that we have the Poisson bracket algebra

$$\{J_1, J_2\} = -J_3 \;, \quad \{J_2, J_3\} = -J_1 \;, \quad \{J_3, J_1\} = -J_2 \;. \tag{8.105}$$

This follows directly from the definition (8.86).

This is an example where the phase space is a sphere, and where no natural split of the coordinates (θ and ϕ) into a canonical pair (q, p) exists. But according to Darboux' theorem canonical coordinates must exist on any open region of the sphere. Indeed it is easy to check that

$$q = \cos\theta \ , \qquad p = \phi \ , \tag{8.106}$$

form a canonical pair. To be precise we find that

$$\{q, p\} = \{\cos\theta, \phi\} = \partial_\mu(\cos\theta)\omega^{\mu\nu}\partial_\nu\phi = \frac{1}{J} \ . \tag{8.107}$$

Moreover q and p are good coordinates on the sphere everywhere except at the poles. In these coordinates the area element on the sphere becomes

$$dA = J\sin\theta d\theta d\phi = Jdqdp \ . \tag{8.108}$$

Hence the canonical coordinates are suitable if you want a map of the sphere to display areas correctly. (By all means draw a picture to illustrate this result!)

Closer examination of this example reveals that generalization to any even-dimensional sphere is non-trivial. The point is that we cannot use any antisymmetric non-degenerate tensor as a symplectic form. According to the definition it must also solve eq. (8.81). In the two-dimensional case this is a trivial point, but as it turns out there is simply no solution for any higher dimensional sphere. Other, non-spherical, examples of phase spaces with non-trivial topologies do exist in higher dimensions.

$\diamond$ **Problem 8.1**   Derive the Hamiltonian corresponding to the Lagrangian for a charged particle in an external field, eq. (2.10).

$\diamond$ **Problem 8.2**   The Lagrangian for a relativistic particle is

$$L = -mc\sqrt{c^2 - \dot{x}^2} \ , \qquad \dot{x}^2 = \dot{x}_i\dot{x}_i \ . \tag{8.109}$$

Show that this reduces to the ordinary free particle when c is large compared to the velocity. Then derive the Hamiltonian for the relativistic particle.

$\diamond$ **Problem 8.3**   Replace Hamilton's equations with the seemingly equally natural

$$\dot{q} = \partial_q H \qquad \dot{p} = \partial_p H \ . \tag{8.110}$$

Is there an analogue of Liouville's theorem? If you choose H to be the harmonic oscillator Hamiltonian, what would the phase space flow look like?

$\diamond$ **Problem 8.4**   Consider Euler's equations for a spinning top. Write them in terms of J_i. There are two conserved quantities of the form $G = G(J_1, J_2, J_3)$ and $H = H(J_1, J_2, J_3)$. Prove that the time development of any function $F = F(J_1, J_2, J_3)$ is given by

$$\dot{F} = \epsilon^{ijk} \partial_i F \partial_j G \partial_k H \ , \tag{8.111}$$

where ϵ^{ijk} is totally anti-symmetric and the partial derivatives are with respect to J_i. Indices run from 1 to 3. For your information, this is called Nambu mechanics.

$\diamondsuit$ **Problem 8.5**   Evaluate the mutual Poisson brackets enjoyed by the three conserved charges (2.32) of the particle on the sphere, using polar coordinates for the calculation. Do the same for the three conserved charges of the particle on the hyperboloid, considered in problem 2.8.

$\diamondsuit$ **Problem 8.6**   Consider two Lagrangians related by a total derivative,

$$L'(q_i, \dot{q}_i) = L(q_i, \dot{q}_i) + \frac{d}{dt} \Lambda(q_i) = \frac{1}{2} \dot{q}_i \dot{q}_i - V(q_i) + \frac{d}{dt} \Lambda(q_i) \ , \tag{8.112}$$

for some function Λ. Derive the Hamiltonian formulations of these two Lagrangians, and show that there is a canonical transformation relating them. Show explicitly that Hamilton's equations are equivalent in the two cases. (Have you seen this Lagrangian before?)

$\diamondsuit$ **Problem 8.7**   Consider the Hamiltonian

$$H = \frac{1}{2m} \big(p_i - A_i(q) \big) \big(p_i - A_i(q) \big) \ . \tag{8.113}$$

Find the conditions on the 'vector potential' $A_i(q)$ ensuring that the transformation $Q_i = q_i$, $P_i = p_i - A_i(q)$ is canonical.

$\diamondsuit$ **Problem 8.8**   Transform from Cartesian coordinates (x, y) to polar coordinates (r, ϕ). Do the functions (x, r) coordinatize the plane? The upper half plane?

$\diamondsuit$ **Problem 8.9**   Show that the transformation

$$Q = \ln \left(\frac{\sin p}{q} \right) \qquad P = q \cot p \tag{8.114}$$

is canonical, and determine the generating functions F_1 and F_2.

$\diamondsuit$ **Problem 8.10**   Find the canonical transformation generated by

$$F_1 = kq^2 \cot Q \ , \tag{8.115}$$

where k is a constant.

$\diamondsuit$ **Problem 8.11**   Using a generating function $F_2 = F_2(q, P)$, extend Bohlin's transformation (3.12) to a canonical transformation. Can you see, using eq. (4.26), how this relates the oscillator Hamiltonian to the Kepler Hamiltonian? (It is allowed to add constants to Hamiltonians.)

$\diamondsuit$ **Problem 8.12**   Let the coordinates of a space transform according to $x^\mu \to A^\mu{}_\nu x^\nu$. Prove that this implies that the gradient of a function transforms like a covariant vector, $\partial_\mu f \to \partial_\mu f (A^{-1})^\nu{}_\mu$.

◇   **Problem 8.13**   Show that, given any metric $g_{\mu\nu}$ (with positive eigenvalues) on a vector space, one can always change the coordinates x^μ to new coordinates $X^\mu = A^\mu{}_\nu x^\nu$, in such a way that

$$x^\mu g_{\mu\nu} y^\nu = X^\mu G_{\mu\nu} Y^\nu \ , \tag{8.116}$$

and such that the new metric tensor $G_{\mu\nu}$ becomes equal to a Kronecker delta. Show that if we instead require

$$y^\mu \omega_{\mu\nu} x^\nu = Y^\mu \Omega_{\mu\nu} X^\nu \tag{8.117}$$

then the symplectic form in the new coordinates can always be made to assume the standard form (8.82).

◇   **Problem 8.14**   Use eq. (8.81) to prove the Jacobi identity (8.24) for an arbitrary symplectic form (i.e. with components that may be non-constant phase space functions).

9 Hamilton-Jacobi theory

Let us return to the motivating example of Fermat's Principle and Snell's Law. The wave theory of light offers a simple way to understand the angles that arise—staring at Fig. 9.1 should be enough to derive Snell's Law, once it is understood how the index of refraction is related to the velocity of propagation. The light rays arise in the *geometrical optics* limit of the wave theory. This raises the question whether analytical mechanics can be understood in the same way, as the geometrical optics limit of an underlying wave theory. Interestingly, already in the nineteenth century Hamilton and Jacobi understood how to do this—although their main motivation was to develop effective ways to solve the equations of motion, rather than to go beyond the classical theory.
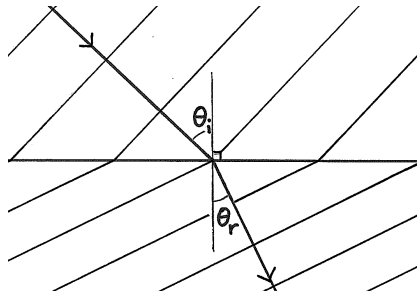
Figure 9.1. Snell's Law derived at a glance, from wave physics.

9.1 Geometrical optics

We begin with some quick (and non-rigourous) remarks on geometrical optics. This is an approximation to the wave theory of light in which diffraction is ignored—the one surviving notion from the wave theory is that of the *wave front*. But this notion is just what we need here.

We assume that we have an *inhomogeneous* and *anisotropic* optical medium, in which the velocity of light depends on both position and direction. At each point $\mathbf{q}_0$ in the medium we define an ellipsoid called the *indicatrix*, which describes the velocity vectors of light at the given point. See Fig. 9.2. To each
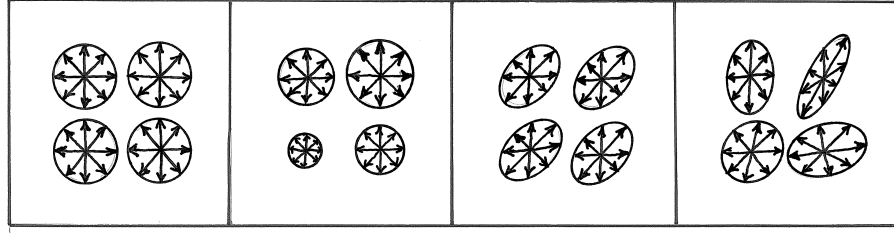
Figure 9.2. Indicatrices (formed from velocity vectors) in media that are, from left to right, homogeneous and isotropic, inhomogeneous and isostropic, homogeneous and anisotropic, inhomogenous and anisotropic.
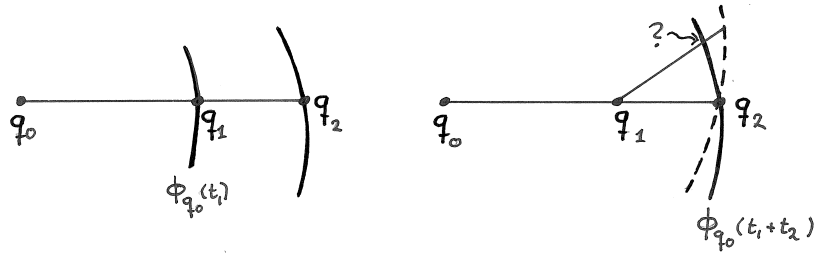

Figure 9.3. The proof of Huygens' principle: If the wave front from $\mathbf{q}_1$ (dashed) fails to be tangent to that from $\mathbf{q}_0$, how fast can light propagate from $\mathbf{q}_0$ to the point marked with a question mark?

point we associate a wave front $\Phi_{\mathbf{q}_0}(t)$, consisting of all the points that light emitted from $\mathbf{q}_0$ can reach in time t but not faster. Now consider the wave front at two successive times t_1 and $t_1 + t_2$. At time t_1 we have a wave front $\Phi_{\mathbf{q}_0}(t_1)$. It is a remarkable result due to Huygens that the wave front $\Phi_{\mathbf{q}_0}(t_1 + t_2)$ is the envelope of all the wave fronts $\Phi_{\mathbf{q}}(t_2)$ emerging from points $\mathbf{q}$ on $\Phi_{\mathbf{q}_0}(t_1)$, meaning that the former is at every point tangent to one of the latter. This is called *Huygens' principle*. The two steps in the proof are illustrated in Fig. 9.3. First we consider a point $\mathbf{q}_1$ on the front $\Phi_{\mathbf{q}_0}(t_1)$ and another point $\mathbf{q}_2$ on $\Phi_{\mathbf{q}_0}(t_1 + t_2)$. The latter must lie on $\Phi_{\mathbf{q}_1}(t_2)$, otherwise light can reach it from $\mathbf{q}_0$ in less time than $t_1 + t_2$. Next suppose that the wave front $\Phi_{\mathbf{q}_1}(t_2)$ fails to be tangent to $\Phi_{\mathbf{q}_0}(t_1 + t_2)$, in which case they must cross. But it then follows that there is a point on the latter which can be reached from $\mathbf{q}_1$ in less time than t_2, and hence from $\mathbf{q}_0$ in less time than $t_1 + t_2$, which is a contradiction.

Each individual wave front can be described as the level set of a function $S_{\mathbf{q}_0}(\mathbf{q})$, that is to say it consists of all points such that

$$S_{\mathbf{q}_0}(\mathbf{q}) = t \ . \tag{9.1}$$

The function is called the *optical length of the path* from $\mathbf{q}_0$ to $\mathbf{q}$, and is in fact equal to the least time it takes light to propagate from the source to the front.

Serious complications with these pictures can arise, because it can happen

(depending on the medium we are looking at) that light rays emerging from a point refocus and cross each other. We can ignore this if we agree to follow the wave front for a sufficiently short time only.

In Fig. 9.1 the light rays are orthogonal to the wave fronts, but this is so because we are considering two homogeneous and isotropic media. To characterize the motion of a general wave front Hamilton introduced the *vector of normal slowness*

$$p_i = \frac{\partial S}{\partial q_i} \ .$$

(9.2)

It is normal to the front. The front moves fast if the vector is short, and conversely, which explains its name. Now one can show that the vector of normal slowness will be normal to a plane that is tangential to the indicatrix at the point where it meets the vector tangent to the ray. To see this, consider Fig. 9.4, and an indicatrix at a point on the ray shifted an amount ϵ into the interior of the front. The wave front $\Phi_{\mathbf{q}_{t-\epsilon}}(\epsilon)$ differs by terms of order ϵ^2 from the indicatrix at $\mathbf{q}_{t-\epsilon}$. By Huygens' principle it is tangential to the front. Taking the limit $\epsilon \to 0$ the result follows.
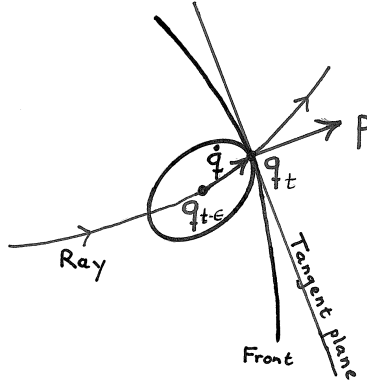
Figure 9.4. The relation between the vector $\dot{\mathbf{q}}$, tangent to a light ray, and the vector of normal slowness $\mathbf{p}$.

Let us see how the formulas work out. Choose a point $\mathbf{q}$ sitting on the wave front. Denote the velocity vectors there by $\dot{\mathbf{q}}$. Then the indicatrix is described by the equation

$$T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \dot{q}^i g_{ij} \dot{q}^j = 1 \ .$$

(9.3)

If the symmetric matrix g_{ij} equals δ_{ij} everywhere the medium is homogeneous and isotropic, but it may have a more complicated form, and it may depend on the point $\mathbf{q}$.

At the point where the velocity vector $\dot{\mathbf{q}}$ meets the ellipsoid the latter has a tangent plane with normal

$$p_i = \frac{\partial T}{\partial \dot{q}_i} = g_{ij}\dot{q}^j \; . \tag{9.4}$$

Unless the indicatrix is a circle the direction of the normal vector will differ from that of the light ray. Indeed—going back to chapter 7— we see that they are related by a Legendre transformation. The formalism of analytical mechanics is beginning to emerge, and we take leave of geometrical optics.

9.2 Hamilton's Principal Function

What function in analytical mechanics can serve the role of the optical length of the path? We have one very interesting function on the configuration space available, namely the function

$$S(q,t) = \int_{q_0(t_0)}^{q(t)} L(q,\dot{q},t)dt \; . \tag{9.5}$$

(We allow an explicit time dependence of the Lagrangian.) This is no longer to be regarded as the action functional $S[q(t)]$, which was a functional of all possible paths connecting two chosen endpoints. Instead we assume that the function $q(t)$ is a solution of the Euler-Lagrange equations such that it goes from q_0 at time t_0 to q at time t, and then we do the integral. The result is a function of q and t (also if the Lagrangian lacks an explicit time dependence). This function is called *Hamilton's Principal Function*. To see where we are heading, let us quote Hamilton himself: "Lagrange's function *states*, Mr. Hamilton's function would *solve* the problem."[1]

As a matter of fact we dealt with Hamilton's Principal Function in section 2.3, in the course of the derivation of Noether's theorem. In particular eq. (2.34) together with the definition of the canonical momenta implies that

$$\frac{\partial S}{\partial q_i} = \frac{\partial L}{\partial \dot{q}_i} \equiv p_i \; . \tag{9.6}$$

This is the partial derivative of S with time t kept fixed. The Lagrangian is analogous to the indicatrix in geometrical optics. We also wish to compute the partial derivative of S with respect to t, keeping q_i fixed. The easy way to do this is to observe that an expression for the total time derivative is known,

$$\frac{dS}{dt} = L \; . \tag{9.7}$$

On the other hand

[1] W. R. Hamilton, Report of the Fourth Meeting of the British Association for the Advancement of Science, Edinburgh 1834.

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial q_i}\dot{q}_i = \frac{\partial S}{\partial t} + p_i \dot{q}_i \; . \tag{9.8}$$

We can then read off that

$$\frac{\partial S}{\partial t} = L - p_i \dot{q}_i = -H \; , \tag{9.9}$$

where H is the canonical Hamiltonian function. We can summarize our calculation of the derivatives by writing down the one-form

$$dS = p_i dq_i - H dt \; . \tag{9.10}$$

We would like to have a way of finding Hamilton's Principal Function without having to solve the equations of motion first. Fortunately there is a way. Given the Hamiltonian function $H = H(q, p)$ we can replace the momenta by partial derivatives of S using eq. (9.6), and then rewrite eq. (9.9) as

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}\right) = 0 \; . \tag{9.11}$$

This famous partial differential equation goes under the name of the *Hamilton-Jacobi equation*. Hamilton's Principal function provides a solution that depends on $n+1$ arbitrary constants, namely the values assigned as initial values to the configuration space variables q_i at some arbitrarily chosen time t_0. The question is whether we can find such solutions directly from the Hamilton-Jacobi equation.

Suppose we can. If so we obtain a Principal function of the form

$$S = S(q_i, P_i, t) + \alpha \; . \tag{9.12}$$

The constant α is harmless. The n constants P_i on the other hand are interesting. We will regard them as being members of new canonical pairs (Q_i, P_i), connected by a canonical transformation to the original pairs. The generating function for this canonical transformation is going to be the Principal function itself. According to eq. (8.56) we obtain

$$Q_i = \frac{\partial S}{\partial P_i} \; . \tag{9.13}$$

We now take the time derivative of the new coordinates. Using the Hamilton-Jacobi equation we obtain

$$\dot{Q} = \frac{\partial}{\partial P}\left(\frac{\partial S}{\partial t} + \frac{\partial S}{\partial q}\dot{q}\right) = -\frac{\partial}{\partial P}\left(H\left(q, \frac{\partial S}{\partial q}\right) - \frac{\partial S}{\partial q}\dot{q}\right) \; . \tag{9.14}$$

Now the dependence of the new momenta P is purely in the function S, so we continue the calculation and obtain

$$\dot{Q} = -\frac{\partial H}{\partial \left(\frac{\partial S}{\partial q}\right)} \frac{\partial}{\partial P}\left(\frac{\partial S}{\partial q}\right) + \frac{\partial^2 S}{\partial P \partial q}\dot{q} = -\frac{\partial^2 S}{\partial P \partial q}\left(\frac{\partial H}{\partial \left(\frac{\partial S}{\partial q}\right)} - \dot{q}\right) \ . \qquad (9.15)$$

Finally we use Hamilton's equations of motion to conclude that

$$\dot{Q}_i = 0 \ . \qquad (9.16)$$

Since we knew from the start that $\dot{P}_i = 0$ we have found a (time dependent) canonical transformation to new canonical coordinates that are independent of time. The transformed Hamiltonian vanishes. In effect we are now using the initial values of the canonical coordinates q and p as our new canonical coordinates Q and P.

We have, however, not explained how to find solutions of the Hamilton-Jacobi equation taking the required form. In fact, if we were faced by the partial differential equation to start with, we could find it very helpful to observe that we can obtain solutions by integrating Hamilton's equations of motion, thus turning the problem of solving a partial differential equation into the presumably easier task of integrating a set of ordinary differential equation. Still, as Hamilton said, the reduction of the most complex dynamical problem to the study of one characteristic function may result in "intellectual pleasure".

If the Hamiltonian has no explicit time dependence the equation separates. There will exist solutions of the form $S(q, t) = S_1(q) + S_2(t)$. In fact there will be solutions of the form

$$S(q_i, P_i, t) = W(q_i, P_i) - ct \ . \qquad (9.17)$$

The constant $c = c(P_i)$ is some function of the constant momenta, and can be chosen largely at will. The function W is known as *Hamilton's characteristic function*. It obeys the equation

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = c \ . \qquad (9.18)$$

In geometrical optics this would be called an *eikonal* equation. From the point of view of solving the Hamilton-Jacobi equation it is at least a step in the right direction.

9.3 Soluble examples

In order to see concretely how the Hamilton-Jacobi equation works we turn to a simple soluble example: the harmonic oscillator. We move directly to equation for the characteristic function, and choose the constant $c = P$. Thus

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{kq^2}{2} = P \ . \tag{9.19}$$

It follows that

$$W = \sqrt{mk}\int \sqrt{\frac{2P}{k} - q^2}\ dq \ . \tag{9.20}$$

We can use this as it stands in order to calculate

$$Q = \frac{\partial S}{\partial P} = \frac{\partial W}{\partial P} - t = \sqrt{\frac{m}{k}}\int \frac{dq}{\sqrt{\frac{2P}{k} - q^2}} - t \ . \tag{9.21}$$

Performing the integral, and inverting the resulting function, we obtain

$$q = \sqrt{\frac{2P}{k}}\cos\sqrt{\frac{k}{m}}(Q + t) \ . \tag{9.22}$$

Clearly this is the solution to the dynamical problem posed by the harmonic oscillator.

A slightly more involved case is that of the gravitational two-body problem. We know the form of the Hamiltonian, and the equation to be solved is

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial W}{\partial \phi}\right)^2\right] - \frac{k}{r} = E \ . \tag{9.23}$$

The constant on the right hand side is identified as the energy. This equation is fully separable, that is to say that there exists solutions of the form

$$W(r,\theta,\phi) = W_r(r) + W_\theta(\theta) + W_\phi(\phi) \ . \tag{9.24}$$

Inserting this Ansatz into the equation we realize that the only dependence on the coordinate ϕ can come through the third term within brackets. Hence

$$\frac{\partial W}{\partial \phi} = \alpha_\phi \ , \tag{9.25}$$

where α_ϕ is a constant. Inserting this expression into the equation we find that the only dependence on θ comes from the last two terms in the equations, and we conclude that

$$\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{\alpha_\phi^2}{\sin^2\theta} = \alpha_\theta^2 \ , \tag{9.26}$$

where the right hand side is another constant. Making use of this expression too in eq. (9.23) leads to

$$\left(\frac{\partial W}{\partial r}\right)^2 + \frac{\alpha_\theta^2}{r^2} = 2m\left(E + \frac{k}{r}\right) \ . \tag{9.27}$$

Hence we can get a solution with the desired number of undetermined constants by solving three ordinary differential equations of the first order. The full solution of the two-body problem can be recovered from this. In fact all soluble examples arise because the partial differential equation separates into a set of ordinary differential equations depending on only one variable each.

The degree of sophistication that we introduced to solve the harmonic oscillator may appear rather excessive. However, there do exist a number of interesting problems where this way of doing it is the only feasible way. Hamilton's own judgment on his reinterpretation of the expression for the action S was this: "even if it should be thought that no practical facility is gained, yet an intellectual pleasure may result from the reduction of the most complex, probably, of all researches respecting the forces and motions of body, to the study of one characteristic function, the unfolding of one central relation." [2] In retrospect, perhaps the most interesting point to make is that his reformulation of classical mechanics directly suggests how to go beyond the classical theory to quantum mechanics—although it took some time, and considerable pressure from experiments, to take this step. Now try to invent a similarly deep reformulation of quantum mechanics!

◇ **Problem 9.1**
In exercise 2.13 you calculated Hamilton's Principal function for a free particle and for a harmonic oscillator in terms of the standard configuration space variables. Check that these expressions do solve the Hamilton-Jacobi equation.

◇ **Problem 9.2**
Can you solve the Hamilton-Jacobi equation for the gravitational two-body problem if you use Cartesian coordinates?

---

[2] W. R. Hamilton, *On a General Method in Dynamics; by which the Study of the Motions of all free systems of attracting or repelling Points is reduced to the Search and Differentiation of one central Relation, or characteristic Function*, Phil. Trans. R. Soc. Lond. **124** (1834) 247.

10 Integrable and chaotic motion

We have arrived at the Hamiltonian form of the equations of motion, as well as the most general transformations in phase space leaving that form invariant. Can these results be used to describe the properties of the solutions with any generality? Partial answers to this question can in fact be given, and this is the topic of this chapter—which will, however, end in chaos.

10.1 Can chaos occur?

Although we know that the solution to a generic dynamical system—a system of first order coupled ordinary differential equations—exists and is unique, given the initial data, it is not clear that we can expect to actually find these solutions with a reasonable expense of effort. This depends to a large extent on the dimension of phase space. If the dimension is two it is clear that the solutions cannot look too bad. As long as the system is autonomous—i.e. when the right hand side of eq. (1.39) does not depend explicitly on time— the key observation is that the flow lines in phase space never cross. If we fill a two dimensional plane (or some other two dimensional surface) with non-intersecting flow lines we necessarily get a fairly orderly pattern, and hence the solutions of the equations of motion will be orderly too. For a Hamiltonian system, with a Hamiltonian of the form

$$H = \frac{p^2}{2} + V(q) \ , \tag{10.1}$$

we obtain a differential equation that determines the shape of the flowlines:

$$\begin{cases} \dot{q} = p \\ \dot{p} = -\partial_q V \end{cases} \quad \Rightarrow \quad \frac{dp}{dq} = -\frac{\partial_q V}{p} \ . \tag{10.2}$$

The solution for the curves $p = p(q)$ is found to be

$$p = \pm\sqrt{2(E - V(q))} \ , \tag{10.3}$$

where the energy E is an integration constant that labels the individual flow lines. And this is almost the end of the story.

Note that for non-autonomous systems the evolution equations depend explicitly on t, as in eq. (1.38). This includes systems that are driven by some time-dependent external force. For them it is no longer true that the phase space flow lines are non-intersecting. In this case we expect that trouble will arise already in two dimensional phase spaces.

If the dimension of phase space is three things can become much more complex. Once the flow lines can move in a third dimension there is absolutely no guarantee that they form an orderly pattern. It is instructive to compare two three dimensional systems that we have encountered, namely the Lorenz and Euler equations. In the Euler equations (6.57) the phase space coordinates can be taken to be J_1, J_2, J_3. We can solve the equations because any solution $J_i(t)$ is a curve that lies on two well behaved surfaces

$$J_1^2 + J_2^2 + J_3^2 = \text{constant} , \qquad \frac{J_1^2}{I_1} + \frac{J_2^2}{I_2} + \frac{J_3^2}{I_3} = 2E . \qquad (10.4)$$

Hence any solution curve is the intersection of a sphere and an ellipsoid, and such a curve is not difficult to describe, nor is the set of all such curves hard to understand. The lesson is that the solutions of the Euler equations are easy to describe precisely because there exist two constants of the motion, each describing a well behaved surface in phase space. With the Lorenz equations (1.45) things are very different: in this case there simply does not exist any well behaved constants of the motion, and hence there is no reason to believe that the set of all solutions can be described in any easy and comprehensive manner. And indeed the solutions are chaotic, in particular they turn out to be chaotic in the sense that the smallest change in the initial data will change the long time behaviour of the solution in a dramatic fashion. The point here is simply that one expects this to happen if the dimension of phase space is three or more—unless there are special reasons to think otherwise.

Still, the Lorenz equations are not Hamiltonian. They cannot be, because their phase space has odd dimension. Now Hamiltonian phase space flows have some rather special properties, such as that uncovered by Liouville's theorem—they are like the flow of an incompressible fluid, and certain special kinds of fixed points cannot occur. Hence there is still some hope that the long time behaviour of Hamiltonian systems may be significantly simpler than that of a typical dynamical system. This is what we will look into next. (Eventually our hope of understanding the detailed behaviour of the general solution of a general Hamiltonian system will be completely dashed, but there will be many compensations.)

To study the kind of phase space flows that can occur we must be able to picture them effectively. If the space to which the flow is confined is three dimensional—either because the phase space has three dimensions or because the flow is confined to a three dimensional energy surface in a four dimensional phase space—the *Poincaré section* solves the problem for us. The idea is to
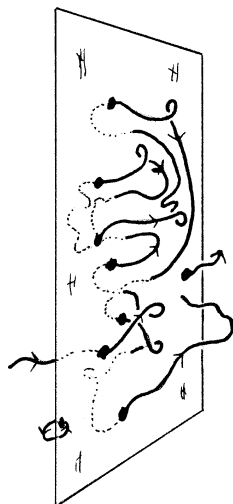
Figure 10.1. A Poincaré section displaying an orbit in three dimensions as a set of dots on a plane; a periodic orbit giving rise to just one dot occurs in the corner of the plane.

study a two dimensional cross section of the three dimensional space, through which the orbits pass in some definite direction. Each time the orbit passes out through the section one marks the corresponding point with a dot. If the orbit is periodic, the number of dots one obtains is finite. If the orbit is highly irregular, the pattern of dots will be highly irregular too.

10.2   Integrable systems

The harmonic oscillator is a very simple example of a soluble system. We will now discuss it using the whole apparatus of canonical transformations, in the hope that this will lead to a general strategy for how to deal with more complicated examples. The phase space flow takes place on circles surrounding a single fixed point, and the idea is to introduce new canonical coordinates (Q, P) such that P labels the flow line we are on, while Q tells us where we are on the given flow line. We insist that the transformation to these coordinates should be canonical because we want to make maximum use of the special structure of Hamiltonian dynamics.

The Hamiltonian of the harmonic oscillator is

$$H = \frac{1}{2}p^2 + \frac{\omega^2}{2}q^2 \ , \tag{10.5}$$

and the explicit solution of Hamilton's equations is given by

$$q = A\sin\left(\omega t + \delta\right) \qquad p = \omega A\cos\left(\omega t + \delta\right) \ , \tag{10.6}$$

where the amplitude A is given in terms of the energy E by

$$E = \frac{\omega^2}{2} A^2 \quad \Leftrightarrow \quad A = \frac{\sqrt{2E}}{\omega} \; . \tag{10.7}$$

We define the transformation to new phase space coordinates (Q, P) by

$$q = \sqrt{\frac{2P}{\omega}} \sin Q \qquad p = \sqrt{2\omega P} \cos Q \; . \tag{10.8}$$

It follows that

$$E = H = \omega P \; , \tag{10.9}$$

so indeed—up to a constant factor—one of the new coordinates is just the energy, and we already know that the energy serves to label the flow lines in phase space. Moreover, using the tricks introduced in section 8.2, a calculation verifies that

$$\{q, p\} = \frac{\partial q}{\partial Q} \frac{\partial p}{\partial P} - \frac{\partial q}{\partial P} \frac{\partial p}{\partial Q} = \cdots = 1 = \{Q, P\} \; . \tag{10.10}$$

The transformation is indeed canonical. After the transformation Hamilton's equations take the simple form

$$\dot{Q} = \{Q, H\} = \omega \qquad \dot{P} = \{P, H\} = 0 \; . \tag{10.11}$$

In these coordinates the solution is trivial.

This simple example suggests a general strategy for solving Hamilton's equations of motion for n degrees of freedom. In the first step we must find n constants of the motion I_i, such that

$$\{I_i, I_j\} = 0 \; . \tag{10.12}$$

The idea is that each of these constants of the motion will serve as a member of a canonical pair, whose other member will be called θ_i. In the second step we must devise a canonical transformation

$$q_i = q_i(\theta, I) \qquad p_i = p_i(\theta, I) \; , \tag{10.13}$$

such that the Hamiltonian has no dependence on the θ,

$$H = H(I) \; . \tag{10.14}$$

If we succeed in this Hamilton's equations will take the soluble form

$$\dot{\theta}_i = \partial_{I_i} H(I) \qquad \dot{I}_i = -\partial_{\theta_i} H(I) = 0 \; . \tag{10.15}$$

The new canonical variables I_i and θ_i are called *action-angle variables*. If such

variables exist the Hamiltonian system is said to be *integrable*. Since every $\dot{\theta}_i$ is constant the motion is associated to n constant frequencies ω_i.

In an integrable system every trajectory in the $2n$ dimensional phase space is confined to an n dimensional submanifold, defined by the n conditions $I_i =$ constant. We also have available n vector fields

$$I_i^\mu = \omega^{\mu\nu}\partial_\nu I_i \ . \tag{10.16}$$

The normal vector of any hypersurface defined by the equation $I_i =$ constant is

$$n_\mu^i = \partial_\mu I_i \ , \tag{10.17}$$

and it is clear that

$$\{I_i, I_j\} = 0 \quad \Leftrightarrow \quad n_\mu^i I_j^\mu = 0 \ . \tag{10.18}$$

Thus these vector fields are tangential to all the n hypersurfaces, and therefore tangential to the n-dimensional submanifold. Moreover they must be everywhere non-vanishing, because

$$\partial_\mu\theta_1 I_1^\mu = \partial_\mu\theta_1 \omega^{\mu\nu}\partial_\nu I_1 = \{\theta_1, I_1\} = 1 \neq 0 \tag{10.19}$$

and so on for all the n vector fields. We have found n everywhere non-vanishing vector fields pointing along the surface of the n dimensional submanifold defined by the n equations $I_i = 0$. Let us assume for simplicity that these submanifolds are closed and bounded. Then our conclusion is rather remarkable, because very few closed and bounded manifolds admit even a single everywhere non-vanishing vector field. (The circle does, but the sphere does not—you cannot comb a sphere.) Closer inspection shows that eq. (10.12) means that an additional technical condition on these vector fields is obeyed, namely that the vector fields commute. Without bothering too much about what this means, we can then rely on the following mathematical theorem:

The only closed and bounded n dimensional manifold admitting n everywhere non-vanishing commuting vector fields is the n dimensional torus.

A one dimensional torus is a circle, and a two dimensional torus is an ordinary torus, or more abstractly it is like a square with periodic boundary conditions. A three dimensional torus is like a cube with periodic boundary conditions, and so on. And the conclusion is that in an integrable Hamiltonian system all trajectories are confined to tori with half the dimension of the entire phase space.

The motion on a torus depends on the n frequencies that characterize the motion on the given torus. The trajectories will be open or closed depending on whether the frequencies are rationally related (as in the Kepler problem), or not (as in the case of a Lissajous figure that never repeats, see section 1.3).

Meanwhile, our emphasis on conserved quantities is now understandable. We also see why the three components of the conserved angular momentum vector really contributed only to the decoupling of two of the degrees of freedom in the central force two body problem. What we need for the general strategy are constants of the motion that "Poisson commute", as in eq. (10.12). From the angular momentum vector we can construct two such constants of the motion, namely (say) L_3 and L^2. We know that

$$\{L_i, L_j\} = \epsilon_{ijk} L_k \quad \Rightarrow \quad \{L_3, L^2\} = 0 \ . \tag{10.20}$$

Therefore L_3 and L^2 can serve as action variables. Clearly this is reminiscent of quantum mechanics, with its emphasis on commuting operators.

10.3 Canonical perturbation theory

The canonical transformation to action-angle variables will be difficult to find in general, and we will have to resort to *canonical perturbation theory* to construct it order by order in some parameter. There is a well developed technology for doing this. In 1969 it took us to the Moon.

The first step then is to split the Hamiltonian function into two parts,

$$H(\theta, I) = H_0(I) + \epsilon H_{\text{int}}(\theta, I) \ , \tag{10.21}$$

where the split is defined in such a way that the equations of motion coming from H_0 alone are integrable. It is assumed that ϵ is some small parameter, and the full problem is to be solved as a power series expansion in ϵ.

Already in two dimensions we can see some problems with this idea. Consider the case of a pendulum, with the Hamiltonian

$$H = \frac{1}{2} p^2 - \frac{\alpha^2}{2} \cos q \ . \tag{10.22}$$

The problem is not the non-linearity of the equations of motion as such. The problem is that, as discussed in section 1.4, phase space splits into three regions separated by separatrices. The two separatrices, and the hyperbolic fixed point, all have the same energy E, so the conserved energy will not do as a coordinate uniquely labelling the orbits. A related problem is that motion along a separatrix is not periodic; formally it corresponds to zero frequency, because it takes an infinite amount of time to traverse it.

When the energy is large the effect of gravity is small (the pendulum is rotating almost freely). Thus we can choose

$$H_0 = \frac{1}{2} p^2 \ , \tag{10.23}$$

and treat α as a small parameter. For small oscillations on the other hand we can set
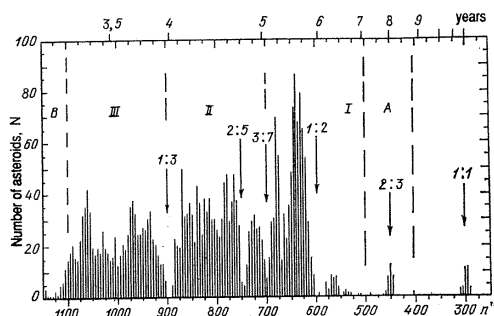
Figure 10.2. The Kirkwood gaps. The asteroids clustered around 1 : 1 are not a contradiction—they are the Greek and Trojan asteroids at the Lagrange points.

$$H_0 = \frac{1}{2}p^2 + \frac{\alpha^2}{2}q^2 \ . \tag{10.24}$$

The unperturbed system behaves as a harmonic oscillator. Supposing that we obtain the solution as two different power series based on these two different ways of splitting the Hamiltonian, how can we "join" them to obtain the behaviour at the separatrix? Actually what will happen is that the question of the convergence of the series will raise its ugly head, and the complete solution for the pendulum will escape us.

In a two dimensional phase space there is trouble only close to the separatrix, where the frequency vanishes. What can go wrong in higher dimensions? In four dimensions it turns out that there will be trouble in the neighbourhood of those tori where the frequencies obey

$$n_1\omega_1 + n_2\omega_2 = 0 \tag{10.25}$$

for some integers n_1, n_2. Moreover the trouble is most severe if these integers are small.

So are almost all Hamiltonian systems integrable? No. But if the Hamiltonian is of the form (10.21), with an ϵ that is not too large, then large regions of phase space will still be filled with tori. In between there will be regions where the tori have been destroyed and the motion is chaotic. The chaotic regions grow in size with ϵ, and the tori that disappear first are those tori for which the unperturbed motion obeys eq. (10.25) for small integers n_1, n_2. The full story here is known as the *KAM theorem*, for Kolmogorov, Arnold, and Moser.

A beautiful example of the KAM theorem in action is provided by the distribution of asteroids inside Jupiter's orbit. Presumably the asteroid belt was originally created in such a way that the number of asteroids as a function of their angular frequencies could be approximated by a fairly smooth function. Then, as time goes on, an asteroid whose trajectory was on a torus in

phase space will remain there, while an asteroid on a chaotic orbit behaves differently—it may, for instance, suffer sudden changes in eccentricity and crash into a planet. Let ω_1 be the frequency of the asteroid's unperturbed motion, and ω_2 that of Jupiter's. The theory then suggests that there will be "gaps" in the asteroid distribution, so that asteroids with an ω_1 obeying eq. (10.25) for small values of n_1 and n_2 are missing. Observation bears this out, and the gaps are known as the Kirkwood gaps for their discoverer. An exception is that the number of asteroids with $\omega_1 = \omega_2$ is particularly large—but these are not an exception to the KAM theorem, rather they are the Trojan asteroids that we discussed in section 4.6.
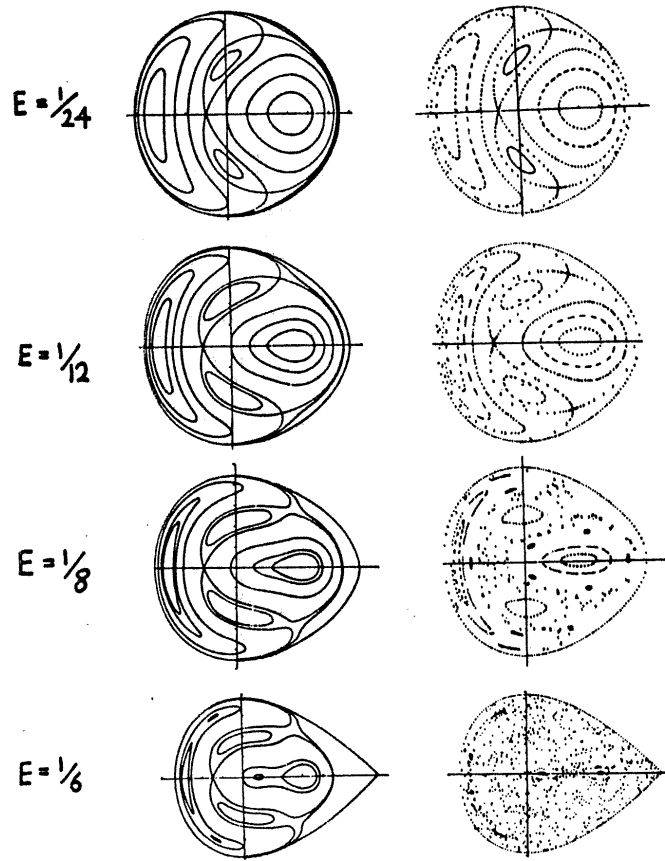
Figure 10.3. Poincaré sections, in the $(y, \dot{y})$ plane, for the Hénon-Heiles Hamiltonian. To the left you see the prediction from eighth order perturbation theory, to the right numerical results. The chaotic dots in the lower two figures come from a single orbit.

Another beautiful illustration is provided by the Hénon-Heiles Hamiltonian

$$H = \frac{1}{2}(p_x^2 + p_y^2) + V(x,y) = \frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{2}x^2 + \frac{1}{2}y^2 + x^2 y - \frac{y^3}{3} \ . \quad (10.26)$$

This toy model was originally inspired by the study of stars moving in the gravitational field of a galaxy. The potential is not bounded from below, but there is a potential well near the origin. Note that

$$V(x,y) - \frac{1}{6} = -\frac{1}{3}\left(y + \frac{1}{2}\right)(y - 1 - \sqrt{3}x)(y - 1 + \sqrt{3}x) \ . \quad (10.27)$$

Setting this to zero defines three straight lines that bound a triangular well. Provided the energy does not exceed $E = 1/6$ motion can be confined inside this triangle, and the energy surface is bounded. The influence of the cubic terms in the potential becomes more pronounced as the energy goes up.

To picture the dynamics it is convenient to study a Poincaré section based on some two dimensional cross section of phase space, say the plane spanned by the coordinates y and p_y. When studying a trajectory (presumably generated by a computer) one makes a dot on the two dimensional plane whenever the trajectory passes through it. If the trajectory is confined to a torus in phase space one should then see these dots lining up along some one dimensional curve in the resulting picture. If this does not happen one concludes that the tori—which would be present if the cubic terms in the Hamiltonian could be ignored—have been destroyed by the perturbation.

What Hénon and Heiles found illustrates the KAM theorem quite well. For $E = 1/24$ and $E = 1/12$ the actual trajectories do give rise to closed curves on the Poincaré section, and these curves lie exactly where they should lie according to an eighth order canonical perturbation theory calculation performed by Gustavson. Above $E = 1/9$, when the importance of the cubic terms is larger, one finds orbits that definitely do not lie on tori since they fill large parts of the picture with chaotic dots. But there will also be orbits that do lie on tori, so the picture becomes that of a mixture of order and chaos.

10.4   Stability of the Solar System

Since Newton many of these developments were driven by the wish to know whether the Solar System is stable in the long run. Its Hamiltonian includes eight planets and the Sun, mutually attracting each other according to Newton's law of gravity, as well as numerous other bodies whose masses were not well known to the pioneers. A brief account of what has been learned seems an appropriate way to end this story.

To lowest order we simply ignore the attractions between planets, and the system is integrable as we have seen. However, the planets in fact do not move exactly on Kepler ellipses. Indeed it was realised early on—by Halley, using observational data gathered in the sixteenth and seventeenth centuries—that Jupiter is accelerating and Saturn is decelerating. The effect is not small. If

it were extrapolated linearly one would conclude that these two planets were at the same distance from the Sun six million years ago, raising the issue of a possible collision between them. The problem arises because the ratio of the periods of these two planets is close to 2:5. If one tries to include the interplanetary forces in perturbation theory the question becomes whether it is a periodic perturbation, or whether it is proportional to time, in which case it is said to be a *secular* perturbation. Laplace answered this question to lowest order in perturbation theory: the observed variations in the orbits of Jupiter and Saturn are periodic, with a period of about 900 years. Thus they do not threaten the long term stability of the solar system. Moreover Laplace noted that his result is consistent with observations going back to Ptolemy and the Chaldeans.

In the nineteenth century it was realised that higher order perturbation theory leads to terms containing expressions such as $n_1\omega_1 + n_2\omega_2$ in their denominators. This means that higher order terms can become very large when the frequencies involved are in resonance. Eventually Poincaré dashed all hopes that the series will converge for arbitrary intitial data. The KAM theorem on the other hand clearly suggests the possibility that planets nevertheless lie on regular orbits, since there will be regions in phase space with this behaviour.

In the twentieth century it became clear that the question is relevant only over a timescale of about 10^{10} years, since this is the expected lifetime of the Sun. At the end of the twentieth century extensive calculations were made on computers especially designed for the purpose. (Such a computer was known as a "digital orrery"—orreries were first constructed using cogwheels, but the clockwork metaphor of mechanics has by now been replaced by another.) What one finds is that trajectories arising from near-by initial data may diverge exponentially in time, with an *e*-folding time of the order of $5 \cdot 10^6$ years. The conclusion therefore seems to be that accurate prediction in the form of ephemeridis tables is possible over timescales of 10^7 years, but not possible— at least not for the inner planets—over 10^8 years. On the other hand the qualitative properties of the Solar System do seem to be stable over much longer time scales.

Qualitative questions one can ask concern, for instance, the bounds within the eccentricities of the orbits are likely to vary. If the eccentricity of the orbit of Mars reaches 0.3 and that of the Earth reaches 0.1 a collision might ensue. The most recent results are not altogether reassuring.[1] Over a timescale of 10^{10} years variations in the eccentricities can be ignored as far as the four major planets are concerned, but there is a potential problem with Mercury. The Newtonian theory gives a 60 % probability that its eccentricity will grow to such an extent that it causes a dramatic rearrangement of the inner Solar System sometime during the next $5 \cdot 10^9$ years. Mercury may collide with Venus, and Mars may be ejected from the system. Fortunately this probability drops to about 1 % if corrections from General Relativity are taken into account, because the increase in the perihelion precession of Mercury turns out to be

---

[1] J. Laskar, *Is the Solar System Stable?*, Poincaré Séminar, Paris 2010.

helpful. There is no room for the addition of a fifth inner planet—if there ever was one it was ejected from the system or disappeared in a collision (perhaps giving rise to the Moon). The Solar System seems to be at most marginally stable over this timescale.

The Solar System would not be stable if the mass of its Sun were significantly less than its actual mass. Hence the Solar System as we know it is unlikely to survive the post-main-sequence mass loss of the Sun, which is expected to occur in about $5 \cdot 10^9$ years. And this is the end of the story.

$\diamondsuit$ **Problem 10.1**   Consider the Hamiltonian (10.1), and let $V(q)$ be a fourth order polynomial. Draw the phase space flow for all qualitatively different choices of this polynomial. Pay special attention to the transition cases, when a maximum and a minimum "merge" to produce an inflection point on the graph of $V(q)$.

$\diamondsuit$ **Problem 10.2**   Solve the evolution equation

$$\dot{x} = x + \epsilon x^2 \tag{10.28}$$

as a series expansion in ϵ, to second order in ϵ. Then solve the equation exactly, and compare the results. As initial condition, set $x(0) = k$.

$\diamondsuit$ **Problem 10.3**   For the purposes of canonical perturbation theory it is convenient to write the Hénon-Heiles Hamiltonian as

$$H = \frac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2) + \lambda(x^2 y - \frac{y^3}{3}) \ . \tag{10.29}$$

Explain why small coupling constants λ correspond to small values of the total energy in eq. (10.26). What is the nature of the energy surface when λ is small? When it is large?

$\diamondsuit$ **Problem 10.4**   Once we know that integrable motion takes place on tori on phase space we become interested in doubly periodic functions. Show that the elliptic function $\theta = \theta(t)$ which solves the equation for the pendulum has this property. Do this indirectly by showing that the equations make sense also if we make t purely imaginary. Then let the physics tell you that if t is taken to be a complex variable $\theta(t)$ must be periodic under purely real and purely imaginary shifts in t.

$\diamondsuit$ **Problem 10.5**   The orbital period of Jupiter is 11.86 years, that of Saturn 29.46 years. There exists something called the continued fraction expansion, which allows you to approximate any real number with rational numbers of progressively increasing denominators. Look this up. Then try to see why Laplace found a periodicity close to 900 years for the Sun-Jupiter-Saturn system.

Appendix 1 Books

I recommend the following books:

- V. I. Arnold: Mathematical Methods of Classical Mechanics, Springer 1978. Written by the person who understood the subject best.
- C. Lanczos: The Variational Principles of Mechanics, Dover 1986. A beautiful book with a philosophical slant.
- L. D. Landau and E. M. Lifshitz: Mechanics; Pergamon 1960. Traditional, and perfectly clear.
- I. Percival and D. Richards: Introduction to Dynamics, Cambridge 1982. Dynamical systems point of view; selective coverage.
- J. R. Taylor: Classical Mechanics, University Science Books, 2005. Helpful as an introduction at a somewhat lower level.

You may also enjoy:

- Ernst Mach: Die Mechanik in ihrer Entwicklung, many editions (avoid the ninth!). For the prehistory of mechanics.
- Max Born, Cause, Purpose and Economy in Natural Laws, in: Physics in My Generation, Springer 1969. The text of a lecture on variational principles.
- A. Dahan Dalmedico et al.: Chaos et déterminisme, Éditions du Seuil 1992. An interesting collection of articles about chaos.
- Julian Barbour: The End of Time, Weidenfeld and Nicolson 1999. A stimulating account of philosophical issues raised by Newton's laws.
- H. Scott Dumas: The KAM Story, World Scientific 2014. A fantastically well written account of the mathematical side of the subject.

Index