Classical Mechanics

LECTURE 9

5 Hamiltonian mechanics

We now leave behind the world of pendulums, springs, pulleys and tops, and develop the Hamiltonian formulation of classical mechanics. This is equivalent to the Lagrangian formulation but provides a different perspective, giving yet more insight into the structure of the theory. In particular the geometry underlying Hamiltonian mechanics is distinct from that in Lagrangian mechanics. Mathematically this leads on to the more abstract 20$^{\text{th}}$ century subjects of (for example) symplectic geometry and integrable systems, both still very active areas of research. On the other hand, from a physics viewpoint Hamiltonian mechanics was very important in the development of quantum mechanics. For those who have studied the latter topic some of the structures we shall encounter may look familiar.

5.1 The Legendre transformation

In the Lagrangian formulation of classical mechanics we have the Lagrangian function $L(\mathbf{q}, \mathbf{v}, t)$, and the principle of least action leads to Lagrange's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \mathbf{v}}\right) - \frac{\partial L}{\partial \mathbf{q}} \;=\; \mathbf{0}\,, \qquad \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} \;=\; \mathbf{v}\,. \tag{5.1}$$

Here $\mathbf{q} = (q_1, \ldots, q_n)$ are generalized coordinates, and we have written the second order Lagrange equations in a first order form by writing $\mathbf{v} = \dot{\mathbf{q}}$ for the generalized velocity.

The form of Lagrange's equations is invariant under coordinate transformations $\mathbf{q} \to \tilde{\mathbf{q}} = \tilde{\mathbf{q}}(\mathbf{q}, t)$, as demonstrated in equation (2.13). Notice in particular that the generalized velocity $\mathbf{v} = (v_1, \ldots, v_n)$ correspondingly transforms as $v_a \to \tilde{v}_a = \tilde{v}_a(\mathbf{q}, \mathbf{v}, t) = \sum_{b=1}^{n} \frac{\partial \tilde{q}_a}{\partial q_b} v_b + \frac{\partial \tilde{q}_a}{\partial t}$. However, the form of the equations of motion is *not* preserved by more general transformations with $\tilde{\mathbf{q}} = \tilde{\mathbf{q}}(\mathbf{q}, \mathbf{v}, t)$, $\tilde{\mathbf{v}} = \tilde{\mathbf{v}}(\mathbf{q}, \mathbf{v}, t)$. This is related to the fact that the coordinates $\mathbf{q}$ and velocities $\mathbf{v} = \dot{\mathbf{q}}$ are treated quite differently in the Lagrangian approach.

The Hamiltonian formulation rewrites the dynamical equations so that the variables appear on a more equal footing. The form of (5.1) suggests we might achieve this by introducing the *generalized momenta*

$$\mathbf{p} \;\equiv\; \frac{\partial L}{\partial \mathbf{v}}\,. \tag{5.2}$$

Here $\mathbf{p} = \mathbf{p}(\mathbf{q}, \mathbf{v}, t)$, and we saw this definition before in equation (2.48). The Lagrange equations (5.1) are then

$$\dot{\mathbf{p}} \;=\; \frac{\partial L}{\partial \mathbf{q}}\,, \qquad \dot{\mathbf{q}} \;=\; \mathbf{v}\,. \tag{5.3}$$

We aren't quite there yet though, because now we have $\mathbf{q}$, $\mathbf{v}$ *and* $\mathbf{p}$. What we'd like to do is eliminate $\mathbf{v}$ in favour of $\mathbf{p}$. This involves regarding L as a function of $\mathbf{q}$ and the partial derivative $\partial L/\partial \mathbf{v}$ (and time t). Mathematically the transformation that does this is called the *Legendre transform*.

Classical Mechanics

LECTURE 9

The Legendre transformation

For simplicity we focus first on a function of a single variable $f(x)$, so that $f:\mathbb{R}\to\mathbb{R}$. We also introduce its derivative

$$s(x) \equiv \frac{\mathrm{d}f}{\mathrm{d}x} = \partial_x f . \tag{5.4}$$

Here we've used the partial derivative notation, even though we only have one variable, as this will generalize below. We'd like to regard f as a function of its derivative $s=\partial_x f$, rather than as a function of x. In particular this involves inverting the map $s:\mathbb{R}\to\mathbb{R}$. From first year Analysis we know a sufficient condition to be able to do this: if $\partial_x s>0$ in some interval $[a,b]\subset\mathbb{R}$ then s is monotonic increasing on $[a,b]$ and will have a differentiable inverse that we call $x(s)$. (Of course we could replace this by $\partial_x s<0$ so that s is monotonic decreasing, but it is the case $\partial_x s>0$ that arises in classical mechanics, as we shall see shortly.)

We then define the *Legendre transform* of f to be the function

$$g(s) \equiv s\cdot x(s) - f(x(s)) . \tag{5.5}$$

There are various ways to understand why this is a natural definition. First notice that

$$\partial_s g = \frac{\mathrm{d}g}{\mathrm{d}s} = x(s) + s\frac{\mathrm{d}x}{\mathrm{d}s} - \frac{\mathrm{d}f}{\mathrm{d}x}(x(s))\cdot\frac{\mathrm{d}x}{\mathrm{d}s} = x(s) . \tag{5.6}$$

Thus we have symmetry between $s(x)=\partial_x f$ and $x(s)=\partial_s g$. Another important fact is that the Legendre transform is its own inverse. That is, if we perform another Legendre transformation on $g(s)$, we simply recover $f(x)$. We may also more informally rewrite (5.5) as

$$f+g = s\cdot x , \tag{5.7}$$

where this is understood to mean either $f(x(s))+g(s)=s\cdot x(s)$, or the equivalent $f(x)+g(s(x))=s(x)\cdot x$. Equation (5.7) then has a simple geometric interpretation, shown in Figure 17.

Example: Taking $f(x)=x^\alpha/\alpha$, one finds that the Legendre transform is $g(s)=s^\beta/\beta$, where $\frac{1}{\alpha}+\frac{1}{\beta}=1$ $(\alpha,\beta>1)$.

The multivariable case is similar. Now $f=f(x_1,\ldots,x_n)$, and the dual variables are $\mathbf{s}=\partial_{\mathbf{x}}f$, where $\mathbf{s}=(s_1,\ldots,s_n)$. Then $\mathbf{s}=\mathbf{s}(x_1,\ldots,x_n)$ will be invertible in a neighbourhood of the point $(x_1,\ldots,x_n)\in\mathbb{R}^n$ provided the Jacobian determinant $\det(\partial_{x_a}\partial_{x_b}f)$ is non-zero (this is the *inverse function theorem* again). In particular this is true if the $n\times n$ symmetric matrix $(\partial_{x_a}\partial_{x_b}f)$, called the *Hessian* of f, is positive definite. There is then an inverse $\mathbf{x}=\mathbf{x}(s_1,\ldots,s_n)$ and the Legendre transform of $f=f(\mathbf{x})$ is

$$g(\mathbf{s}) = \sum_{a=1}^{n} s_a\cdot x_a(\mathbf{s}) - f(\mathbf{x}(\mathbf{s})) . \tag{5.8}$$

LECTURE 9

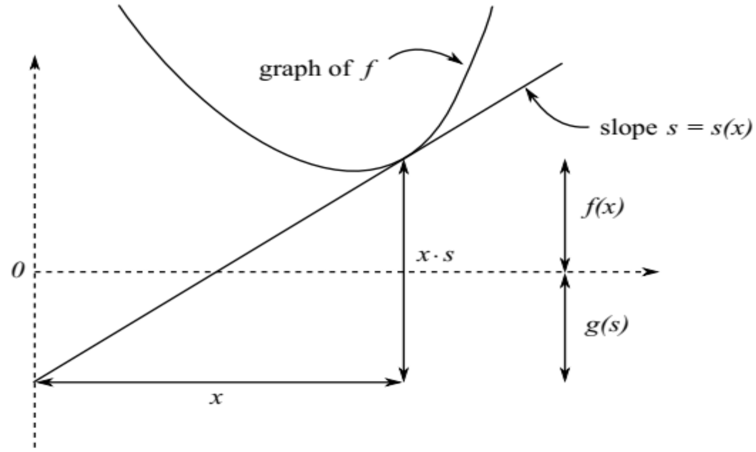

Figure 17: The Legendre transform $g(s)$ of the function $f(x)$.

Using the chain rule we compute

$$\partial_{\mathbf{s}} g \quad = \quad \mathbf{x}(\mathbf{s}) \ , \tag{5.9}$$

analogously to (5.6).

We may now go back to our classical mechanics problem. We define the *Hamiltonian function* $H = H(\mathbf{q}, \mathbf{p}, t)$ to be the Legendre transform of the Lagrangian with respect to the velocities $\mathbf{v} = \dot{\mathbf{q}}$. Thus

$$H(\mathbf{q}, \mathbf{p}, t) \quad \equiv \quad \sum_{a=1}^{n} p_a \dot{q}_a - L(\mathbf{q}, \dot{\mathbf{q}}, t) \Big|_{\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)} \ . \tag{5.10}$$

Comparing to (5.8) we are taking $f = L$ (regarded as a function of $\mathbf{v} = \dot{\mathbf{q}}$ – the dependence on $\mathbf{q}$ and t goes along for the ride), $\mathbf{x} = \dot{\mathbf{q}}$ and hence $\mathbf{s} = \mathbf{p} = \partial_{\dot{\mathbf{q}}} L$. In particular in regarding the right hand side as a function of $\mathbf{p}$, we must invert $\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)$. In classical mechanics notice that the $n \times n$ symmetric matrix $\partial_{\dot{q}_a} \partial_{\dot{q}_b} L$ is usually positive definite. For example for the Lagrangian (3.1) this is simply the kinetic energy tensor T_{ab} (whose positivity played a role in small perturbations of equilibria in section 3). The inverse function theorem then guarantees we can invert $\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)$.

5.2 Hamilton's equations

It is now straightforward to derive *Hamilton's equations*:

$$\dot{\mathbf{p}} \quad = \quad -\frac{\partial H}{\partial \mathbf{q}} \ , \qquad \dot{\mathbf{q}} \quad = \quad \frac{\partial H}{\partial \mathbf{p}} \ . \tag{5.11}$$

The cleanest derivation uses the notation of differentials. We compute

$$\mathrm{d}L \quad = \quad \sum_{a=1}^{n} \left(\frac{\partial L}{\partial q_a} \mathrm{d}q_a + \frac{\partial L}{\partial \dot{q}_a} \mathrm{d}\dot{q}_a \right) + \frac{\partial L}{\partial t} \mathrm{d}t \quad = \quad \sum_{a=1}^{n} (\dot{p}_a \mathrm{d}q_a + p_a \mathrm{d}\dot{q}_a) + \frac{\partial L}{\partial t} \mathrm{d}t \ , \tag{5.12}$$

LECTURE 9

where we have used Lagrange's equations in the form (5.3) for the first term. Thus we find

$$\mathrm{d}H \;=\; \sum_{a=1}^{n}(\dot{q}_a\mathrm{d}p_a + p_a\mathrm{d}\dot{q}_a) - \mathrm{d}L \;=\; \sum_{a=1}^{n}(\dot{q}_a\mathrm{d}p_a - \dot{p}_a\mathrm{d}q_a) - \frac{\partial L}{\partial t}\mathrm{d}t\ . \tag{5.13}$$

But of course also

$$\mathrm{d}H \;=\; \sum_{a=1}^{n}\left(\frac{\partial H}{\partial q_a}\mathrm{d}q_a + \frac{\partial H}{\partial p_a}\mathrm{d}p_a\right) + \frac{\partial H}{\partial t}\mathrm{d}t\ . \tag{5.14}$$

Comparing these last two equations we thus deduce (5.11), together with the relation $\partial L/\partial t = -\partial H/\partial t$.

As an alternative derivation, notice that the second equation in (5.11) is precisely (5.9), where we identify the Legendre transformed function $g = H$, and the variables $\mathbf{x} = \dot{\mathbf{q}}$ and $\mathbf{s} = \mathbf{p} = \partial_{\dot{\mathbf{q}}}L$. Thus it is this particular property of the Legendre transform that has led to such a simple form for Hamilton's equations. The first equation in (5.11) is then simply the Lagrange equation written in the Hamiltonian variables.

Hamilton's equations (5.11) are a set of $2n$ coupled first order differential equations for the two sets of functions $\mathbf{q}(t)$, $\mathbf{p}(t)$. They are equivalent to Lagrange's equations, but are much more symmetric – in particular it is the property (5.9) of the Legendre transform that leads to this symmetry. The coordinates $q_1,\ldots,q_n,p_1,\ldots,p_n$ are often referred to as *canonical coordinates*, and Hamilton's equations (5.11) are sometimes also called the *canonical equations*.

Example: Consider a particle of mass m moving in $\mathbb{R}^3$ under the influence of a potential $V(\mathbf{r})$. The Lagrangian is

$$L \;=\; L(\mathbf{r},\dot{\mathbf{r}}) \;=\; \frac{1}{2}m|\dot{\mathbf{r}}|^2 - V(\mathbf{r})\ . \tag{5.15}$$

The (generalized) momentum conjugate to $\mathbf{r}$ is

$$\mathbf{p} \;=\; \frac{\partial L}{\partial \dot{\mathbf{r}}} \;=\; m\dot{\mathbf{r}}\ , \tag{5.16}$$

which is the usual linear momentum of the particle. It is particularly trivial to invert this as $\dot{\mathbf{r}} = \mathbf{p}/m$, so that the Hamiltonian is

$$H \;=\; H(\mathbf{r},\mathbf{p}) \;=\; \mathbf{p}\cdot\dot{\mathbf{r}} - L \;=\; \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{r})\ . \tag{5.17}$$

Of course this is $H = T + V$, where the kinetic energy is $T = |\mathbf{p}|^2/2m$. Hamilton's equations (5.11) read

$$\dot{\mathbf{p}} \;=\; -\frac{\partial V}{\partial \mathbf{r}}\ , \qquad \dot{\mathbf{r}} \;=\; \frac{\mathbf{p}}{m}\ , \tag{5.18}$$

which we recognize as Newton's second law, together with the standard relation between velocity and momentum.

LECTURE 9

Example: More generally we might consider Lagrangians of the form

$$L = L(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2}\sum_{a,b=1}^{n} T_{ab}(\mathbf{q}, t)\dot{q}_a\dot{q}_b + \sum_{a=1}^{n} C_a(\mathbf{q}, t)\dot{q}_a - V(\mathbf{q}, t) , \qquad (5.19)$$

where $T_{ab} = T_{ba}$ is symmetric. This is a general quadratic function of the generalized velocities $\dot{\mathbf{q}}$. We compute

$$p_a = \frac{\partial L}{\partial \dot{q}_a} = \sum_{b=1}^{n} T_{ab}\dot{q}_b + C_a , \qquad (5.20)$$

and the Hamiltonian is hence

$$\begin{aligned} H &= \sum_{a=1}^{n} p_a\dot{q}_a - L = \frac{1}{2}\sum_{a,b=1}^{n} T_{ab}\dot{q}_a\dot{q}_b + V \\ &= \frac{1}{2}\sum_{a,b=1}^{n} (T^{-1})_{ab}(p_a - C_a)(p_b - C_b) + V , \qquad (5.21) \end{aligned}$$

where $(T^{-1})_{ab}$ is the matrix inverse of T_{ab}. Notice that the last line of (5.21) is a function of $\mathbf{q}, \mathbf{p}, t$, as it should be.

As an example of this, recall from equation (2.84) that the Lagrangian of a particle of mass m and charge e moving in an electromagnetic field is

$$L = \frac{1}{2}m|\dot{\mathbf{r}}|^2 - e\,(\phi - \dot{\mathbf{r}}\cdot\mathbf{A}) . \qquad (5.22)$$

Here ϕ is the electric scalar potential and $\mathbf{A}$ is the magnetic vector potential. The momentum canonically conjugate to $\mathbf{r}$ is

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} + e\,\mathbf{A} . \qquad (5.23)$$

The Hamiltonian is hence

$$\begin{aligned} H(\mathbf{r}, \mathbf{p}) &= \mathbf{p}\cdot\dot{\mathbf{r}} - L \\ &= \frac{1}{m}\mathbf{p}\cdot(\mathbf{p} - e\,\mathbf{A}) - \left[\frac{1}{2m}(\mathbf{p} - e\,\mathbf{A})^2 - e\,\phi + \frac{e}{m}(\mathbf{p} - e\,\mathbf{A})\cdot\mathbf{A}\right] \\ &= \frac{1}{2m}(\mathbf{p} - e\,\mathbf{A})^2 + e\,\phi , \qquad (5.24) \end{aligned}$$

where we have inverted (5.23) to give $\dot{\mathbf{r}} = \frac{1}{m}(\mathbf{p} - e\,\mathbf{A})$. Comparing to (5.19), (5.21) we have $T = \mathrm{diag}(m, m, m)$, $\mathbf{C} = e\,\mathbf{A}$, $V = e\,\phi$. The Hamilton equation $\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{1}{m}(\mathbf{p} - e\,\mathbf{A})$ simply re-expresses $\dot{\mathbf{r}}$ in terms of $\mathbf{p}$, while the other Hamilton equation reads (in components)

$$\dot{p}_a = -\frac{\partial H}{\partial r_a} = -e\frac{\partial \phi}{\partial r_a} + \frac{e}{m}\sum_{b=1}^{3}(p_b - e\,A_b)\frac{\partial A_b}{\partial r_a} . \qquad (5.25)$$

Classical Mechanics

LECTURE 9

Phase space

The canonical coordinates $q_1, \ldots, q_n, p_1, \ldots, p_n$ parametrize the $2n$-dimensional *phase space* $\mathscr{P}$ of the system. Notice this has twice the dimension of the configuration space $\mathcal{Q}$, which is parametrized by $q_1, \ldots, q_n$. For example, a single particle moving in $\mathbb{R}^n$ has phase space $\mathscr{P} = \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$, with the points labelled by $(\mathbf{r}, \mathbf{p})$. In particular a particle moving in one dimension along the x-axis has phase space $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, with coordinates (x, p). A more interesting example is the simple pendulum we studied in section 2. Recall here that the configuration space is $\mathcal{Q} = S^1$, parametrized by the angle θ the pendulum makes with the vertical. Then $\theta \in [-\pi, \pi)$, with $\theta = -\pi$ identified with $\theta = \pi$. On the other hand the conjugate momentum $p = ml^2\dot{\theta}$ can take any real value, so the phase space is $\mathscr{P} = S^1 \times \mathbb{R}$.

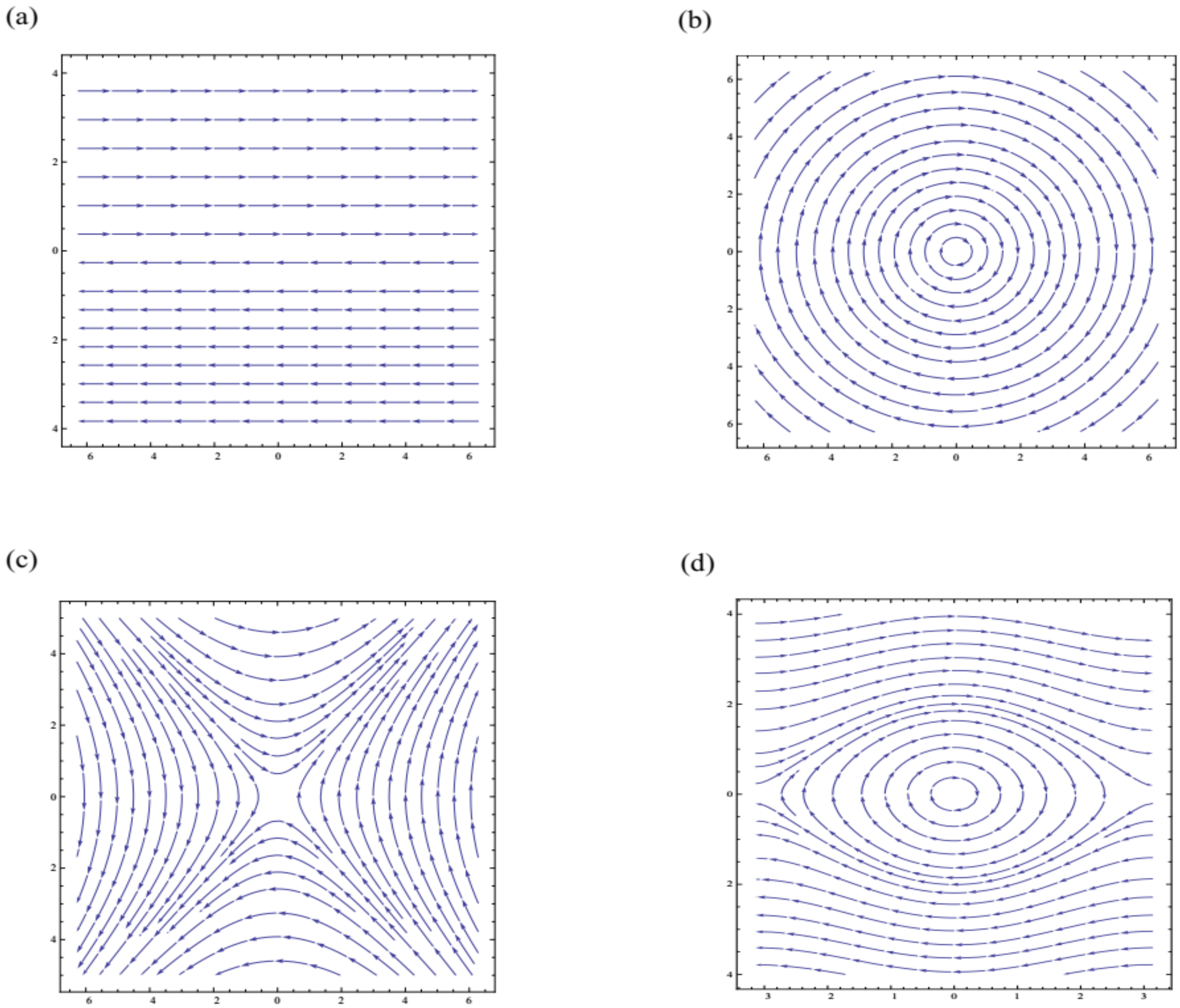

Figure 18: Hamiltonian flows on the phase spaces of a number of one-dimensional systems: (a) a free particle, (b) the harmonic oscillator, (c) the inverted harmonic oscillator, (d) the simple pendulum.

Classical Mechanics

LECTURE 9

One can think of Hamilton's equations (5.11) as first order equations for a trajectory $\gamma(t) = (q_1(t), \ldots, q_n(t), p_1(t), \ldots, p_n(t))$ in $\mathscr{P}$. If we pick a point $(\mathbf{q}^{(0)}, \mathbf{p}^{(0)}) \in \mathscr{P}$, then general results from the theory of differential equations guarantee that under suitable conditions on the Hamiltonian function H we will have a unique solution for the path $\gamma(t) = (\mathbf{q}(t), \mathbf{p}(t))$ satisfying Hamilton's equations with initial condition $\gamma(0) = (\mathbf{q}^{(0)}, \mathbf{p}^{(0)})$. The set of all trajectories $\{\gamma(t)\}$ is called the *Hamiltonian flow* on phase space.

In order to make this more concrete, and to give some examples, let's focus on the case of dimension $n = 1$. The phase space is then two-dimensional, parametrized by (q, p). In Figure 18 we have shown parts of the phase spaces and Hamiltonian flow trajectories for 4 systems, where the spatial variable q is plotted horizontally and the momentum p vertically.

(a) For a free particle of mass m the Hamiltonian is simply $H = \frac{p^2}{2m}$. If the particle moves on the x-axis then we may use $q = x$ as a generalized coordinate, and the Hamiltonian flow equations are then $(\dot{x}(t), \dot{p}(t)) = (\partial_p H, -\partial_x H) = (\frac{1}{m}p(t), 0)$. In Figure 18(a) we have plotted some of the trajectories (setting $m = 1$). Notice that the entire x-axis at $p = 0$ is *fixed*.

(b) Using the same spatial coordinate $q = x$, the Hamiltonian for the harmonic oscillator is $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$, where ω is the frequency. The flow equations are hence $(\dot{x}(t), \dot{p}(t)) = (\partial_p H, -\partial_x H) = (\frac{1}{m}p(t), -m\omega^2 x)$. Some trajectories are plotted in Figure 18(b), where we have set $m = 1 = \omega$ so that $(\dot{x}, \dot{p}) = (p, -x)$. The origin $(x, p) = (0, 0)$ is a stable equilibrium point.

(c) The *inverted harmonic oscillator* has minus the potential energy of the harmonic oscillator, *i.e.* the Hamiltonian is $H = \frac{p^2}{2m} - \frac{1}{2}m\omega^2 x^2$. With $m = 1 = \omega$ the trajectories now obey $(\dot{x}, \dot{p}) = (p, x)$ – see Figure 18(c). The origin $(x, p) = (0, 0)$ is now an *unstable* equilibrium point.

(d) Finally recall that the simple pendulum has phase space $S^1 \times \mathbb{R}$, with the angle $\theta = q$ being the generalized coordinate on S^1. The Hamiltonian corresponding to the Lagrangian (2.24) is $H = \frac{p^2}{2ml^2} - mgl\cos\theta$, giving trajectories obeying $(\dot{\theta}, \dot{p}) = (\frac{1}{ml^2}p, -mgl\sin\theta)$. Some trajectories are shown in Figure 18(d), with $m = 1 = l = g$. The θ-axis is horizontal, with the left and right limits being $\theta = -\pi$, $\theta = \pi$, respectively, which should be identified. Notice that the trajectories around the origin (a stable equilibruim point) resemble those for the harmonic oscillator (as happens near any stable equilibrium point). As $|p|$ increases we begin to see trajectories that wind around the circle S^1, *i.e.* the pendulum swings all the way around. The point $(\theta, p) = (-\pi, 0) \sim (\pi, 0)$ is an unstable equilibrium point.