Chapter 6

Hamilton's Equations

We discussed the generalized momenta

$$p_i = \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i},$$

and how the canonical variables $\{q_i, p_j\}$ describe phase space. One can use phase space rather than $\{q_i, \dot{q}_j\}$ to describe the state of a system at any moment. In this chapter we will explore the tools which stem from this phase space approach to dynamics.

6.1 Legendre transforms

The important object for determining the motion of a system using the Lagrangian approach is not the Lagrangian itself but its variation, under arbitrary changes in the variables q and $\dot{q}$, treated as independent variables. It is the vanishing of the variation of the action under such variations which determines the dynamical equations. In the phase space approach, we want to change variables $\dot{q} \to p$, where the p_i are components of the gradient of the Lagrangian with respect to the velocities. This is an example of a general procedure called the Legendre transformation. We will discuss it in terms of the mathematical concept of a differential form.

Because it is the variation of L which is important, we need to focus our attention on the differential dL rather than on L itself. We first want to give a formal definition of the differential, which we will do first for a function $f(x_1, ..., x_n)$ of n variables, although for the Lagrangian we will

later subdivide these into coordinates and velocities. We will take the space in which x takes values to be some general n-dimensional space we call $\mathcal{M}$, which might be ordinary Euclidean space but might be something else, like the surface of a sphere[1]. Given a differentiable function f of n independent variables x_i, the differential is

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i. \tag{6.1}$$

What does that mean? As an approximate statement, this can be regarded as saying

$$df \approx \Delta f \equiv f(x_i + \Delta x_i) - f(x_i) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i + \mathcal{O}(\Delta x_i \Delta x_j),$$

with some statement about the Δx_i being small, followed by the dropping of the "order $(\Delta x)^2$" terms. Notice that df is a function not only of the point $x \in \mathcal{M}$, but also of the small displacements Δx_i. A very useful mathematical language emerges if we formalize the definition of df, extending its definition to arbitrary Δx_i, even when the Δx_i are not small. Of course, for large Δx_i they can no longer be thought of as the difference of two positions in $\mathcal{M}$ and df no longer has the meaning of the difference of values of f at two different points. Our formal df is now defined as a linear function of these Δx_i variables, which we therefore consider to be a vector $\vec{v}$ lying in an n-dimensional vector space $\mathbb{R}^n$. Thus $df : \mathcal{M} \times \mathbb{R}^n \to \mathbb{R}$ is a real-valued function with two arguments, one in $\mathcal{M}$ and one in a vector space. The dx_i which appear in (6.1) can be thought of as operators acting on this vector space argument to extract the $i'th$ component, and the action of df on the argument $(x, \vec{v})$ is $df(x, \vec{v}) = \sum_i (\partial f / \partial x_i) v_i$.

This differential is a special case of a 1-form, as is each of the operators dx_i. All n of these dx_i form a basis of **1-forms**, which are more generally

$$\omega = \sum_i \omega_i(x) dx_i,$$

where the $\omega_i(x)$ are n functions on the manifold $\mathcal{M}$. If there exists an ordinary function $f(x)$ such that $\omega = df$, then ω is said to be an **exact** 1-form.

[1]Mathematically, $\mathcal{M}$ is a manifold, but we will not carefully define that here. The precise definition is available in Ref. [16].

Consider $L(q_i, v_j, t)$, where $v_i = \dot{q}_i$. At a given time we consider q and v as independant variables. The differential of L on the space of coordinates and velocities, at a fixed time, is

$$dL = \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i \frac{\partial L}{\partial v_i} dv_i = \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i p_i dv_i.$$

If we wish to describe physics in phase space (q_i, p_i), we are making a change of variables from v_i to the gradient with respect to these variables, $p_i = \partial L/\partial v_i$, where we focus now on the variables being transformed and ignore the fixed q_i variables. So $dL = \sum_i p_i dv_i$, and the p_i are functions of the v_j determined by the function $L(v_i)$. Is there a function $g(p_i)$ which reverses the roles of v and p, for which $dg = \sum_i v_i dp_i$? If we can invert the functions $p(v)$, we can define $g(p_i) = \sum_i p_i v_i(p_j) - L(v_i(p_j))$, which has a differential

$$\begin{aligned} dg &= \sum_i p_i\, dv_i + \sum_i v_i\, dp_i - dL = \sum_i p_i\, dv_i + \sum_i v_i\, dp_i - \sum_i p_i\, dv_i \\ &= \sum_i v_i dp_i \end{aligned}$$

as requested, and which also determines the relationship between v and p,

$$v_i = \frac{\partial g}{\partial p_i} = v_i(p_j),$$

giving the inverse relation to $p_k(v_\ell)$. This particular form of changing variables is called a **Legendre transformation**. In the case of interest here, the function g is called $H(q_i, p_j, t)$, the **Hamiltonian**,

$$H(q_i, p_j, t) = \sum_k p_k \dot{q}_k(q_i, p_j, t) - L(q_i, \dot{q}_j(q_\ell, p_m, t), t). \tag{6.2}$$

Then for fixed time,

$$dH = \sum_k \left(dp_k\, \dot{q}_k + p_k d\dot{q}_k\right) - dL = \sum_k \left(\dot{q}_k\, dp_k - p_k dq_k\right),$$

$$\left.\frac{\partial H}{\partial p_k}\right|_{q,t} = \dot{q}_k, \qquad \left.\frac{\partial H}{\partial q_k}\right|_{p,t} = -p_k. \tag{6.3}$$

Other examples of Legendre transformations occur in thermodynamics. The energy change of a gas in a variable container with heat flow is sometimes written

$$dE = đQ - pdV,$$

where $đQ$ is not an exact differential, and the heat Q is not a well defined system variable. Though Q is not a well defined state function, the differential $đQ$ is a well defined 1-form on the manifold of possible states of the system. It is not, however, an exact 1-form, which is why Q is not a function on that manifold. We can express $đQ$ by defining the entropy and temperature, in terms of which $đQ = TdS$, and the entropy S and temperature T are well defined state functions. Thus the state of the gas can be described by the two variables S and V, and changes involve an energy change

$$dE = TdS - pdV.$$

We see that the temperature is $T = \partial E/\partial S|_V$. If we wish to find quantities appropriate for describing the gas as a function of T rather than S, we define the **free energy** F by $-F = TS - E$ so $dF = -SdT - pdV$, and we treat F as a function $F(T, V)$. Alternatively, to use the pressure p rather than V, we define the **enthalpy** $X(p, S) = Vp + E$, $dX = Vdp + TdS$. To make both changes, and use (T, p) to describe the state of the gas, we use the **Gibbs free energy** $G(T, p) = X - TS = E + Vp - TS$, $dG = Vdp - SdT$. Each of these involves a Legendre transformation starting with $E(S, V)$.

Unlike Q, E is a well defined property of the gas when it is in a volume V if its entropy is S, so $E = E(S, V)$, and

$$T = \left.\frac{\partial E}{\partial S}\right|_V, \qquad p = \left.\frac{\partial E}{\partial V}\right|_S.$$

As $\dfrac{\partial^2 E}{\partial S \partial V} = \dfrac{\partial^2 E}{\partial V \partial S}$ we can conclude that $\left.\dfrac{\partial T}{\partial V}\right|_S = \left.\dfrac{\partial p}{\partial S}\right|_V$. We may also consider the state of the gas to be described by T and V, so

$$dE = \left.\frac{\partial E}{\partial T}\right|_V dT + \left.\frac{\partial E}{\partial V}\right|_T dV$$

$$dS = \frac{1}{T} dE + \frac{p}{T} dV = \frac{1}{T} \left.\frac{\partial E}{\partial T}\right|_V dT + \left[\frac{1}{T}\left(p + \left.\frac{\partial E}{\partial V}\right|_T\right)\right] dV,$$

from which we can conclude

$$\left.\frac{\partial}{\partial V}\left(\frac{1}{T} \left.\frac{\partial E}{\partial T}\right|_V\right)\right|_T = \left.\frac{\partial}{\partial T}\left[\frac{1}{T}\left(p + \left.\frac{\partial E}{\partial V}\right|_T\right)\right]\right|_V,$$

and therefore

$$T\left.\frac{\partial p}{\partial T}\right|_V - p = \left.\frac{\partial E}{\partial V}\right|_T.$$

This is a useful relation in thermodynamics.

Let us get back to mechanics. Most Lagrangians we encounter have the decomposition $L = L_2 + L_1 + L_0$ into terms quadratic, linear, and independent of velocities, as considered in 2.4.2. Then the momenta are linear in velocities, $p_i = \sum_j M_{ij}\dot q_j + a_i$, or in matrix form $p = M\cdot\dot q + a$, which has the inverse relation $\dot q = M^{-1}\cdot(p-a)$. As $H = L_2 - L_0$, $H = \frac{1}{2}(p-a)\cdot M^{-1}\cdot(p-a) - L_0$. As a simple example, with $a = 0$ and a diagonal matrix M, consider spherical coordinates, in which the kinetic energy is

$$T = \frac{m}{2}\left(\dot r^2 + r^2\dot\theta^2 + r^2\sin^2\theta\dot\phi^2\right) = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}\right).$$

Note that the generalized momenta are not normalized components of the ordinary momentum, as $p_\theta \neq \vec p\cdot\hat e_\theta$, in fact it doesn't even have the same units.

The equations of motion in Hamiltonian form (6.3),

$$\dot q_k = \left.\frac{\partial H}{\partial p_k}\right|_{q,t}, \qquad \dot p_k = -\left.\frac{\partial H}{\partial q_k}\right|_{p,t},$$

are almost symmetric in their treatment of q and p. If we define a $2N$ dimensional coordinate η for phase space,

$$\left.\begin{aligned}\eta_i &= q_i\\ \eta_{N+i} &= p_i\end{aligned}\right\}\qquad \text{for } 1\le i\le N,$$

we can write Hamilton's equation in terms of a particular matrix J,

$$\dot\eta_j = \sum_{k=1}^{2N} J_{jk}\frac{\partial H}{\partial\eta_k}, \quad \text{where } J = \begin{pmatrix} 0 & \mathbb{1}_{N\times N}\\ -\mathbb{1}_{N\times N} & 0\end{pmatrix}.$$

J is like a multidimensional version of the $i\sigma_y$ which we meet in quantum-mechanical descriptions of spin 1/2 particles. It is real, antisymmetric, and because $J^2 = -\mathbb{1}$, it is orthogonal. Mathematicians would say that J describes the **complex structure on phase space**, also called the **symplectic structure**.

In Section 2.1 we discussed how the Lagrangian is unchanged by a change of generalized coordinates used to describe the physical situation. More precisely, the Lagrangian transforms as a scalar under such point transformations, taking on the same value at the same physical point, described in the new coordinates. There is no unique set of generalized coordinates which describes the physics. But in transforming to the Hamiltonian language, different generalized coordinates may give different momenta and different Hamiltonians. An nice example is given in Goldstein, a mass on a spring attached to a "fixed point" which is on a truck moving at uniform velocity v_T, relative to the Earth. If we use the Earth coordinate x to describe the mass, the equilibrium position of the spring is moving in time, $x_{eq} = v_T t$, ignoring a negligible initial position. Thus $U = \frac{1}{2}k(x - v_T t)^2$, while $T = \frac{1}{2}m\dot x^2$ as usual, and $L = \frac{1}{2}m\dot x^2 - \frac{1}{2}k(x-v_T t)^2$, $p = m\dot x$, $H = p^2/2m + \frac{1}{2}k(x-v_T t)^2$. The equation of motion is $\dot p = m\ddot x = -\partial H/\partial x = -k(x - v_T t)$, of course. This shows that H is not conserved, $dH/dt = (p/m)dp/dt + k(\dot x - v_T)(x - v_T t) = -(kp/m)(x - v_T t) + (kp/m - kv_T)(x - v_T t) = -kv_T(x - v_T t) \neq 0$. Alternatively, $dH/dt = -\partial L/\partial t = -kv_T(x - v_T t) \neq 0$. This is not surprising; the spring exerts a force on the truck and the truck is doing work to keep the fixed point moving at constant velocity.

On the other hand, if we use the truck coordinate $x' = x - v_T t$, we may describe the motion in this frame with $T' = \frac{1}{2}m\dot x'^2$, $U' = \frac{1}{2}kx'^2$, $L' = \frac{1}{2}m\dot x'^2 - \frac{1}{2}kx'^2$, giving the correct equations of motion $p' = m\dot x'$, $\dot p' = m\ddot x' = -\partial L'/\partial x' = -kx'$. With this set of coordinates, the Hamiltonian is $H' = \dot x'p' - L' = p'^2/2m + \frac{1}{2}kx'^2$, which is conserved. From the correspondence between the two sets of variables, $x' = x - v_T t$, and $p' = p - mv_T$, we see that the Hamiltonians at corresponding points in phase space differ, $H(x,p) - H'(x',p') = (p^2 - p'^2)/2m = 2mv_T p - \frac{1}{2}mv_T^2 \neq 0$.

Thus the Hamiltonian is not invariant, or a scalar, under change of generalized coordinates, or point transformations. We shall see, however, that it is invariant under time independent transformations that are more general than point transformations, mixing coordinates and momenta.

6.2 Variations on phase curves

In applying Hamilton's Principle to derive Lagrange's Equations, we considered variations in which $\delta q_i(t)$ was arbitrary except at the initial and final times, but the velocities were fixed in terms of these, $\delta\dot q_i(t) = (d/dt)\delta q_i(t)$.

In discussing dynamics in terms of phase space, this is not the most natural variation, because this means that the momenta are not varied independently. Here we will show that Hamilton's equations follow from a modified Hamilton's Principle, in which the momenta are freely varied.

We write the action in terms of the Hamiltonian,

$$I = \int_{t_i}^{t_f} \left[\sum_i p_i \dot{q}_i - H(q_j, p_j, t) \right] dt,$$

and consider its variation under arbitrary variation of the path in phase space, $(q_i(t), p_i(t))$. The $\dot{q}_i(t)$ is still dq_i/dt, but the momentum is varied free of any connection to $\dot{q}_i$. Then

$$\delta I = \int_{t_i}^{t_f} \left[\sum_i \delta p_i \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) - \sum_i \delta q_i \left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \right] dt + \sum_i p_i \delta q_i \Big|_{t_i}^{t_f},$$

where we have integrated the $\int \sum p_i d\delta q_i/dt$ term by parts. Note that in order to relate stationarity of the action to Hamilton Equations of Motion, it is necessary only to constrain the $q_i(t)$ at the initial and final times, without imposing any limitations on the variation of $p_i(t)$, either at the endpoints, as we did for $q_i(t)$, or in the interior (t_i, t_f), where we had previously related p_i and $\dot{q}_j$. The relation between $\dot{q}_i$ and p_j emerges instead among the equations of motion.

The $\dot{q}_i$ seems a bit out of place in a variational principle over phase space, and indeed we can rewrite the action integral as an integral of a 1-form over a path in extended phase space,

$$I = \int \left(\sum_i p_i dq_i - H(q, p, t) dt \right).$$

We will see, in section 6.6, that the first term of the integrand leads to a very important form on phase space, and that the whole integrand is an important 1-form on extended phase space.

6.3 Canonical transformations

We have seen that it is often useful to switch from the original set of coordinates in which a problem appeared to a different set in which the problem

became simpler. We switched from cartesian to center-of-mass spherical coordinates to discuss planetary motion, for example, or from the Earth frame to the truck frame in the example in which we found how Hamiltonians depend on coordinate choices. In all these cases we considered a change of coordinates $q \to Q$, where each Q_i is a function of all the q_j and possibly time, but not of the momenta or velocities. This is called a **point transformation**. But we have seen that we can work in phase space where coordinates and momenta enter together in similar ways, and we might ask ourselves what happens if we make a change of variables on phase space, to new variables $Q_i(q, p, t)$, $P_i(q, p, t)$, which also cover phase space, so there is a relation $(q, p) \leftrightarrow (Q, P)$ given by the two descriptions of the same point in phase space. We should not expect the Hamiltonian to be the same either in form or in value, as we saw even for point transformations, but there must be a new Hamiltonian $K(Q, P, t)$ from which we can derive the correct equations of motion,

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}.$$

The analog of η for our new variables will be $\zeta = \begin{pmatrix} Q \\ P \end{pmatrix}$, and the relation $\eta \leftrightarrow \zeta$ means each can be viewed as a function of the other. Hamilton's equations for ζ are $\dot{\zeta} = J \cdot \dfrac{\partial K}{\partial \zeta}$. If this exists, we say the new variables (Q, P) are **canonical variables** and the transformation $(q, p) \to (Q, P)$ is a **canonical transformation**. Note that the functions Q_i and P_i may depend on time as well as on q and p.

These new Hamiltonian equations are related to the old ones, $\dot{\eta} = J \cdot \partial H/\partial \eta$, by the function which gives the new coordinates and momenta in terms of the old, $\zeta = \zeta(\eta, t)$. Then

$$\dot{\zeta}_i = \frac{d\zeta_i}{dt} = \sum_j \frac{\partial \zeta_i}{\partial \eta_j} \dot{\eta}_j + \frac{\partial \zeta_i}{\partial t}.$$

Let us write the Jacobian matrix $M_{ij} := \partial \zeta_i / \partial \eta_j$. In general, M will not be a constant but a function on phase space. The above relation for the velocities now reads

$$\dot{\zeta} = M \cdot \dot{\eta} + \left.\frac{\partial \zeta}{\partial t}\right|_{\eta}.$$

The gradients in phase space are also related,

$$\left.\frac{\partial}{\partial \eta_i}\right|_{t,\eta} = \sum_j \left.\frac{\partial \zeta_j}{\partial \eta_i}\right|_{t,\eta} \left.\frac{\partial}{\partial \zeta_j}\right|_{t,\zeta}, \quad \text{or } \nabla_\eta = M^T \cdot \nabla_\zeta.$$

Thus we have

$$\begin{aligned}\dot\zeta &= M\cdot\dot\eta + \frac{\partial \zeta}{\partial t} = M\cdot J\cdot\nabla_\eta H + \frac{\partial\zeta}{\partial t} = M\cdot J\cdot M^T\cdot\nabla_\zeta H + \frac{\partial\zeta}{\partial t}\\ &= J\cdot\nabla_\zeta K.\end{aligned}$$

Let us first consider a canonical transformation which does not depend on time, so $\partial\zeta/\partial t|_\eta = 0$. We see that we can choose the new Hamiltonian to be the same as the old, $K = H$ (*i.e.* $K(\zeta,t) := H(\eta(\zeta),t)$), and get correct mechanics, provided

$$M\cdot J\cdot M^T = J. \tag{6.4}$$

We will require this condition even when ζ does depend on t, but then we need to revisit the question of finding K.

The condition (6.4) on M is similar to, and a generalization of, the condition for orthogonality of a matrix, $\mathcal{O}\mathcal{O}^T = 1\!\!1$, which is of the same form with J replaced by $1\!\!1$. Another example of this kind of relation in physics occurs in special relativity, where a Lorentz transformation $L_{\mu\nu}$ gives the relation between two coordinates, $x'_\mu = \sum_\nu L_{\mu\nu}x_\nu$, with x_ν a four dimensional vector with $x_4 = ct$. Then the condition which makes L a Lorentz transformation is

$$L\cdot g\cdot L^T = g, \quad \text{with } g = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}.$$

The matrix g in relativity is known as the indefinite metric, and the condition on L is known as pseudo-orthogonality. In our current discussion, however, J is not a metric, as it is antisymmetric rather than symmetric, and the word which describes M is **symplectic**.

So far the only examples of canonical transformation which we have discussed have been point transformations. Before we continue, let us recall that such transformations can be discrete, *e.g.* $(x,y,z)\to(r,\theta,\phi)$, but they can also be continuous, or depending on a parameter. For example, with rotating coordinates we considered $x = r\cos(\theta+\phi), y = r\sin(\theta+\phi)$, which can be viewed as a set of possible discrete transformations $(r,\theta)\leftrightarrow(x,y)$ depending on a parameter ϕ, which might even be a function of time, $\phi(t) = \Omega t$. For each fixed ϕ this transformation $\mathcal{P}_\phi : (x,y)\mapsto(r,\theta)$ is an acceptible point transformation, which can be used to describe phase space at any time t. Thus we may also consider a composition, the point tranformation $\mathcal{P}_{\phi(t+\Delta t)}\mathcal{P}^{-1}_{\phi(t)}$ from the polar coordinate system at time t to that at $t+\Delta t$. This is an infinitesimal point transformation, and is what we used in discussing the angular velocity in rigid body motion.

Just as for orthogonal transformations, symplectic transformations can be divided into those which can be generated by infinitesimal transformations (which are connected to the identity) and those which can not. Consider a transformation M which is almost the identity, $M_{ij} = \delta_{ij} + \epsilon G_{ij}$, or $M = 1\!\!1 + \epsilon G$, where ϵ is considered some infinitesimal parameter while G is a finite matrix. As M is symplectic, $(1+\epsilon G)\cdot J\cdot(1+\epsilon G^T) = J$, which tells us that to lowest order in ϵ, $GJ + JG^T = 0$. Comparing this to the condition for the generator of an infinitesimal rotation, $\Omega = -\Omega^T$, we see that it is similar except for the appearence of J on opposite sides, changing orthogonality to symplecticity. The new variables under such a canonical transformation are $\zeta = \eta + \epsilon G\cdot\eta$.

The condition (6.4) for a transformation $\eta\to\zeta$ to be canonical does not involve time — each canonical transformation is a fixed map of phase-space onto itself, and could be used at any t. We might consider a set of such maps, one for each time, giving a time dependant map $g(t):\eta\to\zeta$. Each such map could be used to transform the trajectory of the system at any time. In particular, consider the set of maps $g(t,t_0)$ which maps each point η at which a system can be at time t_0 into the point to which it will evolve at time t. That is, $g(t,t_0):\eta(t_0)\mapsto\eta(t)$. If we consider $t = t_0+\Delta t$ for infinitesimal Δt, this is an infinitesimal transformation. As $\zeta_i = \eta_i + \Delta t\dot\eta_i = \eta_i + \Delta t\sum_k J_{ik}\partial H/\partial\eta_k$, we have $M_{ij} = \partial\zeta_i/\partial\eta_j = \delta_{ij} + \Delta t\sum_k J_{ik}\partial^2 H/\partial\eta_j\partial\eta_k$, so $G_{ij} = \sum_k J_{ik}\partial^2 H/\partial\eta_j\partial\eta_k$,

$$\begin{aligned}(GJ+JG^T)_{ij} &= \sum_{k\ell}\left(J_{ik}\frac{\partial^2 H}{\partial\eta_\ell\partial\eta_k}J_{\ell j} + J_{i\ell}J_{jk}\frac{\partial^2 H}{\partial\eta_\ell\partial\eta_k}\right)\\ &= \sum_{k\ell}\left(J_{ik}J_{\ell j} + J_{i\ell}J_{jk}\right)\frac{\partial^2 H}{\partial\eta_\ell\partial\eta_k}.\end{aligned}$$

The factor in parentheses in the last line is $(-J_{ik}J_{j\ell} + J_{i\ell}J_{jk})$ which is antisymmetric under $k \leftrightarrow \ell$, and as it is contracted into the second derivative, which is symmetric under $k \leftrightarrow \ell$, we see that $(GJ + JG^T)_{ij} = 0$ and we have an infinitesimal canonical transformation. Thus the infinitesimal flow of phase space points by the velocity function is canonical. As compositions of canonical transformations are also canonical[2], the map $g(t, t_0)$ which takes $\eta(t_0)$ into $\eta(t)$, the point it will evolve into after a finite time increment $t - t_0$, is also a canonical transformation.

Notice that the relationship ensuring Hamilton's equations exist,

$$M \cdot J \cdot M^T \cdot \nabla_\zeta H + \frac{\partial \zeta}{\partial t} = J \cdot \nabla_\zeta K,$$

with the symplectic condition $M \cdot J \cdot M^T = J$, implies $\nabla_\zeta(K - H) = -J \cdot \partial\zeta/\partial t$, so K differs from H here. This discussion holds as long as M is symplectic, even if it is not an infinitesimal transformation.

6.4 Poisson Brackets

Suppose I have some function $f(q, p, t)$ on phase space and I want to ask how f, evaluated on a dynamical system, changes as the system evolves through phase space with time. Then

$$\begin{aligned} \frac{df}{dt} &= \sum_i \frac{\partial f}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial f}{\partial p_i}\dot{p}_i + \frac{\partial f}{\partial t} \\ &= \sum_i \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \sum_i \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial t}. \end{aligned} \tag{6.5}$$

The structure of the first two terms is that of a **Poisson bracket**, a bilinear operation of functions on phase space defined by

$$[u, v] := \sum_i \frac{\partial u}{\partial q_i}\frac{\partial v}{\partial p_i} - \sum_i \frac{\partial u}{\partial p_i}\frac{\partial v}{\partial q_i}. \tag{6.6}$$

Thus Eq. 6.5 may be rewritten as

$$\frac{df}{dt} = [f, H] + \frac{\partial f}{\partial t}. \tag{6.7}$$

[2] If $M = M_1 \cdot M_2$ and $M_1 \cdot J \cdot M_1^T = J$, $M_2 \cdot J \cdot M_2^T = J$, then $M \cdot J \cdot M^T = (M_1 \cdot M_2) \cdot J(\cdot M_2^T \cdot M_1^T) = M_1 \cdot (M_2 \cdot J \cdot M_2^T) \cdot M_1^T = M_1 \cdot J \cdot M_1^T = J$, so M is canonical.

The Poisson bracket is a fundamental property of the phase space. In symplectic language,

$$[u, v] = \sum_{ij} \frac{\partial u}{\partial \eta_i} J_{ij} \frac{\partial v}{\partial \eta_j} = (\nabla_\eta u)^T \cdot J \cdot \nabla_\eta v. \tag{6.8}$$

If we describe the system in terms of a different set of canonical variables ζ, we should still find the function $f(t)$ changing at the same rate. We may think of u and v as functions of ζ as easily as of η. Really we are thinking of u and v as functions of points in phase space, represented by $u(\eta) = \tilde{u}(\zeta)$ and we may ask whether $[\tilde{u}, \tilde{v}]_\zeta$ is the same as $[u, v]_\eta$. Using $\nabla_\eta = M^T \cdot \nabla_\zeta$, we have

$$\begin{aligned} [u, v]_\eta &= \left(M^T \cdot \nabla_\zeta \tilde{u}\right)^T \cdot J \cdot M^T \nabla_\zeta \tilde{v} = (\nabla_\zeta \tilde{u})^T \cdot M \cdot J \cdot M^T \nabla_\zeta \tilde{v} \\ &= (\nabla_\zeta \tilde{u})^T \cdot J \nabla_\zeta \tilde{v} = [\tilde{u}, \tilde{v}]_\zeta\,, \end{aligned}$$

so we see that the Poisson bracket is independent of the coordinatization used to describe phase space, as long as it is canonical.

The Poisson bracket plays such an important role in classical mechanics, and an even more important role in quantum mechanics, that it is worthwhile to discuss some of its abstract properties. First of all, from the definition it is obvious that it is antisymmetric:

$$[u, v] = -[v, u]. \tag{6.9}$$

It is a linear operator on each function over *constant* linear combinations, but is satisfies a Leibnitz rule for non-constant multiples,

$$[uv, w] = [u, w]v + u[v, w], \tag{6.10}$$

which follows immediately from the definition, using Leibnitz' rule on the partial derivatives. A very special relation is the **Jacobi identity**,

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0. \tag{6.11}$$

We need to prove that this is true. To simplify the presentation, we introduce some abbreviated notation. We use a subscript $_{,i}$ to indicate partial derivative with respect to η_i, so $u_{,i}$ means $\partial u/\partial \eta_i$, and $u_{,i,j}$ means $\partial(\partial u/\partial \eta_i)/\partial \eta_j$. We will assume all our functions on phase space are suitably differentiable, so $u_{,i,j} = u_{,j,i}$. We will also use the **summation convention**, that any index which appears twice in a term is assumed to be summed over[3]. Then $[v,w] = v_{,i}J_{ij}w_{,j}$, and

$$\begin{aligned}[u,[v,w]] &= [u, v_{,i}J_{ij}w_{,j}] \\ &= [u,v_{,i}]J_{ij}w_{,j} + v_{,i}J_{ij}[u,w_{,j}] \\ &= u_{,k}J_{k\ell}v_{,i,\ell}J_{ij}w_{,j} + v_{,i}J_{ij}u_{,k}J_{k\ell}w_{,j,\ell}.\end{aligned}$$

In the Jacobi identity, there are two other terms like this, one with the substitution $u \to v \to w \to u$ and the other with $u \to w \to v \to u$, giving a sum of six terms. The only ones involving second derivatives of v are the first term above and the one found from applying $u \to w \to v \to u$ to the second, $u_{,i}J_{ij}w_{,k}J_{k\ell}v_{,j,\ell}$. The indices are all dummy indices, summed over, so their names can be changed, by $i \to k \to j \to \ell \to i$, converting this second term to $u_{,k}J_{k\ell}w_{,j}J_{ji}v_{,\ell,i}$. Adding the original term $u_{,k}J_{k\ell}v_{,i,\ell}J_{ij}w_{,j}$, and using $v_{,\ell,i} = v_{,i,\ell}$, gives $u_{,k}J_{k\ell}w_{,j}(J_{ji}+J_{ij})v_{,\ell,i} = 0$ because J is antisymmetric. Thus the terms in the Jacobi identity involving second derivatives of v vanish, but the same argument applies in pairs to the other terms, involving second derivatives of u or of w, so they all vanish, and the Jacobi identity is proven.

This argument can be made more elegantly if we recognize that for each function f on phase space, we may view $[f,\cdot]$ as a differential operator on functions g on phase space, mapping $g \to [f,g]$. Calling this operator D_f, we see that

$$D_f = \sum_j \left(\sum_i \frac{\partial f}{\partial \eta_i} J_{ij}\right) \frac{\partial}{\partial \eta_j},$$

which is of the general form that a differential operator has,

$$D_f = \sum_j f_j \frac{\partial}{\partial \eta_j},$$

where f_j are an arbitrary set of functions on phase space. For the Poisson bracket, the functions f_j are linear combinations of the $f_{,j}$, but $f_j \neq f_{,j}$. With this interpretation, $[f,g] = D_f g$, and $[h,[f,g]] = D_h D_f g$. Thus

$$\begin{aligned}[h,[f,g]] + [f,[g,h]] &= [h,[f,g]] - [f,[h,g]] = D_h D_f g - D_f D_h g \\ &= (D_h D_f - D_f D_h) g, \qquad (6.12)\end{aligned}$$

[3] This convention of understood summation was invented by Einstein, who called it the "greatest contribution of my life".

and we see that this combination of Poisson brackets involves the commutator of differential operators. But such a commutator is always a linear differential operator itself,

$$\begin{aligned}D_h D_f &= \sum_{ij} h_i \frac{\partial}{\partial \eta_i} f_j \frac{\partial}{\partial \eta_j} = \sum_{ij} h_i \frac{\partial f_j}{\partial \eta_i}\frac{\partial}{\partial \eta_j} + \sum_{ij} h_i f_j \frac{\partial^2}{\partial \eta_i \partial \eta_j} \\ D_f D_h &= \sum_{ij} f_j \frac{\partial}{\partial \eta_j} h_i \frac{\partial}{\partial \eta_i} = \sum_{ij} f_j \frac{\partial h_i}{\partial \eta_j}\frac{\partial}{\partial \eta_i} + \sum_{ij} h_i f_j \frac{\partial^2}{\partial \eta_i \partial \eta_j}\end{aligned}$$

so in the commutator, the second derivative terms cancel, and

$$\begin{aligned}D_h D_f - D_f D_h &= \sum_{ij} h_i \frac{\partial f_j}{\partial \eta_i}\frac{\partial}{\partial \eta_j} - \sum_{ij} f_j \frac{\partial h_i}{\partial \eta_j}\frac{\partial}{\partial \eta_i} \\ &= \sum_{ij}\left(h_i \frac{\partial f_j}{\partial \eta_i} - f_i \frac{\partial h_j}{\partial \eta_i}\right)\frac{\partial}{\partial \eta_j}.\end{aligned}$$

This is just another first order differential operator, so there are no second derivatives of g left in (6.12). In fact, the identity tells us that this combination is

$$D_h D_f - D_f D_h = D_{[h,f]} \qquad (6.13)$$

An antisymmetric product which obeys the Jacobi identity is what makes a **Lie algebra**. Lie algebras are the infinitesimal generators of Lie groups, or continuous groups, one example of which is the group of rotations $SO(3)$ which we have already considered. Notice that the "product" here is not *associative*, $[u,[v,w]] \neq [[u,v],w]$. In fact, the difference $[u,[v,w]] - [[u,v],w] = [u,[v,w]] + [w,[u,v]] = -[v,[w,u]]$ by the Jacobi identity, so the Jacobi identity replaces the law of associativity in a Lie algebra. Lie groups play a major role in quantum mechanics and quantum field theory, and their more extensive study is highly recommended for any physicist. Here we will only mention that infinitesimal rotations, represented either by the $\omega\Delta t$ or $\Omega\Delta t$ of Chapter 4, constitute the three dimensional Lie algebra of the rotation group (in three dimensions).

Recall that the rate at which a function on phase space, evaluated on the system as it evolves, changes with time is

$$\frac{df}{dt} = -[H,f] + \frac{\partial f}{\partial t}, \qquad (6.14)$$

where H is the Hamiltonian. The function $[f,g]$ on phase space also evolves that way, of course, so

$$\begin{aligned}\frac{d[f,g]}{dt} &= -[H,[f,g]] + \frac{\partial[f,g]}{\partial t}\\ &= [f,[g,H]] + [g,[H,f]] + \left[\frac{\partial f}{\partial t}, g\right] + \left[f, \frac{\partial g}{\partial t}\right]\\ &= \left[f, \left(-[H,g] + \frac{\partial g}{\partial t}\right)\right] + \left[g, \left([H,f] - \frac{\partial f}{\partial t}\right)\right]\\ &= \left[f, \frac{dg}{dt}\right] - \left[g, \frac{df}{dt}\right].\end{aligned}$$

If f and g are conserved quantities, $df/dt = dg/dt = 0$, and we have the important consequence that $d[f,g]/dt = 0$. This proves **Poisson's theorem**: The Poisson bracket of two conserved quantities is a conserved quantity.

We will now show an important theorem, known as **Liouville's theorem**, that the volume of a region of phase space is invariant under canonical transformations. This is not a volume in ordinary space, but a $2n$ dimensional volume, given by integrating the volume element $\prod_{i=1}^{2n} d\eta_i$ in the old coordinates, and by

$$\prod_{i=1}^{2n} d\zeta_i = \left|\det \frac{\partial \zeta_i}{\partial \eta_j}\right| \prod_{i=1}^{2n} d\eta_i = |\det M| \prod_{i=1}^{2n} d\eta_i$$

in the new, where we have used the fact that the change of variables requires a Jacobian in the volume element. But because $J = M \cdot J \cdot M^T$, $\det J = \det M \det J \det M^T = (\det M)^2 \det J$, and J is nonsingular, so $\det M = \pm 1$, and the volume element is unchanged.

In statistical mechanics, we generally do not know the actual state of a system, but know something about the probability that the system is in a particular region of phase space. As the transformation which maps possible values of $\eta(t_1)$ to the values into which they will evolve at time t_2 is a canonical transformation, this means that the volume of a region in phase space does not change with time, although the region itself changes. Thus the probability density, specifying the likelihood that the system is near a particular point of phase space, is invariant as we move along with the system.

6.5 Higher Differential Forms

In section 6.1 we discussed a reinterpretation of the differential df as an example of a more general differential 1-form, a map $\omega : \mathcal{M} \times \mathbb{R}^n \to \mathbb{R}$. We saw that the $\{dx_i\}$ provide a basis for these forms, so the general 1-form can be written as $\omega = \sum_i \omega_i(x)\, dx_i$. The differential df gave an example. We defined an *exact* 1-form as one which is a differential of some well-defined function f. What is the condition for a 1-form to be exact? If $\omega = \sum \omega_i dx_i$ is df, then $\omega_i = \partial f/\partial x_i = f_{,i}$, and

$$\omega_{i,j} = \frac{\partial \omega_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \omega_{j,i}.$$

Thus one *necessary* condition for ω to be exact is that the combination $\omega_{j,i} - \omega_{i,j} = 0$. We will define a 2-form to be the set of these objects which must vanish. In fact, we define a **differential k-form** to be a map

$$\omega^{(k)} : \mathcal{M} \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k \text{ times}} \to \mathbb{R}$$

which is linear in its action on each of the $\mathbb{R}^n$ and totally antisymmetric in its action on the k copies, and is a smooth function of $x \in \mathcal{M}$. At a given point, a basis of the k-forms is[4]

$$dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k} := \sum_{P \in S_k} (-1)^P dx_{i_{P1}} \otimes dx_{i_{P2}} \otimes \cdots \otimes dx_{i_{Pk}}.$$

For example, in three dimensions there are three independent 2-forms at a point, $dx_1 \wedge dx_2$, $dx_1 \wedge dx_3$, and $dx_2 \wedge dx_3$, where $dx_1 \wedge dx_2 = dx_1 \otimes dx_2 - dx_2 \otimes dx_1$, which means that, acting on $\vec{u}$ and $\vec{v}$, $dx_1 \wedge dx_2(\vec{u}, \vec{v}) = u_1 v_2 - u_2 v_1$. The product $\wedge$ is called the **wedge product** or **exterior product**, and can be extended to act between k_1- and k_2-forms so that it becomes an associative distributive product. Note that this definition of a k-form agrees, for $k = 1$,

[4]Some explanation of the mathematical symbols might be in order here. S_k is the group of permutations on k objects, and $(-1)^P$ is the **sign of the permutation** P, which is plus or minus one if the permutation can be built from an even or an odd number, respectively, of transpositions of two of the elements. The tensor product $\otimes$ of two linear operators into a field is a linear operator which acts on the product space, or in other words a bilinear operator with two arguments. Here $dx_i \otimes dx_j$ is an operator on $\mathbb{R}^n \times \mathbb{R}^n$ which maps the pair of vectors $(\vec{u}, \vec{v})$ to $u_i v_j$.

with our previous definition, and for $k = 0$ tells us a 0-form is simply a function on $\mathcal{M}$. The general expression for a k-form is

$$\omega^{(k)} = \sum_{i_1<\ldots<i_k} \omega_{i_1\ldots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

Let us consider some examples in three dimensional Euclidean[5] space E^3, where there is a correspondance we can make between vectors and 1- and 2-forms. In this discussion we will not be considering how the objects change under changes in the coordinates of E^3, to which we will return later.

$k = 0$: As always, 0-forms are simply functions, $f(x)$, $x \in E^3$.

$k = 1$: A 1-form $\omega = \sum \omega_i dx_i$ can be thought of, or associated with, a vector field $\vec{A}(x) = \sum \omega_i(x)\hat{e}_i$. Note that if $\omega = df$, $\omega_i = \partial f/\partial x_i$, so $\vec{A} = \vec{\nabla} f$. The 1-form is not actually the vector field, as it is a function that depends on two arguments, x and $\vec{v}$. If the 1-form ω is associated with the vector field $\vec{A} = \sum A_i(x)\hat{e}_i$, acting on the vector field $\vec{B} = \sum B_i(x)\hat{e}_i$, we have $\omega(x, \vec{B}) = \sum A_i B_i = \vec{A} \cdot \vec{B}$.

$k = 2$: A general two form is a sum over the three independent wedge products with independent functions $B_{12}(x), B_{13}(x), B_{23}(x)$. Let us extend the definition of B_{ij} to make it an antisymmetric matrix, so

$$B = \sum_{i<j} B_{ij} dx_i \wedge dx_j = \sum_{i,j} B_{ij} dx_i \otimes dx_j.$$

As we did for the angular velocity matrix Ω in (4.2), we can condense the information in the antisymmetric matrix B_{ij} into a vector field $\vec{B} = \sum B_i \hat{e}_i$, with $B_{ij} = \sum \epsilon_{ijk} B_k$. Note that this step requires that we are working in E^3 rather than some other dimension. Thus $B = \sum_{ijk} \epsilon_{ijk} B_k dx_i \otimes dx_j$. Also note $\frac{1}{2}\sum_{ij} \epsilon_{ijk} B_{ij} = \frac{1}{2}\sum_{ij\ell} \epsilon_{ijk}\epsilon_{ij\ell} B_\ell = B_k$. The 2-form B takes two vector arguments, so acting on vectors $\vec{A} = \sum A_i(x)\hat{e}_i$ and $\vec{C} = \sum C_i(x)\hat{e}_i$, we have $B(x, \vec{A}, \vec{B}) = \sum B_{ij} A_i C_j = \sum \epsilon_{ijk} B_k A_i C_j = (\vec{A}(x) \times \vec{C}(x)) \cdot \vec{B}(x)$.

[5]Forms are especially useful in discussing more general manifolds, such as occur in general relativity. Then one must distinguish between covariant and contravariant vectors, a complication we avoid here by treating only Euclidean space.

$k = 3$: There is only one basis 3-form available in three dimensions, $dx_1 \wedge dx_2 \wedge dx_3$. Any other 3-form is proportional to this one, though the proportionality can be a function of $\{x_i\}$. In particular $dx_i \wedge dx_j \wedge dx_k = \epsilon_{ijk} dx_1 \wedge dx_2 \wedge dx_3$. The most general 3-form C is simply specified by an ordinary function $C(x)$, which multiplies $dx_1 \wedge dx_2 \wedge dx_3$.

Having established, in three dimensions, a correspondance between vectors and 1- and 2-forms, and between functions and 0- and 3-forms, we can ask to what the wedge product corresponds in terms of these vectors. If $\vec{A}$ and $\vec{C}$ are two vectors corresponding to the 1-forms $A = \sum A_i dx_i$ and $C = \sum C_i dx_i$, and if $B = A \wedge C$, then

$$B = \sum_{ij} A_i C_j dx_i \wedge dx_j = \sum_{ij} (A_i C_j - A_j C_i) dx_i \otimes dx_j = \sum_{ij} B_{ij} dx_i \otimes dx_j,$$

so $B_{ij} = A_i C_j - A_j C_i$, and

$$B_k = \frac{1}{2}\sum \epsilon_{kij} B_{ij} = \frac{1}{2}\sum \epsilon_{kij} A_i C_j - \frac{1}{2}\sum \epsilon_{kij} A_j C_i = \sum \epsilon_{kij} A_i C_j,$$

so

$$\vec{B} = \vec{A} \times \vec{C},$$

and the wedge product of two 1-forms is the cross product of their vectors.

If A is a 1-form and B is a 2-form, the wedge product $C = A \wedge B = C(x) dx_1 \wedge dx_2 \wedge dx_3$ is given by

$$\begin{aligned}
C &= A \wedge B = \sum_i \sum_{j<k} A_i \underbrace{B_{jk}}_{\epsilon_{jk\ell} B_\ell} \underbrace{dx_i \wedge dx_j \wedge dx_k}_{\epsilon_{ijk} dx_1 \wedge dx_2 \wedge dx_3} \\
&= \sum_{i\ell} A_i B_\ell \sum_{j<k} \underbrace{\epsilon_{jk\ell}\epsilon_{ijk}}_{\text{symmetric under } j \leftrightarrow k} dx_1 \wedge dx_2 \wedge dx_3 \\
&= \frac{1}{2}\sum_{i\ell} A_i B_\ell \sum_{jk} \epsilon_{jk\ell}\epsilon_{ijk} dx_1 \wedge dx_2 \wedge dx_3 = \sum_{i\ell} A_i B_\ell \delta_{i\ell} dx_1 \wedge dx_2 \wedge dx_3 \\
&= \vec{A} \cdot \vec{B}\, dx_1 \wedge dx_2 \wedge dx_3,
\end{aligned}$$

so we see that the wedge product of a 1-form and a 2-form gives the dot product of their vectors.

If A and B are both 2-forms, the wedge product $C = A \wedge B$ must be a 4-form, but there cannot be an antisymmetric function of four dx_i's in three dimensions, so $C = 0$.

The exterior derivative

We defined the differential of a function f, which we now call a 0-form, giving a 1-form $df = \sum f_{,i} dx_i$. Now we want to generalize the notion of differential so that d can act on k-forms for arbitrary k. This generalized differential

$$d : k\text{-forms} \to (k+1)\text{-forms}$$

is called the **exterior derivative**. It is defined to be linear and to act on one term in the sum over basis elements by

$$\begin{aligned} d\left(f_{i_1\dots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}\right) &= \left(df_{i_1\dots i_k}(x)\right) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= \sum_j f_{i_1\dots i_k,j} dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \end{aligned}$$

Clearly some examples are called for, so let us look again at three dimensional Euclidean space.

$k = 0$: For a 0-form f, $df = \sum f_{,i} dx_i$, as we defined earlier. In terms of vectors, $df \sim \vec{\nabla} f$.

$k = 1$: For a 1-form $\omega = \sum \omega_i\, dx_i$, $\; d\omega = \sum_i d\omega_i \wedge dx_i = \sum_{ij} \omega_{i,j} dx_j \wedge dx_i = \sum_{ij} \left(\omega_{j,i} - \omega_{i,j}\right) dx_i \otimes dx_j$, corresponding to a two form with $B_{ij} = \omega_{j,i} - \omega_{i,j}$. These B_{ij} are exactly the things which must vanish if ω is to be exact. In three dimensional Euclidean space, we have a vector $\vec{B}$ with components $B_k = \frac{1}{2} \sum \epsilon_{kij} (\omega_{j,i} - \omega_{i,j}) = \sum \epsilon_{kij} \partial_i \omega_j = (\vec{\nabla} \times \vec{\omega})_k$, so here the exterior derivative of a 1-form gives a curl, $\vec{B} = \vec{\nabla} \times \vec{\omega}$.

$k = 2$: On a two form $B = \sum_{i<j} B_{ij} dx_i \wedge dx_j$, the exterior derivative gives a 3-form $C = dB = \sum_k \sum_{i<j} B_{ij,k} dx_k \wedge dx_i \wedge dx_j$. In three-dimensional Euclidean space, this reduces to

$$C = \sum_{k\ell} \sum_{i<j} \left(\partial_k \epsilon_{ij\ell} B_\ell\right) \epsilon_{kij} dx_1 \wedge dx_2 \wedge dx_3 = \sum_k \partial_k B_k dx_1 \wedge dx_2 \wedge dx_3,$$

so $C(x) = \vec{\nabla} \cdot \vec{B}$, and the exterior derivative on a 2-form gives the divergence of the corresponding vector.

$k = 3$: If C is a 3-form, dC is a 4-form. In three dimensions there cannot be any 4-forms, so $dC = 0$ for all such forms.

We can summarize the action of the exterior derivative in three dimensions in this diagram:

$$f \xrightarrow[\nabla f]{d} \omega^{(1)} \sim \vec{A} \xrightarrow[\nabla \times A]{d} \omega^{(2)} \sim \vec{B} \xrightarrow[\nabla \cdot B]{d} \omega^{(3)} \xrightarrow{d} 0$$

Now that we have d operating on all k-forms, we can ask what happens if we apply it twice. Looking first in three dimenions, on a 0-form we get $d^2 f = dA$ for $\vec{A} \sim \nabla f$, and $dA \sim \nabla \times A$, so $d^2 f \sim \nabla \times \nabla f$. But the curl of a gradient is zero, so $d^2 = 0$ in this case. On a one form $d^2 A = dB$, $\vec{B} \sim \nabla \times \vec{A}$ and $dB \sim \nabla \cdot B = \nabla \cdot (\nabla \times \vec{A})$. Now we have the divergence of a curl, which is also zero. For higher forms in three dimensions we can only get zero because the degree of the form would be greater than three. Thus we have a strong hint that d^2 might vanish in general. To verify this, we apply d^2 to $\omega^{(k)} = \sum \omega_{i_1\dots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k}$. Then

$$\begin{aligned} d\omega &= \sum_j \sum_{i_1<i_2<\cdots<i_k} \left(\partial_j \omega_{i_1\dots i_k}\right) dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ d(d\omega) &= \sum_{\ell j} \sum_{i_1<i_2<\cdots<i_k} \Big(\underbrace{\partial_\ell \partial_j}_{\text{symmetric}} \omega_{i_1\dots i_k}\Big) \underbrace{dx_\ell \wedge dx_j}_{\text{antisymmetric}} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= 0. \end{aligned}$$

This is a very important result. A k-form which is the exterior derivative of some $(k-1)$-form is called **exact**, while a k-form whose exterior derivative vanishes is called **closed**, and we have just proven that all exact k-forms are closed.

The converse is a more subtle question. In general, there are k-forms which are closed but not exact, given by harmonic functions on the manifold $\mathcal{M}$, which form what is known as the cohomology of $\mathcal{M}$. This has to do with global properties of the space, however, and locally every closed form can be written as an exact one.[6] The precisely stated theorem, known as **Poincaré's**

[6] An example may be useful. In two dimensions, irrotational vortex flow can be represented by the 1-form $\omega = -yr^{-2}dx + xr^{-2}dy$, which satisfies $d\omega = 0$ wherever it is well defined, but it is not well defined at the origin. Locally, we can write $\omega = d\theta$, where θ is the polar coordinate. But θ is not, strictly speaking, a function on the plane, even on the plane with the origin removed, because it is not single-valued. It is a well defined function on the plane with a half axis removed, which leaves a simply-connected region, a region with no holes. In fact, this is the general condition for the exactness of a 1-form — a closed 1-form on a simply connected manifold is exact.

Lemma, is that if ω is a closed k-form on a coordinate neighborhood U of a manifold M, and if U is contractible to a point, then ω is exact on U. We will ignore the possibility of global obstructions and assume that we can write closed k-forms in terms of an exterior derivative acting on a $(k-1)$-form.

Coordinate independence of k-forms

We have introduced forms in a way which makes them appear dependent on the coordinates x_i used to describe the space $\mathcal{M}$. This is not what we want at all[7]. We want to be able to describe physical quantities that have intrinsic meaning independent of a coordinate system. If we are presented with another set of coordinates y_j describing the same physical space, the points in this space set up a mapping, ideally an isomorphism, from one coordinate system to the other, $\vec{y}=\vec{y}(\vec{x})$. If a function represents a physical field independent of coordinates, the actual function $f(x)$ used with the x coordinates must be replaced by another function $\tilde{f}(y)$ when using the y coordinates. That they both describe the physical value at a given physical point requires $f(x)=\tilde{f}(y)$ when $y=y(x)$, or more precisely[8] $f(x)=\tilde{f}(y(x))$. This associated function and coordinate system is called a scalar field.

If we think of the differential df as the change in f corresponding to an infinitesimal change dx, then clearly $d\tilde{f}$ is the same thing in different coordinates, provided we understand the dy_i to represent the same physical displacement as dx does. That means

$$dy_k = \sum_j \frac{\partial y_k}{\partial x_j} dx_j.$$

As $f(x)=\tilde{f}(y(x))$ and $\tilde{f}(y)=f(x(y))$, the chain rule gives

$$\frac{\partial f}{\partial x_i} = \sum_j \frac{\partial \tilde{f}}{\partial y_j}\frac{\partial y_j}{\partial x_i}, \qquad \frac{\partial \tilde{f}}{\partial y_j} = \sum_i \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial y_j},$$

[7]Indeed, most mathematical texts will first define an abstract notion of a vector in the tangent space as a directional derivative operator, specified by equivalence classes of parameterized paths on $\mathcal{M}$. Then 1-forms are defined as duals to these vectors. In the first step any coordinatization of $\mathcal{M}$ is tied to the corresponding basis of the vector space $\mathbb{R}^n$. While this provides an elegant coordinate-independent way of defining the forms, the abstract nature of this definition of vectors can be unsettling to a physicist.

[8]More elegantly, giving the map $x \to y$ the name ϕ, so $y=\phi(x)$, we can state the relation as $f=\tilde{f}\circ\phi$.

so

$$\begin{aligned} d\tilde{f} &= \sum_k \frac{\partial \tilde{f}}{\partial y_k} dy_k = \sum_{ijk} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial y_k}\frac{\partial y_k}{\partial x_j} dx_j \\ &= \sum_{ij} \frac{\partial f}{\partial x_i}\delta_{ij} dx_j = \sum_i f_{,i} dx_i = df. \end{aligned}$$

We impose this transformation law in general on the coefficients in our k-forms, to make the k-form invariant, which means that the coefficients are covariant,

$$\begin{aligned} \tilde{\omega}_j &= \sum_i \frac{\partial x_i}{\partial y_j}\omega_i \\ \tilde{\omega}_{j_1\ldots j_k} &= \sum_{i_1,i_2,\ldots,i_k} \left(\prod_{\ell=1}^k \frac{\partial x_{i_\ell}}{\partial y_{j_l}}\right)\omega_{i_1\ldots i_k}. \end{aligned}$$

Integration of k-forms

Suppose we have a k-dimensional smooth "surface" S in $\mathcal{M}$, parameterized by coordinates $(u_1,\cdots,u_k)$. We define the integral of a k-form

$$\omega^{(k)} = \sum_{i_1<\ldots<i_k} \omega_{i_1\ldots i_k} dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

over S by

$$\int_S \omega^{(k)} = \int \sum_{i_1,i_2,\ldots,i_k} \omega_{i_1\ldots i_k}(x(u)) \left(\prod_{\ell=1}^k \frac{\partial x_{i_\ell}}{\partial u_\ell}\right) du_1 du_2\cdots du_k.$$

We had better give some examples. For $k=1$, the "surface" is actually a path $\Gamma : u \mapsto x(u)$, and

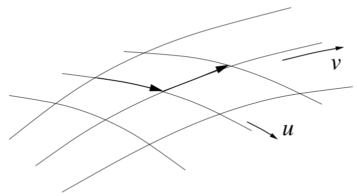
$$\int_\Gamma \sum \omega_i dx_i = \int_{u_{\min}}^{u_{\max}} \sum \omega_i(x(u)) \frac{\partial x_i}{\partial u} du,$$

which seems obvious. In vector notation this is $\int_\Gamma \vec{A}\cdot d\vec{r}$, the path integral of the vector $\vec{A}$.

For $k=2$,

$$\int_S \omega^{(2)} = \int B_{ij} \frac{\partial x_i}{\partial u}\frac{\partial x_j}{\partial v} du dv.$$

In three dimensions, the parallelogram which is the image of the rectangle $[u, u+du] \times [v, v+dv]$ has edges $(\partial\vec{x}/\partial u)du$ and $(\partial\vec{x}/\partial v)dv$, which has an area equal to the magnitude of



$$“d\vec{S}” = \left(\frac{\partial\vec{x}}{\partial u} \times \frac{\partial\vec{x}}{\partial v}\right) du dv$$

and a normal in the direction of “$d\vec{S}$”. Writing B_{ij} in terms of the corresponding vector $\vec{B}$, $B_{ij} = \epsilon_{ijk}B_k$, so

$$\begin{aligned}\int_S \omega^{(2)} &= \int_S \epsilon_{ijk}B_k \left(\frac{\partial\vec{x}}{\partial u}\right)_i \left(\frac{\partial\vec{x}}{\partial v}\right)_j du dv \\ &= \int_S B_k \left(\frac{\partial\vec{x}}{\partial u} \times \frac{\partial\vec{x}}{\partial v}\right)_k du dv = \int_S \vec{B}\cdot d\vec{S},\end{aligned}$$

so $\int \omega^{(2)}$ gives the flux of $\vec{B}$ through the surface.

Similarly for $k = 3$ in three dimensions,

$$\sum \epsilon_{ijk} \left(\frac{\partial\vec{x}}{\partial u}\right)_i \left(\frac{\partial\vec{x}}{\partial v}\right)_j \left(\frac{\partial\vec{x}}{\partial w}\right)_k du dv dw$$

is the volume of the parallelopiped which is the image of $[u, u+du] \times [v, v+dv] \times [w, w+dw]$. As $\omega_{ijk} = \omega_{123}\epsilon_{ijk}$, this is exactly what appears:

$$\int \omega^{(3)} = \int \sum \epsilon_{ijk}\omega_{123}\frac{\partial x_i}{\partial u}\frac{\partial x_j}{\partial v}\frac{\partial x_k}{\partial w} du dv dw = \int \omega_{123}(x) dV.$$

Notice that we have only defined the integration of k-forms over submanifolds of dimension k, not over other-dimensional submanifolds. These are the only integrals which have coordinate invariant meanings. Also note that the integrals do not depend on how the surface is coordinatized.

We state[9] a marvelous theorem, special cases of which you have seen often before, known as **Stokes' Theorem**. Let C be a k-dimensional submanifold

9 For a proof and for a more precise explanation of its meaning, we refer the reader to the mathematical literature. In particular [14] and [3] are advanced calculus texts which give elementary discussions in Euclidean 3-dimensional space. A more general treatment is (possibly???) given in [16].

of $\mathcal{M}$, with ∂C its boundary. Let ω be a $(k-1)$-form. Then Stokes' theorem says

$$\int_C d\omega = \int_{\partial C} \omega. \qquad (6.15)$$

This elegant jewel is actually familiar in several contexts in three dimensions. If $k = 2$, C is a surface, usually called S, bounded by a closed path $\Gamma = \partial S$. If ω is a 1-form associated with $\vec{A}$, then $\int_\Gamma \omega = \int_\Gamma \vec{A}\cdot d\vec{\ell}$. Now $d\omega$ is the 2-form $\sim \vec{\nabla}\times\vec{A}$, and $\int_S d\omega = \int_S \left(\vec{\nabla}\times\vec{A}\right)\cdot d\vec{S}$, so we see that this Stokes' theorem includes the one we first learned by that name. But it also includes other possibilities. We can try $k = 3$, where $C = V$ is a volume with surface $S = \partial V$. Then if $\omega \sim \vec{B}$ is a two form, $\int_S \omega = \int_S \vec{B}\cdot d\vec{S}$, while $d\omega \sim \vec{\nabla}\cdot\vec{B}$, so $\int_V d\omega = \int \vec{\nabla}\cdot\vec{B} dV$, so here Stokes' general theorem gives Gauss's theorem. Finally, we could consider $k = 1$, $C = \Gamma$, which has a boundary ∂C consisting of two points, say A and B. Our 0-form $\omega = f$ is a function, and Stokes' theorem gives[10] $\int_\Gamma df = f(B) - f(A)$, the “fundamental theorem of calculus”.

6.6 The natural symplectic 2-form

We now turn our attention back to phase space, with a set of canonical coordinates (q_i, p_i). Using these coordinates we can define a particular 1-form $\omega_1 = \sum_i p_i dq_i$. For a point transformation $Q_i = Q_i(q_1, \ldots, q_n, t)$ we may use the same Lagrangian, reexpressed in the new variables, of course. Here the Q_i are independent of the velocities $\dot{q}_j$, so on phase space[11] $dQ_i = \sum_j(\partial Q_i/\partial q_j)dq_j$. The new velocities are given by

$$\dot{Q}_i = \sum_j \frac{\partial Q_i}{\partial q_j}\dot{q}_j + \frac{\partial Q_i}{\partial t}, \qquad \text{so} \qquad \frac{\partial\dot{Q}_i}{\partial\dot{q}_j} = \frac{\partial Q_i}{\partial q_j}.$$

10 Note that there is a direction associated with the boundary, which is induced by a direction associated with C itself. This gives an ambiguity in what we have stated, for example how the direction of an open surface induces a direction on the closed loop which bounds it. Changing this direction would clearly reverse the sign of $\int \vec{A}\cdot d\vec{\ell}$. We have not worried about this ambiguity, but we cannot avoid noticing the appearence of the sign in this last example.

11 We have not included a term $\frac{\partial Q_i}{\partial t}dt$ which would be necessary if we were considering a form in the $2n+1$ dimensional extended phase space which includes time as one of its coordinates.

Thus the old canonical momenta,

$$p_i = \left.\frac{\partial L(q,\dot q,t)}{\partial \dot q_i}\right|_{q,t} = \sum_j \left.\frac{\partial L(Q,\dot Q,t)}{\partial \dot Q_j}\right|_{q,t} \left.\frac{\partial \dot Q_j}{\partial \dot q_i}\right|_{q,t} = \sum_j P_j \frac{\partial Q_j}{\partial q_i}.$$

Thus the form ω_1 may be written

$$\omega_1 = \sum_i \sum_j P_j \frac{\partial Q_j}{\partial q_i} dq_i = \sum_j P_j dQ_j,$$

so the form of ω_1 is invariant under **point** transformations. This is too limited, however, for our current goals of considering general canonical transformations on phase space, under which ω_1 will not be invariant. However, its exterior derivative

$$\omega_2 := d\omega_1 = \sum_i dp_i \wedge dq_i$$

is invariant under all canonical transformations, as we shall show momentarily. This makes it special, the **natural symplectic structure** on phase space. We can reexpress ω_2 in terms of our combined coordinate notation η_i, because

$$-\sum_{i<j} J_{ij} d\eta_i \wedge d\eta_j = -\sum_i dq_i \wedge dp_i = \sum_i dp_i \wedge dq_i = \omega_2.$$

We must now show that the natural symplectic structure is indeed form invariant under canonical transformation. Thus if Q_i, P_i are a new set of canonical coordinates, combined into ζ_j, we expect the corresponding object formed from them, $\omega_2' = -\sum_{ij} J_{ij} d\zeta_i \otimes d\zeta_j$, to reduce to the same 2-form, ω_2. We first note that

$$d\zeta_i = \sum_j \frac{\partial \zeta_i}{\partial \eta_j} d\eta_j = \sum_j M_{ij} d\eta_j,$$

with the same Jacobian matrix M we met in section 6.3. Thus

$$\begin{aligned}
\omega_2' &= -\sum_{ij} J_{ij} d\zeta_i \otimes d\zeta_j = -\sum_{ij} J_{ij} \sum_k M_{ik} d\eta_k \otimes \sum_\ell M_{j\ell} d\eta_\ell \\
&= -\sum_{k\ell} \left(M^T \cdot J \cdot M\right)_{k\ell} d\eta_k \otimes d\eta_\ell.
\end{aligned}$$

Things will work out if we can show $M^T \cdot J \cdot M = J$, whereas what we know for canonical transformations from Eq. (6.4) is that $M \cdot J \cdot M^T = J$. We also know M is invertible and that $J^2 = -1$, so if we multiply this known equation from the left by $-J \cdot M^{-1}$ and from the right by $J \cdot M$, we learn that

$$\begin{aligned}
-J \cdot M^{-1} \cdot M \cdot J \cdot M^T \cdot J \cdot M &= -J \cdot M^{-1} \cdot J \cdot J \cdot M \\
&= J \cdot M^{-1} \cdot M = J \\
= -J \cdot J \cdot M^T \cdot J \cdot M &= M^T \cdot J \cdot M,
\end{aligned}$$

which is what we wanted to prove. Thus we have shown that the 2-form ω_2 is form-invariant under canonical transformations, and deserves its name.

One important property of the 2-form ω_2 on phase space is that it is **non-degenerate**. A 2-form has two slots to insert vectors — inserting one leaves a 1-form. Non-degenerate means there is no non-zero vector $\vec v$ on phase space such that $\omega_2(\cdot, \vec v) = 0$, that is, such that $\omega_2(\vec u, \vec v) = 0$, for all $\vec u$ on phase space. This follows simply from the fact that the matrix J_{ij} is non-singular.

Extended phase space

One way of looking at the evolution of a system is in phase space, where a given system corresponds to a point moving with time, and the general equations of motion corresponds to a velocity field. Another way is to consider **extended phase space**, a $2n+1$ dimensional space with coordinates (q_i, p_i, t), for which a system's motion is a path, monotone in t. By the modified Hamilton's principle, the path of a system in this space is an extremum of the action $I = \int_{t_i}^{t_f} \sum p_i dq_i - H(q,p,t)dt$, which is the integral of the one-form

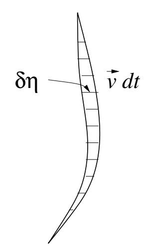
$$\omega_3 = \sum p_i dq_i - H(q,p,t)dt.$$

The exterior derivative of this form involves the symplectic structure, ω_2, as $d\omega_3 = \omega_2 - dH \wedge dt$. The 2-form ω_2 on phase space is nondegenerate, and every vector in phase space is also in extended phase space. On such a vector, on which dt gives zero, the extra term gives only something in the dt direction, so there are still no vectors in this subspace which are annihilated by $d\omega_3$. Thus there is at most one direction in extended phase space which is annihilated by $d\omega_3$. But any 2-form in an odd number of dimensions must annihilate some vector, because in a given basis it corresponds to an antisymmetric matrix B_{ij}, and in an odd number of dimensions $\det B = \det B^T = \det(-B) = (-1)^{2n+1} \det B = -\det B$, so $\det B = 0$ and the matrix is singular, annihilating some vector ξ. In fact, for $d\omega_3$ this annihilated vector ξ is the tangent to the path the system takes through extended phase space.

One way to see this is to simply work out what $d\omega_3$ is and apply it to the vector ξ, which is proportional to $\vec{v} = (\dot{q}_i, \dot{p}_i, 1)$. So we wish to show $d\omega_3(\cdot, \vec{v}) = 0$. Evaluating

$$\begin{aligned}
\sum dp_i \wedge dq_i(\cdot, \vec{v}) &= \sum dp_i\, dq_i(\vec{v}) - \sum dq_i\, dp_i(\vec{v}) = \sum dp_i \dot{q}_i - \sum dq_i \dot{p}_i \\
dH \wedge dt(\cdot, \vec{v}) &= dH\, dt(\vec{v}) - dt\, dH(\vec{v}) \\
&= \left(\sum \frac{\partial H}{\partial q_i} dq_i + \sum \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt\right) 1 \\
&\quad -dt \left(\sum \dot{q}_i \frac{\partial H}{\partial q_i} + \sum \dot{p}_i \frac{\partial H}{\partial p_i} + \frac{\partial H}{\partial t}\right) \\
&= \sum \frac{\partial H}{\partial q_i} dq_i + \sum \frac{\partial H}{\partial p_i} dp_i - dt \sum \left(\dot{q}_i \frac{\partial H}{\partial q_i} + \dot{p}_i \frac{\partial H}{\partial p_i}\right) \\
d\omega_3(\cdot, \vec{v}) &= \sum \left(\dot{q}_i - \frac{\partial H}{\partial p_i}\right) dp_i - \left(\dot{p}_i + \frac{\partial H}{\partial q_i}\right) dq_i \\
&\quad + \sum \left(\dot{q}_i \frac{\partial H}{\partial q_i} + \dot{p}_i \frac{\partial H}{\partial p_i}\right) dt \\
&= 0
\end{aligned}$$

where the vanishing is due to the Hamilton equations of motion.



There is a more abstract way of understanding why $d\omega_3(\cdot, \vec{v})$ vanishes, from the modified Hamilton's principle, which states that if the path taken were infinitesimally varied from the physical path, there would be no change in the action. But this change is the integral of ω_3 along a loop, forwards in time along the first trajectory and backwards along the second. From Stokes' theorem this means the integral of $d\omega_3$ over a surface connecting these two paths vanishes. But this surface is a sum over infinitesimal parallelograms one side of which is $\vec{v}\,\Delta t$ and the other side of which[12] is $(\delta\vec{q}(t), \delta\vec{p}(t), 0)$. As this latter vector is an arbitrary function of t, each parallelogram must independently give 0, so that its contribution to the integral, $d\omega_3((\delta\vec{q}, \delta\vec{p}, 0), \vec{v})\Delta t =$

[12]It is slightly more elegant to consider the path parameterized independently of time, and consider arbitrary variations $(\delta q, \delta p, \delta t)$, because the integral involved in the action, being the integral of a 1-form, is independent of the parameterization. With this approach we find immediately that $d\omega_3(\cdot, \vec{v})$ vanishes on all vectors.

0. In addition, $d\omega_3(\vec{v}, \vec{v}) = 0$, of course, so $d\omega_3(\cdot, \vec{v})$ vanishes on a complete basis of vectors and is therefore zero.

6.6.1 Generating Functions

Consider a canonical transformation $(q, p) \to (Q, P)$, and the two 1-forms $\omega_1 = \sum_i p_i dq_i$ and $\omega_1' = \sum_i P_i dQ_i$. We have mentioned that the difference of these will not vanish in general, but the exterior derivative of this difference, $d(\omega_1 - \omega_1') = \omega_2 - \omega_2' = 0$, so $\omega_1 - \omega_1'$ is an closed 1-form. Thus it is exact[13], and there must be a function F on phase space such that $\omega_1 - \omega_1' = dF$. We call F the **generating function of the canonical transformation**[14] If the transformation $(q, p) \to (Q, P)$ is such that the old q's alone, without information about the old p's, do not impose any restrictions on the new Q's, then the dq and dQ are independent, and we can use q and Q to parameterize phase space[15]. Then knowledge of the function $F(q, Q)$ determines the transformation, as

$$\begin{aligned}
\omega_1 - \omega_1' &= \sum_i (p_i dq_i - P_i dQ_i) = dF = \sum_i \left(\left.\frac{\partial F}{\partial q_i}\right|_Q dq_i + \left.\frac{\partial F}{\partial Q_i}\right|_q dQ_i \right) \\
&\implies p_i = \left.\frac{\partial F}{\partial q_i}\right|_Q, \quad -P_i = \left.\frac{\partial F}{\partial Q_i}\right|_q .
\end{aligned}$$

If the canonical transformation depends on time, the function F will also depend on time. Now if we consider the motion in extended phase space, we know the phase trajectory that the system takes through extended phase space is determined by Hamilton's equations, which could be written in any set of canonical coordinates, so in particular there is some Hamiltonian $K(Q, P, t)$ such that the tangent to the phase trajectory, $\vec{v}$, is annihilated by $d\omega_3'$, where $\omega_3' = \sum P_i dQ_i - K(Q, P, t)dt$. Now in general knowing that two 2-forms both annihilate the same vector would not be sufficient to identify them, but in this case we also know that restricting $d\omega_3$ and $d\omega_3'$ to their action on the $dt = 0$ subspace gives the same 2-form ω_2. That is to say, if

[13]We are assuming phase space is simply connected, or else we are ignoring any complications which might ensue from F not being globally well defined.

[14] This is not an infinitesimal generator in the sense we have in Lie algebras — this generates a finite canonical transformation for finite F.

[15]Note that this is the opposite extreme from a point transformation, which is a canonical transformation for which the Q's depend *only* on the q's, independent of the p's.

$\vec{u}$ and $\vec{u}'$ are two vectors with time components zero, we know that $(d\omega_3 - d\omega_3')(\vec{u}, \vec{u}') = 0$. Any vector can be expressed as a multiple of $\vec{v}$ and some vector $\vec{u}$ with time component zero, and as both $d\omega_3$ and $d\omega_3'$ annihilate $\vec{v}$, we see that $d\omega_3 - d\omega_3'$ vanishes on all pairs of vectors, and is therefore zero. Thus $\omega_3 - \omega_3'$ is a closed 1-form, which must be at least locally exact, and indeed $\omega_3 - \omega_3' = dF$, where F is the generating function we found above[16]. Thus $dF = \sum pdq - \sum PdQ + (K - H)dt$, or

$$K = H + \frac{\partial F}{\partial t}.$$

The function $F(q, Q, t)$ is what Goldstein calls F_1. The existence of F as a function on extended phase space holds even if the Q and q are not independent, but in this case F will need to be expressed as a function of other coordinates. Suppose the new P's and the old q's are independent, so we can write $F(q, P, t)$. Then define $F_2 = \sum Q_i P_i + F$. Then

$$\begin{aligned} dF_2 &= \sum Q_i\, dP_i + \sum P_i\, dQ_i + \sum p_i\, dq_i - \sum P_i\, dQ_i + (K - H)dt \\ &= \sum Q_i\, dP_i + \sum p_i\, dq_i + (K - H)dt, \end{aligned}$$

so

$$Q_i = \frac{\partial F_2}{\partial P_i}, \qquad p_i = \frac{\partial F_2}{\partial q_i}, \qquad K(Q, P, t) = H(q, p, t) + \frac{\partial F_2}{\partial t}.$$

The generating function can be a function of old momenta rather than the old coordinates. Making one choice for the old coordinates and one for the new, there are four kinds of generating functions as described by Goldstein. Let us consider some examples. The function $F_1 = \sum_i q_i Q_i$ generates an interchange of p and q,

$$Q_i = p_i, \qquad P_i = -q_i,$$

which leaves the Hamiltonian unchanged. We saw this clearly leaves the form of Hamilton's equations unchanged. An interesting generator of the second type is $F_2 = \sum_i \lambda_i q_i P_i$, which gives $Q_i = \lambda_i q_i$, $P_i = \lambda_i^{-1} p_i$, a simple change in scale of the coordinates with a corresponding inverse scale change

[16] From its definition in that context, we found that in phase space, $dF = \omega_1 - \omega_1'$, which is the part of $\omega_3 - \omega_3'$ not in the time direction. Thus if $\omega_3 - \omega_3' = dF'$ for some other function F', we know $dF' - dF = (K' - K)dt$ for some new Hamiltonian function $K'(Q, P, t)$, so this corresponds to an ambiguity in K.

in momenta to allow $[Q_i, P_j] = \delta_{ij}$ to remain unchanged. This also doesn't change H. For $\lambda = 1$, this is the identity transformation, for which $F = 0$, of course.

Placing point transformations in this language provides another example. For a point transformation, $Q_i = f_i(q_1, \ldots, q_n, t)$, which is what one gets with a generating function

$$F_2 = \sum_i f_i(q_1, \ldots, q_n, t) P_i.$$

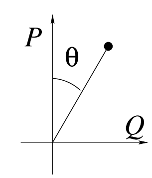
Note that

$$p_i = \frac{\partial F_2}{\partial q_i} = \sum_j \frac{\partial f_j}{\partial q_i} P_j$$

is at any point $\vec{q}$ a linear transformation of the momenta, required to preserve the canonical Poisson bracket, but this transformation is $\vec{q}$ dependent, so while $\vec{Q}$ is a function of $\vec{q}$ and t only, independent of $\vec{p}$, $\vec{P}(q, p, t)$ will in general have a nontrivial dependence on coordinates as well as a linear dependence on the old momenta.

For a harmonic oscillator, a simple scaling gives

$$H = \frac{p^2}{2m} + \frac{k}{2} q^2 = \frac{1}{2}\sqrt{k/m}\left(P^2 + Q^2\right),$$



where $Q = (km)^{1/4} q$, $P = (km)^{-1/4} p$. In this form, thinking of phase space as just some two-dimensional space, we seem to be encouraged to consider a second canonical transformation $Q, P \xrightarrow[F_1]{} \theta, \mathcal{P}$, generated by $F_1(Q, \theta)$, to a new, polar, coordinate system with $\theta = \tan^{-1} Q/P$ as the new coordinate, and we might hope to have the radial coordinate related to the new momentum, $\mathcal{P} = -\left.\partial F_1/\partial\theta\right|_Q$. As $P = \left.\partial F_1/\partial Q\right|_\theta$ is also $Q \cot\theta$, we can take $F_1 = \frac{1}{2} Q^2 \cot\theta$, so

$$\mathcal{P} = -\frac{1}{2} Q^2(-\csc^2\theta) = \frac{1}{2} Q^2(1 + P^2/Q^2) = \frac{1}{2}(Q^2 + P^2) = H/\omega.$$

Note as F_1 is not time dependent, $K = H$ and is independent of θ, which is therefore an ignorable coordinate, so its conjugate momentum $\mathcal{P}$ is conserved. Of course $\mathcal{P}$ differs from the conserved Hamiltonian H only by the factor $\omega = \sqrt{k/m}$, so this is not unexpected. With H now linear in the new momentum $\mathcal{P}$, the conjugate coordinate θ grows linearly with time at the fixed rate $\dot{\theta} = \partial H/\partial\mathcal{P} = \omega$.

Infinitesimal generators, redux

Let us return to the infinitesimal canonical transformation

$$\zeta_i = \eta_i + \epsilon g_i(\eta_j).$$

$M_{ij} = \partial\zeta_i/\partial\eta_j = \delta_{ij} + \epsilon\partial g_i/\partial\eta_j$ needs to be symplectic, and so $G_{ij} = \partial g_i/\partial\eta_j$ satisfies the appropriate condition for the generator of a symplectic matrix, $G \cdot J = -J \cdot G^T$. For the generator of the canonical transformation, we need a perturbation of the generator for the identity transformation, which can't be in F_1 form (as (q, Q) are not independent), but is easily done in F_2 form, $F_2(q, P) = \sum_i q_i P_i + \epsilon G(q, P, t)$, with $p_i = \partial F_2/\partial q_i = P_i + \epsilon\partial G/\partial q_i$, $Q_i = \partial F_2/\partial P_i = q_i + \epsilon\partial G/\partial P_i$, or

$$\zeta = \begin{pmatrix} Q_i \\ P_i \end{pmatrix} = \begin{pmatrix} q_i \\ p_i \end{pmatrix} + \epsilon \begin{pmatrix} 0 & 1\!\!1 \\ -1\!\!1 & 0 \end{pmatrix} \begin{pmatrix} \partial G/\partial q_i \\ \partial G/\partial p_i \end{pmatrix} = \eta + \epsilon J \cdot \nabla G,$$

where we have ignored higher order terms in ϵ in inverting the $q \to Q$ relation and in replacing $\partial G/\partial P_i$ with $\partial G/\partial p_i$.

The change due to the infinitesimal transformation may be written in terms of Poisson bracket with the coordinates themselves:

$$\delta\eta = \zeta - \eta = \epsilon J \cdot \nabla G = \epsilon[\eta, G].$$

In the case of an infinitesimal transformation due to time evolution, the small parameter can be taken to be Δt, and $\delta\eta = \Delta t\,\dot\eta = \Delta t[H, \eta]$, so we see that the Hamiltonian acts as the generator of time translations, in the sense that it maps the coordinate η of a system in phase space into the coordinates the system will have, due to its equations of motion, at a slightly later time.

This last example encourages us to find another interpretation of canonical transformations. Up to now we have taken a **passive** view of the transformation, as a change of variables describing an unchanged physical situation, just as the passive view of a rotation is to view it as a change in the description of an unchanged physical point in terms of a rotated set of coordinates. But rotations are also used to describe changes in the physical situation with regards to a fixed coordinate system[17], and similarly in the case of motion through phase space, it is natural to think of the canonical transformation generated by the Hamiltonian as describing the actual motion of a system

[17]We leave to Mach and others the question of whether this distinction is real.

through phase space rather than as a change in coordinates. More generally, we may view a canonical transformation as a **diffeomorphism**[18] of phase space onto itself, $g : \mathcal{M} \to \mathcal{M}$ with $g(q, p) = (Q, P)$.

For an infinitesimal canonical transformation, this **active** interpretation gives us a small displacement $\delta\eta = \epsilon[\eta, G]$ for every point η in phase space, so we can view G and its associated infinitesimal canonical transformation as producing a flow on phase space. G also builds a finite transformation by repeated application, so that we get a sequence of canonical transformations g^λ parameterized by $\lambda = n\Delta\lambda$. This sequence maps an initial η_0 into a sequence of points $g^\lambda\eta_0$, each generated from the previous one by the infinitesimal transformation $\Delta\lambda G$, so $g^{\lambda+\Delta\lambda}\eta_0 - g^\lambda\eta_0 = \Delta\lambda[g^\lambda\eta_0, G]$. In the limit $\Delta\lambda \to 0$, with n allowed to grow so that we consider a finite range of λ, we have a one (continuous) parameter family of transformations $g^\lambda : \mathcal{M} \to \mathcal{M}$, satisfying the differential equation

$$\frac{dg^\lambda(\eta)}{d\lambda} = \left[g^\lambda\eta, G\right].$$

This differential equation defines a phase flow on phase space. If G is not a function of λ, this has the form of a differential equation solved by an exponential,

$$g^\lambda(\eta) = e^{\lambda[\cdot, G]}\eta,$$

which means

$$g^\lambda(\eta) = \eta + \lambda[\eta, G] + \frac{1}{2}\lambda^2[[\eta, G], G] + \ldots.$$

In the case that the generating function is the Hamiltonian, $G = H$, this phase flow gives the evolution through time, λ is t, and the velocity field on phase space is given by $[\eta, H]$. If the Hamiltonian is time independent, the velocity field is fixed, and the solution is formally an exponential.

Let us review changes due to a generating function considered in the passive and alternately in the active view. In the passive picture, we view η and $\zeta = \eta + \delta\eta$ as alternative coordinatizations of the same physical point A in phase space. For an infinitesimal generator $F_2 = \sum_i q_i P_i + \epsilon G$, $\delta\eta = \epsilon J\nabla G = \epsilon[\eta, G]$. A physical scalar defined by a function $u(\eta)$ changes its functional

[18]An isomorphism $g : \mathcal{M} \to \mathcal{N}$ is a 1-1 map with an image including all of $\mathcal{N}$ (onto), which is therefore invertible to form $g^{-1} : \mathcal{N} \to \mathcal{M}$. A diffeomorphism is an isomorphism g for which both g and g^{-1} are differentiable.

form to $\tilde{u}$, but not its value at a given physical point, so $\tilde{u}(\zeta_A) = u(\eta_A)$. For the Hamiltonian, there is a change in value as well, for $\tilde{H}$ or $\tilde{K}$ may not be the same as H, even at the corresponding point,

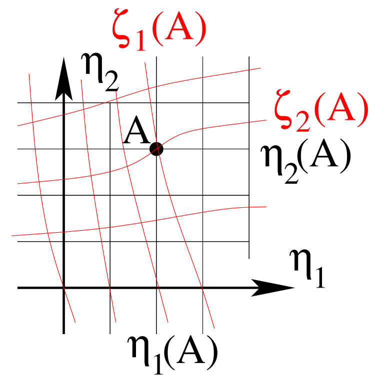
$$\tilde{K}(\zeta_A) = H(\eta_A) + \frac{\partial F_2}{\partial t} = H(\eta_A) + \epsilon \left.\frac{\partial G}{\partial t}\right|_A .$$

Now consider an active view. Here a canonical transformation is thought of as moving the point in phase space, and at the same time changing the functions $u \to \tilde{u}$, $H \to \tilde{K}$, where we are focusing on the form of these functions, on how they depend on their arguments. We think of ζ as representing the η coordinates of a different point B of phase space, although the coordinates $\eta_B = \zeta_A$. We ask how $\tilde{u}$ and $\tilde{K}$ differ from u and H at B, evaluated at the same values of their arguments, not at what we considered the same physical point in the passive view. Let[19]

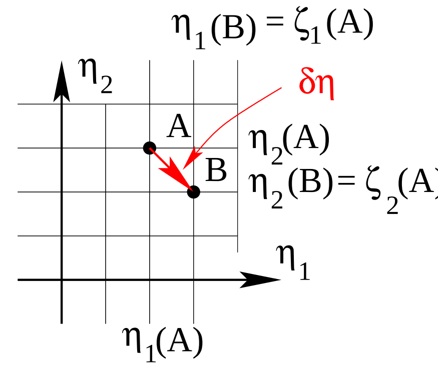
$$\begin{aligned}
\Delta u &= \tilde{u}(\eta_B) - u(\eta_B) = \tilde{u}(\zeta_A) - u(\zeta_A) = u(\eta_A) - u(\zeta_A) = -\delta\eta_i \frac{\partial u}{\partial \eta_i} \\
&= -\epsilon \sum_i [\eta_i, G] \frac{\partial u}{\partial \eta_i} = -\epsilon[u, G]
\end{aligned}$$

$$\begin{aligned}
\Delta H &= \tilde{K}(\eta_B) - H(\eta_B) = \tilde{K}(\zeta_A) - H(\eta_B) \\
&= H(\eta_A) + \epsilon \left.\frac{\partial G}{\partial t}\right|_A - H(\eta_A) - \delta\eta \cdot \nabla_\eta H = \epsilon \left(\frac{\partial G}{\partial t} - [H, G] \right) \\
&= \epsilon \frac{dG}{dt}.
\end{aligned}$$

[19] We differ by a sign from Goldstein in the definition of Δu.



Passive view of the canonical transformation. Point A is the same point, whether expressed in coordinates η_j or ζ_j, and scalar functions take the same value there, so $u(\eta_A) = \tilde{u}(\zeta_A)$.



Active view of the transformation, moving from point A in phase space to point B. So $\eta_B = \zeta_A$. The function $\tilde{u}$ then differs from u when evaluated on the same coordinates η.

Note that if the generator of the transformation is a conserved quantity, the Hamiltonian is unchanged, in that it is the same function after the transformation as it was before. That is, the Hamiltonian is **form invariant**.

So we see that conserved quantities are generators of symmetries of the problem, transformations which can be made without changing the Hamiltonian. We saw that the symmetry generators form a closed algebra under Poisson bracket, and that finite symmetry transformations result from exponentiating the generators. Let us discuss the more common conserved quantities in detail, showing how they generate symmetries. We have already seen that ignorable coordinates lead to conservation of the corresponding momentum. Now the reverse comes if we assume one of the momenta, say p_I, is conserved. Then from our discussion we know that the generator $G = p_I$ will generate canonical transformations which are symmetries of the system. Those transformations are

$$\delta q_j = \epsilon[q_j, p_I] = \epsilon\delta_{jI}, \qquad \delta p_j = \epsilon[p_j, p_I] = 0.$$

Thus the transformation just changes the one coordinate q_I and leaves all the other coordinates and all momenta unchanged. In other words, it is a translation of q_I. As the Hamiltonian is unchanged, it must be independent of q_I, and q_I is an ignorable coordinate.

Second, consider the angular momentum component $\vec{\omega}\cdot\vec{L} = \epsilon_{ijk}\omega_i r_j p_k$ for a point particle with $q = \vec{r}$. As a generator, $\epsilon\vec{\omega}\cdot\vec{L}$ produces changes

$$\begin{aligned}\delta r_\ell &= \epsilon[r_\ell, \epsilon_{ijk}\omega_i r_j p_k] = \epsilon\epsilon_{ijk}\omega_i r_j [r_\ell, p_k] = \epsilon\epsilon_{ijk}\omega_i r_j \delta_{\ell k} = \epsilon\epsilon_{ij\ell}\omega_i r_j \\ &= \epsilon(\vec{\omega}\times\vec{r})_\ell,\end{aligned}$$

which is how the point moves under a rotation about the axis $\vec{\omega}$. The momentum also changes,

$$\begin{aligned}\delta p_\ell &= \epsilon[p_\ell, \epsilon_{ijk}\omega_i r_j p_k] = \epsilon\epsilon_{ijk}\omega_i p_k [p_\ell, r_j] = \epsilon\epsilon_{ijk}\omega_i p_k(-\delta_{\ell j}) = -\epsilon\epsilon_{i\ell k}\omega_i p_k \\ &= \epsilon(\vec{\omega}\times\vec{p})_\ell,\end{aligned}$$

so $\vec{p}$ also rotates in the same way.

By Poisson's theorem, the set of constants of the motion is closed under Poisson bracket, and given two such generators, the bracket is also a symmetry, so the symmetries form a Lie algebra under Poisson bracket. For a free particle, $\vec{p}$ and $\vec{L}$ are both symmetries, and we have just seen that $[p_\ell, L_i] = \epsilon_{ik\ell}p_k$, a linear combination of symmetries, while of course $[p_i, p_j] = 0$ generates the identity transformation and is in the algebra. What about $[L_i, L_j]$? As $L_i = \epsilon_{ik\ell}r_k p_\ell$,

$$\begin{aligned}[L_i, L_j] &= [\epsilon_{ik\ell}r_k p_\ell, L_j] \\ &= \epsilon_{ik\ell}r_k[p_\ell, L_j] + \epsilon_{ik\ell}[r_k, L_j]p_\ell \\ &= -\epsilon_{ik\ell}r_k\epsilon_{j\ell m}p_m + \epsilon_{ik\ell}\epsilon_{jmk}r_m p_\ell \\ &= (\delta_{ij}\delta_{km} - \delta_{im}\delta_{jk})\, r_k p_m - (\delta_{ij}\delta_{m\ell} - \delta_{im}\delta_{j\ell})\, r_m p_\ell \\ &= (\delta_{ia}\delta_{jb} - \delta_{ib}\delta_{ja})\, r_a p_b \\ &= \epsilon_{kij}\epsilon_{kab}r_a p_b = \epsilon_{ijk}L_k. \end{aligned} \tag{6.16}$$

We see that we get back the third component of $\vec{L}$, so we do not get a new kind of conserved quantity, but instead we see that the algebra closes on the space spanned by the momenta and angular momenta. We also note that it is impossible to have two components of $\vec{L}$ conserved without the third component also being conserved. Note also that $\vec{\omega}\cdot\vec{L}$ does a rotation the same way on the three vectors $\vec{r}$, $\vec{p}$, and $\vec{L}$. Indeed it will do so on any vector composed from $\vec{r}$, and $\vec{p}$, rotating all of the physical system[20].

[20]If there is some rotationally non-invariant property of a particle which is not built out of $\vec{r}$ and $\vec{p}$, it will not be suitably rotated by $\vec{L} = \vec{r}\times\vec{p}$, in which case $\vec{L}$ is not the full angular momentum but only the **orbital angular momentum**. The generator of a rotation of all of the physics, the full angular momentum $\vec{J}$, is then the sum of $\vec{L}$ and another piece, called the **intrinsic spin** of the particle.

The use of the Levi-Civita ϵ_{ijk} to write L as a vector is peculiar to three dimensions; in other dimensions $d \neq 3$ there is no ϵ-symbol to make a vector out of L, but the angular momentum can always be treated as an antisymmetric tensor, $L_{ij} = x_i p_j - x_j p_i$. There are $D(D-1)/2$ components, and the Lie algebra again closes

$$[L_{ij}, L_{k\ell}] = \delta_{jk}L_{i\ell} - \delta_{ik}L_{j\ell} - \delta_{j\ell}L_{ik} + \delta_{i\ell}L_{jk}.$$

We have related conserved quantities to generators of infinitesimal canonical transformation, but these infinitesimals can be integrated to produce finite transformations as well. As we mentioned earlier, from an infinitesimal generator G we can exponentiate to form a one-parameter set of transformations

$$\begin{aligned}\zeta_\alpha(\eta) &= e^{\alpha[\cdot,G]}\eta \\ &= \left(1 + \alpha[\cdot,G] + \frac{1}{2}\alpha^2[[\cdot,G],G] + ...\right)\eta \\ &= \eta + \alpha[\eta,G] + \frac{1}{2}\alpha^2[[\eta,G],G] + ...\end{aligned}$$

In this fashion, any Lie algebra, and in particular the Lie algebra formed by the Poisson brackets of generators of symmetry transformations, can be exponentiated to form a continuous group of finite transformations, called a Lie Group. In the case of angular momentum, the three components of $\vec{L}$ form a three-dimensional Lie algebra, and the exponentials of these form a three-dimensional Lie group which is $SO(3)$, the rotation group.

6.7 Hamilton–Jacobi Theory

We have mentioned the time dependent canonical transformation that maps the coordinates of a system at a given fixed time t_0 into their values at a later time t. Now let us consider the reverse transformation, mapping $(q(t), p(t)) \to (Q = q_0, P = p_0)$. But then $\dot{Q} = 0$, $\dot{P} = 0$, and the Hamiltonian which generates these trivial equations of motion is $K = 0$. We denote by $S(q, P, t)$ the generating function of type 2 which generates this transformation. It satisfies

$$K = H(q,p,t) + \frac{\partial S}{\partial t} = 0, \qquad \text{with } p_i = \frac{\partial S}{\partial q_i},$$

so S is determined by the differential equation

$$H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0, \tag{6.17}$$

which we can think of as a partial differential equation in $n+1$ variables q, t, thinking of P as fixed and understood. If H is independent of time, we can solve by separating the t from the q dependence, we may write $S(q, P, t) = W(q, P) - \alpha t$, where α is the separation constant independent of q and t, but not necessarily of P. We get a time-independent equation

$$H\left(q, \frac{\partial W}{\partial q}\right) = \alpha. \tag{6.18}$$

The function S is known as **Hamilton's principal function**, while the function W is called **Hamilton's characteristic function**, and the equations (6.17) and (6.18) are both known as the **Hamilton-Jacobi** equation. They are still partial differential equations in many variables, but under some circumstances further separation of variables may be possible. We consider first a system with one degree of freedom, with a conserved H, which we will sometimes specify even further to the particular case of a harmonic oscillator. Then we we treat a separable system with two degrees of freedom.

We are looking for new coordinates (Q, P) which are time independent, and have the differential equation for Hamilton's principal function $S(q, P, t)$:

$$H\left(q, \frac{\partial S}{\partial q}\right) + \frac{\partial S}{\partial t} = 0.$$

For a harmonic oscillator with $H = p^2/2m + \frac{1}{2}kq^2$, this equation is

$$\left(\frac{\partial S}{\partial q}\right)^2 + kmq^2 + 2m\frac{\partial S}{\partial t} = 0. \tag{6.19}$$

For any conserved Hamiltonian, we can certainly find a separated solution of the form

$$S = W(q, P) - \alpha(P)t,$$

and then the terms in (6.17) from the Hamiltonian are independent of t. For the harmonic oscillator, we have an ordinary differential equation,

$$\left(\frac{dW}{dq}\right)^2 = 2m\alpha - kmq^2,$$

which can be easily integrated

$$\begin{aligned} W &= \int_0^q \sqrt{2m\alpha - kmq^2}\, dq + f(\alpha) \\ &= f(\alpha) + \frac{\alpha}{\omega}\left(\theta + \frac{1}{2}\sin 2\theta\right), \end{aligned} \tag{6.20}$$

where we have made a substitution $\sin\theta = q\sqrt{k/2\alpha}$, used $\omega = \sqrt{k/m}$, and made explicit note that the constant (in q) of integration, $f(\alpha)$, may depend on α. For other hamiltonians, we will still have the solution to the partial differential equation for S given by separation of variables $S = W(q, P) - \alpha t$, because H was assumed time-independent, but the integral for W may not be doable analytically.

As S is a type 2 generating function,

$$p = \frac{\partial F_2}{\partial q} = \frac{\partial W}{\partial q}.$$

For our harmonic oscillator, this gives

$$p = \frac{\partial W}{\partial \theta} \bigg/ \frac{\partial q}{\partial \theta} = \frac{\alpha}{\omega}\frac{1 + \cos 2\theta}{\sqrt{2\alpha/k}\,\cos\theta} = \sqrt{2\alpha m}\cos\theta.$$

Plugging into the Hamiltonian, we have

$$H = \alpha(\cos^2\theta + \sin^2\theta) = \alpha,$$

which will always be the case when (6.18) holds.

We have not spelled out what our new momentum P is, except that it is conserved, and we can take it to be α. The new coordinate $Q = \partial S/\partial P = \partial W/\partial\alpha\big|_q - t$. But Q is, by hypothesis, time independent, so

$$\left.\frac{\partial W}{\partial \alpha}\right|_q = t + Q.$$

For the harmonic oscillator calculation (6.20),

$$\left.\frac{\partial W}{\partial \alpha}\right|_q = f'(\alpha) + \frac{1}{\omega}(\theta + \frac{1}{2}\sin 2\theta) + \frac{\alpha}{\omega}\left.\frac{\partial \theta}{\partial \alpha}\right|_q (1 + \cos 2\theta) = f'(\alpha) + \frac{\theta}{\omega}$$

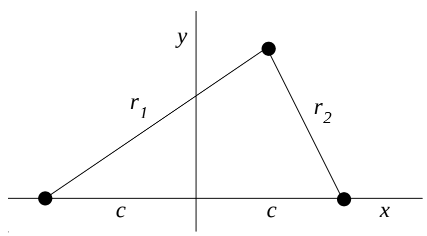
where for the second equality we used $\sin\theta = q\sqrt{k/2\alpha}$ to evaluate

$$\left.\frac{\partial\theta}{\partial\alpha}\right|_q = \frac{-q}{2\alpha\cos\theta}\sqrt{\frac{k}{2\alpha}} = -\frac{1}{2\alpha}\tan\theta.$$

Thus $\theta = \omega t + \delta$, for δ some constant.

As an example of a nontrivial problem with two degrees of freedom which is nonetheless separable and therefore solvable using the Hamilton-Jacobi method, we consider the motion of a particle of mass m attracted by Newtonian gravity to two equal masses fixed in space. For simplicity we consider only motion in a plane containing the two masses, which we take to be at $(\pm c, 0)$ in cartesian coordinates x, y. If r_1 and r_2 are the distances from the particle to the two fixed masses respectively, the gravitational potential is $U = -K(r_1^{-1} + r_2^{-1})$, while the kinetic energy is simple in terms of x and y, $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$. The relation between these is, of course,

$$\begin{aligned} r_1^2 &= (x+c)^2 + y^2 \\ r_2^2 &= (x-c)^2 + y^2 \end{aligned}$$



Considering both the kinetic and potential energies, the problem will not separate either in terms of (x, y) or in terms of (r_1, r_2), but it does separate in terms of elliptical coordinates

$$\begin{aligned} \xi &= r_1 + r_2 \\ \eta &= r_1 - r_2 \end{aligned}$$

From $r_1^2 - r_2^2 = 4cx = \xi\eta$ we find a fairly simple expression $\dot{x} = (\xi\dot{\eta} + \dot{\xi}\eta)/4c$. The expression for y is more difficult, but can be found from observing that $\frac{1}{2}(r_1^2 + r_2^2) = x^2 + y^2 + c^2 = (\xi^2 + \eta^2)/4$, so

$$y^2 = \frac{\xi^2 + \eta^2}{4} - \left(\frac{\xi\eta}{4c}\right)^2 - c^2 = \frac{(\xi^2 - 4c^2)(4c^2 - \eta^2)}{16c^2},$$

or

$$y = \frac{1}{4c}\sqrt{\xi^2 - 4c^2}\sqrt{4c^2 - \eta^2}$$

and

$$\dot{y} = \frac{1}{4c}\left(\xi\dot{\xi}\sqrt{\frac{4c^2 - \eta^2}{\xi^2 - 4c^2}} - \eta\dot{\eta}\sqrt{\frac{\xi^2 - 4c^2}{4c^2 - \eta^2}}\right).$$

Squaring, adding in the x contribution, and simplifying then shows that

$$T = \frac{m}{8}\left(\frac{\xi^2 - \eta^2}{4c^2 - \eta^2}\dot{\eta}^2 + \frac{\xi^2 - \eta^2}{\xi^2 - 4c^2}\dot{\xi}^2\right).$$

Note that there are no crossed terms $\propto \dot{\xi}\dot{\eta}$, a manifestation of the orthogonality of the curvilinear coordinates ξ and η. The potential energy becomes

$$U = -K\left(\frac{1}{r_1} + \frac{1}{r_2}\right) = -K\left(\frac{2}{\xi + \eta} + \frac{2}{\xi - \eta}\right) = \frac{-4K\xi}{\xi^2 - \eta^2}.$$

In terms of the new coordinates ξ and η and their conjugate momenta, we see that

$$H = \frac{2/m}{\xi^2 - \eta^2}\left(p_\xi^2(\xi^2 - 4c^2) + p_\eta^2(4c^2 - \eta^2) - 2mK\xi\right).$$

Then the Hamilton-Jacobi equation for Hamilton's characteristic function is

$$\frac{2/m}{\xi^2 - \eta^2}\left((\xi^2 - 4c^2)\left(\frac{\partial W}{\partial\xi}\right)^2 + (4c^2 - \eta^2)\left(\frac{\partial W}{\partial\eta}\right)^2 - 2mK\xi\right) = \alpha,$$

or

$$\begin{aligned} &(\xi^2 - 4c^2)\left(\frac{\partial W}{\partial\xi}\right)^2 - 2mK\xi - \frac{1}{2}m\alpha\xi^2 \\ &\quad + (4c^2 - \eta^2)\left(\frac{\partial W}{\partial\eta}\right)^2 + \frac{1}{2}\alpha m\eta^2 = 0. \end{aligned}$$

If W is to separate into a ξ dependent piece and an η dependent one, the first line will depend only on ξ, and the second only on η, so they must each be constant, with $W(\xi, \eta) = W_\xi(\xi) + W_\eta(\eta)$, and

$$\begin{aligned} (\xi^2 - 4c^2)\left(\frac{dW_\xi(\xi)}{d\xi}\right)^2 - 2mK\xi - \frac{1}{2}\alpha m\xi^2 &= \beta \\ (4c^2 - \eta^2)\left(\frac{dW_\eta(\eta)}{d\eta}\right)^2 + \frac{1}{2}\alpha m\eta^2 &= -\beta. \end{aligned}$$

These are now reduced to integrals for W_i, which can in fact be integrated to give an explicit expression in terms of elliptic integrals.

6.8 Action-Angle Variables

Consider again a general system with one degree of freedom and a conserved Hamiltonian. Suppose the system undergoes periodic behavior, with $p(t)$ and $\dot{q}(t)$ periodic with period τ. We don't require q itself to be periodic as it might be an angular variable which might not return to the same value when the system returns to the same physical point, as, for example, the angle which describes a rotation.

If we define an integral over one full period,

$$J(t) = \frac{1}{2\pi}\int_t^{t+\tau} p\, dq,$$

it will be time independent. As $p = \partial W/\partial q = p(q, \alpha)$, the integral can be defined without reference to time, just as the integral $2\pi J = \int pdq$ over one orbit of q, holding α fixed. Then J becomes a function of α alone, and if we assume this function to be invertible, $H = \alpha = \alpha(J)$. We can take J to be our canonical momentum P. Using Hamilton's Principal Function S as the generator, we find $Q = \partial S/\partial J = \partial W(q, J)/\partial J - (d\alpha/dJ)t$. Alternatively, we might use Hamilton's Characteristic Function W by itself as the generator, to define the conjugate variable $\phi = \partial W(q, J)/\partial J$, which is simply related to $Q = \phi - (d\alpha/dJ)t$. Note that ϕ and Q are both canonically conjugate to J, differing at any instant only by a function of J. As the Hamilton-Jacobi Q is time independent, we see that $\dot{\phi} = d\alpha/dJ = dH/dJ = \omega(J)$, which is a constant, because while it is a function of J, J is a constant in time. We could also derive $\dot{\phi}$ from Hamilton's equations considering W as a genenerator, for W is time independent, the therefore the new Hamiltonian is unchanged, and the equation of motion for ϕ is simply $\dot{\phi} = \partial H/\partial J$. Either way, we have $\phi = \omega t + \delta$. The coordinates (J, ϕ) are called **action-angle** variables. Consider the change in ϕ during one cycle.

$$\Delta\phi = \oint \frac{\partial \phi}{\partial q}\, dq = \oint \left(\frac{\partial}{\partial q}\frac{\partial W}{\partial J}\right) dq = \frac{d}{dJ}\oint pdq = \frac{d}{dJ} 2\pi J = 2\pi.$$

Thus we see that in one period τ, $\Delta\phi = 2\pi = \omega\tau$, so $\omega = 1/\tau$.

For our harmonic oscillator, of course,

$$2\pi J = \oint pdq = \sqrt{2\alpha m}\sqrt{\frac{2\alpha}{k}}\int_0^{2\pi}\cos^2\theta d\theta = \frac{2\alpha\pi}{\sqrt{k/m}}$$

so J is just a constant $1/\sqrt{k/m}$ times the old canonical momentum α, and thus its conjugate $\phi = \sqrt{k/m}(t + \beta)$, so $\omega = \sqrt{k/m}$ as we expect. The important thing here is that $\Delta\phi = 2\pi$, even if the problem itself is not solvable.

Exercises

6.1 In Exercise 2.7, we discussed the connection between two Lagrangians, L_1 and L_2, which differed by a total time derivative of a function on extended configuration space,

$$L_1(\{q_i\}, \{\dot{q}_j\}, t) = L_2(\{q_i\}, \{\dot{q}_j\}, t) + \frac{d}{dt}\Phi(q_1, ..., q_n, t).$$

You found that these gave the same equations of motion, but differing momenta $p_i^{(1)}$ and $p_i^{(2)}$. Find the relationship between the two Hamiltonians, H_1 and H_2, and show that these lead to equivalent equations of motion.

6.2 A uniform static magnetic field can be described by a static vector potential $\vec{A} = \frac{1}{2}\vec{B} \times \vec{r}$. A particle of mass m and charge q moves under the influence of this field.
(a) Find the Hamiltonian, using inertial cartesian coordinates.
(b) Find the Hamiltonian, using coordinates of a rotating system with angular velocity $\vec{\omega} = -q\vec{B}/2mc$.

6.3 Consider a symmetric top with one point on the symmetry axis fixed in space, as we did at the end of chapter 4. Write the Hamiltonian for the top. Noting the cyclic (ignorable) coordinates, explain how this becomes an effective one-dimensional system.

6.4 (a) Show that a particle under a central force with an attractive *potential* inversely proportional to the distance *squared* has a conserved quantity $D = \frac{1}{2}\vec{r} \cdot \vec{p} - Ht$.
(b) Show that the infinitesimal transformation generated by $G := \frac{1}{2}\vec{r} \cdot \vec{p}$ scales $\vec{r}$ and $\vec{p}$ by opposite infinitesimal amounts, $\vec{Q} = (1 + \frac{\epsilon}{2})\vec{r}$, $\vec{P} = (1 - \frac{\epsilon}{2})\vec{p}$, or for a finite transformation $\vec{Q} = \lambda\vec{r}$, $\vec{P} = \lambda^{-1}\vec{p}$. Show that if we describe the motion in terms of a scaled time $T = \lambda^2 t$, the equations of motion are invariant under this combined transformation $(\vec{r}, \vec{p}, t) \to (\vec{Q}, \vec{P}, T)$.

6.5 We saw that the Poisson bracket associates with every differentiable function f on phase space a differential operator $D_f := [f, \cdot]$ which acts on functions g

on phase space by $D_f g = [f, g]$. We also saw that every differential operator is associated with a vector, which in a particular coordinate system has components f_i, where

$$D_f = \sum f_i \frac{\partial}{\partial \eta_i}.$$

A 1-form acts on such a vector by

$$dx_j(D_f) = f_j.$$

Show that for the natural symplectic structure ω_2, acting on the differential operator coming from the Poisson bracket as its first argument,

$$\omega_2(D_f, \cdot) = df,$$

which indicates the connection between ω_2 and the Poisson bracket.

6.6 Give a complete discussion of the relation of forms in cartesian coordinates in four dimensions to functions, vector fields, and antisymmetric matrix (tensor) fields, and what wedge products and exterior derivatives of the forms correspond to in each case. This discussion should parallel what is done in section 6.5 for three dimensions. [Note that two different antisymmetric tensors, $B_{\mu\nu}$ and $\tilde{B}_{\mu\nu} = \frac{1}{2}\sum_{\rho\sigma} \epsilon_{\mu\nu\rho\sigma} B_{\rho\sigma}$, can be related to the same 2-form, in differing fashions. They are related to each other with the four dimensional $\epsilon_{jk\ell m}$, which you will need to define, and are called **duals** of each other. Using one fashion, the two different 2-forms associated with these two matrices are also called duals.
(b) Let $F_{\mu\nu}$ be a 4×4 matrix defined over a four dimensional space (x, y, z, ict), with matrix elements $F_{jk} = \epsilon_{jk\ell} B_\ell$, for j, k, ℓ each $1, 2, 3$, and $F_{4j} = iE_j = -F_{j4}$. Show that the statement that F corresponds, by one of the two fashions, to a closed 2-form F, constitutes two of Maxwell's equations, and explain how this implies that 2-form is the exterior derivative of a 1-form, and what that 1-form is in terms of electromagnetic theory described in 3-dimensional language.
(c) Find the 3-form associated with the exterior derivative of the 2-form dual to F, and show that it is associated with the 4-vector charge current density $J = (\vec{j}, ic\rho)$, where $\vec{j}$ is the usual current density and ρ the usual charge density.

6.7 Consider the following differential forms:

$$\begin{aligned}
A &= y\,dx + x\,dy + dz \\
B &= y^2\,dx + x^2\,dy + dz \\
C &= xy(y-x)\,dx \wedge dy + y(y-1)\,dx \wedge dz + x(x-1)\ dy \wedge dz \\
D &= 2(x-y)\,dx \wedge dy \wedge dz \\
E &= 2(x-y)\,dx \wedge dy
\end{aligned}$$

Find as many relations as you can, expressible without coordinates, among these forms. Consider using the exterior derivative and the wedge product.

6.8 We have considered k-forms in 3-D Euclidean space and their relation to vectors expressed in cartesian basis vectors. We have seen that k-forms are invariant under change of coordinatization of $\mathcal{M}$, so we can use them to examine the forms of the gradient, curl, divergence and laplacian in general coordinates in three dimensional space. We will restrict our treatment to *orthogonal curvilinear coordinates* (q_1, q_2, q_3), for which we have, at each point $\mathbf{p} \in \mathcal{M}$, a set of **orthonormal** basis vectors $\hat{e}_i$ directed along the corresponding coordinate, so that $dq_i(\hat{e}_j) = 0$ for $i \neq j$. We assume they are right handed, so $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$ and $\hat{e}_i \times \hat{e}_j = \sum_k \epsilon_{ijk} \hat{e}_k$. The dq_i are not normalized measures of distance, so we define $h_i(\mathbf{p})$ so that $dq_i(\hat{e}_j) = h_i^{-1} \delta_{ij}$ (no sum).

(a) For a function $f(q_1, q_2, q_3)$ and a vector $\vec{v} = \sum v_i \hat{e}_i$, we know that $df(\vec{v}) = \vec{v} \cdot \vec{\nabla} f$. Use this to find the expression for $\vec{\nabla} f$ in the basis $\hat{e}_i$.
(b) Use this to get the general relation of a 1-form $\sum \omega_i dq_i$ to its associated vector $\vec{v} = \sum v_i \hat{e}_i$.
(c) If a 1-form $\omega^{(a)}$ is associated with $\vec{v}^{(a)}$ and 1-form $\omega^{(b)}$ is associated with $\vec{v}^{(b)}$, we know the 2-form $\omega^{(a)} \wedge \omega^{(b)}$ is associated with $\vec{v}^{(a)} \times \vec{v}^{(b)}$. Use this to find the general association of a 2-form with a vector.
(d) We know that if a 1-form ω is associated with a vector $\vec{v}$, then $d\omega$ is associated with $\vec{\nabla} \times \vec{v}$. Use this to find the expression for $\vec{\nabla} \times \vec{v}$ in orthogonal curvilinear coordinates.
(e) If the 1-form ω is associated with $\vec{v}$ and the 2-form Ω is associated with $\vec{F}$, we know that $\omega \wedge \Omega$ is associated with the scalar $\vec{v} \cdot \vec{F}$. Use this to find the general association of a 3-form with a scalar.
(f) If the 2-form Ω is associated with $\vec{v}$, we know that $d\Omega$ is associated with the divergence of $\vec{v}$. Use this to find the expression for $\vec{\nabla} \cdot \vec{v}$ in orthogonal curvilinear coordinates.
(g) Use (a) and (f) to find the expression for the laplacian of a scalar, $\nabla^2 f = \vec{\nabla} \cdot \vec{\nabla} f$, in orthogonal curvilinear coordinates.

6.9 Consider the unusual Hamiltonian for a one-dimensional problem

$$H = \omega(x^2 + 1)p,$$

where ω is a constant.

(a) Find the equations of motion, and solve for $x(t)$.