6.3 Canonical transformations

We have seen that it is often useful to switch from the original set of coordinates in which a problern 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 Q, where each Qi is a function of all the qj 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 sirnilar ways, and we rnight ask ourselves what happens if we make a change of variables on phase space, to new variables Qi(q,p, t), t). 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 r/ for our new variables will be called (, so

C-'

$$\dot{\zeta} = J \cdot \frac{\partial K}{\partial \zeta}.$$

If this exists, we say the new variables (Q, P) are canonical variables and the transformation (q, p) —+ (Q, P) is a canonical transformation. Note that the functions Q and Pi may depend on time as well as on q and p.

These new Hamiltonian equations are related to the old ones, i) = J .DI-I/DQ, by the function which gives the new coordinates and momenta in terms of the old, (= ((n, t). Then

Let us write the Jacobian matrix Miy DCi/Dr)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

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

$$= J\cdot\nabla_\zeta K.$$

vcH + —

6.3. CANONICAL TRANSFORMATIONS

Let us first consider a canonical transformation which does not depend on time, so DC/DtlQ —— O. We see that we can choose the new Hamiltonian to be the same as the old, K = H, and get correct mechanics, if

$$M\cdot J\cdot M^T = J. \tag{6.3}$$

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

The condition (03) on M is similar to, and a generalization of, the condition for orthogonality of a matrix, OO^T = 11, which is of the same form with J replaced by ll. Another example of this kind of relation in physics occurs in special relativity, where a Lorentz transformation Lev gives the relation between two coordinates, = $\sum_\nu L_{\mu\nu}x_\nu$, with a four dimensional vector with $4 = ct. Then the condition which makes L a Lorentz transformation

L g • L^T = g, 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 symplcctic.

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, Mlj = -F J , or M — 11 -h EG, where e is considered some infinitesimal parameter while G is a finite matrix. As M is symplectic, (1 + G) • J • (1 -F EG^T) — J, which tells us that to lowest order in E, GJ + JGT = O. Comparing this to the condition for the generator of an infinitesimal rotation, Q —Q^T , we see that it is sirnilar 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 63) for a transformation T) —i (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) : r/ C. 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, to) which maps each point T) at which a system can be at time to into the point to which it will evolve at time t. That is, g(t, to) : rl(to) 1-+ n(t). If we consider t — to -E- At for infinitesimal At, this is an infinitesimal transformation. As (i = Atüi — 7b -F At Ek JikDH/Dr)k, we have IVfij =$\partial\zeta_i/\partial\eta_j = \delta_{ij}$+ AtEk JikD^2 H/DTbDQk, so Go Jik0^2+1/DQjDQk,

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

The factor in parentheses in the last line is (—JikJje -F JilJjk) which is antisymmetric under k C, and as it is contracted into the second derivative, which is symmetric under k C, we see that (GJ -l- JG^T)iJ = O 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 canonicafl the map g(t, to) which takes "(to) into the point it will evolve into after a finite time increment t—to, is also a canonical transformation.

Notice that the relationship ensuring Hamilton's equations exist,

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

with the symplectic condition $M \cdot J \cdot M^T = J$, implies $\nabla_\zeta(K - H) = -$ — J•DC/Dt, 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

If M and • J . M/ — J, J, then M • MT —
(Ml • J(.Mi[T] • MT) — Ml Ml • J • MT J, so M is canonical.

6.4. POISSON BRACKETS

phase space with time. Then

$$\begin{aligned}\frac{df}{dt} &= \sum_i \frac{\partial f}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial f}{\partial p_i}\dot{p}_i + \frac{\partial f}{\partial t} \\ &= \sum_i \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \sum_i \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial t}.\end{aligned} \quad (6.4)$$

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

$$[u, v] := \sum_i \frac{\partial u}{\partial q_i}\frac{\partial v}{\partial p_i} \quad \text{-E\_\_\_ \_\_\_} \quad . \quad (6.5)$$

Dpi Dqz

Thus Eq. may be rewritten as

$$\frac{df}{dt} = [f, H] + \frac{\partial f}{\partial t}. \quad (6.6)$$

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

$$[u, 'v] = - \text{Jij} \frac{\partial v}{\partial \eta_j} = (\nabla_\eta u)^T \cdot J \cdot \nabla_\eta v. \quad (6.7)$$

. .Dr)i

If we describe the system in terms of a different set of canonical variables C, we should still find the function f (t) changing at the same rate. We may think of u and v as functions of (as easily as of r). Really we are thinking of u and v as functions of points in phase space, represented by u(q) = ü(C) and we may ask whether [i, is the same as [u, v] n. Using $M^T \cdot \nabla_\zeta$, 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:

$$[u,v] = -[v,u]. \tag{6.8}$$

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, -Fulv, (6.9)

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

$$[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0. \tag{6.10}$$

We need to prove that this is true. To simplify the presentation, we introduce some abbreviated notation. We use a subscript to indicate partial derivative with respect to m, so means Du/Drji, and means $\partial(\partial u/\partial \eta_i)/\partial \eta_j$. We will assume all our functions on phase space are suitably differentiable, so $u_{,i,j}$ — z. We will also use the summation convention, that any index which appears twice in a term is assumed to be summed ovelff 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 —i v —+ w —4 u and the other with u —+ w —+ v —i 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 —i W —+ —+ 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 —+ k —+ j —+ C —+ i, converting this second term to Adding the original term $u_{,k} J_{k\ell} v_{,i,\ell} J_{ij} w_{,j}$, and using v c z — t',iß, gives O 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 recogmize that for each function f on phase space, we may view [f, •J as a differential operator on

---

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

functions g on phase space, mapping g —+ [f, g]. Calling this operator Df, 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 % are an arbitrary set of functions on phase space. For the Poisson bracket, the functions % are linear combinations of the fo, but % f,j. With this interpretation, [f, g] = Dfg, and [h, [f, g]] = DhDfg. Thus

$$[h,[f,g]] + [f,[g,h]] = [h,[f,g]] - [f,[h,g]] - D_hD_fg - D_fD_hg$$

$$= (D_hD_f - D_fD_h)g, \qquad (6.11)$$

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,

$$D_hD_f = \sum h_i \frac{\partial}{\partial \eta_i} f_j \frac{\partial}{\partial \eta_j} = \sum_{ij} h_i \frac{\partial f_j}{\partial \eta_i} \frac{\partial}{\partial \eta_j} + \sum_{ij} h_i f_j \frac{\partial^2}{\partial \eta_i \partial \eta_j}$$

$$D_h = \sum f_j \frac{\partial}{\partial \eta_j} h_i \frac{\partial}{\partial \eta_i} = \sum_{ij} f_j \frac{\partial h_i}{\partial \eta_j} \frac{\partial}{\partial \eta_i} + \sum_{ij} h_i f_j \frac{\partial^2}{\partial \eta_i \partial \eta_j}$$

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

$$D_hD_f - D_fD_h = \sum_{ij} h_i \frac{\partial f_j}{\partial \eta_i} \frac{\partial}{\partial \eta_j} - \sum_{ij} f_j \frac{\partial h_i}{\partial \eta_j} \frac{\partial}{\partial \eta_i}$$

$$= \sum_{ij} \left(h_i \frac{\partial f_j}{\partial \eta_i} - f_i \frac{\partial h_j}{\partial \eta_i} \right) \frac{\partial}{\partial \eta_j}.$$

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

$$DhDf — DfDh = D_{[h,f]} \quad (6.12)$$

An antisymmetric product which obeys the Jacobi identity is what makes a Lio 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]] [IV, v], w]. In fact, the difference [u, [v, w]] — [u v w] =[u, [v, w]] + w, [u, v]] = —tv, [w, u]] by the Jacobi identity, so the Jacobi identity replaces the law Of associativity in n Lie algebra. Lie groups play a major role in quantum rnecllanics 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 wAt or QAt of Chapter 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

$$\frac{df}{dt} = -[H, f] + \frac{\partial f}{\partial t}, \quad (6.13)$$

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 = O, and we have the important consequence that Illf, g]/dt O. 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 Liouvillc's theorom, 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 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 • J • M^T det J = det M det J det M^T = (det det J, and J is nonsingular, so det M = +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 to the values into which they will evolve at time t2 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 prob-ability density, specifying the likelihood that the system is near a particular point of phase space, is invariant as we move along with the system.