

Review for Hamiltonian Mechanics

1. Basic Geometry

We begin with some particular, physical system, and suppose that it is appropriately described by some set, $\{q^a\}_1^n$, of *generalized coordinates*; as well of course it has something which characterizes how it changes as the time, t , changes. We suppose that they constitute one set of coordinates for some (smooth, differential) manifold, Q . [In principle, this manifold has been constructed in some way by considering the problem, originally, or even hypothetically, in ordinary Euclidean space, with all its parts, and its constraints, so that we could determine a desirable set of generalized coordinates. On the other hand, this has already been done, and don't do it again now.] We now propose to put fibers over each point, q , of Q , with coordinates \dot{q}^a . In the general case, these fibers may be referred to as “*the first jet space*” over $q \in Q$, and the full *bundle* of those fibers, and the underlying manifold, as the first jet bundle, $J^{(1)}(Q)$. This is a useful place, for instance, to study the behavior of the solutions of first-order differential equations, since any (set of) first-order differential equation may be considered as defining a surface within the jet; i.e., the differential equation(s) may be considered as a functional relationship between the \dot{q}^a and the q^a , which then describes this surface in the space of those variables. [Actually it's a variety, i.e., a surface that may well have self-intersections, but locally that doesn't matter much.] The general geometry of jet spaces is a useful, standard place to study differential equations: for a k -th order equation one would look at it in the k -th jet bundle, $J^{(k)}(Q)$, as a surface there. For further purposes toward the study of such an equation, one would need to ensure that vector fields and/or differential forms are restricted to that surface. The simplest way to do this is to have a projection operator that is then pulled forward or backward to the tangent or cotangent spaces. It is also quite common to, instead, take the inverse limit on k , and look at the equation on the surface which the pde generates which lies within the infinite jet bundle, $J^{(\infty)}$. We will not go further in that direction now, since it's not needed for basic classical mechanics.

In mechanics, we may characterize a trajectory of our system as a curve on Q , parameterized by the time. The tangent vector to that curve would simply be $\partial/\partial t = \dot{q}^a \partial_{q^a}$, with the \dot{q}^a coming from the solutions to the dynamical problem. As this is a vector in the tangent bundle, which is the same as the first jet bundle above, then it now has this additional structure so that each fiber is a vector space. Therefore we will refer to the individual fiber over a point $q \in Q$ by the symbol $T_q Q$, and refer to the entire bundle simply by TQ ! On the other hand, since Newton's equations involve accelerations, a more reasonable approach to a description of our dynamical problem, i.e., the time evolution of this system, would be to think of it as a trajectory on the tangent bundle itself. This has the property that any individual point in this TQ would have a unique trajectory, or flow, running through it, for which it constitutes the “initial conditions,” at some particular choice of initial time t_0 . The associated curve in TQ then constitutes a unique *flow*, corresponding to the choice of initial condition, the “dynamics” having already been specified when we chose Q itself. On this trajectory in TQ , we may also talk about tangent vectors, in the tangent bundle to TQ , rather like a second-order jet space. An arbitrary vector in that (larger) tangent bundle would be of the form

$$\tilde{X} = X_1^a \partial_{q^a} + X_2^a \partial_{\dot{q}^a} ,$$

where the quantities X_1^a and X_2^a are sets of functions over TQ . Any particular (smooth) choice of these functions is usually referred to as a vector field, or, sometimes, a section of the bundle. The

tangent vector to the trajectory in TQ then has the particular form

$$\frac{d}{dt} = \dot{q}^a \partial_{q^a} + \ddot{q}^a \partial_{\dot{q}^a} \equiv \tilde{\Delta}, \quad (1.1)$$

where the second derivatives are to be determined by the solution to the dynamical problem, including the initial conditions. Therefore José and Saletan refer to this particular vector field as *the dynamical vector field*, or, simply *the dynamics*. The t in question is of course just the parameter along the curve at this point, so it must be distinguished from some more abstract notion of time. If we have some interesting function over TQ , i.e., a function that depends on q^a and \dot{q}^a , then we may ask how that function changes as we follow along this trajectory, by asking for

$$\text{along a trajectory, with parameter } t: \quad \dot{F} = \frac{d}{dt}F = \dot{q}^a F_{,q^a} + \ddot{q}^a F_{,\dot{q}^a} = \tilde{\Delta}(F) \equiv \mathcal{L}_{\tilde{\Delta}}F, \quad (1.2)$$

where $\mathcal{L}_{\tilde{\Delta}}$ indicates the Lie derivative in the direction $\tilde{\Delta}$. On the other hand, one may evaluate this in a different way by first introducing the 1-form,

$$dF \equiv F_{,q^a} dq^a + F_{,\dot{q}^a} d\dot{q}^a, \quad (1.3a)$$

and then evaluating this 1-form against the tangent vector:

$$\text{along a trajectory, with parameter } t: \quad \dot{F} = dF\left(\frac{d}{dt}\right) = \dot{q}^a F_{,q^a} + \ddot{q}^a F_{,\dot{q}^a}, \quad (1.3b)$$

remembering the rules for the action of 1-forms on vector fields. [We should note that if F also depends on t directly, one must also add to the right-hand side a partial derivative with respect to t : $\partial F/\partial t \equiv F_{,t}$.]

The dynamics, as above, can be determined by knowing the quantities \ddot{q}^a as functions over TQ . These are given as solutions of some set of differential equations. The Lagrangian method involves setting up a scalar function, L as a function over TQ , and possibly also the time t , and also requiring that the dynamics is determined by satisfying the Euler-Lagrange equations for L , namely

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^a} \right) - \frac{\partial L}{\partial q^a} = 0. \quad (1.4a)$$

A standard phraseology for those equations in this language is, instead, given as follows. We first introduce the **canonical one-form**, which of course depends on the Lagrangian:

$$\boldsymbol{\theta} \equiv \frac{\partial L}{\partial \dot{q}^a} dq^a. \quad (1.5)$$

We then require that the Lie derivative (in the direction $\tilde{\Delta}$ of this canonical 1-form should equal the exterior derivative of the Lagrangian:

$$dL = \mathcal{L}_{\tilde{\Delta}} \boldsymbol{\theta}. \quad (1.4b)$$

That this works is straightforward, provided one recalls that the Lie derivative and the exterior derivative commute, i.e.,

$$\mathcal{L}_{\tilde{\Delta}} \boldsymbol{\theta} = \left(\mathcal{L}_{\tilde{\Delta}} \frac{\partial L}{\partial \dot{q}^a} \right) dq^a + \frac{\partial L}{\partial \dot{q}^a} d \left(\mathcal{L}_{\tilde{\Delta}} q^a \right) = \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^a} \right) dq^a + \frac{\partial L}{\partial \dot{q}^a} d\dot{q}^a. \quad (1.4c)$$

That the canonical 1-form is important in this way suggests that its coefficients are also of especial importance. We therefore follow Hamilton and define the associated, canonical momenta for our problem:

$$p_a \equiv \frac{\partial L}{\partial \dot{q}^a}, \quad (1.6)$$

which are clearly coefficients of a 1-form, i.e., they transform in that way under some coordinate transformation. One may then re-consider the problem of the dynamics from the Hamiltonian point of view, via $H = H(q^a, p_a, t)$. By solving Eqs. (1.6) for \dot{q}^a as functions of the q^a and the p_a , we may re-consider our problem in terms of q^a and p_a , instead of q^a and \dot{q}^a . As the p_a are components of an arbitrary 1-form, at an arbitrary point on Q , we will take them as coordinates for the fibers of the cotangent bundle, T^*Q , over Q , also usually referred to as *phasespace*. It is then relevant to re-express the canonical 1-form in a form appropriate for this bundle, namely as

$$\boldsymbol{\theta} = p_a dq^a. \quad (1.5')$$

Since we are now describing the dynamics for our problem over the cotangent bundle, one should have an alternate description for the trajectories in phasespace, i.e., trajectories over T^*Q with parameter t , alternate to Eq. (1.1). José and Saletan distinguish our first one by giving it the index L , for Lagrangian, and this next one, below, with the index H , for Hamiltonian; I will usually not bother putting an index on it, and figure that we can decide which one we mean:

$$\left. \begin{array}{l} \text{on the cotangent bundle,} \\ \text{a trajectory has tangent vector} \end{array} \right\} \quad \frac{d}{dt} = \dot{q}^a \partial_{q^a} + \dot{p}_b \partial_{p_b} \equiv \tilde{\Delta}_H \equiv \tilde{\Delta}. \quad (1.7)$$

Expressing the Hamiltonian via the Legendre transform, as a function over phasespace:

$$H = H(q, p, t) \equiv p_b \dot{q}^b(q, p, t) - L[q, \dot{q}(q, p, t), t], \quad (1.8)$$

which gives us the version of the dynamical equations in the cotangent bundle:

$$\frac{d}{dt} q^a = \frac{\partial H}{\partial p_a}, \quad \frac{d}{dt} p_b = -\frac{\partial H}{\partial q^b}. \quad (1.9)$$

[Note that there is a very nice description of the geometrical meaning of a Legendre transform given in p. 212-215 of José and Saletan, which explains how a curve may be described by giving either the set of all points through which it passes or by the set of all its tangent planes. Basically the Legendre transform switches between these two descriptions.] If we want to re-describe the tangent vector for a trajectory, on the cotangent bundle, in terms of the explicit dynamics given by the Hamiltonian, we may insert Hamilton's equations into that tangent vector:

$$\tilde{\Delta} = \frac{\partial H}{\partial p_a} \partial_{q^a} - \frac{\partial H}{\partial q^b} \partial_{p_b}. \quad (1.10)$$

The form of this presentation of the tangent vector, with alternating p 's and q 's, and a minus sign, induces us to consider a generic mapping that performs such a role: it takes an arbitrary function over phasespace and associates with it a vector field of that form, so that it would map the Hamiltonian into this tangent vector field for some trajectory. However, before doing that, as we will from now on be performing all our operations over the phasespace, let us first introduce more

general coordinates for that space, T^*Q . We introduce coordinates ξ^α to cover both the q^a 's and the p_b 's, as follows:

$$\xi^\alpha \implies \begin{pmatrix} q^a \\ p_b \end{pmatrix}, \quad \alpha = 1, 2, \dots, 2n, \quad a, b = 1, 2, \dots, n. \quad (1.11)$$

A basis set for the space of vectors over T^*Q is then just $\{\partial_{\xi^\alpha}\}_1^{2n}$, a basis set for 1-forms over T^*Q is $d\xi^\alpha\}_1^{2n}$, and then a basis set for 2-forms is $\{d\xi^\alpha \wedge d\xi^\beta\}$, where of course the skew-symmetry of the wedge product gives us a total of $\frac{1}{2}(2n)(2n-1)$ of them. In this notation our tangent vector to a trajectory, as given in Eq. (1.7), and the (exterior) derivative of the Hamiltonian take the simple forms:

$$\begin{aligned} \tilde{\Delta} &= \dot{q}^a \partial_{q^a} + \dot{p}_b \partial_{p_b} = \dot{\xi}^\alpha \partial_{\xi^\alpha} = \frac{d}{dt}, \\ dH &= H_{,q^a} dq^a + H_{,p_b} dp_b = H_{,\xi^\alpha} d\xi^\alpha. \end{aligned} \quad (1.12)$$

We may now introduce the notion of a **Hamiltonian vector field**, which is a very important generalization of Eq. (1.10), for the Hamiltonian function itself. Our generalization of that arrangement is a mapping that takes any arbitrary (nonconstant) function, say f , on phasespace, T^*Q , and maps it into a vector field, \tilde{X}_f , referred to as the Hamiltonian vector field generated by f :

$$\tilde{X}_f \equiv f_{,p_b} \partial_{q^b} - f_{,q^a} \partial_{p_a} \equiv (P^{\alpha\beta} f_{,\xi^\beta}) \partial_{\xi^\alpha}, \quad P^{\alpha\beta} \implies \begin{pmatrix} \mathbf{0}_n & +\mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0}_n \end{pmatrix}, \quad (1.13)$$

where the matrix $P^{\alpha\beta}$ allows us to consider this as a linear mapping from the cotangent bundle, over T^*Q , to the tangent bundle, presented in the current coordinates by this matrix. As intended, this mapping clearly arranges for

$$\tilde{X}_H = \tilde{\Delta} \implies \frac{d}{dt} F = \tilde{X}_H(F) + F_{,t}. \quad (1.14)$$

It is important to point out that not all vector fields are Hamiltonian vector fields for some function, as will be discussed later.

We may also use this matrix to create a **Poisson structure** over the cotangent space over T^*Q , i.e., a (skew-symmetric) bilinear map of two 1-form fields into a function, which is then the same as the Poisson bracket when the 1-forms are exterior derivatives of functions:

$$\begin{aligned} P(\alpha, \beta) &\equiv P^{\mu\nu} \alpha_\mu \beta_\nu \\ \implies P(df, dg) = \{f, g\} &= -f_{,p_b} g_{,q^b} + f_{,q^a} g_{,p_a} = -\tilde{X}_f(g) + \tilde{X}_g(f) = df(\tilde{X}_g), \\ \implies \tilde{X}_f = P(\bullet, df) &, \text{ and in particular } \frac{d}{dt} \xi^\alpha = \tilde{X}_H(\xi^\alpha) = \{\xi^\alpha, H\}. \end{aligned} \quad (1.15)$$

We also use this same matrix, but now with lower indices, to create the dual object, ϖ , the canonical **symplectic 2-form**, which maps a pair of vector fields into a function, and gives us yet other ways to characterize the Poisson bracket operation:

$$\begin{aligned} \varpi &\equiv \left. \begin{aligned} &\frac{1}{2} \omega_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta, \\ &\varpi(\tilde{u}, \tilde{v}) \equiv \omega_{\alpha\beta} u^\alpha v^\beta, \end{aligned} \right\}, \quad \omega_{\alpha\beta} \implies \begin{pmatrix} \mathbf{0}_n & +\mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0}_n \end{pmatrix} \\ \implies \varpi(\tilde{X}_f, \tilde{X}_g) &= \{f, g\} = df(\tilde{X}_g) = -dg(\tilde{X}_f). \end{aligned} \quad (1.16)$$

We may then use this 2-form, having been given only one vector, to map vectors from the tangent bundle down to 1-forms in the cotangent bundle; however, we want to be sure and arrange this so that if we used the Poisson structure to move some 1-form up to the tangent bundle that this operation moves it back to the same 1-form; therefore, we re-define our earlier mapping in the following way:

$$\varpi(\tilde{X}_f, \bullet) = df \iff \omega_{\alpha\beta}(\tilde{X}_f)^\alpha = (df)_\beta . \quad (1.17)$$

(Note that the summation is on the “other index” than was the case when we used the Poisson structure in Eqs. (1.13).] This mapping takes a Hamiltonian vector field, and gives us an exact differential, i.e., an exact 1-form. As an exact 1-form is closed, this allows us a useful “integrability condition” for whether or not a given vector field is a Hamiltonian one; i.e.,

$$d\left(\varpi(\tilde{Y}, \bullet)\right) = 0 \iff \exists f \ni \tilde{Y} = \tilde{X}_f . \quad (1.17a)$$

This now allows us to use the canonical symplectic form to give an explicit, coordinate-invariant expression for Hamilton’s equations, an analogue to Eqs. (1.4b) for the Lagrangian approach, namely

$$dH = \varpi(\bullet, \tilde{\Delta}) = \varpi(\bullet, \tilde{X}_H) . \quad (1.18)$$

When the expression in Eq. (1.16) for ϖ is written out explicitly in the $\{q^a, p_b\}$ variables it takes the following form:

$$\varpi = dq^a \wedge dp_a \implies \varpi = -d\boldsymbol{\theta} . \quad (1.19)$$

On the other hand, the converse may also be shown; i.e., if we state that Eq. (1.18) is Hamilton’s equations, and that ϖ is skew-symmetric and non-degenerate, we may show that there exist coordinates such that our matrix presentation for the components is true. It will also be useful to have a more compact method of describing the canonical 1-form in terms of the generic coordinates ξ^α . Therefore, we “divide in half” the matrix that presents the components of ϖ ; we create a matrix that picks out the p_b ’s and the q^a ’s separately, from a given set of ξ^α ’s:

$$\Gamma = ((\gamma_{\alpha\beta})) \implies \begin{pmatrix} \mathbf{0}_n & \mathbf{I}_n \\ \mathbf{0}_n & \mathbf{0}_n \end{pmatrix} \Rightarrow \omega_{\alpha\beta} = \gamma_{\alpha\beta} - \gamma_{\beta\alpha} \implies \boldsymbol{\theta} = p_a dq^a = \{\gamma_{\alpha\beta} \xi^\beta\} d\xi^\alpha . \quad (1.20)$$

Looking now at the properties of the Poisson bracket, in whichever form we write it, there are various important facts that are worth noting:

$$\left. \begin{array}{l} \{q^r, p_s\} = \delta_s^r, \\ \{q^r, q^s\} = 0, \\ \{p_r, p_s\} = 0, \end{array} \right\} ; \quad \{\xi^\alpha, \xi^\beta\} = P^{\alpha\beta} . \quad (1.21)$$

It’s also worth remembering that the Poisson bracket has a number of useful properties: it is obvious that it is linear, and skew symmetric. That it satisfies the product rule is somewhat less clear, and also the Jacobi identity:

$$\begin{aligned} \text{linearity: } & \{\alpha f + t, g\} = \alpha\{f, g\} + \{t, g\}, \quad \alpha \text{ constant} \\ \text{skew-symmetry: } & \{f, g\} = -\{g, f\}, \\ \text{derivation: } & \{t, fg\} = \{t, f\}g + f\{t, g\}, \\ \text{Jacobi identity: } & \{r, \{s, t\}\} + \{s, \{t, r\}\} + \{t, \{r, s\}\} = 0 . \end{aligned} \quad (1.22)$$

A more interesting property is the one that tells us how it changes along a trajectory:

$$\frac{d}{dt}\{f, g\} = \left\{ \frac{df}{dt}, g \right\} + \left\{ f, \frac{dg}{dt} \right\}, \quad \begin{cases} \frac{d}{dt}f = \{f, H\} + f_{,t}, \\ \frac{d}{dt}g = \{g, H\} + g_{,t}, \end{cases} \quad (1.23)$$

which is true for general, possibly time-dependent functions, as in Eq. (1.14). Although a proof is perfectly straightforward by calculation, it could also be noted that this equation follows from the invariance of the symplectic form under dragging along the trajectory, as will be shown below.

At this point it is interesting to discuss an interesting and important fact about the symplectic form itself: it is invariant when dragged along the flow generated by the Hamiltonian:

$$\mathcal{L}_{\tilde{X}_H} \omega = 0. \quad (1.24)$$

This invariance is important because it leads to a proof of the invariance of a set of $2p$ -forms referred to as the Poincaré invariants. Therefore, we will give a quick proof here, which will also have the advantage that it helps us to understand (or recall) some of the important properties of the Lie derivative operator. As mentioned at Eq. (1.4c), the Lie derivative, along the flow of some vector field, when acting on a function gives simply the action of that vector field on the function. Also the Lie derivative commutes with exterior differentiation; therefore, we may write

$$\begin{aligned} \mathcal{L}\{f dg\} &= (\mathcal{L}f) dg + f d(\mathcal{L}g) \\ \implies \mathcal{L}\{df \wedge dg\} &= [d(\mathcal{L}f)] \wedge dg + df \wedge [d(\mathcal{L}g)]. \end{aligned} \quad (1.25)$$

We may then apply this to the form for ω given in Eq. (1.19):

$$\begin{aligned} \mathcal{L}_{\tilde{X}_H} \omega &= \mathcal{L}_{\tilde{X}_H} \{dq^a \wedge dp_a\} = d[\mathcal{L}_{\tilde{X}_H} q^a] \wedge dp_a + dq^a \wedge d[\mathcal{L}_{\tilde{X}_H} p_a] \\ &= d\left[\frac{d}{dt}q^a\right] \wedge dp_a + dq^a \wedge d\left[\frac{d}{dt}p_a\right] = d\left(\frac{\partial H}{\partial p_a}\right) \wedge dp_a - dq^a \wedge d\left(\frac{\partial H}{\partial q^a}\right) \\ &= \frac{\partial^2 H}{\partial p_a \partial q^b} dq^b \wedge dp_a + \frac{\partial^2 H}{\partial p_a \partial p_b} dp_b \wedge dp_a - \frac{\partial^2 H}{\partial q^a \partial q^b} dq^a \wedge dq^b - \frac{\partial^2 H}{\partial q^a \partial p_b} dq^a \wedge dp_b \\ &= \frac{\partial^2 H}{\partial p_a \partial q^b} (dq^b \wedge dp_a - dq^b \wedge dp_a) = 0. \end{aligned} \quad (1.26)$$

It should be noted that one may give much more abstract, elegant proofs of this if they are really desired; they all depend on first proving, or believing, ‘‘Cartan’s magic formula’’ for the general action of a Lie derivative on a p -form, β : $\mathcal{L}_{\tilde{Y}} \beta = \tilde{Y} \lrcorner d\beta + d(\tilde{Y} \lrcorner \beta)$. Here the definition of the symbol \lrcorner is the following: $\tilde{Y} \lrcorner \beta \equiv \beta(\tilde{Y}, \bullet, \bullet, \dots)$; i.e., \lrcorner inserts the vector on its left into the first available slot in the object on its right. [Note that José and Saletan use another standard symbol for this operation; they use $i_{\tilde{Y}}$ for $\tilde{Y} \lrcorner$, although there are occasional inconsistencies in their usage.] The concept of step, or insertion, may be useful later one; however, the use, and proof, of this more abstract formula did not seem to be necessary here. Such a proof may be found, for instance, on p. 278 of José and Saletan.

Now let us consider the Poincaré invariants. They say that all the wedge products of ω with itself are invariant, as one moves along any given trajectory. This is more or less obvious, given the

above: we simply note that $\mathcal{L}\{(\varpi)^{\wedge n}\} = n(\varpi)^{\wedge(n-1)}\mathcal{L}\varpi$. Nonetheless, the entire subject deserves some more attention: the set of all these Grassmann powers of ϖ are referred to as the *differential form of the Poincaré invariants*. In particular, we begin by looking at the n -th Grassmann product of ϖ with itself. (Note that since ϖ is a 2-form, its products with itself are symmetric, rather than skew-symmetric as it would be for 1-forms.) As this is a $2n$ -dimensional space, the n th power is a $2n$ -form, and therefore may be taken as a choice for the volume element, \mathcal{V} , on the phasespace. Therefore, the volume element is invariant; i.e., the volume element is preserved as one moves along any given trajectory. If, **in particular**, we consider some dense set of initial points in T^*Q , which constitute initial conditions for some set of nearby trajectories, then they fill some initial volume in phasespace, say V_0 . One would obtain the value of V_0 by picking out a set of $2n$ vectors, in the tangent space(s) over this volume, and integrating the volume form over it. (We omit many details concerning integration over manifolds, considering them as mathematical details not needed for our own concerns.) Our theorem then tells us that if we move along all these trajectories, that begin in V_0 , following each one of these dense set of trajectories, then at any later time t , the volume filled by the points on all those trajectories, at that time, will be the same as when we began. Of course the shape may change quite substantially. [Also see Ch. 4 of Percival and Richards for rather a yet more pedestrian approach to a proof of the theorem concerning conservation of volume, for the case when $n = 1$ and pictures may be drawn.]

While, in various cases, many of the different powers of the invariant ϖ may be important, we will concentrate here only on two, that are most useful. We have already mentioned that the total phasespace volume, a $2n$ -dimensional quantity, is invariant along trajectories. The other most important one is ϖ itself, i.e., the first power. For it in particular I invoke Stoke's theorem for integration: for a given exact 1-form, $\beta = d\alpha$, integrated over a $2(n - m)$ -dimensional "area" Σ , with boundary $\partial\Sigma$, this theorem tells us that

$$\int_{\Sigma} d\alpha = \oint_{\partial\Sigma} \alpha . \quad (1.27)$$

Applying this theorem to ϖ , and noting that any integral of it over whatever region will be invariant along trajectories (with the same Hamiltonian, i.e., on the same phasespace), we label it as I_1 , and say that

$$I_1 \equiv \int_{\Sigma} \varpi = - \int_{\Sigma} d\boldsymbol{\theta} = - \oint_{\partial\Sigma} \boldsymbol{\theta} = - \oint_{\partial\Sigma} p_a dq^a . \quad (1.28)$$

This invariant will be important when we consider perturbation theory for mechanical systems.