H

Hamiltonian Systems

J.M. Sanz-Serna
Departamento de Matemática Aplicada, Universidad de Valladolid, Valladolid, Spain

Synonyms

Canonical systems

Definition

Let I be an open interval of the real line $\mathbb{R}$ of the variable t (time) and Ω a domain of the Euclidean space $\mathbb{R}^d \times \mathbb{R}^d$ of the variables (p, q), $p = (p_1, \dots, p_d)$, $q = (q_1, \dots, q_d)$. If $H(p, q; t)$ is a real smooth function defined in $\Omega \times I$, the *canonical* or *Hamiltonian* system associated with H is the system of $2d$ scalar ordinary differential equations

$$\frac{d}{dt} p_i = -\frac{\partial H}{\partial q_i}(p, q; t),$$
$$\frac{d}{dt} q_i = +\frac{\partial H}{\partial p_i}(p, q; t), \qquad i = 1, \dots, d. \quad (1)$$

The function H is called the *Hamiltonian*, d is the *number of degrees of freedom,* and Ω the *phase space*. Systems of the form (1) (which may be generalized in several ways, see below) are ubiquitous in the applications of mathematics; they appear whenever dissipation/friction is absent or negligible.

It is sometimes useful to rewrite (1) in the compact form

$$\frac{d}{dt} y = J^{-1} \nabla H(y; t), \qquad (2)$$

where $y = (p, q)$, $\nabla H = (\partial H/\partial p_1, \dots, \partial H/\partial p_d; \partial H/\partial q_1, \dots, \partial H/\partial q_d)$ and

$$J = \begin{bmatrix} 0_{d\times d} & I_{d\times d} \\ -I_{d\times d} & 0_{d\times d} \end{bmatrix}. \qquad (3)$$

Origin of Hamiltonian Systems

Newton's Second Law in Hamiltonian Form

Consider the motion of a system of N point masses in three-dimensional space (cases of interest range from stars or planets in celestial mechanics to atoms in molecular dynamics). If $\mathbf{r}_j$ denotes the radius vector joining the origin to the j-th point, Newton's equations of motion read

$$m_j \ddot{\mathbf{r}}_j = \mathbf{F}_j, \qquad j = 1, \dots, N. \qquad (4)$$

In the conservative case, where the force $\mathbf{F}_j$ is the gradient with respect to $\mathbf{r}_j$ of a scalar potential V, that is,

$$\mathbf{F}_j = -\nabla_{\mathbf{r}_j} V(\mathbf{r}_1, \dots, \mathbf{r}_N; t), \qquad j = 1, \dots, N, \qquad (5)$$

the system (4) may be rewritten in the Hamiltonian form (1) with $d = 3N$ by choosing, for $j = 1, \dots, N$, $(p_{3j-2}, p_{3j-1}, p_{3j})$ as the cartesian components of the momentum $\mathbf{p}_j = m_j \dot{\mathbf{r}}_j$ of the j-th mass, $(q_{3j-2}, q_{3j-1}, q_{3j})$ as the cartesian components of $\mathbf{r}_j$, and setting

B. Engquist (ed.), *Encyclopedia of Applied and Computational Mathematics*,
DOI 10.1007/978-3-540-70529-1

$$H = T + V, \qquad T = \frac{1}{2}\sum_{j=1}^{N}\frac{1}{m_j}\mathbf{p}_j^2 = \frac{1}{2}p^T M^{-1} p \tag{6}$$

(here M is the $3N \times 3N$ diagonal mass matrix $\operatorname{diag}(m_1, m_1, m_1; \ldots; m_N, m_N, m_N)$). The Hamiltonian H coincides with the total, kinetic + potential, mechanical energy.

Lagrangian Mechanics in Hamiltonian Form

Conservative systems $\mathcal{S}$ more complicated than the one just described (e.g., systems including rigid bodies and/or constraints) are often treated within the Lagrangian formalism [1,3], where the configuration of $\mathcal{S}$ is (locally) described by d (independent) *Lagrangian coordinates* q_i. For instance, the motion of a point on the surface of the Earth – with two degrees of freedom – may be described by the corresponding longitude and latitude, rather than by using the three (constrained) cartesian coordinates. The movements are then governed by the coupled second-order differential equations

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \qquad i = 1, \ldots, d, \tag{7}$$

where $\mathcal{L} = \mathcal{L}(q, \dot{q}; t)$ is the Lagrangian function of $\mathcal{S}$. For each $i = 1, \ldots, d$, $p_i = \partial\mathcal{L}/\partial\dot{q}_i$ represents the *generalized momentum* associated with the coordinate q_i. Under not very demanding hypotheses, the transformation $(\dot{q}, q) \mapsto (p, q)$ may be inverted (i.e., it is possible to retrieve the value of the velocities $\dot{q}$ from the knowledge of the values of the momenta p and coordinates q) and then (7) may be rewritten in the form (1) with $H(p, q; t) = p^T\dot{q} - \mathcal{L}(q, \dot{q}; t)$, where, in the right-hand side, it is understood that the velocities have been expressed as functions of p and q (this is an instance of a *Legendre's transformation,* see [1], Sect. 14). The function H often corresponds to the total mechanical energy in the system $\mathcal{S}$.

Calculus of Variations

According to *Hamilton's variational principle of least action* (see, e.g., Sect. 13 in [1] or Sects. 2-1–2-3 in [3]), the motions of the mechanical system $\mathcal{S}$, we have just described, are extremals of the functional (*action*)

$$\int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t); t)\, dt; \tag{8}$$

in fact (7) are just the Euler-Lagrange equations associated with (8). The evolutions of many (not necessarily mechanical) systems are governed by variational principles for functionals of the form (8). The corresponding Euler-Lagrange equations (7) may be recast in the first order format (1) by following the procedure we have just described ([2], Vol. I, Sects. IV and 9). In fact Hamilton first came across differential equations of the form (1) when studying Fermat's variational principle in geometric optics.

The Hamiltonian Formalism in Quantum and Statistical Mechanics

In the context of classical mechanics the transition from the Lagrangian format (7) to the Hamiltonian format (1) is mainly a matter of mathematical convenience as we shall discuss below. On the contrary, in other areas, including for example, quantum and statistical mechanics, the elements of the Hamiltonian formalism are essential parts of the physics of the situation. For instance, the statistical canonical ensemble associated with (4)–(5) possesses a density proportional to $\exp(-\beta H)$, where H is given by (6) and β is a constant.

First Integrals

Assume that, for some index i_0, the Hamiltonian H is independent of the variable q_{i_0}. It is then clear from (1) that, for any solution $(p(t), q(t))$ of (1) the value of p_{i_0} remains constant; in other words the function p_{i_0} is a *first integral or conserved quantity* of (1). In mechanics, this is expressed by saying that the momentum p_{i_0} conjugate to the *cyclic coordinate* q_{i_0} is a *constant of motion;* for instance, in the planar motion of a point mass in a central field, the polar angle is a cyclic coordinate and this implies the conservation of angular momentum (second Kepler's law). Similarly q_{i_0} is a first integral whenever H is independent of p_{i_0}.

In the *autonomous* case where the Hamiltonian does not depend explicitly on t, that is, $H = H(p, q)$ a trivial computation shows that for solutions of (1) $(d/dt)H(p(t), q(t)) = 0$, so that H is a constant of motion. In applications this often expresses the *principle of conservation of energy*.

Canonical Transformations

The study of Hamiltonian systems depends in an essential way on that of canonical or symplectic transformations.

Definition

With the compact notation in (2), a differentiable transformation $y^* = (p^*, q^*) = \Psi(y)$, $\Psi : \Omega \to \mathbb{R}^d \times \mathbb{R}^d$ is called *canonical (or symplectic)* if its Jacobian matrix $\Psi'(y)$, with (i, j) entry $\partial y_i^*/\partial y_j$, satisfies, for each $y = (p, q)$ in Ω,

$$\Psi'(y)^T J\, \Psi'(y) = J. \tag{9}$$

The composition of canonical transformations is canonical; the inverse of a canonical transformation is also canonical.

By equating the entries of the matrices in (9) and taking into account the skew-symmetry, one sees that (9) amounts to $d(2d-1)$ independent scalar equations for the derivatives $\partial y_i^*/\partial y_j$. For instance, for $d = 1$, (9) is equivalent to the single relation

$$\frac{\partial p^*}{\partial p}\frac{\partial q^*}{\partial q} - \frac{\partial p^*}{\partial q}\frac{\partial q^*}{\partial p} \equiv 1. \tag{10}$$

Simple examples of canonical transformations with $d = 1$ are the rotation

$$p^* = \cos(\theta) p - \sin(\theta) q, \qquad q^* = \sin(\theta) p + \cos(\theta) q \tag{11}$$

and the hyperbolic rotation $p^* = \exp(\theta) p$, $q^* = \exp(-\theta) q$ (θ is an arbitrary constant).

Geometric Interpretation

Consider first the case $d = 1$ where, in view of (10), canonicity means that the Jacobian determinant $\Delta = \det\big(\partial(p^*, q^*)/\partial(p, q)\big)$ takes the constant value 1. The fact $|\Delta| = 1$ entails that for any bounded domain $D \subset \Omega$, the areas of D and $\Psi(D)$ coincide. Furthermore $\Delta > 0$ means that Ψ is orientation preserving. Thus, the triangle with vertices $A = (0, 0)$, $B = (1, 0)$, $C = (0, 1)$ cannot be symplectically mapped onto the triangle with vertices $A^* = (0, 0)$, $B^* = (1, 0)$, $C^* = (0, -1)$ in spite of both having the same area, because the boundary path $A^* \to B^* \to C^* \to A^*$ is oriented clockwise and $A \to B \to C \to A$ has the opposite orientation. One may say that, when $d = 1$, a transformation is canonical if and only if it preserves oriented area.

For $d > 1$ the situation is similar, if slightly more complicated to describe. It is necessary to consider two-dimensional bounded surfaces $D \subset \Omega$ and orient them by choosing one of the two orientations of the boundary curve ∂D. The surface D is projected onto each of the d two-dimensional planes of the variables (p_i, q_i) to obtain d two-dimensional domains $\Pi_i(D)$ with oriented boundaries; then we compute the number $S(D) = \sum_i \pm\text{Area}(\Pi_i(D))$, where, when summing, a term is taken with the $+$ (resp. with the $-$) sign if the orientation of the boundary of $\Pi_i(D)$ coincides with (resp. is opposite to) the standard orientation of the (p_i, q_i) plane. Then a transformation Ψ is canonical if and only if $S(D) = S(\Psi(D))$ for each D.

In Euclidean geometry, a planar transformation that preserves distances automatically preserves areas. Similarly, it may be shown that the preservation of the sum $S(D)$ of oriented areas implies the preservation of similar sums of oriented 4–, 6–, ..., $2d$-dimensional measures (the so-called *Poincaré integral invariants*). In particular a symplectic transformation preserves the orientation of the $2d$-dimensional phase space (i.e., its Jacobian determinant Δ is > 0) and also preserves volume: for any bounded domain $V \subset \Omega$, the volumes (ordinary Lebesgue measures) of V and $\Psi(V)$ coincide.

The preceding considerations (and for that matter most results pertaining to the Hamiltonian formalism) are best expressed by using the language of differential forms. Lack of space makes it impossible to use that alternative language here and the reader is referred to [1], Chap. 8 (see also [6], Sect. 2.4).

Changing Variables in a Hamiltonian System

Assume that in (1) we perform an invertible change of variables $y = \chi(z)$ where χ is *canonical*. A straightforward application of the chain rule shows that the new system, that is, $(d/dt)z = (\chi'(z))^{-1} J^{-1} \nabla_y H(\chi(z); t)$, coincides with the Hamiltonian system $(d/dt)z = J^{-1}\nabla_z K(z; t)$ whose Hamiltonian function $K(z; t) = H(\chi(z); t)$ is obtained by expressing the old H in terms of the new variables. In fact, if one looks for a condition on $y = \chi(z)$ that ensures that in the z-variables (1) becomes the Hamiltonian system with Hamiltonian $H(\chi(z); t)$, one easily discovers the definition of canonicity in (9). The same exercise shows

H

that the matrix J in (9) and that appearing in (2) have to be inverses of one another.

This most important result has of course its counterpart in Lagrangian mechanics or, more generally, in the calculus of variations (see [1], Sect. 12D or [2], Vol. I, Sects. IV and 8): to change variables in the Euler-Lagrange equations for (8) it is enough to first change variables in $\mathcal{L}$ and then form the Euler-Lagrange equations associated with the new Lagrangian function. However, in the Lagrangian case, only the change of the d coordinates $q = \Xi(w)$ is at our disposal; the choice of Ξ determines the corresponding formulae for the velocities $\dot{q} = \Xi'(w)\dot{w}$. In the Hamiltonian case, the change $y = \chi(z)$ couples the $2d$-dimensional $y = (p, q)$ with the $2d$-dimensional z and the class of possible transformations is, therefore, much wider. Jacobi's method (see below) takes advantage of these considerations.

Exact Symplectic Transformations

A transformation $(p^*, q^*) = \Psi(p, q)$, $(p, q) \in \Omega$ is said to be *exact symplectic* if

$$pdq - p^* dq^* = \sum_{i=1}^{d} \left(p_i dq_i - p_i^* \sum_{j=1}^{d} \left(\frac{\partial q_i^*}{\partial p_j} dp_j + \frac{\partial q_i^*}{\partial q_j} dq_j \right) \right) \tag{12}$$

is the differential of a real-valued function $S(p, q)$ defined in Ω.

For (12) to coincide with dS it is necessary but not sufficient to impose the familiar $d(2d - 1)$ relations arising from the equality of mixed second order derivatives of S. It is trivial to check that those relations coincide with the $d(2d - 1)$ relations implicit in (9) and therefore exact symplectic transformations are always symplectic. In a simply connected domain Ω, symplectic transformations are also exact symplectic; in a general Ω, a symplectic transformation is not necessarily exact symplectic and, when it is not, the function S only exists locally.

Generating Functions: Hamilton-Jacobi Theory

Generating functions provide a convenient way of expressing canonical transformations.

Generating Function S_1

Given a canonical transformation $(p^*, q^*) = \Psi(p, q)$, let us define locally a function S such that dS is given by (12) and *assume* that $\partial(q^*, q)/\partial(p, q)$ is non-singular. Then, in lieu of (p, q), we may take (q^*, q) as independent variables and express S in terms of them to obtain a new function $S_1(q^*, q) = S(p(q^*, q), q)$, called the *generating function (of the first kind)* of the transformation. From (12)

$$\frac{\partial S_1}{\partial q_i} = p_i, \quad \frac{\partial S_1}{\partial q_i^*} = -p_i^*, \qquad i = 1, \ldots, d; \tag{13}$$

the relations in the first group of (13) provide d coupled equations to find the q_i^* as functions of (p, q) and those in the second group then allow the explicit computation of p^*. For (11) the preceding construction yields $S_1 = -(\cot(\theta)/2)(q^2 - 2\sec(\theta)qq^* + q^{*2})$, (provided that $\sin(\theta) \neq 0$), an expression that, via (13) leads back to (11).

Conversely, if $S_1(q^*, q)$ is any given function and the relations (13) define uniquely (p^*, q^*) as functions of (p, q), then $(p, q) \mapsto (p^*, q^*)$ is a canonical transformation ([1], Sect. 47A).

Other Generating Functions

The construction of S_1 is only possible under the assumption that (q^*, q) may be taken as independent variables. This assumption does not hold in many important cases, including that where Ψ is the identity transformation (with $q^* = q$). It is therefore useful to introduce a new kind of generating function as follows: If (12) is the differential of $S(p, q)$ (perhaps only locally), then $d(p^{*T} q^* + S) = q^{*T} dp^* + p^T dq$. If $\partial(p^*, q)/\partial(p, q)$ is non-singular, (p^*, q) may play the role of independent variables and if we set $S_2(p^*, q) = p^{*T} q^*(p^*, q) + S(p(p^*, q), q)$ it follows that

$$\frac{\partial S_2}{\partial q_i} = p_i, \quad \frac{\partial S_2}{\partial p_i^*} = q_i^*, \qquad i = 1, \ldots, d; \tag{14}$$

here the first equations determine the p_i^* as functions of (p, q) and the second yield the q_i^* explicitly. The function S_2 is called the *generating function of the 2nd kind* of Ψ. The identity transformation is generated by $S_2 = p^{*T} q$. For (11) with $\cos(\theta) \neq 0$ (which ensures that (p^*, q) are independent) we find:

$$S_2 = \frac{\tan(\theta)}{2} \left(q^2 + 2\csc(\theta) q p^* + p^{*2} \right). \tag{15}$$

Conversely if $S_2(p^*, q)$ is any given function and the relations (14) define uniquely (p^*, q^*) as functions of (p, q), then $(p, q) \mapsto (p^*, q^*)$ is a canonical transformation ([1], Sect. 48B).

Further kinds of generating functions exist ([3], Sect. 9-1, [1], Sect. 48).

The Hamilton-Jacobi Equation

In Jacobi's method to integrate (1) (see [1], Sect. 47 and [3], Sect. 10-3) *with time-independent* H, a canonical transformation (14) is sought such that, in the new variables, the Hamiltonian $K = H(p(p^*, q^*), q(p^*, q^*))$ is a function $K = K(p^*)$ of the new momenta alone, that is, all the q_i^* are cyclic. Then in the new variables – as pointed out above – all the p_i^* are constants of motion and therefore the solutions of the canonical equations are given by $p_i^*(t) = p_i^*(0)$, $q_i^*(t) = q_i^*(0) + t(\partial K/\partial p_i^*)_{p^*(0)}$. Inverting the change of variables yields of course the solutions of (1) in the originally given variables (p, q).

According to (14) the required $S_2(q^*, q)$ has to satisfy the *Hamilton-Jacobi* equation

$$H\left(\frac{\partial S_2}{\partial q_1}, \ldots, \frac{\partial S_2}{\partial q_d}, q_1, \ldots, q_d\right) = K(p_1^*, \ldots, p_d^*).$$

This is a first-order partial differential equation ([2], Vol. II, Chap. II) for the unknown S_2 called *the characteristic function;* the independent variables are $(q_1, \ldots, q_d)$ and it is required to find *a particular* solution that includes d independent integration constants $p_1^*, \ldots, p_d^*$ (a *complete* integral in classical terminology). Jacobi was able to identify, via separation of variables, a complete integral for several important problems unsolved in the Lagragian format. His approach may also be used with S_1 and the other kinds of generating functions.

Time-dependent Generating Functions

So far we have considered time-independent canonical changes of variables. It is also possible to envisage changes $(p^*, q^*) = \Psi(p, q; t)$, where, for each fixed t, Ψ is canonical. An example is afforded by (14) if the generating function includes t as a parameter: $S_2 = S_2(p^*, q; t)$. In this case, the evolution of (p^*, q^*) is governed by the Hamiltonian equations (1) associated with the Hamiltonian $K = H + \partial S_2/\partial t$, where in the right-hand side it is understood that the arguments (p, q) of H and (p^*, q) of S_2 have been expressed as functions of the new variables (p^*, q^*) with the help of formulae (14). Note the contribution $\partial S_2/\partial t$ that arises from the time-dependence of the change of variables.

If $S_2 = S_2(p^*, q; t)$ satisfies the Hamilton-Jacobi equation

$$H\left(\frac{\partial S_2}{\partial q_1}, \ldots, \frac{\partial S_2}{\partial q_d}, q_1, \ldots, q_d; t\right) + \frac{\partial S_2}{\partial t} = 0, \quad (16)$$

then the new Hamiltonian K vanishes identically and all p_i^* and q_i^* remain constant; this trivially determines the solutions $(p(t), q(t))$ of (1). In (16) the independent variables are t and the q_i and it is required to find *a* complete solution, that is, *a* solution S_2 that includes d independent integration constants p_i^*. It is easily checked that, conversely, (1) is the *characteristic system* for (16), so that it is possible to determine all solutions of (16) whenever (1) may be integrated explicitly ([2], Vol. II, Chap. II).

Hamiltonian Dynamics

Symplecticness of the Solution Operator

We denote by Φ_{t,t_0}^H the solution operator of (1) (t, t_0 are real numbers in the interval I). By definition, Φ_{t,t_0}^H is a transformation that maps the point (p^0, q^0) in Ω into the value at time t of the solution of (1) that satisfies the initial condition $p(t_0) = p^0$, $q(t_0) = q^0$. Thus, if in $\Phi_{t,t_0}^H(p^0, q^0)$ we keep t_0, p^0, and q^0 fixed and let t vary, then we recover the solution of the initial-value problem given by (1) in tandem with $p(t_0) = p^0$, $q(t_0) = q^0$. However, we shall be interested in seeing t and t_0 as fixed parameters and (p^0, q^0) as a variable so that Φ_{t,t_0}^H represents a transformation mapping the phase space into itself. (It is possible for Φ_{t,t_0}^H not to be defined in the whole of Ω; this happens when the solutions of the initial value problem do not exist up to time t.) Note that $\Phi_{t_2,t_0}^H = \Phi_{t_2,t_1}^H \circ \Phi_{t_1,t_0}^H$ for each t_0, t_1, t_2 (the circle $\circ$ means composition of mappings). In the autonomous case where $H = H(y)$, Φ_{t,t_0}^H depends only on the difference $t - t_0$ and we write $\phi_{t-t_0}^H$ instead of Φ_{t,t_0}^H; then the *flow* ϕ_t^H has the group property: $\phi_{t+s}^H = \phi_t^H \circ \phi_s^H$, for each t and s.

The key geometric property of Hamiltonian systems is that Φ_{t,t_0}^H is, for each fixed t_0 and t, a canonical transformation ([1], Sect. 44). In fact the canonicity of the solution operator is also sufficient for the system to be Hamiltonian (at least locally).

The simplest illustration is provided by the harmonic oscillator: $d = 1$, $H = (1/2)(p^2 + q^2)$. The t-flow $(p^*, q^*) = \phi_t(p, q)$ is of course given by (11) with $\theta = t$, a transformation that, as remarked earlier, is canonical. The group property of the flow is the statement that rotating through θ radians and then through θ' radians coincides with a single rotation of amplitude $\theta + \theta'$ radians.

The Generating Function of the Solution Operator

Assume now that we subject (1) to the t-dependent canonical change of variables $(p^*, q^*) = \Psi^H_{t_0,t}(p, q)$ with t_0 fixed. The new variables (p^*, q^*) remain constant: $(p^*(t), q^*(t)) = \Psi^H_{t_0,t}(p(t), q(t)) = \Psi^H_{t_0,t} \circ \Psi^H_{t,t_0}(p(t_0), q(t_0)) = (p(t_0), q(t_0))$. Therefore, the new Hamiltonian $K(p^*, q^*; t)$ must vanish identically and the generating function S_2 of $\Psi^H_{t_0,t}$ must satisfy (16). Now, as distinct from the situation in Jacobi's method, we are interested in solving the initial value problem given by Hamilton-Jacobi equation (16) and the initial condition $S(p^*, q; t_0) = p^{*T} q$ (for $t = t_0$ the transformation $\Psi^H_{t_0,t}$ is the identity).

As an illustration, for the harmonic oscillator, as noted above, $\Psi^H_{t_0,t}$ is given by (11) with $\theta = t_0 - t$; a simple computation shows that its generating function found in (15) satisfies the Hamilton-Jacobi equation $(1/2)\big((\partial S_2/\partial q)^2 + q^2\big) + \partial S_2/\partial t = 0$.

Symplecticness Constrains the Dynamics

The canonicity of Φ^H_{t,t_0} has a marked impact on the long-time behavior of the solutions of (1). As a simple example, consider a system of two scalar differential equations $\dot{p} = f(p, q)$, $\dot{q} = g(p, q)$ and assume that (p^0, q^0) is an equilibrium where $f = g = 0$. Generically, that is, in the "typical" situation, the equilibrium is hyperbolic: the real parts of the eigenvalues λ_1 and λ_2 of the Jacobian matrix $\partial(f, g)/\partial(p, q)$ evaluated at (p^0, q^0) have nonzero real part and the equilibrium is a sink ($\Re\lambda_1 < 0$, $\Re\lambda_2 < 0$), a source ($\Re\lambda_1 > 0$, $\Re\lambda_2 > 0$), or a saddle ($\Re\lambda_1 > 0$, $\Re\lambda_2 < 0$). The situation where λ_1 and λ_2 are conjugate purely imaginary numbers does not arise typically: small perturbations change it into either a sink or a source. However, if we restrict the attention to Hamiltonian systems the situation changes completely: sinks and sources cannot appear, because in their neighborhood the flow contracts (expands) area. The case $\Re\lambda_1 = 0$, $\Re\lambda_2 = 0$ is now not exceptional: it persists under small Hamiltonian perturbations.

Similar considerations apply to periodic orbits, invariant tori, etc. To sum up, thanks to symplecticness, dynamical features that are exceptional for general systems become the rule for Hamiltonian systems. Conversely features that are typical for general systems cannot arise at all in Hamiltonian problems.

Poisson Brackets

Let us present yet another useful tool of the Hamiltonian formalism. Although some of the results to be discussed are valid for general Hamiltonians $H = H(y; t)$, for simplicity, we shall *assume in the rest of this Encyclopedia entry that all Hamiltonians are autonomous* $H = H(y)$.

Definition

If F, G are smooth real functions defined in the phase space Ω, their *Poisson bracket* is the real function

$$\{F, G\} = \nabla F^T J^{-1} \nabla G, \quad \text{i.e.,}$$

$$\{F, G\} = \sum_{i=1}^{d} \left(\frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i} \right). \qquad (17)$$

Clearly the operation $\{\cdot, \cdot\}$ is bilinear and skew-symmetric, that is, $\{F, G\} = -\{G, F\}$. It furthermore satisfies *Jacobi's identity:* if F, G, and H are smooth functions then $\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$.

Canonical changes of variables do not alter the value of the Poisson bracket: if $y = \chi(z)$ is canonical, then the Poisson bracket of the functions $F(\chi(z))$, $G(\chi(z))$ may be obtained by first computing $\{F, G\}$ and then substituting $y = \chi(z)$. In fact a transformation is canonical if and only if it does not change the value of the Poisson bracket ([6], Remark 12.1).

Poisson Brackets and Hamiltonian Systems

From (17), (1) may be rewritten as $\dot{y}_i = \{y_i, H\}$, $i = 1, \dots, 2d$. More generally, if F is any smooth real function defined in Ω, the value at a point $y^0 \in \Omega$ of $\{F, H\}$ coincides with the rate of change $(d/dt) F(\phi^H_t(y^0))_{t=0}$. This has two important implications ([1], Sect. 40):

(a) F is a first integral of (1) if and only if $\{F, H\} \equiv 0$.
(b) The differential operator $L_{J^{-1}\nabla H}$ associated with the vector field $J^{-1}\nabla H$ in (1) coincides with $F \mapsto \{F, H\}$. (Recall that, given the system

$\dot{y} = f(y)$ with vector field f and flow ϕ_t^f, L_f is, by definition, the differential operator that maps each real function F into the real function that at y takes the value $(d/dt)F(\phi_t^f(y))_{t=0}$. By the chain rule, $L_f F = \sum_i f_i(y)(\partial F/\partial y_i)$.)

In turn, (a) together with Jacobi's identity yield immediately *Poisson's theorem:* The Poisson bracket of two first integrals of (1) is again a first integral. An example: if two of the cartesian components of the angular momentum of a mechanical system are conserved, so is the third.

Assume next that H is kept invariant by a Hamiltonian flow ϕ_t^F, that is, $H \circ \phi_t^F \equiv H$. According to (a), $\{H, F\} \equiv 0$, and by skew-symmetry $\{F, H\} \equiv 0$. A new application of (a) shows that F is a first integral of (1). In this way we have obtained a generalization of a well-known theorem of Noether [1]: to each group of symmetries that leave invariant a mechanical system there corresponds a constant of motion. Here is the simplest example. The flow of $F = p_1$ is given by the translations along the q_1 axis $(p, q) \mapsto (p, q_1 + t, q_2, \ldots, q_d)$, so that H is invariant if and only if q_1 is cyclic. The general result in this paragraph yields, once again, the known statement "the momentum conjugate to a cyclic coordinate is a first integral."

Before we point out some consequences of (b), we recall ([1], Sect. 39C), that, if $f(y)$ and $g(y)$ are vector fields on the same phase space with operators L_f and L_g, then $L_g L_f - L_f L_g$ is the operator L_h associated with a new vector field h, denoted by $h = [f, g]$ and called the *Lie bracket or commutator* of f and g. This notion is relevant in view of the following result: $[f, g]$ vanishes identically if and only if the flows ϕ_t^f and ϕ_t^g commute, that is, $\phi_t^f \circ \phi_s^g = \phi_s^g \circ \phi_t^f$, for each t and s.

From the Jacobi identity and (b) it is easily concluded that the commutator of the Hamiltonian vector fields with Hamiltonian functions F, G is again a Hamiltonian vector field and that the corresponding Hamiltonian is $\{F, G\}$. In particular the flows ϕ_t^F and ϕ_t^G commute if and only if the Hamiltonian vector field associated with $\{F, G\}$ vanishes, that is, if and only if $\{F, G\}$ is (locally) constant.

Integrability: Perturbation Theory

As we have seen in connection with Jacobi's method, the possibility of integrating effectively Hamiltonian system is closely related to the existence of sufficiently many conserved quantities.

The *integrability theorem of Liouville and Arnold* ([1], Sect. 49, [4], Chap. X), that we sketch next, addresses this issue. It is assumed that the system (1) has d (independent) conserved quantities F_i and that these are in involution, that is, $\{F_i, F_j\} = 0$ if $i \neq j$. Each level set of the form $M(a_1, \ldots, a_d) = \{y : F_1(y) = a_1, \ldots, F_d(y) = a_d\}$ is a smooth manifold invariant by the flow ϕ_t^H; furthermore, it may be proved that if the level sets $M(a_1, \ldots, a_d)$ are compact and connected, then each of them will be (diffeomorphic to) a d-dimensional torus. In that case it is possible to compute explicitly (in terms of quadratures) a canonical change of variables $p = p(I, \alpha)$, $q = q(I, \alpha)$ to the so-called *action/angle variables* (I, α) so that the new Hamiltonian K is independent of the α_i and therefore the equations of motion read

$$\dot{I}_i = 0, \quad \dot{\alpha}_i = \frac{\partial K}{\partial I_i}, \qquad i = 1, \ldots, d.$$

The actions I_i are first integrals; their level sets $\{y : I_1(y) = b_1, \ldots, I_d(y) = b_d\}$ coincide with the invariant tori of the dynamics. Each invariant torus is parameterized by the d variables α_i that are angles (increasing them by 2π leads to the starting point in (p, q)). On any fixed torus each α_i varies at a constant angular velocity $\partial K/\partial I_i$, so that the motion is quasi-periodic.

For the harmonic oscillator in non-dimensional form $H = (1/2)(p^2 + q^2)$ the invariant sets are the circles $p^2 + q^2 =$ constant; the canonical change of variables is given by $p = \sqrt{2I}\cos\alpha$, $q = \sqrt{2I}\sin\alpha$, so that $I = H$. (In dimensional variables the action I would be the ratio of the energy H to the frequency of oscillation.)

When the hypotheses of the Arnold-Liouville theorem hold, the dynamics of (1) are perfectly understood. At the other end of the spectrum, the behavior of the solutions of Hamiltonian systems away from integrability may be bewildering complicated. An intermediate situation is that where the system, without being integrable, may be seen as a small perturbation of an integrable one. The literature contains many important results on perturbation theory. The most celebrated is the *Kolmogorov-Arnold-Moser (KAM) theorem* ([1], Sect. 49, [4], Chap. X) that ensures that, under suitable hypotheses, most invariant tori of the unperturbed case do not disappear under perturbation. The book [5] gathers a number of important contributions to the study of Hamiltonian dynamics.

Extensions

The canonical format (1) is only the simplest and historically first of a series of Hamiltonian formats that appear in the applications. Here are more formats:

Changing the Structure Matrix

It is possible ([4], Chap. VII), while keeping the form in (2), to replace the so-called structure matrix J defined in (3) by a more general *invertible*, skew-symmetric matrix $\widetilde{J}(y)$ (note the dependence on y and that the dimension of the phase space is still necessarily even as skew-symmetric matrices of odd dimension are singular). Most of the theory goes through provided that the associated Poisson bracket (defined as in the first equality in (17)) satisfies the Jacobi identity. In this setup it is also possible to define the symplecticness of a transformation via (9). The matrix $\widetilde{J}(y)$ defines then a *noncanonical symplectic structure*.

Poisson Structures

Another possibility ([6], Sect. 14.5, [4], Chap. VII) is to use (2) with J^{-1} replaced by a *non-invertible*, skew-symmetric matrix $B(y)$. Again $B(y)$ has to be chosen in such a way that the Jacobi identity for the Poisson bracket (defined by the first equality in (17) with $B(y)$ in lieu of J^{-1}) holds. Here it is not possible to generalize the definition in (9), which would require the inverse of the non-invertible $B(y)$; there is no symplectic structure and one speaks of a Poisson structure. Note that the dimension of the phase space is not necessarily even. A salient feature of Poisson structures is the existence of *Casimir functions* C such that $\nabla C(y)^T B(y) \equiv 0$. Since $\{C, H\} = 0$ if C is a Casimir function and H arbitrary, Casimir functions are constants of motion for all systems of the form $\dot{y} = B(y)\nabla H(y)$, regardless of the choice of Hamiltonian H.

Differential Geometry

So far all variables have been points in Euclidean spaces. However, symplectic and Poisson structures may be defined on manifolds [1] and in fact, in many applications, the problems investigated appear naturally in a manifold context and only a, more or less arbitrary, choice of local coordinates allows to rephrase them in a Euclidean setting.

Hamiltonian Partial Differential Equations

Many evolutionary partial differential equations may also be understood as (infinite dimensional) Hamiltonian systems. Typically, each point u in phase space is a smooth real or vector-valued function of one or more spatial variables. The real functions F, H, ... defined in phase space are functionals and the operator ∇ in (2) is replaced by the variational derivative $\delta/\delta u$. An example follows, but very many other exist including the Korteweg-de Vries equation, linear and nonlinear Schroedinger equations, etc. (see [6], Sect. 14.7). Assume that $u = (p, q)$ with p, q smooth real functions of the variable x, $0 \le x \le 1$, satisfying homogeneous Dirichlet boundary conditions. If H is the functional

$$H(u) = \frac{1}{2}\int_0^1 \left(p(x)^2 + q_x(x)^2\right) dx$$

then (q_{xx} appears after integrating by parts)

$$\begin{aligned} &H(u + \epsilon\widetilde{u}) \\ &\quad = H(u) + \epsilon \int_0^1 \left(p(x)\widetilde{p}(x) - q_{xx}(x)\widetilde{q}(x)\right) dx \\ &\qquad + \mathcal{O}(\epsilon^2). \end{aligned}$$

Therefore, $\delta H/\delta p = p$, $\delta H/\delta q = -q_{xx}$ and we have the following Hamiltonian system (note the analogy with (1) with i replaced by x)

$$\frac{\partial}{\partial t} p = -\frac{\delta H}{\delta q} = q_{xx}, \qquad \frac{\partial}{\partial t} q = \frac{\delta H}{\delta p} = p,$$

where, after eliminating p, we recognize the familiar wave equation.

References

1. Arnold, V.I.: Mathematical Methods of Classical Mechanics, 2nd edn. Springer, New York (1989)
2. Courant, R., Hilbert, D.: Methods of Mathematical Physics. Wiley, New York (1989)
3. Goldstein, H.: Classical Mechanics, 2nd edn. Addison-Wesley, Reading (1980)
4. Hairer, E., Lubich, Ch., Wanner, G.: Geometric Numerical Integration, 2nd edn. Springer, Berlin (2006)

5. MacKay, R.S., Meiss, J.D.: Hamiltonian Dynamical Systems. Adam Hilger, Bristol (1987)
6. Sanz-Serna, J.M., Calvo, M.P.: Numerical Hamiltonian Problems. Chapman, London (1994)

Hamilton–Jacobi Equations

Emiliano Cristiani
Istituto per le Applicazioni del Calcolo "Mauro Picone", Consiglio Nazionale delle Ricerche, Rome, RM, Italy

Mathematics Subject Classification

35F21; 49Lxx; 70H20

Synonyms

HJ equations

Definition

The Hamilton–Jacobi equation (HJE) is a first-order nonlinear partial differential equation. The HJE first appeared in the studies of W. R. Hamilton (1805–1865) and C. G. J. Jacobi (1804–1851) in the field of classical mechanics [7]. The interest of mathematicians started in the 1950s and grew considerably since the 1980s with the introduction of the theory of *viscosity solutions* [2,3]. Nowadays, it is encountered in problems of mechanics, geometry, optics, front propagation, computer vision, optimal control, and differential games. The general form of the HJE is

$$\frac{\partial u}{\partial t}(x,t) + H(x,t,u(x,t),D_x u(x,t)) = 0, \quad x \in \Omega,\ t > 0\,,$$

where Ω is an open domain of $\mathbb{R}^n$, $x = (x_1, \ldots, x_n)$, $u : \Omega \times (0, +\infty) \to \mathbb{R}$ is the unknown, the *Hamiltonian* $H : \mathbb{R}^n \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is given, and $D_x = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right)$.

The HJE can be also written in an equivalent time-independent form

$$\widehat{H}(y, u(y), D_y u(y)) = 0, \quad y \in \widehat{\Omega}\,,$$

defining $y := (x,t)$, $\widehat{\Omega} := \Omega \times (0, +\infty)$, and, for any $p \in \mathbb{R}^{n+1}$,

$$\widehat{H}(y,u,p) := p_{n+1} + H(y_1, \ldots, y_n, y_{n+1}, u, p_1, \ldots, p_n)\,, \quad y \in \widehat{\Omega}.$$

Original Formulation in Classical Mechanics

Consider a system described by the generalized coordinates $q = q(t) \in \mathbb{R}^n$, the generalized velocities $\dot{q}(t)$, and the Lagrangian function $L(q, \dot{q}, t)$. The *Hamiltonian* H of the system is

$$H(q,p,t) := p \cdot \dot{q}(q,p,t) - L(q, \dot{q}(q,p,t), t)$$

where $p_i(t) = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t)$, $i = 1, \ldots, n$ are the coordinates of the generalized momentum and $\dot{q}$ is written as a function of (q,p,t). For any (t_0, q_0), define

$$S(x,t) := \inf \left\{ \int_{t_0}^{t} L(q(s), \dot{q}(s), s)ds \right\}$$

where the infimum is taken over all C^1 trajectories $q(\cdot)$ starting from q_0 at time t_0 and ending at x at time t. Then, the function $S(x,t)$ is solution of the HJE [5,7]

$$\frac{\partial S}{\partial t}(x,t) + H(x, D_x S(x,t), t) = 0.$$

Theoretical Results

It is easy to see that the HJE equation can lack of classical solutions (i.e., of class C^1) while can have multiple *weak* solutions (i.e., solutions which are a.e. differentiable and satisfy the equation where differentiable). Consider, for example, the one-dimensional *eikonal equation* $|D_x u| = 1$, $x \in [-1,1]$, complemented with boundary conditions $u(-1) = u(1) = 0$. Both functions $u_1(x) = -|x| + 1$ and $u_2(x) = |x| - 1$ are weak solutions.

Existence and uniqueness results can be achieved by means of the notion of *viscosity solution* [2, 3].