Lecture 1: Hamiltonian systems

Table of contents

The main topic of this lecture[1] is a deeper understanding of Hamiltonian systems

$$\dot p = -\nabla_q H(p,q), \qquad \dot q = \nabla_p H(p,q). \tag{1}$$

Here, p and q are vectors in $\mathbb{R}^d$, and $H(p,q)$ is a scalar sufficiently differentiable function. It is called the 'Hamiltonian' or the 'total energy'.

1 Derivation from Lagrange's equation

Suppose that the position of a mechanical system with d degrees of freedom is described by $q = (q_1, \ldots, q_d)^T$ as *generalized coordinates* (this can be for example Cartesian coordinates, angles, arc lengths along a curve, etc.). Consider the *Lagrangian*

$$L = T - U, \tag{2}$$

where $T = T(q,\dot q)$ denotes the kinetic energy and $U = U(q)$ the potential energy. The motion of the system is described by Lagrange's equation[2]

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q}\right) = \frac{\partial L}{\partial q}\ , \tag{3}$$

[1] Most parts of this manuscript are taken from the monograph *Geometric Numerical Integration* by Hairer, Lubich & Wanner (2nd edition, Springer Verlag 2006).

[2] Lagrange, *Applications de la méthode exposée dans le mémoire précédent a la solution de différents problèmes de dynamique*, 1760, Oeuvres Vol. 1, 365–468.

which are just the Euler–Lagrange equations of the variational problem $S(q) = \int_a^b L\big(q(t), \dot q(t)\big)\,\mathrm{d}t \to \min.$

Hamilton[3] simplified the structure of Lagrange's equations and turned them into a form that has remarkable symmetry, by

* introducing Poisson's variables, the conjugate *momenta*

$$p_k = \frac{\partial L}{\partial \dot q_k}(q, \dot q) \qquad \text{for} \quad k = 1, \ldots, d, \tag{4}$$

* considering the *Hamiltonian*

$$H := p^T \dot q - L(q, \dot q) \tag{5}$$

  as a function of p and q, i.e., taking $H = H(p, q)$ obtained by expressing $\dot q$ as a function of p and q via (4).

Here it is, of course, required that (4) defines, for every q, a continuously differentiable bijection $\dot q \leftrightarrow p$. This map is called the *Legendre transform*.

Theorem 1 *Lagrange's equations (3) are equivalent to Hamilton's equations*

$$\dot p_k = -\frac{\partial H}{\partial q_k}(p, q), \qquad \dot q_k = \frac{\partial H}{\partial p_k}(p, q), \qquad k = 1, \ldots, d. \tag{6}$$

Proof. The definitions (4) and (5) for the momenta p and for the Hamiltonian H imply that

$$\begin{aligned} \frac{\partial H}{\partial p} &= \dot q^T + p^T \frac{\partial \dot q}{\partial p} - \frac{\partial L}{\partial \dot q}\frac{\partial \dot q}{\partial p} = \dot q^T, \\ \frac{\partial H}{\partial q} &= p^T \frac{\partial \dot q}{\partial q} - \frac{\partial L}{\partial q} - \frac{\partial L}{\partial \dot q}\frac{\partial \dot q}{\partial q} = -\frac{\partial L}{\partial q}. \end{aligned}$$

The Lagrange equations (3) are therefore equivalent to (6). □

[3] Sir W.R. Hamilton, *On a general method in dynamics; by which the study of the motions of all free systems of attracting or repelling points is reduced to the search and differentiation of one central relation, or characteristic function*, Phil. Trans. Roy. Soc. Part II for 1834, 247–308; Math. Papers, Vol. II, 103–161.

Case of quadratic kinetic energy. If $T(q, \dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q}$, where $M(q)$ is a symmetric and positive definite matrix, we have $p = M(q)\dot{q}$. Replacing the variable $\dot{q}$ by $M(q)^{-1}p$ in the definition (5) of $H(p, q)$, we obtain

$$\begin{aligned} H(p,q) &= p^T M(q)^{-1}p - L\big(q, M(q)^{-1}p\big) \\ &= p^T M(q)^{-1}p - \frac{1}{2}\,p^T M(q)^{-1}p + U(q) \;=\; \frac{1}{2}\,p^T M(q)^{-1}p + U(q) \end{aligned}$$

and the Hamiltonian is $H = T + U$, which is the *total energy*.

2 Energy conservation and first integrals, examples

Definition 1 A non-constant function $I(y)$ is a *first integral* of $\dot{y} = f(y)$ if

$$I'(y)f(y) = 0 \qquad \text{for all } y. \tag{7}$$

This is equivalent to the property that *every* solution $y(t)$ of $\dot{y} = f(y)$ satisfies $I\big(y(t)\big) = Const$.

Example 1 (Conservation of the total energy) For Hamiltonian systems (1) the Hamiltonian function $H(p, q)$ is a first integral.

Example 2 (Conservation of the total linear and angular momentum) *We consider a system of N particles interacting pairwise with potential forces depending on the distances of the particles. This is a Hamiltonian system with total energy*

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{m_i}\,p_i^T p_i + \sum_{i=2}^{N}\sum_{j=1}^{i-1} V_{ij}\Big(\|q_i - q_j\|\Big)\ .$$

Here $q_i, p_i \in \mathbb{R}^3$ represent the position and momentum of the ith particle of mass m_i, and $V_{ij}(r)$ ($i > j$) is the interaction potential between the ith and jth particle. The equations of motion read

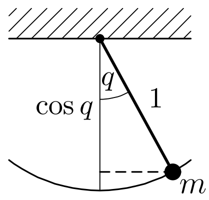
$$\dot{q}_i = \frac{1}{m_i}\,p_i\ , \qquad \dot{p}_i = \sum_{j=1}^{N}\nu_{ij}\,(q_i - q_j)$$

where, for $i > j$, we have $\nu_{ij} = \nu_{ji} = -V'_{ij}(r_{ij})/r_{ij}$ with $r_{ij} = \|q_i - q_j\|$. The conservation of the total linear and angular momentum

$$P = \sum_{i=1}^{N} p_i, \qquad L = \sum_{i=1}^{N} q_i \times p_i$$

is a consequence of the symmetry relation $\nu_{ij} = \nu_{ji}$:

Example 3 (Mathematical pendulum) The mathematical pendulum (mass $m = 1$, massless rod of length $\ell = 1$, gravitational acceleration $g = 1$) is a system with one degree of freedom having the Hamiltonian



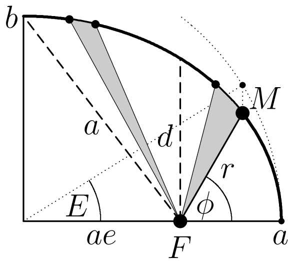
$$H(p, q) = \frac{1}{2} p^2 - \cos q,$$

so that the equations of motion (1) become

$$\dot{p} = -\sin q, \qquad \dot{q} = p. \tag{8}$$

Figure 3 below shows some level curves of $H(p, q)$. By Example 1, the solution curves of the problem (8) lie on such level curves.

Example 4 (Two-body problem or Kepler problem) For computing the motion of two bodies (planet and sun) which attract each other, we choose one of the bodies (sun) as the centre of our coordinate system; the motion will then stay in a plane and we can use two-dimensional coordinates $q = (q_1, q_2)$ for the position of the second body. Newton's laws, with a suitable normalization, then yield the following differential equations



$$\ddot{q}_1 = -\frac{q_1}{(q_1^2 + q_2^2)^{3/2}}, \qquad \ddot{q}_2 = -\frac{q_2}{(q_1^2 + q_2^2)^{3/2}}. \tag{9}$$

This is equivalent to a Hamiltonian system with the Hamiltonian

$$H(p_1, p_2, q_1, q_2) = \frac{1}{2}\left(p_1^2 + p_2^2\right) - \frac{1}{\sqrt{q_1^2 + q_2^2}}, \qquad p_i = \dot{q}_i. \tag{10}$$

The planet moves in *elliptic orbits* with the sun at one of the foci (Kepler's[4] first law). With initial values

$$q_1(0) = 1 - e, \quad q_2(0) = 0, \quad \dot{q}_1(0) = 0, \quad \dot{q}_2(0) = \sqrt{\frac{1+e}{1-e}} \tag{11}$$

the solution is an ellipse with eccentricity e ($0 \le e < 1$), $a = 1$, $b = \sqrt{1 - e^2}$, $d = 1 - e^2$, and period 2π. The total energy is $H_0 = -1/2$, and the angular momentum is $L_0 = \sqrt{1 - e^2}$.

[4] J. Kepler, *Astronomia nova* $\alpha\iota\tau\iota o\lambda o\gamma\eta\tau \acute{o} \varsigma$ *seu Physica celestis, traditia commentariis de motibus stellae Martis, ex observationibus G. V. Tychonis Brahe*, Prague 1609.

Example 5 (Hénon–Heiles problem) The polynomial Hamiltonian in two degrees of freedom[5]

$$H(p,q) = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3 \tag{12}$$

is a Hamiltonian differential equation that can have chaotic solutions. Figure 1 shows a regular behaviour of solutions when the value of the Hamiltonian is small, and a chaotic behaviour for large Hamiltonian.

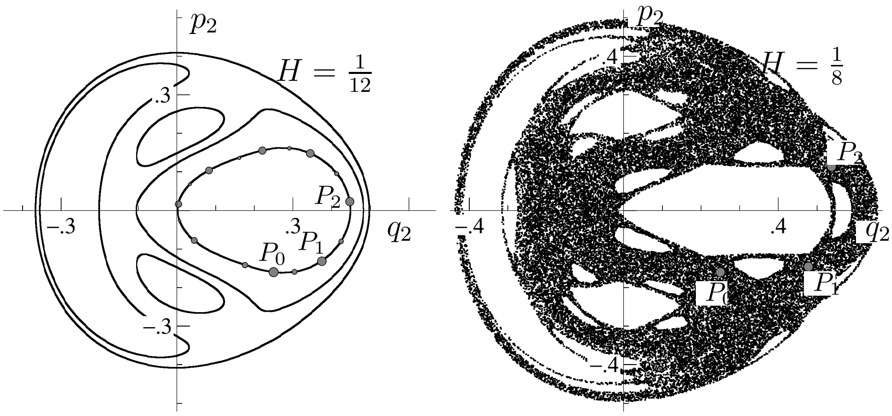


Figure 1: Poincaré cuts for $q_1 = 0$, $p_1 > 0$ of the Hénon–Heiles model for $H = \frac{1}{12}$ (6 orbits, left) and $H = \frac{1}{8}$ (1 orbit, right).

3 Symplectic transformations

The basic objects to be studied are two-dimensional parallelograms lying in $\mathbb{R}^{2d}$. We suppose the parallelogram to be spanned by two vectors

$$\xi = \begin{pmatrix} \xi^p \\ \xi^q \end{pmatrix}, \qquad \eta = \begin{pmatrix} \eta^p \\ \eta^q \end{pmatrix}$$

in the (p,q) space ($\xi^p, \xi^q, \eta^p, \eta^q \in \mathbb{R}^d$) as $P = \{ t\xi + s\eta \mid 0 \le t \le 1,\ 0 \le s \le 1 \}$. In the case $d = 1$ we consider the *oriented area*

$$\text{or.area}\,(P) = \det \begin{pmatrix} \xi^p & \eta^p \\ \xi^q & \eta^q \end{pmatrix} = \xi^p \eta^q - \xi^q \eta^p \tag{13}$$

[5] M. Hénon & C. Heiles, *The applicability of the third integral of motion: some numerical experiments*, Astron. J. 69 (1964) 73–79.

(left picture of Fig. 2). In higher dimensions, we replace this by the *sum of the oriented areas of the projections of P onto the coordinate planes* (p_i, q_i), i.e., by

$$\omega(\xi,\eta) := \sum_{i=1}^{d} \det \begin{pmatrix} \xi_i^p & \eta_i^p \\ \xi_i^q & \eta_i^q \end{pmatrix} = \sum_{i=1}^{d} (\xi_i^p \eta_i^q - \xi_i^q \eta_i^p). \tag{14}$$

This defines a bilinear map acting on vectors of $\mathbb{R}^{2d}$, which will play a central role for Hamiltonian systems. In matrix notation, this map has the form

$$\omega(\xi,\eta) = \xi^T J \eta \qquad \text{with} \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{15}$$

where I is the identity matrix of dimension d.

Definition 2 A linear mapping $A : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called *symplectic* if

$$A^T J A = J$$

or, equivalently, if $\omega(A\xi, A\eta) = \omega(\xi,\eta)$ for all $\xi, \eta \in \mathbb{R}^{2d}$.

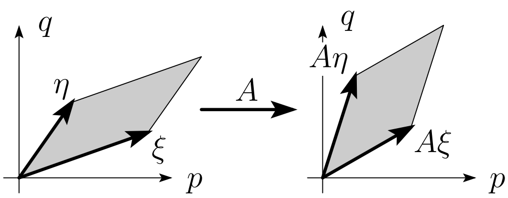


Figure 2: Symplecticity (area preservation) of a linear mapping.

In the case $d = 1$, where the expression $\omega(\xi,\eta)$ represents the area of the parallelogram P, symplecticity of a linear mapping A is therefore the *area preservation* of A (see Fig. 2). In the general case ($d > 1$), symplecticity means that the sum of the oriented areas of the projections of P onto (p_i, q_i) is the same as that for the transformed parallelograms $A(P)$.

Definition 3 (Symplectic mappings) A differentiable map $g : U \to \mathbb{R}^{2d}$ (where $U \subset \mathbb{R}^{2d}$ is an open set) is called *symplectic* if the Jacobian matrix $g'(p,q)$ is everywhere symplectic, i.e., if

$$g'(p,q)^T J\, g'(p,q) = J \qquad \text{or} \qquad \omega(g'(p,q)\xi, g'(p,q)\eta) = \omega(\xi,\eta).$$

Since a 2-dimensional sub-manifold M of the $2d$-dimensional set U can be approximated by a union of small parallelograms, the above preservation property carries over to nonlinear manifolds.

4 Theorem of Poincaré

We are now able to prove the main result[6] of this lecture. We use the notation $y = (p, q)$, and we write the Hamiltonian system (6) in the form

$$\dot{y} = J^{-1} \nabla H(y), \tag{16}$$

where J is the matrix of (15) and $\nabla H(y) = H'(y)^T$.

Recall that the flow $\varphi_t : U \to \mathbb{R}^{2d}$ of a Hamiltonian system is the mapping that advances the solution by time t, i.e., $\varphi_t(p_0, q_0) = (p(t, p_0, q_0), q(t, p_0, q_0))$, where $p(t, p_0, q_0)$, $q(t, p_0, q_0)$ is the solution of the system corresponding to initial values $p(0) = p_0$, $q(0) = q_0$.

Theorem 2 (Poincaré 1899) *Let $H(p, q)$ be a twice continuously differentiable function on $U \subset \mathbb{R}^{2d}$. Then, for each fixed t, the flow φ_t is a symplectic transformation wherever it is defined.*

Proof. The derivative $\partial\varphi_t/\partial y_0$ (with $y_0 = (p_0, q_0)$) is a solution of the variational equation which, for the Hamiltonian system (16), is of the form $\dot{\Psi} = J^{-1}\nabla^2 H\big(\varphi_t(y_0)\big)\Psi$, where $\nabla^2 H(p, q)$ is the Hessian matrix of $H(p, q)$ ($\nabla^2 H(p, q)$ is symmetric). We therefore obtain

$$\begin{aligned}
\frac{d}{dt}\Big(\Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)\Big) &= \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)'^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big) + \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)' \\
= \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T \nabla^2 H\big(\varphi_t(y_0)\big) J^{-T} J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big) &+ \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T \nabla^2 H\big(\varphi_t(y_0)\big)\Big(\frac{\partial\varphi_t}{\partial y_0}\Big) = 0,
\end{aligned}$$

because $J^T = -J$ and $J^{-T}J = -I$. Since the relation

$$\Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big) = J \tag{17}$$

is satisfied for $t = 0$ (φ_0 is the identity map), it is satisfied for all t and all (p_0, q_0), as long as the solution remains in the domain of definition of H. $\square$

We illustrate this theorem with the pendulum problem (Example 3) using the normalization $m = \ell = g = 1$. We have $q = \alpha$, $p = \dot{\alpha}$, and the Hamiltonian is given by

$$H(p, q) = p^2/2 - \cos q.$$

[6]H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste. Tome III*, Gauthiers-Villars, Paris, 1899.

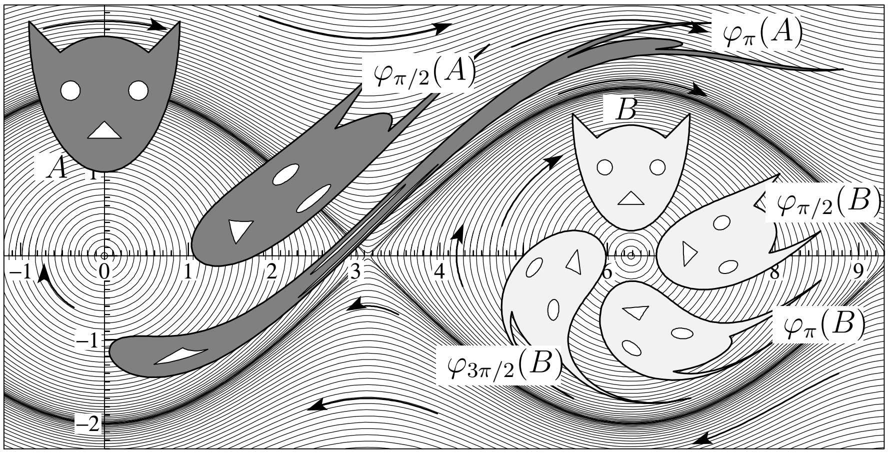


Figure 3: Area preservation of the flow of Hamiltonian systems

Figure 3 shows level curves of this function, and it also illustrates the area preservation of the flow φ_t. Indeed, by Theorem 2, the areas of A and $\varphi_t(A)$ as well as those of B and $\varphi_t(B)$ are the same, although their appearance is completely different.

We next show that symplecticity of the flow is a characteristic property for Hamiltonian systems. We call a differential equation $\dot{y} = f(y)$ *locally Hamiltonian*, if for every $y_0 \in U$ there exists a neighbourhood where $f(y) = J^{-1}\nabla H(y)$ for some function H.

Theorem 3 *Let $f : U \to \mathbb{R}^{2d}$ be continuously differentiable. Then, $\dot{y} = f(y)$ is locally Hamiltonian if and only if its flow $\varphi_t(y)$ is symplectic for all $y \in U$ and for all sufficiently small t.*

Proof. The necessity follows from Theorem 2. We therefore assume that the flow φ_t is symplectic, and we have to prove the local existence of a function $H(y)$ such that $f(y) = J^{-1}\nabla H(y)$. Differentiating (17) and using the fact that $\partial\varphi_t/\partial y_0$ is a solution of the variational equation $\dot{\Psi} = f'\big(\varphi_t(y_0)\big)\Psi$, we obtain

$$\frac{d}{dt}\left(\Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)\right) = \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)\Big(f'\big(\varphi_t(y_0)\big)^T J + J f'\big(\varphi_t(y_0)\big)\Big)\Big(\frac{\partial\varphi_t}{\partial y_0}\Big) = 0.$$

Putting $t = 0$, it follows from $J = -J^T$ that $Jf'(y_0)$ is a symmetric matrix for all y_0. The Integrability Lemma below shows that $Jf(y)$ can be written as the gradient of a function $H(y)$. $\square$

Lemma 1 (Integrability Lemma) *Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^n$ be continuously differentiable, and assume that the Jacobian $f'(y)$ is symmetric for all $y \in D$. Then, for every $y_0 \in D$ there exists a neighbourhood and a function $H(y)$ such that*

$$f(y) = \nabla H(y) \tag{18}$$

on this neighbourhood.

Proof. Assume $y_0 = 0$, and consider a ball around y_0 which is contained in D. On this ball we define

$$H(y) = \int_0^1 y^T f(ty)\, dt + Const.$$

Differentiation with respect to y_k, and using the symmetry assumption $\partial f_i / \partial y_k = \partial f_k / \partial y_i$ yields

$$\frac{\partial H}{\partial y_k}(y) = \int_0^1 \Big(f_k(ty) + y^T \frac{\partial f}{\partial y_k}(ty)t \Big)\, dt = \int_0^1 \frac{d}{dt}\Big(t f_k(ty) \Big)\, dt = f_k(y),$$

which proves the statement. □

For $D = \mathbb{R}^{2d}$ or for star-shaped regions D, the above proof shows that the function H of Lemma 1 is globally defined. Hence the Hamiltonian of Theorem 3 is also globally defined in this case. This remains valid for simply connected sets D.

5 Generating functions

Like Hamiltonian systems are described by only one scalar function (the total energy or Hamiltonian), also symplectic mappings can be described by only one scalar function (the generating function).

The following results are conveniently formulated in the notation of differential forms. For a function $F(y)$ we denote by $dF = dF(y)$ its (Fréchet) derivative. It is the linear mapping

$$dF(y)(\xi) = F'(y)\xi = \sum_{i=1}^n \frac{\partial F}{\partial y_i}(y)\, \xi_i$$

For the special case $F(y) = y_k$ we denote the derivative by dy_k, so that $dy_k(\xi) = \xi_k$ is the projection to the kth component. With $dy = (dy_1, \ldots, dy_n)^\mathsf{T}$ we thus have $dF = \sum_{i=1}^n \frac{\partial F}{\partial y_i}(y)\, dy_i$.

For a function $S(p,q)$ we use the notation

$$dS(p,q) = dS = S_p dp + S_q dq = \sum_{i=1}^{d} \Big(\frac{\partial S}{\partial p_i}(p,q) dp_i + \frac{\partial S}{\partial q_i}(p,q) dq_i \Big),$$

where we use the notation S_p, S_q for the row vectors consisting of partial derivatives, and $dp = (dp_1, \dots, dp_d)^\mathsf{T}$, $dq = (dq_1, \dots, dq_d)^\mathsf{T}$.

Theorem 4 *A mapping $\varphi : (p,q) \mapsto (P,Q)$ is symplectic if and only if there exists locally a function $S(p,q)$ such that*

$$P^\mathsf{T} dQ - p^\mathsf{T} dq = dS. \tag{19}$$

This means that $P^\mathsf{T} dQ - p^\mathsf{T} dq$ is a total differential.

Proof. We split the Jacobian of φ into the natural 2×2 block matrix

$$\frac{\partial(P,Q)}{\partial(p,q)} = \begin{pmatrix} P_p & P_q \\ Q_p & Q_q \end{pmatrix}.$$

Inserting this into (17) and multiplying out shows that the three conditions

$$P_p^T Q_p = Q_p^T P_p, \qquad P_p^T Q_q - I = Q_p^T P_q, \qquad Q_q^T P_q = P_q^T Q_q \tag{20}$$

are equivalent to symplecticity. We now insert $dQ = Q_p\, dp + Q_q\, dq$ into the left-hand side of (19) and obtain

$$\Big(P^T Q_p,\ P^T Q_q - p^T \Big) \begin{pmatrix} dp \\ dq \end{pmatrix} = \begin{pmatrix} Q_p^T P \\ Q_q^T P - p \end{pmatrix}^\mathsf{T} \begin{pmatrix} dp \\ dq \end{pmatrix}.$$

To apply the Integrability Lemma 1, we just have to verify the symmetry of the Jacobian of the coefficient vector,

$$\begin{pmatrix} Q_p^T P_p & Q_p^T P_q \\ Q_q^T P_p - I & Q_q^T P_q \end{pmatrix} + \sum_i P_i \, \frac{\partial^2 Q_i}{\partial (p,q)^2}. \tag{21}$$

Since the Hessians of Q_i are symmetric anyway, it is immediately clear that the symmetry of the matrix (21) is equivalent to the symplecticity conditions (20). □

Mixed-Variable Generating Functions. The relation (19) suggests to use (q,Q) as independent coordinates of the mapping S. When working with mappings that are close to the identity (numerical integrators), it is more convenient to use mixed variables like (P,q) or (p,Q) or $\big((P+p)/2, (Q+q)/2\big)$.

Theorem 5 *Let $(p,q) \mapsto (P,Q)$ be a smooth transformation, close to the identity. It is symplectic if and only if one of the following conditions holds locally:*

- $Q^\mathsf{T} dP + p^\mathsf{T} dq = d(P^\mathsf{T} q + S^1)$ *for some function $S^1(P,q)$;*
- $P^\mathsf{T} dQ + q^\mathsf{T} dp = d(p^\mathsf{T} Q - S^2)$ *for some function $S^2(p,Q)$;*
- $(Q-q)^\mathsf{T} d(P+p) - (P-p)^\mathsf{T} d(Q+q) = 2\, dS^3$
  for some function $S^3\big((P+p)/2, (Q+q)/2\big)$.

Proof. The first characterization follows from $d(Q^\mathsf{T} P) = Q^\mathsf{T} dP + P^\mathsf{T} dQ$ and (19), if we put S^1 such that $P^\mathsf{T} q + S^1 = Q^\mathsf{T} P - S$. For the second characterization we use $d(p^\mathsf{T} q) = p^\mathsf{T} dq + q^\mathsf{T} dp$ and the same arguments as before. The last one follows from the fact that (19) is equivalent to $(Q-q)^\mathsf{T} d(P+p) - (P-p)^\mathsf{T} d(Q+q) = d\big((P+p)^\mathsf{T}(Q-q) - 2S\big)$. □

The generating functions S^1, S^2, and S^3 have been chosen such that we obtain the identity mapping when they are replaced with zero. Comparing the coefficient functions of dq and dP in the first characterization of Theorem 5, we obtain

$$p = P + \nabla_q S^1(P,q), \qquad Q = q + \nabla_P S^1(P,q). \tag{22}$$

Whatever the scalar function $S^1(P,q)$ is, the relation (22) defines a symplectic transformation $(p,q) \mapsto (P,Q)$. Similar relations are obtained from the other two characterizations.

6 Hamilton–Jacobi partial differential equation

We know from Theorem 2 that the exact flow of a Hamiltonian differential equation

$$\dot p = -\nabla_q H(p,q), \qquad \dot q = \nabla_p H(p,q). \tag{23}$$

is a symplectic transformation. With the notation $P(t) = P(t,p,q)$, $Q(t) = Q(t,p,q)$ for the solution corresponding to initial values (p,q) at $t=0$, we known from Theorem 5 that there exists a scalar function $S^1(P,q,t)$ such that the flow $\big(P(t), Q(t)\big)$ is a solution of the equations

$$p = P(t) + \nabla_q S^1\big(P(t),q,t\big), \qquad Q(t) = q + \nabla_P S^1\big(P(t),q,t\big). \tag{24}$$

Our aim is to find the generating function $S^1(P,q,t)$ directly from $H(p,q)$. The result is the Hamilton–Jacobi differential equations.[7]

Theorem 6 *If $S^1(P,q,t)$ is a solution of the partial differential equation*

$$\frac{\partial S^1}{\partial t}(P,q,t) \;=\; H\Big(P, q + \frac{\partial S^1}{\partial P}(P,q,t)\Big), \qquad S^1(P,q,0) = 0, \tag{25}$$

then the mapping $(p,q) \mapsto \big(P(t),Q(t)\big)$, defined by (24), is the exact flow of the Hamiltonian system (23).

Proof. Let $S^1(P,q,t)$ be given by the Hamilton–Jacobi equation (25), and assume that $(p,q) \mapsto (P,Q) = \big(P(t),Q(t)\big)$ is the transformation given by (24). Differentiation of the first relation of (24) with respect to time t and using (25) yields[8]

$$\Big(I + \frac{\partial^2 S^1}{\partial P \partial q}(P,q,t)\Big)\dot P \;=\; -\frac{\partial^2 S^1}{\partial t \partial q}(P,q,t) \;=\; -\Big(I + \frac{\partial^2 S^1}{\partial P \partial q}(P,q,t)\Big)\frac{\partial H}{\partial Q}(P,Q).$$

Differentiation of the second relation of (24) gives

$$\begin{aligned}
\dot Q \;&=\; \frac{\partial^2 S^1}{\partial t \partial P}(P,q,t) + \frac{\partial^2 S^1}{\partial P^2}(P,q,t)\dot P \\
&=\; \frac{\partial H}{\partial P}(P,Q) + \frac{\partial^2 S^1}{\partial P^2}(P,q,t)\Big(\frac{\partial H}{\partial Q}(P,Q) + \dot P\Big).
\end{aligned}$$

Consequently, $\dot P = -\frac{\partial H}{\partial Q}(P,Q)$ and $\dot Q = \frac{\partial H}{\partial P}(P,Q)$, so that $\big(P(t),Q(t)\big)$ is the exact flow of the Hamiltonian system. □

Other choices of independent variables lead to different but closely related Hamilton–Jacobi differential equations.

---

[7] Sir W.R. Hamilton, *On a general method in dynamics; by which the study of the motions of all free systems of attracting or repelling points is reduced to the search and differentiation of one central relation, or characteristic function,* Phil. Trans. Roy. Soc. Part II for 1834, 247–308; Math. Papers, Vol. II, 103–161.

C.G.J. Jacobi, *Vorlesungen über Dynamik* (1842-43), Reimer, Berlin 1884.

[8] Due to an inconsistent notation of the partial derivatives $\frac{\partial H}{\partial Q}$, $\frac{\partial S^1}{\partial q}$ as column or row vectors, this formula may be difficult to read. Use indices instead of matrices in order to check its correctness.

A formal solution of the Hamilton–Jacobi partial differential equation can be obtained by inserting the ansatz

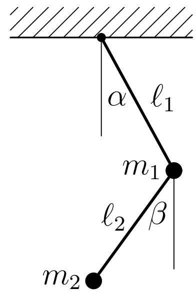
$$S^1(P,q,t) = t\,G_1(P,q) + t^2\,G_2(P,q) + t^3\,G_3(P,q) + \ldots \tag{26}$$

into (25), and by comparing like powers of t. This yields

$$\begin{aligned}
G_1(P,q) &= H(P,q),\\
G_2(P,q) &= \frac{1}{2}\Big(\frac{\partial H}{\partial P}\frac{\partial H}{\partial q}\Big)(P,q),\\
G_3(P,q) &= \frac{1}{6}\bigg(\frac{\partial^2 H}{\partial P^2}\Big(\frac{\partial H}{\partial q}\Big)^2 + \frac{\partial^2 H}{\partial P\partial q}\frac{\partial H}{\partial P}\frac{\partial H}{\partial q} + \frac{\partial^2 H}{\partial q^2}\Big(\frac{\partial H}{\partial P}\Big)^2\bigg)(P,q).
\end{aligned}$$

If we use a truncated series of (26) in the relations (24), the resulting mapping $(p,q) \mapsto \big(P(t),Q(t)\big)$, which is symplectic by Theorem 5, defines an excellent approximation to the exact solution of the Hamiltonian system.

7 Exercises



1. Let α and β be the generalized coordinates of the double pendulum, whose kinetic and potential energies are

$$\begin{aligned}
T &= \frac{m_1}{2}(\dot x_1^2 + \dot y_1^2) + \frac{m_2}{2}(\dot x_2^2 + \dot y_2^2)\\
U &= m_1 g y_1 + m_2 g y_2.
\end{aligned}$$

   Determine the generalized momenta of the corresponding Hamiltonian system.

2. Prove that a linear transformation $A : \mathbb{R}^2 \to \mathbb{R}^2$ is symplectic, if and only if $\det A = 1$.

3. Consider the transformation $(r,\varphi) \mapsto (p,q)$, defined by

$$p = \psi(r)\cos\varphi, \qquad q = \psi(r)\sin\varphi.$$

   For which function $\psi(r)$ is it a symplectic transformation?

4. Consider the Hamiltonian system $\dot y = J^{-1}\nabla H(y)$ and a variable transformation $y = \varphi(z)$. Prove that, for a symplectic transformation $\varphi(z)$, the system in the z-coordinates is again Hamiltonian with $\widetilde H(z) = H\big(\varphi(z)\big)$.

5. Write Kepler's problem with Hamiltonian

$$H(p,q) \;=\; \frac{1}{2}(p_1^2+p_2^2) - \frac{1}{\sqrt{q_1^2+q_2^2}}$$

in polar coordinates $q_1 = r\cos\varphi$, $q_2 = r\sin\varphi$. What are the conjugated generalized momenta p_r, p_φ? What is the Hamiltonian in the new coordinates.

6. Let $Q = \chi(q)$ be a change of position coordinates. Show that the mapping $(p,q) \mapsto (P,Q)$ is symplectic if $p = \chi'(q)^\mathsf{T} P$.

   Hint. Consider the mixed-variable generating function $S(P,q) = P^\mathsf{T}\chi(q)$.

7. On the set $U = \{(p,q)\,;\,p^2+q^2>0\}$ consider the differential equation

$$\begin{pmatrix}\dot p\\ \dot q\end{pmatrix} = \frac{1}{p^2+q^2}\begin{pmatrix}p\\ q\end{pmatrix}. \tag{27}$$

   Prove that
   a) its flow is symplectic everywhere on U;
   b) on every simply-connected subset of U the vector field (27) is Hamiltonian (with $H(p,q) = \mathrm{Im}\,\log(p+\mathrm{i}\,q) + \mathit{Const}$);
   c) it is not possible to find a differentiable function $H : U \to \mathbb{R}$ such that (27) is equal to $J^{-1}\nabla H(p,q)$ for all $(p,q)\in U$.

   Remark. The vector field (27) is locally (but not globally) Hamiltonian.

8. Find the Hamilton–Jacobi equation (cf. Theorem 6) for the generating function $S^2(p,Q)$ of Theorem 5.