The generating function of a canonical transformation

G.F. Torres del Castillo
Departamento de Física Matemática, Instituto de Ciencias, Universidad Autónoma de Puebla, 72570 Puebla, Pue., México.

Recibido el 27 de junio de 2011; aceptado el 3 de octubre de 2011

An elementary proof of the existence of the generating function of a canonical transformation is given. A shorter proof, making use of the formalism of differential forms is also given.

Keywords: Canonical transformations; generating function.

Se da una prueba elemental de la existencia de una función generatriz de una transformación canónica. Se da también una prueba más corta, usando el formalismo de formas diferenciales.

Descriptores: Transformaciones canónicas; función generatriz.

PACS: 45.20.Jj

1. Introduction

One of the main reasons why the Hamiltonian formalism is more useful than the Lagrangian formalism is that the set of coordinate transformations that leave invariant the form of the Hamilton equations is much wider than the set of coordinate transformations that leave invariant the form of the Lagrange equations. Furthermore, each of the so-called canonical transformations leaves invariant the form of the Hamilton equations and can be obtained from a single real-valued function of $2n+1$ variables, where n is the number of degrees of freedom of the system, which is therefore called the generating function of the transformation.

The proof of the existence of a generating function for an arbitrary canonical transformation given in most standard textbooks is usually based on the calculus of variations (see, *e.g.*, Refs. 1 to 6), which allows one to obtain the basic relations quickly. The aim of this paper is to give a straightforward, elementary derivation of the existence of the generating function of a canonical transformation, not based on the calculus of variations. One of the advantages of the proof given here is that it allows one to see clearly the assumptions involved, by contrast with the more diffuse proof usually given in the textbooks, and to realize that the canonical transformations *are not* the most general transformations that leave invariant the form of the Hamilton equations. In Sec. 2, the definition of a canonical transformation is briefly reviewed in order to derive the basic equations that lead to the existence of the generating function of the transformation. In Sec. 3 we point out some of the frequent errors contained in the proofs given in some of the standard textbooks. For those readers acquainted with the formalism of (exterior) differential forms, a considerably shorter proof is given in the appendix. The simplicity of this latter proof may serve as an invitation to learn the language of differential forms for those not already familiar with it.

2. Canonical transformations

In order to present the ideas in a simple way, it is convenient to consider firstly the case where there is only one degree of freedom, which greatly simplifies the derivations.

2.1. Systems with one degree of freedom

We shall consider a system with one degree of freedom, described by a Hamiltonian function $H(q,p,t)$. This means that the time evolution of the phase space coordinates q and p is determined by the Hamilton equations

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}. \tag{1}$$

We want to find the coordinate transformations, $Q{=}Q(q,p,t)$, $P{=}P(q,p,t)$, that maintain the form of the Hamilton equations (1). That is, we want that Eqs. (1) be equivalent to

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\partial K}{\partial P}, \qquad \frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial K}{\partial Q}, \tag{2}$$

where K may be the original Hamiltonian H expressed in terms of the new coordinates or another function. (The last possibility is relevant since it turns out that the new Hamiltonian can be made equal to zero by means of a suitable transformation.)

Assuming that the transformation $Q = Q(q,p,t)$, $P = P(q,p,t)$ is differentiable and can be inverted (that is, it is possible to find q and p in terms of Q, P, and t and, therefore, H can be viewed also as a function of Q, P, and t), making use repeatedly of the chain rule and of Eqs. (1) and (2) we find that

$$\frac{\partial K}{\partial P} = \frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\partial Q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial Q}{\partial t} \tag{3}$$

$$= \frac{\partial Q}{\partial q}\left(\frac{\partial H}{\partial Q}\frac{\partial Q}{\partial p} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial p}\right) \tag{4}$$

$$-\frac{\partial Q}{\partial p}\left(\frac{\partial H}{\partial Q}\frac{\partial Q}{\partial q}+\frac{\partial H}{\partial P}\frac{\partial P}{\partial q}\right)+\frac{\partial Q}{\partial t} \tag{5}$$

$$=\frac{\partial H}{\partial P}\left(\frac{\partial Q}{\partial q}\frac{\partial P}{\partial p}-\frac{\partial Q}{\partial p}\frac{\partial P}{\partial q}\right)+\frac{\partial Q}{\partial t}$$

$$=\frac{\partial H}{\partial P}\{Q,P\}+\frac{\partial Q}{\partial t}, \tag{6}$$

where we have made use of the definition of the Poisson brackets

$$\{f,g\}\equiv\frac{\partial f}{\partial q}\frac{\partial g}{\partial p}-\frac{\partial f}{\partial p}\frac{\partial g}{\partial q}. \tag{7}$$

In a similar way, we obtain

$$\frac{\partial K}{\partial Q}=\frac{\partial H}{\partial Q}\{Q,P\}-\frac{\partial P}{\partial t}. \tag{8}$$

Now we have two possibilities: either the Hamiltonian K is essentially the original Hamiltonian H, expressed in terms of the new variables [that is, $K\big(Q(q,p,t), P(q,p,t),t\big)=H(q,p,t)$], or K differs from H. In the first case, Eqs. (6) and (8) will hold, independent of the Hamiltonian H, if and only if

$$\{Q,P\}=1 \tag{9}$$

and the coordinate transformation does not involve the time, $Q=Q(q,p)$, $P=P(q,p)$. Then, Eq. (9) is the necessary and sufficient condition for the *local* existence of a function F such that

$$P\mathrm{d}Q-p\mathrm{d}q=\mathrm{d}F. \tag{10}$$

(That is, the function F may not be defined in all the phase space, we can only ensure its existence in some neighborhood of each point of the phase space.) In fact, writing the left-hand side of Eq. (10) in the equivalent form

$$\left(P\frac{\partial Q}{\partial q}-p\right)\mathrm{d}q+P\frac{\partial Q}{\partial p}\mathrm{d}p,$$

one finds that the condition

$$\frac{\partial}{\partial q}\left(P\frac{\partial Q}{\partial p}\right)=\frac{\partial}{\partial p}\left(P\frac{\partial Q}{\partial q}-p\right)$$

is equivalent to Eq. (9) [1].

Even though more general transformations are also possible (see below), attention is restricted to the transformations satisfying Eq. (9), also when the coordinate transformation involves the time explicitly. The coordinate transformations satisfying Eq. (9) are called *canonical transformations*. One good reason to consider only canonical transformations is that the Poisson brackets (7) are invariant under these transformations, in the sense that

$$\frac{\partial f}{\partial q}\frac{\partial g}{\partial p}-\frac{\partial f}{\partial p}\frac{\partial g}{\partial q}=\frac{\partial f}{\partial Q}\frac{\partial g}{\partial P}-\frac{\partial f}{\partial P}\frac{\partial g}{\partial Q},$$

for any pair of functions f,g, if and only if Eq. (9) holds. In fact, making use of the chain rule, one can readily show that

$$\frac{\partial f}{\partial q}\frac{\partial g}{\partial p}-\frac{\partial f}{\partial p}\frac{\partial g}{\partial q}=\{Q,P\}\left(\frac{\partial f}{\partial Q}\frac{\partial g}{\partial P}-\frac{\partial f}{\partial P}\frac{\partial g}{\partial Q}\right). \tag{11}$$

Thus, restricting ourselves to coordinate transformations satisfying Eq. (9), but allowing them to involve the time explicitly, Eqs. (6) and (8) yield

$$\frac{\partial Q}{\partial t}=-\frac{\partial(H-K)}{\partial P},\qquad \frac{\partial P}{\partial t}=\frac{\partial(H-K)}{\partial Q}. \tag{12}$$

Now, it turns out that Eqs. (9) and (12) are necessary and sufficient conditions for the local existence of a function F such that

$$P\mathrm{d}Q-K\mathrm{d}t-p\mathrm{d}q+H\mathrm{d}t=\mathrm{d}F, \tag{13}$$

as can be seen writing the left-hand side of the last equation as

$$\left(P\frac{\partial Q}{\partial q}-p\right)\mathrm{d}q+P\frac{\partial Q}{\partial p}\mathrm{d}p+\left(P\frac{\partial Q}{\partial t}+(H-K)\right)\mathrm{d}t$$

and applying the standard criterion for a linear (or Pfaffian) differential form to be exact. For instance, by considering the coefficients of $\mathrm{d}q$ and $\mathrm{d}t$ (recalling that q, p, and t are treated as three independent variables), we have

$$\frac{\partial}{\partial q}\left(P\frac{\partial Q}{\partial t}+(H-K)\right)-\frac{\partial}{\partial t}\left(P\frac{\partial Q}{\partial q}-p\right)$$

$$=\frac{\partial P}{\partial q}\frac{\partial Q}{\partial t}-\frac{\partial P}{\partial t}\frac{\partial Q}{\partial q}+\frac{\partial(H-K)}{\partial q}$$

$$=-\frac{\partial P}{\partial q}\frac{\partial(H-K)}{\partial P}-\frac{\partial Q}{\partial q}\frac{\partial(H-K)}{\partial Q}+\frac{\partial(H-K)}{\partial q}=0.$$

If q and Q are functionally independent, then the function F appearing in Eq. (13) can be expressed in terms of q, Q, and t (in a unique way), and from Eq. (13) it follows that

$$P=\frac{\partial F}{\partial Q},\qquad p=-\frac{\partial F}{\partial q},\qquad H-K=\frac{\partial F}{\partial t}, \tag{14}$$

and, necessarily, $\partial^2 F/\partial q\partial Q\neq 0$ (otherwise q and p would not be independent). Conversely, given a function $F(q,Q,t)$ such that $\partial^2 F/\partial q\partial Q\neq 0$, Eqs. (14) can be locally inverted to find Q and P in terms of q, p, and t. In this way, F is a generating function of a canonical transformation.

If q and Q are functionally dependent (that is, Q can be expressed as a function of q and t only, or q can be expressed as a function of Q and t only), the function F appearing in Eq. (13) can be written in infinitely many ways in terms of q, Q, and t, and the first two equations in (14) make no sense (since, *e.g.*, keeping q and t constant in the partial differentiation with respect to Q, would make Q also constant). In such a case, the variables p and Q (as well as P and q) are necessarily functionally independent (otherwise q and p would be dependent). Then, writing Eq. (14) in the equivalent form

$$P\mathrm{d}Q-K\mathrm{d}t+q\mathrm{d}p+H\mathrm{d}t=\mathrm{d}F', \tag{15}$$

where $F' \equiv F + pq$, it follows that the generating function F' can be expressed in a unique way as a function of Q, p, and t, and the canonical transformation is determined by

$$P = \frac{\partial F'}{\partial Q}, \qquad q = \frac{\partial F'}{\partial p}, \qquad H - K = \frac{\partial F'}{\partial t} \tag{16}$$

and, necessarily, $\partial^2 F'/\partial p \partial Q \neq 0$. Conversely, a given function $F'(p, Q, t)$ such that $\partial^2 F'/\partial p \partial Q \neq 0$, defines a canonical transformation by means of the first two equations in (16). In a similar way, one can consider generating functions depending on (q, P, t), or (p, P, t) (see, *e.g.*, Refs. 1 to 6).

It should be clear, from the derivation above, that the coordinate transformations satisfying condition (9) are not the most general coordinate transformations that leave invariant the form of the Hamilton equations and, by contrast to what is claimed in some textbooks (*e.g.*, Refs. 3 and 4), the Poisson bracket $\{Q, P\}$ needs not be a (trivial) constant. (By a trivial constant we mean a function whose value is the same at all points of its domain or, equivalently, a function whose partial derivatives are all identically equal to zero.) A simple example is given by the transformation

$$Q = \arctan \frac{q}{p}, \qquad P = \sqrt{p^2 + q^2}.$$

One readily finds that the Poisson bracket $\{Q, P\}$ is equal to $(p^2 + q^2)^{-1/2}$, which is not a trivial constant, but is a *constant of the motion* if the Hamiltonian is, for instance, $H = (1/2)(p^2 + q^2)$ (corresponding to a harmonic oscillator). Then, the Hamilton equations (1) yield $\mathrm{d}q/\mathrm{d}t = p$, and $\mathrm{d}p/\mathrm{d}t = -q$; therefore, we have, $\mathrm{d}Q/\mathrm{d}t = 1$ and $\mathrm{d}P/\mathrm{d}t = 0$, which can be expressed as the Hamilton equations (2) if the transformed Hamiltonian is chosen as $K = P$. In place of an equation of the form (13), in this case one finds the relation

$$P\mathrm{d}Q - K\mathrm{d}t = 2(p^2+q^2)^{-1/2}\big[p\mathrm{d}q - H\mathrm{d}t - \mathrm{d}(pq/2)\big]. \tag{17}$$

A second example, related to the previous one, is given by the coordinate transformation

$$Q = \left(t - \arctan \frac{q}{p}\right)^2, \qquad P = \frac{1}{2}(p^2 + q^2).$$

Now $\{Q, P\} = -2(t - \arctan q/p)$, which is also a constant of motion if $H = (1/2)(p^2 + q^2)$, as above. Furthermore, $\mathrm{d}Q/\mathrm{d}t = 0$, $\mathrm{d}P/\mathrm{d}t = 0$, which can be written in the form (2) with a new Hamiltonian $K = 0$. This is not strange, since in the Hamilton–Jacobi method one finds a transformation leading to a new Hamiltonian equal to zero, but this is usually done with the aid of canonical transformations (the solution of the Hamilton–Jacobi equation is the generating function of a canonical transformation to a new set of variables corresponding to a Hamiltonian equal to zero). For this transformation we obtain the relation

$$P\mathrm{d}Q = -2\left(t - \arctan \frac{q}{p}\right)\big[p\mathrm{d}q - H\mathrm{d}t - \mathrm{d}(pq/2)\big]$$

[cf. Eqs. (13) and (17)].

The most general coordinate transformation that preserves the form of the Hamilton equations (1) corresponds to $\{Q, P\}$ being a constant of the motion. Indeed, making use of the definition of the Poisson bracket (7), Eqs. (6), (8), the chain rule, and Eqs. (1)

$$\begin{aligned}
\frac{\partial}{\partial t}\{Q, P\} &= \frac{\partial Q}{\partial q}\frac{\partial}{\partial t}\frac{\partial P}{\partial p} + \frac{\partial P}{\partial p}\frac{\partial}{\partial t}\frac{\partial Q}{\partial q} - \frac{\partial Q}{\partial p}\frac{\partial}{\partial t}\frac{\partial P}{\partial q} - \frac{\partial P}{\partial q}\frac{\partial}{\partial t}\frac{\partial Q}{\partial p} = \frac{\partial Q}{\partial q}\frac{\partial}{\partial p}\left(\frac{\partial H}{\partial Q}\{Q, P\} - \frac{\partial K}{\partial Q}\right) \\
&+ \frac{\partial P}{\partial p}\frac{\partial}{\partial q}\left(-\frac{\partial H}{\partial P}\{Q, P\} + \frac{\partial K}{\partial P}\right) - \frac{\partial Q}{\partial p}\frac{\partial}{\partial q}\left(\frac{\partial H}{\partial Q}\{Q, P\} - \frac{\partial K}{\partial Q}\right) - \frac{\partial P}{\partial q}\frac{\partial}{\partial p}\left(-\frac{\partial H}{\partial P}\{Q, P\} + \frac{\partial K}{\partial P}\right) \\
&= \frac{\partial\{Q, P\}}{\partial p}\left(\frac{\partial H}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial q}\right) - \frac{\partial\{Q, P\}}{\partial q}\left(\frac{\partial H}{\partial P}\frac{\partial P}{\partial p} + \frac{\partial H}{\partial Q}\frac{\partial Q}{\partial p}\right) \\
&+ \{Q, P\}\left(\frac{\partial Q}{\partial q}\frac{\partial}{\partial p}\frac{\partial H}{\partial Q} - \frac{\partial P}{\partial p}\frac{\partial}{\partial q}\frac{\partial H}{\partial P} - \frac{\partial Q}{\partial p}\frac{\partial}{\partial q}\frac{\partial H}{\partial Q} + \frac{\partial P}{\partial q}\frac{\partial}{\partial p}\frac{\partial H}{\partial P}\right) \\
&- \frac{\partial Q}{\partial q}\frac{\partial}{\partial p}\frac{\partial K}{\partial Q} + \frac{\partial P}{\partial p}\frac{\partial}{\partial q}\frac{\partial K}{\partial P} + \frac{\partial Q}{\partial p}\frac{\partial}{\partial q}\frac{\partial K}{\partial Q} - \frac{\partial P}{\partial q}\frac{\partial}{\partial p}\frac{\partial K}{\partial P} = -\frac{\partial\{Q, P\}}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} - \frac{\partial\{Q, P\}}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} \\
&+ \{Q, P\}\left(\left\{Q, \frac{\partial H}{\partial Q}\right\} + \left\{P, \frac{\partial H}{\partial P}\right\}\right) - \left\{Q, \frac{\partial K}{\partial Q}\right\} - \left\{P, \frac{\partial K}{\partial P}\right\}.
\end{aligned}$$

Now, according to Eq. (11) we have, for instance,

$$\left\{Q, \frac{\partial H}{\partial Q}\right\} = \{Q, P\}\left(\frac{\partial Q}{\partial Q}\frac{\partial}{\partial P}\frac{\partial H}{\partial Q} - \frac{\partial Q}{\partial P}\frac{\partial}{\partial Q}\frac{\partial H}{\partial Q}\right) = \{Q, P\}\frac{\partial}{\partial P}\frac{\partial H}{\partial Q}$$

and

$$\left\{P, \frac{\partial H}{\partial P}\right\} = \{Q, P\}\left(\frac{\partial P}{\partial Q}\frac{\partial}{\partial P}\frac{\partial H}{\partial P} - \frac{\partial P}{\partial P}\frac{\partial}{\partial Q}\frac{\partial H}{\partial P}\right) = -\{Q, P\}\frac{\partial}{\partial Q}\frac{\partial H}{\partial P};$$

therefore
$$\left\{Q, \frac{\partial H}{\partial Q}\right\} + \left\{P, \frac{\partial H}{\partial P}\right\} = 0$$
and, similarly,
$$\left\{Q, \frac{\partial K}{\partial Q}\right\} + \left\{P, \frac{\partial K}{\partial P}\right\} = 0,$$
thus showing that $\{Q, P\}$ is a constant of motion (cf. Ref. 1). (A shorter proof is given in the appendix.)

2.2. Systems with an arbitrary number of degrees of freedom

When the number of degrees of freedom is greater than 1, the existence of a generating function of any canonical transformation can be demonstrated following essentially the same steps as in the preceding subsection. We start assuming that the set of Hamilton equations
$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q^i} \tag{18}$$
($i = 1, 2, \ldots, n$), is equivalent to the set
$$\frac{\mathrm{d}Q^i}{\mathrm{d}t} = \frac{\partial K}{\partial P_i}, \qquad \frac{\mathrm{d}P_i}{\mathrm{d}t} = -\frac{\partial K}{\partial Q^i}, \tag{19}$$
where the new coordinates Q^i and P_i are functions of q^i, p_i, and possibly also of the time. Then, by virtue of the chain rule and Eqs. (18) and (19) we obtain (here and in what follows there is summation over repeated indices)
$$\begin{aligned}
\frac{\partial K}{\partial P_i} &= \frac{\mathrm{d}Q^i}{\mathrm{d}t} \frac{\partial Q^i}{\partial q^j}\frac{\partial H}{\partial p_j} - \frac{\partial Q^i}{\partial p_j}\frac{\partial H}{\partial q^j} + \frac{\partial Q^i}{\partial t} \\
&= \frac{\partial Q^i}{\partial q^j}\left(\frac{\partial H}{\partial Q^k}\frac{\partial Q^k}{\partial p_j} + \frac{\partial H}{\partial P_k}\frac{\partial P_k}{\partial p_j}\right) \\
&\quad - \frac{\partial Q^i}{\partial p_j}\left(\frac{\partial H}{\partial Q^k}\frac{\partial Q^k}{\partial q^j} + \frac{\partial H}{\partial P_k}\frac{\partial P_k}{\partial q^j}\right) + \frac{\partial Q^i}{\partial t} \\
&= \frac{\partial H}{\partial Q^k}\left(\frac{\partial Q^i}{\partial q^j}\frac{\partial Q^k}{\partial p_j} - \frac{\partial Q^i}{\partial p_j}\frac{\partial Q^k}{\partial q^j}\right) \\
&\quad + \frac{\partial H}{\partial P_k}\left(\frac{\partial Q^i}{\partial q^j}\frac{\partial P_k}{\partial p_j} - \frac{\partial Q^i}{\partial p_j}\frac{\partial P_k}{\partial q^j}\right) + \frac{\partial Q^i}{\partial t} \\
&= \frac{\partial H}{\partial Q^k}\{Q^i, Q^k\} + \frac{\partial H}{\partial P_k}\{Q^i, P_k\} + \frac{\partial Q^i}{\partial t},
\end{aligned} \tag{20}$$
with the Poisson brackets being now defined by
$$\{f, g\} \equiv \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}, \tag{21}$$
and, similarly,
$$-\frac{\partial K}{\partial Q^i} = \frac{\partial H}{\partial Q^k}\{P_i, Q^k\} + \frac{\partial H}{\partial P_k}\{P_i, P_k\} + \frac{\partial P_i}{\partial t} \tag{22}$$
[cf. Eqs. (6) and (8)].

By analogy with the case where the number of degrees of freedom is 1, the canonical transformations are *defined* by the conditions
$$\{Q^i, Q^k\} = 0, \quad \{P_i, P_k\} = 0, \quad \{Q^i, P_k\} = \delta^i_k. \tag{23}$$
Then, Eqs. (20) and (22) yield
$$\frac{\partial Q^i}{\partial t} = -\frac{\partial(H-K)}{\partial P_i}, \qquad \frac{\partial P_i}{\partial t} = \frac{\partial(H-K)}{\partial Q^i} \tag{24}$$
[cf. Eqs. (12)]. As is well known, Eqs. (23) imply that
$$\begin{aligned}
\frac{\partial Q^k}{\partial q^m}\frac{\partial P_k}{\partial q^j} - \frac{\partial P_k}{\partial q^m}\frac{\partial Q^k}{\partial q^j} &= 0, \\
\frac{\partial Q^k}{\partial q^m}\frac{\partial P_k}{\partial p_j} - \frac{\partial P_k}{\partial q^m}\frac{\partial Q^k}{\partial p_j} &= \delta^j_m, \\
\frac{\partial Q^k}{\partial p_m}\frac{\partial P_k}{\partial p_j} - \frac{\partial P_k}{\partial p_m}\frac{\partial Q^k}{\partial p_j} &= 0
\end{aligned} \tag{25}$$
(as a matter of fact, Eqs. (25) are equivalent to Eqs. (23) [1,4]). Indeed, assuming that Eqs. (23) hold, we have
$$\begin{aligned}
\frac{\partial Q^i}{\partial q^m} &= \frac{\partial Q^k}{\partial q^m}\{Q^i, P_k\} - \frac{\partial P_k}{\partial q^m}\{Q^i, Q^k\} \\
&= \frac{\partial Q^k}{\partial q^m}\left(\frac{\partial Q^i}{\partial q^j}\frac{\partial P_k}{\partial p_j} - \frac{\partial Q^i}{\partial p_j}\frac{\partial P_k}{\partial q^j}\right) \\
&\quad - \frac{\partial P_k}{\partial q^m}\left(\frac{\partial Q^i}{\partial q^j}\frac{\partial Q^k}{\partial p_j} - \frac{\partial Q^i}{\partial p_j}\frac{\partial Q^k}{\partial q^j}\right) \\
&= \frac{\partial Q^i}{\partial q^j}\left(\frac{\partial Q^k}{\partial q^m}\frac{\partial P_k}{\partial p_j} - \frac{\partial P_k}{\partial q^m}\frac{\partial Q^k}{\partial p_j}\right) \\
&\quad - \frac{\partial Q^i}{\partial p_j}\left(\frac{\partial Q^k}{\partial q^m}\frac{\partial P_k}{\partial q^j} - \frac{\partial P_k}{\partial q^m}\frac{\partial Q^k}{\partial q^j}\right)
\end{aligned}$$
and, in a similar manner,
$$\begin{aligned}
\frac{\partial Q^i}{\partial p_m} &= \frac{\partial Q^i}{\partial q^j}\left(\frac{\partial Q^k}{\partial p_m}\frac{\partial P_k}{\partial p_j} - \frac{\partial P_k}{\partial p_m}\frac{\partial Q^k}{\partial p_j}\right) \\
&\quad - \frac{\partial Q^i}{\partial p_j}\left(\frac{\partial Q^k}{\partial p_m}\frac{\partial P_k}{\partial q^j} - \frac{\partial P_k}{\partial p_m}\frac{\partial Q^k}{\partial q^j}\right), \\
\frac{\partial P_i}{\partial q^m} &= \frac{\partial P_i}{\partial p_j}\left(\frac{\partial P_k}{\partial q^m}\frac{\partial Q^k}{\partial q^j} - \frac{\partial Q^k}{\partial q^m}\frac{\partial P_k}{\partial q^j}\right) \\
&\quad - \frac{\partial P_i}{\partial q^j}\left(\frac{\partial P_k}{\partial q^m}\frac{\partial Q^k}{\partial p_j} - \frac{\partial Q^k}{\partial q^m}\frac{\partial P_k}{\partial p_j}\right), \\
\frac{\partial P_i}{\partial p_m} &= \frac{\partial P_i}{\partial p_j}\left(\frac{\partial P_k}{\partial p_m}\frac{\partial Q^k}{\partial q^j} - \frac{\partial Q^k}{\partial p_m}\frac{\partial P_k}{\partial q^j}\right) \\
&\quad - \frac{\partial P_i}{\partial q^j}\left(\frac{\partial P_k}{\partial p_m}\frac{\partial Q^k}{\partial p_j} - \frac{\partial Q^k}{\partial p_m}\frac{\partial P_k}{\partial p_j}\right),
\end{aligned}$$
and this set of relations implies Eqs. (25).

Equations (24) and (25) are necessary and sufficient conditions for the local existence of a function F such that
$$P_i \mathrm{d}Q^i - K\mathrm{d}t - p_i \mathrm{d}q^i + H\mathrm{d}t = \mathrm{d}F, \tag{26}$$

as can be readily verified writing the left-hand side of the last equation in terms of the original variables

$$\left(P_j \frac{\partial Q^j}{\partial q^i} - p_i\right) \mathrm{d}q^i + P_j \frac{\partial Q^j}{\partial p_i} \mathrm{d}p_i + \left(P_i \frac{\partial Q^i}{\partial t} + (H - K)\right) \mathrm{d}t$$

and applying again the standard criterion for the local exactness of a linear differential form.

If the $2n$ variables q^i, Q^i are functionally independent (which is not necessarily the case), Eq. (26) implies that F can be expressed as a function of q^i, Q^i, and t, in a unique way, and

$$P_i = \frac{\partial F}{\partial Q^i}, \qquad p_i = -\frac{\partial F}{\partial q^i}, \qquad H - K = \frac{\partial F}{\partial t}. \quad (27)$$

The independence of the $2n$ variables q^i, p_i requires that $\det(\partial^2 F/\partial q^i \partial Q^j) \neq 0$. Conversely, given a function $F(q^i, Q^i, t)$ satisfying this condition, Eqs. (27) define a local canonical transformation.

For the canonical transformations such that the set q^i, Q^i is functionally dependent, one can employ generating functions that depend on other combinations of old and new variables; some or all of the q^i can be replaced by their conjugates p_i and, similarly, some or all of the Q^i can be replaced by their conjugates P_i, giving a total of 2^{2n} possibilities (not only the four cases considered, *e.g.*, in Ref. 3).

3. Comparison with other treatments

The presence of the combinations $p_i \mathrm{d}q^i - H\mathrm{d}t$ and $P_i \mathrm{d}Q^i - K\mathrm{d}t$ in Eq. (26) is not accidental. It is related to the fact that one obtains the Hamilton equations looking for the path in phase space, $q^i = q^i(t)$, $p_i = p_i(t)$, along which the integral

$$\int_{t_1}^{t_2} (p_i \mathrm{d}q^i - H\mathrm{d}t) \quad (28)$$

has a stationary value (usually a minimum) when compared with neighboring paths with the same end points in phase space for $t = t_1$ and $t = t_2$. Since the addition of the differential of *any* differentiable function $F(q^i, p_i, t)$ to the integrand in (28) changes the value of the integral by a term that is the same for all the paths with the same end points in phase space for $t = t_1$ and $t = t_2$, it is right to say that if

$$P_i \mathrm{d}Q^i - K\mathrm{d}t = p_i \mathrm{d}q^i - H\mathrm{d}t + \mathrm{d}F,$$

[which is Eq. (26)] then the Hamilton equations (18) will be equivalent to Eqs. (19). What is wrong to say is that the converse is also true (see, *e.g.*, Refs. 2, 5, and 6), or that $P_i \mathrm{d}Q^i - K\mathrm{d}t$ and $p_i \mathrm{d}q^i - H\mathrm{d}t$ can only differ by a trivial constant factor and the differential of a function (see, *e.g.*, Refs. 3 and 4) if Eqs. (18) are equivalent to (19).

Even though Eq. (26) implies that there exists a functional relation among F, Q^i, q^i, and t, another frequent error is to conclude that this implies that the $2n$ variables q^i, Q^i, are functionally independent (see, *e.g.*, Refs. 4 to 6).

Since Eq. (26) does not necessarily hold [see, *e.g.*, Eq. (17)], in the case of a non-canonical transformation that preserves the form of the Hamilton equations, the integrals

$$\int_{t_1}^{t_2} (p_i \mathrm{d}q^i - H\mathrm{d}t)$$

and

$$\int_{t_1}^{t_2} (P_i \mathrm{d}Q^i - K\mathrm{d}t)$$

do not coincide nor are simply related. However, the actual path followed by the system in phase space corresponds to stationary values of both functionals (this is analogous, for instance, to the fact that the point $x = 0$ is a local minimum for the functions $f(x) = x^4$ and $g(x) = 1 - \cos x$, despite the fact that these functions are not the same).

Acknowledgements

The author would like to thank the referees for helpful comments.

Appendix

A. Derivation using exterior forms

Making use of the properties of the contraction (or interior product) of a vector field with a differential form (see, *e.g.*, Refs. 7 to 11), one finds that there is only one vector field of the form

$$\mathbf{X} = \frac{\partial}{\partial t} + A^i \frac{\partial}{\partial q^i} + B_i \frac{\partial}{\partial p_i} \quad (A.1)$$

whose contraction with the 2-form

$$\omega \equiv \mathrm{d}p_i \wedge \mathrm{d}q^i - \mathrm{d}H \wedge \mathrm{d}t \quad (A.2)$$

is equal to zero (that is, $\mathbf{X} \lrcorner\, \omega = 0$). In fact, making use of the expressions (A.1) and (A.2), one finds that $\mathbf{X} \lrcorner\, \omega = 0$ is equivalent to

$$A^i = \frac{\partial H}{\partial p_i}, \qquad B_i = -\frac{\partial H}{\partial q^i}.$$

Hence, the integral curves of $\mathbf{X}$ correspond to the solutions of the Hamilton equations (18).

Equations (19) are then equivalent to the condition $\mathbf{X} \lrcorner\, \Omega = 0$, where

$$\Omega \equiv \mathrm{d}P_i \wedge \mathrm{d}Q^i - \mathrm{d}K \wedge \mathrm{d}t. \quad (A.3)$$

If we restrict ourselves to canonical transformations, then $\Omega = \omega$, or, equivalently, $\mathrm{d}(P_i \mathrm{d}Q^i - K\mathrm{d}t - p_i \mathrm{d}q^i + H\mathrm{d}t) = 0$, which implies the local existence of a function F such that Eq. (26) holds. However, there is an infinite number of 2-forms Ω of the form (A.3), that do not differ by a trivial multiplicative constant from ω such that, simultaneously, $\mathbf{X} \lrcorner\, \omega = 0$ and $\mathbf{X} \lrcorner\, \Omega = 0$ [11].

Only in the case of systems with one degree of freedom, any two such 2-forms must be related by $\Omega = f\omega$, where f is some, nowhere vanishing, real-valued function [11]. Among other things, from $\Omega = f\omega$ it follows that $\{Q, P\} = f$ [with the Poisson brackets defined by Eq. (7)]. Since ω and Ω are both closed (that is, their exterior derivatives are equal to zero), equation $\Omega = f\omega$ implies that f must obey the condition

$$\mathrm{d}f \wedge \omega = 0, \tag{A.4}$$

that is

$$0 = \left(\frac{\partial f}{\partial q}\mathrm{d}q + \frac{\partial f}{\partial p}\mathrm{d}p + \frac{\partial f}{\partial t}\mathrm{d}t\right) \wedge \left(\mathrm{d}p \wedge \mathrm{d}q - \frac{\partial H}{\partial q}\mathrm{d}q \wedge \mathrm{d}t - \frac{\partial H}{\partial p}\mathrm{d}p \wedge \mathrm{d}t\right) = \left(\frac{\partial f}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial f}{\partial t}\right)\mathrm{d}p \wedge \mathrm{d}q \wedge \mathrm{d}t.$$

By virtue of the Hamilton equations (1), this equation holds if and only if f is a constant of the motion, that is $\mathbf{X}f = 0$ (see the examples at the end of Sec. 2.1).

1. M.G. Calkin, *Lagrangian and Hamiltonian Mechanics* (World Scientific, Singapore, 1996). Chap. VII.
2. H.C. Corben and P. Stehle, *Classical Mechanics*, 2nd ed. (Wiley, New York, 1960). Sec. 58.
3. H. Goldstein, C. Poole and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, San Francisco, 2002). Chap. 9.
4. D.T. Greenwood, *Classical Dynamics* (Prentice-Hall, Englewood Cliffs, NJ, 1977). Chap. 6.
5. C. Lanczos, *The Variational Principles of Mechanics*, 4th ed. (University of Toronto Press, Toronto, 1970). Chap. VII.
6. D. ter Haar, *Elements of Hamiltonian Mechanics*, 2nd ed. (Pergamon, Oxford, 1971). Chap. 5 § 2.
7. V.I. Arnold, *Mathematical Methods of Classical Mechanics* 2nd ed. (Springer, New York, 2010).
8. M. Crampin and F.A.E. Pirani, *Applicable Differential Geometry* (Cambridge University Press, Cambridge, 1986).
9. L.H. Loomis and S. Sternberg, *Advanced Calculus* (Addison-Wesley, Reading, MA, 1968).
10. S. Sternberg, *Lectures on Differential Geometry* (Chelsea, New York, 1983).
11. G.F. Torres del Castillo, *Differentiable Manifolds: A Theoretical Physics Approach* (Birkhäuser Science, New York, 2012).