Canonical Generating Functions for Hamiltonians Using Structured Differentiation

P. Reany

July 4, 2021

Abstract

Structured Differentiation is used to present the treatment of generating functions in the theory of canonical transformations in the Hamiltonian formalism of mechanics.

1 Introduction

> Matters of notation play a considerable role in connection with the chain rule. Wide varieties of usage exist in mathematical writing where the chain rule is concerned.
>
> —Taylor & Mann

I have two main goals in this note: First, to develop the relevant mathematics of the canonical generating functions in Hamiltonian using SD, and, second, to make the logic flow as clear as possible. This note should be considered a mere introduction to the subject. Variable transformation will usually go from $q, p \to Q, P$, that is, going from the "old variables" to the "new variables."

2 Some Background for Comparison

I want to show that canonical transformations in Hamiltonian mechanics are just specialized forms of a standard change-of-variable problem in multivector calculus. Let's preface all that by observing two closely related problems of change of variables, the first in 2-d coordinatization of a plane.

1. Find a good, simple equation that relates all the variables of interest in the given problem.

2. Take a total derivative across this equation.

3. Take the parametric split[1] of the total derivative where appropriate.

4. Solve for what you need.

Example Problem 1.) Coordinate Transformations

Given the system

$$x = u\cos v, \tag{1}$$
$$y = u\sin v \tag{2}$$

find $\partial u/\partial x$, $\partial u/\partial y$, $\partial v/\partial x$, $\partial v/\partial y$.

We are going to leverage what we have already learned about doing problems of this sort by forcing the solution into the form we have already used successfully.

First, we rewrite the last system as

$$\begin{aligned} u\cos v - x &= 0, \\ u\sin v - y &= 0 \end{aligned} \tag{3}$$

and then make the definitions

$$\begin{aligned} \mathbf{F} &\equiv (F_1, F_2)^t = (u\cos v - x, u\sin v - y)^t \\ \mathbf{u} &= (u, v)^t, \quad \mathbf{x} = (x, y)^t \end{aligned} \tag{4}$$

Since $\mathbf{F} = \mathbf{0}$ then $\delta\mathbf{F}/\delta\mathbf{x} = \mathbf{0}$, or

$$\frac{\partial \mathbf{F}}{\partial \mathbf{x}} + \frac{\not\partial \mathbf{F}}{\not\partial \mathbf{x}} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} + \frac{\partial \mathbf{F}}{\partial \mathbf{u}}\frac{\delta \mathbf{u}}{\delta \mathbf{x}} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} + \frac{\partial \mathbf{F}}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \mathbf{0}\,, \tag{5}$$

where we used that

$$\frac{\delta \mathbf{u}}{\delta \mathbf{x}} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}}\,, \tag{6}$$

because $\mathbf{u}$ has no implicit dependence on $\mathbf{x}$.

Solving for $\partial\mathbf{u}/\partial\mathbf{x}$ we have

$$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} = -\left[\frac{\partial \mathbf{F}}{\partial \mathbf{u}}\right]^{-1}\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\,. \tag{7}$$

Now

$$\frac{\partial \mathbf{F}}{\partial \mathbf{u}} = \begin{bmatrix} \cos v & -u\sin v \\ \sin v & u\cos v \end{bmatrix}, \qquad \det\left[\frac{\partial \mathbf{F}}{\partial \mathbf{u}}\right] = u\,. \tag{8}$$

So,

$$\left[\frac{\partial \mathbf{F}}{\partial \mathbf{u}}\right]^{-1} = \frac{1}{u}\begin{bmatrix} u\cos v & u\sin v \\ -\sin v & \cos v \end{bmatrix} = \begin{bmatrix} \cos v & \sin v \\ -\dfrac{\sin v}{u} & \dfrac{\cos v}{u} \end{bmatrix}\,. \tag{9}$$

[1] A *parametric split* expands the total derivative between explicit and implicit derivative parts.

$$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{bmatrix} = -\begin{bmatrix} \cos v & \sin v \\ -\dfrac{\sin v}{u} & \dfrac{\cos v}{u} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos v & \sin v \\ -\dfrac{\sin v}{u} & \dfrac{\cos v}{u} \end{bmatrix} . \tag{10}$$

Equation (7) works so simply because of our demand that the partial derivative is an explicit derivative.[2] If we could not restrict which variables are to be differentiated one at a time, then Equation (7) would be impossible to write!

Note that the function $\mathbf{F}$ was initially extraneous to the problem, but introduced to much benefit to get the answer we wanted in terms of explicit derivatives, which are easy to interpret and generally easy to perform.

Rules for Taking Total Derivatives

I'll often use the symbol $\boldsymbol{\eta}$ to represent the ordered list of independent variables of a given problem. I think of $\boldsymbol{\eta}$ as a vector; its ith component is η_i. It often occurs in problems that one has to take the total derivative of cofundamental variables, and the following are the rules for doing so.

Rule 1) When one fundamental variable η_i is totally differentiated by a cofundamental variable η_j, the result is

$$\frac{\delta \eta_i}{\delta \eta_j} = \delta_{ij} \, , \tag{11}$$

where δ_{ij} is the Kronecker delta, meaning that when a fundamental variable is totally differentiated by itself, the result is unity, but when it is totally differentiated by a cofundamental variable, the result is zero; hence, the notion of *variable independence* is captured in this rule.

Rule 2) In all other cases, i.e., when a new fundamental variable differentiates an old fundamental variable, or vice versa, the derivative usually reduces to an explicit derivative. This is the standard thing to do in thermodynamics, say, because, by default, we regard state variables as having no implicit dependence on other variables (at least this is my knowledge regarding them).[3]

3 Some Hamiltonian Mechanics

Since the functions here to be considered are dependent only on the scalar fundamental of time, we will use the d-differential instead of the δ-differential

[2]Explicit derivatives are total derivatives, as well, when the implicit part of the derivative is zero This is a subtlety worth making special note of.

[3]When a state variable **does** have an implicit dependence on some variable, then we cannot just drop the implicit derivative as we have done in this case.

for the total derivative.

We will begin with Hamiltonian dynamics, in which the equations of motion are given by

$$\dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}\,, \tag{12a}$$

$$\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}\,. \tag{12b}$$

Now, introduce some function $F = F(t, \mathbf{q}(t), \mathbf{p}(t))$, then

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial t} + \frac{\not\partial F}{\not\partial t} \\ &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial \mathbf{q}}\frac{d\mathbf{q}}{dt} + \frac{\partial F}{\partial \mathbf{p}}\frac{d\mathbf{p}}{dt} \\ &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial F}{\partial \mathbf{p}}\dot{\mathbf{p}}\,. \end{aligned} \tag{13}$$

Using Hamilton's equations above this becomes

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial \mathbf{q}}\frac{\partial \mathcal{H}}{\partial \mathbf{p}} - \frac{\partial F}{\partial \mathbf{p}}\frac{\partial \mathcal{H}}{\partial \mathbf{q}} \\ &= \frac{\partial F}{\partial t} + \{F, \mathcal{H}\}\,, \end{aligned} \tag{14}$$

where we have used the Poisson bracket of F and $\mathcal{H}$. Therefore, we may write

$$\frac{\not\partial F}{\not\partial t} = \{F, \mathcal{H}\}\,. \tag{15}$$

Thus, if F is not explicitly a function of time, then F is a constant of the motion when $\{F, \mathcal{H}\} = 0$, because, then

$$\frac{dF}{dt} = \{F, \mathcal{H}\} = 0\,. \tag{16}$$

4 Canonical Transformations

From the theory of Lagrangian formalism, we know that two Lagrangians can represent the same physical system if they differ by only a total time derivative of some scalar function F:

$$\mathcal{L} = \mathcal{L}' + \frac{dF}{dt}\,. \tag{17}$$

But we also know that the Hamiltonian is a derived scalar function of a Lagrangian, given by

$$\mathcal{H} = \dot{q}p - \mathcal{L}\,. \tag{18}$$

Or, turned around

$$\mathcal{L} = \dot{q}p - \mathcal{H}\,. \tag{19}$$

Hence, we can write

$$\dot{q}p - \mathcal{H} = \dot{q}'p' - \mathcal{H}' + \frac{dF}{dt}\,. \tag{20}$$

But I don't like to use so many primes in these expressions, so, I will follow the usual convention of representing the new coordinate by Q and the new momentum by P. And while we're at it, we'll replace $\mathcal{H}'$ by $\mathcal{K}$. Thus (20) becomes

$$\dot{q}p - \mathcal{H} = \dot{Q}P - \mathcal{K} + \frac{dF}{dt}\,. \tag{21}$$

We will show that for a canonical transformation of variables, the Jacobian matrix of the transformation – that is, the matrix formed by the derivatives of the old variables with respect to the new variables – will have determinant equal to unity. This can be expressed in terms of the Poisson Bracket as

$$J = \{q,p\} = \det\begin{bmatrix}\partial q/\partial Q & \partial q/\partial P\\ \partial p/\partial Q & \partial p/\partial P\end{bmatrix} = 1\,. \tag{22}$$

The determinant of the Jacobuan matrix is called the *Jacobian* of the transformation.

And, as a consequence of (22),

$$J^{-1} = \{Q,P\} = \det\begin{bmatrix}\partial Q/\partial q & \partial Q/\partial p\\ \partial P/\partial q & \partial P/\partial p\end{bmatrix} = 1\,. \tag{23}$$

For a proof of the sufficiency of the condition $J = 1$ to ensure that the form of Hamilton's equations be preserved in the new variables, see Appendix 1.

5 Type 1 Transformation

Our general idea is to find a new Hamiltonian $\mathcal{K}$ that will be easier to solve for its variables $Q = Q(t)$ and $P = P(t)$ and then pass that information back to q and p by using algebraic relationships between the old and new coordinates. To that end, let's see what we can get by introducing the three vectors

$$\mathbf{x} \equiv \begin{bmatrix}q\\ p\end{bmatrix}\,, \quad \mathbf{X} \equiv \begin{bmatrix}Q\\ P\end{bmatrix}\,, \quad \mathbf{Y} \equiv \begin{bmatrix}p\\ P\end{bmatrix}\,. \tag{24}$$

By the way, our choice of the letter $\mathbf{Y}$ is arbitrary. Continuing, we setup the functional dependencies

$$\mathbf{Y}(\mathbf{x}) = \mathbf{Y}((\mathbf{X}(\mathbf{x}))\,, \tag{25}$$

and then form total derivatives using the chain rule[4]:

[4]SD maintains that the ONLY derivative one can take across an equality, and guarantee that the equality will still hold, is a total derivative, regardless of how one denotes a total derivative. In SD, the total derivative is denoted by either a 'δ' or a 'd'.

$$\frac{\delta \mathbf{Y}}{\delta \mathbf{x}} = \frac{\delta \mathbf{Y}}{\delta \mathbf{X}} \frac{\delta \mathbf{X}}{\delta \mathbf{x}}\,, \tag{26}$$

and we will evaluate/simplify these total derivatives according to the rules set down earlier on page (3). Making this last equation explicit, gives us

$$\begin{bmatrix} \delta p/\delta q & \delta p/\delta p \\ \delta P/\delta q & \delta P/\delta p \end{bmatrix} = \begin{bmatrix} \delta p/\delta Q & \delta p/\delta P \\ \delta P/\delta Q & \delta P/\delta P \end{bmatrix} \begin{bmatrix} \delta Q/\delta q & \delta Q/\delta p \\ \delta P/\delta q & \delta P/\delta p \end{bmatrix}\,. \tag{27}$$

But, in the matrix

$$\begin{bmatrix} \delta p/\delta q & \delta p/\delta p \\ \delta P/\delta q & \delta P/\delta p \end{bmatrix}\,, \tag{28}$$

variables p and q are mutually independent, therefore $\delta p/\delta q = 0$ and $\delta p/\delta p = 1$. And in the matrix

$$\begin{bmatrix} \delta p/\delta Q & \delta p/\delta P \\ \delta P/\delta Q & \delta P/\delta P \end{bmatrix}\,, \tag{29}$$

variables P and Q are mutually independent, so $\delta P/\delta Q = 0$ and $\delta P/\delta P = 1$. Therefore, on substituting in these results into (27), and making simplifications, we get

$$\begin{bmatrix} 0 & 1 \\ \partial P/\partial q & \partial P/\partial p \end{bmatrix} = \begin{bmatrix} \partial p/\partial Q & \partial p/\partial P \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \partial Q/\partial q & \partial Q/\partial p \\ \partial P/\partial q & \partial P/\partial p \end{bmatrix}\,. \tag{30}$$

So, now we note that

$$\begin{bmatrix} \partial Q/\partial q & \partial Q/\partial p \\ \partial P/\partial q & \partial P/\partial p \end{bmatrix} \tag{31}$$

is the inverse of the Jacobian matrix of the transformation of variables $q, p \to Q, P$, and setting its determinant equal to unity guarantees a canonical transformation, which we'll do. Thus, after taking determinants on (30), we get

$$\begin{vmatrix} 0 & 1 \\ \partial P/\partial q & \partial P/\partial p \end{vmatrix} = \begin{vmatrix} \partial p/\partial Q & \partial p/\partial P \\ 0 & 1 \end{vmatrix}\,. \tag{32}$$

On simplifying this we get what I refer to as the *adjoint* equation or condition or constraint:

$$-\frac{\partial P}{\partial q} = \frac{\partial p}{\partial Q}\,. \tag{33}$$

Equation (33) will guarantee that the transformation of variables is canonical.

Just to be clear about this, my interest at the moment is to analyze conservative potentials and fixed total energy E of the system, where

$$\mathcal{H} = E \tag{34}$$

and $\partial F/\partial t = 0$.

Anyway, we add to our system of equations in the new coordinates (under a canonical transformation)

$$\dot{Q} = \frac{\partial \mathcal{K}}{\partial P} \quad \text{and} \quad \dot{P} = -\frac{\partial \mathcal{K}}{\partial Q}\,. \tag{35}$$

Of course, our intention of performing this transformation of variables is to find that in the resultant system, (35) system of equations is easier to solve in Q, P as functions of time than are (12a) and (12b) to solve for q, p as functions of time.

Before proceeding to the Type 1 transformation, it's well to deal with an arbitrariness of my choice of Eq. (25) and then Eq. (26). Why couldn't we have instead chosen to work with

$$\mathbf{Y}(\mathbf{X}) = \mathbf{Y}((\mathbf{x}(\mathbf{X}))\,, \tag{36}$$

and then formed total derivatives using the chain rule as

$$\frac{\delta \mathbf{Y}}{\delta \mathbf{X}} = \frac{\delta \mathbf{Y}}{\delta \mathbf{x}}\frac{\delta \mathbf{x}}{\delta \mathbf{X}} = \frac{\delta \mathbf{Y}}{\delta \mathbf{x}}\frac{\delta \mathbf{x}}{\delta \mathbf{X}}\,? \tag{37}$$

Actually, we could have and we would have gotten the same adjoint constraint (33). If we do proceed by this alternative course, then we can deal directly with the Jacobian of the transformation $\dfrac{\partial \mathbf{x}}{\partial \mathbf{X}}$, though I don't think this has any particular advantage to the previous approach.

Now, at this point, the ansatz transformation we proffered is likely to be insufficient to solve for the new variables in terms of the old ones (unless we are really good guessers!). Therefore we need a way to add in the missing information that is consistent with the transformation being canonical. To that end, we ask if we can find a smooth differentiable function

$$F = F(q, Q)\,. \tag{38}$$

(In the literature, this is referred to as a Type 1 Transformation, and denoted as F_1. We'll go into this more in a later section.) In order to make $\mathcal{K} = \mathcal{H}$, we need to set

$$\frac{dF}{dt} = p\dot{q} - P\dot{Q}\,, \tag{39}$$

where we used Eq. (21). Then we differentiate the F in (38) to get

$$\frac{dF}{dt} = \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}\dot{Q}\,. \tag{40}$$

On comparing the last two equations we get

$$\frac{\partial F}{\partial Q} = -P \quad \text{and} \quad \frac{\partial F}{\partial q} = p\,. \tag{41}$$

So, how do we take advantage of the equations in (41)? Generally speaking, we'll integrate one of them to get F, and then differentiate F to get a new algebraic relationship in the other variable. (This will be made clear in the following example.)

The purpose of F is to assist in finding an additional algebraic relationship in the variables p, q, P, Q so that we can solve for

$$p = p(P, Q)\,, \quad \text{and} \quad q = q(P, Q)\,. \tag{42}$$

Why do we need this? We need this because the point of the transformation of variables is to attempt to solve Hamilton's equations in P, Q, and to do that we need to get $\mathcal{K}$ such that

$$\mathcal{K}(Q, P) = \mathcal{H}(q(Q, P), p(Q, P))\,. \tag{43}$$

Once we have $\mathcal{K}(Q, P)$, we can then hope to solve the coupled equations

$$\dot{Q} = \frac{\partial \mathcal{K}}{\partial P} \quad \text{and} \quad \dot{P} = -\frac{\partial \mathcal{K}}{\partial Q}\,. \tag{44}$$

Now, if we combine this last equation with (21), we find that

$$\mathcal{H} = \mathcal{K}\,, \tag{45}$$

the importance of which is that, using (34), now $\mathcal{K} = E$, though this fact may not always be needed to solve the problem at hand.[5]

6 Harmonic Oscillator in 1-D Solved

The following uses two problems and their solutions given by Masahiro Morii of Harvard, found in
http://users.physics.harvard.edu/˜morii/phys151/lectures/Lecture20.pdf

First Problem: The Hamiltonian for the 1-dimensional harmonic oscillator is given by

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + \frac{kq^2}{2} = \frac{1}{2m}\left(p^2 + m^2\omega^2 q^2\right)\,, \tag{47}$$

where $\omega^2 = \dfrac{k}{m}$.

Overview: We'd like to find a canonical transformation from the old variables into new variables, and then solve Hamilton's equations of motion in the new variables, and hope we can invert our solution back into the old variables.

Suppose we choose the ansatz for this problem as[6]

$$p = f(P) \cos Q\,, \tag{48a}$$

$$q = \frac{f(P)}{m\omega} \sin Q\,, \tag{48b}$$

---

[5] As an aside, if we define, for the moment, $\nabla = (\partial_q, \partial_Q)$, then

$$\nabla F = (p, -P)\,, \tag{46a}$$

and

$$(P, p) \cdot \nabla F = 0\,. \tag{46b}$$

Another symplectic relationship?

[6] This is our coupled system of priming functions, $p = p(Q, P)$ and $q = q(Q, P)$.

expecting the coordinate Q to drop out of the expression. On substituting these into (47) and using that $E = \mathcal{H} = \mathcal{K}$, we get that

$$\mathcal{K}(Q, P) = E = \frac{(f(P))^2}{2m} \,. \tag{49}$$

Clearly, since E is a constant of the motion, then so is $f(P)$, and this implies that P is also a constant of the motion.

Solving for $f(P)$ in (49), and using this value in (48a) and (48b), our ansatz becomes, after taking square roots:

$$p = \sqrt{2mE}\cos Q \,, \tag{50a}$$

$$q = \frac{\sqrt{2mE}}{m\omega}\sin Q \,. \tag{50b}$$

From these two equations we now know our way back into (q, p)-space from (Q, P)-space. But first we must find the functional value of $Q(t)$. However, we adopt the rule of thumb that we should eliminate the 'constant' E wherever is shows up before we have found Eq. (43). In accordance with this rule, we eliminate $\sqrt{2mE}$ between (50a) and (50b):

$$p = m\omega q \cot Q \,. \tag{51}$$

We have not as yet forced our ansatz transformation to be canonical. To accomplish this, we force the variables q, p, Q, P to conform to (33), and this is where the generating function $F = F(q, Q)$ comes in by use of the relations found in (41). So, we have a choice between two possible ways to solve for F: Either by integrating $\dfrac{\partial F}{\partial Q} = -P$ or integrating $\dfrac{\partial F}{\partial q} = p$. The latter is the logical choice, since we have a specific relation in (51) for p, but not a specific relation (yet) for P. Therefore, integrating (in q) over the equation

$$p = \frac{\partial F}{\partial q} = m\omega q \cot Q \,, \tag{52}$$

we get (remembering that $F(q, Q)$ treats q and Q as mutually independent!)

$$F(q, Q) = \frac{m\omega q^2}{2}\cot Q \,. \tag{53}$$

And now we complete the enforcement of the canonical requirement by insisting that $\dfrac{\partial F}{\partial Q} = -P$, which provides for us the additional algebraic relation

$$P = -\frac{\partial F}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q} \,. \tag{54}$$

We've now reached the point where we are ready to see if all this work has produced for us a simpler Hamiltonian in $\mathcal{K}(Q, P)$.

$$E = \mathcal{K}(Q, P) = \mathcal{H}(q(Q, P), p(Q, P)) \,, \tag{55}$$

where, remember, $\mathcal{H}$ is given by (47). Now, we can solve (52) and (54) for $q(Q,P)$ and $p(Q,P)$, or, rather, for q^2 and p^2, yielding

$$q^2 = \frac{2P\sin^2 Q}{m\omega}\,, \tag{56a}$$

$$p^2 = 2m\omega P\cos^2 Q\,. \tag{56b}$$

And plugging these results into (47), yields

$$\mathcal{K}(Q,P) = \omega P\,. \tag{57}$$

However, we must have that

$$\dot{Q} = \frac{\partial\mathcal{K}}{\partial P} = \omega\,, \tag{58}$$

and thus by integrating by time, we get

$$Q(t) = \omega t + \alpha\,, \tag{59}$$

where α is a constant. Using $Q(t)$ from (59), we get our final results

$$p = \sqrt{2mE}\,\cos(\omega t + \alpha)\,, \tag{60a}$$

$$q = \frac{\sqrt{2mE}}{m\omega}\,\sin(\omega t + \alpha)\,. \tag{60b}$$

7 The Four Generating Functions

There are only four ways to set up a generating function for use in canonical transformations. We already saw the first one, namely, $F = F_1(q,Q,t)$. Altogether, the four are

$$F_1(q,Q,t)\,,\quad F_2(q,P,t)\,,\quad F_3(p,Q,t)\,,\quad F_4(p,P,t)\,. \tag{61}$$

In accordance with Eq. (27), we will find that for each $\mathbf{Y}_i$, we need only set the constraint

$$\begin{vmatrix}\delta p/\delta q & \delta p/\delta p\\ \delta P/\delta q & \delta P/\delta p\end{vmatrix} = \begin{vmatrix}\delta p/\delta Q & \delta p/\delta P\\ \delta P/\delta Q & \delta P/\delta P\end{vmatrix}\,, \tag{62}$$

since the determinant of the rightmnost factor on the RHS is always unity. And, as before, we will use Rules 1 and 2 to evalute the deltal derivatives in the determinants.

Type	Generating Function	Adjoint	Partials of F
1	$F_1 = F_1(q, Q, t)$	$-\dfrac{\partial P}{\partial q} = \dfrac{\partial p}{\partial Q}$	$\dfrac{\partial F_1}{\partial Q} = -P\,, \quad \dfrac{\partial F_1}{\partial q} = p$
2	$F_2 = F_2(q, P, t)$	$\dfrac{\partial Q}{\partial q} = \dfrac{\partial p}{\partial P}$	$\dfrac{\partial F_2}{\partial P} = Q\,, \quad \dfrac{\partial F_2}{\partial q} = p$
3	$F_3 = F_3(p, Q, t)$	$\dfrac{\partial P}{\partial p} = \dfrac{\partial q}{\partial Q}$	$\dfrac{\partial F_3}{\partial Q} = -P\,, \quad \dfrac{\partial F_3}{\partial p} = -q$
4	$F_4 = F_4(p, P, t)$	$\dfrac{\partial Q}{\partial p} = -\dfrac{\partial q}{\partial P}$	$\dfrac{\partial F_4}{\partial P} = Q\,, \quad \dfrac{\partial F_4}{\partial p} = -q$

Table 1. Generating Functions, Adjoints, and Partials of F. One use of this information is to decide if a proposed transformation is canonical or not.

To find the type-2 transformation, take the variable $\mathbf{Y}_2(p, Q)$, given as

$$\mathbf{Y}_2(p, Q) = \begin{bmatrix} p \\ Q \end{bmatrix}, \tag{63}$$

and put it into Equation (24); then perform the computations with the new values, and find the generating function $F_2(q, P, t)$ with canonical constraint, similar to (33), yielding

$$\frac{\partial Q}{\partial q} = \frac{\partial p}{\partial P}\,. \tag{64}$$

with

$$\frac{\partial F_2}{\partial P} = Q \quad \text{and} \quad \frac{\partial F_2}{\partial q} = p\,. \tag{65}$$

For the third generating function, use $\mathbf{Y}_3(q, P)$ and find the generating function $F_3(p, Q, t)$ with canonical constraint:

$$\frac{\partial P}{\partial p} = \frac{\partial q}{\partial Q}\,. \tag{66}$$

with

$$\frac{\partial F_3}{\partial Q} = -P \quad \text{and} \quad \frac{\partial F_3}{\partial p} = -q\,. \tag{67}$$

For the fourth and last generating function, use $\mathbf{Y}_4(q, Q)$ and find the generating function $F_4(p, P, t)$ with canonical constraint:

$$\frac{\partial Q}{\partial p} = -\frac{\partial q}{\partial P}\,. \tag{68}$$

with

$$\frac{\partial F_4}{\partial P} = Q \quad \text{and} \quad \frac{\partial F_4}{\partial p} = -q\,. \tag{69}$$

If we try either of the other two combinations of variables for Y, such as $\mathbf{Y}_5(q,p)$ or $\mathbf{Y}_6(Q,P)$, we won't produce anything new when expanding the derivatives in (26).

This information is collected into a Table 1 above. To see a computational example on how to find the first partials of F_3, see Appendix 2.

8 Is this Transformation Canonical?

Suppose we start with the Hamiltonian given by

$$\mathcal{H}(q,p) = \frac{1}{2q^2} + \frac{p^2q^4}{2}\,. \tag{70}$$

Suppose we choose the following ansatz for the new variables in terms of the old:

$$P = pq^2\,. \tag{71}$$

Then we get the new Hamiltonian

$$\mathcal{K}(Q,P) = \frac{1}{2}Q^2 + \frac{1}{2}P^2\,. \tag{72}$$

We have two ways to proceed here, but they have one thing in common: We will have to find out which of the four generating functions is used and then apply its adjoint condition as a test. I'll spare the reader unnecessary stress and just declare the appropriate generating function to be $F_2(q,p)$.

Method 1: Here, we'll ignore Eq. (72) and discover Q on our own by use the generating function F_2. First we integrate, then we differentiate. From Table 1 we have that

$$\frac{\partial F_2}{\partial P} = Q \quad \text{and} \quad \frac{\partial F_2}{\partial q} = p\,. \tag{73}$$

So, from this last equation and from (71), we have that

$$\frac{\partial F_2}{\partial q} = p = \frac{P}{q^2}\,. \tag{74}$$

On integrating this, we get

$$F_2 = -\frac{P}{q}\,. \tag{75}$$

Then,

$$Q = \frac{\partial F_2}{\partial P} = -\frac{1}{q}\,, \tag{76}$$

Finally we get that

$$\frac{\partial Q}{\partial q} = \frac{1}{q^2} = \frac{\partial p}{\partial P} \,. \tag{77}$$

Thus, the adjoint equation to F_2 holds, making this a cononical transformation.

Method 2: Here, we'll use Eq. (72) and discover Q algebraically, without the use of the generating function F_2. On comparing Eqs. (70) and (72), we get that

$$\frac{1}{2q^2} = \tfrac{1}{2}Q^2 \,, \tag{78}$$

Solving for Q, we get

$$Q = \pm\frac{1}{q} \,. \tag{79}$$

Finally, on applying the adjoint codition, we find that if we choose $Q = -1/q$ then we have a canonical transformation.

Method 3: Our last method is to the Poisson bracket condition

$$\{p, q\} = 1 \,. \tag{80}$$

Of course we have to make sure this is true. It expands to

$$\frac{\partial p}{\partial P}\frac{\partial q}{\partial Q} - \frac{\partial q}{\partial P}\frac{\partial q}{\partial Q} = 1 \,. \tag{81}$$

With

$$q = \pm\frac{1}{Q} \quad \text{and} \quad p = \frac{P}{Q^2} \,, \tag{82}$$

and we'll decide the sign on q after we find the derivative and substitute them into (83). Then

$$(Q^2)(\pm\frac{1}{Q^2}) - (0)(2PQ) = 1 \,, \tag{83}$$

if we choose the negative sign for q in (82). And with that choice we've shown that the transformation is canonical.

9 Another Transformation Problem

The Hamiltonian for a conservative system is given by

$$\mathcal{H}(q, p) = \frac{p^2}{2} + \frac{1}{2q^2} \,, \tag{84}$$

where $m \equiv 1$. Suppose we choose the ansatz

$$P = pq \quad \text{and} \quad Q = ? \,, \tag{85}$$

and employ a Type-2 transformation

$$\frac{\partial F_2}{\partial P} = Q \quad \text{and} \quad \frac{\partial F_2}{\partial q} = p\,, \tag{86}$$

with adjoint constraint equation

$$\frac{\partial p}{\partial P} = \frac{\partial Q}{\partial q}\,. \tag{87}$$

As before, we look for a way to obtain $F_2(q, P)$ by integrating one of the equations in (86). Now, since we have a readymade candidate for p in (85) (but do not have one for Q), let's choose that, yielding

$$\frac{\partial F_2}{\partial q} = p = \frac{P}{q}\,. \tag{88}$$

Integrating this, we get

$$F_2(q, P) = P \log q\,, \tag{89}$$

where we assume that $q > 0$. Returning to (86), we can immediately compute Q:

$$Q = \frac{\partial F_2}{\partial P} = \log q\,, \tag{90}$$

which easily inverts to

$$q = e^Q\,. \tag{91a}$$

To this we add

$$p = Pe^{-Q}\,. \tag{91b}$$

Now, all we have to do is to solve for Q and P as functions of time in the new domain, and then just plug those functions of time into (91a) and (91b) and we've solved for q and p as functions of time.

We now have q, p as functions of Q, P, by which we can find $\mathcal{K}(Q, P)$

$$E = \mathcal{K}(Q, P) = \mathcal{H}(q(Q, P), p(Q, P))\,, \tag{92}$$

where, remember, $\mathcal{H}$ is given by (84). Then we solve the equations of motion in the new domain. Thus,

$$\mathcal{K}(Q, P) = \frac{1}{2}P^2 e^{-2Q} + \frac{1}{2}e^{-2Q}\,, \tag{93a}$$

or, since $E = \mathcal{K}(Q, P)$,

$$E = \frac{1}{2}(P^2 + 1)\, e^{-2Q}\,, \tag{93b}$$

Now, we solve the equations of motion in this system.

$$\dot{Q} = \frac{\partial \mathcal{K}}{\partial P} = Pe^{-2Q}\,, \tag{94a}$$

$$\dot{P} = -\frac{\partial \mathcal{K}}{\partial Q} = (P^2 + 1)e^{-2Q}\,. \tag{94b}$$

Using E in (93b), we can write

$$\dot{P} = 2E\,, \tag{95a}$$

$$P = 2Et + C\,, \tag{95b}$$

where C is a constant.

Instead of solving (94a) for Q, we can more easily arrive at it from (93b) and (95b), yielding

$$e^{2Q} = \frac{P^2 + 1}{2E} = \frac{(2Et + C)^2 + 1}{2E}\,, \tag{96}$$

We're now ready to present our final form solutions for q, p:

$$q = e^{Q} = \sqrt{\frac{(2Et + C)^2 + 1}{2E}}\,, \tag{97a}$$

$$p = Pe^{-Q} = (2Et + C)\sqrt{\frac{2E}{(2Et + C)^2 + 1}}\,. \tag{97b}$$

10 Appendix 1: Proof that We Must Set $J = 1$

Here we show the sufficiency of the condition $J = 1$ for Hamilton's equations to have the same form in the new variables Q, P as it had in the old variables q, p. As a reminder, the Jacobian matrix is composed of the partial derivatives of the old variables in terms of the new variables.

We start with the assumption that we can solve the new variables in terms of the old, and differentiate them. Then, from

$$\mathcal{H}(q, p) = \mathcal{K}(Q(q, p), P(q, p))\,, \tag{98}$$

we differentiate to get (leaving out a few steps this time)

$$\frac{\partial \mathcal{H}}{\partial(q, p)} = \frac{\partial \mathcal{K}}{\partial(Q, P)} \frac{\partial(Q, P)}{\partial(q, p)}\,. \tag{99}$$

Now, the matrix $\dfrac{\partial(Q, P)}{\partial(q, p)}$ is the inverse of the Jacobian matrix, so we know that it has an inverse. Applying the inverse of this matrix on both sides and transposing sides, we get

$$\begin{aligned}
\frac{\partial \mathcal{K}}{\partial(Q,P)} &= \frac{\partial \mathcal{H}}{\partial(q,p)} \left[\frac{\partial(Q,P)}{\partial(q,p)}\right]^{\text{-1}} \\
&= \frac{\partial \mathcal{H}}{\partial(q,p)} \begin{bmatrix} \partial Q/\partial q & \partial Q/\partial p \\ \\ \partial P/\partial q & \partial P/\partial p \end{bmatrix}^{\text{-1}} \\
&= \frac{\partial \mathcal{H}}{\partial(q,p)} \begin{bmatrix} \partial P/\partial p & -\partial Q/\partial p \\ \\ -\partial P/\partial q & \partial Q/\partial q \end{bmatrix} \frac{1}{J^{-1}}\,, \qquad (100)
\end{aligned}$$

where we have used the standard form for the inverse of a 2×2 matrix, and used that

$$\det \begin{bmatrix} \partial Q/\partial q & \partial Q/\partial p \\ \\ \partial P/\partial q & \partial P/\partial p \end{bmatrix} = J^{-1}\,. \qquad (101)$$

We'll see in a moment that we don't want this factor of $\frac{1}{J^{-1}}$ tagging along, so we'll get rid of it by setting $J^{-1} = 1$, but $J^{-1} = 1$ if and only if $J = 1$. On expanding, substituting from the Hamilton equations of motion in the old variables, and simplifying (100), we get, progressively

$$\begin{aligned}
\left[\frac{\partial \mathcal{K}}{\partial Q}, \frac{\partial \mathcal{K}}{\partial P}\right] &= \left[\frac{\partial \mathcal{H}}{\partial q}, \frac{\partial \mathcal{H}}{\partial p}\right] \begin{bmatrix} \partial P/\partial p & -\partial Q/\partial p \\ \\ -\partial P/\partial q & \partial Q/\partial q \end{bmatrix} \\
&= [-\dot{p}, \dot{q}] \begin{bmatrix} \partial P/\partial p & -\partial Q/\partial p \\ \\ -\partial P/\partial q & \partial Q/\partial q \end{bmatrix} \\
&= \left[-\dot{p}\,\frac{\partial P}{\partial p} - \dot{q}\,\frac{\partial P}{\partial q}, \dot{p}\,\frac{\partial Q}{\partial p} + \dot{q}\,\frac{\partial Q}{\partial q}\right] \\
&= \left[-\dot{P}, \dot{Q}\right] \quad , \qquad (102)
\end{aligned}$$

or, more simply,[7]

$$\frac{\partial \mathcal{K}}{\partial Q} = -\dot{P}\,, \tag{105a}$$

$$\frac{\partial \mathcal{K}}{\partial P} = \dot{Q}\,. \tag{105b}$$

So, we would not have gotten these pretty Hamilton's equations in the new variables if we had not set $J = 1$. Therefore, the constraint $J = 1$ will be an implicit constraint on all canonical transformations of the Hamiltonian.

11 Appendix 2: The First Partial Derivatives of $F_3(p, Q)$

This section serves to show how to find the first partial derivatives of the F generating functions, using $F_3(p, Q)$ as a particular example. So, we begin with the total derivative of $F_3(p, Q)$ by time:

$$\begin{aligned}\frac{dF_3}{dt} &= \frac{\partial F_3}{\partial p}\dot{p} + \frac{\partial F_3}{\partial Q}\dot{Q} \\ &= p\dot{q} - P\dot{Q}\,.\end{aligned} \tag{106}$$

By comparing the rightmost terms on the right sides, it's already apparent that

$$\frac{\partial F_3}{\partial Q} = -P\,. \tag{107}$$

To find $\dfrac{\partial F_3}{\partial p}$, we need to be a bit clever. Using that

$$\frac{d}{dt}(pq) = \dot{p}q + p\dot{q}\,, \tag{108}$$

we can replace $p\dot{q}$ by $\dfrac{d}{dt}(pq) - \dot{p}q$ in (106) to get

$$\begin{aligned}\frac{dF_3}{dt} &= \frac{\partial F_3}{\partial p}\dot{p} + \frac{\partial F_3}{\partial Q}\dot{Q} \\ &= \frac{d}{dt}(pq) - \dot{p}q - P\dot{Q} \\ &= -\dot{p}q - P\dot{Q}\,,\end{aligned} \tag{109}$$

[7] One may rightly ask why the total derivative of P by time is here

$$\dot{P} = \dot{p}\,\frac{\partial P}{\partial p} + \dot{q}\,\frac{\partial P}{\partial q} \tag{103}$$

instead of the full expression

$$\dot{P} = \dot{p}\,\frac{\partial P}{\partial p} + \dot{q}\,\frac{\partial P}{\partial q} + \frac{\partial P}{\partial t}? \tag{104}$$

The reason is because P is not explicitly dependent on time therefore $\dfrac{\partial P}{\partial t} = 0$. And a similar argument is to made for $\dot{Q}$.

where we ignored the total derivative term. This leaves us to conclude that

$$\frac{\partial F_3}{\partial p} = -q\,. \tag{110}$$

References

[1] R.C. Buck, *Advanced Calculus*, 3rd ed. McGraw-Hill Book Co. (1978).

[2] A. Taylor and W.R. Mann. *Advanced Calculus,* 2nd ed. John Wiley & Sons. New York (1972).