# Review for Hamiltonian Mechanics

## 2. Canonical Transformations

Another less obvious property of the Poisson bracket is that it satisfies the "chain rule": We let $\{\eta^\sigma\}_1^{2n}$ be an alternative set of coordinates on phase space; then we have

$$\{R, S\} = R_{,\xi^\alpha}\, P^{\alpha\beta}\, S_{,\xi^\beta} = R_{,\eta^\rho}\frac{\partial\eta^\rho}{\partial\xi^\alpha}\, P^{\alpha\beta}\, \frac{\partial\eta^\sigma}{\partial\xi^\beta} S_{,\eta^\sigma} = R_{,\eta^\rho}\{\eta^\rho, \eta^\sigma\}S_{,\eta^\sigma}\ . \tag{2.1}$$

Those coordinate transformations that "preserve" the canonical form, i.e., such that $\{\eta^\rho, \eta^\sigma\} = P^{\rho\sigma}$ are referred to as ***canonical transformations*** of the phasespace. In that case, we see that the Poisson bracket of arbitrary objects has been preserved as well: If we denote the bracket using coordinates $\xi^\alpha$ with a subscript $\xi$, and do likewise for the $\eta^\sigma$, then we have the result that

$$\text{for a canonical transformation,}\ \ \eta^\sigma = \eta^\sigma(\xi^\alpha, t), \qquad \{R, S\}_\xi = \{R, S\}_\eta\ . \tag{2.2}$$

Defining the Jacobian matrix for this transformation as $J$, then the requirement above for a canonical transformation is that

$$J\, P\, J^T = P\ , \qquad J^\rho{}_\alpha \equiv \frac{\partial\eta^\rho}{\partial\xi^\alpha}\ , \tag{2.3}$$

where the superscript $T$ indicates the transposed matrix. As these are all $2n \times 2n$ matrices, the set of all matrices $J$ which "preserve" the Poisson structure, or the symplectic structure, in this way is referred to as the group $\boldsymbol{Sp}$(n). One method to study this group is via its infinitesimal generators: Let $W$ be a $2n \times 2n$ matrix such that $J = e^W$. We then determine the following condition on $W$ which insists that $J$ preserves $P$, where we look at matrices $W$ that are not too far from the identity:

$$\begin{aligned} J\, P\, J^T = P \implies \quad & J^{-1} = P\, J^T\, P^{-1} \implies e^{-W} = P\, e^{W^T}\, P^{-1} = e^{PW^TP^{-1}} \\ \implies \quad & W = -P\, W^T\, P^{-1} \implies WP = -PW^T = +(WP)^T \\ \implies \quad & W^\rho{}_\alpha\, P^{\alpha\beta} = (WP)^{\rho\beta} = (WP)^{\beta\rho}\ . \end{aligned} \tag{2.4}$$

This tells us that the matrix $W$ must be traceless and such that the matrix $WP$ is symmetric. Therefore, the number of degrees of freedom of such a matrix $W$ is $\{\frac{1}{2}(2n)(2n+1)\} = n(2n+1)$, which then also gives us the dimension of the Lie group in question.

A rather different way to think about the symplectic group is first to consider the fact that the transformation from the coordinates $\xi(0)$ to some "later set," $\xi(t)$ is surely a canonical transformation, for every value of $t$. This transformation is of course the solution to the equations $d\xi^\alpha/dt = \widetilde{X}_H\xi^\alpha$, as for instance phrased at Eqs. (1.15). We may describe this by saying that the Hamiltonian is the generating function for the one-parameter family of canonical transformations, parameterized by $t$, which amount to the flow along the curves with tangent vector $\widetilde{X}_H$. On the other hand, from this point of view the Hamiltonian is not a particularly special function. We could begin with any function, $g$, defined on $T^*Q$, and consider the flow along the curves with tangent vector $\widetilde{X}_g$; in particular then, for some parameter $w$ along that flow, we have a one-parameter family of canonical transformations, with parameter $w$. Since the flow parameter is additive, the transformations are often written in the form

$$\eta^\rho \equiv \exp(w, g)\,\xi^\rho \equiv e^{w\widetilde{X}_g}\,\xi^\rho \equiv \sum_{m=0}^{\infty} \frac{w^n}{n!}(\widetilde{X}_g)^n\,\xi^\rho\ . \tag{2.5}$$

Since the set of all (smooth) functions over $T^*Q$ is of infinite dimension, this generates an infinite-dimensional Lie group of canonical transformations on phasespace. In general, one does not study this entire group, although there are times when that is appropriate. Instead, a given problem has a Hamiltonian, and one then looks at the largest, Abelian subgroup containing that Hamiltonian. The Lie algebra for this subgroup is generated by the "constants of the motion" for the problem.

In principle canonical transformations should be sought so that they can resolve some given dynamical problem. A plausible approach would be to find a canonical transformation where the new Hamiltonian is independent of all the new coordinates, the $Q^\alpha$'s. This would make all of them cyclic, so that all the corresponding momenta would be constant. (An even better alternative might be to simply make the new Hamiltonian to be zero, so that not only the momenta but also the coordinates would be constant.) Then performing the inverse canonical transformation brings us to the desired solution of the problem. As it turns out, it is usually just as difficult to find the appropriate canonical transformation of this type as it would be to solve the problem in the first place. Therefore, it is more or less true that one can find such a canonical transformation only for those problems that you could have solved anyway.

Rather, the real value of performing canonical transformations is that it creates a "venue" where it is much easier to make perturbations away it, to make approximate resolutions of problems that in fact you could never have solved exactly. As just noted, it is common to consider families of canonical transformations, rather than an individual one, where there is a parameter that varies from 0 to some larger value; these may be described in terms of that parameter and a generator, or generating function. The Hamiltonian was already noted as being such a generator, with the parameter being the time, starting from some initial condition. We therefore now consider ways to find such generators, and to use them to find the new Hamiltonians. We will see that each (locally-defined) canonical transformation has associated with it a unique function on $T^*Q$ that "generates it," via solutions of certain differential equations. On the other hand, while the mapping is unique in that direction, it is not unique in the other direction. A given generating function usually will generate more than one canonical transformation. We will resolve this by specifying an additional property of the generating function, referred to as *the type of the generating function.*

The general canonical transformation is a $t$-dependent transformation on phasespace

$$\eta^\alpha = \eta^\alpha(\xi, t) = \begin{pmatrix} Q^a(q,p,t) \\ P_b(q,p,t) \end{pmatrix}\ . \tag{2.6}$$

We then remember that the symplectic form is invariant under such a transformation; therefore, we may write the following, writing $\boldsymbol{\theta}$ for the canonical 1-form relative to $(q,p)$ and $\boldsymbol{\Theta}$ for the canonical 1-form relative to $(Q,P)$:

$$\begin{aligned} d\underset{\sim}{\boldsymbol{\theta}} = -\underset{\sim}{\omega} = d\underset{\sim}{\boldsymbol{\Theta}} \quad &\Longrightarrow \quad d(\boldsymbol{\theta} - \boldsymbol{\Theta}) = 0 \quad \Longrightarrow \quad \exists\, F \ni\ \boldsymbol{\theta} - \boldsymbol{\Theta} = dF\ , \\ &\text{or}\ \ p_a\,dq^a - P_a\,dQ^a = dF\ . \end{aligned} \tag{2.7}$$

By choosing any pair of the four sets of $n$ quantities as independent variables, as appropriate for the transformation in question, one may write out these differential forms as pde's. For instance, simply take the original $(q,p)$ as independent; then the Eqs. (2.7) are written out explicitly as

$$\left(p_b - P_a\frac{\partial Q^a}{\partial q^b}\right) dq^b - P_a\frac{\partial Q^a}{\partial p_c} dp_c = \frac{\partial F}{\partial q^b} dq^b + \frac{\partial F}{\partial p_c} dp_d \ . \tag{2.8a}$$

Equating, separately, the coefficients of the two sets of independent variables on the two sides of the equation we have

$$\frac{\partial F}{\partial q^b} = p_b - P_a\frac{\partial Q^a}{\partial q^b} \quad \text{and} \quad \frac{\partial F}{\partial p_c} = -P_a\frac{\partial Q^a}{\partial p_c} \ . \tag{2.8b}$$

An alternative version is obtained by using the matrix $\Gamma$ from Eq. (1.20), which helps us to pick out the $p_b$'s or the $q^a$'s from the $\xi^\alpha$'s. In that notation the previous pair of formulae just takes the form

$$F_{,\xi^\alpha} \equiv \frac{\partial F}{\partial \xi^\alpha} = \gamma_{\beta\mu}\left(\delta^\beta_\alpha\,\xi^\mu - \eta^\beta_{,\xi^\alpha}\eta^\mu\right) \ . \tag{2.8c}$$

These are pde's to be solved for $F$; that they have a solution is guaranteed by the derivation of the equations above, i.e., that the integrability conditions are satisfied because the transformation is canonical. The solution is not completely uniquely determined, since the derivatives with respect to $t$ are not given; therefore, the solution is ambiguous to within an arbitrary, additive function of $t$ only. This (generating) function may be used to determine the (desired) new Hamiltonian, which we call $K = K(\eta, t)$. There are two different ways to write out the (total) time derivatives of the coordinates $\eta^\alpha$, i.e., to find the trajectories of the system in terms of these (new) coordinates; we may write out the equations using either the new Hamiltonian or using the old one, treating them as functions of the old coordinates:

$$\begin{aligned} \frac{d}{dt}\eta^\alpha &= P^{\alpha\beta}\frac{\partial K}{\partial \eta^\beta} \ , \\ \frac{d}{dt}\eta^\alpha &= \{\eta^\alpha, H\}_\xi + \eta^\alpha_{,t} \\ &\Longrightarrow \eta^\alpha_{,t} = P^{\gamma\beta}\left\{\delta^\alpha_\gamma\frac{\partial K}{\partial \eta^\beta} - \frac{\partial \eta^\alpha}{\partial \xi^\gamma}\frac{\partial H}{\partial \xi^\beta}\right\} \ . \end{aligned} \tag{2.9}$$

Having these equations for those explicit time derivatives, we now retreat to the equations that determine the generating function $F$, to acquire more information: we take the partial derivative with respect to $t$ of Eqs. (2.8c), remembering that while the $\eta^\alpha$'s depend on $t$ the original $\xi^\alpha$'s do not:

$$F_{,\xi^\alpha,t} = -\gamma_{\beta\mu}\left(\eta^\beta_{,\xi^\alpha}\eta^\mu_{,t} + \eta^\beta_{,\xi^\alpha,t}\eta^\mu\right) = -\partial_{\xi^\alpha}\left(\gamma_{\beta\mu}\eta^\beta_{,t}\eta^\mu\right) - \gamma_{\beta\mu}\left(\eta^\beta_{,\xi^\alpha}\eta^\mu_{,t} - \eta^\beta_{,t}\eta^\mu_{,\xi^\alpha}\right) \ . \tag{2.10}$$

Now we take the first term in the last right-hand side above and just introduce a definition for the scalar quantity that appears in it:

$$\psi \equiv \gamma_{\beta\mu}\eta^\beta_{,t}\eta^\mu \ = \ P_r\,Q^r_{,t} \ , \tag{2.11}$$

so that the first term is now just $-\partial\psi/\partial\xi^\alpha$. However, we need to perform considerable more manipulation on the second term from that equation. Firstly it may be rewritten as follows:

$$-\gamma_{\beta\mu}\left(\eta^\beta_{,\xi^\alpha}\eta^\mu_{,t} - \eta^\beta_{,t}\eta^\mu_{,\xi^\alpha}\right) = -\left(\gamma_{\beta\mu} - \gamma_{\mu\beta}\right)\eta^\beta_{xi^\alpha}\eta^\mu_{,t} = -\omega_{\beta\mu}\eta^\beta_{,\xi^\alpha}\eta^\mu_{,t} \ . \tag{2.12}$$

Now we insert the required value of $\partial\eta/\partial t$ into this last form, taking it from Eqs. (2.9):

$$\begin{aligned}-\omega_{\beta\mu}\eta^\beta_{,\xi^\alpha}\eta^\mu_{,t} &= -\omega_{\beta\mu}\eta^\beta_{,\xi^\alpha}P^{\gamma\nu}\left\{\delta^\mu_\gamma\frac{\partial K}{\partial\eta^\nu} - \frac{\partial\eta^\mu}{\partial\xi^\gamma}\frac{\partial H}{\partial\xi^\nu}\right\}\\ &= -\eta^\beta_{,\xi^\alpha}\omega_{\beta\mu}P^{\mu\nu}\frac{\partial K}{\partial\eta^\nu} + \left(\eta^\beta_{,\xi^\alpha}\omega_{\beta\mu}\eta^\mu_{,\xi^\gamma}\right)P^{\gamma\nu}\frac{\partial H}{\partial\xi^\nu}\end{aligned} \tag{2.13}$$

In the second term the quantity in the parentheses is sometimes called the Lagrange bracket, in the $\eta^\alpha$ coordinates. On the other hand, since $\underset{\sim}{\omega}$ is a 2-form, under any transformation of coordinates, its components must transform appropriately for such a tensor:

$$\omega'_{\alpha\beta} = \omega_{\rho\sigma}\frac{\partial\xi^\rho}{\partial\eta^\alpha}\frac{\partial\xi^\sigma}{\partial\eta^\beta}\ ,$$

where $\omega'_{\alpha\beta}$ denotes the components in the new coordinate system. However, the importance of a canonical transformation is that it leaves this set of components invariant; therefore, we may make a much stronger statement:

$$\omega_{\alpha\beta} = \omega_{\rho\sigma}\frac{\partial\xi^\rho}{\partial\eta^\alpha}\frac{\partial\xi^\sigma}{\partial\eta^\beta} \implies \omega_{\rho\sigma} = \omega_{\alpha\beta}\frac{\partial\eta^\alpha}{\partial\xi^\rho}\frac{\partial\eta^\eta}{\partial\xi^\sigma}\ . \tag{2.14}$$

Inserting this for the quantity in parentheses, in both terms we now have a product of the form $\omega_{\beta\mu}P^{\mu\nu} = -\delta^\nu_\beta$. Thus our expression simplifies immensely, leaving us with the following simple form for the second term:

$$-\gamma_{\beta\mu}\left(\eta^\beta_{,\xi^\alpha}\eta^\mu_{,t} - \eta^\beta_{,t}\eta^\mu_{,\xi^\alpha}\right) = \eta^\beta_{,\xi^\alpha}\frac{\partial K}{\partial\eta^\beta} - \frac{\partial H}{\partial\xi^\alpha} = \partial_{\xi^\alpha}(K-H)\ .$$

Inserting this back into our original equation, Eq. (2.10), and remembering the first term, we have the simple result that

$$\partial_{\xi^\alpha}\left(F_{,t} + \psi - K + H\right) = 0\ . \tag{2.15}$$

This equation simply says that the object within the parentheses may only depend upon $t$; however, we recall that $F$ is ambiguous to within the choice of a function of $t$. Therefore we now use that ambiguity to say that the object within the parentheses is simply zero, which, recalling the definition of $\psi$ above in Eq. (2.11), gives us the desired relationship:

$$K = H + F_{,t} + P_r\, Q^r_{,t}\ . \tag{2.16}$$

This says that the desired new Hamiltonian, $K$, is simply the old Hamiltonian plus some terms that are generated by the time dependence of the transformation. In particular, were this canonical transformation to be independent of time, then the two Hamiltonians would be the same. However, that statement is to be taken carefully, since they are functions of different variables; a more accurate statement would be

$$K(\eta) = H[\xi(\eta)] + F_{,t}[\xi(\eta),\eta] + P_r\, Q^r_{,t}\ ,\qquad \eta \Longrightarrow \begin{pmatrix} Q\\ P\end{pmatrix}\ . \tag{2.16$'$}$$

It is possible to acquire yet another useful relationship between these quantities by calculating the total time derivative of $F$, where we take the partial time derivative from Eq. (2.16) and the derivatives with respect to the coordinates from Eqs. (2.8c):

$$\begin{aligned}\frac{d}{dt}F = F_{,\xi^\alpha}\dot{\xi}^\alpha + F_{,t} &= \gamma_{\alpha\mu}\dot{\xi}^\alpha\xi^\mu + K - H - \gamma_{\beta\mu}\left(\eta^\beta_{,\xi^\alpha}\dot{\xi}^\alpha + \eta^\beta_{,t}\right)\eta^\mu\\ &= p_a\frac{d}{dt}q^a + K - H - \gamma_{\beta\mu}\eta^\mu\frac{d}{dt}\eta^\beta = p_a\frac{d}{dt}q^a + K - H - P_r\frac{d}{dt}Q^r\ .\end{aligned} \tag{2.17}$$

A different way to write this same expression is the following, where we re-introduce the Lagrangians associated to the problem:

$$\frac{d}{dt}F = (p_a\,\dot{q}^a - H) - (P_r\,\dot{Q}^r - K) = L_\xi - L_\eta \ . \tag{2.17$'$}$$

There are in principle two rather different ways to deal with the time in these dynamical problems. So far we have been using it as a parameter along the trajectories, considered as curves in phasespace. We will continue to do that often. Nonetheless, a different approach is to expand the size of the phasespace by (at least) one more dimension, that one being the time. This expands the trajectories outward along the time axis. For dynamical situations with some sort of external driving, surely time dependent, it is possible for phasespace trajectories to cross themselves, but only at different times. Therefore, a very useful aspect of increasing the size of the phasespace, under these circumstances, is to ensure that the trajectories no longer have self intersections. Now, if we re-consider Eq. (2.17) in such an extended phase space, where the coordinates are $\{q^a, p_b, t\}$, then it may be re-thought in the form

$$dF = \underset{\sim}{\boldsymbol{\theta}} - \underset{\sim}{\boldsymbol{\Theta}} + (K - H)\,dt \ . \tag{2.17$''$}$$

If we now compare this equation to Eq. (2.7), which gives a different form for $dF$, we should ask why there is a difference. We may answer that by re-writing the one above just a little:

$$dF = \Big(p_a\,dq^a - H\,dt\Big) - \Big(P_r\,dQ^r - K\,dt\Big) \ . \tag{2.17$'''$}$$

This tells us that in such an enlarged phase space, not only should be have a new coordinate, $t$, but that it also should have a canonical momentum associated with it, which is $-H$, and we would re-write the canonical 1-form so as to include that extra **pair** of variables.

Having shown that a given canonical transformation will indeed determine a unique generating function and, thereby, a unique new Hamiltonian, we now return to the consideration of the problem associated with the mapping in the opposite direction. Indeed it does turn out that a given canonical transformation can generate several different canonical transformations, and associated (new) Hamiltonians. A fairly simple example can be used to show this fact, although while looking at more interesting cases we will also return to it from time to time. We consider two canonical transformations defined by the following equations, with one denoted with a caret:

$$\begin{cases} Q = p,\ P = -q, & \text{Case I;} \\ \hat{Q} = \tfrac{1}{2}p^2,\ \hat{P} = -q/p, & \text{Case II.} \end{cases}$$

We check that they are indeed canonical by calculating the Poisson brackets, only one being relevant, in each case:

$$\begin{cases} \{Q, P\} = -\{p, q\} = +1\,, & \text{Case I;} \\ \{\hat{Q}, \hat{P}\} = \tfrac{1}{2}\{p^2, -q\}/p = -\{p, q\} = +1\,, & \text{Case II.} \end{cases}$$

We then insert them into Eqs. (2.8b), to determine the pde's that determine the generating function:

$$F_{,q} = p\,,\ F_{,p} = q\,, \quad \Longrightarrow \quad F = pq\ , \quad \text{Case I and Case II,}$$

showing that a given $F$ does not uniquely determine a canonical transformation.

In order to have a given $F$ determine a unique canonical transformation, we will now consider separating generating functions according to their *type*. Given the valid coordinate system, on $T^*Q$, of $(q, p)$, and also the valid coordinate system, $(Q, P)$, we may suppose that there must be a way of choosing one set of variables from each of these systems so that that pair **also** furnishes a valid coordinate system, over some neighborhood at least, for the phasespace. There are of course 4 different ways to do this, and label them:

1. Type 1: (q,Q) are independent,
2. Type 2: (q,P) are independent,
3. Type 3: (p,Q) are independent,
4. Type 4: (p,P) are independent.

Referring back quickly to our example above, we see that the first transformation was of type 1 and also type 4, while the second transformation was of type 2, type 3, and also type 4. (It is quite common for a transformation to be of more than one type.)

**Canonical Transformations of Type 1**: We now construe everything to be functions of $q^a$ and $Q^r$; therefore, our earlier function $F = F(q, p, t)$, we now take in the form $F^1 = F^1(q, Q, t) = F[q, p(q, Q, t), t]$. We may then retreat back to Eq. (2.7) and rewrite it as follows:

$$p_a\,dq^a - P_r\,dQ^r = dF = F^1_{,q^a}dq^a + F^1_{,Q^r}dQ^r \quad\Longrightarrow\quad \frac{\partial F^1}{\partial q^a} = p_a\,,\ \frac{\partial F^1}{\partial Q^r} = -P_r\,. \tag{2.18a}$$

To work with these to determine $K$, we set $F = F^1[q, Q(q, p, t), t]$ and use this to determine the quantity $F_{,t}$ that we need:

$$F_{,t} = F^1_{,t} + Q^r_{,t}F^1_{,Q^r} = F^1_{,t} - Q^r_{,t}P_r\,. \tag{2.18b}$$

We now insert this information into Eq. (2.16), which gives us

$$K = H + F^1_{,t}\,. \tag{2.18c}$$

The difference between this equation and Eq. (2.16) is of course caused by the fact that $F_{,t}$ holds $q$ and $p$ fixed, while $F^1_{,t}$ holds $q$ and $Q$ fixed.

Given a function $F^1 = F^1(q, Q)$ such that the determinant of its second derivatives with respect to all its arguments, i.e., its Hessian determinant, is not zero, and also noting that the quantity $\partial F^1/\partial q^a$ in the first resulting equation in Eqs. (2.18a) is a function of $q$ and $Q$, we set it equal to $p_a$, as that equation insists, and solve it for $Q = Q(q, p, t)$. We then insert that value of $Q$ into the derivative, $\partial F^1/\partial Q^r$, that occurs in the second of those equations, which determines $P$. Lastly, we may insert all of this into Eq. (2.18c) to obtain the form of the new Hamiltonian, $K$. Since this clearly works provided the inversion is possible, we may say that **all** canonical transformations of type 1 of parameterized by those functions on $T^*Q$ that satisfy this constraint.

**Canonical Transformations of Type 2**: If, instead, we want to consider transformations of type 2, then we begin by defining $f^2(q, P, t) = F(q, p(q, P, t), t)$ and rewriting Eq. (2.7) as before. (However, we note in passing that Eq. (2.7) was set up so that Type 1 would be simplest, since it already had differentials with respect to $dq^a$ and $dQ^r$.) Nonetheless, we proceed ahead:

$$\begin{gathered} f^2_{,q^a}dq^a + f^2_{,P_s}dP_s = df^2 = dF = p_a\,dq^a - P_r\,dQ^r = p_a\,dq^a - P_r\,\left(Q^r_{,q^a}dq^a + Q^r_{,P_s}dP_s\right) \\ \Longrightarrow \begin{cases} f^2_{,q^a} = p_a - P_rQ^r_{,q^a}\,, \\ f^2_{,P_s} = \ -P_rQ^r_{,P_s}\,. \end{cases} \end{gathered} \tag{2.19a}$$

We may now define a quantity $F^2$ and write for it equations very similar to those for $F^1$:

$$\begin{aligned} F^2 \ &\equiv f^2(q, P, t) + P_r Q^r(q, P, t) \ , \\ p_a = \frac{\partial F^2}{\partial q^a} \ , \quad Q^r &= \frac{\partial F^2}{\partial P_r} \ , \quad K = H + \frac{\partial F^2}{\partial t} \ . \end{aligned} \tag{2.19b}$$

Because of the fact that our method equations, before we introduced type, picked out $q$ and $Q$ somewhat specially, we see that the generating function $F^2$ is more complicated than simply inserting the correct choice of variables into $F(q, p, t)$. The same would be true for $F^3$ and $F^4$ as well. This lack of symmetry comes, however, by the original choice we made for the value of the the canonical 1-form, such that $\underset{\approx}{\omega} = -d(p_a\, dq^a$. Clearly one could also have chosen $\underset{\approx}{\omega} = +d(q^a\, dp_a)$. Making that choice for the new canonical 1-form, i.e., $\underset{\approx}{\omega} = +d(Q^r\, dP_r)$, would have altered the form of our beginning equation, Eq. (2.7), and made the further discussions such that $F^2$ would have been the simplest one. By making these choices in 4 different ways, this could happen for any one of the four Types as "special." On the other hand, it is reasonable also to simply stick with our version of Eq. (2.7) and work out the explicit forms for each of the remaining two Types. I will leave those similar calculations for $F^3(p, Q, t)$ and $F^4(p, P, t)$ for homework problems, except to note here that

$$\begin{aligned} F^3(Q, p, t) &= F[q(Q, p, t), p, t) - p_a\, q^a(Q, p, t) \ , \\ F^4(p, P, t) &= F[q(p, P, t), p, t] + P_r\, Q^r(p, P, t) - p_a\, q^a(p, P, t) \ . \end{aligned} \tag{2.20}$$