Lecture 6: The Methods of Lagrange II – More on Coordinate Transformations and Constraints

As we have discussed in previous lectures the laws of Newton are truly simple only in Cartesian coordinates in an inertial frame. Here we want to generalize the discussion of the previous lectures to "generalized coordinates" to obtain a formulation of Newtonian dynamics valid even in non-inertial frames. We consider a general coordinate transformation described in terms of three holonomic (ordinary) functions (*i.e.*, the symbol $\vec{R}_j$ is to remind us that the original Cartesian coordinates are individually functions of the generalized coordinates)

$$\begin{aligned}\vec{r}_j &= \left(x_j, y_j, z_j\right) = \vec{R}_j\left(q_1, \ldots q_f, t\right) \\ &= \left(f_j\left(q_1, \ldots q_f, t\right), g_j\left(q_1, \ldots q_f, t\right), h_j\left(q_1, \ldots q_f, t\right)\right).\end{aligned} \tag{6.1}$$

Here we start with 3*N* ($j = 1, \ldots, N$) coordinates describing our *N*-particle system in 3 Cartesian dimensions with *n* constraints and we transform to $f = 3N-n$ generalized coordinate q_j ($j = 1, \ldots, 3N - n = f$). Note that we are explicitly assuming that, by an appropriate choice of coordinates, which are orthogonal to any constraint surfaces, we have eliminated the constraints from the analysis. We return to the variational form at the end of Lecture 5, Eq. (5.44), but without the constraint forces,

$$\sum_{j=1}^{N}\left(\dot{\vec{p}}_j - \vec{F}_j\right)\cdot\delta\vec{r}_j = 0, \tag{6.2}$$

where the forces included may be either conservative or dissipative. We can deduce the form of the transformed force from the relation equating the virtual work done as described in the two coordinate systems

$$\sum_{l=1}^{f} Q_l \delta q_l = \sum_{j=1}^{N}\vec{F}_j\cdot\delta\vec{r}_j. \tag{6.3}$$

We find the generalized force acting on the generalized coordinate q_j to be

$$Q_l = \sum_{j=1}^{N}\vec{F}_j\cdot\frac{\partial\vec{r}_j}{\partial q_l} = \sum_{j=1}^{N}\vec{F}_j\cdot\frac{\partial\vec{R}_j}{\partial q_l}, \tag{6.4}$$

i.e., the expected Jacobian ($\partial\vec{r}_j/\partial q_l$) from the change of variables appears. Thus the variational equation in the generalized coordinates looks like $\left(\delta\vec{r}_j \to \sum_l \left(\partial\vec{r}_j/\partial q_l\right)\delta q_j\right)$

$$\begin{aligned}\sum_{j=1}^{N}\left(\dot{\vec{p}}_j - \vec{F}_j\right)\cdot\delta\vec{r}_j &= \sum_{l=1}^{f}\left(\sum_{j=1}^{N} m_j \frac{d^2\vec{r}_j}{dt^2}\cdot\frac{\partial\vec{r}_j}{\partial q_l} - \vec{F}_l\cdot\frac{\partial\vec{r}_j}{\partial q_l}\right)\delta q_l \\ &= \sum_{l=1}^{f}\left(\sum_{j=1}^{N} m_j \frac{d^2\vec{r}_j}{dt^2}\cdot\frac{\partial\vec{r}_j}{\partial q_l} - Q_l\right)\delta q_l = 0.\end{aligned} \tag{6.5}$$

Next consider the transformation of the acceleration term. We start with

$$\dot{\vec{r}}_j = \frac{d\vec{r}_j}{dt} = \frac{\partial\vec{r}_j}{\partial q_l}\dot{q}_l + \frac{\partial\vec{r}_j}{\partial t}, \tag{6.6}$$

and note that

$$\frac{\partial\dot{\vec{r}}_j}{\partial\dot{q}_l} = \frac{\partial}{\partial\dot{q}_l}\left(\frac{\partial\vec{r}_j}{\partial q_l}\dot{q}_l + \frac{\partial\vec{r}_j}{\partial t}\right) = \frac{\partial\vec{r}_j}{\partial q_l}. \tag{6.7}$$

Then it follows from

$$\begin{aligned}&\frac{d}{dt}\left(\frac{\partial\vec{r}_j}{\partial q_l}\cdot\dot{\vec{r}}_j\right) = \frac{d^2\vec{r}_j}{dt^2}\cdot\frac{\partial\vec{r}_j}{\partial q_l} + \dot{\vec{r}}_j\cdot\frac{d}{dt}\left(\frac{\partial\vec{r}_j}{\partial q_l}\right) \\ &\Rightarrow \frac{d^2\vec{r}_j}{dt^2}\cdot\frac{\partial\vec{r}_j}{\partial q_l} = \frac{d}{dt}\left(\frac{\partial\vec{r}_j}{\partial q_l}\cdot\dot{\vec{r}}_j\right) - \dot{\vec{r}}_j\cdot\frac{d}{dt}\left(\frac{\partial\vec{r}_j}{\partial q_l}\right) \\ &\qquad = \frac{d}{dt}\left(\frac{\partial\vec{r}_j}{\partial q_l}\cdot\dot{\vec{r}}_j\right) - \dot{\vec{r}}\cdot\left(\frac{\partial\dot{\vec{r}}_j}{\partial q_l}\right),\end{aligned} \tag{6.8}$$

that we can write, using Eqs. (6.7) and (6.8), that

$$
\begin{aligned}
m_j \frac{d^2\vec{r}_j}{dt^2}\cdot\frac{\partial\vec{r}_j}{\partial q_l} &= \frac{d}{dt}\left(m_j \dot{\vec{r}}_j \cdot \frac{\partial\dot{\vec{r}}_j}{\partial \dot{q}_l}\right) - m_j \dot{\vec{r}}_j \cdot \left(\frac{\partial\dot{\vec{r}}_j}{\partial q_l}\right) \\
&= \frac{1}{2}\frac{d}{dt}\left(m_j \frac{\partial\dot{\vec{r}}_j^{\,2}}{\partial \dot{q}_l}\right) - \frac{1}{2} m_j \frac{\partial\dot{\vec{r}}_j^{\,2}}{\partial q_l} \\
&= \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l},
\end{aligned}
\tag{6.9}
$$

where we have recognized the kinetic energy in this expression. Finally we see that Eq. (6.5) becomes the transformed version of Hamilton's principle

$$
\sum_{l=1}^{f}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} - Q_l\right)\delta q_l = 0, \tag{6.10}
$$

which we might have expected from the Euler-Lagrange equations. Since (for now we are assuming that) the generalized coordinates were chosen to eliminate the constraints, the variations of these variables, the δq_l, can be chosen independently. Hence Eq. (6.10) can be satisfied for all such variations if and only if

$$
\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = Q_l. \tag{6.11}
$$

This last relation is called Lagrange's equation of the first kind. For the familiar case of Cartesian coordinates in an inertial frame ($T = m_j \dot{\vec{r}}_j^{\,2}/2$, $Q_l = F_l$), Eq. (6.11) clearly becomes just Newton's second law. On the other hand Eq. (6.11) applies for both dissipative *and* conservative systems in *all* generalized coordinates, both inertial *and* accelerated. As we will see, the less familiar term on the left-hand-side, $\partial T/\partial q_l$, will generate the extra terms needed, for example, when we transform to curvilinear coordinates (the ones we observed in lecture 4).

Two familiar and useful examples of generalized coordinates are cylindrical and spherical coordinates, which are especially relevant when the constraints exhibit the corresponding symmetry structure. For the former we have

$$
\begin{aligned}
\vec{r} &= \rho\hat{\rho} + z\hat{z}, \\
d\vec{r} &= d\rho\hat{\rho} + \rho d\phi\hat{\phi} + dz\hat{z}, \\
\dot{\vec{r}} &= \dot{\rho}\hat{\rho} + \rho\dot{\phi}\hat{\phi} + \dot{z}\hat{z}, \\
q_1 &= \rho, q_2 = \phi, q_3 = z, \\
T &= \frac{m}{2}\left(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2\right) \\
&= \frac{m}{2}\left(\dot{q}_1^2 + q_1^2\dot{q}_2^2 + \dot{q}_3^2\right).
\end{aligned}
\tag{6.12}
$$

Lagrange's equations then read (as we saw in Lecture 4)

$$
\begin{aligned}
&\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\rho}}\right) - \frac{\partial T}{\partial \rho} = m\left(\ddot{\rho} - \rho\dot{\phi}^2\right) = Q_\rho, \\
&\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\phi}}\right) - \frac{\partial T}{\partial \phi} = m\frac{d}{dt}\left(\rho^2\dot{\phi}\right) = \frac{d}{dt}\left(L_z\right) = m\left(2\rho\dot{\rho}\dot{\phi} + \rho^2\ddot{\phi}\right) = Q_\phi, \\
&\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{z}}\right) - \frac{\partial T}{\partial z} = m\ddot{z} = Q_z.
\end{aligned}
\tag{6.13}
$$

In spherical coordinates we have

$$
\begin{aligned}
\vec{r} &= r\hat{r}, \\
d\vec{r} &= dr\hat{r} + rd\theta\hat{\theta} + r\sin\theta d\phi\hat{\phi}, \\
\dot{\vec{r}} &= \dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + r\sin\theta\dot{\phi}\hat{\phi}, \\
q_1 &= r, q_2 = \theta, q_3 = \phi, \\
T &= \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right) \\
&= \frac{m}{2}\left(\dot{q}_1^2 + q_1^2\dot{q}_2^2 + q_1^2\sin^2 q_2\dot{q}_3^2\right).
\end{aligned}
\tag{6.14}
$$

So in spherical coordinates Lagrange tells us to write the equations of motion as

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}}\right)-\frac{\partial T}{\partial r}=m\left(\ddot{r}-r\dot{\theta}^2-r\sin^2\theta\dot{\phi}^2\right)=Q_r,$$
$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\theta}}\right)-\frac{\partial T}{\partial \theta}=m\left\{\frac{d}{dt}\left(r^2\dot{\theta}\right)-r^2\sin\theta\cos\theta\dot{\phi}^2\right\}$$
$$=m\left(2r\dot{r}\dot{\theta}+r^2\ddot{\theta}-r^2\sin\theta\cos\theta\dot{\phi}^2\right)=Q_\theta, \tag{6.15}$$
$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\phi}}\right)-\frac{\partial T}{\partial \phi}=m\frac{d}{dt}\left(r^2\sin^2\theta\dot{\phi}\right)=\frac{d}{dt}\left(L_z\right)$$
$$=\left(2r\dot{r}\sin^2\theta\dot{\phi}+2r^2\sin\theta\cos\theta\dot{\theta}\dot{\phi}+r^2\sin^2\theta\ddot{\phi}\right)=Q_\phi.$$

The slick thing here is that all the extra terms, which we laboriously obtained in Lecture 4 by careful analysis of the accelerating frame, appear naturally from Lagrange's equation.

As usual we now look at the case for conservative forces taken to be defined by scalar functions of both the new and old coordinates.

$$U\left(\vec{r}_1,\ldots,\vec{r}_N\right)=\tilde{U}\left(q_1,\ldots,q_f\right), \tag{6.16}$$

so that

$$Q_l\delta q_l=-\frac{\partial\tilde{U}}{\partial q_l}\delta q_l. \tag{6.17}$$

Now we can define a Lagrangian as we did in Lecture 5 and the equations of motion for a conservative system are the generalized version of Eq. (5.22), modulo the sign to match our current notation, $L=T-U$,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}-\frac{\partial L}{\partial q_k}=0, \tag{6.18}$$

which is Lagrange's equation of the second kind. Consider, for example, a central force potential in spherical coordinates, $\tilde{U}=\tilde{U}\left(r\right)$, for which Lagrange's equations yield (just Eq. (6.15) with explicit forms for the generalized forces)

$$
\begin{aligned}
&m\left(\ddot{r} - r\dot{\theta}^2 - r\sin^2\theta\dot{\phi}^2\right) = -\frac{\partial \tilde{U}}{\partial r}, \\
&2r\dot{r}\dot{\theta} + r^2\ddot{\theta} - r^2\sin\theta\cos\theta\dot{\phi}^2 = 0, \\
&\left(r^2\sin^2\theta\dot{\phi}\right) = L_z = \text{constant.}
\end{aligned}
\tag{6.19}
$$

The last line is, of course, the conservation of the z component of angular momentum.

It is customary and useful to make the following definition,

$$
p_j \equiv \frac{\partial T}{\partial \dot{q}_j}, \tag{6.20}
$$

which is called the canonical momentum, conjugate to the generalized coordinate q_j. In Cartesian coordinates this is just the mass times the velocity, but this is not the case more generally. For example, in spherical coordinates we have

$$
\begin{aligned}
p_r &= \frac{\partial T}{\partial \dot{r}} = m\dot{r}, \\
p_\theta &= \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta}, \\
p_\phi &= mr^2\sin^2\theta\dot{\phi}\left[= L_z\right].
\end{aligned}
\tag{6.21}
$$

This definition will be very helpful when we study the role of symmetries and conservation laws in the Lagrangian formalism. It also leads naturally to the appearance of extra terms in Newton's laws inherent in accelerating frames. We can rewrite Lagrange's equations of the first kind, Eq. (6.11), as

$$
\frac{dp_j}{dt} = \frac{\partial T}{\partial q_j} + Q_j. \tag{6.22}
$$

The first term on the right-hand-side, the position dependence of the kinetic energy, is exactly the extra term we found earlier when transforming to curvilinear coordinates. In a conservative system with $\tilde{U}$ typically not a function of the velocities we have

($L = T - \tilde{U}$)

$$p_j = \frac{\partial L}{\partial \dot{q}_j},$$
$$\frac{dp_j}{dt} = \frac{\partial L}{\partial q_j}, \tag{6.23}$$

where again we expect any peculiarities of the coordinates to appear as position dependence in the kinetic energy and contribute to the right-hand-side.

Finally we note that in a linearly accelerated frame, $\vec{r} = \vec{r}_0 + \vec{r}', \ddot{\vec{r}}_0 \neq 0$, we have

$$\vec{p} = \frac{\partial T}{\partial \dot{\vec{r}}'} = m\left(\dot{\vec{r}}' + \dot{\vec{r}}_0\right)$$
$$\Rightarrow \frac{d\vec{p}}{dt} = m\left(\ddot{\vec{r}}' + \ddot{\vec{r}}_0\right) = \vec{F}'. \tag{6.24}$$

The use of the canonical momentum preserves the simple form of Newton, but with the *correct* answer, *i.e.*, including the acceleration of the frame.

Now let us reconsider again the question of constraints. What if we cannot eliminate all constraints by our choice of generalized coordinates? Or, what if we want to learn something about the constraint forces? Assume that we are dealing with $\bar{N}$ generalized coordinates and n constraints, which we can (always) express in differential form

$$a_{jk}\delta q_j + a_{tk}\delta t = 0 \left[j = 1, \ldots, \bar{N}, k = 1, \ldots, n \right]. \tag{6.25}$$

If a constraint is holonomic, *i.e.*, the constraint is integrable, we can simplify the analysis by writing the constraint in terms of a point function of the generalized coordinates,

$$a_{jk} = \frac{\partial}{\partial q_j} \phi_k \left(q_1, \ldots, q_{\bar{N}} \right). \tag{6.26}$$

Note that, even if the constraint is nonholonomic, either because we cannot integrate Eq. (6.25) or because the constraint is really an inequality, we can still proceed using Lagrange's method of undetermined multipliers as introduced in Lecture 5.

Minimizing the action yields the familiar result

$$\left(\frac{\partial L}{\partial q_j} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j}\right)\delta q_j = 0, \tag{6.27}$$

where the variations are now not all arbitrary (*i.e.*, only $\bar{N} - n$ can be varied freely). To address this issue we introduce *n* undetermined parameters and express the non-variation of the constraints under the virtual variations of the coordinates as

$$\lambda_k a_{jk} \delta q_j = 0. \tag{6.28}$$

Combining Eqs. (6.27) and (6.28) we have

$$\left(\frac{\partial L}{\partial q_j} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} + \lambda_k a_{jk}\right)\delta q_j = 0, \tag{6.29}$$

for all values of *j* and where the index *k* is summed over. Now, after agreeing that we will eventually determine the λ_k so that the constraints are satisfied, we treat the variations as independent. Thus we now solve for $\bar{N} + n$ unknowns, the q_j and λ_k, using the $\bar{N} + n$ equations

$$\begin{aligned} &\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \lambda_k a_{jk}, \\ &a_{jk}\dot{q}_j + a_{tk} = 0. \end{aligned} \tag{6.30}$$

As an example, consider a point particle of mass m that slides without friction on the inside of a smooth paraboloid of revolution defined by $\rho^2 = az$. The paraboloid is oriented vertically in the earth's (uniform) gravitational field. If the particle is initially moving in a horizontal trajectory with velocity v_0 at radius ρ_0, what must the magnitude of v_0 be in order that the particle moves neither up nor down. Clearly we should choose cylindrical coordinates where the Lagrangian is

$$L = \frac{m}{2}\left(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2\right) - mgz. \tag{6.31}$$

The constraint of moving on the specified surface is expressed by

$$\begin{aligned} &az - \rho^2 = 0, \\ &-2\rho d\rho + adz = 0, \\ &a_{11} = -2\rho, a_{21} = 0, a_{31} = a. \end{aligned} \tag{6.32}$$

We note that the initial conditions include $z_0 = \rho_0^2 / a$. Lagrange's equations (Eq. (6.30)) yield

$$\begin{aligned} &m\ddot{\rho} - m\rho\dot{\phi}^2 = -2\rho\lambda_1, \\ &\frac{d}{dt}\left(m\rho^2\dot{\phi}\right) = 0 \Rightarrow m\rho^2\dot{\phi} = m\rho_0 v_0 = \text{constant}, \\ &m\ddot{z} + mg = a\lambda_1, \\ &\dot{z} = \frac{2\rho\dot{\rho}}{a} \Rightarrow \ddot{z} = \frac{2\rho\ddot{\rho}}{a} + \frac{2\dot{\rho}^2}{a}. \end{aligned} \tag{6.33}$$

Clearly for no *z* motion we want $\lambda_1 = mg/a$ from the third equation and $\ddot{\rho} = -\dot{\rho}^2/\rho = 0$ from the last equations. Thus it follows from the first two equations that

$$\begin{aligned} &\lambda_1 = \frac{m\dot{\phi}^2}{2} = \frac{mv_0^2}{2\rho_0^2} = \frac{mg}{a} \\ &\Rightarrow v_0 = \rho_0\sqrt{\frac{2g}{a}}. \end{aligned} \tag{6.34}$$

The two components of the constraint force, in this special case, are

$$
\begin{aligned}
f_z &= a\lambda_1 = mg, \\
f_\rho &= -2\rho_0\lambda_1 = -\frac{2\rho_0}{a}mg, \\
f_\phi &= 0.
\end{aligned}
\tag{6.35}
$$

The interested student is encouraged to consider the general problem, *i.e.*, motion not confined to the horizontal plane, and determine, for example, the general form for the constraint force. Here we simply note that, for total (conserved) energy E and (conserved) angular momentum L_z, the Lagrange multiplier has the form

$$
\lambda_1 = \frac{\frac{4E}{a^2} + \frac{8L_z^2}{ma^4} + \frac{mg}{a}}{\left(1 + \frac{4\rho^2}{a^2}\right)^2}. \tag{6.36}
$$

Another class of related and interesting problems is associated with round bodies rolling on other round bodies. The example of a cylinder rolling on another (fixed) cylinder is discussed in detail in Chapter 3 of F&W, and will appear in a slightly different form in HW IV. We can see the relevant issues by returning to the problem of the hoop on the inclined plane we discussed in Lecture 5 (see pages 12 and 13). In that problem there are formally 2 constraints, rolling on an inclined plane, which was built into the choice of coordinates (x along the plane), and rolling without slipping, which leads to the constraint in Eq. (5.36). Keeping this latter constraint in the Lagrange multiplier formalism allowed us to determine the tangential (frictional) force along the plane in Eq. (5.43). On the other hand, since we eliminated the other constraint by our choice of coordinates, we did not determine the normal force to the inclined plane. To connect to the class of problems being described here, consider the case where we replace the inclined plane by the surface of a sphere and the hoop by a disk, *i.e.*, the disk rolls (upright) on the surface of a sphere (where the starting point is the disk at rest at the top of the sphere). Since we typically want to determine when the disk flies off the sphere, we want to determine the normal force (it vanishes as the disk flies off) and so we want to explicitly keep both constraints. So consider a sphere of radius R_1 and a disk of radius R_2 as in Figure 19.2 in F&W (note that the figure looks the same even though the setup is slightly different). Further we define other variables as in the figure: r is the distance between the center of the sphere and the center of the disk; θ_1 is the polar angle of the contact point (between the sphere and the disk); θ_2 is the angle through which the disk has rotated since it started rolling down the sphere. The constraint of rolling on the sphere is (see Eq. (19.37) in F&W)

$$\phi_1(r) = r - R_1 - R_2 = 0. \tag{6.37}$$

The rolling without slipping constraint (replacing Eq. (5.36)) looks like (see Eq. (19.38) in F&W)

$$\phi_2(\theta_1,\theta_2) = R_1\theta_1 - R_2(\theta_2 - \theta_1) = 0. \tag{6.38}$$

Note in particular the final term ($+R_2\theta_1$), which might be surprising (recall that the corresponding expression for the inclined plane was $\phi(\theta,x) = x - r\theta = 0$ and we might have expected that we should just make the changes $x \to R_1\theta_1$, $r\theta \to R_2\theta_2$). The last term accounts for the fact that we are using curvilinear coordinates and the disk rotates as moves along the sphere, even if it is not rolling at all, *i.e.*, even if it is slipping with the same point always in contact with the sphere. Another way to look at Eq. (6.38) is that it guarantees that the point on the disk that is instantaneously in contact with the sphere has vanishing instantaneous velocity, $v = (R_1 + R_2)\dot{\theta}_1 - R_2\dot{\theta}_2 = 0$, which must be true if the point is not slipping on the sphere.

As discussed in F&W we can proceed to solve this problem using 2 Lagrange multipliers for Eqs. (6.37) and (6.38), and the Lagrangian

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}_1^2\right) + \frac{m}{2}\left(\frac{R_2^2}{2}\right)\dot{\theta}_2^2 - mgr\cos\theta_1, \tag{6.39}$$

i.e., find an extremal path for $L - \lambda_1\phi_1 - \lambda_2\phi_2$. In particular, we find that the multiplier corresponding to Eq. (6.37) has the form (see Eq. (19.54) in F&W modulo a minus sign)

$$\lambda_1 = \frac{mg}{3}\left(4 - 7\cos\theta_1\right). \tag{6.40}$$

Hence the normal force vanishes and the disk flies off the sphere for $\cos\theta_1 = 4/7$, $\theta_1 \simeq 55.2^\circ$.