LAGRANGE EQUATIONS AND D'ALEMBERT'S PRINCIPLE

Newton's equations are the fundamental laws of non-relativistic mechanics but their vector nature makes them simple to use only in Cartesian coordinates. The Lagrange equations represent a reformulation of Newton's laws to enable us to use them easily in a general coordinate system which is not Cartesian. Important examples are polar coordinates in the plane, spherical or cylindrical coordinates in three dimensions. The great power of the Lagrange method is that its basic equations take the same form in all coordinate systems.

In the Newtonian description we start from the second law applied to each of the particles in an N-particle system,

$$m_i \frac{d^2 r_i}{dt^2} = m_i \ddot{r}_i = F_i \quad , \quad N = 1, \ldots, N \quad . \tag{1}$$

To get away from a vector description, we can transform these N equations into a single statement about a scalar quantity which is equivalent to all N equations. To do this we introduce the concept of a *virtual displacement*. A virtual displacement of the system is defined as an arbitrary displacement δr_i of each particle but with the time frozen. In other words, it is not the physical displacement that would happen in a time δt, rather it is a mathematical displacement which we can carry out conceptually at a frozen instant of time. It then follows from (1) that for any choice of displacements,

$$\sum_{i=1}^{N} (m_i \ddot{r}_i - F_i) \cdot \delta r_i = 0 \quad ! \quad . \tag{2}$$

If we start from (1), then (2) is trivially obvious. What is significant about (2), however, is the statement that if (2) holds for *arbitrary virtual displacements* δr_i, then all N Newton equations (1) must follow. Thus the entire set of Newton equations is equivalent to the statement that (2) is true at each instant of time for any choice of the virtual displacements. Interpreted in this sense, (2) is called D'Alembert's principle.

Our aim is to find a way to write Newton's laws (1) in a way that is valid for any coordinate system. We can use (2) to see how to do this. We now define a set of *generalised coordinates* as any $3N$ numbers $q_i(t)$ whose values at time t uniquely specify the position (configuration) of all N particles in the system. Any set of numbers with this property, no matter how outlandish , will be acceptable as a coordinate system. For example, the $3N$ components x_i, y_i, z_i of the N three-vectors r_i are one possible choice of generalised coordinates. Equally acceptable are the $3N$ numbers r_i, θ_i, φ_i describing the position of the N particles in spherical polar coordinates. What we want to do then is transform from the vector coordinates r_i, $i = 1,...,N$, to the q_j, $j = 1,...,3N$. In mathematical terms each of the vectors r_i can be regarded as a function of the new coordinates q_j,

$$r_1 = r_1(q_1,q_2,...,q_{3N}) \quad , r_2 = r_1(q_1,q_2,...,q_{3N}),$$

$$... \tag{3}$$

$$r_N = r_1(q_1,q_2,...,q_{3N}) \quad .$$

Just as we can express the r_i in terms of the q_j, we assume that it is possible, given the r_i, to go back uniquely to find the q_j. In other words this transformation of coordinates must be invertible, i.e., we can go both ways.

To give a very simple example, suppose we have only one particle with position vector $r = (x,y,z)$ in Cartesian coordinates. If we describe its position in spherical polar coordinates r,θ,φ, we have

$$r = (x,y,z) = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta) \quad .$$

Thus the three q_j would be $q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$, and for any set of values of the q_j we can calculate the value of the vector r. Likewise we can find the q_j given x,y,z as $r = "x^2 + y^2 + z^2$, $\theta = \arccos(z/"x^2 + y^2 + z^2)$, $\varphi = \arctan(y/x)$.

The Lagrange equations arise by simply carrying out the above change of variables in D'Alembert's principle (2). The details of this are a bit tedious but the final result is impressive and easy to remember. No new physics is being introduced in this process so the final result is exactly equivalent to Newton's laws but it is in an extremely useful form. We begin by noting that since the r_i are functions of the q_j we can use the basic definition of partial differentiation to express the virtual displacements as

$$\delta r_i = \sum_{j=1}^{3N} \frac{\partial r_i}{\partial q_j}\delta q_j \quad , \tag{4}$$

where the δq_j describe the virtual displacement expressed as changes in the variables q_j. This enables us to transform the force term in (2) as

$$\sum_{i=1}^{N} F_i \cdot \delta r_i = \sum_{i=1}^{N} F_i \cdot \sum_{j=1}^{3N} \frac{\partial r_i}{\partial q_j}\delta q_j = \sum_{j=1}^{3N}\left(\sum_{i=1}^{N} F_i \cdot \frac{\partial r_i}{\partial q_j}\right)\delta q_j = \sum_{j=1}^{3N} Q_j \delta q_j \quad !\quad ,$$

where I have simply interchanged the summations over i and j and

$$Q_j = \sum_{i=1}^{N} F_i \cdot \frac{\partial r_i}{\partial q_j}$$

is called the *generalised force* associated with coordinate q_j.

Next we transform the acceleration terms in(2) by the use of (4).

$$\sum_{i=1}^{N} m_i \ddot{r}_i \cdot \delta r_i = \sum_{j=1}^{3N} \left(\sum_{i=1}^{N} m_i \ddot{r}_i \cdot \frac{\partial r_i}{\partial q_j} \right) \delta q_j \, .$$

To simplify the inner sum here we write

$$\sum_{i=1}^{N} m_i \ddot{r}_i \cdot \frac{\partial r_i}{\partial q_j} = \frac{d}{dt} \left(\sum_{i=1}^{N} m_i \dot{r}_i \cdot \frac{\partial r_i}{\partial q_j} \right) - \sum_{i=1}^{N} m_i \dot{r}_i \cdot \frac{d}{dt} \left(\frac{\partial r_i}{\partial q_j} \right) , \qquad (5)$$

where we use the product rule for derivatives. Since both $r_i(t)$ and $q_j(t)$ vary with time we can use the chain rule to write

$$\dot{r}_i = \sum_{j=1}^{3N} \frac{\partial r_i}{\partial q_j} \dot{q}_j \, . \qquad (6)$$

This tells us mathematically that $\dot{r}_i$ is a function that depends on both the q_j and the $\dot{q}_k$ separately. Since the $\dot{q}_k$ appear linearly it is easy to see that

$$\frac{\partial \dot{r}_i}{\partial \dot{q}_j} = \frac{\partial r_i}{\partial q_j} \, . \qquad (7)$$

We get a second identity by again applying the chain rule

$$\frac{d}{dt} \left(\frac{\partial r_i}{\partial q_j} \right) = \sum_{p=1}^{3N} \frac{\partial^2 r_i}{\partial q_p \partial q_j} \dot{q}_p = \frac{\partial}{\partial q_j} \left(\sum_{p=1}^{3N} \frac{\partial r_i}{\partial q_p} \dot{q}_p \right) = \frac{\partial \dot{r}_i}{\partial q_j} \, . \qquad (8)$$

Here in the last step we used (6). Finally we use the results (7) and (8) to transform the two terms on the right hand side of (5). In the first term of (5) we use (7) to write

$$\frac{d}{dt} \left(\sum_{i=1}^{N} m_i \dot{r}_i \cdot \frac{\partial r_i}{\partial q_j} \right) = \frac{d}{dt} \left(\sum_{i=1}^{N} m_i \dot{r}_i \cdot \frac{\partial \dot{r}_i}{\partial \dot{q}_j} \right) = \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_j} \left(\sum_{i=1}^{N} \frac{1}{2} m_i \dot{r}_i^2 \right) \right) = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) ,$$

where T is the total kinetic energy of the system. In the second term of (5) we use (8) to obtain

$$\sum_{i=1}^{N} m_i \dot{r}_i \cdot \frac{d}{dt} \left(\frac{\partial r_i}{\partial q_j} \right) = \sum_{i=1}^{N} m_i \dot{r}_i \cdot \left(\frac{\partial \dot{r}_i}{\partial q_j} \right) = \frac{\partial}{\partial q_j} \left(\sum_{i=1}^{N} \frac{1}{2} m_i \dot{r}_i^2 \right) = \frac{\partial T}{\partial q_j} \, .$$

Putting all this together gives D'Alembert's principle now in the form

$$\sum_{j=1}^{3N} \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - Q_j \right) \delta q_j = 0 \quad .$$

However, this equation must hold for all possible virtual displacements δq_j in the new variables. That is possible only if for each value of j we have

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \quad . \tag{9}$$

These equations for $j = 1,\ldots 3N$ are one form of the Lagrange equations.

For conservative forces $F_i = -\nabla_i V$, this simplifies further. In that case the generalised force (3) becomes (remember the chain rule)

$$Q_j = \sum_{i=1}^{N} F_i \cdot \frac{\partial r_i}{\partial q_j} = -\sum_{i=1}^{N} \nabla_i V \cdot \frac{\partial r_i}{\partial q_j} = -\frac{\partial V}{\partial q_j} \quad ,$$

so that Lagrange's equations (9) become

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = -\frac{\partial V}{\partial q_j} \quad . \tag{10}$$

Since the potential energy is a function of position we can regard it as a function $V\ (q_1,q_2,\ldots,q_N)$ which depends on the q_j but not on the $\dot{q}_j$. Thus $\partial V/\partial \dot{q}_j = 0$, and for conservative systems Lagrange's equations finally take the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \quad , \quad j = 1,\ldots,3N \quad , \tag{11}$$

where the *Lagrangian* L is defined to be

$$L = T - V \tag{12}$$

In the Lagrangian formulation of mechanics we may use any coordinate system we please and the equations of motion look the same. The only requirement is that we must express the kinetic energy T and the potential energy V in terms of the q_j and $\dot{q}_j$ which we have chosen. In practice, the safe way to do this is to first write T and V in Cartesian coordinates, and then use the transformation equations (3) to get T and V in terms of the new coordinates. Each of the coordinates q_j is said to describe a *degree of freedom* of the system, i.e., each q_j describes an independent way in which the system can move. We now have generalised coordinates q_j and generalised forces Q_j so it is natural to introduce *generalised momenta* p_j defined as

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \quad , \quad j = 1,\ldots,3N \quad . \tag{13}$$

The Lagrange equations now look like

$$\frac{dp_j}{dt} = \frac{\partial L}{\partial q_j} \quad .$$

Note that if L does not depend upon the particular variable q_k, we then have $\partial L/\partial q_k = 0$ and we say that the variable q_k is a *cyclic* or *ignorable* coordinate. From Lagrange's equations in the form (13) we see that $dp_k/dt = 0$ so that p_k is constant in time, i.e., it is conserved. This is the basis of all conservation laws in Lagrangian mechanics.