Appendix D: Lagrange Multipliers for problems with Constraints

Our main focus in this document is the use of Lagrange Multipliers to find candidates $\mathbf{r}_i$ for which a scalar function f($\mathbf{r}$) is stationary, df = 0, subject to constraints. In this Appendix, however, we consider a different application of Lagrange Multipliers that relates to classical mechanics for a system of N particles in the presence of constraints. A certain amount of background information is required to put this application into context, which we now present.

Constraints and virtual displacements

Let $\mathbf{r}_k(t)$ be the position of the kth particle. At any fixed time t, if the particles are unencumbered with constraints, the variables ($\mathbf{r}_1$, $\mathbf{r}_2$..... $\mathbf{r}_k$) can be regarded as a single point in the space E^{3N} and the "configuration space" of the system indicated by ($\mathbf{r}_1$, $\mathbf{r}_2$..... $\mathbf{r}_k$) is all of this space E^{3N}. All points in E^{3N} are "legal" points for the system.

A **holonomic** constraint has the form

$$a(\mathbf{r}_1, \mathbf{r}_2 ..... \mathbf{r}_N, t) = 0. \qquad (D.1)$$

At a given time t, this equation represents a surface of dimension 3N-1 in E^N. If there are s such holonomic constraints

$$f_i(\mathbf{r}_1, \mathbf{r}_2 ..... \mathbf{r}_N, t) = 0 \qquad i = 1,2...s \qquad (D.2)$$

then at time t the system vector ($\mathbf{r}_1$, $\mathbf{r}_2$..... $\mathbf{r}_N$) is constrained to lie on a surface of dimension 3N-s in E^N, as discussed in Appendix C, so each extra constraint lowers the operating surface dimension by 1. One can think of the term holonomic as meaning that the constraints are functions of "whole" coordinates like $\mathbf{r}_2$ as opposed to differential coordinates like $d\mathbf{r}_2$.

A **non-holonomic** constraint is one that is not holonomic. An inequality constraint like $a(\mathbf{r}_1, \mathbf{r}_2 ..... \mathbf{r}_N, t) < 0$ which might bound the particles to one side of a surface is an example of a nonholonomic constraint, but our interest lies more in nonholonomic constraints of the form

$$\Sigma_{k=1}^{N} \mathbf{b}_k \bullet d\mathbf{r}_k + b_t dt = 0 \qquad (D.3)$$

in which differentials rather than "whole" coordinates appear. The "coefficients" $\mathbf{b}_k$ and b_t can in general be functions of ($\mathbf{r}_1$, $\mathbf{r}_2$..... $\mathbf{r}_k$, t). If it happens that such a constraint can be integrated by some hook or crook to give an equation of the form $a(\mathbf{r}_1, \mathbf{r}_2 ..... \mathbf{r}_k, t) = 0$, then it is regarded as being a holonomic constraint, so a non-holonomic constraint must be non-integrable. If there are m such non-holonomic constraints imposed on a system, one could write them as

$$\Sigma_{k=1}^{N} \mathbf{A}_{ik} \bullet d\mathbf{r}_k + A_{it} dt = 0 \qquad i = 1,2...m$$

or (D.4)

$$\Sigma_{k=1}^{N} \mathbf{A}_{ik} \bullet \dot{\mathbf{r}}_k + A_{it} = 0 \qquad i = 1,2...m \ .$$

Here bolded $\mathbf{A}_{ik}$ represents 3 matrices each of which is m x N (rows x columns). Since these constraints are non-integrable, one cannot really described them in terms of surfaces in E^{3N}.

If one differentiates the holonomic constraints (D.2) one gets

$$\Sigma_{k=1}^{N} [\nabla^{(k)} f_i] \bullet d\mathbf{r}_k + [\partial_t f_i]\, dt = 0 \qquad i = 1,2...s$$

or (D.5)

$$\Sigma_{k=1}^{N} [\nabla^{(k)} f_i] \bullet \dot{\mathbf{r}}_k + [\partial_t f_i] = 0 \qquad i = 1,2...s$$

where $\nabla^{(k)} \equiv (\partial/\partial(\mathbf{r}_k)_1, \partial/\partial(\mathbf{r}_k)_2, \partial/\partial(\mathbf{r}_k)_3)$ is a gradient with respect to the components of $\mathbf{r}_k$. These equations have the same form as those in (D.4), but these are holonomic because they are integrable to give $f_i(\mathbf{r}_1, \mathbf{r}_2 ..... \mathbf{r}_N, t) = 0$.

For convenience below, we shall now combine (D.4) and (D.5) into a single set of equations:

$$\Sigma_{k=1}^{N} \mathbf{B}_{ik} \bullet d\mathbf{r}_k + B_{it} dt = 0 \qquad i = 1,2....C \qquad // \ C = s+m$$

or (D.6)

$$\Sigma_{k=1}^{N} \mathbf{B}_{ik} \bullet \dot{\mathbf{r}}_k + B_{it} = 0 \qquad i = 1,2....C$$

where

$$\begin{array}{lllll} \mathbf{B}_{ik} = \mathbf{A}_{ik} & B_{it} = \mathbf{A}_{it} & \text{for } i = 1,2...s & \text{non-holonomic} & \\ B_{it} = \nabla^{(k)} f_i & B_{it} = \partial_t f_i & \text{for } i = s+1, s+2....C & \text{holonomic} \ . & (D.7) \end{array}$$

In general, the functions f_i, $\mathbf{B}_{ik}$ and B_{it} are all functions of $(\mathbf{r}_1, \mathbf{r}_2 ..... \mathbf{r}_k, t)$.

In the above we have followed the notation of Ray and Shamanna (2006) and we continue our path in the manner they propose.

Suppose the set of N differentials $\{d\mathbf{r}_k\}$ satisfies all of equations (D.6) for some time interval dt. Such a set $\{d\mathbf{r}_k\}$ is then called a set of **allowed displacements**. Imagine that $\{d\mathbf{r'}_k\}$ is some other distinct allowed displacement set. Elements of the difference set defined as

$$\{\delta\mathbf{r}_k\} = \{d\mathbf{r}_k\} - \{d\mathbf{r'}_k\} \qquad k = 1,2...N \qquad (D.8)$$

are then called **virtual displacements**. This is how Ray and Shamanna define the vague term virtual displacement which appears in many texts (about which they have much to say in their Section I A). It follows then that for any virtual displacement set $\{\delta\mathbf{r}_k\}$ equations (D.6) take this form

$$\Sigma_{k=1}^{N} \mathbf{B}_{ik} \bullet \delta\mathbf{r}_k = 0 \qquad (D.9)$$

simply because the last term in (D.6) cancels out when one writes $\delta\mathbf{r}_k = d\mathbf{r}_k - d\mathbf{r'}_k$. Notice that the $\delta\mathbf{r}_k$ are not arbitrary differential displacements. They are differences of allowed displacements.

Meanwhile, Newton's Law for a system of N particles subject to constraints has the form

$$m\ddot{\mathbf{r}}_k = \mathbf{F}_k + \mathbf{R}_k \qquad k = 1,2...N \qquad \text{(D.10)}$$

where $\mathbf{R}_k$ is the sum of all constraint forces applied to particle k and $\mathbf{F}_k$ is the sum of all *other* forces applied to particle k.

Assumption that constraints do no virtual work

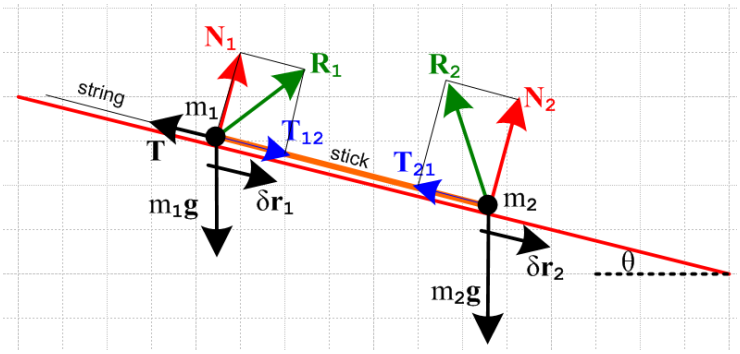
Next, *assume* that (R&S refer to this as being an "ideal constraint" situation),

$$0 = \Sigma_{k=1}^{N} \mathbf{R}_k \bullet \delta\mathbf{r}_k \ . \qquad \text{(D.11)}$$

In Goldstein this equation corresponds to setting the second term in (1-40) to zero. The main justification for making this assumption is that it is valid for particles which form a rigid body and for many other constraint situations. There seems to be no general justification of this equation for *all* constraint situations, and it is certainly not valid if friction is present. The equation can be interpreted as saying that the **total virtual work** done by all the constraint forces on all the particles of the system vanishes.

As a first simple example, consider two particles 1 and 2 independently sliding down a frictionless static ramp. In this case for particle 1 the constraint force $\mathbf{R}_1$ is the normal force of the ramp acting on the particle, and this is clearly perpendicular to $d\mathbf{r}_1$ for any allowed displacement set $\{d\mathbf{r}_k\}$, and thus $\mathbf{R}_1$ is also perpendicular to any $\delta\mathbf{r}_1$ which is part of any virtual displacement set $\{\delta\mathbf{r}_k\}$. In this case we have $\mathbf{R}_1 \bullet \delta\mathbf{r}_1 = 0$ and $\mathbf{R}_2 \bullet \delta\mathbf{r}_2 = 0$ so the terms in (D.11) are separately zero.

A second example is more interesting. Particles 1 and 2 of arbitrary masses slide down a one dimensional ramp but are glued to the two ends of a massless stick so they comprise a rigid body. The upper mass is attached to a string which provides a constant tension T. Here is a picture,



(D.12)

Particle 2 exerts a constraint force $\mathbf{T}_{12}$ on particle 1 through the stick, and Particle 1 exerts a constraint force $\mathbf{T}_{21}$ on particle 2. Since there is only one force of tension (or compression) in such a stick, $T_{12} = T_{21}$. All allowed displacement sets $\{d\mathbf{r}_1, d\mathbf{r}_2\}$ are of the form $\{d\mathbf{r}, d\mathbf{r}\}$, and therefore all *virtual* displacement sets are of the form $\{\delta\mathbf{r}_1, \delta\mathbf{r}_2\} = \{\delta\mathbf{r}, \delta\mathbf{r}\}$. In (D.12) we have drawn the two normal forces $\mathbf{N}_i$ and the two *total* constraint forces $\mathbf{R}_i$. It is no longer true that $\mathbf{R}_1 \bullet \delta\mathbf{r}_1 = 0$ and $\mathbf{R}_2 \bullet \delta\mathbf{r}_2 = 0$ as the

picture shows. However, it is true that $\Sigma_{k=1}^{2}\ \mathbf{R}_k \bullet \delta\mathbf{r}_k = [\ \Sigma_{k=1}^{2}\ \mathbf{R}_k\ ] \bullet \delta\mathbf{r} = 0$ because $\mathbf{R}_1$ and $\mathbf{R}_2$ have equal and opposite components along the ramp, so equation (D.11) is valid for this example.

Now suppose the ramp is translating in some direction not along the ramp. The displacements $d\mathbf{r}_1 = d\mathbf{r}_2$ are then no longer along the ramp so that $\Sigma_{k=1}^{2}\ \mathbf{R}_k \bullet d\mathbf{r}_k \neq 0$. However, any virtual displacement set would still have $\delta\mathbf{r}_1 = \delta\mathbf{r}_2$ along the ramp and then again $\Sigma_{k=1}^{2}\ \mathbf{R}_k \bullet \delta\mathbf{r}_k = 0$. Remember that a virtual displacement set is the difference of any two allowed displacement sets, and in this subtraction the effect of the ramp translation cancels out. So equation (D.11) is still valid.

Other examples are given in Ray and Shamanna Section III.

Reader Exercise: show in the last example that $T_{12} = T\ m_2 / (m_1+m_2)$ and that this makes sense if $m_1 >> m_2$ and vice versa.

We proceed then assuming our system of interest respects equation (D.11).

Comment on internal forces: In (D.10) we have classified the forces acting on particle k into two groups and we write $\mathbf{R}_k + \mathbf{F}_k$, where $\mathbf{R}_k$ are "constraint" forces and the $\mathbf{F}_k$ are any "other" forces. Goldstein refers to the $\mathbf{F}_k$ forces as "applied" forces in (1-39), while Ray and Shamanna call them "external" in Sec IV. Consider a system consisting of N masses all connected by a network of massless springs. The force acting on particle k due to all the springs to which it is attached would be best classified in the $\mathbf{F}_k$ "other" category. Were we to classify this force into the $\mathbf{R}_k$ group, we could not use the no-constraint-work property (D.12) since spring forces do work as the springs compress and expand. An example of such a system is the Solar System where the spring forces are represented by gravity. It would seem a misnomer to classify the internal forces in this case as "applied" or "external". However, if all the springs are slowly adjusted until they become infinitely stiff, in that limit it would be better to classify the internal spring forces into the $\mathbf{R}_k$ constraint forces group, since these springs do no work and (D.11) applies. The system is then a "rigid body" and in this case, with the internal forces classified as $\mathbf{R}_k$ constraint forces, one benefits from the many simplifications which arise for rigid body mechanics. The internal forces in our spring example fall into the category where $\mathbf{F}_{ij} = -\ \mathbf{F}_{ji}$ and $\mathbf{F}_{ij}$ is in the direction $\mathbf{r}_i - \mathbf{r}_k$. When this is not the case, the rigid body limit does not apply, but the internal forces can still be classified as "constraint" plus "other".

An Application of Lagrange Multipliers

Recall the C = m+s constraint equations (D.9),

$$\Sigma_{k=1}^{N}\ \mathbf{B}_{ik} \bullet \delta\mathbf{r}_k = 0 \qquad i = 1,2...C \qquad C \equiv m+s = \text{total number of constraints} \qquad \textit{(D.9)}$$

Multiply each of these C equations by a Lagrange Multiplier λ_i and then add the equations to get

$$0 = \Sigma_{i=1}^{C}\ \lambda_i\ [\Sigma_{k=1}^{N}\ \mathbf{B}_{ik} \bullet \delta\mathbf{r}_k] = 0$$

$$= \Sigma_{k=1}^{N}\ [\ \Sigma_{i=1}^{C}\ \lambda_i\ \mathbf{B}_{ik}\ ] \bullet\ \delta\mathbf{r}_k\ . \qquad \text{(D.13)}$$

Now subtract equation (D.13) from (D.11) to get

$$0 = \Sigma_{k=1}^{N}\ [\ \mathbf{R}_k - \Sigma_{i=1}^{C}\ \lambda_i\ \mathbf{B}_{ik}] \bullet \delta\mathbf{r}_k\ . \qquad \text{(D.14)}$$

We know that the N vectors in the set $\{\delta \mathbf{r}_k\}$ are not linearly independent because they must satisfy the equations (D.9), so we cannot claim from (D.14) that $[\ \mathbf{R}_k - \Sigma_{i=1}^{C} \lambda_i\ \mathbf{B}_{ik}] = 0$ for each k. But now write out equation (D.14) in full detail, showing all vector components,

$$0 = \Sigma_{k=1}^{N}\ \Sigma_{j=1}^{3}\ [\ (\mathbf{R}_k)_j - \Sigma_{i=1}^{C} \lambda_i\ (\mathbf{B}_{ik})_j]\ (\delta \mathbf{r}_k)_j\ . \qquad \text{(D.15)}$$

Next, relabel the 3N terms in this sum using a single index n defined in this way

$$\begin{array}{llllllllll} (k,j) & = (1,1), & (1,2), & (1,3), & (2,1), & (2,2), & (2,3), & \ldots\ldots\ldots(N,1), & (N,2), & (N,3) \\ n & = 1, & 2, & 3, & 4, & 5, & 6\ , & 3N\text{-}2, & 3N\text{-}1, & 3N \end{array}$$

so (D.16)

$$n = 3(k-1)+j \qquad k = 1+ \mathrm{Int}[(n-1)/3] \qquad j = 1 + \mathrm{Rem}[(n-1)/3]\ .$$

Then define (here the n value is implied by the k,j labels as per above)

$$x_n \equiv (\mathbf{r}_k)_j \qquad R_n \equiv (\mathbf{R}_k)_j \qquad B_{in} \equiv (\mathbf{B}_{ik})_j\ . \qquad \text{(D.17)}$$

Equation (D.15) can then be expressed as

$$0 = \Sigma_{n=1}^{3N}\ [\ R_n - \Sigma_{i=1}^{C} \lambda_i\ B_{in}]\ \delta x_n \qquad \text{(D.18)}$$

and (D.9) becomes $0 = \Sigma_{k=1}^{N}\ \Sigma_{j=1}^{3}\ (\mathbf{B}_{ik})_j\ (\delta \mathbf{r}_k)_j$ or

$$0 = \Sigma_{n=1}^{3N} B_{in}\ \delta x_n\ . \qquad i = 1,2...C\ . \qquad \text{(D.19)}$$

Once again, we cannot set the square brackets in (D.18) to 0 because the 3N coordinates δx_n are not linearly independent due to the C equations (D.19).

At any given time t, due to these equations (D.19), there must exist at least one subset of the 3N variables δx_n which are independent and the remaining C δx_n are dependent. Assume that the dependent variables have indices $n = n_1, n_2,...n_C$. Consider then the following set of C equations involving the dependent-term square brackets in (D.18),

$$[\ R_n - \Sigma_{i=1}^{C} \lambda_i\ B_{in}] = 0 \qquad n = n_1, n_2,...n_C\ . \qquad \text{(D.20)}$$

This represents C linear equations in the C parameters λ_i so, when all is said and done, there must exist a set of values $\{\lambda_i\}$ which make these C equations be true. We assume these values will be $\lambda_i(t)$ and this is justified below when the full set of equations is considered.

Now the terms in (D.18) indicated by $n = n_1, n_2,...n_C$ all vanish due to (D.20), so (D.18) can be written,

$$0 = \Sigma_{n \neq n_1, n_2, \ldots n_C}\ [\ R_n - \Sigma_{i=1}^{C} \lambda_i\ B_{in}]\ \delta x_n\ . \qquad \text{(D.21)}$$

But all the δx_n appearing in (D.21) are independent, so the square brackets here must also vanish. We then end up with *all* the square brackets in (D.18) being 0 (assuming the correct solution values of λ_i),

$$[\, R_n - \Sigma_{i=1}^{C} \lambda_i B_{in}] = 0 \qquad n = 1,2....3N$$

or

$$R_n = \Sigma_{i=1}^{C} \lambda_i \, B_{in} \qquad n = 1,2....3N$$

or

$$(\mathbf{R_k})_j = \Sigma_{i=1}^{C} \lambda_i \, (\mathbf{B_{ik}})_j \qquad k = 1,2..N \qquad j = 1,2,3$$

or

$$\mathbf{R_k} = \Sigma_{i=1}^{C} \lambda_i \, \mathbf{B_{ik}} \qquad k = 1,2...N \ . \qquad \text{(D.22)}$$

Thus the constraint forces are certain linear combinations of the $\mathbf{B_{ik}}$ constraint functions shown in (D.7).

Meanwhile, Newton's Law (D.10) says that $m_k \mathbf{a_k} = \mathbf{F_k} + \mathbf{R_k}$ where the $\mathbf{F_k}$ are the non-constraint forces for the problem. Inserting (D.22) gives

$$m_k \ddot{\mathbf{r}}_k = \mathbf{F_k} + \Sigma_{i=1}^{C} \lambda_i \mathbf{B_{ik}} \qquad \mathbf{B_{ik}} = \begin{cases} \mathbf{A_{ik}} & i = 1 \text{ to m (non-holonomic)} \\ [\nabla^{(k)} f_i] & i = m+1 \text{ to s (holonomic)} \end{cases} \qquad \text{(D.23)}$$

The problem is then summarized in the following set of equations from (D.23) and (D.6),

$$m_k \dot{\mathbf{v}}_k(t) = \mathbf{F_k}(\mathbf{r_1}, \mathbf{r_2}..... \mathbf{r_N}, t) + \Sigma_{i=1}^{C} \lambda_i(t) \mathbf{B_{ik}}(\mathbf{r_1}, \mathbf{r_2}..... \mathbf{r_N}, t) \qquad k = 1..N \qquad 3N \text{ equations}$$

$$\dot{\mathbf{r}}_k(t) = \mathbf{v_k}(t) \qquad k = 1..N \qquad 3N \text{ equations}$$

$$\Sigma_{k=1}^{N} \mathbf{B_{ik}}(\mathbf{r_1}, \mathbf{r_2}..... \mathbf{r_N}, t) \bullet \dot{\mathbf{r}}_k(t) + B_{it}(\mathbf{r_1}, \mathbf{r_2}..... \mathbf{r_N}, t) = 0 \qquad i = 1,2....C \ . \qquad C \text{ equations}$$

(D.24)

This can be regarded as a set of 6N+C scalar first-order ODEs in the single variable t for the 6N+C unknown scalar functions $\mathbf{r_k}(t)$, $\mathbf{v_k}(t)$ and $\lambda_i(t)$. Although the first 3N equations happen to be linear in the $\lambda_i(t)$, the equations are in general non-linear in the $\mathbf{r_k}(t)$ so this is then a *non-linear* system of ODE's. Nevertheless, a well known existence theorem (e.g., Ince Section 3.3) states that, provided the $\mathbf{F_k}$, $\mathbf{B_{ik}}$ and B_{it} functions are continuous and differentiable in all their arguments, a unique solution exists for any set of initial values $\mathbf{r_k}(0)$, $\mathbf{v_k}(0)$ and $\lambda_i(0)$. As (D.22) shows, the $\lambda_i(0)$ are related to the constraint forces at time 0.

The uniqueness of the solution agrees with our intuition that a physical system will evolve in a unique manner. Finding an analytic form for the solution might not be possible, but a numeric form can always be obtained.

Summary of the above Lagrange Multiplier Application

We have used a set of Lagrange Multipliers as an aid to solving a mechanics problem involving N particles with C constraints. As already noted, the above presentation is based on Section IV of Ray and Shamanna. The steps were as follows, considered more generically:

1. One has a sum $\Sigma_{n=1}^{N} R_n \delta x_n = 0$. (D.25)

2. One also has a set of sums $\Sigma_{n=1}^{N} B_{in} \delta x_n = 0$ for i = 1,2...C with C < N (constraints in the above case). This implies that one can find C of the δx_n which are linearly dependent on the other δx_n.

3. If λ_i are C arbitrary factors, certainly $\Sigma_{i=1}^{C} \lambda_i (\Sigma_{n=1}^{N} B_{in} \delta x_n) = 0$ based on 2. These λ_i are "Undetermined Lagrange Multipliers".

4. One constructs from steps 1 and 3 the new sum $\Sigma_{n=1}^{N} [R_n - \Sigma_{i=1}^{C} \lambda_i B_{in}] \delta x_n = 0$.

5. The C λ_i are selected so that [..] = 0 in the C dependent δx_n terms of this sum. The remaining sum has only independent δx_n terms so [..] = 0 for all those terms as well. Then *all* [...] = 0 in the sum of step 4.

6. The conclusion is that $[R_n - \Sigma_i \lambda_i B_{in}] = 0$ for all n, assuming the solution λ_i values of step 5.

In contrast with our main topic of finding stationary points df = 0 of a function f(**r**), in the Lagrange Multiplier application just presented there is no explicit function f(**r**) we are making stationary. On the other hand, one can regard (D.11) that $\Sigma_k \mathbf{R}_k \bullet \delta \mathbf{r}_k = 0$ as the statement $\delta W = 0$ where $\delta W = \Sigma_k \mathbf{R}_k \bullet \delta \mathbf{r}_k$ is the total differential "virtual work" done by the constraint forces during a differential virtual displacement $\{\delta \mathbf{r}_k\}$. Using Newton's law (D.10) this equation can then be written

$$0 = \delta W = \Sigma_k [m_k \mathbf{a}_k - \mathbf{F}_k] \bullet \delta \mathbf{r}_k \qquad \text{(D.26)}$$

in which form it is known as **D'Alembert's Principle** (1743, Goldstein (1-42)). The analog of f(**r**) is then an action integral W of the differential virtual work δW which integral one renders stationary to get δW = 0. The notion of "the action" is mentioned at the end of this appendix.)

We now present another application of the Lagrange Multiplier technique outlined in (D.25). As before, a certain amount of background is needed before the application can be demonstrated.

Generalized coordinates, generalized forces, and Lagrange's Equations

Usually the Cartesian components of the particle positions $\mathbf{r}_k$ are not the most convenient variables to use in solving a problem. Angles are often more useful. With C holonomic constraints one can replace the 3N variables $\{x_n\}$ in (D.17) with a set of 3N-C *independent* variables $\{q_n\}$ where

$$x_n = x_n(q_1, q_2, ...q_{3N-C}, t) \qquad \text{for n=1 to 3N} \qquad \text{(D.27)}$$

and one can then rework the above presentation in these new **generalized coordinates** q_j which then won't always have the dimensions of length. For example, one then has, for virtual displacements,

$$\delta x_n = \Sigma_{j=1}^{3N-C}(\partial x_n/\partial q_j)\,\delta q_j \qquad n = 1,2..3N$$

or

$$\delta \mathbf{r}_k = \Sigma_{j=1}^{3N-C}(\partial \mathbf{r}_k/\partial q_j)\,\delta q_j \qquad k = 1,2..N \qquad \text{// as in Goldstein (1-44).} \qquad \text{(D.28)}$$

The usual presentation of Lagrangian dynamics with holonomic constraints *avoids* dealing with the forces of constraint and also avoids the need for using a set of Lagrange Multipliers. That presentation appears in Section V of the Ray and Shamanna and Goldstein p 17 (and many other places) and proceeds like this:

1. Start with $\Sigma_k \mathbf{R}_k \bullet \delta \mathbf{r}_k = 0$ in (D.11) which we write in the form of D'Alembert's Principle (D.26) already mentioned above,

$$0 = \Sigma_{k=1}^{M} [m_k \mathbf{a}_k - \mathbf{F}_k] \bullet \delta \mathbf{r}_k \qquad \textit{(D.26)}$$

where $\mathbf{F}_k$ is the total non-constraint force acting on particle k.

2. Install (D.28) into D'Alembert's Principle to get

$$0 = \Sigma_k [m_k \mathbf{a}_k - \mathbf{F}_k] \bullet \Sigma_j \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right)\delta q_j$$

$$= \Sigma_j \{ \Sigma_k m_k \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_j} - \Sigma_k \mathbf{F}_k \bullet \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right) \}\ \delta q_j$$

$$= \Sigma_j \{ \Sigma_k m_k \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_j} - Q_j \}\ \delta q_j \qquad \text{(D.29)}$$

where

$$Q_j \equiv \Sigma_{k=1}^{M} \mathbf{F}_k \bullet \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right) . \qquad \text{(D.30)}$$

Q_j defined above are the **generalized forces** associated with the "generalized coordinates" q_j .

3. Make use of the following non-obvious result (derived below),

$$\Sigma_k m_k \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} \qquad \text{(D.31)}$$

where T is the total kinetic energy of the N particles,

$$T \equiv (1/2)\Sigma_k m_k (\mathbf{v}_k \bullet \mathbf{v}_k) . \qquad \text{(D.32)}$$

Then (D.29) reads,

$$0 = \Sigma_j \{ \frac{d}{dt}(\frac{\partial T}{\partial \dot{q}_j}) - \frac{\partial T}{\partial q_j} - Q_j \} \delta q_j . \qquad (D.33)$$

Since the δq_j are independent variables, {..} = 0 and we end up with one form of "Lagrange's Equations"

$$\frac{d}{dt}(\frac{\partial T}{\partial \dot{q}_j}) - \frac{\partial T}{\partial q_j} = Q_j \qquad j = 1,2....3N-C \quad . \qquad \text{// as in Goldstein (1-50)} \qquad (D.34)$$

If the applied forces can be derived from a potential $\mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, ... \mathbf{r}_N)$ such that

$$\mathbf{F}_k = - \nabla^{(k)} \mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, ... \mathbf{r}_N) \qquad \text{// no t or } \dot{\mathbf{r}}_i \text{ dependence in } \mathcal{V} \qquad (D.35)$$

then from (D.30),

$$Q_j \equiv \Sigma_k \mathbf{F}_k \bullet (\frac{\partial \mathbf{r}_k}{\partial q_j}) = - \Sigma_k (\frac{\partial \mathcal{V}}{\partial \mathbf{r}_k}) \bullet (\frac{\partial \mathbf{r}_k}{\partial q_j}) = - \frac{\partial V}{\partial q_j} \qquad (D.36)$$

where $V(q_i) \equiv \mathcal{V}(\mathbf{r}_k(q_i))$ is a function of the generalized coordinates q_i. Since V is not a function of the $\dot{q}_i$ we know that $\frac{\partial V}{\partial \dot{q}_j} = 0$. Using this fact, and (D.36) that $Q_j = - \frac{\partial V}{\partial q_j}$, (D.34) can be written

$$\frac{d}{dt}(\frac{\partial (T-V)}{\partial \dot{q}_j}) - \frac{\partial (T-V)}{\partial q_j} = 0 . \qquad (D.37)$$

The last step is to define the classical **Lagrangian**

$$L \equiv T - V \qquad (D.38)$$

to obtain

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_j}) - \frac{\partial L}{\partial q_j} = 0 \qquad j = 1,2....3N-C \quad . \qquad \text{// as in Goldstein (1-53)} \qquad (D.39)$$

which is the more conventional form of "Lagrange's Equations" with C holonomic constraints.

One can define **generalized (canonical, conjugate) momenta** $p_j \equiv \frac{\partial L}{\partial \dot{q}_j}$ so the Lagrange equations (D.39) are just $\dot{p}_j = \frac{\partial L}{\partial q_j}$ or $\dot{\mathbf{p}} = \nabla^{(q)} L$. If $V = V(q_i)$, then $p_j = \frac{\partial T}{\partial \dot{q}_j}$ and then if $T = \frac{1}{2} \Sigma_k m_k \dot{q}_k^2$ one has the familiar result $p_j = m_j q_j$. When $T = T(\dot{q}_i)$ one finds $Q_j \equiv - \frac{\partial V}{\partial q_j} = \frac{\partial L}{\partial q_j} = \dot{p}_j$ so $\dot{\mathbf{p}} = \mathbf{Q}$, which is the **generalized Newton's Law** in terms of generalized coordinates and generalized forces.

ok to here

Footnote: Derivation of the result (D.31) used above

Define the 3N-C independent generalized coordinates q_i as in (D.27) according to

$$\mathbf{r}_k = \mathbf{r}_k(q_1, q_2 \ldots q_{3N-C}, t) \equiv \mathbf{r}_k(q_i, t) \,. \tag{D.40}$$

Compute the total time derivative of $\mathbf{r}_k$,

$$\mathbf{v}_k = \dot{\mathbf{r}}_k = \frac{d\mathbf{r}_k}{dt} = \Sigma_j \frac{\partial \mathbf{r}_k}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial \mathbf{r}_k}{\partial t} = \Sigma_j \frac{\partial \mathbf{r}_k}{\partial q_j}\dot{q}_j + \frac{\partial \mathbf{r}_k}{\partial t} \qquad // = \mathbf{v}_k(q_i, \dot{q}_i, t) \,. \tag{D.41}$$

Then based on this result compute these two partial derivatives of $\mathbf{v}_k$,

$$\frac{\partial \mathbf{v}_k}{\partial q_i} = \Sigma_j \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial t} \tag{D.42}$$

$$\frac{\partial \mathbf{v}_k}{\partial \dot{q}_i} = \frac{\partial \mathbf{r}_k}{\partial q_i} \,. \tag{D.43}$$

Next, the total kinetic energy of all k particles is given by (D.32),

$$T \equiv (1/2)\,\Sigma_k\, m_k\, (\mathbf{v}_k \bullet \mathbf{v}_k) \tag{*D.32*}$$

so that

$$\frac{\partial T}{\partial \dot{q}_i} = \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{\partial \mathbf{v}_k}{\partial \dot{q}_i} = \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_i} \,. \qquad // \text{ using (D.43)} \tag{D.44}$$

Apply d/dt to the above, with $\mathbf{a}_k = \dot{\mathbf{r}}_k$:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = \Sigma_k\, m_k\, \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_i} + \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{d}{dt}\left(\frac{\partial \mathbf{r}_k}{\partial q_i}\right) \,. \tag{D.45}$$

Compute the total time derivative appearing in the second term,

$$\frac{d}{dt}\left(\frac{\partial \mathbf{r}_k}{\partial q_i}\right) = \Sigma_j \frac{\partial^2 \mathbf{r}_k}{\partial q_j \partial q_i}\dot{q}_j + \frac{\partial^2 \mathbf{r}_k}{\partial t\, \partial q_i} \,. \tag{D.46}$$

Meanwhile,

$$\frac{\partial T}{\partial q_i} = \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{\partial \mathbf{v}_k}{\partial q_i}$$

$$= \Sigma_k\, m_k\, \mathbf{v}_k \bullet [\, \Sigma_j \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial q_j} \dot{q}_j + \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial t} \,] \qquad // \text{(D.42)}$$

$$= \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{d}{dt}(\frac{\partial \mathbf{r}_k}{\partial q_i}) \quad . \qquad // \text{(D.46)} \qquad \text{(D.47)}$$

Subtract (D.47) from (D.45) to get

$$\frac{d}{dt}(\frac{\partial T}{\partial \dot{q}_i}) - \frac{\partial T}{\partial q_i} = \Sigma_k\, m_k\, \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_i} \qquad \text{(D.48)}$$

which is the result quoted above in (D.31).

Lagrange Equations with non-holonomic constraints

This is our final Lagrange Multiplier example, and it exactly follows the generic prescription of (D.25). The general topic is well described in Goldstein Chapter 2.

As the starting point we invoke **Hamilton's Principle** which is this :

$$\delta S = 0 \qquad \text{where} \qquad S = \int_{t1}^{t2} dt\, L(q_n(t), \dot{q}_n(t), t) \qquad \text{(D.49)}$$

where there are M generalized coordinates q_n. Here L = T - V is the Lagrangian (D.38) with $T(\dot{q}_i)$ and $V(q_i)$ being the kinetic and potential energies associated with the M generalized coordinates. We shall assume that M = N-D where there were D holonomic constraints on N initial problem coordinates. As shown in Goldstein p 41 (2-20') based on p 37 (2-15), the equation δS = 0 results in the following sum being 0,

$$1. \quad \Sigma_{n=1}^{M} [\, \int_{t1}^{t2} dt\, \{ \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_n}) - \frac{\partial L}{\partial q_n}) \,\} \,]\, \delta q_n(t) = 0 \quad . \qquad \text{(D.50)}$$

The number 1. on the left refers to the first step outlined in (D.25). In this equation the $\delta q_n(t)$ are M independent functions of t ("path variations"), restricted only by $\delta q_n(t_1) = \delta q_n(t_2) = 0$. In order for (D.50) to be true for M arbitrary independent *functions* $\delta q_n(t)$, one must have {} = 0,

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_n}) - \frac{\partial L}{\partial q_n} = 0 \qquad \text{(D.51)}$$

which are the Lagrange Equations (D.39). Thus it is that the Lagrange Equations can be derived from Hamilton's Principle in the case of only holonomic constraints.

But suppose that in addition to the D holonomic constraints there are C non-holonomic constraints, so only M-C of the δq_n are independent. In this case the above conclusion is incorrect.

2. The C non-holonomic constraints result in C sums $\Sigma_n A_{ni} \delta q_n = 0$ i = 1,2...C analogous to (D.9) above.

3. If λ_i are C arbitrary factors, certainly $\Sigma_i \lambda_i (\Sigma_n A_{in} \delta q_n) = 0$ based on 2. These λ_i are "Undetermined Lagrange Multipliers". Then $\int_{t1}^{t2} dt \ (\Sigma_i \lambda_i \Sigma_n A_{in} \delta q_n) = 0$ as well.

4. Construct from 1 and 3 the new sum,

$$\Sigma_{n=1}^{M} \left[\int_{t1}^{t2} dt \left\{ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_n} \right) - \frac{\partial L}{\partial q_n} - \Sigma_i \lambda_i A_{in} \right\} \right] \delta q_n(t) = 0 \quad \text{// as in Goldstein (2-26)} \qquad (D.52)$$

5. The C λ_i are selected so that {..} = 0 in the C dependent $\delta q_n(t)$ terms of this sum. The remaining sum has only independent $\delta q_n(t)$ terms so {..} = 0 for all those terms as well. Then *all* {...} = 0 in the sum of 4.

6. The conclusion is that

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_n} \right) - \frac{\partial L}{\partial q_n} = \Sigma_i \lambda_i A_{in} \ . \qquad (D.53)$$

This then is the corrected form of the Lagrange Equations in the presence of non-holonomic constraints, as in Goldstein (2-30). Comparing (D.53) with (D.34), one may *interpret* $Q'_n \equiv \Sigma_i \lambda_i A_{in}$ as the generalized forces that would have to be present to simulate the effect of the non-holonomic constraints were they not present. The forces for any holonomic constraints are already accounted for by L as in (D.39).

The integral S shown in (D.49) is called "the action" and Hamilton's Principle is the classical mechanics incarnation of the **Principle of Least Action**. In relativistic classical and quantum field theory the action is a spacetime integral of the Lagrangian density $\mathcal{L}(\varphi_n(x^\mu), \partial_\mu \varphi_n(x^\mu), x^\mu)$ where the coordinates q_n of L are replaced by fields $\varphi_n(x^\mu)$, and t is replaced by x^μ, a point in spacetime.

****

E. L. Ince, *Ordinary Differential Equations* (Longmans, Green & Co. (Ltd.), London, 1927). This classic text became a Dover book in 1956 and is available online used for $4, shipping included (576p).

S. Ray and J. Shamanna, "On Virtual Displacement and Virtual Work in Lagrangian Dynamics", European Journal of Physics, 27 (2006) 311-329. Also https://arxiv.org/abs/physics/0510204v2 .

Ray and Shamanna