Proofed in full at 3PM 9.26.16

Appendix D: Lagrange Multipliers for problems with Holonomic Constraints

In the Comment 4 at the end of Section 2 we note that sometimes the constraints on a system with "generalized coordinates" q_k can be written $\Sigma_k a_{ik} dq_k + a_{it} dt = 0$. Such constraints written in terms of differentials of the coordinates are examples of **non-holonomic** constraints. This situation arises typically when constraints act on velocities rather than positions. An example is a vertical disk rolling on a horizontal plane as shown in Goldstein p 13. Such constraints are difficult to handle because the coordinates q_k are not generally independent, meaning that the values of all the q_k cannot be set independently. There is no constraint equation of the form $a(q_1,q_2.....) = 0$ which is a function of the whole coordinates q_i; differential constraints are "non-whole". Inequality constraints are also non-holonomic, such as for particles bounded inside a box.

Nevertheless, as shown in our Comment 4, there is a role to be played by "Lagrange Multipliers". There is one such multiplier λ_i for each constraint and the values of the λ_i are determined as part of the problem solution.

Our main focus in this document is the use of Lagrange Multipliers to find candidates $\mathbf{r}_i$ for which a scalar function f(**r**) is stationary, df = 0, subject to constraints. This is application of Lagrange multipliers is distinct from that outlined in Comment 4 where there is no obvious function f(**r**) which we are trying make stationary.

However, by extending the notation of stationarity one can first define a certain function of the possible solution trajectories $q_k(t)$ for a problem. That function is called "the action" S which is the time integral of the Lagrangian $L(q_k, \dot{q}_k, t)$. The Lagrange equations of motion (including constraints) can be generated by setting $\delta S = 0$, a fact known as **Hamilton's Principle** (Goldstein Chapter 2 or the web). This notion is also used in the path integral formulations of quantum mechanics and quantum field theory. But these action applications are really not the same as the relatively simple problem addressed in this document concerning finding extrema of some scalar function f(**r**) .

A constraint of the form $a(q_1,q_2....q_N, t) = 0$ is referred to as a **holonomic** constraint, and our present document deals only with such constraints.

We shall now outline another mechanics application of Lagrange multipliers, again unrelated to making some obvious function f(**r**) be stationary. The presentation below follows Section IV of a 2003 paper by Ray and Shamanna (R&S). Usually one wants to remove constraint forces from an analysis, but here they are instead the focus of the approach. In any event, this is just another example of the use of Lagrange Multipliers.

Note: In this appendix subscripts are just labels, they do not indicate differentiation as we used elsewhere in this document. For example, f_i does not mean $\partial f/\partial x_i$, it just refers to the i^{th} constraint function. Similarly $\mathbf{r}_k$ is the 3D coordinate of the k^{th} particle.

An application of Lagrange Multipliers

Consider a system of M particles each having a 3D coordinate $\mathbf{r_k}$, so there are N = 3M scalar variables. Assume there is a set of C holonomic constraints $f_i(\mathbf{r_1}, \mathbf{r_2},,,,\mathbf{r_M}, t) = 0$ for i = 1,2...C. These C constraint surfaces each of dimension 3M-1 in E^{3M} intersect in some surface of dimension 3M-C in E^{3M} which then restricts the positions of the particles, as outlined in Appendix C: only certain $\{\mathbf{r_1}, \mathbf{r_2}...\mathbf{r_M}\}$ are allowed in any solution of the problem.

For example, suppose there are M=3 particles constrained to lie at the vertices of an equilateral triangle of edge 3, and in addition particle 1 is only allowed to be on a sphere of radius 5. The four constraints then are $| \mathbf{r_1} - \mathbf{r_2} | - 3 = 0$, $| \mathbf{r_2} - \mathbf{r_3} | - 3 = 0$, $| \mathbf{r_3} - \mathbf{r_1} | - 3 = 0$, and $(\mathbf{r_1})_1^2 + (\mathbf{r_1})_2^2 + (\mathbf{r_1})_3^2 - 5^2 = 0$. Each such constraint fits the holonomic template form $f_i(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, t) = 0$. Each constraint is a surface of dimension 8 in E^9 (there are many extrusions as discussed in Appendix C), and the intersection surface has dimension 5 in E^9. One could count these 5 degrees of freedom as follows: location of particle 1 on the sphere (2) [perhaps θ.φ] and orientation of the triangle in space (3) [three Euler angles].

In Newtonian mechanics, the equation of motion of a particle may be written $\mathbf{F_k} + \mathbf{G_k} = m_k\mathbf{a_k}$ where $\mathbf{G_k}$ is the total force of the constraints acting on particle k and $\mathbf{F_k}$ is the total of all *other* forces acting on this particle. Parameter m_k is the particle's mass and $\mathbf{a_k}$ is its vector acceleration.

A simple example of a holonomic constraint for one particle (particle 1) has that particle sliding down a frictionless curvy water-slide type surface (but no water) in the presence of gravity. In that case, $\mathbf{F_1}$ is the force of gravity acting on the particle, while $\mathbf{G_1}$ is the force of the slide surface acting on the particle. This force is normal to the instantaneous path of the particle so then $\mathbf{G_1} \bullet d\mathbf{r_1} = 0$, which can also be written in terms of the particle's velocity $\mathbf{G_1} \bullet \mathbf{v_1} = 0$.

Now differentiate the C constraint equations $f_i(\mathbf{r_1}, \mathbf{r_2},,,,\mathbf{r_M}, t) = 0$ to get

$$df_i = 0 = \Sigma_{k=1}^{M} \nabla^{(k)} f_i \bullet d\mathbf{r_k} + (\partial_t f_i)\, dt \qquad i = 1,2...C \qquad (D.1)$$

where $\nabla^{(k)} \equiv (\partial/\partial[(\mathbf{r_k})_1], \partial/\partial[(\mathbf{r_k})_2], \partial/\partial[(\mathbf{r_k})_3])$ is the 3D gradient with respect to the vector coordinate $\mathbf{r_k}$. Let $\{d\mathbf{r_k}\}$ for k = 1,,M be a set of differential displacements which satisfy equations (D.1), and call this a set of **allowed displacements**. Let $\{d\mathbf{r'_k}\}$ be some other set of allowed displacements so that

$$df_i = 0 = \Sigma_{k=1}^{M} \nabla^{(k)} f_i \bullet d\mathbf{r'_k} + (\partial_t f_i)\, dt \qquad i = 1,2...C \ . \qquad (D.2)$$

Now define $\delta\mathbf{r_k} \equiv d\mathbf{r_k} - d\mathbf{r'_k}$ to obtain a set $\{\delta\mathbf{r_k}\}$ of so-called **virtual displacements**. Such a set is the difference of any two different sets of allowed displacements. Then subtracting the above two equations gives

$$0 = \Sigma_{k=1}^{M} \nabla^{(k)} f_i \bullet \delta\mathbf{r_k} \qquad i = 1,2...C \ . \qquad (D.3)$$

This equation is true only as a sum, the individual terms in general do not vanish. Now multiply each of the C sums in (D.3) by a separate (and arbitrary) Lagrange Multiplier λ_i and add the sums together to get

$$0 = \Sigma_{i=1}^{C} \lambda_i \left[\Sigma_{k=1}^{M} \nabla^{(k)} f_i \bullet \delta\mathbf{r_k}\right]$$

or

$$0 = \Sigma_{k=1}^{M} \left[\Sigma_{i=1}^{C} \lambda_i \nabla^{(k)} f_i\right] \bullet \delta\mathbf{r_k} \ . \qquad (D.4)$$

This method of defining the meaning of a virtual displacement set follows the prescription (5) of R&S.

Assumption that constraints do no virtual work

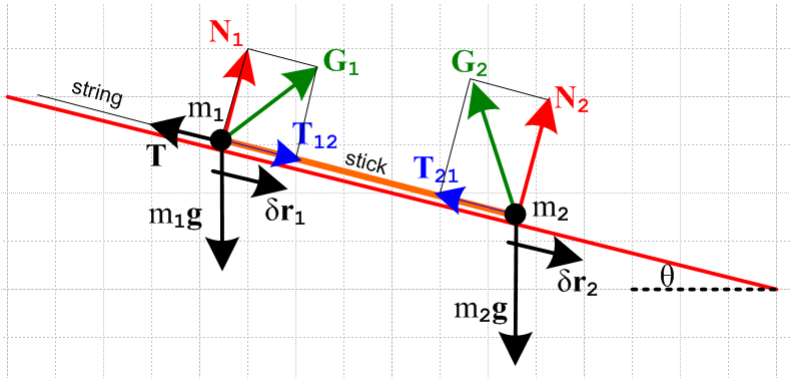
Next, *assume* that

$$0 = \Sigma_{k=1}^{M} \mathbf{G}_k \bullet \delta\mathbf{r}_k \,. \qquad \text{(D.5)}$$

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 the constraints 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 $\mathbf{G}_1$ 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 $\mathbf{G}_1$ 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 $\mathbf{G}_1 \bullet \delta\mathbf{r}_1 = 0$ and $\mathbf{G}_2 \bullet \delta\mathbf{r}_2 = 0$ so the terms in (D.5) 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.6)

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.6) we have drawn the two normal forces $\mathbf{N}_i$ and the two total constraint forces $\mathbf{G}_i$. It is no longer true that $\mathbf{G}_1 \bullet \delta\mathbf{r}_1 = 0$ and $\mathbf{G}_2 \bullet \delta\mathbf{r}_2 = 0$ as the picture shows. However, it is true that $\Sigma_{k=1}^{2} \mathbf{G}_k \bullet \delta\mathbf{r}_k = [\, \Sigma_{k=1}^{2} \mathbf{G}_k \,] \bullet \delta\mathbf{r} = 0$ because $\mathbf{G}_1$ and $\mathbf{G}_2$ have equal and opposite components along the ramp, so equation (D.5) 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 $\Sigma_{k=1}^{2} \mathbf{G}_k \bullet d\mathbf{r}_k \neq 0$. However, any virtual displacement set would still have $\delta\mathbf{r}_1 = \delta\mathbf{r}_2$ along the ramp and then $\Sigma_{k=1}^{2} \mathbf{G}_k \bullet \delta\mathbf{r}_k = 0$. 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.5) is still valid.

Other examples are given in R&S 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.5).

Resume Lagrange Mulitpliers application

Now subtract equation (D.4) from (D.5) to get,

$$0 = \Sigma_{k=1}^{M} [\, \mathbf{G}_k - \Sigma_{i=1}^{C} \lambda_i \ \nabla^{(k)} f_i] \bullet \delta \mathbf{r}_k \,. \qquad \text{(D.7)}$$

We know that the M vectors in the set $\{\delta \mathbf{r}_k\}$ are not linearly independent because they must satisfy the equations (D.3), so we cannot claim from (D.7) that $[\, \mathbf{G}_k - \Sigma_{i=1}^{C} \lambda_i \ \nabla^{(k)} f_i] = 0$ for each k. But now write out equation (D.7) in full detail, showing all vector components,

$$0 = \Sigma_{k=1}^{M} \ \Sigma_{j=1}^{3} [\, (\mathbf{G}_k)_j - \Sigma_{i=1}^{C} \lambda_i \ \frac{\partial f_i}{\partial (\mathbf{r}_k)_j}] \, (\delta \mathbf{r}_k)_j \,. \qquad \text{(D.8)}$$

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

$$\begin{array}{lccccccccccc} (k,j) & = & (1,1), & (1,2), & (1,3), & (2,1), & (2,2), & (2,3), & \ldots\ldots\ldots & (M,1), & (M,2), & (M,3) \\ n & = & 1, & 2, & 3, & 4, & 5, & 6\,, & & 3M\text{-}2, & 3M\text{-}1, & 3M \end{array} \qquad \text{(D.9)}$$

so that n = 3(k-1)+j. Then define

$$G_n \equiv (\mathbf{G}_k)_j \qquad x_n \equiv (\mathbf{r}_k)_j \,. \qquad // \ n = 3(k-1)+j \qquad \text{(D.10)}$$

Equation (D.8) can then be expressed as

$$0 = \Sigma_{n=1}^{3M} [\, G_n - \Sigma_{i=1}^{C} \lambda_i \frac{\partial f_i}{\partial x_n}] \, \delta x_n \,. \qquad \text{(D.11)}$$

Once again, we cannot set the square brackets to 0 because the 3M coordinates δx_n are not linearly independent.

]

Using the same notation we can write (D.3) as

$$0 = \ \Sigma_{k=1}^{M} \Sigma_{j=1}^{3} \ \frac{\partial f_i}{\partial (\mathbf{r}_k)_j} (\delta \mathbf{r}_k)_j \qquad\qquad i = 1,2...C \,.$$

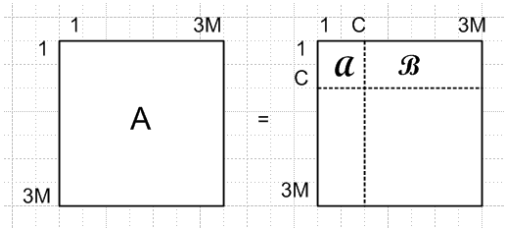
or

$$0 = \ \Sigma_{n=1}^{3M} \frac{\partial f_i}{\partial x_n} \delta x_n \qquad\qquad i = 1,2...C \qquad \text{(D.12)}$$

It will be convenient to define the following C x (3M) matrix A

$$A_{in} = \frac{\partial f_i}{\partial x_n} \qquad i = 1,2...C \qquad n = 1,2...3M \tag{D.13}$$

We then define two submatrices $\mathcal{A}$ and $\mathcal{B}$ of matrix A as follows,



(D.14)

so that C x C matrix $\mathcal{A}$ is the upper left corner of A and C x 3M matrix $\mathcal{B}$ is the rest of the top strip. In terms of matrix A equations ** and ** become

$$0 = \Sigma_{n=1}^{3M} [\, G_n - \Sigma_{i=1}^{C} \lambda_i \, A_{in} \,] \, \delta x_n \, . \tag{D.15}$$

$$0 = \Sigma_{n=1}^{3M} A_{in} \delta x_n \qquad\qquad i = 1,2...C \tag{D.16}$$

For simplicity, suppose it *happens* that the last 3M-C of the δx_n can be taken to be the independent variables. We can then write (D.15) as

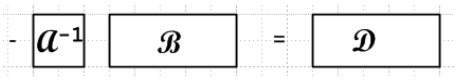
$$\begin{aligned}
&\Sigma_{n=1}^{C} A_{1n} \delta x_n \;\; = \; - \; \Sigma_{n=C+1}^{3M} A_{1n} \delta x_n \\
&\Sigma_{n=1}^{C} A_{2n} \delta x_n \;\; = \; - \; \Sigma_{n=C+1}^{3M} A_{2n} \delta x_n \\
&..... \\
&\Sigma_{n=1}^{C} A_{Cn} \delta x_n \;\; = \; - \; \Sigma_{n=C+1}^{3M} A_{Cn} \delta x_n
\end{aligned} \tag{D.17}$$

where the independent variables appear on the right side and the dependent ones on the left. Using the above defined matrices $\mathcal{A}$ and $\mathcal{B}$ we write the above equations in matrix form as

$$\mathcal{A} \begin{pmatrix} \delta x_1 \\ \delta x_2 \\ ... \\ \delta x_C \end{pmatrix} = \begin{pmatrix} - \; \Sigma_{n=C+1}^{3M} A_{1n} \delta x_n \\ - \; \Sigma_{n=C+1}^{3M} A_{2n} \delta x_n \\ ... \\ - \; \Sigma_{n=C+1}^{3M} A_{Cn} \delta x_n \end{pmatrix} = - \, \mathcal{B} \begin{pmatrix} \delta x_{C+1} \\ \delta x_{C+2} \\ ... \\ \delta x_{3M} \end{pmatrix} \tag{D.18}$$

Assuming $\det(\mathcal{A}) \neq 0$ we solve, introducing another C x 3M matrix $\mathcal{D}$, to get

$$\begin{pmatrix} \delta x_1 \\ \delta x_2 \\ ... \\ \delta x_C \end{pmatrix} = - \, \mathcal{A}^{-1} \mathcal{B} \begin{pmatrix} \delta x_{C+1} \\ \delta x_{C+2} \\ ... \\ \delta x_{3M} \end{pmatrix} \equiv \mathcal{M} \begin{pmatrix} \delta x_{C+1} \\ \delta x_{C+2} \\ ... \\ \delta x_{3M} \end{pmatrix} \qquad \mathcal{D} \equiv - \, \mathcal{A}^{-1} \mathcal{B} \tag{D.19}$$



so that

$$\delta x_i = \Sigma_{n=C+1}^{3M} D_{ij} \delta x_j \qquad i = 1,2...C \qquad (D.20)$$

so that each dependent δx_i on the left is some linear combination of the independent δx_n on the right, which we abbreviate by writing the above as

$$\delta x_i = \Sigma_{n=C+1}^{3M} B_{in} \delta x_n$$

If we regard the independent variables as "known" and the dependent ones as "unknown", we can solve the above equation as a Cramer's Rule problem to get expressions for the dependent variables

$$\delta x_i =$$

$$\delta x_j = - \left(\frac{\partial f_i}{\partial x_j}\right)^{-1} \Sigma_{n \neq j} \frac{\partial f_i}{\partial x_n} \delta x_n \qquad j = 1,2...C \text{ and } i = 1,2...C \qquad (D.13)$$

or

$$\delta x_j = \Sigma_{n \neq j} \left\{ - \left(\frac{\partial f_i}{\partial x_j}\right)^{-1} \frac{\partial f_i}{\partial x_n} \right\} \delta x_n = \Sigma_{n \neq j} A_{ijn} \delta x_n$$

Now we shall determine the Lagrange Multipliers. Consider this set of C equations,

$$\left[G_n - \Sigma_{i=1}^{C} \left(\frac{\partial f_i}{\partial x_n}\right) \lambda_i \right] = 0 \qquad n = 1,2...C \qquad (D.12a)$$

which involve the first C square-bracket quantities in (D.11).

Define the CxC matrix A by $A_{ni} \equiv \frac{\partial f_i}{\partial x_n}$ so the above may be written as a matrix equation,

$$\begin{pmatrix} G_1 \\ G_2 \\ \dots \\ G_C \end{pmatrix} = A \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_C \end{pmatrix} \qquad \Rightarrow \qquad \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_C \end{pmatrix} = A^{-1} \begin{pmatrix} G_1 \\ G_2 \\ \dots \\ G_C \end{pmatrix} \equiv B \begin{pmatrix} G_1 \\ G_2 \\ \dots \\ G_C \end{pmatrix} . \qquad \text{(D.12b)}$$

Assuming detA ≠ 0 we have found solution values for the C Lagrange Multipliers λ_i which are linear functions of the lowest C of the G_i

$$\lambda_i = \Sigma_{j=1}^{C} B_{ij} G_j = \lambda_i(G_1, G_2 \dots G_C)$$

Install these λ_i into (D.12a) to get

$$[\, G_n - \Sigma_{i=1}^{C} \left(\frac{\partial f_i}{\partial x_n}\right) \Sigma_{j=1}^{C} B_{ij} G_j \,] = 0 \qquad n = 1,2 \dots C \qquad (**)$$

or

$$[\, G_n - \Sigma_{j=1}^{C} \{\Sigma_{i=1}^{C} \left(\frac{\partial f_i}{\partial x_n}\right) B_{ij}\} G_j \,] = 0 \quad n = 1,2 \dots C \qquad (**)$$

or

$$[\, G_n - \Sigma_{j=1}^{C} Q_{nj} G_j \,] = 0 \qquad n = 1,2 \dots C \qquad Q_{nj} \equiv \Sigma_{i=1}^{C} \left(\frac{\partial f_i}{\partial x_n}\right) B_{ij} \qquad (**)$$

Suppose C = 2. The above equations are then

$$G_1 - [\, Q_{11} G_1 + Q_{12} G_2 \,] = 0$$
$$G_2 - [\, Q_{21} G_1 + Q_{22} G_2 \,] = 0$$

or

$$(1-Q_{11}) G_1 - Q_{12} G_2 = 0$$
$$-Q_{21} G_1 - (1-Q_{22}) G_2 = 0$$

In general, the solution of these two equations is $G_1 = 0$ and $G_2 = 0$, so then $\lambda_i = 0$ and all $G_i = 0$.

I have assumed a special case where the first C of the δx_i are dependent and the last 3M-c are independent. Seems to me that could happen. If it does, then there must be no constraint forces!

Since the first C terms in (D.11) now vanish, that equation becomes,

$$0 = \Sigma_{n=C+1}^{3M} [\, G_n - \Sigma_{i=1}^{C} \lambda_i \frac{\partial f_i}{\partial x_n} \,] \, \delta x_n \ . \qquad \text{(D.13)}$$

But the 3M-C variables δx_n appearing in (D.13) *are* independent, so each [...] in (D.13) must vanish. We then end up with *all* the coefficients of δx_n vanishing,

$$[\, G_n - \Sigma_{i=1}^{C} \lambda_i \frac{\partial f_i}{\partial x_n} \,] = 0 \qquad\qquad n = 1,2....3M$$

or

$$G_n = \Sigma_{i=1}^{C} \lambda_i(G_1, ...G_C) \frac{\partial f_i}{\partial x_n} \qquad\qquad n = 1,2....3M \ . \qquad\qquad (D.14)$$

or

$$G_n = \Sigma_{i=1}^{C} [\Sigma_{j=1}^{C} B_{ij} G_j] \frac{\partial f_i}{\partial x_n} \qquad\qquad n = 1,2....3M \ . \qquad\qquad (D.14)$$

$$= \Sigma_{j=1}^{C} \{ \Sigma_{i=1}^{C} B_{ij} \frac{\partial f_i}{\partial x_n} \} G_j$$

$$= \Sigma_{j=1}^{C} Q_{nj} G_j$$

Suppose C = 2, and M = a large number. Then we have

$$G_1 = \Sigma_{j=1}^{C} Q_{1j} G_j = Q_{11}G_1 + Q_{12}G_2$$
$$G_2 = \Sigma_{j=1}^{C} Q_{2j} G_j = Q_{21}G_1 + Q_{22}G_2$$
$$G_3 = \Sigma_{j=1}^{C} Q_{3j} G_j = Q_{31}G_1 + Q_{32}G_2$$
$$G_4 = \Sigma_{j=1}^{C} Q_{4j} G_j = Q_{41}G_1 + Q_{42}G_2$$
$$... \qquad = \text{etc}$$

The first two equations

The first two equations

or

(D.14) represents 3M easily solvable linear equations in the 3M unknown constraint forces G_n.

Something wrong. Only solution would be all $G_i = 0$!!

In this manner we have managed to determine all the constraint forces $G_n \equiv (\mathbf{G_k})_j$ in the problem. If the constraints f_i have explicit time dependence, then $F_{ni} \equiv \frac{\partial f_i}{\partial x_n}$ will be time dependent and so will the $\mathbf{G_k}$. From Newton's Law $\mathbf{F_k} + \mathbf{G_k} = m_k\mathbf{a_k}$ where $\mathbf{F_k}$ are any "applied" forces, one then knows all the $\mathbf{a_k}(t)$ which can then be time-integrated to obtain $\mathbf{v_k}(t)$ and then $\mathbf{r_k}(t)$.

To summarize, we have used an application of Lagrange Multipliers as an aid to solving a mechanics problem involving N particles with C holonomic constraints. As already noted, the above presentation is based on Section IV of R&S. That paper comments (at great length) on the often vague term "virtual displacement" as it appears in Goldstein and other standard texts, and we have followed their suggestion for defining this term as a difference of "allowed displacements" which then removes the dt time term in (D.1).

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.5) that $\Sigma_k\mathbf{G_k} \bullet \delta\mathbf{r_k} = 0$ as the statement $\delta W = 0$ where $\delta W = \Sigma_k\mathbf{G_k} \bullet \delta\mathbf{r_k}$ is the differential "virtual work" done by the constraint forces during a differential virtual displacement. This equation can then be written $0 = \delta W = \Sigma_k[m_k\mathbf{a_k} - \mathbf{F_k}] \bullet \delta\mathbf{r_k}$ in which form it is known as **D'Alembert's Principle** (1743, Goldstein (1-42)). The analog of f(**r**) is then presumably some action-like integral W of the virtual work δW.

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. For example, in the triangle example noted above one could use 5 angles to describe the state ("configuration") of the constrained system. One can replace the 3M non-independent variables $\{x_n\}$ in (D.10) with a set of 3M-C independent variables $\{q_n\}$ where $x_n = x_n(q_1, q_2....q_{3M-C}, t)$ for n=1 to 3M, 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 $\delta x_n = \Sigma_{j=1}^{3M-C}(\partial x_n/\partial q_j)\,\delta q_j$ for n = 1 to 3M, or alternatively, $\delta\mathbf{r_k} = \Sigma_{j=1}^{3M-C}(\partial\mathbf{r_k}/\partial q_j)\,\delta q_j$ as in Goldstein (1-44).

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 R&S and Goldstein p 17 (and many other places) and proceeds like this:

1. Start with $\Sigma_k\mathbf{G_k} \bullet \delta\mathbf{r_k} = 0$ in (D.5) which we write in the form of D'Alembert's Principle already mentioned above,

$$0 = \Sigma_{k=1}^{M}\,[m_k\mathbf{a_k} - \mathbf{F_k}] \bullet \delta\mathbf{r_k} \qquad \text{(D.15)}$$

where $\mathbf{F_k}$ is the total non-constraint force acting on particle k (sometimes called the "applied" force).

2. Make the replacement just noted above to convert to 3M-C independent generalized coordinates,

$$\delta \mathbf{r}_k = \Sigma_{j=1}^{3M-C} \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right) \delta q_j \quad \text{where} \quad \mathbf{r}_k = \mathbf{r}_k(q_1, q_2 \ldots q_{3M-C}, t) \tag{D.16}$$

3. Insert this into (D.15) to get

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

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

$$= \Sigma_j \left\{ \Sigma_k m_k \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_j} - Q_j \right\} \delta q_j \qquad Q_j \equiv \Sigma_{k=1}^{M} \mathbf{F}_k \bullet \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right) \tag{D.17}$$

where Q_j is the "generalized force" associated with the generalized coordinate q_j .

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

$$\Sigma_k m_k \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} \tag{D.34}$$

where

$$T \equiv (1/2)\Sigma_k m_k \left(\mathbf{v}_k \bullet \mathbf{v}_k\right) \tag{D.18}$$

is the total kinetic energy of the M particles. Then (D.17) reads,

$$0 = \Sigma_j \left\{ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - Q_j \right\} \delta q_j \, . \tag{D.19}$$

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

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \qquad j = 1,2....3M\text{-}C \; . \qquad \text{// as in Goldstein (1-50)} \tag{D.20}$$

If the applied forces can be derived from a potential $\mathcal{V}$ such that

$$\mathbf{F}_k = -\nabla^{(k)} \mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_M) \tag{D.21}$$

then from (D.17),

$$Q_j \equiv \Sigma_k \mathbf{F}_k \bullet \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right) = -\Sigma_k \left(\frac{\partial \mathcal{V}}{\partial \mathbf{r}_k}\right) \bullet \left(\frac{\partial \mathbf{r}_k}{\partial q_j}\right) = -\frac{\partial V}{\partial q_j} \tag{D.22}$$

where $V(q_i) = \mathcal{V}(\mathbf{r}_k(q_i))$ 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 $Q_j = -\frac{\partial V}{\partial q_j}$ in (D.20) gives

$$\frac{d}{dt}\left(\frac{\partial(T-V)}{\partial \dot{q}_j}\right) - \frac{\partial(T-V)}{\partial q_j} = 0 \ . \tag{D.23}$$

The last step is to define the classical **Lagrangian** $L \equiv T - V$ to obtain

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad j = 1,2....3M-C \ . \qquad \text{// as in Goldstein (1-53)} \tag{D.24}$$

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

Derivation of the result (D.34) used above

Define the 3M-C independent generalized coordinates q_i according to

$$\mathbf{r}_k = \mathbf{r}_k(q_1, q_2....q_{3M-C}, t) \equiv \mathbf{r}_k(q_i, t) \ . \tag{D.25}$$

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

$$\mathbf{v}_k = \dot{\mathbf{r}}_k = \frac{d\mathbf{r}_k}{dt} = \Sigma_j \frac{\partial \mathbf{r}_k}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial \mathbf{r}_k}{\partial t} = \Sigma_j \frac{\partial \mathbf{r}_k}{\partial q_j}\dot{q}_j + \frac{\partial \mathbf{r}_k}{\partial t} = \mathbf{v}_k(q_i, \dot{q}_i, t) \ . \tag{D.26}$$

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

$$\frac{\partial \mathbf{v}_k}{\partial q_i} = \Sigma_j \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial t} \tag{D.27}$$

$$\frac{\partial \mathbf{v}_k}{\partial \dot{q}_i} = \frac{\partial \mathbf{r}_k}{\partial q_i} \ . \tag{D.28}$$

Next, define total kinetic energy of all k particles,

$$T \equiv (1/2)\,\Sigma_k\, m_k\,(\mathbf{v}_k \bullet \mathbf{v}_k) \tag{D.29}$$

so that

$$\frac{\partial T}{\partial \dot{q}_i} = \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{\partial \mathbf{v}_k}{\partial \dot{q}_i} = \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_i} \ . \qquad \text{// (D.28)} \tag{D.30}$$

Applying d/dt to the above then gives,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = \Sigma_k\, m_k\, \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_i} + \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{d}{dt}\left(\frac{\partial \mathbf{r}_k}{\partial q_i}\right) \quad . \tag{D.31}$$

Computing the total time derivative in the second term gives,

$$\frac{d}{dt}\left(\frac{\partial \mathbf{r}_k}{\partial q_i}\right) = \Sigma_j \frac{\partial^2 \mathbf{r}_k}{\partial q_j \partial q_i}\dot{q}_j + \frac{\partial^2 \mathbf{r}_k}{\partial t\, \partial q_i} \quad . \tag{D.32}$$

Meanwhile,

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

$$= \Sigma_k\, m_k\, \mathbf{v}_k \bullet \left[\, \Sigma_j \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial^2 \mathbf{r}_k}{\partial q_i \partial t}\,\right] \qquad \text{// (D.27)}$$

$$= \Sigma_k\, m_k\, \mathbf{v}_k \bullet \frac{d}{dt}\left(\frac{\partial \mathbf{r}_k}{\partial q_i}\right) \quad . \qquad \text{// (D.32)} \tag{D.33}$$

Then subtracting (D.33) from (D.31) we get our desired result,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} = \Sigma_k\, m_k\, \mathbf{a}_k \bullet \frac{\partial \mathbf{r}_k}{\partial q_i} \quad . \tag{D.34}$$