# Physics 106a, Caltech — 15 October, 2019

# Lecture 5: Hamilton's Principle with Constraints

We have been avoiding forces of constraint, because in many cases they are uninteresting, and the constraints can be built in to carefully chosen generalized coordinates that can be varied independently.

Motion along the constraint directions are zero so we simply leave those coordinates out. If we start with M particles in 3 dimensions, or $3M$ degrees of freedom, but can make use of $N_c$ constraints, we can reduce the problem to $N_{dof} = 3M - N_c$ degrees of freedom and thus $N_{dof}$ equations of motion. This is "method 1" discussed below.

But sometimes, it's not possible (or desirable) to do this:

- The constraints may be too complicated to allow for "carefully chosen generalized coordinates."

- Sometimes, it's *important* to know the forces of constraints; for example, we may need to know the tension on a pendulum, to choose a strong enough material. Or, we need to know the forces of constraint that keep a roller coaster car on its track, to ensure that the track is strong enough to not break, and to ensure that the car doesn't fly off the track.

In these cases, choose coordinates that are easy to identify, even if they don't vary independently (because of the constraints); eg, the 3M Cartesian coordinates of M bodies in 3 dimensions.

To this one adds $N_c$ constraint equations, and, as we'll see, $N_c$ *Lagrange multipliers.* We will end up with $N_{dof} = 3M + N_c$ equations of motion. It will thus be more complicated, but we will be able to address the two concerns above.

### Examples

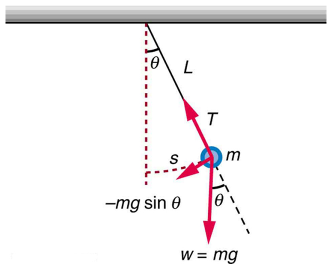
We'll illustrate the following ideas with four examples:

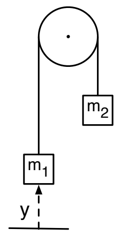
- The simple pendulum in 2-dimensions. Here, the constraint $\sqrt{x^2 + y^2} = \ell$ reduces the problem from two cartesian coordinates ($x$ and $y$) to one generalized coordinate $\phi$ satisfying the constraint. Or, if we need to know the tension on the pendulum string, we will use Lagrange multipliers.

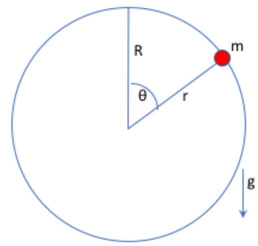
- The Atwood machine (two masses at the two ends of a massless rope slung over a massless, frictionless pulley) is similarly simple to solve by choosing the right generalized coordinate, but again, if we need to know the tension on the pendulum string, we will use Lagrange multipliers.

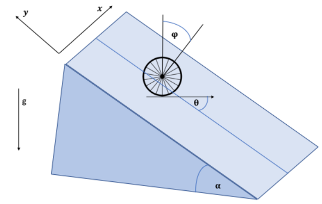
- A mass sliding on a circular wire (constraint: $r = R$), or a mass rolling or sliding on a bowling ball (constraint: $r \geq R$), or a mass being swung around on a string (constraint: $r \leq R$).

- The bicycle wheel on an inclined plane (Hand & Finch Fig 2.6). Here, the constraints are *non-holonomic*; they are constraints on the velocity, not position of the wheel, so they can't be integrated without first knowing the full solution. Here, we need to use Lagrange multipliers whether we're interested in the constraint forces or not.









## Holonomic constraints

First set the problem up in terms of $3M$ coordinates $q_k$ that completely specify the system *before* constraints are introduced. (These might be the Cartesian components of the $M$ position vectors $\mathbf{r}_i$ or might be some other choice of coordinates.) Now introduce $N_c$ holonomic constraints

$$G_j(q_1, q_2 \dots q_{3M}, t) = 0 \qquad j = 1 \dots N_c. \tag{1}$$

In Hamilton's principle we choose to look at path variations that are consistent with the constraints, so that

$$S = \int L dt \quad \Rightarrow \quad \delta S = \int \sum_{k=1}^{3M} \frac{\delta L}{\delta q_k} \delta q_k \, dt = 0 \qquad \text{for } \{\delta q_k\} \text{ consistent with constraints} \tag{2}$$

with $\delta L / \delta q_k$ the variational derivative

$$\frac{\delta L}{\delta q_k} \equiv \frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right). \tag{3}$$

We *cannot* put $\delta L / \delta q_k$ to zero, since the $\delta q_k$ are not independent.

There are two methods to do the constrained optimization.

### Method 1: Reduced number of coordinates

To do the constrained optimization we can find some reduced number $N = 3M - N_c$ of generalized coordinates $\{\bar{q}_k, k = 1 \dots N\}$ such that we can vary them independently and each variation is consistent with the constraints. These are the "generalized coordinates consistent with the constraints" we have used before. One choice might be a reduced set $N$ of the original $q_k$ with the other $N_c$ of the $\{q_k\}$ varying to maintain the constraints. Varying with respect to these coordinates gives

$$\delta S = \int \sum_{k=1}^{N} \frac{\delta L}{\delta \bar{q}_k} \delta \bar{q}_k \, dt = 0 \tag{4}$$

and since the $\delta \bar{q}_k$ are independent and may be chosen arbitrarily each $\delta L / \delta \bar{q}_k = 0$, giving the $N$ equations of motion for the constrained problem as before

$$\frac{\partial L}{\partial \bar{q}_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\bar{q}}_k}\right) = 0. \tag{5}$$

Also, since variations of the $\bar{q}_k$ are consistent with the constraints, the Lagrangian can be evaluated without including the constraint forces.

### Method 2: Lagrange multipliers

Alternatively we can do the constrained optimization using the *method of Lagrange multipliers.* In this approach we find the stationary value of a *modified* action $\bar{S}$ with the constraints added using Lagrange multipliers $\lambda_j(t)$, which may depend on $t$, but not on any of the coordinates $\{q_k\}$. If the $\lambda_j$'s are time-independent, this is called "static equilibrium".

$$\bar{S} = \int \left[ L + \sum_{j=1}^{N_c} \lambda_j(t) G_j \right] dt\,, \tag{6}$$

now treating all $q_k$ as effectively independent. Requiring $\bar{S}$ to be stationary,

$$\delta \bar{S} = \int_{t_i}^{t_f} \sum_{k=1}^{3M} \left[ \frac{\delta L}{\delta q_k} + \sum_{j=1}^{N_c} \lambda_j \frac{\partial G_j}{\partial q_k} \right] \delta q_k \, dt = 0; \tag{7}$$

then gives the $3M + N_c$ equations

$$\frac{\delta L}{\delta q_k} + \sum_{j=1}^{N_c} \lambda_j \frac{\partial G_j}{\partial q_k} = 0 \qquad k = 1 \ldots 3M \tag{8}$$

$$\frac{\partial \bar{S}}{\partial \lambda_j} = 0 \quad \Rightarrow \quad G_j = 0, \qquad j = 1 \ldots N_c \tag{9}$$

for the same number $3M + N_c$ of unknowns at each time: $\ddot{q}_k(t)$, $\lambda_j(t)$.

Why does this work? Note that $\bar{S} = S$ for paths satisfying the constraints ($G = 0$): making $S$ stationary for path variations satisfying the constraints is certainly the same as making $\bar{S}$ stationary for such variations. How about the Lagrange multiplier terms? Here's an outline of the argument, using the example of a pendulum in 2-dimensions.

In Cartesian coordinates, the 2D pendulum has Lagrangian, and constraint:

$$L = \frac{1}{2} m \left( \dot{x}^2 + \dot{y}^2 \right) - mgy, \quad \text{and} \quad G(x, y) = \sqrt{x^2 + y^2} - \ell = 0. \tag{10}$$

One constraint means we have one Lagrange multiplier, $\lambda$. Requiring $\bar{S}$ to be stationary,

$$\delta \bar{S} = \int_i^f \left[ \left( \frac{\delta L}{\delta x} + \lambda \frac{\partial G}{\partial x} \right) \delta x + \left( \frac{\delta L}{\delta y} + \lambda \frac{\partial G}{\partial y} \right) \delta y \right] dt = 0. \tag{11}$$

Because of the constraint, $\delta x$ and $\delta y$ cannot be varied independently, and thus the two terms in parentheses can't be set to zero separately. Or can they? The trick: we haven't yet specified $\lambda$. Let's *choose* it to be such that the first term in parentheses is equal to zero. We are left only with the term containing $\delta y$. Now we can assert that the path $y$ *can* be varied independently (so that $\delta x$ is constrained, but it no longer appears in the integral!). Thus the term multiplying $\delta y$ (which also contains the $\lambda$ that is now specified from the first term) must be equal to zero. Got it?

Continuing with this example, we have:

$$\delta x : \quad \frac{\delta L}{\delta x} = -m\ddot{x}; \quad \frac{\partial G}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} = \frac{x}{\ell} \tag{12}$$

$$\delta y : \quad \frac{\delta L}{\delta y} = -m\ddot{y} - mg; \quad \frac{\partial G}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}} = \frac{y}{\ell} \tag{13}$$

where we have used the constraint for the last steps in the above two lines. We thus have two equations of motion:

$$m\ddot{x} = \lambda \frac{x}{\ell} \tag{14}$$

$$m\ddot{y} = -mg + \lambda \frac{y}{\ell} \tag{15}$$

Now think about how Newton would describe this. The (massless) pendulum string is under tension $T$ (the force along the string, pulling up) to hold up the mass against gravity. (If the string is massive, the tension can vary along the string). That constraining force has components $F_x^c = -T\sin\theta = -T(x/\ell)$ and $F_y^c = -T\cos\theta = -T(y/\ell)$. We thus have exactly the same equations as above, with the identification $\lambda - -T$. Thus, we interpret $\lambda$ to be (proportional to) the constraining force. This (a) justifies the Lagrange multiplier method, and (b) gives us a method for identifying and quantifying the required constraining force.

Exercise: go through this argument with the Atwood machine, derive a formula for the tension on the massless string holding up the two masses.

You can refer to textbooks (eg. Hand and Finch §2.7) for more discussion of the method. Hand and Finch §2.6 also discusses the application of Lagrange multipliers to the simpler problem of the constrained min/maximization of *functions* of variables ("static equilibrium"; $\lambda_j(t) \to \lambda_j$), rather than *functionals* of functions — read these discussions if you find the use of Lagrange multipliers unclear.

Equation (8) has a geometrical interpretation that the gradient of $L$ is in the "plane" formed by the normal derivatives to the constraint surfaces, and so has zero component in directions of variation consistent with the constraints.

The Lagrange multipliers are related to the (generalized) components of the total constraint force

$$\mathcal{F}_k^{(c)} = \sum_j \lambda_j \frac{\partial G_j}{\partial q_k} \tag{16}$$

i.e. the second term on the left hand side of Eq. (8) is $\mathcal{F}_k^{(c)}$. I find Hand and Finch a little confusing in the discussion of this point, so here is my version.

Derive the generalized equations of motion (Golden rule #1) – see Lecture 4 – for the $3M$ coordinates $q_k$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} = \frac{\delta W}{\delta q_k} = \mathcal{F}_k^{(\mathrm{nc})} + \mathcal{F}_k^{(\mathrm{c})} \tag{17}$$

where we include *all* the forces in the virtual work, including the constraint forces, since variations of the $3M$ coordinates do not necessarily give virtual displacements carefully arranged to be "perpendicular" to the constraint forces. The non-constraint forces derive from the "external" potential $V$ that we know $\mathcal{F}_k^{(\mathrm{nc})} = -\partial V/\partial q_k$: these terms are transferred to the left hand side and form part of the conventional Lagrangian to give

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = \mathcal{F}_k^{(\mathrm{c})}. \tag{18}$$

Now compare with Eq. (8).

## Nonholonomic constraints

The inclusion of constraints in the Lagrange multiplier approach Eq. (8) involves only the *differential* form of the constraints, and so *nonintegrable differential nonholonomic constraints* can also be implemented. For $N_c$ constraints of the form

$$\sum_{k=1}^{3M} g_{jk}\delta q_k = 0 \qquad \text{equivalent to} \qquad \sum_{k=1}^{3M} g_{jk}\dot{q}_k = 0, \qquad j = 1 \dots N_c, \tag{19}$$

where the coefficients $g_{jk}$ may depend on $\{q_l\}$ and $t$, the constrained optimization is given by setting

$$\int \sum_{k=1}^{3M} \left( \frac{\delta L}{\delta q_k} + \sum_{j=1}^{N_c} \lambda_j(t) g_{jk} \right) \delta q_k dt = 0 \tag{20}$$

where again the second term in the brackets give zero total contribution to the expression for variations consistent with the constraints. As before, with the extra freedom introduced by the $\lambda_j$, the $3M$ values of $\delta q_k$ can be treated as effectively independent.

## Examples of the use of Lagrange multipliers

**1. Mass $m$ confined to vertical circle of radius $R$:** This might be a particle sliding on a ball in a vertical plane through the center, a mass twirling in a vertical plane on the end of a string, or a bead on a hoop, etc. Use polar coordinates $(r, \theta)$ with $\theta$ measured from the vertical. Then

$$T = \frac{1}{2} m (r^2 \dot{\theta}^2 + \dot{r}^2) \qquad \text{kinetic energy} \tag{21}$$

$$V = -mg(R - r\cos\theta) \qquad \text{potential energy } (V = 0 \text{ at top}) \tag{22}$$

$$G(r) = r - R = 0 \qquad \text{constraint} \tag{23}$$

Find the stationary value of the effective action $\bar{S}$ including the constraint with Lagrange multiplier $\lambda$

$$\bar{S} = \int \left[ \frac{1}{2} m (r^2 \dot{\theta}^2 + \dot{r}^2) + mg(R - r\cos\theta) + \lambda(t)(r - R) \right] \tag{24}$$

taking $\delta\theta, \delta r$ as independent

$$\frac{\delta \bar{S}}{\delta \theta} : \qquad mgr\sin\theta - \frac{d}{dt}(mr^2\dot{\theta}) = 0 \tag{25}$$

$$\frac{\delta \bar{S}}{\delta r} : \qquad mr\dot{\theta}^2 - mg\cos\theta + \lambda - \frac{d}{dt}(m\dot{r}) = 0 \tag{26}$$

These are to be solved together with the constraint

$$r - R = 0. \tag{27}$$

Equation (25) is the equation of motion for $\theta$, and Eq. (26) with the constraint ()$r$ is constant, $\dot{r} = 0$, and $\ddot{r} = 0$) gives $\lambda$:

$$\lambda = mg\cos\theta - mr\dot{\theta}^2 \tag{28}$$

which is indeed the radial constraint force $F_r^{(c)}$, since $r\dot{\theta}^2$ is the radial acceleration for the circular motion. (In this case $\partial G/\partial r = 1$, and so $\lambda$ is equal to the component of the constraint force. More generally the constraint force is $\lambda$ multiplied by the derivative of the constraint function.)

This constraining force is supplied by the material properties of the hoop, bowling ball, rope, or roller coaster. At least, we hope. For a mass sliding on a ball we require $F_r^{(c)} > 0$ (the ball can only push) and for a mass on a string we require $F_r^{(c)} < 0$ (the string can only pull), so this sets limits on when the constrained solution actually matches the physical solution (ie, tells us how strong our constraining material needs to be before it fails). For a bead on a wire $F_r$c) can have either sign, so the constrained solution is valid for all times. For a roller coaster, the engineer must ensure that the required constraining force, supplied by the track, is sufficient to keep the cart on the track and keep the track from collapsing.

2. **Nonholonomic constraints: vertical rotating wheel on a sloping plane:** See Hand and Finch §2.8 and Fig. 2.6 for the setup, although the $y$ coordinate should point up hill, not down. Note that the $xy$ plane is tilted, not horizontal. I'll do the case of a wheel with all the mass at the rim; Hand and Finch look at a disk, which changes the moment of inertia calculation. The wheel is assumed not to tilt in the motion. See the figure a few pages back.

   The rolling-without-slipping constraints are

$$\delta x - R\sin\theta\delta\phi = 0 \qquad \text{or} \qquad \dot{x} = R\sin\theta\dot{\phi} \tag{29}$$

$$\delta y - R\cos\theta\delta\phi = 0 \qquad \text{or} \qquad \dot{y} = R\cos\theta\dot{\phi} \tag{30}$$

   where the second pair are called *velocity constraints*. Then the kinetic energy is

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}mR^2\dot{\phi}^2 + \tfrac{1}{4}mR^2\dot{\theta}^2 \tag{31}$$

   If you substitute $\dot{x}$ and $\dot{y}$ using the velocity constraints, as H&F does, you get:

$$T = mR^2\dot{\phi}^2 + \tfrac{1}{4}mR^2\dot{\theta}^2 \tag{32}$$

   This actually works; but in general it will *not* work (see below); you should not make such substitutions in the Lagrangian, but only in the equations of motion obtained from the Lagrangian. The potential energy is

$$V = mgf(y) \qquad \text{with } f(y) = y\sin\alpha\ . \tag{33}$$

   The Lagrangian is

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}mR^2\dot{\phi}^2 + \tfrac{1}{4}mR^2\dot{\theta}^2 - mgf(y)\,. \tag{34}$$

   Finding the stationary point of the action for variations subject to the constraints introducing the Lagrange multipliers $\lambda_x, \lambda_y$ gives

$$\int \frac{\delta L}{\delta x}\delta x + \frac{\delta L}{\delta y}\delta y + \frac{\delta L}{\delta \phi}\delta\phi + \frac{\delta L}{\delta\theta}\delta\theta + \lambda_x\left(\delta x - R\sin\theta\delta\phi\right) + \lambda_y\left(\delta y - R\cos\theta\delta\phi\right) = 0. \tag{35}$$

   Collecting terms

$$\int\left(\frac{\delta L}{\delta x} + \lambda_x\right)\delta x + \left(\frac{\delta L}{\delta y} + \lambda_y\right)\delta y + \left(\frac{\delta L}{\delta\phi} - \lambda_x R\sin\theta - \lambda_y R\cos\theta\right)\delta\phi + \left(\frac{\delta L}{\delta\theta}\right)\delta\theta = 0 \tag{36}$$

Now $\delta x, \delta y, \delta\phi, \delta\theta$ can be treated as effectively independent so each () must be zero. This gives

$$-\frac{d}{dt}(m\dot{x}) + \lambda_x = 0 \tag{37}$$

$$-\frac{d}{dt}(m\dot{y}) - mgf'(y) + \lambda_y = 0 \tag{38}$$

$$-\frac{d}{dt}(mR^2\dot{\phi}) - \lambda_x R\sin\theta - \lambda_y R\cos\theta = 0 \tag{39}$$

$$-\frac{d}{dt}\left(\frac{1}{2}mR^2\dot{\theta}\right) = 0 \tag{40}$$

These are to be solved together with the constraint equations

$$\dot{x} = R\sin\theta\dot{\phi} \tag{41}$$

$$\dot{y} = R\cos\theta\dot{\phi} \tag{42}$$

Using Eqs. (37,38) to evaluate the Lagrange multipliers eliminating $\dot{x}, \dot{y}$ with Eqs. (41,42) gives

$$\lambda_x = \frac{d}{dt}(mR\sin\theta\dot{\phi}) \tag{43}$$

$$\lambda_y = \frac{d}{dt}(mR\cos\theta\dot{\phi}) + mgf'(y) \tag{44}$$

When these are substituted into Eq. (39) some miraculous cancellation gives

$$2mR^2\ddot{\phi} = mgf'(y)R\cos\theta \tag{45}$$

where the factor of 2 on the left hand side comes from the $x, y$ kinetic energy adding to the $\phi$ kinetic energy. This is to be solved with Eq. (40) or

$$mR^2\dot{\theta} = \text{constant} \tag{46}$$

which is the conservation of angular momentum about the axis normal to the plane ($\theta$ is ignorable). For a uniform slope $f(y) = y\sin\alpha$, $f'(y) = \sin\alpha$, so that $y$ drops out of these two equations, and the constraint equation is not needed. For a ramp of varying slope where $y$ is needed to evaluate the right hand side of Eq. (45), we would also have include the constraint equation

$$\dot{y} = R\cos\theta\dot{\phi} \tag{47}$$

in the list of equations to be solved together, i.e. Eqs. (45-47).

There is a subtlety in such calculations including nonintegrable differential constraints. You might be tempted, as were Hand and Finch, to use the constraint equations Eqs. (41,42) to simplify the kinetic energy to

$$T = mR^2\dot{\phi}^2 + \frac{1}{4}mR^2\dot{\theta}^2 \tag{48}$$

Using this, the Lagrangian would depend just on $\phi, \theta, y$ and their time derivatives, and we might plan to find the stationary value of the action varying these variables subject to the single constraint Eq. (42). This procedure — using the *velocity* constraints to express the Lagrangian in terms of fewer variables and using this to form the action — is in general *incorrect* for nonholonomic constraints (it is fine for the *coordinate* constraints of a holonomic system). You can find a discussion of this rather subtle point in the book *A Mathematical Introduction to Robot Manipulation* by Murray, Li, and Shastry, pp 274-6. This is available on Richard Murray's Caltech website and the direct link to a pdf file of the book is here. The development is quite mathematical, so you may find it an effort to follow. The book is also a good indication of the importance of the Lagrangian formulation of mechanics, and of nonholonomic constraints in robotics. For reasons I am not clear about, the reduction procedure (as used by Hand and Finch) gives the correct answer for the rolling wheel problem just considered.