

Graduate Classical Mechanics

Lecture 6: Generalized Coordinates, Constraints, and Virtual Work

Let the geometry remove the redundant coordinates

1 Why change coordinates?

The Euler–Lagrange equations from Lecture 5 were written in the form

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0.$$

We deliberately used the symbol q_j rather than x_j . The reason is that the variables used to describe a mechanical system need not be Cartesian coordinates. In many problems, Cartesian coordinates contain more information than the system is actually free to choose.

Consider a particle constrained to move on a circle of radius R :

$$x^2 + y^2 = R^2.$$

The pair (x, y) uses two numbers, but the constraint ties those numbers together. Only one independent number is needed. A natural choice is

$$q = \theta, \quad x = R \cos \theta, \quad y = R \sin \theta.$$

The particle has one degree of freedom, even though its Cartesian description uses two coordinates.

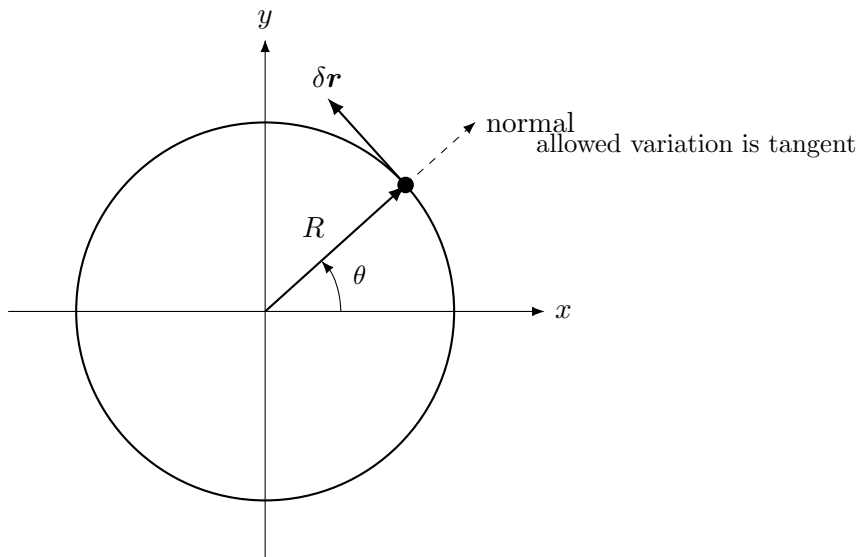


Figure 1: The constraint $x^2 + y^2 = R^2$ restricts the particle to a one-dimensional configuration space. The coordinate θ labels points on that space. At fixed time, an allowed virtual displacement is tangent to the circle.

The important lesson is not that polar coordinates are convenient. It is deeper:

A constraint often tells us that our original coordinates are redundant. Generalized coordinates label the genuinely independent directions in which the system can move.

2 Constraints and degrees of freedom

For N particles in three-dimensional space, the unconstrained Cartesian description contains $3N$ coordinates. Suppose the motion is subject to m independent *holonomic* constraints

$$f_\alpha(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0, \quad \alpha = 1, \dots, m.$$

Locally, these constraints remove m independent degrees of freedom, leaving

$$s = 3N - m$$

independent generalized coordinates

$$q_1, \dots, q_s.$$

This counting assumes that the constraints are independent and regular in the region of interest.

The Cartesian positions can then be written as functions of the generalized coordinates and time:

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \dots, q_s, t) = \mathbf{r}_i(\mathbf{q}, t).$$

This map contains the geometry of the constraints.

2.1 Configuration space

The set of all allowed configurations is called the *configuration space*. The generalized coordinates $q_1, \dots, q_s$ label points in this space.

For the particle on a circle, the allowed Cartesian points satisfy

$$x^2 + y^2 = R^2.$$

The configuration space is therefore one-dimensional, and θ is a coordinate on it.

For a double pendulum, two angles are enough. For a rigid body, three coordinates are required to specify orientation in addition to the coordinates locating the body. The useful coordinates depend on the geometry of the problem.

Some terminology

A holonomic constraint with no explicit time dependence,

$$f_\alpha(\mathbf{r}_1, \dots, \mathbf{r}_N) = 0,$$

is called *scleronomous*. If the constraint itself moves or changes with time,

$$f_\alpha(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0,$$

it is called *rheonomous*.

There are also *nonholonomic* constraints, often written as relations involving velocities that cannot be integrated into equations of the form $f_\alpha(\mathbf{r}, t) = 0$. We will not develop their full theory here. Our main focus is on holonomic constraints and coordinates adapted to them.

3 Velocities follow from the coordinate map

Once

$$\mathbf{r}_i = \mathbf{r}_i(\mathbf{q}, t)$$

is known, the velocity follows from the ordinary chain rule:

$$\dot{\mathbf{r}}_i = \sum_{j=1}^s \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \mathbf{r}_i}{\partial t}.$$

Thus the coordinate map immediately gives the sequence

$$\mathbf{q} \longrightarrow \mathbf{r}_i(\mathbf{q}, t) \longrightarrow \dot{\mathbf{r}}_i(\mathbf{q}, \dot{\mathbf{q}}, t) \longrightarrow T(\mathbf{q}, \dot{\mathbf{q}}, t).$$

For point particles,

$$T = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i^2.$$

Therefore kinetic energy becomes a function of the generalized coordinates and generalized velocities.

3.1 Example: particle on a circle

Write

$$\mathbf{r}(\theta) = R \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

Then

$$\dot{\mathbf{r}} = R\dot{\theta} \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix},$$

so

$$T = \frac{1}{2} m R^2 \dot{\theta}^2.$$

The generalized coordinate θ is dimensionless, and the coefficient multiplying $\dot{\theta}^2$ is mR^2 , a moment-of-inertia-like quantity. There is no reason for kinetic energy in generalized coordinates to have the form $\frac{1}{2} m \dot{q}^2$.

3.2 A first glimpse of the mass matrix

If the coordinate map has no explicit time dependence, then

$$\dot{\mathbf{r}}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j.$$

Substituting this into T gives

$$T = \frac{1}{2} \sum_{j,k} M_{jk}(\mathbf{q}) \dot{q}_j \dot{q}_k,$$

where

$$M_{jk}(\mathbf{q}) = \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k}.$$

In matrix notation,

$$T = \frac{1}{2} \dot{\mathbf{q}}^T M(\mathbf{q}) \dot{\mathbf{q}}.$$

The matrix $M(\mathbf{q})$ contains both masses and geometry. We will see this structure repeatedly in more complicated systems.

4 Actual displacement versus virtual displacement

This distinction is central.

For

$$\mathbf{r}_i = \mathbf{r}_i(\mathbf{q}, t),$$

the actual differential of the position is

$$d\mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} dq_j + \frac{\partial \mathbf{r}_i}{\partial t} dt.$$

This compares the system at two nearby times along an actual motion.

A *virtual displacement* does something different. It compares two nearby allowed configurations at the *same time*. By definition,

$$\delta t = 0.$$

Therefore

$$\delta \mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j.$$

So, in general,

$$d\mathbf{r}_i \neq \delta \mathbf{r}_i.$$

For a time-independent constraint the distinction can be easy to overlook, because both actual and virtual displacements lie tangent to the same fixed constraint surface. For a time-dependent constraint the difference is essential: the actual displacement also contains the motion of the constraint itself through $\partial \mathbf{r}_i / \partial t$.

4.1 Virtual displacements are tangent to the constraint surface

Suppose one particle satisfies

$$f(\mathbf{r}, t) = 0.$$

At fixed time, vary the position while remaining on the constraint surface:

$$\delta f = \nabla f \cdot \delta \mathbf{r} = 0.$$

Hence

$$\nabla f \cdot \delta \mathbf{r} = 0.$$

The gradient ∇f is normal to the surface, so $\delta \mathbf{r}$ is tangent to the allowed configuration space.

For the circle,

$$f(x, y) = x^2 + y^2 - R^2,$$

and

$$\delta f = 2x \delta x + 2y \delta y = 0.$$

Therefore

$$x \delta x + y \delta y = 0,$$

which says exactly that the allowed virtual displacement is perpendicular to the radial direction.

5 Constraint forces and ideal constraints

A particle constrained to a smooth rigid surface generally experiences a force that keeps it on the surface. Call this force $\mathbf{R}$, for *reaction* or constraint force.

For a smooth frictionless surface, $\mathbf{R}$ is normal to the surface while the allowed virtual displacement is tangent. Therefore

$$\mathbf{R} \cdot \delta \mathbf{r} = 0.$$

For many particles and many ideal constraints, the corresponding statement is

$$\sum_i \mathbf{R}_i \cdot \delta \mathbf{r}_i = 0.$$

This is the defining virtual-work property of an *ideal constraint*.

The meaning should be kept precise:

Constraint forces do not disappear because they are unimportant or because they are zero. They disappear from the virtual-work equations because, for ideal constraints, they do no work on the allowed virtual displacements.

This is one of the main advantages of coordinates adapted to the constraints: forces that are needed physically to enforce the geometry may not need to be solved for in order to obtain the equations of motion.

6 Generalized forces

Let $\mathbf{F}_i$ denote the applied forces, excluding the ideal constraint forces for the moment. Their virtual work is

$$\delta W = \sum_i \mathbf{F}_i \cdot \delta \mathbf{r}_i.$$

Use

$$\delta \mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j.$$

Then

$$\delta W = \sum_i \mathbf{F}_i \cdot \left(\sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \right).$$

Rearranging the sums,

$$\delta W = \sum_j \left[\sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right] \delta q_j.$$

This motivates the definition

$$Q_j \equiv \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j},$$

so that

$$\delta W = \sum_j Q_j \delta q_j.$$

The quantity Q_j is the *generalized force conjugate to q_j* .

A generalized force need not have the dimensions of an ordinary force. Since each term $Q_j \delta q_j$ has the dimensions of energy, the units of Q_j depend on the units of the corresponding generalized coordinate.

6.1 Example: polar coordinates

For a particle in the plane, use

$$(q_1, q_2) = (r, \theta), \quad \mathbf{r} = r \hat{\mathbf{e}}_r.$$

Then

$$\frac{\partial \mathbf{r}}{\partial r} = \hat{\mathbf{e}}_r,$$

so

$$Q_r = \mathbf{F} \cdot \hat{\mathbf{e}}_r = F_r.$$

But

$$\frac{\partial \mathbf{r}}{\partial \theta} = r \hat{\mathbf{e}}_\theta,$$

so

$$Q_\theta = \mathbf{F} \cdot r \hat{\mathbf{e}}_\theta = r F_\theta.$$

The generalized force conjugate to an angle is a torque. This is exactly what dimensional analysis predicts:

$$Q_\theta \delta \theta$$

must have units of work, and $\delta \theta$ is dimensionless.

7 Newton's law projected onto allowed motion

Now return to Newton's law for particle i :

$$\boxed{m_i \mathbf{a}_i = \mathbf{F}_i + \mathbf{R}_i.}$$

Rearrange:

$$\mathbf{F}_i + \mathbf{R}_i - m_i \mathbf{a}_i = 0.$$

Dot with an allowed virtual displacement and sum over all particles:

$$\sum_i (\mathbf{F}_i + \mathbf{R}_i - m_i \mathbf{a}_i) \cdot \delta \mathbf{r}_i = 0.$$

For ideal constraints,

$$\sum_i \mathbf{R}_i \cdot \delta \mathbf{r}_i = 0,$$

so

$$\boxed{\sum_i (\mathbf{F}_i - m_i \mathbf{a}_i) \cdot \delta \mathbf{r}_i = 0.}$$

This is *d'Alembert's principle*. It is Newton's law projected onto the allowed virtual displacements.

Substitute

$$\delta \mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j.$$

Then

$$0 = \sum_j \left[Q_j - \sum_i m_i \mathbf{a}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right] \delta q_j.$$

The remaining task is to rewrite the inertial term in terms of the kinetic energy.

8 The kinetic-energy identity

Let

$$T = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i^2.$$

From

$$\dot{\mathbf{r}}_i = \sum_k \frac{\partial \mathbf{r}_i}{\partial q_k} \dot{q}_k + \frac{\partial \mathbf{r}_i}{\partial t},$$

we have

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

Therefore

$$\frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

Take a total time derivative:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \mathbf{a}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} + \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_j} \right).$$

Now note that

$$\frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_j} \right) = \frac{\partial \dot{\mathbf{r}}_i}{\partial q_j}.$$

Hence

$$\frac{\partial T}{\partial q_j} = \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial q_j} = \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_j} \right).$$

Subtracting gives the key identity

$$\boxed{\sum_i m_i \mathbf{a}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j}.$$

This is the bridge from Newton's vector equations to equations written entirely in generalized coordinates.

9 Lagrange's equations with generalized forces

Return to d'Alembert's principle:

$$0 = \sum_j \left[Q_j - \sum_i m_i \mathbf{a}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right] \delta q_j.$$

Use the kinetic-energy identity:

$$0 = \sum_j \left[Q_j - \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} + \frac{\partial T}{\partial q_j} \right] \delta q_j.$$

The generalized coordinates are independent, so the virtual variations δq_j can be chosen independently. Therefore each coefficient must vanish:

$$\boxed{\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j, \quad j = 1, \dots, s.}$$

This is Lagrange's equation written in terms of kinetic energy and generalized applied forces.

Notice what is absent: the ideal constraint forces $\mathbf{R}_i$. They were removed by the geometry of the allowed virtual displacements.

10 Conservative forces and the return of $L = T - V$

If the applied forces come from a potential $V(\mathbf{q}, t)$, then

$$\delta W = -\delta V.$$

At fixed time,

$$\delta V = \sum_j \frac{\partial V}{\partial q_j} \delta q_j.$$

Therefore

$$\delta W = - \sum_j \frac{\partial V}{\partial q_j} \delta q_j.$$

Comparing with

$$\delta W = \sum_j Q_j \delta q_j,$$

we find

$$Q_j = - \frac{\partial V}{\partial q_j}.$$

Thus

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} + \frac{\partial V}{\partial q_j} = 0.$$

For an ordinary position-dependent potential,

$$\frac{\partial V}{\partial \dot{q}_j} = 0.$$

Define

$$L = T - V.$$

Then

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0.$$

We have now reached the same Euler–Lagrange equations from a second direction:

$$\boxed{\text{Newton} \longrightarrow \text{virtual work} \longrightarrow \text{generalized coordinates} \longrightarrow \text{Lagrange.}}$$

Lectures 4 and 5 reached them through stationary action. The agreement of these routes is one reason the Lagrangian formulation is so powerful.

If nonconservative applied forces are also present, it is often useful to write

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = Q_j^{(\text{nc})}.$$

11 Running example: the simple pendulum

A pendulum bob of mass m is attached to a massless rigid rod of length ℓ . In Cartesian coordinates,

$$x^2 + y^2 = \ell^2.$$

The tension in the rod is an unknown constraint force. We can avoid solving for it by choosing the angle θ as the generalized coordinate.

Take

$$\mathbf{r}(\theta) = \begin{pmatrix} \ell \sin \theta \\ -\ell \cos \theta \end{pmatrix}.$$

Then

$$\frac{\partial \mathbf{r}}{\partial \theta} = \ell \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix},$$

and the allowed virtual displacement is

$$\delta \mathbf{r} = \frac{\partial \mathbf{r}}{\partial \theta} \delta \theta.$$

It is tangent to the circular path.

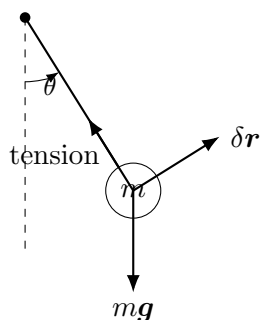


Figure 2: For the pendulum, the tension is radial while the allowed virtual displacement is tangent. Their dot product vanishes, so the tension does no virtual work.

The velocity is

$$\dot{\mathbf{r}} = \ell \dot{\theta} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix},$$

so

$$T = \frac{1}{2} m \ell^2 \dot{\theta}^2.$$

Now compute the generalized force due to gravity directly. With

$$\mathbf{F}_g = \begin{pmatrix} 0 \\ -mg \end{pmatrix},$$

we have

$$Q_\theta = \mathbf{F}_g \cdot \frac{\partial \mathbf{r}}{\partial \theta} = -mg\ell \sin \theta.$$

Thus the generalized-force form of Lagrange's equation gives

$$\frac{d}{dt} (m \ell^2 \dot{\theta}) = -mg\ell \sin \theta,$$

or

$$\ddot{\theta} + \frac{g}{\ell} \sin \theta = 0.$$

The tension never had to be calculated.

We can obtain the same result from a potential. Choosing zero gravitational potential at the bottom,

$$V = mg\ell(1 - \cos \theta).$$

Hence

$$L = \frac{1}{2}m\ell^2\dot{\theta}^2 - mg\ell(1 - \cos \theta).$$

Euler–Lagrange again gives

$$\ddot{\theta} + \frac{g}{\ell} \sin \theta = 0.$$

The key question is not merely “Can we derive the pendulum equation?” Newton’s laws can certainly do that. The structural question is:

Where did the tension go?

It is still physically present, but it is absent from the reduced equation because its virtual work vanishes on the allowed tangent displacement.

12 Generalized momentum and a first symmetry clue

For any generalized coordinate, define the conjugate generalized momentum

$$p_j \equiv \frac{\partial L}{\partial \dot{q}_j}.$$

The Euler–Lagrange equation can then be written

$$\dot{p}_j = \frac{\partial L}{\partial q_j}.$$

Now suppose a coordinate q_j does not appear explicitly in the Lagrangian:

$$\frac{\partial L}{\partial q_j} = 0.$$

Then

$$\dot{p}_j = 0,$$

so

$$p_j = \text{constant}.$$

Such a coordinate is called *cyclic* or *ignorable*.

This is our first clear glimpse of a much broader idea: when the physics does not change under a continuous change of some coordinate, a corresponding quantity is conserved. The next lecture will turn that observation into a systematic connection between symmetry and conservation.

13 Summary

The central chain of ideas in this lecture is

$$\begin{array}{c}
 \text{constraints reduce the allowed configurations} \\
 \Downarrow \\
 \mathbf{r}_i = \mathbf{r}_i(\mathbf{q}, t) \\
 \Downarrow \\
 \delta \mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \\
 \Downarrow \\
 \delta W = \sum_j Q_j \delta q_j \\
 \Downarrow \\
 \text{ideal constraint forces do no virtual work} \\
 \Downarrow \\
 \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j \\
 \Downarrow \\
 L = T - V \quad \text{for ordinary conservative forces.}
 \end{array}$$

The distinction to remember is

$$d\mathbf{r} \quad \text{versus} \quad \delta \mathbf{r}.$$

An actual displacement follows motion in time. A virtual displacement compares neighboring allowed configurations at fixed time.

The practical payoff is equally important:

Choose coordinates adapted to the constraints. Then the geometry is built into the coordinates, and ideal constraint forces can disappear from the equations without disappearing from the physics.