

# Chapter 1

## Newtonian Mechanics

### 1.1 Definitions

Classical Mechanics is the theory governing the motion of particles. The theory is unchanged since its discovery by Newton.

**Point particles** are idealized particles whose internal dimensions and properties can be neglected. The motion of a point particle can be completely described by the particle's position as a function of time in some coordinate system. The coordinates of the particle's position at time  $t$  is denoted by a vector  $\mathbf{r}(t)$ .

The particle's instantaneous velocity  $\mathbf{v}(t)$  and acceleration  $\mathbf{a}(t)$  in this coordinate system are defined by

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt}(t) \quad (1.1)$$

$$\mathbf{a}(t) = \frac{d^2\mathbf{r}}{dt^2}(t). \quad (1.2)$$

I use MKS units where distance is measured in meters, mass is measured in kilograms and time is measured in seconds.

### 1.2 Experimental observations

#### 1.2.1 Principle of Newtonian determinacy and Newton's second law

A fundamental observation (recognized by Newton and nicely discussed in Arnold's book) is that the particle's position,  $\mathbf{r}(t)$ , is a function that depends only on the time  $t$  and the coordinate and velocity of the particle at some earlier initial, time  $t = t_0$ . Arnold refers to this observation as the **principle of Newtonian determinacy**. This can be written mathematically as

$$\mathbf{r}(t) = \mathbf{R}(t, t_0, \mathbf{r}(t_0), \mathbf{v}(t_0)). \quad (1.3)$$

That  $\mathbf{r}(t)$  does *not* depend on higher order initial derivatives of the position with respect to time or the past history of the particle's position is a profound observation.

If I differentiate this equation twice with respect to  $t$  and set the initial time to the current time,  $t_0 = t$ , I get

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{a}(\mathbf{r}(t), \mathbf{v}(t), t) := \frac{\partial^2}{\partial t^2} \mathbf{R}(t, t, \mathbf{r}(t), \mathbf{v}(t)) \quad (1.4)$$

where the partial derivative in the term on the right acts only on the first  $t$  variable. Expressing the instantaneous velocity as the derivative of the particle's position gives the second order differential equation:

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{a}(\mathbf{r}(t), \frac{d\mathbf{r}(t)}{dt}, t). \quad (1.5)$$

Thus the principle of Newtonian determinacy indicates that **motion of the particle is governed by a second order differential equation in time!**

The acceleration function is not completely arbitrary. Since real particles do not spontaneously disappear, **physically acceptable** acceleration functions lead to second order differential equations whose solutions can be extended for **all** time.

Next I use the concept of inertial mass to decompose the acceleration function into a part that depends on the particle and a part that is independent of the particle.

If I consider two different point particles connected to identical springs, the two point particles will experience different accelerations. This means that in addition to the initial position and velocity, the acceleration function depends on additional properties of the particle. A second important observation is that in some *preferred coordinate systems* the acceleration functions for different particles are related by a multiplicative constant. The preferred coordinate systems are called **inertial coordinate systems** and in these systems the acceleration function can be separated into a product of a constant that is characteristic of the particle and a vector-valued function that is independent of the choice of particle. In inertial coordinate systems the second order differential equation can be put in the form

$$m \frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}(\mathbf{r}(t), \dot{\mathbf{r}}(t), t). \quad (1.6)$$

The constant  $m$ , characteristic of the particle, is called the **inertial mass** of the particle and the quantity  $\mathbf{F}(\mathbf{r}(t), \dot{\mathbf{r}}(t), t)$ , that only depends on the particle's coordinate and velocity, is called the **Force** on the particle. This equation is the familiar form of **Newton's second law**. The MKS unit of mass is a kilogram and MKS unit of force is a Newton ( $kg \cdot m/s^2$ ).

For example, for a fixed linear spring with force constant  $k$ , the angular frequency  $\omega := \sqrt{\frac{k}{m}}$  of oscillation depends on the inertial mass of the particle and can be used to distinguish particles with different inertial masses.

The inertial mass permits a clean separation of the acceleration function into a part that depends on properties of the particle and a part that depends on the initial conditions and the rest of the system on the right.

These considerations generalize to systems of  $N$  interacting point particles. The result is a set of coupled second order ordinary differential equations of the form

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_1 \cdots \mathbf{r}_N; \frac{d\mathbf{r}}{dt_1} \cdots \frac{d\mathbf{r}}{dt_N}, t) \quad (1.7)$$

In this case the particle's coordinate as a function of time,  $\mathbf{r}_i(t)$ , is the solution of  $3N$  coupled second-order differential equations. The solution requires specifying the initial coordinates and velocities of all  $N$  particles.

### 1.2.2 Local solutions

The equations of motion (1.6) can be equivalently expressed as a system of  $6N$  coupled first order differential equations

$$\frac{d\mathbf{v}_i}{dt} = \frac{1}{m_i} \mathbf{F}_i(\mathbf{r}_1, \cdots, \mathbf{r}_N, \mathbf{v}_1, \cdots, \mathbf{v}_N, t) \quad (1.8)$$

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i. \quad (1.9)$$

Integrating these equations from  $t_0$  to  $t$  gives an equivalent system of integral equations that incorporate the initial coordinate and velocity:

$$\mathbf{v}_i(t) = \mathbf{v}_i(t_0) + \int_{t_0}^t \frac{1}{m_i} \mathbf{F}_i(\mathbf{r}_1(t'), \cdots, \mathbf{r}_N(t'), \mathbf{v}_1(t'), \cdots, \mathbf{v}_N(t'), t') dt' \quad (1.10)$$

$$\mathbf{r}_i(t) = \mathbf{r}_i(t_0) + \int_{t_0}^t \mathbf{v}_i(t') dt'. \quad (1.11)$$

For small  $t - t_0$  this system can be solved by iteration. The iterative solution is given by

$$\mathbf{v}_i(t) = \lim_{k \rightarrow \infty} \mathbf{v}_i^k(t) \quad (1.12)$$

$$\mathbf{r}_i(t) = \lim_{k \rightarrow \infty} \mathbf{r}_i^k(t) \quad (1.13)$$

where the initial values are

$$\mathbf{v}_i^0(t) = \mathbf{v}_i(t_0) \quad (1.14)$$

$$\mathbf{r}_i^0(t) = \mathbf{r}_i(t_0) \quad (1.15)$$

and the  $k$ -th approximations can be expressed in terms of the  $k - 1$ -th approximation as

$$\mathbf{v}_i^k(t) = \mathbf{v}_i(t_0) + \int_{t_0}^t \frac{1}{m_i} \mathbf{F}_i(\mathbf{r}_1^{k-1}(t'), \cdots, \mathbf{r}_N^{k-1}(t'), \mathbf{v}_1^{k-1}(t'), \cdots, \mathbf{v}_N^{k-1}(t'), t') dt' \quad (1.16)$$

$$\mathbf{r}_i^k(t) = \mathbf{r}_i(t_0) + \int_{t_0}^t \mathbf{v}_i^{k-1}(t') dt'. \quad (1.17)$$

This is just the iteration used by Picard to prove the existence of local solutions of systems of differential equations. The convergence of this method has only been established for sufficiently small  $t - t_0$ . It is much more difficult to find stable computational methods for finding solutions that are valid for all time, however independent of mathematical considerations, physics considerations require that acceptable physical forces must be of the type that have solutions that are valid for all time.

### 1.2.3 Galilean invariance

A third important observation is called the **Principle of Galilean Relativity**. It states that **for isolated systems** the form of the dynamical equations is the same in all inertial coordinate systems.

For the special case of a single point particle in the absence of forces the equation of motion in an inertial coordinate system is

$$\frac{d^2 \mathbf{r}}{dt^2} = \mathbf{0}. \quad (1.18)$$

The form of this equation must be the same in any inertial coordinate system.

The form of equation (1.18) is preserved by (1) changing the origin of the coordinate system by a fixed constant vector (2) changing the origin of the coordinate system by a fixed constant velocity (3) rotating the coordinate system about a fixed point (4) or changing the time by a fixed amount. Combining these transformations defines a new inertial coordinate system:

$$\mathbf{r}_i \rightarrow \mathbf{r}'_i = R\mathbf{r}_i + \mathbf{v}_0 t + \mathbf{r}_0 \quad (1.19)$$

$$t \rightarrow t' = t + c. \quad (1.20)$$

Transformations generated by these elementary transformations are called **Galilean transformations**. Here  $\mathbf{r}_0$  and  $\mathbf{v}_0$  are constant vectors and  $R$  is a constant  $3 \times 3$  orthogonal matrix with unit determinant. The orthogonality ensures that it preserves the length of vectors and the condition on the determinant ensures that it does not include transformations that involve space reflection. Later I will show that these transformations are rotations.

Equation (1.18) are also invariant under changes of length or time scale, but these correspond to changes in distance or time units. When we consider interactions which have dimensions, they will also scale according to their dimensions.

It is an easy exercise to show (1) the composition of two Galilean transformations is a Galilean transformation (2) the identity is a Galilean transformation (3) every Galilean transformation has an inverse (4) the composition of Galilean transformations is associative.

These properties imply that the Galilean transformations define a mathematical structure called a **group** under composition. This is called the Galilean

group. The elements of this group are coordinate transformations that relate different inertial coordinate systems.

The Galilean transformations can be expressed in matrix form as

$$\begin{pmatrix} \mathbf{r}' \\ t' \\ 1 \end{pmatrix} = \begin{pmatrix} R & \mathbf{v} & \mathbf{r}_0 \\ 0 & I & t_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix}. \quad (1.21)$$

The requirement that the equations of motion for an isolated system have the same form in any inertial coordinate system also restricts the structure of the allowed forces,  $\mathbf{F}_i(\mathbf{r}_1 \cdots \mathbf{r}_N; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N, t)$ :

Invariance with respect to spatial translations means

$$\mathbf{F}_i(\mathbf{r}_1 \cdots \mathbf{r}_N; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N, t) = \mathbf{F}_i(\mathbf{r}_1 - \mathbf{a} \cdots \mathbf{r}_N - \mathbf{a}; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N, t). \quad (1.22)$$

Choosing  $\mathbf{a} = \mathbf{r}_1$  means that the force for an isolated system in an inertial coordinate system only depends on **coordinate differences**.

Invariance with respect to time shifts requires

$$\mathbf{F}_i(\mathbf{r}_1 \cdots \mathbf{r}_N; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N, t) = \mathbf{F}_i(\mathbf{r}_1 \cdots \mathbf{r}_N; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N, t - c). \quad (1.23)$$

Setting  $c = -t$  implies that the forces have **no explicit dependence on time**.

Since coordinate differences are preserved under shifts by constant velocity, invariance with respect to shifts by constant velocity implies

$$\mathbf{F}_i(\mathbf{r}_1 - \mathbf{r}_i \cdots \mathbf{r}_N - \mathbf{r}_i; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N) = \mathbf{F}_i(\mathbf{r}_1 - \mathbf{r}_i \cdots \mathbf{r}_N - \mathbf{r}_i; \dot{\mathbf{r}}_1 - \mathbf{v} \cdots \dot{\mathbf{r}}_N - \mathbf{v}) \quad (1.24)$$

Setting  $\mathbf{v} = \dot{\mathbf{r}}_1$  gives

$$\mathbf{F}_i(\mathbf{r}_1 \cdots \mathbf{r}_N; \dot{\mathbf{r}}_1 \cdots \dot{\mathbf{r}}_N, t) = \mathbf{F}_i(\mathbf{r}_1 - \mathbf{r}_i \cdots \mathbf{r}_N - \mathbf{r}_i; \dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_i \cdots \dot{\mathbf{r}}_N - \dot{\mathbf{r}}_i) \quad (1.25)$$

which means that the forces depend only on velocity differences. Invariance of the form of the equations of motion with respect to rotations means

$$\begin{aligned} & \mathbf{F}_i(\mathbf{r}_1 - \mathbf{r}_i \cdots \mathbf{r}_N - \mathbf{r}_i; \dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_i \cdots \dot{\mathbf{r}}_N - \dot{\mathbf{r}}_i) = \\ & R^{-1} \mathbf{F}_i(R(\mathbf{r}_1 - \mathbf{r}_i) \cdots R(\mathbf{r}_N - \mathbf{r}_i); R(\dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_i) \cdots R(\dot{\mathbf{r}}_N - \dot{\mathbf{r}}_i)) \end{aligned} \quad (1.26)$$

or

$$\begin{aligned} & R \mathbf{F}_i(\mathbf{r}_1 - \mathbf{r}_i \cdots \mathbf{r}_N - \mathbf{r}_i; \dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_i \cdots \dot{\mathbf{r}}_N - \dot{\mathbf{r}}_i) = \\ & \mathbf{F}_i(R(\mathbf{r}_1 - \mathbf{r}_i) \cdots R(\mathbf{r}_N - \mathbf{r}_i); R(\dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_i) \cdots R(\dot{\mathbf{r}}_N - \dot{\mathbf{r}}_i)) \end{aligned} \quad (1.27)$$

Thus forces consistent with the principle of Galilean relativity in an inertial coordinate system requires the the forces have no explicit time dependence, are functions of coordinate and velocity differences, and are vectors constructed out of the coordinate and velocity differences.

This principle applies to the universe or isolated subsystems in inertial coordinate systems. This principle does not apply to subsystems that are not isolated from their environment, and requires modification in non-inertial coordinate systems.

**End of first lecture?**

### 1.2.4 Inertial reference frames

The group of Galilean transformations tells one how to transform from one inertial coordinate system to another one, but it does not give any indication of how to experimentally determine if a given coordinate system is inertial. To answer this question it is useful to consider how the equations of motion are modified in a non-inertial coordinate system.

### 1.2.5 Accelerated reference frames

To understand this problem it is useful to postulate a dynamics formulated in a fixed inertial reference frame and consider how the dynamics looks in an arbitrary frame.

Consider general coordinates  $\mathbf{y}_i$  related to inertial coordinates  $\mathbf{r}_i$  by

$$\mathbf{y}_i = \mathbf{y}_i(\mathbf{r}_i, t). \quad (1.28)$$

To use Newton's laws in an inertial coordinate system I compute the first and second time derivatives of these coordinates:

$$\frac{d\mathbf{y}_i}{dt} = \sum_j \frac{\partial \mathbf{y}_i}{\partial r_i^j} \frac{dr_i^j}{dt} + \frac{\partial \mathbf{y}_i}{\partial t} \quad (1.29)$$

$$\begin{aligned} \frac{d^2\mathbf{y}_i}{dt^2} &= \sum_j \frac{\partial \mathbf{y}_i}{\partial r_i^j} \frac{d^2 r_i^j}{dt^2} + \sum_{jk} \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial r_i^k} \frac{dr_i^j}{dt} \frac{dr_i^k}{dt} + \\ &2 \sum_j \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial t} \frac{dr_i^j}{dt} + \frac{\partial^2 \mathbf{y}_i}{\partial t^2}. \end{aligned} \quad (1.30)$$

Multiplying by  $m_i$  and using Newton's second laws in an inertial coordinate system gives

$$\begin{aligned} m_i \frac{d^2\mathbf{y}_i}{dt^2} &= \sum_j \frac{\partial \mathbf{y}_i}{\partial r_i^j} F_i^j + m_i \sum_{kj} \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial r_i^k} \frac{dr_i^j}{dt} \frac{dr_i^k}{dt} + \\ &2m_i \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial t} \frac{dr_i^j}{dt} + m_i \frac{\partial^2 \mathbf{y}_i}{\partial t^2}. \end{aligned} \quad (1.31)$$

It is desirable to express these equations in terms of the new coordinates. This can be done using

$$\frac{dr_i^j}{dt} = \sum_k \frac{\partial r_i^j}{\partial y_i^k} \left( \frac{dy_i^k}{dt} - \frac{\partial y_i^k}{\partial t} \right). \quad (1.32)$$

Using (1.32) in (1.33) gives

$$m_i \frac{d^2\mathbf{y}_i}{dt^2} = \sum_j \frac{\partial \mathbf{y}_i}{\partial r_i^j} F_i^j + m_i \sum_{kj} \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial r_i^k} \left[ \sum_k \frac{\partial r_i^j}{\partial y_i^m} \left( \frac{dy_i^m}{dt} - \frac{\partial y_i^m}{\partial t} \right) \right] \left[ \sum_n \frac{\partial r_i^k}{\partial y_i^n} \left( \frac{dy_i^n}{dt} - \frac{\partial y_i^n}{\partial t} \right) \right] +$$

$$2m_i \sum_{jk} \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial t} \frac{\partial r_i^j}{\partial y_i^k} \left( \frac{dy_i^k}{dt} - \frac{\partial y_i^k}{\partial t} \right) + m_i \frac{\partial^2 \mathbf{y}_i}{\partial t^2}. \quad (1.33)$$

This equation looks like Newton's second law with a transformed force

$$\sum_j \frac{\partial \mathbf{y}_i}{\partial r_i^j} F_i^j \quad (1.34)$$

and three additional **inertial forces**:

$$\mathbf{F}_1 := m_i \sum_{kj} \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial r_i^k} \left[ \sum_k \frac{\partial r_i^j}{\partial y_i^m} \left( \frac{dy_i^m}{dt} - \frac{\partial y_i^m}{\partial t} \right) \right] \left[ \sum_n \frac{\partial r_i^k}{\partial y_i^n} \left( \frac{dy_i^n}{dt} - \frac{\partial y_i^n}{\partial t} \right) \right] \quad (1.35)$$

$$\mathbf{F}_2 := 2m_i \sum_{jk} \frac{\partial^2 \mathbf{y}_i}{\partial r_i^j \partial t} \frac{\partial r_i^j}{\partial y_i^k} \left( \frac{dy_i^k}{dt} - \frac{\partial y_i^k}{\partial t} \right) \quad (1.36)$$

and

$$\mathbf{F}_3 := m_i \frac{\partial^2 \mathbf{y}_i}{\partial t^2}. \quad (1.37)$$

The inertial forces can be distinguished from the transformed force by the appearance of the inertial mass  $m_i$  in these forces. In addition they do not vanish when the applied forces vanish, i.e.  $F_i^j = 0$ . This means that particles in non-inertial coordinate systems will experience spontaneous acceleration in the absence of applied forces.

These inertial forces are familiar. They include the force that pushes you back in you seat when an airplane takes off.

This equation suggest that a coordinate system for an isolated system is inertial if in that coordinate system a particle's acceleration is inversely proportional to its inertial mass. Clearly, for the inertial forces the acceleration is independent of the mass, so non-inertial coordinate systems do not have this property.

### 1.2.6 Gravity

The problem with the test described above is that while it works fine for electrical forces, it fails for gravitational forces. This is because of the remarkable observation that *the gravitational "charge" of a particle is identical to its inertial mass*. As a consequence of the **equivalence of the inertial and gravitational mass** the masses cancel on both sides of the equation and the particle's acceleration is independent of its mass in all coordinate systems. The equivalence of the gravitational and inertial mass is not explained by classical mechanics. In classical mechanics both masses have very difference origins. This observation is a consequence of the theory of general relativity.

### 1.2.7 Newton's first law

If the force on a particle in an inertial coordinate system vanishes, then the particle's acceleration is zero. Integrating the differential equation (1.6) with no force term gives

$$\mathbf{r}(t) = \mathbf{v}(t_0)(t - t_0) + \mathbf{r}(t_0) \quad (1.38)$$

which means that the particle moves with constant velocity. This is the content of Newton's first law.

### 1.2.8 Conservative forces

Many forces in nature are independent of velocities and are derivable from a single-valued potential function. I consider potential functions that do not depend on the particle velocities. A set of forces are **conservative** if they can be written in the form

$$\mathbf{F}_i = -\frac{\partial}{\partial \mathbf{r}_i} V(\mathbf{r}_1 \cdots \mathbf{r}_N) \quad (1.39)$$

for some single valued potential function  $V(\mathbf{r}_1 \cdots \mathbf{r}_N)$ .

The work done,  $W_{A,B}[\boldsymbol{\gamma}]$ , by a force along some path  $\boldsymbol{\gamma}(\lambda)$  between  $\mathbf{r}_A$  and  $\mathbf{r}_B$  is

$$W_{A,B}[\boldsymbol{\gamma}] := \int_{\boldsymbol{\gamma}} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 \mathbf{F} \cdot \frac{d\boldsymbol{\gamma}}{d\lambda} d\lambda \quad (1.40)$$

where  $\boldsymbol{\gamma}(0) = \mathbf{r}_A$  and  $\boldsymbol{\gamma}(1) = \mathbf{r}_B$ .

For a system of particles this is replaced by

$$W_{A,B}[\boldsymbol{\gamma}] := \sum_i \int_{\boldsymbol{\gamma}_i} \mathbf{F}_i \cdot d\mathbf{r}_i = \int_0^1 \sum_i \mathbf{F}_i \cdot \frac{d\boldsymbol{\gamma}_i}{d\lambda} d\lambda. \quad (1.41)$$

where  $\boldsymbol{\gamma}_i(0) = \mathbf{r}_{iA}$  and  $\boldsymbol{\gamma}_i(1) = \mathbf{r}_{iB}$ .

The work done by a conservative force is independent of the paths used to get from the initial to the final points.

$$W_{A,B}[\boldsymbol{\gamma}] := \int_0^1 \sum_i \mathbf{F}_i \cdot \frac{d\boldsymbol{\gamma}_i}{d\lambda} d\lambda = \int (-\nabla_i V(\boldsymbol{\gamma}_1(\lambda), \dots, \boldsymbol{\gamma}_N(\lambda))) \cdot \frac{d\boldsymbol{\gamma}_i}{d\lambda} d\lambda = \quad (1.42)$$

$$\begin{aligned} & \int_0^1 \frac{d}{d\lambda} V(\boldsymbol{\gamma}_1(\lambda) \cdots \boldsymbol{\gamma}_N(\lambda)) d\lambda = \\ & V(\boldsymbol{\gamma}_1(1) \cdots \boldsymbol{\gamma}_N(1)) - V(\boldsymbol{\gamma}_1(0) \cdots \boldsymbol{\gamma}_N(0)) = \\ & V(\mathbf{r}_{1A}, \dots, \mathbf{r}_{NA}) - V(\mathbf{r}_{1B}, \dots, \mathbf{r}_{NB}) \end{aligned} \quad (1.43)$$

which only depends on the endpoints of the paths, not the specific path.

For a closed path,  $\mathbf{r}_{iB} = \mathbf{r}_{iA}$

$$\boldsymbol{\gamma}_-(1) = \boldsymbol{\gamma}_i(0) \quad (1.44)$$



and for a single-valued potential

$$\begin{aligned} V(\boldsymbol{\gamma}_1(1) \cdots \boldsymbol{\gamma}_N(1)) - V(\boldsymbol{\gamma}_1(0) \cdots \boldsymbol{\gamma}_N(0)) = \\ V(\mathbf{r}_{1A}, \cdots, \mathbf{r}_{NA}) - V(\mathbf{r}_{1A}, \cdots, \mathbf{r}_{NA}) = 0 \end{aligned} \quad (1.45)$$

This can be expressed in the form

$$\sum_i \oint_{\boldsymbol{\gamma}_i} \mathbf{F}_i \cdot d\mathbf{r}_i = 0 \quad (1.46)$$

for any closed path  $\gamma = (\boldsymbol{\gamma}_1(\lambda), \dots, \boldsymbol{\gamma}_N(\lambda))$ .

### 1.2.9 Newton's third law

For a Galilean invariant systems interacting with velocity-independent conservative forces the potential only depends on the coordinate differences. This means that the potential functions satisfy

$$V(\mathbf{r}_1, \cdots, \mathbf{r}_N, t) = \tilde{V}(\mathbf{r}_1 - \mathbf{r}_N, \cdots, \mathbf{r}_{N-1} - \mathbf{r}_N, t). \quad (1.47)$$

It follows that the net force on this system is

$$\sum_{i=1}^N \mathbf{F}_i = - \sum_{i=1}^N \frac{\partial \tilde{V}}{\partial \mathbf{r}_i} = - \sum_{i=1}^{N-1} \left( \frac{\partial \tilde{V}}{\partial \mathbf{r}_i} - \frac{\partial \tilde{V}}{\partial \mathbf{r}_i} \right) = \mathbf{0}. \quad (1.48)$$

Which shows that the net force on a translationally invariant conservative system with velocity independent forces is zero. When applied to isolated systems of two particles it means that the force on particle 1 due to particle 2 is the negative of the force on particle 2 due to article 1. This is the usual form of Newton's third law.

### 1.2.10 Macroscopic particles - elementary conservation laws

Consider a macroscopic particle consisting of many point particles. Assume that each point particle experiences (1) an applied external force and (2) Galilean invariant conservative forces due the other point particles in the system.

Define the **total inertial mass**

$$M := \sum_i m_i \quad (1.49)$$

and the **center of mass coordinate**

$$\mathbf{R} := \sum_i \frac{m_i}{M} \mathbf{r}_i. \quad (1.50)$$

In terms of these quantities

$$M \frac{d^2 \mathbf{R}}{dt^2} = \sum_i \frac{M m_i}{M} \frac{d^2 \mathbf{r}_i}{dt^2} = \sum_i m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \sum_i \mathbf{F}_i \quad (1.51)$$

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i} + \mathbf{F}_{i \text{ ext}} \quad (1.52)$$

The sum of the derivatives of the potential vanishes (1.48) due to the Galilean invariance of the conservative inter-particle forces. It follows that the total force on the macroscopic particle is the vector sum of the applied external forces on the system. Summing over the particles gives

$$\sum_i \mathbf{F}_i = -\sum_i \frac{\partial V}{\partial \mathbf{r}_i} + \sum_i \mathbf{F}_{i \text{ ext}} \quad (1.53)$$

It follows that

$$M \frac{d^2 \mathbf{R}}{dt^2} = \sum_i \mathbf{F}_{i \text{ ext}} \quad (1.54)$$

which as the form of Newton's second law for a point particle of mass  $M$  and coordinate  $\mathbf{R}$  being acted on by a force  $\mathbf{F}_{\text{ext}} := \sum_i \mathbf{F}_{i \text{ ext}}$ .

This means the center of mass coordinate of the system behaves like a point particle being acted on by the sum of the external forces on the system. This justifies treating a macroscopic system of point particles by an idealized point particle. This equation holds independent of all of relative motion of the constituent point particles.

### 1.2.11 Conservation Laws

Symmetries in classical mechanics are normally associated with conserved quantities.

In the above analysis the translational invariance of conservative force ensures that the corresponding potential was only a function of coordinate differences. This in turn showed that the sum of all of the inter-particle forces vanish.

In the absence of external forces equation (1.54) becomes

$$M \frac{d^2 \mathbf{R}}{dt^2} = \frac{d}{dt} \left( M \frac{d\mathbf{R}}{dt} \right) = \mathbf{0} \quad (1.55)$$

The conserved vector quantity

$$\mathbf{P} = M \frac{d\mathbf{R}}{dt} \quad (1.56)$$

is called the **linear momentum of the system**. The above equation shows that all three components are conserved in the absence of external forces.

From the definitions it follows that

$$\mathbf{P} = M \frac{d\mathbf{R}}{dt} = \sum_i m_i \frac{d\mathbf{r}_i}{dt} := \sum_i \mathbf{p}_i \quad (1.57)$$

where  $\mathbf{p}_i := m_i \frac{d\mathbf{r}_i}{dt}$  is the **linear momentum of particle  $i$** .

The quantity

$$\mathbf{L} := \mathbf{R} \times \mathbf{P} \quad (1.58)$$

is called the **angular momentum of the system**. Note that its value depends on the position of the origin of the coordinate system. Note that for an isolated Galilean-invariant system

$$\frac{d\mathbf{L}}{dt} := \frac{d\mathbf{R}}{dt} \times \mathbf{P} + \mathbf{R} \times \frac{d\mathbf{P}}{dt} = \frac{1}{M} \mathbf{P} \times \mathbf{P} + \mathbf{R} \times \mathbf{F}_{ext} = \mathbf{R} \times \mathbf{F}_{ext}. \quad (1.59)$$

this vanishes in the absence of external forces or more generally when the **external torque**

$$\mathbf{T} := \mathbf{R} \times \mathbf{F}_{ext} \quad (1.60)$$

vanishes. It follows that in the absence of external torques the angular momentum of the system

$$\mathbf{L} := \mathbf{R} \times \mathbf{P} \quad (1.61)$$

is conserved. The last conservation law follows from

$$\mathbf{0} = M \frac{d^2\mathbf{R}}{dt^2} \cdot \frac{d\mathbf{R}}{dt} = \frac{1}{2} M \frac{d}{dt} \left( \frac{d\mathbf{R}}{dt} \cdot \frac{d\mathbf{R}}{dt} \right) = \frac{1}{2M} \mathbf{P} \cdot \mathbf{P} \quad (1.62)$$

which shows that in the absence of external forces the total kinetic energy is conserved.

These conservation laws hold for isolated systems or isolated point particles. We will discuss a more general treatment of conservation laws in the next section.

**End second lecture**



## Chapter 2

# Lagrangian Dynamics

### 2.1 Problems with constraints

Many mechanical systems involve constraints. In addition to explicit forces, the system also has forces due to the constraints that are often not explicitly known.

There are many different kinds of constraints that may be relevant for an isolated system. A **holonomic constraint** is a relation of the form

$$f(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0 \quad (2.1)$$

that constrains the particle's coordinates. The time dependence means that these constraints can depend on time. Not all constraints are holonomic. Some elementary examples of non-holonomic constraints are

$$r^2 + y^2 \leq L^2$$

$$f\left(\frac{d\mathbf{r}_1}{dt}, \dots, \frac{d\mathbf{r}_N}{dt}, t\right) = 0$$

$$\sum_u c_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \cdot d\mathbf{r}_i = 0 \quad (\text{unless } c_i = \frac{\partial g}{\partial \mathbf{r}_i} \text{ for some } g).$$

The method that I discuss for treating holonomic constraints can also be applied to the third kind of constraint listed above, called differential constraints.

There are two problems when a system is subject to holonomic constraints.

1. Because of the constraints, not all of the coordinates are independent.
2. The forces of constraint are not explicitly known.

An important observation that helps to solve both of these problems is that the motion is normally perpendicular to the forces of constraint, so the constraint forces do no work. For example, the normal force on an inclined plane keeps the particle on the plane, but it does no work as the particle slides down the plane.

## 2.2 Principle of virtual work

I begin by considering a system of  $N$  point particles. Newton's second law implies

$$\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i = 0 \quad (2.2)$$

Here  $\mathbf{F}_i$  represents the sum of all of the forces on particle  $i$  including the constraint forces. It follows that for an arbitrary infinitesimal displacement,  $\delta\mathbf{r}_i$  that

$$\left(\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i\right) \cdot \delta\mathbf{r}_i = 0 \quad (2.3)$$

Next I write the force on particle  $i$  as the sum of an applied force,  $\mathbf{F}_i^a$ , and a constraint force,  $\mathbf{F}_i^c$

$$\left(\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i^a - \mathbf{F}_i^c\right) \cdot \delta\mathbf{r}_i = 0. \quad (2.4)$$

I also restrict the infinitesimal displacements so they are **consistent with the constraints**. This requirement, along with the observation that the forces of constraint are perpendicular to the displacement, implies

$$\mathbf{F}_i^c \cdot \delta\mathbf{r}_i = 0. \quad (2.5)$$

Thus, for infinitesimal displacements consistent with the constraints, equation (2.4) becomes

$$\left(\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i^a\right) \cdot \delta\mathbf{r}_i = 0. \quad (2.6)$$

This step has eliminated the forces of constraint from the problem.

If the system has  $K$  holonomic constraints then it is possible, by eliminating variables, to express coordinates consistent with the constraints in terms of  $3N - K$  **generalized coordinates**,  $q_1 \cdots q_{m=3N-K}$ :

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \cdots, q_m) \quad (2.7)$$

Arbitrary small displacements consistent with the constraints can be expressed as

$$\delta\mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \quad (2.8)$$

where the small displacements in the  $3N - K$  generalized coordinates are independent. Thus for each  $i$

$$\left(\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i^a\right) \cdot \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j = 0. \quad (2.9)$$

If I sum (2.9) over  $i$ , because of the independence of the  $\delta q_i$ , the coefficient of each  $\delta q_i$  must vanish. This gives the following  $3N - K = m$  equations:

$$\sum_i \left(\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i^a\right) \frac{\partial \mathbf{r}_i}{\partial q_j} = 0 \quad (2.10)$$

for  $1 \leq j \leq m$ .

Now I use some elementary calculus to express (2.9) and (2.10) in a more useful form:

$$\begin{aligned} \sum_i \frac{d\mathbf{p}_i}{dt} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} &= \sum_i \frac{d}{dt} \left( m_i \frac{d\mathbf{r}_i}{dt} \right) \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = \\ \sum_i \frac{d}{dt} \left( m_i \frac{d\mathbf{r}_i}{dt} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} \right) &- \sum_i m_i \frac{d\mathbf{r}_i}{dt} \cdot \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_k} \right) \end{aligned} \quad (2.11)$$

Note that

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_k} \right) &= \frac{\partial^2 \mathbf{r}_i}{\partial q_k \partial q_l} \frac{dq_l}{dt} + \frac{\partial^2 \mathbf{r}_i}{\partial q_k \partial t} = \\ \frac{\partial}{\partial q_k} \left( \frac{\partial \mathbf{r}_i}{\partial q_l} \frac{dq_l}{dt} + \frac{\partial \mathbf{r}_i}{\partial t} \right) &= \frac{\partial}{\partial q_k} \left( \frac{d\mathbf{r}_i}{dt} \right) \end{aligned} \quad (2.12)$$

I also observe that because

$$\frac{d\mathbf{r}_i}{dt} = \frac{\partial \mathbf{r}_i}{\partial q_l} \frac{dq_l}{dt} + \frac{\partial \mathbf{r}_i}{\partial t} \quad (2.13)$$

it follows that

$$\frac{\partial \mathbf{r}_i}{\partial q_l} = \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_l}, \quad (2.14)$$

where  $\dot{q}_l = \frac{dq_l}{dt}$  and  $\dot{\mathbf{r}}_i = \frac{d\mathbf{r}_i}{dt}$ .

Using (2.11-2.14) in (2.10) gives

$$\begin{aligned} 0 &= \sum_i \left( \frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i^a \right) \cdot \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} = \\ \sum_i m_i \left( \frac{d}{dt} \left( \frac{d\mathbf{r}_i}{dt} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right) - \frac{d\mathbf{r}_i}{dt} \cdot \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) \right) &- \sum_i \mathbf{F}_i^a \cdot \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} = \\ \sum_i m_i \left( \frac{d}{dt} \left( \frac{d\mathbf{r}_i}{dt} \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} \right) - \frac{d\mathbf{r}_i}{dt} \cdot \left( \frac{\partial}{\partial \dot{q}_j} \left( \frac{d\mathbf{r}_i}{dt} \right) \right) \right) &- \sum_i \mathbf{F}_i^a \cdot \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} = \\ \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} \left( \sum_i \frac{1}{2} m_i \frac{d\mathbf{r}_i}{dt} \cdot \frac{d\mathbf{r}_i}{dt} \right) - \frac{\partial}{\partial \dot{q}_j} \left( \sum_i \frac{1}{2} m_i \frac{d\mathbf{r}_i}{dt} \cdot \frac{d\mathbf{r}_i}{dt} \right) &- \sum_i \mathbf{F}_i^a \cdot \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} = \\ \left( \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} - \frac{\partial}{\partial \dot{q}_j} \right) T - \sum_i \mathbf{F}_i^a \cdot \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \end{aligned} \quad (2.15)$$

where  $T$  is the **kinetic energy** of the system:

$$T := \sum_i \frac{1}{2} m_i \frac{d\mathbf{r}_i}{dt} \cdot \frac{d\mathbf{r}_i}{dt} \quad (2.16)$$

With these substitutions equation (2.9) becomes

$$\sum_j \left( \left( \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} - \frac{\partial}{\partial \dot{q}_j} \right) T - Q_j \right) \delta q_j = 0 \quad (2.17)$$

and equation (2.10) becomes

$$\left( \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} - \frac{\partial}{\partial q_j} \right) T = Q_j \quad (2.18)$$

where

$$Q_j := \sum_i \mathbf{F}_i^a \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (2.19)$$

is called called a **generalized force**. When the applied force is conservative then the generalized force takes on a simple form

$$\begin{aligned} Q_j &:= \sum_i \mathbf{F}_i^a \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \\ &= - \sum_i \frac{\partial V}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = - \frac{\partial V}{\partial q_j} \end{aligned} \quad (2.20)$$

Since the potential is independent of velocities I can replace (2.20) by

$$Q_j = \left( \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} - \frac{\partial}{\partial q_j} \right) V \quad (2.21)$$

Using this in (??) gives the following system of equations

$$\left( \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} - \frac{\partial}{\partial q_j} \right) (T - V) = 0 \quad (2.22)$$

The quantity  $L = T - V$  is called the **Lagrangian** of this system. The differential equations (2.22) are called **Lagrange's equations**. The standard form of Lagrange's equations is

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0. \quad (2.23)$$

For a conservative force equation (2.9) becomes

$$\sum_j \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} \right) \delta q_j = 0. \quad (2.24)$$

For conservative forces these equations are equivalent to Newton's second law. One advantage of Lagrange's equations is that the dynamical input is a single scalar quantity  $L$  rather than a number of vector forces. A second advantage is that in the Lagrangian formulation it is not necessary to know the forces of constraint to find the equation of motion of the particle. Finally, the Lagrangian allows one to use any set of convenient generalized coordinates. This is true even when there are no constraints on the system.

**End third lecture?**



### 2.2.1 Forces of constraint - Lagrange multipliers

Lagrange's equations have the feature that the forces of constraint never appear in the equations of motion. Sometimes it is desirable to know the forces of constraint. For example, in designing a roller coaster it is important to know that the cart stays on the track. The signal for failure is the normal force, which is a force of constraint, vanishes.

If I return to the principle of virtual work but replace the independent generalized coordinates by the full set of generalized coordinates, still assuming that the motion is consistent with the constraints, equation (2.9) is replaced by

$$\sum_{j=1}^{3N} \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} \right) \delta q_j = 0. \quad (2.25)$$

This differs from equation (2.9) in that the sum runs over all  $3N$  generalized coordinates, however in this case we cannot demand that the coefficient of each  $\delta q_i$  vanish because they are no longer independent.

The holonomic constraints can be expressed in terms of the full set of generalized coordinates as:

$$f_i(q_1 \cdots q_{3N}) = 0 \quad 1 \leq i \leq k. \quad (2.26)$$

Because there are  $k$  constraints we expect that there are only  $3N - k$  independent generalized coordinates. The constraints imply the  $k$  additional relations

$$\sum_{j=1}^{3N} \frac{\partial f_i(q_1 \cdots q_{3N})}{\partial q_j} \delta q_j = 0 \quad (2.27)$$

for any displacements  $\delta q_j$  consistent with the constraints.

I add zero to equation (2.25) using some undetermined coefficients  $\lambda_i$  multiplied by (2.27) for each  $1 \leq i \leq k$  I get

$$\sum_{j=1}^{3N} \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} - \sum_{i=1}^k \lambda_i \frac{\partial f_i(q_1 \cdots q_{3N})}{\partial q_j} \right) \delta q_j = 0. \quad (2.28)$$

This is valid for any  $\lambda_i$  as long as the particles move in a manner that is consistent with the constraints. I choose the  $k$   $\lambda_i$ 's so that for each time the coefficient of  $\delta q_j$  for  $3N - k + 1 \leq j \leq 3N$  is zero. It follows that

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} - \sum_{i=1}^k \lambda_i \frac{\partial f_i(q_1 \cdots q_{3N})}{\partial q_j} = 0 \quad (2.29)$$

for  $j = 3N - k + 1 \cdots 3N$  by the choice of  $\lambda$ . For this choice of the **Lagrange multipliers**  $\lambda_i$  the last  $k$  terms in the sum in (2.28) do not appear. The remaining  $3N - k$   $\delta q_i$  in the sum can be taken as the independent generalized

coordinates It follows that the coefficients of each of them also vanish. This gives

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} - \sum_{i=1}^k \lambda_i \frac{\partial f_i(q_1, \dots, q_{3N})}{\partial q_j} = 0 \quad (2.30)$$

for  $j = 1 \dots 3N - k$ .

While the interpretation of the first  $3N - k$  and last  $k$  equations differ, the result is the system of  $3N$  equations:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} - \sum_{i=1}^k \lambda_i \frac{\partial f_i(q_1 \dots q_{3N})}{\partial q_j} = 0 \quad (2.31)$$

for  $1 \leq j \leq 3N$ . When the  $k$  constraints are used in these equations we get equations that we can solve for the **Lagrange multipliers**,  $\lambda_i$ . Comparing (2.31 to (2.18) gives explicit expressions for the **generalized forces of constraint**:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \sum_{i=1}^k \lambda_i \frac{\partial f_i(q_1 \dots q_{3N})}{\partial q_j} = Q_{j\text{constraint}} \quad 1 \leq j \leq 3N - k \quad (2.32)$$

Note that unlike equation (2.24) equations (2.32) involve all  $3N$  generalized coordinates with explicit constraint forces.

This method is called the method of Lagrange undetermined multipliers. Note that generalized forces can depend on both coordinates and velocities.

### 2.3 Dissipative forces

Many dissipative forces in nature are velocity dependent. In many cases of interest it is possible express all of the dissipative forces of a system in terms of a single scalar function,  $P$ , of the coordinates and velocities

$$\mathbf{F}_i^d = \frac{\partial P}{\partial \dot{\mathbf{r}}_i} \quad (2.33)$$

This is true, for example when the dissipative force on particle  $i$  depends only the velocity of particle  $i$  and coordinates of all of the particles. Even when the dissipative forces are not velocity-dependent they can be put in this form.

For this type of dissipative force

$$Q_j := \sum_i \mathbf{F}_i^d \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \sum_i \frac{\partial P}{\partial \dot{\mathbf{r}}_i} \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} \quad (2.34)$$

Using (2.14) this becomes

$$Q_j = \sum_i \frac{\partial P}{\partial \dot{\mathbf{r}}_i} \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \frac{\partial P}{\partial \dot{q}_j}. \quad (2.35)$$

It follows that the equations of motion for combined conservative and dissipative forces become

$$\left( \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} - \frac{\partial}{\partial q_j} \right) L = \frac{\partial P}{\partial \dot{q}_j} \quad (2.36)$$

An example of a useful class of dissipation functions  $P$  are

$$P = - \sum_i \frac{\alpha_i}{n} (\dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i)^{n/2} \quad (2.37)$$

where  $n = 1$  gives static friction forces and  $n = 2$  gives viscous forces.

### 2.3.1 Principle of stationary action

There is an alternate way to derive Lagrange's equations. The derivation is based on a variational principle that has applications that go beyond Newtonian mechanics.

A **functional** is a mapping from a space of functions with certain properties to the space of real or complex numbers. A simple example is

$$F[f] = f(0) + \int (f^2(r) + \frac{df}{dr}(r)) dr. \quad (2.38)$$

For mechanics applications we consider a system that can be described by specifying the values of  $N$  generalized coordinates,  $\{q_i\}_{i=1}^N$ . Let  $\gamma_i(t) = q_i$  be a path that gives the value of the  $i$ -th generalized coordinate of the system as a function of time subject to the initial and final conditions:  $\gamma_i(t_0) = q_i^A$  and  $\gamma_i(t_f) = q_i^B$ .

The **Action functional**,  $A[\boldsymbol{\gamma}]$ , assigns a real number to a fixed collection of paths  $\{\gamma_i(t)\}$ . It is defined as

$$A[\boldsymbol{\gamma}] := \int_{t_0}^{t_f} L(\gamma_1(t'), \dots, \gamma_N(t'), \dot{\gamma}_1(t'), \dots, \dot{\gamma}_N(t'), t') dt' \quad (2.39)$$

where  $L(q_1, \dots, q_N, \dot{q}_1, \dots, \dot{q}_N, t)$  is the Lagrangian of the system. In general, the path  $\boldsymbol{\gamma}(t)$  may have no relation to the solution of the equations of motion, except that it has the same initial and final coordinates.

A path  $\boldsymbol{\gamma}_0$  is an **extremal path** or **stationary point** of  $A[\boldsymbol{\gamma}]$  if the **first variation** of  $A$  at  $\boldsymbol{\gamma}$ :

$$\delta A[\boldsymbol{\gamma}_0; \delta \boldsymbol{\gamma}] := \frac{dA[\boldsymbol{\gamma}]}{d\lambda} [\boldsymbol{\gamma}_0 + \lambda \delta \boldsymbol{\gamma}]_{\lambda=0} = 0 \quad (2.40)$$

vanishes for all displacement paths,  $\delta \boldsymbol{\gamma}(t)$ , that vanish at  $t = t_0$  and  $t = t_f$ . This means that is extremal with respect to all paths that have the same initial and final positions for given initial and final times.

It is instructive to compare this condition to the following formulation of a partial derivative of a function of  $N$  variables in the  $\hat{\mathbf{n}}$  direction.

$$\frac{dF(\mathbf{r} + \lambda \hat{\mathbf{n}})}{d\lambda} = \frac{\partial F}{\partial \mathbf{r}} \cdot \hat{\mathbf{n}}. \quad (2.41)$$

This is stationary at  $\mathbf{r} = \mathbf{r}_0$  if the partial derivatives of  $F$  in all directions vanish at  $\mathbf{r}_0$ . In the functional case vectors  $\mathbf{r}$  are replaced by functions, and the direction  $\hat{\mathbf{n}}$  is replaced by  $\delta\boldsymbol{\gamma}(t)$ .

### End fourth lecture

Next I show that the curves that leave the Action functional stationary are solutions of Lagrange's equations with prescribed initial and final conditions. The condition for the action to be stationary is

$$0 = \frac{dA[\boldsymbol{\gamma}_0 + \lambda\delta\boldsymbol{\gamma}]}{d\lambda} \Big|_{\lambda=0} \quad (2.42)$$

where

$$A[\boldsymbol{\gamma}_0 + \lambda\delta\boldsymbol{\gamma}] = \int_{t_0}^{t_f} L(\boldsymbol{\gamma}_{10}(t') + \lambda\delta\boldsymbol{\gamma}_1(t'), \dots, \boldsymbol{\gamma}_{N0}(t') + \lambda\delta\boldsymbol{\gamma}_N(t'), \dot{\boldsymbol{\gamma}}_{10}(t') + \lambda\delta\dot{\boldsymbol{\gamma}}_1(t'), \dots, \dot{\boldsymbol{\gamma}}_{N0}(t') + \lambda\delta\dot{\boldsymbol{\gamma}}_N(t'), t') dt' \quad (2.43)$$

where

$$\delta\dot{\boldsymbol{\gamma}}_i(t') := \frac{d}{dt'} \delta\boldsymbol{\gamma}_i(t') \quad (2.44)$$

Differentiating with respect to  $\lambda$  and setting  $\lambda$  to zero gives:

$$\sum_{i=1}^N \int_{t_i}^{t_f} \left( \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \frac{d\delta q_i}{dt} \right) dt \quad (2.45)$$

Next I integrate the second term by parts to get

$$0 = \sum_{i=1}^N \int_{t_i}^{t_f} \left( \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right) \delta q_i dt + \left( \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) (t_f) - \left( \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) (t_i). \quad (2.46)$$

The boundary term vanishes because  $\delta q_i(t_f) = \delta q_i(t_i) = 0$ .

Since the  $\delta q_i(t)$  are arbitrary independent functions, the coefficient of each  $\delta q_i$  in the integral must vanish, giving

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = 0 \quad (2.47)$$

which are identical to differential equations that I derived using the principle of virtual work.

The argument that the variational principle leads to the differential equations proceeds as follows. If I assume by contradiction that the differential equation is not satisfied for a point on the stationary curve, then there is a small time interval containing that point where  $\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right)$  are strictly positive or strictly negative. We can choose the  $\delta q_i$  to vanish outside of this region and

have a signs in this region so  $(\frac{\partial L}{\partial q_i} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}))\delta q_i$  are all strictly positive in this region. This particular variation of the action about stationary curve does not give zero, which contradicts the requirement that  $\gamma_0(t)$  makes the action stationary.

There is one difference between Lagrange's equations derived using the principle of virtual work compared to the principle of stationary action. That is that the principle of virtual work treats Lagrange's equations as an initial value problem. The motion is specified by the differential equations and the initial generalized coordinates and velocities. The principle of stationary action treats Lagrange's equations as a boundary value problem. The motion is specified by the differential equations, the initial and final generalized coordinates, and the initial and final times.

The variational principle can be used to find extremal solutions to any functional. Before we discuss some examples it is useful to discuss when the extreme points are local minima.

### 2.3.2 The second variation

The principle of stationary action is sometimes incorrectly called the principle of least action. When I consider a function of several variables, a stationary point is a local minimum if all of the first partial derivatives of the function vanish at the stationary point and the matrix of second partial derivatives evaluated at the stationary point is a real symmetric matrix with positive eigenvalues:

$$f(\mathbf{r}) = f(\mathbf{r}_0) + \sum_i \underbrace{\frac{\partial f}{\partial r_i}}_{\text{vanishes}}(\mathbf{r}_0)(r^i - r_0^i) + \frac{1}{2!} \sum_{ij} \underbrace{\frac{\partial^2 f}{\partial r_i \partial r_j}}_{\text{positive eigenvalues}}(\mathbf{r}_0) (r^i - r_0^i)(r^j - r_0^j) + \dots \quad (2.48)$$

A similar condition is used to determine whether a stationary point (curve) of a functional is a local minima. In this case equation (2.48) is replaced by

$$A[\gamma] = A[\gamma_0] + \frac{dA[\gamma_0 + \lambda\delta\gamma]}{d\lambda} \Big|_{\lambda=0} + \frac{1}{2} \frac{d^2A[\gamma_0 + \lambda\delta\gamma]}{d\lambda^2} \Big|_{\lambda=0} + \dots = A[\gamma_0] + \int \sum_i \frac{\delta A}{\delta \gamma_i(t)}[\gamma_0] \delta \gamma_i(t) dt + \frac{1}{2} \int \sum_{ij} \frac{\delta^2 A}{\delta \gamma_i(t) \delta \gamma_j(t')} [\gamma_0] \gamma_i(t) \gamma_j(t') dt dt' + \dots \quad (2.49)$$

where the second term vanishes if the action is stationary at  $\gamma_0(t)$  and the third term is positive for non-zero  $\delta\gamma_0$ . The quantities

$$\delta A[\gamma_0; \delta\gamma_0] = \frac{dA[\gamma_0 + \lambda\delta\gamma]}{d\lambda} \Big|_{\lambda=0} \quad (2.50)$$

and

$$\delta^2 A[\boldsymbol{\gamma}_0; \delta\boldsymbol{\gamma}] = \frac{d^2 A[\boldsymbol{\gamma}_0 + \lambda \delta\boldsymbol{\gamma}]}{d\lambda^2} \Big|_{\lambda=0} \quad (2.51)$$

are called the **first and second variation** of  $A[\boldsymbol{\gamma}]$  at  $\boldsymbol{\gamma}_0$ . The quantities

$$\frac{\delta A}{\delta\gamma_i(t)}[\boldsymbol{\gamma}_0] \quad (2.52)$$

and

$$\frac{\delta^2 A}{\delta\gamma_i(t)\delta\gamma_j(t')}[\boldsymbol{\gamma}_0] \quad (2.53)$$

are called the first and second functional derivatives of  $A$  at  $\boldsymbol{\gamma}_0$ . In order to investigate whether the action is a minimum at a path  $\boldsymbol{\gamma}_0$  that makes the action stationary, I express the second variation explicitly in terms of the Lagrangian as

$$\begin{aligned} \delta^2 A[\boldsymbol{\gamma}_0; \delta\boldsymbol{\gamma}] = & \\ & \int_{t_i}^{t_f} \left\{ \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}(\boldsymbol{\gamma}_0(t)) \delta \dot{\gamma}_i(t) \delta \dot{\gamma}_j(t) + 2 \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}(\boldsymbol{\gamma}_0(t)) \delta \dot{\gamma}_i(t) \delta \gamma_j(t) + \right. \\ & \left. \frac{\partial^2 L}{\partial q_i \partial \dot{q}_j}(\boldsymbol{\gamma}_0(t)) \delta \gamma_i(t) \delta \dot{\gamma}_j(t) \right\} dt. \end{aligned} \quad (2.54)$$

For fixed  $\boldsymbol{\gamma}_0(t)$  this becomes a quadratic functional in  $\delta\boldsymbol{\gamma}$ . The strategy used to determine if the stationary point  $\boldsymbol{\gamma}_0(t)$  is a local minimum of the action is to look for stationary solutions of the new functional of  $\delta\boldsymbol{\gamma}$ ,

$$F[\delta\boldsymbol{\gamma}] := \delta^2 A[\boldsymbol{\gamma}_0; \delta\boldsymbol{\gamma}] \quad (2.55)$$

If this has a minimum, the minimum will be a stationary  $\delta\boldsymbol{\gamma}(t)$ . If these stationary  $\delta\boldsymbol{\gamma}(t)$  all make this functional positive, then the solution  $\boldsymbol{\gamma}_0(t)$  of the original problems is a local minimum of the action.

The difficulty with this result is because this functional is homogeneous of degree 2 in  $\delta\boldsymbol{\gamma}$ ,

$$F[\lambda\delta\boldsymbol{\gamma}] = \lambda^2 F[\delta\boldsymbol{\gamma}] \quad (2.56)$$

the best that can be hoped for is a minimum of zero. This problem can be avoided by requiring a normalization condition

$$1 = \int \sum_i \delta\gamma_i(t) \delta\gamma_i(t) dt \quad (2.57)$$

which simply fixes a scale. Rather than using this constraint to eliminate degrees of freedom, it is more efficient to introduce this constraint using a Lagrange multiplier. This involves adding the following term in the integrand of (2.54)

$$\lambda \left( \sum_i \delta q_i(t) \delta \dot{q}_i(t) - \frac{1}{t_f - t_i} \right) \quad (2.58)$$

I define the-time dependent matrices

$$A_{ij}(t) := \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}(\boldsymbol{\gamma}_0(t)) \quad (2.59)$$

$$B_{ij}(t) := \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}(\boldsymbol{\gamma}_0(t)) \quad (2.60)$$

$$C_{ij}(t) := \frac{\partial^2 L}{\partial q_i \partial \dot{q}_j}(\boldsymbol{\gamma}_0(t)) \quad (2.61)$$

Note that  $A_{ij}$  and  $C_{ij}$  are symmetric while  $B_{ij}$  is not. In this notation the functional  $F[\delta\boldsymbol{\gamma}]$ , including the Lagrange multiplier, becomes

$$\int_{t_i}^{t_f} \sum_{ij} \{A_{ij}(t)\delta\dot{\gamma}_i(t)\delta\dot{\gamma}_j(t) + 2B_{ij}(t)\delta\dot{\gamma}_i(t)\delta\gamma_j(t) + C_{ij}(t)\delta\gamma_i(t)\delta\dot{\gamma}_j(t) - \lambda(\sum_i \delta\gamma_i(t)\delta\gamma_i(t) - \frac{1}{t_f - t_i})\} dt \quad (2.62)$$

Lagrange's equations for the variation  $\delta\gamma_i(t)$  are

$$2\frac{d}{dt}(A_{ij}(t)\delta\dot{\gamma}_j(t)) + 2\frac{d}{dt}(B_{ji}(t)\delta\gamma_j(t)) - 2B_{ij}(t)\delta\dot{\gamma}_j(t) - 2C_{ij}(t)\delta\gamma_j(t) + 2\lambda\delta_{ij}\delta\gamma_j(t) = 0. \quad (2.63)$$

This has the form of an eigenvalue problem

$$\sum_j \frac{d}{dt}(A_{ij}(t)\delta\dot{\gamma}_j(t) + B_{ij}(t)\delta\gamma_j(t)) - \sum_j (B_{ij}(t)\dot{\gamma}_j(t) + C_{ij}(t)\delta\gamma_j(t)) = -\lambda\delta_{ij}\delta\gamma_j(t). \quad (2.64)$$

This differential equation is a linear equation of the form

$$\mathbf{D}\delta\boldsymbol{\gamma}(t) = -\lambda\delta\boldsymbol{\gamma}(t) \quad (2.65)$$

where  $D$  is a second order linear differential operator. It satisfies

$$\int \boldsymbol{\xi}(t)D\boldsymbol{\gamma}(t)dt = \int \boldsymbol{\gamma}(t)D\boldsymbol{\xi}(t)dt \quad (2.66)$$

which can be seen by integrations by parts using the fact the that  $\delta\boldsymbol{\gamma}(t)$  vanish at the endpoints of the integral.

It only has solutions for certain eigenvalues  $\lambda$ . the solution  $\delta\boldsymbol{\gamma}_i(t)$  for the  $i$ -th eigenvalue is called the  $i^{\text{th}}$  eigenfunction. Integration by parts and the boundary conditions imply

$$\int_{t_i}^{t_f} \delta\boldsymbol{\gamma}_i(t)\mathbf{D}(t)\delta\boldsymbol{\gamma}_j(t)dt = -\lambda_i \int_{t_i}^{t_f} \delta\boldsymbol{\gamma}_i(t)\delta\boldsymbol{\gamma}_j(t)dt = -\lambda_j \int_{t_i}^{t_f} \delta\boldsymbol{\gamma}_i(t)\delta\boldsymbol{\gamma}_j(t)dt \quad (2.67)$$

Subtracting this equation from itself leads to the orthogonality condition

$$0 = (\lambda_i - \lambda_j) \int_{t_i}^{t_f} \delta\gamma_i(t) \delta\gamma_j(t) dt \quad (2.68)$$

from which one learns that the eigenfunctions are orthogonal. In the case that there are several eigenfunctions with the same eigenvalue it is possible to construct linear combinations of these functions that have the same eigenvalue and are orthogonal.

This class of differential equations are called Sturm-Liouville equations. They have the following properties. There are an infinite number of discrete eigenvalues  $\lambda_n$  with  $|\lambda_n| \rightarrow \infty$ . The eigenvalues are real and the eigenfunctions are a basis for functions  $\delta\gamma(t)$  on the interval  $[t_i, t_f]$ .

Since the equation is homogeneous we can normalize the eigenfunctions so they are orthonormal

$$\int_{t_i}^{t_f} \delta\gamma_i(t) \delta\gamma_j(t) dt = \delta_{ij} \quad (2.69)$$

The basis property means that an arbitrary  $\delta\gamma(t)$  can be expressed as

$$\delta\gamma(t) = \sum_n c_n \delta\gamma_n(t) \quad (2.70)$$

where

$$c_n = \int_{t_i}^{t_f} dt \delta\gamma_n(t) \delta\gamma(t) \quad (2.71)$$

Evaluating  $F[\delta\gamma]$  gives

$$F[\delta\gamma] = F\left[\sum_n c_n \delta\gamma_n\right] = - \sum_{mn} c_m c_n \int_{t_i}^{t_f} \delta\gamma_m(t) \mathbf{D}(t) \delta\gamma_n(t) dt \quad (2.72)$$

$$\sum_n \lambda_n c_n^2 \quad (2.73)$$

which is a sum of positive constants (note the functions and eigenvalues are all real) multiplied by eigenvalues  $\lambda_n$ . Thus, a necessary and sufficient condition for the second variation of  $A$  about the stationary  $\gamma_0(t)$  to be positive is that all of the eigenvalues  $\lambda_n > 0$  of the Sturm-Liouville eigenvalue equation are positive. The eigenvalues are the Lagrange multipliers and they represent the value of the second variation on the  $n$ -th normalized eigenfunction.

Next we consider the simplest one degree of freedom case and argue that initially the stationary solution is a minimum of the action. For sufficiently small  $t$  the motion is determined by the initial coordinate and velocity. The force only contributes to the acceleration, which is the second time derivative:

$$r(t) = r(0) + v(0)t + \dots \quad (2.74)$$

This is the solution of

$$\frac{dT}{dt} = 0 \quad (2.75)$$



which corresponds to  $A = m$ ,  $B = C = 0$ , The boundary value problem for the  $\delta r$  is

$$m \frac{d^2}{dt^2} \delta r = -\lambda \delta r \quad (2.76)$$

which has the solutions

$$\delta r(t) = c \sin\left(\sqrt{\frac{\lambda}{m}}(t_f - t_i)\right) \quad (2.77)$$

with eigenvalues

$$\lambda_n = \frac{m\pi}{t_f - t_i} n^2 \quad (2.78)$$

These are positive for all  $n$ , which means that for small time the path followed by a particle is always the path that minimizes the action functional.

If one investigates what happens for longer times the eigenvalues change with time depending on the potential. What can happen is that the eigenvalues can pass through zero and change sign. A point  $t_f$  where the second variation has a zero eigenvalue is called a conjugate point to  $t_i$ . These are characterized by having a non-trivial solution to

$$\frac{d}{dt}(A_{ij}(t)\delta\dot{\gamma}_j(t)) + \left(\frac{d}{dt}(B_{ij}(t) - (C_{ij}(t))\delta\gamma_j(t))\right) = 0. \quad (2.79)$$

which is called the Jacobi equation.

To understand the role played by the Jacobi equation consider the simple case of one degree of freedom and let  $\gamma_0(\dot{q}_0, t)$  be a solution of Lagrange's equation with initial condition  $\gamma_0(t) = q_0$  and  $\dot{\gamma}_0(0) = \dot{q}_0$ . Define

$$J(\dot{q}_0, t) := \frac{\partial \gamma_0(\dot{q}_0, t)}{\partial \dot{q}_0} \quad (2.80)$$

By definition

$$\gamma_0(\dot{q}_0 + \epsilon, t) - \gamma_0(\dot{q}_0, t) = \epsilon J(\dot{q}_0, t) + o(\epsilon)^2 \quad (2.81)$$

Since independent of the choice of  $\dot{q}_0$ ,  $\gamma_0(\dot{q}_0, t)$  is a solution of Lagrange's equation with the same initial coordinate

$$\frac{d}{d\dot{q}_0} \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{\gamma}} - \frac{\partial L}{\partial \gamma} \right) = 0e \quad (2.82)$$

Changing the order of the derivatives gives

$$\left( \frac{d}{dt} \left( \frac{\partial^2 L}{\partial \dot{\gamma}^2} \frac{d\dot{\gamma}}{d\dot{q}_0} + \frac{\partial^2 L}{\partial \dot{\gamma} \partial \gamma} \frac{d\gamma}{d\dot{q}_0} - \frac{\partial L^2}{\partial \dot{\gamma} \partial \gamma} \frac{d\dot{\gamma}}{d\dot{q}_0} - \frac{\partial L^2}{\partial \gamma^2} \frac{d\gamma}{d\dot{q}_0} \right) \right) = 0 \quad (2.83)$$

This can be rewritten in the form

$$\left( \frac{d}{dt} \left( \frac{\partial^2 L}{\partial \dot{\gamma}^2} \frac{d\dot{\gamma}}{d\dot{q}_0} \right) + \left( \frac{d}{dt} \left( \frac{\partial^2 L}{\partial \dot{\gamma} \partial \gamma} \right) - \frac{\partial L^2}{\partial \gamma^2} \right) \right) \quad (2.84)$$

Inspection shows that this is the one dimensional version of

$$\frac{d}{dt}(C(t)\frac{d\dot{\gamma}}{d\dot{q}_0}) - (A(t) - \frac{dB}{dt}(t))\frac{d\gamma}{d\dot{q}_0} = 0 \quad (2.85)$$

which demonstrates that  $\frac{d\gamma}{d\dot{q}_0}$  satisfies the Jacobi equation.

Note that

$$J(\dot{q}_0, 0) = 0 \quad (2.86)$$

because the he initial coordinate does not depend on  $\dot{q}_0$ . If

$$J(\dot{q}_0, t) = 0 \quad (2.87)$$

also vanishes then the final point is also independent of

$$\dot{q}_0 \quad (2.88)$$

to leading order. This corresponds to a conjugate point and leads to the interpretation of the conjugate points as focal points of solutions that start at the same point with different velocities.

### 2.3.3 Noether's theorem

One of the advantages of the Lagrangian formulation of mechanics is that there is a connection between symmetries of the action and conserved quantities. This connection is often used in field theories, but the principle also applies to systems of particles.

To formulate Noether's theorem consider the following transformations on the generalized coordinates as time

$$t \rightarrow t' = t + \delta t(\epsilon, t) \quad (2.89)$$

$$q_i(t) \rightarrow q'_i(t') + \delta q_i(\epsilon, t) \quad (2.90)$$

where both  $\delta t(\epsilon, t)$  and  $\delta q_i(\epsilon, t)$  vanish in the limit that  $\epsilon \rightarrow 0$ .

Now we assume that the combined transformation leaves the action invariant

$$A[\boldsymbol{\gamma}, t_f, t_i] = A[\boldsymbol{\gamma}', t'_f, t'_i] = \quad (2.91)$$

**Noether's theorem** states that the invariance of the action, (2.91), implies that the following quantity is conserved for solutions of Lagrange's equations:

$$(L - \sum_i \frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{dt}) \frac{\partial \delta t}{\partial \epsilon}(0, t) + \sum_i \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \delta \gamma_i}{\partial \epsilon}(0, t) \quad (2.92)$$

To prove this theorem we only need to require that the action is conserved to leading order in  $\epsilon$ . This gives

$$\frac{d}{d\epsilon} A[\boldsymbol{\gamma}', t'_f, t'_i]_{\epsilon=0} = 0 \quad (2.93)$$

which can be manipulated to get the above theorem.

To evaluate this we need to expand the coordinate and time changes to leading order in  $\epsilon$ :

$$t \rightarrow t' = t + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t) + o(\epsilon^2) \quad (2.94)$$

$$q_i(t) \rightarrow q'_i(t') = q_i(t) + \epsilon \frac{\partial \delta q_i}{d\epsilon}(0, t) + o(\epsilon^2) \quad (2.95)$$

Since the time is a dummy integration variable in the action, it is useful to use the first equation to express the right side of the second equation in the same variable,  $t'$ , that appears on the left. In doing these we only retain the terms that are linear in  $\epsilon$ . Thus (2.95) becomes

$$\begin{aligned} q'_i(t') &= q_i(t' - \epsilon \frac{\partial \delta t}{d\epsilon}(0, t)) + \epsilon \frac{\partial \delta q_i}{d\epsilon}(0, t) + o(\epsilon^2) = \\ q_i(t') - \epsilon \frac{dq_i}{dt} \frac{\partial \delta t}{d\epsilon}(0, t') + o(\epsilon) &+ \epsilon \frac{\partial \delta q_i}{d\epsilon}(0, t) + o(\epsilon^2) = \\ q_i(t') - \epsilon \frac{dq_i}{dt} \frac{\partial \delta t}{d\epsilon}(0, t') + \epsilon \frac{\partial \delta q_i}{d\epsilon}(0, t) &+ o(\epsilon^2) \end{aligned} \quad (2.96)$$

Next we use these to express the difference between the transformed action and original action to leading order in  $\epsilon$

$$\begin{aligned} 0 &= A[\gamma', t'_f, t'_i] - A[\gamma, t_f, t_i] = \int_{t_i + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t_i) + \dots}^{t_f + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t_f) + \dots} \\ L\left(\frac{d}{dt}(q_i(t') - \epsilon \frac{dq_i}{dt} \frac{\partial \delta t}{d\epsilon}(0, t') + \epsilon \frac{\partial \delta q_i}{d\epsilon}(0, t) + o(\epsilon^2)), q_i(t')\right. \\ &\quad \left. - \epsilon \frac{dq_i}{dt} \frac{\partial \delta t}{d\epsilon}(0, t') + \epsilon \frac{\partial \delta q_i}{d\epsilon}(0, t') + o(\epsilon^2), t') dt' \right. \\ &\quad \left. - \int_{t_i}^{t_f} L\left(\frac{d}{dt}(q_i(t'), q_i(t'), t') dt' \right) \end{aligned} \quad (2.97)$$

To expand this out we first write

$$\begin{aligned} &\int_{t_i + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t_i) + \dots}^{t_f + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t_f) + \dots} = \\ &\int_{t_i + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t_i) + \dots}^{t_i} + \int_{t_i}^{t_f} + \int_{t_f}^{t_f + \epsilon \frac{\partial \delta t}{d\epsilon}(0, t_f) + \dots} \end{aligned} \quad (2.98)$$

The integrand can be expanded in a Taylor series in  $\epsilon$ . Since the width of the domain of integration in the first and last integrals is of order  $\epsilon$  we only need to pick up the  $\epsilon$  independent term in those factors, and we only need the first order term in the middle integral, because the 0 -  $th$  order term is subtracted off.

Combining everything gives

$$\begin{aligned}
&= \epsilon \left( \frac{\partial \delta t}{\partial \epsilon}(0, t_f) L(\dot{\mathbf{q}}, \mathbf{q}, t_f) - \frac{\partial \delta t}{\partial \epsilon}(0, t_i) L(\dot{\mathbf{q}}, \mathbf{q}, t_i) \right) + \\
&\quad + \epsilon \sum \int \left( \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \frac{d}{dt} \left( -\dot{\mathbf{q}} \frac{\partial \delta t}{\partial \epsilon}(0, t) + \frac{\partial \delta \mathbf{q}}{\partial \epsilon}(0, t) \right) \right) \\
&\quad + \epsilon \sum \int \left( \frac{\partial L}{\partial \mathbf{q}} \cdot \left( -\dot{\mathbf{q}} \frac{\partial \delta t}{\partial \epsilon}(0, t) + \frac{\partial \delta \mathbf{q}}{\partial \epsilon}(0, t) \right) \right) \quad (2.99)
\end{aligned}$$

Now we assume that  $\mathbf{q}(t)$  is a solution of Lagrange's equation so

$$\frac{\partial L}{\partial \mathbf{q}} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \quad (2.100)$$

Using this and writing the first term in (??) as the integral of a derivative we get:

$$\begin{aligned}
0 &= \epsilon \sum \int_{t_i}^{t_f} \frac{d}{dt} \frac{\partial \delta t}{\partial \epsilon}(0, t) L(\dot{\mathbf{q}}, \mathbf{q}, t) + \\
&\quad + \epsilon \sum \int_{t_i}^{t_f} \frac{d}{dt} \left( \frac{\partial L}{\partial \mathbf{q}} \cdot \left( -\dot{\mathbf{q}} \frac{\partial \delta t}{\partial \epsilon}(0, t) + \frac{\partial \delta \mathbf{q}}{\partial \epsilon}(0, t) \right) \right) dt + o(\epsilon^2) \quad (2.101)
\end{aligned}$$

Differentiating with respect to  $\epsilon$  and setting  $\epsilon \rightarrow 0$  gives the conservation law

$$0 = \left( L - \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \dot{\mathbf{q}} \right) \frac{\partial \delta t}{\partial \epsilon}(0, t) + \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \frac{\partial \delta \mathbf{q}}{\partial \epsilon}(0, t)$$

is independent of time. It is important to realize (1) this only works if the action is unchanged to leading order and (2) only for solutions of Lagrange's equations.

This completes the proof of Noether's theorem. Now I present some example showing the implications for Galilean invariance of the action.