Classical mechanics of particles and fields — pt 1

1. Overview

We begin by reformulating the familiar vectorial laws of Newtonian dynamics — the result is the analytical mechanics introduced by Joseph Louis Lagrange (Fig. 1) in his *Mécanique analytique*, published in 1788. Sir William Rowan Hamilton described this as a scientific poem, on account of the elegance of the analysis.

This is a powerful approach for complex dynamical problems. Whereas the Newtonian approach places great emphasis on external agents via the forces they exert, the Lagrangian approach focuses specifically on the behaviour of the dynamical objects themselves. We will see how it is applied to some discrete systems, but our major focus will be methods and formulations, not specific answers, since fundamental to solving difficult problems is normally to formulate them correctly using appropriate methods.

Next we will see how the approach can be couched within the language of variational calculus, and through Hamilton's principle of least action see the mathematics behind what Pierre Louis Maupertuis articulated as "*Nature is thrifty in all its actions*":

> *The laws of motion and equilibrium derived from this principle are exactly those observed in Nature. We may admire the applications of this principle in all phenomena: the movement of animals, the growth of plants, the revolutions of the planets, all are consequences of this principle. The spectacle of the universe seems all the more grand and beautiful and worthy of its Author, when one considers that it is all derived from a small number of laws laid down most wisely*.[1]

The close connection between symmetry and conservation laws in mechanics will be revealed through further analysis. Indeed, a further reformulation to Hamiltonian dynamics will shed more light on dynamical systems, and we will develop a concept of energy conservation that is more general than the conventional one with which you are familiar. Finally, by making a transition from discrete to continuous systems we will extend our theory to `fields`, in doing so making some small steps in developing the mathematical language used to understand what are now considered the most fundamental objects in nature.

Figure 1: L-R: Isaac Newton, 1643-1727. Pierre-Louis Moreau de Maupertuis, 1698-1759. Giuseppe Lodovico Lagrangia, 1736-1813. William Rowan Hamilton, 1805-1865. Four grumpy old men.

[1] Pierre Louis Maupertuis, Accord de différentes loix de la nature qui avoient jusqu'ici paru incompatibles, 1744.

Except when their solution shows something of interest, we will generally consider problems solved when we have stated differential equations describing their evolution — at the very least these can, if required, be solved numerically using methods taught elsewhere. We will, however, consider the description of systems close to stable equilibrium, since this opens up a powerful description in terms of normal modes.

Quantum mechanics: we've heard of that! Although slightly off-topic, as a nod to what we know to be a more complete theory of Nature than is classical mechanics, we will at points explore the connection between what we do and what we should do, and for example recognise that actually *Nature is not necessarily thrifty in all its actions.*

2. Lagrangian dynamics

We all have a basic understanding of how to analyse (simple) dynamical systems using Newtonian mechanics. We draw an illustrative diagram showing a particular configuration of the particles or bodies[2] that make up the system. We indicate forces that act on the various part(icle)s, and the accelerations. We then apply the equation $F = ma$ in one, two or three different directions, depending upon the dimensionality of the problem. If torques are present, we use $\tau = I\ddot{\vartheta}$. (If masses are varying, then $F = \dot{p}, \tau = \dot{L}$ is more correct.) These equations of motion are then integrated up, to yield the solution.

It sounds straightforward, but we all know it isn't always such in practice. If for no other reason than that, a reformulation should be of interest.

2.1 Generalised coordinates

We normally use Cartesian coordinates when describing the position of a particle in a dynamical system: in 3D (x, y, z), or (x_1, x_2, x_3). They are straightforward, and familiar. However, we might have a system where symmetry or the nature of the motion make it more appropriate to use cylindrical polars (ϱ, φ, z), spherical polars (r, ϑ, φ), or one of the other (16?) 3D orthogonal curvilinear coordinate systems. For example, if we are describing the dynamics of atoms in a molecule, we might instead specify the configuration of the system in terms of bond lengths, and bond angles.

In reformulating our approach to mechanics we seek to move away from Cartesians, and instead use `generalised coordinates`, denoted q_i. The generalised coordinates of a system is a natural, minimal complete set of parameters with which one can completely specify the configuration of a system. They may be lengths, angles, or some other. We assume that an instantaneous configuration of our dynamical system is specified by a set of n independent generalised coordinates: q_i for $i = 1,2, \ldots n$. Such a system has n `degrees of freedom`. The instantaneous position of a particle freely moving in three dimensions is given by three independent coordinates, such as $(q_1, q_2, q_3) = (x, y, z)$; the system has three degrees of freedom. Two particles freely moving in three dimensions has 6 degrees of freedom: three for each particle. If there are N particles, we need $3N$ coordinates, and there are $3N$ degrees of freedom — unless there are constraints (see below). The evolution of a system in time then corresponds to the variation of the generalised coordinates in an n-dimensional space, called `configuration space`.

Consider a system of N freely moving particles in 3D – a specific configuration of the system can be described by $n = 3N$ Cartesian coordinates, $x_i,\ i = 1,\ \ldots n$. The first three are the Cartesian coordinates of particle 1; the next three, the Cartesian coordinates of particle 2, etc. Now suppose the system can also be specified in terms of n generalised coordinates, $q_i,\ i = 1 \ldots n$. For concreteness, let us think of using

[2] From now on, we stick to the term particles. But they may be bodies.

cylindrical polar coordinates instead, so the first three q_i's are ϱ, φ, z of particle 1; the next three ϱ, φ, z of particle 2, etc. We know we can express the x_i in terms of the q_i, and in general

$$x_i = x_i(q_1, q_2, \dots, q_n, t) \qquad i = 1, \dots n. \qquad (1)$$

For example, we have expressions like $x = \varrho\cos\phi$, $y = \varrho\sin\phi$ and $z = z$ relating cartesian and cylindricals. In Eqn. (1) we have indicated an explicit dependence of time t. For example, the two coordinate systems may be in different inertial reference frames moving relative to one another; or the particles may be constrained to move on a surface, which is itself moving.

2.2 Constraints

A particle moving in three dimensions may have fewer than three degrees of freedom, because it is subject to `constraints`. For example, a ball bearing set in motion inside a hemispherical bowl of radius a has (x, y, z) coordinates that change with time, but we can uniquely specify the position by the two spherical polar angles (ϑ, φ). Gravity constrains the ball bearing to move over the surface of the bowl, so that $x^2 + y^2 + z^2 = a^2$, or $r = a$.

This is an example of a `holonomic constraint`, which is one which can be expressed as an equation of the form

$$f(q_1, q_2, \dots, q_m, t) = 0. \qquad (2)$$

A system of N particles moving in three dimensions subject to k holonomic constraints has $3N - k$ degrees of freedom. $3N - k$ generalised coordinates are needed to describe the system. For the ball bearing in the hemispherical bowl, we have one particle and one constraint – and can use the spherical polar angles as the two generalised coordinates. For another example, consider a system in which a bead threaded onto a wire twisted into a circular helix moves under the action of some force. It moves in three dimensions (x, y, z), but only requires one generalised coordinate to specify its location. This is because the parametric equations of a circular helix of radius a, pitch b with axis along the z-axis are $x(s) = a\cos s$, $y(s) = a\sin s$, $z(s) = (b/2\pi)s$, so that the motion of the bead is constrained to always satisfy the two holonomic constraints

$$\begin{aligned} f_1(x, y, z, t) &= x - a\cos(2\pi z/b) = 0 \\ f_2(x, y, z, t) &= y - a\sin(2\pi z/b) = 0. \end{aligned}$$

$3 - 2 = 1$. In this case we could use for the one generalised coordinate z, or s.

Not all constraints are holonomic. A particle is placed on the surface of a solid sphere. Its subsequent motion satisfies $|\vec{r} - \vec{r_0}|^2 \geq a^2$, since the distance of the particle from the centre of the sphere, at $\vec{r}_0$, cannot be less than the radius of the sphere. A ball moves within a squash court: $|x| \leq w/2$, $|y| \leq l/2$. Non-holonomic constraints cannot generally be dealt with as easily as holonomic constraints.

Physically, constraints are generally maintained by forces, whose values are not known initially, and knowledge of which is not necessarily significant. The solid sphere is in reality an elastic solid, whose atoms resist the inward motion of the particle. The Newtonian approach typically requires these forces be identified and included in the analysis; the analytical approach to dynamics circumvents this need.

Class example

2.3 Generalised Forces

When the cartesian coordinates specifying a dynamical system change by δx_i, work is done:

$$\delta W = \sum_{i=1}^{n} F_i \, \delta x_i,$$

where the F_i are the cartesian components of the forces acting upon the particles: F_1, F_2 and F_3 are the force components acting upon particle 1 at (x_1, x_2, x_3), F_4, F_5 and F_6 act upon particle 2 at (x_4, x_5, x_6) etc.

Since the cartesian components can be expressed in terms of generalised coordinates, Eqn. (1), the changes in cartesian coordinates can be expressed in terms of changes in generalised coordinates, δq_i. In particular we get from Eqn. (1)

$$\delta x_i = \sum_{j=1}^{n} \frac{\partial x_i}{\partial q_j} \delta q_j.$$

So

$$\delta W = \sum_{i=1}^{n} F_i \sum_{j=1}^{n} \frac{\partial x_i}{\partial q_j} \delta q_j$$

which can be reorganised into the expression

$$\delta W = \sum_{j=1}^{n} \mathcal{F}_j \, \delta q_j$$

by introducing as a `generalised force`

$$\mathcal{F}_j = \sum_{i=1}^{n} F_i \frac{\partial x_i}{\partial q_j}. \qquad (3)$$

In the same way that the work required to change cartesian coordinate x_i by δx_i is $F_i \delta x_i$, the work required to change generalised coordinate q_i by δq_i is $\mathcal{F}_i \delta q_i$.

Since q_i need not be a length, $\mathcal{F}_i$ need not have the dimensions of a normal force – for example, if q_i is an angle, $\mathcal{F}_i$ is a torque. Let $[h]$ denote the dimension of quantity h. Then

$$[\mathcal{F}_i] = \frac{M \, L^2}{T^2} \frac{1}{[q_i]}.$$

Dimensional analysis can be a valuable tool in confirming the correctness of expressions (but you knew that already because you are a *physicist*).

Let us consider the case of a `conservative` system. In conservative systems, the work done by forces can be expressed as a perfect differential, independent of the paths followed by the particles and depending

only upon their initial and final positions. A conservative force[3] depends only on position, and can be written as the negative gradient of a potential $\mathcal{V} = \mathcal{V}(x_1, x_2, \ldots x_n, t)$

$$F_i = -\frac{\partial \mathcal{V}}{\partial x_i}.$$

It then follows from Eqn. (3) that

$$\mathcal{F}_j = -\sum_{i=1}^{n} \frac{\partial \mathcal{V}}{\partial x_i}\frac{\partial x_i}{\partial q_j} \equiv -\frac{\partial \mathcal{V}}{\partial q_j},$$

recognising the chain rule.

2.4 Lagrange's equation(s) of motion

We now derive equations of motion with our generalised coordinates – stick with it, the result is worth it.

In cartesians, we have the empirical fact (Newton's second law) that

$$m_i \ddot{x}_i = F_i$$

where m_1, m_2 and m_3 are the mass of particle 1, whose position is (x_1, x_2, x_3), m_4, m_5 and m_6 are the mass of particle 2 etc. Then from Eqn. (3)

$$\mathcal{F}_i = \sum_j F_j \frac{\partial x_j}{\partial q_i} = \sum_j m_j \ddot{x}_j \frac{\partial x_j}{\partial q_i}.$$

But

$$\frac{d}{dt}\left(\dot{x}_j \frac{\partial x_j}{\partial q_i}\right) = \ddot{x}_j \frac{\partial x_j}{\partial q_i} + \dot{x}_j \frac{d}{dt}\frac{\partial x_j}{\partial q_i}$$

so we can write

$$\mathcal{F}_i = \sum_j \left[m_j \frac{d}{dt}\left(\dot{x}_j \frac{\partial x_j}{\partial q_i}\right) - m_j \dot{x}_j \frac{d}{dt}\frac{\partial x_j}{\partial q_i}\right]. \qquad (4)$$

We also have for the kinetic energy

$$\mathcal{T} = \frac{1}{2}\sum_j m_j \dot{x}_j^2$$

so that

$$\frac{\partial \mathcal{T}}{\partial q_i} = \sum_j m_j \dot{x}_j \frac{\partial \dot{x}_j}{\partial q_i}, \quad \text{and} \quad \frac{\partial \mathcal{T}}{\partial \dot{q}_i} = \sum_j m_j \dot{x}_j \frac{\partial \dot{x}_j}{\partial \dot{q}_i}. \qquad (5)$$

---

[3] The gravitational and electrostatic forces are conservative, as is the force due to a spring or the force binding molecules. The velocity-dependent forces of friction or aerodynamic drag, and the force on a charged particle in a magnetic field $q\vec{v}_\times\vec{B}$ are not.

The cartesian coordinates are functions of the generalised coordinates, $x_i = x_i(q_1, q_2, \ldots q_n, t)$, so that

$$\dot{x}_j = \frac{dx_j}{dt} = \sum_i \frac{\partial x_j}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial x_j}{\partial t} = \sum_i \frac{\partial x_j}{\partial q_i}\dot{q}_i + \frac{\partial x_j}{\partial t} \qquad (6)$$

and the cartesian *velocities* are seen to be functions of both the generalised coordinates (through $\partial x_j / \partial q_i$ and $\partial x_j / \partial t$) *and* the generalised velocities $\dot{q}_i = dq_i/dt$. From this last equation we see

$$\frac{\partial \dot{x}_j}{\partial \dot{q}_i} = \frac{\partial x_j}{\partial q_i}, \qquad (7)$$

treating the q_i and $\dot{q}_i$ as independent variables. We get a second useful result by noting that

$$\frac{d}{dt}\frac{\partial x_i}{\partial q_j} = \sum_k \frac{\partial^2 x_i}{\partial q_k \partial q_j}\dot{q}_k + \frac{\partial^2 x_i}{\partial t \partial q_j}$$

and, from Eqn. (6), that

$$\frac{\partial \dot{x}_i}{\partial q_j} = \frac{\partial}{\partial q_j}\left(\sum_k \frac{\partial x_i}{\partial q_k}\dot{q}_k + \frac{\partial x_i}{\partial t}\right) = \sum_k \frac{\partial^2 x_i}{\partial q_j \partial q_k}\dot{q}_k + \frac{\partial^2 x_i}{\partial q_j \partial t}.$$

This is the same as the preceding expression, and so we can equate left hand sides

$$\frac{\partial \dot{x}_i}{\partial q_j} = \frac{d}{dt}\frac{\partial x_i}{\partial q_j}. \qquad (8)$$

Using equations (7) and (8) in (5) gives

$$\frac{\partial \mathcal{T}}{\partial q_i} = \sum_j m_j \dot{x}_j \frac{d}{dt}\frac{\partial x_j}{\partial q_i}, \quad \text{and} \quad \frac{\partial \mathcal{T}}{\partial \dot{q}_i} = \sum_j m_j \dot{x}_j \frac{\partial x_j}{\partial q_i}.$$

Hence (4) is seen to be

$$\mathcal{F}_i = \frac{d}{dt}\frac{\partial \mathcal{T}}{\partial \dot{q}_i} - \frac{\partial \mathcal{T}}{\partial q_i} \qquad (9)$$

which gives the generalised force associated with a given generalised coordinate. For conservative systems this can be written

$$\frac{d}{dt}\frac{\partial \mathcal{T}}{\partial \dot{q}_i} - \frac{\partial \mathcal{T}}{\partial q_i} = -\frac{\partial \mathcal{V}}{\partial q_i}. \qquad (10)$$

Introducing the `Lagrangian`

$$\mathcal{L} = \mathcal{T} - \mathcal{V}, \qquad (11)$$

this can be expressed as

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0 \qquad \text{Lagrange's equation} \qquad (12)$$

since $\partial \mathcal{V}/\,\partial \dot{q}_i = 0$ for conservative systems. Eqn. (10) is sometimes referred to as Lagrange's equation, but more commonly the form in Eqn. (12). These are key results. If we can express the kinetic energy and the potential energy in a system in terms of generalised coordinates, we can use these to immediately write down equations of motion for these coordinates. The first form of Lagrange's equation is often used in solving actual discrete dynamical problems. The second is more appropriate for further theoretical analysis.

2.4.1 Non-conservative forces

If a system also contains non-conservative forces, we can decompose the total generalised force into conservative and non-conservative contributions: $\mathcal{F}_i = \mathcal{F}_i^c + \mathcal{F}_i^{nc}$. Then Lagrange's equation (12) generalises to

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = \mathcal{F}_i^{nc},$$

where $\mathcal{L} = \mathcal{T} - \mathcal{V}$ with $\mathcal{V}$ the potential corresponding to the conservative forces in the system.

2.5 Some examples

To solve problems using Lagrangian dynamics

- Draw a diagram indicating particles/bodies and their motion
- Identify a set of generalised coordinates $q_i,\ i = 1,2,\ldots n$
- Find the kinetic energy $\mathcal{T}$, the potential energy $\mathcal{V}$ and the Lagrangian $\mathcal{L} = T - V$. For $\mathcal{T}$, it is often advisable to first express in cartesians, and then convert to generalised coordinates.
- Evaluate $\partial \mathcal{L}/\,\partial q_i$ and $\partial \mathcal{L}/\,\partial \dot{q}_i$ and apply the Lagrange equation.

Class examples

2.6 Hamilton's principle

The form of Eqn. (12) is suggestive. We now look at formulating our new perspective on dynamics in the framework of variational calculus. We start with a brief recap of some basic elements from the calculus of variations.

2.6.1 Calculus of variations - a brief recap/introduction

Consider a functional $\mathcal{J}[y]$ which depends on the function $y(t)$ through an integral of the form

$$\mathcal{J}[y] = \int_{t_1}^{t_2} f\,(y, \dot{y}, t)dt, \qquad \dot{y} = dy/dt. \tag{13}$$

Inserting different curves $y(t)$ into the integrand and evaluating gives different results: $\mathcal{J}[y]$. We ask ourselves, for what `extremal` curve(s) is the value of the integral stationary, that is, a minimum, a maximum

or a saddle point, when $y(t)$ is allowed to vary over all possible smooth curves that pass through fixed end points at $y(t_1)$ and $y(t_2)$.

To answer this, we consider the change in $\mathcal{J}$ if $y(t) \to y(t) + \delta y(t)$. Treating y and $\dot{y}$ as independent variables,

$$f(y + \delta y, \dot{y} + \delta\dot{y}, t) = f(y, \dot{y}, t) + \frac{\partial f(y, \dot{y}, t)}{\partial y}\delta y + \frac{\partial f(y, \dot{y}, t)}{\partial \dot{y}}\delta\dot{y} + \mathcal{O}(\delta y^2, \delta\dot{y}^2, \delta y\delta\dot{y})$$

where we only consider first order terms explicitly, since they are all that is required to establish the existence of a stationary point. Substituting into Eqn. (13) and simplifying notation in an obvious manner, if $y \to y + \delta y$ then $\mathcal{J} \to \mathcal{J} + \delta\mathcal{J}$ where

$$\delta\mathcal{J} = \int_{t_1}^{t_2} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial \dot{y}}\delta\dot{y}\right) dt.$$

Noting $\delta\dot{y} = d(\delta y)/dt$, we can integrate the second term by parts, getting

$$\left[\frac{\partial f}{\partial \dot{y}}\delta y\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left(\frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right)\delta y dt.$$

The first term vanishes as $\delta y(t_1) = 0 = \delta y(t_2)$ to ensure the curve $y(t) + \delta y(t)$ passes through the fixed end points. Hence

$$\delta\mathcal{J} = \int_{t_1}^{t_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\right)\delta y dt \equiv \int_{t_1}^{t_2} \frac{\delta\mathcal{J}}{\delta y}\delta y dt.$$

The last expression introduces notation that is used to indicate the `variational derivative` or `functional derivative` of the functional $\mathcal{J}$. For $\mathcal{J}$ to be stationary, $\delta\mathcal{J}$ must vanish for all possible δy, which we see requires

$$\frac{\delta\mathcal{J}}{\delta y} = \frac{\partial f}{\partial y} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}} = 0. \qquad \text{Euler-Lagrange equation} \qquad (14)$$

Hence the function $y(t)$ which makes the functional $\mathcal{J}[y]$ in Eqn. (13) stationary is the solution of the Euler-Lagrange equation, named after its co-discoverers Leonhard Euler (Fig.2) and Lagrange.

Figure 2. Leonhard Euler, 1707-1783. A grumpy old man.

2.6.2 Principle of least action

Extension to functionals of more variables $y_i,\ i = 1,2, \ldots n$ is straightforward, as is recognising the similarity between Eqns. (14) and (12). Introducing the `action` functional

$$\mathcal{S} = \int_{t_1}^{t_2} \mathcal{L}\, dt, \qquad \mathcal{L} = T - V, \qquad (15)$$

we deduce that the motion of a mechanical system coincides with extremals of $\mathcal{S}$, i.e.

$$\delta\mathcal{S} = \delta \int_{t_1}^{t_2} \mathcal{L}\, dt = 0 \qquad \text{Hamilton's principle.} \qquad (16)$$

The value of the functional is usually a minimum, leading to the alternative name: Hamilton's *principle of least action*.

We know that in a dynamical system, energy flows between kinetic and potential forms. Kinetic energy $\mathcal{T}$ measures the extent to which things are "happening", the extent to which the system is in motion. Potential energy $\mathcal{V}$ measures the extent to which things *could* happen, hence the name. The Lagrangian $\mathcal{L} = T - V$ is large when $\mathcal{T}$ is large, and small when the energy is in the form of potential. So it quantifies in some sense the activity in the system. The principle of least action is the mathematical statement that nature prefers systems to evolve by means that minimise activity over time. (Maupertuis: "*Nature is thrifty in all its actions*").

We have obtained Eqn. (16) starting from Newton's laws, introducing generalised coordinates and recognising the form of Lagrange's equation. Alternatively, one might consider (16) as reflecting some fundamental statement of Nature, from which Lagrange's equations and Newton's laws follow as consequences. Here, we envisage some system of particles described by a set of generalised coordinates $q_j,\ j = 1, \ldots n$ which evolves along some trajectory in configuration space. Vibrating atoms in a molecule, maybe, or planets in a distant solar system, or some Heath Robinson-esque contraption (Fig. 3) comprised of wheels, rods, pulleys, springs, hemispherical bowls, slopes and ladders leaning against walls. Hamilton's principle $\delta\mathcal{S} = 0$ implies the system somehow "knows" that the trajectory to follow is that which makes the action integral stationary. The motion *locally* ensures that the *global* action be stationary.

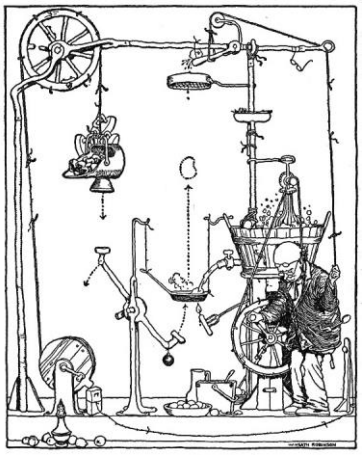

Figure 3: Professor Branestawm's pancake making machine (Heath-Robinson): $\delta\mathcal{S} = 0$.

Note that $[\mathcal{S}] = ML^2T^{-1}$, S.I. units: J s, the same as Planck's constant. The path integral formulation of quantum mechanics developed by Richard Feynman (Fig. 6) is based upon the three postulates

- The probability for an event is given by the squared length of a complex number called the "probability amplitude".
- The probability amplitude is given by adding together the contributions of all paths in configuration space.
- The contribution of a path is proportional to $\exp i\mathcal{S}/\hbar$ where $\mathcal{S}$ is the action given by the time integral of the Lagrangian along the path.

A quantum system is thus seen to explore *all* trajectories, which acquire a phase proportional to the corresponding action (for a ball thrown from A to B, this includes paths that loop the loop, or encircle Jupiter). Summing, there will be interference between contributions from different trajectories, leaving the dominant contributions those for which the phases add constructively. This is when $\delta\mathcal{S} = 0$. See Appendix A.

2.7 Charged particle in an electromagnetic field

Suppose more generally than before (section 2.3), that the components of some generalised force are not simply the negative gradient of some potential but in fact satisfy

$$\mathcal{F}_i = -\frac{\partial\tilde{\mathcal{V}}}{\partial q_i} + \frac{d}{dt}\frac{\partial\tilde{\mathcal{V}}}{\partial\dot{q}_i} \qquad (17)$$

where $\tilde{\mathcal{V}} = \tilde{\mathcal{V}}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)$ is a function of time, the coordinates *and* their derivatives. Then from Eqn. (9)

$$\frac{d}{dt}\frac{\partial\mathcal{T}}{\partial\dot{q}_i} - \frac{\partial\mathcal{T}}{\partial q_i} = \mathcal{F}_i = -\frac{\partial\tilde{\mathcal{V}}}{\partial q_i} + \frac{d}{dt}\frac{\partial\tilde{\mathcal{V}}}{\partial\dot{q}_i}$$

and

$$\frac{d}{dt}\frac{\partial(\mathcal{T}-\tilde{\mathcal{V}})}{\partial\dot{q}_i} - \frac{\partial(\mathcal{T}-\tilde{\mathcal{V}})}{\partial q_i} = 0$$

and we once again have the Lagrange equation with $\mathcal{L} = \mathcal{T} - \tilde{\mathcal{V}}$, but valid for a more general class of systems, those where the forces can be expressed as in Eqn. (17).

Can this really be a useful generalisation? Indeed, the answer is yes, because Eqn. (17) includes the important case of the Lorentz force acting upon a charged particle in an electromagnetic field. Recall that this force is $\vec{F} = q(\vec{E} + \vec{v}\times\vec{B})$ or

$$F_x = qE_x + q(\dot{y}B_z - \dot{z}B_y) \qquad + \quad \text{cyclic permutations} \quad x \mapsto y \mapsto z \mapsto x.$$

Furthermore, according to electromagnetic theory a scalar potential $\varphi(\vec{r}, t)$ and vector potential $\vec{A}(\vec{r}, t)$ exist such that

$$\vec{E} = -\vec{\nabla}\varphi - \frac{\partial\vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla}\times\vec{A}.$$

Now consider $\tilde{V} = q(\varphi - \vec{v}\cdot\vec{A}) = q\varphi - q(\dot{x}A_x + \dot{y}A_y + \dot{z}A_z)$. Clearly

$$-\frac{\partial\tilde{V}}{\partial x} = -q\frac{\partial\varphi}{\partial x} + q\left(\dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x}\right)$$

And

$$\frac{d}{dt}\frac{\partial\tilde{V}}{\partial\dot{x}} = \frac{d}{dt}(-qA_x) = -q\left(\frac{\partial A_x}{\partial t} + \frac{\partial A_x}{\partial x}\dot{x} + \frac{\partial A_x}{\partial y}\dot{y} + \frac{\partial A_x}{\partial z}\dot{z}\right)$$

(since $\vec{A}$ varies both due to the explicit t-dependence and the variation in time of the particle's position). Combining

$$\begin{aligned} -\frac{\partial\tilde{V}}{\partial x} + \frac{d}{dt}\frac{\partial\tilde{V}}{\partial\dot{x}} &= -q\frac{\partial\varphi}{\partial x} - q\frac{\partial A_x}{\partial t} + q\left(\dot{y}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) + \dot{z}\left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}\right)\right) \\ &= -qE_x + q(\dot{y}B_z - \dot{z}B_y) = F_x^{\text{Lorentz}}. \end{aligned}$$

Evidently, the Lorentz force can be written in the form of Eqn. (17), and so the equations of motion of a charged particle in an electromagnetic field are given by the Lagrange equations with the Lagrangian

$$\mathcal{L} = T - q(\varphi - \vec{v}\cdot\vec{A}).$$

3. Symmetry and conservation laws

Symmetry properties are an important aspect of physical systems and recognising them often permits a simplified mathematical description. There is also a fundamental relationship between symmetry and conservation laws, which we explore within the Lagrangian formalism. We will see the basis of Noether's theorem, named after Emmy Noether (Fig. 6).

Figure 4. Emmy Amelie Noether, 1882-1935. Very probably the greatest female mathematician to ever live.

Consider a one-parameter coordinate transformation $q_i(t) \to \tilde{q}_i(\varepsilon, t)$ where ε is a continuous parameter and when $\varepsilon = 0$ the transformation is the identity. Examples might be a translation in some direction $\hat{n}$,

$$\vec{r} \to \vec{r} + \varepsilon\hat{n},$$

or a rotation about an axis $\hat{n}$:[4]

$$\vec{r} \to \vec{r} + \varepsilon\hat{n}_\times\vec{r}.$$

For small ε

$$\tilde{q}_i(\varepsilon) = q_i + \varepsilon\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} + \cdots, \qquad \begin{aligned}\dot{\tilde{q}}_i(\varepsilon) &= \dot{q}_i + \varepsilon\frac{\partial\dot{\tilde{q}}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} + \cdots \\ &= \dot{q}_i + \varepsilon\frac{d}{dt}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} + \cdots\end{aligned}$$

and the Lagrangian to leading order is

$$\begin{aligned}L(\varepsilon) = L\big(\tilde{q}_1(\varepsilon), \dots, \tilde{q}_n(\varepsilon), \dot{\tilde{q}}_1(\varepsilon), \dots, \dot{\tilde{q}}_n(\varepsilon), t\big) &= L(0) + \sum_i\left(\frac{\partial L}{\partial q_i}\epsilon\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} + \frac{\partial L}{\partial\dot{q}_i}\varepsilon\frac{\partial\dot{\tilde{q}}_i}{\partial\varepsilon}\Big|_{\varepsilon=0}\right) + \cdots \\ &= L(0) + \varepsilon\sum_i\left(\frac{d}{dt}\frac{\partial L}{\partial\dot{q}_i}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} + \frac{\partial L}{\partial\dot{q}_i}\frac{d}{dt}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0}\right) + \cdots\end{aligned} \tag{18}$$

where in the last step it has been used that along the path followed by the motion, the Lagrange equation (12) holds, i.e. $d/dt(\partial\mathcal{L}/\partial\dot{q}_i) = \partial\mathcal{L}/\partial q_i$.

Now suppose that the Lagrangian is invariant under the coordinate transformation. This means that $\mathcal{L}(\varepsilon) = \mathcal{L}(0)$, and requires that the last term in Eqn. (18) vanishes. Recognising this as the derivative of a product, we deduce

$$0 = \sum_i\left(\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} + \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{d}{dt}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0}\right) = \frac{d}{dt}\sum_i\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} \quad \Rightarrow \quad \Lambda = \sum_i\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{\partial\tilde{q}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} = \text{const.} \tag{19}$$

So, if $\mathcal{L}$ is invariant with respect to a set of coordinate transformations $\{q_i(t) \to q_i(\varepsilon, t)\}$, then Λ is a constant of motion, and (19) is a conservation law. Λ is often termed a charge, in a more general sense than electric charge.[5]

This is an example of Noether's theorem, which states that *for each continuous symmetry property of the Lagrangian, there corresponds a conservation law, and vice versa*.

---

[4] To recognise this is a rotation, note that the change in $\vec{r}$, $\delta\vec{r} = \varepsilon\hat{n}_\times\vec{r}$, is perpendicular to both $\hat{n}$ and $\vec{r}$. Now draw a diagram. ε is the angle of rotation.

[5] Electric charge can be shown to be the conserved quantity Λ corresponding to invarience with respect to phase transformation $\psi \to \psi e^{i\vartheta}$ in the Lagrangian arising in (complex) quantum field theory.

For example, consider a Lagrangian

$$\mathcal{L} = \frac{1}{2}m|\vec{v}|^2 - U(\vec{r}) \equiv \frac{1}{2}m\sum_i \dot{x}_i^2 - U(x_1, x_2, x_3)$$

describing a particle moving in a potential U. *If* $\mathcal{L}$ is invariant with respect to coordinate transformation $\vec{r} \to \vec{r} + \varepsilon\hat{n}$, i.e. $\partial\tilde{q}_i/\partial\varepsilon \equiv \partial\tilde{x}_i/\partial\varepsilon = \hat{e}_i \cdot \hat{n}$, then

$$\sum_{i=1}^{3} \frac{\partial\mathcal{L}}{\partial\dot{x}_i}\frac{\partial\tilde{x}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} = \sum_{i=1}^{3} m\,\dot{x}_i(\hat{e}_i \cdot \hat{n}) = \hat{n} \cdot \sum_{i=1}^{3} m\,(\dot{x}_i\hat{e}_i) = \hat{n} \cdot m\vec{v} = \text{const.}$$

This is the well-known result that momentum is conserved in a direction of translational invariance — e.g. if U is constant in the z-direction, $m\dot{z}$ is conserved.

Now consider when the Lagrangian is invariant under a rotation: $\partial\tilde{q}_i/\partial\varepsilon \equiv \partial\tilde{x}_i/\partial\varepsilon = \hat{e}_i \cdot (\hat{n}\times\vec{r})$. We then have that

$$\sum_{i=1}^{3} \frac{\partial\mathcal{L}}{\partial\dot{x}_i}\frac{\partial\tilde{x}_i}{\partial\varepsilon}\Big|_{\varepsilon=0} = \sum_{i=1}^{3} m\,\dot{x}_i(\hat{e}_i \cdot (\hat{n}\times\vec{r})) = m\vec{v} \cdot (\hat{n}\times\vec{r}) = \hat{n} \cdot (\vec{r}\times m\vec{v}) = \text{const.}$$

(using $\vec{a} \cdot (\vec{b}\times\vec{c}) = \vec{b} \cdot (\vec{c}\times\vec{a})$). We see that $\hat{n} \cdot (\vec{r}\times\vec{p})$ is a constant. In other words, the component of the angular momentum in the direction $\hat{n}$ is a conserved quantity.

Class example

We have considered here single-particle cases. The symmetry/conservation law relationship holds for the many-particle case too.

3.1 Time invariance

What if $\mathcal{L}$ does not explicitly depend upon time? By this, we mean functionally $\mathcal{L}$ does not contain the variable t, but may still vary with time through the variation in time of q_i, $\dot{q}_i$. It means $\mathcal{L}$ is invariant with respect to time translations $t \to t + \varepsilon$.

So $\mathcal{L} = \mathcal{L}(q_1, \ldots, q_n, \dot{q}, \ldots, \dot{q}_n)$, and $\partial\mathcal{L}/\partial t = 0$. But

$$\frac{d\mathcal{L}}{dt} = \sum_i \left(\frac{\partial\mathcal{L}}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{d\dot{q}_i}{dt}\right).$$

Again, along the classical path followed by the system, the Lagrange equation (12) holds, so $d/dt(\partial\mathcal{L}/\partial\dot{q}_i) = \partial\mathcal{L}/\partial q_i$. Therefore

$$\frac{d\mathcal{L}}{dt} = \sum_i \left(\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right)\dot{q}_i + \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{d\dot{q}_i}{dt}\right) = \sum_i \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\dot{q}_i\right)$$

or

$$\frac{d}{dt}\left[\sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L}\right] = 0 \qquad \Rightarrow \qquad \mathcal{H} = \sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} = \text{const.}$$

The quantity $\mathcal{H}$ is called the *Hamiltonian*; it is a generalised energy. $\mathcal{H}$ is conserved if $\mathcal{L}$ does not change under time translations, i.e., does not explicitly depend upon t.

Class example

3.2 Gauge symmetry

Consider two Lagrangians that differ by the addition of a *total* time derivative of a function of generalised cordinates and time i.e., if $f = f(q_1, q_2, \ldots q_i, t)$ and

$$\mathcal{L}' = \mathcal{L} + \frac{df}{dt}. \qquad (20)$$

Note that f varies with t both through its explicit t-dependence *and* through the t-dependence of the coordinates $\{q_i\}$:

$$\frac{df}{dt} = \sum_i \frac{\partial f}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial f}{\partial t} = \sum_i \frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial t}.$$

Then

$$\mathcal{S}' = \int_{t_1}^{t_2} \mathcal{L}'\, dt = \int_{t_1}^{t_2} \mathcal{L}\, dt + \int_{t_1}^{t_2} \frac{df}{dt}\, dt = \mathcal{S} + f(q_1, \ldots, q_n, t)\big|_{t_1}^{t_2}$$

The additional term added to the action is a constant: Hamilton's principle considers variations in the path between fixed end points, $\delta q_i(t_1) = 0 = \delta q_i(t_2)$, so $\mathcal{S}'$ and $\mathcal{S}$ differ by a constant, and we see that $\mathcal{S}'$ is stationary when $\mathcal{S}$ is stationary. Adding f has no consequential effect on Hamilton's principle of least action.

More explicitly,

$$\mathcal{L}' = \mathcal{L} + \sum_i \frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial t}$$

so

$$\frac{\partial \mathcal{L}'}{\partial q_j} = \frac{\partial \mathcal{L}}{\partial q_j} + \sum_i \frac{\partial^2 f}{\partial q_j\, \partial q_i}\dot{q}_i + \frac{\partial^2 f}{\partial q_j\, \partial t}$$

and

$$\frac{\partial \mathcal{L}'}{\partial \dot{q}_j} = \frac{\partial \mathcal{L}}{\partial \dot{q}_j} + \frac{\partial f}{\partial q_j} \qquad \Rightarrow \qquad \begin{aligned} \frac{d}{dt}\frac{\partial \mathcal{L}'}{\partial \dot{q}_j} &= \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} + \frac{d}{dt}\frac{\partial f}{\partial q_j} \\ &= \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} + \sum_i \frac{\partial^2 f}{\partial q_i\, \partial q_j}\dot{q}_i + \frac{\partial^2 f}{\partial t\, \partial q_j}. \end{aligned}$$

Thus

$$\frac{d}{dt}\frac{\partial \mathcal{L}'}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}'}{\partial q_j} = \left[\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} + \sum_i \frac{\partial^2 f}{\partial q_i \partial q_j}\dot{q}_i + \frac{\partial^2 f}{\partial t \partial q_j}\right] - \left[\frac{\partial \mathcal{L}}{\partial q_j} + \sum_i \frac{\partial^2 f}{\partial q_j \partial q_i}\dot{q}_i + \frac{\partial^2 f}{\partial q_j \partial t}\right] = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0$$

on the classical path followed by the motion. Both $\mathcal{L}'$ and $\mathcal{L}$ satisfy the same Lagrange equation of motion.

The transformation in Eqn. (20), which leaves the dynamics unaffected, is called a gauge transformation. An important example of this arises in the case of the Lagrangian for a charged particle in an electromagnetic field.

3.3 The Lagrangian from symmetry arguments

We have seen that Newtonian mechanics follows from the stationary principle applied to the action

$$\mathcal{S} = \int \mathcal{L}dt = \int \mathcal{L}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)dt$$

which gives rise to equations of motion

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i}.$$

This stationary principle means that the dynamics are invariant to changes in $\mathcal{L}$ in the form that take the form of a multiplicative constant, and/or by the addition of the total time derivative of a function of coordinates and time. Explicitly, if

$$\mathcal{L}' = \gamma\mathcal{L} + \frac{dg}{dt}$$

where $g = g(q_1, \dots, q_n, t)$ then (see Section 3.2)

$$\mathcal{S}' = \int_1^2 \mathcal{L}'dt = \gamma\int_1^2 \mathcal{L}\,dt + \int_1^2 \frac{dg}{dt}dt = \gamma\mathcal{S} + g_2 - g_1.$$

g_1 and g_2 are fixed by initial and final conditions, and have no influence on the dynamics between them, whilst $\gamma\mathcal{S}$ is stationary when $\mathcal{S}$ is.

Other symmetries follow from our observations of Nature. Consider the Lagrangian describing a free particle, $\mathcal{L}_0 = \mathcal{L}_0(\vec{r}, \vec{v}, t)$. Free space is *isotropic*. The motion cannot depend upon the orientation of our coordinate system. Thus $\mathcal{L}_0$ cannot depend upon the direction of the velocity, but only its magnitude. For simplicity we write

$$\mathcal{L}_0 = \mathcal{L}_0(\vec{r}, v^2, t).$$

Space-time is *homogeneous*. The motion cannot depend upon from where we choose to measure it, or from when we choose to time it. Thus

$$\mathcal{L}_0 = \mathcal{L}_0(v^2).$$

Immediately we see from the Lagrange equation

$$\frac{d}{dt}\frac{\partial \mathcal{L}_0}{\partial v_i} = \frac{\partial \mathcal{L}_0}{\partial x_i} \qquad \rightarrow \qquad \frac{d}{dt}\left(\frac{d\mathcal{L}_0}{d(v^2)}.2v_i\right) = 0$$

and hence v_i is a constant. A free particle moves at a constant speed. We have deduced the *Law of Inertia* from the stationary principle and the symmetry of space-time.

Nature presents further symmetry. At least at low speeds, we observe Galilean relativity (Fig. 5). If a free particle is observed to be moving at a velocity $\vec{v}$ by a stationary observer, then the same particle observed by an observer from a vantage point moving at velocity $\vec{u}$ is $\vec{v} - \vec{u}$. This further constrains the Lagrangian. For the stationary observer, $\mathcal{L} = \mathcal{L}_0(v^2)$. For the moving observer

$$\mathcal{L}'_0 = \mathcal{L}_0(|\vec{v} - \vec{u}|^2).$$

Assume u is infinitesimal, so

$$\mathcal{L}'_0 \simeq \mathcal{L}_0(v^2) - 2\vec{u} \cdot \vec{v} \frac{\partial \mathcal{L}_0}{\partial v^2}.$$

Galilean relativity requires this extra term must give the same particle motion, which requires it to be at most a total time derivative of function of $\vec{r}$ and t. Since $\vec{v} = d\vec{r}/dt$ is itself a total time derivative, it must be that

$$\frac{\partial \mathcal{L}_0}{\partial v^2} = \text{const.} \qquad \Rightarrow \qquad \mathcal{L}_0(v^2) = \alpha v^2 + \beta.$$

With this form, for non-infinitesimal u

$$\mathcal{L}'_0(|\vec{v} - \vec{u}|^2) = \alpha(v^2 - 2\vec{v}\vec{u} + u^2) + \beta = \alpha v^2 + \beta + \frac{d}{dt}(tu^2 + 2\vec{r} \cdot \vec{u})$$

where the extra term is a total time derivative and so has no effect on the motion, as required.

The additive constant β has no effect on equations of motion derived from the Lagrangian, since Lagrange's equation only contains derivatives. The prefactor α appears arbitrary, but if we now consider a system of two free particles, each with its own Lagrangian $\mathcal{L}_1$ and $\mathcal{L}_2$, then the simultaneous motion of the particles comes from the Lagrangian

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2.$$

and we see the ratio of the two prefactors is *not* arbitrary. We can define the prefactor for one particle in the Universe, and the prefactors for the rest then follow. An overall factor remains arbitrary, which is equivalent to choosing a unit in which to measure α for the reference particle. We choose to write $\alpha = m/2$, take it to be positive (so that our stationary principle is a minimum principle, and call m mass. We use units such that $m = 1$ for a reference particle,[6] which we take to be a right-circular cylinder of PtIr alloy of height 39.17 mm, stored in Sèvres, France[7]. Then the free Lagrangian, from symmetry principles, is

$$\mathcal{L}_0 = \frac{1}{2} m v^2.$$

This process of using symmetry to constrain and indeed pin down the Lagrangian works more generally, and indeed all fundamental *interactions* can be derived from symmetry principles.

---

[6] Clearly the particle here is a body. See footnote 2.

[7] Or did. In a recent change, the definition of the kilogram now in terms of Planck's constant.

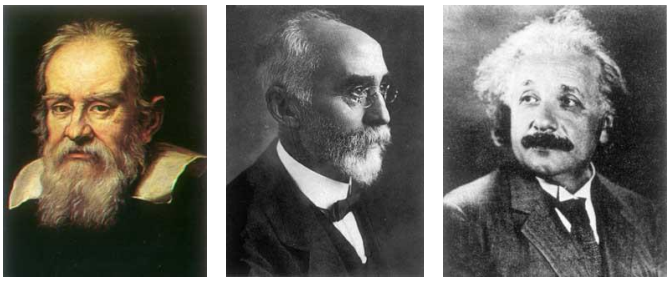

Figure 5: L-R: Galileo Galilei, 1564-1642. Hendrik Antoon Lorentz, 1853-1928. Albert Einstein, 1879-1955. Yet more grumpy old men.

3.4 Relativistic Lagrangian

In deriving $\mathcal{L}_0 = (1/2)mv^2$ we used the fact that the laws of nature are the same to all inertial observers, and demanded that under a Galilean transformation relating the inertial frames in which two observers are located, the equations of motion from $d/dt(\partial\mathcal{L}_0/\partial\dot{x}_i) = \partial\mathcal{L}_0/\partial x_i$ have the same form. We assumed in doing so that the time variable appearing in the Lagrange equation is the same in both frames. Time is universal.

Special relativity tells us this is not true: time is not invariant but differs for observers in different frames. Numerous observations tell us this, such as the observation that muons from cosmic rays reach the surface of the Earth. In the laboratory stationary muons decay in 2.2 μs, and so energetic muons from cosmic rays ($v \simeq 0.98c$) should travel ~650 metres through the atmosphere. In fact, they are observed to travel much further. The 2.2μs that passes in the rest frame of the muon is dilated to something much longer to an observer whose vantage point is on the rapidly moving (to the muon) Earth. Moving clocks run slower. The correct description of motion at speeds close to the speed of light c is one of the major achievements of Albert Einstein (Fig. 5). The transformations relating distances and times separating events as seen by observers if reference frames move at relative speeds close to c are the Lorentz transformations (Fig. 5).

So can we derive a Lagrangian for a relativistic free particle, using symmetry? The only (?) scalar quantity that is relativistically-invariant, for which all observers irrespective of frame of reference agree, is the space-time interval between events 1 and 2:

$$s_{21}^2 = c^2(t_2 - t_1)^2 - |\vec{r}_2 - \vec{r}_1|^2 \qquad s_{12} > 0.$$

All observers agree light travels at the same speed, c, although how far it went or for how long it was travelling is up for debate. If the action $\mathcal{S}$ is formulated in terms of the interval, all observers should draw similar conclusions regarding it.

Assuming Nature keeps things simple, we might consider then the simplest possible relativistic action:

$$\mathcal{S} = \alpha \int_1^2 ds,$$

where $ds = \sqrt{c^2 dt^2 - \sum_i (dx_i)^2}$ is the differential space-time interval, and α a constant. In a given frame, we can parameterise the integral by the proper time variable between when the two events are observed in that frame,

$$\mathcal{S} = \alpha c \int_{t_1}^{t_2} \sqrt{1 - v^2/c^2}\, dt, \qquad \text{with} \quad \frac{dx_i}{dt} = v_i, \quad \sum_i v_i^2 = v^2.$$

Comparing with $\mathcal{S} = \int_{t_1}^{t_2} \mathcal{L}\, dt$ we identify

$$\mathcal{L}_0^{rel} = \alpha c\sqrt{1 - v^2/c^2}.$$

In the low velocity limit, $\mathcal{L}_0^{rel} \to \alpha c(1 - v^2/(2c^2) + \cdots)$ and the requirement that Newtonian mechanics be recovered in that limit indicates $\alpha = -mc$. Thus

$$\mathcal{L}_0^{rel} = -mc^2\sqrt{1 - v^2/c^2}.$$

The equations of motion derived from this Lagrangian (plus extension to the case when electromagnetic fields are present) are indeed consistent with the dynamics of fast-moving particles. And we have derived it from space-time symmetry alone.

Appendix

A. From path integrals to the Schrodinger equation

According to Feynman (Fig. 6), the conditional probability amplitude of finding in final state 2 a system that was initially in state 1 is

$$K(2|1) = \sum_j \phi\,[X_j]$$

where the sum is in practice an integral over paths $X(t)$ from 1 to 2, and the probability amplitude associated with path $X(t)$ is

$$\phi[X] \propto \exp i\mathcal{S}[X]/\hbar$$

with

$$\mathcal{S}[X] = \int_{t_1}^{t_2} \mathcal{L}\,(X, \dot{X}, t)dt,$$

the *classical* action for path $X(t)$.

Figure 6: L-R: Richard Phillips Feynman, 1918-1988. Erwin Rudolf Joseph Alexander Schrödinger, 1887-1961. Two grumpy old men.

What is K? Let us consider a system in which a particle moves in a potential $V(x,t)$ and consider $1=(x_1,t_2)$ and $2=(x_2,t_2)=(x_1+\Delta x,t_1+\Delta t)$ to be infinitesimally close. In moving between 1 and 2 a classical particle will do so at velocity $\dot{x}=\Delta x/\Delta t$. Expanding V about the midpoint $V(x_1+\Delta x/2,t_1+\Delta t/2)$ then gives

$$\mathcal{S}[x]=\frac{m(\Delta x)^2}{2\Delta t}-V\left(x_1+\frac{\Delta x}{2},t_1+\frac{\Delta t}{2}\right)\Delta t$$

as the action for the classical path. In the classical limit ($\hbar\to 0$), paths that deviate significantly from the classical path $x(t)$ give changes $\delta\mathcal{S}$ for which $\delta\mathcal{S}/\hbar$ is large. Such paths result in highly oscillatory contributions to the sum over paths in the conditional probability K, which tend to average to zero.[8] On the other hand paths very close to the classical path deviate from $\mathcal{S}[x]$ by an amount $\delta\mathcal{S}\simeq 0$ to first order. So we see it is the paths for which $\delta\mathcal{S}$ is within $\hbar$ of the classical action that will contribute significantly. In the strict limit $\hbar\to 0$ the only path making a non-vanishing contribution to the probability amplitude will be the classical path.

Inserting $\mathcal{S}[x]$ we can therefore write as an approximation for K for initial and finial states that are infinitesimally close as

$$K(x_1+\Delta x,t_1+\Delta t|x_1,t_1)=\frac{1}{A}\exp\frac{i}{\hbar}\left[\frac{m(\Delta x)^2}{2\Delta t}-V\left(x_1+\frac{\Delta x}{2},t_1+\frac{\Delta t}{2}\right)\Delta t\right].$$

where A normalises.

We call the probability amplitude for a particle to be at position x at time t the quantum wave function $\psi(x,t)$. If we know ψ for all positions at some time t, we can use the conditional probability K to determine ψ at some later time $t+\epsilon$:

$$\psi(x,t+\epsilon)=\int_{-\infty}^{+\infty}dx_1\,K(x,t+\epsilon|x_1,t)\psi(x_1,t)=\int_{-\infty}^{+\infty}d\eta\,K(x,t+\epsilon|x+\eta,t)\psi(x+\eta,t),$$

where in the second form we have chosen to express x_1 relative to x as $x_1=x+\eta$. Taking ϵ to be infinitesimal we can use our expression for K above, giving

$$\psi(x,t+\epsilon)=\int_{-\infty}^{+\infty}d\eta\,\frac{1}{A}\exp\left[\frac{i}{\hbar}\left(\frac{m\eta^2}{2\epsilon}\right)-\epsilon V\left(x+\frac{\eta}{2},t+\frac{\epsilon}{2}\right)\right]\psi(x+\eta,t). \qquad (21)$$

To make progress, we look to expand to first order in ϵ. In the following we use notation $\psi_t=\partial\psi/\partial t$, $\psi_x=\partial\psi/\partial x$ and $\psi_{xx}=\partial^2\psi/\partial x^2$. Then on the left-hand side we have

$$\psi(x,t+\epsilon)=\psi(x,t)+\epsilon\psi_t(x,t).$$

In the integrand on the right-hand side,

$$\exp\left[\left(\frac{im}{2\hbar}\right)\frac{\eta^2}{\epsilon}\right]$$

[8] Mathematically, this is the method of stationary phase/steepest descent at work – see texts such as Arfken.

will oscillate rapidly except for values of η for which $m\eta^2/2\hbar\epsilon$ is less than 1 or so, i.e., a range of η that scales like $\epsilon^{1/2}$. So expanding up to η^2, to be consistent with a first order expansion in ϵ,

$$\psi(x+\eta,t) = \psi(x,t) + \eta\psi_x(x,t) + (\eta^2/2)\psi_{xx}(x,t) + \mathcal{O}(\eta^3)$$

and further Taylor expansions give

$$\exp\left[-\frac{i\epsilon}{\hbar}V(x+\eta/2, t+\epsilon/2)\right] = 1 - \frac{i\epsilon}{\hbar}V(x,t) + \mathcal{O}(\epsilon\eta, \epsilon^2).$$

Introducing $a = m/(2i\hbar\epsilon)$ we then have that to first order on either side

$$\begin{aligned}\psi + \epsilon\psi_t &= \frac{1}{A}\int d\eta\, e^{-a\eta^2}\left(1 - \frac{i\epsilon}{\hbar}V\right)\left(\psi + \eta\psi_x + \frac{\eta^2}{2}\psi_{xx}\right) \\ &= \frac{1}{A}\sqrt{\frac{\pi}{a}}\left(1 - \frac{i\epsilon}{\hbar}V\right) + 0 + \frac{\sqrt{\pi}}{4Aa\sqrt{a}}\psi_{xx}\end{aligned}$$

using standard results $\int \exp(-a\eta^2)d\eta = \sqrt{\pi/a}$, $\int \eta\exp(-a\eta^2)d\eta = 0$ and $\int \eta^2\exp(-a\eta^2)d\eta = (1/2a)\sqrt{\pi/a}$. Equating coefficients of ϵ (recall $a \propto \epsilon^{-1}$):

$$\begin{aligned}\epsilon^0: \quad \psi &= \frac{1}{A}\sqrt{\frac{\pi}{a}}\psi \\ \epsilon^1: \quad \psi_t &= -\frac{1}{A}\sqrt{\frac{\pi}{a}}\frac{i}{\hbar}V\psi + \frac{1}{4Aa}\sqrt{\frac{\pi}{a}}\psi_{xx}\end{aligned}$$

Thus $A = \sqrt{\pi/a}$, and $\psi_t = -i/\hbar V\psi + (i\hbar/2m)\psi_{xx}$.

This last equation relates the temporal and spatial changes in the quantum wave function, assumed to evolve according to a conditional probability amplitude $K(2|1)$ which is given by a path integral over paths $X(t)$ from 1 to 2 which contribute with probability $\phi[X]$ containing an action-dependent phase: $\phi[X] \propto \exp i\mathcal{S}[X]/\hbar$. It is more commonly written as

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi \qquad \text{The Schrödinger equation (Fig. 6).}$$

4. Further reading

Classical mechanics is a classic topic in mathematical physics – you will find many good textbooks on the subject in the library under class mark 531.01 plus elsewhere. The following have good reputations:

Landau, L D and Lifshitz, E M, *Mechanics (Vol. 1, Course of Theoretical Physics)*, Elsevier, Butterworth-Heinemann, Oxford, 1976. [3rd ed.]

Kibble, T W B and Berkshire, F H, *Classical mechanics*, Longman, Harlow, 1996. [4th ed.]

Goldstein H, Poole C and Safko J, *Classical Mechanics*, Addison Wesley, San Francisco, 2002. [3rd ed.]

Classical mechanics of particles and fields — pt 2

5. Hamiltonian dynamics

The Lagrange equations for a dynamical system containing n degrees of freedom represent a set of n second-order differential equations based upon the Lagrangian $\mathcal{L} = \mathcal{L}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)$, a function of generalised coordinates, velocities, and time. The evolution of the system is considered in the n-dimensional configuration space.

We now use Hamilton's principle to develop an alternative description of the dynamics. In particular, we will obtain a set of $2n$ first-order differential equations based upon the Hamiltonian $\mathcal{H} = \mathcal{H}(q_1, \dots, q_n, p_1, \dots, p_n, t)$, a function of generalised coordinates, *generalised momenta*, and time. In Hamiltonian dynamics, the evolution of the system is considered in a $2n$-dimensional space called *phase space*.

5.1 Generalised momenta

The Lagrangian $\mathcal{L} = T - V$ has dimensions of energy: $[\mathcal{L}] = M\,L^2\,T^{-2}$. If q_i is a length, then $\partial\mathcal{L}/\,\partial\dot{q}_i$ has dimensions of linear momentum. If q_i is an angle, $\partial\mathcal{L}/\,\partial\dot{q}_i$ has dimensions of angular momentum. For this reason, we can consider the derivative $\partial\mathcal{L}/\,\partial\dot{q}_i$ as a `generalised momentum`

$$p_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i}, \qquad (22)$$

also referred to as the momentum `conjugate` to coordinate q_i. The Lagrange equation then reads[9]

$$\frac{dp_i}{dt} = \frac{\partial\mathcal{L}}{\partial q_i}.$$

We note that if the Lagrangian for some system is found to be independent of some generalised coordinate q_i, then from Eqn. (22)

$$\frac{dp_i}{dt} = \frac{\partial\mathcal{L}}{\partial q_i} = 0 \qquad \Rightarrow \qquad p_i = \text{const.}$$

In this case the coordinate q_i is refered to as an `ignorable coordinate`, or a `cyclic coordinate`. This is something of a misnomer: the presence of an ignorable coordinate is something to take particular note of, as it reveals a `constant of motion`. Recognising the existence of a constant of motion often opens up a deeper understanding of a system.

[9] The obvious similarity to Newton's equation of motion $dp/dt = F$ means that $\partial\mathcal{L}/\,\partial q_i$ can also be considered as a 'generalised force' but be aware it is not the same as the generalised force introduced previously that relates a change in generalised coordinate δq_i to the work done. $\partial\mathcal{L}/\,\partial q_i$ contains the contribution from the potential, $\partial\mathcal{V}/\,\partial q_i$, but also $\partial\mathcal{T}/\,\partial q_i$ arising from any curvilinear nature of the coordinate system.

5.2 Hamilton's equations

We define the `Hamiltonian` $\mathcal{H}$ as

$$\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L} \qquad (23)$$

or more fully[10]

$$\mathcal{H}(q_1, \dots, q_n, p_1, \dots, p_n, t) = \sum_i p_i \dot{q}_i - \mathcal{L}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t).$$

Note the arguments. It is assumed that the velocities $\dot{q}_i$, $i = 1,2, \dots, n$ in the right-hand side have been substituted for in terms of the positions and momenta by making use of the n equations $p_i = \partial L / \partial \dot{q}_i$, $i = 1,2, \dots, n$. We will assume this can always be done — see graduate texts for proof (*Legendre transformation*).

So if $\mathcal{H} = \sum_i p_i \dot{q}_i - \mathcal{L}$, Hamilton's principle of least action can be written

$$\begin{aligned} \delta\mathcal{S} &= \delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - \mathcal{H} \right) dt = 0, \qquad \dot{q}_i = \dot{q}_i(q_1, \dots, q_n, p_1, \dots, p_m, t) \\ &= \int_{t_1}^{t_2} \left(\sum_i (\delta p_i \dot{q}_i + p_i \delta \dot{q}_i) - \sum_i \left(\frac{\partial \mathcal{H}}{\partial p_i} \delta p_i + \frac{\partial \mathcal{H}}{\partial q_i} \delta q_i \right) \right) dt. \end{aligned}$$

We integrate the second term by parts

$$\int p_i \delta \dot{q}_i dt = \int p_i \frac{d(\delta q_i)}{dt} dt = [p_i \delta q_i] - \int \delta q_i \frac{dp_i}{dt} dt,$$

and by noting $\delta q_i(t_1) = 0 = \delta q_i(t_2)$ so that the contribution from the ends of the integration range vanish, we obtain

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \sum_i \left[\left(\dot{q}_i - \frac{\partial \mathcal{H}}{\partial p_i} \right) \delta p_i - \left(\dot{p}_i + \frac{\partial \mathcal{H}}{\partial q_i} \right) \delta q_i \right] dt = 0.$$

The variations δp_i and δq_i are arbitrary — hence the following must hold:

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}. \qquad (24)$$

These $2n$ first-order differential equations are `Hamilton's equations`.

In this Hamiltonian-dynamics formulation, the system is described by n generalised coordinates and n conjugate momenta, with the $2n$ first-order Hamilton's equations describing the time-evolution of the system. The solution is a trajectory in a $2n$-dimensional space called `phase space` – see Fig. 7.

[10] The quantity $\mathcal{H}$ was introduced earlier in the context of time-invariant Lagrangians. Here we consider it to be a function of generalised coordinates and conjugate momenta, not generalised coordinates and generalised velocities as before.

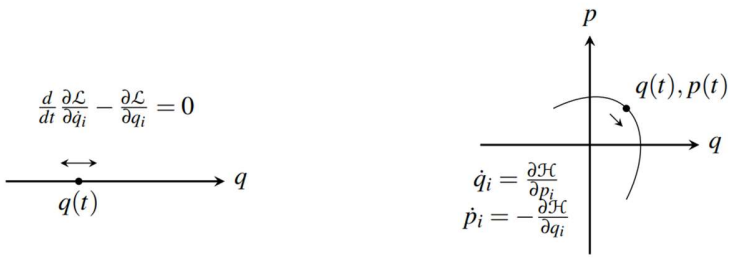

Figure 7: Configuration space versus phase space for a system with n = 1 degrees of freedom. L: n-dimensional configuration space; the dynamics are governed by n second-order Lagrange equations. R: 2n-dimensional phase space; the dynamics are governed by 2n first-order Hamilton's equations.

Immediately we see an advantage of the Hamiltonian formalism: if $\mathcal{L}$ (and hence $\mathcal{H}$) does not depend explicitly upon some coordinate q_i, it follows from Eqn. (24) that $\dot{p}_i = 0$ so p_i =constant (conserved) and the effective number of degrees of freedom of the system is reduced by one.

5.3 Constructing $\mathcal{H}$ via $\mathcal{L}$

Given a Lagrangian $\mathcal{L}$ in terms of generalised coordinates $q_i, i = 1, \dots, n$ and generalised velocities $\dot{q}_i, i = 1, \dots, n$, we can construct the Hamiltonian $\mathcal{H}$ as follows:

- Using $p_i = \partial\mathcal{L}/\,\partial\dot{q}_i$ obtain conjugate momenta as functions of the q_i, $\dot{q}_i$, and possibly t.
- Construct $\mathcal{H} = \sum_i p_i\,\dot{q}_i - \mathcal{L}$. At this stage $\mathcal{H}$ is a mixed expression containing various q_i's, $\dot{q}_i$'s, p_i's and t.
- Invert the $p_i = \partial\mathcal{L}/\,\partial q_i$ relations to obtain the $\dot{q}_i$ as functions of the q_i's, p_i's and t.
- Substitute for the $\dot{q}_i$'s in $\mathcal{H}$, leaving it solely in terms of q_i's, p_i's, and possibly t.

Class examples:

To be honest, few problems are more easily solved explicitly using $\mathcal{H}$ rather than $\mathcal{L}$; the true power of the Hamiltonian reformulation is its application to systems involving extremely large numbers of degrees of freedom (→ statistical mechanics), systems where the action is of the order of $\hbar$ (→ quantum mechanics) and the establishing of general results such as Liouville's theorem, important among other things in the description of chaotic dynamics.

5.4 Liouville's theorem

Consider a 1D system describing a particle moving in some potential. Then $\mathcal{H} = \mathcal{H}(q, p, t)$. The initial conditions (q_0, p_0) plus Hamilton's equations uniquely determine the subsequent path of the system through phase space, since $\dot{q} = \partial\mathcal{H}/\,\partial p$ and $\dot{p} = -\partial\mathcal{H}/\,\partial q$ determine how the path flows away from any instantaneous state (q, p).

To leading order in small time interval δt, the system evolves from (q_0, p_0) to

$$(q_0 + \dot{q}\delta t, p_0 + \dot{p}\delta t) = \left(q_0 + \frac{\partial\mathcal{H}(q_0, p_0)}{\partial p}\delta t, p_0 - \frac{\partial\mathcal{H}(q_0, p_0)}{\partial q}\delta t\right) \qquad (25)$$

— see Fig. 8L. Now consider the set of initial values within the range $q_0 \le q \le q_0 + \Delta q, p_0 \le p \le p_0 + \Delta p$, Fig. 8R. These occupy an area $\mathcal{A} = \Delta q \Delta p$ in phase space. After time δt, they will have evolved. Where to?

Using Eqn. (25), after time δt the corner trajectory $(q_0 + \Delta q, p_0)$ flows to

$$\left(q_0 + \Delta q + \frac{\partial \mathcal{H}(q_0 + \Delta q, p_0)}{\partial p}\delta t, p_0 - \frac{\partial \mathcal{H}(q_0 + \Delta q, p_0)}{\partial q}\delta t\right)$$

and the edge vector $\vec{a}$, the difference in this location and that which the state (q_0, p_0) evolved to, Eqn. (25), is

$$\begin{aligned}\vec{a} = (a_q, a_p) &= \left(\Delta q + \left(\frac{\partial \mathcal{H}(q_0 + \Delta q, p_0)}{\partial p} - \frac{\partial \mathcal{H}(q_0, p_0)}{\partial p}\right)\delta t, -\left(\frac{\partial \mathcal{H}(q_0 + \Delta q, p_0)}{\partial q} - \frac{\partial \mathcal{H}(q_0, p_0)}{\partial q}\right)\delta t\right) \\ &= \left(\Delta q\left(1 + \frac{\partial^2 \mathcal{H}}{\partial q\,\partial p}\delta t\right), -\frac{\partial^2 \mathcal{H}}{\partial q^2}\Delta q \delta t\right)\end{aligned}$$

to leading order in Δq. Similarly (check!)

$$\vec{b} = (b_q, b_p) = \left(\frac{\partial^2 \mathcal{H}}{\partial p^2}\Delta p \delta t, \Delta p\left(1 - \frac{\partial^2 \mathcal{H}}{\partial p\,\partial q}\delta t\right)\right).$$

The area spanned by $\vec{a}$ and $\vec{b}$ is $|\vec{a}_\times \vec{b}| = |a_q b_p - a_p b_q|$, and so the new area spanned by the trajectories is

$$\mathcal{A} + \delta A = \Delta q \Delta p\left[\left(1 + \frac{\partial^2 \mathcal{H}}{\partial q\,\partial p}\delta t\right)\left(1 - \frac{\partial^2 \mathcal{H}}{\partial p\,\partial q}\delta t\right) + \frac{\partial^2 \mathcal{H}}{\partial q^2}\frac{\partial^2 \mathcal{H}}{\partial p^2}\delta t^2\right]$$

and

$$\delta \mathcal{A} = \Delta q \Delta p\left(\frac{\partial^2 \mathcal{H}}{\partial q\,\partial p} - \frac{\partial^2 \mathcal{H}}{\partial p\,\partial q}\right)\delta t + \mathcal{O}(\delta t)^2.$$

Dividing by δt and taking the limit $\delta t \to 0$

$$\frac{d\mathcal{A}}{dt} = \lim_{\delta t \to 0}\frac{\delta \mathcal{A}}{\delta t} = \Delta q \Delta p\left(\frac{\partial^2 \mathcal{H}}{\partial q\,\partial p} - \frac{\partial^2 \mathcal{H}}{\partial p\,\partial q}\right) = 0$$

using the fact that the partial derivatives commute.

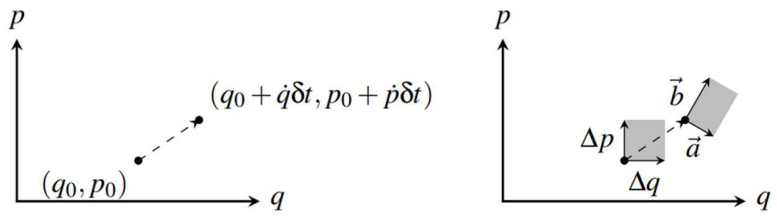

Figure 8: Left (L): In time δt, the system evolves from (q_0, p_0) to $(q_0 + \dot{q}\,\delta t, p_0 + \dot{p}\delta t)$. Right (R): In time δt, the trajectories in the initial rectangular phase space area ΔqΔp evolve to the area spanned by vectors $\vec{a}$ and $\vec{b}$. For a Hamiltonian system, the area is constant.

A similar result can be shown to hold in the case of systems with more than one generalised coordinate/degree of freedom, when the phase space *volume* of trajectories is shown to be conserved, and is known as Liouville's theorem: *as (representative) states of a system evolve in time, their density in phase space remains constant*. They flow like an incompressible fluid.

Liouville's theorem is key to the fluctuation dissipation theorem of statistical mechanics, from which the second law of thermodynamics can be derived.

5.5 Poisson brackets

A Poisson bracket is an important mathematical operation which plays an important role e.g. in celestial mechanics, when deriving *Lagrange's planetary equations* for *osculating orbital elements* due to deviations from two-body interactions, and in quantum mechanics, where they are intimately related to commutators.

Let $f = f(q_1, q_2, \dots, q_n, p_1, p_2, \dots, p_n, t)$ be *any* quantity of a system described by the conjugate variables $q_1, \dots, q_n, p_1, \dots, p_n$. Then

$$\frac{df}{dt} = \sum_i \left(\frac{\partial f}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} \right) + \frac{\partial f}{\partial t}.$$

Using Hamilton's equations, this becomes

$$\frac{df}{dt} = \sum_i \left(\frac{\partial f}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right) + \frac{\partial f}{\partial t}.$$

We define the `Poisson bracket` of any two such dynamical quantities f and g as[11]

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \qquad (26)$$

It turns out that pairs of functions for which $\{f, g\} = 0$, which are said to `commute`, are of special importance. And that pairs of functions for which $\{f, g\} \neq 0$, which do not commute, are also of special importance.

Using our newly introduced Poisson bracket, we can write

$$\frac{df}{dt} = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t}. \qquad (27a)$$

It is clear from the definition of the Poisson bracket that $\{f, g\} = -\{g, f\}$, and so equivalently

$$\frac{df}{dt} = -\{\mathcal{H}, f\} + \frac{\partial f}{\partial t}. \qquad (27b)$$

---

[11] Some authors use [,]; others { , }. I was taught to use { , }, so I teach { , }. Using [,] may cause confusion with the quantum mechanical commutator, to which the Poisson bracket is intimately connected, although there will be no such confusion here. What everyone agrees on, though, is that they are called Poisson brackets. Not Poisson braces. Even if your preferred notation is { , }. Note also that some authors swap the two terms in (26), in effect defining $\{ , \}_{\text{them}} = -\{ , \}_{\text{here}}$.

We note that if f does not contain t explicitly, so $\partial f / \partial t = 0$, it is sufficient for its Poisson bracket with the Hamiltonian to vanish for it to be a constant of motion. This is independent of whether $\mathcal{H}$ is constant or not.

Let us consider some special cases. We have $f = f(q_1, q_2, \dots, q_n, p_1, p_2, \dots, p_n, t)$. Then since $q_i = q_i$, we recognise q_i does not *explicitly* depend upon t (but of course it does, in general, vary in time!). Similarly, p_i does not explicitly depend upon t. Then[12]

$$\dot{q}_i = \{q_i, \mathcal{H}\}, \qquad \dot{p}_i = \{p_i, \mathcal{H}\}, \tag{28}$$

and we see that the generalised coordinates and conjugate momenta do not commute with the Hamiltonian unless constant in time. These two equations are Hamilton's equations in Poisson bracket form.

What about Poisson brackets not involving the Hamiltonian? Consider a system described by two generalised coordinates and momenta: q_1, q_2, p_1, p_2. Then explicitly

$$\{f, g\} = \frac{\partial f}{\partial q_1}\frac{\partial g}{\partial p_1} - \frac{\partial f}{\partial p_1}\frac{\partial g}{\partial q_1} + \frac{\partial f}{\partial q_2}\frac{\partial g}{\partial p_2} - \frac{\partial f}{\partial p_2}\frac{\partial g}{\partial q_2}.$$

So

$$\begin{array}{rcccccccc}
\{q_1, q_2\} & = & 0 & - & 0 & + & 0 & - & 0 & = & 0 \\
\{p_1, p_2\} & = & 0 & - & 0 & + & 0 & - & 0 & = & 0 \\
\{q_1, p_2\} & = & 0 & - & 0 & + & 0 & - & 0 & = & 0 \\
\{p_1, q_2\} & = & 0 & - & 0 & + & 0 & - & 0 & = & 0 \\
\{q_1, p_1\} & = & 1 & - & 0 & + & 0 & - & 0 & = & 1 \\
\{q_2, p_2\} & = & 0 & - & 0 & + & 1 & - & 0 & = & 1.
\end{array}$$

More succinctly, and generalising to more than two generalised coordinates/momenta,

$$\{q_i, q_j\} = 0 = \{p_i, p_j\} \qquad \text{and} \qquad \{q_i, p_j\} = \delta_{i,j} = -\{p_j, q_i\} \tag{29}$$

with $\delta_{i,j}$ the Kronecker delta symbol.

We have previously noted (Sec. 5.1) that when $\dot{p}_i = 0$ we have an ignorable, or cyclic coordinate, and p_i is a constant of motion. One approach to dynamics is to seek transformations amongst the independent variables, turning the set of q_i's and p_i's into q'_i's and p'_i's, such that one of the $\dot{p}'_i = 0$. Each time we do so, we identify a constant of motion, and we reduce the dimensionality of the remaining problem by 1 degree of freedom. Do this often enough, and the problem becomes solved. This is a motivation behind *transformation theory*. Eqn. (29) shows that we should be careful when we do this. These relations *must* be satisfied by *any* set of coordinates and momenta specifying the system, including those that we transform to. Take for example, the angular momentum components $l_x = yp_z - zp_y$, $l_y = zp_x - xp_z$ and $l_z = xp_y -$

---

[12] For an explicit demonstration of this, consider a system described by a single generalised coordinate and its conjugate momentum, (q, p). In this case we have that

$$\{f, g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q},$$

and so

$$\{q, \mathcal{H}\} = \frac{\partial q}{\partial q}\frac{\partial \mathcal{H}}{\partial p} - \frac{\partial q}{\partial p}\frac{\partial \mathcal{H}}{\partial q}.$$

Since q and p are independent variables, $\partial q / \partial q = 1$ and $\partial q / \partial p = 0$. Hence $\{q, \mathcal{H}\} = \partial \mathcal{H} / \partial p$, which is $\dot{q}$ by Hamilton's equations. Similarly we can show explicitly (do it!) that $\{p, \mathcal{H}\} = \dot{p}$.

yp_z. One can show that e.g. $\{l_x, l_y\} = l_z$, and cyclic permutations:[13] no two components of *angular momentum* can simultaneously act as generalised momenta. Transformation theory taken to its ultimate limit is the Hamilton-Jacobi method, which identifies a transformation whereby all transformed momenta are constants.

5.6 Complex variable formulation

In quantum mechanics, we are familiar with the idea that some pairs of operators commute, and others do not. Eqn. (29) hints at a possible connection between Poisson brackets involving certain observables, and the corresponding quantum mechanical operators: the Poisson bracket of x and p_x is non-vanishing, they do not commute. In quantum mechanics, this is elevated on a pedestal and lies at the heart of the mystery behind Heisenberg's Uncertainty principle. But as we have just seen, in reality there is nothing whatsoever mysterious about it.

Problems in classical mechanics, be it masses on springs, pulleys or pendula, celestial mechanics, fluid dynamics or other, are naturally formulated in real variables. On the other hand, problems in quantum mechanics and electromagnetism require or benefit from the use of complex variables. To explore the classical–quantum connection in more detail, we consider a complex variable formulation of Hamiltonian dynamics.

We deal explicitly with a system with one degree of freedom: $\mathcal{H} = \mathcal{H}(q, p, t)$, so that

$$\dot{q} = \frac{\partial \mathcal{H}}{\partial p} \equiv \mathcal{H}_p, \qquad \dot{p} = -\frac{\partial \mathcal{H}}{\partial q} \equiv -\mathcal{H}_q,$$

defining a shorthand notation for derivatives with respect to q and p. We now introduce

$$z = (q + ip)/\sqrt{2}, \qquad z^* = (q - ip)/\sqrt{2}.$$

z and z^* are considered independent variables – Fig. 9. Then using Hamilton's equations

$$\dot{z} = \frac{1}{\sqrt{2}}(\mathcal{H}_p - i\mathcal{H}_q), \qquad \dot{z}^* = \frac{1}{\sqrt{2}}(\mathcal{H}_p + i\mathcal{H}_q).$$

Since $q = (z + z^*)/\sqrt{2}$ and $p = (z - z^*)/\sqrt{2}$, we can consider $\mathcal{H} = \mathcal{H}(z, z^*, t)$. Then

$$\mathcal{H}_p = \frac{\partial \mathcal{H}}{\partial p} = \frac{\partial \mathcal{H}}{\partial z}\frac{\partial z}{\partial p} + \frac{\partial \mathcal{H}}{\partial z^*}\frac{\partial z^*}{\partial p} = \frac{i}{\sqrt{2}}\left(\frac{\partial \mathcal{H}}{\partial z} - \frac{\partial \mathcal{H}}{\partial z^*}\right), \qquad \mathcal{H}_q = \frac{\partial \mathcal{H}}{\partial q} = \frac{\partial \mathcal{H}}{\partial z}\frac{\partial z}{\partial q} + \frac{\partial \mathcal{H}}{\partial z^*}\frac{\partial z^*}{\partial q} = \frac{1}{\sqrt{2}}\left(\frac{\partial \mathcal{H}}{\partial z} + \frac{\partial \mathcal{H}}{\partial z^*}\right)$$

and hence $\mathcal{H}_p + i\mathcal{H}_q = \sqrt{2}i\,\partial\mathcal{H}/\partial z$, $\mathcal{H}_p - i\mathcal{H}_q = -\sqrt{2}i\,\partial\mathcal{H}/\partial z^*$, and

$$i\dot{z} = \frac{\partial \mathcal{H}}{\partial z^*}, \qquad i\dot{z}^* = -\frac{\partial \mathcal{H}}{\partial z}.$$

More generally, for a system with n degrees of freedom

$$i\dot{z}_j = \frac{\partial \mathcal{H}}{\partial z_j^*}, \qquad i\dot{z}_j^* = -\frac{\partial \mathcal{H}}{\partial z_j}, \qquad \text{for} \qquad j = 1,2,\ldots n; \qquad z_j = (q_j + ip_j)/\sqrt{2}.$$

---

[13] *Which may look vaguely familiar to you, from your studies of quantum mechanics.*

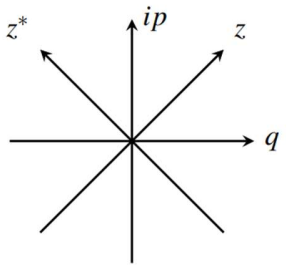

Figure 9: z and z^* are treated as independent variables, just as q and p are.

If a system is described by equations of motion of this form, it must be that $\mathcal{H}$ is a Hamiltonian.

Now let f be some dynamical quantity. Instead of $f = f(q,p,t)$ we can write $f = f(z,z^*,t)$ and

$$\frac{df}{dt} = \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} + \frac{\partial f}{\partial z^*}\frac{\partial z^*}{\partial t} + \frac{\partial f}{\partial t} = \frac{\partial f}{\partial z}\left(-i\frac{\partial \mathcal{H}}{\partial z^*}\right) + \frac{\partial f}{\partial z^*}\left(i\frac{\partial \mathcal{H}}{\partial z}\right) + \frac{\partial f}{\partial t}$$

or

$$\frac{df}{dt} = -i\{f,\mathcal{H}\} + \frac{\partial f}{\partial t}.$$

Comparing with Eqn. (27a), we see that $\{f,\mathcal{H}\}_{\text{real}} \to -i\{f,\mathcal{H}\}_{\text{complex}}$.

We are now in a position to state the connection between these classical results and quantum mechanics:

$$\{\,,\,\}_{\text{classical/complex}} \to \frac{1}{\hbar}[\,,\,]_{\text{quantum}} \qquad\qquad \{\,,\,\}_{\text{classical/real}} \to \frac{1}{i\hbar}[\,,\,]_{\text{quantum}}.$$

If the quantum mechanical operator $\hat{f}$ corresponds to the classical quantity $f(q,p,t)$, then the classical equation of motion, Eqn. (27), corresponds to the quantum operator equation

$$\frac{d\hat{f}}{dt} = \frac{i}{\hbar}\left[\hat{\mathcal{H}},\hat{f}\right] + \frac{\partial \hat{f}}{\partial t}. \qquad (30)$$

$\hat{\mathcal{H}}$ is the Hamiltonian operator, and $[\hat{f},\hat{g}] = \hat{f}\hat{g} - \hat{g}\hat{f}$. Equation (30) is the equation of motion underpinning the *Heisenberg picture* of quantum mechanics (or Heisenberg representation), introduced by Werner Heisenberg, Fig. 10, and for which he received the 1932 Nobel prize — although he believed he should have shared it with Max Born and Pascual Jordan, who recognised in what he did the non-commutative algebra of matrices. At the time, matrices were not part of the standard mathematical repertoire of physicists. Equation (30) was first written down by Paul Dirac — who shared the 1933 Nobel prize with Schrödinger. He called it Heisenberg's equation of motion. It can be derived starting from the Schrödinger equation, showing these to be two different formulations of quantum mechanics, as Lagrangian and Hamiltonian dynamics are two different formulations of classical mechanics.

> From Dirac's diary: *I went back to Cambridge at the beginning of October 1925, and resumed my previous style of life, intense thinking about these problems during the week and relaxing on Sunday, going for a long walk in the country alone. The main purpose of these long walks was to have a rest so that I would start refreshed on the following Monday. It was during one of the Sunday walks in October 1925, when I was thinking about this (uv - vu), in spite of my intention to relax, that I*

Figure 10: L-R:Werner Karl Heisenberg, 1901-1976. Max Born, 1882-1970. Ernst Pascual Jordan, 1902-1980. Paul Adrien Maurice Dirac, 1902-1984. Two grumpy old men, a Nazi, and the greatest scientist name-checked on the wikipedia page for the University of Bath.

thought about Poisson brackets. I remembered something which I had read up previously, and from what I could remember, there seemed to be a close similarity between a Poisson bracket of two quantities and the commutator. The idea came in a flash, I suppose, and provided of course some excitement, and then came the reaction – No, this is probably wrong. I did not remember very well the precise formula for a Poisson bracket, and only had some vague recollections. But there were exciting possibilities there, and I thought that I might be getting to some big idea. It was really a very disturbing situation, and it became imperative for me to brush up on my knowledge of Poisson brackets. Of course, I could not do that when I was right out in the countryside. I just had to hurry home and see what I could find about Poisson brackets. I looked through my lecture notes, the notes that I had taken at various lectures, and there was no reference there anywhere to Poisson brackets. The textbooks which I had at home were all too elementary to mention them. There was nothing I could do, because it was Sunday evening then and the libraries were all closed. I just had to wait impatiently through that night without knowing whether this idea was really any good or not, but I still think that my confidence gradually grew during the course of the night. The next morning I hurried along to one of the libraries as soon as it was open, and then I looked up Poisson brackets in Whittacker's Analytical Dynamics, and I found that they were just what I needed.

6. Small oscillations

The study of systems close to stable equilibrium has some unifying themes. We consider conservative systems in which the potential is a function of position only, and that the equations defining the generalised coordinates do not depend upon time explicitly: $x_i = x_i(q_1, q_2, \dots, q_n)$. Such systems are termed `natural`, or `autonomous`, as against `forced` systems with time-dependent potentials or constraints.

The system is in `equilibrium` when the generalised forces vanish:

$$\mathcal{F}_i = -\frac{\partial \mathcal{V}}{\partial q_i}\bigg|_0 = 0, \qquad (31)$$

which identifies the locations of equilibrium configurations of the system $q_1^0, q_2^0, \dots q_n^0$ as extrema of the potential. A pendulum hanging vertically, an egg stood on end. If small displacements from equilibrium result in bounded motion about q_i^0 the equilibrium is `stable`. It is `unstable` if infinitesimal displacements

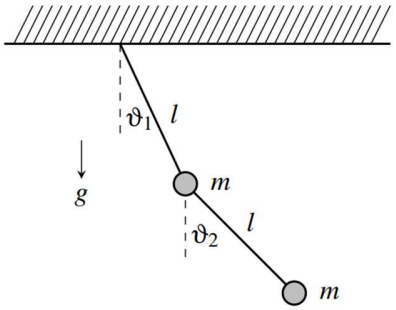

Figure 11: The double planar pendulum consists of two (here identical) pendula attached to one another. Motion in the plane.

produce unbounded motion. Stable equilibrium corresponds to the extremum of $\mathcal{V}$ being minimum, as conservation of energy then dictates displacements from equilibrium increase $\mathcal{V}$ and so reduce speeds, eventually stopping movement and the motion is bound. If $\mathcal{V}$ decreases moving from equilibrium, speeds will increase and the motion is unbound. The limp pendulum is stable; the egg stood on end is unstable.

We illustrate the description of systems close to stable equilibrium with the double planar pendulum shown in Fig. 11. It is straightforward [go on, do it!] to obtain

$$\begin{aligned} \mathcal{T} &= ml^2\dot{\vartheta}_1^2 + \frac{1}{2}ml^2\dot{\vartheta}_2^2 + ml^2\cos(\vartheta_1 - \vartheta_2)\dot{\vartheta}_1\dot{\vartheta}_2 \\ \mathcal{V} &= -2mgl\cos\vartheta_1 - mgl\cos\vartheta_2. \end{aligned}$$

Obtaining exact solution of the equations of motion from the Lagrangian $\mathcal{L} = T - V$ is a formidable problem (indeed not possible analytically; they can be solved numerically, and the resulting behaviour is *chaotic* for some initial conditions). Instead, we seek to describe the system for **small** amplitude motion about the stable equilibrium. Physically we know stable equilibrium corresponds to $\vartheta_1^0 = 0$ and $\vartheta_2^0 = 0$ (corresponding to q_1^0, q_2^0), and we can confirm this — but not here — by evaluating first and second derivatives of $\mathcal{V}$. For small displacements from equilibrium which we denote $\eta_i = \vartheta_i - \vartheta_i^0$, effectively introducing a new set of generalised coordinates (here η_i and ϑ_i are identical, but we use the new variable η_i both to remind us that the following assumes *small* displacements, and for consistency with the later generalised treatment) and by using $\cos x = 1 - x^2/2 + \cdots$ we get to second order the Lagrangian

$$\mathcal{L} = ml^2\left(\dot{\eta}_1^2 + \frac{1}{2}\dot{\eta}_2^2 + \dot{\eta}_1\dot{\eta}_2\right) - mgl\left(\eta_1^2 + \frac{1}{2}\eta_2^2\right)$$

where we have ignored constant terms that have no significance for the resulting equations of motion. In the usual way using the Lagrange equations we get equations of motion:

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\eta}_i} = \frac{\partial\mathcal{L}}{\partial\eta_i} \quad \Rightarrow \quad \begin{cases} ml^2(2\ddot{\eta}_1 + \ddot{\eta}_2) &= -2mgl\eta_1 \\ ml^2(\ddot{\eta}_1 + \ddot{\eta}_2) &= -mgl\eta_2. \end{cases}$$

They are coupled, due to the cross term in the original kinetic energy, but we can write the two simultaneous linear differential equations as a single matrix equation

$$\mathbf{T}\ddot{\mathbf{x}} = -\mathbf{V}\mathbf{x} \qquad (32)$$

where

$$\mathbf{T} = ml^2\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}, \qquad \mathbf{V} = mgl\begin{pmatrix}2 & 0\\ 0 & 1\end{pmatrix}.$$

Note that $\mathbf{T}$ and $\mathbf{V}$ are constant matrices, and $\mathbf{x} = \mathbf{x}(t)$, whilst using the same notation the Lagrangian can be written

$$\mathcal{L} = \frac{1}{2}\dot{\mathbf{x}}^T\mathbf{T}\dot{\mathbf{x}} - \frac{1}{2}\mathbf{x}^T\mathbf{V}\mathbf{x}. \qquad (33)$$

To solve the matrix equation of motion Eqn. (32) we assume oscillatory solutions letting

$$\mathbf{x}(t) = \mathbf{a}e^{i\omega t}, \qquad \mathbf{a} \text{ constant.} \qquad (34)$$

At the end we will use the linearity of the equation to construct overall real solutions to the physical problem from suitable linear combinations of the possible solutions. Substituting into Eqn. (32) and cancelling exponential time factors gives a matrix eigenvalue problem:

$$\mathbf{V}\mathbf{a} = \omega^2\mathbf{T}\mathbf{a}. \qquad (35)$$

Non-trivial solution ($\mathbf{a} \neq 0$) requires

$$\det(\mathbf{V} - \omega^2\mathbf{T}) = 0 \qquad \Rightarrow \qquad \left|\begin{matrix}2(\omega_0^2 - \omega^2) & -\omega^2\\ -\omega^2 & \omega_0^2 - \omega^2\end{matrix}\right| = 0, \quad \omega_0^2 = \frac{g}{l}$$

The characteristic polynomial $2(\omega_0^2 - \omega^2)^2 - \omega^4 = (\omega^2)^2 - 4\omega_0^2(\omega^2) + 2\omega_0^4 = 0$ is a quadratic in ω^2 and yields

$$\omega^2 = \frac{-(-4\omega_0^2) \pm \sqrt{(-4\omega_0^2)^2 - 4.(2\omega_0^4)}}{2} = (2 \pm \sqrt{2})\omega_0^2$$

There are two ω^2 values, $\omega_\pm^2 = (2 \pm \sqrt{2})\omega_0^2$ and four ω values, which come in plus-minus pairs: $\omega = \pm\omega_\pm$. Inserting ω_+^2 and ω_-^2 into Eqn. (35) we obtain the corresponding (unnormalised) eigenvectors $\mathbf{a}_\pm$:

$$\mathbf{V}\mathbf{a}_\pm = \omega_\pm^2\mathbf{T}\mathbf{a}_\pm \qquad \Rightarrow \qquad \mathbf{a}_\pm = \begin{pmatrix}1\\ \mp\sqrt{2}\end{pmatrix}.$$

We can now write down the general real solution describing small amplitude oscillations of the double pendulum:

$$\begin{aligned}\mathbf{x}(t) &= c_+\mathbf{a}_+e^{i\omega_+t} + c_+^*\mathbf{a}_+e^{-i\omega_+t} + c_-\mathbf{a}_-e^{i\omega_-t} + c_-^*\mathbf{a}_-e^{-i\omega_-t}\\ &= C_+\mathbf{a}_+\cos(\omega_+t + \phi_+) + C_-\mathbf{a}_-\cos(\omega_-t + \phi_-).\end{aligned} \qquad (36)$$

The two complex coefficients c_+ and c_-, or equivalently the four real amplitudes and phases $C_\pm$ and $\phi_\pm$ in the second form, are determined by the initial conditions (position and velocity) of the two masses.

We recognise in Egn. (36) two **independent** motions, or `normal modes`. These are characterised by the `eigen vectors` $\mathbf{a}_\pm$ and corresponding `eigen frequencies` $\omega_\pm$. The general motion is a superposition of the two normal motions. In one ($-$) the two angular displacements from equilibrium have similar sign (see the eigen vector entries) — they are *in phase*, and the two masses are always displaced to the same side of equilibrium. This oscillation has a frequency $\omega_- = \sqrt{2 - \sqrt{2}}\omega_0$ which is less than ω_0, the natural frequency of the single pendulum. The second normal mode ($+$) is motion in which the two masses are displaced to opposite sides of equilibrium, and oscillate at a frequency $\omega_+ = \sqrt{2 + \sqrt{2}}\omega_0$, greater than that of the single pendulum. See Fig. 12.

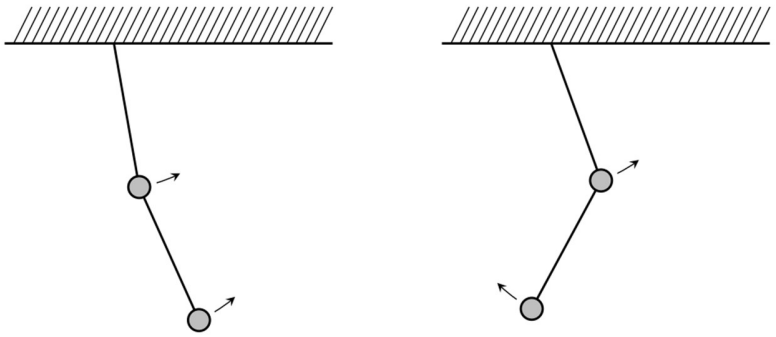

Figure 12: The two normal modes of the double pendulum: left, in-phase; right anti-phase.

Now let us introduce

$$Q_{\pm} = c_{\pm}e^{i\omega_{\pm}t} + c_{\pm}^{*}e^{-i\omega_{\pm}t}$$

so that

$$\mathbf{x} = \mathbf{a}_{+}Q_{+} + \mathbf{a}_{-}Q_{-} \equiv \begin{pmatrix} 1 & 1 \\ -\sqrt{2} & +\sqrt{2} \end{pmatrix}\begin{pmatrix} Q_{+} \\ Q_{-} \end{pmatrix} = \mathbf{AQ}$$

defining $\mathbf{A}$ as the matrix of eigen vectors and $\mathbf{Q}$ as the vector containing the corresponding Q's. Inverting, $\mathbf{Q} = \mathbf{A}^{-1}\mathbf{x}$ which can be expanded out to show that Q_{+} and Q_{-} can be expressed in terms of the generalised coordinates η_1 and η_2; hence the Q's are seen to be a new set of generalised coordinates, the `normal coordinates` of the system. These normal coordinates do not describe the positions of the individual masses, but are collective coordinates describing coupled displacements of both.

The normal coordinates satisfy the two independent equations of motion

$$\ddot{Q}_{\pm} = -\omega_{\pm}^{2}Q_{\pm}$$

describing oscillators with frequency $\omega_{\pm}$. We know the Lagrangian for an oscillator.[14] Hence the dynamics of the double pendulum near stable equilibrium can be written as

$$\mathcal{L} = \left(\frac{1}{2}\dot{Q}_{+}^{2} - \frac{1}{2}\omega_{+}^{2}Q_{+}^{2}\right) + \left(\frac{1}{2}\dot{Q}_{-}^{2} - \frac{1}{2}\omega_{-}^{2}Q_{-}^{2}\right).$$

This Lagrangian, expressed in normal coordinates, has no crossed terms.

Class example:

[14] Previously we have written the Lagrangian of a mass on a spring as $\mathcal{L} = (m/2)\dot{x}^2 - (k/2)x^2$ where m is the mass and k the stiffness of the spring; it is an oscillator that oscillates at frequency $\omega = \sqrt{k/m}$. Since here we are talking about normal modes of vibration, for which only the frequency is defined and not a particular mass or stiffness, we set $m = 1$ and $k = \omega^2$.

6.1 General treatment

It is not difficult to see that this transformation to normal coordinates is more generally possible. We start by expanding the potential about a point of stable equilibrium $(q_1^1, \ldots, q_n^0)$ as

$$\mathcal{V}(q_1, \ldots, q_n) = \mathcal{V}|_0 + \sum_j (q_j - q_j^0) \left.\frac{\partial \mathcal{V}}{\partial q_j}\right|_0 + \frac{1}{2}\sum_{jk} (q_j - q_j^0)(q_k - q_k^0) \left.\frac{\partial^2 \mathcal{V}}{\partial q_j \, \partial q_k}\right|_0 + \cdots.$$

Letting $\eta_i = q_i - q_i^0$, so $\partial / \partial q_i = \partial / \partial \eta_i$, and noting $\partial \mathcal{V} / \partial q_j \big|_0 = 0$ at equilibrium, see Eqn. (31), we can write

$$\mathcal{V} = \frac{1}{2}\sum_{jk} \eta_j \, \eta_k V_{jk} + \mathcal{O}(\eta^3) \qquad V_{jk} = \left.\frac{\partial^2 \mathcal{V}}{\partial \eta_j \, \partial \eta_k}\right|_0 = V_{kj}, \tag{37}$$

ignoring the constant term $\mathcal{V}|_0$ since it is of no consequence. We also express the kinetic energy in terms of the new generalised coordinates η_i, the distances from stable equilibrium. With $x_i = x_i(\eta_1, \ldots, \eta_n)$

$$\dot{x}_i = \sum_j \frac{\partial x_i}{\partial \eta_j} \dot{\eta}_i.$$

Expanding the derivatives about the position of equilibrium,

$$\frac{\partial x_i}{\partial \eta_j} = \left.\frac{\partial x_i}{\partial \eta_j}\right|_0 + \sum_k \eta_k \left.\frac{\partial^2 x_i}{\partial \eta_k \, \partial \eta_j}\right|_0 + \cdots$$

and hence

$$\begin{aligned} \mathcal{T} = \frac{1}{2}\sum_i m_i \, \dot{x}_i^2 = \frac{1}{2}\sum_i m_i \, \dot{x}_i \cdot \dot{x}_i \quad &= \quad \frac{1}{2}\sum_i m_i \left(\sum_j \left.\frac{\partial x_i}{\partial \eta_j}\right|_0 \dot{\eta}_j \right) \cdot \left(\sum_k \left.\frac{\partial x_i}{\partial \eta_k}\right|_0 \dot{\eta}_k \right) + \mathcal{O}(\eta\dot{\eta}^2) \\ &= \quad \frac{1}{2}\sum_{jk} T_{jk} \, \dot{\eta}_j \dot{\eta}_k + \mathcal{O}(\eta\dot{\eta}^2), \qquad T_{jk} = \sum_i m_i \left.\frac{\partial x_i}{\partial \eta_j}\right|_0 \left.\frac{\partial x_i}{\partial \eta_k}\right|_0 = T_{kj}. \end{aligned} \tag{38}$$

So in the vicinity of stable equilibrium, the system is described by the Lagrangian

$$\begin{aligned} \mathcal{L} = \mathcal{L}(\eta_1, \ldots, \eta_n, \dot{\eta}_1, \ldots, \dot{\eta}_n) \quad &= \quad \frac{1}{2}\sum_{jk} T_{jk} \, \dot{\eta}_j \dot{\eta}_k - \frac{1}{2}\sum_{jk} \eta_j \, \eta_k V_{jk} \\ &= \quad \frac{1}{2}\dot{\mathbf{x}}^T \mathbf{T} \dot{\mathbf{x}} - \frac{1}{2}\mathbf{x}^T \mathbf{V} \mathbf{x} \end{aligned}$$

where

$$\mathbf{x} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

and the matrix elements of **V** and **T** are given in Eqns. (37) and (38) respectively.

We now see the formal similarity between this general case, and the double pendulum problem just considered, and the subsequent analysis follows the same steps. The equations of motion from the Lagrange

equations have the form of Eqn. (32), and assuming oscillator solutions of the form (34) yields the eigenvalue problem (35). Solution of this yields n eigen frequencies and corresponding eigen vectors defining the normal modes. Defining a set of normal coordinates Q_j, $j = 1, \dots, n$ we can express the Lagrangian as a sum of oscillators:

$$\mathcal{L} = \sum_{j=1}^{n} \left(\frac{1}{2} \dot{Q}_j^2 - \frac{1}{2} \omega_j^2 Q_j^2 \right). \qquad (39)$$

We have arrived at this result truncating the expansion of the potential at quadratic terms. Going beyond this to include third derivative and higher-order terms is necessary if we wish to study *anharmonic* effects associated with larger displacements from equilibrium — thermal expansion is a consequence of anharmonic effects in the motion of atoms in a crystal. Retaining higher-order terms prevents a description in terms of independent normal modes, and restores the complexity of the problem. One approach is to assume these as secondary effects that cause interactions between the normal modes, with the general motion represented as a superposition of normal mode oscillations with time dependant coefficients, whose variation is determined by the anharmonic terms.

7. Continuous systems and fields

The Lagrangian and Hamiltonian formalism can be extended to continuous systems, which have an infinite number of degrees of freedom. The development is straightforward, but the result has some unexpected aspects.

We consider explicitly a mechanical system, an elastic rod whose dynamics are governed by Newton's laws of motion. The rod comprises a great number of interacting particles, which are not recognisable at the macroscopic level, where it appears continuous. A limiting procedure allows us to develop an description in terms of a continuous function of space and time — a field. We thus arrive at an approximate continuous description of a system that at a microscopic level is made up of particles. The Lagrangian and Hamiltonian formalisms are thus seen to encompass such systems, governed by Newtonian laws.

However, it turns out that the formalism is more generally applicable to almost any field system, including those that cannot be considered the continuum limit of an underlying many-particle system, or to be associated with Newtonian dynamics — the electromagnetic field is a particular example. Indeed field theory forms the basis of modern fundamental theories describing elementary particles and their interactions. Described as fields. Which are then quantised. Yielding particles.

7.1 Field equations from a continuous limit

Consider a chain of $n + 2$ identical masses separated by distance a and connected by identical springs, stiffness k. The first and last mass are secured to immovable supports, Fig. 13, so that the total length of the system is $\ell = (n + 1)a$.

Displacing the masses from equilibrium will extend or compress the springs, creating forces that will act upon the masses, accelerating them. Motion will result. Thus the displacement of mass *j* from its equilibrium position, φ_j, is a function of time, and the Lagrangian[15] is

[15] L not $\mathcal{L}$ is not a typo, but intentional.

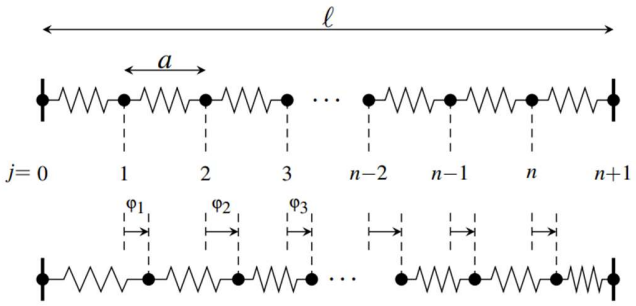

Figure 13: A system of discrete masses connected by springs, displaced from equilibrium by $\varphi_j(t)$. A limiting procedure taking $n \to \infty$ yields a continuum description of an elastic rod in terms of a displacement field $\varphi(x,t)$.

$$L = T - V = \sum_{j=1}^{n} \frac{1}{2} m \left(\frac{d\varphi_j}{dt}\right)^2 - \sum_{j=0}^{n} \frac{1}{2} k \left(\varphi_{j+1} - \varphi_j\right)^2$$

It is easy to confirm[16] that the usual Lagrange equation gives the expected equations of motion for the masses.

To arrive at a continuum description as an elastic rod, we increase the number of particles to infinity ($n \to \infty$) keeping constant the total length ℓ, the linear mass density $\varrho = m/a$, and the elastic modulus $K = ka$. With these in mind, we write

$$\begin{aligned} L &= \frac{1}{2} \sum_{j=1}^{n} \frac{m}{a} \left(\frac{d\varphi_j}{dt}\right)^2 a - \frac{1}{2} \sum_{j=0}^{n} (ka) \left(\frac{\varphi_{j+1} - \varphi_j}{a}\right)^2 a \\ &\to \frac{1}{2} \int_0^{\ell} \varrho \left(\frac{\partial \varphi}{\partial t}\right)^2 dx - \frac{1}{2} \int_0^{\ell} K \left(\frac{\partial \varphi}{\partial x}\right)^2 dx. \end{aligned}$$

As a becomes the infinitesimal dx, the index j is replaced by the continuous coordinate x, $\varphi_j(t)$ becomes also a function of x, i.e. $\varphi_j(t) \to \varphi(x,t)$, and

$$\lim_{a \to 0} \frac{\varphi_{j+1} - \varphi_j}{a} = \lim_{dx \to 0} \frac{\varphi(x + dx) - \varphi(x)}{dx} = \frac{\partial \varphi}{\partial x}.$$

So introducing the shorthand subscript-derivative notation $\partial_x \varphi \equiv \partial \varphi / \partial x$, $\partial_t \varphi \equiv \partial \varphi / \partial t$, $\partial_x^2 \varphi = \partial^2 \varphi / \partial x^2$ etc[17]

$$L \to \int \frac{1}{2} (\rho (\partial_t \varphi)^2 - K (\partial_x \varphi)^2)\, dx = \int \mathcal{L} dx\,, \qquad (40)$$

and we see that the Lagrangian for a continuous one-dimensional rod is itself a functional of position. The integrand

[16] You know what this footnote is going to say! "Do it!"

[17] Another widely used notation is the subscript notation $\varphi_x \equiv \partial \varphi / \partial x$, $\varphi_t \equiv \partial \varphi / \partial t$, $\varphi_{xx} \equiv \partial^2 \varphi / \partial x^2$ etc

$$\mathcal{L} = \frac{1}{2}(\rho(\partial_t\varphi)^2 - K(\partial_x\varphi)^2) \qquad (41)$$

is the `Lagrangian density`, usually referred to simply as the Lagrangian[18], since whether a Lagrangian or Lagrangian density is meant is usually obvious from the context. Mathematically, when dealing with fields we will use roman script for integrated quantities and calligraphic script for densities.

Let us stress here that x is *not* a generalised coordinate (i.e. *not* the same as $x(t)$ for a single mass on a spring). It is a continuous independent variable that has entered as a replacement for the discrete index j that labelled the masses in the chain. There are no dynamics associated with the variable x. At the risk of repetition, the displacement $\varphi_j \to \varphi(x)$. In fact, $\varphi = \varphi(x,t)$ since the displacements vary in time too (originally, $\varphi_j = \varphi_j(t)$). φ is a field, or field variable, a function of independent position and time[19].

Here, the Lagrangian (density) in (41) has the form $\mathcal{L} = \mathcal{L}(\partial_t\varphi, \partial_x\varphi)$. More generally, in addition to $\partial_t\varphi$, and $\partial_x\varphi$, both x and t might appear in the Lagrangian[20], as well as φ.[21] So in one space-dimension let us consider Lagrangians $\mathcal{L} = \mathcal{L}(\varphi, \partial_t\varphi, \partial_x\varphi, x, t)$. What do we find if we assume Hamilton's principle $\delta\mathcal{S} = 0$ applies? Then

$$\delta\mathcal{S} = \delta\int L\,dt = 0 \qquad \Rightarrow \qquad \delta\mathcal{S} = \delta\iint \mathcal{L}dxdt\ = 0.$$

This suggests x and t are to be treated on an equal footing, and we consider variations in the field φ which vanish at the endpoints of both x and t integrals[22].

$$\mathcal{S}[\varphi + \delta\varphi] = \int dx \int dt \left(\mathcal{L} + \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\delta(\partial_t\varphi) + \frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\delta(\partial_x\varphi) + \cdots\right).$$

The first term gives $\mathcal{S}[\varphi]$. Since $\delta(\partial_t\varphi) = \partial_t(\delta\varphi)$, and $\delta(\partial_x\varphi) = \partial_x(\delta\varphi)$, the third and fourth terms can be integrated by parts. Doing so, and grouping terms containing $\delta\varphi$ gives

$$\delta\mathcal{S} = \int dx \int dt\,\delta\varphi\left[\frac{\partial\mathcal{L}}{\partial\varphi} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)} - \frac{d}{dx}\frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\right] + \int dx\,\delta\varphi\,\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\Big|_{t_1}^{t_2} + \int dt\,\delta\varphi\,\frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\Big|_{x_1}^{x_2}.$$

The final two terms are zero for variations vanishing at $x = x_1, x_2$, and at $t = t_1, t_2$. Hence the stationary principle for arbitrary variations $\delta\varphi$ requires

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)} + \frac{d}{dx}\frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)} - \frac{\partial\mathcal{L}}{\partial\varphi} = 0. \qquad (42)$$

By extension, for a Lagrangian density that varies in x, y and z, the field φ that makes stationary the action

---

[18] Hence the change in notation noted in footnote 15!

[19] Note that since t and x are independent variables, $\partial\varphi/\,\partial x \equiv d\varphi/dx$ and $\partial\varphi/\,\partial t \equiv d\varphi/dt$.

[20] Either explicitly, or disguised e.g. in the Lagrangian in Eqn. (41) if the mass density were spatially varying, $\varrho = \varrho(x)$

[21] Obviously, problems in more spatial dimensions will potentially introduce $y, z, \partial_y\varphi, \partial_z\varphi, \ldots$. We might also have the situation where the Lagrangian contains more than one field variable, φ_s, $s = 1,2,\ldots,n$. φ may also be complex (e.g. $\psi(x,y,z,t)$, a wave function for a quantum mechanical system; we could think of the real and imaginary parts as separate field variables $\psi = \varphi_1 + i\varphi_2$ and proceed normally, or alternatively/equivalently treat $\varphi_1 = \psi$ and $\varphi_2 = \psi^*$ as the two independent field variables).

[22] Alternatively, we could apply periodic boundary conditions, and require variations to coincide at both "endpoints".

$$\delta\mathcal{S} = \delta \iint \mathcal{L} d\vec{r} dt \; = 0$$

satisfies the field equation

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial(\partial_t \varphi)} + \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial(\partial_x \varphi)} + \frac{d}{dy}\frac{\partial \mathcal{L}}{\partial(\partial_y \varphi)} + \frac{d}{dz}\frac{\partial \mathcal{L}}{\partial(\partial_z \varphi)} - \frac{\partial \mathcal{L}}{\partial \varphi} = 0 \qquad (43)$$

or

$$\sum_{\mu=t,x,y,z} \frac{d}{d\mu}\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} - \frac{\partial \mathcal{L}}{\partial \varphi} = 0. \qquad (44)$$

Eqns. (42), (43) and (44) are continuum Lagrange equations for one- and three-dimensional fields.[23]

If the Lagrangian (density) contains more than one field, Eqns. (42)–(44) apply to each field separately. If the field variable is complex, then they apply to the real and imaginary parts of the field separately, or alternatively to the field variable and the complex conjugate of the field variable separately, with the two treated as independent fields.

Returning to our one-dimensional rod, with the Lagrangian density of Eqn. (41), using the continuum Lagrange equation gives

$$\rho\, \partial_t^2 \varphi - K\, \partial_x^2 \varphi = 0$$

or

$$\frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} - \frac{\partial^2 \varphi}{\partial x^2} = 0,$$

a wave equation with speed $c = \sqrt{K/\rho}$. Experiment tells us this is the correct description of longitudinal displacements in an elastic rod. Comparison of the predictions of the stational action principle to other continuum systems shows that it provides a valid basis for the theoretical description of physical systems.

Class example:

[23] This last equation is sometimes written as

$$\frac{d}{d\mu}\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} - \frac{\partial \mathcal{L}}{\partial \varphi} = 0$$

with an implied summation over repeated indices. This is the Einstein summation convention introduced — you guessed it — by the great man himself.

7.2 Hamiltonian formulation

Returning to the discrete system of the chain of masses connected by springs, the momentum conjugate to the displacement variable of mass j is $p_j = \partial L/\,\partial\dot{\varphi}_j$. Summing over masses,

$$\mathcal{P} = \sum_j \frac{\partial}{\partial\dot{\varphi}_j}\left[\frac{1}{2}\frac{m}{a}\left(\frac{d\varphi_j}{dt}\right)^2 - \frac{1}{2}(ka)\left(\frac{\varphi_{j+1}-\varphi_j}{a}\right)^2\right]a$$

and taking the continuum limit ($m/a \to \varrho$, $ka \to K$, $\varphi_j \to \varphi(x,t)$, $(\varphi_{j+1}-\varphi_j)/a \to \partial_x\varphi$) we find

$$\mathcal{P} = \int \frac{\partial}{\partial(\partial_t\varphi)}\left[\frac{1}{2}\varrho(\partial_t\varphi)^2 - \frac{1}{2}K(\partial_x\varphi)^2\right]dx = \int \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}dx.$$

We identify a momentum density

$$\pi = \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}. \qquad (45)$$

Similarly the Hamiltonian $H = \sum_i p_i\,\dot{q}_i - L$, or in the current notation $H = \sum_i p_i\,\dot{\varphi}_i - L$, becomes

$$H \to \int \mathcal{H}dx\,, \qquad \mathcal{H} = \pi\,\partial_t\varphi - \mathcal{L},$$

with $\mathcal{H}$ the Hamiltonian density (or energy density[24]), and it is assumed that the $\partial_t\varphi$ dependence has been replaced by a dependence upon π: $\mathcal{H} = \mathcal{H}(\varphi,\pi,\partial_x\varphi,x,t)$ for one spatial dimension. For the one-dimensional rod, for which the Lagrangian density is given by Eqn. (41), $\mathcal{L} = (1/2)\varrho(\partial_t\varphi)^2 - (1/2)K(\partial_x\varphi)^2$, we have

$$\pi = \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)} = \varrho\,\partial_t\varphi$$

and so

$$\mathcal{H} = (\varrho\,\partial_t\varphi)\,\partial_t\varphi - \left(\frac{1}{2}\varrho(\partial_t\varphi)^2 - \frac{1}{2}K(\partial_x\varphi)^2\right) = \frac{\pi^2}{2\varrho} + \frac{1}{2}K(\partial_x\varphi)^2.$$

Not derived here (why not try it; apply $\delta\mathcal{S} = 0$ to $\mathcal{S} = \iint[\pi\,\partial_t\varphi - \mathcal{H}(\pi,\varphi)]dxdt$) one has Hamilton's field equations

$$\partial_t\varphi = \frac{\partial\mathcal{H}}{\partial\pi}, \qquad \partial_t\pi = -\frac{\partial\mathcal{H}}{\partial\varphi} + \frac{d}{dx}\frac{\partial\mathcal{H}}{\partial(\partial_x\varphi)} \qquad \text{one-dimension} \qquad (46)$$

$$\partial_t\varphi = \frac{\partial\mathcal{H}}{\partial\pi}, \qquad \partial_t\pi = -\frac{\partial\mathcal{H}}{\partial\varphi} + \sum_{\mu=x,y,z}\frac{d}{d\mu}\frac{\partial\mathcal{H}}{\partial(\partial_\mu\varphi)} \qquad \text{three-dimensions.} \qquad (47)$$

Comparing Hamilton's field equations with Lagrange's field equations e.g. Eqn. (46) with (42), we see that in contrast to the Lagrangian formalism, in which space and time enter symmetrically, the Hamiltonian formalism singles out the t-variable for special treatment. As a result, the Lagrangian formalism lends itself

[24] Since $\mathcal{L} \to \alpha\mathcal{L}$ leaves the equations of motion unchanged, we need to be careful in identifying $\mathcal{H}$ as the actual energy density, and not just proportional to it. Also, as in the case of discrete systems, there are cases where the Hamiltonian density is definitely not the same as the energy density, due to external influences acting as energy sources or sinks.

better to the incorporation of relativistic effects (which unfortunately we do not have time to pursue in any depth), and has proven to be more useful for fundamental field theories than the Hamiltonian formalism.

Class example:

7.3 Conservation laws

Consider[25] a one-dimensional system, so $\mathcal{L} = \mathcal{L}(\varphi, \partial_t\varphi, \partial_x\varphi, x, t)$, and

$$\frac{d\mathcal{L}}{dx} = \frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial x} + \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\frac{d(\partial_t\varphi)}{dx} + \frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\frac{d(\partial_x\varphi)}{dx} + \frac{\partial\mathcal{L}}{\partial x}.$$

Using the continuum Lagrange equation to substitute for $\partial\mathcal{L}/\,\partial\varphi$ in the first term,

$$\frac{d\mathcal{L}}{dx} = \left(\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)} + \frac{d}{dx}\frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\right)\frac{\partial\varphi}{\partial x} + \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\frac{d(\partial_t\varphi)}{dx} + \frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\frac{d(\partial_x\varphi)}{dx} + \frac{\partial\mathcal{L}}{\partial x},$$

and recognising terms that can be combined as the derivative of products, we have (note: $\partial(\partial_t\varphi)/\,\partial x = \partial_x\,\partial_t\varphi \equiv \partial_t\,\partial_x\varphi = \partial(\partial_x\varphi)/\,\partial t$)

$$\frac{d\mathcal{L}}{dx} = \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\partial_x\varphi\right) + \frac{d}{dx}\left(\frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\partial_x\varphi\right) + \frac{\partial\mathcal{L}}{\partial x},$$

or

$$\frac{\partial\mathcal{L}}{\partial x} = \frac{d}{dx}\left(\mathcal{L} - \frac{\partial\mathcal{L}}{\partial(\partial_x\varphi)}\partial_x\varphi\right) - \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\partial_x\varphi\right).$$

In three-dimensions ($x, y, z \equiv x_1, x_2, x_3$) this becomes

$$\frac{\partial\mathcal{L}}{\partial x_i} = \sum_j \frac{d}{dx_j}\left(\mathcal{L}\delta_{ij} - \frac{\partial\mathcal{L}}{\partial(\partial_{x_j}\varphi)}\partial_{x_i}\varphi\right) - \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\partial_{x_i}\varphi\right).$$

If $\mathcal{L}$ is independent of the variable x_i (or stated alternatively, if $\mathcal{L}$ is symmetric under translations in x_i), then $\partial\mathcal{L}/\,\partial x_i = 0$ and the right-hand side here must vanish. Defining

$$T^i_j = \frac{\partial\mathcal{L}}{\partial(\partial_{x_j}\varphi)}\partial_{x_i}\varphi - \mathcal{L}\delta_{ij}, \quad T^i_0 = \frac{\partial\mathcal{L}}{\partial(\partial_t\varphi)}\partial_{x_i}\varphi,$$

and considering the vector $\vec{T}^i = (T^i_1, T^i_2, T^i_3)$, then

$$\vec{\nabla}\cdot\vec{T}^i + \frac{d\,T^i_0}{dt} = 0 \qquad \text{if} \qquad \frac{\partial\mathcal{L}}{\partial x_i} = 0.$$

The symmetry of the Lagrangian is seen to result in continuity equations, or *conservation laws*, satisfied by the quantities T^i_j, which form what is called a *stress tensor*. Why does a continuity equation imply a

[25] This section looks cumbersome, and it is. It would be nicer using the Einstein notation of footnote 23, but to use that requires familiarity with it which only comes from repeated exposure — which you have not had.

conservation law? Integrating over all space, and using that the fields vanish at infinity (or satisfy periodic boundary conditions as an alternative), then we have that

$$\frac{d}{dt}\int T_0^i dx = 0,$$

and we see that the integrated quantity is a conserved (Noether) charge. Time invariance similarly gives rise to an energy continuity equation,

$$\vec{\nabla}\cdot\vec{S} + \frac{d\,\mathcal{H}}{dt} = 0 \qquad \text{if} \qquad \frac{\partial \mathcal{L}}{\partial t} = 0,$$

where $S_i = (\partial\mathcal{L}/\,\partial(\partial_{x_i}\varphi))\,\partial_t\varphi$ is a vector representing the flow of energy. In this case, if the Lagrangian density is explicitly independent of time, we have the conserved (Noether) charge

$$\Lambda = \int \mathcal{H} dx.$$

In a complex field theory, analysis along the lines above can be performed to show that continuous *phase* symmetry of the wave function of an electron implies conservation of electrical *charge*.

Class example:

8. Further reading

Classical mechanics is a classic topic in mathematical physics – you will find many good textbooks on the subject in the library under class mark 531.01 plus elsewhere. The following have good reputations:

Landau, L D and Lifshitz, E M, *Mechanics (Vol. 1, Course of Theoretical Physics)*, Elsevier, Butterworth-Heinemann, Oxford, 1976. [3rd ed.]

Kibble, T W B and Berkshire, F H, *Classical mechanics*, Longman, Harlow, 1996. [4th ed.]

Goldstein H, Poole C and Safko J, *Classical Mechanics*, Addison Wesley, San Francisco, 2002. [3rd ed.]