Following the shocking and tragic loss of Professor Alessandro De Vita in Oct 2018 I was asked to take over his second year lecture course in Electromagnetism. I spent some time looking at the College web pages of his course and I found some powerpoint slides (possibly rather ancient), lots of problems and some screen captures of notes on his tablet. I knew from talking to him that his preferred medium for lecturing is via his tablet which he connects up to the data projector. (I remember his amazement at the College's incredulity that he wished to make this connection over the wireless network—doubtless so that he could move freely around the lecture theatre.) I realised that he was teaching the students in the language of four-vectors, retarded potentials and Lagrangian mechanics. I was told by students that his lectures were fluent and often very wide ranging and I can believe that he was reluctant to adhere to the very pedestrian syllabus laid down in the "Module Description". I also realised that I lack his unique brilliance to carry this off. I recall from many conversations with Sandro that his mind was ceaslessly and turbulently leaping from topic to topic as he would turn over a question and explore connections across all branches of physics to find viewpoints that would never occur to me. So I decided to re-write the course from scratch. However I found about three powerpoint slides which exposed for me his oblique insight: these were about how to postulate a Lagrangian in the electromagnetic field that is consistent with the Lorentz force. He ended with a box on the third slide inspired by Feynman's lecture on the principle of least action. I decided to build a lecture around these three slides and to add a section on how the vector potential may in this way be allowed to enter the hamiltonian in quantum mechanics which is outside the classical electromagnetism but which I thought the students may find useful since they are also studying quantum mechanics in another course.

I will not include this in the following years since I wished this to be a unique reference to the students to a departed colleague, and my tribute to Sandro De Vita.

Tony Paxton, 5 June 2019

Hamiltonian and Lagrangian—equations of motion and potentials

After Professor Alessandro De Vita

1. Hamiltonian formulation of mechanics

The potential energy is usually defined in mechanics as a function, $U(\mathbf{r})$, of position only; that is to say that we address the case of a *time-independent* potential energy function.† To begin with, we treat just a single particle moving in one dimension, so its position is x, its velocity is $\dot{x}$ and its acceleration is $\ddot{x}$. The total energy, or Hamiltonian, function for the particle is then defined by adding the kinetic energy,

$$H(x, \dot{x}) = K(\dot{x}) + U(x)$$

where

$$K = \frac{1}{2} m \dot{x}^2$$

and m is the mass of the particle. We impose the condition that the total energy is conserved along the trajectory of the particle. This leads to

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} H(x, \dot{x}) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{2} m \dot{x}^2 + U(x) \right) \tag{1}$$

How do we take such a derivative in general? Well, if we have a function of time, position and velocity, $f(x, \dot{x}, t)$ then, treating x, $\dot{x}$ and t as *independent* variables,‡

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x, \dot{x}, t) = \left(\dot{x} \frac{\partial}{\partial x} + \ddot{x} \frac{\partial}{\partial \dot{x}} + \frac{\partial}{\partial t} \right) f(x, \dot{x}, t) \tag{2}$$

It follows that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} H(x, \dot{x}) &= m \dot{x} \ddot{x} + \frac{\mathrm{d}U}{\mathrm{d}x} \dot{x} \\ &= \left(m \ddot{x} + \frac{\mathrm{d}U}{\mathrm{d}x} \right) \dot{x} = 0 \end{aligned}$$

† Try to think of some mechanical devices, *e.g.,* a pendulum, for which the potential energy is time-independent; and try to think of some which have a time dependent potential, for example a charge being driven by an oscillating electric field.

‡ See I. S. Sokolnikoff and R. M. Redheffer, *Mathematics of Physics and Modern Engineering,* second edition (1968, McGraw Hill) chap. 5, sec. 4. If $u = f(x_1, x_2 \ldots x_n)$ is a function of n variables, then the *total derivative* is

$$\mathrm{d}u = \frac{\partial f}{\partial x_1} \mathrm{d}x_1 + \frac{\partial f}{\partial x_2} \mathrm{d}x_2 + \cdots + \frac{\partial f}{\partial x_n} \mathrm{d}x_n$$

You use this all the time without remark in thermodynamics.

expresses energy conservation; leading to the well-known equation of motion (Newton's second law),

$$F = m\ddot{x} \tag{3}$$

where

$$F = -\frac{\mathrm{d}U}{\mathrm{d}x}$$

is a conservative force, namely the negative gradient of the potential energy. If I take it that the linear momentum is $p = m\dot{x}$ then Newton's second law is

$$F = \dot{p}$$

In words, "rate of change of momentum is force." As long as forces *are* conservative as in elementary classical mechanics and in electrostatics, this if fine; but it is insufficient to describe the magnetic force, which is *not* conservative—meaning that it cannot be derived from the gradient of a potential energy.

2. Lagrangian formulation of mechanics

The way to generalise the equation of motion is to introduce a function called the Lagrangian, $L(x, \dot{x}, t)$. The form of this function depends on the nature of the system under study. In the case that we have considered above the Lagrangian is

$$L(x, \dot{x}, t) = K(\dot{x}) - U(x)$$

independent of time. Suppose we *define* the linear momentum as

$$p \stackrel{\mathrm{def}}{=} \frac{\partial L}{\partial \dot{x}} \tag{4a}$$

and the Hamiltonian as†

$$H \stackrel{\mathrm{def}}{=} p\dot{x} - L \tag{4b}$$

We can immediately check consistency with (1) as follows.

$$p = \frac{\partial L}{\partial \dot{x}} = \frac{\partial}{\partial \dot{x}} \left(\frac{1}{2} m\dot{x}^2 - U(x) \right) = m\dot{x}$$

---

† In three dimensions we'd generalise this to

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} \;\; ; \quad H = \mathbf{p} \cdot \dot{\mathbf{r}} - L$$

or in index notation with the summation convention,

$$p_i = \frac{\partial L}{\partial \dot{x}_i} \;\; ; \quad H = p_k \dot{x}_k - L$$

as expected; and the Hamiltonian becomes

$$\begin{aligned} H &= p\dot{x} - L \\ &= m\dot{x}\dot{x} - \frac{1}{2}m\dot{x}^2 + U(x) \\ &= \frac{1}{2}m\dot{x}^2 + U(x) \end{aligned}$$

as expected. In conclusion, in "simple" systems we can define a Hamiltonian and a momentum consistent with a Lagrangian

$$L(x, \dot{x}, t) = K(\dot{x}) - U(x)$$

3. Time-dependent potentials and energy conservation

Let us now examine energy conservation in the Lagrangian formulation of mechanics. We start with the time variation of the Hamiltonian, using (4),

$$\begin{aligned} \frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}}\dot{x} - L\right) \\ &= \dot{x}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} + \ddot{x}\frac{\partial L}{\partial \dot{x}} - \dot{x}\frac{\partial L}{\partial x} - \ddot{x}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial t} \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x}\right)\dot{x} - \frac{\partial L}{\partial t} \end{aligned} \tag{5}$$

In the second line we used (2). We now address two cases.

1. The Lagrangian is not time-dependent. Energy conservation, $\dot{H} = 0$, leads to, using (5),

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0 \tag{6}$$

This is called the *Euler–Lagrange* equation.

2. The Lagrangian is time-dependent. In this case the potential is time-dependent and the correct equation of motion is

$$m\ddot{x} = -\frac{\partial U(x,t)}{\partial x}$$

while

$$L = \frac{1}{2}m\dot{x}^2 - U(x,t)$$

leads to

$$H = \frac{1}{2}m\dot{x}^2 + U(x,t)$$

The rate of energy change along the particle's trajectory is

$$\begin{aligned}\frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m\dot{x}^2 + U(x,t)\right) \\ &= m\dot{x}\ddot{x} + \dot{x}\frac{\partial U}{\partial x} + \frac{\partial U}{\partial t} \\ &= \left(m\ddot{x} + \frac{\partial U}{\partial x}\right)\dot{x} + \frac{\partial U}{\partial t} \\ &= -\frac{\partial L}{\partial t} \neq 0 \end{aligned} \tag{7}$$

because the two terms in large parentheses cancel by virtue of Newton's second law (3). So if I put this result,

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{\partial L}{\partial t}$$

into (5) then I recover

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$$

so the Euler–Lagrange equation applies also to the time-dependent case, even though as (7) insists, energy is *not* conserved.

4. Electrodynamics and the Lorentz force

Experimentally, the force on a particle carrying charge, q, is the Lorentz force,

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v}\times\mathbf{B}\right)$$

If $\mathbf{B} = 0$ then the force is conservative and is given by $-\boldsymbol{\nabla}U(\mathbf{r},t) = -q\boldsymbol{\nabla}V(\mathbf{r},t)$, where $V(\mathbf{r},t)$ is the electric potential. If $\mathbf{B}$ is present then the force depends on the velocity of the particle as well as its position; so the force is non conservative and so cannot be obtained as (minus) the gradient of a potential:

$$\mathbf{F} \neq -\boldsymbol{\nabla}U(\mathbf{r},t)$$

which means that

$$L \neq \frac{1}{2}mv^2 - U(\mathbf{r},t)$$

We can show, however, that the Lorentz force can be derived from the Euler–Lagrange equation by introducing a more general Lagrangian which includes both the scalar potential, $V(\mathbf{r},t)$, and a new term involving the vector potential $\mathbf{A}(\mathbf{r},t)$. If you like, we *postulate* the Lagrangian

$$L = \frac{1}{2}mv^2 - qV + q\mathbf{A}\cdot\mathbf{v}$$

Or, in index notation using the summation convention,

$$L = \frac{1}{2}m\dot{x}_k\dot{x}_k - qV + qA_k\dot{x}_k$$

and see where it takes us. Firstly, the i^{th}-component of momentum is now

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + qA_i \tag{8}$$

The three Euler–Lagrange equations are ($i = 1 \ldots 3$),

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{x}_i + qA_i\right) - \frac{\partial}{\partial x_i}\left(\frac{1}{2}m\dot{x}_k\dot{x}_k - qV + qA_k\dot{x}_k\right) \\ &= m\ddot{x}_i + q\frac{\mathrm{d}A_i}{\mathrm{d}t} - q\frac{\partial}{\partial x_i}\left(A_k\dot{x}_k - V\right) \end{aligned} \tag{9}$$

Examine the term

$$\frac{\mathrm{d}A_i}{\mathrm{d}t}$$

There are two reasons why the particle sees $\mathbf{A}$ to be changing with time: (i) the magnetic field may itself be time-varying, and (ii) the particle is moving and the field may be *spatially* varying. We must include both those possibilities by writing down

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \frac{\partial \mathbf{A}}{\partial t} + \frac{\partial \mathbf{A}}{\partial \mathbf{r}} \cdot \frac{\partial \mathbf{r}}{\partial t} = \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \cdot \frac{\partial \mathbf{A}}{\partial \mathbf{r}}$$

This is the total, or "convective" derivative. In index notation using the summation convention,

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = \frac{\partial A_i}{\partial t} + \dot{x}_k\frac{\partial A_i}{\partial x_k}$$

And so (9) is

$$m\ddot{x}_i = q\left(-\frac{\partial V}{\partial x_i} - \frac{\partial A_i}{\partial t}\right) + q\left(-\dot{x}_k\frac{\partial A_i}{\partial x_k} + \dot{x}_k\frac{\partial A_k}{\partial x_i}\right) \tag{10}$$

Now, because the electric field is

$$\mathbf{E} = -\boldsymbol{\nabla} V - \frac{\partial \mathbf{A}}{\partial t} \qquad \text{or,} \qquad E_i = -\frac{\partial V}{\partial x_i} - \frac{\partial A_i}{\partial t}$$

the first term in parentheses is obviously qE_i. With a little thought you can convince yourself that in the second parentheses two terms cancel and the remaining four are the i^{th}-component of $\mathbf{v} \times \boldsymbol{\nabla} \times \mathbf{A} = \mathbf{v} \times \mathbf{B}$. It therefore follows that (10) is nothing else but the Lorentz force.

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)$$

5. The Hamiltonian in electric and magnetic fields

Now that we have a Lagrangian that is consistent with the Lorentz force we can construct the Hamiltonian (4b)

$$\begin{aligned} H &= p_i\dot{x}_i - L \\ &= p_i\dot{x}_i - \frac{1}{2}m\dot{x}_i\dot{x}_i - q\dot{x}_iA_i + qV \\ &= \frac{1}{2}m\dot{x}_i\dot{x}_i + qV \end{aligned}$$

where in the last step I used (8). It makes sense. It's kinetic plus potential energy; but I want it in terms of the momentum and the vector and scalar potentials. So I note that, using (8) again, (the summation convention will apply after expanding the parentheses)

$$\frac{1}{2m}\left(p_i - qA_i\right)\left(p_i - qA_i\right) = \frac{1}{2}m\dot{x}_i\dot{x}_i$$

so I can write the Hamiltonian in this way,

$$\begin{aligned} H &= \frac{1}{2m}\left(p_i - qA_i\right)\left(p_i - qA_i\right) + qV \\ &= \frac{1}{2m}\left(\mathbf{p} - q\mathbf{A}\right)^2 + qV \end{aligned}$$

6. Magnetic field in quantum mechanics—the Zeeman effect

This is of particular importance in quantum mechanics when dealing with particles in magnetic fields. As always, to quantise the classical Hamiltonian we make the canonical substitution,

$$p \to -i\hbar\frac{\partial}{\partial x}$$

or in three dimensions

$$p_i \to -i\hbar\frac{\partial}{\partial x_i} \quad \text{or} \quad \mathbf{p} \to -i\hbar\boldsymbol{\nabla}$$

and we respect the canonical commutation rules. So the Hamiltonian operator for a quantum particle in electric and magnetic fields is

$$\begin{aligned} \hat{H} &= \frac{1}{2m}\left(-i\hbar\boldsymbol{\nabla} - q\mathbf{A}\right)^2 + qV \\ &= -\frac{\hbar^2}{2m}\nabla^2 + \frac{i\hbar q}{2m}\left(\boldsymbol{\nabla}\cdot\mathbf{A} + \mathbf{A}\cdot\boldsymbol{\nabla}\right) + \frac{q^2}{2m}A^2 + qV \end{aligned}$$

Be a bit careful over $\mathbf{A}\cdot\boldsymbol{\nabla}$. $\hat{H}$ is an operator and this particular term displays its nakedness. It needs to act on a function; then you can see how the next step works. For any arbitrary function ψ, using the differentiation of a product rule,

$$\boldsymbol{\nabla}\cdot(\mathbf{A}\psi) = \psi\boldsymbol{\nabla}\cdot\mathbf{A} + \mathbf{A}\cdot(\boldsymbol{\nabla}\psi)$$

In the Lorenz gauge

$$\boldsymbol{\nabla}\cdot\mathbf{A}+\frac{1}{c^2}\frac{\mathrm{d}V}{\mathrm{d}t}=0$$

and so

$$\boldsymbol{\nabla}\cdot(\mathbf{A}\psi)=\mathbf{A}\cdot(\boldsymbol{\nabla}\psi)-\psi\frac{1}{c^2}\frac{\mathrm{d}V}{\mathrm{d}t}$$

Since ψ is arbitrary we may interpret this as an operator identity in the Lorenz gauge,

$$\boldsymbol{\nabla}\cdot\mathbf{A}=\mathbf{A}\cdot\boldsymbol{\nabla}-\frac{1}{c^2}\frac{\mathrm{d}V}{\mathrm{d}t}$$

and this casts the Hamiltonian into the following form,

$$\hat{H}=-\frac{\hbar^2}{2m}\nabla^2+\frac{i\hbar q}{m}\mathbf{A}\cdot\boldsymbol{\nabla}+\frac{q^2}{2m}A^2+qV-\frac{i\hbar q}{2mc^2}\frac{\mathrm{d}V}{\mathrm{d}t}$$

In the absence of an electric field and in a weak magnetic field for which A^2 may be neglected, only the first two terms survive.

$$\hat{H}=-\frac{\hbar^2}{2m}\nabla^2+\frac{i\hbar q}{m}\mathbf{A}\cdot\boldsymbol{\nabla}$$

contains just the standard quantum kinetic energy operator and an operator that describes the *Zeeman shift* of the energy of an electron in a weak, uniform, static magnetic field chosen along the z-direction. The electron has charge $-e$ and mass m_e and the Zeeman term in the Hamiltonian is†

$$-\frac{i\hbar e}{m_e}\mathbf{A}\cdot\boldsymbol{\nabla}=-\frac{eB}{2m_e}i\hbar\left(-y\frac{\partial}{\partial x}+x\frac{\partial}{\partial y}\right)=\frac{eB}{2m_e}\hat{\ell}_z$$

where B is the magnetic field and

$$\hat{\ell}_z=-i\hbar\frac{\partial}{\partial\phi}$$

is the orbital angular momentum operator with ϕ the azimuthal angle. The eigenvalue of $\hat{\ell}_z$ is $\hbar$ times the "magnetic quantum number", m_ℓ, and the Zeeman energy shift is

$$\frac{e\hbar}{2m_e}\,m_\ell\,B$$

You may draw the analogy of the energy of a magnetic dipole in a magnetic field, which is $-\mu B$ where μ is the dipole moment. Then the electron is acting as if it had a z-component of magnetic moment

$$\mu=-\left(\frac{e\hbar}{2m_e}\right)\,m_\ell$$

† To get $\mathbf{B}=(0,0,B)$ we need $A_x=-yB/2$, $A_y=xB/2$ and $A_z=0$.

The collection of fundamental constants in the parenthesis is called the Bohr magneton μ_B and is the standard unit of magnetic dipole moment of elementary particles. (See J. M. Cassels, *Basic Quantum Mechanics,* second edition (1982, Macmillan), section 20.) Please note, that we are dealing here with the *orbital* magnetic moment of the electron; we are picturing the electron in a hydrogen atom, in orbit about a proton, whose energy is shifted by the magnetic field through the interaction of the magnetic moment associated with its circular motion and the external field. In fact *in this case* the classical picture turns out to be correct: if you pretend that the circulating electron amounts to a loop of current then you find that the gyromagnetic ratio is classical. This situation *definitely does not hold* in relation to the electron's spin magnetic moment—its gyromagnetic ratio is ever so slightly larger, and significantly not exactly, *twice* the classical value.

7. Action and Feynman's path integral formulation

The Lagrangian may be used to define an "action integral" for a particle moving in one dimension,

$$S = \int_{t_1}^{t_2} L\left(x(t), \dot{x}(t)\right) \mathrm{d}t \tag{11}$$

where we assume that the Lagrangian function is itself time-independent. We do a functional minimisation of the action, meaning that we require the first order variation of S to vanish:

$$0 = \delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} \right) \mathrm{d}t \tag{12}$$

Here we are imagining that the particle follows some trajectory in space and time, under the constraint that is starts out at $x = x_1$ at $t = t_1$ and ends up at $x = x_2$ at $t = t_2$. In between those fixed points the function $\delta x(t)$, everywhere small, is the variation from a given trajectory, $x(t)$; and which is zero at the extremes as fixed by the constraint,

$$\delta x(t_1) = \delta x(t_2) = 0 \tag{13}$$

Note also that

$$\delta \dot{x}(t) = \frac{\mathrm{d}}{\mathrm{d}t} \delta x$$

We integrate (12) by parts, the boundary terms vanishing by virtue of (13),

$$0 = \delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}} \right) \delta x(t) \, \mathrm{d}t$$

and find that for δS to vanish for *any* choice of $\delta x(t)$ requires that

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}} = 0$$

which is the Euler–Lagrange equation. In conclusion, the physically observed trajectory, $x(t)$, is the function that minimises (or more accurately, makes stationary) the action integral (11). Professor De Vita writes,

> ... there is something very deep in this. One could construct an *amplitude* $\exp(iS/\hbar)$ for any possible trajectory $x(t)$. Suppose the particle somehow "takes" *all* these paths, but they "interfere," meaning that you have to sum all the amplitudes to discover what the particle actually does. Surely if the quantity $\hbar$ is very small, only the paths close to the one that makes S stationary will not cancel out. The classical trajectory would emerge from this interference in the limit where $\hbar$ goes to zero. It turns out that quantum mechanics can be formulated in this "sum over paths" way, where $\hbar$ is Planck's constant [divided by 2π]. So the least action principle and the classical trajectory emerge as a limit from quantum mechanics (for a brilliant discussion, see Feynman Lectures, Vol II, chapter 19).

Here's a quote from Freeman J. Dyson in a statement in 1980, as reported in *Quantum Reality: Beyond the New Physics,* (1987) by Nick Herbert.

> Thirty-one years ago [1949], Dick Feynman told me about his "sum over histories" version of quantum mechanics. "The electron does anything it likes," he said. "It just goes in any direction at any speed, forward or backward in time, however it likes, and then you add up the amplitudes and it gives you the wave function." I said to him, "You're crazy." But he wasn't.