School Of Mathematics, Statistics, and Operations Research
Te Kura Mātai Tatauranga, Rangahau Pūnaha

MATH 321/322/323	Applied Mathematics	T1 and T2 2013

Math 321/322/323:
Notes on Lagrangian and Hamiltonian mechanics

Matt Visser

School of Mathematics, Statistics, and Operations Research,
Victoria University of Wellington,
Wellington, New Zealand.

E-mail: matt.visser@mcs.vuw.ac.nz
URL: http://www.mcs.vuw.ac.nz/~visser

Version of 26 February 2013; LaTeX-ed Wednesday, February 27, 2013; 01:24pm

URL:

http://msor.victoria.ac.nz/Courses/MATH321_2013T1
http://msor.victoria.ac.nz/Courses/MATH321_2013T2

Warning:

Semi–Private — Not intended for large-scale dissemination.

These notes are provided as a reading course on Lagrangian and Hamiltonian mechanics.

There are still a lot of rough edges.

Semi–Private — Not intended for large-scale dissemination.

Contents

Chapter 1

Introduction to Lagrangian and Hamiltonian mechanics

1.1 General background

Lagrangian and Hamiltonian mechanics was developed in the 1800's in an attempt to move classical Newtonian mechanics beyond the level of just saying "force equals mass times acceleration",

$$\vec{F} = m\,\vec{a}.$$

Of course in a complicated system containing lots of individual bits and pieces no one can stop you from simply adding up the forces, and masses, and calculating the accelerations of all the individual bits and pieces — but at a certain stage this becomes both impractical and inelegant.

In particular during the late 1700s and 1800s one of the biggest scientific questions at the forefront of research had to do with "celestial mechanics" — this is the mechanics of the solar system under Newton's inverse-square law of gravitation *once one takes account of the interactions between the planets* (in addition to the dominant effect due to the Sun on the individual planets, which is easy to deal with).

Lagrangian mechanics (and later on Hamiltonian mechanics) was developed in an attempt to construct a general mathematical framework for handling (in principle) arbitrarily complicated systems, and an attempt at extracting general mathematical theorems based on Newtonian mechanics — such as statements about when exactly energy and momentum are conserved in appropriately defined sub-systems. (As far as we can tell, and with appropriate technical qualifications, the *total* energy and momentum of the universe is conserved, but in smaller sub-systems the presence of friction can cause heat generation, and the sub-system might lose energy to its environment.)

Mathematically, the key tool used in developing this theory of "analytical mechanics" is the so-called "calculus of variations". The "calculus of variations'" asks questions about certain types of integral, and asks when those integrals might take on values that are *maximum*, *minimum*, or more generally *extremal*. (That is, when does the integral have a "point of inflexion" as you vary the things that are being integrated over?) These are purely mathematical questions with purely mathematical answers.

- **Leonhard Euler** (1707–1783) is generally credited with the systematic development of the "calculus of variations", (and in particular discovered of what are now called the Euler–Lagrange equations circa 1750).
- The 19th, 20th, and 23rd "**Hilbert problems**" of 1900 were based on the "calculus of variations" — see the appendix for details.

Physically, the question then arises as to whether or not these "extremality conditions" derived from the "calculus of variations" have anything to do with real world mechanical problems (and optical problems for that matter). The answer of course is yes (otherwise I would not be putting this set of notes together):

- **Joseph-Louis Lagrange**, (born **Giuseppe Lodovico Lagrangia**), developed what is now called "Lagrangian mechanics" during the period 1772–1788. The key insight here was to take a quantity (now called the Lagrangian) that depends on the positions and velocities of the particles you are interested in, integrate that quantity with respect to time, and *extremize* the resulting integral (that integral now being called the "action"). By doing this Lagrange was able to recover Newton's equations and so reformulate classical mechanics in a very elegant and powerful form.[1]
- **William Rowan Hamilton** developed what is now known as (classical) "Hamiltonian mechanics" around 1827–1833. The Hamiltonian formulation extends the Lagrangian formulation and focuses on slightly different issues. The key insight here was to take a different quantity (now called the Hamiltonian, and very closely related to the *energy*) that depends on the positions and *momenta* of the particles you are interested in, integrate that quantity with respect to time, and *extremize* the resulting integral (that integral now being called the "action"). By doing this Hamilton was again able to recover Newton's equations and so reformulate classical mechanics in a very elegant and powerful form.

[1] There are people out there who view Lagrangian mechanics as "proof" of the philosophical assertion by Leibnitz that we live in "the best of all possible worlds". Apart from the sheer lunacy of attempting to use mathematics to "prove" a philosophical point, a quick glance at Lagrangian mechanics should convince you that the very kindest philosophical interpretation (mis-interpretation) of what Lagrange mathematically achieved is that we live in "the most extreme of all possible worlds".

Both of these reformulations of classical mechanics come under the rubric of "the principle of least action" — and there had been a long tradition of ideas along these lines — suggestions that you should be trying to *maximize/ minimize/ extremize* something to give you a handle on classical mechanics and/or ray optics.

- **Pierre de Fermat** developed his "principle of least time" in 1662 — this should more accurately be renamed the "principle of extremal time". According to this principle right rays travel through a refractive medium, and/or are reflected by mirrors in such a way that when you fix the endpoints of the light ray, the actual path followed by the light ray *extremizes* the time taken to traverse that path.
  - You can use this principle of extremal time, for instance, to *derive* both the standard laws of reflection (specular reflection), and **Snell**'s law of refraction.
  - An early version of the principle of extremal time, applying only to mirrors, is sometimes attributed to **Heron of Alexandria** (circa 60 CE).
  - An early version of the principle of extremal time applied to refraction is sometimes attributed to **Ibn al-Haytham**, more commonly known in Europe as **Alhacen** (circa 1021 CE) .
- **Pierre Louis Maupertuis** developed a (rather metaphysical and imprecise) version of the "principle of least action" circa 1741–1746; again, from a modern perspective this should more accurately be described as a "principle of extremal action".

  (Note that at that time, even among leading experts, things were sufficiently confused that the concepts of what we now call *momentum* $\vec{p} = m\vec{v}$ was often confused with the concept of what we now call *kinetic energy* $K = \frac{1}{2}mv^2$ — the Latin phrase "vis viva" was often interchangeably used for both concepts, so in some sense it was no wonder that it was difficult for Maupertuis to come up with a precise formulation of his ideas — for that one had to wait another 40 years for Lagrange and then yet another 50 years for Hamilton.)[2]

In modern theoretical physics:

- **Emmy Noether** developed a mathematical theorem in 1918 that very closely related conservation laws (for example energy conservation and momentum conservation) to symmetries (for example time translation invariance and space translation invariance).
- More radically, essentially all of *quantum field theory* and *classical general relativity* are based on a suitably defined "principle of extremal action".

[2]Oddly enough the Maupertuis form of the variational principle has recently taken on a new life in the form of Julian Barbour's "shape mechanics". This is a seriously new bleeding edge variation on the theme of Einstein's general relativity. You might wish to amuse yourself with a Google search.

– Essentially you will write down some appropriate Lagrangian as the one-formula summary of your theory. **David Hilbert** was able to summarize general relativity in one line — the Lagrangian for general relativity is really absurdly simple (if you understand Riemannian differential geometry). More recently the entire "standard model of particle physics" can be summarized in a $\frac{1}{4}$-page Lagrangian.

– Once you somehow have your Lagrangian:

  * Integrate the Lagrangian to get the action.
  * Write down the Euler-Lagrange equations to investigate the classical limit.
  * Perform a **Feynman** "integral over all possible classical configurations" to investigate the quantum physics.
    (This could be either a "path integral" for quantum mechanics, or more generally an "integral over all possible field configurations" for quantum filed theory)

– It can be shown, (and no, we are not about to do this), that the "stationary phase approximation" to evaluating **Richard Feynman**'s quantum "integral over all possible classical configurations" automatically leads to the "principle of extremal action".

– That is, from a modern perspective you can *derive* the "principle of extremal action" from quantum physics.

In closing, let me point out that many of the names you have encountered above should be recognizable to you — many of these mathematicians/physicists contributed to *many* fields in both mathematics and physics, all the way from abstract number theory (and the distribution of prime numbers) all the way to very applied questions in planetary orbits, the motion of fluids, optics, etc, etc...

1.2 Plan of this module

- First there will be a quite general and purely mathematical chapter on the "calculus of variations".
- The next chapter will be devoted specifically to Lagrangian mechanics.
- I will then present a chapter specifically on Hamiltonian mechanics.
- Finally, there will be a short chapter on symmetries and conservation laws leading up to Noether's theorem.

Assignments based on these notes can be downloaded from the website:

http://msor.victoria.ac.nz/Courses/MATH322_2013T1/
http://msor.victoria.ac.nz/Courses/MATH322_2013T2/

Chapter 2

Elements of the calculus of variations

2.1 Euler–Lagrange equation

The "calculus of variations" has to do with the study of integrals (defined on some suitable set of functions) and the conditions under which the integral is "extremal"; meaning that the value of the integral is a [local] maximum, minimum, or a "point of inflexion".

The canonical example is to suppose we have a (sufficiently smooth) function

$$L(\cdot, \cdot, \cdot) \tag{2.1}$$

which itself depends on a function $x(t)$, its first derivative $\dot{x}(t) = \mathrm{d}x(t)/\mathrm{d}t$, and might (possibly) also explicitly depend on the parameter t. That is:

$$L = L\left(\dot{x}(t), x(t), t\right). \tag{2.2}$$

Now consider the integral

$$S[a, b; x(t)] = \int_a^b L\left(\dot{x}(t), x(t), t\right) \mathrm{d}t. \tag{2.3}$$

This is a *functional* mapping some suitable set $\{x(t)\}$ of functions $x(t)$ into the real numbers $\mathbb{R}$. Under what conditions is this integral extremal?

Notation: Once we get around to doing mechanics, the function $L\left(\dot{x}(t), x(t), t\right)$ will be called the **Lagrangian**, and the functional $S[a, b; x(t)]$ will be called the **action**. The symbols L and S have become quite standard over the last few centuries. ♢

Theorem 1 (Euler–Lagrange equation)
The integral

$$S[a, b; x(t)] = \int_a^b L\left(\dot{x}(t), x(t), t\right) \mathrm{d}t \tag{2.4}$$

is extremal *(maximum/ minimum/ point of inflexion)* if and only if *the function* $x(t)$ *satisfies the differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)}\right] = \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)}. \tag{2.5}$$

This equation was first derived by Euler, and is now called the Euler–Lagrange equation. It is the *basic equation of the calculus of variations.*

Proof: To analyse the question of extremality, write

$$x(t) \to x(t) + \delta x(t), \tag{2.6}$$

and note that by definition

$$S[a, b; x(t) + \delta x(t)] = \int_a^b L\left(\dot{x}(t) + \frac{\mathrm{d}}{\mathrm{d}t}[\delta x(t)], x(t) + \delta x(t), t\right) \mathrm{d}t. \tag{2.7}$$

Now expand $L(\cdot, \cdot, \cdot)$ as a Taylor series in its first two arguments, so that

$$\begin{aligned} S[a, b; x(t) + \delta x(t)] &= S[a, b; x(t)] \\ &+ \int_a^b \left\{\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)} \frac{\mathrm{d}}{\mathrm{d}t}[\delta x(t)] + \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)} \delta x(t)\right\} \mathrm{d}t \\ &+ O\left([\delta x(t)]^2\right). \end{aligned} \tag{2.8}$$

Now integrate by parts.

Then

$$\begin{aligned} \delta S[a, b; x(t)] &= \left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)} \delta x(t)\right]\Bigg|_a^b \\ &+ \int_a^b \left\{-\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)}\right] + \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)}\right\} \delta x(t) \, \mathrm{d}t \\ &+ O\left([\delta x(t)]^2\right). \end{aligned} \tag{2.9}$$

Now let us restrain the set of functions $\{x(t)\}$ to consist only of functions $x(t)$ that are fixed at the end-points a and b.

That is, we consider

$$\delta x(a) \equiv 0 \equiv \delta x(b). \tag{2.10}$$

For this set of functions the integral $S[a, b; x(t)]$ is extremal, (meaning $\delta S[a, b; x(t)] = 0$ for *all* $\delta x(t)$ satisfying the endpoint constraints), *if and only if* for arbitrary $\delta x(t)$ satisfying the endpoint constraints we have the intermediate result

$$\int_a^b \left\{-\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)}\right] + \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)}\right\} \delta x(t) \, \mathrm{d}t = 0. \tag{2.11}$$

We now have to invoke the fundamental theorem of variational calculus.

Theorem 2 (Fundamental theorem of variational calculus)
Let $f(t)$ be a sufficiently smooth function on the closed interval $[a, b]$. Suppose that for every smooth function $\delta x(t)$ satisfying $\delta x(a) = 0 = \delta x(b)$ we know

$$\int_a^b f(t)\ \delta x(t)\ \mathrm{d}t = 0. \tag{2.12}$$

Then we can conclude

$$f(t) = 0 \qquad (\forall t \in [a, b]). \tag{2.13}$$

Proof: Note that the function $-(t-a)(t-b)$ is positive on (a, b) and zero at the end-points a and b. Then let

$$\delta x(t) = -(t-a)(t-b) f(t) \tag{2.14}$$

This particular $\delta x(t)$ satisfies the hypotheses of the theorem. But then

$$0 = \int_a^b f(t)\ \delta x(t)\ \mathrm{d}t = \int_a^b -(t-a)(t-b) f(t)^2\ \mathrm{d}t. \tag{2.15}$$

Since $-(t-a)(t-b)$ is positive on (a, b), and $f(t)^2$ is non-negative on (a, b), this can only be true of $f(t) = 0$ on (a, b). $\diamondsuit$

Now applying the "fundamental theorem of variational calculus" to the intermediate result we have already derived

$$\int_a^b \left\{ -\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)} \right] + \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)} \right\}\ \delta x(t)\ \mathrm{d}t = 0, \tag{2.16}$$

we now see that we have derived the result we were aiming for:

- **The Euler–Lagrange equation.**

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)} \right] = \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)}. \tag{2.17}$$

- It is relatively common to suppress the argument t of the functions $\dot{x}(t)$ and $x(t)$, and merely write

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial L\left(\dot{x}, x, t\right)}{\partial \dot{x}} \right] = \frac{\partial L\left(\dot{x}, x, t\right)}{\partial x}. \tag{2.18}$$

  If you ever run across this form of the equations you are supposed to be bright enough to realize that the mathematically precise statement requires $x \to x(t)$ and $\dot{x} \to \dot{x}(t)$.

- It is also reasonably common to suppress all the arguments of the integrand L and even more simply write:
$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}}\right] = \frac{\partial L}{\partial x}. \tag{2.19}$$
Note that this will only make sense if you are implicitly requiring
$$L \leftrightarrow L(\dot{x}, x, t) \leftrightarrow L(\dot{x}(t), x(t), t). \tag{2.20}$$
- For the time being, I will keep all function arguments and integrand arguments *explicit*; I'll let you know if it ever becomes useful to suppress them.

2.2 Functional derivative

It is quite common to *define* the notion of a *functional derivative* (or *functional gradient*)
$$\frac{\delta S}{\delta x(t)}, \tag{2.21}$$
in terms of the so-called Frechet derivative[1], by the equation
$$\int_a^b \frac{\delta S}{\delta x(t)}\; \delta x(t)\; \mathrm{d}t = \lim_{\epsilon\to 0}\left\{\frac{S[a,b,x(t)+\epsilon\delta x(t)] - S[a,b,x(t)]}{\epsilon}\right\}. \tag{2.22}$$
This is essentially the derivative of the functional $S[a,b,x(t)]$ in the "direction" $\delta x(t)$.

Adopting this *definition* of functional derivative it is a brief calculation, using logic entirely equivalent to the derivation of the Euler–Lagrange equation above, to see that
$$\frac{\delta S}{\delta x(t)} = -\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\,(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\right] + \frac{\partial L\,(\dot{x}(t), x(t), t)}{\partial x(t)}, \tag{2.23}$$
or equivalently
$$\frac{\delta S}{\delta x} = -\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\,(\dot{x}, x, t)}{\partial \dot{x}}\right] + \frac{\partial L\,(\dot{x}, x, t)}{\partial x}, \tag{2.24}$$
or even
$$\frac{\delta S}{\delta x} = -\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}}\right] + \frac{\partial L}{\partial x}. \tag{2.25}$$
Adopting this formalism, the Euler–Lagrange equation becomes the "obvious" statement that at any extremal of the action we must have
$$\frac{\delta S}{\delta x(t)} = 0. \tag{2.26}$$
That is, the "functional gradient" of the action S vanishes at any extremum $x(t)$.

[1] Some of you may previously have seen the Frechet derivative, most likely in Math301 (Calculus 3).

2.3 Notes

Note: There are enormously many generalizations and applications of the Euler–Lagrange equation — for instance:

- to functions $x^i(t)$ in $I\!R^n$;
- to higher derivatives;
- and to fields $\phi(t, x^i)$ defined over $I\!R^{d+1}$.

$\diamondsuit$

Note: The use of t as a parameter and $x(t)$ as the function is just a convention, we could just as well use x as the parameter and $f(x)$ as the function, and so write

$$L = L\left(f'(x), f(x), x\right). \tag{2.27}$$

Now consider the integral

$$S[a, b; f(x)] = \int_a^b L\left(f'(x), f(x), x\right) \mathrm{d}x, \tag{2.28}$$

which would then lead to the Euler–Lagrange equation in the form

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial L\left(f'(x), f(x), x\right)}{\partial f'(x)}\right] = \frac{\partial L\left(f'(x), f(x), x\right)}{\partial f(x)}. \tag{2.29}$$

Mathematically this $f(x)$ form of the Euler–Lagrange equation is completely equivalent to the $x(t)$ form.

(*Physically*, once we turn to applications in classical mechanics, this choice of notation would make it more difficult to understand what is going on.) $\diamondsuit$

Note: In a mechanics context, the integrand $L\left(\dot{x}(t), x(t), x\right)$ is typically referred to as the **Lagrangian**; while the functional $S[a, b; x(t)]$ is typically referred to as the **action**. $\diamondsuit$

Example: For our first and simplest example, consider the special case

$$L\left(\dot{x}(t), x(t)\right) = \frac{1}{2}\, m \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2 - V\left(x(t)\right). \tag{2.30}$$

Then

$$\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)} = m\, \frac{\mathrm{d}x}{\mathrm{d}t}, \tag{2.31}$$

and

$$\frac{\partial L\left(\dot{x}(t), x(t)\right)}{\partial x(t)} = -\frac{\mathrm{d}V}{\mathrm{d}x}. \tag{2.32}$$

So the Euler–Lagrange equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m\,\frac{\mathrm{d}x}{\mathrm{d}t}\right] = -\frac{\mathrm{d}V}{\mathrm{d}x}. \tag{2.33}$$

If m is time independent then

$$m\,\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\frac{\mathrm{d}V}{\mathrm{d}x}. \tag{2.34}$$

If we choose to interpret m as mass, x as position, t as time, and $V(x)$ as a potential, then this is just Newton's second law, $\vec{F} = m\,\vec{a}$ for a force

$$\vec{F} = -\frac{\mathrm{d}V}{\mathrm{d}x}. \tag{2.35}$$

◇

2.4 Euler–Lagrange equations for a function defined in $\mathbb{R}^n$

Suppose we have a collection of functions $x^i(t)$ with $i \in \{1, 2, 3, \ldots, n\}$, that is $x^i(t) \in \mathbb{R}^n$. Then we are interested in the integrand

$$L(\dot{x}^i(t), x^i(t), t), \tag{2.36}$$

and are interested in extremizing the functional

$$S[a, b, x^i(t)] = \int_a^b L(\dot{x}^i(t), x^i(t), t)\ \mathrm{d}t. \tag{2.37}$$

We can now repeat exactly the same analysis as before — with the minor change that we now need to vary all n of functions the $\delta x^i(t)$ independently — to do this we merely need to add a suitable superscript (such as ^{i}) on all the x's. The new Euler–Lagrange equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{x}^j(t), x^j(t), t\right)}{\partial \dot{x}^i(t)}\right] = \frac{\partial L\left(\dot{x}^j(t), x^j(t), t\right)}{\partial x^i(t)}. \tag{2.38}$$

2.5 Euler–Lagrange equations for higher-derivatives

Suppose now that $L(\cdots)$ depends not only on the function $x(t)$ and its first derivative, but also on second and higher derivatives up to order N. That is

$$L = L\left(\frac{\mathrm{d}^N}{\mathrm{d}t^N}x(t), \cdots, \ddot{x}(t), \dot{x}(t), x(t), t\right). \tag{2.39}$$

In this case, by looking at variations $\delta x(t)$ subject to the conditions that at the end points

$$\left.\frac{\mathrm{d}^n[\delta x(t)]}{\mathrm{d}t^n}\right|_a = 0 = \left.\frac{\mathrm{d}^n[\delta x(t)]}{\mathrm{d}t^n}\right|_b \qquad n \in \{0, 1, 2, \dots, N-1\}, \tag{2.40}$$

the Euler–Lagrange equations generalize to

$$\sum_{n=0}^{N}(-)^n \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)}\left(\frac{\mathrm{d}^N}{\mathrm{d}t^N}x(t), \cdots, \ddot{x}(t), \dot{x}(t), x(t), t\right)\right] = 0. \tag{2.41}$$

To see how this comes about we simply copy the original argument above, now noting that in the Taylor expansion of $S[a, b, \delta x(t)]$ there will be terms such as

$$\int_a^b \frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)} \frac{\mathrm{d}^n[\delta x(t)]}{\mathrm{d}t^n}\,\mathrm{d}t \qquad n \in \{0, 1, 2, 3, \dots, N\}. \tag{2.42}$$

Integrating by parts once this becomes

$$\left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)} \frac{\mathrm{d}^{n-1}[\delta x(t)]}{\mathrm{d}t^{n-1}}\right]_a^b - \int_a^b \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)}\right] \frac{\mathrm{d}^{n-1}[\delta x(t)]}{\mathrm{d}t^{n-1}}\,\mathrm{d}t \tag{2.43}$$

In order to be able to safely discard the surface term it is both necessary and sufficient that

$$\left.\frac{\mathrm{d}^n[\delta x(t)]}{\mathrm{d}t^n}\right|_a = 0 = \left.\frac{\mathrm{d}^n[\delta x(t)]}{\mathrm{d}t^n}\right|_b \qquad n \in \{0, 1, 2, \dots, N-1\}, \tag{2.44}$$

in which case we are left with terms of the form

$$-\int_a^b \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)}\right] \frac{\mathrm{d}^{n-1}[\delta x(t)]}{\mathrm{d}t^{n-1}}\,\mathrm{d}t. \tag{2.45}$$

Now integrate by parts a second time. Again discarding the surface term (using exactly the same boundary condition at the endpoints a and b) we have

$$(-)^2 \int_a^b \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)}\right] \frac{\mathrm{d}^{n-2}[\delta x(t)]}{\mathrm{d}t^{n-2}}\,\mathrm{d}t. \tag{2.46}$$

After m iterations we have terms of the form

$$(-)^m \int_a^b \frac{\mathrm{d}^m}{\mathrm{d}t^m} \left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)} \right] \frac{\mathrm{d}^{n-m}[\delta x(t)]}{\mathrm{d}t^{n-m}} \, \mathrm{d}t. \tag{2.47}$$

The integration by parts process stops once $m = n$, at which stage we have

$$(-)^n \int_a^b \frac{\mathrm{d}^n}{\mathrm{d}t^n} \left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)} \right] [\delta x(t)] \, \mathrm{d}t. \tag{2.48}$$

Adding all such terms we see

$$\delta S = \int_a^b \left\{ \sum_{n=0}^{N} (-)^n \frac{\mathrm{d}^n}{\mathrm{d}t^n} \left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)} \right] \right\} [\delta x(t)] \, \mathrm{d}t. \tag{2.49}$$

If we desire S to be extremal then we must have $\delta S = 0$ for all $\delta x(t)$, (subject to the boundary conditions enunciated above), whence applying the fundamental theorem of the variational calculus we finally have

$$\sum_{n=0}^{N} (-)^n \frac{\mathrm{d}^n}{\mathrm{d}t^n} \left[\frac{\partial L}{\partial(\mathrm{d}^n x/\mathrm{d}t^n)} \right] = 0 \tag{2.50}$$

as desired.

While there is no mathematical problem writing down such "higher-derivative" equations, there are significant technical problems with giving such equations a nice and clean physical interpretation — as far as we can tell, these "higher-derivative" equations do not seem to be fundamental to physical theory.

2.6 Euler–Lagrange equations for a field defined over $\mathbb{R}^{d+1}$

Notation: The symbol ∂_i is an extremely common shorthand for a partial derivative:

$$\partial_i \quad \leftrightarrow \quad \frac{\partial}{\partial x^i} \tag{2.51}$$

This notation cuts down on a lot of writing. ♢

Suppose the integrand $L(\cdots)$ now depends on a quantity $\phi(t, x^j)$ which is a field in the sense of analytical mechanics — that is $\phi(t, x^j)$ is a function of the variables t, and x^i

(typically d-dimensional space x^j and one-dimensional time t).[2] The integrand L depends on first derivatives, $\dot\phi(t,x^j)$ and $\partial_i\phi(t,x_j)$, the field itself $\phi(t,x^j)$, and might also explicitly depend on position x^j and time t.[3] That is

$$L = L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right). \tag{2.52}$$

For simplicity assume the x^j are Cartesian coordinates in a Euclidean geometry. Then integrating over some region $(d+1)$-dimensional region Ω we can write

$$S[\Omega;\phi(t,x^j)] = \int_\Omega L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)\; \mathrm{d}^d x\; \mathrm{d}t. \tag{2.53}$$

Now go through the same sort of steps as previously. Taylor expand the integrand $L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)$, now with respect to its first three arguments $\dot\phi(t,x)$, $\partial_i\phi(t,x)$ and $\phi(t,x)$.

$$\begin{aligned}
\delta S[\Omega;\phi(x)] \;=\; & \int_\Omega \left[\frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial[\dot\phi(t,x^j)]}\right]\; \delta\dot\phi(t,x)\;\; \mathrm{d}^d x \mathrm{d}t \\
+ & \int_\Omega \sum_{k=1}^{n} \left[\frac{\partial L\left(\dot\phi(t,x^i), \partial_i\phi(x^j), \phi(t,x^j), t, x^j\right)}{\partial[\partial_k\phi(x^j)]}\right]\; \partial_k[\delta\phi(t,x)]\;\; \mathrm{d}^d x \mathrm{d}t \\
+ & \int_\Omega \frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial\phi(x^j)}\; \delta\phi(x^j)\;\; \mathrm{d}^d x \mathrm{d}t \\
+ & \; O\left([\delta\phi(t,x^j)]^2\right).
\end{aligned} \tag{2.54}$$

Let $\partial\Omega$ denote the d-dimensional boundary of the $(d+1)$-dimensional region Ω. After an integration by parts using the $(d+1)$-dimensional version of Gauss' theorem, and writing

[2]Within the context of special relativity or general relativity one might group time and space together as space-time, with coordinates $x^a = (t, x^i)$, but we are going to keep space and time separate for now.

[3]For definiteness, you might think of $\phi(t,x^i)$ as the electromagnetic scalar potential, or the Newtonian gravitational potential; though with a little more work and the addition of appropriate extra indices you could think of this $\phi(t,x^j)$ as representing, for instance, electric and magnetic vector fields.

the $(d+1)$-dimensional normal vector as $\hat{n} = (n^t; n^i)$ we have

$$
\begin{aligned}
\delta S[\Omega; \phi(x)] \;\; = \;\; & \int_{\partial\Omega} \left[\frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial[\dot\phi(x^j)]} \; \delta\phi(t,x^j) \right] \; n^t \;\; \mathrm{d}^d(\text{area}) \\
+ \;\; & \int_{\partial\Omega} \sum_{k=1}^{d} \left[\frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial[\partial_k\phi(t,x^j)]} \; \delta\phi(t,x^j) \right] \; n^k \;\; \mathrm{d}^d(\text{area}) \\
+ \;\; & \int_{\Omega} -\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial L\left(\partial_i\phi(x^j), \phi(x^j), x^j\right)}{\partial[\dot\phi(x^j)]} \right] \;\; \delta\phi(t,x^j) \;\; \mathrm{d}^d x \mathrm{d}t \\
+ \;\; & \int_{\Omega} -\sum_{k=1}^{d} \frac{\mathrm{d}}{\mathrm{d}x^k} \left[\frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial[\partial_k\phi(t,x^j)]} \right] \;\; \delta\phi(t,x^j) \;\; \mathrm{d}^d x \mathrm{d}t \\
+ \;\; & \int_{\Omega} \left\{ \frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial\phi(t,x^j)} \right\} \;\; \delta\phi(t,x^j) \;\; \mathrm{d}^d x \mathrm{d}t \\
+ \;\; & O\left([\delta\phi(x^j)]^2\right).
\end{aligned}
\tag{2.55}
$$

Note that I have kept things as explicit as possible, with an exhausting listing of all the indices and arguments to all the functions. Note that I have written $\mathrm{d}/\mathrm{d}x^k$ to emphasize the fact that after the integration by parts you have to use the chain rule to differentiate the entire contents of the square brackets $[\dots]$ with respect to x^k — both the *implicit* and *explicit* occurrences of x^k.

Now assume that $\delta\phi(x) = 0$ on the boundary $\partial\Omega$. This simplifies things so that one has

$$
\begin{aligned}
\delta S[\Omega; \phi(x)] \;\; = \;\; & \int_{\Omega} \Bigg\{ -\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial[\dot\phi(x^j)]} \right] \\
& \qquad -\sum_{k=1}^{n} \frac{\mathrm{d}}{\mathrm{d}x^k} \left[\frac{\partial L\left(\dot\phi(t,x^j, \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial[\partial_k\phi(x^j)]} \right] \\
& \qquad + \frac{\partial L\left(\dot\phi(t,x^j), \partial_i\phi(t,x^j), \phi(t,x^j), t, x^j\right)}{\partial\phi(x^j)} \Bigg\} \; \delta\phi(x^j) \; \mathrm{d}^d x \mathrm{d}t \\
& \qquad + O\left([\delta\phi(x^j)]^2\right).
\end{aligned}
\tag{2.56}
$$

Now apply the obvious multi-dimensional field-theoretic version of the fundamental theorem of variational calculus — it is now easy to see that the relevant field theoretic version of the Euler–Lagrange equations are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{\phi}(t,x^j),\partial_i\phi(t,x^j),\phi(t,x^j),t,x^j\right)}{\partial[\dot{\phi}(t,x^j)]}\right] + \sum_{k=1}^{d}\frac{\mathrm{d}}{\mathrm{d}x^k}\left[\frac{\partial L\left(\dot{\phi}(t,x^j),\partial_i\phi(t,x^j),\phi(t,x^j),t,x^j\right)}{\partial[\partial_k\phi(t,x^j)]}\right] = \frac{\partial L\left(\dot{\phi}(t,x^j,\partial_i\phi(t,x^j),\phi(t,x^j),t,x^j\right)}{\partial\phi(t,x^j)}. \tag{2.57}$$

It is common to adopt the "Einstein summation convention" whereby there is an implicit summation whenever indices are repeated (not indices inside arguments to functions, only repeated indices attached to the functions themselves).

The only such index occurring above is k and using the "Einstein summation convention" one writes the field-theoretic version of the Euler–Lagrange equations as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{\phi}(t,x^j),\partial_i\phi(t,x^j),\phi(t,x^j),t,x^j\right)}{\partial[\dot{\phi}(t,x^j)]}\right] + \frac{\mathrm{d}}{\mathrm{d}x^k}\left[\frac{\partial L\left(\dot{\phi}(t,x^j),\partial_i\phi(t,x^j),\phi(t,x^j),t,x^j\right)}{\partial[\partial_k\phi(t,x^j)]}\right] = \frac{\partial L\left(\dot{\phi}(t,x^j,\partial_i\phi(t,x^j),\phi(t,x^j),t,x^j\right)}{\partial\phi(t,x^j)}. \tag{2.58}$$

I have kept space and time separate so far. Let is now define $(d+1)$-dimensional coordinates $x^a = (t; x^i)$, then we can simplify this a little and write (again with an implied summation, now on the index c):

$$\frac{\mathrm{d}}{\mathrm{d}x^c}\left[\frac{\partial L\left(\partial_a\phi(x^b),\phi(x^b),x^b\right)}{\partial[\partial_c\phi(x^b)]}\right] = \frac{\partial L\left(\partial_a\phi(x^b),\phi(x^b),x^b\right)}{\partial\phi(x^b)}. \tag{2.59}$$

It is quite common to suppress a few indices and write

$$\frac{\mathrm{d}}{\mathrm{d}x^c}\left[\frac{\partial L\left(\partial_a\phi(x),\phi(x),x\right)}{\partial[\partial_c\phi(x^b)]}\right] = \frac{\partial L\left(\partial_a\phi(x),\phi(x),x\right)}{\partial\phi(x)}, \tag{2.60}$$

with the understanding that you *mean* the full version as given above, and that you put appropriate indices back on the x as necessary.

It is also quite common to get a little "fluffy" about the partial derivative notation and to write:

$$\frac{\partial}{\partial x^c}\left[\frac{\partial L\left(\partial_a\phi(x),\phi(x),x\right)}{\partial[\partial_c\phi(x^b)]}\right] = \frac{\partial L\left(\partial_a\phi(x),\phi(x),x\right)}{\partial\phi(x)}, \tag{2.61}$$

with the understanding that the partial derivative acts on all occurrences of x, both explicit and implicit.

Note: The generalization to non–Cartesian coordinates on a Euclidean space is straightforward — try it. $\diamondsuit$

Note: The formalism can be generalized even further by developing the theory of integration on a general curved manifold (as in general relativity for instance). The Euler–Lagrange equations can then be suitably modified to deal with not only special relativity but general relativity as well. $\diamondsuit$

Note: Suppose now that $L(\cdots)$ depends on the field $\phi(x^a)$, plus its first *and* second derivatives. The Euler-Lagrange equations (for Cartesian coordinates in a Euclidean space) are now (suppressing explicit arguments of the function L)

$$\frac{\partial L}{\partial\phi(x^c)} - \frac{\mathrm{d}}{\mathrm{d}x^a}\left[\frac{\partial L}{\partial[\partial_a\phi(x^c)]}\right] + \frac{\mathrm{d}^2}{\mathrm{d}x^a\,\mathrm{d}x^b}\left[\frac{\partial^2 L}{\partial[\partial_a\partial_b\phi(x^c)]}\right] = 0. \tag{2.62}$$

Here we have again adopted the "Einstein summation convention" and there are implied sums on the indices a and b; sums $\sum_{a=1}^{d+1}$ and $\sum_{b=1}^{d+1}$. Similarly it is quite common to write

$$\frac{\partial L}{\partial\phi(x^c)} - \frac{\partial}{\partial x^a}\left[\frac{\partial L}{\partial[\partial_a\phi(x^c)]}\right] + \frac{\partial^2}{\partial x^a\,\partial x^b}\left[\frac{\partial^2 L}{\partial[\partial_a\partial_b\phi(x^c)]}\right] = 0, \tag{2.63}$$

again with the understanding that the partial derivative acts on all occurrences of x, both explicit and implicit. Similarly it is quite common to write

$$\frac{\partial L}{\partial\phi(x)} - \frac{\partial}{\partial x^a}\left[\frac{\partial L}{\partial[\partial_a\phi(x)]}\right] + \frac{\partial^2}{\partial x^a\,\partial x^b}\left[\frac{\partial^2 L}{\partial[\partial_a\partial_b\phi(x)]}\right] = 0. \tag{2.64}$$

The generalization to even higher derivatives is obvious but notationally messy. $\diamondsuit$

2.7 Beltrami's identity

One version of the Beltrami identity is this:

Theorem 3 (Beltrami's identity I)
If the function $x(t)$ satisfies the Euler–Lagrange equation derived from the integrand $L(\dot{x}(t), x(t), t)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot{x}(t), x(t), t) - \dot{x}(t)\,\frac{\partial L(\dot{x}(t), x(t), t)}{\partial\dot{x}(t)}\right] = \frac{\partial L(\dot{x}(t), x(t), t)}{\partial t}. \tag{2.65}$$

Proof: To prove this we first need to use the chain rule to evaluate

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot{x}(t), x(t), t)\right] = \frac{\partial L(\dot{x}(t), x(t), t)}{\partial x(t)}\;\dot{x}(t) + \frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\;\ddot{x}(t) + \frac{\partial L(\dot{x}(t), x(t), t)}{\partial t}. \tag{2.66}$$

Similarly, again using the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\dot{x}(t)\;\frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\right] = \ddot{x}(t)\;\frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)} + \dot{x}(t)\;\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\right]. \tag{2.67}$$

But now, using the Euler–Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\dot{x}(t)\;\frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\right] = \frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\;\ddot{x}(t)\; + \left[\frac{\partial L(\dot{x}(t), x(t), t)}{\partial x(t)}\right]\;\dot{x}(t). \tag{2.68}$$

Combining these two equations there is a significant cancellation and we see

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot{x}(t), x(t), t) - \dot{x}(t)\;\frac{\partial L(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}\right] = \frac{\partial L(\dot{x}(t), x(t), t)}{\partial t}, \tag{2.69}$$

as claimed.

It is common to suppress the arguments of $\dot{x}(t)$ and $x(t)$ and write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot{x}, x, t) - \dot{x}\;\frac{\partial L(\dot{x}, x, t)}{\partial \dot{x}}\right] = \frac{\partial L(\dot{x}, x, t)}{\partial t}, \tag{2.70}$$

or even

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L - \dot{x}\;\frac{\partial L}{\partial \dot{x}}\right] = \frac{\partial L}{\partial t}. \tag{2.71}$$

In particular, if the integrand $L(\dot{x}(t), x(t))$ has no *explicit* time dependence, and $x(t)$ satisfies the Euler–Lagrange equation, then we have a *conservation law.*

Theorem 4 (Beltrami's identity II)
If $L(\dot{x}(t), x(t))$ has no explicit *time dependence, and $x(t)$ satisfies the Euler–Lagrange equation, then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot{x}(t), x(t)) - \dot{x}(t)\;\frac{\partial L(\dot{x}(t), x(t))}{\partial \dot{x}(t)}\right] = 0. \tag{2.72}$$

It is again common to suppress the arguments of $\dot{x}(t)$ and $x(t)$ and write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot{x}, x, t) - \dot{x}\;\frac{\partial L(\dot{x}, x, t)}{\partial \dot{x}}\right] = 0, \tag{2.73}$$

or even

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L - \dot{x}\;\frac{\partial L}{\partial \dot{x}}\right] = 0. \tag{2.74}$$

Once we get around to explicitly discussing mechanics, we will see that Beltrami's identity (which at this stage is a purely *mathematical* result of the *calculus of variations*), will turn into a *physical* statement regarding the *conservation of energy.*

2.8 Beltrami's identity for fields defined over $I\!R^{d+1}$

Let's take the Euler–Lagrange equations for fields $\phi(t,x)$ defined over $I\!R^{d+1}$, and suppress a few of the indices for clarity

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot\phi,\partial\phi,\phi,t,x\right)}{\partial\dot\phi}\right]+\frac{\mathrm{d}}{\mathrm{d}x^k}\left[\frac{\partial L\left(\dot\phi,\partial\phi,\phi,t,x\right)}{\partial[\partial_k\phi]}\right]=\frac{\partial L\left(\dot\phi,\partial\phi,\phi,t,x\right)}{\partial\phi}. \tag{2.75}$$

Can this be used to come up with a field theoretic version of Beltrami's identity? Most definitely yes.

Theorem 5 (Beltrami's identity III)
If the field $\phi(t,x)$ satisfies the Euler–Lagrange equation derived from the field-theory integrand $L(\dot\phi(t),\partial\phi,\phi,t,x)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[L(\dot\phi,\partial\phi,\phi,t,x)-\dot\phi\,\frac{\partial L(\dot\phi,\partial\phi,\phi,t,x)}{\partial\dot\phi}\right]=\frac{\partial L(\dot\phi,\partial\phi,\phi,t,x)}{\partial t}+\frac{\mathrm{d}}{\mathrm{d}x^k}\left[\dot\phi\,\frac{\partial L\left(\dot\phi,\partial_i\phi,\phi,t,x\right)}{\partial[\partial_k\phi]}\right]. \tag{2.76}$$

Proof: Apply the chain rule of differentiation and the Euler–Lagrange equations of motion. $\diamondsuit$

Now integrate over all space — we can use Gauss' theorem to turn the divergence into a surface integral — under the assumption that the field $\phi(t,x)$ has nice falloff behaviour at infinity we can then neglect the surface integral. This yields a spatially integrated version of the Beltrami identity.

Theorem 6 (Beltrami's identity IV)
If the field $\phi(t,x)$ satisfies the Euler–Lagrange equation derived from the field-theory integrand $L(\dot\phi(t),\partial\phi,\phi,t,x)$, and suitable falloff conditions at spatial infinity, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int\left\{L(\dot\phi,\partial\phi,\phi,t,x)-\dot\phi\,\frac{\partial L(\dot\phi,\partial\phi,\phi,t,x)}{\partial\dot\phi}\right\}\mathrm{d}^dx\right]=\int\frac{\partial L(\dot\phi,\partial\phi,\phi,t,x)}{\partial t}\,\mathrm{d}^dx. \tag{2.77}$$

Finally, if there is no explicit time dependence in the integrand L we have the field-theory *conservation law*:

Theorem 7 (Beltrami's identity V)
If the field $\phi(t,x)$ satisfies the Euler–Lagrange equation derived from the field-theory integrand $L(\dot\phi(t), \partial\phi, \phi, x)$, and suitable falloff conditions at spatial infinity, and if L contains no explicit t dependence, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int \left\{ L(\dot\phi, \partial\phi, \phi, x) - \dot\phi\, \frac{\partial L(\dot\phi, \partial\phi, \phi, x)}{\partial\dot\phi} \right\} \mathrm{d}^d x\right] = 0. \tag{2.78}$$

2.9 Summary

The calculus of variations is a *general tool* that has applications in many fields; far beyond the straightforward application to Lagrangian mechanics and its generalizations, the Euler–Lagrange equations are also relevant in classical field theories [such as say Maxwell's electromagnetism or Einstein's general relativity] where they are often the easiest way of obtaining the field equations (equations of motion), and also in *quantum* field theories where (semi-)classical solutions satisfying the Euler–Lagrange equation often dominate the physics.

Chapter 3

Lagrangian mechanics

3.1 Overview

Lagrangian mechanics now takes the ideas of the calculus of variations and specifically applies it to mechanical problems:

- The parameter t is really to be thought of as physical time (typically Newtonian time).
- The function $x(t)$ is to be thought of as position as a function of time — though eventually these will become "generalized position coordinates".
- The integrand $L(x(t), \dot{x}(t), t)$ is now called the "Lagrangian".
- The Euler–Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)}\right] = \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)} \tag{3.1}$$

  can, using the chain rule, be rewritten as

$$\begin{aligned}\left[\frac{\partial^2 L\left(\dot{x}(t), x(t), t\right)}{[\partial \dot{x}(t)]^2}\right]\ddot{x}(t) + \left[\frac{\partial^2 L\left(\dot{x}(t), x(t), t\right)}{[\partial \dot{x}(t)]\,[\partial x(t)]}\right]\dot{x}(t) + \left[\frac{\partial^2 L\left(\dot{x}(t), x(t), t\right)}{[\partial \dot{x}(t)]\partial t}\right] \\ = \frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial x(t)}.\end{aligned} \tag{3.2}$$

  This is now seen to involve *at most* second-order time derivatives of position, and so it is a natural generalization of Newton's second law $m\,\ddot{x} = F$.

As our first slightly nontrivial example, consider the special case

$$L\left(\dot{x}(t), x(t), t\right) = \frac{1}{2}\, m(x(t), t) \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2 - V(x(t), t). \tag{3.3}$$

Then

$$\frac{\partial L\left(\dot{x}(t), x(t), t\right)}{\partial \dot{x}(t)} = m(x(t), t)\, \frac{\mathrm{d}x}{\mathrm{d}t} \tag{3.4}$$

and

$$\frac{\partial L\left(\dot{x}(t), x(t)\right)}{\partial x(t)} = \frac{1}{2}\frac{\partial m(x(t), t)}{\partial x} \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2 - \frac{\partial V(x(t), t)}{\partial x}. \tag{3.5}$$

So the Euler–Lagrange equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m(x(t), t)\, \frac{\mathrm{d}x}{\mathrm{d}t}\right] = \frac{1}{2}\frac{\partial m(x(t), t)}{\partial x} \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2 - \frac{\partial V(x(t), t)}{\partial x}. \tag{3.6}$$

This is just Newton's second law, in the form $\mathrm{d}p/\mathrm{d}t = F$, for a time and position-dependent mass $m(x(t), t)$, a *momentum* $p = m\dot{x}$, and an explicitly time-dependent potential. This is more than just a curiosity:

- Time-dependent masses occur, for instance, in the rocket equation: as the rocket burns it loses mass in the form of exhaust gasses.
- Position-dependent "effective" masses occur, for instance in condensed matter; specifically for the effective masses of electrons and holes in electronic conduction bands.
- Time-dependent potentials are even easier to set up — just apply a time dependent voltage to a particle that has an electric charge.
- The whole point of setting up Lagrangian mechanics is to have a straightforward formalism capable of handling situations just a little more complicated than the easy $F = ma$.

Applying the chain rule to the above we have:

$$m(x(t), t)\, \frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \frac{\partial m(x(t), t)}{\partial x} \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2 + \left[\frac{\partial}{\partial t} m(x(t), t)\right]\, \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{2}\frac{\partial m(x(t), t)}{\partial x} \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2 - \frac{\partial V(x(t), t)}{\partial x}. \tag{3.7}$$

We now rearrange this to:

$$m(x(t), t)\, \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\frac{\partial V(x(t), t)}{\partial x} - \left[\frac{\partial}{\partial t} m(x(t), t)\right]\, \frac{\mathrm{d}x}{\mathrm{d}t} - \frac{1}{2}\frac{\partial m(x(t), t)}{\partial x} \left[\frac{\mathrm{d}x}{\mathrm{d}t}\right]^2. \tag{3.8}$$

So far, everything has been *explicit*. We can get a little sloppy and suppress some of the arguments to obtain the simpler-looking result:

$$m\, \ddot{x}\ = -\frac{\partial V}{\partial x} - \frac{\partial m}{\partial t}\, \dot{x} - \frac{1}{2}\frac{\partial m}{\partial x}\, \dot{x}^2. \tag{3.9}$$

This is enough to tell us that time and position dependent masses effectively mimic velocity dependent pseudo-forces.

3.2 Momentum

Given the above, it is useful to *define* the (generalized) *momentum* as:

$$p(\dot{x}(t), x(t), t) = \frac{\partial L\,(\dot{x}(t), x(t), t)}{\partial \dot{x}(t)}, \tag{3.10}$$

since the Euler–Lagrange equation is then simply

$$\frac{\mathrm{d}}{\mathrm{d}t}\,[p(\dot{x}(t), x(t), t)] = \frac{\partial L\,(\dot{x}(t), x(t), t)}{\partial x(t)}. \tag{3.11}$$

This has a nice simple interpretation in terms of a "generalized Newton's law".

Many people will simply write

$$p(\dot{x}, x, t) = \frac{\partial L\,(\dot{x}, x, t)}{\partial \dot{x}}, \tag{3.12}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\,[p(\dot{x}, x, t)] = \frac{\partial L\,(\dot{x}, x, t)}{\partial x}, \tag{3.13}$$

or even

$$p = \frac{\partial L}{\partial \dot{x}}, \tag{3.14}$$

and

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial L}{\partial x}, \tag{3.15}$$

expecting the reader (that is, you) to apply the relevant context and "fill in the blanks".

3.3 Effective mass

Given the above, it is (sometimes) useful to *define* the (generalized) *effective mass* as:

$$m(\dot{x}(t), x(t), t) = \frac{\partial^2 L\,(x(t), \dot{x}(t), t)}{[\partial \dot{x}(t)]^2}, \tag{3.16}$$

since the Euler–Lagrange equation can then be written as

$$[m(\dot{x}(t), x(t), t)]\;\ddot{x}(t) + \left[\frac{\partial p\,(\dot{x}(t), x(t), t)}{\partial x(t)}\right]\;\dot{x}(t) + \left[\frac{\partial p\,(\dot{x}(t), x(t), t)}{\partial t}\right] = \frac{\partial L\,(\dot{x}(t), x(t), t)}{\partial x(t)}. \tag{3.17}$$

This again has a nice simple interpretation in terms of a "generalized Newton's law". As usual, do not be surprised if you see "simplified" formulae such as

$$[m(\dot{x}, x, t)]\;\ddot{x} + \left[\frac{\partial p\,(\dot{x}, x, t)}{\partial x}\right]\;\dot{x} + \left[\frac{\partial p\,(\dot{x}, x, t)}{\partial t}\right] = \frac{\partial L\,(\dot{x}, x, t)}{\partial x}, \tag{3.18}$$

or even the somewhat ambiguous

$$m\,\ddot{x} + \left[\frac{\partial p}{\partial x}\right]\,\dot{x} + \left[\frac{\partial p}{\partial t}\right] = \frac{\partial L}{\partial x}, \tag{3.19}$$

or

$$m\,\ddot{x} = \frac{\partial L}{\partial x} - \left[\frac{\partial p}{\partial t}\right] - \left[\frac{\partial p}{\partial x}\right]\,\dot{x}. \tag{3.20}$$

Be prepared to "fill in the blanks" as necessary.

3.4 From Beltrami's identity to the conservation of energy

Now consider the special case Lagrangian (suppressing arguments to the functions $\dot{x}(t)$ and $x(t)$ when no confusion can arise):

$$L = \frac{1}{2}\,m\,\dot{x}^2 - V(x). \tag{3.21}$$

Beltrami's identity can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}}\,\dot{x} - L\right] = 0. \tag{3.22}$$

Since for this particular Lagrangian

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}, \tag{3.23}$$

we see

$$\left[\frac{\partial L}{\partial \dot{x}}\,\dot{x} - L\right] = (m\dot{x})\,\dot{x} - \left(\frac{1}{2}m\dot{x}^2 - V(x)\right) = \frac{1}{2}m\dot{x}^2 + V(x). \tag{3.24}$$

That is, Beltrami's identity applied to this particular Lagrangian implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}m\dot{x}^2 + V(x)\right] = 0. \tag{3.25}$$

That is, in this system the *sum* of kinetic energy $\frac{1}{2}m\dot{x}^2$ and potential energy $V(x)$ is conserved; the *total* energy is conserved.

$$E = \frac{1}{2}m\dot{x}^2 + V(x) = (\text{constant}). \tag{3.26}$$

Of course, there are many generalizations of this phenomenon.

3.4.1 Motion in $\mathbb{R}^n$

For motion in $\mathbb{R}^n$ the Euler–Lagrange equations generalize to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}^i}\right] = \frac{\partial L}{\partial x^i}. \tag{3.27}$$

Furthermore, the Beltrami identity now generalizes to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}^i}\, \dot{x}^i - L\right] = 0. \tag{3.28}$$

Now specifically consider a Lagrangian of the form

$$L = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\, \dot{x}^i \dot{x}^j - V(x^i). \tag{3.29}$$

Then the Euler–Lagrange equations become

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{j=1}^{n} m_{ij}\, \dot{x}^j\right] = \frac{\partial V}{\partial x^i}, \tag{3.30}$$

and so

$$\left[\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}^i}\, \dot{x}^i - L\right] = \left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} m_{ij}\, \dot{x}^j\right)\dot{x}^i\right\} - \left\{\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\, \dot{x}^i\dot{x}^j - V(x^i)\right\} \tag{3.31}$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\, \dot{x}^i\dot{x}^j + V(x^i). \tag{3.32}$$

So the Beltrami identity becomes the conservation of energy in $\mathbb{R}^n$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\, \dot{x}^i\dot{x}^j + V(x^i)\right] = 0. \tag{3.33}$$

- If we choose $n = 3$ and $m_{ij} = m\,\delta_{ij}$ then this would correspond to a single particle in 3-dimensional space.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}m\sum_{i=1}^{3} (\dot{x}^i)^2 + V(x^i)\right] = 0. \tag{3.34}$$

  You can again interpret the quantity that is conserved via Beltrami's identity as a *sum* of kinetic energy and potential energy; a *total* energy.

- If we choose $n = 3N$ and choose m_{ij} to be the diagonal matrix

$$m_{ij} = \text{diag}\left\{m_1, m_1, m_1;\ m_2, m_2, m_2;\ \dots;\ m_N, m_N, m_N\right\}, \tag{3.35}$$

then we are describing N particles, of masses $m_1, m_2, \dots, m_N$ respectively, each of which is moving in 3-dimensional space, interacting through some complicated potential $V(x^i)$. You can again interpret the quantity that is conserved via Beltrami's identity as a *sum* of individual kinetic energies for the N particles plus a potential energy. Indeed, let $A \in \{1, 2, \dots, N\}$ label the particle, and let $i \in \{1, 2, 3\}$ label the three dimensions of space, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}\sum_{A=1}^{N}\sum_{i=1}^{3} m_A \ \left(\dot{x}^i_A\right)^2 + V(x^i_A)\right] = 0. \tag{3.36}$$

- For a generic system moving in $I\!R^n$ the x^i are referred to as "generalized coordinates". They may represent positions, angles, whatever...

- The general *definition* of energy for motion of a system in $I\!R^n$ is:

$$E = \sum_{i=1}^{n} \frac{\partial L(\dot{x}, x)}{\partial \dot{x}^i}\ \dot{x}^i - L(\dot{x}, x). \tag{3.37}$$

This purely mathematical definition passes all the sanity checks to correspond to what physicists call the energy.

3.4.2 Fields defined over $I\!R^{d+1}$

For fields $\phi(t, x)$ defined over $I\!R^{d+1}$ we have seen that the Euler–Lagrange equations generalize to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial[\dot{\phi}(x)]}\right] + \sum_{i=1}^{d}\frac{\mathrm{d}}{\mathrm{d}x^i}\left[\frac{\partial L}{\partial[\partial_i\phi(x)]}\right] = \frac{\partial L}{\partial\phi(x)}. \tag{3.38}$$

Furthermore, as we have also already seen, under suitable conditions the Beltrami identity now generalizes to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int\left\{\frac{\partial L}{\partial\dot{\phi}(x)}\ \dot{\phi}(x)\ - L\right\}\mathrm{d}^d x\right] = 0. \tag{3.39}$$

Note that the $\sum_{i=1}^{n}$ that is present for particle motion in $I\!R^n$ has now been replaced by an integral over space, $\int(\dots)\ \mathrm{d}^d x$, for fields defined over $I\!R^{d+1}$. This motivates a definition of (total) energy as

$$E = \int\left\{\frac{\partial L}{\partial\dot{\phi}(x)}\ \dot{\phi}(x)\ - L\right\}\mathrm{d}^d x, \tag{3.40}$$

and strongly suggests it would be a good idea to define an *energy density* as:

$$\rho = \frac{\partial L}{\partial \dot{\phi}(x)} \, \dot{\phi}(x) \; - L, \tag{3.41}$$

so that

$$E = \int \rho \, \mathrm{d}^d x. \tag{3.42}$$

Some further examples along these lines will be part of the assignments.

Chapter 4

Hamiltonian mechanics

Hamiltonian mechanics can be *motivated* on the basis of Lagrangian mechanics; but once you have the motivation in place one can actually discard the original Lagrangian mechanics, and view Hamiltonian mechanics as an independent subject in its own right.

4.1 Hamilton's equations

Start from the Lagrangian-based definition of momentum

$$p(\dot{x}, x, t) = \frac{\partial L(\dot{x}, x, t)}{\partial \dot{x}}. \tag{4.1}$$

In very many interesting cases,[1] but certainly not always, it proves possible to invert this relationship to find $\dot{x}$ as a function of p and x (and possibly t). Let us assume such a inverse function exists and write it as

$$\dot{x}(p, x, t). \tag{4.2}$$

Now construct the quantity

$$H(p, x, t) = p\,\dot{x}(p, x, t) - L(\dot{x}(p, x, t), x, t). \tag{4.3}$$

Of course this definition is ultimately *motivated* by the quantity that we saw occurred in Beltrami's identity, but now we are not (yet) demanding the Euler–Lagrange equations, and are instead *defining* the new quantity $H(p, x, t)$ for arbitrary p and x.

Now compute

$$\frac{\partial H(p, x, t)}{\partial p} = \dot{x} + p\frac{\partial \dot{x}}{\partial p} - \frac{\partial L}{\partial \dot{x}}\frac{\partial \dot{x}}{\partial p} = \dot{x} + p\frac{\partial \dot{x}}{\partial p} - p\frac{\partial \dot{x}}{\partial p} = \dot{x}, \tag{4.4}$$

[1] For example, whenever L is quadratic in $\dot{x}$.

where at this stage all I have used is the various definitions. In counterpoint, consider

$$\frac{\partial H(p,x,t)}{\partial x} = p\frac{\partial \dot{x}}{\partial x} - \frac{\partial L}{\partial \dot{x}}\frac{\partial \dot{x}}{\partial x} - \frac{\partial L}{\partial x} = p\frac{\partial \dot{x}}{\partial x} - p\frac{\partial \dot{x}}{\partial x} - \frac{\partial L}{\partial x} = -\frac{\partial L}{\partial x} = -\frac{\mathrm{d}p}{\mathrm{d}t}, \tag{4.5}$$

where at the last step I have used the Euler–Lagrange equations. Rearranging we see that we have derived Hamilton's equations:

Theorem 8 (Hamilton's equations for Lagragian systems)
For a Lagrangian system define an object that we will call a (classical) Hamiltonian by

$$H(p,x,t) = p\,\dot{x}(p,x,t) - L(\dot{x}(p,x,t),x,t) \tag{4.6}$$

Then, if the underlying system satisfies the Euler–Lagrange equations coming from the Lagrangian $L(\dot{x},x,t)$, we have:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H(p,x,t)}{\partial p}; \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H(p,x,t)}{\partial x}. \tag{4.7}$$

These are known as Hamilton's equations.

4.2 Hamiltonian systems

It is often useful to simply forget the underlying Lagrangian and take the Hamiltonian formalism as *primary*. In this case we adopt a *definition*:

Definition 1 (Hamiltonian system)
A Hamiltonian system is described by two functions $p(t)$ and $x(t)$, called the momentum and position, and characterized by a function $H(p,x,t)$ called the (classical) Hamiltonian, which governs the time evolution of the system via Hamilton's equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H(p,x,t)}{\partial p}; \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H(p,x,t)}{\partial x}. \tag{4.8}$$

In either case, with or without an underlying Lagrangian, we get yet another version of Beltrami's identity. Let us compute:

$$\begin{aligned}
\frac{\mathrm{d}H(p,q,t)}{\mathrm{d}t} &= \frac{\partial H(p,x,t)}{\partial p}\,\dot{p} + \frac{\partial H(p,x,t)}{\partial x}\,\dot{x} + \frac{\partial H(p,x,t)}{\partial t} && (4.9)\\
&= -\frac{\partial H(p,x,t)}{\partial p}\,\frac{\partial H(p,x,t)}{\partial x} + \frac{\partial H(p,x,t)}{\partial x}\,\frac{\partial H(p,x,t)}{\partial p} + \frac{\partial H(p,x,t)}{\partial t} && (4.10)\\
&= \frac{\partial H(p,x,t)}{\partial t}. && (4.11)
\end{aligned}$$

This gives us:

Theorem 9 (Hamiltonian version of the Beltrami identity)
For any system satisfying Hamilton's equations we have

$$\frac{\mathrm{d}H(p,x,t)}{\mathrm{d}t} = \frac{\partial H(p,x,t)}{\partial t}. \tag{4.12}$$

In particular if the Hamiltonian has no explicit *time dependence then*

$$\frac{\mathrm{d}H(p,x)}{\mathrm{d}t} = 0, \tag{4.13}$$

so that

$$H(p,x) = E. \tag{4.14}$$

is a constant of the motion called the energy.

4.3 Variational principle for Hamilton's equations

Let us now construct a suitable variational principle for Hamilton's equations. Consider the functional

$$S[a,b,p(t),x(t)] = \int_a^b \{p(t)\dot{x}(t) - H(p(t),x(t),t)\}\,\mathrm{d}t. \tag{4.15}$$

Now ask that this functional be extremal. We compute

$$\delta S[a,b,p(t),x(t)] = S[a,b,p(t)+\delta p(t),x(t)+\delta x(t)] - S[a,b,p(t),x(t)]. \tag{4.16}$$

Expand as a Taylor series, then

$$\begin{aligned}\delta S[a,b,p(t),x(t)] &= \int_a^b \Big\{\delta p(t)\dot{x}(t) + p(t)\delta\dot{x}(t) \\ &\qquad -\partial_p H(p(t),x(t),t)\;\delta p(t) - \partial_q H(p(t),x(t),t)\;\delta q(t) \\ &\qquad +O[(\delta x)^2] + O[(\delta p)^2]\Big\}\,\mathrm{d}t. \end{aligned} \tag{4.17}$$

Integrate by parts:

$$\begin{aligned}\delta S[a,b,p(t),x(t)] &= [p(t)\,\delta x(t)]_a^b + \int_a^b \Big\{\delta p(t)\,[\dot{x}(t) - \partial_p H(p(t),x(t),t)] \\ &\qquad -\delta q(t)\,[\dot{p}(t) + \partial_q H(p(t),x(t),t)] \\ &\qquad +O[(\delta x)^2] + O[(\delta p)^2]\Big\}\,\mathrm{d}t. \end{aligned} \tag{4.18}$$

As previously, impose boundary conditions

$$\delta x(a) = 0 = \delta x(b), \tag{4.19}$$

so that

$$\delta S[a,b,p(t),x(t)] = \int_a^b \Big\{ \delta p(t)\left[\dot{x}(t) - \partial_p H(p(t),x(t),t)\right] - \delta q(t)\left[\dot{p}(t) + \partial_q H(p(t),x(t),t)\right] + O[(\delta x)^2] + O[(\delta p)^2] \Big\}\, \mathrm{d}t. \tag{4.20}$$

Then

$$\frac{\delta S}{\delta p(t)} = \left[\dot{x}(t) - \partial_p H(p(t),x(t),t)\right]; \qquad \frac{\delta S}{\delta q(t)} = -\left[\dot{p}(t) + \partial_q H(p(t),x(t),t)\right]. \tag{4.21}$$

Consequently $S[a,b,p(t),q(t)]$ is *extremal* if and only if

$$\dot{x}(t) = \partial_p H(p(t),x(t),t); \qquad \dot{p}(t) = -\partial_q H(p(t),x(t),t). \tag{4.22}$$

But these are just another form of Hamilton's equations.

4.4 From Hamiltonian back to Lagrangian

Now suppose someone gives you a Hamiltonian $H(p,x,t)$ as the *primary* definition of your mechanical system. Can you work backwards to find an equivalent Lagrangian? Consider the specific Hamilton equation

$$\dot{x}(t) = \partial_p H(p(t),x(t),t). \tag{4.23}$$

It is relatively common to be able to invert this equation[2] to find

$$p(\dot{x},x,t). \tag{4.24}$$

Let us assume such an inversion exists. Then:

Theorem 10 (Lagrangian reconstruction theorem)

$$L(\dot{x},x,t) = p(\dot{x},x,t)\;\dot{x}(t) - H(p(\dot{x},x,t),x(t),t) \tag{4.25}$$

is a Lagrangian that gives you back the original Euler–Lagrange equation of motion.

Proof: Compute, using Hamilton's equations

$$\frac{\partial L}{\partial \dot{x}} = p(\dot{x},x,t) + \frac{\partial p(\dot{x},x,t)}{\partial \dot{x}}\,\dot{x} - \frac{\partial H}{\partial p}\,\frac{\partial p(\dot{x},x,t)}{\partial \dot{x}} = p(\dot{x},x,t). \tag{4.26}$$

[2] For example, whnever the Hamiltonian H is quadratic in the momentum p.

Similarly, again using Hamilton's equations

$$\frac{\partial L}{\partial x} = \frac{\partial p(\dot{x}, x, t)}{\partial x} \dot{x} - \frac{\partial H}{\partial p} \frac{\partial p(\dot{x}, x, t)}{\partial x} - \frac{\partial H}{\partial x} = -\frac{\partial H}{\partial x}. \tag{4.27}$$

Thus, using Hamilton's equations yet again, we see

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}}\right] - \frac{\partial L}{\partial x} = \frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial H}{\partial x} = 0. \tag{4.28}$$

The net result is that we have now gone back in a big (internally self-consistent) circle, from:

$$\text{Lagrangian} \to \text{Hamiltonian} \to \text{Lagrangian}.$$

Chapter 5

Noether's theorem

We have already seen, in the various forms of the Beltrami identity, how a symmetry of the Lagrangian (time-independence, where by this we mean independence from any *explicit* time dependence), leads to the conservation of total energy. In 1918, **Emmy Noether** extended and generalized this result to raise it to a general principle:

- Each independent symmetry of the Lagrangian (or Hamiltonian) leads to an independent conservation law.

Specifically:

- Via Beltrami's identity we have already seen that time translation invariance leads to conservation of total energy.
- We shall soon see that space translation invariance leads to conservation of total momentum.
- Similarly, symmetry under rotations leads to conservation of angular momentum.
- More radically, symmetry under gauge invariance leads to conservation of electric charge.
- Approximate symmetry under isospin leads to approximate conservation rules in beta-decay of nuclei.
- More generally, any "internal symmetry" leads so some sort of "conserved charge".

5.1 Conservation of momentum

First, let us consider the trivial case where $L(\dot{x})$ is independent of both x and t. Then the Euler–Lagrange equations simplify quite drastically to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}}\right] = 0. \tag{5.1}$$

That is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[p(\dot{x})\right] = 0. \tag{5.2}$$

So momentum is conserved — of course this is the trivial situation where there are no forces at play — this is just Newton's first law in disguise.

A more interesting situation arises when we have a number N of particles, labelled by indices $A \in \{1, 2, \ldots, N\}$, moving in 3-dimensional space $i \in \{1, 2, 3\}$, and interacting with each other through two-body pairwise potentials that depend only on the *relative separation* of the 2 particles:

$$V_{AB}(x^i_A - x^i_B). \tag{5.3}$$

That is, consider the Lagrangian

$$L(\dot{x}, x) = \left[\frac{1}{2}\sum_{A=1}^{N}\sum_{i=1}^{3} m_A\ \left(\dot{x}^i_A\right)^2 - \frac{1}{2}\sum_{A,B,A\neq B} V_{AB}(x^i_A - x^i_B)\right]. \tag{5.4}$$

Now this Lagrangian has a symmetry under a *simultaneous* shift of *all* the positions:

$$x^i_A \to x^i_A + a^i; \qquad\qquad \dot{x}^i_A \to \dot{x}^i_A. \tag{5.5}$$

Under this transformation $\delta L = 0$, so according to Emmy Noether, *something* should be conserved. Let's first use the low-brow approach: Consider the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{x}}\right] = \frac{\partial L}{\partial x}. \tag{5.6}$$

In the current situation these become[1]

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m_A\ \dot{x}^i_A\right] = -\frac{1}{2}\sum_{B,B\neq A}\left[\frac{\partial V_{AB}}{\partial x^i}(x^i_A - x^i_B) - \frac{\partial V_{BA}}{\partial x^i}(x^i_B - x^i_A)\right]. \tag{5.7}$$

Now add up over all the particles to find the time evolution of the total momentum:

$$\sum_A\left\{\frac{\mathrm{d}}{\mathrm{d}t}\left[m_A\ \dot{x}^i_A\right]\right\} = -\frac{1}{2}\sum_A\left\{\sum_{B,B\neq A}\left[\frac{\partial V_{AB}}{\partial x^i}(x^i_A - x^i_B) - \frac{\partial V_{BA}}{\partial x^i}(x^i_B - x^i_A)\right]\right\}. \tag{5.8}$$

[1] If you have any doubt about this step, try the simplest case of just 2 particles.

That is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_A m_A\,\dot{x}^i_A\right] = -\frac{1}{2}\left\{\sum_{A,B,B\neq A}\left[\frac{\partial V_{AB}}{\partial x^i}(x^i_A - x^i_B) - \frac{\partial V_{BA}}{\partial x^i}(x^i_B - x^i_A)\right]\right\}. \tag{5.9}$$

The $\partial_i V$ terms cancel in pairs[2] so we deduce:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_A m_A\,\dot{x}^i_A\right] = 0. \tag{5.10}$$

That is, total momentum is conserved:

$$P^i = \sum_A m_A\,\dot{x}^i_A = (\text{constant}). \tag{5.11}$$

The key point in deriving this conservation law for the total momentum was that the potential only depends on the position *differences* between the individual particles, so that the potential (and also the Lagrangian) was *invariant* under an overall change in the position of the centre of mass.

This conservation law applies for instance to our solar system — the sun, planets, moons, asteroids, etc interact via individual 2-body potentials of Newtonian gravity:

$$V_{AB}(x^i_A - x^i_B) = \frac{G_{\text{Newton}}\,m_A\,m_B}{||x^i_A - x^i_B||}, \tag{5.12}$$

and the result follows.

5.2 Conservation of angular momentum

Conservation of *angular momentum* is a deeper result that requires more information about the 2-body potentials. Suppose the individual 2-body potentials are radially symmetric, so that the potential energy depends only on *distance* but not direction:

$$V_{AB}(x^i_A - x^i_B) \to V_{AB}(||x^i_A - x^i_B||) = V_{AB}(r_{AB}). \tag{5.13}$$

In this situation both the individual 2-body potentials *and* the total potential

$$V(x) = \frac{1}{2}\sum_{A,B,A\neq B} V_{AB}(||x^i_A - x^i_B||) = \frac{1}{2}\sum_{A,B,A\neq B} V_{AB}(r_{AB}) \tag{5.14}$$

are *invariant* under rotation around any arbitrarily chosen axis.

[2] If you have any doubt about this step, try the simplest case of just 2 particles.

Via Noether's theorem this will led to a conserved quantity of some sort.

Use the low-brow approach of explicit computation. Start with the Lagrangian

$$L(\dot{x}, x) = \left[\frac{1}{2} \sum_{A=1}^{N} \sum_{i=1}^{3} m_A \, \left(\dot{x}_A^i\right)^2 - \frac{1}{2} \sum_{A,B,A\neq B} V_{AB}(||x_A^i - x_B^i||) \right]. \tag{5.15}$$

Note that

$$\frac{\partial \, r_{AB}}{\partial x_B^i} = - \frac{x_A^i - x_B^i}{||x_A^i - x_B^i||}, \tag{5.16}$$

which is minus the unit vector from particle B to particle A. Now compute the Euler–Lagrange equations, noting that:

$$\frac{\partial V_{AB}}{\partial x_B^i} = - \frac{\mathrm{d}V_{AB}}{\mathrm{d}r} \, \frac{x_A^i - x_B^i}{||x_A^i - x_B^i||}. \tag{5.17}$$

Therefore[3]

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[m_A \, \dot{x}_A^i \right] = - \sum_{B,B\neq A} \left[\frac{\mathrm{d}V_{AB}}{\mathrm{d}r} \, \frac{x_A^i - x_B^i}{||x_A^i - x_B^i||} \right]. \tag{5.18}$$

Now take the cross product with x_A^i, we see

$$\vec{x}_A \times \frac{\mathrm{d}}{\mathrm{d}t} \left[m_A \, \dot{\vec{x}}_A \right] = - \vec{x}_A \times \sum_{B,B\neq A} \left[\frac{\mathrm{d}V_{AB}}{\mathrm{d}r} \, \frac{\vec{x}_A - \vec{x}_B}{||\vec{x}_A - \vec{x}_B||} \right], \tag{5.19}$$

whence, using the properties of the vector cross product

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[m_A \, \vec{x}_A \times \dot{\vec{x}}_A \right] = - \sum_{B,B\neq A} \left[\frac{\mathrm{d}V_{AB}}{\mathrm{d}r} \, \frac{\vec{x}_A \times \vec{x}_B}{r_{AB}} \right]. \tag{5.20}$$

Now add over all the particles A, we see

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_A m_A \, \vec{x}_A \times \dot{\vec{x}}_A \right] = - \sum_{A,B,B\neq A} \left[\frac{\mathrm{d}V_{AB}}{\mathrm{d}r} \, \frac{\vec{x}_A \times \vec{x}_B}{r_{AB}} \right] = 0, \tag{5.21}$$

where at the last stage we have used the fact that the sum is symmetric in the labels AB while the cross product is antisymmetric in AB. That is:

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_A m_A \, \vec{x}_A \times \dot{\vec{x}}_A \right] = 0. \tag{5.22}$$

So the total angular momentum is conserved:

$$\vec{J} = \sum_A m_A \, \vec{x}_A \times \dot{\vec{x}}_A = (\text{constant}). \tag{5.23}$$

[3] If you have any doubt about this step, try the simplest case of just 2 particles.

Note that the key step in deriving this result is that the individual 2-body potentials are rotationally invariant, depending only on the distance r_{AB} between the two particles labelled A and B. In particular, this means that the total angular momentum of the solar system is conserved, at least as long as we are happy to use Newtonian gravity as a good approximation, and as long as there are no significant external forces due to nearby marauding stars (nemesis hypothesis) or black holes passing by...

5.3 Summary

We have now see three separate conservation laws, for total energy, total momentum, and total angular momentum, and have seen how these are related to time translation invariance, space translation invariance, and rotational invariance respectively. This is part of a general pattern to do with Noether's theorem:

- Each independent symmetry of the Lagrangian (or Hamiltonian) leads to an independent conservation law.

We will not discuss this more fully in these notes, but will ask for some appropriate computations in the assignments.

Chapter 6

Coda

At this stage let us summarize what we have done:

- We have developed a concise introduction to the mathematical formalism known as the "calculus of variations".
- We have then used this "calculus of variations" to address problems of advanced classical mechanics:
  - We have seen how to describe classical mechanics in terms of a Lagrangian $L(\dot{x}, x, t)$.
  - We have seen how to describe classical mechanics in terms of a Hamiltonian $H(p, x, t)$.
  - We have seen how to convert back and forth from the Lagrangian to the Hamiltonian formalism.
  - We have seen several examples of how to relate symmetries to conservation laws.
- For a little more background, the appendix gives you a feel for some of the deeper mathematics associated with this class of problem.
- See also the list of books and relevant websites in the bibliography.

Appendix A

Hilbert's 19th, 20th, and 23rd problems

In the year 1900 Professor **David Hilbert** gave a key-note address at the International Congress of Mathematicians which was that year held in Paris. Hilbert's address set out a list of 23 problems that he thought were important — and much of 20th century mathematics was devoted to solving about half of these problems. Dr. Maby Winton Newson translated this address into English with the author's permission for the "Bulletin of the American Mathematical Society" 8 (1902), 437-479. A reprint of the address appears in "Mathematical Developments Arising from Hilbert Problems", edited by Felix Brouder, American Mathematical Society, 1976. Various versions are also available on the internet; go to Google and search on "**Hilbert problems**". Three of the 23 problems directly involve the calculus of variations, the 19th, 20th, and 23rd problems. Excerpts from the lecture are presented below. Note especially that Hilbert's 23rd problem was somewhat more open-ended than the others ...

A.1 Introduction

Mathematical Problems

Lecture delivered before the International Congress of Mathematicians Paris 1900

By Professor David Hilbert

... lacuna ...

It is difficult and often impossible to judge the value of a problem correctly in advance; for the final award depends upon the gain which science obtains from the problem. Nevertheless we can ask whether there are general criteria which mark a good mathematical problem. An old French mathematician said: "A mathematical theory is not to be considered complete until you have made it so clear that you can explain it to the first man whom you meet on the street." This clearness and ease of comprehension, here insisted on for a mathematical theory, I should still more demand for a mathematical problem if it is to be perfect; for what is clear and easily comprehended attracts, the complicated repels us.

Moreover a mathematical problem should be difficult in order to entice us, yet not completely inaccessible, lest it mock at our efforts. It should be to us a guide post on the mazy paths to hidden truths, and ultimately a reminder of our pleasure in the successful solution.

... lacuna ...

A.2 The 19th problem

19. Are the solutions of regular problems in the calculus of variations always necessarily analytic?

One of the most remarkable facts in the elements of the theory of analytic functions appears to me to be this: That there exist partial differential equations whose integrals are all of necessity analytic functions of the independent variables, that is, in short, equations susceptible of none but analytic solutions. The best known partial differential equations of this kind are the potential equation

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

and certain linear differential equations investigated by Picard;[46] also the equation

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = e^f,$$

the partial differential equation of minimal surfaces, and others. Most of these partial differential equations have the common characteristic of being the Lagrangian differential equations of certain problems of variation, viz., of such problems of variation

$$\int\int F(p, q, z; x, y)\,\mathrm{d}x\,\mathrm{d}y = \text{minimum}$$

$$\left[p = \frac{\partial z}{\partial x},\quad q = \frac{\partial z}{\partial y}\right],$$

as satisfy, for all values of the arguments which fall within the range of discussion, the inequality

$$\frac{\partial^2 F}{\partial p^2} \cdot \left(\frac{\partial^2 F}{\partial p\,\partial q}\right)^2 > 0,$$

F itself being an analytic function. We shall call this sort of problem a regular variation problem. It is chiefly the regular variation problems that play a role in geometry, in mechanics, and in mathematical physics; and the question naturally arises, whether all solutions of regular variation problems must necessarily be analytic functions. In other words, *does every Lagrangian partial differential equation of a regular variation problem have the property of admitting analytic integrals exclusively?* And is this the case even when the function is constrained to assume, as, e.g., in Dirichlet's problem on the potential function, boundary values which are continuous, but not analytic?

I may add that there exist surfaces of constant negative gaussian curvature which are representable by functions that are continuous and possess indeed all the derivatives, and yet are not analytic; while on the other hand it is probable that every surface whose gaussian curvature is constant and positive is necessarily an analytic surface. And we know that the surfaces of positive constant curvature are most closely related to this regular variation problem: To pass through a closed curve in space a surface of minimal area which shall enclose, in connection with a fixed surface through the same closed curve, a volume of given magnitude.

A.3 The 20th problem

20. The general problem of boundary values

An important problem closely connected with the foregoing is the question concerning the existence of solutions of partial differential equations when the values on the boundary of the region are prescribed. This problem is solved in the main by the keen methods of H. A. Schwarz, C. Neumann, and Poincare for the differential equation of the potential. These methods, however, seem to be generally not capable of direct extension to the case where along the boundary there are prescribed either the differential coefficients or any relations between these and the values of the function. Nor can they be extended immediately to the case where the inquiry is not for potential surfaces but, say, for surfaces of least area, or surfaces of constant positive gaussian curvature, which are to pass through a prescribed twisted curve or to stretch over a given ring surface. It is my conviction that it will be possible to prove these existence theorems by means of a general principle whose nature is indicated by Dirichlet's principle. This general principle will then perhaps enable us to approach the question: *Has not every regular variation problem a solution, provided certain assumptions regarding the given boundary conditions are satisfied* (say that the functions concerned in these boundary conditions are continuous and have in sections one or more derivatives)*, and provided also if need be that the notion of a solution shall be suitably extended?*[47]

... lacuna ...

A.4 The 23rd problem

23. Further development of the methods of the calculus of variations

So far, I have generally mentioned problems as definite and special as possible, in the opinion that it is just such definite and special problems that attract us the most and from which the most lasting influence is often exerted upon science. Nevertheless, I should like to close with a general problem, namely with the indication of a branch of mathematics repeatedly mentioned in this lecture—which, in spite of the considerable advancement lately given it by Weierstrass, does not receive the general appreciation which, in my opinion, is its due—I mean the calculus of variations.[50]

The lack of interest in this is perhaps due in part to the need of reliable modern text books. So much the more praiseworthy is it that A. Kneser in a very recently published work has treated the calculus of variations from the modern points of view and with regard to the modern demand for rigor.[51]

The calculus of variations is, in the widest sense, the theory of the variation of functions, and as such appears as a necessary extension of the differential and integral calculus. In this sense, Poincare's investigations on the problem of three bodies, for example, form a chapter in the calculus of variations, in so far as Poincare derives from known orbits by the principle of variation new orbits of similar character.

I add here a short justification of the general remarks upon the calculus of variations made at the beginning of my lecture.

The simplest problem in the calculus of variations proper is known to consist in finding a function y of a variable x such that the definite integral

$$J = \int_a^b F(y_x, y; x)\mathrm{d}x, \qquad \left[y_x = \frac{\mathrm{d}y}{\mathrm{d}x}\right]$$

assumes a minimum value as compared with the values it takes when y is replaced by other functions of x with the same initial and final values.

The vanishing of the first variation in the usual sense

$$\delta J = 0$$

gives for the desired function y the well-known differential equation

$$\frac{\mathrm{d}F_{y_x}}{\mathrm{d}x} - F_y = 0, \qquad \left[F_{y_x} = \frac{\partial F}{\partial y_x}, \qquad F_y = \frac{\partial F}{\partial y}\right] \tag{1}$$

In order to investigate more closely the necessary and sufficient criteria for the occurrence of the required minimum, we consider the integral

$$J^* = \int_a^b \{F(y_x, y; x) + (y_x - p)F_p\}\, \mathrm{d}x,$$

$$\left[F = F(p, y; x), \qquad F_p = \frac{\partial F(p, y; x)}{\partial p}\right].$$

Now we inquire how p is to be chosen as function of x, y in order that the value of this integral J^* shall be independent of the path of integration, i.e., of the choice of the function y of the variable x. The integral J^* has the form

$$J^* = \int_a^b \{A\, y_x - B\}\, \mathrm{d}x,$$

where A and B do not contain y_x, and the vanishing of the first variation

$$\delta J* = 0$$

in the sense which the new question requires gives the equation

$$\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} = 0,$$

i.e., we obtain for the function p of the two variables x, y the partial differential equation of the first order

$$\frac{\partial F_p}{\partial x} + \frac{\partial(pF_p - F)}{\partial y} = 0. \qquad (1*)$$

The ordinary differential equation of the second order (1) and the partial differential equation (1*) stand in the closest relation to each other. This relation becomes immediately clear to us by the following simple transformation

$$\begin{aligned}
\delta J^* &= \int_a^b \{F_y\, \delta y + F_p\, \delta p + (\delta y_x - \delta p)F_p + (y_x - p)\, \delta F_p\}\, \mathrm{d}x \\
&= \int_a^b \{F_y\, \delta y + \delta y_x\, F_p + (y_x - p)\, \delta F_p\}\, \mathrm{d}x \\
&= \delta J + \int_a^b (y_x - p)\, \delta F_p\, \mathrm{d}x
\end{aligned}$$

We derive from this, namely, the following facts: If we construct any simple family of integral curves of the ordinary differential equation (l) of the second order and then form an ordinary differential equation of the first order

$$y_x = p(x, y) \qquad (2)$$

which also admits these integral curves as solutions, then the function $p(x, y)$ is always an integral of the partial differential equation (1*) of the first order; and conversely, if $p(x, y)$ denotes any solution of the partial differential equation (1*) of the first order, all the non-singular integrals of the ordinary differential equation (2) of the first order are at the same time integrals of the differential equation (l) of the second order, or in short if $y_x = p(x, y)$ is an integral equation of the first order of the differential equation (l) of the second order, $p(x, y)$ represents an integral of the partial differential equation (1*) and conversely; the integral curves of the ordinary differential equation of the second order are therefore, at the same time, the characteristics of the partial differential equation (1*) of the first order.

In the present case we may find the same result by means of a simple calculation; for this gives us the differential equations (1) and (1*) in question in the form

$$y_{xx}F_{y_xy_x} + y_xF_{y_xy} + F_{y_xx} - F_y = 0, \qquad (1)$$

$$(p_x + pp_x)F_{pp} + pF_{py} + F_{px} - F_y = 0, \qquad (1*)$$

where the lower indices indicate the partial derivatives with respect to x, y, p, y_x. The correctness of the affirmed relation is clear from this.

The close relation derived before and just proved between the ordinary differential equation (1) of the second order and the partial differential equation (1*) of the first order, is, as it seems to me, of fundamental significance for the calculus of variations. For, from the fact that the integral $J^*$is independent of the path of integration it follows that

$$\int_a^b \{F(p) + (y_x - p)F_p(p)\}\,\mathrm{d}x = \int_a^b F(\bar{y}_x)\,\mathrm{d}x, \qquad (3)$$

if we think of the left hand integral as taken along any path y and the right hand integral along an integral curve of the differential equation

$$\bar{y}_x = p(x, \bar{y}).$$

With the help of equation (3) we arrive at Weierstrass's formula

$$\int_a^b F(y_x)\,\mathrm{d}x - \int_a^b F(\bar{y}_x)\,\mathrm{d}x = \int_a^b E(y_x, p)\,\mathrm{d}x, \qquad (4)$$

where E designates Weierstrass's expression, depending upon y_x, p, y, x,

$$E(y_x, p) = F(y_x) - F(p) - (y_x - p)F_p(p),$$

Since, therefore, the solution depends only on finding an integral $p(x, y)$ which is single valued and continuous in a certain neighborhood of the integral curve , which we are considering, the developments just indicated lead immediately—without the introduction of the second variation, but only by the application of the polar process to the differential equation (1)—to the expression of Jacobi's condition and to the answer to the question: How far this condition of Jacobi's in conjunction with Weierstrass's condition $E > 0$ is necessary and sufficient for the occurrence of a minimum.

The developments indicated may be transferred without necessitating further calculation to the case of two or more required functions, and also to the case of a double or a multiple integral. So, for example, in the case of a double integral

$$J = \int F(z_x, z_y, z; x, y)\mathrm{d}\omega, \qquad \left[z_x = \frac{\partial z}{\partial x}, \;\; z_y = \frac{\partial z}{\partial y}\right]$$

to be extended over a given region ω, the vanishing of the first variation (to be understood in the usual sense)

$$\delta J = 0$$

gives the well-known differential equation of the second order

$$\frac{\mathrm{d}F_{z_x}}{\mathrm{d}x} + \frac{\mathrm{d}F_{z_y}}{\mathrm{d}y} - F_z = 0, \qquad \left[F_{z_x} = \frac{\partial F}{\partial z_x}, \;\; F_{z_y} = \frac{\partial F}{\partial z_y}, \;\; F_z = \frac{\partial F}{\partial z}\right] \qquad (I)$$

for the required function z of x and y.

On the other hand we consider the integral

$$J^* = \int \{F + (z_x - p)F_p + (z_y - q)F_q\}\,\mathrm{d}\omega$$

$$\left[F = F(p, q, z; x, y), \;\; F_p = \frac{\partial F}{\partial p}, \;\; F_q = \frac{\partial F}{\partial q}\right]$$

and inquire, how p and q are to be taken as functions of x, y and z in order that the value of this integral may be independent of the choice of the surface passing through the given closed twisted curve, i.e., of the choice of the function z of the variables x and y.

The integral J* has the form

$$J^* = \int \{Az_x + Bz_y - C\}\,\mathrm{d}\omega$$

and the vanishing of the first variation

$$\delta J* = 0$$

in the sense which the new formulation of the question demands, gives the equation

$$\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z} = 0,$$

i.e., we find for the functions p and q of the three variables x, y and z the differential equation of the first order

$$\frac{\partial F_p}{\partial x} + \frac{\partial F_q}{\partial y} + \frac{\partial (pF_p + qF_q - F)}{\partial z} = 0. \qquad (I)$$

If we add to this differential equation the partial differential equation

$$p_y + qp_x = q_x + pq_z \qquad (I*)$$

resulting from the equations

$$z_x = p(x, y, z),$$
$$z_y = q(x, y, z)$$

the partial differential equation (I) for the function z of the two variables x and y and the simultaneous system of the two partial differential equations of the first order (I*) for the two functions p and q of the three variables x, y, and z stand toward one another in a relation exactly analogous to that in which the differential equations (1) and (1*) stood in the case of the simple integral.

It follows from the fact that the integral J^* is independent of the choice of the surface of integration z that

$$\int \{F(p,q) + (z_x - p)F_p(p,q) + (z_y - q)F_q(p,q)\}\,\mathrm{d}\omega = \int F(\bar{z}_z, \bar{z}_y)\,\mathrm{d}\omega, \qquad (III)$$

if we think of the right hand integral as taken over $\bar{z}$ an integral surface of the partial differential equations

$$\bar{z}_x = p(x, y, \bar{z}), \qquad \bar{z}_y = q(x, y, \bar{z});$$

and with the help of this formula we arrive at once at the formula

$$\int F(z_x, z_y)\mathrm{d}\omega - \int F(\bar{z}_x, \bar{z}_y)\mathrm{d}\omega = \int E(z_x, z_y, p, q)\mathrm{d}\omega,$$

$$[E(z_x, z_y, p, q) = F(z_x, z_y) - F(p,q) - (z_x - p)F_p(p,q) - (z_y - q)F_q(p,q)]\,, \qquad (IV)$$

which plays the same role for the variation of double integrals as the previously given formula (4) for simple integrals. With the help of this formula we can now answer the question how far Jacobi's condition in conjunction with Weierstrass's condition $E > 0$ is necessary and sufficient for the occurrence of a minimum.

Connected with these developments is the modified form in which A. Kneser,[52] beginning from other points of view, has presented Weierstrass's theory. While Weierstrass employed integral curves of equation (1) which pass through a fixed point in order to derive sufficient conditions for the extreme values, Kneser on the other hand makes use of any simple family of such curves and constructs for every such family a solution, characteristic for that family, of that partial differential equation which is to be considered as a generalization of the Jacobi-Hamilton equation.

— # # # —

The problems mentioned are merely samples of problems, yet they will suffice to show how rich, how manifold and how extensive the mathematical science of today is, and the question is urged upon us whether mathematics is doomed to the fate of those other sciences that have split up into separate branches, whose representatives scarcely understand one another and whose connection becomes ever more loose. I do not believe this nor wish it. Mathematical science is in my opinion an indivisible whole, an organism whose vitality is conditioned upon the connection of its parts. For with all the variety of mathematical knowledge, we are still clearly conscious of the similarity of the logical devices, the relationship of the ideas in mathematics as a whole and the numerous analogies in its different departments. We also notice that, the farther a mathematical theory is developed, the more harmoniously and uniformly does its construction proceed, and unsuspected relations are disclosed between hitherto separate branches of the science. So it happens that, with the extension of mathematics, its organic character is not lost but only manifests itself the more clearly.

But, we ask, with the extension of mathematical knowledge will it not finally become impossible for the single investigator to embrace all departments of this knowledge? In answer let me point out how thoroughly it is ingrained in mathematical science that every real advance goes hand in hand with the invention of sharper tools and simpler methods which at the same time assist in understanding earlier theories and cast aside older more complicated developments. It is therefore possible for the individual investigator, when he makes these sharper tools and simpler methods his own, to find his way more easily in the various branches of mathematics than is possible in any other science.

The organic unity of mathematics is inherent in the nature of this science, for mathematics is the foundation of all exact knowledge of natural phenomena. That it may completely fulfill this high mission, may the new century bring it gifted masters and many zealous and enthusiastic disciples!

— # # # —

A.5 Original references

46 — Picard: Jour. de l'Ecole Polytech., 1890.

47 — Cf. D. Hilbert: "Uber das Dirichlet'sche Princip," Jahresber. d. Deutschen Math.-Vereinigung, 8 (1900), 184-188.

50 — Text-books:
Moigno and Lindelof, Lecons du calcul des variations, Mallet-Bachelier, Paris, 1861, and A. Kneser, Lehrbuch der Variations-rechnung, Vieweg, Braunschweig, 1900.

51 — As an indication of the contents of this work, it may here be noted that for the simplest problems Kneser derives sufficient conditions of the extreme even for the case that one limit of integration is variable, and employs the envelope of a family of curves satisfying the differential equations of the problem to prove the necessity of Jacobi's conditions of the extreme. Moreover, it should be noticed that Kneser applies Weierstrass's theory also to the inquiry for the extreme of such quantities as are defined by differential equations.

52 — Cf. Kneser's above-mentioned textbook, §§14, 16, 19 and 20.

A.6 Additional references

1996 — S. Chern: Remarks on Hilbert's 23rd problem. Math. Intelligencer 18, no. 4, 7-8.

1976 — G. Stampacchia: Hilbert's twenty-third problem. Extensions of the calculus of variations. Proceedings of the Symposium in Pure Mathematics of the American Mathematical Society, Held at Northern Illinois University 1974, (edited by F.E.Browder), 611-628.

1900 — D. Hilbert: Weiterfuhrung der Methoden der Variationsrechnung. Akad. Wiss. Gottingen 1900, 291-296.

A.7 Notes

Note: Problems 19 and 20 are now commonly agreed to be "solved". ♢

Note: Problem 23 is sufficiently open-ended that there is no general agreement as to whether or it can ever be "solved". ♢

Exercise: Do a literature survey to judge the extent to which these problems have actually been "solved". If you find something new and interesting, publish. ♢

Appendix B

Bibliography

B.1 Some books on Lagrangian and Hamiltonian mechanics:

[Often with a lot of useful background information.]

- Mechanics
  LD Landau and EM Lifshitz
  [Course in theoretical physics, volume 1, 3rd edition]
  (Butterworth–Heinenann, Oxford, 2000)

- Classical Mechanics
  H Goldstein
  (Addison Wesley, 1959)

- Classical Mechanics (3rd edition)
  H Goldstein, CP Poole, and JL Safko
  (Addison Wesley, 2001)

- Classical Mechanics
  JR Taylor
  (University Science Books, 2005)

- Classical mechanics
  RD Gregory
  (Cambridge University Press, 2006)

- A History of the Calculus of Variations from the 17th through the 19th Century.
  Goldstine, H. H.
  (Springer-Verlag, New York, 1980)

- The variational principles of mechanics
  C Lanczos
  (U Toronto Press, 1970)

B.2 Some websites on Lagrangian and Hamiltonian mechanics:

[Often with a lot of useful background information.]

[On this subject Wikipedia is reasonably reliable.]

- http://en.wikipedia.org/wiki/Lagrange
- http://en.wikipedia.org/wiki/William_Rowan_Hamilton
- http://en.wikipedia.org/wiki/Pierre_de_Fermat
- http://en.wikipedia.org/wiki/Pierre_Louis_Maupertuis
- http://en.wikipedia.org/wiki/Lagrangian_mechanics
- http://en.wikipedia.org/wiki/Hamiltonian_mechanics
- http://en.wikipedia.org/wiki/Lagrangian
- http://en.wikipedia.org/wiki/Fermat's_principle
- http://en.wikipedia.org/wiki/Vis_viva
- http://en.wikipedia.org/wiki/Euler–Lagrange_equation
- http://en.wikipedia.org/wiki/Fundamental_lemma_of_calculus_of_variations
- http://en.wikipedia.org/wiki/Euler
- http://en.wikipedia.org/wiki/Noether's_theorem
- http://arxiv.org/pdf/physics/0503066
  (English translation of Emmy Noether's original article.)
- http://en.wikipedia.org/wiki/Hilbert's_problems

— # # # —