Tecniche variazionali

Alberto Tibaldi

December 15, 2012

Contents

Chapter 1

Introduction to perturbative methods

In this chapter the idea of perturbative method will be introduced. Let's suppose that we need to solve a problem which is in some way similar to an easier problem; we can "perturb" the solution of the easy problem, in order to find the solution of the harder one.

This idea, without examples, may be complicated to understand; therefore, let's consider a very easy mathematical problem: an algebraic equation of third degree:

$$x^3 + 10^{-4} x = 1$$

since 10^{-4} is a very small number, we can say that this problem is similar to:

$$x^3 = 1$$

which can be solved analytically in an easy way: $x = 1$.

1.1 Analytical solution of the problem

Let's consider the more generic problem:

$$x^3 + px = q$$

this has an analytical solution, given by Tartaglia / Scipione del ferro (italian mathematicians). The starting point is to write x as:

$$x = u + v$$

by substituting, we can obtain:

$$u^3 + v^3 + (u + v)(p + 3uv) = q$$

we transformed the problem in one variable to a problem in two variables; therefore, we can assume that u and v follow this constraint:

$$p + 3uv = 0$$

by this way we have two variables and two equations:

$$\begin{cases} u^3 + v^3 = q \\ uv = -\dfrac{p}{3} \end{cases} \tag{1.1}$$

the second equation can be elevated to the third power, obtaining:

$$\begin{cases} u^3 + v^3 = q \\ u^3 v^3 = -\dfrac{p^3}{27} \end{cases} \tag{1.2}$$

now, if we call $u^3 = a$, $v^3 = b$, we have to solve:

$$\begin{cases} a + b = q \\ ab = -\dfrac{p^3}{27} \end{cases} \tag{1.3}$$

the problem now is: which are the solutions of the system? In other words: which are the numbers a and b such that their sum equals q, and their product equals $-\frac{p^3}{27}$? The answer is: the solutions of the second order degree equation in a, b! Let's say that:

$$a = \frac{q + \sqrt{q^2 + \frac{4p^3}{27}}}{2} \qquad b = \frac{q - \sqrt{q^2 + \frac{4p^3}{27}}}{2}$$

So:

$$x = \sqrt[3]{\frac{q + \sqrt{q^2 + 4\frac{p^3}{27}}}{2}} + \sqrt[3]{\frac{q - \sqrt{q^2 + 4\frac{p^3}{27}}}{2}}$$

Under the analytical point of view this solution may be interesting, but the reason why it was very important must be found in the history of mathematics: it was found before the formalization of the idea of complex number; since it is geometrically evident that a cubic parabola has always at least one interception with a straight line, but this formula may not produce any solution (if we have a negative number inside the square root), mathematicians started thinking about the definition of a larger set of number: **complex numbers**.

1.2 Application of a perturbative method to the problem

Let's go back to the former problem, and let's see how we can solve it knowing that the solution is close to 1; the starting equation was:

$$x^3 + 10^{-4}x = 1$$

we know that:

$$x \simeq 1 + \delta, \quad |\delta| \ll 1$$

by substituting this into the equation, it is possible to obtain:

$$(1 + \delta)^3 + 10^{-4}(1 + \delta) = 1$$

it is possible to eliminate small terms: these terms are the ones where δ has a power greater than 1, and when we have 10^{-4} which multiplies δ; so:

$$1 + 3\delta + 3\delta^2 + \delta^3 + 10^{-4} + 10^{-4}\delta = 1 \simeq 3\delta + 10^{-4} = 0$$

so:

$$\delta = -\frac{1}{3}10^{-4}$$

this solution is pretty accurate, and it is very easy to find.

Now, let's consider the original problem in order to solve it in a nicer fashion. If we have something like:

$$x^3 + \varepsilon x = 1, \quad |\varepsilon| \ll 1$$

there are several kinds of approximations; the zeroth-order approximation is to consider $\varepsilon \to 0$, so:

$$x^3 = 1$$

which has trivial solutions.

It is possible to improve the order of the approximation, finding a Taylor expansion in the ε variable. Considering only the first terms, we have:

$$x = 1 + a_1\varepsilon + a_2\varepsilon^2 + a_3\varepsilon^3 + ...$$

then, if we substitute this expression into our equation, we obtain:

$$\left(1 + a_1\varepsilon + a_2\varepsilon^2 + ...\right)^3 + \varepsilon(1 + a_1\varepsilon + a_2\varepsilon^2 + ...) - 1 = 0$$

by doing all calculations it is possible to write that all first and second order terms are:

$$1 + 3a_1\varepsilon + 3a_1^2\varepsilon + 3a_2\varepsilon^2 + ... + \varepsilon + a_1\varepsilon^2 + a_2\varepsilon^3 + ... - 1 = 0$$

we have an equation which equals zero; it is known that it is satisfied if and only if all the coefficients which multiply the main variable (which is ε) are zero; this means that we can group them, truncating to the second power:

$$(3a_1 + 1)\varepsilon + (3a_1^2 + 3a_2 + a_1)\varepsilon^2 = 0$$

so, in order to satisfy the equation, it is necessary to solve the following system:

$$\begin{cases} 3a_1 + 1 = 0 \\ 3a_1^2 + 3a_2 + a_1 = 0 \end{cases}$$

by this way it is possible to find the second order approximation of the solution; increasing the order, it is possible to find even more approximate solutions.

1.3 Classification of perturbation problems - Differential equation example

There is a classification for perturbation problems:

- regular problems: these are problems like

$$x^2 + \varepsilon x = 1$$

- singular problems: these are problems like

$$\varepsilon x^2 + x = 1$$

these, are situations where the case $\varepsilon = 0$ changes the nature of the problem.

In fact, in the first case, the problem remains a second-order problem, while in the second case the nature changes completely: from having two solutions, we have just one.

1.3.1 Example of differential problem: D'Alembert equation

An example of *serious* problem which can be solved is the D'Alembert equation:

$$\nabla^2 \phi - \frac{1}{\text{c}^2} \frac{\partial^2 \phi}{\partial t^2} = 0$$

The first hypothesis which can be introduced is the **monochromatic waves** one: the dependence on time of the solution of the differential equation is:

$$\phi \sim \text{e}^{-\text{j}\omega t}, \quad \omega = 2\pi\nu$$

where ν is the frequency of the oscillation of the wave in time. If we substitute this, we obtain:

$$\nabla^2 \phi + \frac{\omega^2 n^2}{\text{c}^2} \phi = 0$$

but:

$$\frac{\omega}{\text{c}} = \frac{2\pi\nu}{\text{c}} = \frac{2\pi}{\lambda_0} = \frac{1}{k}$$

so:

$$\nabla^2 \phi + k^2 n^2 \phi = 0$$

In a real problem, n is not constant in space; what we can do is to consider it approximately constant, obtaining as a solution a **plane wave**:

$$\phi \propto e^{j\underline{k}\cdot\underline{r}}$$

where $\underline{k}$ is the wavenumber (in three dimensions), while $\underline{r}$ is the vector $\underline{r} = (x, y, z)$. So:

$$\underline{k} = k_0 n \hat{\Omega}$$

where $\hat{\Omega}$ is the unit vector of the propagation direction. In other words, the solution of the differential equation is:

$$\phi = e^{jnk_0\hat{\Omega}\cdot\underline{r}}$$

The presence of n is a limitation to the average speed of light: it is proved that photons can only travel at a velocity equal to c, but since when they meet an atom they are absorbed and re-emitted, there is a *stop* which reduces the average velocity.

How can we solve this equation when λ is very small with respect to the region where it is possible to sense a variation of the refractive coefficient n ?

As first step, it is convenient to re-write this solution, in a dimensionless format: we can consider a change of variables like:

$$\left(\frac{x}{L}, \frac{y}{L}, \frac{z}{L}\right) = (\tilde{x}, \tilde{y}, \tilde{z})$$

where L is the length which should be crossed in order to sense a variation of n. Now, let's define $\tilde{\nabla}^2$ as the spatial differential equation in the new reference system as:

$$\tilde{\nabla}^2 = \sum_i \frac{\partial^2}{\partial \tilde{x}_i^2} \implies \frac{1}{L^2}\tilde{\nabla}^2\phi + k^2 n^2 \phi = 0$$

now, if we divide the equation for k^2, we obtain:

$$\frac{1}{k^2 L^2}\tilde{\nabla}^2\phi + n^2\phi = 0$$

if we define

$$\varepsilon = \frac{1}{kL}$$

we obtain the problem:

$$\varepsilon^2 \tilde{\nabla} \phi + n^2 \phi = 0$$

which is the **perturbative problem** from which we start; in this case, we have a **singular perturbation**, since if $\varepsilon = 0$ the nature and number of the solutions changes.

From now on, in order to obtain a lighter notation, we will consider the following notation for the problem (eliminating all *tilde*):

$$\varepsilon^2 \nabla^2 \phi + n^2 \phi = 0, \quad |\varepsilon| \ll 1$$

ε is the perturbative parameter of the method: it is very small, so it is the one which should be use to formulate the method. In order to solve this method there is a mathematical technique, called **WKB method**. This technique takes its name from the three researchers who formulated it, in the quantum mechanics theory: Wentzel, Kramer, Brillouin.

The idea is: if we have a constant refraction coefficient n, the solution of the PDE is a plane wave, so $e^{jn\hat{\Omega}\cdot\underline{r}}$. The solution of the perturbated problem, where the perturbation is related to the fact that n is variable in space, will be in some way similar, like:

$$\phi(\underline{r}) = e^{j\frac{S(\underline{r})}{\varepsilon}}$$

where $S(\underline{r})$ will be something different from $\underline{k}\cdot\underline{r}$, since we will consider a non-constant n in space. It is important to remark that ε is at the denominator of the exponential; otherwise the method could not be working.

Now, considering the expression of the solution:

$$\phi(\underline{r}) = e^{j\frac{S(\underline{r})}{\varepsilon}}$$

we substitute it in the equation

$$\varepsilon^2 \nabla \phi + n^2 \phi = 0$$

so, let's evaluate every term:

$$\nabla^2 \phi = \nabla \cdot (\nabla \phi)$$

where

$$\nabla \phi = \frac{\mathrm{j}}{\varepsilon} \phi \nabla S$$

so:

$$\nabla \cdot (\nabla \phi) = \nabla(\phi \frac{\mathrm{j}}{\varepsilon}) \cdot \nabla S + \phi \frac{\mathrm{j}}{\varepsilon} \nabla^2 S = \frac{\mathrm{j}}{\varepsilon} \phi \nabla S \cdot \nabla S + \phi \frac{\mathrm{j}}{\varepsilon} \nabla^2 S =$$
$$= \frac{\mathrm{j}}{\varepsilon} \phi \nabla^2 S - \frac{1}{\varepsilon} \phi (\nabla S)^2$$

Now, by substituting in the differential equation, doing some algebra, we can find:

$$\mathrm{j}\varepsilon \nabla^2 S - (\nabla S)^2 + n = 0$$

We started from a linear equation, and now we got a nonlinear equation in S: it seems that, using this method, we obtained a more complicated problem. Anyway, we are going to discover that these ideas lead also to some advantages.

Up to this point, we introduced only a different formulation of the same equation; we are going to approximate $S(\underline{r})$ as follows:

$$S(\underline{r}) = S_0(\underline{r}) + \varepsilon S_1(\underline{r}) + \varepsilon^2 S_2(\underline{r}) + ...$$

We want a first-order approximation; therefore, it is possible to truncate this expression up to the second term: $S_1(\underline{r})$. Let's remark that a good perturbative method should work with **few** terms. Now, we are going to substitute all these expressions in the last differential equation, obtaining:

$$\mathrm{j}\varepsilon \left[\nabla^2 S_0 + \varepsilon \nabla^2 S_1 + ... \right] - \left[\nabla S_0 + \varepsilon \nabla S_1 + ... \right]^2 + n = 0$$

so, let's expand what is inside the second group of parentheses, truncating at first order in ε:

$$\longrightarrow (\nabla S_0)^2 + 2\nabla S_0 \varepsilon \nabla S_1 + ...$$

so, if we collect from both terms constant in ε neglecting all the terms with a power of ε higher than the first one, just like we did for the algebraic

equation, we obtain the following system of equations (by equating all the coefficients of the equation in ε to zero):

$$\begin{cases} (\nabla S_0)^2 - n^2 = 0 \\ \mathrm{j}\nabla^2 S_0 - 2\nabla S_0 \nabla S_1 = 0 \end{cases}$$

Now, we have two equations: the first one is only in $S_0(\underline{r})$, the second one in both $S_0(\underline{r})$ and $S_1(\underline{r})$. So:

$$S(\underline{r}) \simeq S_0(\underline{r}) + \varepsilon S_1(\underline{r})$$

and so:

$$\phi(\underline{r}) = \mathrm{e}^{\mathrm{j}\frac{1}{\varepsilon}[S_0(\underline{r}) + \varepsilon S_1(\underline{r})]} = \mathrm{e}^{\mathrm{j}\frac{1}{\varepsilon}S_0(\underline{r})}\, \mathrm{e}^{\mathrm{j}S_1(\underline{r})}$$

now, some observations: in the first equation, everything will be real, so $S_0(\underline{r})$ will be a real function of $\underline{r}$; in the second equation, there is $S_0(\underline{r})$ which is real (from the first one), and the other term; this means that if ∇S_0 is real, since we have something equal to an imaginary term, S_1 will be imaginary. This means that the first exponential will be complex, while the second one will be real. Moreover, since the first term is divided by ε, the argument of the exponential will be very big; being it complex, this means that the first exponential term is associated to **quick oscillations**. Since the second exponential term has an imaginary $S_1(\underline{r})$, it will be an **amplitude variation term** (small).

This is correct, if we think that this is a perturbated problem: in the non-perturbated problem, which is the study of a homogeneous medium (no variation of n), there is only an oscillation term, with constant amplitude; in this case there is a variation of the amplitude, but it is very small[1]

Since the second exponential is an amplitude term, we define the following term:

$$A = \mathrm{e}^{\mathrm{j}S_1(\underline{r})}$$

so:

---

[1]let's remark that this situation can model a continuous variation of n, but not a steep variation, like what happens when there are two hemi-spaces of different dielectric! In this case there may be also a total reflection phenomenon or something else, which can not be analyzed by means of this kind of approximation.

$$\nabla A = \mathrm{j} A \nabla A$$

so:

$$\nabla S_1 = -\mathrm{j}\frac{\nabla A}{A}$$

This can be substituted in the second equation of the system, obtaining:

$$\mathrm{j}\nabla^2 S_0 + \frac{2\mathrm{j}}{A}\nabla A \cdot \nabla S_0 = 0$$

now, sorting and eliminating common j, we obtain:

$$\nabla^2 S_0 + \frac{2}{A}\nabla S_0 \cdot \nabla A = 0$$

This allows us to re-write the two equations as:

$$\begin{cases} (\nabla S_0)^2 - n^2 = 0 \\ A\nabla^2 S_0 + 2\nabla S_0 \cdot \nabla A = 0 \end{cases}$$

where the first equation is related to **phase** of the wave, while the second equation is related to its **amplitude**.

Now, let's consider the phase equation:

$$(\nabla S_0)^2 - n^2 = 0$$

this equation means that the **norm** of the gradient of S_0 equals n:

$$|\nabla S_0(\underline{r})| = n$$

this is the **Eikonal equation**. Since ∇S_0 is a vector, we know that:

$$\nabla S_0 = n\hat{\Omega}(\underline{r})$$

so, it has a modulus (which is n), and a certain direction $\hat{\Omega}$, which depends on the point of the space, $\underline{r}$, that we are considering. In other words, for each point of the space there is a certain $\hat{\Omega}(\underline{r})$.

Starting from a point of the space it is possible to draw a curved line which is tangent in each point to $\hat{\Omega}(\underline{r})$; this line is called **light ray**. This curve can be parametrized with a s parameter: the curvilinear coordinate, which indicates the length of part of the curve.

We can imagine to follow the light ray: in each point there will be a different amplitude A, governed by the second differential equation:

$$A\nabla^2 S_0 + 2\nabla S_0 \cdot \nabla A = 0$$

So, recalling that $\nabla S_0 = n\Omega(\underline{r})$, we have:

$$\nabla S_0 \cdot \nabla A = n\hat{\Omega} \cdot \nabla A$$

it is known that the scalar product between a gradient and a unit vector is the **directional derivative** along the direction pointed by the unit vector:

$$n\hat{\Omega} \cdot \nabla A = \frac{\partial A}{\partial s}$$

so, it is possible to write, considering the second equation of the system, which was $A\nabla^2 S_0 + 2n\frac{\partial A}{\partial s} = 0$ (substituting directly the term $\nabla A \cdot \nabla S_0$):

$$\frac{\partial A}{\partial s} = -\frac{A}{2n}\nabla^2 S_0$$

this is the law of variation of the amplitude along the path s of the ray of light.

Now, let's assume that there is a plane where we can measure the amplitude of the electromagnetic radiation, and that's say that within a circle there is radiation ($A \neq 0$), while outside of it there is no radiation. If we try drawing a light ray starting from a point which do not belong to the circle, it will carry no radiation power, because outside of the circle radiation equals zero, and the derivative is zero. In other hand, this is a **dark region**.

On the other hand, if we take all the possible rays starting from a point belonging to the circle, and we propagate them up to another plane, we can see that on this plane rays will still begin to some surface (maybe different from the starting circle): radiation is confined **within the tube**.

Now, if the size of the circle is very small, we can approximate everything with a **single line**: this is **geometric optics**, which means studying the propagation of light simply by studying a curved trajectory. By describing a classical electromagnetic equation by means of a perturbative expansion, it is possible to obtain geometric optics!

We stated that:

$$\nabla S_0 = n\hat{\Omega}$$

We can deduce a simple equation from a single light ray, simply by considering the derivative of S_0 with respect to its i-th component; this means simply to write the previous equation, which is in a compact, vector form, by **components**: ∇S_0 is a vector, where each i-th component is $\frac{\partial S_0}{\partial x_i}$, where x_i is the i-th component of the reference system; therefore, we have:

$$n\Omega_i = \frac{\partial S_0}{\partial x_i}$$

(in fact, $\hat{\Omega} = (\Omega_1, \Omega_2, ...)$, so the i-th component is Ω_i, just like the derivatives of S_0. The derivative of this quantity along a ray of light is:

$$\frac{\partial}{\partial s}\left[n\Omega_i\right] = \hat{\Omega} \cdot \nabla\left[n\Omega_i\right] = \hat{\Omega} \cdot \nabla\left[\frac{\partial S_0}{\partial x_i}\right]$$

at this point, it is possible to exchange the order of differentiation, obtaining:

$$= \hat{\Omega} \cdot \frac{\partial}{\partial x_i}\nabla S_0$$

but, from the phase equation, we have:

$$\hat{\Omega} = \frac{1}{n}\nabla S_0$$

so:

$$\frac{\partial}{\partial s}\left[n\Omega_i\right] = \left(\frac{1}{n}\nabla S_0\right) \cdot \left(\frac{\partial}{\partial x_i}\nabla S_0\right)$$

by multiplying and dividing by 2 it is possible to obtain:

$$= \frac{1}{2n}2\nabla S_0 \cdot \frac{\partial}{\partial x_i}\nabla S_0$$

now, we can observe that:

$$\frac{\partial}{\partial x_i}\left(\nabla S_0\right)^2 = 2\nabla S_0\frac{\partial}{\partial x_i}\nabla S_0$$

so, this expression will equal:

$$= \frac{1}{2n}\frac{\partial}{\partial x_i}\left(\nabla S_0\right)^2$$

but, from the phase equation:

$$(\nabla S_0)^2 = n^2$$

so:

$$\frac{1}{2n}\frac{\partial}{\partial x_i}\left(n\right)^2 = \frac{1}{2n}2n\frac{\partial n}{\partial x_i} = \frac{\partial n}{\partial x_i}$$

this is a set of equations, one for each component; by collecting all these components in vector form we obtain, finally:

$$\frac{\partial}{\partial s}\left(n\hat{\Omega}\right) = \nabla n$$

This, is the fundamental equation of geometric optics; this means that this is the formula which allows us to deduce the trajectory of a ray of light.

In order to understand the power of this equations, let us consider a simple example where we apply it: given a medium where n is function only of the x variable (coordinate), considering a planar domain (identified by a couple of variables, (x, y), we have something like:

where the differential equation is:

$$\frac{\mathrm{d}}{\mathrm{d}s}(n\Omega_x) = \frac{\partial n}{\partial x}$$

this is just one differential equation: one component of the equation that we have previously written; in this case, $\frac{\partial n}{\partial x}$ is not equal to zero for hypothesis, because we stated that n does not depend on y, but it may depend on x. On the other hand, the second equation will be:

$$\frac{\mathrm{d}}{\mathrm{d}s}(n\Omega_y) = 0$$

this is the y component of the same equation; since we have two coordinates, we have a 2d problem, constituted by a system of two differential equations:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}s}(n\Omega_x) = \dfrac{\partial n}{\partial x} \\ \dfrac{\mathrm{d}}{\mathrm{d}s}(n\Omega_x) = 0 \end{cases}$$

if the light ray is solution of these two equations, given ϑ the angle between the $\hat{\Omega}$ unit vector (which identifies in each point the direction of the curvilinear abscissa) and $\hat{\mathrm{x}}$, we have that:

$$\Omega_x = \cos(\vartheta)$$

with the same idea, we can find, for $\hat{y}$, that the y component of $\hat{\Omega}$ is:

$$\Omega_y = \sin(\vartheta)$$

but, for the second equation, we have that:

$$\frac{\mathrm{d}}{\mathrm{d}s}(n\Omega_y) = \frac{\partial}{\partial s}(n\sin(\vartheta)) = 0$$

therefore, $n\sin(\vartheta)$ is a constant quantity; this is the **Snell's law refraction**. In other words, given a discontinuity of refraction index along x, but not along y, we have that:

$$n_1 \sin(\vartheta_1) = n_2 \sin(\vartheta_2)$$

1.3.2 Fermat's principle

There are several ways to write the Snell's law of refraction; the classical deduction is based on the **Fermat principle**, which allows us to find the geometrical optics theory. Fermat discovered that if we consider a light ray which travels from some point A to some other point B, light *chooses* the trajectory which minimizes the flight time from A to B. What does *flight time* means, exactly? Well, for instance the space can contain some dielectric, therefore the refraction index may be different; light chooses the minimum path, keeping into account also the material in the space: the minimum path, in this sense (which is **not** necessarily the shortest one!). Now we discuss this idea. We have that:

$$\mathrm{d}t = \frac{\mathrm{d}s}{\frac{c}{n}}$$

in fact, light travels at a speed equal to $\frac{c}{n}$, where n as usual is the refraction coefficient. Therefore, when light moves for a space $\mathrm{d}s$, with this speed, it needs a time equal to $\mathrm{d}t$. Inverting this:

$$c\,\mathrm{d}t = n\,\mathrm{d}s$$

we can apply the integrating operator at both terms, obtaining:

$$\int_A^B c\,\mathrm{d}t = \int_A^B n\,\mathrm{d}s$$

but:

$$c\,\mathrm{d}t = ct_{AB}$$

Fermat's principle says that:

$$\int_A^B n\,\mathrm{d}s = ct_{AB} = \text{ minimum}$$

the flight time is minimum, therefore the integral of the refraction coefficient along the path of the ray of light should be minimum; in this case, what it is not known *a priori* is the path: we imagine not-to-know which will be the path taken by light; but what Fermat states is that, knowing n in every point of the space, the path is the one which minimizes the line integral of the reflection coefficient (which is related to the flight time as we have proved).

The reason why we have introduced all this theory is the fact that this is a **variational problem**: the path is the one which minimizes this integral functional! In this case, actually, it is not necessary to apply any classical variational technique, because the solution is already known: it is the equation of the geometrical optics! As a matter of fact, since we know that the trajectory of the ray of light satisfies the geometrical optics equation, we can say that its solution minimizes this integral. We know that:

$$\nabla S_0 = n\hat{\Omega}$$

now, let's focus on one particular light ray, which satisfies this equation; we project this equation on the $\hat{\Omega}$ of this equation, so:

$$\frac{\partial S_0}{\partial s} = \hat{\Omega} \cdot (\nabla S_0) = n$$

if we calculate this integral:

$$\int_A^B n\,\mathrm{d}s = \int_A^B \frac{\mathrm{d}S_0}{\mathrm{d}s}\,\mathrm{d}s = \int_A^B \mathrm{d}S_0 = S_0|_A^B = S_0(B) - S_0(A)$$

we obtain that the integral is just the difference of S_0 in these two points; this value **has to be the minimum possible value**. $\hat{\Omega}$ has been chosen exactly equal to the $\hat{\Omega}$ of the geometric optics equation.

If we choose a different line, identified by a different unit vector $\hat{\Omega}' \neq \hat{\Omega}$, we can prove that the integral will be surely greater. In this case, we have:

$$\frac{\mathrm{d}S_0}{\mathrm{d}s'} = \hat{\Omega}' \cdot \nabla S_0 = n\hat{\Omega}' \cdot \hat{\Omega}$$

in this case, $\hat{\Omega}' \cdot \hat{\Omega} \leq 1$, because they are not parallel! $\hat{\Omega}'$ in fact is a different ray of light, so it is not parallel to the $\hat{\Omega}$ of the geometric optics differential equation. Therefore:

$$\frac{\mathrm{d}S_0}{\mathrm{d}s'} \leq n$$

now, let us calculate the integral:

$$\int_A^B \frac{\mathrm{d}S_0}{\mathrm{d}s'}\mathrm{d}s' \leq \int_A^B n\mathrm{d}s'$$

therefore, in general:

$$S_0|_A^B \leq \int_A^B n\,\mathrm{d}s'$$

but we proved that:

$$S_0|_A^B = \int_A^B n\,\mathrm{d}s$$

so:

$$\int_A^B n\,\mathrm{d}s \leq \int_A^B n\,\mathrm{d}s'$$

this means that, if our ray of light is on a trajectory s' which is different from the s defined from the geometric optics, it will not minimize this integral, therefore this **will not be the actual ray of light**, due to the Fermat's principle! It will not minimize the integral!

Now, let's start from:

$$\frac{\mathrm{d}}{\mathrm{d}s}(n\hat{\Omega}) = \nabla n$$

Let's focus on this, and study its aspects. First of all, let's write it in an expanded fashion, using the Leibnitz's rule:

$$\frac{\mathrm{d}}{\mathrm{d}s}(n\hat{\Omega}) = \frac{\mathrm{d}n}{\mathrm{d}s} + n\frac{\mathrm{d}\hat{\Omega}}{\mathrm{d}s} = \nabla n$$

Now, let's work on the derivative of $\hat{\Omega}$: we have that

$$\hat{\Omega} \cdot \hat{\Omega} = \hat{\Omega}^2 = 1$$

now, let's write the derivative of $\hat{\Omega}^2$:

$$\frac{\mathrm{d}\hat{\Omega}^2}{\mathrm{d}s} = 2\hat{\Omega} \cdot \frac{\mathrm{d}\hat{\Omega}}{\mathrm{d}s} = 0$$

in fact, since $\hat{\Omega}^2 = 1$, so since it is constant, its derivative is zero; in order to prove in a more formal fashion this equation, it is possible as usual to write it by components:

$$\hat{\Omega}^2 = \hat{\Omega} \cdot \hat{\Omega} = \sum_{i=1}^{n} \Omega_i \Omega_i$$

so, exploiting the linearity of the derivative and the Leibnitz's rule:

$$\frac{\mathrm{d}}{\mathrm{d}s}\sum_{i=1}^{n}\Omega_i\Omega_i = \sum_{i=1}^{n}\Omega_i\frac{\mathrm{d}\Omega_i}{\mathrm{d}x_i}+\Omega_i\frac{\mathrm{d}\Omega_i}{\mathrm{d}s} = \sum_{i=1}^{n}\Omega_i\frac{\mathrm{d}\Omega_i}{\mathrm{d}x_i}+\sum_{i=1}^{n}\Omega_i\frac{\mathrm{d}\Omega_i}{\mathrm{d}x_i} = 2\sum_{i=1}^{n}\Omega_i\frac{\mathrm{d}\Omega_i}{\mathrm{d}x_i} = 2\hat{\Omega}\cdot\frac{\partial\hat{\Omega}}{\mathrm{d}s}$$

This equation says that the derivative of $\hat{\Omega}$ with respect to s is orthogonal to $\hat{\Omega}$ itself; this means that:

$$\frac{\mathrm{d}\hat{\Omega}}{\mathrm{d}s} = \frac{\hat{N}}{R}$$

where $\hat{N}$ is the unit vector normal to $\hat{\Omega}$, while R is the **curvature radius**: this is the radius of the circle tangent to the curve in each point; in other words, it is the **normal unit vector**, which is parallel to the centripetal acceleration vector; given $\hat{\Omega}$ parallel to the velocity (in a classical mechanics framework), $\hat{N}$ is normal to it. This is a classical formula of the differential geometry of curves.

We know that the directional derivative along s of the refraction coefficient is:

$$\frac{\mathrm{d}n}{\mathrm{d}s} = \hat{\Omega} \cdot \nabla n$$

therefore, it is possible to write the equation of geometrical optics in the following way:

$$\frac{\hat{N}}{R} = \frac{\nabla n - (\hat{\Omega} \cdot \nabla n)\hat{\Omega}}{n}$$

this means that, from ∇n, we subtract its projection along $\hat{\Omega}$; by this way, since this is a vector subtraction, we eliminate the component along $\hat{\Omega}$, obtaining only a component along $\hat{N}$.

In fact, given ∇n, we have that $\hat{\Omega} \cdot \nabla n$ is the projection of the gradient towards $\hat{\Omega}$; gradient minus projection means that the resulting vector is directed towards $\hat{N}$; so, we can synthetically write this equation as follows, recalling that $\frac{\mathrm{d}n}{\mathrm{d}s} + n\frac{\mathrm{d}\hat{\Omega}}{\mathrm{d}s} = \nabla n$:

$$\frac{\hat{N}}{R} = \frac{(\nabla n)_\perp}{n}$$

therefore, taking the norm of this equation, we obtain:

$$\frac{1}{R} = \frac{|(\nabla n)_\perp|}{n}$$

this is the norm of the perpendicular component of the gradient of the refraction coefficient. This allows us to calculate the radius of curvature of the ray of light.

Now, let's stop focusing on optics for a moment, and let's focus on **classical mechanics**: this equation is not very different from the equation of an object moving in a potential field: given $U(\underline{x})$ a potential field, the equation of the motion is the Newton's las:

$$m\underline{a} = \underline{F}$$

where $\underline{F}$ can be found as the gradient of the potential, changing the sign:

$$\underline{F} = -\nabla U$$

therefore, given the trajectory of a point, the acceleration can be decomposed in two contributions: one is parallel to the tangent line, while the other one is normal to it; this second contribution is called **centripetal acceleration**:

$$\underline{a} = \underline{a}_\perp + \underline{a}_\|$$

but from basic physics we know that:

$$\underline{a}_\perp = \frac{v^2}{R}\hat{N}$$

where $\hat{N}$ has the very same meaning as before. Therefore, since we are just dividing acceleration in two parts, we can state that Newton's law is valid for both parts: we just have to divide this equation in two equations; considering the equation related to the centripetal acceleration, we have:

$$m\underline{a}_\perp = \underline{F}_\perp$$

therefore:

$$m\frac{v^2}{R}\hat{N} = [-\nabla U]_\perp$$

this, considering only the normal component of the gradient. Here, we have mass, which is a constant; velocity, which is not constant; therefore, we can write that:

$$\frac{1}{2}mv^2 + U = E$$

where E is the total energy, given by the sum of kinetic energy and potential energy. Inverting this formula:

$$mv^2 = 2(E - U)$$

but, from previous equation, if we calculate the gradient of the right term, and take the normal component:

$$\frac{\hat{N}}{R} = \frac{[(\nabla(-U)]_\perp}{2(E - U)}$$

Now, if we consider the fact that E is a constant, so that $\nabla E = 0$, we can add it inside the gradient sign, and obtain:

$$\frac{\hat{N}}{R} = \frac{[(E - \nabla(-U)]_\perp}{2(E - U)}$$

this seems to be a *close relative* to the geometrical optics equation previously written; we found that:

$$\frac{\hat{N}}{R} = \frac{(\nabla n)_\perp}{n}$$

except for the two, it is the same. But, if we define:

$$n \longrightarrow \sqrt{E - U}$$

we obtain:

$$\frac{[\nabla(\sqrt{E - U})]_\perp}{\sqrt{E - U}} = \frac{[(E - \nabla(-U)]_\perp}{2(E - U)}$$

in fact:

$$\nabla(\sqrt{E - U}) = \frac{1}{2}\frac{1}{\sqrt{E - U}}\nabla(E - U)$$

therefore, one equation becomes the other.

This means that the trajectory of the light ray equals the one of the photon, simply substituting to n this combination of the total energy E and of the potential U; this happens because light and matter follow similar rules, therefore the trajectory of photons is the same of the light ray one. This means that this variational principle known as Fermat's principle has a *cousin* in classical mechanics: also from a material point of view it is possible to establish a variational principle which states that the trajectory of a point comes from the minimization of an integral functional. In this case we just have to perform the substitution to n in the integral, obtaining:

$$\int_A^B \sqrt{E - U}\, \mathrm{d}s = \text{ minimum}$$

therefore, the actual trajectory of the particles is the one which minimizes this integral functional. This principle is called **Maupertuis Principle**.

This principle is very important, for this reason; we started from the wave equation:

$$\nabla^2 \phi + \frac{n^2 \omega^2}{c^2}\phi = 0$$

we obtained, from its approximation, geometrical optics, which can be summarized by means of the Fermat principle, which states that it can be found by minimizing an integral functional. On the other side, we introduced classical mechanics, which can be summarized by the Maupertuis principle; what is missing, is the equation *equivalent* to the wave equation, for **classical mechanics**. In fact, at the beginning of the 20th century, scientists discovered that, for elementary particles, it is necessary to use a wave approach also for particles; therefore, we can ask to ourselves: is there some wave equation describing how matter *propagate* in space? The idea is to write something like:

$$\nabla^2 \Psi + C(E - U)\Psi = 0$$

the $(E - U)$ term comes from the Maupertuis principle; C is some constant which should be determined. This equation is different from a *point* propagation: in the usual meaning, a wave propagates in space; here, the evolution of this wave is representable in a different way: now, we can analyze this wave with a line, with a trajectory, not with a space surface.

Let us find this constant C: given a free particle (therefore, potential $U = 0$), we have:

$$\nabla^2 \Psi + CE\Psi = 0$$

the solution for this equation is very easy: just like before, *plane waves*:

$$\Psi = e^{jkx}$$

from where, we can say, substituting this:

$$\nabla^2 \Psi = \nabla^2 e^{jkx} = -k^2 e^{jkx} = -k^2 \Psi$$

therefore:

$$-k^2 \Psi + CE\Psi = 0$$

but, **from experiments**, we know that:

$$p = \hbar k$$

this is simply the De Broglie formula, which comes also on experiments on diffraction of particles. On the other hand, since in this case there is no

potential energy, the total energy equals the kinetic energy; moreover, $p = mv$:

$$E = \frac{1}{2}mv^2 = \frac{1}{2m}m^2v^2 = \frac{p^2}{2m}$$

finally, from the equation, simplifying Ψ:

$$k^2 = CE \implies \frac{p^2}{\hbar^2} = C\frac{p^2}{2m}$$

which leads to:

$$C = \frac{2m}{\hbar^2}$$

therefore, the equation is, by applying the superposition of a potential U:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + U\Psi = E\Psi$$

this is called **Schrödinger equation**, and this is the equation of matter waves; for $U = 0$, which is the case of free particles, matter follow the geometrical optics, since the Schrödinger equation degenerates in the wave equation.

Chapter 2

Calcolo variazionale

2.1 Panoramica sul calcolo variazionale

Il calcolo variazionale si può vedere come una sorta di calcolo dei minimi di una funzione di una o più variabili; in verità però, invece che una funzione, si cerca di minimizzare un **funzionale**: questo significa che si cerca sempre di minimizzare un numero, dove però quelli che si variano non sono punti, ma funzioni: si cerca una certa funzione tale per cui, elaborata in qualche modo (per esempio, integrata), essa restituisca un numero che sia **il minimo possibile**. Si consideri dunque una certa funzione $L(y, y', t)$, dove essa è funzione di una certa funzione $y(t)$, della sua derivata rispetto a t, e del parametro t (per esempio t può assumere il significato di **tempo**). Si definisce un funzionale F come:

$$F(y) = \int_a^b L\left(y, \frac{\mathrm{d}y}{\mathrm{d}t}, t\right) \, \mathrm{d}t$$

il nostro obiettivo è minimizzare $F(y)$: con $y(t)$ si ricava L, la si integra, e ciò produce un numero: il risultato del funzionale applicato a $y(t)$.

Esempio: curva di minima lunghezza

Per capire qual è il senso di questo tipo di formulazione, si immagini di voler trovare la funzione $y(x)$ tale per cui, fissati due punti A e B, l_{AB} è minima. Per fare ciò, prima di tutto si calcoli il differenziale:

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = \mathrm{d}x^2 \left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 \right]$$

dunque:

$$\mathrm{d}s = \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2} \, \mathrm{d}x$$

$$l_{AB} = \int_A^B \mathrm{d}s = \int_A^B \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2} \, \mathrm{d}x$$

Cenni ad altri problemi

Un altro problema che si può studiare con queste tecniche è la determinazione della **brachistocrona**: dati due punti A e B, si può cercare la curva $y(x)$ tale per cui, se una pallina scende dalla curva definita da questa, il tempo che essa impiega per arrivare da A a B deve essere minimo. Questo è una sorta di piano inclinato generalizzato, dove l'obiettivo è proprio determinare l'andamento dell'inclinazione del piano: più si *attribuisce una curvatura* alla curva, più essa diventa lunga; d'altro canto, se essa è rettilinea, allora ci mette più tempo a prendere velocità, e quindi si perde tempo in questo modo. Il risultato sarà dunque un trade-off tra la lunghezza della curva e il tempo che la pallina ci impiega a prendere velocità.

Un altro esempio di problema è lo studio della forma ottimale dei fili della luce: la **catenaria**. In questo caso il problema è ancora più complicato, dal momento che il problema è la **ricerca di un minimo con vincolo**: si vuole la curva che minimizza l'energia potenziale, dove però la lunghezza ha un valore fisso L: questo introduce un vincolo aggiuntivo al problema precedente: la lunghezza del filo.

2.2 Equazioni di Eulero-Lagrange

Le leggi della meccanica possono essere scritte a partire dal **principio di minima azione**: se un oggetto in un tempo t_a si trova in un punto x_a, e questo in t_b deve trovarsi in x_b, il principio di minima azione afferma che esiste un S tale per cui:

$$S = \int_{t_a}^{t_b} L\left(\vec{x}(t), \frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t}, t\right) \mathrm{d}t = \text{ minimo}$$

di conseguenza, di tutte le leggi orarie del moto possibili $\vec{x}(t)$, questa L è quella relativa alla $\vec{x}(t)$ tale per cui l'integrale S sia minimo.

A questo punto, si proporrà un esempio di ciò. Si consideri il seguente problema:

$$\begin{cases} F(y) = \int_a^b L\left(\vec{x}(t), \dfrac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t}, t\right) \mathrm{d}t \\ y(a) = y_a \\ y(b) = y_b \end{cases}$$

si supponga che questo problema sia già risolto: ossia, si immagini di aver trovato una funzione, $\overline{y}$, tale per cui questo funzionale sia minimo. Di conseguenza, si ha che:

$$y = \overline{y} \ : \ F(\overline{y}) = \text{ minimo}$$

A questo punto, si vogliono studiare le proprietà di questa, utilizzando il seguente stratagemma: si supponga di prendere una funzione $y(t) \neq \overline{y}(t)$, e la si pensi in questo modo:

$$y(t) = \overline{y}(t) + \alpha\eta(t)$$

$y(t)$ deve comunque *avere senso*, ossia essa deve comunque soddisfare i vincoli: $y(a) = y_a, y(b) = y_b$. Questo significa che $\eta(a) = \eta(b) = 0$: questa è l'unica condizione che poniamo sulla funzione η; $\alpha \in \mathbb{R}$. A questo punto, si valuti il funzionale sulla funzione $y(t)$ appena definita:

$$F(y) = F(\overline{y} + \alpha\eta)$$

questo significa che, fissato η, si ha solo più dipendenza da α, che è uno scalare; questo permette di definire come segue $f(\alpha)$:

$$f(\alpha) = F(\overline{y} + \alpha\eta) = \int_a^b L\left(\overline{y} + \alpha\eta, \frac{\mathrm{d}\overline{y}}{\mathrm{d}t} + \alpha\frac{\mathrm{d}\eta}{\mathrm{d}t}, t\right) \mathrm{d}t$$

variando α, si avrà un diverso valore di questo integrale; tuttavia, è evidente che, per $\alpha = 0$, questo integrale sia minimo, dal momento che in questo caso

si ha che $y = \overline{y}$. Se si ha un minimo di $f(\alpha)$, dal momento che questa è una funzione scalare, si avrà che la derivata rispetto ad α di $f(\alpha)$ dovrà essere nulla: essendo un minimo, questo è un punto stazionario, dunque:

$$\left.\frac{\mathrm{d}f(\alpha)}{\mathrm{d}\alpha}\right|_{\alpha=0} = 0$$

A questo punto, si calcoli la derivata, e si ottenga:

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\int_a^b L\left(\overline{y} + \alpha\eta, \frac{\mathrm{d}\overline{y}}{\mathrm{d}t} + \alpha\frac{\mathrm{d}\eta}{\mathrm{d}t}, t\right)\mathrm{d}t$$

sfruttando il **teorema della Bontà Divina in forma forte**, si può dire che si può sempre commutare il segno di derivata e integrale[1]. Si ottiene dunque:

$$= \int_a^b \frac{\mathrm{d}}{\mathrm{d}\alpha} L\left(\overline{y} + \alpha\eta, \frac{\mathrm{d}\overline{y}}{\mathrm{d}t} + \alpha\frac{\mathrm{d}\eta}{\mathrm{d}t}, t\right)\mathrm{d}t$$

ora, si ha che:

$$\frac{\mathrm{d}L}{\mathrm{d}\alpha} = \frac{\partial L}{\partial y}\frac{\partial(\overline{y} + \alpha\eta)}{\partial\alpha} + \frac{\partial L}{\partial y'}\frac{\partial(\overline{y}' + \alpha\eta')}{\partial\alpha} =$$
$$= \frac{\partial L}{\partial y}\eta + \frac{\partial L}{\partial y'}\eta'$$

dove $L = L\left(\overline{y} + \alpha\eta, \frac{\mathrm{d}\overline{y}}{\mathrm{d}t} + \alpha\frac{\mathrm{d}\eta}{\mathrm{d}t}, t\right) = (\overline{y} + \alpha\eta, \overline{y}' + \alpha\eta', t)$. Quindi, si ha che:

$$\int_a^b \left\{\frac{\partial L}{\partial y}\eta(t) + \frac{\partial L}{\partial y'}\eta'(t)\right\}\mathrm{d}t = 0$$

questa equazione deve essere verificata $\forall\eta(t)$, dove $\eta(t)$ deve solamente essere nulla ai bordi. Si valuti ora questa espressione in $\alpha = 0$: si ottiene

$$\int_a^b \left\{\frac{\partial L}{\partial y}(\overline{y}, \overline{y}', t)\eta(t) + \frac{\partial L}{\partial y'}(\overline{y}, \overline{y}', t)\eta'(t)\right\}\mathrm{d}t = 0$$

---

[1] il teorema della Bontà Divina in forma forte dice che si può sempre scambiare il segno di integrale e quello di limite, sommatoria, derivata, e così via; quello in forma debole, dice che quando non si può fare, viene fuori qualche boiata clamorosa e dunque ci si ferma prima di proseguire; in verità servirebbero i teoremi di passaggio al limite di Teoria della Misura, ma detto così è molto più bello.

come già detto, $\eta(t)$ è una funzione qualsiasi, quindi deve sempre essere verificata. Questo integrale dunque deve sempre annullarsi. A questo punto, si vuole scrivere questa in modo più semplice da capire; per fare ciò, si effettui l'integrazione per parti del secondo membro:

$$= \int_a^b \frac{\partial L}{\partial y}(\overline{y}, \overline{y}', t)\eta(t)\mathrm{d}t - \int_a^b \frac{\partial^2 L}{\partial y' \partial t}(\overline{y}, \overline{y}', t)\eta(t)\mathrm{d}t + \frac{\partial L}{\partial y'}\eta(t)\Big|_a^b = 0$$

ma $\eta(a) = \eta(b) = 0$, quindi:

$$\Longrightarrow \int_a^b \left\{ \frac{\partial L}{\partial y}(\overline{y}, \overline{y}', t) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y'}\frac{\partial L}{\partial y}(\overline{y}, \overline{y}', t) \right\} \eta(t)\,\mathrm{d}t = 0$$

dal momento che questa deve valere $\forall \eta$, questo è possibile solo se l'altro termine è nullo; si ottiene dunque:

$$\frac{\partial L}{\partial y}(\overline{y}, \overline{y}', t) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y'}(\overline{y}, \overline{y}', t) = 0$$

questa è la **equazione di Eulero-Lagrange**, e rappresenta la soluzione del problema variazionale appena affrontato come discuteremo tra breve.

2.2.1 Esempio applicativo: deduzione della legge di Newton

A questo punto si applica questa teoria al fine di vedere come essa può essere efficace. Si consideri come funzione $y(t)$ (per come è stata introdotta nella precedente sezione) la $x(t)$, ossia la **legge oraria** di un moto. Dunque:

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

dove $v(t)$ sarà la **velocità** di un oggetto nello spazio. Dunque, si avrà una lagrangiana funzione di $x(t)$, di $v(t)$, e di t:

$$L = L\left(x(t), \frac{\mathrm{d}x(t)}{\mathrm{d}t}, t\right) = L(x, v, t)$$

quindi, si vuole cercare la situazione tale per cui:

$$\int_a^b L(x, v, t)\,\mathrm{d}t = \text{ minimo}$$

A questo punto, applicando tutta la teoria precedentemente vista, è possibile scrivere per questa situazione l'equazione di Eulero-Lagrange:

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

ora dimostreremo che le equazioni della meccanica si possono ottenere minimizzando un certo funzionale. Per fare ciò, è necessario definire la lagrangiana. Questa viene definita come differenza tra l'energia cinetica del sistema e l'energia potenziale $U(x)$:

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

S, ossia l'integrale della lagrangiana, viene detto **azione**:

$$S=\int_a^b L(x,v,t)\,\mathrm{d}t$$

il **principio di minima azione** afferma che la legge oraria del moto è quella che minimizza l'azione S. Dato dunque un oggetto che si muove nel campo di potenziale $U(x)$, la legge di minima azione è quella che minimizza l'integrale, ossia che minimizza la differenza tra energia cinetica e potenziale. Partendo dalle equazioni di Eulero-Lagrange, e minimizzando, si trovano le leggi di Newton; infatti, si calcolino i vari termini dell'equazione di Eulero-Lagrange:

$$\frac{\partial L}{\partial v}=mv$$

$$\frac{\partial L}{\partial x}=-\frac{\partial U}{\partial x}$$

infatti, l'energia cinetica $(\frac{1}{2}mv^2)$ non è funzione della posizione x, mentre si ha che il potenziale U è funzione solo della posizione; tutto ciò è stato dunque fatto derivando la lagrangiana rispetto a queste variabili. Dunque, per l'equazione di Eulero-Lagrange, si ha che:

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=-\frac{\partial U}{\partial x}$$

ossia:

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

questa è semplicemente la legge di Newton!

In questo caso, questa formulazione è stata utilizzata al fine di ricavare un risultato banale, dunque per ora non sembrerebbe molto interessante. Tuttavia, lo è: questa formulazione è estremamente utile qualora si voglia studiare il **moto di oggetti vincolati**; volendo studiare ciò con un approccio generale, si può arrivare a espressioni molto complicate; al contrario, partendo da queste espressioni, è possibile ottenere una formulazione molto più lineare e trovare risultati in modo molto più diretto. Un esempio di ciò verrà proposto in seguito.

2.3 Formulazione Hamiltoniana della meccanica classica

Abbiamo appena dimostrato che, definita una lagrangiana $L(x, v, t)$, è stato possibile, mediante le equazioni di Eulero-Lagrange, ottenere la legge di Newton. A questo punto però si vuole fornire una formulazione alternativa a tutto ciò. Per fare ciò, si introduce una grandezza p, detta **momento coniugato**, definita come la derivata rispetto alla velocità della lagrangiana:

$$\frac{\partial L}{\partial v}(x, v, t) \triangleq p$$

di conseguenza, dal momento che, nello specifico caso della meccanica classica,

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

si ha che:

$$\frac{\partial L}{\partial v} = p = mv$$

dunque, **nel caso della meccanica classica**, l'impulso coniugato p coincide con la quantità di moto.

A questo punto, dal momento che la velocità v non sarà una grandezza primaria, nell'ambito della formulazione hamiltoniana, si vuole ricavare la velocità come funzione della posizione, del momento coniugato, e del tempo:

$$v = \mathcal{V}(x, p, t)$$

in generale, questa può essere un'espressione molto complicata; in questo caso semplice, però, da quello che si è appena visto:

$$\mathcal{V}(x, p, t) = \frac{p}{m}$$

si riprenda a questo punto l'equazione di Eulero-Lagrange, e si sostituisca quanto visto finora; si ha:

$$\frac{\partial p}{\partial t} = \frac{\partial L}{\partial x}$$

di conseguenza:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial L}{\partial x}(x, \mathcal{V}(x, p, t), t)$$

inoltre, si ha che la velocità può essere scritta come segue:

$$\mathcal{V}(x, p, t) = \frac{\mathrm{d}x}{\mathrm{d}t}$$

dunque, abbiamo spezzato le equazioni di Eulero-Lagrange in due equazioni:

$$\begin{cases} \dfrac{\mathrm{d}p}{\mathrm{d}t} = \dfrac{\partial L}{\partial x}(x, \mathcal{V}(x, p, t), t) \\ \dfrac{\mathrm{d}x}{\mathrm{d}t} = \mathcal{V}(x, p, t) \end{cases}$$

dunque, sostituendo:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial U}{\partial x}$$

e

$$\frac{\partial x}{\partial t} = \frac{p}{m}$$

in questo caso, tutto è stato molto semplice, dal momento che l'espressione di $\mathcal{V}(x, p, t)$ è molto semplice; in generale, però, questa può essere molto più complicata. Tutti questi passaggi possono essere utilizzati al fine di scrivere la meccanica mediante la formulazione di Hamilton: si definisce la **hamiltoniana** come la funzione $H(x, p, t)$:

$$H(x, p, t) = p\mathcal{V}(x, p, t) - L(x, \mathcal{V}(x, p, t), t)$$

dove il secondo termine è la **lagrangiana**, definita nelle sezioni precedenti. Data H, il sistema si può scrivere come segue: date x, t le variabili indipendenti:

$$\frac{\partial H}{\partial x} = p\frac{\partial \mathcal{V}}{\partial x} - \frac{\partial L}{\partial x} - \frac{\partial L}{\partial v}\frac{\partial \mathcal{V}}{\partial x}$$

però:

$$\frac{\partial L}{\partial v} = p$$

quindi, i due termini si compensano, e si ottiene:

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}$$

invece, derivando H rispetto a p, si ottiene:

$$\frac{\partial H}{\partial p} = \mathcal{V}(x, p, t) + p\frac{\partial \mathcal{V}}{\partial p} - \frac{\partial L}{\partial v}\frac{\partial \mathcal{V}}{\partial p}$$

ma:

$$\frac{\partial L}{\partial v} = p$$

di nuovo, e dunque:

$$\frac{\partial H}{\partial p} = \mathcal{V}(x, p, t)$$

effettuando queste sostituzioni, il sistema di equazioni di prima può essere **in generale** scritto come:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{\partial H}{\partial p} \\ \dfrac{\mathrm{d}p}{\mathrm{d}t} = -\dfrac{\partial H}{\partial x} \end{cases}$$

Questo, è un risultato **generale**: è vero per ogni lagrangiana e per ogni $\mathcal{V}$. Volendo analizzare il caso della meccanica classica, si trovano i risultati appena visti. Il procedimento dunque da seguire è qui riportato.

1. Si parte da una lagrangiana.

2. Dalla lagrangiana, mediante derivazione rispetto alla velocità, si arriva a definire il momento p.

3. Si scrive la lagrangiana H note x,p, $\mathcal{V}$, L.

4. A questo punto il sistema appena scritto è valido, e quindi la formulazione è terminata.

2.3.1 Riscrittura delle equazioni di Eulero-Lagrange mediante formulazione *veloce*

Si vuole a questo punto riscrivere un risultato precedentemente ottenuto, utilizzando però una notazione più *veloce*; in questo modo, la formulazione di problemi più complicati sarà comunque più immediata. Per chiarire questa notazione con la precedente, è possibile confrontare i vari passaggi.

Si definisce l'azione $S(x)$ come:

$$S(x) = \int L(x, x', t)\mathrm{d}t = \min$$

invece di considerare x soluzione esatta, si considera $x + \delta x$, per ripetere il ragionamento di prima:

$$S(x + \delta x) = \int L(x + \delta x, x' + \delta x', t)\mathrm{d}t =$$

$$\int L(x, x', t)\mathrm{d}t + \int \left(\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial x'}\delta x' \right)\mathrm{d}t = S + \delta S$$

dove:

$$\delta S = S(x + \delta x) - S(x) = \int_a^b \left(\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial x'}\delta x' \right)\mathrm{d}t$$

a questo punto si effettua direttamente l'integrazione per parti:

$$= \int_a^b \left[\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial x'} \right) \right](\delta x)\mathrm{d}t = 0$$

questo è il risultato ottenuto prima. Ora, questa notazione più *snella* verrà adottata per risolvere problemi nuovi.

2.3.2 Esempio di problema di minimizzazione in più variabili

A questo punto, si vuole applicare questa notazione più *snella* a un problema un po' più complicato del precedente: si considererà ora un funzionale che dipende da due funzioni. In questo caso nel dettaglio si hanno due funzioni: $x(t)$ e $y(t)$; questo significa che la legge oraria riguarda un **piano**, o una superficie. Di conseguenza:

$$S(x, y) = \int_a^b L(x, y, x', y', t) \, \mathrm{d}t$$

a questo punto, si vuole trovare una coppia di funzioni (x, y) che minimizzi l'azione S, dove in questo caso l'azione S è un funzionale di due funzioni. Di conseguenza, si prendano due variabili indipendenti, δx e δy, tali da annullarsi sugli estremi a e b:

$$\delta S = \int_a^b \left[\frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial y} \delta y + \frac{\partial L}{\partial x'} \delta x' + \frac{\partial L}{\partial y'} \right] \delta y' \mathrm{d}t$$

applicando l'integrazione per parti, questo diventa:

$$= \int_a^b \left(\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial x'} \right) \delta x \mathrm{d}t + \int_a^b \left(\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right) \delta y \mathrm{d}t = 0$$

per avere che questa sia uguale a zero, con δx e δy arbitrari, entrambi i membri che moltiplicano δx e δy devono tendere a zero; di conseguenza, date due funzioni, in questo caso si avranno ancora le equazioni di Eulero-Lagrange, ma questa volta se ne avranno due, per giunta accoppiate tra loro:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \left(\dfrac{\partial L}{\partial x'} \right) = \dfrac{\partial L}{\partial x} \\ \dfrac{\mathrm{d}}{\mathrm{d}t} \left(\dfrac{\partial L}{\partial y'} \right) = \dfrac{\partial L}{\partial y} \end{cases}$$

questo significa che, se il problema di minimizzazione riguarda un funzionale di due funzioni, allora si hanno due equazioni. Ovviamente, così come il formalismo lagrangiano può essere esteso, allo stesso modo può essere fatto per quanto concerne il formalismo hamiltoniano; si definisce ancora una volta il momento coniugato, per componenti:

$$\frac{\partial L}{\partial x'} = p_x \qquad \frac{\partial L}{\partial y'} = p_y$$

di conseguenza, facendo la stessa cosa fatta prima, si ottiene tutto come prima:

$$x' = \mathcal{V}_x(x, y, p_x, p_y, t)$$

$$y' = \mathcal{V}_y(x, y, p_x, p_y, t)$$

di conseguenza, l'hamiltoniana sarà più complicata, ma nella stessa forma della precedente:

$$H(x, y, p_x, p_y, t) = p_x \mathcal{V}_x + p_y \mathcal{V}_y - L(x, y, \mathcal{V}_x, \mathcal{V}_y, t)$$

e, infine, valgono le equazioni di Hamilton; riportando gli stessi ragionamenti:

$$\begin{cases} \dfrac{\partial x}{\partial t} = \dfrac{\partial H}{\partial p_x} \\[2mm] \dfrac{\partial y}{\partial t} = \dfrac{\partial H}{\partial p_y} \\[2mm] \dfrac{\partial p_x}{\partial t} = -\dfrac{\partial H}{\partial x} \\[2mm] \dfrac{\partial p_y}{\partial t} = -\dfrac{\partial H}{\partial y} \end{cases}$$

Esempio

Per concludere questo argomento, si consideri ora un esempio. Si consideri:

$$L(\vec{x}, \vec{v}, t) = \frac{1}{2}mv^2 - U(\vec{x}, t)$$

dove:

$$v = |\vec{v}|$$

$$p = |\vec{p}|$$

Si ha:

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = m\vec{v}$$

quindi:

$$\vec{v} = \frac{\vec{p}}{m}$$

quindi:

$$H(\vec{x}, \vec{v}, t) = \vec{p} \cdot \frac{\vec{p}}{m} - \left(\frac{1}{2}m\frac{p^2}{m^2} - U(\vec{x}, t)\right)$$

quindi:

$$H(\vec{x}, \vec{v}, t) = \vec{p} \cdot \frac{\vec{p}}{2m} + U(\vec{x}, t)$$

Da qui:

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}}{m}$$

e

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \vec{x}} = -\frac{\partial U}{\partial \vec{x}}$$

2.4   Use of polar coordinates

Now, the hamiltonian formulation will be applied in order to find a particular kind of motion equation. Up to this moment, we only use cartesian coordinates; in this situation, problems are very easy, therefore the use of hamiltonian formulations may seem useless; in this section, a more complicated example will be introduced. More in details, in this case we want to study the motion of a particle on a plain, where the potential energy is only function of r.

Knowing x and y components from a polar reference system (r, ϑ) is easy; it is possible to use the following transformation:

$$x = r\sin\vartheta$$
$$y = r\cos\vartheta$$

however, **it is better to work on polar coordinates**, avoiding the intro-
duction of conversions from polar to cartesian. It is possible to define the
velocity as:

$$v^2 = \left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}t}\right)^2$$

this is some kind of Pitagora theorem: we add the square of the velocity
toward the radial direction r and the one toward the angular direction ϑ.
This means that:

$$v_r = \frac{\mathrm{d}r}{\mathrm{d}t}$$

$$v_\vartheta = r\frac{\mathrm{d}\vartheta}{\mathrm{d}t}$$

Now, let us look at the lagrangian of this problem:

$$L = L\left(r, \vartheta, \frac{\mathrm{d}r}{\mathrm{d}t}, \frac{\mathrm{d}\vartheta}{\mathrm{d}t}, t\right)$$

if we define as usual the lagrangian as kinetic energy minus potential energy,
we obtain:

$$L = \frac{1}{2}mv^2 - U(r) = \frac{1}{2}m\left[\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}t}\right)^2\right] - U(r)$$

now, we introduce a faster notation: to indicate the derivative of r with
respect to time, we use $\dot{r}$, and the same for ϑ:

$$\dot{r} = \frac{\mathrm{d}r}{\mathrm{d}t}$$

$$\dot{\vartheta} = \frac{\mathrm{d}\vartheta}{\mathrm{d}t}$$

Therefore:

$$L(r, \vartheta, \dot{r}, \dot{\vartheta}, t) = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\vartheta}^2\right)$$

This problem has two variables, therefore we have two momenta:

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}$$

$$p_\vartheta = \frac{\partial L}{\partial \dot{\vartheta}} = mr^2\dot{\vartheta}$$

we can therefore find the functions $\mathcal{V}_r$ and $\mathcal{V}_\vartheta$ inverting these expressions:

$$\mathcal{V}_r = \dot{r} = \frac{p_r}{m}$$

$$\mathcal{V}_\vartheta = \dot{\vartheta} = \frac{p_\vartheta}{mr^2}$$

This means that we can write the hamiltonian as:

$$
\begin{aligned}
H(r, \vartheta, p_r, p_\vartheta, t) &= p_r \mathcal{V}_r + p_\vartheta \mathcal{V}_\vartheta - L(r, \vartheta, \dot{r}, \dot{\vartheta}, t) = \\
&= p_r \dot{r} + p_\vartheta \dot{\vartheta} - L(r, \vartheta, \dot{r}, \dot{\vartheta}, t) = \\
&= \frac{p_r^2}{m} + \frac{p_\vartheta^2}{mr^2} - \left[\frac{1}{2}m\frac{p_r^2}{m^2} + \frac{1}{2}mr^2\frac{p_\vartheta^2}{m^2r^4} - U(r) \right] = \\
&= \frac{p_r^2}{2m} + \frac{p_\vartheta^2}{2mr^2} + U(r)
\end{aligned}
$$

Now, from the hamiltonian, the deduction of the equation of this motion are straightforward; in fact:

$$
\begin{cases}
\dfrac{\mathrm{d}r}{\mathrm{d}t} = \dfrac{\partial H}{\partial p_r} = \dfrac{p_r}{m} \\[2mm]
\dfrac{\mathrm{d}\vartheta}{\mathrm{d}t} = \dfrac{\partial H}{\partial p_\vartheta} = \dfrac{p_\vartheta}{mr^2} \\[2mm]
\dfrac{\mathrm{d}p_r}{\mathrm{d}t} = -\dfrac{\partial H}{\partial r} = -\dfrac{\partial U}{\partial r} - \dfrac{\partial p_\vartheta^2}{mr^3} \\[2mm]
\dfrac{\mathrm{d}p_\vartheta}{\mathrm{d}t} = -\dfrac{\partial H}{\partial \vartheta} = -\dfrac{\partial U}{\partial \vartheta} = 0
\end{cases}
$$

the last equation states that p_ϑ is a constant of the motion; this means that, thanks to this equation, p_ϑ is a constant in every equation. Therefore, because of the fact that p_ϑ is constant, we have two equations and two unknowns:

$$\begin{cases} \dfrac{\mathrm{d}r}{\mathrm{d}t} = \dfrac{\partial H}{\partial p_r} = \dfrac{p_r}{m} \\ \dfrac{\mathrm{d}p_r}{\mathrm{d}t} = -\dfrac{\partial U}{\partial r} - \dfrac{\partial p_\vartheta^2}{mr^3} \end{cases}$$

once that we know $r(t)$, so the radial position as a function of time, it is possible to solve the remaining equation, and find:

$$\frac{\mathrm{d}\vartheta}{\mathrm{d}t} = \frac{p_\vartheta}{mr^2}$$

The use of the hamiltonian formulation allowed us to write the equations of motions with a very easy notation, also in this case, where the use of polar coordinates was recommended. Moreover, there is a very important simplification: in general, if the hamiltonian does not depend on one coordinate (just like in this case, where there was no dependance on ϑ, the momentum in that direction is constant.

2.5 Motion of a particle in an electromagnetic field

2.5.1 Law of motion (reference result)

Up to this moment we have analyzed situations related to classical mechanics; now, we are going to study the motion of a particle with charge q in an electromagnetic field. In this situation, the particle will move following the Lorentz's force:

$$\vec{F} = q\left(\vec{E} - \frac{\vec{v}}{\mathrm{c}} \times \vec{B}\right)$$

instead of using the usual electromagnetic quantities $\vec{E}$ and $\vec{B}$, we are going to formulate this problem by means of a potential formulations; therefore, let us define the scalar potential Φ and the vector potential $\vec{A}$ as:

$$\vec{B} = \nabla \times \vec{A}$$

$$\vec{E} = -\nabla\Phi - \frac{1}{\mathrm{c}}\frac{\partial \vec{A}}{\partial t}$$

The first equation states that the magnetic field can be written as the curl of another vector field; this is possible because the field $\vec{B}$ is solenoidal, which means that its divergence equals zero, and it is possible to prove that this admits this representation. Similar considerations for the curl of $\vec{E}$, which equals zero in the electrostatic case, therefore this allows to write $\vec{E}$ as a gradient of the scalar potential, plus another term (since in general the electric field is not static). Given these two formulas, it is possible to rewrite $\vec{F}$, using the following vector identity:

$$\vec{v} \times (\nabla \times \vec{A}) = \nabla(\vec{v} \cdot \vec{A}) - (\vec{v} \cdot \nabla)\vec{A}$$

this can be found for instance by writing component-by-component the vector:

$$\vec{v} \times (\nabla \times \vec{A}) = \frac{\partial}{\partial x_i}\left(\sum_{j=1}^{3} v_j A_j\right) - \left(\sum_{j=1}^{3} v_j \frac{\partial}{\partial x_j}\right) A_j$$

then from where we can exploit the fact that $\vec{v}$ does not depend on position, but only on time. Therefore, we can obtain that:

$$\vec{F} = m\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = q\left[-\nabla\Phi - \frac{1}{\mathrm{c}}\frac{\partial\vec{A}}{\partial t} + \nabla\left(\frac{\vec{v}}{\mathrm{c}} \cdot \vec{A}\right) - \frac{1}{\mathrm{c}}(\vec{v} \cdot \nabla)\vec{A}\right] =$$

$$= -q\left[\nabla\left(\Phi - \frac{\vec{v}}{\mathrm{c}} \cdot \vec{A}\right)\right] - \frac{q}{\mathrm{c}}\left[\frac{\partial}{\partial t} + \vec{v} \cdot \nabla\right]\vec{A}$$

the term:

$$\frac{\partial}{\partial t} + \vec{v} \cdot \nabla$$

is called **lagrangian derivative**; it evaluates, starting from a particular point in space, the total derivative. In fact, this is also called **total derivative** or **convective derivative**.

At this point, we can move on, and obtain, applying the lagrangian derivative definition (this means simply grouping the second term!):

$$m\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -q\nabla\left(\Phi - \frac{\vec{v}}{\mathrm{c}} \cdot \vec{A}\right) - \frac{q}{\mathrm{c}}\frac{\mathrm{d}\vec{A}}{\mathrm{d}t}$$

now, we can move the last term at the left-hand side member of the equation:

$$m\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} + \frac{q}{\mathrm{c}}\frac{\mathrm{d}\vec{A}}{\mathrm{d}t} = -q\nabla\left(\Phi - \frac{\vec{v}}{\mathrm{c}}\cdot\vec{A}\right)$$

now, using the derivative:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m\vec{v} + \frac{q}{\mathrm{c}}\vec{A}\right] = -q\nabla\left(\Phi - \frac{\vec{v}}{\mathrm{c}}\cdot\vec{A}\right)$$

This is the equation of the motion of a particle with mass m and charge q in an electromagnetic field. This equation can be also deduced from a particular form of the lagrangian.

2.5.2 Deduction of the equation of motion with the hamiltonian formulation

Now, we are going to deduce the equation of motion starting from another point. In an electrostatic field, we can write the lagrangian $L(\vec{x}, \vec{v}, t)$ as:

$$L(\vec{x}, \vec{v}, t) = \frac{1}{2}mv^2 - q\Phi(\vec{x})$$

this is trivial: it is simply the difference between the kinetic energy of the particle and the potential energy. However, this is true only for the static case; if the electric field and the magnetic field are not constant in time and space, this lagrangian should be written as follows:

$$L(\vec{x}, \vec{v}, t) = \frac{1}{2}mv^2 - q\Phi(\vec{x}) + \frac{q}{c}\vec{v}\cdot\vec{A}$$

This is our starting point; now, we can substitute this in the Euler-Lagrange equations, and obtain all our results. So, first of all, we recall the Euler-Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \vec{v}}\right) = \frac{\partial L}{\partial \vec{x}}$$

we know that the momentum $\vec{p}$ can be evaluated as:

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = m\vec{v} + \frac{q}{c}\vec{A}$$

(this is not immediate; in order to do this, it should be necessary to write by components the lagrangian, so in x_i, and then to perform the differentiation!). In this case, $\vec{p} \neq m\vec{v}$! About the right-hand side member:

$$\frac{\partial L}{\partial \vec{x}} = -q\nabla\Phi + \frac{q}{c}\nabla\left(\vec{v} \cdot \vec{A}\right)$$

$\vec{p}$ is exactly what is differentiated with respect to time in the left-hand side of the motion equation found in the previous subsection, while $\frac{\partial L}{\partial \vec{x}}$ is the right-hand side; therefore, the lagrangian is correct.

As well as we have used the Euler-Lagrange formulation, we can also use the Hamilton one. To this extent, the first step is to find $\mathcal{V}$:

$$\vec{v} = \frac{1}{m}\left(\vec{p} - \frac{q}{c}\vec{A}\right)$$

so:

$$H(\vec{x}, \vec{p}, t) = \vec{p} \cdot \vec{v} - L(\vec{x}, \vec{v}, t)$$

so:

$$H(\vec{x}, \vec{p}, t) = \frac{\vec{p}}{m} \cdot \left(\vec{p} - \frac{q}{c}\vec{A}\right) - \left[\frac{1}{2}m\frac{\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2}{m^2} - q\Phi + \frac{q}{c}\frac{\left(\vec{p} - \frac{q}{c}\vec{A}\right)}{m}\vec{A}\right]$$

now, with some arrangements:

$$H(\vec{x}, \vec{p}, t) = \frac{1}{m}\left(\vec{p} - \frac{q}{c}\vec{A}\right) \cdot \left(\vec{p} - \frac{q}{c}\vec{A}\right) - \frac{m}{2}\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2 + q\Phi =$$
$$= \frac{1}{2m}\left[\vec{p} - \frac{q}{c}\vec{A}\right]^2 + q\Phi$$

now, by writing again everything by components, it is possible to write (we write directly the result):

$$\frac{\partial \vec{x}}{\partial t} = \frac{\partial H}{\partial \vec{p}} = \frac{1}{m}\left(\vec{p} - \frac{q}{c}\vec{A}\right)$$

and (we write it well):

$$-\frac{\partial H}{\partial x_i} = \frac{1}{m}\left(\vec{p} - \frac{q}{c}\vec{A}\right) \cdot \frac{q}{c}\frac{\partial A}{\partial x_i}$$

which becomes:

$$-\frac{\partial H}{\partial \vec{x}} = \frac{1}{m}\left(\vec{p} - \frac{q}{c}\vec{A}\right) \cdot \nabla\vec{A}$$

but: $\frac{1}{m}\left(\vec{p} - \frac{q}{c}\vec{A}\right) = \vec{v}$; moreover, $\vec{v}$ is not dependent on position ($\vec{v}$ and $\vec{x}$ are independent variables), therefore the operator ∇ does not act on it; therefore:

$$= \nabla\left(\vec{v} \cdot \vec{A}\right)$$

so:

$$\frac{\partial \vec{p}}{\partial t} = -\frac{\partial H}{\partial \vec{x}} = \nabla\left(\vec{v} \cdot \vec{A}\right) - q\nabla\Phi$$

and this shows that the lagrangian/hamiltonian formulations can be used also in electromagnetics.

2.6 Minimization of a functional with constraints

We already discussed the problem of minimizing a functional, where some boundary conditions are introduced; now, we are going to solve a more complicated problem: the minimization of a functional where the only constraints are not only put on boundary conditions, but also on some other functional. For instance, let's consider the following problem: given A, B two points, we want the curve which connects the two points, **with a specified area beyond the curve**. This means that: $y(a)$, $y(b)$ are fixed, but it is not enough: we have to search the solution in the space of functions which have the integral from a to b equal to some constant. A similar problem is the determination of the curve which minimizes some energy functional (like the lagrangian), where the length is fixed.

The determination of the minimum length is the minimization of the following integral (we already proved it):

$$\int_a^b \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = \text{ minimum}$$

but we have to restrict the research among all the curves which satisfy also the following condition:

$$\int_a^b y\,\mathrm{d}x = S$$

where S is some fixed constant.

This was an example of problem. Now, we will focus on the general formulation of this kind of problems, where one constraint is present. In general, we want to minimize a functional:

$$\int_a^b L(y, y', x)\,\mathrm{d}x = \text{ minimum}$$

where there is a constraint about one functional $M(y, y', x)$:

$$\int_a^b M(y, y', x)\,\mathrm{d}x = \text{ constant}$$

How can we proceed? Well, to minimize the integral, what we had done previously was writing a generic function y as:

$$y = \overline{y} + \delta y = \overline{y} + \varepsilon\eta(x)$$

where $\overline{y}$ was the solution of the problem (the function which minimizes the functional), η some function satisfying the constraints about $y(a)$ and $y(b)$. Now, this is not enough: $\eta(x)$ should satisfy also the constraint about the functional M. Therefore, we have to add some degrees of freedom. In order to operate in this direction, we can tale:

$$y = \overline{y} + \varepsilon_1\eta_1(x) + \varepsilon_2\eta_2(x)$$

therefore, in this case:

$$\delta y = \varepsilon_1\eta_1(x) + \varepsilon_2\eta_2(x)$$

in this case, η_1, η_2 are two arbitrary functions, such that:

44

$$\eta_1(a) = \eta_2(a) = \eta_1(b) = \eta_2(b) = 0$$

ε_1 is an arbitrary coefficient; ε_2 will be chosen in order to satisfy the integral condition on M. Let us call the functional $B(y)$ as the functional which should be equal to a constant (the **constraint functional**):

$$B(y) = \int_a^b M(y, y', x) \, \mathrm{d}x = \text{ constant}$$

therefore, since $\overline{y}$ is the solution of the problem (and this means that it keeps also into account the constraint):

$$B(\overline{y}) = \int_a^b M(\overline{y}, \overline{y}', x) \, \mathrm{d}x = \text{ constant}$$

but we have a generic y:

$$B(y) = B(\overline{y} + \delta y) = \text{ constant}$$

this means that:

$$\delta B = 0$$

this should be found by applying the previous theory:

$$\delta B = B(\overline{y} + \delta y) - B(\overline{y})$$

again, by taking the previous theory, it is straightforward to obtain:

$$\delta B = \int_a^b \left[\frac{\partial M}{\partial y} \delta y + \frac{\partial M}{\partial y'} \frac{\mathrm{d}(\delta y)}{\mathrm{d}x} \right] \mathrm{d}x =$$

$$= \int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \delta y \, \mathrm{d}x = 0$$

Now, we can substitute δy to this, and obtain:

$$\delta B = \varepsilon_1 \int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_1(x) \mathrm{d}x + \varepsilon_2 \int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_2(x) \mathrm{d}x = 0$$

Now, it is necessary to exploit the degree of freedom that we introduced: ε_2; this means that now we have to calculate ε_2 in order to satisfy this constraint; this can be easily done, isolating it:

$$\varepsilon_2 = -\varepsilon_1 \frac{\int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_1(x)\mathrm{d}x}{\int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_2(x)\mathrm{d}x}$$

If we use the ε_2 defined here, the constraint will be kept into account.

This is the starting point; now, we have to apply this result on the functional that we are going to minimize. So, from the theory:

$$\delta A = \int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \delta y\, \mathrm{d}x = 0$$

the difference with respect to the previous case is that δy is not completely arbitrary: now, δy has to keep into account the constraint on the functional M. This can be done as follows:

$$\delta A = \varepsilon_1 \int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \eta_1(x)\mathrm{d}x + \varepsilon_2 \int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \eta_2(x)\mathrm{d}x = 0$$

but we have proved that, in order to keep into account the constraint, ε_2 has a particular expression, which is going to be substituted here now:

$$\varepsilon_1 \int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \eta_1(x)\mathrm{d}x - \varepsilon_1 \frac{\int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_1(x)\mathrm{d}x}{\int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_2(x)\mathrm{d}x} \int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \eta_2(x)\mathrm{d}x = 0$$

therefore, the final expression will not depend on ε_1 (which is arbitrary):

$$\frac{\int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \eta_1(x)\mathrm{d}x}{\int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_1(x)\mathrm{d}x} = \frac{\int_a^b \left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial y'} \right] \eta_2(x)\mathrm{d}x}{\int_a^b \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right] \eta_2(x)\mathrm{d}x}$$

This equation is satisfied only if the ratio between $\left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial y'} \right]$ and $\left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right]$ equals a constant λ:

$$\left[\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial y'} \right] = \lambda \left[\frac{\partial M}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial M}{\partial y'} \right]$$

exploiting linearity, this condition can be re-written as:

$$\frac{\partial}{\partial y}(L - \lambda M) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial}{\partial y'}(L - \lambda M)\right]$$

This equation is an Euler-Lagrange equation, if we consider a lagrangian $\tilde{L}$ defined as:

$$\tilde{L} = L - \lambda M$$

This is the origin of the method of the **Lagrange multipliers**: starting from a lagrangian L which has to satisfy some constraint defined on a function M, it is possible to define another lagrangian, $\tilde{L}$, which do not have to satisfy any constraint. λ is an additional degree of freedom: since two degrees of freedom come from the solution of the Euler-Lagrange equation, which is a second-order ODE, it is necessary a third degree of freedom, to comply also with the integral constraint on M. This is λ.

Example

Now, we will propose a practical example of application of this method. Given the functional to be minimized:

$$\int_a^b \sqrt{1 + (y')^2}\mathrm{d}x = \text{ minimum}$$

with the integral constraint:

$$\int_a^b y\,\mathrm{d}x = S$$

where S is a constant, we want to solve this problem. First of all: $M(y, y', x) = y$, and $L(y, y', x') = \sqrt{1 + (y')^2}$; therefore, it is possible to write the non-constrained lagrangian as:

$$\tilde{L} = \sqrt{1 + (y')^2} - \lambda y$$

then, it is possible to write the Euler-Lagrange equation for this lagrangian:

$$\frac{\partial \tilde{L}}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}$$

therefore: the Euler-Lagrange equation for this problem is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{y'}{\sqrt{1+(y')^2}}\right] = -\lambda$$

this can be solved directly, integrating both members with respect to x:

$$\frac{y'}{\sqrt{1+(y')^2}} = \mu - \lambda x$$

where λ is the Lagrange multiplier, and μ is a constant which comes from the integration. If we elevate to square both members, we obtain:

$$\frac{(y')^2}{1+(y')^2} = (\mu - \lambda x)^2$$

so:

$$(y')^2 = (1 + (y')^2)(\mu - \lambda x)^2$$

so:

$$(y')^2\left[1 - (\mu - \lambda x)^2\right] = (\mu - \lambda x)^2$$

so:

$$(y')^2 = \frac{(\mu - \lambda x)^2}{1 - (\mu - \lambda x)^2}$$

so:

$$y'(x) = \pm\frac{\mu - \lambda x}{\sqrt{1 - (\mu - \lambda x)^2}}$$

Therefore, finally:

$$y(x) = \pm\int \frac{\mu - \lambda x}{\sqrt{1 - (\mu - \lambda x)^2}}\,\mathrm{d}x = K\sqrt{1 - (y - \lambda x)^2}$$

where K is some constant. K, λ, μ should be determined in order to satisfy the boundary conditions **and** the integral constraint of the problem: three conditions, three degrees of freedom!

This is the equation of a circle: this says, anyway, that the circle is the shortest curve which satisfies all the constraints (in fact, this comes from

Euler-Lagrange, which is what guarantees that the functional is minimum; since L was defined as the functional for the determination of the length of the curve $\sqrt{1 + (y')^2}$, this is the research of a minimum length).

2.6.1 Extension to the case of multiple integral constraints

We introduced the method of Lagrange multipliers for one integral constraint, so to solve a problem like:

$$
\begin{cases}
\displaystyle\int_a^b L(y, y', x)\mathrm{d}x = \text{ minimum} \\[2mm]
\displaystyle\int_a^b M(y, y', x)\mathrm{d}x = \text{ constant}
\end{cases}
$$

What happens if there are multiple constraints? For instance, if the problem is:

$$
\begin{cases}
\displaystyle\int_a^b L(y, y', x)\mathrm{d}x = \text{ minimum} \\[2mm]
\displaystyle\int_a^b M(y, y', x)\mathrm{d}x = \text{ constant} \\[2mm]
\displaystyle\int_a^b N(y, y', x)\mathrm{d}x = \text{ constant}
\end{cases}
$$

The extension is very easy: it is straightforward to prove that the lagrangian $\tilde{L}$ has to be satisfied as:

$$
\tilde{L} = L - \lambda M - \mu N
$$

so, we have two extra degrees of freedom: λ and μ. In total, there are four constants: two from the ODE, and two Lagrange multipliers, which can be used to satisfy the four constraints ($(y(a)$, $y(b)$, the one on M and the one on N).

Example of problem

We are going to introduce an example, related to Quantum Mechanics. Given $\Psi(x)$ a wave function satisfying the following problem:

$$\begin{cases} \dfrac{\int_{-\infty}^{+\infty} x^2 \Psi^2(x)\,\mathrm{d}x}{\int_{-\infty}^{+\infty} \Psi^2(x)\,\mathrm{d}x} = \sigma^2 \\[4mm] \dfrac{\int_{-\infty}^{+\infty} \left(\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x}\right)^2 \mathrm{d}x}{\int_{-\infty}^{+\infty} \Psi^2(x)\,\mathrm{d}x} = \ \text{minimum} \end{cases}$$

we want to find Ψ. The first equation is a constraint, while the second one is the one which should be minimized. Ψ, which is the solution of the problem, is the wave function which has a certain minimum momentum, with a given position. We can notice that if Ψ is a solution, also $C\Psi$, with C constant, is a solution; therefore, it is possible to choose one between infinite solutions, where each solution is equal to another, multiplied by a certain constant. Therefore, let's introduce the following normalization:

$$\int_{-\infty}^{+\infty} \Psi^2(x)\,\mathrm{d}x = 1$$

this is one example of condition which can be enforced in order to give a certain value to the constant C. Keeping into account this condition, our problem becomes:

$$\begin{cases} \displaystyle\int_{-\infty}^{+\infty} x^2 \Psi^2(x)\,\mathrm{d}x = \sigma^2 \\[4mm] \displaystyle\int_{-\infty}^{+\infty} \left(\frac{\mathrm{d}\Psi(x)}{\mathrm{d}x}\right)^2 \mathrm{d}x = \ \text{minimum} \\[4mm] \displaystyle\int_{-\infty}^{+\infty} \Psi^2(x)\,\mathrm{d}x = 1 \\[2mm] \displaystyle\lim_{x\to\pm\infty} \Psi(x) = 0 \end{cases}$$

This is a problem where two integral constraints are defined, and one function has to be minimized. Then, a boundary condition for the problem has been inserted. Now, let us apply the method of Lagrange multipliers: the lagrangian of the problem is

$$L = \left(\frac{\mathrm{d}\Psi}{\mathrm{d}x}\right)^2 - \lambda\Psi^2 - \mu x^2 \Psi^2$$

50

therefore, starting from this it is possible to write the Euler-Lagrange equation:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial L}{\partial \Psi'}\right] = \frac{\partial L}{\partial \Psi}$$

so:

$$\frac{\partial L}{\partial \Psi'} = 2\Psi'$$

and

$$\frac{\partial L}{\partial \Psi} = -2\lambda\Psi - 2\mu x^2 \Psi$$

so:

$$\frac{\mathrm{d}^2 \Psi}{\mathrm{d}x^2} = -(\lambda + \mu x^2)\Psi$$

This equation has many solutions; a possible solution, which satisfies the boundary condition of the problem and which is a stationary minimum of the lagrangian integral, is:

$$\Psi(x) = Ae^{-kx^2}$$

the Gaussian function. As we are going to prove, A and k are two constants, just line λ and μ; actually, they are related to them. We have:

$$\Psi'(x) = -2kxAe^{-kx^2}$$

and:

$$\Psi''(x) = -(2k - 4k^2x^2)\Psi(x)$$

therefore:

$$\lambda = 2k \qquad \mu = -4k^2$$

It is possible to observe that this equation is close to the quantum harmonic oscillator:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\Psi}{\mathrm{d}x^2} + \frac{1}{2}k^2\Psi = E\Psi$$

where $F = -kx$: here, k is just the elastic constant of the spring.

2.6.2 Extension to a 3d problem

We are going to propose the solution of a 3d problem by means of this formulation. We are going to minimize the following lagrangian:

$$\int_{\mathcal{D}} = L(\Phi, \nabla\Phi, \vec{x})(\mathrm{d}\vec{x}) = \text{ minimum}$$

where $\mathcal{D}$ is some domain. The constraint, in this case, is that:

$$\Phi(\vec{x}) = f(\vec{x}), \forall \vec{x} \in \partial\mathcal{D}$$

where $f(\vec{x})$ is a known function. This means that, in this case, we enforce the values of the solution $\Phi(x)$ on a certain boundary $\mathcal{D}$, to equal this function. As usual, we have that:

$$\Phi = \overline{\Phi} + \delta\Phi$$

therefore, applying the usual theory, it is possible to obtain:

$$\delta \left(\int_{\mathcal{D}} L(\Phi, \nabla\Phi, \vec{x})(\mathrm{d}\vec{x}) \right) = 0$$

which is:

$$\int_{\mathcal{D}} \left\{ \frac{\partial L}{\partial \Phi}\delta\Phi + \sum_{i=1}^{3} \frac{\partial L}{\partial \left(\frac{\partial \Phi}{\partial x_i} \right)} \left[\frac{\partial}{\partial x_i}(\delta\Phi) \right] \right\} \mathrm{d}\vec{x}$$

this is simply obtained considering, as usual, the derivative with respect to $\vec{x}$ as the derivative with respect to each i-th component of $\vec{x}$ or of Φ. In the second term, the one inside the sum, L is differentiated with respect to the sum of the derivatives of Φ with respect to each component. First of all, let us define:

$$\frac{\partial L}{\partial \left(\frac{\partial \Phi}{\partial x_i} \right)} = p_i$$

therefore, the second term becomes:

$$\vec{p} \cdot \nabla(\delta\Phi)$$

so, the integral, is:

$$\int_{\mathcal{D}} \left\{ \frac{\partial L}{\partial \Phi} \delta \Phi + \vec{p} \cdot \nabla(\delta \Phi) \right\} d\vec{x}$$

Now we can apply some kind of integration by parts to the ∇ operator; this can be done by means of Gauss Theorem (Divergence Theorem), obtaining:

$$\int_{\mathcal{D}} \left\{ \frac{\partial L}{\partial \Phi} \delta \Phi + \vec{p} \cdot \nabla(\delta \Phi) \right\} d\vec{x} = \int_{\mathcal{D}} \left(\frac{\partial L}{\partial \Phi} - \nabla \cdot \vec{p} \right) \delta \Phi d\vec{x} + \int_{\partial \mathcal{D}} \vec{p} \cdot \hat{n} \delta \Phi \, d\sigma$$

but we enforce (from the theory, see previous sections) $\Delta \Phi$ to be zero on the boundary, therefore we have:

$$\int_{\mathcal{D}} \left(\frac{\partial L}{\partial \Phi} - \nabla \cdot \vec{p} \right) \delta \Phi \, d\vec{x} = 0$$

which leads to:

$$\frac{\partial L}{\partial \Phi} = \nabla \cdot \vec{p}$$

this is the generalized Euler-Lagrange equation; this can be written, component by component, as:

$$\sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left[\frac{\partial L}{\partial \left(\frac{\partial \Phi}{\partial x_i} \right)} \right] = \frac{\partial L}{\partial \Phi}$$

Example

Let us consider:

$$L(\phi, \nabla \phi, \vec{x}) = \frac{1}{2} \alpha(\vec{x}) \left[\nabla \phi \right]^2 + \frac{1}{2} \beta(\vec{x}) \phi^2 - \gamma(\vec{x}) = 0$$

this is a quadratic function in the unknown; α, β, γ are known coefficients; we can use the previously-introduced theory, and find:

$$\frac{\partial L}{\partial \phi} = \beta \phi - \gamma$$

moreover:

$$[\nabla\phi]^2 = \left[\frac{\partial\phi}{\partial x}\right]^2 + \left[\frac{\partial\phi}{\partial y}\right]^2 + \left[\frac{\partial\phi}{\partial z}\right]^2$$

so:

$$\frac{\partial L}{\partial\left(\frac{\partial\phi}{\partial x}\right)} = \alpha\frac{\partial}{\partial x} \quad \frac{\partial L}{\partial\left(\frac{\partial\phi}{\partial y}\right)} = \alpha\frac{\partial}{\partial y} \quad \frac{\partial L}{\partial\left(\frac{\partial\phi}{\partial z}\right)} = \alpha\frac{\partial}{\partial z}$$

so:

$$\frac{\partial}{\partial x}\left(\alpha\frac{\partial\phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\alpha\frac{\partial\phi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\alpha\frac{\partial\phi}{\partial z}\right) = \beta\phi - \gamma$$

which can be resumed as:

$$\nabla\cdot(\alpha\nabla\phi) - \beta\phi + \gamma = 0$$

this is the **diffusion equation for neutrons**, where the solution, ϕ, is called **neutron flux**. In this context, β is the absorption function of neutrons, γ is the external neutrons source, α is the diffusion coefficient. Then, this equation can be solved by means of some finite differences scheme, or in some other numerical way.