Theoretical Physics: A Primer for Philosophers of Science

Francisco Antonio Doria
Advanced Studies Group, PEP–COPPE, UFRJ

Abstract. We give a overview of the main areas in theoretical physics, with emphasis on their relation to Lagrangian formalism in classical mechanics. This review covers classical mechanics; the road from classical mechanics to Schrödinger's quantum mechanics; electromagnetism, special and general relativity, and (very briefly) gauge field theory and the Higgs mechanism. We shun mathematical rigor in favor of a straightforward presentation.

Keywords: Theoretical physics, lagrangian, classical mechanics, quantum mechanics, electromagnetism, relativity, Higgs mechanism.

1. Introduction

The present text is intended as a very short primer on theoretical physics for philosophers of science. We give here a necessarily brief albeit self–contained overview of theories such as classical mechanics, electromagnetism, quantum mechanics at the first–quantization level and its relation to classical mechanics, plus special and general relativity and an addendum on gauge field theory.

It is stressed throughout the text that classical mechanics in its Lagrangian formulation is the formal backbone of theoretical physics, where we use the Lagrangian formalism for field theories. Physicists use to quote the slogan:

> *To formulate a theory is to build a Lagrangian for it.*

(We will soon elaborate on that.) So, classical mechanics is a kind of unifying tool for most of physics (domains that do not fit into the picture are thermodynamics and statistical mechanics).

While we mention gauge theory in its relation to general relativity, as general relativity is a very singular example of gauge theory, we do not delve on more recent areas such as the bulk of quantum field theory, or string theory (we make a brief remark on the Higgs mechanism). A much wider survey is now being prepared with N. C. A. da Costa.

This text was the basis for two courses given at the IEA–USP in 2002 and 2003, and (in modified form) it was taught at the Philosophy of Science Program at COPPE–UFRJ. Prerequisites are: calculus, differential equations, linear algebra, some group

Principia **13**(2): 195–232 (2009).
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

theory. Plus that rather mysterious condition, "mathematical maturity," that is, the ability to read mathematics as if one were going through a musical score.

I must acknowledge my huge, constant debt to Newton da Costa, with whom I've been discussing topics in the philosophy of physics for a long time.

Sources

We will only give sources for classical mechanics, as we've followed the presentation of Goldstein 1950 and Leech 1965, with some input from the Lanczos treatise (Lanczos 1970). Presentation of electromagnetism and of general relativity is standard, as well as that of gauge theory, and there aren't many variations in the way those areas are introduced in the literature.

2. Classical mechanics

This is intended as a summary of the main approaches and formalisms in classical (or analytical) mechanics. There are actually three such viewpoints:

1. The *analytical or Lagrangian formalism*, where everything starts from the action integral, $S = \int L dt$ whose variation $\delta S = 0$ gives us Lagrange's equations.

2. The *canonical* or Hamiltonian formalism, an intermediate step (of a deep meaning of its own, since it originates symplectic geometry theory).

3. *Hamilton–Jacobi theory*, where we conclude that Hamilton's principal function S for material bodies is an (idealized?) wavefunction, and coincides with the action integral $S = \int L dt$ (modulo a constant), so that we are back to step 1 and close the circle.

Legendre transformations take the Lagrangian formalism onto the Hamiltonian formalism. Then, a particular kind of the so–called canonical transformations is used to reach the Hamilton–Jacobi equation which Cornelius Lanczos called the sacred burning bush of the theory of mechanics.

Motivation for Lagrange's equations

We start from Newton's Second Law, which is usually written for point–like objects as:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt},$$

where $\mathbf{p} = m\mathbf{v}$, m being the object's mass and $\mathbf{v}$ its velocity vector. If m is constant, we get:

$$\mathbf{F} = m\mathbf{a},$$

and that means: the force $\mathbf{F}$ that acts upon a point–like object of mass m imparts on it an acceleration $\mathbf{a}$. (We are using boldface characters for vector objects.)

The intuition[1] is:

> *if let by itself, a body will either be at rest or in rectilinear, uniform motion. A force which is applied to it will only change its state of motion. Forces do not originate the motion; forces change the motion, that's the point.*

If we use components (vector indices appear here "upstairs"):

$$F^i = ma^i.$$

Or,

$$m\frac{d^2x^i}{dt^2} = F^i.$$

Example 2.1. Suppose that we are dealing with a one–dimensional situation (some phenomenon, say, with radial symmetry). Suppose moreover that the force is a gradient force, and decreases as the variable increases. The preceding equation becomes, in this specific case,

$$m\frac{d^2r}{dt^2} = -\frac{\partial V(r)}{\partial r}.$$

(Use of ∂ is somewhat abusive—but see what we get now.) For we have:

$$\frac{d}{dt}[m\frac{dr}{dt}] = -\frac{\partial V(r)}{\partial r},$$

and for $\dot{r} = dr/dt$,

$$\frac{d}{dt}[m\dot{r}] = -\frac{\partial V(r)}{\partial r}.$$

We immediately conclude:

$$\frac{d}{dt}\frac{\partial}{\partial \dot{r}}[\frac{1}{2}m(\dot{r})^2] = -\frac{\partial V(r)}{\partial r}.$$

Now recall the energy theorem:

$$\int_{x_0}^{x_1} \mathbf{F}.d\mathbf{x} = (1/2)mv_1^2 - (1/2)mv_0^2,$$

(v_i, $i = 0, 1$, is the length of vector $\mathbf{v}_i$.) It is easily proved:

$$Fdx = m\frac{dv}{dt}dx = mdv\frac{dx}{dt} = mvdv,$$

and integration gets our result. We note $T(\dot{r}) = \frac{1}{2}m(\dot{r})^2$; T is the *kinetic energy* of the object with mass m. Or,

$$\frac{d}{dt}\frac{\partial}{\partial \dot{r}}T(\dot{r}) = -\frac{\partial V(r)}{\partial r}.$$

Now, as we can define a function $L(r, \dot{r}) = T(\dot{r}) - V(r)$, we can write the above equation as:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0.$$

This is the blueprint for Lagrange's equations. □

We can argue in a similar vein for several coordinates, whereas we get *Lagrange's equations*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0.$$

L is the *Lagrangian function.*

Motivation for Hamilton's equations

Example 2.2. Let us be given Lagrange's equations,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i}.$$

Write down the total energy of the system as the sum of kinetic energy T and potential energy V, $H = T + V$. Here,

$$T = \frac{1}{2}m(\dot{q}^i)^2 = \frac{1}{2m}(p_i)^2$$

and

$$V = V(q^i).$$

(Recall that the momenta are defined as $p^i = m\dot{q}^i$). See that one gets:

$$p_i = \frac{\partial L}{\partial q^i}.$$

Of course we can write H as a function of the positions q^i and of the momenta p_i, and also a function of t, both explicitly and through the q_i and p_i. Then it is immediately seen that Newton's Second Law becomes:

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}.$$

There is still a second set of equations. We'll meet them below. In this specific case, they simply mean that $\dot{q}^i = p_i/m$:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}. \square$$

3. Variational calculus in a nutshell

We will now give some rigor (and a more precise meaning) to the δt symbolism we will use as our main tool in the Lagrangian formulation.

Variations

Consider the integral

$$J = \int_{x_0}^{x_1} f(y, y', x) dx,$$

where $y(x)$ is seen as a curve. Consider the integral as a functional of the possible curves y;.we ask, for which curve y is the integral a maximum or a minimum?

The idea is to examine a family of curves $y(x, \alpha)$, where α is a parameter, and to compute $\partial J/\partial \alpha = 0$. We impose that:

1. $y(x, 0) = y(x)$, for $y(x)$ the desired solution.
2. $y(x_0, \alpha) = y(x_0)$.
3. $y(x_1, \alpha) = y(x_1)$.

That is, $\alpha = 0$ gives us our extremal value, and there is no variation at the integral's endpoints. Let's compute $\partial J/\partial \alpha = 0$:

$$\frac{\partial J}{\partial \alpha} = \int_{x_0}^{x_1} \{\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\} dx = 0.$$

Now the integrand equals,

$$\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial^2 y}{\partial x \partial \alpha}.$$

For the second term,

$$\int_{x_0}^{x_1} \frac{\partial f}{\partial y'} \frac{\partial^2 y}{\partial x \partial \alpha} dx = \frac{\partial f}{\partial y'} \frac{\partial y}{\partial \alpha}|_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx}(\frac{\partial f}{\partial y'})\frac{\partial y}{\partial \alpha} dx.$$

We can multiply $\partial J/\partial \alpha$ throughout by $d\alpha$. Notice that at the extremes the variation of the curve is 0. Then we get, while defining the operator δ:

$$\delta J =_{\text{Def}} \frac{\partial J}{\partial \alpha} d\alpha = \int_{x_0}^{x_1} \{\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\}\frac{\partial y}{\partial \alpha} d\alpha dx.$$

If we equate it to 0, as $d\alpha$ is arbitrary we get:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0.$$

Those are the *Euler*, or *Euler–Lagrange variational equations.*

Example 3.1. Which is the minimum length curve? Put:

$$J = \int_1^2 ds = \int_1^2 [1 + (\frac{dy}{dx})^2]^{1/2} dx.$$

The integrand doesn't depend on y, just on y'. Then we have $\partial f/\partial y = 0$, and

$$\frac{\partial f}{\partial y'} = \frac{y'}{[1 + (y')^2]}.$$

Therefore,

$$\frac{d}{dx}[\frac{y'}{[1 + (y')^2]}] = 0,$$

or

$$\frac{y'}{[1 + (y')^2]} = A,$$

A a constant. It is immediate to see that $y' = C$, a constant, or that $y = Cx + C'$, a straight line. □

4. The Lagrangian formulation

Physicists like to quote the slogan:

> *To formulate a theory is to write down a Lagrangian for it.*

Let's see why. Let us be given a collection of N material points of mass m_P each, and let x^{1P}, x^{2P}, x^{3P}, $1 \leq P \leq N$, be the corresponding coordinates of the position vectors in real Euclidean $3N$ space. Let moreover $\dot{x}^{1P}, \dot{x}^{2P}, \dot{x}^{3P}$ be the corresponding coordinates of the velocity vectors, that is,

$$\dot{\mathbf{x}}^P = \frac{d\mathbf{x}^P}{dt}.$$

Form (locally at least) the product space spanned by the coordinates $\langle q^i, \dot{q}^i \rangle$, where the q^i range over the position coordinates and the $\dot{q}^i$ over the velocities. Since we are dealing with time–varying objects we moreover suppose:

- The functions $q^i = q^i(t)$, $\dot{q}^i = \dot{q}^i(t)$.
- Those functions are of class C^r, $r > 1$.

We define the (general) *Lagrangian function*

$$L = L(\ldots q^i \ldots, \ldots \dot{q}^i \ldots, t) = L(q^i, \dot{q}^i, t).$$

The Lagrange function is usually constructed as $L = T - V$, the 'spare' or net amount of energy which is available for the motion of the system. Hamilton's variational principle (the variation of the action integral) can be understood as the following:

> *The system's motion is such that the net energy $L = T - V$ is an extremum (minimum or maximum).*

Lagrange's equations out of a variational principle

Notice that δ can be seen as a differential operator (differential operators satisfy the condition $\delta(fg) = (\delta f)g + f(\delta g)$, besides being linear, commuting with the integral sign and with other differential operators). Impose and compute:

$$\delta S = \delta \int_{t_0}^{t_1} L dt = 0. \tag{1}$$

We also have that $\delta t|_{t_0} = \delta t|_{t_1} = \delta t = \delta q^i|_{t_0} = \delta q^i|_{t_1} = 0$. From the property of commutation with $\int$ and with the other derivative operators, we get:

$$\delta S = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i\right) dt = 0. \tag{2}$$

Now,

$$\delta\dot{q}^i = \delta(\frac{dq^i}{dt}) = \frac{d\delta q^i}{dt}.$$

Given that, compute (out of integration by parts):

$$\int_{t_0}^{t_1} \frac{\partial L}{\partial q^i}\frac{d\delta q^i}{dt}dt = [\frac{\partial L}{\partial q^i}\delta q^i]_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i})\delta q^i dt.$$

The term without an integral sign vanishes because of the conditions on δq^i at the integration limits. If we go back to eq. (2),

$$\delta S = \int_{t_0}^{t_1} [\frac{\partial L}{\partial q^i}\delta q^i - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i})\delta q^i]dt = 0, \tag{3}$$

or

$$\delta S = \int_{t_0}^{t_1} [\frac{\partial L}{\partial q^i} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i})]\delta q^i dt = 0. \tag{4}$$

Due to the arbitrariness of the variations δq^i, this implies:

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i}) = 0. \tag{5}$$

Eqs. (5) are *Lagrange's equations* for the considered system.

Non–potential forces, constraints

The Lagrangian technique can be extended to include some quite general situations. Suppose that we add the following differential constraints, not necessarily integrable, to our system given by a Lagrangian L:

$$\sum_k a_{ik}dq^k + a_{it}dt = 0.$$

We have that $0 \leq i \leq m$, where $m \leq n$, n dimension of the configuration space (space of the q^k). The a_{ik} are functions of the q^k. Of course we can start from a system of equations,

$$f_i(\ldots, q^k, \ldots) = 0,$$

which act as geometrical constraints upon the system. We differentiate the f_i and get our previous set of equations; in that case the differential system of constraints is integrable.

We now show that geometry leads to physics, that is, the existence of those constraints can be described as a set of forces that act upon the system.

Lagrange multipliers

From our differential condition,

$$\sum_k a_{ik} dq^k + a_{it} dt = 0,$$

we get our *variational* constraints,

$$\sum_k a_{ik} \delta q^k = 0.$$

(Recall that $\delta t = 0$.) Let λ^i be (so far) undetermined constants, and form the linear combination,

$$\sum_{ik} \lambda^i a_{ik} \delta q^k = 0.$$

Get the system's variational principle:

$$\int_0^1 dt \sum_k (\frac{\partial L}{\partial q^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^k}) \delta q^k = 0.$$

Now as the δq^k are *not* independent due to the constraints. Then we may add the condition above,

$$\sum_{ik} \lambda^i a_{ik} \delta q^k = 0,$$

duly integrated, to the variational principle,

$$\int_0^1 dt \sum_k (\frac{\partial L}{\partial q^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^k} + \sum_i \lambda^i a_{ik}) \delta q^k = 0.$$

Now the situation we have is the following:

- There are n δq^ks.
- There are m relations that connect them.
- Therefore, we can independently choose $n - m$ variations δq^ks.
- Once made that choice, the remaining m δq^k are determined by the constraint equations.

We then impose the following $n - m$ conditions on the λ^i:

$$\frac{\partial L}{\partial q^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^k} + \sum_i \lambda^i a_{ik} = 0.$$

Now: the first term in the above equations gives us potential forces for the system; second term is the kinetic term. So the linear combination involving the λ^i can be understood as a system of non–potential forces, if the differentials aren't exact. (If they are, we may add them as extra potential forces.)

5. Hamilton's equations

Reread the motivation at the introduction, before going through the whole argument below.

Hamiltonian formulation

Define a function:

$$H(q^i, p_i, t) = \sum_i \dot{q}^i p_i - L(q^i, \dot{q}^i, t), \tag{6}$$

where we must put:

$$p_i = \frac{\partial L(q^i, \dot{q}^i, t)}{\partial \dot{q}^i}.$$

Each p_i is the *momentum component* associated to the velocity components q^i.

Remark 5.1. This is just a notational convention: objects that sort of behave like tangent vectors to a curve of coordinates $q^i(t)$ with components

$$A^i = \frac{dq^i}{dt}$$

have upper indices; those of the form $\partial X / \partial q^i$, also called gradient–like objects, have lower (or "downstairs") indices.

Objects with upper indices are called *contravariant* vectors; with lower indices, they are called *covariant* vectors. Those names come from the late 19th–century linear transformation theory on vector spaces. □

Suppose that the transformation given by eq. (6) is in fact a map that changes from coordinates (or variables) $\langle q^i, \dot{q}^i \rangle$ to the set $\langle q^i, p_i \rangle$. So, the left side depends on the q, p. Compute its differential:

$$dH = \sum_i [\frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i] + \frac{\partial H}{\partial t} dt. \tag{7}$$

Now compute the differential of the rhs:

$$dH = \sum_i [\dot{q}^i dp_i + p_i d\dot{q}^i - \frac{\partial L}{\partial \dot{q}^i} s\dot{q}^i - \frac{\partial L}{\partial q} dq^i] - \frac{\partial L}{\partial t} dt. \tag{8}$$

Notice that Lagrange's equations (eqs. (5) can be written as:

$$\dot{p}_i = \frac{\partial L}{\partial q^i}.$$

If we substitute and cancel terms in eqs. (8) (and take into account the definition for p_i and the value of $\dot{p}_i$ from Lagrange's equations), we get:

$$dH = \sum_i [\dot{q}^i dp_i - \dot{p}_i dq^i] - \frac{\partial L}{\partial t} dt. \tag{9}$$

Now compare to eq. (7):

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \tag{10}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{11}$$

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{12}$$

Eqs. (10) and (11) are *Hamilton's equations*.

Meaning of H

Example 5.2. Let $L = T - V$. T is taken to be quadratic homogeneous on the $\dot{q}^i$, and L doesn't explicitly depend on t. Compute in that case:

$$\frac{dL}{dt} = \sum_i [\frac{\partial L}{\partial q^i} \frac{dq^i}{dt} + \frac{\partial L}{\partial \dot{q}^i} \frac{d\dot{q}^i}{dt}].$$

From Lagrange's equations, eqs. (5),

$$\frac{dL}{dt} = \sum_i [\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\frac{dq^i}{dt} + \frac{\partial L}{\partial \dot{q}^i}\frac{d\dot{q}^i}{dt}],$$

or

$$\frac{dL}{dt} = \sum_i \frac{d}{dt}(\dot{q}^i \frac{\partial L}{\partial \dot{q}^i}).$$

Follows (also from the definition of p_i),

$$\frac{d}{dt}(L - \sum_i \dot{q}^i p_i) = 0.$$

Thus, in this specific case (L independent of t, and T homogeneous quadratic, a property to be used below),

$$L - \sum_i \dot{q}^i p_i = -H = \text{ constant.}$$

Now recall from Euler's theorem that if f is homogeneous on variable x of degree n,

$$\sum_i \frac{\partial f}{\partial x^i} x^i = nf.$$

From Lagrange's equations, as V is independent of $\dot{q}^i$,

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = \frac{\partial T}{\partial \dot{q}^i}.$$

Then, in this case:

$$\sum_i \dot{q}^i p_i = \sum \dot{q}^i \frac{\partial T}{\partial \dot{q}^i} = 2T.$$

Thus,

$$L - 2T = -H = T - V - 2T = -(T + V),$$

or

$$H = T + V.$$

That is, in this particular case, H is the sum of the kinetic energy and the potential energy. This means, the system's total energy—a constant here, as $\partial H/\partial t = 0$ (H doesn't explicitly depend on t), and:

$$\frac{dH}{dt} = \sum_i (\frac{\partial H}{\partial q^i}\dot{q}^i + \frac{\partial H}{\partial p_i}\dot{p}_i) + \frac{\partial H}{\partial t},$$

and, with the help of Hamilton's equations:

$$\frac{dH}{dt} = \sum_i (\frac{\partial H}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q^i}) + \frac{\partial H}{\partial t} = 0. \square$$

6. Hamilton–Jacobi theory

Canonical transformations

Supose that, from a domain spanned by the coordinates

$$(\ldots, q^i, \ldots; \ldots, p_i, \ldots),$$

we make a coordinate transformation spanned by the new coordinates

$$(\ldots, Q^i, \ldots; \ldots, P_i, \ldots),$$

so that, for a new Hamiltonian function K, we have the (transformed) Hamilton–like equations:

$$\dot{Q}^i = \frac{\partial K}{\partial P_i}, \tag{13}$$

$$\dot{P}_i = \frac{\partial K}{\partial Q^i}. \tag{14}$$

Thus, as the variational principle (where we abbreviate the limits, t_0 as 0, t_1 as 1),

$$\delta L = \delta \int_0^1 (\sum_j p_j \dot{q}^j - H(q,p,t))dt = 0,$$

must hold, so does the principle:

$$\delta L' = \delta \int_0^1 (\sum_j P_j \dot{Q}^j - K(Q,P,t))dt = 0.$$

Thus

$$L' = L + \frac{dF}{dt},$$

some F. For recall that, for any function F,

$$\delta \int_0^1 \frac{dF}{dt} dt = \delta(F_1 - F_0) = 0,$$

the integrands in the action integral must differ by a total time derivative dF/dt.

Consider a specific case: let $F_2 = F_2(q, P, t)$. (Reasons for that notation follow from traditional usage.) F should in principle depend on q, Q, p, P, but as there is a 1–1 map between the two sets of coordinates (q, p) and (Q, P), F can only depend on $2n$ variables. If we equate the integrands in the two action integrals, we get, after expanding dF_2/dt,

$$K = H + \frac{\partial F}{\partial t}.$$

The Hamilton–Jacobi equation

The Hamilton–Jacobi equation arises out of a singular Hamiltonian. If we ask for $K = 0$:

$$K = H + \frac{\partial F}{\partial t} = 0,$$

we trivialize everything, since the new Hamiltonian is constant and equal to zero:

$$H(q, p, t) + \frac{\partial S}{\partial t} = 0.$$

If $S = F_2$ (S is called Hamilton's principal function) we get that $p_i = \partial S/\partial q^i$. Then,

$$H(\ldots, q^i, \ldots, \frac{\partial S}{\partial q^i}, \ldots, t) + \frac{\partial S}{\partial t} = 0,$$

is the *Hamilton–Jacobi equation.* Its solution trivializes as desired the transformed Hamiltonian. (This is a particular albeit very important example of a very general theorem in dynamical systems theory: locally, that is, in a convenient open domain were there are no singularities, every vector field can be "rectified.")

7. Back to where the action is

Let $S = S(q^i, p_i, t)$. Since:

$$\frac{dS}{dt} = \sum_i \frac{\partial S}{\partial q^i}\dot{q}^i + \sum_i \frac{\partial S}{\partial p_i}\dot{p}_i + \frac{\partial S}{\partial t},$$

as the momenta are constant once we solve the Hamilton–Jacobi equation, and from

$$p_i = \frac{\partial S}{\partial q^i},$$

we get that

$$\frac{dS}{dt} = \sum_i p_i \dot{q}^i - H = L,$$

from the $K = 0$ condition in the Hamilton–Jacobi equation. Therefore,

$$S = \int L dt + C,$$

C an integration constant. We are back to our starting point; Hamilton's principal function S is in fact (but for a constant) the Action Integral.

This closes the circle. René Thom calls S the "entropy of mechanics," since that function tells us the direction of time evolution in a mechanical system.

Geodesics as the path of a system

Example 7.1. *A particular Lagrangian and its consequences.* Consider the Lagrangian:

$$L = \sum_{i,j} \frac{1}{2} a_{ij}(q)\dot{q}^i \dot{q}^j + \sum_i b_i(q) q^i - h(q). \tag{15}$$

Here $a_{ij}(q)$ means dependence on the position variables $\ldots q^i \ldots$. From:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \sum_{j,k} \frac{\partial a_{ij}}{\partial q^k} \dot{q}^k \dot{q}^j + \sum_j a_{ij}(q)\ddot{q}^j + \sum_k \frac{\partial b_i(q)}{\partial q^k} \dot{q}^k, \tag{16}$$

and from:

$$\frac{\partial L}{\partial q^k} = \sum_{i,j} \frac{\partial a_{ij}(q)}{\partial q^k} \dot{q}^i \dot{q}^j + \sum_i \frac{\partial b_i(q)}{\partial q^k} \dot{q}^i - \frac{\partial h(q)}{\partial q^k}, \tag{17}$$

we get, after some manipulations,

$$\sum_j a_{kj}\ddot{q}^j + \sum_{j,m} \left(\frac{\partial a_{kj}}{\partial q^m} + \frac{\partial a_{km}}{\partial q^j} - \frac{\partial a_{mj}}{\partial q^k}\right)\dot{q}^m \dot{q}^j = \sum_m \left(\frac{\partial b_m}{\partial q^k} - \frac{\partial b_k}{\partial q^m}\right)\dot{q}^m - \frac{\partial h}{\partial q^k}. \tag{18}$$

The l.h.s. of this equation is pretty well-known in general relativity: if we define the a^{ij} through $\sum_k a^{ik} a_{kj} = \delta^i_j$, where δ^i_j is the Kronecker delta, or simply the elements of the unit matrix written (δ^i_j) and if we abbreviate the parentheses with the derivatives of the a_{ij} as $[k, mj]$, and write $\{^k_{mn}\} = \sum_j a^{kj}[j, mn]$, we get:

$$\ddot{q}^j + \sum_{mk} \{^j_{mk}\} \dot{q}^m \dot{q}^k = F^j. \tag{19}$$

This is the equation of motion of a body along a geodesic in Riemannian space with metric tensor a_{ij} but for the external force of components F^j. This force is formally akin to the Lorentz force in electromagnetism, and is a *Coriolis–type force*.

The space whose metric tensor (see below, the section on tensor calculus) is a_{ij} and which is spanned by the coordinates $(\ldots, q^i, \ldots)$ is called *configuration space*. If the $F^j = 0$, the path of the system is a geodesic in configuration space with respect to the metric tensor a_{ij}. □

The harmonic oscillator; planetary motion

There are two classical examples of solvable (integrable) mechanical problems that can be easily settled with the analytical–canonical formalism: the harmonic oscillator (periodic motion) and planetary motion restricted to the two–body problem.

Example 7.2. *Harmonic oscillator.* The Lagrangian is:

$$L = (1/2)m\dot{x}^2 - (1/2)kx^2.$$

The total energy of the system is a constant, and equals the Hamiltonian,

$$E = H = (1/2)m\dot{x}^2 + (1/2)kx^2.$$

Motion equation is immediately derived:

$$m\ddot{x} + kx = 0.$$

General solution is an oscillator of the form $x = Ae^{iBt} + Ce^{-iDt}$, A, B, C, D constants, as can be easily checked. □

Example 7.3. *Two–body central motion.* We can reduce the two–body, central–force motion to a single particle, central force problem, whose Lagrangian is given (in polar coordinates, due to the problem's symmetry), by:

$$L = [1/2]m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r).$$

One of the two equations of motion is:

$$\dot{p}_\theta = \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = \frac{d}{dt}[mr^2\dot{\theta}] = 0,$$

since the Lagrangian is independent of $\dot{r}$. The integral of motion is an angular momentum,

$$mr^2\dot{\theta} = l,$$

of absolute value l.

Areal velocity

Imagine a triangle of height r and of infinitesimal basis $rd\theta$. Its area is $(1/2)r^2d\theta$. Then $(1/2)r\dot{\theta}$ is the so–called *areal velocity*. Follows from the conservation of l that the areal velocity is constant, that is, we get *Kepler's Second Law* in a more general setting: *the radius vector sweeps out equal areas in equal times.*

The equation for r

The other Lagrangian equation, for r, is:

$$\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V}{\partial r} = 0.$$

An obvious substitution of the gradient of V by $-f(r)$ gives,

$$m\ddot{r} - mr\dot{\theta}^2 = f(r).$$

θ can be eliminated by the conservation law for angular momentum:

$$m\ddot{r} - \frac{l}{mr^3} = f(r).$$

One immediately sees that this last equation is equivalent to,

$$m\ddot{r} = -\frac{\partial}{\partial r}(V + [1/2]\frac{l^2}{mr^2}).$$

Multiplication by $\dot{r}$ gives,

$$m\ddot{r}\dot{r} = \frac{d}{dt}(\frac{1}{2}m\dot{r}^2) = \frac{dr}{dt}\frac{\partial}{\partial r}(V + \frac{1}{2}\frac{l^2}{mr^2}).$$

Or,

$$m\ddot{r}\dot{r} = \frac{d}{dt}(V + \frac{1}{2}\frac{l^2}{mr^2}).$$

Or even,

$$\frac{d}{dt}[\frac{1}{2}m\dot{r}^2 + V + \frac{1}{2}\frac{l^2}{mr^2}] = 0.$$

This is the Law of Energy Conservation. The last term gives the angular kinetic energy. Thus, if E is the system's energy,

$$E = [\frac{1}{2}m\dot{r}^2 + V + \frac{1}{2}\frac{l^2}{mr^2}].$$

If we solve for $\dot{r}$, we get:

$$\frac{dr}{dt} = [\frac{2}{m}(E - V - \frac{l^2}{2mr^2})]^{1/2},$$

which gives us the variables already separated, so that we can integrate for dt and $f(r)dr$. □

8. From classical mechanics to quantum mechanics

Quantum mechanics cozily fits within classical mechanics. This is how it is done: take a single particle in 3–space; its Hamilton–Jacobi equation is, from the general Hamilton–Jacobi equation:

$$H(q, \frac{\partial S}{\partial q}, t) + \frac{\partial S}{\partial t} = 0.$$

If we restrict our attention to "almost plane–wave" solutions $S(q,t) = W(q) - Et$, we get the "characteristic function" equation:

$$H(q, \frac{\partial W}{\partial q}) = E.$$

It is immediate to see that, for a single particle under the action of a potential V in 3–space, the Hamilton–Jacobi equation becomes:

$$(1/2m)\left\{\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 + \left(\frac{\partial W}{\partial z}\right)^2\right\} + V = E.$$

We can condense that as:

$$(\nabla W)^2 = 2m(E - V) = 2mT,$$

where momentum $\mathbf{p}$ is given by:

$$\mathbf{p} = \nabla W.$$

Formally, this is the same as the *eikonal equation* of geometrical optics, whose solutions are trajectories of light rays:

$$(\nabla L)^2 = n^2.$$

Wave motion associated to W

L is a wave front, and the light rays are orthogonal to it. We can look at the equation $(\nabla W)^2 = 2mT$ from the same perspective. In that case, when W moves to $W + dW$, because of the wave motion given by $S = W - Et$, we have, at the same time

$$dW = Edt,$$

and

$$dW = |\nabla W| ds,$$

as it is the increment along the gradient path. Therefore we get,

$$u = \frac{ds}{dt} = \frac{E}{|\nabla W|}.$$

From the value of the gradient, given by the Hamilton–Jacobi equation for the single particle,

$$u = \frac{E}{\sqrt{2m(E - V)}}.$$

Now, for the single particle, $T = E - V = (1/2m)p^2$. Thus,

$$u = \frac{E}{p} = \frac{E}{mv}.$$

(Pay attention to what we have shown: the *wavefront velocity* u for W is inversely proportional to the *particle velocity* v.)

So, classical mechanics allows us to associate a wavefront (given by W) to a single particle system. The same argument is valid in a more general context, whenever we can associate the length element ds to kinetic energy, that is, whenever we have $ds^2 = 2Tdt^2$.

Now, if W corresponds to L, then $S = W - Et$ must be a *phase* in a wave equation, and must correspond to

$$k_0(L - ct) = 2\pi\left(\frac{L}{k_0} - \nu t\right).$$

When we move from the W–eikonal equation to the L–eikonal equation, we see that W and L must be proportional; we cannot simply equal them. Therefore we cannot equate both phases, but have to relate them through a proportionality constant. Call it h. Then we get,

$$E = h\nu,$$

which is *Planck's relation.*

Schrödinger's time–independent equation

The eikonal equation $(\nabla L)^2 = n^2$ is derived from the full wave equation,

$$\nabla^2\phi - \frac{1}{u^2}\frac{\partial^2\phi}{\partial t^2} = 0,$$

for the solution

$$\phi = \exp(A(\mathbf{r}) + ik_0(L(\mathbf{r} - ct)),$$

with $\mathbf{r}$ the position vector and the extra assumption that n varies very slowly with distance.

Let's go backwards and obtain the wave–equation associated to W, that is, to the ψ out of which we obtained W as the object that corresponds to the eikonal L.

Eliminate the time from the wave equation, that is, place time as an independent factor $e^{-i\omega t}$. From the relation

$$-\frac{\omega^2}{u^2} = \frac{4\pi^2}{\lambda^2},$$

we get the time–independent wave–equation:

$$\nabla^2\psi + \frac{4\pi^2}{\lambda^2}\psi = 0.$$

But $\lambda = h/p$, and $p = \sqrt{2m(E - V)}$. Then,

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0.$$

That is: *Schrödinger's time–independent equation.*

9. Lagrangian densities, field theory

Suppose that we define a Lagrangian as $L = \int_V \mathscr{L}(\psi, \psi_\mu)d^k x$, where the μ range over $0, 1, 2, \ldots, k-1$, and $\psi_\mu = \partial\psi/\partial x^\mu$; we abbreviate $\partial_\mu = \partial/\partial x^\mu$. $\mathscr{L}$ is a *Lagrangian density*, since we integrate it over a hypervolume V to obtain the full Lagrangian L; the $\psi, \ldots$ are the *field variables*. If we impose that $\delta L = 0$ we get:

$$\delta\mathscr{L} = \frac{\partial\mathscr{L}}{\partial\psi}\delta\psi + \sum_\mu \frac{\partial L}{\partial\psi_\mu}\delta\psi_\mu = 0. \tag{20}$$

Now $\delta\psi = \delta\partial_\mu\psi = \partial(\delta\psi)$, where we suppose commutativity of the differential operators δ and ∂_μ. Now,

$$\partial_\mu[\frac{\partial\mathscr{L}}{\partial\psi_\mu}\delta\psi] = \sum_\mu(\partial_\mu\frac{\partial\mathscr{L}}{\partial\psi_\mu})\delta\psi + \sum_\mu\frac{\partial L}{\partial\psi_\mu}\partial_\mu\delta\psi. \tag{21}$$

The l.h.s. is a divergence; if we impose that it be 0 — a conservation condition — we get, from eqs. (21) and (20),

$$\delta\mathscr{L} = 0 = \{\frac{\partial\mathscr{L}}{\partial\psi} - (\sum_\mu\partial_\mu\frac{\partial\mathscr{L}}{\partial_\mu})\}\delta\psi. \tag{22}$$

For an arbitrary variation $\delta\psi$, follow the *Lagrangian field equations for* $\mathscr{L}$:

$$\frac{\partial\mathscr{L}}{\partial\psi} - \sum_\mu\partial_\mu\frac{\partial\mathscr{L}}{\partial\psi_\mu} = 0. \tag{23}$$

We are going to use those Lagrangian–density field equations to obtain the electromagnetic field equations and the Einstein gravitational equations.

10. The electromagnetic field

We restrict our attention to the so–called vacuum fields. We start from another sacred spot, *Maxwell's equations* :

1. *Gauss' Law.* It shows how the electric field "emanates" from a charge distribution. It is:

   $$\nabla\cdot\mathbf{E} = 4\pi\rho.$$

   3–vectors are given in boldface. **E** is the electric field; ρ is the charge density. Coulomb's Law is one of the solutions of the preceding equation.

   (More elaborately, we know that the *electrostatic field* can be derived from a *scalar potential* Φ, that is, $\mathbf{E} = -\nabla\Phi$.)

2. *Ampère's Law.* It gives us the magnetic field generated by the flow of an electric current through a wire:

   $$\nabla\times\mathbf{B} = (4\pi/c)\mathbf{J} + (1/c)\frac{\partial\mathbf{E}}{\partial t}.$$

   B is the magnetic field, and **J** is the current density. The last term is Maxwell's *displacement current.*

3. *Nonexistence of magnetic monopoles.* This is:

$$\nabla \cdot \mathbf{B} = 0.$$

(This implies that there is a vector $\mathbf{A}$ so that, in an adequate domain, $\mathbf{B} = \nabla \times \mathbf{A}$.)

4. *Faraday's Law,* or how to generate an electric field out of a time–varying magnetic field:

$$\nabla \times \mathbf{E} + (1/c)\frac{\partial \mathbf{B}}{\partial t} = 0.$$

5. *Charge conservation.* Also called, the *continuity equation* (this is essentially a balance equation):

$$\nabla \cdot \mathbf{J} + (1/c)\frac{\partial \rho}{\partial t} = 0.$$

Covariant or 4–dimensional formulation, first version

Consider the 4–dimensional space spanned by the following coordinate system, $(\mathbf{x}, ict)$.

We represent a 4–coordinate by Greek indices, e.g., x_μ. Whenever a Latin index appears, e.g. x_k, it is supposed to range over the space coordinates, that is, $k = 1, 2, 3$. We won't bother for a while with "upstairs" and "downstairs" indices.

Einstein's summation convention

We agree that whenever a Greek index appears twice in an expression, it is understood that the index is to be summed from 0 to 3. That is,

$$a_\mu b_\mu = \sum_{\mu=0}^{\mu=3} a_\mu b_\mu.$$

Covariant current, covariant potential

Define the *covariant current* J as follows:

$$J = (J_\mu) = (ic\rho, \mathbf{J}).$$

Then the continuity equation becomes,

$$\partial_\mu J_\mu = 0.$$

(We have $\partial_\mu = \partial/\partial x_\mu$.) Now for the potentials: we put,

$$A = (A_\mu) = (i\Phi, \mathbf{A})$$

(We'll soon see the usefulness of that formulation; name "covariant" will be explained below.)

The electromagnetic field $F_{\mu\nu}$

Form the skew–symmetrical array:

$$F = (F_{\mu\nu}) = \begin{pmatrix} 0 & B_3 & -B_2 & -iE_1 \\ -B_3 & 0 & B_1 & -iE_2 \\ B_2 & -B_1 & 0 & -iE_3 \\ iE_1 & iE_2 & iE_3 & 0 \end{pmatrix}$$

Then Maxwell's equations can be thus written:

1. *Inhomogeneous set*, or Gauss' Law and Ampère's Law:

   (24) $$\partial_\mu F_{\mu\nu} = (4\pi/c)J_\nu.$$

2. *Homogeneous set*, or Faraday's Law and the nonexistence of magnetic monopoles:

   (25) $$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0.$$

   This is the so–called *circular*, or *Bianchi linear condition.* As it will be seen, it implies that, locally at least, there is a potential $A = (A_\mu)$ such that:

   (26) $$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

The Bianchi conditions are the expression of a very deep mathematical fact, the "double differential" condition $\partial^2 = 0$, or, simply stated in a geometrical translation, "the boundary of a boundary is empty."

The *wave equation* is obtained when we substitute the definition of $F_{\mu\nu}$ in terms of the potentials A_μ into the inhomogeneous equation:

$$\partial_\mu[\partial_\mu A_\nu - \partial_\nu A_\mu] = (4\pi/c)J_\nu,$$

or

$$\partial_\mu \partial_\mu A_\nu - \partial_\nu(\partial_\mu A_\mu) = (4\pi/c)J_\nu.$$

We now impose the so–called *gauge condition* $\partial_\mu A_\mu = 0$. This is no constraining condition on the potentials A_μ, for notice that, if we write:

$$A'_\mu = A_\mu + \partial_\mu \phi, \tag{27}$$

for any adequately differentiable ϕ, it is immediately seen that this transformed A'_μ gives the same field $F_{\mu\nu}$ as A_μ. So we have some freedom here.

The transformation $A_\mu \mapsto A'_\mu = A_\mu + \partial_\mu \phi$ is called a *gauge transformation*, and the value of A_μ given by $\partial_\mu A_\mu = 0$ is the *Lorentz gauge*. Given condition $\partial_\mu A_\mu = 0$ we get the (inhomogeneous) *wave equation*:

$$\partial_\mu \partial_\mu A_\nu = (4\pi/c) J_\nu,$$

and for $\Box = \partial_\mu \partial_\mu$,

$$\Box A_\nu = (4\pi/c) J_\nu.$$

For $J_\nu = 0$, all ν, this reduces to the well–known homogeneous wave equations for the scalar and vector potentials:

$$\nabla^2 \Phi - (1/c^2)\frac{\partial \Phi}{\partial t} = 0,$$

$$\nabla^2 \mathbf{A} - (1/c^2)\frac{\partial \mathbf{A}}{\partial t} = 0.$$

Dirac–like formulation

It is known that Maxwell's equations are equivalent to the Dirac–like set

$$\nabla \varphi = \iota,$$

where

$$\varphi = (1/2) F_{\mu\nu} \gamma^{\mu\nu},$$

and

$$\iota = j_\mu \gamma^\mu,$$

$$\nabla = \gamma^\rho \partial_\rho,$$

(where the $\{\gamma^\mu : \mu = 0, 1, 2, 3\}$ are the Dirac gamma matrices with respect to η). They satisfy:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu}$$

and $\eta = \mathrm{diag}(-1, +1, +1, +1)$ is a diagonal matrix that represents the metric tensor over Minkowski space.

Those equation systems are to be considered together with boundary conditions that specify a particular field tensor $F_{\mu\nu}$ "out of" the source j^ν.

A Lagrangian density for the empty–space Maxwell equations

The sourceless Maxwell equations can be derived through the Lagrangian field equations, eq. (23) from the Lagrangian density:

$$\mathscr{L} = \frac{1}{2}F_{\mu\nu}F_{\mu\nu} - F_{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu). \tag{28}$$

11. From electromagnetism to special relativity

In a nutshell: special relativity is a theory that aims to make compatible particle mechanics and classical electromagnetic theory. The basic criterion is: we keep electromagnetic theory as it is, and make the required changes in particle mechanics.

The main postulate in the theory of special relativity is: the speed of light c is constant in the vacuum, independent of the (linear) reference system. More precisely, if observer O is at rest with respect to reference system R_O, and observer O' is at rest with respect to reference system $R_{O'}$, and $R_{O'}$ is uniform motion with respect to R_O, then both O and O' will see the same light ray approaching them with speed c. This means:

$$\Box A_\mu = 0$$

holds for R_O, while

$$\Box A'_\mu = 0$$

holds for $R_{O'}$, where A_μ and A'_μ are measured with respect to the corresponding reference systems.

The requirement that the vacuum speed of light c be a constant follows from electromagnetic theory, as it can be easily shown that c onle depends on the constants ϵ_0, μ_0, that is, the vacuum electric permissitiy and magnetic permeability.

Given that requirement, the linear transformations that relate R_O and $R_{O'}$ are those that keep invariant the *metric form*

$$ds^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 - c^2(dt^2).$$

If $(\eta_{\mu\nu})$ denotes the diagonal matrix with diagonal $(-1,+1,+1,+1)$, then, for $dx_0 = cdt$, we can write:

$$ds^2 = \eta_{\mu\nu}dx_\mu dx_\nu. \tag{29}$$

(We call the *signature* of η the algebraic sum of the elements in the diagonal $(-1,+1,+1,+1)$; here η has signature $+2$.)

The *Lorentz transformations* are the homogeneous linear transformations $(L_{\mu\rho})$ that satisfy:

$$\eta_{\rho\sigma} = L_{\mu\rho}L_{\nu\sigma}\eta_{\mu\nu}.$$

The well–known transformations:

$$x' = \frac{x - vt}{\sqrt{(1 - v^2/c^2)}}, \quad y' = y, \tag{30}$$

$$t' = \frac{t - (vx/c^2)}{\sqrt{(1 - v^2/c^2)}}, \quad z' = z, \tag{31}$$

are a particular case of the above.

This is what we require about special relativity. Special relativity follows from the condition: *the equations of physics should be covariant with respect to the linear transformations that keep ds^2 invariant.* Covariant means: their form must be unchanged.

Best example are Maxwell's equations themselves; for Lorentz–transformed $F'_{\mu\nu}$ and J'_μ, we get, again,

$$\partial_\mu F'_{\mu\nu} = (4\pi/c)J'_\mu,$$

$$\partial_\mu F'_{\nu\rho} + \partial_\rho F'_{\mu\nu} + \partial_\nu F'_{\rho\mu} = 0.$$

The Lorentz–transformed current and field are:

$$J'_\rho(x') = L_{\rho\mu} J(x)_\mu,$$

$$F'_{\rho\sigma}(x') = L_{\rho\mu} L_{\sigma\nu} F_{\mu\nu}(x).$$

One also requires the transformation $x \mapsto x'$ to transform current and field.

12. General Relativity

General relativity is covariant gravitation theory; it is an extension of Newton's classical–mechanics version of gravitation.

Tensor calculus

We need here a rather extended geometrical introduction, in order to develop a language that will be the setting of general relativity.

Paths, contravariant vectors

Let M be some kind of n–dimensional, real, curved space (we think of M as a n–dimensional real differentiable manifold in the mathematical sense, but we'll require

very little of its properties here). Let $U \subset M$ be a domain over which range coordinates x^μ, $\mu = 0, 1, \ldots, n-1$. From here on until further notice we suppose that $\mu, \nu, \rho, \ldots$ range over $0, 1, \ldots, n-1$.

We will also use here indices that are "upstairs," as in the coordinates x^μ, and "downstairs," as in the partial differential operators ∂_μ. We again use here Einstein's summation convention, with a slight modification that will be specified below.

A *path* γ in U is a n–ple $x^\mu(t)$, with real parameter t ranging over a (possibly improper) real interval. A *covariant vector tangent to path* γ has as components the n derivatives dx^μ/dt, or $\dot{x}^\mu$.

Suppose that one makes a coordinate change $x^\mu \mapsto x'^\mu$ over domain U. Then the components of the vector tangent to γ will change accordingly:

$$\frac{dx^\rho}{dt}(x') = \frac{\partial x'^\rho}{\partial x^\mu}\frac{dx^\mu}{dt}(x).$$

Let now $A^\mu(x)$ be a set of functions that transform as above, that is:

$$A'^\rho(x') = \frac{\partial x'^\rho}{\partial x^\mu}A^\mu(x).$$

They are said to be the components of a *contravariant vector* on U.

Gradients, covariant vectors

Let $f(x)$ be a function defined on U. The *gradient of* f has components $\partial f/\partial x^\mu$, or as abbreviated above, $\partial_\mu f$. Given a transformation $x \mapsto x'$, the gradient's components transform as:

$$\frac{\partial}{\partial x'^\rho}f(x') = \frac{\partial x^\mu}{\partial x'^\rho}\frac{\partial}{\partial x^\mu}f(x).$$

Given any set of functions $B_\mu(x)$ that transform as a gradient,

$$B'_\rho(x') = \frac{\partial x^\mu}{\partial x'^\rho}B_\mu(x),$$

we say that the B_μ are the components of a *covariant vector* on U.

Summation convention

We can now state a new version of Einstein's summation convention, now for contravariant and covariant objects:

> *Repeated indices, one upstairs and the other one downstairs, are to be summed from* 0 *to* $n-1$.

The Kronecker delta δ^α_β

It is given by:

$$\frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta,$$

or $\delta^\alpha_\beta = 1$, if $\alpha = \beta$, and 0 otherwise. It's in fact a convenient representation for the elements of the $n \times n$ identity matrix. The Kronecker delta is our first example of a *mixed tensor of order* 2.

Tensors, contravariant, covariant and mixed

Objects with several indices such as $g_{\mu\nu}$, $T^{\alpha\beta}$, and A^α_β that transform according to the covariant rule of transformation, for downstairs indices, and the contravariant rule, for upstairs indices, are called *higher order tensors*; vectors and gradient–like objects are *first–order tensors*. The $g_{\mu\nu}$ are the components of a *covariant tensor of order* 2; the $T^{\mu\nu}$ are those of a *contravariant tensor of order* 2; and the A^α_β are those of a *2nd–order mixed tensor*, and so on.

The covariant derivative: contravariant vectors

We need a derivative operator that produces a tensor when applied to another tensor. That concept is given by the *covariant derivative*.

Let the $\mathbf{e}_\mu(x)$ denote a vector basis that spans U as above. For a contravariant vectorfield one has, $\mathbf{A} = A^\mu \mathbf{e}_\mu$. Let's compute the differential $d\mathbf{A}$:

$$d\mathbf{A} = dA^\mu \mathbf{e}_\mu + A^\mu d\mathbf{e}_\mu.$$

We now postulate that the differential $d\mathbf{e}_\mu$ is:

$$d\mathbf{e}_\mu = \Gamma^\rho_{\mu\nu} dx^\nu \mathbf{e}_\rho.$$

where the $\Gamma^\alpha_{\beta\gamma}$ account (in a way to be seen) for the effects of curvature. Into the equation upstairs,

$$d\mathbf{A} = dA^\mu \mathbf{e}_\mu + A^\mu \Gamma^\rho_{\mu\nu} dx^\nu \mathbf{e}_\rho.$$

We can change summed indices, as they are only bound variables:

$$d\mathbf{A} = dA^\mu \mathbf{e}_\mu + A^\rho \Gamma^\mu_{\rho\nu} dx^\nu \mathbf{e}_\mu.$$

Or, still,

$$d\mathbf{A} = \partial_\nu A^\mu dx^\nu \mathbf{e}_\mu + A^\rho \Gamma^\mu_{\rho\nu} dx^\nu \mathbf{e}_\mu.$$

If we now put $\mathbf{e}_\mu, dx^\nu$ in evidence,

$$d\mathbf{A} = (\partial_\nu A^\mu + \Gamma^\mu_{\rho\nu} A^\rho) dx^\nu \mathbf{e}_\mu.$$

We define: the covariant derivative $A^\mu_{;\nu}$ of A^μ is:

$$A^\mu_{;\nu} = \partial_\nu A^\mu + \Gamma^\mu_{\rho\nu} A^\rho = A^\mu_{,\nu} + \Gamma^\mu_{\rho\nu} A^\rho.$$

Since we have constructed $d\mathbf{A}$ as a vector, the $A^\mu_{;\nu}$ must transform as a 2nd–order mixed tensor. The effects of that on the $\Gamma^\rho_{\mu\nu}$ will be seen below.

The covariant derivative: covariant vectors

Consider a covariant vector $\mathbf{B} = B_\mu \mathbf{e}^\mu$. We can relate covariant basis $\mathbf{e}^\mu$ to the contravariant one by a duality induced by a scalar product:

$$\mathbf{e}_\mu \cdot \mathbf{e}^\nu = \delta^\nu_\mu.$$

(This is the usual duality operation in linear algebra.) Now, notice that δ^ν_μ is a constant object, and so its differential $d\delta^\nu_\mu = 0$. Then,

$$d\mathbf{e}_\mu \cdot \mathbf{e}^\nu + \mathbf{e}_\mu \cdot \mathbf{e}^\nu = 0.$$

As,

$$d\mathbf{e}_\mu = \Gamma^\rho_{\mu\nu} dx^\nu \mathbf{e}_\rho,$$

a simple computation shows that:

$$d\mathbf{e}^\mu = -\Gamma^\mu_{\rho\sigma} dx^\sigma \mathbf{e}^\rho.$$

Follows that, for a covariant vector $\mathbf{B} = B_\mu \mathbf{e}^\mu$,

$$B_{\mu;\nu} = \partial_\nu B_\mu - \Gamma^\rho_{\mu\nu} B_\rho = B_{\mu,\nu} - \Gamma^\rho_{\mu\nu} B_\rho.$$

Again the $B_{\mu;\nu}$ must transform as a tensor; this time as a 2nd–order covariant tensor.

More notational conventions

Let's write for the sake of convenience:

$$A^\mu_{;\nu} = \partial_\nu A^\mu + \Gamma^\mu_{\rho\nu} A^\rho \tag{32}$$

as:

$$D_\mu A = (\partial_\mu + \Gamma_\mu) A = \partial_\mu A + \Gamma_\mu A. \tag{33}$$

Here A is a "vertical" vector of components A^μ, and the n $\Gamma_\mu = (\Gamma^\rho_\sigma)_\mu$ are seen as matrices that act upon A.

Suppose now that we can detach the transformation of the "inner" indices in $A = (A^\mu)$ from those of the "outer" indices like the μ in Γ_μ. Let us consider a mapping $D_\mu A \mapsto U^{-1}(D_\mu A)$, where U is a matrix that acts on the space of the A. (We'll soon relate U to the preceding coordinate transformations; the use of an inverse matrix is both by convenience and to emphasize that we deal with a nonsingular map.)

Then:

$$U^{-1}(D_\mu A) = U^{-1}(D_\mu U U^{-1} A), \tag{34}$$

or

$$U^{-1} D_\mu A = [U^{-1}\partial U + U^{-1}\Gamma_\mu U] U^{-1} A. \tag{35}$$

That is, the maps $A \mapsto U^{-1}A, D_\mu A \mapsto U^{-1}D_\mu A$ induce (and are induced by) the transformation:

$$\Gamma_\mu \mapsto [U^{-1}\partial_\mu U + U^{-1}\Gamma_\mu U].$$

For $U = e^{\Lambda(x^\mu)}$, and Abelian,

$$\Gamma_\mu \mapsto \Gamma_\mu + \partial_\mu \Lambda,$$

that is, a gauge transformation like those in electromagnetic theory; see eq. (27).

One sometimes call the Γ_ρ an *affinity*.

Properties of the affinity $\Gamma^\mu_{\nu\rho}$

Out of the preceding remarks it is easy but cumbersome to check that the $\Gamma^\mu_{\nu\rho}$ transform as follows:

$$\bar{\Gamma}^\alpha_{\beta\gamma} = \frac{\partial \bar{x}^\alpha}{\partial x^\mu}\frac{\partial x^\nu}{\partial \bar{x}^\beta}\frac{\partial x^\rho}{\partial \bar{x}^\gamma}\Gamma^\mu_{\nu\rho} + \frac{\partial \bar{x}^\alpha}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial \bar{x}^\beta \bar{x}^\gamma}.$$

Moreover:

- The sum $\Gamma^\mu_{\nu\rho} + A^\mu_{\nu\rho}$ of an affinity Γ and a tensor A, one time contravariant and two times covariant, is another affinity.
- The difference of two affinities $\Gamma'^\mu_{\nu\rho} - \Gamma^\mu_{\nu\rho}$ is a tensor.

The curvature tensor

The preceding trick also simplifies the computations that lead to the curvature tensor. The partial derivative operator commutes, that is, $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$, for well–behaved points. This isn't the case for the covariant derivative D_μ:

$$(D_\mu D_\nu - D_\nu D_\mu)A = R_{\mu\nu} A, \tag{36}$$

where:

(37) $$R_{\mu\nu} = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + \Gamma_\mu \Gamma_\nu - \Gamma_\nu \Gamma_\mu,$$

or even, if we write the commutator $[A, B] = AB - BA$,

(38) $$R_{\mu\nu} = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + [\Gamma_\mu, \Gamma_\nu].$$

Again, if we deal with Abelian Γ_μ, we get an electromagnetic–like field: $R_{\mu\nu} = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\nu$.

Interpretation of the $R_{\mu\nu}$ is simple: infinitesimally parallel transport A along coordinate x^μ, and then along x^ν; get A'. Now do the same along x^ν and then along $x^{\mu\prime\prime}$. Get A''. If there is no curvature, $A' - A'' = 0$. This is not the case when there is some curvature present, and the nonzero term $R_{\mu\nu}A$ gives a measure of how far from flatness is this region U of our original space.

Riemann's tensor

Riemann's tensor is then, simply:

$$R^\mu_{\nu\rho\sigma} = \partial_\rho \Gamma^\mu_{\nu\sigma} - \partial_\sigma \Gamma^\mu_{\nu\rho} + \Gamma^\mu_{\lambda\rho} \Gamma^\lambda_{\nu\sigma} - \Gamma^\mu_{\lambda\sigma} \Gamma^\lambda_{\nu\rho}.$$

Metric tensor, Christoffel symbols, Riemann–Christoffel tensor

Let $d\mathbf{s} = dx^\mu \mathbf{e}_\mu$ be an infinitesimal displacement; we have that:

$$ds^2 = d\mathbf{s} \cdot d\mathbf{s} = (\mathbf{e}_\mu \cdot \mathbf{e}_\nu) dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu.$$

The $g_{\mu\nu} = (\mathbf{e}_\mu \cdot \mathbf{e}_\nu)$ are the components of the *Riemannian metric tensor.* Notice the symmetry $g_{\mu\nu} = g_{\nu\mu}$. We now impose that the $g_{\mu\nu}$ be *covariantly constant*, that is, for all ρ,

$$D_\rho g_{\mu\nu} = g_{\mu\nu;\rho} = 0.$$

This means that, if we view the D_ρ as a kind of infinitesimal displacement operators — which they are — then the components of the Riemannian metric tensor are kept unchanged throughout those displacements.

The fact that the $g_{\mu\nu}$ are covariantly constant leads (after some computations) to the relation:

$$\{^\mu_{\nu\rho}\} = \Gamma^\mu_{\nu\rho} = \frac{1}{2} g^{\mu\sigma} (\partial_\nu g_{\rho\sigma} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\nu\rho}).$$

Here we have used the defining relation $g^{\mu\rho} g_{\rho\nu} = \delta^\mu_\nu$. The Γ thus defined are the *Christoffel symbols of the second kind*; those of the first kind are the expression above without the $g^{\mu\nu}$ in front of it.

The fully covariant Riemann curvature tensor defined out of the Christoffel symbols as affinities is the *Riemann–Christoffel tensor.* Formally,

$$R_{\mu\nu\rho\sigma} = g_{\mu\lambda} R^{\lambda}_{\nu\rho\sigma}. \tag{39}$$

The following symmetry properties are established by direct calculation:

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu}, \quad R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma}. \tag{40}$$

Geodesics

Let $x^{\mu}(\lambda)$ be a curve parametrized by λ. Let ds be an infinitesimal arc along that curve. We know that the curve thus given is extremal if and only if $\delta \int_{\sigma} ds = 0$, that is,

$$\delta \int_{\sigma} g_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda} d\lambda^2. \tag{41}$$

Compare with the deduction of eq. (19); we get:

$$\frac{d^2 x^{\mu}}{d\lambda^2} + \{^{\mu}_{\rho\sigma}\} \frac{dx^{\rho}}{d\lambda} \frac{dx^{\sigma}}{d\lambda} = 0. \tag{42}$$

The Bianchi identities

They are easy to check; they only require a few straightforward calculations:

$$R^{\mu}_{\nu\rho\sigma;\lambda} + R^{\mu}_{\nu\lambda\rho;\sigma} + R^{\mu}_{\nu\sigma\lambda;\rho} = 0. \tag{43}$$

Compare eq. (43) to eq. (25). The earlier, linear set locally determines potentials for the fields $F_{\mu\nu}$; also, the similarity between eqs. (26) and (38) is striking. However the above nonlinear conditions *do not determine* a single "potential" (in our case, the affinity) for the curvature seen as a "field"; not even modulo a gauge transformation.

Ricci and Ricci–Einstein tensors

The *Ricci tensor* is:

$$R_{\mu\nu} = R^{\rho}_{\mu\nu\rho}. \tag{44}$$

The Ricci tensor is symmetrical in its two indices. Out of it we obtain the *Ricci–Einstein tensor*, also symmetrical in its two indices:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \tag{45}$$

where:

$$R = g^{\mu\nu} R_{\mu\nu} \tag{46}$$

is the *curvature scalar*.

The Einstein field equations

Now let our space M be 4–dimensional, real, and endowed with a (pseudo)–Riemannian metric tensor of signature +2 as in eq. (29). M is *Einstein's spacetime*. We can finally state the *Einstein field equations for the gravitational field*:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa T_{\mu\nu}. \tag{47}$$

$T_{\mu\nu}$ is the *energy–momentum tensor*; it is the material source for the geometrical l.h.s. of the field equations. κ includes the contribution of Newton's gravitational constant G. We can also write the Einstein equations as:

$$R_{\mu\nu} = \kappa(T_{\mu\nu} - \frac{1}{2} T),$$

for $T = g^{\mu\nu} T_{\mu\nu}$. Then, in empty space (that is, $T_{\mu\nu} = 0$, absence of matter), eq. (47) gives:

$$R_{\mu\nu} = 0. \tag{48}$$

A Lagrangian density for the Einstein field equations

Let g denote the determinant of the matrix $(g_{\mu\nu})$. The Lagrangian density for the empty–space Einstein field equations will be:

$$\mathscr{L} = \sqrt{-g} R. \tag{49}$$

Here,

$$R = g^{\mu\nu} [\Gamma^{\rho}_{\mu\rho,\nu} - \Gamma^{\rho}_{\mu\nu,\rho} + \Gamma^{\rho}_{\mu\sigma} \Gamma^{\sigma}_{\rho\nu} - \Gamma^{\rho}_{\mu\nu} \Gamma^{\sigma}_{\rho\sigma}], \tag{50}$$

can be seen as defined out of two independent variables, the $g_{\mu\nu}$ and the $\Gamma^{\rho}_{\mu\nu}$. If we treat them independently we use the so–called Palatini formalism. The relation between metric tensor and affinity will be derived through the Euler–Lagrange equations.

We sketch the derivation of the field equations. For the variations we require:

$$\delta\sqrt{-g} = \frac{\partial\sqrt{-g}}{\partial g_{\mu\nu}} \delta g_{\mu\nu} = \frac{1}{2}\sqrt{-g}\frac{1}{g}\frac{\partial g}{\partial g_{\mu\nu}} \delta g_{\mu\nu} = \frac{1}{2}\sqrt{-g} g^{\mu\nu} \delta g_{\mu\nu}. \tag{51}$$

Also, from $g^{\mu\nu} g_{\nu\sigma} = \delta^{\mu}_{\sigma}$ we get that $\delta g^{\rho\sigma} = -g^{\rho\mu} g^{\sigma\nu} \delta g_{\mu\nu}$. This leads to:

$$(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R) \delta g_{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu} = 0. \tag{52}$$

Since the $g_{\mu\nu}$ and $\Gamma^{\rho}_{\mu\nu}$ are independently varied, the term within parentheses gives the field equations for empty space; the second term, after several manipulations, implies $g^{\mu\nu}_{;\sigma} = 0$, that is the relation that leads to the relation $\Gamma^{\rho}_{\mu\nu} = \{^{\rho}_{\mu\nu}\}$, or: the affinities are the Christoffel symbols.

13. Gauge field theories

Gauge field theories are very similar to both electromagnetic theory and gravitation theory; in fact they are a kind of nonlinear version of electromagnetism, that includes gravitation as a particular case. We start from a finite–dimensional, semi–simple Lie group G, finite–dimensional, usually identified to one of its (simple) matrix representations; the Lie algebra of G is noted $L(G)$, and if we impose that G be connected, $L(G)$ "generates" G in the following sense: if $A_g \in L(G)$ then $g = e^{A_g} \in G$. So, we can look at the elements of $L(G)$ as a set of matrices that generate G. The x^μ range over spacetime, as in the case of general relativity and gravitation theory.

We will proceed exactly as in eqs. (33)–(38), and get, for a covariant vector $(A^a_b)_\mu = A_\mu$, where the a, b are matrix indices to indicate that the A_μ are matrices in the Lie algebra $L(G)$:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu], \tag{53}$$

where the $L(G)$–valued $F_{\mu\nu}$ are the components of the *gauge field*. The U in eq. (35) take values in elements of G, or can be seen as elements of G parametrized by spacetime variables x^μ. The *gauge field equations* are:

$$D_\mu F^{\mu\nu} = j^\mu, \tag{54}$$

while we again have a set of Bianchi–like identities,

$$D_\rho F_{\mu\nu} + D_\nu F_{\rho\mu} + D_\mu F_{\nu\rho} = 0. \tag{55}$$

The gauge field equations are eqs. (54), and (55), this last one coupled with eq. (53), as they are *not* equivalent (over an open ball in spacetime), as it is the case for the electromagnetic field.

A Lagrangian density for the empty–space gauge field equations

The sourceless gauge field equations can be derived (through the Lagrangian field equations, eq. (23)) from the Lagrangian density:

$$\mathscr{L} = \frac{1}{2} F^a_{\mu\nu} F^{a\mu\nu} - F^{a\mu\nu}(D_\mu A^a_\nu - D_\nu A^a_\mu). \tag{56}$$

(Summation is also to be done over the gauge index a.) This density is the nonlinear version of the electromagnetic Lagrangian density (56).

Gauge transformations of the first and second kind

Consider a Lagrangian density which is quadratic in all fields in the following sense: it is a sum of terms of the form $\phi^*\chi$, where * denotes complex conjugation or Hermitian conjugation. Thus the transformation (for α a real constant):

$$\phi \mapsto e^{i\alpha}\phi,$$

$$\phi^* \mapsto e^{-i\alpha}\phi^*,$$

keeps invariant the Lagrangian density, since the exponential is a constant and moves inwards and outwards differential operators.

This is called a *gauge transformation of the first kind.*

Now suppose that the transformation is pointwise or locally defined, that is, for a function $\chi(x)$:

$$\phi \mapsto e^{i\chi(x)}\phi,$$

$$\phi^* \mapsto e^{-i\chi(x)}\phi^*.$$

The transformation doesn't commute with the differential operator ∂_μ, and so the Lagrangian density isn't invariant with respect to it. This is a *gauge transformation of the second kind.*

We give an example for Abelian fields.

We use here a trick. We add a field A_μ that will compensate for the "bad" behavior of the transformation under ∂_μ: if $\psi' = e^{i\chi(x)}\psi$, then:

$$(\partial_\mu + A'_\mu)\psi' = e^{i\chi}(\partial_\mu + A_\mu)\psi,$$

provided that:

$$A'_\mu = A_\mu - i\partial_\mu\chi,$$

which is a gauge transformation as we know it from electromagnetic theory (for Abelian fields in general).

Generalization to the nonabelian case is straightforward. This means: *the imposition of a gauge invariance of the second kind can be compensated by the addition of a gauge potential A_μ to the original Lagrangian density in an adequate covariant derivative operator.*

That is to say, where we had an ordinary differential operator ∂_μ, imposition of a local gauge symmetry makes it into a covariant derivative, viz. $\partial_\mu \mapsto D_\mu = \partial_\mu + A_\mu$.

A brief overview of the Higgs mechanism

We will restrict our consideration to Abelian fields, but the extension to the non-abelian case is straightforward.

We start from a massive, spin 0 field with a nonlinear interaction:

$$\mathscr{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi^* \phi),$$

$$V(\phi^* \phi) = \frac{m^2}{2v^2}(\phi^* \phi - v^2)^2.$$

(Squared mass is $m^2/2v^2$.) m, v are constants.The extra terms in V are the non-linear self–interaction of the field ϕ. Now impose local gauge symmetry, that is, add a gauge field to countermand the symmetry condition. The Lagrangian density becomes:

$$\mathscr{L} = D_\mu \phi^* D^\mu \phi - V(\phi^* \phi) - (1/4) F_{\mu\nu} F^{\mu\nu},$$

and with a slight change,

$$D_\mu = \partial_\mu + iqA_\mu.$$

Notice that the gauge field is a massless field, of course.

Now suppose that $v > 0$ and consider a perturbed field $\phi' = v + (1/\sqrt{2})h$, ϕ a real field. The idea is that such a v is a ground state for the potential energy that (in fact) breaks the gauge symmetry of the system. The original Lagrangian density becomes:

$$\mathscr{L} = \partial_\mu h \partial^\mu h - m^2 h^2 - (1/4) F_{\mu\nu} F^{\mu\nu} + q^2 v^2 A'_\mu A'^\mu + \text{other terms}.$$

We have here two massive fields that interact: the Higgs field h and the gauge field $F_{\mu\nu}$, with mass term $q^2 v^2 A'_\mu A'^\mu$. (A' is the gauge transform of A.) We say that *spontaneously broken symmetry leads to the appearance of mass,* as if mass were known through interaction with the Higgs field.

This concludes our survey, that went from the 18th century beginnings of mechanics to techniques developed in the 60s in the 20h century, such as the Higgs mechanism. We conclude with a brief discussion of axiomatic tools in theoretical physics.

14. A note on axiomatizations of physics

An axiomatics for physics is the query in Hilbert's 6th Problem. It is considered to be an ugly duckling among the Hilbert Problems, as many authors dispute the actual usefulness of placing physics in an axiomatic framework. We will briefly mention a few results of our own (with da Costa) on the axiomatics of physics:

Proposition 14.1. *Classical mechanics, Schrödinger's quantum mechanics, electromagnetism, general relativity and gauge field theory can all be axiomatized within Zermelo–Fraenkel set theory.* □

Therefore all known results in theoretical physics become ZF theorems, as long as they can be given rigorous formulations, let us say, by the usual mathematical standards and without the help of "very large" objects (that last condition is given within quotation marks to mean that we exclude large cardinal extensions of ZF).

Proposition 14.2. *There are "physically meaningful" undecidable sentences in any consistent* ZF *axiomatization of physics.* □

By "physically meaningful" we mean sentences that describe actual situations in physics, e.g., for a particular Lagrangian L, the formalized version of the sentence "L describes a harmonic oscillator" is formally undecidable within ZF. Obviously,

Proposition 14.3. *Any two* ZF *axiomatizations of physics have the same undecidable sentences.* □

This means that, for a wide variety of constructions, choosing a particular axiomatics for physics is just a matter of taste. It won't affect the results we can derive (or that we can't derive) from our axiomatic framework. For more details see da Costa & Doria 2006.[2]

References

da Costa, N. C. A. and Doria, F. A. 2006. Janus–faced physics: on Hilbert's 6th Problem. In C. Calude (ed.) *Randomness and Complexity: From Leibniz to Chaitin*. New York: World Scientific, pp. 25–68.

Goldstein, H. 1950. *Classical Mechanics*. New York: Addison–Wesley.

Lanczos, C. 1970. *The Variational Principles of Mechanics*. Toronto: Univ. of Toronto Press.

Leech, J. W. 1965. *Classical Mechanics*. Methuen: Methuen.

FRANCISCO ANTONIO DORIA
Advanced Studies Research Group, Production Engineering Program, COPPE–UFRJ
fadoria@gmail.com, fadoria@pep.ufrj.br

Resumo. Fazemos um sumário das principais áreas em física teórica, enfatizando–se sua relação ao formalismo lagrangeano em mecânica clássica. Este sumário cobre: mecânica clássica, o caminho da mecânica clássica à mecânica quântica de Schrödinger, o eletromagnetismo, relativade restrita e geral, e, rapidamente, teorias de gauge, inclusive o mecanismo de Higgs. Deixamos de lado o rigor matemático em favor de uma apresentação mais simples.

Palavras-chave: Física teórica, lagrangeana, mecânica clássica, mecânica quântica, eletromagnetismo, relatividade, mecanismo de Higgs.

Notes

[1] Actually an assertion out of Newton's First and Second Laws.

[2] The present paper is part of an ongoing research program with N. C. A. da Costa. The author thanks CNPq, Philosophy Section, for support, as well as PEP–COPPE, at Rio's Federal University in the person of its chairman Professor R. Bartholo. Thanks are also due to Professor C. A. Cosenza, chairman of the Fuzzy Sets Research Group. FAD also stresses the interest shown by the History & Philosophy of Science Program at COPPE in his research activities and wants to thank Professor S. Fuks for it.