The Physics behind Genetic Programming.

James Cunha Werner (wernerjc@sbu.ac.uk)

SCISM
South Bank University
103 Borough Road
London SE1 0AA

Abstract. This paper presents the historic scenery of calculus of variations (CV), one of the central tools of theoretical physics, and its relashionship with genetic programming (GP) algorithms, a search method with would be considered a numerical solution for the method of variations.
Keywords: Genetic Programming, Calculus of Variations, Euler-Lagrange Equation.

Introduction

In elementary physics, equations are usually applied establishing relationships among variables involved in the problem. The origins of these equations are based on nature fundamental laws of minimum effort, were enunciated by Euler in 1753, with a spectacular contribution of Lagrange.
Joseph Louis Lagrange descended of a ownerships family from Turin, and its first scientific contribution was a letter to Euler, with 19 years old in 1755, in which solved the isoperimetric problem that for more than half century had been a subject of discussion in scientific community.
The problem consists of determining the form of a function so that a formula in which it entered should satisfy a certain condition. To solve this problem, it was necessary to enunciate the principles of the calculus of variations.
Euler recognized the generality of the method adopted, and the superiority in relation to its proposal, what took him to withheld a paper he had previously written, with covered some of the same ground, given time for the young mathematician to finish his work and publish it in Miscellanea Taurinensia's volumes ([1]).
Inside of this historical scenery, we identified the genetic programming as a solution method to the problem of calculus of variations, difficult of being accomplished in an analytic or numeric way.
This paper sections deals with the historical point-of-view, followed by a mathematical formulation, and its relation with genetic programming, with will solve the least distance between two points problem.

Maupertuis laws.

On 20 February 1740 ([2]), M. de Maupertuis presented in the Real Academy of Science of Paris, the universal principle of rest and equilibrium, supplying a great universality to all the equilibrium cases. There was several laws of nature:

1. "In any assemblage of bodies their common centre of gravity is as low as possible".
2. Conservation of living forces, a brief enumeration of Bernoulli's conservatio virium vivarum, that kinetic energy cannot be destroyed without causing a change, that is, without being conserved into another form. At 1750, this statement was expressed as *conservation of energy* and influenced all classical mechanics structure: "In every system of particles which mutually interact, the sum of the products of each mass by the square of its velocity, which is called living force, must remain invariant".
3. Law of equilibrium: "Consider a system of heavy particles which are attracted to centres by forces, each of which acts towards its centre, as the nth power of its distance to the centre. In order that all the particles remain in equilibrium, it is necessary that the sum of the products of each mass by the intensity of its force and by the (n+1)st power of its distance to the centre of its force (which is called the sum of the equilibrium forces) must be a maximum or minimum".

In 1744, he proposed the principle of least action into the paper "Harmony between different laws of nature which have, up to now, appeared incompatible", as "Nature in the production of its effects, acts always by the simplest means", with means that the action is the least.
The concept is clear, but the mathematical tool to obtain it was work field for Euler and Lagrange during many years.

Euler's Principio Minimae Actionis ([3] and [4]).

Euler considered that the equilibrium state mainly in the mechanical systems were certain obtained with easiness. It also verified that the action sum of each system force in each time interval is minimum, termed "Principle of less action".
Euler considered the problem of finding all curves $y=y(x)$, with $0<=x<=a$, for what the action

$$\int_0^a Z\left(x, y, p = \frac{dy}{dx}, q = \frac{dp}{dx}, r = \frac{dq}{dx}, ... \right) dx$$

is a maximum or a minimum.

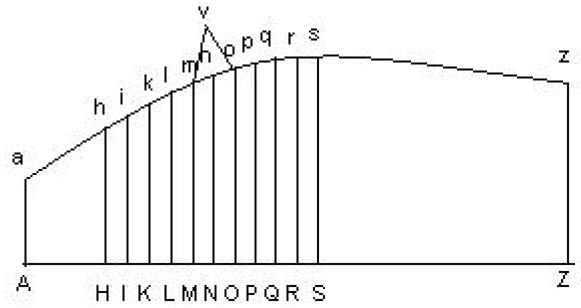


Fig. 1 Interval division in segments.

Let the differential displacements from fig. 1:

$$dx = HI = IK = KL = LM = MN = NO = OP = PQ = QR = RS$$

Euler introduces the nomenclature of index superscript and subscript to represent positions before and after the point he was working. Then:

$$\begin{aligned} &... \\ &Kk = y_{ii} \\ &Ll = y_i \\ &Mm = y \\ &Nn = y^i \\ &Oo = y^{ii} \\ &... \end{aligned}$$

At Euler's time, the geometry had a fundamental paper in the modelling of the physical systems. In this context, the extreme occurs when is considered that the first differential of a variable tends to zero without second differential vanishes.

In the case of second differential be negative, we will have a maximum, and in positive case, a minimum. If it tends also to zero, and the third differential don't tend, it will be a maximum if fourth differential is negative and a minimum if positive. In general, to a point be a maximum or a minimum is necessary that the successive orders of differentials that tend to zero happen in odd number, and the last differential sign had defined the type.

The derivative is calculated as:

$$...;\ p_{ii} = \frac{y_i - y_{ii}}{dx};\ p_i = \frac{y - y_i}{dx};\ p = \frac{y^i - y}{dx};\ p^i = \frac{y^{ii} - y^i}{dx};\ p^{ii} = \frac{y^{iii} - y^{ii}}{dx} ...$$

Let the limit where dx→0, the action integral by the sum:

$$\int_0^{x_i} Zdx + Z^i dx + Z^{ii} dx + Z^{iii} dx + ...$$

where Z=Z(x,y,p,q,…), $Z^i=Z(x^i,y^i,p^i,q^i,\ldots)$, $Z^{ii}=Z(x^{ii},y^{ii},p^{ii},q^{ii},\ldots)$,…

The derivative of this expression with respect of y' must vanish at the value y^i=Nn.

Euler start to analyse Z of the form:

$$dZ = Mdx + Ndy + Pdp$$

and the integral of Z(x,y,p) is an extremum. Then

$$Zdx + Z^i dx = Z(x, y, p)dx + Z(x^i, y^i, p^i)dx$$

This equation is a function of y', and if this expression is a extremum, its derivation with y' vanish. The change in Z dx caused by varying y' is P.nv, where nv=y'-y. and Z' dx is N' nv – P' nv. This implies that for an extremum

$$P + N^i dx - P^i = 0$$

$$N' - \frac{(P' - P)}{dx} = 0$$

$$N - \frac{dP}{dx} = 0$$

For the case where Z(x,y,p,q) Euler obtained

$$N - \frac{dP}{dx} + \frac{d^2 Q}{dx^2} = 0$$

with contains the same structure of later equation, plus a higher order term.

Lagrange's letter to Euler.

At 1755, Louis de La Grange wrote Euler a letter showing how to eliminate from Euler's methods the tedium and need for geometrical insight and to reduce the entire process to a quite analytical technique, with is the calculus of variations. Euler said that he "had meditated a long time and had revealed to friends his desire yet the glory of first discovery was reserved to the very penetrating geometer of Turin La Grange, who having used analysis alone, has clearly attained the very same solution which he had deduced by geometrical considerations".
To obtain this, Lagrange ([4], [6], and [7]) developed a new form of differential, acting on functions, which came to be know as variations. To distinguish this operator from usual differential he wrote it as δ instead of d. The variable x is kept fixed and the ordinary differential is calculated, but with the differential arguments δy, δy', …
This operator has the commutative properties:

$$\delta\alpha\left(f_1 + g_1\right) = \alpha\left(\delta f_1 + \delta g_1\right)$$

$$d\delta = \delta d$$

$$\frac{d}{dt}\delta f_1 = \delta\left(\frac{df_1}{dt}\right)$$

$$\delta\int \; = \int\delta$$

Then

$$\delta\int Z = \int \delta Z$$

The $\int Z$ could be understood as evaluated along a family of arcs y=y(x,α) and variation process consists of differentiation with respect to the parameter α. Instead of Euler's highly specialized comparison curves, Lagrange has recourse to very general ones.
Then if the given arc corresponds to

$$\alpha = 0 \rightarrow \delta y(x) = \alpha\left(\left.\frac{\partial y(x,\alpha)}{\partial \alpha}\right|_{\alpha=0}\right)$$

Lagrange defined Z as Euler:

$$dZ = Mdx + Ndy + Pd^2y + Qd^3y + Rd^4y + \ldots$$

$$\delta Z = N\delta y + P\delta dy + Q\delta d^2 y + R\delta d^3 y + \ldots$$

$$\delta\int Z = \int N\delta y + \int P\delta dy + \int Q\delta d^2 y + \int R\delta d^3 y + \ldots$$

$$= \int N\delta y + \int Pd\delta y + \int Qd^2\delta y + \int Rd^3\delta y + \ldots$$

$$= \int N\delta y + P\delta y - \int dP\delta y + Qd\delta y - dQ\delta y + \int d^2 Q\delta y + \ldots$$

$$= \int\left(N - dP + d^2Q - \ldots\right)\delta y + (P - dQ + \ldots)\delta y + (Q - \ldots)d\delta y + \ldots$$

If the comparison curves are such that 0=δy =dδy =… at x=0 and x=a then:

$$\delta\int Z = \int\left(N - dP + d^2Q - \ldots\right)\delta y$$

Since δy is arbitrary on the interval from 0 to a, it follows that

$$N - dP + d^2Q - \ldots = 0$$

Lagrange assumes that the integrand is zero due to the definition obtained using differentials to obtain function extremum.

Modern approach of the calculus of variation ([5] and [7]).

To formulate this concept mathematically, be

$$L = L(r, v ; t)$$

a continuous differentiable function over a points set with

$$r = x_i(t)$$

$$v = \frac{dr}{dt}$$

A field is a bijection space point-real number where in each point is associated a function L, called Lagrangean of the field (or Lagrange function). Lets define action S of the considered field as the function:

$$S = \int_b^a L\, dt$$

Let variation (or in fact, first variation) of a function f_1 the result of the mathematical operation of an operator δ under f_1 that

$$d f_1 = f_1 - f_2$$

The operator variation is symbolized by δ to emphasize the difference between it and the differential operator d, although, as it will be seen, they can made a mistake formally: while d reflects the lineal part of the f_i increment for a given trajectory, δ refers to the f_i variation on two different trajectories for the same value of t, or different trajectories between points A and B (see figure 2).

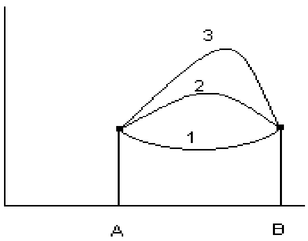


Figure 2: Different trajectories between A and B with different values of action.

The fundamental problem of calculus of variations is: Let S the action and its boundaries conditions, obtain the trajectory

$$r(t)$$

where S is stationary (usually a minimal value) into [a,b] interval, and the boundary conditions be respected. The S variation, when occurs a variation of the trajectory

$$d(r)$$

is:

$$\int_a^b L(r + dr, v + dv; t)dt - \int_a^b L(r, v; t)dt$$

We would develop using the mean value theorem (generalized Taylor theorem). Let

$$D_k = \left(dr \frac{\partial}{\partial r} + \partial v \frac{\partial}{\partial v} \right)^k$$

the variation of S will be:

$$\int_a^b \left(\sum_k \frac{1}{k!} D_k L \right) dt + \frac{1}{(k+1)!} D_{k+1} L_r\, dt \qquad \text{Eq. 1}$$

We call k-esima variation of S the k-esima part of expression 1. The necessary condition to S be stable is that the first variation $\delta S=0$.

The first term of equation 1 is:

$$\delta S = \delta\int_a^b L\, dt = \int_a^b \left(\delta r \frac{\partial L}{\partial r} + \delta v \frac{\partial L}{\partial v} + \delta t \frac{\partial L}{\partial t} \right) dt$$

Then the problem of variation calculus is obtain $\delta S=0$, with satisfy boundaries conditions.
When occurs the variation of S, $\delta t=0$ by variation definition, then:

$$\delta S = \int_a^b \left(\delta r \frac{\partial L}{\partial r} + \delta v \frac{\partial L}{\partial v} \right) dt$$

Let

$$\frac{d}{dt}\delta r = \delta \frac{dr}{dt} = \delta v$$

$$\delta S = \int_a^b \left(\delta r \frac{\partial L}{\partial r} + \frac{\partial L}{\partial v}\frac{d}{dt}\delta r \right) dt$$

that with part integration results:

$$\int_a^b \frac{\partial L}{\partial v} d(\delta r) = \frac{\partial L}{\partial v}\delta r \Big|_a^b - \int_a^b d\left(\frac{\partial L}{\partial v}\right)\delta r = -\int_a^b d\left(\frac{\partial L}{\partial v}\right)\delta r$$

or

$$\delta S = \int \left[\frac{\partial L}{\partial r} - \frac{d}{dt}\left(\frac{\partial L}{\partial v}\right)\right] \delta r\, dt = 0$$

If δr is arbitrary, we would conclude that the hook term will be null, otherwise we would have vectors r with δr times the hook vector be always a positive number, and by integral definition, it will be not null. The orthogonallity condition between δr and the other vector result into the same result.
Then:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) = \frac{\partial L}{\partial r} \qquad \text{Eq. 2}$$

This equation is called Euler-Lagrange Equation.
Dynamic systems can be modelled obtaining the Lagrangean function that satisfies the minimum condition of S. Minimum (or maximum) is symmetric extreme conditions (min [-f] = max [f]) and usually termed functional extremes.
A very interesting result is the application to show that the least distance between two points is a line. The distance between two points [a,b] under a generic curve is:

$$ds = \sqrt{dx^2 + dy^2} = dx\sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \sqrt{1 + (y')^2}\, dx$$

$$s = \int_a^b \sqrt{1 + (y')^2}\, dx$$

$$L(x) = \sqrt{1 + (y')^2}$$

Then

$$\frac{\partial L}{\partial y} = \frac{y' y''}{\sqrt{1 + (y')^2}} \;;\; \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}} \;;\; \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = \frac{y' y''}{\sqrt{\left(1 + (y')^2\right)^3}}$$

The solution occurs when y’=0, with integration results y=cte, or when y’’=0, with integration is a line [y = cx + d].

The basis of Analytical Mechanics [5].

The application of Euler-Lagrange equation in mechanical problems consists into the definition of a origin for a reference system, obtain the potential and kinetics energy, replace into the Lagragean definition and obtain the differential equation model of the system through Euler Lagrange equation. The integration of these coupled equations are the solution of the problem.
The main problem here is the solution of the system of coupled equations by analytical methods, usually difficult or impossible, where the numerical solution would solve, when convergence is obtained. Fig. 3 shows two examples of application in mechanical systems.

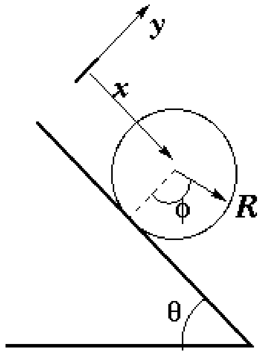
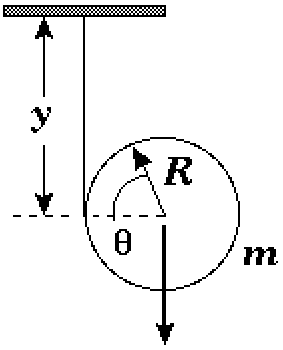
	 	
Potential Energy	U = -Mg x sin θ	U=-mgy
Constraints	$x = R f$ $\dot{x} = R \dot{f}$	f=y-Rθ=0
Kinetics Energy	$T = \frac{1}{2} M \dot{x}^2 + \frac{1}{2} I \dot{f}^2 =$ $\frac{1}{2} M \dot{x}^2 + \frac{1}{2}\left(\frac{2}{5} MR^2\right)\frac{\dot{x}^2}{R^2} =$ $\frac{7}{10} M \dot{x}^2$	$T = \frac{1}{2} M \dot{y}^2 + \frac{1}{2} I_{CM} \dot{q}^2 =$ $\frac{1}{2} M \dot{y}^2 + \frac{1}{2}\left(\frac{1}{2} mR^2\right)\dot{q}^2 =$ $\frac{1}{2} M \dot{y}^2 + \frac{1}{4} mR^2 \dot{q}^2$
Lagrangean	$L = T - V = \frac{7}{10} M \dot{x}^2 + Mgx\ sin(q)$	$L = T - V =$ $\frac{1}{2} M \dot{y}^2 + \frac{1}{4} mR^2 \dot{q}^2 + mgy$
Euler equations	$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}$ $\frac{d}{dt}\left(\frac{7}{5} M \dot{x}\right) = Mg\ sin(q)$ $\ddot{x} = \frac{5}{7} g sin(q)$	$\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) + l \frac{\partial f}{\partial y} = 0$ $mg - m\ddot{y} + l = 0$ $\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) + l \frac{\partial f}{\partial q} = 0$ $-\frac{1}{2} mR^2 \ddot{q} - lR = 0$

Fig. 3 Application of Euler-Lagrange Equation to mechanical engineering problems.

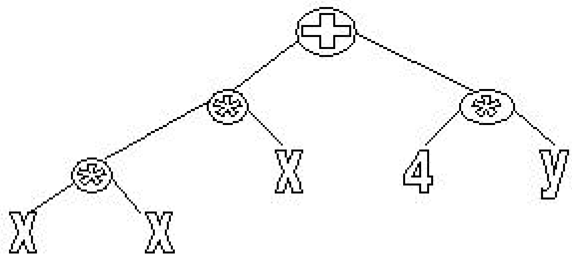
Understanding Genetic Programming as a Variation Methods.

The genetic programming (GP) algorithm ([8] and [9]) mimics the evolution and improvement of life through reproduction, when each individual contributes with its own genetic information to building new individuals with greater fitness to the environment and higher chances of survival.

Each ‘individual’ in a generation represents, with its chromosome, a feasible solution to the problem: a mathematical function that is the goal of calculus of variations.

A mathematical function would be wrote as a tree (Fig. 4) or be coded into the chromosome in infix format like, for example:

Mathematical format: x^3+4y



Infix format: or (+ (* (* x x) x) (* 4 y))

Fig. 4 Tree representation of mathematical function.

A set of random functions (a population) is created and for each one fitness function is measured. It can be observed that the action has the same role of the fitness function, measured in the function generated by the genetic programming, calculated along an interval of time or metric dimension.
The algorithm would be synthesises as (see fig.5):

- Generation of random elements trees (different guests to the mathematical shape of the solution).
- Fitness evaluation of each individual (evaluation of action for each individual).
- Selection of best individuals, Crossover and Mutation (genetic operators searching for the action extreme as numerical solution).
- Go to step 2, until convergence condition be find.

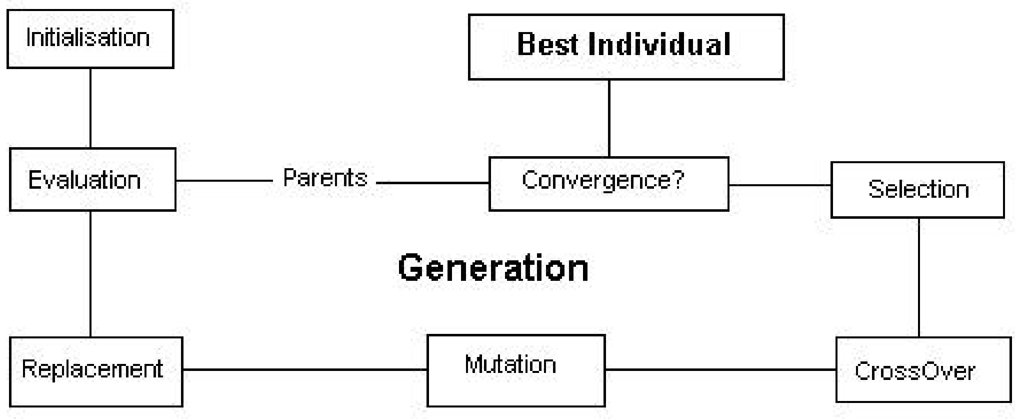


Fig 5. Genetic programming diagram.

The best individuals are continuously being selected, and crossover and mutation take place. Following a number of generations, the population converges to the optimal solution, with action extreme. There are two kinds of information defined for the algorithm: terminals (variable values and random numbers) and functions (mathematical functions used in the generated model). Depending its definition, a different solution space is mapped.
Lets apply genetic programming to the problem of least distance between two points. The functions defined are multiplication, sum and subtraction, and the terminal is the x coordinate. Genetic Programming goal is obtain an equation with the least length between the points [0,20] and [19,40]. Figure 6 shows the generated curves of first generation and the 12^{th} generation.

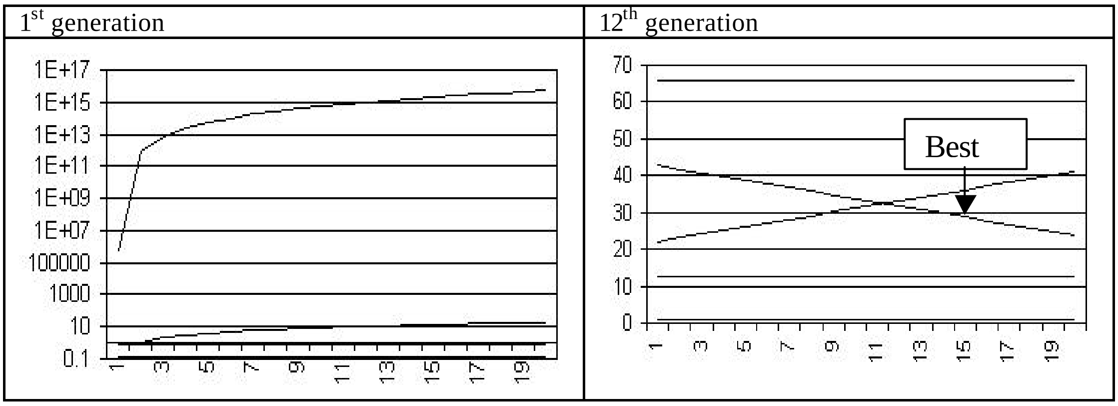


Fig. 6 Evolution of genetic programming.

The results show a very spread set of curves in first generation, with the convergence to 5 curves (one of them the best one). The formula obtained is:

(+ x 22.16)

or

y=x+22.16

The exact solution is y=x+20, the line equation between the two points specified.

We would enunciate the solution as: the equation y=a x+b solve the problem of less distance between two points without external potential, if we abstract the numerical values of the result. The interesting point is that GP found the model and its realization to the specific case. We would divide the output into two parts: the model (with we termed Cognitive Structure – CS, because it represents a learn result of the algorithm) and the realization (the CS + initial conditions is a specific solution).

Conclusion.

This paper establishes a relationship between the CV and GP as its numerical methods.

The central goal of CV is determining the functional that attend to some constraints solving fundamentals differential equations by analytical methods while GP try to obtain the solution applying genetic operators in tree coded chromosomes. The differential displacement in analytical solution assumes the format of a change into the functional structure through the application of crossover and mutation operators.

The action integral has its similar in the fitness function, with in the same way is obtained during all solution interval time.

The initial condition appears in both approaches defining a realisation of an intrinsic solution (we termed Cognitive Structure) that is holistic, i.e., complete and self-contained. It's a solution not for one single problem, but for a large class of similar problems.

Under this point of view, we would divide any problem solution as two different levels: one for the CS search, and the other to its adaptation to one realization.

A general overview is: the information available of the problem feeds GP software, with after some generations obtain the cognitive structure of the problem, or the best available at this moment with minimise the fitness function, i.e., the action for the system. This structure needs to be adapted to the real conditions of the system.

References:

1. BALL, WALTER WILLIAM ROUSE; "*A Short Account of the History of Mathematics*" Macmillan & Co.: London, 1888.
2. EULERO, LEONHARDO; "*Methodus inveniendi lineas curves maximi minimive proprietates gaudentes, sive solutio problematic isoperimetrici latissimo sensu accepti*" Lausannæ et Genevæ, 1744
3. EULERO, LEONHARDO; "*Dissertatio de principio minimœ actionis, una cum examine objectionum … Prof. Koenigii contra hoc principium factarum.*" Berolini, 1753
4. LA GRANGE,M. LOUIS DE; "*Recherches sur la methode de maximis et minimis*" from Melanges de Philosophie et de Mathematique de la Societe Royale de Turin pour les annees de 1760-1761
5. T.W.B. KIBBLE; "*Classical mechanics*" London. Longman. 1985
6. GOLDSTINE, HERMAN H; "A history of the calculus of variations from the 17th through the 19th century", ISBN 3540905219.v, Studies in the history of mathematics and physics sciences; New York; Springer, 1980.
7. J.L.LA GRANGE; "Mecanique analytique" – Paris 1811.
8. HOLLAND,J.H. "*Adaptation in natural and artificial systems: na introductory analysis with applications to biology, control and artificial intelligence.*" Cambridge: Cambridge press 1992 reedição 1975.
9. KOZA,J.R. "*Genetic programming: On the programming of computers by means of natural selection.*" Cambridge,Mass.: MIT Press, 1992.