International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 8 Issue 7 Ser. I ǁ 2020 ǁ PP. 17-42

Characterization Of Maxima Entropy Functions Using Calculus Of Variations

ANDRÉS PÉREZ
A. Pérez, Escuela De Matemática, Fac. Ciencias, Universidad Central De Venezuela

ABSTRACT. *Shannon definition of entropy is considered, and then Shannon formula is deducted based on the unexpected value of uncertainty function. Later, Euler-Lagrange equation is extended to an unbounded interval, and using this extension we can prove some results for characterizing maxima entropy functions and some examples are shown.*

Date of Submission: 03-08-2020 Date of acceptance: 18-08-2020

I. INTRODUCTION

In origins, entropy was matched to thermodynamics second principle or also named energy degradation principle, as heat is considered a form more down or "degraded" of energy, understanding by degradation, when a form of energy passed to another form, but it is impossible being accompanied by the complete inversion process.

During the 19th century, first works about entropy emerged by Carnot, Clausius and then Boltzmann. Boltzmann achieved a deep law interpretation, as he showed the link between entropy and the distribution of the energetic of a system in thermodynamic equilibrium, called the Boltzmann probability distribution. Hence, he concluded that the entropy of an isolated system is linked to the probability of its current state [4].

A simple definition of entropy is, the number of distinct forms in that we could dispose of the particles a system, mathematically is the proportional quantity to a logarithm of a number of possible distributions of the states of the system. Formally, Claude Elwood Shannon (1916–2001) electric engineer and mathematician (he is remembered as information theory father), in his paper "A Mathematical Theory of Communication" ([24]), defined entropy as a function that permits a measure of information quantity associated to a random process, and there he established that to major probability produced, less information contributed.

Shannon took a random variable X and from it defines a new variable$I[X] = -\log_2 p_i$ namedquantity of information, where pi is the probability of i-th event. Then, he defines entropy asexpected value of information quantity, that is to say,

$$H[X] = E[I[X]] = -\sum_{i=1}^{n} p_i \log_2 p_i$$

If in Shannon entropy we consider X as a random variable, which takes values on $\mathbb{R}$, and f is thedensity function associated to X, Shannon entropy could be defined by the following functional

$$H[f] = \int_{-\infty}^{\infty} f(x) \log f(x)\, dx$$

In the present paper, I want to determine conditions under which it is possible to guarantee existenceand unicity of density functions, that maximize the functional gave above, submitting it to diverserestrictions. For it, I do use technics of the calculus of variations.

The paper is organized as follows: In Section 2, I give some preliminary results and definitions,in Section 3, I obtain some results about extension of Euler-Lagrange equation to a unboundedinterval, in Section 4, I prove some results about Euler-Lagrange equation with functionals, inSection 5 I prove general result for extreme values and finally, in section 6 it achieves to characterizesome maxima entropy functions.

II. PRELIMINARES

The basic concept of entropy in Theory of Information is related to the uncertainty that existsin any experiment or random signal. Inclusive it is possible to visualize it as the noise quantity orsystems disorder. In this way, we can talk about information quantity contained in a signal. Often,when we talk about entropy from the theory of information point of view, we find that this entropyis named Shannon entropy in honor to Claude E. Shannon.

2.1. **Shannon Entropy.** Shannon offers an entropy definition that satisfies the following statements:

- The information measure must be proportional (continuous). That is to say a small change in the probabilities of the occurrence of any signal element, must produce a small change in its entropy.
- If all signal elements are equiprobable at the moment to appear, then the entropy will be maximum.

In an intuitive way, the entropy can be viewed as a state function S of the system, that is asystem state function. Usually, when you link the entropy to a state function, this means thatyou can characterize the system by S. For example, you can establish if a system will go on aspontaneous (irreversible) change analyzing its entropy. Electing of this unexpected function mustsatisfy some basic axioms, those are:

Axiom 1. $S(1) = 0$

Namely, if an event is sure, the unexpected value is zero.

Axiom 2. *If* $p > q$,*then*$S(p) < S(q)$

Hence, S is a decreasing strictly function. Then, when an event is more improbable than other,the unexpected must be bigger.

Axiom 3. $S(p)$*is a continuous function in* p.

Little changes in probability, determines little variations in unexpected value. Now, let $E$and Fbe two events with probabilities $p$and $q$respectively. Hence, surprise for both $E$and $F$is given by$S(pq)$, but unexpected value from $E$is $S(p)$, and unexpected value from $F$is $S(pq) - S(p)$. As $E$and F are independent events, the last surprise is just $S(q)$. Let's see,

Axiom 4. $S(pq) = S(p) + S(q)$,*for*p, q *in* $(0, 1)$.

Using these axioms, we can find the functional form of $S(p)$. We can remark that using Axiom4 in an inductively way, we obtain

$$S(px) = xS(p)\,, x \in \mathbb{Q}^+$$

Consider a sequence $\{q_n\}$such that converges to q, and we take $q = p_x$, for any $x > 0$and arbitrary.Considering $q_n = p^{\lfloor nx \rfloor/n}$, and the continuity of function S, we can obtain that

$$S(px) = xS(p)\,, x \in \mathbb{R}^+$$

Furthermore, if $x = 0$by Axiom 1, we have to $S(p^0) = S(1) = 0 = 0.S(p)$. Considering $0 < p \leq 1$and any $a > 0$, we define$x = -\log_a p$, such as $p = \left(\frac{1}{a}\right)^x$
, then

$$S(p) = S\left(\frac{1}{a}\right)^x = xS\left(\frac{1}{a}\right) = -C \log_a p$$

where$C = S\left(\frac{1}{a}\right) > S(1) = 0$(by Axioms 1 and 2). Thus, these axioms determine that th unexpected response function must be

$$S(p) = -S\left(\frac{1}{a}\right)\log_a p$$

If we denote by $S_a(p)$ to the function that results of choosing$S\left(\frac{1}{a}\right) = 1$, we have that

$$S_a(p) = -\log_a p$$

For a discrete random variable X, which has probability $p_i$when it takes the value xi_i, the entropyis calculated using the probability of the mass function of the random variable, that is

$$H_a(X) = E(S_a(p(X))) = -\sum_i p(x_i)\log_a p(x_i)$$

Comparisons between entropies its our interest. Thus, if X and Y are random variables, then

$$\frac{H_a(X)}{H_a(Y)} = \frac{H_b(X)}{H_b(Y)}, \text{for} a, b > 0$$

Remark 5. Another way to see this constant factor is, thinking in the relation with measure unit,that is to say, the logarithm base just specifies the units in which we are measuring the entropy.Then, we can establish that measure unit is: *bit*, *nat*, *dec*, etc., depending it is $2, e, 10$, etc.

Definition 6. Let $x_i$be the i-th event of a random variable X, then the information quantity thatthis event supply is determined by

$$I_i = -\log_2 p(x_i)$$

where$p(x_i)$ is the probability of this event.

Definition 7. Let X be a random variable, which takes values $x_1, x_2, \dots, x_n$with probabilities$p_1, p_2, \dots, p_n$that apport information quantities $I_1, I_2, \dots, I_n$. We established the random variable$I[X]$ as the information quantity associated to X, and we consider its expected value $E[I[X]]$ thatwe name $H[X]$. This value is named *Shannon entropy*.

Hence, the real number $H[X]$, is the expected value of information quantity that we will obtainin an experiment result expressed by said random variable. All in all, Shannon entropy analyticexpression is given by

$$H[X] = E[I[X]] = \sum_{i=1}^{n} p(x_i)I(x_i) \Longrightarrow H[X] = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$$

In fact, if we consider X as a random variable that it takes values in $\mathbb{R}$, with probability density$f : \mathbb{R} \longrightarrow \mathbb{R}$, we have that the Shannon entropy of X is just

$$\int_{-\infty}^{\infty} f(x)\log f(x)\, dx$$

with the convention that $0 \log 0 = 0$.

III. EULER-LAGRANGE EQUATION IN A BOUNDED INTERVAL

In this section, we focus in all the needed requirements to analyze the fundamental problem ofthe calculus of variations. First, we will aboard the problem in the finite-dimensional case, thenfrom this information, we will consider the infinite-dimensional case.

3.1. **Finite-dimensional case.** Let En be a n-dimensional Euclidean space and let $D \subset E_n$bean open set. Consider $\boldsymbol{x} \in D$, such that: $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$. Let $f$be a function defined in D.We know that if $D$is a closed and bounded set, and if f is continuous, then Bolzano-Weierstrasstheorem guarantees the existence of extremes values of$f$in D.

We consider a function $f$defined in Ω, such that $\boldsymbol{x_0} \in \Omega$is an extreme point ($\Omega$is a neighborhoodof $\boldsymbol{x_0}$). We suppose that f is differentiable in Ω, so $\frac{\partial f}{\partial x_i}$exist for all$i = 1, 2, \dots, n$. If $\boldsymbol{x_0}$is an extremepoint of f, then it must satisfy that $\frac{\partial f}{\partial x_i}$ $(\boldsymbol{x_0})$ = 0, for all $i = 1, 2, \dots, n$.This result, is known as*necessary condition of extreme*.

To points that satisfying the necessary condition of extreme, are known as *critical points*, andpoints $\boldsymbol{x_0}$such that df($\boldsymbol{x_0}$) = 0, are known as stationary points of the function f. Last condition isequivalent to$\frac{\partial f}{\partial x_i}$ ($\boldsymbol{x_0}$) = 0, for all $i = 1, 2, \ldots, n$.

Remark 8. The existence of a critical point does not a guarantee the existence of an extreme. Tosolve this problem, let's see the sufficient conditions for strict extremes.

Definition 9. We say that the quadratic form

$$A(\boldsymbol{x}) = A(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{i,j} x_i x_j, \qquad a_{ij} = a_{ji}, i, j = 1, 2, \ldots, n$$

is*positive definite* if $A(\boldsymbol{x}) > 0$, for all $\boldsymbol{x} \in E^n$, $\boldsymbol{x} \neq 0$, and is null only when $\boldsymbol{x} = 0$, that is to say,when $x_i = 0$for all $i = 1, 2, \ldots, n$. (Analogous for negative definite).

Let $f$be a function of class C^2 in $\Omega$and let $\boldsymbol{x^0} \in \Omega$a stationary point of f. If the quadratic form

$$A(dx_1, dx_2, \ldots, dx_n) = \sum_{i,j=1}^{n} \frac{\partial^2 f(\boldsymbol{x^0})}{\partial x_i \partial x_j} dx_i dx_j$$

is positive definite, the critical point $\boldsymbol{x^0}$is a strict minimum. (Analogous for strict maximum). Ifthe quadratic form is undefined, then$\boldsymbol{x^0}$is not an extreme point of f. These conditions are knownas *sufficient conditions of strict extreme*.

3.2. **Conditioned extremes.** Let $z = f(x_1, x_2, \ldots, x_n)$ be a $n$variables function defined in anopen set $D \subset E^n$. Suppose that the $n$variables are linked by $m$complementary conditions, where$m < n$and, $c_i \in \mathbb{R}$ for $i = 1, \ldots, m$, hence

$$\begin{cases} \varphi_1(x_1, x_2, \ldots, x_n) = c_1 \\ \vdots \\ \varphi_m(x_1, x_2, \ldots, x_n) = c_m \end{cases}$$

These equations are named *link equations*. Let $\boldsymbol{x^0} = (x_1^0, x_2^0, \ldots, x_n^0)$ be an internal point of D. We say that f has a conditioned maximum (conditioned minimum respectively) in $\boldsymbol{x^0}$if the inequality

$$f(\boldsymbol{x}) \leq f(\boldsymbol{x^0}) \text{or} f(\boldsymbol{x}) \geq f(\boldsymbol{x^0})$$

is satisfied into some neighborhood of $\boldsymbol{x^0}$, always that $\boldsymbol{x^0}$and $\boldsymbol{x}$verify the link equations.

We consider the case $n = 2$to illustrate the process. Suppose that we need to find the extremevalue of $z = f(x, y)$ subject to the restriction $\varphi(x, y) = c$. By the implicit function theorem,the link equation $\varphi(x, y) = c$, determines to $y$as a differentiable function univocally definite, so$y = \psi(x)$. Hence$z = f(x, \psi(x)) = F(x)$, where the unconditioned extreme of F, will be just, theextreme sought of function$f$with the respective link function. To solve the originalproblema for
$z = f(x_1, x_2, \ldots, x_n)$subject to the link equations, is enough to combine $\nabla f(\boldsymbol{x}) = 0$with the implicitfunction theorem, obtaining as consequence Lagrange multipliers method.

3.3. **Lagrange multipliers method.** To illustrate the Lagrange multipliers method (the proofof this theorem will be accomplished later using more general results), we consider a couple ofconditions which must be satisfied to find extremes of given functions, from these conditions wecould give a scheme of the process.

(1) Partial derivatives of functions $f(x_1, x_2, \ldots, x_n)$and $\varphi_i(x_1, x_2, \ldots, x_n)$with $i = 1, 2, \ldots, m$, are continuous in D.

(2) The range of matrix$\left(\frac{\partial \varphi_i}{\partial x_j}\right)$, with $i = 1, \ldots, m$and $j = 1, \ldots, n$is equal to $m < n$in everypoint of$D$.

Consider a new function

$$\Phi(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{k=1}^{n} \lambda_k \varphi_k(\boldsymbol{x})$$

known as *Lagrange function*, where λ_k's are undetermined constant factors (Lagrange multipliers).
Following, we analyze the unconditioned extremes of function Φ, that is to say, we formed thesystem of equations (necessary conditions of extreme)

$$\frac{\partial \Phi}{\partial x_1} = 0, \frac{\partial \Phi}{\partial x_2} = 0, \ldots, \frac{\partial \Phi}{\partial x_n} = 0 \tag{3.1}$$

and from both this system and$m$linked equations, we determine both parameter values$\lambda_1, \lambda_2, \ldots, , \lambda_m$and coordinates $(x_1, x_2, \ldots, x_n)$ of possible extreme points, that is $\frac{\partial \Phi}{\partial x_j} = 0$, with $j = 1, \ldots, m$. Toanalyze the stationary point $\boldsymbol{x^0}$as a conditioned extreme of Φwe must consider the quadratic form

$$B(dx_1, dx_2, \ldots, dx_{n-m}) = \sum_{i,j=1}^{n-m} b_{ij}\, dx_i dx_j$$

and from the possible definitions of this quadratic form, we will find the possible strict extremes ofΦ.

IV. GENERAL CASE

In this section, we do a formal entry to Calculus of variations, having as a first variant $C[a, b]$which is an infinite-dimensional space. Hence, we work with functions of functions, that is tosay, *functionals*. It is important to highlight that our functionals are defined over continuou trajectories, in consequence, the partial derivatives are not to be taken as the above case, withrespect to an independent variable of Euclidean space, but with respect to a regular trajectory.

4.1. **Functionals extremes.** Suppose that $E$is a linear manifold into $C[a, b]$. Analogously to thefinite-dimensional case, maximums, minimums, extremes, are defined.

Definition 10. Let $(E, \|\cdot\|_E)$ be a normed vectorial space and consider $M \subset E$. A *functional*, isa function $F : M \longrightarrow \mathbb{R}$. The set M formed by functions $\psi$where the functional is defined, it isnamed *definition's field of functional*.

On the other hand, It is called *variation* or *increment* $\Delta\psi$of the plot $\psi(t)$of the functiona $F(\psi)$, to the difference between the functions$\psi(t)$ and $\varphi(t)$ that belong to subset M, namely

$$\Delta\psi = \psi(t) - \varphi(t)$$

Suppose then that the curves $\psi(t)$ and $\varphi(t)$ are defined in the interval $[a, b]$. From increment offunctional definition, we can study the proximity between curves of a certain family, that is to say,we will say that $\psi(t)$ and $\varphi(t)$ have approach order zero, if over the interval $[a, b]$, the magnitude$\|\psi(t) - \varphi(t)\|_E$is very small from a geometric point of view.
Suppose that the curves $\psi, \varphi \in C^1[a, b]$ at least. Then, $\psi(t)$ and $\varphi(t)$ have approach order one,if the magnitudes

$$\|\psi(t) - \varphi(t)\|_E \text{ and } \|\psi'(t) - \varphi'(t)\|_E \tag{4.1}$$

are small over$[a, b]$. It is called *order k distance*,$k = 0, 1$to the major of expressions given in (4.1). Then, werepresent this distance as follow

$$\rho_k = \rho_k[\psi(t), \varphi(t)] = \max_{0 \le i \le k} \max_{a \le t \le b} \left\| \psi^{(i)}(t) - \varphi^{(i)}(t) \right\|_E, \quad k = 0, 1$$

It is called ε *−neighborhood*of order $k$of curve$\psi(t), a \le t \le b$to the set of all curves $\varphi(t)$ suchthat their distances of k-*th* order to curve $\psi(t)$ are less than ε. Now, the ε *−neighborhood*of orderzero is named *strong* ε *−neighborhood*of $\psi(t)$, and the ε *−neighborhood*of first order is called *weak*ε *−neighborhood*of $\psi(t)$.

Definition 11. A functional $F(\psi)$ defined into the set M of functions $\psi(t)$ is called continuousin $\varphi(t)$ in the proximity of k-*th* order sense, if for all $\varepsilon > 0$exist $\eta > 0$such that it verifies that$|F(\psi) - F(\varphi)| < \varepsilon$, for all functions $\psi \in M$ that they satisfy $\rho_k[\psi(t), \varphi(t)] < \eta$.Suppose that we have a functional defined over M, the magnitude

$$\Delta F = \Delta F(\psi) = F(\psi + \Delta\psi) - F(\psi)$$

is named *increment or variation* of functional $F(\psi)$, corresponding to increment $\Delta\psi$of the plot.Another form to measure the increment of a functional, is considering the derivative of the functional$F$in the $\Delta\psi$direction over the point ψ_0, this is

$$\Delta F = \frac{\partial}{\partial \alpha} F(\psi_0 + \alpha \Delta\psi) \Big|_{\alpha=0}$$

Definition 12. A functional Freaches its *local maximum* on the curve φ, if values that $F(\varphi)$ takesover any curve close to φ(in some order), are not major that $F(\varphi)$, such that

$$\Delta F = F(\psi(t)) - F(\varphi(t)) \le 0 \qquad (4.2)$$

If $\Delta F = 0$, only when $\psi(t) = \varphi(t)$, we say that the functional reaches a *strict maximum* on thecurve φ. Analogous for minimum.

On the other hand, we say that Freaches a *strong local maximum* on ψ, if it satisfies (4.2) for allcurves that belong to ε *−neighborhood*of order zero on the curve φ. F reaches a *weak local máximum* if it satisfies the same inequality but in a ε *−neighborhood*of order one. They are analogous for theminimum.

Remark 13. It is important to note that all strong extreme is at the same time a weak extreme.The functional extreme that refers to the total of functions that are defined over the same curve, iscalled *absolute extreme*, and hence all absolute extreme is at the same time a strong local extreme.

We need to evaluate the conditions for the existence of the extremes of the functionals, for it, wewill observe the derivative definition over normed spaces and see that is the same that Euclideanspaces, with the exception of the requirement that the approximation must be made by a linearfunctional continuous.

Definition 14. Let $(E, \|\cdot\|_E)$ and $(H, \|\cdot\|_H)$be normed spaces. The functional $F : E \longrightarrow H$ isdifferentiable in $x \in E$, if and only if, exist a linear functional continuous $L : E \longrightarrow H$, such that

$$\lim_{h \to 0} \frac{F(x + h) - F(x) - L(h)}{\|h\|_E} = 0 \qquad (4.3)$$

If $F : E \longrightarrow H$ is differentiable in $x \in E$, then the continuous linear functional that satisfies (4.3), iscalled differential of Fin $x$and it is denoted by $dF_x : E \longrightarrow H$.

Remark 15. In the finite-dimensional case, that is $E = \mathbb{R}^n$and $H = \mathbb{R}^m$this differential coincideswith the matrix formed by the partial derivatives, expressed on the canonical basis, but ou main concern the correspondences between infinite-dimensional spaces, where there is not a matrixrepresentation.

Theorem 16. *Let* $L : E \longrightarrow H$*be a linear functional, where* $(E, \|\cdot\|_E)$*and*$(H, \|\cdot\|_H)$*are normedvector spaces. Then, the following statements over* L*are equivalents*

1) *There is a number* $c > 0$, *such that* $\|L(v)\|\_H \le c\|v\|_E$, *for all* $v \in E$.
2) L*is continuous in everywhere.*

3) $\quad$ *$L$is continuous in $0 \in E$.*

Proof. Suppose that the first condition is satisfied. Then, given $x_0 \in E$and $\varepsilon > 0$, we have that

$$\|x - x_0\|_E < \frac{\varepsilon}{c} \Rightarrow \|L(x) - L(x_0)\|_H = \|L(x - x_0)\|_H \leq c\|x - x_0\|_E < \varepsilon$$

then, *(1)* implies *(2)*. Obviously, *(2)* implies *(3)*. Then *(3)* implies *(1)*.
As $L$is continuous in $0 \in E$, we can choose $\delta > 0$such $that\, \|x\|_E \leq \delta \Rightarrow \|L(x)\|_H < 1$. Hence,given $v \neq 0 \in E$, we obtain that

$$\begin{aligned} \|L(x)\|_H &= L\left(\frac{\|v\|_E}{\delta}.\frac{\delta}{\|v\|_E}v\right) \\ &= \frac{\|v\|_E}{\delta}L\left(\frac{\delta}{\|v\|_E}v\right) \\ &\leq \frac{\|v\|_E}{\delta} \end{aligned}$$

taking$x = \frac{\delta}{\|v\|_E}v$ and $c = \frac{1}{\delta}$, we get what we want ■

The following theorem is given without proofbecause it is analogous to the finite-dimensionalcase.

Theorem 17. *Let $U$and Vbe open subsets of the normed spaces $(E, \|\cdot\|_E)$and $(H, \|\cdot\|_H)$respectively. If the functionals $F : U \longrightarrow H$and $g : U \longrightarrow G$(a third normed vector space), aredifferentiable in $x \in U$and $F(x) \in V$respectively, then the composition $h = g \circ F$is differentiable in$x$and it satisfies that*

$$dh_x = dg_{F(x)} \circ dF_x$$

Consider a subset M of normed vector space E, the tangent space of M in $x \in M$, is the setof all vectors $v \in E$, for which there is a differentiable trajectory $\varphi : \mathbb{R} \longrightarrow M$, such that $\varphi(0) = x$ and $\varphi'(0) = v$. It is important to highlight that, if M is open, then $TM_x = E$for all $x \in M$.

Theorem 18. *Let $(E, \|\cdot\|_E)$a normed space and $F : E \longrightarrow \mathbb{R}$a differentiable functional and let$M \subset E$and consider $x \in M$. If $F$is differentiable in $x$and $F|_M$reaches a minimum in x, then*

$$dF_x|_{TM_x} = 0$$

Proof. Given $v \in TM_x$let $\varphi : \mathbb{R} \longrightarrow M$ be a differentiable trajectory on E, such that $\varphi(0) = x$and$\varphi'(0) = v$. Then, the function $g : \mathbb{R} \longrightarrow \mathbb{R}$, defined by $g = F \circ\varphi$has a local minimum in 0, suchthat $g'(0) = 0$. Hence, using the chain rule we get

$$0 = g^{'}(0) = dg_0(1) = dF_{\varphi(0)}\big(d\varphi_0(1)\big) = dF_x\Big(\varphi^{'}(0)\Big) = dF_x(v) ■$$

Remark 19. If M is an open set, then directly $dF_x x = 0$. If M is any set, for finding the extremes isnecessary to find the tangent space to M.

Theorem 20 (Lagrange)**.** *Let $D \subset E^n$be an open set and let $f, \varphi_k: D \longrightarrow \mathbb{R}$,$k = 1,\dots,m$ smooth functions with $m < n$. Let$x_0 \in D$, such that $\varphi_k(x_0) = c_k$, with $c_k =$ constant, for each$k = 1,\dots,m$and let $S = \{x \in D: \varphi_k(x) = c_k\}$. Suppose that $\nabla\varphi_k(x_0)$ are linearly independent for$k = 1,\dots,m$. If $f|_S$has a maximum or a minimum in S, reached inx_0, then there real numbers $\lambda_k, k = 1,\dots,m$such that*

$$\nabla f(x_0) = \sum_{k=1}^{n} \nabla\varphi_k(x_0)$$

Proof. Let $G : D \longrightarrow \mathbb{R}^m$, the function defined by $G(x) = (\varphi_1(x),\dots,\varphi_m(x))$, then $G(x_0) = (c_1,\dots, c_m)$ and $S = \{x \in D : G(x) = C\}$, where $C = (c_1,\dots, c_m)$. The tangent space to Sinx_0, isthe intersection between tangent hyperplane to each$n$-1dimensional manifolds defined by$\{x \in D : \varphi_k(x) = c_k\}$, $k = 1,\dots,m$.

Hence, S_{x^0} the tangent space to $S$is the set's orthogonal complement $\{\nabla\varphi_1(x^0),\dots,\nabla\varphi_m(x^0)\}$. If$f|_S$reaches a maximum in x^0, then $df_{x^0}|_{S_{x^0}} = 0 \Longrightarrow \nabla f(x^0)\cdot x = 0$, as long as $x \in S_{x^0}$, that is to say$x \in \{\nabla\varphi_1(x^0),\dots,\nabla\varphi_m(x^0)\}^{\perp}$, in consequence, $\nabla f(x^0) \in \{\{\nabla\varphi_1(x^0),\dots,\nabla\varphi_m(x^0)\}^{\perp}\}^{\perp}$and this isjustly the manifold generated by $\nabla\varphi_1(x^0),\dots,\nabla\varphi_m(x^0)$, then exist real numbers $\lambda_1,\dots,\lambda_m$ suchthat

$$\nabla f(x^0) = \sum_{k=1}^{m} \lambda_k \nabla\varphi_k(x^0) \blacksquare$$

From this moment, we will consider $E = C^1[a,b]$, that is to say we will consider the normedspace $(C^1[a,b], \|\cdot\|_{C^1})$, where

$$\|\varphi\|_{C^1} = \max_{a\le t\le b}|\varphi(t)| + \max_{a\le t\le b}|\varphi'(t)|$$

Now, let $M$the subset composed by the functions $\psi \in C^1[a,b]$ that satisfy edge conditions $\psi(a) = \alpha$ and $\psi(b) = \beta$. If $F$is differentiable on $\varphi \in M$ and $F|_M$ has a local extreme on φ, then by theorem18 we have

$$dF\varphi\,|_{TM_\varphi} = 0 \tag{4.4}$$

Hence, we say that $\varphi \in M$ is an extremal function of Fover M, if it satisfies the necessary condition(4.4). We need to calculate explicitly $dF_\varphi$and determining the tangent space TM_φ.Last problem is easy to solve, if we consider a fixed trajectory $\varphi_0 \in M$, and given $\varphi \in M$ thedifference $\varphi - \varphi_0$belongs to $C_0^1[a,b]$, that is

$$C_0^1[a,b] = \{\psi \in C^1[a,b] : \psi(a) = \psi(b) = 0\}$$

Conversely, if $\psi \in C_0^1[a,b]$ we have that $\varphi_0 + \psi \in M$. Hence, M is a hyperplane on $C^1[a,b]$,more exactly translation by $\varphi_0$in $C_0^1[a,b]$, but hyperplane's tangent plane in any point, is just thesubspace of which is a translation. Finally,

$$TM_\varphi = C_0^1[a,b]\blacksquare$$

To calculate dF_φ, we need to prove the following theorem.

Theorem 21. *Let* $F : C^1[a,b] \longrightarrow \mathbb{R}$, *defined by*

$$F(\varphi) = \int_a^b f(\varphi(t),\varphi'(t),t)dt$$

with$f : \mathbb{R}^3 \longrightarrow \mathbb{R}$, *a* C^2*class function. Then,* F*is differentiable and*

$$dF_\varphi(h) = \int_a^b \left[\frac{\partial f}{\partial x}(\varphi(t),\varphi'(t),t)h(t) + \frac{\partial f}{\partial y}(\varphi(t),\varphi'(t),t)h(t)h'(t)\right]dt \tag{4.5}$$

for all $\varphi, h \in C^1[a,b]$.

Proof. As $F$is differentiable on $\varphi \in C^1[a,b]$, $dF_\varphi(h)$ must be the linear part of the difference$F(\varphi+h) - F(\varphi)$. We use Taylor expansion of second order of $f$over the point $(\varphi(t),\varphi'(t),t) \in \mathbb{R}^3$, let's see

$$\begin{aligned} f(\varphi(t) + h(t),\varphi'(t) + h'(t),t) - f(\varphi(t),\varphi'(t),t) &= \frac{\partial f}{\partial x}(\varphi(t),\varphi'(t),t)h(t) \qquad (4.6)\\ &+ \frac{\partial f}{\partial y}(\varphi(t),\varphi'(t),t)h'(t) + r(h(t)) \end{aligned}$$

Where

$$r\big(h(t)\big) = \frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}\big(\xi(t)\big)\big(h(t)\big)^2 + 2\frac{\partial^2 f}{\partial x \partial y}\big(\xi(t)\big)h(t)h'(t) + \frac{\partial^2 f}{\partial y^2}\big(\xi(t)\big)h(t)h'(t)\right]$$

]

for some point $\xi(t)$, over the line segment in $\mathbb{R}^3$from the point $(\varphi(t), \varphi'(t), t)$ to the point$(\varphi(t) + h(t), \varphi'(t) + h'(t), t)$. If B is a ball with enough radius such that contains the continuous trajectory image $t \mapsto (\varphi(t), \varphi'(t), t)$, for $t \in [a, b]$, and $P$is the maximum of absolute valuesof partial derivatives of second order of $f$on B points, then we obtain

$$|r(h(t))| \leq \frac{P}{2}(|h(t)| + |h'(t)|)^2 \quad (4.7)$$

for all $t \in [a, b]$, if $||h||_{C^1}$ is small enough.

From (4.6), we have left that $F(\varphi + h) - F(\varphi) = L(h) + R(h)$, where

$$L(h) = \int_a^b \left[\frac{\partial f}{\partial x}(\varphi(t), \varphi'(t), t)h(t) + \frac{\partial f}{\partial y}(\varphi(t), \varphi'(t), t)h'(t)\right] dt$$

and

$$R(h) = \int_a^b r(h(t))dt$$

We need to prove that $dF_\varphi(h) = L(h)$, where $L : C^1[a, b] \longrightarrow \mathbb{R}$ is obviously continuous and linear.Now we need to verify that

$$\lim_{||h||_{C^1}\to 0} \frac{|R(h)|}{||h||_{C^1}} = 0 \quad (4.8)$$

where

$$||h||_{C^1} = \max_{a\leq t\leq b}\{|h(t)|, |h'(t)|\}$$

From (4.7), immediately we obtain

$$|R(h)| \leq \int_a^b |r\big(h(t)\big)|dt \leq \int_a^b \frac{P}{2}(2||h||_{C^1})^2 dt = 2P(b-a)(||h||_{C^1})^2$$

and this prove (4.8). ■

Remark 22. This theorem shows that $b - a < \infty$is necessary to Fwill be differentiable.

Corollary 23. *Let $F : C^1[a, b] \longrightarrow \mathbb{R}$be defined by*

$$F(\varphi) = \int_a^b f(\varphi(t), \varphi'(t), t)dt$$

with$f : \mathbb{R}^3 \longrightarrow \mathbb{R}$, a C^2class function. If $\varphi$is a C^2class functionin $[a, b]$and$h \in C_0^1[a,b]$, then

$$dF\varphi(h) = \int_a^b \left[\frac{\partial f}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{\partial f}{\partial y}(\varphi(t), \varphi'(t), t)\right] h(t)dt \quad (4.9)$$

Proof. Given that $\varphi$is *a C^2class function*, then $\frac{\partial f}{\partial y}(\varphi(t), \varphi'(t), t)$ is a C^1class function in $[a, b]$.Integrating by parts, we obtain

$$\int_a^b \left[\frac{\partial f}{\partial y}(\varphi(t),\varphi'(t),t)\right] h'(t)dt \quad = \quad -\int_a^b \frac{d}{dt}\left[\frac{\partial f}{\partial y}(\varphi(t),\varphi'(t),t)\right] h(t)dt$$

as$h(a) = 0 = h(b)$, therefore (4.9) could be obtained directly from (4.5).■

From this moment, M_0denotes a linear manifold into the set

$$\{\psi \in C^1[a,b] : \psi(a) = 0 = \psi(b)\}.$$

Lemma 24. *If $M_0$is dense in $L^1[a,b]$and, $\varphi : [a,b] \longrightarrow \mathbb{R}$is a continuous function such that*

$$\int_a^b \int \varphi(t)h(t)dt = 0$$

for any $h \in M_0$, then $\varphi$is identically zero in $[a,b]$.

Proof. First we consider when $M_0 = C_0^1[a,b]$. Suppose that $\varphi(t_0) \neq 0$, for some $t_0 \in [a,b]$. Then, $\varphi$is not null in an interval that contains to t_0, because $\varphi$is continuous. If $\varphi(t) > 0$ for$t \in [t_1,t_2] \subset [a,b]$ and we define the function $h$as follows

$$h(t) = \begin{cases}(t-t_1)^2(t-t_2)^2, & t \in [t_1,t_2] \\ 0, & \text{another case}\end{cases}$$

then$h \in C_0^1[a,b]$ in addition

$$\int_a^b \varphi(t)h(t)dt = \int_{t_1}^{t_2} \varphi(t)\,(t-t_1)^2(t-t_2)^2 > 0$$

which is a contradiction, hence $\varphi \equiv 0$in $[a,b]$. As $M_0$is dense in $L^1[a,b]$, there exist a sequenceof functions $\{h_n\} \subset M_0$such that $h_n \to \varphi$into $L^1[a,b]$, then

$$\int \varphi h_n \longrightarrow \int \varphi^2$$

which implies that $\varphi = 0$almost everywhere, because$\varphi h_n = 0$ for all n.■

4.2. **Euler equation.** Suppose that there is a function $f(x,y,t)$ which has partial derivativescontinuous until second order, with respect to all its plots. Between all possible functions $\varphi(t)$ withcontinuous derivative, we must seek the function that offers the weak extreme to the functional

$$F(\psi) = \int_a^b f(\psi(t),\psi'(t),t)dt \tag{4.10}$$

subject to the boundary conditions

$$\psi(a) = \delta \text{ and } \psi(b) = \gamma \tag{4.11}$$

Again, it is necessary to establish conditions under which is possible to seek extremal functionsof the functional defined in (4.10). Suppose that $\quad_1$is contained in the set

$$\{\psi \in C^1[a,b] : \psi(a) = \delta \text{ and} \psi(b) = \gamma\}$$

and there is a linear manifold $c = C_0^1[a,b]$, such that $M_0 + M_1 \subset M_1$and $M_0$is dense in $L^1[a,b]$.

Theorem 25 (Necessary condition)**.** *Under the above conditions, for the functional given in* (4.10),*defined by the set of all functions $\psi \in M_1$reaches its maximum value, is necessary that the function$\psi(t)$verifies the Euler equation, that is*

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial y} = 0 \tag{4.12}$$

Proof*.* Consider $h \in M_0$. By corollary 23,

$$dF\varphi(h) = \int_a^b \left[\frac{\partial f}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial f}{\partial y}(\varphi(t), \varphi'(t), t)\right] h(t)dt$$

and as $TM_1 \supset M_0$, by theorem 18 the above differential is null, for all $h \in M_0$. Now, by lemma 24 we obtain

$$\frac{\partial f}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial f}{\partial y}(\varphi(t), \varphi'(t), t) = 0 \qquad \blacksquare$$

Remark 26. The integral curves that satisfy (4.12) are called *Lagrange curves*, because of this, theequation is usually called *Euler-Lagrange equation.*Developing (4.12), we obtain

$$\psi''(t)f_{yy}(\psi(t), \psi'(t), t) + \psi'(t)f_{yy}(\psi(t), \psi'(t), t) + f_{ty}(\psi(t), \psi'(t), t) - f_x(\psi(t), \psi'(t), t) = 0 \qquad (4.14)$$

which represents a second order differential equation, and its general solution has two arbitrary constants,whose values will be determined from boundary conditions above mentioned. The developgiven in (4.14) is possible, if the following theorem is considered.

Theorem 27. *Let $\psi(t)$be a solution of Euler-Lagrange equation. If the function $f(\psi(t), \psi'(t), t)$has until second order continuous partial derivatives , then the function $\psi(t)$has continuous secondorder derivative in every point for which*

$$f_{yy}(\psi(t), \psi'(t), t) \neq 0$$

Proof. Consider the difference

$$\Delta f_y = f_y(x + \Delta x, y + \Delta y, t + \Delta t) - f(x, y, t) = \Delta t\bar{f}_{yt} + \Delta x\bar{f}_{yx} + \Delta y\bar{f}_{yy}$$

where “overlines” indicate that the corresponding derivatives are evaluated over certain meancurves. If this difference is divided by $\Delta t$and we take the limit when Δtapproach 0, we have

$$\bar{f}_{yt} + \frac{\Delta x}{\Delta t}\bar{f}_{yx} + \frac{\Delta y}{\Delta t}\bar{f}_{yy}$$

The last limit exist, as $f_y$has derivative with respect t, and according to Euler equation is equalto f_x. Additionally, f has continuous second order derivatives, then as Δtapproach 0, $\bar{f}_{yt}$approachf_{yt}. From both $y$existence and second derivative continuity f_{yy}, we have that the limit of secondterm exist when $\Delta t \to 0$, but the third term also has limit, because the limit of the sum of threeterms exist. When $\Delta t \to 0$, $f_{yy} \longrightarrow \bar{f}_{yy} \neq 0$, hence

$$\lim_{\Delta t \to 0}\frac{\Delta y}{\Delta t} = y(t)\text{exist}$$

All in all, from Euler equation we can find an expression to y'which is continuous when $f_{yy} \neq 0$.

$\blacksquare$

1.3. **Isoperimetric problem.** Around IX century, the princess Dido solved a problem that we can summarize as follows: “Finding between all closed curves with fixed length, the one that delimits the greater surface”.

If we consider two points $A(a, 0)$ and $B(b, 0)$ over x-axis where the distance between them isgiven, that is to say, $d(A, B) = l$, the problem to find a curve that maximize the area between itand x-axis will be:

To find a function $f(x)$ such that

$$I[f] = \int_a^b f(x)dx = max$$

subject to the restriction

$$G[f] = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx = l$$

where $f(a) = 0 = f(b)$.

From a general point of view, our problem consists in finding extremes of the functional givenin (4.10), subject to edge conditions given in (4.11) and with the additional restriction

$$G(\psi) = \int_a^b g(\psi(t), \psi'(t), t)dt = c\,, \quad c \in \mathbb{R} \tag{4.15}$$

with $f, g : \mathbb{R}^3 \longrightarrow \mathbb{R}, C^2[a, b]$ class functions. Let M be the hyperplane in $C^1[a, b]$ that containsthose C^1class functions $\psi : [a, b] \longrightarrow \mathbb{R}$, such that they satisfy (4.11), then our problem consistsin locate the local extremes of the functional F over the set $M \cap G^{-1}(c)$. The principal idea is togeneralize to the Lagrange multipliers method to infinite-dimension spaces.

We consider the following simpler case: Let both $F, G : \mathbb{R}^n \longrightarrow \mathbb{R}$ be functions of class C^1, suchthat $G(\mathbf{0}) = \mathbf{0}$and $\nabla G(\mathbf{0}) \neq \mathbf{0}$. If Fhas a local maximum or minimum in $\mathbf{0}$, subject to restriction $G(\boldsymbol{x}) = 0$, then there must a number λ, such that

$$\nabla F(\mathbf{0}) = \lambda \nabla G(\mathbf{0}) \tag{4.16}$$

Furthermore, the differentials $dF_{\mathbf{0}}, dG_{\mathbf{0}}: \mathbb{R}^n \longrightarrow \mathbb{R}$ given by

$$dF_{\mathbf{0}}(\boldsymbol{v}) = \nabla F(\mathbf{0}) \cdot \boldsymbol{v} \quad \text{and} dG_{\mathbf{0}}(\boldsymbol{v}) = \nabla G(\mathbf{0})$$

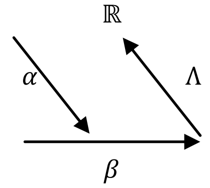
then(4.16), as

$$dF_{\mathbf{0}} = \Lambda \circ dG_{\mathbf{0}} \tag{4.17}$$

where $\Lambda : \mathbb{R} \longrightarrow \mathbb{R}$ is the linear function defined by $\Lambda(t) = \lambda t$.

Lemma 28. *Let αandβbe two linear functions to real values in vector space Esuch that $\boldsymbol{Ker}\,\alpha \supset \boldsymbol{Ker}\,\beta$and $\boldsymbol{Image}\,\beta = \mathbb{R}$.Then there exist $\lambda \in \mathbb{R}$, such that $\alpha = \lambda\beta$. That is to say, exist a linear function $\Lambda : \mathbb{R} \longrightarrow \mathbb{R}$, such*
that $\alpha = \Lambda \circ \beta$, that is the same to say that the diagram

E $\mathbb{R}$



commutes.

Proof. Given $t \in \mathbb{R}$, we take $x \in E$such that $\beta(x) = 1$and we define $\Lambda(t) = \alpha(x)$. To provethat $\Lambda$is well-defined, we must see if $y \in E$, such that $x \neq y$with the condition $\beta(y) = t$, the $\alpha(x) = \alpha(y)$. If $\beta(x) = \beta(y) = t$, then $x - y \in \boldsymbol{Ker}\,\beta \subset \boldsymbol{Ker}\,\alpha$, that is $\alpha(x - y) = 0$, by linearity $\alpha(x) = \alpha(y)$. Now, if $\beta(x) = s$and $\beta(y) = t$, then

$$\Lambda(as + bt) = \alpha(ax + by) = a\alpha(x) + b\alpha(y) = a\Lambda(s) + b\Lambda(t)$$

hence, $\Lambda$is linear. ■

Let $E, F$and Gbe three full vector normed spaces. Consider the differentiable application$f : E \times F \longrightarrow G$, then for each $a \in E$and each $b \in F$, the applications $\varphi : E \longrightarrow G$and $\psi : F \longrightarrow G$ defined by $\varphi(x) = f(x, b)$ and $\psi(y) = f(a, y)$ are differentiable in $a \in E$ and $b \in F$respectively.

Then the partial differentials $d_x f(a, b)$ and $d_y f(a, b)$, are defined by

$$d_x f(a, b) = d\varphi_a \text{ and } d_y f(a, b) = d\psi_b$$

Hence, $d_x f(a, b)$ is the differential of the application from Eon Gwhich was obtained from application$f : E \times F \longrightarrow G$, fixing $y = b$, and analogously for $dyf(a, b)$, but fixing $x = a$.

Definition 29. Let $E$and Fbe normed vector spaces. The application $g : E \longrightarrow F$is said to becontinuously differentiable or C^1class, if it is differentiable and $dg_x(v)$ is a continuous function of(x, v), that is to say the application $(x, y) \longmapsto dg_x(v)$ from $E \times E$to $F$is continuous.

Remark 30. The next result will be presented without proof, for more details you can consult [5].

Theorem 31 (Implicit function theorem over normed spaces)**.** *Let $E, F$and Gbe three full normedvector spaces and let $f : E \times F \longrightarrow G$be a C^1class application. Suppose that $f(a, b) = 0$andthat $d_y f(a, b) : F \dashrightarrow G$is an isomorphism. Then, there exist a neighborhood $U$of $a$in Eanda neighborhood $W$of (a, b)in $E \times F$and an application $\varphi : U \longrightarrow F$of class C^1, such that: If$(x, y) \in W$and $x \in U$, then $f(x, y) = 0$if only if $y = \varphi(x)$.*

Theorem 32. *Let $F$and Gbe C^1class functions to real values, defined in a full normed vectorspace $E$with $G(\mathbf{0}) = \mathbf{0}$and $dG_{\mathbf{0}} \neq \mathbf{0}$. If $F : E \dashrightarrow \mathbb{R}$has a local extreme in $\mathbf{0}$, subject to restriction$G(x) = \mathbf{0}$, then there is a linear function$\Lambda : \mathbb{R} \longrightarrow \mathbb{R}$such that satisfies the equation $dF_{\mathbf{0}} = \Lambda \circ dG_{\mathbf{0}}$.*

Proof. We can use the lemma 28 with $\alpha = dF_{\mathbf{0}}$and $\beta = dG_{\mathbf{0}}$, always that$\boldsymbol{Ker}\, dF_{\mathbf{0}}. \supset \boldsymbol{Ker}\, dG_{\mathbf{0}}$. For it we suppose that given $v \in \boldsymbol{Ker}\, dG_{\mathbf{0}}$, there is a differentiable trajectory$\gamma : (-\varepsilon, \varepsilon) \longrightarrow E$, where its image is a subset of $G^{-1}(\mathbf{0})$ such that, $\gamma(0) = \mathbf{0}$and $\gamma'(0) = v$(for justifying this existence, see the next remark).

Hence, the composition $h = F \circ \gamma : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}$ has a local extreme in 0, where$h'(0) = 0$, thenthe chain rule gives us the following result

$$0 = h'(0) = dh_0(1) = dF_{\gamma(0)}(d\gamma_0(\mathbf{1})) = dF_{\mathbf{0}}(\gamma'(0)) = dF_{\mathbf{0}}(v)$$

as we wanted to show.■

Remark 33. Now we will justify the existence of a differentiable trajectory γ in theorem 32, for it,we will use Implicit function's theorem 31.

If $X = \boldsymbol{Ker}\, dG_{\mathbf{0}}$, then $dG_{\mathbf{0}}: E \longrightarrow \mathbb{R}$ is a continuous function and X is a closed subset of Eandtherefore complete. Choosing $w \in E$, such that $dG_{\mathbf{0}}(w) = 1$and denoting by Y a closed subsetthat contains all scalar multiples of w, we can say that Y is a copy of $\mathbb{R}$.

On the other hand, $X \cap Y = \mathbf{0}$and if $e \in E$, $a \in dG_{\mathbf{0}}(e) \in \mathbb{R}$, then

$$dG_{\mathbf{0}}(e - aw) = dG_{\mathbf{0}}(e) - a dG_{\mathbf{0}}(w) = 0$$

Hence, $e - aw \in X$ and in consequence, $e = x + y$, where $x \in X$ and $y = aw \in Y$. Therefore, Eisthe direct algebraic sum of subspaces X and Y, in addition the norm over the space $E$is equivalent to the product norm over $X \times Y$ so, we can write $E = X \times Y$. To apply the implicit function theoremwe need to see that $d_y G_{\mathbf{0}}: Y \longrightarrow \mathbb{R}$ is an isomorphism. From $Y \equiv \mathbb{R}$, we must show that $d_y G_{\mathbf{0}} \neq 0$but given $(r, s) \in X \times Y = E$we obtain that

$$dG_{\mathbf{0}}(r, s) = dG_{\mathbf{0}}(r, \mathbf{0}) + dG_{\mathbf{0}}(\mathbf{0}, s) = dG_{\mathbf{0}}(\mathbf{0}, s) = d_y G_{\mathbf{0}}(s)$$

So if we suppose that $d_y G_{\mathbf{0}} = 0$, we obtain $dG_{\mathbf{0}} = 0$, which is a contradiction.

Consequently, the implicit function theorem gives a differential function$\varphi : X \longrightarrow Y$, which graph $y = \varphi(x)$ coincides with $G^{-1}(\mathbf{0})$ in any neighborhood of0. If $H(x) = G(x, \varphi(x))$ so, $H(x) = 0$ for $x$closed to $\mathbf{0}$meaning

$$0 = dH_0(u) = d_x G_{\mathbf{0}}(u) + d_y G_{\mathbf{0}}(d\varphi_0(u)) = d_y G_{\mathbf{0}}(d\varphi_0(u))$$

for all $u \in X$, hence we get that $d\varphi_0 = 0$, as $d_y G_{\mathbf{0}}$is an isomorphism.

Finally, given $v = (u, 0) \in \boldsymbol{Ker}\, dG_{\mathbf{0}}$, we define $\gamma : \mathbb{R} \longrightarrow E$ by$\gamma(t) = (tu, \varphi(tu))$. So $\gamma(0) = \mathbf{0}$ and $\gamma(t) \in G^{-1}(0)$, for $t$enough small and additionally

$$\gamma'(0) = (u, d\varphi_0(u)) = (u, 0) = v$$

Theorem 34. *Let* F, G_1*and* G_2*be functionals over* $C^1[a,b]$*defined by*

$$F(\psi) = \int_a^b f(\psi(t), \psi'(t), t)dt \text{ and} G_i(\psi) = \int_a^b g_i(\psi(t), \psi'(t), t)dt - c_i$$

where$i = 1, 2$*and additionally*f*and* g_i*are functions of* C^2*class in* $\mathbb{R}^3$, *for* $i = 1, 2$. *Let* $\varphi \in M$*afunction of* C^2*class which is not a extreme function of* G_i, *for*$i = 1, 2$. *If* F*has a local extreme in*φ*subject to the conditions*$\psi(a) = \alpha$, $\psi(b) = \beta$*and* $G_i(\psi) = 0$; $i = 1, 2$. *Then there exist real numbers* λ_1*and* λ_2*such that* φ*satisfies the Euler-Lagrange equation for thefunction* $h = f - \lambda_1 g_1 - \lambda_2 g_2$, *that is to say*

$$\frac{\partial h}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t), \varphi'(t), t) = 0 \quad (4.18)$$

for all $t \in [a, b]$.

Proof. Suppose that $\varphi \in C^1[a,b]$and it is a class function C^2, such that Fhas a local extreme over$M \cap [G_1^{-1}(\mathbf{0}) \cap G_2^{-1}(\mathbf{0})]$. We will consider the functions to real values $F \circ T, Gi \circ T$, with $i = 1, 2$over$C_0^1[a,b]$ and from the fact that Fhas a local extreme over $M \cap [G_1^{-1}(\mathbf{0}) \cap G_2^{-1}(\mathbf{0})]$in φ, we get that$F \circ T$has a local extreme in $\mathbf{0}$, subject to the conditions $Gi \circ T(\psi) = 0$for $i = 1, 2$, so $d(Gi \circ T)_{\mathbf{0}} \neq 0$.Then, if we consider a linear function $\Lambda : \mathbb{R} \longrightarrow \mathbb{R}$, by theorem 32

$$d(F \circ T)_{\mathbf{0}} = \Lambda \circ d([G_1 \circ T] + [G_2 \circ T])_{\mathbf{0}}$$

for each $i = 1, 2$and where $dT_{\mathbf{0}}$is the identity application over $C_0^1[a,b]$. The chain rule gets us inconsequence that

$$dF_\varphi = \Lambda \circ (dG_{1_\varphi} + dG_{2_\varphi})$$

in$C_0^1[a,b]$. Writing $\Lambda(t) = (\lambda_1 + \lambda_2)t$and applying corollary 23 for $dF_\varphi$and dG_{i_φ}we concludethat

$$\int_a^b \left[\frac{\partial f}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial f}{\partial y}(\varphi(t), \varphi'(t), t)\right] u(t)dt$$
$$= \lambda_1 \int_a^b \left[\frac{\partial g_1}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial g_1}{\partial y}(\varphi(t), \varphi'(t), t)\right] u(t)dt$$
$$+\lambda_2 \int_a^b \left[\frac{\partial g_2}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial g_2}{\partial y}(\varphi(t), \varphi'(t), t)\right] u(t)dt$$

for all $u \in C_0^1[a,b]$. If $h : \mathbb{R}^3 \longrightarrow \mathbb{R}$ is defined by

$$h(x,y,t) = f(x,y,t) - \lambda_1 g_1(x,y,t) - \lambda_2 g_2(x,y,t)$$

we obtain that

$$\int_a^b \left[\frac{\partial h}{\partial x}(\varphi(t),\varphi^{'}(t),t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t),\varphi^{'}(t),t)\right] u(t)dt = 0$$

for all $u \in C_0^1[a,b]$. As a consequence of lemma 24, (4.18) is satisfied.■

Theorem 35. *Let $M_1 \subset \{\psi \in C^1[a,b] : \psi(a) = \alpha \text{ y } \psi(b) = \beta\}$and, let $F$and Gbe functionalsover $C^1[a,b]$defined by*

$$F(\psi) = \int_a^b f(\psi(t),\psi'(t),t)dt \text{ and} G(\psi) = \int_a^b g(\psi(t),\psi'(t),t)dt - c$$

where$f$and $g$are functions of C^2class in $\mathbb{R}^3$class in $\mathbb{R}^3$. Let $\varphi \in M_1$be a function of C^2class that is not anextremal of G. If Fhas a local extreme in φsubject to the conditions$\psi(a) = \alpha$,$\psi(b) = \beta$ and$G(\psi) = 0$. Then there exist a real number λsuch that φsatisfies the Euler-Lagrange equation for the function$h = f - \lambda g$, that is to say

$$\frac{\partial h}{\partial x}(\varphi(t),\varphi'(t),t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t),\varphi'(t),t) = 0$$

for all $t \in [a,b]$.

Proof. Suppose that $\varphi \in M_1$and it is of class C^2, such that Fhas a local extreme over $M_1 \cap G^{-1}(\mathbf{0})$.We know that $M_0 \subset C_0^1[a,b]$and $M_1$is the traslation by any fixed element of this subspace. Let$T : M_0 \longrightarrow M_1$the traslation defined by$T(\psi) = \psi + \varphi$, again $T(\mathbf{0}) = \varphi$, where

$$dT_{\mathbf{0}}: M_0 \longrightarrow M_0 = TM_{1_\varphi}$$

is justly identity mapping.

Now, consider the functions of real values $F \circ T$and $G \circ T$ over M_0. From the fact that Fhasa local extreme over $M_1 \cap G^{-1}(\mathbf{0})$. in φ, we obtain that $F \circ T$ has a local extreme in $\mathbf{0}$, subjectthe condition $G \circ T(\psi) = 0$. By hypothesis, $\varphi$is not an extremal for Gover M_1, that is to say$dG_\varphi\big|_{TM_{1_\varphi}} \neq 0$, so $d(G \circ T)0 \neq 0$. Then, if we consider a linear function $\Lambda : \mathbb{R} \longrightarrow \mathbb{R}$, by theorem32, we have to

$$d(F \circ T)_0 = \Lambda \circ d(G \circ T)_0$$

where$dT_0$is the identity application over M_0. The chain rule gives us that$dF_\varphi = \Lambda \circ dG_\varphi$, in M_0.Taking $\Lambda(t) = \lambda t$and applying corollary 23 to differentials $dF_\varphi$and dG_φ, we conclude to

$$\int_a^b \left[\frac{\partial f}{\partial x}(\varphi(t),\varphi^{'}(t),t) - \frac{d}{dt}\frac{\partial f}{\partial y}(\varphi(t),\varphi^{'}(t),t)\right] u(t)dt$$
$$= \lambda \int_a^b \left[\frac{\partial g}{\partial x}(\varphi(t),\varphi^{'}(t),t) - \frac{d}{dt}\frac{\partial g}{\partial y}(\varphi(t),\varphi^{'}(t),t)\right] u(t)dt$$

for all $u \in M_0$.If $h : \mathbb{R}^3 \longrightarrow \mathbb{R}$ is defined by

$$h(x,y,t) = f(x,y,t) - \lambda g(x,y,t)$$

we obtain that

$$\int_a^b \left[\frac{\partial h}{\partial x}(\varphi(t),\varphi'(t),t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t),\varphi'(t),t)\right] u(t)dt = 0$$

for all $u \in M_0$. By lemma 24 we have what we wanted. ■

We can enunciate a similar theorem over M_1, using both theorems 34 and 35 directly for itsproof. Let's see

Theorem 36. *Let $F, G_1$and G_2be functionals over $C^1[a,b]$defined by*

$$F(\psi) = \int_a^b f(\psi(t),\psi'(t),t)dt \text{ and} G_i(\psi) = \int_a^b g_i(\psi(t),\psi'(t),t)dt - c_i$$

where$i = 1,2$and additionally$f$and $g_i$are functions of C^2class in $\mathbb{R}^3$, for $i = 1,2$. Let $\varphi \in M_1$afunction of C^2class which is not a extreme function of G_i, for$i = 1,2$. If Fhas a local extreme inφsubject to the conditions$\psi(a) = \alpha$, $\psi(b) = \beta$and $G_i(\psi) = 0$; $i = 1,2$.Then there exist real numbers $\lambda_1$and $\lambda_2$such that φsatisfies the Euler-Lagrange equation for thefunction $h = f - \lambda_1 g_1 - \lambda_2 g_2$, that is to say

$$\frac{\partial h}{\partial x}(\varphi(t),\varphi'(t),t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t),\varphi'(t),t) = 0$$

for all $t \in [a,b]$.

2. EXTENSIONOF EULER-LAGRANGEEQUATION TO A NON-BOUNDED INTERVAL

Suppose that we have a functional defined over a subset which contains C^1class functions in asemi-infinite interval like $(-\infty,a]$ or $[b,\infty)$, or inclusive, over the real line, and we need to find theextremal functions of this functional. For it, we will use the fact that Euler-Lagrange equation isinvariant under regular transformations.

5.1. **Invariance of Euler-Lagrange equation.** If the functional given in (4.10) is transformdoing an independent variable substitution or a simultaneous substitution of unknown function and the independent variable, the extremals can be determined by Euler equation from integrandtransformed.

For it, we consider $t = t(u,v)$ and $x = x(u,v)$, where

$$\begin{vmatrix} t_u & t_v \\ x_u & x_v \end{vmatrix} \neq 0$$

Consider additionally $u$as an independent variable, then by chain rule we have

$$y = \frac{dx}{dt} = \frac{x_u + x_v v'}{t_u + t_v v'}$$

Then

$$\int f(x,y,t)dt = \int f\left(x(u,v), \frac{x_u + x_v v'}{t_u + t_v v'}, t(u,v)\right)(t_u + t_v v')du = \int h(v,v',u)du$$

and the extremals of the initial equation are determined from Euler's equation for the functional$\int h(v,v',u)du$, that is to say it must satisfy equation (4.12),

$$h_v - \frac{d}{du}h_{v'} = 0$$

Remark 37. For more details, you can see [9].

5.2. **Variables change in Euler-Lagrange equation.** Let's see the fundamental result of theorem34: If the functional

$$F(\psi) = \int_a^b f(\psi(t), \psi'(t), t)dt \tag{5.1}$$

subject to the restriction

$$G(\psi) = \int_a^b g(\psi(t), \psi'(t), t)dt = k, k \text{ is a constant} \tag{5.2}$$

reaches a maximum or minimum in $\varphi(t)$, then there exist $\lambda \in \mathbb{R}$ such that $\varphi(t)$ satisfies the equation

$$\frac{\partial h}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t), \varphi'(t), t) = 0$$

for$h = f - \lambda g$(Lagrange function), where $f$and $g$are C^2class functions in $\mathbb{R}^3$and $\varphi \in M$ is a C^2class function that is not an extremal of G, in addition we always consider that $-\infty < a < b < \infty$.

Now is necessary to precise the invariance of Euler-Lagrange equations, for it we consider thefollowing problem: To find the maximums or minimums of functional (5.1), subject to the restriction(5.2) with the particularity that a and b can take the values $-\infty$and/or ∞respectively.

We consider the following set to prove the next theorem

$$D = \{\psi \in C^1(\mathbb{R}): \lim_{|t|\to\infty} |t|^n \psi(t) = 0, \text{for all } n \in \mathbb{N}\}$$

Theorem 38. *Let $F$and Gbe the functionals defined by*

$$F(\psi) = \int_{-\infty}^{\infty} f(\psi(t), \psi'(t), t)dt \tag{5.3}$$

and

$$G(\psi) = \int_a^b g(\psi(t), \psi'(t), t)dt - c \tag{5.4}$$

with$\psi \in D$and f and$g$are C^2class functions. There exist constants α_{1_f}, $\alpha_{1_g} \in (0,1)$and$\alpha_{2_f}, \alpha_{2_g} \in [1, \infty)$ and $\beta_{1_f}, \beta_{1_g}, \beta_{2_f}, \beta_{2_g}$, k_1, k_2, k_3, $k_4 > 0$such that

$$|f(x,y,t)| \leq k_1|x|^{\alpha_{1f}}|t|^{\beta_{1f}} \text{and} |g(x,y,t)| \leq k_2|x|^{\alpha_{1g}}|t|^{\beta_{1g}}$$

if$|x| < 1$and$t \in \mathbb{R}$and additionally

$$|f(x,y,t)| \leq k_3|x|^{\alpha_{2f}}|t|^{\beta_{2f}} \text{and} |g(x,y,t)| \leq k_4|x|^{\alpha_{2g}}|t|^{\beta_{2g}}$$

if$|x| \geq 1$ and$t \in \mathbb{R}$. Let φbe a C^2class function, such that

$$\varphi(a) = \varphi(b) = \varphi'(a) = \varphi'(b) = 0$$

with$-\infty < a < b < \infty$that is not an extremal of G. If Fhas a local extreme in φ, then there exista real number λsuch that φsatisfies the Euler-Lagrange equation for the function $h = f - \lambda g$, thatis to say

$$\frac{\partial h}{\partial x}(\varphi(t),\varphi'(t),t)-\frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t),\varphi'(t),t)= 0 \tag{5.5}$$

for all $t \in \mathbb{R}$.

Proof. Let (a,b) be a finite interval and let $u:(a,b)\longrightarrow(-\infty,\infty)$ a C^1 class increasing mapping,such that

$$\lim_{s\to a^+} u(s)=-\infty \quad , \quad \lim_{s\to b^-} u(s)=\infty$$

Let $s$be an interior point of (a,b). Considering the change of variables $t = u(s)$, the expressions(5.3) and (5.4), give us a new variational problem, given by

$$F(\psi) = \int_a^b f(\psi(u(s)),\psi'(u(s)),u(s))u'(s)ds \tag{5.6}$$

and

$$G(\psi) = \int_a^b g(\psi(u(s)),\psi'(u(s)),u(s))u'(s)ds = K, K \text{ is a constant} \tag{5.4}$$

Is important to highlight that

$$\begin{aligned}\int_{-\infty}^{\infty} f(\psi(t),\psi'(t),t)dt &= \int_{|\psi(t)|<1} f(\psi(t),\psi'(t),t)dt + \int_{|\psi(t)|\geq 1} f(\psi(t),\psi'(t),t)dt \\ &\leq \int_{-\infty}^{\infty} k_1|\psi(t)|^{\alpha_{1f}}|t|^{\beta_{1f}}dt + \int_{-\infty}^{\infty} k_3|\psi(t)|^{\alpha_{2f}}|t|^{\beta_{2f}}dt\end{aligned}$$

and both integrals converge, as: $\lim_{|t|\to\infty}|t|^n\psi(t)=0$, for all $n\in\mathbb{N}$.

Hence, the integral that defines $F(\psi)$ in (5.6) always converges. Analogously, the integral thatdefines $G(\psi)$ in (5.7) also converges. Now, we consider γas a function defined in (a,b) and suchthat $\gamma(s) = \psi(u(s))$, then we will have

$$\gamma'(s) = \psi'(u(s))u'(s) \Rightarrow \psi'(u(s)) = \gamma'(s)u'(s)$$

Furthermore, the problem was reduced to seek extremes of the functional

$$F(\psi)=\int_a^b f\left(\gamma(s),\frac{\gamma'(s)}{u'(s)},u(s)\right)u'(s)ds \tag{5.8}$$

subject to the restriction

$$G(\psi)=\int_a^b g\left(\gamma(s),\frac{\gamma'(s)}{u'(s)},u(s)\right)u'(s)ds = K, K\text{is a constant} \tag{5.9}$$

Let $\omega(s)$ be an extreme, then ωmust satisfy the equation

$$\frac{\partial\tilde{h}}{\partial x}(\omega(s),\omega'(s),s)-\frac{d}{ds}\frac{\partial\tilde{h}}{\partial y}(\omega(s),\omega'(s),s)=0 \tag{5.10}$$

where$\tilde{h} = \tilde{f} - \lambda\tilde{g}$and

$$\tilde{f}(x,y,s) = f\left(x, \frac{y}{u'(s)}, u(s)\right)u'(s) \text{and} \tilde{g}(x,y,s) = g\left(x, \frac{y}{u'(s)}, u(s)\right)u'(s)$$

If we define $\varphi : (a,b) \longrightarrow \mathbb{R}$ by $\varphi(t) = \omega(u-1(t))$, then $\varphi$is an extreme of initial problem and$\varphi(u(s)) = \omega(s)$, where$\omega'(s) = \varphi'(u(s)).u'(s)$.

Let $h = f - \lambda g$, where

$$\tilde{h}(x,y,s) = h\left(x, \frac{y}{u'(s)}, u(s)\right)u'(s)$$

Then rewriting (5.10), we have

$$\frac{\partial \tilde{h}}{\partial x}(\omega(s), \omega'(s), s) = \frac{\partial h}{\partial x}\Big(\varphi\big(u(s)\big), \varphi'\big(u(s)\big), u(s)\Big)u'(s) \tag{5.11}$$

and

$$\frac{\partial \tilde{h}}{\partial y}(\omega(s), \omega'(s), s) = \frac{\partial h}{\partial y}\left(x, \frac{y}{u'(s)}, u(s)\right) \tag{5.12}$$

then taking the derivative with respect to s in (5.12), we obtain

$$\begin{aligned}
\frac{d}{ds}\frac{\partial \tilde{h}}{\partial y}(\omega(s), \omega'(s), s) &= \frac{d}{ds}\frac{\partial h}{\partial y}\left(\omega(s), \frac{\omega'(s)}{u'(s)}, u(s)\right) \\
&= \frac{d}{ds}\frac{\partial h}{\partial y}(\varphi(u(s)), \varphi'(u(s)), u(s)) \\
&= \nabla\frac{\partial h}{\partial y}\Big(\varphi\big(u(s)\big), \varphi'\big(u(s)\big), u(s)\Big).\Big(\varphi'\big(u(s)\big)u'(s), \varphi''\big(u(s)\big)u'(s), u'(s)\Big) \\
&= u'(s)\nabla\frac{\partial h}{\partial y}\Big(\varphi\big(u(s)\big), \varphi'\big(u(s)\big), u(s)\Big).\Big(\varphi'\big(u(s)\big), \varphi''\big(u(s)\big), 1\Big) \qquad (5.13)
\end{aligned}$$

matching(5.11) and (5.13), we have

$$\frac{\partial h}{\partial x}\Big(\varphi\big(u(s)\big), \varphi'\big(u(s)\big), u(s)\Big)u'(s) = u'(s)\nabla\frac{\partial h}{\partial y}\Big(\varphi\big(u(s)\big), \varphi'\big(u(s)\big), u(s)\Big).\Big(\varphi'\big(u(s)\big), \varphi''\big(u(s)\big), 1\Big)$$

dividing by $u'(s)$ and taking $t = u(s)$, we finally obtain

$$\begin{aligned}
\frac{\partial h}{\partial x}(\varphi(t), \varphi'(t), t) &= \nabla\frac{\partial h}{\partial y}(\varphi(t), \varphi'(t), t).(\varphi'(t), \varphi''(t), 1) \\
&= \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t), \varphi'(t), t)
\end{aligned}$$

which is the same equation of a finite interval.■

Doing a proof totally analogous, we can prove a result with multiple restrictions. We will proceedto enunciate this result without proof.

Theorem 39. *Let $F$and $G_j$be the functionals defined by*

$$F(\psi) = \int_{-\infty}^{\infty} f(\psi(t), \psi'(t), t)dt$$

and

$$G_j(\psi) = \int_a^b g_j(\psi(t), \psi'(t), t)dt - c_j$$

with $\psi \in D$ *and* f *and* g_j *for each* $j = 1, \dots n$ *are* C^2 *class functions. Then there exist constants* α_{1_f}, $\alpha_{1_{g_j}} \in (0,1), j = 1, \dots, n$ *and* $\alpha_{2_f}, \alpha_{2_{g_j}} \in [1, \infty), j = 1, \dots, n$ *and* β_{1_f}, β_{2_f}, k_1, $k_3 > 0$ *and* $\beta_{1_{g_j}}$, $\beta_{2_{g_j}}$, k_{2_j}, $k_{4_j} > 0$, $j = 1, \dots, n$ *such that*

$$|f(x,y,t)| \leq k_1 |x|^{\alpha_{1f}} |t|^{\beta_{1f}} \text{and} |g_j(x,y,t)| \leq k_{2_j} |x|^{\alpha_{1g}} |t|^{\beta_{1g_j}}$$

if $|x| < 1$ *and* $t \in \mathbb{R}$ *with* $j = 1, \dots, n$ *and additionally*

$$|f(x,y,t)| \leq k_3 |x|^{\alpha_{2f}} |t|^{\beta_{2f}} \text{and} |g_j(x,y,t)| \leq k_{4_j} |x|^{\alpha_{2g_j}} |t|^{\beta_{2g_j}}$$

if $|x| \geq 1$ and $t \in \mathbb{R}$ *with* $j = 1, \dots, n$. *Let* φ *be a* C^2 *class function, such that*

$$\varphi(a) = \varphi(b) = \varphi'(a) = \varphi'(b) = 0$$

with $-\infty < a < b < \infty$ *that is not an extremal of* G_j *for all* $j = 1, \dots, n$. *If* F *has a local extreme in* φ, *then there exista real number* λ_j *with* $j = 1, \dots, n$ *such that* φ *satisfies the Euler-Lagrange equation for the function* $h = f - \sum_{j=1}^{n} \lambda_j g_j$, *thatis to say*

$$\frac{\partial h}{\partial x}(\varphi(t), \varphi'(t), t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\varphi(t), \varphi'(t), t) = 0 \qquad (5.5)$$

for all $t \in \mathbb{R}$.

Remark 40. We can find distribution functions that maximize the functional that defines theentropy, and they are not defined by all line, then in these cases, we can use results exposedin the last two theorems as follows:

(1) If the interval where the functional and restrictions are defined, is of the form (a, ∞) with $a \in \mathbb{R}$, then the set D is defined as follows:

$$D = \{\psi \in C^1(\mathbb{R}): \lim_{t \to \infty} |t|^n \psi(t) = 0, \text{for all } n \in \mathbb{N}\}$$

(2) If the interval where the functional and restrictions are defined, is of the form $(-\infty, a)$ with $a \in \mathbb{R}$, then the set D is defined as follows:

$$D = \{\psi \in C^1(\mathbb{R}): \lim_{t \to -\infty} |t|^n \psi(t) = 0, \text{for all } n \in \mathbb{N}\}$$

V. MAXIMA ENTROPY PROBLEMS

Let X be a random variable which takes values in the real line. The probability that X takesvalues less or equal to the given real number $x$given, it's obtained integrating the density functionρ, namely

$$P(X \leq x) = \int_{-\infty}^{x} \rho(t)dt$$

but as X can take any value, we obtain that

$$\int_{\mathbb{R}} \rho(t)dt = 1$$

In too many problems, the principal interest is focused on achieving determining the density functionρ, having as basis the background about certain expected values. We will do this, using thetheorems of the last section. Now, we will enunciate and prove the necessary theorem which willallow us to characterize the maxima entropy functions.

6.1. **General theorem.**

Theorem 41. *Let* $h_1,\dots,h_n:\mathbb{R}\longrightarrow\mathbb{R}$*be measurable functions, such that* $|h_k(t)|\leq m_k|t|^{n_k}, m_k, n_k > 0$*for each* $k = 1,\dots,n$*and* $c_1,\dots,c_n\in\mathbb{R}$*. Suppose that there exist* $a_0,\dots,a_n\in\mathbb{R}$*suchthat the function* $q:\mathbb{R}\longrightarrow\mathbb{R}$*defined by*

$$q(t) = \exp(a_0 + a_1h_1(t) + \cdots + a_nh_n(t))$$

is a density of probability and each of functions $h_kq, k = 1,\dots,n$*are integrable in* $\mathbb{R}$*and*

$$\int_{-\infty}^{\infty} h_k(t)q(t)dt = c_{k,}\ \ k = 1,\dots,n$$

Then the functional F*defined by*

$$F(p) = \int_{-\infty}^{\infty} p(t)\log p(t)dt,$$

with its domain formed by all measurable functions$p:\mathbb{R}\longrightarrow\mathbb{R}$*such that*

$$p(t)\geq 0,\ \int_{-\infty}^{\infty} p(t)dt = 1 \text{ and } \int_{-\infty}^{\infty} h_k(t)q(t)dt = c_{k,}\ \ k = 1,\dots,n$$

reach an extreme in q*.*

Proof. Consider the functional given by

$$F(p) = \int_{-\infty}^{\infty} p(t)\log p(t)dt \tag{6.1}$$

subject to the restrictions

$$\int_{-\infty}^{\infty} h_k(t)q(t)dt = c_{k,}\ \ k = 1,\dots,n \tag{6.2}$$

Additionally, consider the functions f and g_k, with$k = 0,\dots,n$ defined as follow

$$f(x,y,t) = -x\ \log x\ , g_k(x,y,t) = xh_k(t)\,, k = 0,\dots,n$$

where

$$|f(x,y,t)| = |-x\ log\ x| = |x||\log x| \leq |x|^{\frac{3}{2}}, \text{if} |x| < 1$$

and

$$|f(x,y,t)| \leq |x|^2\,, \text{if} |x| > 1$$

and

$$|g_k(x,y,t)| = |xh_k(t)| \leq |x||t|^{n_k}$$

in addition, $h_0(t)\equiv 1$, for all $t\in\mathbb{R}$ and $c_0 = 1$.

Now, if $q$is an extreme of the functional given in (6.1) subject to the restrictions (6.2), thereare real numbers $\lambda_0,\dots,\ \lambda_n$, such that $q$satisfies the Euler-Lagrange equation for the function $z$defined by

$$z(x,y,t)=f(x,y,t)-\sum_{k=0}^{n}\lambda_k g_k(x,y,t)$$

that is to say

$$\frac{\partial z}{\partial x}(q(t),q^{'}(t),t)-\frac{d}{dt}\frac{\partial z}{\partial y}(q(t),q^{'}(t),t)=\ 0$$

Hence, we have that

$$\frac{\partial z}{\partial x}=-\log x-1-\sum_{k=0}^{n}\lambda_k h_k(t)\ \text{ and }\frac{\partial z}{\partial y}=\ 0$$

then the Euler-Lagrange equation for q, is given by

$$-\log q(t)-1-\sum_{k=0}^{n}\lambda_k h_k(t)=0$$

furthermore

$$\log q(t)=-1-\sum_{k=0}^{n}\lambda_k h_k(t)\Longrightarrow q(t)=\exp\left\{-1-\sum_{k=0}^{n}ch_k(t)\right\}$$

that is what we wanted to prove, where $a_k\ =\ -\lambda_k$, for each $k\ =\ 1,\dots,n$and $a_0\ =\ -1\ -\ \lambda_0$.■

6.2. **Characterization of maxima entropy functions depending on their restrictions.**

Consider X a random variable which takes values in some subset Ion the real line. Let $\rho:\ I\longrightarrow\mathbb{R}$ be the density function of X, that is to say, ρsatisfies the following conditions:

(i) $\rho(t)\ >\ 0$, for all $t\ \in I$ and $\rho(t)\ =\ 0$in another case.
(ii)$\int_I\ \rho(t)dt=1$

From calculus of variations technics, we will try to generate some typical density functions thatmaximize the functional

$$S[\rho]=\int_{-\infty}^{\infty}\rho(t)\log\rho(t)dt \tag{6.3}$$

additionally, we will suppose that the function ρsatisfies conditions of the form:

$$\int_I p_k(t)\rho(t)dt=c_k,\quad c_k\in\mathbb{R},\qquad k=1,2,\dots \tag{6.4}$$

where$p_1(t)\ =\ 1$, for all $t\ \in I$and $c_1\ =\ 1$.

6.2.1. *Exponential case.* Suppose that $I\ =\ (0,\infty)$ and $\rho$is a function that satisfies the conditions (i)and (ii), in addition to $E(X)\ =\ \alpha$. Furthermore, the problem focuses on maximizing the functional(6.3), subject to restrictions:

$$\int_0^\infty \rho(t)dt = 1 \text{ and } \int_0^\infty t\rho(t)dt = \alpha \,, \alpha > 0 \qquad (6.5)$$

where $p_2(t) = t$ and $c_2 = \alpha$.

Consider the functions

$$f(x,y,t) = -x \log x \,,\ g_1(x,y,t) = x \text{ and } g_2(x,y,t) = tx$$

Then the functions g_1 and g_2, satisfy

$$|\, g_1(x,y,t)| \leq |x| \text{ and } |g_2(x,y,t)| \leq |t||x|$$

Then by theorems 36 and 39, there exist $\lambda_1, \lambda_2 \in \mathbb{R}$ such that ρ satisfies the Euler-Lagrange equation for the function h defined by

$$h(x,y,t) = f(x,y,t) - \lambda_1 g_1(x,y,t) - \lambda_2 g_2(x,y,t)$$

that is

$$\frac{\partial h}{\partial x}(\rho(t), \rho'(t), t) - \frac{d}{dt}\frac{\partial h}{\partial y}(\rho(t), \rho'(t), t) = 0 \qquad (6.6)$$

for all $t \in (0, \infty)$.

In virtue of

$$\frac{\partial h}{\partial x} = -\log x - 1 - \lambda_1 - \lambda_2 t \text{ and } \frac{\partial h}{\partial y} = 0$$

we have that (6.6), can be represented by $-\log x - 1 - \lambda_1 - \lambda_2 t = 0$. Hence, if ρ is the extremal function we obtain in consequence

$$\rho(t) = Ce^{-\lambda_2 t} \,, C = e^{1+\lambda_1}$$

Then, using the conditions given in (6.5), we obtain that

$$1 = \int_0^\infty Ce^{-\lambda_2 t}dt = C\left[-\frac{1}{\lambda_2 e^{-\lambda_2 t}}\right]_0^\infty = \frac{C}{\lambda_2} \Rightarrow C = \lambda_2$$

On the other hand,

$$\alpha = \int_0^\infty Cte^{-\lambda_2 t}dt = C\int_0^\infty te^{-\lambda_2 t}dt = C\left[-\frac{t}{\lambda_2 e^{-\lambda_2 t}}\right]_0^\infty + \frac{C}{\lambda_2}\int_0^\infty e^{-\lambda_2 t}dt = \frac{1}{\lambda_2}$$

Then, $\rho(t) = \frac{1}{\alpha}e^{-\frac{1}{\alpha}t}$. Taking $\frac{1}{\alpha} = a$ we obtain the expected result.

6.2.2. *Gamma case.* Suppose again that $I = (0, \infty)$ and ρ maximises (6.3), subject to the restrictions

$$\int_0^\infty \rho(t)dt = 1 \text{ and } \int_0^\infty t\rho(t)dt = \alpha \,, \alpha > 0 \text{ and } \int_0^\infty \log(t)\,\rho(t)dt = \beta \qquad (6.7)$$

Taking now the function

$$h(x,y,t) = f(x,y,t) - \lambda_1 g_1(x,y,t) - \lambda_2 g_2(x,y,t) - \lambda_3 g_3(x,y,t)$$

and considering the functions

$$f(x,y,t) = -x \log x \ , \ g_1(x,y,t) = x \ , g_2(x,y,t) = tx \text{ and } g_3(x,y,t) = x \log t$$

By the exponential case, we know that the functions g_1 and g_2, satisfy the conditions of theorems36 and 39, then

$$|g_3(x,y,t)| = |x \log t| \leq |t||x|, \text{for each } t \in (0,\infty)$$

which verifies again theorems 36 and 39, and doing an analogous calculus, we have to

$$-\log x - 1 - \lambda_1 - \lambda_2 t - \lambda_3 \log t = 0$$

from which we obtain

$$\rho(t) = Ce^{-\lambda_2 t} t^{-\lambda_3}$$

Then, using the conditions given in (6.7), we get

$$1 = \int_0^\infty Ce^{-\lambda_2 t} t^{-\lambda_3} dt = C \int_0^\infty e^{-u} \frac{u^{-\lambda_3}}{\lambda_2^{-\lambda_3}} \frac{du}{\lambda_2}, \text{ if } u = \lambda_2 t \tag{6.8}$$

Now, if $p - 1 = -\lambda_3$ then (6.8) can be wrote as follows

$$1 = \frac{C}{\lambda_2^p} \int_0^\infty e^{-u} u^{p-1} du = \frac{C}{\lambda_2^p} \Gamma(p) \Longrightarrow C = \frac{\lambda_2^p}{\Gamma(p)}$$

from the second condition in (6.7), we have

$$\alpha = \int_0^\infty Cte^{-\lambda_2 t} t^{-\lambda_3} dt = C \int_0^\infty \frac{u}{\lambda_2} e^{-u} \frac{u^{-\lambda_3}}{\lambda_2^{-\lambda_3}} \frac{du}{\lambda_2} = \frac{C}{\lambda_2^{p+1}} \int_0^\infty e^{-u} u^p du = \frac{C}{\lambda_2^{p+1}} p\, \Gamma(p) = \frac{p}{\lambda_2}$$

Hence, $\lambda_2 = \frac{p}{\alpha}$. If $\frac{p}{\alpha} = a$, we will have that ρ can be wrote as

$$\rho(t) = \frac{a^p}{\Gamma\,(p)} e^{-at} t^{p-1}$$

Remark 42. Using the third condition in (6.7), it can be seen that $p = 1$, furthermore the obtainedfunction transforms to the exponential case , because $\Gamma\,(1) = 1$.

6.2.3. *Laplace case.* If we consider $I = \mathbb{R}$ with conditions

$$\int_{-\infty}^\infty \rho(t) dt = 1 \text{ and } \int_{-\infty}^\infty |t| \rho(t) dt = \alpha\,, \alpha > 0 \tag{6.9}$$

we can verify that the density function is given by

$$\rho(t) = Ce^{-\lambda_2 |t|}$$

Then, as the function ρ in this case is symmetric, by the exponential case, it is easy to reach

$$\rho(t) = \frac{1}{2} a e^{-a|t|}$$

6.2.4. *Normal case.* We consider $I = \mathbb{R}$, with conditions

$$\int_{-\infty}^{\infty} \rho(t)dt = 1 \text{ and } \int_{-\infty}^{\infty} t^2\rho(t)dt = \sigma^2 \qquad (6.10)$$

In addition considering the functions $f, g_1$and $g_2$as follow

$$f(x,y,t) = -x \log x \ , \ g_1(x,y,t) = x \ , g_2(x,y,t) = t^2x$$

by taking the function $h$which satisfies the Euler-Lagrange equation like

$$h(x,y,t) = f(x,y,t) - \lambda_1 g_1(x,y,t) - \lambda_2 g_2(x,y,t)$$

we have that

$$\frac{\partial h}{\partial x} = -\log x - 1 - \lambda_1 - \lambda_2 t^2 \text{and} \frac{\partial h}{\partial y} = 0$$

Hence, the Euler-Lagrange equation for the extremal function $\rho$is given by

$$-\log \rho(t) - 1 - \lambda_1 - \lambda_2 t^2 = 0$$

using logarithm properties, we obtain

$$\log[C\rho(t)] = -\lambda_2 t^2 \Rightarrow \rho(t) = \frac{1}{C} e^{-\lambda_2 t^2}$$

where$C = e^{1+\lambda_1}$.

Finally, using the restrictions in (6.10) and the fact that

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

we obtain

$$\rho(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

REFERENCES

[1]. N.I. Akhiezer, *The Calculus of Variations*, Blaisdell Publishing Company, 1962.
[2]. N. I. Akhiezer, *The classical moment problem*, Hafner Publishing, Translation 1965, Oliver and Boyd.
[3]. G.A. Bliss, *Lectures on the Calculus of Variations*, The Univ. of Chicago Press, 1946.
[4]. J.B. Brissaud, *The meanings of Entropy*, Entropy ISNN 1099-4300, 7: 68-96, 2005.
[5]. H. Cartan, *C_alculo Diferencial*, Ediciones Omega, Barcelona, 1972.
[6]. J. Chover, *On Normalized Entropy and the Extensions of a Positive-Definite Functions*, Journal of Mathematicsand Mechanics, Vol. 10, **6**, (1961) 927-945.
[7]. C. Edwards, *Advanced Calculus*, Prentice Hall, Voll. II, 1963.
[8]. J. Figueroa, *Brief Notes on the Calculus of Variations*,http://www.maths.ed.ac.uk/ jmf/Teaching/Lectures/CoV.pdf.
[9]. I. Gelfand, S. Fomin, *Calculus of Variations*, Prentice hall, INC, (1963).
[10]. M. Junk, *Maximum entropy for reduced moment problems*, Mathematical Models and Methods in Applied Sci-ences, 10(7): 1001-1025, 2000.
[11]. A. Kagan, M., Y. Linnik, C. Rao, *Characterization problems in mathematical statistics*, Wiley. New York, (1973).
[12]. D. A. Lavis, *Some Examples of Simple Systems*, http://www.mth.kcl.ac.uk/ dlavis/papers/examples-ver2.pdf.
[13]. R.D. Levine, M. Tribus (editors), *The Maximum Entropy Formalism*, MIT Press, Cambridge, 1979.
[14]. B. Mahan, *Termodinmica Química Elemental*, Ed, Revert, S.A., (1976).
[15]. G. Makarenko, M. Krasnov, A. Kiseliov, *Cálculo Variacional*, Ed. Mir Moscú, (1976).
[16]. J. Marsden, *Elementary Classical Analysis*, W. H. Freeman and Company, (1974).
[17]. J. Marsden, A. Tromba, *C_alculo Vectorial*, Addison - Wesley Iberoamericana, (1991).
[18]. M. Ohya, D. Petz, *Relative Engropy and its Use*, Springer-Verlag, 1993.
[19]. C. Rao, *Linear statistical inference and its applications*, Wiley, 2nd ed. New York, (1973).

[20]. R. Resnick, D. Halliday, *Fsica*, Compaia Editorial Continental, S.A., Parte I, (1980).
[21]. W. Rudin, *Functional Analysis*, 2d ed.,McGraw-Hill, Barcelona, 1974.
[22]. W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw Hill, 1987.
[23]. W. Rudin, *The extension problem for positive-de_nite functions*, Illinois J. Math., 7:532-539, 1963.
[24]. C. Shannon, *A Mathematical Theory of Communication,* The Bell System Technical Journal, **27**, No.2, (1948)623-656.