Far East Journal of Theoretical Statistics
Volume 41, Number 2, 2012, Pages 77-92
Published Online: November 2012
Available online at http://pphmj.com/journals/fjts.htm
Published by Pushpa Publishing House, Allahabad, INDIA

THE EXISTENCE AND UNIQUENESS OF SOLUTION TO AN ISOPERIMETRIC PROBLEM IN THE CLASS OF QUASI-PROBABILITY DENSITY FUNCTIONS

W. R. Javier, H. Muñoz-Barona and K. Prabaharan

Department of Mathematics
Southern University and A & M College
Baton Rouge, LA 70813, U. S. A.

Abstract

Section I of this paper provides a necessary and sufficient condition for the existence and uniqueness of the solution of a particular isoperimetric problem, the minimal arc-length problem defined on the set of Quasi-Probability Density Functions (QPDF). This settles the question of existence or non-existence of solution to the minimal arc-length among the eligible Probability Density Functions (PDF). This is accomplished by using the answer to an ancient isoperimetric problem of Queen Dido. Section II provides an alternative proof of the same conclusion by a direct application of an extremal theorem from the Calculus of Variations.

Introduction

Imagine a set, finite or infinite, of nonnegative continuous functions $y = y(x)$ defined over the unit interval $[0, 1]$, whose graph joins the endpoints of the unit interval; and which traps, with the unit line segment, an

2010 Mathematics Subject Classification: 49J.

Keywords and phrases: calculus of variation, minimal arc-length problem, class of QPDFs, boundary conditions, legend of Queen Dido, Mathematica.

Received June 30, 2012

area exactly equal to one square unit. Assume further that each graph has an arc-length that can be computed by the familiar integral calculus arc-length formula (see [4-5])

$$\int_0^1 \sqrt{1 + (y'(x))^2}\, dx.$$

Is there a member of this set that has the shortest arc-length?

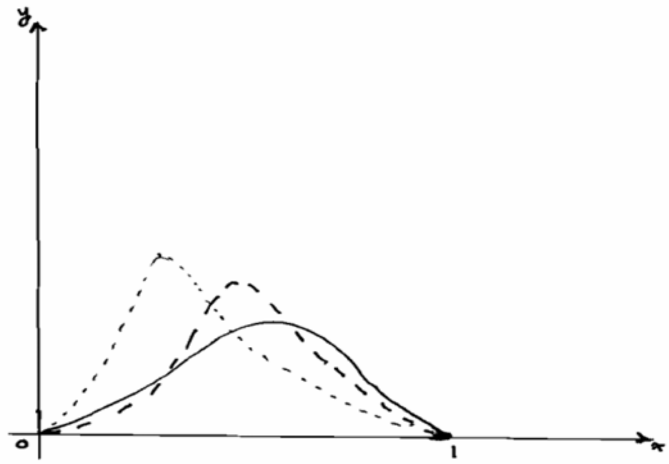

Figure 1. Showing graphs of some contending functions; some may not be unimodal.

Examples of Sets of Contending Functions

(E_1) The set of "Triangular" functions of the form

$$y = f(x;\, a) \equiv \frac{2}{a} x(I_{[0,\, a)}(x)) + \frac{2}{a-1}(x-1)(I_{[a,1]}(x)),$$

where, in general, for a given set of numbers S

$I_S(x) = 1$ whenever x is a number in the set S; otherwise 0 for any other value of x; and a is a parameter value, $0 < a < 1$, that identifies a particular Triangular function.

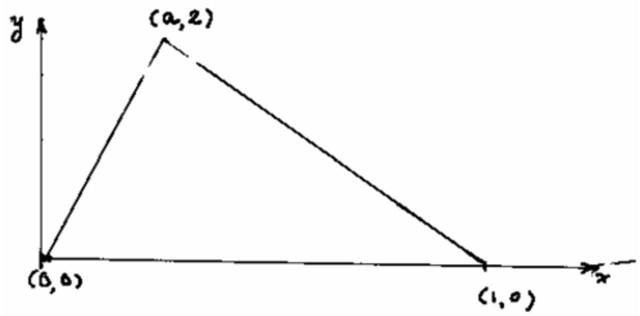

Figure 2. Showing one Triangular function; $0 < a < 1$.

Exercise. Show that this set has a minimal arc-length function, namely the isosceles triangle function, corresponding to $a = 1/2$.

(E_2) The set of "Tent" functions, symmetric about $x = 1/2$, of the form

$$y = g(x;\, a) \equiv \frac{x}{a(1-a)}\left(I_{[0,\, a)}(x)\right) + \frac{1}{1-a}\left(I_{[a,\, 1-a)}(x)\right)$$

$$- \frac{x-1}{a(1-a)}\left(I_{[1-a,\, 1]}(x)\right),$$

where each parameter $0 < a \le 1/2$, identifies a unique member of the set. Figure 3 below shows one such Tent function

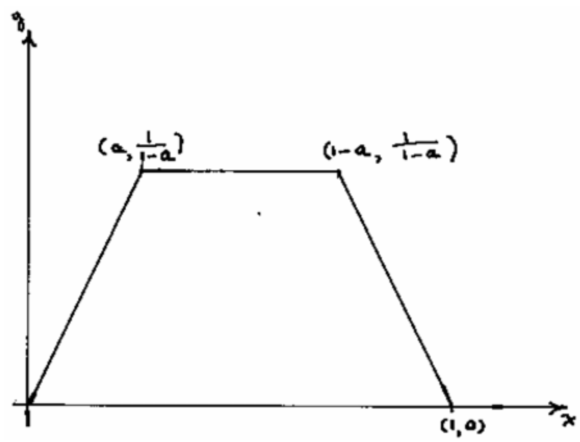

Figure 3. One Tent function; $0 < a \le 1/2$.

Exercises. (i) Find an expression for the arc-length $L(a)$ as a function of a;

(ii) Show that $L(a)$ is monotonically increasing in $(0, 1/2]$;

(iii) Explain why no minimal arc exists in this set of functions.

(E_3) The set of Beta density functions symmetric about $x = 1/2$,

$$y = h(x;\, a) \equiv \frac{(x(1-x))^{a-1}}{B(a)}(I_{[0,1]}(x)),$$

where

$$B(a) = \int_0^1 (x(1-x))^{a-1} dx; \quad a > 1.$$

Is the Beta integral function $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$, of Probability Theory?

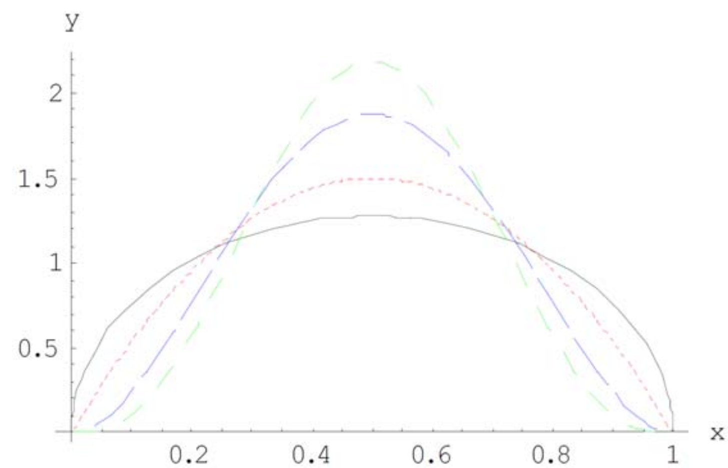

Figure 4. Showing a few Beta probability density functions.

Exercise. The reader is invited to show, numerically, with the use of any scientific calculator, that this class of functions has a minimal arc-length function at the parameter value a in the interval (1.17, 1.18).

One notes that each of the sets in the above examples consists of Probability density functions of Probability theory, see Ross [3]. It is therefore natural to pose the general question:

(GQ) In the class of all probability density functions defined over the unit interval and joining the two endpoints, find one, if it exists, that has the shortest arc-length.

This general question which properly belongs to the Calculus of Variations is treated and answered in Section I, using fairly elementary methods. Section II arrives at the same conclusion by using directly, a result in the Calculus of Variations.

In this paper this general question (GQ) is treated in the wider context of the class of Quasi Probability Density Functions (QPDF) and is stated as

An isoperimetric* problem:

(IP) Among all piecewise continuous non-negative functions $y = f(x)$ defined in the unit interval $[0, 1]$, with piecewise continuous first derivative $y'(x)$, and satisfying the conditions

(C_1) $y(0) = y(1) = 0,$

(C_2) $\int_0^1 y(x) = A_0 \quad (A_0 > 0)$

find the function $y = y(x)$ that has the shortest arc-length, i.e., $y(x)$ minimizes the functional

$$J(y) = \int_0^1 \sqrt{1 + (y'(x))^2}\, dx.$$

Remark 1*. The term "isoperimetric", traditionally used in the Calculus of Variations, is a misnomer in the context of the present general question (GQ), since the feasible functions in this situation, all have equal areas A_0, not equal arcs or perimeters.

Remark 2. Since the positive constant in (C_2) is in general $A_0 \neq 1,$ the function $y(x)$ is termed "quasi" probability density function QPDF; however when $A_0 = 1,$ then the nonnegative function $y = y(x)$ becomes a bonafide probability density function (PDF), see Ross [3], and the problem reverts to the original problem (GQ).

I. Necessary and Sufficient Conditions for the Existence and Uniqueness of Solution to (IP)

First, some history:

The story or legend of Queen Dido lends some insight on this minimal arc-length problem. Queen Dido was sent away from her homeland after her brother murdered her husband. She was exiled to another country, in North Africa, where the local king grudgingly welcomed her. She asked the king for a piece of land to live on. The king gave her a pile of ox hides and informed her that what piece of land she could enclose with those pieces of hide would be her and her progenies to keep. Queen Dido took the ox hides, had them cut into thin strips and had them tied end to end, thereby forming a long closed loop. What should the shape of the loop be, when laid on the ground, so the area of land enclosed is largest? She then guessed quickly, that the optimum shape of the loop should be a circle. That was how she claimed the largest possible piece of prime real estate in the king's domain. The territory she thus claimed later became the city of Carthage, the birthplace of the Carthaginian general Hannibal The Great of the Punic Wars fame, in ancient history.

Queen Dido got her land but she did not live long enough to enjoy her acquisition. She preferred death by her own hand, to marrying the King who fell for her shortly after her display of mathematical acumen. However, the answer that she found by intuition, is a now a celebrated result, also intuitively guessed, just like Queen Dido did more than 2500 years ago, by most alert present-day students. The result was rigorously proved many years later in the Calculus of Variations, a branch of advanced mathematics whose introduction and development are associated with names of prominent modern mathematicians including Newton, Lagrange, Euler, and Bernoulli. We state this result as.

Theorem 1 (Queen Dido's Theorem)**.** *Among all closed plane curves, all having the same arc-length L, the circle encloses maximum area.*

This result has an interesting ancient history or legend connected with Queen Dido of Carthage as recounted in the previous two paragraphs; the result is proved in the Calculus of Variations, see Gelfand and Fomin [1].

As a consequence of Theorem 1, we state and prove the following key result.

Theorem 2. *Among all simple closed loops joining the endpoints of the unit interval, and enclosing the same area* A_0 *with the unit interval and above the x-axis, the unique circular arc passing through the endpoints has the shortest arc-length.*

Proof. The proof consists of two parts: existence and uniqueness of the circular arc that traps the given area A_0 above the unit interval, must first be established. This will be shown in Subsection 2.1. The proof then proceeds as that given in Javier, [2], this is done in Subsection 2.2 below; it is based on Queen Dido's Theorem.

Subsection 2.1

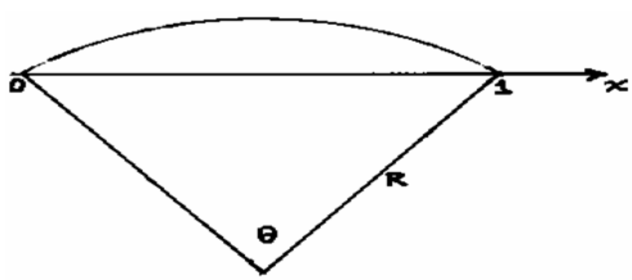

Figure 5. Circular arc passing through the endpoints and with central angle θ and radius R. Area of circular sector is $R^2\theta/2$; area of isosceles triangle is $(1/2)(R)(R\sin\theta)$.

First, observe that there are infinitely many circles that pass through the endpoints of the unit interval, each has center on the vertical line $x = 1/2$. We shall now show that there is only one that traps the specified area A_0. Let θ and R be the central angle subtended by the arc and the radius, respectively, of such a circle. Then, by the Law of cosines applied to the triangle formed by the two endpoints of the unit line segment and the center of the circle, we have for any $0 \leq \theta \leq 2\pi$

$$1 = 2R^2 - 2R^2 \cos\theta. \tag{1.1}$$

The area A, trapped by that circular arc and the unit line segment is the difference between area of the circular sector and the area of the isosceles triangle formed by the unit line segment and the two radii, thus:

$$A = \frac{R^2}{2}(\theta - \sin\theta). \tag{1.2}$$

From (1.1),

$$R^2 = \frac{1}{2(1 - \cos\theta)} \tag{1.3}$$

using (1.3) in (1.2), we have the area as a function of the central angle θ. Thus, (1.2) becomes

$$A = A(\theta) = \frac{[\theta - \sin\theta]}{4(1 - \cos\theta)}. \tag{1.4}$$

The last equation allows us to define the area A as a function of the variable θ on the interval $(0, 2\pi)$. As such, it is a differentiable function with respect to θ in the interval $(0, 2\pi)$. Furthermore, it is a monotonically increasing function in that interval, since its derivative

$$A'(\theta) = \frac{2(1 - \cos\theta) - (\theta\sin\theta)}{4(1 - \cos\theta)^2} \tag{1.5}$$

is positive in the interval $(0, 2\pi)$. In fact, it can be shown that the numerator is always positive in that interval either graphically or analytically using elementary calculus. Furthermore, the range of this area function is $(0, \infty)$. Thus given any area A, equation (1.4) can be solved for a unique value θ; this value of θ, will then give the corresponding radius R, using equation (1.3).

These two values define the unique circular arc, corresponding to the specified area A_0. Equation (1.4) defines a one-one onto map between the intervals $(0, 2\pi)$ and $(0, \infty)$.

Thus, in general, there is a unique circular arc that traps with the unit interval the given area A_0.

Next, we show that this unique circular arc, corresponding to the specified area A_0, has the shortest arc-length among all loops joining the endpoints of the unit interval and which trap the same area A_0.

Subsection 2.2

In Figure 6, we make the following designations:

$$\Gamma_0 \equiv ABC, \quad \Gamma_1 \equiv ABE, \quad \Gamma \equiv ADB.$$

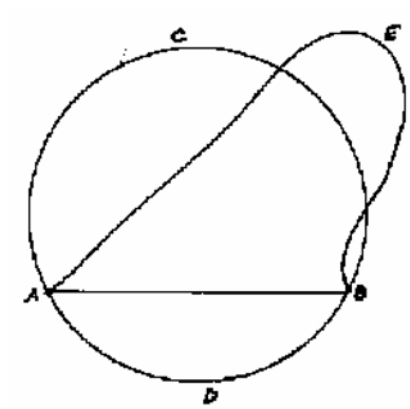

Figure 6. Graphs of the loops $\Gamma_0 \equiv ABC, \quad \Gamma_1 \equiv ABE, \quad \Gamma \equiv ADB.$

Let Γ_0 be the unique circular arc joining the endpoints of the unit interval and has area A_0. In Figure 6, it is the arc ABC; this statement is justified in view of the work done in Subsection 2.1. Let the arc *ABE*, denoted as Γ_1 be any other loop joining the endpoints of the unit interval and has also area A_0. We shall show that the length of Γ_1, $L(\Gamma_1)$ cannot be shorter than the length of Γ_0, $L(\Gamma_0)$, i.e., $L(\Gamma_0) \leq L(\Gamma_1)$.

Let *ADB*, denoted as Γ, be the completion of Γ_0 so that $\Gamma_0 \cup \Gamma$ is a circle, which we denote as C_0, i.e., $C_0 = \Gamma_0 \cup \Gamma$; Γ_0 and Γ are its circular arcs having the unit interval *AB* as their common chord *AB*.

Now consider the loop $\Gamma_1 \cup \Gamma$; *didonize* ** this loop into the circle *C*. In this transformation the length of the perimeter remains unchanged. We note

(i) $Area(C) \geq Area(\Gamma_1 \cup \Gamma) = Area(C_0)$ (the inequality holds by virtue of Dido's Theorem; the equality follows from the choice of the arcs Γ and Γ_1),

(ii) $Circumference(C) = L(\Gamma_1 \cup \Gamma) = L(\Gamma_1) + L(\Gamma)$,

(iii) $Circumference(C_0) = L(\Gamma_0 \cup \Gamma) = L(\Gamma_0) + L(\Gamma)$.

Therefore, equation (i) becomes

$$\pi\left(\frac{L(\Gamma_1) + L(\Gamma)}{2\pi}\right)^2 = Area(C) \geq Area(C_0) = \pi\left(\frac{L(\Gamma_0) + L(\Gamma)}{2\pi}\right)^2;$$

this implies

$$L(\Gamma_1) \geq L(\Gamma_0).$$

This completes the proof of Theorem 2. □

An Application

As an application, equation (1.4) can be solved for θ when $A(\theta) = 1$. In this case, we have

$$\theta - \sin\theta = 4(1 - \cos\theta) \tag{1.6}$$

which gives, using Mathematica, $\theta = 4.37607$, and equation (1.3) gives

$R = 0.613137.$ These values lead to the circle with center above the x-axis, on the vertical line $x = 1/2$ and passing through the endpoints of the unit interval. Its intercepted arc-length is 2.68. According to Theorem 2, among all curves joining the two endpoints of the unit interval with the same unit area, no curve can be shorter than this circular arc. Numerical verifications of this statement can be carried out by doing the exercises in examples (E_1), (E_2) and (E_3).

**Didonize, the term first used in Javier, [2], (2003), in honor of the legendary Queen Dido, means transforming the loop into a circle; the process is assumed to preserve the length of the loop.

Theorem 3. *A circular arc passing through the endpoints of the unit interval* $[0, 1]$ *and above the unit interval*, *represents a non-negative function* $y = y(x)$ *if and only if the circular arc traps area* $A_0 \leq \frac{\pi}{8}$.

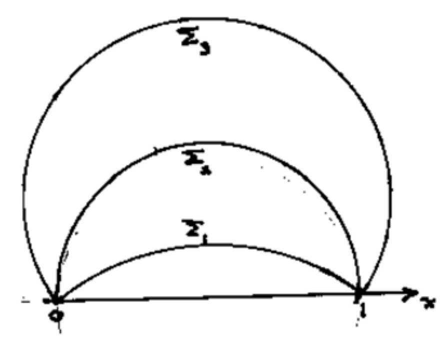

Figure 7. Showing three circular arcs: arc Σ_1 is a function and has area less than $\pi/8$; arc Σ_2 is the limiting circular arc - function, has area exactly equal to $\pi/8$; arc Σ_3 is a circular arc that does not qualify as a function, it has area greater than $\pi/8$.

Proof. Consider the semi-circle, Σ_2, centered at $(1/2, 0)$; it has area equal to $\frac{\pi}{8}$. Any other circular arc with area $\leq \frac{\pi}{8}$ represents a function, since the entire arc is over $[0, 1]$. However, when the area trapped exceeds

$\frac{\pi}{8}$, as in Σ_3, it does not represent a function over the unit interval, since portions of the arc, near the endpoints, spill out of $[0, 1]$; furthermore, it has no unique value at $x = 0$ and at $x = 1$. □

Putting together Theorems 1, 2, and 3 proves the following theorem.

Theorem 4. *The Isoperimetric problem* (*IP*) *has a unique solution if and only if the positive constant* A_0 *does not exceed* $\frac{\pi}{8}$.

Theorem 5. *The Isoperimetric* (*IP*) *has no solution if* $A_0 = 1$.

Proof. Since $A_0 = 1 > \frac{\pi}{8}$, Theorem 4 implies that problem (IP) has no solution when $A_0 = 1$. □

Answer to (GQ)

Thus, in the class of eligible PDF functions, defined in Remark 2 above, there is no minimal arc; hence

(GQ) is answered in the negative

Next, we shall show another way to prove the main results of Section I (Theorems 4 and 5) and the answer to (GQ), by using directly a result in the Calculus of Variations: the Gelfand's necessary condition (Gelfand and Fomin, Theorem 1, p. 43). This is taken up in the next section.

II. Non-existence of the Solution for a Variational Problem with Subsidiary Conditions

Following Gelfand Fomin [1], we state

A Variational-Isoperimetric Problem

Find the curve $y = f(x)$ for which the functional

$$J(y) = \int_a^b F(x, y, y')dx \tag{2.1}$$

attains its smallest value among all admissible curves $y(x)$ that satisfies the boundary conditions

$$y(a) = c, \quad y(b) = d$$

and are such that another functional

$$K(y) = \int_a^b G(x, y, y')dx = l \tag{2.2}$$

is satisfied; l is a fixed value.

To solve this problem, it is assumed that the functions F and G defining the functionals (2.1) and (2.2) have continuous first and second derivatives in $[a, b]$ for arbitrary values of y and y'. The following theorem (Gelfand and Fomin [1]) which we quote without proof, provides the necessary condition for the existence of a solution for this isoperimetric problem.

Theorem 6. *Given the functional* $J(y) = \int_a^b F(x, y, y')dx$ *let the admissible curves satisfy the conditions* $y(a) = c$, $y(b) = d$ *and*

$$K(y) = \int_a^b G(x, y, y')dx = A_0, \quad A_0 > 0, \tag{2.3}$$

where $K[y]$ *is another functional, and let* $J[y]$ *have an extremum for* $y = y(x)$. *Then if* $y = y(x)$ *is not an extremal of* $K[y]$, *there exists a constant* λ *such that* $y = y(x)$ *is an extremal of the functional*

$$\int_a^b (F(x, y, y') + \lambda G(x, y, y'))dx,$$

i.e., $y = y(x)$ *satisfies the differential equation*

$$F_y - \frac{d}{dx}F_{y'} + \lambda\left(G_y - \frac{d}{dx}G_{y'}\right) = 0. \tag{2.4}$$

We apply Theorem 6, to problem (PI) with functions

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

$$G(x, y, y') = y(x), \quad J[y] = \int_0^1 \sqrt{1 + (y'(x))^2}\, dx;$$

boundary conditions $y(0) = 0$ and $y(1) = 0$, and the constraint that $y(x) \geq 0$, x in $[0, 1]$; and $l = A_0$. Then, a solution of problem (PI), $y(x)$, if it exists, satisfies the differential equation, equation (2.4) which reduces to the sequence of equations

$$F_y - \frac{d}{dx}(F_{y'}) + \lambda\left(G_y - \frac{d}{dx}G_{y'}\right) = -\frac{d}{dx}(F_{y'}) + \lambda = 0;$$

$$\frac{d}{dx}(F_{y'}) = \lambda;$$

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

Solving the last equation for $y'(x)$ and integrating with respect to x, we have an expression for the solution function

$$y(x) = -\frac{1}{\lambda}\sqrt{1 - (\lambda x + \alpha)^2} + \beta. \tag{2.5}$$

The constants α, β and λ will be determined from the boundary conditions and the subsidiary condition as follows:

$$y(0) = -\frac{1}{\lambda}\sqrt{1 - (\alpha)^2} + \beta = 0, \tag{2.6}$$

$$y(1) = -\frac{1}{\lambda}\sqrt{1 - (\lambda + \alpha)^2} + \beta = 0, \tag{2.7}$$

$$\int_0^1 y(x)dx = \int_0^1\left(-\frac{1}{\lambda}\sqrt{1 - (\lambda x + \alpha)^2} + \beta\right)dx = A_0 > 0. \tag{2.8}$$

After integrating and substituting the limits of integration, and re-arranging, equation (2.8) becomes

$$\sin^{-1}(\lambda + \alpha) + (\lambda + \alpha)\sqrt{1 - (\lambda + \alpha)^2} - \alpha\sqrt{1 - (\alpha)^2} = 2\lambda^2(\beta - A_0). \quad (2.9)$$

Equations (2.6), (2.7) and (2.9) constitute a system of non-linear equations in the unknown parameter constants α, β and λ.

We now proceed and obtain an expression for A_0. From equations (2.6) and (2.7), we have

$$\sqrt{1 - (\alpha)^2} = \sqrt{1 - (\lambda + \alpha)^2} = \lambda\beta, \quad (2.10)$$

$$\lambda = -2\alpha, \quad \beta = -\frac{\sqrt{1 - \alpha^2}}{2\alpha}. \quad (2.11)$$

Substituting λ and β from (2.13) in (2.11) we have an expression for A_0 in terms of α.

$$A_0 = \frac{1}{4\alpha^2}[\sin^{-1}(\alpha) - \alpha\sqrt{1 - \alpha^2}], \quad (2.12)$$

where $0 < \alpha \leq 1$ and A_0 takes only positive values.

In the last equation, we are forced to restrict α to positive values, since the arc representing the solution equation (2.5) is recognized as a circular arc of a circle with center at $x = 1/2$ (Remark 3).

Equation (2.12) allows us to consider A_0 as a differentiable function of α; it is a monotonically increasing function, concave up, with maximum at $\alpha = 1$; maximum value there is $\frac{\pi}{8}$; it has a removable discontinuity at $\alpha = 0$ by L'Hôpital's rule. In other words, equation (2.12) defines A_0 as a 1-1 function in the closed interval $[0, 1]$; it has a unique solution for α if and only if $0 \leq A_0 \leq \frac{\pi}{8}$. This translates to the following.

Conclusion

The isoperimetric problem (IP) with $A_0 = 1$ has no solution; in fact (IP) has no solution whenever $A_0 > \frac{\pi}{8}$.

Remark 3. The solution $y(x) = -\frac{1}{\lambda}\sqrt{1-(\lambda x+\alpha)^2} + \beta,$ from equation (2.5), can be re-written and be recognized as a part of a circle, a circular arc, passing through the endpoints of the unit interval and centered on the vertical line $x = 1/2.$ Since the first part of equation (2.13) holds, we must have $\alpha > 0.$

Remark 4. This conclusion is a direct confirmation from the Calculus of Variations for the validity of Theorem 4 and Theorem 5 of Section I.

Remark 5. The validity of the results in Section I, also are rooted in the Calculus of Variations since they are based on the original isoperimetric problem of Queen Dido — a result proved in the Calculus of Variations.

References

[1] I. M. Gelfand and S. V. Fomin, Calculus of Variations, 232, Prentice Hall, Inc. Englewood Cliffs, NJ, 1963.

[2] W. R. Javier, Queen Dido's hide and minimal arc-length problems in calculus, paper presented at AMATYC, Conference, Salt Lake City, Utah, 2003.

[3] S. M. Ross, Probability Models, 8th ed., Academic Press, 2003.

[4] J. Stewart, Calculus Early Transcendentals, 5th ed., Thomson, 2010.

[5] Weir, Hass, Giordano: Thomas' Calculus, 11th ed., Pearson, 2004.