The Isoperimetric Problem Revisited: Extracting a Short Proof of Sufficiency from Euler's 1744 Approach to Necessity*

Richard Tapia†

April 16, 2013

Dedicated to Peter Lax in Recognition of his Numerous Mathematical Contributions

Abstract

Our primary objective in this paper is, with the student in mind, to present a short, elementary, and teachable solution of the isoperimetric problem. A secondary objective is to give a brief, but reasonably complete, overview of the remarkable life of the isoperimetric problem. In 1744 Euler constructed multiplier theory to solve the isoperimetric problem. However, contrary to Euler's belief, satisfaction of his multiplier rule is not a sufficient condition to demonstrate that the circle is a solution. In 1995 in a short paper in this Monthly aptly entitled A Short Path to the Shortest Path, Peter Lax constructed what is currently considered to be the shortest and most elementary of all existing proofs. This background material is presented to set the stage for our demonstration that Euler's approach can be extended to give a sufficiency proof that we believe to be short and elementary and therefore competitive with the Lax proof from this point of view.

1 Introduction.

1.1 The Road Ahead.

The primary goal of this paper, as our title states, is to show that Euler's 1744 approach to the isoperimetric problem can be extended to actually establish a proof of sufficiency, i.e., the circle solves the isoperimetric problem. Our sufficiency proof is surprisingly short and elementary.

*This paper was written in support of the author's plenary lecture presented at the SIAM Annual Meeting in Minneapolis, Minnesota in July of 2012. It is a revision of an original draft circulated in July of 2012.

†This work was supported in part by funds associated with the Maxfield-Oshman Chair in Engineering at Rice University.

Mathematical scholars immediately realized that Euler's 1744 approach did not establish sufficiency as he believed, but thought that instead he had established necessity for the circle to solve the isoperimetric problem. However, a point not widely appreciated is that unfortunately this also was not the case as we show in Section 2.3.

A secondary goal of our paper is to give a concise, but reasonably complete, review of the exciting life of the isoperimetric problem. A first by–product of our review is the belief that the isoperimetric problem is arguably the most influential mathematics problem of all time. A second by–product is the realization that the golden era of the construction of elementary optimization theory (the time period that starts with the establishment of Fermat's rule that the derivative must vanish at an extreme point say 1630 or so and ends with the elegant sufficiency theory of Weierstrass say 1890 or so) is completely void of definitions and discussions related to the important optimization notion of functional convexity and the powerful sufficiency theory that follows from this notion. For no good reasons, it just seems to have been left for the 20$^{\text{th}}$ century.

In Section 2 we give a brief historical development of the isoperimetric problem and its solution. This 2000 year historical journey will take us from the early Greeks to the 1995 Peter Lax short proof. In presenting Euler's theory we correct for several mathematical oversights and attempt to put the necessity theory that he is credited with on a solid footing by constructing a mathematically rigorous necessity theory.

In Section 3 we first present our extension of the Euler's approach and obtain sufficiency for the isoperimetric problem. We then present a second proof of sufficiency as a straightforward extension of our observation and the first variation arguments that Lagrange used in his derivation of the Euler-Lagrange equation for the simplest problem in the calculus of variations. We also discuss the likelihood that Euler or Lagrange could have made our observations.

Finally in Section 4 we offer concluding remarks.

1.2 Preliminary Statements.

It is interesting that throughout history mathematicians, and other scientists, who are certainly very aware of the difference between the formal mathematical notions of necessity and sufficiency fall into the trap of using necessity as sufficiency. More specifically in our presentation in Section 2 we will encounter the following rather infamous trap.

Consider the optimization problem

$$\text{maximize the real-valued functional } f \text{ over a set } P.$$

The Use of Necessity as Sufficiency Trap:

If for each p_0 contained in P and not equal to p^*, we can find a p^0 contained in P such that $f(p_0) < f(p^0)$, then p^* maximizes f over P.

Observe that what has been shown is that every member of P that is not p^* can not furnish a maximum to f over P. This is necessity, i.e., if the problem has a solution it must be p^*. However, we have not shown sufficiency, i.e., the problem has a solution. Now, if we know from other considerations that the problem has a solution, then it follows that the solution must be p^*. An example of the use of this trap is the following flawed proof that the number 1 is the largest of the counting numbers $1, 2, \ldots$. For any counting number p not equal to 1 we can construct the larger counting number p^2. Hence, 1 is the largest counting number. Of course no such largest number exists.

Another component in our historical presentation will be the important use of scholia (scholium). Originally scholia consisted of explanatory comments inserted in the margin of a written manuscript. However, they later included more formal extensive comments that were included in the main text of a manuscript. These latter comments could be viewed as what in today's mathematical writings we call remarks, observations, or informal conjectures. Scholia were used extensively by scholars in many disciplines from the early Greek period through the 19^{th} century. Indeed, much of Fermat's work was presented in the form of scholia including what we call today Fermat's last theorem.

2 The Historical Development of the Isoperimetric Problem and its Solution.

The contemporary literature abounds with information on the classical isoperimetric problem and related issues. Sources that have influenced the current presentation are H. H. Goldstine's well-known *History of the Calculus of Variations* [15], the little known masters thesis of Thomas Porter entitled *A History of the Classical Isoperimetric Problem* submitted to the the University of Chicago Mathematics Department in 1931 [27], the article *The Isoperimetric Problem* by Victor Blasjo [4] that appeared in this Monthly, and the various web notes of Jennifer Wiegert. We begin by defining the classical isoperimetric problem and two related problems.

Isoperimetric Problem. *Determine, from all simple closed planar curves of the same perimeter, the one that encloses the greatest area.*

Iso-Area Problem. *Determine from all simple closed planar curves that enclose the same area, the one with the smallest perimeter.*

Isoperimetric Inequality Problem. *Given a simple closed planar curve with perimeter L and enclosed area A show that*

$$4\pi A \leq L^2$$

with equality if and only if the curve is the circle.

It is a well-known fact, and is not difficult to prove, that these three problems are equivalent in the sense that if a given curve, e.g. the circle, solves one it solves the other. Of course isoperimetric problems restricted to a subclass of curves, for example triangles, rectangles, or n-sided polygons are of interest and appear throughout the early Greek literature. It appears to be an ancient bit of human knowledge that the solution of the isoperimetric problem is the circle. When first introduced to this problem, even the less mathematically initiated individuals readily conjecture that the solution is the circle. Perhaps because the isoperimetric problem is so easy to understand and yet mathematically so profound, it is unique among mathematics problems in that has been embraced by poets, historians, and other scholars in both ancient and medieval times and incorporated into their works. Of course the best known example is Virgils' epic Latin poem *The Anied* written in the ten year period 29 BC to 19 BC. Here Queen Dido negotiates for as much land as she can enclose with the hide of an ox. She then has her people cut the hide into thin strips and ties the strips together to form a long cord. Next, using the coastline as part of the boundary she forms as large a semi-circle as possible with the cord. Hence, she solved a version of the isoperimetric problem to obtain as much land as possible. Today many mathematicians call the isoperimetric problem Dido's problem, and this is particularly true for the version of the isoperimetric problem that has the semi-circle as solution.

2.1 The Early Greeks.

Isoperimetry (the study of geometric figures of equal perimeters) was a topic well embraced by the ancient Greeks. Yet the Greeks did not have a clear understanding of the relationship between perimeter and area. Proclus [28] claims that this lack of understanding led to cheating in land dealings. Moreover, the theorem that all triangles formed on the same base and always between the same two parallel lines are equal in area was considered paradoxical by the Greeks since the perimeter could be made as large as possible. In spite of this Greeks were outstanding geometers and proved that the equilateral triangle solved the isoperimetric problem for the triangle and that the square solves the isoperimetric problem for the rectangle; these latter results can be found in the writings of Euclid (330 BC - 260 BC).

While the origin of the isoperimetric problem should be attributed to the early Greeks, it is not known who among them was the first to state the problem, state the solution, or attempt a solution. Some historians claim that Pythagorus (580 BC - 500 BC) knew the maximum principle of the circle. However, Porter [27] claims that Pythagorus' knowledge was no deeper than believing that of all plane figures the circle is the most beautiful. Aristotle (384 BC - 322 BC) [2, p.287] remarks "Now, of lines which return upon themselves, the line which bounds the circle is the shortest." Porter [27] dismisses this statement saying that the subject Aristotle was concerned with was philosophical rather than mathematical. While we accept that this may be the case, we still have some difficulty completely divorcing Aristotle's statement from the iso-area problem. Mathematical historians tend to agree that Archimedes (287 BC - 212 BC) was well aware of the isoperimetric problem and its solution; however there is no agreement as to whether or not he attempted a proof; the literature is deficient on this issue.

2.2 Zenodorus.

We know that Zenodorus (200 BC - 140 BC) authored a book entitled *On Isoperimetric Figures*. This book was unfortunately lost, but the work has been partially preserved by Theon (335 AD - 405 AD) and Pappus (290 AD - 350 AD). See Heath [16]. The preserved work includes the following two theorems and their proofs.

Theorem 2.1. *Among all polygons of equal number of sides and equal perimeters, the regular polygon encloses the greatest area.*

Theorem 2.2. *The circle encloses a greater area than any regular polygon of equal perimeter.*

Porter [27] notes that Zenodorus assumed existence of a solution in his proof of Theorem 2.1 and this gap in his proof was corrected by Weierstrass [30, pp.70-75] some two thousand years later. Historians and mathematicians alike credit Zenodorus with the first attempt to prove that the circle solves the isoperimetric problem, claiming that the proof either contained a flaw or was incomplete. However, that Zenodorus attempted a proof can not be validated from looking at his work preserved by Theon or Pappus. Hence, it could just be that what some are referring to as an incomplete proof is merely the proof of Theorem 2.2 stated above. However, we believe that it is more likely that Zenodorus merely stated that the circle solves the isoperimetric problem in two dimensions and the sphere solves the isoperimetric problem in three dimensions as scholia and did not attempt a formal proof.

2.3 Euler's Multiplier Rule and the Isoperimetric Problem.

To begin with we use the word extremize to mean either minimize or maximize. By the simplest problem in the calculus of variations we mean the optimization problem

$$\begin{aligned} &\text{extremize} \quad F(y) = \int_a^b f(x, y(x), y'(x))\,dx \\ &\text{subject to} \quad y(a) = \alpha \text{ and } y(b) = \beta. \end{aligned} \tag{1}$$

It is standard to assume that $f : \mathbb{R}^3 \to \mathbb{R}$ has continuous first-order partial derivatives. Moreover, we are interested in functions $y : [a, b] \to \mathbb{R}$ which are continously differentiable on the interval $[a, b]$.

First Euler [10] in 1744 and later Lagrange [22] in 1759, in a different manner which we will discuss in Section 3, demonstrated that a solution of problem (1) must satisfy the so-called Euler-Lagrange equation

$$f_y(x, y, y') - \frac{df_{y'}(x, y, y')}{dx} = 0, \;\; y(a) = \alpha \text{ and } y(b) = \beta. \tag{2}$$

In (2) the subscripts denote the respective partial derivatives. Solutions of the Euler-Lagrange equation are called *extremals* of problem (1). In Chapter V of [10] Euler explains how to handle problems where, in addition to the boundary conditions, the solution must satisfy a subsidiary condition (constraint). He considered the problem

$$\begin{aligned} &\text{extremize} \quad F(y) = \int_a^b f(x,y,y')\,dx \\ &\text{subject to} \quad G(y) = \int_a^b g(x,y,y')\,dx = \ell \\ &\qquad\qquad y(a) = \alpha \text{ and } y(b) = \beta. \end{aligned} \tag{3}$$

The standard assumptions made for problem (1) carry over to problem (3). In the case that F is area and G is arc length problem (3) is the standard isoperimetric problem. Historically, problem (3) has been called a general isoperimetric problem, and the constraint $G(y) = \ell$ has been called a general isoperimetric constraint, even though it may not represent arc length. Euler derived the rule which we call

Euler's Multiplier Rule:

If y^* is a solution of the general isoperimetric problem (3), then there exists an associated multiplier λ such that y^* is an extremal of the auxiliary problem

$$\begin{aligned} &\textit{extremize} \quad L(y) = F(y) + \lambda\, G(y) \\ &\textit{subject to} \quad y(a) = \alpha \textit{ and } y(b) = \beta. \end{aligned} \tag{4}$$

We say that y^* is an extremal of problem (3) if there exists an associated multiplier λ so that y^* is an extremal of problem (4) with this choice of λ.

We now consider the isoperimetric problem in Queen Dido form

$$\begin{aligned} &\text{maximize} \quad \int_{-a}^{a} y(x)\,dx \\ &\text{subject to} \quad \int_{-a}^{a} \sqrt{1 + y'(x)^2}\,dx = a\pi \\ &\qquad\qquad y(-a) = y(a) = 0. \end{aligned} \tag{5}$$

Euler observed that the semi-circle

$$y_c(x) = \sqrt{a^2 - x^2},\ -a \le x \le a \tag{6}$$

is an extremal of the Multiplier Rule auxiliary problem

$$\begin{aligned} &\text{maximize} \quad J(y) = \int_{-a}^{a} \left(y(x) - a\sqrt{1 + y'(x)^2} \right) dx \\ &\text{subject to} \quad y(-a) = y(a) = 0. \end{aligned} \tag{7}$$

Problem (7) corresponds to a multiplier choice of $\lambda = -a$ in the Euler auxiliary problem for problem (5). Notice that for problem (7), with f denoting the obvious quantity, we have

$$f_y = 1 \quad \text{and} \quad f_{y'} = -a\frac{y'(x)}{\sqrt{1+y'(x)^2}}. \tag{8}$$

Now evaluating these quantities for the semi-circle (6) we see that

$$\frac{y_c'(x)}{\sqrt{1+y_c'(x)^2}} = -\frac{x}{a}. \tag{9}$$

Hence $f_{y'} = x$ and the semi-circle satisfies the Euler-Lagrange equation (2) for problem (7).

Euler's Genius, Naiveté, and Generousity.

At this juncture some remarks on historical points are in order. First of all Goldstine [15, p.73] states

> "It is interesting that Euler did not completely understand the fact that his condition [satisfaction of the Euler-Lagrange equation (2)] is a necessary but not a sufficient one. In his discussion it is clear the he felt his condition was sufficient to ensure an extremum, and that by evaluating the integral along an extremal [and one other curve] he could decide whether it was a maximum or a minimum."

There is a flaw in both Euler's and Lagrange's derivation of the Euler-Lagrange equation (2). Specifically in writing the total derivative term in (2) they implicitly assumed that a solution was twice differentiable. A rigorous derivation of the Euler-Lagrange equation (2), including the existence of the total derivative, was first given by du Bois-Reymond [7] in 1879, some 135 years later. In the interim practitioners merely assumed that the solution had two derivatives, which it essentially always did.

Above we used the terminology Euler's Multiplier Rule to distinguish it from what today is called Euler's Rule in number theory. Weierstrass [30] simply called it Euler's Rule, as did most authors of the time, while Bolza [5, p.269] called it Euler's Isoperimetric Rule. Concerning this multiplier rule Goldstine [15, p.74] remarks

> "What he [Euler] finds is a simple-but the first-instance of a multiplier rule. Perhaps Lagrange multipliers ought to be called Euler-Lagrange multipliers just as the first necessary condition for general problems is often referred to as the Euler-Lagrange equation."

While Goldstine describes this first instance as simple, Carathéodory who edited Euler's great works of 1744 said that Euler's analysis of the constrained problem is an achievement of first class and

> "A major accomplishment that even an Euler did not achieve very often,"

see [14, p.132]. Unfortunately as Euler moved forward to problems with multiple isoperimetric constraints he found that his geometric techniques were sufficiently challenging to use and Carathéodory identified the flaw in Euler's theory that he had failed to demonstrate the fact that the multipliers were constants, a fact that he later used in his proof. Carathéodory aptly comments in his review

> "it is a pity that this work which contains so many novel ideas would end in this fashion on a discordant note,"

see Goldstine [15, pp.100-101].

Then how did the name Lagrange become attached to the greatly innovative work of Euler? The answer lies with Euler himself. According to Goldstine [15, p.110] in August of 1755 a 19 year old Lagrange wrote Euler a brief letter to which was attached an appendix containing mathematical details of a beautiful and revolutionary idea, the notion of variations. See [21]. This idea could be used to remove Euler's tedium and need for geometrical insight and all could be done with analysis using Lagrange's new notion of variations. Euler was so impressed that he dropped his own methods, expoused those of Lagrange, named the subject the calculus of variations, and called multiplier theory Lagrange multiplier theory. Indeed today the Euler auxiliary functional $L(y)$ given in problem (4) is called the Lagrangian: how fair can that be?

We ask, does Euler have the right to implicitly give up credit for his original contributions simply because Lagrange came up with a cleaner way of driving the theory? Many of my colleagues and students think he does, but we think not. Euler's contribution was immediately etched in the pages of history; no mortal or immortal has the right to make changes to these pages. While it is probably too late to credit correctly, we move to the usage Euler-Lagrange multiplier theory in place of the usage Lagrange multiplier theory.

Euler's Oversight: Promoting Rigor in the Statement of the Isoperimetric Problem.

Today we know that it was Weierstrass who first promoted badly needed rigor in the calculus of variations. Carathéodory makes the following comments concerning Weierstrass' contribution to the promotion of rigor (taken from Giaquinta and Hildebrandt [14, p.51]):

> "In his earlier work, prior to the year 1879, he succeeded in removing all the difficulties that were contained in the old investigations of Euler, Lagrange, Legendre, and Jacobi, simply by stating precisely and analyzing carefully the problems involved. In improving upon the work of these men he did several things of paramount importance . . .
>
> (1) he showed the advantages of parametric representation;

(2) he pointed out the necessity of first defining in any treatment of a problem in the Calculus of Variations the class of curves in which the minimizing curve is to be sought, and of subsequently choosing the curves of variation so that they always belong to this class."

Euler's treatment of the isoperimetric problem is perhaps the prime example of the lack of rigor as described above. Hence, we accept as our first task the introduction of rigor in Euler's treatment of the isoperimetric problem.

Consider the Queen Dido statement of the isoperimetric problem given by (5). We observe that the semi-circle given by (6) has infinite slope at the end points $x = a$ and $x = -a$; so the arc length integral in problem (5) is not defined. Throughout the years, beginning with Euler and including contemporary times, authors have swept this undesirable aspect of the problem under the rug and just ignored it. Weierstrass dealt with the situation by stating the problem in parametric form. His treatment can be found in Bolza [5, pp.210-211]. Many authors, for example Elsgolc [9, pp.139-143] consider the Euler-Lagrange equation for the Euler's auxiliary problem (7) as Euler did, but in order to solve the Euler-Lagrange equation they make a change of variables that leads to the solution in parametric form, and then they argue that elimination of the parameter gives the equation for a circle. Yet still other authors, for example Gelfand and Fomin on page 49 of their very popular text [13] consider the Euler-Lagrange equation as Euler did and then simply infer that integration of the equation leads to a family of circles.

In support of Euler we now define a proper class of functions as the domain of the isoperimetric problems (5) and (7), show that the semi-circle curve belongs to this class, the curves of variation (used for the derivation of the Euler-Lagrange equation) belong to this class, and the class contains a large selection of rectifiable curves. We then extend the definition of Euler's auxiliary problem (7) to this larger class, rigorously derive the Euler-Lagrange equation as a necessary condition for a solution, and show that the semi-circle function (6) satisfies this Euler-Lagrange necessity condition.

An immediate first thought is to conjecture that the arc length integral for the semi-circle exists as an improper Riemann integral. This conjecture is correct as we now demonstrate. We work with the quarter circle in the first quadrant, so all our limits of integration are from 0 to $a-\epsilon$ for small $\epsilon>0$. The quantity ϵ is introduced to make the improper integrals proper as is usually done in these type of arguments. Then we let ϵ go to zero.

To begin with if y_c is the semi-circle curve (6), then

$$\sqrt{1+y_c'(x)^2} = \frac{a}{\sqrt{a^2-x^2}}.$$

Hence,

$$\int_0^{a-\epsilon} \sqrt{1+y_c'(x)^2} = \int_0^{a-\epsilon} \frac{a}{\sqrt{a^2-x^2}} \quad = \int_0^{a-\epsilon} \frac{a}{\sqrt{a-x}\sqrt{a+x}}$$

$$\leq \sqrt{a}\int_0^{a-\epsilon} \frac{1}{\sqrt{a-x}} = 2a - 2\sqrt{a\epsilon}.$$

Letting $\epsilon \to 0$ we have that the arc length is less than or equal to $2a$. The inequality used the facts that $\frac{1}{\sqrt{a+x}} \leq \frac{1}{\sqrt{a}}$ for $x \in [0, a]$ and the fact that the anti-derivative of $\frac{1}{\sqrt{a-x}}$ is $-2\sqrt{a-x}$. So we have demonstrated that the arc length integral exists as an improper integral.

Now since the semi-circle curve is obviously rectifiable, we ask if the arc length is given by the improper arc length integral. It happens to be the case that it is; but we will not prove it directly since it will follow from our subsequent theorem. These preliminary considerations serve as motivation for our construction of the Euler class and the establishment of some of its properties. The fact that the arc length integral for the semi-circle curve exists as an improper Riemann integral also follows from the theorem that we are about to consider. However, we felt that it would be satisfying and instructive to first prove it directly.

Definition 2.3 (The Euler Class)**.** *By $E(-a, a)$, the Euler class of curves for the isoperimetric problem (5), we mean the collection of $y : [-a, a] \to \mathbb{R}$ satisfying the following conditions*

(A) $y(-a) = y(a) = 0$,

(B) y *is continuous on* $[-a, a]$,

(C) y *is differentiable except possibly on a countable subset of* $[-a, a]$,

(D) the arc length integral

$$\int_{-a}^{a} \sqrt{1 + y'(x)^2}\, dx \tag{10}$$

(d_1) exists as a proper Riemann integral or

(d_2) the curve y is rectifiable and the arc length integral exists as an improper Riemann integral.

Theorem 2.4.

(i) If y and η belong to the Euler class, then so do the variations $y+t\eta$ for $t \in [0, 1]$.

(ii) If y is continuous and concave on $[-a, a]$, as is the case with the semi-circle function (6), then y belongs to the Euler class.

(iii) Each member of the Euler class is rectifiable and its arc length is given by its corresponding proper or improper arc length integral (10).

Proof of (i). We are interested in admissible variations with respect to problem (7). Since the boundary conditions are homogenous, the admissible variations are the elements of $E(-a, a)$ themselves. Moreover, each admissible variation at $y_1 \in E(-a, a)$

can be written as $y_2 - y_1$ for some $y_2 \in E(-a,a)$. Hence it is sufficient to show that if $y_1, y_2 \in E(-a,a)$, then

$$y_1 + t(y_2 - y_1) \in E(-a,a) \text{ for } t \in [0,1].$$

Towards this end for $t \in [0,1]$ and real x and y consider the inequality

$$\begin{aligned} \left[1 + (tx + (1-t)y)^2\right]^{1/2} &\leq t(1+x^2)^{1/2} + (1-t)(1+y^2)^{1/2} \\ &\leq (1+x^2)^{1/2} + (1+y^2)^{1/2}. \end{aligned} \tag{11}$$

Straightforward expansion reduces the first inequality in (11) to

$$1 + xy \leq (1+x^2)^{1/2}(1+y^2)^{1/2},$$

and this latter inequality is verified by squaring both sides. The second inequality in (11) is direct.

For the sake of convenience we introduce the notation

$$A(y,\epsilon) = \int_{-a+\epsilon}^{a-\epsilon} \sqrt{1 + y'(x)^2}\, dx \tag{12}$$

for y contained in the Euler class $E(-a,a)$ and $0 \leq \epsilon < a$. The notation $A(y,0)$ denotes the improper Riemann integral. Hence by definition

$$A(y,0) = \lim_{\epsilon \to 0} A(y,\epsilon). \tag{13}$$

Now turning to inequality (11) we see that for any $t \in [0,1]$ and any $\epsilon \in (0,a)$ we have

$$A(y_1 + t(y_2 - y_1), \epsilon) \leq A(y_1,\epsilon) + A(y_2,\epsilon);$$

and letting $\epsilon \to 0$ establishes part (i) of the theorem. What we actually have proved is that the arc length integral is a convex function; hence it has its maximum at the end points.

Proof of (ii). Let y be a continuous concave function which vanishes at $x = -a$ and $x = a$. We need to show that conditions (C) and (D) of Definition 2.3 hold. From [20, p.579] we see that condition (C) holds, and that y' is a decreasing function. Our entry into condition (D) consists of first showing that y is rectifiable. Towards this end observe that there exists $b \in [-a,a]$ such that y is monotone increasing on $[-a,b]$ and monotone decreasing on $[b,a]$. The two regions correspond to subintervals where f' is nonnegative and nonpositive. A monotone function is of bounded variation on its domain. The concave function y is monotone on the subintervals $[-a,b]$ and $[b,a]$, hence of bounded variation on $[-a,a]$ and therefore rectifiable on $[-a,a]$, see [20, p.564]. We introduce the notation $L(y,\epsilon)$ to denote the arc length of the curve y restricted to the domain $[-a+\epsilon, a-\epsilon]$. Now y is absolutely continuous on $[-a+\epsilon, a-\epsilon]$, see [20, p.578] and therefore

$$L(y,\epsilon) = A(y,\epsilon), \tag{14}$$

see [20, p.564]. From [20, p.578] we know that $L(y,\epsilon)$ is continuous in ϵ. Therefore by letting $\epsilon \to 0$ in (14) we see that for a concave curve the arc length is necessarily given by the improper Riemann arc length integral. This establishes part (ii) of our theorem.

Proof of (iii). We first consider the case where the arc length integral exists as a proper Riemann integral. Recall Lebesgue's well-known criterion for Riemann integrability, i.e., a function f defined on an interval $[a, b]$ is Riemann integrable if and only if f is bounded and the set of discontinuities of f has (Lebesgue) measure zero. Since the arc length integral exists as a Riemann integral it follows that y' is Riemann integrable on $[-a, a]$. Moreover, since y is continuous, differentiable except on a countable subset of $[-a, a]$, and y' is Riemann integrable we have from [20, p.556] that y can be written as the indefinite integral of its derivative; hence from [20, p.550] y is absolutely continuous. It follows that y is of bounded variation and therefore rectifiable, see pages 546 and 564 of [20], and that the arc length is given by the arc length integral, see [20, p.564].

We now consider the second case where the curve y is rectifiable and the arc length integral exists as an improper Riemann integral. The argument just given can be used to show that (14) holds. Now since y is rectifiable $L(y,\epsilon)$ is continuous in ϵ, see [20, p.564]. Hence letting $\epsilon \to 0$ in (14) shows that the arc length is given by the improper Riemann integral. We have proved our theorem. ■

Remark. Not all rectifiable curves are members of the Euler class. For example, the well-known Volterra function and Cantor ternary function defined on the interval $[0, 1]$ are known to be continuous and rectifiable, but the arc length integral does not exist either as a proper or improper Riemann integral. Granted these two curves are quite pathological. For the Cantor ternary curve the arc length integral exists as a Lebesgue integral and gives the value 1, but the arc length of the curve is equal to 2. This shows that our theorem is not valid if we replace Riemann with Lebesgue in condition (D) in the definition of the Euler class.

Remark. Jim Case provided us with a very clever proof that a continuous concave function is rectifiable. He first showed that a bounded convex polygon has the property that its perimeter is bounded by twice its diameter. Hence for a continuous and concave curve, twice the diameter is an upper bound for the length of any polygonal approximation to the curve; hence to the arc length of the curve as the supremum of the lengths of such approximations.

We are now in a position to provide a proper statement and valid proof of Euler's statement concerning problem (7) and the semi-circle function (6). But first recall that the standard derivation of the Euler-Lagrange equation requires that the functions under consideration be continuously differentiable on the interval of interest; while members of the Euler class defined in Definition 2.3 are only guaranteed to be continuously differentiable almost everywhere. This follows from our assumption that the arc length integral exists as a proper or improper Riemann integral. These statements serve to qualify the following theorem.

Theorem 2.5. *The functional*

$$J(y) = \int_{-a}^{a} f(x, y, y')\, dx \tag{15}$$

with

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

and the integral interpreted as an improper Riemann integral is well defined for y contained in the Euler class $E(-a, a)$.

Moreover, if $\hat{y}$, a member of the Euler class, extremizes J in this class, then it necessarily satisfies the Euler-Lagrange equation in integral form

$$f_{y'}(x, y, y') - \int_{-a}^{x} f_y(t, y, y')\, dt = c, \ \ y(-a) = y(a) = 0 \tag{16}$$

at points of continuity of $\hat{y}'$ and such points exist almost everywhere on $(-a, a)$. In (16) c represents a constant.

Finally, the semi-circle function $y(x) = \sqrt{a^2 - x^2}$ is a member of the Euler class, has a continuous derivative on $(-a, a)$, and in addition to satisfying (16) at all points in $(-a, a)$ satisfies the Euler-Lagrange equation in derivative form on $(-a, a)$

$$f_y(x, y, y') - \frac{df_{y'}(t, y, y')}{dx} = 0, \ \ y(-a) = y(a) = 0. \tag{17}$$

Proof. The first part of the theorem follows from condition (D) in Definition 2.3. Now, suppose that $\hat{y} \in E(-a, a)$ extremizes J in this class. We are interested in variations $\eta \in E(-a, a)$ which have the property that their arc length integral is given by the proper Riemann arc length integral. It follows that for such a variation η, its derivative η' is Riemann integrable on $[-a, a]$. Hence η is absolutely continuous. Choose such an η and consider

$$\phi(t) = J(\hat{y} + t\eta) \text{ for } t \in [0, 1]. \tag{18}$$

We know that $\hat{y} + t\eta$ is contained in $E(-a, a)$ from (i) of Theorem 2.4, but it will also follow from the analysis we are about to present. Since $\hat{y}$ is an extremizer, we can conclude that $\phi'(0) = 0$ once we demonstrate that such a derivative exists. Towards this end we have

$$\begin{aligned}
\phi'(0) &= \int_{-a}^{a} \eta - a \lim_{t\to 0}\lim_{\epsilon\to 0}\int_{-a+\epsilon}^{a-\epsilon} \frac{1}{t}\left[\sqrt{1+(\hat{y}'+t\eta')^2}-\sqrt{1+(\hat{y}')^2}\right] && (19)\\
&= \int_{-a}^{a} \eta - a \lim_{t\to 0}\lim_{\epsilon\to 0}\int_{-a+\epsilon}^{a-\epsilon} \frac{(\hat{y}'+\theta t\eta')\eta'}{\sqrt{1+(\hat{y}'+\theta t\eta')^2}} \text{ for some } \theta\in(0,1) && (20)\\
&= \int_{-a}^{a} [\eta - a\frac{\hat{y}'\eta'}{\sqrt{1+(\hat{y}')^2}}] && (21)\\
&= -\int_{-a}^{a} [x + a\frac{\hat{y}'}{\sqrt{1+(\hat{y}')^2}}]\eta'. && (22)
\end{aligned}$$

Hence $\phi'(0)$ exists; so we may conclude that $\phi'(0) = 0$. It follows that for any constant c

$$\int_{-a}^{a} [x + a\frac{\hat{y}'}{\sqrt{1+(\hat{y}')^2}} - c]\eta' = 0. \qquad (23)$$

We obtained (20) from (19) by way of the mean-value theorem in the variable t, (21) from (20) by first noticing that the improper integral is actually a proper integral since the integrand is bounded and continuous almost everywhere (recall Lebesgue's criterion for Riemann integrability) and then passing the limit in t under the integral again justified by the fact that the integrand is bounded, and we obtain (22) from (21) by a simple integration by parts justified since η is absolutely continuous (see [20, p.553]).

Let

$$g(x) = x + a\frac{\hat{y}'(x)}{\sqrt{1+(\hat{y}'(x))^2}} \qquad (24)$$

and

$$\hat{c} = \frac{1}{2a}\int_{-a}^{a} g(x)\,dx. \qquad (25)$$

Define

$$\hat{\eta}(x) = \int_{-a}^{a} g(t)\,dt - (x+a)\hat{c}. \qquad (26)$$

Observe that $\hat{\eta}\in E(-a,a)$ and $\hat{\eta}$ is absolutely continuous; hence it is rectifiable and its arc length is given by the arc length integral, see [20, p.564]. So $\hat{\eta}$ qualifies as one of the variations under consideration. Hence making the choices $c = \hat{c}$ and $\eta = \hat{\eta}$ in (23) we obtain

$$\int_{-a}^{a} [x + a\frac{\hat{y}'}{\sqrt{1+(\hat{y}')^2}} - \hat{c}]^2 = 0. \qquad (27)$$

Now, it is well known that if the integral of a non-negative function is zero, then the function must be zero at points of continuity, otherwise we could construct a positive contribution to the integral. Hence recalling the form of the partial derivatives of f from (8) we see that (27) leads to (16).

The last part of our theorem follows by observing that for the semi-circle function (16) holds for all $x \in (-a, a)$. Now, since two terms in (16) are differentiable with respect to x, the third term must also be, and such a differentiation leads to (17). This proves our theorem. ■

Remark. We summarize what we have demonstrated. Euler showed us that the semi-circle function satisfied the Euler-Lagrange equation (17). He believed that such a demonstration established sufficiency, i.e., the semi-circle solved the isoperimetric problem. Scholars immediately recognized that Euler's demonstration did not imply sufficiency, and maintained that it only implied necessity, i.e., if the isoperimetric problem had a solution it must be the semi-circle. However, this conclusion was also deficient, for the objective function was not defined for the semi-circle function. Our Theorem 2.5 shows that indeed satisfaction of the Euler-Lagrange equation is a necessary condition. Moreover, in Section 3 we show essentially that satisfaction of the Euler-Lagrange equation by the semi-circle function does imply sufficiency, and in some sense completely redeem Euler.

2.4 Steiner.

By the 19^{th} century the mathematical community had at its disposal the well-established inquiry tools of synthetic geometry (Euclidean geometry based) and the more recently developed inquire tools of analysis (calculus based). Moreover, the mathematical value of isoperimetry was well-embraced and the isoperimetric problem represented activity from both worlds. Indeed the mathematical world was well aware of the short comings of Zenodorus' incomplete proof using geometry and the shortcomings of Euler's work that used analysis to establish necessity and not sufficiency. At this juncture analysis was sufficiently mature to become a rival, at times bitter, with geometry for providing tools for mathematical thought and proof.

Jakob Steiner (1796 AD -1867 AD) was one of the most brilliant and creative geometers in history. His mathematical inquiries were confined to geometry to the total exclusion of analysis. In fact he hated analysis and doubted if anything that could not be proved with geometry could be proved with analysis. In 1838 Steiner [29] gave the first of his five equivalent proofs that the circle solved the isoperimetric problem. His proofs used synthetic geometry and were mathematically quite elegant. Mathematical historians embrace and promote his proofs and call them models of mathematical ingenuity. They became very visible in the mathematical community. Stenier boasted that he had done with geometry what had not been done, and could not be done, with analysis; i.e., solve the isoperimetric problem. What Steiner proved was that any curve which was not the circle could be modified using a geometric procedure now called *Steiner symmetrization* to obtain a curve with the same perimeter but a larger area. He then concluded that as a consequence the circle must be ths solution to the isoperimetric problem. As such he fell into the use of necessity as sufficiency trap described above and made the trap rather infamous. The analysts of the time, led by Peter Dirichlet, pointed out to Steiner that his proof is not valid unless he assumes that the isoperimeteric problem has a solution, i.e., existence. Steiner did not accept this criticism well, and rebutted with a very superficial argument that he claimed demonstrated that the isoperimetric problem must have a solution. Acutually, at best he demonstrated the existence of an upper bound for the area functional. Now analysts observed that if the Steiner symmetrization process could be applied repeatedly creating a sequence of curves with the same perimeter but increasing area that converges to the area of the circle, the so-called process of *completing Steiner's proof*, then the flaw in Steiner's proof would be removed.

Porter [27, p.12] mentions the paper [1] by an anonymous author that appeared in 1823, fifteen years prior to Steiner, and uses geometric arguments not unlike Steiner arguments, working with the iso-area problem. This proof, like Steiner's, suffers from this flaw of using necessity as sufficiency. We find this quite interesting, since the standard literature seems not to mention this paper. Moreover, we do not comprehend why someone would write an anonymous paper on such an important topic. Of course Newton may have done such a thing, but by this time Newton had been gone for 100 years.

Following Weierstrass' 1879 solution of the isoperimetric problem, that we describe in Section 2.5, the literature exploded with valid proofs of the isoperimetric problem. Some completed Steiner's proof using either geometry or analysis; while others went off in interesting mathematical directions and gave proofs completely unrelated to Steiner's work. We find it quite interesting that while the time period between Steiner's highly visible flawed proof and Weierstrass's valid proof was 40 years, and while incomplete proofs were offered in this time period, no authors provided a complete proof. In passing we comment on a few of the more visible solutions of the isoperimetric problem that followed Weierstrass' solution. Perhaps the first work that completed Steiner's proof was given by Edler [8] in 1882. He used geometry to make the completion. In 1909 in a two-part paper co-authored by Carathéodory and Study [6], Carathéodory completed Steiner's proof employing a most elegant and beautiful analysis argument based on the compactness of sets on the real line. Study completed Steiner using geometrical arguments. A highly visible proof that used Fourier series analysis was given by Hurwitz [17] in 1902 and a similar proof (undoubtedly influenced by Hurwitz' proof) was used by Lebesgue [25] in his 1906 text. For a list of references fairly complete up to the year 1931 see Porter [27].

2.5 Weierstrass' Sufficiency Proof.

Weierstrass, in addition to his numerous mathematical contributions, is known for introducing rigor of proof and cleanliness of definition into the calculus of variations at a time that it was sorely needed. He did not publish his work in this area but developed a complete and polished set of lecture notes that he used in his university courses at the University of Berlin. Today we know about his many contributions in the calculus of variations from his collected works [30] which was constructed primarily from the lecture notes of his numerous students during the time period 1865-1890. It is alleged that Weierstrass had 40 or so students during this time period and many of these students became quite distinguished in their own right; for example Cantor, Frobenius, (Sofia) Kowalewski (as a woman she was not officially accepted as a student at the University of Berlin and was given an honorary degree), Mittag-Leffler, Runge, Schur, and Schwarz.

Some believe that Weierstrass' critical sense and need to base his analysis on such a firm foundation led him to continually revise and perfect his writings to the point that publication was precluded. In spite of this his work became so well known that today he is often called the father of analysis.

Concerning the solution of the isoperimetric problem, Weierstrass was quite aware of the shortcomings of Steiner's proof and somewhat bothered by Steiner's arrogantly promoted negative view of analysis and analysts. Hence he boldly and proudly placed himself in the noble role of defender of analysis and vowed to solve the isoperimetric problem using analysis. He introduces his work with the following statements concerning the solution of the isoperimetric problem [30, p.259]

> "A detailed discussion of this problem is desirable, since Steiner was of the opinion that the methods of the calculus of variations were not sufficient to give a complete proof, but the calculus of variations is in a position to prove all this, as we will show later; furthermore it can show what Steiner could not – that such a maximum really exits."

The following comments are derived essentially from Chapter 5 of Goldstine [15]. Weierstrass first builds an elegant and sophisticated sufficiency theory for the simplest problem from the calculus of variations (2) employing such subtle notions as Jacobi's notion of conjugate points and his own notion of fields of extremals. He then extends this sufficiency theory to the isoperimetric problem (3) by turning to Euler's multiplier rule that we presented in Section 2.3 and applying his sufficiency theory to the Euler auxiliary problem (4). Using this theory he demonstrates that Euler's auxiliary problem for the isoperimetric problem has the circle as solution; hence the isoperimetric problem has the circle as solution. So, according to the literature, some 135 years after Euler's proof of necessity we have the first sufficiency proof.

While this notable work gave the world its first sufficiency proof for the isoperimetric problem, we expect to convey to the reader in Section 4 that Weierstrass really used a sludge hammer to pound a nail. His sophisticated sufficiency theory is not needed to merely demonstrate that the circle solves Euler's auxiliary problem for the isoperimetric problem.

2.6 Peter Lax's Short Proof of Sufficiency.

The following proof is taken verbatim from Lax [24], 1995.

Lax's Proof of the Isoperimetric Inequality.

Let $x(s)$, $y(s)$ be the parametric representation of the curve, s arc length, $0 \le s \le \pi$. Suppose that we have so positioned the curve so that the points $(x(0), y(0))$ and $(x(\pi), y(\pi))$ lie on the x-axis, i.e.,

$$y(0) = y(\pi) = 0.$$

The area enclosed by the curve is given by the formula

$$A = \int_0^\pi y\dot{x}\, ds,$$

where the dot denotes differentiation with respect to s.

According to a basic inequality

$$ab \le \frac{a^2 + b^2}{2};$$

equality holds only when $a = b$. Applying this to $y = a$, $\dot{x} = b$, we get

$$A = \int_0^\pi y\dot{x}\, ds \le \frac{1}{2} \int_0^\pi (y^2 + \dot{x}^2)\, ds. \tag{28}$$

Since s is arc length, $\dot{x}^2 + \dot{y}^2 = 1$; and we can rewrite (28) as

$$A \le \frac{1}{2} \int_0^\pi (y^2 + 1 - \dot{y}^2)\, ds. \tag{29}$$

Since $y = 0$ at $s = 0$ and π, we can factor y as

$$y(s) = u(s) \sin\, s, \tag{30}$$

u bounded and differentiable. Differentiate (30):

$$\dot{y} = \dot{u} \sin s + u \cos s.$$

Setting this into (29) gives

$$A \le \frac{1}{2} \int_0^{2\pi} \left(u^2(\sin^2 s - \cos^2 s) - 2u\dot{u} \sin s \cos s - \dot{u}^2 \sin^2 s + 1\right)\, ds. \tag{31}$$

The product $2u\dot{u}$ is the derivative of u^2; integrating by parts changes (31) into

$$A \le \frac{1}{2} \int_0^{2\pi} (1 - \dot{u}^2 \sin^2 s)\, ds, \tag{32}$$

clearly $\le \frac{\pi}{2}$ and equality holds only if $\dot{u} = 0$, which makes $y(s) \equiv \text{constant} \sin s$. Since equality in (32) holds only if $y = \dot{x} = \sqrt{1 - \dot{y}^2}$, $y(s) \equiv \pm \sin s$, $x(s) \equiv \mp \cos s + \text{constant}$.

This is a semi-circle. ■

The Lax proof is considered to be the shortest and most elementary proof of the isoperimetric inequality. While the proof is certainly ingenious, instructive and serves the purpose of providing a short proof, it is far from intuitive and is void of any flavor of optimization theory. More will be said about this in our concluding section, Section 4.

3 Two New Proofs of Sufficiency Motivated by Euler and Lagrange Respectively.

3.1 A Proof Motivated by Euler.

Theorem 3.1. *The semi-circle curve $y_c(x)$ given by (6) uniquely solves the isoperimetric problem (5), with the arc length integral interpreted as an improper Riemann integral, over $E(-a,a)$, the Euler class of functions given by Definition 2.3.*

Proof. Consider the objective function in Euler's auxiliary problem (7) for the isoperimetric problem (5),

$$J(y)=\int_{-a}^{a}\left(y(x)-a\sqrt{1+y'(x)^2}\right)\,dx, \tag{33}$$

and the semi-circle

$$y_c(x)=\sqrt{a^2-x^2},\ \text{for } -a\le x\le a. \tag{34}$$

For the sake of convenience we will always consider the integral in (33) as an improper integral, and nothing is lost if it exists as a proper integral.

Now consider any $y\neq y_c$ contained in the Euler class and let η denote $y-y_c$. Define

$$\phi_\epsilon(t)=\int_{-a+\epsilon}^{a-\epsilon}\left[y_c+t\eta-a\sqrt{1+(y_c'+t\eta')^2}\right]\,dx \tag{35}$$

for $t\in[0,1]$ and $\epsilon\in(0,a)$.

Straightforward differentiations give

$$\phi_\epsilon'(t)=\int_{-a+\epsilon}^{a-\epsilon}\left[\eta-a\,\frac{(y_c'+t\eta')\eta'}{\sqrt{1+(y_c'+t\eta')^2}}\right]\,dx, \tag{36}$$

and

$$\phi_\epsilon''(t)=-a\int_{-a+\epsilon}^{a-\epsilon}\frac{(\eta')^2}{[1+(y_c'+t\eta')^2]^{3/2}}\,dx. \tag{37}$$

Taylor's Theorem tells us that

$$\phi_\epsilon'(1) = \phi_\epsilon(0) + \phi_\epsilon'(0) + \frac{1}{2}\,\phi_\epsilon''(\theta) \text{ for some } \theta \in (0,1). \tag{38}$$

Recalling (9) we see that (36) gives

$$\phi_\epsilon'(0) = \int_{-a+\epsilon}^{a-\epsilon} (\eta + x\eta')\,dx = -x\eta(x)\Big|_{-a+\epsilon}^{a-\epsilon},$$

and because η is continuous $\phi_\epsilon'(0) \to 0$ as $\epsilon \to 0$. Also observe that $\phi_\epsilon''(0) < 0$ and decreases as ϵ decreases. So letting $\epsilon \to 0$ in (38) gives

$$J(y) < J(y_c). \tag{39}$$

The fourth term in (38) must have a limit since the first three do. Now, restricting our attention to all y in the Euler class which have arc length equal to $a\pi$ (the arc length of the semi-circle) (39) tells us that y_c uniquely solves the isoperimetric problem in the Euler class. ∎

Remark. It is important to realize that our sufficiency proof borrowed only the objective function J of the auxiliary problem (7) from Euler's necessity proof. Hence, it doesn't matter whether Euler's proof of his rule was rigorous or not.

Remark. It is also important to realize that since $\epsilon > 0$, $\phi_\epsilon(t)$ in (35) exists as a Riemann integral and we do not have to appeal to part (i) of Theorem 2.4 for existence.

Could Euler have made our observation at the time of his 1744 writing? Let's pursue this question in some detail. The foundation of our observation is Taylor's theorem with remainder. The literature tells us that Taylor published his theorem in 1715, and that it had been discovered by Gregory, but not published, some 40 years earlier; see [31]. According to Fraser [12], Euler was not only aware, but was influenced by Taylor's early work in the calculus of variations. So Euler most likely was aware of Taylor's theorem in 1744. However, the rub is that Taylor's theorem with remainder was not known at that time. It is somewhat ironic that the form of the remainder that we use in our proof is credited to Lagrange [23] in 1797, and is actually referred to today as the Lagrange form of the remainder; see [31]. So Euler would not have been in good position to make our observation.

3.2 A Proof Motivated by Lagrange.

Theorem 3.2. *The semi-circle curve $y_c(x)$ given by (6) uniquely solves the iso-area problem*

$$\begin{aligned} \textit{minimize} \quad & J(y) = \int_{-a}^{a} \sqrt{1 + y'(x)^2}\, dx \\ \textit{subject to} \quad & \int_{-a}^{a} y(x)\, dx = \frac{\pi}{2} a^2, \\ & y(-a) = y(a) = 0, \end{aligned} \tag{40}$$

with the arc length integral interpreted as an improper Riemann integral, over $E(-a, a)$; the Euler class of functions given by Definition 2.3. Hence it uniquely solves the isoperimetric problem, with the arc length integral interpreted as an improper Riemann integral, over the Euler class.

Proof. Following Lagrange's 1759 [22] derivation of the Euler-Lagrange equation we first consider a class of admissible variations. In dealing with problem (40) our class is

$$S = \left\{ \eta \in E(-a, a) \colon \int_{-a}^{a} \eta(x)\, dx = 0 \right\}. \tag{41}$$

Since members of $E(-a, a)$ are continuous the area integral in problem (40) and in (41) are viewed as proper Riemann integrals. As before let y_c denote the semi-circle (34) and consider any $y \neq y_c$ contained in the Euler class. Let $\eta = y - y_c$ and notice that $\eta \in S$. As in the previous proof define

$$\phi_\epsilon(t) = \int_{-a+\epsilon}^{a-\epsilon} \sqrt{1 + (y_c' + t\eta)^2}\, dx \tag{42}$$

for $t \in [0, 1]$ an $\epsilon \in (0, a)$.

Straightforward differentiations with respect to t give

$$\phi_\epsilon'(t) = \int_{-a+\epsilon}^{a-\epsilon} \frac{(y_c' + t\eta')\eta'}{\sqrt{1 + (y_c' + t\eta')^2}}\, dx \tag{43}$$

and

$$\phi_\epsilon''(t) = \int_{-a+\epsilon}^{a-\epsilon} \frac{(\eta')^2}{(1 + (y_c' + t\eta')^2)^{3/2}}\, dx. \tag{44}$$

Recalling (9) and an integration by parts give

$$\phi_\epsilon'(0) = -\frac{1}{a} \int_{-a+\epsilon}^{a-\epsilon} x\eta'\, dx = \frac{1}{a} \left[-x\eta(x) \Big|_{-a+\epsilon}^{a-\epsilon} + \int_{-a+\epsilon}^{a-\epsilon} \eta\, dx \right]. \tag{45}$$

Hence $\phi_\epsilon'(0) \to 0$ as $\epsilon \to 0$.

Taylor's theorem tells us that

$$\phi_\epsilon(1) = \phi_\epsilon(0) + \phi_\epsilon^{'}(0) + \frac{1}{2}\,\phi_\epsilon^{''}(\theta) \text{ for some } \theta \in (0,1). \tag{46}$$

Observe that $\phi_\epsilon^{''}(\theta) > 0$ and increases as ϵ decreases. So letting $\epsilon \to 0$ in (46) gives

$$J(y_c) < J(y).$$

Since y was an arbitrary member of the Euler class y_c uniquely solves the iso-area problem, hence the isoperimetric problem in the Euler class. ■

Remark. Lagrange could have made this proof because he was familiar with the form of Taylor's theorem that we used, indeed it is due to him. While this hypothesized proof would have been made 50 years after Euler, it would still have been some 80 years before Weierstrass.

4 Concluding Remarks.

With our concluding remarks we promote three general thoughts. The first is that the isoperimetric problem has been a most impactful mathematical problem. The isoperimetric problem, perhaps because it is so easy to state and understand and yet its solution has been so mathematically challenging has influenced the writings of scholars in many diverse areas. Zenodorus' and Steiner's efforts on this problem served to promote the tools of geometry. Euler built multiplier theory specifically to solve this problem. Weierstrass built sufficiency theory in the calculus of variations specifically to solve this problem. Hundreds of papers have been written concerning the solution of the isoperimetric problem or related issues. Porter in his masters thesis [27] gives 75 references to papers related to the isoperimetric problem and this is just up to the year 1931. During the golden era, the formative years of mathematical analysis, say 1630 to 1890, the research path of essentially every mathematician of note intersected some portion of the world of the isoperimetric problem . The isoperimetric problem promoted the development of the calculus of variations which in turn led to the development of mathematical analysis. Hence, the isoperimetric problem is arguably the most impactful mathematical problem of all time.

Our second premise is that we believe our proofs given in Section 3 are short, elementary, and teachable. Our proofs follow from a straightforward application of Taylor's theorem and to our liking they retain a flavor of optimization theory. Our comparison benchmark is the elegant short proof presented by Peter Lax in Section 2.6. He claims quite appropriately that his proof is eminently suitable for presentation in an honors calculus course. The honors designation probably comes from the fact that he uses a parametric representation of the curve in question. Our proofs use standard Cartesian coordinate representation and could be included in standard differential calculus classes in college and perhaps in high school when Taylor's theorem is taught.

Our third premise concerns functional convexity. Euler failed to establish sufficiency for the isoperimetric problem, Lagrange seemed to have not tried to establish sufficiency for this problem, and as we have argued Weierstrass developed a sufficiency theory that was excessively complicated for merely solving the isoperimetric problem; yet solution of the isoperimetric problem was his primary objective. While we have enormous respect for these three giants and the other mathematical pioneers of the golden era, we maintain that they failed to solve the isoperimetric problem in an effective manner because of their failure to pursue the general notion of functional convexity and the powerful optimization sufficiency theory that follows directly from this notion.

Qiaquinta and Hildebrandt [14, pp.248-249] remind us that Euler [10], in addition to the multiplier approach that we presented in Section 2.3, found another way of treating the isoperimetric problem. He worked with the iso-area problem and in a most ingenious manner made a coordinated transformation writing the curve under consideration in parametric form where the independent variable was arc length. When he wrote the transformed problem the area constraint vanished. Hence he arrived at a problem which had no subsidiary constraint and had the form of the simplest problem in the calculus of variations. He then showed that the circle was an extremal of this problem by solving the Euler-Lagrange equation associated with the transformed problem. Now as mentioned above Euler believed that extremals were solutions, so he thought that he had solved the isoperimetric problem in an alternative manner. Giaquinta and Hildebrandt point out that Euler's transformed problem is a convex program.

In conclusion, on one hand we are satisfied with our demonstrations that an optimization theory approach to solving the isoperimetric problem is not as difficult as once believed; while on the other hand we are happy that the lack of solution of this problem led to such creative times in mathematics.

Acknowledgments.

My interest in the topic of the historical development of the isoperimetric problem has been honed over the years as I searched for elementary teachable proofs for my Rice optimization theory class. As I constructed the two proofs that are the subject of the current paper I shared them with students and colleagues alike. When I received an invitation to give a plenary talk at the 2012 SIAM Annual Meeting I was undecided about speaking on the topic of the isoperimetric problem revisited. I thank Mark Embree for encouraging me to present this topic. The current paper was written as a formalization of the SIAM address and has been a rewarding experience.

During a presentaton at Rice Steve Cox highlighted the need for understanding along the lines of Theorem 2.4. I thank him for this.

Jim Case did a masterful job of summarizing my plenary talk presentation for SIAM News. During his writing process we had several discussions that contributed positively to the content of Theorem 2.4. I thank him for this and acknowledge his writing and mathematical talent.

Roland Glowinski and I had many interesting and enjoyable converstations on the material in the paper. He made me think more each time that we talked. I thank him and acknowledge his broad understanding and knowledge of mathematics and mathematics history.

In preparation for Theorem 2.4 I visited my Rice colleague Frank Jones with some specific questions. Frank not only answered my questions, he directed me to Chapter 16 of his book [20] which contains a wealth of material that I had not been able to previously find. As the reader can see from the proof of Theorem 2.4 Frank's book is exactly what I needed and is an impressive collection of material. I thank Frank for this valuable help.

When I first learned about Peter Lax's wonderful short proof of the isoperimetric inequality I reflected back on how much Peter has given to the mathematics community over his life time; so I decided to acknowledge his contribution by dedicating this paper to him, my way of saying thank you Peter.

Finally, I thank the SIAM Annual Conference plenary talk selection committee for allowing me to give the talk that this paper represents, and in particular (at that time) SIAM President Nick Trefethan for his comments on the material and for his support.

References

[1] ANONYMOUS, Annals de Mathématiques Pures et Appliquées, 13 (1822-1823), pp. 132-139, 14 (1823-1824), pp. 316-317.

[2] ARISTOTLE, D.C., *The Work of Aristotle*, translated from the Greek by J. A. Smith and W. D. Rose, Clarendon Press, Oxford, 1908, p. 287.

[3] BECKENBACH, E. F., An address delivered before the annual meeting of the American Mathematical Society in Athens, Georgia, on December 30, 1947.

[4] BLASJO, V., *The Isoperimetric Problem*, The American Mathematical Monthly, 112.6 (2005), pp. 526-566.

[5] BOLZA, O., *Lectures on the Calculus of Variations*, The University of Chicago Press, Chicago, IL, 1904.

[6] CARATHÉODORY, C. and STUDY, E., *Zwei Beewise des Satzes, das der Kreis unter allen Figuren gleichen Umfanges den grossten Inhalt hat*, Math. Annalen, 68 (1909), pp. 133-140.

[7] du BOIS-REYMOND, P., *Erlauterungen Zu den Anfangsgrunden der Variationsrechnung*, Math. Annalen, 15 (1879), pp. 283-314.

[8] EDLER, F., *Vervollständigung der Steiner'schen elementargeometrischen Beweise für den Satz, das der Kreis grösseren Flächeninhalt besizt, als jede andere ebene Figur gleich grossen Umfanges*, Nachrichten von der Königl. Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen, (1882), pp. 73-80.

[9] ELSGOLC, L.E., *Calculus of Variations*, Pergamon Press Ltd., London, 1962.

[10] EULER, L., *Methodus inveniendi linas curvas maximi minimive proprietate gaudentes sive Solutio Problematis Isoperimetrici Latissimo Sensu Accepti*, Lausanne and Geneva, 1744. E65A. Opera Omnia.

[11] FENCHEL, W., Translation of a talk given at the celebration of the centenary of the Danish Mathematical Society in 1973. The original version appeared in Dansk Matematisk Forening, 1923-1973, Copenhagen (1973), pp. 103-116.

[12] FRASER, C. G., *The Origins of Euler's Variational Calculus*, Archive for History of Exact Science, 47 (1994), pp. 103-141.

[13] GELFAND, I. M. and FOMIN, S. V., *Calculus of Variations*, translated from the Russian by R. A. Silverman, Prentice-Hall Inc., Englewood Cliffs, NJ, 1963.

[14] GIAQUINTA, M. and HILDEBRANDT, S., *Calculus of Variations I*, Springer-Verlag, Berlin, 2004.

[15] GOLDSTINE, H. H., *A History of the Calculus of Variations from the 17th Through the 19th Century*, Springer-Verlag, New York, NY, 1980.

[16] HEATH, T. L., *A History of Greek Mathematics: From Thales to Euclid*, 2 (1921), Clarendon Press, Oxford.

[17] HURWITZ, A., *Sur quelques applications géométriques des séries de Fourier*, Ann. Sci. Ec. Norm., Paris, 19 (1902), pp. 357-408.

[18] JENSEN, J. L. W. V., *Om Konvexe Funktioner og Uligheder mellem Middelvaerdier*, Nyt. Tidsskr. Math., 16B (1905), pp. 49-69.

[19] JENSEN, J. L. W. V., *Sur les Fonctions Convexes et les Inegalites entre les Valeurs Moyennes*, Acta Math., 30 (1905), pp. 175-193.

[20] JONES, F., *Lebesgue Integration on Eucliden Space*, Jones and Bartlett, Boston, 1993.

[21] LAGRANGE, J. L., *Lettre de Lagrange á Euler*, August 12, 1755. Oeuvre 14 (1892), pp. 138-144. pp. 144-146 contain Euler's response to Lagrange.

[22] LAGRANGE, J. L., *Recherches sur la Méthode de Maximis et Minimis*, Misc. Soc. Tur., I (1759), pp. 3-20.

[23] LAGRANGE, J. L., *Théorie des Fonctions Analytiques*, L'Imprimerie de la République, Prairial an V, Paris, 1797. Nouvelle édition: Paris, Courcier 1813.

[24] LAX, P. D., *A Short Path to the Shortest Path*, The American Mathematical Monthly, 102 (1995), pp. 158-159.

[25] LEBESGUE, J. L., *Lecons sur les Series Trigonometriques Professées au Collége de France*, Gauthier-Villars, Paris, 1906, pp. 107-110.

[26] MAYER, A., *Die Kriterien des Maximums und Minimums der einfachen Integrale in den isoperimetrischen Problemen*, (Aus den Ber. der Kgl. Sächs. Ges. d. Wiss. Juli 1877), Math. Annalen., 13 (1878), pp. 53-68.

[27] PORTER, T. I., *A History of the Classical Isoperimetric Problem*, Masters Dissertation, August 1931, Mathematics Department, University of Chicago.

[28] PROCLUS, L., *The Commentaries of Proclus on the Timaeus of Plato*, English translation by T. Taylor, Book 1, 1820, p. 450.

[29] STEINER, J., *Einfache Beweise der Isoeperimetrische Hauptsätze*, J. Reine Agnew. Math., 18 (1838), pp. 281-296.

[30] WEIERSTRASS, K., *Mathematische Werke*, Mayer and Muller, Berlin 7 (1927).

[31] WEISSTEIN, E. W., *Taylor's Theorem.* From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/TaylorsTheorem.html`.