The Report Committee for Sarah Elise White
Certifies that this is the approved version of the following report:

Optimization Methods

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Edward Odell

Mark Daniels

Optimization Methods

by

Sarah Elise White, BSMath

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

August 2012

Abstract

Optimization Methods

Sarah Elise White, MA

The University of Texas at Austin, 2012

Supervisor: Edward Odell

This report articulates some of the recent research on different methods of optimization. Topics discussed include an implicit differentiation process in which the primary substitution method is not used and a relationship among variables method is introduced. In addition, a finding extrema without limits method is explored. Also included is a discussion on the depth of optimization taught in secondary schools and the different methods and levels of instruction on this topic.

Table of Contents

List of Figures

Chapter 1: An Introduction

As students progress through school, they encounter increasingly higher expectations from their teachers. Students answer problem solving questions with help and guidance from their teachers at early ages, but the help and guidance typically diminishes as the students continue through the school year and the next grade. Once teachers remove the aid of prior scaffolding methods, students are expected to use their problem solving skills themselves. When problem solving individually, students find that there are multiple strategies, or methods, with which to approach each question.

A major topic in high school Calculus classes is that of *optimization*. After learning about derivatives, students study problems in which something is to be maximized or minimized given different parameters. Optimization is the process by which functions are maximized or minimized within a given domain to yield a specific value. In calculus, to optimize, the derivative of the function is calculated, followed by a test of the extrema to determine where on the domain a function might have a maximum, minimum, or both. Finally, this critical domain value is substituted back into the function to yield the solution.

These types of problems are interesting and can be used in both science and mathematics courses, but not all students are fortunate enough to take Calculus. Though traditional calculus textbook methods work, there are other methods which students could use; both with and without the aid of calculus. Young uses an original Calculus approach which employs implicit differentiation in place of the substitution step previously mentioned **[1]**. Additionally, Suzuki introduces an approach to finding extrema using general algorithms based on the theory of equations and the geometric properties of curves **[2]**.

Chapter 2: Constrained Optimization with Implicit Differentiation

Young presents a creative Calculus, approach to optimization problems that employs implicit differentiation. By using implicit differentiation, students discover relationships among the variables, those of which would not have been seen using the standard method. Young walks the reader through three common problems demonstrating the implicit differentiation method in each. Discussed here is one of the problems solved using the implicit differentiation method.

One problem Young presents is: Using a fixed length of fence, find the maximum area enclosed in a rectangular field that is to be subdivided into rectangular plots bounded by fencing. The sides of the field may also be a river bank, cliff walls, or etc.

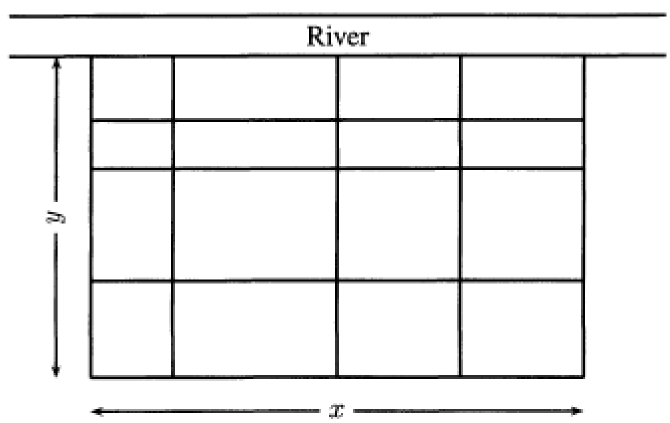

Figure 1. River [**1**, p. 2]

The problem is to maximize $A = xy$ subject to $L = mx + ny$, where n is the number of vertical sections of fence, m is the number horizontal sections of fence, and L is the fixed length of fence available. Young begins by using implicit differentiation to compute dA/dx which is set equal to 0,

$$0 = \frac{dA}{dx} = y + x\frac{dy}{dx}.$$

Since, $L = mx + ny$

$$\frac{dy}{dx} = \frac{-m}{n}$$

and substituting dy/dx into the first equation yields

$$ny = mx.$$

Next, utilizing the previous equation and $L = mx + ny$, it follows that

$$y = \frac{L}{2n}$$

$$x = \frac{L}{2m}$$

producing a maximum area of

$$A = \frac{L}{2m} \times \frac{L}{2n} = \frac{L^2}{4mn}.$$

Finally, by computing the second derivative using implicit differentiation, one sets

$$\frac{d^2A}{dx^2} = \frac{-2m}{n}.$$

Young illustrates a visual of the result by graphing $L = mx + ny$ and $ny = mx$, where the coordinates of the intersection point provide the dimensions of the field.

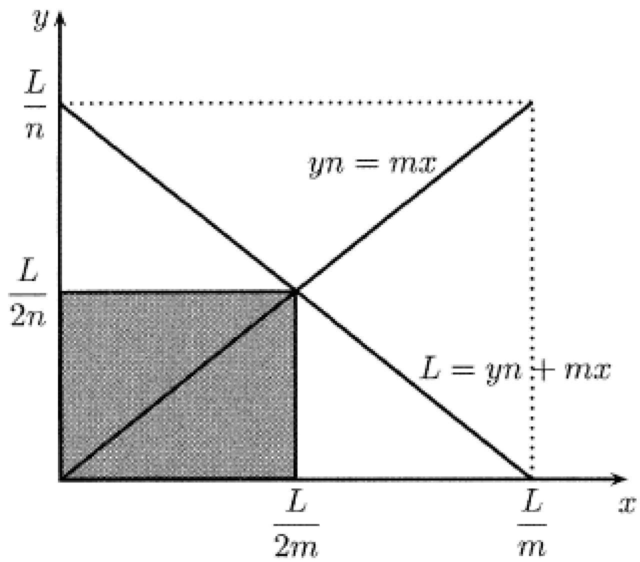

Figure 2. Illustration of Result [**1**, p. 3]

Young's method provides first year Calculus students with a way to optimize without having learned about the concavity of a function or the meaning of roots of a function when it comes to optimization. Also, this approach sidesteps the algebraic mistakes usually made by students with the standard Calculus approach.

Chapter 3: The Lost Calculus: Tangency and Optimization without Limits

An approach to finding extrema without limits was discovered by Decartes and Jan Hudde, a Dutch mathematician. This method is known as the “lost calculus” because the work of Newton and Leibniz in the 1670s relegated these techniques to the role of misunderstood historical curiosities **[2]**. The “lost calculus” method can be tracked back to Descartes’ method of finding tangents to algebraic curves, published in *La Geometrie* in 1637. Hudde further completed the method with his resulting algorithms in the years 1657-1658 **[2]**.

Descartes’ method is as follows: Suppose one wished to find a circle that is tangent to the curve OC at some point C (as shown in Figure 2). Consider a circle with center P on some convenient reference axis and suppose this circle passes through C. This circle may pass through another nearby point E on the curve; in this case, the circle is not tangent to the curve. On the other hand, if C is the only point of contact between the circle and curve, then the circle will be tangent to the curve. Thus, the goal is to find P so that the circle with center P and radius CP will meet the curve OC only at the point C.

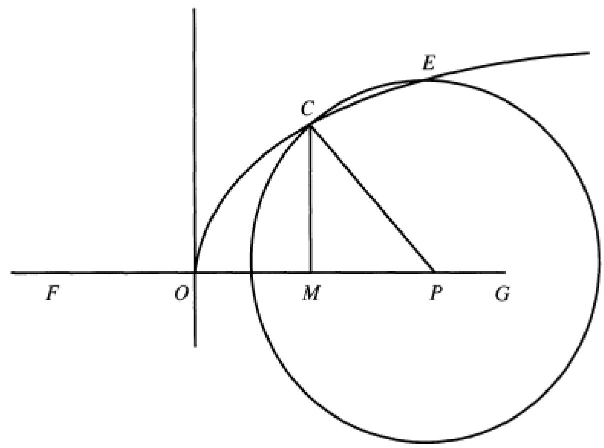

Figure 3. Descartes's Method of Tangents [**2**, p. 2]

Any points that the circle and curve have in common correspond to the solution of the system of equations representing the curve and circle. For example, if there are two distinct intersections, then there are two solutions to the system. Thus, in order for the circle and curve to be tangent and have only one point in common, the representative system of equations must have one solution, a root with multiplicity of two corresponding to the common point C.

Although Descartes did not supply his readers with examples, Suzuki does in his article, *The Lost Calculus: Tangency and Optimization without Limits* **[2]**. Suppose *OC* is the curve $y = \sqrt{x}$, and let *C* be the point (a^2, a). Imagine a circle passing through the point (a^2, a), with radius *r* centered on the *x*-axis at the point $(h, 0)$, with *h* and *r* to be determined. Then, the circle has equation

$$(x-h)^2 + y^2 = r^2.$$

Expanding and setting the equation equal to zero one obtains

$$y^2 + x^2 - 2hx + h^2 - r^2 = 0.$$

The solution, or intersection point(s) of the circle and curve, can be found by solving the following system of equations

$$y^2 + x^2 - 2hx + h^2 - r^2 = 0$$

$$y = \sqrt{x}.$$

The substitution method with $y = \sqrt{x}$, yields

$$x^2 + (1 - 2h)x + (h^2 - r^2) = 0,$$

which is a quadratic function with two solutions. Using the previous assumption that the circle and curve have one point in common, (a^2, a), we have $x = a^2$ as a root of this equation, but not necessarily the only one. To ensure that this will be the only root, it is essential that

$$x^2 + (1 - 2h)x + (h^2 - r^2) = (x - a^2)^2.$$

Expanding the right-hand side and comparing coefficients, one finds that

$$1 - 2h = -2a^2.$$

It follows that

$$h = a^2 + \frac{1}{2}.$$

Therefore, the circle with center $(a^2 + \frac{1}{2}, 0)$ will be tangent to the graph of $y = \sqrt{x}$ at the point (a^2, a) **[2]**.

Descartes approached the problem of tangents by locating the center of the tangent circle, whereas, today, finding the slope of the tangent line is used. However, there is a relationship between these two ideas. From Euclidean Geometry, it is known that the radius through a point C is perpendicular to the tangent line of the circle though C. In this particular example, the radius PC will lie on a line with slope $-2a$. Thus, the tangent line through C will have slope $\frac{1}{2a}$.

In *La Geometrie*, Descartes states that this method works well with all quadratic forms. Suppose one were to find the tangent to the curve $y = x^3$ using Descartes' method. Start with the center of the circle at $(h, 0)$. In order for this method to work, one wants the system of equations

$$x^2 2hx + h^2 + y^2 - r^2 = 0$$

$$y = x^3.$$

to have a double root at the point of tangency (a, a^3). Substituting x^3 for y yields

$$x^6 + x^2 - 2hx + h^2 - r^2 = 0.$$

To find the tangent at the point (a, a^3), this equation should have a double root at $x = a$. Since the left hand side is a sixth degree polynomial, it must factor as the product of $(x - a)^2$ and a fourth degree polynomial such as:

$$x^6 + x^2 - 2hx + h^2 - r^2 = (x - a)^2(x^4 + Bx^3 + Cx^2 + Dx + F).$$

By expanding the right hand side of the equation

$$x^6 + x^2 - 2hx + h^2 - r^2$$

$$= x^6 + (B - 2a)x^5 + (a^2 - 2aB + C)x^4 + (a^2B - 2aC + D)x^3$$

$$+ (a^2C - 2aD + F)x^2 + (a^2D - 2aF)x + a^2F.$$

Then, comparing the coefficients yields the following system:

$$B - 2a = 0$$

$$a^2 - 2aB + C = 0$$

$$a^2B - 2aC + D = 0$$

$$a^2C - 2aD + F = 1$$

$$a^2D - 2aF = -2h$$

$$a^2F = h^2 - r^2.$$

After multiple substitutions it is found that

$$a^2(4a^3) - 2a(1 + 5a^4) = -2h.$$

Thus, $h = a + 3a^4$ and the center of the tangent circle will be at $(a + 3a^5, 0)$. As before, the perpendicular to t curve will have a slope of

$$-\frac{a^3}{3a^5} = -\frac{1}{3a^2}.$$

Thus, the slope of the line tangent to the curve $y = x^3$ at $x = a$ will be $3a^2$.

As previously stated, the Descartes approach relies heavily on the understanding of geometric properties of circles and curves; whereas today, using basic principles of Calculus, the same problem would be solved using the derivative to find the slope of the tangent line. Although the geometric circle method appears simple, it can easily become cumbersome. As the tangent curve increases in degree, the polynomial being solved has a larger degree and the system of equations grows, but the results yielded will be the same as found using calculus.

Descartes later simplified his method by replacing the circle with a line and using the idea of slope from the ratio between the sides of similar triangles **[2]**. In our time, Descartes' simplified method is as follows:

The equation of a line that touches the curve $f(x, y) = 0$ at (a, b) is

$$y = m(x - a) + b,$$

where *m* denotes a parameter to be determined by differentiation. In order for the line to be tangent to $f(x, y)$, it is necessary for the system of equations

$$f(x, y) = 0$$

$$y = m(x - a) + b$$

to have a double root at $x = a$.

Hudde further simplified this method by discovering an algorithm that distinguishes roots of multiplicity two for polynomials. The algorithm Hudde uses is known as "factoring", which allows the ability to see all roots of the polynomial. From the factored form of the polynomial, it can be seen if the roots have a multiplicity of 1 or more. Since Descartes' method of tangents relies on finding multiple roots, Hudde's discovery and clarification of multiplicity became of utmost importance in elevating the original Descartes method to a more reliable form of optimization. Hudde's results are: Given any polynomial

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0,$$

1. If $f(x)$ has a root of multiplicity 2 or more at $x = a$, then the polynomial $f'(a) = 0$, and

2. If $f(x)$ has an extreme value at $x = a$, then $f'(a) = 0$. **[2,** p. 1]

Although factoring is an essential topic in today's Algebra classrooms, Hudde approaches the topic in a slightly different manner. The most popular extension to finding roots of multiplicity greater than one is to find the greatest common divisor, GCD, of each polynomial and determine what factors they have in common. However, this does not work for all polynomials; thus, the use of the Euclidean algorithm is necessary. For example, let $f(x) = x^3 - 4x^2 + 10x - 7$ and $g(x) = x^2 - 2x + 1$. To apply the Euclidean algorithm divide $f(x)$ by $g(x)$ to obtain a quotient and a remainder, resulting in the following

$$x^3 - 4x^2 + 10x - 7 = (x^2 - 2x + 1)(x - 2) + (5x - 5).$$

Now, divide the previous divisor, $x^2 - 2x + 1$ by the remainder, $5x - 5$, to obtain a new quotient and remainder:

$$x^2 - 2x + 1 = (5x - 5)\left(\frac{1}{5}x - \frac{1}{5}\right) + 0.$$

The last nonzero remainder, $5x - 5$, is the GCD of the polynomials.

Hudde, however, presented an alternative to the Euclidean algorithm. Instead of dividing the polynomials, Hudde found the remainder modulo the divisor. Hudde started by setting each polynomial equal to zero and then solved for the highest power term. Using the previous example, along with Hudde's method

$$x^2 = 2x - 1$$

$$x^3 = 4x^2 - 10x + 7.$$

Having an expression for x^2 allows for itself and the higher degree terms of the other factor to be eliminated. Using substitution in this system, the following sequence of steps are obtained:

$$x^3 = 4x^2 - 10x + 7$$

$$x(x^2) = 4(2x - 1) - 10x + 7$$

$$x(2x - 1) = 4(2x - 1) - 10x + 7$$

$$2x^2 - x = 8x - 4 - 10x + 7$$

$$2(2x - 1) - x = -2x + 3$$

$$3x - 2 = -2x + 3$$

$$5x = 5$$

$$x = 1.$$

The second equation, $x^3 = 4x^2 - 10x + 7$, has been reduced to $x = 1$. This is comparable to performing

$$x^3 - 4x^2 - 10x + 7 = x - 1 \; mod \; x^2 - 2x + 1.$$

Then, $x = 1$ can be used to eliminate all the first or higher degree terms of the other equation:

$$x^2 = 2x - 1$$

$$1^2 = 2(1) - 1$$

$$1 = 1.$$

This being true, the GCD is the factor corresponding to the last substitution; here $x = 1$ corresponds to the factor $(x - 1)$ [**2**, p. 7]. The value of finding the GCD is made apparent in Hudde's *tenth rule*. The rule states that if the equation has two equal roots, one can multiply by whatever arithmetic progression and set the product equal to zero.

Then, with the two equations that have been found, find their GCD and divide the original equation by that quantity as many times as possible. However, being the GCD does not mean that the associated root has multiplicity more than one. Thus, Hudde's Theorem is as follows:

> **Hudde's Theorem**. *Let* $f(x) = \sum_{k=0}^{n} a_k x^k$, *and let* $\{b_k\}_{k=0}^{n}$ *be any arithmetic progression. If* $x = r$ *is a root of* $f(x)$ *with multiplicity two or greater, then* $x = r$ *will be a root of* $g(x) = \sum_{k=0}^{n} b_k a_k x^k$.

A polynomial $g(x)$ created in this way from $f(x)$ is called a *Hudde polynomial* [**2**, p. 7]. For example, using the arithmetic sequence 3, 2, 1, 0, and tabular array, the roots of $x^3 - 4x^2 + 5x - 2$ can be found.

$$x^3 - 4x^2 + 5x - 2$$

$$\underline{3 \qquad 2 \qquad 1 \qquad 0}$$

$$3x^3 - 8x^2 + 5x$$

Since the GCD of $x^3 - 4x^2 + 5x - 2$ and $3x^3 - 8x^2 + 5x$ is $(x - 1)$, if the original polynomial has a repeated factor, then the repeated factor can only be $(x - 1)$. After factoring out $(x - 1)$, it is found that

$$x^3 - 4x^2 + 5x - 2 = (x - 1)^2(x - 2).$$

Thus, the roots of the polynomial are 1, 1, and 2. Suzuki further explains, and illustrates using an example, that any arithmetic sequence will work using the same polynomial but a different sequence.

This method can be used to solve the problem that Descartes introduced, as Suzuki shows. Recall the earlier case of finding the tangent to $y = x^3$ to a circle with center $(h, 0)$ that passed through the point (a, a^3). The corresponding system equations

$$y = x^3$$

$$(x - h)^2 + y^2 = r^2$$

could be reduced by substituting x^3 for *y* in the second equation, resulting in

$$x^6 = x^2 - 2hx + h^2 - r^2 = 0.$$

In order for the circle to be tangent to the curve at (a, a^3), the equation must have a double root at $x = a$, thus, the corresponding Hudde polynomial will have a root at $x = a$. Thus, a Hudde polynomial can be constructed through multiplication by an arithmetic sequence ending in zero:

$$x^6 + 0x^5 + 0x^4 + 0x^3 + x^2 - 2hx + h^2 - r^2$$

$$\underline{6 \quad\; 5 \quad\; 4 \quad\; 3 \qquad\quad 2 \qquad\quad 1 \qquad\qquad 0}$$

$$6x^6 + 2x^2 - 2hx.$$

Since the Hudde polynomial has a root at $x = a$, then *h* must satisfy

$$6a^6 + 2a^2 - 2ha = 0.$$

Thus, $h = a + 3a^5$, and the center of the tangent circle will be located at $(a + 3a^5, 0)$.

This method can also be applied to Descartes' simplified method. By finding the expression for the slope of the tangent line to a curve, Hudde's method would yield an expression equivalent to the derivative, as found in Calculus. For the problem of $y = x^3$, the system to be solved is

$$y = x^3$$

$$y = m(x - a) + a^3$$

having a double root at $x = a$. Substituting $y = x^3$ into the second equation and setting it equal to zero yields

$$x^3 - mx + am - a^3 = 0.$$

From the assumption, $x = a$ is a double root, the corresponding Hudde polynomial will have $x = a$ as a root. Multiplying the equation by an arithmetic sequence yields

$$x^3 + 0x^2 - mx + (am - a^3)$$

$$\underline{3 \qquad 2 \qquad 1 \qquad\qquad\qquad 0}$$

$$3x^3 - mx.$$

Since $x = a$ is a root of the Hudde polynomial, then m must satisfy

$$3a^3 - ma = 0.$$

Thus, $m = 3a^2$, which is the derivative of $y = x^3$.

Hudde claimed that this same method can be used to find the extrema **[2]**. Suzuki published the proof given by Hudde, but in a simplified form. Suppose a polynomial $P(x)$ is the product of the third-degree polynomial $x^3 + px^2 + qx + r$ and a second-degree polynomial with $x^2 - 2xy + y^2$ as a root with multiplicity two (where $x = y$ is the root of multiplicity two). Hence the roots of $P(x)$ satisfy

$$P(x) = (x^2 - 2xy + y^2)x^3$$
$$+ (x^2 - 2xy + y^2)px^2$$
$$+ (x^2 - 2xy + y^2)qx$$
$$+ (x^2 - 2xy + y^2)r = 0,$$

where, for convenience, the polynomial $x^2 - 2xy + y^2$ is designated as the coefficients of the terms of the cubic polynomial. Notice that the coefficients of the terms are the three terms of the second-degree polynomial. Thus, when the Hudde polynomial is formed, the coefficients will be multiplied by successive terms in the arithmetic sequence a, $a + b$, $a + 2b$, to become

$$ax^2 - (a + b)2yx + (a + 2b)y^2.$$

If $x = y$, this coefficient will be

$$ay^2 - (a + b)2y^2 + (a + 2b)y^2,$$

which is equal to zero. Thus, $x = y$ will be a root of the Hudde polynomial. This proves Hudde's Theorem [**2**, p. 10].

Geometrically, Hudde's method for finding the extreme value of a polynomial function is clear. Suppose $f(x)$ has an extremum at $x = a$, with $f(a) = Z$, where Z is the extreme value. Then $f(x) - Z$ will have a root with multiplicity of two at $x = a$, and the corresponding Hudde Polynomial will have a root of $x = a$. For example, Suzuki applies this concept to $x^3 - 10x^2 - 7x + 346$. Suppose $x^3 - 10x^2 - 7x + 346 = Z$ has an extreme value, which occurs at $x = a$; then $x^3 - 10x^2 - 7x + 346 - Z$ will have a root with multiplicity two at $x = a$. By multiplying by an arithmetic sequence ending in zero, Z can be eliminated:

$$x^3 - 10x^2 - 7x + 346 - Z$$

$$\underline{3 \qquad 2 \qquad\quad 1 \qquad 0}$$

$$3x^3 - 20x^2 - 7x.$$

The result of this operation is simply $xf'(x)$. The next step is to set the resulting polynomial equal to zero. By the assumption, $x = a$ is a root of multiplicity two of the original polynomial, then by Hudde's Theorem, $x = a$ is a solution to $3x^3 - 20x^2 - 7x$. Thus, the solutions are

$$x = 0, x = -\frac{1}{3}, \text{and } x = 7.$$

By this assumption, at least one of the roots has a multiplicity of two for the original polynomial. It must be verified which of the solutions correspond to a maximum or minimum, and if there are any extraneous solutions. This can be checked by substituting the values into the original polynomial, yielding $x = -\frac{1}{3}$ as the local maximum, $x = 7$ as the local minimum, and $x = 0$ as an extraneous solution.

Descartes and Hudde's approaches to finding the extrema of polynomials present alternatives to optimization problems that do not require calculus. Their methods rely on basic algebraic and geometric skills and allow students at lower levels in their mathematics careers to tackle optimization problems with the tools that they possess before entering into Calculus.

Chapter 4: Conclusion

The study of optimization continues to provide opportunities for mathematicians to introduce new approaches to solving such problems. Young lays out the steps to a popular optimization problem using implicit differentiation. Finally, Suzuki explains a geometric approach to optimization, first outlined by Decartes, and algebraically confirmed and clarified by Hudde.

These two mutually exclusive approaches show that optimization is a topic which can be attacked in many ways, but also highlight a more important conclusion: determining multiple methods to solve problems is a valuable by-product of limiting an approach to non-calculus techniques in solving optimization problems. According to the National Council of Teachers in Mathematics, students in grades nine through twelve need to apply and adapt a variety of appropriate strategies to solve problems **[3]**. This statement supports the common educational philosophy that students must learn how to apply their individual prerequisite skills, to creatively approach problem solving opportunities as their mathematical maturity evolves. Optimization problems create opportunities for students to break free of their earlier reliance on teacher scaffolding techniques, and to apply their knowledge of Algebra, Geometry, and Precalculus to develop problem solving skills and ameliorate their mathematical foundations.

Perhaps the most exciting aspect of optimization problems is that they can be solved using a multitude of approaches. If students ultimately gain the calculus skills required to solve optimization problems using traditional calculus methods, their earlier attempts and successes will augment their final and lasting understanding of the often confusing, yet limitless topic of optimization.

References

1. Gary W. De Young. Constrained Optimization with Implicit Differentiation. *The College Mathematics Journal* 34, No. 2 (March 2003) 148-152.

2. Jeff Suzuki. The Lost Calculus (1637-1670): Tangency and Optimization without Limits. *Mathematics Magazine* 78, No. 5 (2005) 339-353.

3. Principles and Standards for School Mathematics. 2011. Retrieved October 28, 2011. http://standards.nctm.org/.

Vita

Sarah Hunter was born and raised in Dallas, TX. She attended and graduated from Bryan Adams High School in 2006. She then continued her education by earning her BS in Mathematics at The University of Texas at Austin in May of 2010. She currently teaches Algebra 1, Algebra 2, and Functions, Statistics and Trigonometry and St. Dominic Savio Catholic High School in Austin, TX. In the summer of 2010, she entered Graduate School at The University of Texas at Austin.

Email Address: white.sarah.elise@gmail.com

This report was typed by the author.