Tangents and Subtangents Used to Calculate Areas

Tom M. Apostol and Mamikon A. Mnatsakanian

1. INTRODUCTION: SUBTANGENTS. The subtangent associated with any differentiable function f is the function s defined by the equation

$$s(x) = \frac{f(x)}{f'(x)} \tag{1}$$

at each point x where $f'(x)$ is nonzero. When $f(x) > 0$, $s(x)$ represents the base of a right triangle of altitude $f(x)$ and hypotenuse of slope $f'(x)$. Equation (1) can be regarded as a linear first-order differential equation $f'(x) = f(x)/s(x)$ whose general solution (in terms of indefinite integrals) is

$$f(x) = K \exp\left(\int dx/s(x)\right).$$

Thus, the subtangent determines $f(x)$ up to a constant factor. Therefore it is not surprising that information about $f(x)$ can be extracted from a knowledge of $s(x)$. For example, if b is a nonzero constant, then $s(x) = b$ if and only if $f(x) = f(0)e^{x/b}$; and $s(x) = bx$ if and only if $f(x) = f(1)x^{1/b}$. In particular, the parabola $f(x) = x^2$ has subtangent $s(x) = x/2$, and the hyperbola $f(x) = 1/x$ has subtangent $s(x) = -x$. In fact, if $r \neq 0$ the function $f(x) = x^r$ can be defined as that function with $f(1) = 1$ whose subtangent is $s(x) = x/r$.

In [**1**] we used subtangents as an aid to draw tangent lines. Here we use them to calculate areas in a natural and intuitive geometric fashion rather than analytically.

2. TANGENT SWEEP AND TANGENT CLUSTER. A geometric method for calculating areas of many classical regions without using integral calculus has been described in [**1**] through [**5**]. It is based on a result called Mamikon's theorem, illustrated in Figure 1.

The lower curve on the left of Figure 1 is a more or less arbitrary smooth plane curve. A set of tangent segments to the lower curve defines a region that is bounded by the lower curve and an upper curve traced out by the other extremity of the tangent segments. This set is called a *tangent sweep*. It can be visualized dynamically as the region swept out by a tangent segment moving along the lower curve. When a parallel translation brings each point of tangency to a common point, as shown on the right of Figure 1, the set of translated segments is called the corresponding *tangent cluster*. Mamikon's theorem, which is proved in [**1**] and [**2**], states that *the area of the tangent cluster is equal to that of its tangent sweep*. One application given in [**2**] and [**3**] yields the area of the region below the graph of the power function $y = x^r$ and above the interval $[0, x]$ for positive r. This paper presents alternative ways for treating all real exponents r. It uses the following property of tangent sweeps:

> *If each tangent segment of a tangent sweep is scaled (expanded or contracted) by the same positive factor t, then the area of the tangent sweep is multiplied by t^2.*

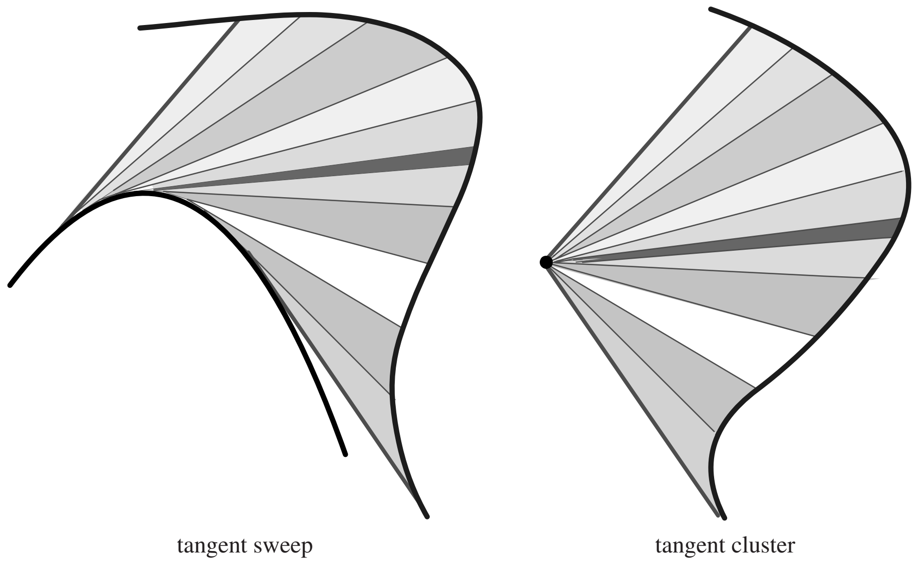

Figure 1. The tangent sweep of a curve and its tangent cluster have equal areas.

This follows from the fact that the corresponding tangent cluster of the scaled tangent sweep is also scaled radially by the factor t producing a similar figure with area multiplied by t^2.

3. PARABOLIC SEGMENT. We begin with one of the oldest problems in mathematics—finding the area of a parabolic segment, the shaded region shown in Figure 2a. The segment is inscribed in a rectangle of base x and altitude x^2. The area R of the rectangle is x^3, but we will not need this explicit formula for R. From the figure it is clear that the area of the parabolic segment is less than $R/2$. Archimedes made the stunning discovery that *the area is exactly* $R/3$, *one-third that of the rectangle*. This was deduced in [**2**] and [**3**] by a method that is simpler than that of Archimedes and more powerful because it also treats more general real powers x^r. This section gives another simple geometric approach, again using tangent sweeps.

The parabola has the equation $y = x^2$, but we shall not need this formula. We use only the fact that the subtangent is $x/2$, so the tangent line above any point x cuts

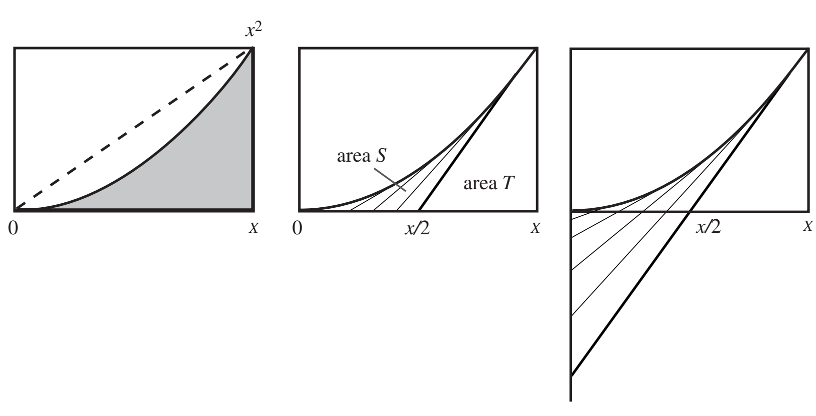

Figure 2. (a) A parabolic segment. (b) A tangent sweep of a parabolic segment cut off by the x-axis. (c) The region obtained by doubling the lengths of the tangent segments in (b).

off a segment of length $x/2$, as in Figure 2b. The shaded portion in Figure 2b is the tangent sweep obtained by drawing all tangent lines to the parabola between 0 and x and cutting them off at the x-axis.

In Figure 2b the parabolic segment is divided into two regions, the tangent sweep (whose area we call S), and a right triangle (of area T). We will prove that $S = T/3$, so the area of the parabolic segment, $S + T$, is equal to $4T/3$. Because $4T = R$, the area of the parabolic segment is $R/3$, as asserted.

To prove that $S = T/3$, refer to Figure 2c, where each tangent segment from Figure 2b has been doubled in length to reach the y-axis, as shown. The area of the scaled tangent sweep is $4S$. But the expanded region consists of two parts, the portion above the x-axis with area S, and the right triangle with area T below the x-axis. Hence $4S = S + T$, so $S = T/3$.

4. GENERAL POSITIVE POWERS. The argument in Section 3 can be extended to more general powers, in which x^2 is replaced by any power x^r with $r > 0$. When $f(x) = x^r$ the subtangent is x/r, so the tangent above any point x cuts off a segment on the x-axis of length x/r. An example with $r = 3$ is shown in Figure 3a, where the tangent above x cuts off a segment of length $x/3$. We will show that the cubic segment below the curve $y = x^3$ and above the interval $[0, x]$ has area $R/4$, one-fourth that of the circumscribed rectangle.

In Figure 3a the cubic segment is divided into two regions, the tangent sweep (of area S) and a right triangle (of area T). In this case we will prove that $S = T/2$. We expand each tangent segment in Figure 3a by a factor 3 so that it reaches the y-axis as shown in Figure 3b. The enlarged tangent sweep in Figure 3b has area $9S$. The portion of the enlarged region above the x-axis has area S and the portion below the x-axis is a right triangle with edges twice those in Figure 3a, so its area is $4T$. Therefore $9S = S + 4T$, hence $S = T/2$. The area of the cubic segment is $S + T = 3T/2 = 6T/4$. But $6T = R$, the area of the circumscribed rectangle, showing that *the area of the cubic segment is* $R/4$, as claimed.

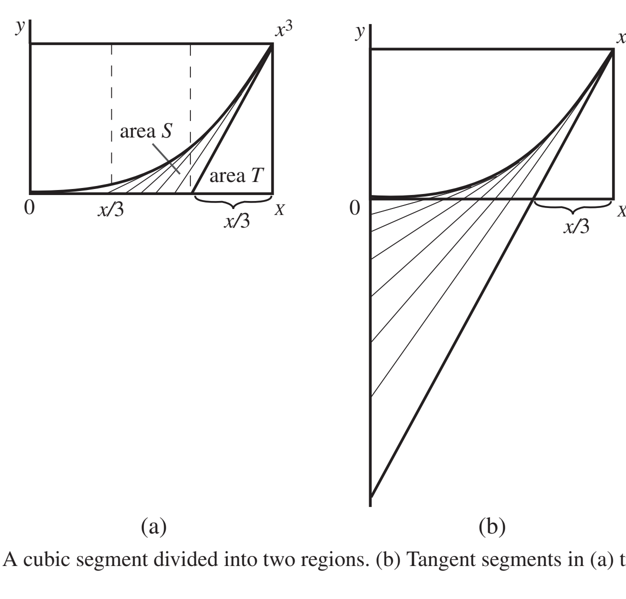

Figure 3. (a) A cubic segment divided into two regions. (b) Tangent segments in (a) tripled in length.

For a general power $r > 1$ the argument is similar. The general segment can be divided into two regions, a tangent sweep of area S cut off by the x-axis, and a right triangle of area T. We expand each tangent segment by a factor r to obtain an enlarged tangent sweep of area r^2S cut off by the y-axis. The portion of the enlarged region above the x-axis has area S, and the portion below the x-axis is a right triangle of area $(r-1)^2T$. Therefore $r^2S = S + (r-1)^2T$, so that $S = (r-1)^2T/(r^2-1) = (r-1)T/(r+1)$, and the area of the general segment is $S + T = 2rT/(r+1)$. But $2rT = R$, the area of the circumscribed rectangle, so *the area of the general segment is* $R/(r+1)$. In the language of integral calculus this is equivalent to

$$\int_0^x t^r\,dt = \frac{x^{r+1}}{r+1}.$$

The method also works if $r = 1$, giving $S = 0$; in this case the segment in question is a right triangle with area half that of the circumscribed rectangle.

If $0 < r < 1$ the graph changes shape from convex to concave, and the tangent sweep lies above the curve rather than below it. The analysis can be modified to cover this case, but the problem for $r < 1$ is easily reduced to that for exponents greater than 1. The example with $r = 1/2$ in Figure 4 shows us how to proceed in general. The portion of the rectangle *above* the graph is congruent to a parabolic segment with area $R/3$, so the portion below the graph has area $2R/3 = R/(1+1/2)$. (In calculus, this method amounts to integration by parts.)

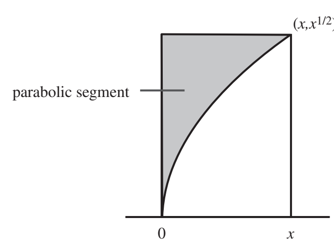

Figure 4. The area for $r = 1/2$ can be derived from that of the parabolic segment.

For the general case $0 < r < 1$ the portion of the rectangle above the graph of $y = x^r$ is congruent to a general segment with exponent $1/r > 1$, so its area is equal to $R/(1+1/r)$. Therefore the portion below the graph has area $R - R/(1+1/r) = R/(r+1)$, the same formula obtained for the case $r > 1$.

5. HYPERBOLIC SEGMENT. We turn next to negative exponents, starting with the function x^r for $r = -1$ whose graph is the rectangular hyperbola shown in Figure 5. We know from integral calculus that the area $A(x)$ of the hyperbolic segment lying below the graph of $y = 1/x$ and above the interval $[1, x]$ is $\log x$, the natural logarithm of x. We will deduce the same result without using integral calculus.

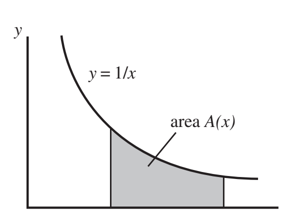

Figure 5. $A(x)$ is the area of the region below the hyperbola $y = 1/x$ and above the interval $[1, x]$.

In some treatments $A(x)$ is taken as the definition of the logarithm function, and then the exponential function is defined to be the inverse of the logarithm. Here we start by defining the exponential as that function with constant subtangent 1 and value 1 at 0 (see [**3**] or [**5**]). Then we show that *the area function* $A(x)$ *is the inverse of the exponential*, which gives us $A(x) = \log x$.

To do this we use a general property relating subtangents of inverse functions. Suppose that a differentiable function A has an inverse B, so that $y = A(x)$ if and only if $x = B(y)$. Then we have $B[A(x)] = x$ and $B'[A(x)]A'(x) = 1$. The subtangent function s associated with B is given by

$$s(y) = \frac{B(y)}{B'(y)},$$

so when $y = A(x)$ we have

$$s[A(x)] = \frac{B[A(x)]}{B'[A(x)]} = xA'(x).$$

The property is illustrated in Figure 6. The tangent line to the graph of $y = A(x)$ at $(x, A(x))$ cuts off a segment on the y-axis of length $xA'(x)$, and this segment is the subtangent of the inverse function B.

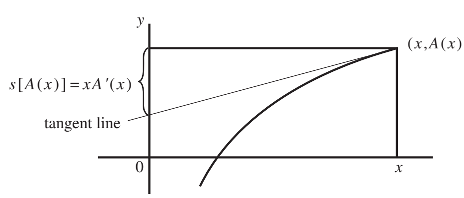

Figure 6. The tangent line to $y = A(x)$ cuts off a segment on the y-axis of length $xA'(x)$.

In particular, if f is a positive continuous function and if $A(x)$ is defined by the integral

$$A(x) = \int_a^x f(t)\,dt$$

from some fixed point a to an arbitrary x greater than a, then A is an increasing function with derivative $A'(x) = f(x)$, whence the tangent line to the curve $y = A(x)$ cuts off a segment on the y-axis of length $xf(x)$. This is the projection on the y-axis of the tangent to the curve $y = A(x)$, and as noted earlier, it is the subtangent of the inverse of A. If f is decreasing, as in Figure 5, the product $xf(x)$ is equal to the area of a rectangle of base x and altitude $f(x)$ inscribed under the graph of f. In particular, when $f(x) = 1/x$ the area of the rectangle above the interval $[0, x]$ inscribed under the hyperbola in Figure 5 is $xf(x) = 1$. Hence in this case the inverse of A has constant subtangent 1, so the inverse must be a constant times e^x. The initial value $A(1) = 0$ implies that the inverse of A is e^x, so $A(x) = \log x$.

6. GENERAL NEGATIVE POWERS. We turn next to the graph of $y = x^{-r}$, where $r > 1$, and we determine the area of the region over the interval $[x, \infty]$ for any $x > 0$.

Integral calculus tells us that the area is given by

$$\int_x^\infty t^{-r}\,dt = \frac{x^{1-r}}{r-1}. \tag{2}$$

We will prove this geometrically without using integral calculus. The general segment in question consists of two parts shown in Figure 7a, a right triangle of area $T = x^{1-r}/(2r)$, and a tangent sweep of area S, hence the segment has area $S + T$. This region is adjacent to a rectangle of base x and altitude x^{-r} whose area $R = x^{1-r} = 2rT$. We will show that $S + T = R/(r-1)$, the same result given by (2). Incidentally, this shows that S cannot be infinite.

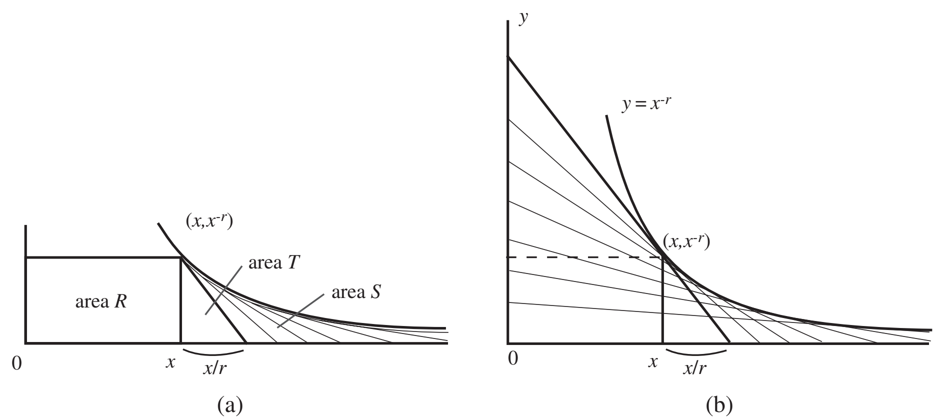

Figure 7. (a) A general segment divided into a tangent sweep of area S and a triangle of area T. (b) The tangent sweep cut off by the y-axis has area $r^2 S$.

Figure 7b shows the region obtained by drawing the tangent segments from each point of tangency (t, t^{-r}) to the y-axis for all $t \geq x$. The length of each tangent segment is r times that of the tangent segment at the same point cut off by the x-axis in Figure 7a, so the tangent sweep in Figure 7b has area $r^2 S$. It consists of two parts, a right triangle of area $(r+1)^2 T$, and the original tangent sweep of area S adjacent to it. Therefore $r^2 S = (r+1)^2 T + S$, which gives $S = (r+1)^2 T/(r^2 - 1) = (r+1)T/(r-1)$. Hence

$$S + T = \left(\frac{r+1}{r-1} + 1\right) T = \frac{2rT}{r-1} = \frac{R}{r-1}, \tag{3}$$

as required.

An approach using tangent clusters. Figure 8 illustrates another geometric approach that leads to the same result. Figure 8a shows a tangent cluster of the tangent sweep in Figure 7a obtained by translating each tangent segment parallel to itself so the point of tangency is brought to the origin. By Mamikon's theorem, this tangent cluster has area S. The tangent cluster and the right triangle adjacent to it in Figure 8a have area $S + T$. Next, reflect this region (tangent cluster plus triangle) through the x-axis, giving a congruent region of area $S + T$, then stretch it horizontally by a factor r (that is, multiply the x-coordinate of each point in the reflected region by r) to get a new region of area $r(S + T)$ indicated by the horizonal shading in Figure 8b. The stretched region is made up of two parts, a rectangle of area R, and the original region with area $S + T$ in Figure 8a. Therefore $r(S + T) = R + (S + T)$, so $S + T = R/(r-1)$.

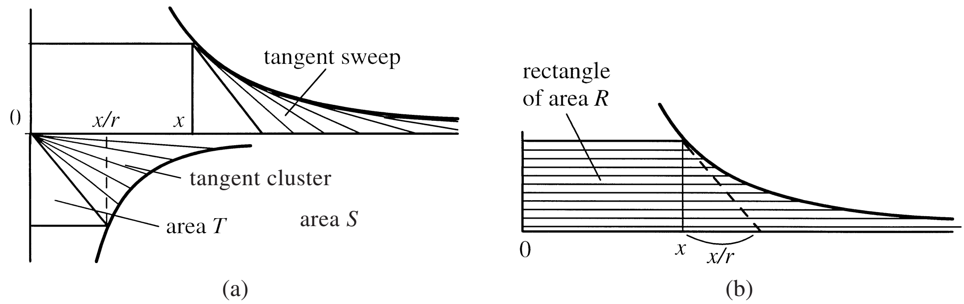

Figure 8. (a) Tangent cluster of the tangent sweep in Figure 7a adjacent to a triangle of area T. (b) Reflection of the lower region in (a) and horizontal stretching by r gives a region of area $r(S+T)$.

If $0 < r < 1$ the integral in (2) diverges, but in this case we have the formula

$$\int_0^x t^{-r}\,dt = \frac{x^{1-r}}{r-1}$$

for the area of the region under the graph of $y = x^{-r}$ and above the interval $[0, x]$. We leave it as an exercise for the reader to find a geometric proof of this result.

7. SUBTANGENTS AND MEAN VALUES. We turn now to a surprising relation between subtangents and mean values. The mean-value theorem for integrals states that, if f is continuous on an interval $[a, x]$, then for some c in $[a, x]$ we have

$$\int_a^x f(t)\,dt = f(c)(x-a). \tag{4}$$

The number $f(c)$ defined by (4) is called the mean value of f on the interval $[a, x]$. If f is nonnegative, equation (4) tells us that the area of the ordinate set of f over $[a, x]$ is equal to the area of a rectangle with base $x - a$ and altitude $f(c)$. If we knew the value of c, this would give a very simple way to calculate the area in question. In general there is no easy way to determine c, so (4) is not a useful formula for calculating areas.

If we replace $f(c)$ by $f(x)$ in (4) the formula is no longer correct, but it can be adjusted to make it so by replacing $x - a$ with another factor, say $h(x)$. Thus, in place of (4) we can write

$$\int_a^x f(t)\,dt = f(x)h(x) \tag{5}$$

for some $h(x)$. If $f(x) \neq 0$ we can define $h(x)$ by this equation, and (5) will be trivially correct, provided that $h(a) = 0$. We show now that $h(x)$ is our old friend the subtangent function associated with the area function

$$A(x) = \int_a^x f(t)\,dt.$$

The first fundamental theorem of calculus tells us that $A'(x) = f(x)$, implying that the subtangent S associated with A is given by

$$S(x) = \frac{A(x)}{A'(x)} = \frac{A(x)}{f(x)},$$

from which we find that $A(x) = f(x)S(x)$ or, in other words,

$$\int_a^x f(t)\,dt = f(x)S(x). \tag{6}$$

This shows that $A(x)$ is equal to the area of a rectangle with base $S(x)$ and altitude $f(x)$. Equation (6) can be regarded as an alternate form of the mean-value theorem in which $S(x)$ *represents a mean value of lengths of intervals on the* x*-axis*. If there is an easy way to determine $S(x)$, then we can find the area $A(x)$ by simply multiplying $S(x)$ by $f(x)$.

For example, let $f_0, f_1, f_2, \ldots$ be a given sequence of positive continuous functions on $[0, x]$ and let $A_1, A_2, A_3, \ldots$ denote the corresonding area functions, so that

$$A_{n+1}(x) = \int_0^x f_n(t)\,dt.$$

The mean-value property in (6) tells us that

$$A_{n+1}(x) = f_n(x)S_{n+1}(x), \tag{7}$$

where $S_{n+1}(x)$ is the subtangent of $A_{n+1}(x)$. We will now use (7) to show that the area function $A_{n+1}(x)$ of the power function $f_n(x) = x^n$ is $x^{n+1}/(n+1)$. We also use the fact that a function f and any constant times f have the same subtangent, because the constant factor cancels in the quotient $f(x)/f'(x)$ that defines the subtangent in (1).

Start with the constant function $f_0(x) = 1$. Its ordinate set over $[0, x]$ is a rectangle of base x and altitude 1, so its area function is $A_1(x) = x$. The ordinate set of $f_1(x) = x$ over $[0, x]$ is a right triangle with base x and altitude x, so its area is half the area of the rectangle, or $x^2/2$. In other words, the area function $A_2(x) = x^2/2$. By (7) the subtangent to $A_2(x)$ is $S_2(x) = A_2(x)/f_1(x) = x/2$. But $f_2(x) = 2A_2(x)$, hence $x/2$ is also the subtangent of the quadratic function $f_2(x) = x^2$. Once we know this, the method of tangent sweeps in Section 3 gives us $A_3(x) = x^3/3$ because that method depended only on the fact that $f_2(x)$ had subtangent $x/2$. From (7) we see that the subtangent to $A_3(x)$ is $x/3$, making this also the subtangent to the cubic function $f_3(x) = x^3$. Therefore the method of tangent sweeps in Section 4 gives us $A_4(x) = x^4/4$. By induction we see that $A_{n+1}(x) = x^{n+1}/(n+1)$ for all integers $n \geq 0$.

Exactly the same argument shows that, if we know the area function of some particular power function $f_r(x) = x^r$ with real $r > -1$, then we can determine the area functions of all the successive higher-power functions $f_{r+1}, f_{r+2}, \ldots$ for $r \neq 0$ from the recursion

$$A_{r+1}(x) = x\frac{r}{r+1}A_r(x).$$

REFERENCES

1. T. M. Apostol and M. A. Mnatsakanian, Subtangents—an aid to visual calculus, *Amer. Math. Monthly* **109** (2002) 525–533.
2. M. A. Mnatsakanian, On the area of a region on a developable surface, *Doklad. Armenian Acad. Sci.* **73** (1981) 97–101 (Russian); communicated by Academician V. A. Ambartsumian.
3. T. M. Apostol, A visual approach to calculus problems, *Engineering and Science* **LXIII** No. 3 (2000) 22–31. (An online version of this article can be found on the web site http://www.its.caltech.edu/~mamikon/calculus.html, which also contains animations displaying the method and its applications.)
4. M. Mnatsakanian, Annular rings of equal area, *Math Horizons* (November, 1997) 5–8.
5. T. M. Apostol and M. A. Mnatsakanian, Surprising geometric properties of exponential functions, *Math Horizons* (September, 1988) 27–29.

TOM M. APOSTOL joined the Caltech mathematics faculty in 1950 and became professor emeritus in 1992. He is director of *Project MATHEMATICS!* (http://www.projectmathematics.com) an award-winning series of videos he initiated in 1987. His long career in mathematics is described in the September 1997 issue of *The College Mathematics Journal*. He is currently working with colleague Mamikon Mnatsakanian to produce materials demonstrating Mamikon's innovative and exciting approach to mathematics.
California Institute of Technology, 1-70 Caltech, Pasadena, CA 91125
apostol@caltech.edu

MAMIKON A. MNATSAKANIAN received a Ph.D. in physics in 1969 from Yerevan University, where he became professor of astrophysics. As an undergraduate he began developing innovative geometric methods for solving many calculus problems by a dynamic and visual approach that makes no use of formulas. He is currently working with Tom Apostol under the auspices of *Project MATHEMATICS!* to present his methods in a multimedia format.
California Institute of Technology, 1-70 Caltech, Pasadena, CA 91125
mamikon@caltech.edu