An Application of Abstract Nonsense to Surface Area

Harriet Lord **Mathematics**

In this paper we present a brief history of attempts to calculate surface area for those surfaces whose area cannot be calculated using standard techniques of integral calculus, and we show the role played by abstract category theory, affectionately known as "abstract nonsense", in this very concrete subject. We then show that the technique used to find the length of a curve, or "arc length" does not work when generalized to find surface area, give a valid technique that is based on a "parametrizaion" of the surface, and investigate the problem of whether different parametrizaions result in different values for the surface area. It is in this that "abstract nonsense" plays an important role.

Historical Background

Archimedes and π

The fact that the ratio of the circumference of a circle to its diameter is constant was known well before the time of Archimedes. He was able to show that $223/71 < \pi < 22/7$ by calculating the perimeters of an inscribed regular polygon of 96 sides and a circumscribed regular polygon of 96 sides. The proof is based on previous results which show that the perimeter of the inscribed polygon is less than the circumference of the circle, which is less than the perimeter of the circumscribed polygon. The techniques used by Archimedes were that of geometric construction using straight edge and compass, and ratios of lengths. In the beginning of the proof he constructs an angle that is one third of a right angle, i.e., 30° and uses the fact that the ratio we call the cotangent of 30° is $\sqrt{3}$. He makes use of this with the following inequalities involving ratios.

$$1351:780 > \sqrt{3} : 1 \text{ and } \sqrt{3} : 1 > 265:153.$$

He then shows that the ratio of the diameter of the circle to the perimeter of the inscribed 96-gon is less than $22/7 = 3\frac{1}{7}$ while the ratio of the diameter of the circle to the perimeter of the circumscribed 96-gon is greater than $223/71 = 3\frac{10}{71}$. It is precisely his use of inscribed and circumscribed polygons that was to play an important role in the study of arc length.

Arc Length and Surface Area

With the advent of the Calculus Archimedes' method was adapted to find *arc length*. The term *arc length* is used for the length of a curve that has finite length. The primary difference between the method used by Archimedes and the practitioners of calculus is that Archimedes' calculation showed that the value of the circumference was greater than the perimeter of an inscribed regular polygon and less than the perimeter of a circumscribed one. By using polygons with a larger number of sides Archimedes and his followers were able to obtain better and better estimates. The technique used in the Calculus is to find the limit of the lengths of the perimeters. That the circumference of the circle is equal to this limit means that we can have the value of the perimeter of an n-gon as close as we wish to that of the circumference if we choose the value of n to be sufficiently large.

Various formulas for calculating arc length can be found in any standard calculus text book. These formulas are only valid for those curves with finite length that have the additional property that to each point on the curve there is a tangent line that varies continuously.

Formulas in calculus textbooks for the calculation of surface area have a similar restriction. Consequently there was a need for other methods for the calculation of arc length and surface area. The search for a method to find arc length was successful, and is illustrated in the next section for the curve that is the graph of $y = x^2$ for $0 \leq x \leq 1$. This method uses the limit of the length of inscribed polygonal lines.

It was J. A. Serret who in 1868 stated the “obvious” generalization of this method to the problem of calculating surface area.

> Given a portion of a curved surface bounded by a curve C, we call the area of this surface the limit S towards which the area of an inscribed polyhedral surface tends, where the inscribed polyhedral surface is formed by triangular faces and is bounded by the polygonal curve G, whose limit is the curve C.

Serret continued:

> One must show that the limit S exists and that it is independent of the way in which the faces of the inscribed surface decrease.

That Serret’s method does not work is illustrated by the counterexample in the section entitled “Surface Area”.

The surface used for this example is the lateral surface of a circular cylinder. It is interesting to note that the value for the actual area was first found by Archimedes. (See Proposition 13 in Book 1 of *On the Sphere and Cylinder* in Heath (c. 1950).)

The problem with Serret’s method was first discovered by H. A. Schwarz, who wrote to Gennochi, mathematics professor in Turin, about this in December of 1880. In the spring of 1882 Peano, a student of Gennochi, made the same discovery and announced it in a course at the University in Turin that semester. In the spring of 1882 Schwarz wrote to Hermite, mathematics professor in Paris, about his example. The published course notes of Hermite contain Schwarz’s letter. Since Peano’s course notes were published before those of Hermite, Peano claimed that he was the first one with the now famous counterexample.

Abstract Nonsense

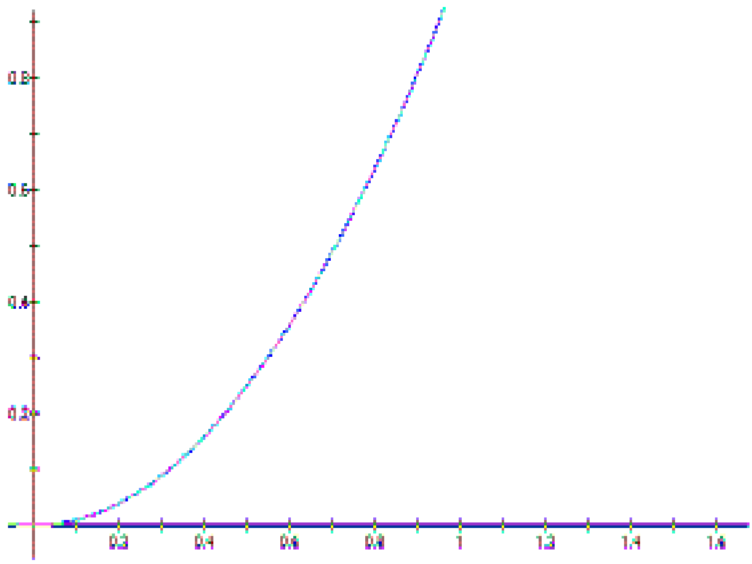
The term “abstract nonsense” was coined by the mathematician Norman Steenrod. It was used to point out the role of *morphisms*, which in our example are simply functions, in the new field of mathematics called Category Theory. According to Colin McLarty (McLarty 1998),

> Norman Steenrod first hung this tag on category theory. He had spent years trying to axiomatize homology, encouraged by Solomon Lefschetz. Lefschetz had also backed the young topologist Sammy Eilenberg, and encouraged Eilenberg’s collaboration with the algebraist Mac Lane explicating certain calculations in homology. When Eilenberg and Mac Lane created category theory, Steenrod saw he could use their way of emphasizing morphisms at least as much as objects. He happily said this “abstract nonsense” was the key to solving his problem.

Arc Length

We begin with a brief example of arc length.

Consider the graph of the parabola whose equation is $y = x^2$, where $0 \le x \le 1$.

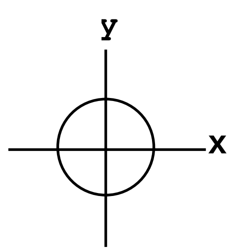


Before looking at the length of this curve (also called the *arc length*) we need to look at what is called a *parametrization* of this curve. By this we mean that we will write the x and y coordinates of each point on the curve as functions of a third variable t, which is called a *parameter*. This will be vital to our study of surface area. The simplest parametrization for this curve is the following:

$$\begin{cases} x = t \\ y = t^2 \end{cases} \quad \text{where } 0 \le t \le 1.$$

This can also be written as $f:[0,1] \to C$, where [0,1] consists of all t such that $0 \le t \le 1$ and C is the circle.

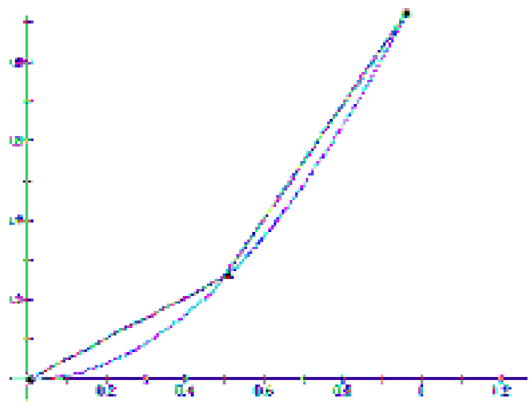
Another example of a parametrization of a curve is the following parametrization of a circle. The circle below is the graph of the equation $x^2+y^2 = 1$.



The most common parametrization is

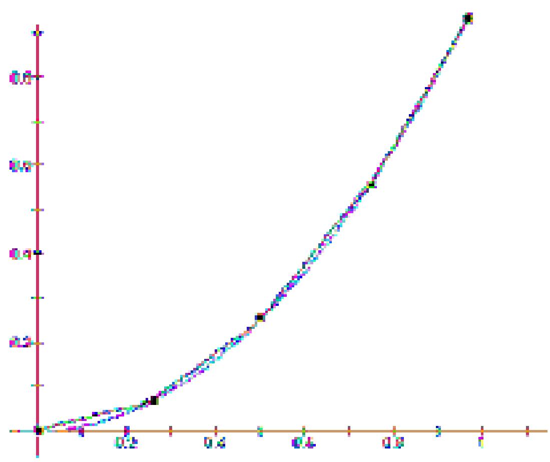
$\begin{cases} x = \cos t \\ y = \sin t \end{cases}$, where $0 \leq t \leq 2\pi$. (We recall from trigonometry that $\sin^2(t) + \cos^2(t) = 1$.) If we think of a point on the circle as the position of a particle at time t, then at time $t = 0$ the particle is at the point (1,0); it then travels around the circle exactly once in a counter-clockwise direction.

We begin our estimate of the length of the parabolic arc by dividing [0,1], the interval of values for t, into two subintervals of equal length, and thus obtain the subintervals [0,0.5] and [0.5,1]. We then find the points on the curve that correspond to the endpoints of these two subintervals and use these points to form an inscribed polygonal line.



Note that if n is the number of subintervals of the interval [0,1] then in this case $n = 2$.

For n =4 we have



What are the lengths of these inscribed polygonal lines?

For $n=2$ the length is $\sqrt{\frac{5}{16}}+\sqrt{\frac{13}{16}}=1.4604\ldots$.

For $n=4$ the length is $\sqrt{\frac{17}{256}}+\sqrt{\frac{25}{256}}+\sqrt{\frac{41}{256}}+\sqrt{\frac{65}{256}}=1.4743\ldots$.

The actual length of a curve is $\lim_{n\to\infty}\left\{l\left(P_n\right)\right\}$, where $\left\{P_n\right\}$ is a sequence of inscribed polygonal lines that converge to the curve and $l\{P_n\}$ is the length of the polygonal line.

The actual length of our parabola is $\frac{1}{4}(2\sqrt{5}+\ln(2\sqrt{5}))=1.4789\ldots$.

It is not difficult to show that if *f:[a,b]* $\to$ *C* and *g: [c,d]* $\to$ *C* are two different continuous parametrizations of the same curve *C* then if there exists a "well behaved function" (called a *homeomorphism) h:[a,b]* $\to$ *[c,d]* such that if *g(h(t)) = f(t)* for all *f* in the interval *[a,b]*, then the value obtained for the arc length using *f* is identical to that obtained using *g*.

Surface Area

The Counterexample

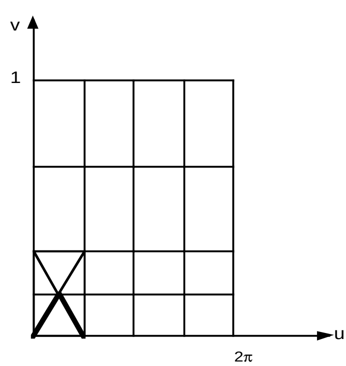
The following counterexample is that of Schwarz and Peano mentioned in the Historical Background section.

Consider a right circular cylinder of height $h=1$ and radius $r=1$ parametrized by

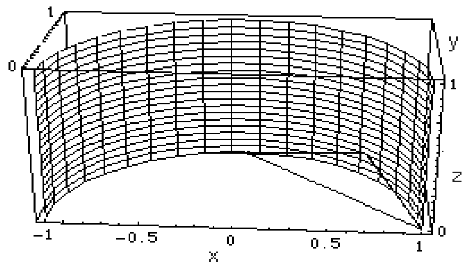
$$\begin{cases} x=\cos u \\ y=\sin u \\ z=v \end{cases} \quad \text{where} \quad \begin{array}{l} O\le u\le 2\pi \\ O\le v\le 1 \end{array}$$

In the *u-v* plane, the set of points (*u,v*) that satisfy both the above inequalities forms a rectangle. Suppose we divide the interval $[0,2\pi]$ into *m* subintervals and the interval [0,1] into *n* subintervals. This induces a division of the rectangle into *mn* subrectangles. If we then draw the two diagonals in each of these subrectangles we will have 4*mn* triangles. We will use the vertices of these triangles to construct the inscribed polyhedral surface.

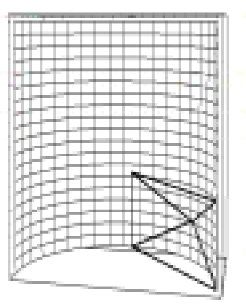
We illustrate the method with $m=4$ and $n=3$.



The triangle in the lower left corner outlined with heavy dark sides has vertices (0,0), $(\frac{\pi}{2},0)$, and $(\frac{\pi}{4},\frac{1}{6})$. These three points in the *uv* plane correspond to the following three points on the surface of the cylinder: (1,0,0), (0,1,0), and $(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},\frac{1}{6})$. We form one face of the polyhedral surface by joining each pair of these points with a straight line segment in 3-dimensional space. (Note that these line segments do NOT lie on the surface of the cylinder.)



If we do the same construction for each of the four triangles in the lower left rectangle we obtain four faces of the inscribed polyhedral surface.



If we continue in this manner for each of the 48 triangles we will obtain the inscribed polyhedral surface that corresponds to the case $m = 4$, $n = 3$.

In the general case we let P be the inscribed polyhedral surface of $4mn$ faces T ($m \geq 4$, $n \geq 1$) with vertices (x,y,z) images of the points:

$(\frac{2\mu\pi}{m},\frac{\upsilon}{n})$, where $\mu = 0, 1, ..., m-1$, $\upsilon = 0, 1, ..., n$ or

$(\frac{(2\mu+1)\pi}{m},\frac{(2\upsilon+1)}{n})$ where $\mu = 0, 1, ..., m-1$, $\upsilon = 0, 1, ..., n-1$.

Each of the resulting faces is an isosceles triangle and, in addition, these triangles are congruent to one another.

Each inscribed triangle has base $2\sin\left(\frac{\pi}{m}\right)$ and height

$$\sqrt{(1-\cos(\frac{\pi}{m}))^2 + (\frac{1}{2n})^2}.$$

The area of P, denoted $a(P)$, is

$$4mn \sin(\frac{\pi}{m})\sqrt{(1-\cos(\frac{\pi}{m}))^2 + (\frac{1}{2n})^2}.$$

If $m \to \infty$, $n \to \infty$, and $\frac{n}{m} = \alpha$ then

$$\lim_{m,n\to\infty} a(P) = 2\pi$$

which is the value of the surface area of the cylinder.

If $m \to \infty$, $n \to \infty$, and $\frac{n}{m^2} = \alpha$ then

$$\lim_{m,n\to\infty} a(P) = 2\pi\sqrt{\alpha\pi^4 + 1}.$$

<u>For any $\alpha > 0$ this value is larger than the surface area of the cylinder.</u>

Both limits can be calculated by use of L'Hôpital's Rule. This result remains valid if we replace the condition $\frac{n}{m^2} = \alpha$ $\lim_{m,n\to\infty} \frac{n}{m^2} = \alpha$. In this case, if $n = m$ we have that $\lim_{m,n\to\infty} \frac{n}{m^2} = \lim_{n\to\infty} \frac{n}{n^2} = 0$. Consequently $\alpha=0$ and $\lim_{m,n\to\infty} a(P) = 2\pi$ which is the correct answer.

Lebesgue Area

The correct method for calculating surface area was discovered by Henri Lebesgue in 1902 (Lebesgue 1902). Suppose that the function $f:A \to S$ is a parametrization of the surface S. (In the case where S is the circular cylinder of radius 1 and height 1 we saw that $f(u, v)$ = (cos u , sin v)). Lebesgue defined the surface to be $L(f)$, where

$$L(f) = \inf \underline{\lim}_{n\to\infty} \alpha(P_n)$$

Each $\{P_n\}$ is a sequence of polyhedral surfaces with the property that $P_n \to S$, and $a(P_n)$ denotes the area of the polyhedral surface P_n. The polyhedral surfaces in the sequence $\{P_n\}$ are those that can parametrized by quasi linear functions $f_n: A_n \to P_n$, where $\{A_n\}$ is a sequence of polygonal regions interior to A that conververges to A.

Since the value of $L(f)$ seems to depend on the parametrization used it is natural to ask if two different parametrizations of the same surface can give us two different values for the surface area. It is this question that will be (partially) answered in the next section.

Monotone-Light Factorizations

Suppose $f{:}A \to S$ is a function from a "nice" region in the plane onto a surface S. We called such a function a *parametrizaion* of the surface S. For example, if A is the rectangle in the uv-plane with $0 \le u \le 2\pi$, $0 \le v \le 1$, *and* $f(u,v) = (\cos\ u, \sin\ u, v)$ then S is the circular cylinder of radius 1 and height 1.

For any other parametrization $g{:}B \to S$ of the same surface we will give a condition that f and g must satisfy in order to guarantee that $L(f) = L(g)$.

A *fibre* of f is a set of the form $f^{-1}(S)$ for some point s on the surface of the cylinder.

In order to define the concepts of *monotone* and *light* we must first define the concept of a *connected subset of the plane*.

Two subsets of the plane are called *separated* if they both contain their boundaries and they are disjoint. A subset A of the plane is called *connected* if it is not the union of two disjoint separated sets.

For example if A is the set of points in the xy-plane that satisfy the inequality $x^2 + y^2 < 1$ then A consists of the points inside or interior to a circle of radius 1. The boundary of A is the circumference of the circle. Since the points on the circumference are not points in the set A we can say that A does not contain its boundary. That two sets are disjoint means that they have no points in common.

A *component* of the set A is a maximal connected subset; i.e., a connected subset of A that is not a subset of any other connected subset of A.

f is called *monotone* if its fibres are connected.

f is called *light* if the components of its fibres are sets containing exactly one element.

$f = lom$ is called a *monotone-light* factorization of f if m is monotone and l is light. If $f{:}A \to S$, $m{:}A \to T$, $l{:}T_{\leftarrow} S$, then $f = lom$ can be illustrated by the following diagram which is called a commutative triangle.

$$\begin{tikzcd}
A \arrow[rr, "f"] \arrow[dr, "m"'] & & S \\
& T \arrow[ur, "l"'] &
\end{tikzcd}$$

If $g{:}B \to S$ is another parametrization of the surface *S,* then *g* and *f* are said to be Keréjártó, or K-equivalent if the monotone-light factorization of *g* is

$$\begin{tikzcd} B \arrow[rr, "g"] \arrow[dr, "n"'] & & S \\ & T \arrow[ur, "l"'] & \end{tikzcd}$$

This means the *f* and *g* are K-equivalent if both functions have the same *light* factor in their monotone-light factorization.

> If *A*, *B*, and *T* are sufficiently “nice” (for example, they are all 2 cells) then the value obtained for the surface area of *S* using the parametrization *f* is equal to the value obtained using the parametrization *g*.

Rectangles, circles, and triangles are examples of 2-cells.

References

Archimedes περ‘ι σφαι′ρας κα‘ι κυλι′νδρου α′ ß´ (On the Sphere and Cylinder)

Cassina, H. (1950) L'Area di una superficie curva nel carteggio inedito di Genocchi con Schwartz ed Hermité, *Rendiconti Instituto Lombardo di Scienze e Lettere, 83*, 311-328.

Cesari, L. (1956) *Surface area*, Princeton: Princeton University Press.

Dijksterhuis, E. J. (1956), *Archimedes*, Copenhagen: Ejnar Munksgaard.

Heath, T. L. (c. 1950) *Works* , New York: Dover Publications, Inc.

Hermite, C. (1883) *Cours de M. Hermite, professé pendant le 2e semestre 1881-82,* Paris.

Kennedy, H. C. (1973) *Selected works of Giuseppe Peano*, Toronto and Buffalo: University of Toronto Press.

Lebesgue, H. (1902**)** Intégrale, longuer, aire, *Annali di Matematica pura ed applicata*, 7, 231-359.

McLarty, C. (1998, March 15) FOM: affectionate jokes, reply to Simpson, *Foundations of Mathematics e-mail list* Retrieved June 30, 1998 from the World Wide Web: http://www.math.psu.edu/simpson/fom/postings/9803.213.

Peano, G. (1882) *Lezioni di calcolo infinitesomale*, Torino.

Schwarz, H. A. (1972) *Gesammelte Mathematische Abhandlungen*, 2nd edition, Bronx NY: Chelsea Publishing Company.

Serret, J.A. (1868) *Cours de calcul différentiel et intégral*, two volumes, Paris, 1st edition 1868, 2nd edition 1880.