gerange–

Proceedings of the 2nd ASCM, –
Aug. 20–22, 1996. Kobe, Japan

Using a Natural Constraint to Approximate Area and Volume

Yongjian Ye[1)], Yiming Ye[2)] Xiang Zhang and QinShi Gao[3)]

Abstract. This paper presents methods to calculate the 2D area (or the 3D volume) enclosed by a plane curve (or space surface) without analytical expression. One property of our methods is that the error of the measurement is proportional to the square of the measuring unit when the unit is small enough. Finally, we point out a property common to calculating area and volume using the corresponding algorithm. This property can be used as a constraint to increase the accuracy of the estimation.

1. Background

In many engineering fields, we often need to calculate the area of a plane region or the volume of a 3D space region. Generally speaking, if we know the analytical expression of the curve bounding the plane region, we can easily calculate the area of the region by integration. The same applies to the calculation of the volume of a 3D region. But in many fields, such as computer vision or surveying and mapping, we often do not know the analytical expressions of the bounding curve for the plane region or the bounding surface for the 3D region. In this situation, we must use other methods to approximately calculate the area and volume. One simple method for calculating the plane area is the "square grid method" [4] [7]. The "square grid method" has the following property: the error of the "square grid" method is bounded by the length of the boundary of the region. Let $l \geq 1$ be the length of the curve bounding the region in the number plane, A be the true area of the enclosed region, N be the total number of integer points within the region (which is the approximate area obtained from the "square grid" method), we have: $\mid A - N \mid < l$. Hua and Wang [4] also described three practical ways of measuring the volume: the Baymah Formula, the truncated-cone method, and the trapezoid-method. These volumetric methods are simple but the errors are large and not evaluated.

In this paper, we propose much more accurate and efficient methods of measuring the area and volume. One property of our methods is that the error of the measurement is proportional to the square of the measuring unit when the unit is small enough. Another property is that our methods can be arbitrarily accurate. That is, we can set the value of the measuring unit in order to make the error of the measurement smaller than any given value ϵ. At the end of the paper, we point out an interesting property common to calculating the area of the plane region and the volume of the 3D region.

2. Basic Propositions

Let $G : \vec{\mathbf{r}} = \vec{\mathbf{r}}(t)$ be a curve, where t is the arc parameter belonging to $[T_1, T_2]$, and $\vec{r}$ is the

---

[1)] Institute of Mathematical Geology and Remote Sensing Geology, China Geoscience University, Wuhan, P.R. China, 430074

[2)] Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 1A4, Email: yiming@vis.toronto.edu. Yiming Ye is also a research fellow at Centre for Advanced Studies, IBM Canada from September 1995 to September 1996

[3)] Chinese Academy of Sciences, ZhongGuanCuan, Beijing, P.R.China 100081

corresponding radius with respect to t. We shall say G satisfies the fundamental assumption if $\frac{d^3\vec{\mathbf{r}}}{dt^3}$ exists everywhere and is uniformly bounded over $t \in [T_1, T_2]$,

Proposition 2.1 *Denote by s and l the length of an arc of the curve G and the length of the corresponding secant, respectively. Suppose G satisfies the fundamental assumption. Then the bound*

$$\frac{s-l}{l} \leq El^2 \tag{1}$$

holds if l is small enough. Here the constant E is given by

$$E = 8(\frac{1}{4}K_{max}^2 + \frac{1}{3}max_{t\epsilon[T_1,T_2]}\{\mid r^{'''}(t) \mid\}$$

and K_{max} is maximum curvature on G.

This proposition is first proved in [6] and [9], where the authors derive an important property in calculating the length of a space curve without analytical expression. However, the error of the method is not evaluated, because the condition that s must satisfy for 2.1 to hold is unknown. By careful analysis and calculation, we obtain the following result: the conclusion of Proposition 2.1 is true if

$$s < min\{\sqrt{\frac{1}{6U_m}}, \sqrt[3]{\frac{1}{6V_m}}, \sqrt[4]{\frac{1}{6W_m}}, \frac{E}{32V_m}, \sqrt{\frac{E}{32W_m}}, \frac{1}{3U_m}\sqrt{\frac{E}{8\sqrt{2}}}\sqrt{\frac{1}{3V_m}\sqrt{\frac{E}{8\sqrt{2}}}}, \sqrt[3]{\frac{1}{3W_m}\sqrt{\frac{E}{8\sqrt{2}}}}\},$$

where

$$U_m = \frac{1}{8}E; \quad V_m = \frac{1}{2}max_{t\in[T_1,T_2]}\{\mid \mathbf{r}^{''}(t) \mid\}max_{t\in[T_1,T_2]}\{\mid \mathbf{r}^{'''}(t) \mid\}; \quad W_m = \frac{1}{12}\{max_{t\in[T_1,T_2]}\{\mid \mathbf{r}^{'''}(t) \mid\}\}^2$$

The above result will be used to evaluate the error of our methods in calculating the area and volume. It can also be used to evaluate the error of the method in calculating the curve length proposed in [6] and [9].

3. The Area

3.1. Propositions

Proposition 3.1 *Let Ω be a region in the plane bounded by a straight line segment AB of length l and a curve $\widehat{AB}$ of length $P-l$, where P is assumed to be a constant satisfying $P > 2l$. Then the area of Ω is maximum when $\widehat{AB}$ is a circle arc.*

Proposition 3.2 *Consider a circle cut by secant AB into two arcs of length s and $s^{'}$. Let the length of AB be l. For l small enough, if $s - l \leq Cl^3$ then $s < s^{'}$, where C is a constant.*

Proposition 3.3 *Consider a circle with secant AB of length l (not the diameter). Let s be the length of arc $\widehat{AB}$, the shorter of the two arcs created by AB. Then the area A^* of the region enclosed by AB and $\widehat{AB}$ satisfies*

$$A^* < l\sqrt{(\frac{s}{2})^2 - (\frac{l}{2})^2}$$

.

Proposition 3.4 *Consider a curve G that satisfies the fundamental assumption. Let $\widehat{AB}$ be an arc on G, and AB be the corresponding secant line. Denote the length of $\widehat{AB}$ by s, the length of AB by l, and the area enclosed by AB and $\widehat{AB}$ by A^*. Then when l is small enough, $A^* \leq \sqrt{E}l^3$ uniformly on G, where E is the same as in Propositional 2.1. Equality holds when G is a straight line: $A^* = \sqrt{E}l^3 = 0$.*

Proof:

When G is a straight line, we have $A^* = 0$ and $E = 0$. Of course, $A^* = \sqrt{E}l^3 = 0$.

Otherwise, from Proposition 2.1, 3.1, 3.2, when l is small enough that these propositions hold and $l < \sqrt{\frac{2}{E}}$), we have:

$$A^* < l\sqrt{(\frac{s}{2})^2 - (\frac{l}{2})^2} \leq \frac{l}{2}\sqrt{(l + El^3)^2 - l^2} < \frac{l^3}{2}\sqrt{2E + E^2(\sqrt{\frac{2}{E}})^2} = \sqrt{E}l^3$$

3.2. The Algorithm

Now suppose we want to find the area A_0 enclosed by a plane curve G. First, we divide the region that contains G with a grid of squares with side length d. There are three classes of squares in the grid: (1) Squares such that all the points in the square belong to the region enclosed by G; (2) Squares such that some of the points in the square belong to the region enclosed by G; (3) Squares lying totally outside region enclosed by G.

Of course, the area A_0 enclosed by G is:

A_0= [the number of square grids of the first class] $\times$ d^2 + [the area of the region enclosed by G and falling in squares of the second class].

Suppose G intersects with the sides of the grid squares of the second class at points $P_1, ..., P_t$, numbering in a counterclockwise direction from an arbitrary starting point on G. Connect the points in a cycle $P_1, ..., P_t$, P_1 to obtain the closed polygonal curve L_d. Suppose the area enclosed by L_d is A_d. We take A_d as the approximation of A_0. We call this method the "improved square grid method".

Obviously, we have: A_d = [the number of grid squares of first class] $\times$ d^2 + [the area of the region enclosed by L_d and the corresponding sides of the grid squares of the second class].

Proposition 3.5 *Let G be a closed plane curve of length S and A_0 the area of the enclosed region. Let d be the grid dimension used in the improved square grid method, and A_d the estimated area. If G satisfies the fundamental assumption and d is small enough, then:*

$$| A_d - A_0 | \leq 2\sqrt{E}Sd^2,$$

where E is the same as in Lemma 1. Thus the error of the area measured by the improved square grid method is proportional to the square of d.

Proof:

From Proposition 3.4, when d and $\sqrt{2}d$ are small enough, we have:

$$\begin{aligned} | A_d - A_0 | \leq\ & \text{The area enclosed by } P_1P_2 \text{ and the corresponding arc} \\ & + \text{the area enclosed by } P_2P_3 \text{ and the corresponding arc} \\ & + \ldots + \text{the area enclosed by } P_tP_1 \text{ and the corresponding arc} \\ \leq\ & \sqrt{E}{l_1}^3 + \sqrt{E}{l_2}^3 + \ldots + \sqrt{E}{l_t}^3 \leq \sqrt{E}{l_1}^2 s_1 + \ldots \sqrt{E}{l_t}^2 s_t \leq (\sqrt{2}d)^2\sqrt{E}(s_1 + ... + s_t) \\ =\ & 2d^2\sqrt{E}S \end{aligned}$$

Where $l_i = \mid P_iP_{i+1} \mid, l_t = \mid P_tP_1 \mid, s_i = \mid \widehat{P_iP_{i+1}} \mid, s_t = \mid \widehat{P_tP_1} \mid$

Theorem 1. *Let $Y(d)$ be the area measured by the improved square grid method with grid dimension d. Obviously, $Y(d) = A_d$, and $lim_{d\to 0}Y(d) = A_0$. Define $Y(0) = A_0$. If G satisfies the fundamental assumption, then the derivative of $Y(d)$ at point $d = 0^+$ is 0.*

Proof:

$$Y^{'}(0^+) = lim_{d\to 0^+} \mid \frac{Y(d) - Y(0)}{d} \mid \leq lim_{d\to 0^+} \frac{2d^2\sqrt{E}S}{d} = 0$$

Theorem 1 can be used as a constraint when increasing the accuracy of the area calculation.

3.3. Setting the Precision

The "improved square grid method" described in Section 3.2. is based on discretization of the plane curve with line segments to approximately measure the area of the geometrical objects surrounded by the curve. Thus, if the area exists mathematically, the approximation should improve as the such discretization gets finer. Hence computationally significant results would be how small the error is with the knowledge of computed values about the geometrical objects, e.g., the grid dimension d used in the measurement. In the following, we give the relation that the grid dimension d should satisfy when the measurement error is required to be less than ϵ.

Suppose that the boundary G of a region satisfies the fundamental assumption. Let l be the length of a secant line, s the length of the corresponding arc and $M > 1$ be a constant. Assume that when $l < l^*$, then $s < Ml$ uniformly holds. For convenience, let $M = 10$. This is quite easily satisfied, and l^* will not be too small.

Recall the proofs of the above propositions. It is obvious that, in order for the conclusion of Proposition 4.1 to hold, l must satisfy:

$$l < min\{l^*, r_{min}, \frac{1}{10}\sqrt[3]{\frac{1}{6V_m}}, \frac{1}{10}\sqrt[4]{\frac{1}{6W_m}}, \frac{E}{320V_m}, \frac{1}{10}\sqrt{\frac{E}{32W_m}}, \frac{1}{30}\sqrt{\sqrt{\frac{8}{\sqrt{2}E}}},$$
$$\frac{1}{10}\sqrt{\frac{1}{3V_m}\sqrt{\frac{E}{8\sqrt{2}}}}, \frac{1}{10}\sqrt[3]{\frac{1}{3W_m}\sqrt{\frac{E}{8\sqrt{2}}}}\} = N$$

where r_{min} is the minimum value of the radius of curvature.

Obviously calculation of N requires knowledge of K_{max}, l^*, and the maximum value of the third derivative of radius with respect to arc. In practice, if the value of these parameters can be calculated or evaluated, then, the value of N can be calculated.

When the error in area measurement is required to be less than ϵ, i.e., $2\sqrt{E}Sd^2 < \epsilon$, then the grid dimension d can be set to any value that is less than $min\{\frac{1}{\sqrt{N}}, \sqrt{\frac{\epsilon}{2\sqrt{E}S}}\}$, where S is the length or an upper bound of the length of the boundary of the region.

4. The Volume

4.1. Propositions

Consider a $3D$ region Z bounded by a closed surface. Let A be the cross section plane formed by truncating Z with any plane, and let L be the boundary of A. Denote the diameter of A by d_A, which is the maximum distance between any two points of L.

In the following discussion, Z is required to satisfy the following condition: the closed plane curve L formed by truncating Z with any plane must meet the fundamental assumption. For convenience, we assume that E_{sup}, V_{sup}, W_{sup} are the maxima of E, V_m, W_m respectively for any L.

Proposition 4.1 *Consider a plane curve G satisfying the fundamental assumption. Let s be the length of an arc on G, and l be the length of the corresponding secant. Let H be the maximum distance from the arc to the secant. Then, when l is small enough, $H \leq \sqrt{E}l^2$ holds uniformly on G.*

Proposition 4.2 *Let Z,A,L, d_A have the same meaning as above. Suppose that region Z is such that any L formed by truncating Z with any plane meets the fundamental assumption. Region Z is divided by A into 2 parts. Let H be the maximum distance from the points on the surface of the smaller part of Z to A. Then when d_A is small enough, we have: $H \leq \sqrt{E_{sup}}d_A^2$.*

4.2. The Algorithm

Suppose the curve L obtained by truncating region Z with any plane meet the fundamental assumption. The algorithm to calculate the volume of Z is as follows. Build a $3D$ grid from a large cube that contains Z totally, where the length of the small grid cubes is d. The small cubes can be grouped into three classes: (1) cubes totally contained in Z; (2) cubes partially contained in Z; and (3) cubes outside of Z. Obviously, the volume of Z is given by

V_Z= [the number of first class cubes] $\times$ d^3 $+$ [the region enclosed by Z and falling in cubes of the second class].

The volume of the part of any second class cube that belongs to Z is calculated as follows.

Let $S^{*1}, S^{*2}, ..., S^{*f}$ be the surfaces formed by truncating the surface of Z with a cube. Generally, f =1, but sometimes there may be many little surfaces. For any such surface S^{*i}, its boundary L^{*i} is formed by the intersection S^{*i} and the surfaces of this cube. Of course, L^{*i} is a closed space curve. Going along L^{*i} counter-clockwise, we obtain the points $P_1, ..., P_g$, where $P_j (0 \leq j \leq g)$ is the intersection point of L^{*i} and a side of the cube. When $g = 0$, $g = 1$, or $g = 2$, we will not count the volume that belongs to Z and that is enclosed by the appropriate part of the surface of the cube and the corresponding part of S^{*i}. When $g \geq 3$, connect the points P_1, ..., P_g, P_1 in sequence to obtain a closed space polygonal line. Then connect P_1 and P_3, P_1 and P_4, ..., P_1 and P_{g-1} to obtain a series of triangles $\triangle P_1P_2P_3, \triangle P_1P_3P_4, ..., \triangle P_1P_{g-1}P_g$. Finally, the volume enclosed by S^{*i} and the corresponding surfaces of the cube will be replaced by the volume bounded by the plane surface formed by this series of triangles and the corresponding surfaces of the cube.

We call the above method the "Cube Cutting Method".

Proposition 4.3 *When the grid dimension d is small enough, the error δ of the volume of Z in any second class cube calculated with our "Cube Cutting Method" satisfies $\delta \leq (9\sqrt{E_{sup}}\sqrt{3}d^3(T-2)M + S_0{}^*)6\sqrt{E_{sup}}d^2$. Here M is the number of the surfaces inside this cube, T is the number of the intersection points of any surface with the side of the cube, and $S_0{}^*$ is the total area of the surfaces inside this cube.*

Proposition 4.4 *Suppose the region Z satisfies the conditions above. Let δ^* be the error of the volume calculated for Z with the "Cube Cutting Method". When d is small enough, we have: $\delta^* \leq (9\sqrt{E_{sup}}\sqrt{3}V^*(T^*-2)M^* + S)6\sqrt{E_{sup}}d^2$, where V^* is the volume of the large cube containing Z, S is the area of the surface of Z, T^* is the maximum of T (in Proposition 4.3) for any second class small cube, M^* is the maximum of M (in Proposition 4.3) for any second class small cube.*

Theorem 2: Suppose the $3D$ region Z satisfies the conditions above. Let $Y(d)$ be the volume calculated with the "Cube Cutting Method", where d is the length of the side of the small cubes of the grid. Define $Y(0) = V_0$, where V_0 is the real volume of Z. Then the derivative of $Y(d)$ at the point $d = 0^+$ is zero.

4.3. Setting the Precision

We can calculate volume in a manner similar to that used for area as described in Section 3.3.. Here, we omit the details.

5. Calculating Area and Volume Using Natural Constraint

The function $Y(x)$ defined in Theorem 1 and Theorem 2 gives the area or volume measured with the grid dimension $d = x$. These theorem imply the natural constraint that $Y'(x) = 0$ at $x = o^+$. In this section, we illustrate how this constraint can improve the accuracy of the calculation of area or volume.

For different grid dimensions d_i, we can use the method illustrated in Section 3.2 to obtain the corresponding estimates of area A_i and method illustrated in Section 4.2 to obtain the corresponding estimates of volume V_i. Thus, we can get a series of ordered pairs: (x_i, y_i), $1 \le i \le n$, where $x_i = d_i$, and $y_i = Y(x_i) = A_i$ if the area is to be measured, or $y_i = Y(x_i) = V_i$ if volume is to be measured. The function $Y(x)$ is interesting because when $x = 0$, it gives the exact area or volume of the geometric object. If we can obtain the function $Y(x)$, then the exact area or volume of the geometric object can be obtained. This is not possible in general, so replace $Y(x)$ by another analytic expression which can be handled as if it were the original function. This is called analytic substitution. Two considerations are involved in analytic substitution. First, what class of approximating functions shall we use? Second, how shall we select the particular member of the class?

It is not easy to obtain a perfect class of approximating functions. We should use all the information we know to determine the most promising class. One way of determining the class is to draw the points (x_i, y_i) on the coordinate system and analyze the pattern of these points and use the class of functions that can form a similar pattern. To select a particular member from the class, we need to answer the following two questions. The first is "what samples shall we use in obtaining the function?". The second is "How should we restrict the function class so as to get a better fit when x approaches 0?". To select the samples, we should analyze the nature of the problem and figure out where does the information lie. Since we want to find the value of $Y(0)$, we should use those samples where x is small and approaching 0. But on the other hand, if x is too small, then the noise in the measurement will prevent an accurate estimation from being made. In practise, such as the measuring of the area of land or the volume of a smooth mountain, it is impossible to use the infinite small grid dimension. Because the work required to obtain a measurement of area or volume will increase dramatically as the grid dimension decreases. Thus, we should obtain samples where x is small enough and the work required to measure the area or volume with this grid dimension is not too labor intensive. Since our task is to extrapolate the value of $Y(0)$ based on the given samples, the properties of function $Y(x)$ at point 0 have a great influence on the final result. Since $Y'(x)\mid_{x=0}= 0$ is true for the measurement of area and volume when the fundamental assumption is satisfied, this property should be used in the estimation of $Y(x)$ and the final result for the area and volume calculation should be more accurate in most cases when this property is used. We will illustrate the advantage of using $Y(x)\mid_{x=0}= 0$ in the analytic substitution in the following.

In general, the plane region or the space region can be in any form as long as the fundamental assumption is satisfied. In order to make the explanation and comparison easier, we only consider the area calculation, and we use a circle as the plane region to illustrate the idea. The radius of the circle is 116. Thus the exact area is 42251.9. This area will be used to compare the accuracy of different measurements. The different dimensions d and the areas measured with these dimensions are listed in the following table and will be used throughout the experiments.

Dimension d	36	32	28	21	16	9
Area measured with d	36496.4	41472	41957.40	42051.2	42106.4	42220.8

5.1. Polynomial Extrapolation

Because of simplicity, polynomials are often used in the process of analytic substitution of a tractable function $y(x)$ for an intractable one $Y(x)$. In general, high degree polynomials are not suitable for extrapolation, only lower degree polynomials seem to be safer. Here we compare the results of polynomial extrapolation when the natural constraint is used and when it is not used.

Suppose from the measurement we get n pairs of data (x_i, y_i), $(i = 1, 2, \ldots, n)$. These data can determine a polynomial of degree $n-1$

$$y(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} = \sum_{k=0}^{n-1} a_k x^k$$

by requiring $y(x)$ to pass through the n points (x_i, y_i), $(i = 1, 2, \ldots, n)$

$$y_i = \sum_{k=0}^{n-1} a_k x_i^k$$

Let

$$\Delta = \begin{vmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ & \ldots & & \\ 1 & x_n & \ldots & x_n^{n-1} \end{vmatrix}$$

$$\Delta_{a_0} = \begin{vmatrix} y_1 & x_1 & \ldots & x_1^{n-1} \\ & \ldots & & \\ y_n & x_n & \ldots & x_n^{n-1} \end{vmatrix}$$

Then

$$y(0) = a_0 = \frac{\Delta_{a_0}}{\Delta}$$

can be taken as the value of the area or volume.

When the natural constraint is used, the n pairs of data can determine a polynomial of degree n

$$y(x) = a_{0c} + a_{2c} x^2 + \ldots + a_{nc} x^n = \sum_{k=0, k\neq 1}^{n} a_{kc} x^k$$

by requiring $y(x)$ to pass through the n points (x_i, y_i), $(i = 1, 2, \ldots, n)$

$$y_i = \sum_{k=0, k\neq 1}^{n} a_{kc} x_i^k$$

Let

$$\Delta_c = \begin{vmatrix} 1 & x_1^2 & \ldots & x_1^n \\ & \ldots & & \\ 1 & x_n^2 & \ldots & x_n^n \end{vmatrix}$$

$$\Delta_{a_{0c}} = \begin{vmatrix} y_1 & x_1^2 & \ldots & x_1^{n-1} \\ & \ldots & & \\ y_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix}$$

Then

$$y(0) = a_{0c} = \frac{\Delta_{a_{0c}}}{\Delta_c}$$

can be taken as the value of the area or volume. Since the constraint expresses more accurate information about $y(x)$, the area calculated by a_{0c} should be more accurate in general.

The following table lists some results when two data points (x_i, y_i) and (x_j, y_j), or three data points (x_i, y_i), (x_j, y_j), and (x_k, y_k), are used. Where $no - C$ means the natural constraint is not used and C means it is used. From the table we can see that generally the error of area calculated when the natural constraint is used is smaller than the that when the constraint is not used.

Grid Dimensions	$Area_{no-C}$	$Area_C$	$Error_{no-C}$	$Error_C$
$32, 21$	43156.94	42489.32	905.04	237.4
$32, 16$	42740.79	42317.86	488.89	65.96
$32, 9$	42513.80	42285.12	261.91	33.22
$28, 21$	42332.60	42171.80	80.70	80.09
$28, 9$	42345.57	42251.14	93.67	0.75
$21, 9$	42348.00	42258.96	96.10	7.06
$16, 9$	42367.89	42273.75	115.99	21.85
$32, 21, 9$	42031.46	42221.53	220.43	30.35
$32, 16, 9$	42221.96	42265.85	29.92	13.95
$28, 21, 9$	42355.29	42276.65	103.39	24.75

5.2. Non-Linear Extrapolation

It frequently happens that polynomials cannot correctly express the function $Y(x)$. In this situation, we need to find a non-polynomial function $y(x)$ to replace $Y(x)$. In this section, we consider the class of $y(x)$ when there are nonlinear parameters and linear parameters mixed together. In particular, we consider the function $y(x)$ in the following form:

$$y(x) = de^{cx} + a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

Here, we assume that two estimations (x_1, y_1) and (x_2, y_2) are available and $c = -1$.

When the natural constraint is not used, we can determine two parameters. Thus $y(x) = de^{-x} + a_0$. Let $y(x)$ pass through (x_1, y_1), (x_2, y_2), then we have $y_1 = de^{-x_1} + a_0$ and $y_2 = de^{-x_2} + a_0$. Thus, we can get the value of d and a_0. Finally, we get the area

$$y(0) = d + a_0 = \frac{y_2 e^{-x_1} - y_1 e^{-x2} + y_1 - y_2}{e^{-x_1} - e^{-x_2}}$$

When the natural constraint is used, we can determine three parameters. Thus $y(x) = de^{-x} + a_1 x + a_0$. Let $y(x)$ pass through (x_1, y_1), (x_2, y_2), and let $y^{'}(0) = 0$. Finally we can get the area

$$y(0) = d + a_0 = \frac{y_2(e^{-x_1} + x_1) - y_1(e^{-x2} + x_2) + y_1 - y_2}{e^{-x_1} + x_1 - e^{-x_2} - x_2}$$

Some of the experimental results are listed in the following. The results show that the error calculated by using the natural constraint is smaller than the error calculated without it (Note, during the experiment, the values of x_1 and x_2 are divided by 10 to avoid overflow when running the C code. This operation does not influence the final value of $y(0)$).

.

Grid Dimensions	$Area_{no-C}$	$Area_C$	$Error_{no-C}$	$Error_C$
$28, 21$	43386.46	42230.82	1134.56	21.07
$28, 16$	42949.26	42219.23	697.36	32.66
$28, 9$	42672.88	42272.75	420.98	20.85
$21, 16$	42660.96	42211.65	409.06	40.24
$21, 9$	42552.50	42291.60	300.60	39.70

5.3. Least Squares Method

The least-squares method of fitting a curve is frequently used when there are more conditions to be satisfied than there are parameters to adjust. Here we use the least squares method to extrapolate the area. Since we only need to approximate the value of $Y(0)$, the data points available are not equally reliable. Typically, the data will be more reliable when the grid dimension approaches 0. Thus we attach suitable weights w_i to each data point. The closer the grid dimension is to 0, the more the weight is attached. Suppose the function used is a second degree polynomial, then, $y(x) = ax^2+bx+c$, when the constraint is not used, and $y(x) = ax^2 + c$, when the constraint is used.

When $y(x) = ax^2 + bx + c$, we wish to minimize $m(a, b, c) = \sum_{i=1}^{n}\{w_i(y(x_i) - y_i)^2\}$ with respect to a, b, c. Differentiating $m(a, b, c)$ with respect to a, b, and c and setting the results equal to zero, we get three equations:

$$\begin{cases} a\sum_{i=1}^{n} w_i x_i^4 + b\sum_{i=1}^{n} w_i x_i^3 + c\sum_{i=1}^{n} w_i x_i^2 = \sum_{i=1}^{n} w_i x_i^2 y_i \\ a\sum_{i=1}^{n} w_i x_i^3 + b\sum_{i=1}^{n} w_i x_i^2 + c\sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i x_i y_i \\ a\sum_{i=1}^{n} w_i x_i^2 + b\sum_{i=1}^{n} w_i x_i + c\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} w_i y_i \end{cases}$$

Solving the above equation for c, we can get the approximate value for the area.

When $y(x) = ax^2 + c$, we wish to minimize $m(a, c) = \sum_{i=1}^{n}\{w_i(y(x_i) - y_i)^2\}$ with respect to a, c. Similar to above, we can get the following two equations:

$$\begin{cases} a\sum_{i=1}^{n} w_i x_i^4 + c\sum_{i=1}^{n} w_i x_i^2 = \sum_{i=1}^{n} w_i x_i^2 y_i \\ a\sum_{i=1}^{n} w_i x_i^2 + c\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} w_i y_i \end{cases}$$

Solving the above equation for c, we can get the approximate value for the area when the constraint is used.

The table on the next page lists the experimental results of the above two method with different weights $(w_1, w_2, w_3, w_4, w_5, w_6)$ assigned to grid dimensions $(9, 16, 21, 28, 32, 36)$ respectively. We can see that the accuracy is greatly increased when the natural constraint is used.

6. Discussion

Theorem 1 and Theorem 2 provided natural constraint when using our method to calculate area and volume. This constraint can be used in practice to increase the accuracy of the measurement. It is interesting to point out that a similar result can be obtained for the method proposed in [5] and [6] for calculating the length of a space curve.

We describe the result here. Suppose OP is a space curve. We measure the length of OP with the line segment l, obtaining a series of points $P_0, P_1, ..., P_n, P_t$, where $P_iP_{i+1} = l$, for $0 \le i \le n$, and $P_nP_t = l^* \le l$. Let

$$Y(l) = \mid P_0P_1 \mid + \ldots + \mid P_{n-1}P_n \mid + \mid P_nP_t \mid = nl + l^*$$

Of course when l approaches 0, $Y(l)$ will be the length of the curve. Let S be the length of the curve and define $Y(0) = S$. We have the following result:

$$\mid Y'(l) \mid_{l=0+} \mid = 0$$

Weights $(w_1, w_2, w_3, w_4, w_5, w_6)$	$Area_{no-C}$	$Area_C$	$Error_{no-C}$	$Error_C$
$(701, 701, 401, 401, 1, 1)$	42284.66	42237.27	32.76	14.62
$(751, 751, 421, 421, 1, 1)$	42298.53	42236.79	46.64	15.10
$(801, 801, 451, 451, 1, 1)$	42292.86	42236.15	40.96	15.74
$(771, 771, 401, 401, 2, 2)$	42205.09	42246.99	46.80	4.90
$(891, 891, 401, 401, 2, 2)$	42220.50	42246.41	31.39	5.48
$(741, 741, 461, 461, 2, 2)$	42211.20	42245.14	40.69	6.75
$(811, 811, 401, 401, 3, 3)$	42134.80	42255.98	117.08	4.05
$(851, 851, 451, 451, 3, 3)$	42146.84	42253.46	105.05	1.57
$(811, 811, 501, 501, 3, 3)$	42149.78	42252.04	102.11	0.14
$(791, 791, 551, 551, 5, 5)$	42016.76	42266.64	235.11	14.74
$(891, 891, 581, 581, 5, 5)$	42056.78	42263.60	195.10	11.70
$(1700, 1700, 411, 411, 2, 2)$	42277.47	42245.22	25.57	6.67
$(1100, 1100, 491, 491, 3, 3)$	42189.13	42250.36	62.76	1.53
$(1900, 1900, 401, 401, 4, 4)$	42178.95	42257.99	72.94	6.09
$(1600, 1600, 511, 511, 5, 5)$	42129.79	42259.86	122.10	7.96

Acknowledge: The authors wish to thank the reviewers and Eric Harley for their valuable comments and to thank Eric Harley for correcting English.

References

[1] Richard F. Demar, *A simple approach to isoperimetric problems in the plane* , Math. Magazine, 48(1975), 1-12.

[2] Richard W. Hamming, *Introduction to Applied Numerical Analysis*, Hemisphere Publishing Corporation, U.S.A., 1971.

[3] Richard W. Hamming, *Numerical Methods for Scientists and Engineers*, McGraw-Hill Book Campany, Inc., U.S.A., 1962.

[4] LuoGen Hua and Yuan Wang, *Numerical integral and its applications*, China Science Publications, Beijing, P.R. China, 1963

[5] Wenqi Huang, *A method to increase the accuracy of predicating curve length*, Technique report, Huazhong Univ. of Sci. and Tech., P.R.China. 1974.1

[6] Wenqi Huang and Yiming Ye, *The length of space curve without analytical expression*, Applied Mathematics and Mechanics, (in Chinese), SiChuan, P.R.China, accepted.

[7] H. Steinhaus and Sur, *theoreme de M.V.Jarnik*, Colloquium Mathematicum, 1(1948), 1-5.

[8] Hermann Weyl, *Symmetry*, Princeton University Press, 1952.

[9] Yiming Ye, *The length of space curve without analytical expression*, B.S. Thesis, Huazhong Univ. of Sci. and Tech., Wuhan, P.R.China, 1985.