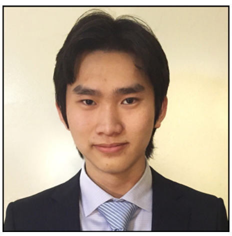
# How to Approximate the Volume of a Lake

Robert L. Foote and Han Nie

**Robert Foote** (footer@wabash.edu) received his B.A. from Kalamazoo College, and his M.S. and Ph.D. from the University of Michigan. He has been at Wabash College since 1989. His research interests include differential geometry and unexplored connections between Euclidean, spherical, and hyperbolic geometries. Besides mathematics, Foote enjoys cycling, playing trumpet, square dancing, and woodturning, a.k.a. $S^1$-Actions in Wood.



**Han Nie** (hn2275@columbia.edu) is in his senior year of a dual-degree program with Wabash College and the Columbia University School of Engineering. At Wabash, he majored in theater and minored in mathematics. At Columbia, he is majoring in operations research. In his spare time, Nie enjoys acting, working on theater productions, and seeking out authentic Chinese food.

To approximate the volume of a lake, limnologists (scientists who study lakes) use a curious, nonlinear numerical integration method based on the volume of a cone frustum. In this article, we give an error estimate for this method and empirically compare it to the trapezoid and Simpson's rules.

A detailed map of a lake typically shows depth contours—the depth of the lake is the same at two points on a given contour. Let $A(x)$ be the area contained in the contour at depth $x$. (One can find the area of a contour by using a planimeter.) The function $A : [0, d] \to \mathbb{R}$ is strictly decreasing and $A(d) = 0$, where $d$ is the maximum depth of the lake. See Figure 2 for examples from actual lake data.

The volume of the lake is given by $V = \int_0^d A(x)\,dx$. Of course, the function $A$ is never known exactly, and so the integral must be approximated numerically. The standard formula used by limnologists is

$$V \approx \sum_{k=1}^{n} \frac{A_{k-1} + A_k + \sqrt{A_{k-1}A_k}}{3}\,\Delta x \qquad (1)$$

where $A_k = A(x_k)$ for $k = 0, 1, \ldots, n$ are the areas of equally spaced depth contours and $\Delta x = d/n$ is the depth difference between consecutive contours [**11, 15**]. The rationale for this formula is that a slice of a lake between nearby horizontal planes is like a cone frustum, the volume of which is given by $V = (A_1 + A_2 + \sqrt{A_1A_2})h/3$ where $A_1$ and $A_2$ are the areas of the bases and $h$ is the height [**6**]. The volume of a cone frustum should, of course, be its height times some mean of $A_1$ and $A_2$. The quantity $(A_1 + A_2 + \sqrt{A_1A_2})/3$ is known as Heron's mean of $A_1$ and $A_2$ [**8, 13**]. The error made by the approximation (1) is zero when the lake is a cone, in which case the area function is $A(x) = K(x - d)^2$ where $K = A_0/d^2$ with $A_0$ being the area of the lake at the surface.

http://dx.doi.org/10.4169/college.math.j.47.3.162
MSC: 65D30

Our purpose here is to study the approximation (1). Given a continuous, nonnegative function $f : [a, b] \to \mathbb{R}$ and a positive integer $n$, define the *cone sum* as

$$C_n(f) = \sum_{k=1}^{n} \frac{f(x_{k-1}) + f(x_k) + \sqrt{f(x_{k-1}) f(x_k)}}{3} \Delta x \tag{2}$$

where $\Delta x = (b - a)/n$ and $x_k = a + k\,\Delta x$ for $k = 0, \ldots, n$. Let $L_n(f) = \sum_{k=1}^{n} m_k\,\Delta x$ be the lower Riemann sum with $m_k = \min f([x_{k-1}, x_k])$ and $T_n(f)$ the trapezoid sum. Since $L_n(f) \le C_n(f) \le T_n(f)$, it follows immediately that $\lim_{n\to\infty} C_n(f) = \int_a^b f(x)\,dx$.

Perhaps the most striking feature of (2) is its nonlinearity. It is positive homogeneous, i.e., $C_n(\alpha f) = \alpha C_n(f)$ for $\alpha \ge 0$, but not additive, i.e., $C_n(f + g)$ is generally not equal to $C_n(f) + C_n(g)$. If $g$ is nonvanishing, the reader can easily verify that $C_n(f + g) \ge C_n(f) + C_n(g)$, with equality for all $n$ if and only if $f/g$ is constant on $[a, b]$.

Our main result is that, under appropriate circumstances, there is a constant $K_f$ depending on $f$ such that

$$\left| \int_a^b f(x)\,dx - C_n(f) \right| \le \frac{K_f (b - a)^3}{n^2}, \tag{3}$$

so the error goes to zero at a rate comparable to that of the trapezoid and midpoint sums. The appropriate circumstances are, however, more restrictive than for the other sums, as will become clear in the theorem, its corollary, and the ensuing discussion.

Following the proof, we review the comparison of the cone and trapezoid sums and recommendations for their use by a noted limnologist, Lars Håkanson. We then compare these to Simpson's sum and an integrated cubic spline. (Note that Simpson's sum also gives the exact value for the volume of a cone since the area function is quadratic.) We suggest these latter standard tools are more accurate and may be easier to use than Håkanson's methods.

## The cone sum error estimate

We now turn to the proof of (3). It turns out that the constant $K_f$ is expressed more easily in terms of the function $g = \sqrt{f}$.

**Theorem (Cone Sum Error Estimate).** *Suppose that $f : [a, b] \to \mathbb{R}$ is nonnegative and $C^2$ (twice continuously differentiable). Let $g(x) = \sqrt{f(x)}$. Then*

$$\left| \int_a^b f(x)\,dx - C_n(f) \right| \le \frac{k_0 k_2 (b - a)^3}{6n^2} \tag{4}$$

*where $C_n(f)$ is defined in (2) and $k_0$ and $k_2$ are, respectively, the maximum values of $g(x)$ and $|g''(x)|$ for $x \in [a, b]$.*

*Proof.* First, we prove (4) for $n = 1$, which is equivalent to

$$\left| \int_a^b g^2(x)\,dx - \frac{g^2(b) + g^2(a) + g(b)g(a)}{3}(b - a) \right| \le \frac{k_0 k_2 (b - a)^3}{6}. \tag{5}$$

To prove this, we define the error function $E : [a, b] \to \mathbb{R}$ by

$$E(x) = \int_a^x g^2(t)\,dt - \frac{g^2(x) + g^2(a) + g(x)g(a)}{3}(x-a).$$

Then we have

$$\begin{aligned} E'(x) &= g^2(x) - \tfrac{1}{3}\big(2g(x)g'(x) + g(a)g'(x)\big)(x-a) \\ &\qquad - \tfrac{1}{3}\big(g^2(a) + g^2(x) + g(a)g(x)\big) \\ &= \tfrac{1}{3}\big(2g^2(x) - g^2(a) - g(a)g(x)\big) - \tfrac{1}{3}\big(g'(x)(2g(x) + g(a)\big)(x-a) \\ &= \tfrac{1}{3}\big(2g(x) + g(a)\big)\big(g(x) - g(a) - g'(x)(x-a)\big). \end{aligned} \tag{6}$$

Applying Lagrange's form of the remainder for Taylor's theorem to $g$ with the roles of $a$ and $x$ switched, there is a $c$ between $a$ and $x$ such that

$$g(a) = g(x) + g'(x)(a-x) + \tfrac{1}{2}g''(c)(a-x)^2.$$

It follows that

$$E'(x) = -\tfrac{1}{3}\big(2g(x) + g(a)\big) \cdot \tfrac{1}{2}g''(c)(x-a)^2. \tag{7}$$

By assumption, $\frac{1}{3}\big(2g(x) + g(a)\big) \le k_0$ and $\left|g''(c)\right| \le k_2$, so we have

$$|E'(x)| \le \tfrac{1}{2}k_0k_2(x-a)^2$$

for all $x \in [a, b]$. Integrating from $a$ to $b$ and noting that $E(a) = 0$ yields (5).

To see that the general case follows from the case for one interval, we restate (5) for the subinterval $[x_{k-1}, x_k]$:

$$\left|\int_{x_{k-1}}^{x_k} g^2(x)\,dx - \frac{g^2(x_{k-1}) + g^2(x_k) + g(x_{k-1})g(x_k)}{3}\Delta x\right| \le \frac{k_0k_2\Delta x^3}{6}.$$

Since $\Delta x = (b-a)/n$, this becomes

$$\left|\int_{x_{k-1}}^{x_k} g^2(x)\,dx - \frac{g^2(x_{k-1}) + g^2(x_k) + g(x_{k-1})g(x_k)}{3}\Delta x\right| \le \frac{k_0k_2(b-a)^3}{6n^3}.$$

Summing this inequality as $k$ goes from 1 to $n$ gives

$$\left|\int_a^b g^2(x)\,dx - C_n(g^2)\right| \le \frac{k_0k_2(b-a)^3}{6n^2},$$

which is equivalent to (4). ■

**Corollary.** *The cone sum error $\int_a^b f(x)\,dx - C_n(f)$ is zero for all $n$ if and only if $g = \sqrt{f}$ is linear, that is, $f$ has the form $f(x) = (cx+d)^2$ for $c, d \in \mathbb{R}$.*

*Proof.* If $g$ is linear, then $k_2 = 0$, which implies the error is zero. On the other hand, if the error is zero for all $n$, then $E'(x) \equiv 0$ and (7) implies that $g''(x) \equiv 0$. ■

Roughly speaking, the cone sum approximates the graph of $f$ on each subinterval using a quadratic, similar to Simpson's sum. In contrast to Simpson's sum, the quadratic is not arbitrary; it must have a double root. The corollary shows that the cone sum has zero error only for constant functions and quadratics with a double root. In particular, the error is nonzero for nonconstant linear functions and other quadratics.

Note that when $f(x) = A(x)$ is the cross-sectional area of a lake (the area of the contour at depth $x$), then the function $g$ has a geometric interpretation: $g(x)/\sqrt{\pi}$ is the radius of the circle of area $A(x)$, so the graph of $g/\sqrt{\pi}$ (turned so the negative $x$-axis points upward) is the profile of the solid of revolution (about the $x$-axis) with the same area versus depth function as the lake. It follows from the corollary that when $f$ is nonconstant, the error is zero if and only if this solid is a cone frustum, as in Figure 1 left. Graphs of area vs. depth and $\sqrt{\text{area}}$ vs. depth are common in limnology [**6, 9**].

**Trouble for the cone sum.** If $f(c) = 0$ for some $c \in (a, b)$, then the hypothesis in the theorem implies that $f'(c) = 0$, which may be more restrictive than desirable. (For the purpose of approximating lake volumes, this is not an issue since the area function is zero only at the right endpoint.) One might want, for example, to approximate $\int_a^b |F(x)|\, dx$, where $F$ has zero at $c$. In this case, the theorem can be applied to the intervals $[a, c]$ and $[c, b]$ separately, assuming $F$ has no other interior zeros. More generally, we could approximate $\int_a^b f(x)\, dx$ when $f$ is $C^2$ away from its zero set, assuming it has finitely many zeros.

This added flexibility does not, however, add to the usefulness of the theorem. In order for the error estimate to be useful, the quantity $|g''| = \frac{1}{4}|2ff'' - (f')^2|f^{-3/2}$ must be bounded. This can be problematic near places where $f = 0$, for example, at the bottom of the lake. It can be shown that if $g''$ is bounded near an isolated zero $c$ of $f$, then $f(x) \le k(x - c)^2$ for some constant $k$. Thus, assuming $g''$ is bounded is nearly as restrictive on $f$ as the assumption that $g''$ is zero.

In general, the cone sum error bound can be large for a function $f$ for which the trapezoid sum error bound (determined by $f''$) is small, suggesting that the cone sum may have difficulty with some functions that other methods may handle better.

**The worst case for the cone sum.** The well-known error estimates for the trapezoid and midpoint sums are sharp in the sense that the error inequalities become equalities for quadratic functions. These functions are then the worst case for these approximating sums. Similarly, quartic functions are the worst case for Simpson's sum. What is the worst case function for the cone sum?

Examining the proof of the theorem following (6), we see that the only way to have $(2g(x) + g(a))/3 \equiv k_0$ is for $g$ to be constant. In this case, $g'' \equiv 0$, so we may take $k_2 = 0$ and the error is zero. It follows that (4) is a strict inequality except when the error is zero, so $K_f = k_0k_2/6$ need not be the optimal constant for which an error estimate of the form (3) holds.

Suppose that $f$ is a function and $K \neq 0$ is a constant such that (3) is an equality for $n = 1$, $K_f = |K|$, and arbitrary $b > a$ in the domain of $f$. From (6), such a function would satisfy

$$E'(x) = \tfrac{1}{3}\big(2g(x) + g(a)\big)\big(g(x) - g(a) - g'(x)(x - a)\big) \equiv 3K(x - a)^2.$$

This differential equation does not appear to be solvable in closed form, however, numerical work suggests that functions satisfying this have a cone sum error strictly less than $|K|(b - a)^3/n^2$ when $n > 1$ and considerably less than that predicted with $K_f = k_0k_2/6$ as in (4). Thus, we conjecture that no constant $K_f$ exists for which (3) is an equality for $n > 1$ except when $g'' \equiv 0$ (in which case we can take $K_f = 0$).

It remains an open question to find a value of $K_f$ smaller than $k_0 k_2/6$, depending on $f$ in some reasonable way, for which (3) holds. From the error bound given in the theorem and numerical work, we conjecture that (3) holds for $K_f = k_{0,2}/6$ where $k_{0,2}$ is the maximum value of $g(x)|g''(x)|$ on $[a, b]$.

## Basin shape and cone versus trapezoid sums

It appears to be standard in limnology to use the cone sum to approximate lake volumes without considering its accuracy. Of the sources we found on computing lake volumes (e.g., [**1, 16**]), all but two use the cone sum with little or no motivation and with no suggestion about its limitations relative to basin geometry. Cole [**6**] is one of the exceptions, but he does not suggest an alternative volume formula. The other is Håkanson [**9**], who gives a detailed comparison of the cone and trapezoid sums for several lakes. One of his main points is that neither sum is an adequate numerical method for approximating all lake volumes; each does a good job for some basin geometries and a poor job for others. He gives recommendations about which to use based on convexity considerations.

It is not surprising that the basin shape of a lake will determine whether its volume is approximated better by $C_n(f)$ or $T_n(f)$. By the corollary, the cone sum is exact when $g'' = 0$, that is, when $f$ is a quadratic with a double root, i.e., $f(x) = A_q(x) \equiv A_0(d - x)^2/d^2$ for $0 \le x \le d$ where $A_0$ is the area of the lake at the surface and $d$ is the depth. Of course, the trapezoid sum is exact when $f$ is linear, $f(x) = A_\ell(x) \equiv A_0(d - x)/d$ for $0 \le x \le d$. Figure 1 left and center show circular lake basins for these two special cases. (As noted in the discussion following the corollary, the function $g/\sqrt{\pi}$ gives the profile of a circular lake with area function $f$.) When $g'' = 0$, the lake is a cone (as expected), and when $f'' = 0$, it is a paraboloid.

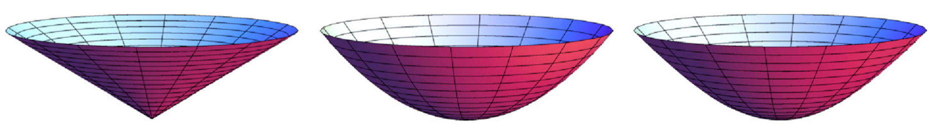


**Figure 1.** Basin shapes with $g'' = 0$, $f'' = 0$, and for the sinusoid model.

Håkanson uses the linear area function $f = A_\ell$ as the border between convex and concave lakes. He calls a lake convex if its area function lies below this line and concave if its area function lies above this line. He recommends using the cone sum for convex lakes and the trapezoid sum for concave lakes. For lakes that are very convex or concave, he describes a qualitative convexity index and uses it to determine a correction factor. The majority of the lakes in his study (73%) are convex, but he notes that some lakes are neither concave nor convex, and he gives some qualitative means to decide what to do in the latter case [**9, 10**]. His method is quite technical and would be difficult to automate. It also seems somewhat subjective and perhaps cumbersome to implement by hand.

We take some issue with Håkanson about when to use $C_n(f)$ vs. $T_n(f)$. Since $T_n(f)$ is exact when $f'' = 0$, the linear area function should not be the transition point, as the trapezoid sum will have less error than the cone sum for functions with $f''$ close to 0. For example, Neumann [**12**] has suggested an elliptic sinusoid as a model for an ideal lake basin. This basin is elliptical in cross section with the bottom having the shape of

$y = -\cos x$ for $-\pi/2 \le x \le \pi/2$. A circular version of this is shown in Figure 1 right. The lake profile and area functions for this model are given by $g(x) = C \arccos(x/d)$ and $A(x) = f(x) = C^2 \arccos^2(x/d)$ for $0 \le x \le d$ where $C = 2\sqrt{A_0}/\pi$. This is slightly convex in Håkanson's scheme; however, $T_n(f)$ has a smaller error than $C_n(f)$ (as we see below), and $g''$ is unbounded for this model.

The transition between using $C_n(f)$ and $T_n(f)$ lies somewhere between $A_\ell$ and $A_q$. However, to explore this is to miss an important point, namely that some other method might typically have less error than either the cone or trapezoid sum, making Håkanson's correction factor unnecessary. In the next section, we empirically compare the cone and trapezoid sums with Simpson's sum and an integrated cubic spline.

We should also note that limnologists classify basin shapes using a quantity called volume development $D_V$, the ratio of a lake's volume to that of a conical lake with the same area and depth [**6, 16**]. Generally speaking, when $D_V > 1$, the basin is U-shaped; when $D_V = 1$, the basin is V-shaped; and when $D_V < 1$, the basin shape is a pinched V. Volume development appears to correlate somewhat with Håkanson's convexity index, but he does not mention this in his papers.

## Numerical comparison with Simpson's sum

In this section, we empirically compare the cone sum with the trapezoid and Simpson's sums. From our experience in calculus, it is easy to simply reach for Simpson's sum as an easy improvement over the trapezoid sum. However, there are two things to keep in mind before doing this.

- The error estimate for Simpson's sum depends on the fourth derivative of $f$. If we have trouble estimating $f''$ and $g''$ for a lake profile, then we certainly cannot estimate $f^{(4)}$. The situation is actually not this bad—a more refined analysis [**14**] shows that the error estimate depends on the average value of $f^{(4)}$ on $[a, b]$. While this is also unknowable, the fluctuations of $f^{(4)}$ likely cause the average to be small.
- The main evidence that Simpson's sum is better than the trapezoid and midpoint sums (or at least the reason we tell our students) is that its error estimate depends inversely on $n^4$, and $1/n^4$ goes to zero much faster than $1/n^2$ as $n$ increases. This is useful only if $n$ is "large enough" or if it is easy to increase $n$. For a limnologist, increasing $n$ is a *very* expensive task! The density and shape of the depth contours depend on a grid of regularly spaced depth measurements followed by the process of piecing the contours together.

Figure 2 gives area vs. depth data for four lakes: Lake 230 from a collection of small lakes in Ontario [**6**], Lakes Ekoln and Vänern in Sweden [**9**], and Seattle's Lake Washington [**7**]. (In the references, the data are given in tabular form; the graphs in Figure 2 represent our interpolation with cubic splines, which connect data points with a sequence of cubic polynomials in such a way that the resulting function is $C^2$ [**4**].) The number of depth contours, here ranging from 11 to 19, is typical for examples in our references.

We approximate the volumes of these lakes in four ways: cone sum (2), trapezoid sum, a modified Simpson's sum (detailed in the appendix), and directly integrating the cubic spline. The results are shown in Table 1, which gives the relative errors of the first three methods compared to the last.

We see that the trapezoid sum does better than the cone sum in three of the four lakes. In Figure 2, we see that the one exception, Lake Vänern, has an area function

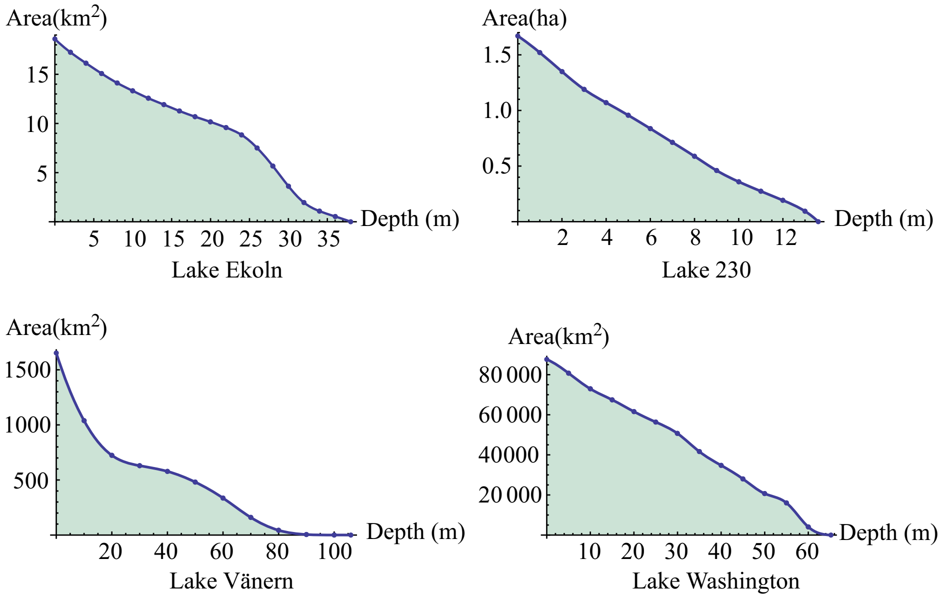



**Figure 2.** Depth vs. area for four lakes.

**Table 1.** Relative errors compared to integrated cubic spline.

| Lake\Method | cone | trapezoid | mod. Simpson |
|---|---|---|---|
| 230 | $-2.318\times10^{-3}$ | $-1.470\times10^{-4}$ | $-1.057\times10^{-4}$ |
| Ekoln | $-1.025\times10^{-3}$ | $4.375\times10^{-4}$ | $7.047\times10^{-5}$ |
| Vänern | $6.265\times10^{-3}$ | $1.370\times10^{-2}$ | $-3.012\times10^{-3}$ |
| Washington | $-2.376\times10^{-3}$ | $5.854\times10^{-4}$ | $1.793\times10^{-3}$ |

visually closer to a parabola tangent to the $x$-axis than the others, so it is quite reasonable that the cone sum does better than the trapezoid sum in this case. We also see that the area functions for Lakes 230 and Washington are visually closer to being linear than the others and that the trapezoid sum does much better than the cone sum for these. Furthermore, we see that the modified Simpson's sum does better than the cone sum in all four lakes and better than the trapezoid sum in three of the four. We hasten to point out that no firm conclusions can be made from four examples and that we have taken the integral of the cubic spline as the "exact" value in each case. The cubic spline is, however, a reasonable choice to represent the area function since it minimizes the amount of bending necessary to fit the data (see [**3**, p. 400] for a precise definition).

Table 2 gives the relative errors of the cone sum, trapezoid sum, Simpson's sum, and integrated cubic spline compared to the exact volume for the sinusoid model (introduced in the section on basin shapes). The sums and cubic spline are based on ten equal subintervals. We see that, in this order, they give increasingly better approximations to the exact value, with the last two being considerably better than the first two.

**Table 2.** Relative errors compared to exact value for sinusoid model.

| | cone | trapezoid | Simpson | cubic spline |
|---|---|---|---|---|
| sinusoid model | $-4.281\times10^{-3}$ | $8.325\times10^{-4}$ | $7.035\times10^{-7}$ | $5.505\times10^{-7}$ |

## Conclusions and further research

While the cone and trapezoid sums are useful for rough approximations and have pedagogical value, the error estimate in the theorem and the empirical work of Håkanson [**9, 10**] show that neither method works well in all cases. Håkanson recommends using a convexity index to decide between the cone and trapezoid sums and to determine a correction factor. Our numerical work suggests this could be replaced with either a modified Simpson's sum or an integrated cubic spline fitted to the data. If limnologists rely on software packages for lake volume approximations, then these standard tools could be incorporated into the software, while making it clear to the users that they are similar to the cone and trapezoid sums, but more sophisticated.

There is other interesting mathematics in limnology. The way that lake surface area varies with volume was treated by Cass and Wildenberg [**5**]. Another metric used by limnologists is the *shoreline development* of a lake, defined as $\mathrm{SLD} = L/(2\sqrt{\pi A})$ where $L$ is the length of the shoreline and $A$ is the area of the lake at the surface [**16**]. This is a measure of how much the shoreline deviates from being a circle: $\mathrm{SLD} \geq 1$ with equality precisely when the lake is a circle. Of course, this is known to mathematicians as the isoperimetric inequality! [**2**]

We close with some questions for further research, which might focus on approximating $\int_a^b f(x)\,dx$ in which $f$ is strictly decreasing and $f(b) = 0$.

- Does some particular weighted average of $C_n(f)$ and $T_n(f)$ approximate the integral at least as well as the modified Simpson's sum or perhaps some adaptive weighted average $\lambda C_n(f) + (1-\lambda)T_n(f)$ in which $\lambda$ depends somehow on $f$?
- Which does a better job, the modified Simpson's sum or integrated cubic spline, especially for a limited number of subintervals?
- Is there any benefit to applying one of the standard numerical methods to $f^{-1}$?

**Acknowledgment.** The authors would like to thank Chad Westphal for his helpful comments.

**Summary.** We study a nonlinear numerical integration method used by limnologists (scientists who study lakes) to approximate the volume of a lake. After proving an error estimate and making empirical comparisons, we suggest using Simpson's sum or integrated cubic splines.

### References

1. *A Beginner's Guide to Water Management—Lake Morphometry*. Second ed. Florida Lakewatch, 2001, https://edis.ifas.ufl.edu/fa081.
2. V. Blåsjö, The isoperimetric problem, *Amer. Math. Monthly* 112 (2005) 526–566, http://dx.doi.org/10.2307/30037526.
3. B. Bradie, *A Friendly Introduction to Numerical Analysis*. Pearson Prentice Hall, Upper Saddle River, NJ, 2006.
4. R. L. Burden, J. D. Faires, *Numerical Analysis*. Eighth ed. Thomson Brooks/Cole, Belmont, CA, 2005.
5. D. Cass, G. Wildenberg, Relations between surface area and volume in lakes, *College Math. J.* **21** (1990) 384–389, http://dx.doi.org/10.2307/2686904.
6. Gerald A. Cole, *Textbook of Limnology*. Third ed. Mosby, St. Louis, 1983.

7. W. T. Edmondson, Lake Washington, in *The Lakes Handbook*. Vol. 2. Eds. P. E. O'Sullivan, C. S. Reynolds. Blackwell, Malden, MA, 2005. 96–116.
8. M. K. Faradj, *Which Mean Do You Mean? An Exposition on Means*, Master's thesis, Louisiana State University, Baton Rouge, 2004, http://etd.lsu.edu/docs/available/etd-07082004-091436/.
9. L. Håkanson, On lake form, lake volume and lake hypsographic survey, *Geogr. Ann. A* **59** (1977) 1–30, http://dx.doi.org/10.2307/520579.
10. L. Håkanson, *A Manual of Lake Morphometry*. Springer, New York, 1981.
11. G. E. Hutchinson, *A Treatise on Limnology*. Vol. I. Wiley, New York, 1957.
12. J. Neumann, Maximum depth and average depth of lakes, *J. Fish. Res. Board Can.* **16** (1959) 923–927, http://dx.doi.org/10.1139/f59-065.
13. E. Neuman, J. Sándor, Generalized Heronian means, *Math. Pannon.* **19** (2008) 57–70.
14. E. Rozema, Estimating the error in the trapezoidal rule, *Amer. Math. Monthly* **87** (1980) 124–128, http://dx.doi.org/10.2307/2321991.
15. P.S. Welch, *Limnological Methods*. Blakiston, Philadelphia, 1948.
16. R. G. Wetzel, *Limnology*. Saunders, Philadelphia, 1975.

## Appendix: Modified Simpson's sum

To employ the numerical methods, we note that the partition points along the horizontal axis in Figure 2 are not quite equally spaced. The last subinterval (at the bottom of the lake) is smaller than the others for most of the lakes. This causes no problem for the cone and trapezoid sums since each term of those sums depends only on one subinterval. Simpson's sum, on the other hand, must be modified to accommodate the last subinterval and the possibility that the number of subintervals is odd.

**Modified Simpson's sum for an even number of subintervals.** In this case, if the last subinterval is smaller than the others, then apply the usual Simpson's sum on $[x_0, x_{n-2}]$ and use the integral of the quadratic Lagrange polynomial on the last two subintervals, $[x_{n-2}, x_n]$. For notational simplicity, denote $x_{n-2}, x_{n-1}, x_n$ by $a, b, c$. The quadratic Lagrange polynomial for $f$ on $[a, c]$ with partition points $a < b < c$ is

$$L_2(x) = f(a)\frac{(x-b)(x-c)}{(a-b)(a-c)} + f(b)\frac{(x-a)(x-c)}{(b-a)(b-c)} + f(c)\frac{(x-a)(x-b)}{(c-a)(c-b)}.$$

Use $\int_a^c f(x) \approx \int_a^c L_2(x)\,dx$. The error in this approximation has order $\mathcal{O}(h^4)$, where $h = \max_k(x_k - x_{k-1})$. When the partition points are equally spaced, this is the same as Simpson's sum [**4**].

**Modified Simpson's sum for an odd number of subintervals.** In this case, apply the usual Simpson's sum on $[x_0, x_{n-3}]$ and use $\int_{x_{n-3}}^{x_n} f(x) \approx \int_{x_{n-3}}^{x_n} L_3(x)\,dx$ where $L_3$ is the cubic Lagrange polynomial on the last three subintervals, $[x_{n-3}, x_n]$. The error in this approximation has order $\mathcal{O}(h^5)$. When the partition points are equally spaced, this is the same as Simpson's three-eighths rule [**4**].