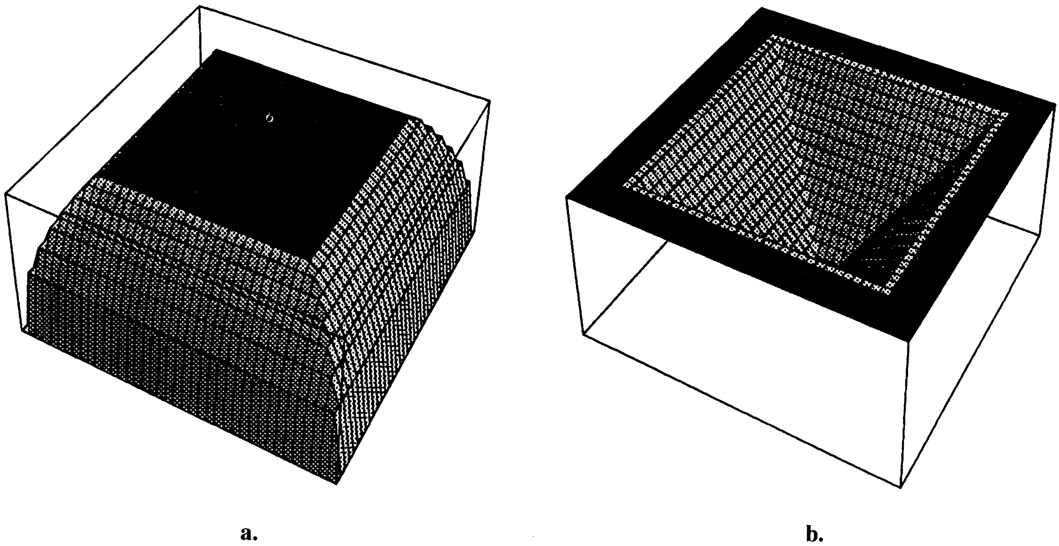

Figure 4
The black square and the black square-shaped ring cut by a plane at height z have the same area.

The horizontal plane at height z intersects this solid in a square-shaped ring of area $(2r)^2 - (2z)^2$. Since both of the regions cut by the plane have the same area, namely $4r^2 - 4z^2$, the two solids have equal volume by Cavalieri's Principle. The pyramid has one third of the volume of the corresponding rectangular solid, so we conclude that the volume of the bicylinder is two thirds that of its circumscribing cube.

Can anyone refer me to a good tombstone engraver? No hurry, of course.

Round-off, Batting Averages, and Ill-Conditioning

Edward Rozema, University of Tennessee, Chattanooga, TN 37403-2598

One summer day, as I was watching a televised Atlanta Braves baseball game, Mark Lemke, the Braves' second baseman, got 5 hits in 6 official at bats.* The television announcer, Skip Carey, commented that Lemke raised his batting average from .182 to .210; however, he didn't give Lemke's total number of hits or his total number of official at bats. This raised a question which I thought would be suitable for my precalculus classes: How many total hits and official at bats did Lemke have at the beginning of the game? This simple question will throw us a few curves and lead to an investigation of round-off errors, ill-conditioned systems, and interval analysis.

*For this article, you will need to know that a player's *batting average* is computed by dividing the total number of hits by the total number of official at bats (both are integers). The batting average is always reported after rounding the value to three correct decimal digits. For example, if a player has 20 hits in 70 official at bats, then the batting average is $20/70 = 0.286$.

Let's begin with a straightforward model: Let

x = Number of official at bats at the beginning of the game

y = Number of hits at the beginning of the game.

Then

$$\frac{y}{x} = 0.182 \qquad \text{and} \qquad \frac{y+5}{x+6} = 0.210. \tag{1}$$

These equations appear to have a unique solution. Clearing denominators yields the linear system

$$\begin{aligned} y &= 0.182x \\ y+5 &= 0.210(x+6). \end{aligned} \tag{2}$$

Substitution yields $x = 3.74/0.028 = 133.57\ldots$ and $y = 0.182(133.57\ldots) = 24.3097\ldots$.

The problem is beginning to look interesting, since neither of these numbers is particularly close to an integer. Rounding to the nearest integers yields $(x, y) = (134, 24)$. However, $y/x = 24/134 = 0.17910\ldots \approx 0.179 \neq 0.182$; so this can't be right. It is tempting to try all combinations of the nearest integer values: $x = 133$ or 134 and $y = 24$ or 25; then $24/133 = 0.180$, $25/133 = 0.188$, and $25/134 = 0.187$. This doesn't seem to be getting anywhere. Although knowing that a solution exists somewhere nearby may motivate us to continue searching, this is not very satisfying. It appears that the seemingly innocent act of rounding the batting average to three decimal digits has caused a relatively large change in the solution of system (1) or (2); such systems are known as *ill-conditioned* systems.

Let's start over and incorporate the round-off errors into the equations (this is the beginning of *interval analysis*). Suppose that $y/x = 0.182 \pm 0.0005$ and $(y+5)/(x+6) = 0.210 \pm 0.0005$. This gives the linear system

$$y = (0.182 \pm 0.0005)x \qquad \text{and} \qquad y + 5 = (0.210 \pm 0.0005)(x+6). \tag{3}$$

Substitution then yields $x = (3.74 \pm 0.003)/(0.028 \pm 0.001)$. The smallest value for x is $x = 3.737/0.029 = 128.86\ldots$ and the largest value is $x = 3.743/0.027 = 138.6296\ldots$. This is a remarkable spread of nearly 10 at bats; the three correct significant digits that we started with have disappeared (along with any hope for using this problem in precalculus).

Using the formula $y = (0.182 \pm 0.0005)x$, we obtain a minimal value for y of $y = 0.1815(128.86\ldots) = 23.38\ldots$ and a maximal value of $y = 0.1825(138.6296\ldots) = 25.29\ldots$. Well, that's a little better: There are only two choices for y, namely 24 or 25. We could now check (1) using all pairs (x, y) where x is any integer from 129 to 138 and y is 24 or 25. This is a little tedious; besides, if y is known, then $x = y/(0.182 \pm 0.0005)$. Thus $y = 24$ implies $x \in [131.5\ldots, 132.2\ldots]$ and $y = 25$ implies $x \in [136.98\ldots, 137.7\ldots]$. We are down to only two possibilities:

$$(x, y) = (132, 24) \qquad \text{and} \qquad (x, y) = (137, 25).$$

Checking the original equations in (1) shows us that both of these pairs are solutions. Unfortunately, I have no idea whether Lemke had 24 or 25 hits at the beginning of the day—this is a contingent fact of history beyond the scope of mathematics.

Now, can we see where the significant digits were lost in the computations? Glancing through the calculations shows that the digits were lost in the subtraction when substitution was used to solve system (3). The coefficient of x was $(0.210 \pm 0.0005) - (0.182 \pm 0.0005) = 0.028 \pm 0.001$. One significant digit was completely lost and the uncertainty in the last digit was doubled. As is so often the case, subtraction of almost equal approximations resulted in a loss of significant digits. In this example, the loss was unavoidable; in many situations, there are strategies for avoiding such subtractions (such as pivoting in Gaussian elimination; see [**3**]).

More insight into the problem can be given through geometry. In Figure 1 the graphs of the linear equations in (2) show us that the system is ill-conditioned: The two lines are nearly parallel so that the uncertainty in the point of intersection is quite large. Figure 2 shows the graph of the inequalities implied by (3):

$$0.1815x \le y \le 0.1825x$$
$$0.2095x - 3.737 \le y \le 0.2105x - 3.743.$$

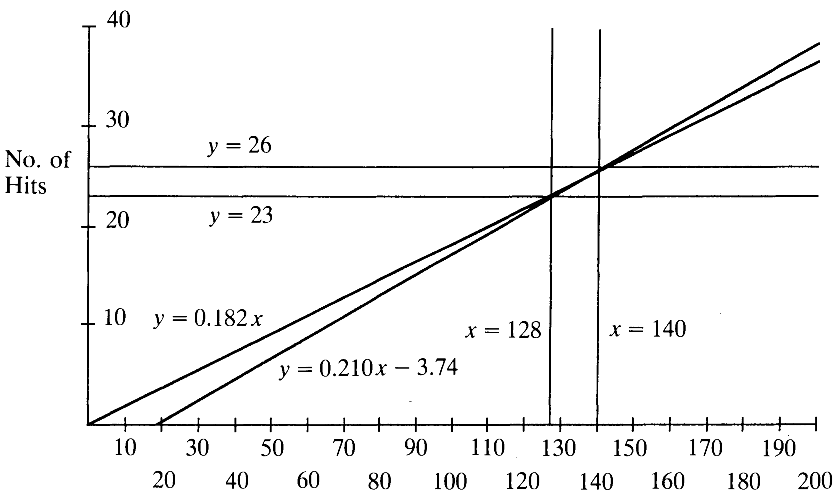

Figure 1

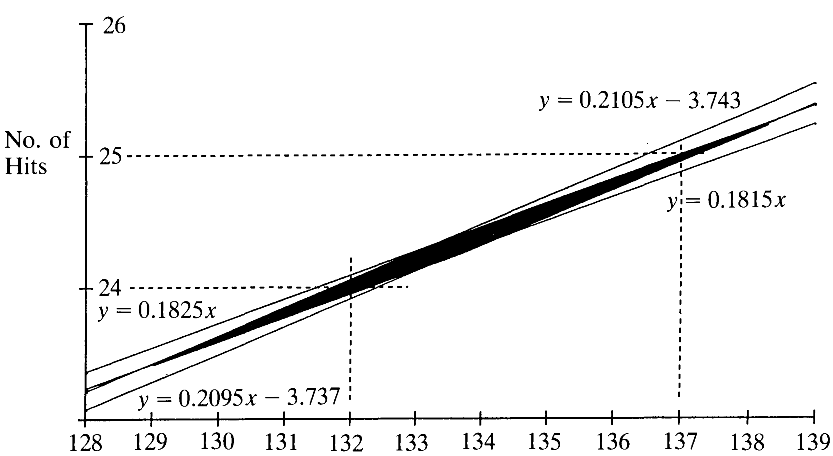

Figure 2

This is the set of all real number pairs (x, y) that satisfy (3). It is most remarkable and fortuitous that there are two points with integer coordinates lying in this region.

Although this example is probably too difficult for a precalculus class, it is quite useful for a numerical analysis or linear algebra course. It points out that a little round-off error can strongly affect the solution to a problem, when the solution procedure involves subtraction of almost equal quantities. It also provides a concrete example of an ill-conditioned linear system in which an input error of less than 0.3% yielded an output error over 10 times as large. Moreover, since the solutions are integers, students are forced to think about whether their solutions make sense in the context of the problem. Most of my students find only one of the solutions unless they are told that there are two possibilities.

This problem illustrates both the usefulness and the difficulty of interval analysis. Ramon E. Moore [**2**] discusses the solution of the linear system $Ax = b$ and distinguishes two cases: the first in which the coefficients of A and b are exactly representable by machine numbers, and the second in which the coefficients are only known to lie in certain intervals. He points out that the second case is "much more difficult" and that "the exact set of solutions... may be a complicated set." Moore cites a 2×2 example [**1**] for which the exact set of solutions is a nonconvex, eight-sided polygon. My example is somewhat easier to use in class, since the solution set is a convex polygon.

References

1. Elrod Hansen, On the solution of linear algebraic equations with interval coefficients, *Linear Algebra and Applications* 2 (1969) 153–165.
2. Ramon E. Moore, *Methods and Applications of Interval Analysis*, SIAM, Philadelphia, 1979.
3. Edward Rozema, Why do we pivot in Gaussian elimination? *College Mathematics Journal* 19 (1988) 63–72.

On the Distance from a Point to a Curve

Mark Schwartz, Ohio Wesleyan University, Delaware, OH 43015

Let C be a smooth curve (in the x-y plane) parametrized by $\mathbf{r}(t)$ and let q be a point not on C. Assume q is the terminal point of the vector $\mathbf{q}$ in standard position, and let $f(t) = |\mathbf{r}(t) - \mathbf{q}|$, the distance from the terminal point P of $\mathbf{r}(t)$ on C to q. In this note, we shall determine the local extrema of f. The critical points of f occur at the values of t for which $(\mathbf{r}(t) - \mathbf{q}) \cdot \mathbf{r}'(t) = 0$. At such a point, if q is on the convex side of C (or C is a line) then f has a local minimum. Of greater interest is the case when q is on the concave side of C; the result furnishes a nice application of the curvature and evolute of a curve. This case, depicted in Figure 1, is assumed in the following result.

Theorem. *Let t_0 be a critical point for the distance function f, and let P_0 be the corresponding point on C. Then f has a local minimum (maximum) at t_0 if the distance from q to P_0 is less (greater) than the radius of curvature at P_0.*

To establish the theorem, we start by recalling some formulas from vector calculus. A recommended reference is *Calculus with Analytic Geometry* by G. F.