Dynamic Function Visualization

Mark Bridger

The College Mathematics Journal, Vol. 27, No. 5. (Nov., 1996), pp. 361-369.

Dynamic Function Visualization

Mark Bridger

Mark Bridger (bridger@neu.edu) received his B.A. (1963) from Columbia University and his M.A. and Ph.D. (1967) from Brandeis University, where he was a student of Maurice Auslander. His doctoral thesis was in commutative algebra, and he has published articles in that field. He is an associate professor of mathematics at Northeastern University (Boston) where he is currently working on constructive analysis, issues in the philosophy of science, and applications of technology to mathematics education. His other interests include music, bicycling, and gardening.

The problem first appears in my multivariable calculus course. As soon as the discussion moves to functions $\mathbb{R}^1 \to \mathbb{R}^3$ and $\mathbb{R}^2 \to \mathbb{R}^3$, a slightly puzzled look comes over the faces of many students. When I refer to the "image" of such a function and try to draw it on the board, I can see that all is not well. Finally, a student may ask: "Is this, like, the *graph*?" Once again it becomes clear that most students do not know the difference between a function and its graph, or that a function can be pictured as a mapping of points.

The problem is not simply that students have never seen these distinctions made—though for many this is the case. Even after examples are given and definitions presented, the confusion persists, sometimes through the entire course. The problem, which has been studied fairly extensively [**2**], results from a confused or imperfect conception of *function*. I believe that the difficulty my students have results from their identification of a function with its graph. Since this identification presents a function as a single completed entity, I call it the *static* view of functions.

There is an alternative to this perspective, in which one sees a function as making an association: $x \to f(x)$ for each allowable x. Such a *dynamic* view of functions captures what we really mean by a function, the way functions are actually constructed, and the uses to which they are put.

In this article, I show how mapping diagrams (using software to implement and animate them) can foster the dynamic perspective. This method of visualizing functions leads to its own collection of conjectures and theorems which, like those suggested by classical cartesian "graphs," are rich and varied in difficulty and complexity.

Mapping Diagrams

A mapping diagram is formed from two parallel lines, one containing the domain of a function f, the other containing the range. We shall take these to be vertical lines in the xy-plane, with the domain lying in the line $x = 0$ and the range in the line $x = 1$. Function evaluation is represented by drawing a line from a number s on the domain line to its image $f(s)$ on the range line. We call this the *mapping line at* s, and denote it $\mathcal{L}_s$. Since $\mathcal{L}_s$ passes through the points $(0, s)$ and $(1, f(s))$, we have the equation of $\mathcal{L}_s$:

$$y = \big(f(s) - s\big)x + s. \tag{1}$$

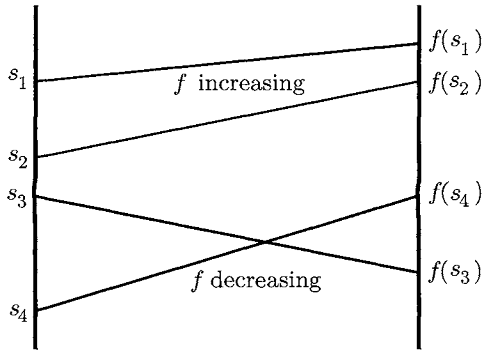

Figure 1

Figure 1 shows some mapping lines for a function f. Note that where the function is *decreasing*, say between domain points s_3 and s_4, the mapping lines *must cross* somewhere between the domain and range lines, since $s_4 < s_3 \Rightarrow f(s_4) > f(s_3)$.

Some properties of mapping diagrams have been described in Richmond [**6**] and Brieske [**1**]. More recently, Goldenberg et al. [**4**] implemented a variation of this idea on a computer, to examine how students can visualize the dependence of one variable on another. These authors all used parallel *horizontal* lines to represent the domain and range. Subsequent testing on students indicates that drawing arrows from left to right between two vertical lines is a more natural setup. Flashman [**3**] makes extensive use of this kind of mapping diagram in his forthcoming calculus book. Hubert Hohn and I recently developed a software program, the *Function Visualizer*, which not only produces mapping diagrams but animates them. The *Visualizer* is described in more detail in the Software Review column of this issue. Here we will describe how the *Visualizer* can provide insights into functional behavior.

Mapping Diagrams for Affine Functions

Among the simplest functions are the *affine* ones, whose form is $s \mapsto as + b$. Such functions are often called "linear" due to the appearance of their graphs; I do not use this terminology because it can confuse students later in linear algebra. With the *Function Visualizer*, students can discover the characteristic properties of the mapping diagrams for these functions by exploring several examples.

The first thing that students usually observe is that affine functions send intervals of equal length in the domain to intervals of equal length in the range. That is, if $|s_2 - s_1| = |s_4 - s_3|$, then

$$|A(s_2) - A(s_1| = |a|\,|s_2 - s_1| = |a|\,|s_4 - s_3| = |A(s_4) - A(s_3)|.$$

Students observe geometrically from the *Visualizer* and algebraically from this formula that the coefficient a acts as a scaling factor for lengths. More about this soon.

Another fact that surfaces after experimenting with the mapping diagram for an affine function is that it always has a *focal point*. That is, all of the mapping lines meet at a single point; see Figure 2. Students can, and should, prove this, using the equations of the mapping lines given by (1). For the affine function $A : s \mapsto as + b$, the focal point is given by

$$F = \left(\frac{1}{1-a}, \frac{b}{1-a} \right). \qquad (2)$$

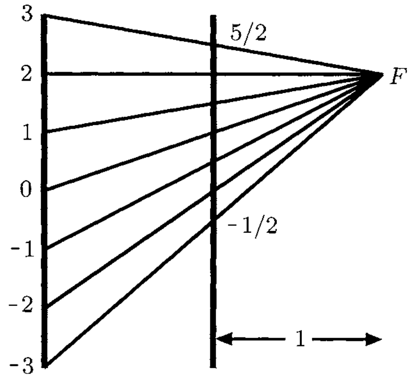

Figure 2a. $s \mapsto \frac{1}{2}s + 1$

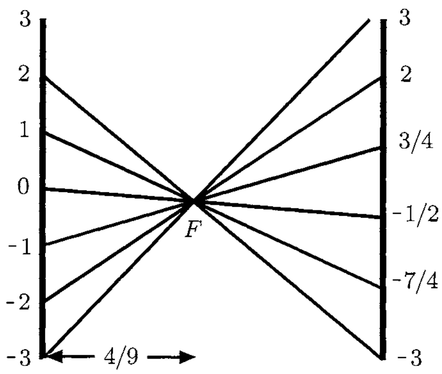

Figure 2b. $s \mapsto -\frac{5}{4}s - \frac{1}{2}$

This formula, of course, requires that $a \neq 1$. When $a = 1$ the mapping lines are all parallel and the focal point is "the point at infinity."

Of particular interest is the horizontal mapping line through the focal point. This is the mapping line $\mathcal{L}_c$ for the unique fixed point $c = b/(1-a)$ of the affine function A. The point c is called the *center* of A [**5**]. The formula for the affine function can be rewritten in "slope-center" form as $A : s \mapsto a(s-c) + c$.

As in photography, the position of the focal point determines whether the image is larger or smaller than its source. Since scaling is also determined by the coefficient a, students may derive, both geometrically and analytically, the following relationships. If $0 < a < 1$ the focal point F is to the right of the range line, and the affine map is a contraction (shrinking). Values of a greater than 1 yield focal points to the left of the domain line and maps that are dilations (magnifications) by the factor a. When $a = 1$ the focal point is at infinity, the mapping lines are parallel, and the affine map is a simple translation. Negative values of a lead to focal points between the domain and range lines. These are decreasing functions or *orientation-reversing* maps: they flip intervals upside down and scale them by a factor of $|a|$. For example, $a = -1$ gives a focal point at $x = 1/2$, halfway between domain and range. Focal points closer to the domain are dilations; those closer to the range are contractions. Finally, $a = 0$ puts the focal point *on* the range line, and the map collapses the entire domain line to the focal point (constant map).

The following theorem, which characterizes affine functions completely by their focal points, is easily proved using elementary analytic geometry.

Theorem. *A function is affine if and only if its mapping diagram has a focal point. When this is the case, the function is given by*

$$s \mapsto \frac{F_x - 1}{F_x}s + \frac{F_y}{F_x}$$

for finite focal points (F_x, F_y). *If the focal point is at infinity, the function is just a translation.*

In Figure 2 we see two examples. The diagram for the function $s \mapsto \frac{1}{2}s + 1$ (Figure 2a) has the focal point $(2, 2)$. Since a is positive with magnitude less than 1, the focal point lies to the right of the range line. In Figure 2b, the diagram for $s \mapsto -\frac{5}{4}s - \frac{1}{2}$, a is negative, so the focal point lies between the domain and range lines. Since the magnitude of a is greater than 1, the focal point lies closer to the domain line than to the range line. Here we have an orientation-reversing magnification.

Differentiable Functions

Since differentiable functions are characterized by being locally affine, their mapping diagrams, after repeated zooming, resemble those of their local affine approximations. That is, f is differentiable at s_0 if and only if

$$\lim_{s\to s_0} \frac{f(s)-[f(s_0)+f'(s_0)(s-s_0)]}{s-s_0}=0. \tag{3}$$

This means that the affine function $A_{s_0} : s \mapsto f(s_0)+f'(s_0)(s-s_0)$ is a very good approximation to f near s_0. According to formula (2), the focal point of this local affine approximation is

$$F_{s_0}=\left(\frac{1}{1-f'(s_0)}, \frac{f(s_0)-s_0 f'(s_0)}{1-f'(s_0)}\right). \tag{4}$$

Just as a close-up view of the graph of f near a point $(x_0, f(x_0))$ looks like a line with slope $f'(x_0)$, a close-up view of the mapping diagram of f near the domain point s_0 looks like the mapping diagram of the affine function A_{s_0}. Thus, the derivative $f'(s_0)$ can be interpreted as the *factor by which the mapping magnifies or shrinks (oriented) segments of the domain line* in the mapping diagram for f near s_0. That is, the derivative measures the local dilation or contraction of the mapping.

By examining examples using the *Visualizer* software, students can develop a clear mental picture of this mapping interpretation of functions and their derivatives, and this picture will make the later study of multivariable calculus, linear algebra, and complex analysis seem more natural and intelligible.

Animated Mapping Diagrams

The *Visualizer* animates and colors the mapping diagrams to enhance the dynamic view they give of functions as mappings. In animation mode, each point $P_s = (0, s)$ on the domain interval moves toward its image $(1, f(s))$ on the range line, along the line $\mathcal{L}_s$ of the mapping diagram. But we adjust the speeds according to the magnitude of the derivative $|f'(s)|$, so that P_s moves with horizontal speed $k + K|f'(s)|$. That is, we choose a positive number k as the *minimal horizontal speed*; points P_s for which $f'(s) = 0$ move at this speed. Points s in the domain interval for which $|f'(s)|$ is maximal will move at horizontal speed $k + K$ along their mapping lines. Thus, the parameter K represents the *spread between the minimal and maximal speeds.* If we set $v(s) = k + K|f'(s)|$, then the horizontal coordinate of P_s at time t is simply $v(s)t$. Putting this in (1), the equation of $\mathcal{L}_s$, we get the vertical coordinate. So the location of P_s at time t in its journey along $\mathcal{L}_s$ toward its image $(1, f(s))$ is given by

$$\big(x(t), y(t)\big) = \Phi(s,t) = \big(v(s)t, s + [f(s) - s]\, v(s)t\big). \tag{5}$$

(*Visualizer* stops the motion of each point when it reaches its destination on the range line.)

Thus, the domain interval moves as a wave front from left to right, as in Figure 3. Its position, frozen at time t, is given by the parametrized curve

$$\phi_t(s) = \Phi(s,t).$$

For additional clarity, points P_s are color-coded, depending on whether the derivative $f'(s)$ is positive or negative.

The pictures produced by this scheme are striking and show characteristic features that one learns by experience to recognize and interpret. A verbal picture can hardly be adequate, but I will try to describe, with "snapshot" diagrams, some of the things we have observed.

Moving walls. An affine function is easily recognized since the derivative is constant; all points move at the same speed. One observes a vertical wave front moving from left to right. If domain and range have the same scale, then the front stretches or contracts vertically depending on whether the derivative is greater or less than 1 in absolute value. When the derivative is negative, the wave front actually flips over vertically, since lower points must move up and higher points move down. This happens at the focal point, which lies between the domain and range lines. The inversion is invisible, however, because the front is vertical. Only with non-affine functions does this flipping become apparent and striking.

Bulges. An interval where the function increases monotonically, with nonconstant derivative, will move rightward, but not as a vertical wave front. Suppose that s, on the domain line, is such that $f'(s)$ is larger than the value of f' at points below s on the domain. Then P_s will move rightward faster than points below it. Thus, the wave front will lag for points below s. If, in fact, $f'(s_0)$ is a local maximum for f', then P_{s_0} will lead points both above and below it: the moving wave front forms a bulge whose leading or rightmost point is P_{s_0}. In Figure 3a, look at the sequence of snapshots of the motion of the domain interval $[-1, 1]$, using the function $s \mapsto (12s - s^3)/13$. (The divisor 13 gives the domain and range roughly the same scale.) The midpoint of the domain, $s_0 = 0$, produces a local maximum for the derivative $(12 - 3s^2)/13$; consequently, P_0 moves to the right faster than points above or below it. The snapshots show P_0 increasing its lead over its neighboring points and arriving first at the finish (range) line.

What if $f'(s_0)$ represents a local *minimum* of the derivative? The moving wave front will lag at P_{s_0}, causing a bulge pointing leftward. The snapshots in Figure 3b for the function $s \mapsto (12s + s^3)/13$ show this behavior. In general, if the derivative is nonconstant, bulges occur with tips at places where the derivative has an extreme value (i.e., where f'' changes sign).

Bowties. The mapping diagram for a *decreasing* function is more interesting since the mapping lines cross each other between the domain and range lines, as we saw

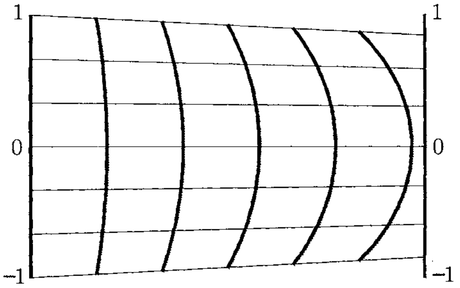

Figure 3a. $s \mapsto (12s - s^3)/13$

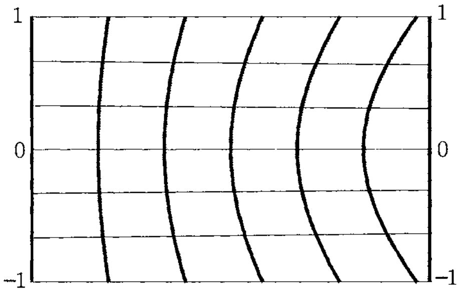

Figure 3b. $s \mapsto (12s + s^3)/13$

in Figure 1. Thus, the paths of some points will cross the paths of others: smaller values of s yield larger values of $f(s)$. However, the wave front $\phi_t(s)$ does not necessarily cross itself! In fact, if the derivative is also a monotone function, we know that the wave front will not cross itself; for instance, with an increasing derivative, points starting higher up on the domain line move at strictly greater speeds than ones starting lower down, so they pass through the crossings of the mapping lines sooner.

If, however, the derivative has a maximum or minimum, then the wave front will cross itself. Consider the wave front snapshots for the function $s \mapsto (-12s + s^3)/13$ in Figure 4a. The derivative has a maximum at 0, so P_0 moves most quickly. There are points a, b such that $a < 0 < b$ and $f'(a) = f'(b)$ (in this easy example we can take them to be $\pm s$ for any s). These points move at the same horizontal speed. However, because f is decreasing, P_a must move upward and P_b downward. They must meet sometime, and the wave front will cross itself there. In fact, the wave front first forms a cusp at P_0, then a loop forms, with the crossing sliding down the interval through the pairs $P_{\pm s}$. These crossings occur at the crossings of mapping lines.

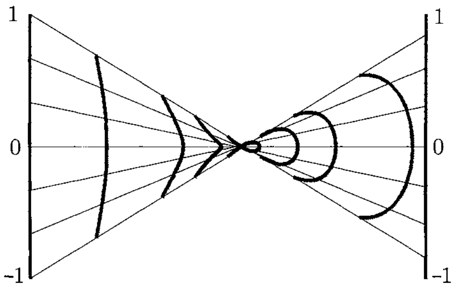

Figure 4a. $s \mapsto (-12s + s^3)/13$

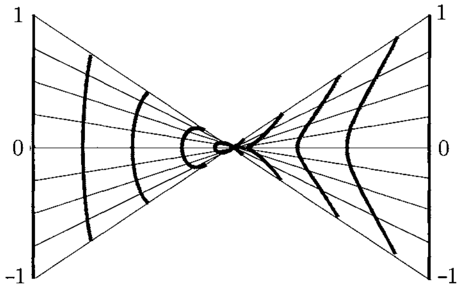

Figure 4b. $s \mapsto (-12s - s^3)/13$

In Figure 4b we can see what happens for a decreasing function whose derivative has a local *minimum*.

When the function is not monotone, the *Visualizer* color-codes intervals of increase and decrease. The animation shows vividly how the large-scale behavior can be broken down into intervals of monotonicity and how the function becomes a severalfold covering of various segments of its range. Where the function is one-to-one the *Visualizer* can reverse the motion of the points, showing how the inverse function behaves.

A Duality Between Mapping Diagrams and Graphs

In addition to providing a useful pedagogical tool for studying functions, mapping diagrams also have interesting mathematical properties. There is an intriguing duality between certain elements of mapping diagrams and cartesian graphs. Note that a mapping line is determined by a pair: $s \to f(s)$, so it corresponds to a point on the graph of f. This establishes a correspondence:

$$\text{mapping } \textit{lines} \leftrightarrow \textit{points} \text{ on the graph.}$$

So a pair of mapping lines corresponds to two points on the graph. The relation between the *point* of intersection of the mapping lines and the *secant line* on the graph is described by the following lemma.

Lemma. (T. Gaffney) *Suppose the mapping lines $\mathcal{L}_{s_0}$ and $\mathcal{L}_{s_1}$ in the diagram for f intersect at the point (x, y). Then the secant line connecting the points $(s_0, f(s_0))$ and $(s_1, f(s_1))$ on the graph of f has slope $(x-1)/x$. Conversely, for any two distinct points $(s_0, f(s_0))$ and $(s_1, f(s_1))$ on the graph of f, the mapping lines $\mathcal{L}_{s_0}$ and $\mathcal{L}_{s_1}$ intersect at a point (x, y) where*

$$\frac{x-1}{x} = \frac{f(s_1) - f(s_0)}{s_1 - s_0},$$

provided this last slope is not 1.

The proof of this fact is entirely straightforward, using (1) to find the equation of the mapping lines and then solving for their intersection. We see that

$$x = \frac{s_1 - s_0}{(s_1 - s_0) - [f(s_1) - f(s_0)]} = \frac{\Delta s}{\Delta s - \Delta f}$$

from which we get $\Delta f / \Delta s = (x-1)/x$. Note that if $\mathcal{L}_{s_0}$ and $\mathcal{L}_{s_1}$ are parallel, so they "meet at infinity," the slope of the corresponding secant line to the graph is $\lim_{x \to \infty} (x-1)/x = 1$.

This establishes a further correspondence:

points of intersection of mapping lines ↔ secant *lines* on the graph.

Now, of course, *tangent* lines are limiting cases of secant lines, so if we fix s_0 and let $s_1 \to s_0$, the secant lines approach the tangent to the graph at $(s_0, f(s_0))$. What happens to the intersecting mapping lines $\mathcal{L}_{s_1}$ and $\mathcal{L}_{s_0}$ as $s_1 \to s_0$ (see Figure 5)?

Since the abscissa of the point of intersection of the mapping lines is given by $x = \Delta s / (\Delta s - \Delta f)$, we see that $x \to 1/[1 - f'(s_0)]$ as $s_1 \to s_0$. (The ordinate y, of course, is just the point on either mapping line determined by x.) But we have seen this before. It is exactly the focal point of the best affine approximation at s_0, given in equation (4) above. Thus, if we let $s_1 \to s_0$, then the intersection of $\mathcal{L}_{s_1}$ and $\mathcal{L}_{s_0}$ approaches the focal point of the best affine approximation to f at s_0.

There is one more piece to this interesting puzzle. Anyone who has looked at "string art" has noticed that curves appear for which the strings form families of "tangents." The curve is called the *envelope* of its family of tangent lines. Similar envelopes

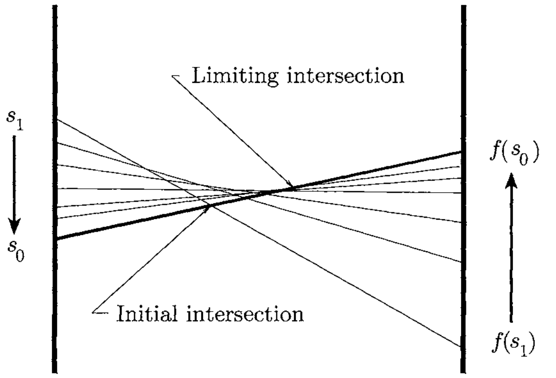

Figure 5

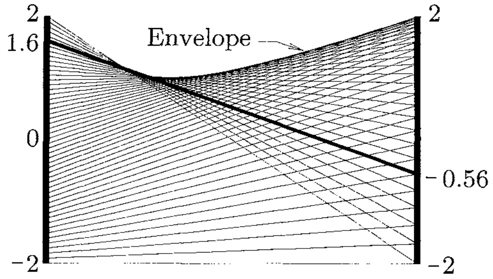

Figure 6. Mapping diagram for $s \mapsto 2 - s^2$ and its envelope.

occur in mapping diagrams. The envelope for a mapping diagram is precisely the set of limiting points of intersections of mapping lines we have just constructed. See the top of Figure 6 for an example of such an envelope. Combining this observation with our characterization of the focal points of the best affine approximation to f, we obtain the following proposition.

Proposition. *The envelope of the mapping diagram for f is exactly the set of focal points for the best affine approximations to f at the points s in the domain of f. This set forms a curve parametrized by*

$$s \mapsto \left(\frac{1}{1 - f'(s)}, \frac{f(s) - sf'(s)}{1 - f'(s)} \right).$$

Note that in the parametrization of the envelope there is no point for any parameter value making $f'(s) = 1$. Near such a value of s the mapping lines are almost parallel (no dilation or contraction), so there can be neither a point on the envelope—a limit of intersecting secant lines—nor a focal point.

This gives us the last part of our correspondence:

points on the envelope ↔ tangent *lines* to the graph.

The *Visualizer* can display only those parts of the mapping lines that lie between the domain line ($x = 0$) and the range line ($x = 1$). Consequently, the intersection of mapping lines will be visible only for *decreasing* functions. This is not a serious limitation; we need only replace f by $-f$ if necessary.

Figure 6 shows an example for the function $s \mapsto 2 - s^2$. The mapping line $\mathcal{L}_{1.6}$ has been darkened (the *Visualizer* colors it). It is tangent to the envelope at a point whose abscissa can easily be estimated (using a mouse-sensitive coordinate display) to be $x \approx 2.4$. The expression $(x - 1)/x$ gives an estimate for the derivative of -3.167, which is fairly close to the correct value of -3.2. Since $f'(s) = -2s$, the parametrization of the envelope becomes

$$s \mapsto \left(\underbrace{\frac{1}{1+2s}}_{x}, \underbrace{\frac{s^2}{1+2s}}_{y} \right).$$

Here it is possible to eliminate the parameter s, which yields the explicit equation for the envelope: $y = (1/x - 2 - x)/4$ (a hyperbola).

To conclude, the *Function Visualizer* presents new and illuminating pictures of functions considered not as static graphs but as animated mappings of points. As the above examples show, this picture is rich enough to provide insight at levels of mathematical sophistication ranging from high school algebra through calculus. Like the traditional graph, the *Visualizer* mappings can provide an intuitive and geometric language for describing many aspects of functional behavior.

Acknowledgments. Hubert Hohn has been my collaborator in the construction of the *Function Visualizer.* His programming skills, artistic sensibility, and pedagogical judgment converted a bare-bones idea into a useful mathematical tool. I appreciate also the help of T. Gaffney and R. Porter in the development of the *Visualizer* and the preparation of this article.

Work on the *Visualizer* was mostly supported by the National Science Foundation. As all works developed under public support should be, the *Visualizer* is in the public domain for all nonprofit purposes. Copies for the Mac and DOS environments can be downloaded via ftp from www.math.neu.edu. I welcome comments and suggestions concerning its use.

References

1. T. Brieske, Mapping diagrams, continuous functions and derivatives, *College Mathematics Journal* 9 (1978) 67–72. (Reprinted in *A Century of Calculus,* part 2, Mathematical Association of America, Washington, DC, 1992, 113–118.)
2. E. Dubinsky and G. Harel, eds., *The Concept of Function* (MAA Notes, vol. 5), Mathematical Association of America, Washington, DC, 1992.
3. M. Flashman, *Sensible Calculus*, in press.
4. P. Goldenberg, P. Lewis, and J. O'Keefe, Representation and the development of a process understanding of function, in E. Dubinsky and G. Harel, eds., *The Concept of Function* (MAA Notes, vol. 5), Mathematical Association of America, Washington, DC, 1992, 235–260.
5. J. E. Graver, A geometric approach to linear functions, *College Mathematics Journal* 26:5 (1995) 389–394.
6. D. E. Richmond, Calculus—A new look, *American Mathematical Monthly* 70 (1963) 415–423. (Reprinted in *A Century of Calculus*, part 1, Mathematical Association of America, Washington, DC, 1992, 84–92.)

———— o ————

Balm of Gilead

In our greatly disturbed times, when aggression and personal ambition are the basic modes of life, a healthy antidote can be found by plunging into the cool waters of the intellect, observing how in every period of history the human genius has tried to come to terms with that objective world that surrounds us, looking at it "sub specie aeternitatis", unhampered by any personal desires.

Cornelius Lanczos, *Space Through the Ages,*
Academic Press, London, 1970, page 309.