Strang's Strange Figures

Norman Richert

Pictures are playing an increasingly important role both in mathematical research and in the teaching of mathematics. Consider the current interest in fractals and in computer programs such as *Mathematica*. Textbooks, particularly at the beginning undergraduate level, need to provide more hooks into this world of pictures. A browse through current calculus texts reveals many computer generated images. In most cases these are not really "interesting" pictures, but only pictures now generated by computer that were formerly created by skilled artists—for example, surfaces. But the revolution in graphics is not the ability to draw pictures that once were very difficult to draw by hand, but rather the ability to draw pictures that were effectively *impossible* to draw by hand.

Why be timid? Let us challenge the students to confront really interesting problems with pictures. For example, in third-semester calculus a battery of techniques for describing the behavior of functions of two variables is developed. The application of these techniques could be viewed as a way of answering questions about pictures. However, their application tends to be trivial. Why? Partly because of the traditional emphasis of quadric surfaces—which happen to be easy to sketch.

A pair of interesting pictures is presented by Gilbert Strang on the cover of his new calculus book [**10**]. Professor Strang presented these plots during the panel discussion, "Calculus for the Twenty-First Century," at the 1990 AMS/MAA meeting in Louisville. They were created by Doug Hardin of Vanderbilt University and they are easy to define yet impossible to draw by hand (no one has the patience). They present mysterious behavior and their "solution" is not really calculus, at least not traditional calculus. Perhaps a "Lean" calculus should stick to business, but part of a "Lively" calculus should be interesting problems.

Figure 1 shows the sine function plotted at integers $n = 1, \dots, 10{,}000$. Figure 2 is an enlarged piece of the same plot, with $n = 1, \dots, 1{,}000$. The first figure seems to be sinusoidal, but it is not $\sin x$. There are too many curves, and their period is wildly wrong (over 15,000). Why doesn't the second plot look more like the first? After all, it is the same function, with the x-scale enlarged.

What is seen could be passed off as the effect of discretization. Is discussion of these effects important in a calculus course? Certainly it cannot be a major part (it takes two pages in Strang's book). On the other hand, part of the philosophy of the current calculus curriculum initiatives implies breaking out of some of the old ruts about the proper content of calculus. Discussion of images generated by computers is an appealing way to implement this philosophy. This note will explore one line of explanation of the plots.

Some very interesting questions can be raised as to what it *means* to plot a function, questions that traditionally have been brushed aside, with cases such as

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is rational,} \\ 1 & \text{if } x \text{ is irrational,} \end{cases}$$

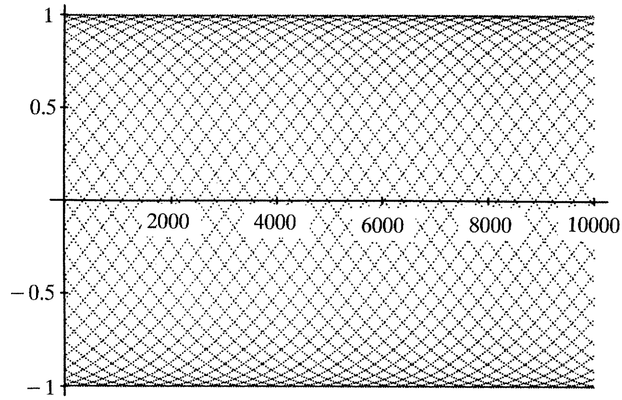

FIG. 1. 10,000 points of sin n, $n = 1, 2, 3, \ldots$. What is the explanation of the many periodic curves?

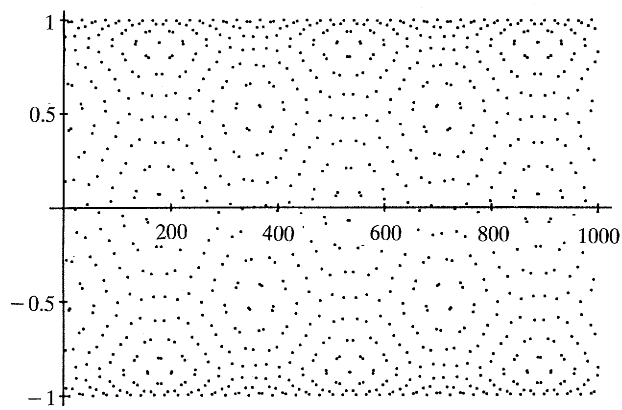

FIG. 2. 1,000 points of sin n. What happened to the periodic curves? Why the hexagonal patterns?

simply viewed as pathological. Yet that example is not so far from what would happen to Figure 2 if more points were plotted: at the scale and dot size of these plots, the graph would look completely black. Computer plotting has made hand plots of points passé in the search for information about a graph. But it has made the *meaning* of plots more pertinent, not less so. Thoughtful students have always had nagging doubts as to why we can blithely play "connect the dots" after a few measly points are plotted. Computer generated plots simply up the ante on these doubts, as these pictures nicely illustrate. This should be a powerful new incentive to ask more interesting questions about functions and their graphs. Many current programs will do a nice job with these plots.

It is not hard to see that the difference between the two figures has to do with *scale*. The current interest in fractals makes such issues topical. Computer programs have become increasingly sophisticated in dealing with scaling issues automatically. Yet scaling issues will not simply go away, as every user of graphical software knows. It seems quite appropriate to begin to discuss rescaling in the context of calculus.

How does the apparently nested family of sine curves arise in Figure 1? They are in some sense optical illusions, like the apparent spirals of seeds in the sunflower. The seeds actually grow in a single tight spiral out of the center. The adjacencies in these figures and in sunflowers defeat any mental effort to see the "real" curve. The key to Figure 1 must lie in a consideration of adjacency, or "nearness." The subinterval size of 1 is a substantial fraction of the total period of 2π, so $\sin k$ and $\sin(k+1)$ are usually quite far apart (when are they close?) in the y-direction. How small can $\sin(k+p) - \sin k$ get for p an integer? Because sine is periodic, this is directly related to how close p is to a multiple of 2π. That is, how small can $|p - q(2\pi)|$ be for p and q positive integers? Because the sine function is continuous and periodic with period 2π, for $|p - q(2\pi)|$ small enough, $|\sin(k+p) - \sin k|$ will be small, independently of k. In fact, a little extra attention paid to the derivative will show that $|\sin(k+p) - \sin k| \leq |p - q(2\pi)|$. The graph points $(k, \sin k)$ and $(k+p, \sin(k+p))$ will then be close, so the collection of points $(k+mp, \sin(k+mp))$, $m = 0, 1, 2\dots$, will appear to form a curve. For $k = 0$ this is the curve that corresponds to a rescaling of the x-axis by a factor of $|(p - q(2\pi))/p|$, namely,

$$f_0(x) = \sin\left(\frac{p - q(2\pi)}{p}x\right).$$

The whole family of curves is

$$f_k(x) = \sin((p - q(2\pi))x/p + 2\pi qk/p), \qquad k = 0, \dots, p-1.$$

But how small is "small"? This is where scale comes in. The figures at Louisville measured roughly 10×14 cm. At this scale, small would seem to be less than 1.0 mm, which is to say less than 70 (Figure 1 units) on the horizontal scale and less than 0.02 on the vertical scale. So we now have detailed specifications: find positive integers $p \leq 70$ and q so that $|p - q(2\pi)| \leq 0.02$.

Now we arrive at a question that is not really calculus, but number theory. How small can we make $|p - q(2\pi)|$ for p and q integral? A related, though weaker question is how small we can make $|p/q - 2\pi|$, a question probably more suggestive to most students. We have meandered into the area called *diophantine approximation*, a piece of which is the study of rational approximations to irrational numbers. Most students know that $22/7$ is a good approximation to π, so they all know a tiny bit of diophantine approximation. It is not implausible to suppose that $44/7$ is a good approximation to 2π. In fact, $44 - 7(2\pi) \approx 0.018$, meeting both the smallness measures estimated above. So $\sin(k+44) - \sin k$ will be relatively small for each k. In fact, a new plot of $\sin(44n)$ quickly shows that one of the family of nested curves has been identified. It is the increasing curve through the origin in Figure 1.

But there are lots of other good approximations to 2π, say $628/100 = 157/25$. Why don't they show up in the picture? We shouldn't stray too far from calculus, except to point out that there are not *lots* of other good approximations to 2π, at least not with numerators less than 70, or what comes to the same thing, denominators less than 11. (The references **[5, 6, 9]** contain further reading.) A lot can be learned with a calculator by hunting for values p and q to make $|p - q(2\pi)|$ small. In particular, $|157 - 25(2\pi)| = 0.08$, so we do worse by a factor of 3.6 horizontally and a factor of as much as 4.9 vertically. This fraction simply does not yield the strong adjacency patterns that $44/7$ does. In fact, considering all values of p and q, the next smallest value of $|p - q(2\pi)|$ is 0.009, using the fraction $333/53$. But 333 is way off the "smallness" scale in the x-direction.

Finding the Good Approximations to 2π

In the discussion of the figures, we have used particular approximations to 2π, the *best approximations*. To be precise, the best approximations to a real number x are the rationals p/q so that $|p - qx| < |p' - q'x|$ for $p', q' \in \mathbf{Z}$, $0 < q' \leq q$ and $p'/q' \neq p/q$. Clearly, given q, we can take p to be the nearest integer to qx. For x irrational, this uniquely defines a sequence $p_0/q_0, p_1/q_1, p_2/q_2, \ldots$, of best approximations to x, with $q_0 < q_1 < q_2 < \cdots$. In fact, an initial segment of the sequence can be calculated by trial and error from the definition simply by considering increasing q. The table illustrates this procedure for $x = 2\pi$.

q	1	2	3	4	5	6	7
qx	6.283	12.566	18.850	25.133	31.416	37.699	43.982
$p - qx$	0.283	−0.434	−0.150	0.133	0.416	−0.301	−0.018

Examining the table, we see that $p_0/q_0 = 6/1$, $p_1/q_1 = 19/3$, $p_2/q_2 = 25/4$, and $p_3/q_3 = 44/7$.

The *continued fraction* algorithm provides a mechanism for directly calculating these approximations, without the need for trial and error calculations. A number of recent Monthly pieces have treated continued fractions, for example [**4**, **7**]. A very beautiful older piece by L. R. Ford is [**3**]. Let $[x]$ denote the *greatest integer* of x, the largest integer not larger than x. Set $x_0 = x$ and $a_0 = [x_0]$. Then recursively calculate a pair of sequences: $\{x_k\}$ of real numbers and $\{a_k\}$ of positive integers, $k = 1, 2, 3, \ldots$, with

$$x_k = \frac{1}{x_{k-1} - a_{k-1}} \quad \text{and} \quad a_k = [x_k].$$

The a_k's are the *partial quotients*. Then the sequence of *convergents* $\{p_k/q_k\}$ is calculated recursively by

$$\begin{aligned} &p_{-1} = 1, \quad p_0 = a_0, \quad q_{-1} = 0, \quad q_0 = 1, \quad \text{and} \\ &p_k = a_k p_{k-1} + p_{k-2}, \\ &q_k = a_k q_{k-1} + q_{k-2}, \quad \text{for } k = 1, 2, 3, \ldots . \end{aligned}$$

This algorithm can easily be programmed. The convergents converge fairly rapidly to x (hence the name). This set of convergents is, with the possible exception of p_0/q_0, identical to the set of best approximations to x. For $x = 2\pi$, we have $[a_0; a_1, a_2, a_3, \ldots] = [6; 3, 1, 1, 7, 2, 146, 3, \ldots]$. Complete understanding of these figures, particularly Figures 4 and 5, requires the use of a larger set of approximating fractions, the *best approximations of the first kind*. Those which are not best approximations are *intermediate fractions*, that is, of the form

$$\frac{p}{q} = \frac{a p_{k-1} + p_{k-2}}{a q_{k-1} + q_{k-2}}, \qquad a = 1, 2, \ldots, a_k - 1.$$

Where did the 44 sine curves go in Figure 2? One answer involves the scale change. The "small" distance in Figure 1 now corresponds to a range of 7 in n values, so 44 has become "large." The scale can be restored by tilting the page and viewing from the side so as to compress x distances. Violà, the curves reappear. But what about the hexagonal pattern?

In Figure 2 we are seeing an interference pattern created by the interaction of three distinct approximations to 2π: 44/7, 25/4 and 19/3. Such patterns result when regular patterns of dots are overlaid on each other and are known as Moiré patterns [**1**, **2**, **8**]. Of interest here is the case when the patterns are *regular screens of dots*, that is, lattices of points formed by the vertices of a tessellation of the plane by regular polygons. Regular screens of dots produce these Moiré interference patterns in a limited number of ways, one of which consists of roughly hexagonal regions. This is an important issue in printing and textile manufacture. These patterns are the telltale sign that a printed photograph, as in a newspaper, has been reproduced from another printed photograph, rather than from an original. To uncover the roughly regular screens of points producing hexagonal Moiré patterns in Figure 2, we must do a bit more analysis.

As is discussed in the accompanying box, there is an infinite sequence of *best approximations* to 2π (or any irrational number). The first few terms in the sequence are 6/1, 19/3, 25/4, 44/7 and 333/53. We have seen why 333/53 is not prominent in Figure 1. It will simplify the discussion if we limit our analysis to points near the x-axis, where $|\Delta \sin x| \approx |\Delta x|$, so that we may take the $|p - q(2\pi)|$ values as exact vertical displacements. The features being discussed are most influenced by the periodicity and the period, rather than the shape of the graph. (We could even substitute a sawtooth function for $\sin x$. The sine function has the conceptual appeal that the value of its period is implicit rather than explicit.)

The previous discussion implicitly used the metric $d_{\max}(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$. Let us measure actual distances on a graph of N points. The plotting rectangle has height v and width av, with $v = 100$ mm and $a = 1.4$ on the Louisville figures. Then the distance d (in graph measurement units) between points on a member of the curve family associated with p/q is

$$d = \sqrt{\left(\frac{p - q(2\pi)}{2/v}\right)^2 + \left(\frac{p}{N/av}\right)^2} = \frac{v}{2}\sqrt{(p - q(2\pi))^2 + \left(\frac{2a}{N}\right)^2 p^2}\,.$$

In Figure 1, with $N = 10\,000$, the values associated with 44/7, 25/4 and 19/3 respectively are roughly 1.1, 6.6 and 7.5 mm. So the eye finds the 44/7 family to be a clear feature. In Figure 2 with $N = 1000$, these values become 6.2, 7.5 and 8.0, so three distances are comparable.

Consider a point $P_1(n_1, \sin n_1)$ near the x-axis. The next point to the right on the family of curves associated with 44/7 is $P_2(n_1 + 44, \sin(n_1 + 44))$. Suppose that $\sin x$ is increasing at n_1, so that $\sin(n_1 + 44) - \sin n_1 \approx 44 - 7(2\pi) > 0$. The next point to the right on the family associated with 25/4 is $P_3(n_1 + 25, \sin(n_1 + 25))$, for which $\sin(n_1 + 25) - \sin n_1 \approx 25 - 4(2\pi)) < 0$. Hence P_3 is downhill from P_1. Finally, since $44 - 25 = 19$, the $\overline{P_2P_3}$ segment is associated with the 19/3 family. Because the distances, calculated above, are roughly equal, the triangle $P_1P_2P_3$ is roughly equilateral. This forms the template of a regular screen with hexagonal symmetry, that is, six-fold rotational symmetry. The overlay of two hexagonal screens produces a Moiré pattern with hexagonal symmetry, which we see in Figure 2.

Because the Moiré pattern dominates our view of Figure 2, it is hard to see the regular screens which produce it. Separate the sine function into two functions: the function of increasing pieces, and that for decreasing pieces. Let

$$\operatorname{SinUp}(x) = \begin{cases} \sin x & \text{if } \sin x \text{ is increasing at } x, \\ 0 & \text{otherwise.} \end{cases}$$

Define SinDown similarly. If $p - q(2\pi) > 0$ then plotting SinUp instead of sine yields the increasing pieces of the corresponding family of curves. If $p - q(2\pi) < 0$ it yields the decreasing pieces.

In the case of $N = 1000$, plotting SinUp reveals the roughly regular screen of points, as Figure 3 shows. Plotting SinDown yields essentially the mirror image of this lattice across a vertical line. Plotting the sine function overlays these two lattices, and produces the hexagonal Moiré pattern of Figure 2. This can be checked directly by making a transparency of Figure 3, flipping it left to right, and overlaying it on Figure 3. The families of curves corresponding to 25/4 and 19/3 which are clear in Figure 3 can be seen in Figure 2 by tilting the page: by viewing at an angle roughly 10° away from the y-axis. The screens of SinUp and SinDown when $N = 10\,000$ are too far from regular for Moiré interference to be noticeable in Figure 1.

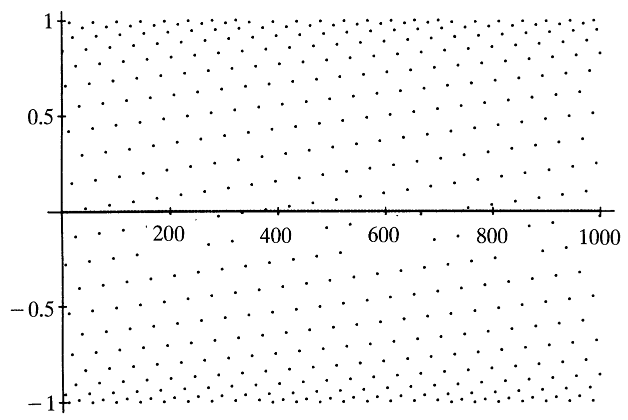

FIG. 3. 1,000 points of SinUp n. Hexagons vanish.

A good test of these observations might be to modify the sine function to yield a period different than 2π, and hence (presumably) different approximating fractions, and make some more plots. Plotting $\sin((2\pi)n/(6 + \varepsilon))$ for various ε is one way to modify the period. Setting $\varepsilon = \phi - 1 = (1 + \sqrt{5})/2 \approx 0.618034$, the fractional part of the golden ratio ϕ, yields some plots quite different than Figures 1 and 2, as Figure 4 illustrates. They are relatively unaffected by changes in scale. The continued fraction expansion of this period differs from that of 2π beginning with the first order partial quotient. In fact, the features of these figures are quite sensitive to the exact value of the period. Figure 5 illustrates what a more substantial modification of the period can produce, with period $2\pi - 5$.

The importance of computing in calculus cannot be overstated. The symbolic capabilities are forcing us to reevaluate the importance of rote techniques in differentiation and integration. Simultaneously the graphical capabilities allow us to discuss genuinely interesting graphs. I applaud Professor Strang for a step in the direction of interesting graphs.

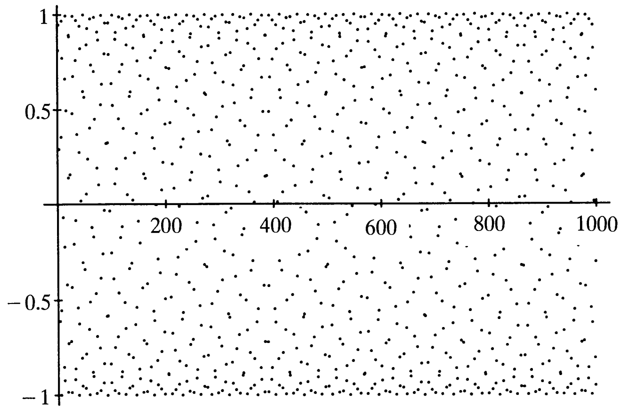

FIG. 4. No hexagons. No periodic curves. 1,000 points of $\sin(n(2\pi/(5+\phi))$.

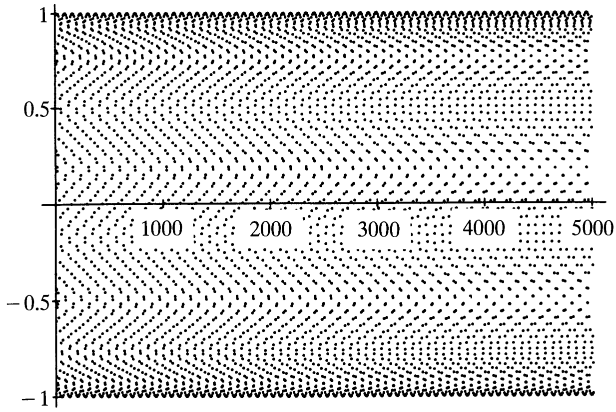

FIG. 5. A very different effect. 5,000 points of $\sin(n(2\pi/(2\pi-5))$.

REFERENCES

1. P. A. Firby, Interference in printing and in textile manufacture, *Mathematical Spectrum*, 20 (87/88) 14–17.
2. P. A. Firby, Controlling interference in graphics, *Math. Gazette*, 71 (1987) 119–125.
3. L. R. Ford, Fractions, *Amer. Math. Monthly*, 45 (1938) 586–601.
4. M. C. Irwin, Geometry of continued fractions, *Amer. Math. Monthly*, 96 (1989) 696–703.
5. A. Ya. Khinchin, *Continued Fractions*, Phoenix Books, Chicago, 1964.
6. W. J. LeVeque, *Fundamentals of Number Theory*, Addison-Wesley, Reading, Mass., 1977.
7. J. Mathews, Gear trains and continued fractions, *Amer. Math. Monthly*, 97 (1990) 505–510.
8. G. Oster and Y. Nishijima, Moiré patterns, *Scientific American*, 208 (1963) 54–63.
9. O. Perron, *Kettenbrüche*, Chelsea, New York, 1950.
10. G. Strang, *Calculus*, Wellesley-Cambridge Press, Wellesley, Mass., 1991.

Department of Mathematics
University of Houston-Clear Lake
Houston, TX 77058