MATHEMATICAL GAMES

In which "monster" curves force redefinition of the word "curve"

by Martin Gardner

When Zulus cannot smile, they frown,
To keep an arc before the eye.
Describing distances to town,
They say, "As flies the butterfly."
—JOHN UPDIKE, "Zulus Live in Land without a Square"

A fascinating aspect of the history of mathematics is the way that the definitions of names for classes of mathematical objects are continually revised. The process usually goes like this: The objects are given a name, *x*, and defined in a rough way that conforms to intuition and usage. Then someone discovers an exceptional object that meets the definition but clearly is not what everyone has in mind when he calls an object *x*. A new and more precise definition is then proposed that either includes the exceptional object or excludes it. The new definition "works" as long as no new exceptions arise. If they do, the definition has to be revised again, and the process may continue indefinitely.

If the exceptions are strongly counter to intuition, they are sometimes called monsters. The adjective pathological is often attached to them. This month we consider the word "curve," describe a few monsters that have forced redefinitions of the term and introduce a frightening new monster captured last year by William Gosper, a brilliant young computer scientist now living in Los Altos Hills, Calif. Readers of this department have met Gosper before in connection with the cellular-automata game Life. It was Gosper who constructed the "glider gun" that made it possible to "universalize" Life's cellular space.

Ancient Greek mathematicians had several definitions for curves. One was that they are the intersection of two surfaces. The conic-section curves, for instance, are generated when a cone is cut at certain angles by a plane. Another was that they are the locus of a moving point. A circle is traced by a rotating compass leg, an ellipse by a moving stylus that is stretching a closed loop of string around two pins, and so on for other curves generated by more complicated mechanisms.

Seventeenth-century analytic geometry made possible a more precise definition. Familiar curves became the diagrams of algebraic equations. Could a plane curve be defined as the locus of points on the Cartesian plane that satisfy any two-variable equation? No, because the diagrams of some equations emerge as disconnected points or lines, and no one wanted to call such diagrams a curve. Calculus suggested a way out. The word "curve" was limited to the loci of points that satisfy equations that are continuous functions.

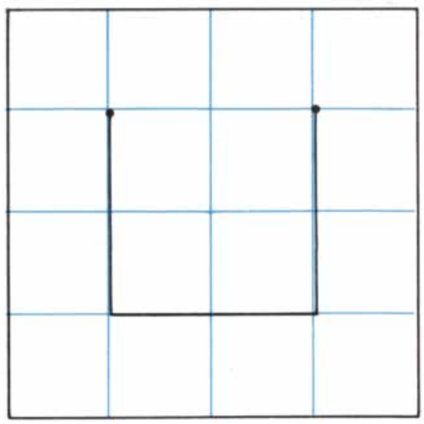
It seems intuitively obvious that if a curve diagrams a continuous function, it should be possible to differentiate the function or, what amounts to the same thing, to draw a tangent to any point on the curve. In the second half of the 19th century, however, mathematicians began to find all kinds of monster curves that had no unique tangent at any point. One of the most disturbing of such monsters was described in 1890 by the Italian mathematician and logician Giuseppe Peano. He showed how a single point, moving continuously over a square, could (in a finite time) pass at least once through every point on the square and its boundary! (Actually any such curve must go through an infinity of points more than once.) Peano's curve is a legitimate diagram of a continuous function. Yet nowhere on it can a unique tangent be drawn because at no instant can we specify the direction in which a point is moving.

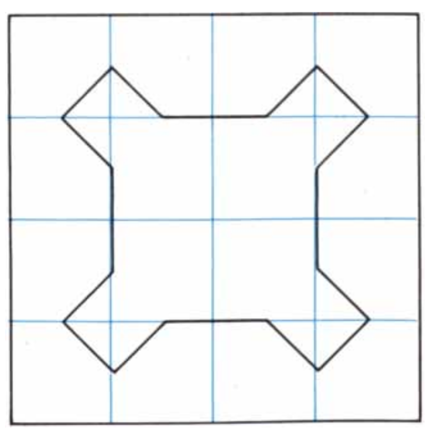
David Hilbert proposed a simple way to generate a Peano curve with two end points. The first four steps of his recursive procedure should be clear from the pictures at the top of the illustration at the right. At the limit the curve begins and ends at the square's top corners. The four steps at the bottom of the illustration show how Waclaw Sierpinski generated a closed Peano curve.

In both versions think of the successive graphs as approximations approaching the graph of the limit curve. This limit curve in each version is infinitely long and completely fills the square even though each approximation misses an uncountable infinity of points both of whose coordinates are irrational. (In general the limit of a sequence of approximation curves may go through many points that are not on any of the approximations.) Sierpinski's curve bounds an area 5/12 that of the square. Well, not exactly. The constructions approach this fraction as a limit, but the curve itself, the diagram of the limiting function, abolishes the distinction between inside and outside!

Peano curves were a profound shock to mathematicians. Their paths seem to be one-dimensional, yet at the limit they occupy a two-dimensional area. Should they be called curves? To make things worse, Peano curves can be drawn just as easily to fill cubes and hypercubes.

Helge von Koch, a Swedish mathematician, proposed in 1904 another delightful monster now called the snowflake curve. We start with an equilateral triangle and apply the simple recursive construction shown in the top illustration on page 126 to generate a crinkly





curve resembling a snowflake. At the limit it is infinite in length; indeed, the distance is infinite between any two arbitrary points on the curve! The area bounded by the curve is exactly 8/5 that of the initial triangle. Like a Peano curve, its points have no unique tangents, which means that the curve's generating function, although continuous, has no derivative.

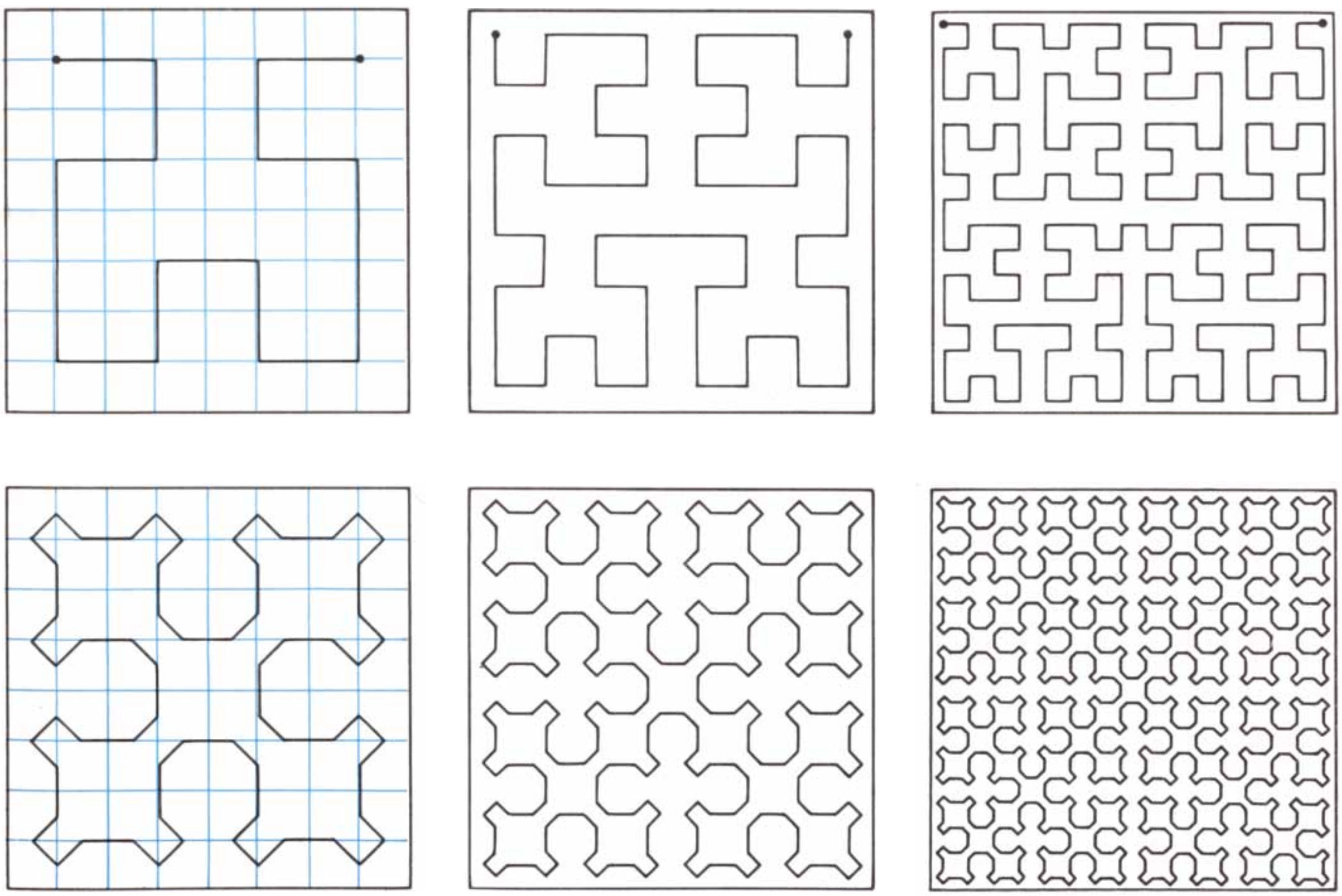
If the triangles are constructed inward instead of outward, one gets the anti-snowflake curve. Its perimeter is also infinite, and it bounds an infinity of disconnected regions with a total area equal to 2/5 that of the original triangle. One can start with regular polygons of more than three sides and erect similar polygons on the middle third of each side. A square, with the added squares projecting outward, produces the cross-stitch curve of infinite length that bounds an area twice the original square. (See my *Sixth Book of Mathematical Games from Scientific American,* Chapter 22.) If the added squares go inward, they produce the anti-cross-stitch, an infinite curve that bounds no area. Similar constructions, starting with polygons of more than four sides, produce curves that self-intersect.

A 3-space analogue of the snowflake is constructed by dividing each face of a regular tetrahedron into four equilateral triangles, erecting a smaller tetrahedron on the central triangle and continuing the procedure indefinitely. At the limit the prickly surface is infinite in area, yet it bounds a finite volume. The cube produces a similar analogue of the cross-stitch.

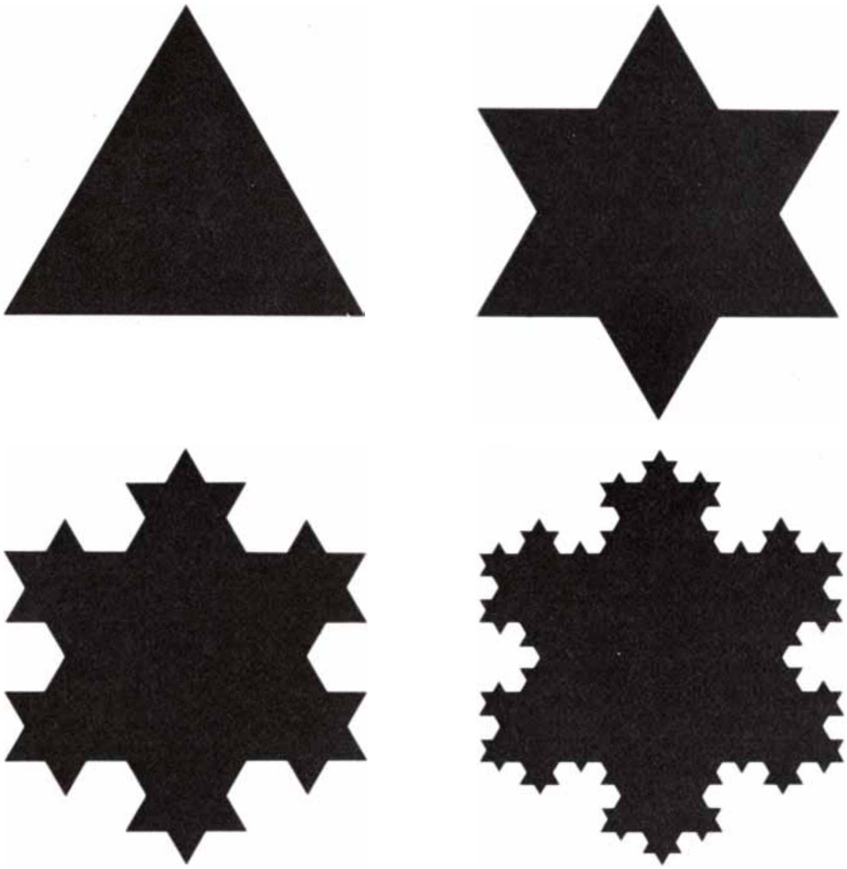
We can generalize further by dividing the sides of a regular polygon into more than three parts. For example, divide each side of an equilateral triangle into five parts, erect smaller triangles on the second and fourth sections and repeat to the limit. For an ultimate generalization begin with any closed curve that can be divided into congruent segments, then alter the segments any way you like, provided the alteration is segmented so that the change can be repeated on the smaller segments and carried to the limit. Analogous constructions can be made on the surfaces of solids. Of course, the results may be messy, self-intersecting curves or surfaces of no special interest.

A book could be written about other kinds of pathological planar monsters. The Dutch topologist L. E. J. Brouwer published in 1910 a recursive construction for cutting a region into three subregions in such an insane way that at the limit all three subregions touch at every point [see "Geometry and Intuition," by Hans Hahn; SCIENTIFIC AMERICAN, April, 1954]. Brouwer's construction generalizes to divide a region into *n* subregions, all meeting at every point. A more recently discovered family of monsters, the dragon curves, were introduced in this department in March, 1967, and were later analyzed by Chandler Davis and Donald E. Knuth in a two-part article in *Journal of Recreational Mathematics* (Vol. 3, April and July, 1970).

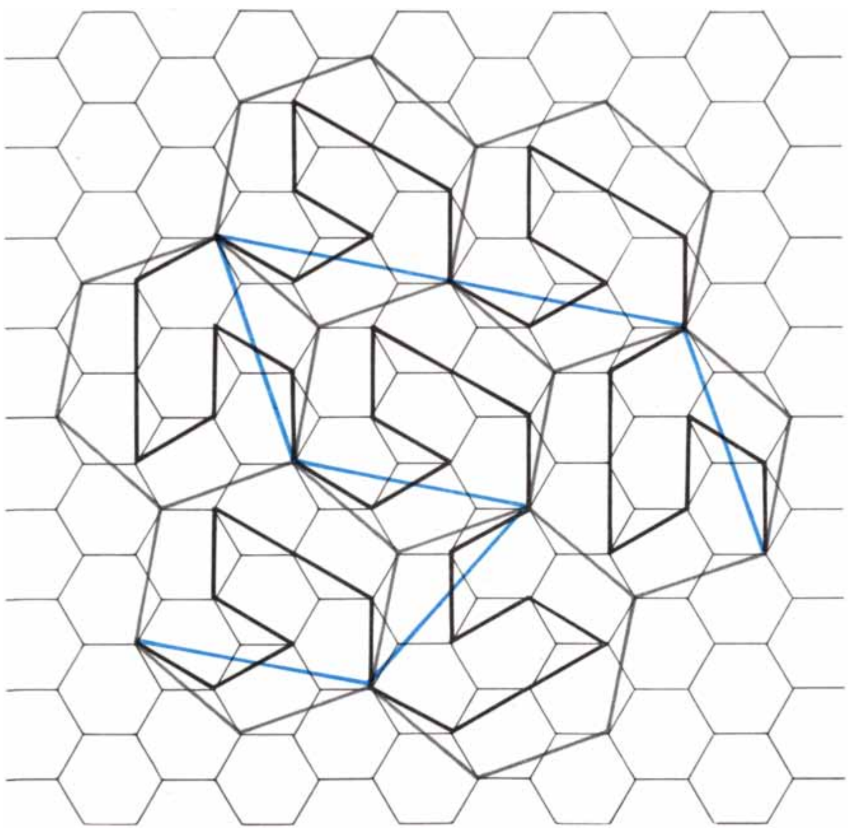
It is now my privilege to present Gosper's new monster, a beautiful space-filling curve he calls the flowsnake. Its construction starts with a pattern of seven regular hexagons [*see bottom illustration on next page*]. Eight vertexes are joined as shown by the colored line, made up of seven segments of equal length. The colored line is order 1 of the flowsnake. Order 2, shown in black, is obtained by replacing each colored segment with a similar twisted line of seven segments. Each segment of the black line is $1/\sqrt{7}$ the length of a colored segment; this



***Peano curves: David Hilbert's open one* (top) *and Waclaw Sierpinski's closed one* (bottom)**



The first four orders of Helge von Koch's snowflake



Order 1 **(color)** *and order 2* **(black)** *of William Gosper's "flowsnake"*

proportion holds at every stage of the construction.

The recursive procedure is continued to produce flowsnakes of higher orders. The illustration on the opposite page shows two computer drawings of flowsnakes of orders 3 and 4. By dividing the plane into black and white, with the bifurcating line passing through the flowsnake's end points, we see how the curve cuts the plane into two regions that twist about in almost, but not quite, the same pattern.

The curve that diagrams the limit of the successive flowsnake functions passes through every point of its region at least once, completely filling the space. The curve is infinite and nondifferentiable. Like the straight line, it is self-similar in the sense that if you enlarge any portion of it, the pattern always looks the same. Snowflake curves have the same property.

"Of course we have no physical snowflake curves," Philip Morrison has written. "Nature gives no infinities, not even within molecular collisions. There is a cutoff at the angstrom level. Still, surprises abound." By surprises Morrison means those random natural patterns that have, in a statistical sense, the property of self-similarity as successive enlargements are made. His remarks appear in a review [SCIENTIFIC AMERICAN, November, 1975] of a remarkable French book, *Les Objets Fractals: Forme, Hasard et Dimensions,* by Benoît Mandelbrot. A much-expanded version in English will be published next year.

Mandelbrot is a Polish-born French mathematician who is currently an IBM Fellow at the Thomas J. Watson Research Center at Yorktown Heights, N.Y. Like Stanislaw Ulam and many other eminent Polish mathematicians, Mandelbrot has had a career involving a marvelous mixture of creative work in both pure and applied mathematics, notably in physics and economics. His teacher, the French mathematician Paul Pierre Lévy, made the first systematic study of statistically self-similar curves, but they were regarded as useless, bizarre curiosities until Mandelbrot recognized them as being a basic tool for analyzing an enormous variety of physical phenomena.

Mandelbrot's forthcoming book is filled with pictures of just such phenomena. Consider coastlines. Their butterfly-flight irregularity is statistically self-similar. A coastline looks the same from a high altitude as it does from a low one. It is meaningless to speak of a coastline's "length" because it all depends on the precision of measurement. As Morrison puts it, "a coastline on maps at varying scales obeys a power law like the snowflake curve's, from a scale of hundreds of kilometers down to one of perhaps meters, where geography stops and pebbles begin."

The surface of the moon is another

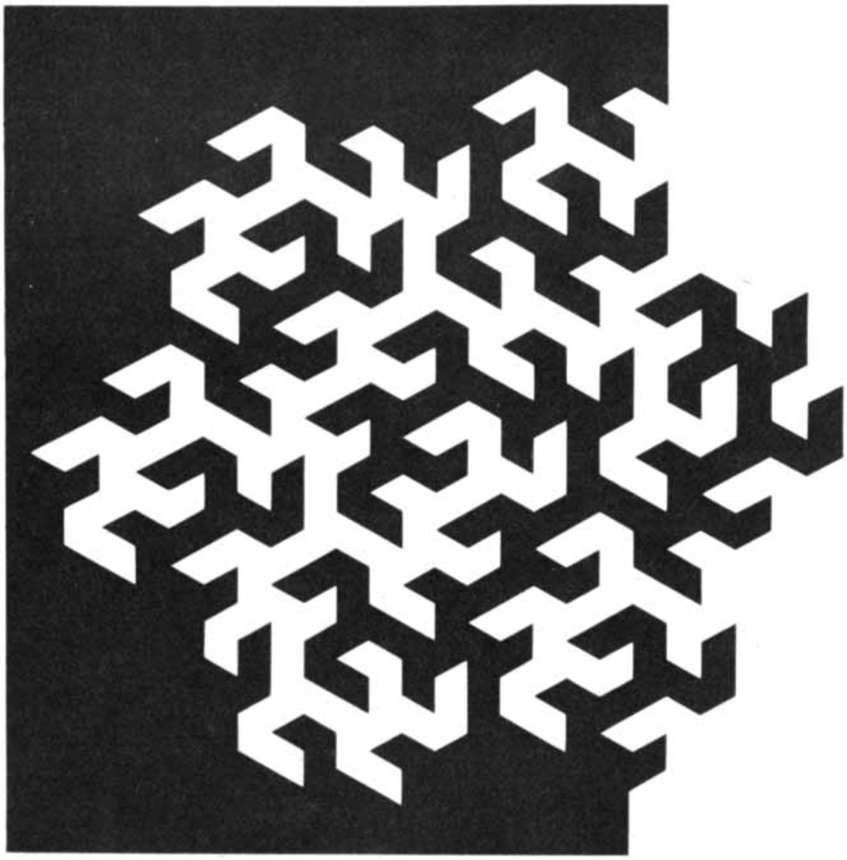
example. Remember your surprise on seeing the first closeup photographs of the moon made from a satellite in orbit around it? The moon's pocked surface looked basically as it did in photographs made with telescopes on the earth. Only the crater sizes were different. The same random self-similarity is found on the surface of certain cheeses, in the scattering of stars in the sky, in the contours of mountains, in atmospheric turbulence, in auditory noise and in countless other natural patterns. The Brownian motion of suspended particles approximates a statistically self-similar curve that (at the limit) has infinite length and no tangents.

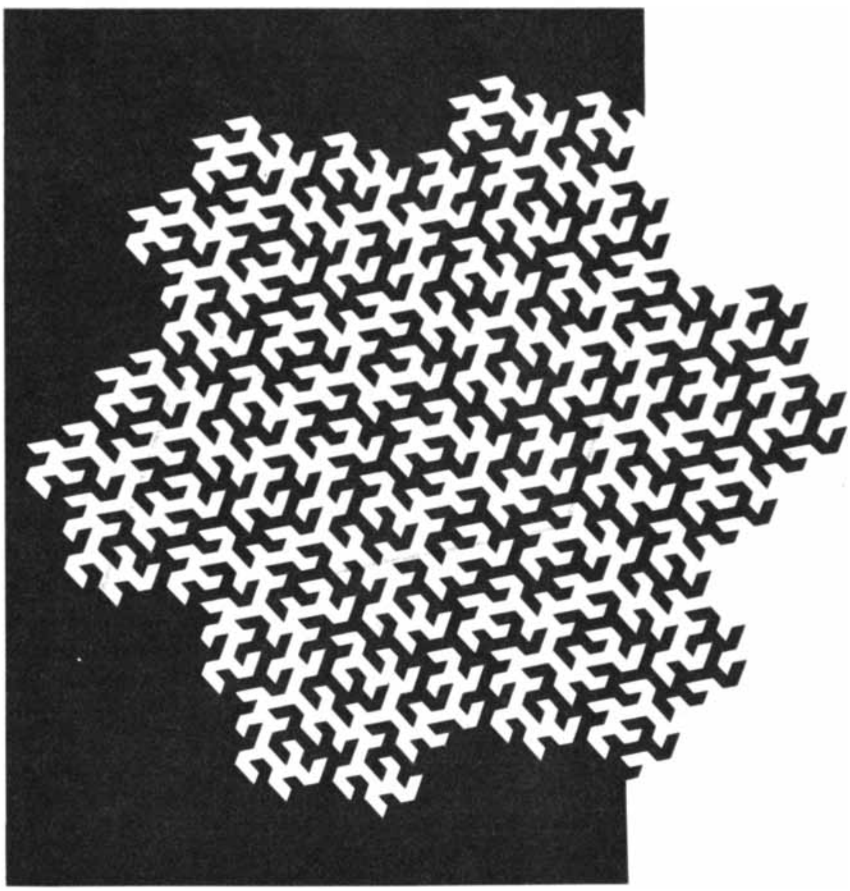
Let us go back to the flowsnake for a close look at its perimeter and at an amazing paradox. The perimeter can be constructed by a recursive procedure much simpler than the one used to get the flowsnake itself. The top illustration on page 133 shows how it works. Start with a regular hexagon, then replace each side with a zigzag line [*color*] of three equal segments, each $1/\sqrt{7}$ the original side. The result is a nonconvex 18-gon. Since the zigzag line adds the same amount of area as it takes away, the 18-gon obviously has the same area as the original hexagon. Repeat the construction on each of the 18 sides to produce a 54-gon, and imagine that the recursive procedure is continued to the limit. At each step the number of sides triples, but the area never changes. At the limit the area filled by the flowsnake is exactly the same as the area of the original hexagon.

The entire region has an astounding property. It can be dissected, as is shown in the illustration on the next page, into seven subregions, each of which is an exact copy of the entire region.

Now for the paradox. What is the ratio of the area of a subregion to the entire region? Clearly it is 1/7, since seven identical subregions make up the whole. But let us approach it from another angle, remembering that the areas of similar figures are proportional to the square of their linear dimensions. If the boundaries of three subregions are bisected, as is shown by the line *AB* in the illustration, the six segments exactly fit the perimeter of the entire region. Clearly the boundary of a subregion must be enlarged by a linear factor of 3 to fit the boundary of the entire region. But if this is true, the areas must be in a ratio of $(1/3)^2 = 1/9$. We seem to have proved that the ratio of the areas is both 1/7 and 1/9. As Gosper asked when he first sent the paradox, *Vas ist los?*

The answer lies in the peculiar, counterintuitive character of the pathological boundary. There is no fuzziness about the area of the region it bounds. It is indeed seven times the area of a subregion. The boundary is not so well behaved. It is true that the boundary of a subregion is exactly similar to the over-





***Flowsnakes of order 3* (top) *and order 4* (bottom)**

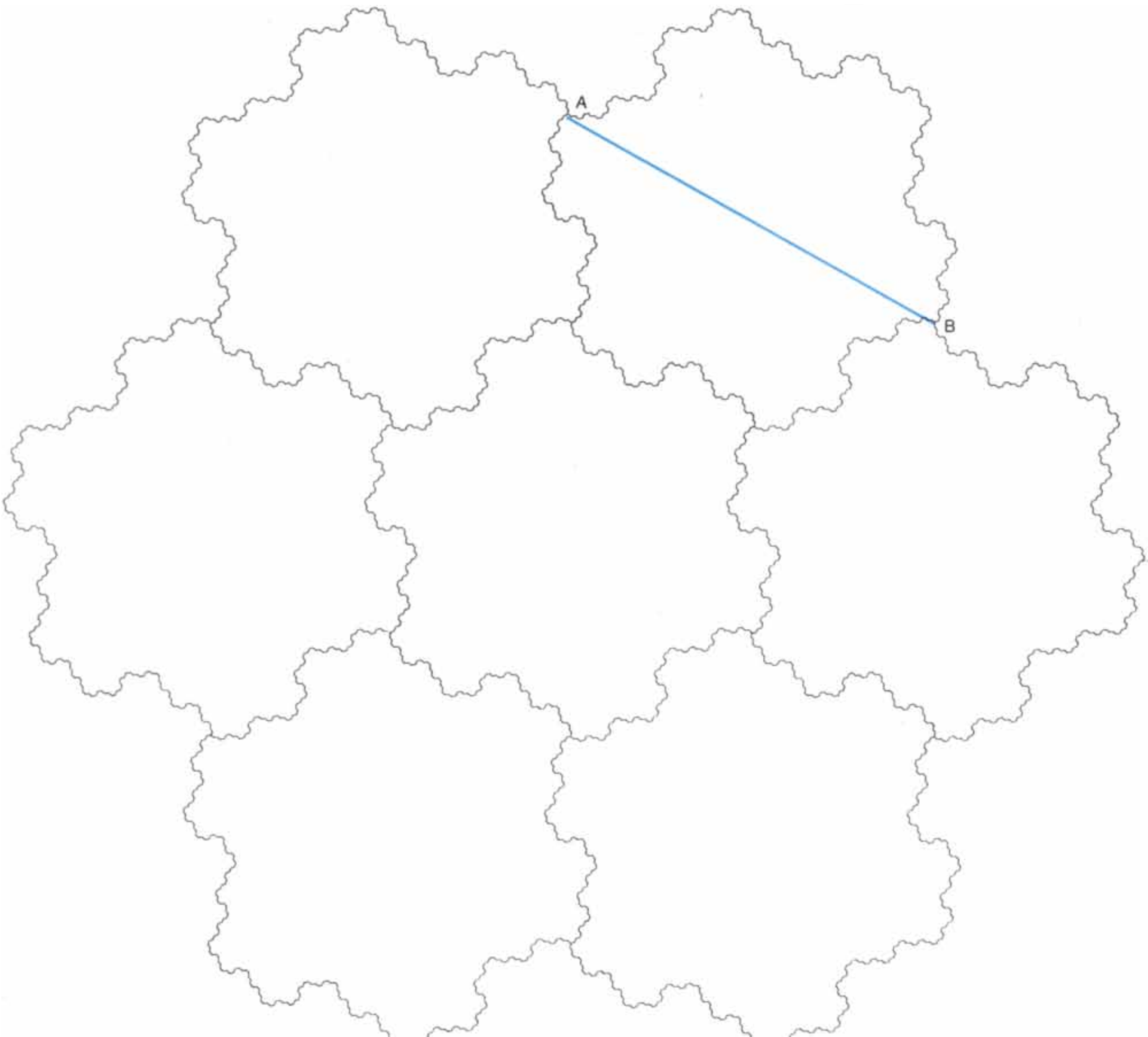
all boundary, but if the two are to be made congruent, the subregion must be magnified by a linear factor of $\sqrt{7} = 2.645\ldots$, not by a factor of 3 as it would appear.

A deep question now arises. What "dimension" should be assigned to the flowsnake's boundary? Like the snowflake, it lies in a strange twilight zone between one dimension and two dimensions. In 1919 a German point-set topologist, Felix Hausdorff, resolved the difficulty by giving fractional dimensions to such curves, or what Mandelbrot in his book calls "fractal" dimensions. They should not be confused with what are known as Hausdorff spaces, which are topological structures that mercifully we do not have to go into here.

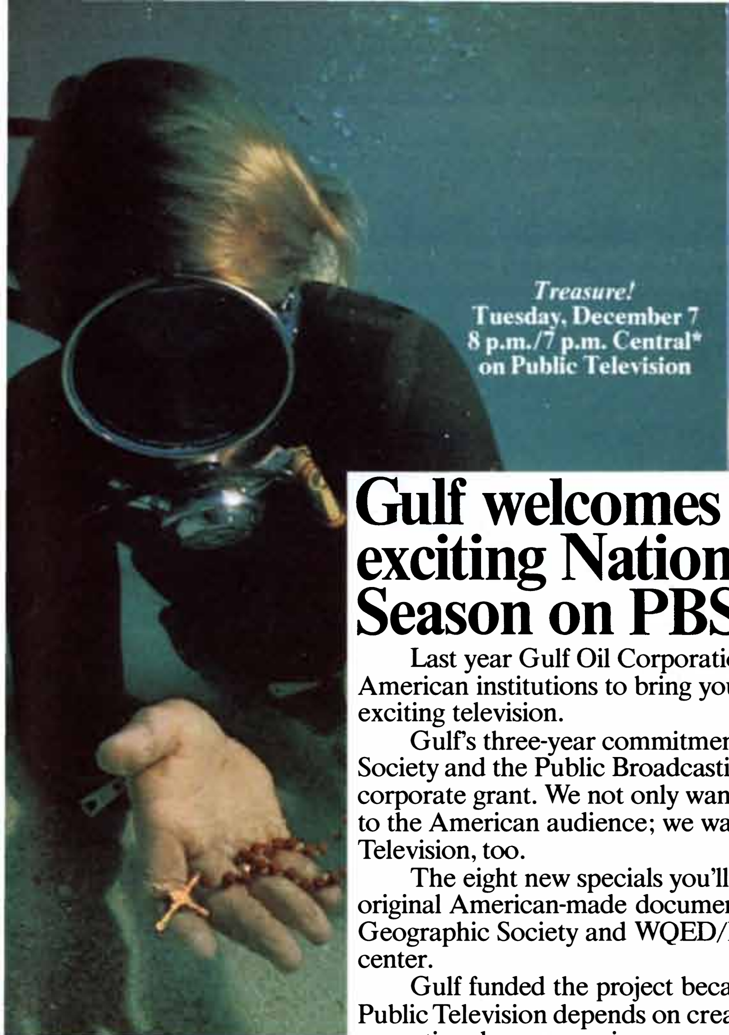
To grasp how fractal dimensions are calculated consider first a line segment. If we magnify it by a factor x, the magnified line can be cut into y copies of the original. The dimension of the line is the exponent of x that gives y. For the line segment in this case x equals y. For example, doubling the line produces two copies. The exponent is log (base 2) of 2, or $\log 2/\log 2 = 1$.

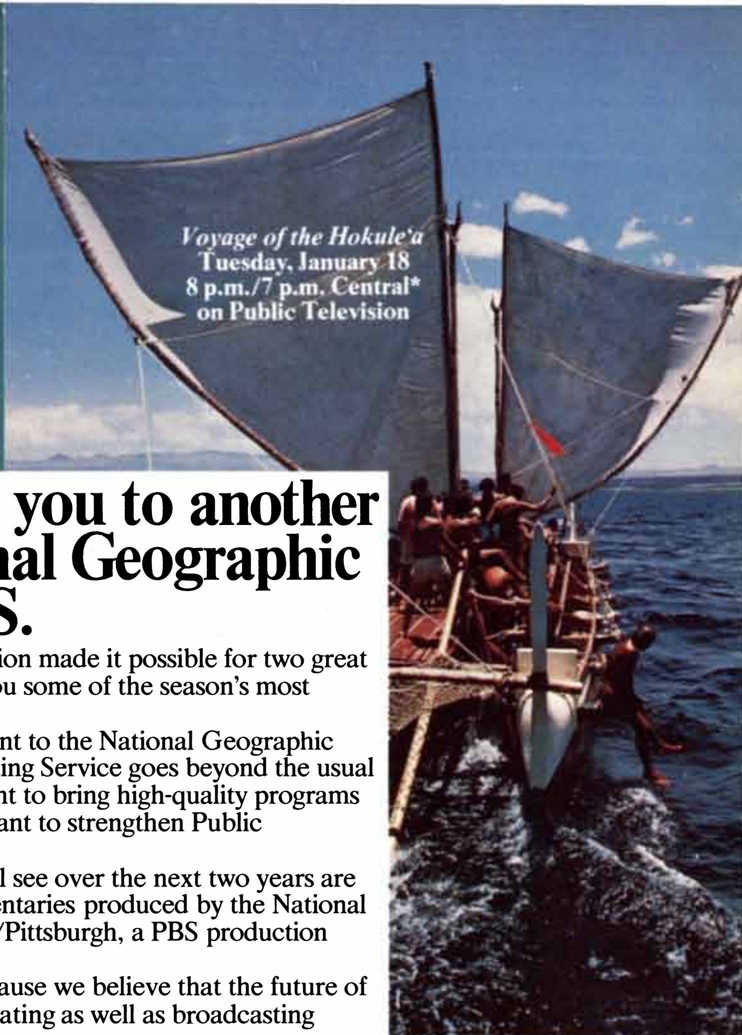
Magnify a square so that its edge doubles; the enlarged square can be cut into four copies of the original. In general if you magnify a plane figure by a linear factor x, its area increases by a factor of x^2. Its dimension is log (base 2) of 4, or $\log 4/\log 2 = 2$. If you double a cube's edge, the larger cube can be cut into eight copies of the original. Its dimension is log (base 2) of 8, or $\log 8/\log 2 = 3$. And so it goes for hypercubes in higher Euclidean spaces.

The snowflake is made by repeatedly replacing a line with one that is 4/3 as long. Thus it is reasonable to assign to the limit curve a dimension, called the Hausdorff dimension, that is log (base 3) of 4, or $\log 4/\log 3 = 1.26181\ldots$. The boundary of the flowsnake is constructed by repeatedly replacing a line with a zigzag path that is $3/\sqrt{7}$ as large. Its Hausdorff dimension is log (base $\sqrt{7}$) of 3, or $\log 3/\log\sqrt{7} = 1.12915\ldots$. It has infinite length, no area in units squared and a certain finite "size" expressed in units to the power of $1.12915\ldots$. Like all space-filling plane curves, the flowsnake itself has a dimension of 2 at the



A flowsnake paradox





Gulf welcomes you to another exciting National Geographic Season on PBS.

Last year Gulf Oil Corporation made it possible for two great American institutions to bring you some of the season's most exciting television.

Gulf's three-year commitment to the National Geographic Society and the Public Broadcasting Service goes beyond the usual corporate grant. We not only want to bring high-quality programs to the American audience; we want to strengthen Public Television, too.

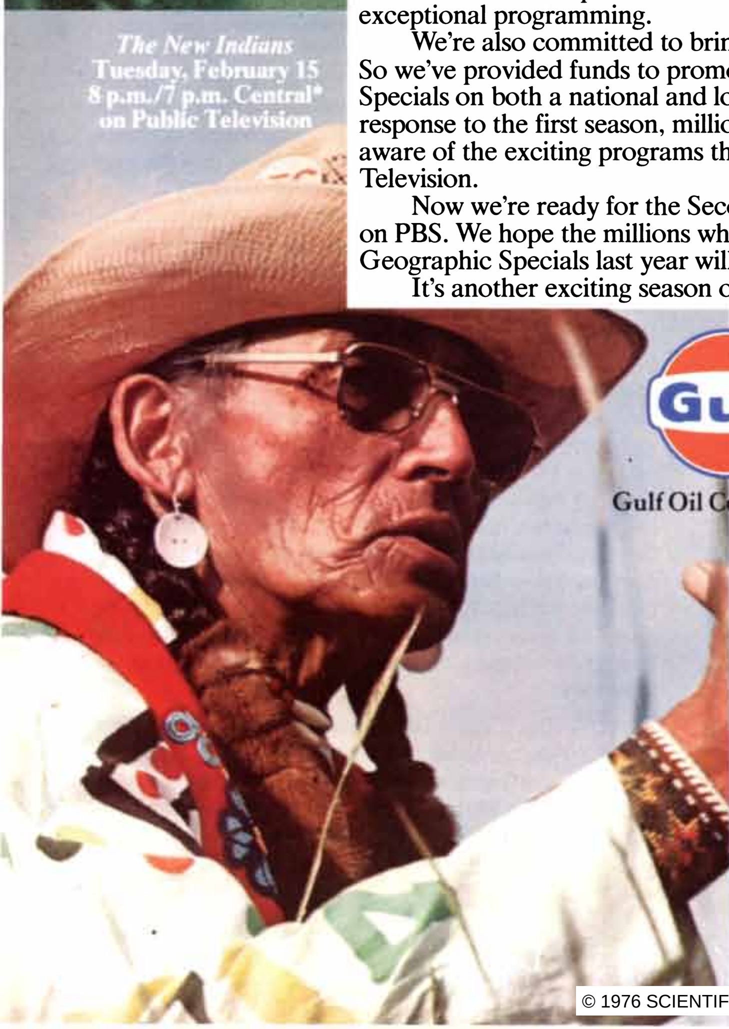
The eight new specials you'll see over the next two years are original American-made documentaries produced by the National Geographic Society and WQED/Pittsburgh, a PBS production center.

Gulf funded the project because we believe that the future of Public Television depends on creating as well as broadcasting exceptional programming.

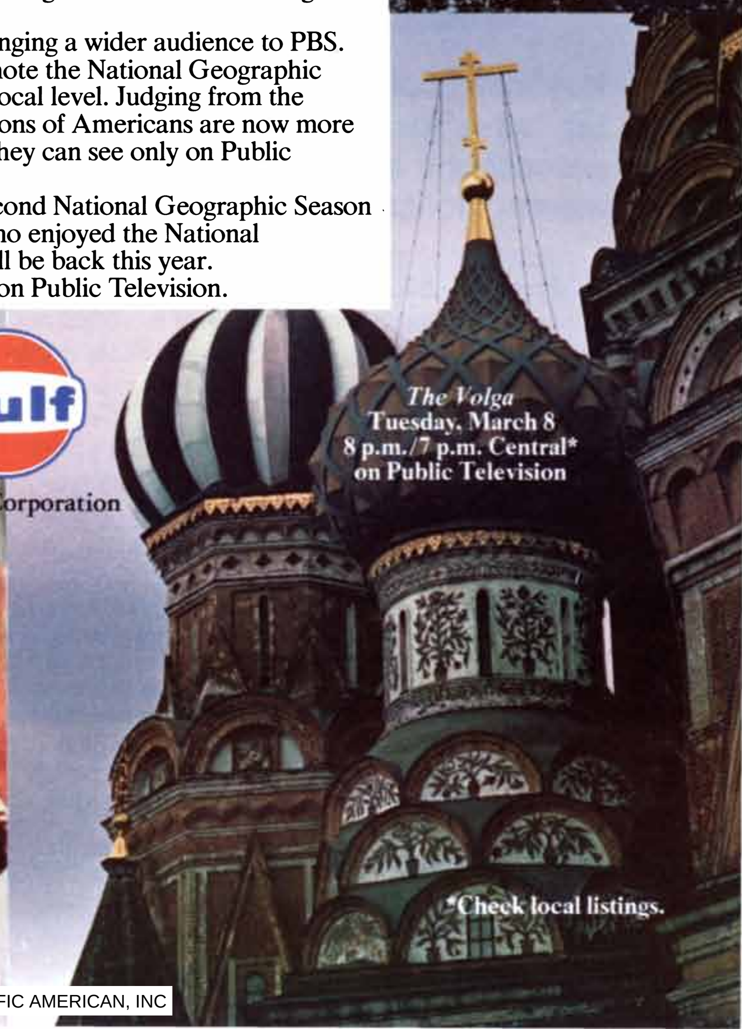
We're also committed to bringing a wider audience to PBS. So we've provided funds to promote the National Geographic Specials on both a national and local level. Judging from the response to the first season, millions of Americans are now more aware of the exciting programs they can see only on Public Television.

Now we're ready for the Second National Geographic Season on PBS. We hope the millions who enjoyed the National Geographic Specials last year will be back this year.

It's another exciting season on Public Television.



Gulf Oil Corporation



Hewlett-Packard offers for the first time (ever) full-power APL on a relatively small general purpose computer.

The recently introduced HP 3000 Series II computer, whose powerful data entry and data base management signaled a price/performance advance in data processing, now adds APL to its language repertoire, accompanied by a new CRT terminal especially designed to operate with APL.

The advantages of APL among computer languages are becoming increasingly apparent. APL is a general purpose programming language, rich in primitive operators and formal identities, that uses powerful symbols in shorthand fashion to define complete functions in very few statements or characters.

APL offers highly beneficial shortcuts to data manipulation in scientific and engineering applications, where it can bring to bear its ability to express complex calculations in a concise way, and to operate on groups of numbers as easily as on single ones. Because APL normally operates directly on data without special commands, the novice can do useful work at once, freed from the necessity of learning complicated procedures that stymie nonspecialists.

By making an unabridged and enriched APL available on a relatively small computer, HP fills a price/performance gap that has frustrated potential APL users in the past: the full power of APL was available only on a massive computer or through costly service bureau time—or one settled for the limited APL capability of a "portable" computer.

Hewlett-Packard's enriched version of the language, APL/3000, actually broadens APL's capability to include the handling of large data bases, file manipulation, and production of reports in desired formats—placing APL squarely in the decision-maker's realm. And since the 3000 Series II computer treats APL/3000 as a standard language subsystem, any of its five other programming languages (FORTRAN, COBOL, RPG, BASIC, and SPL) can be used concurrently with APL in batch or interactive modes on up to 12 terminals.

Hewlett-Packard has developed a CRT terminal especially designed to handle APL symbols: the new HP 2641. Its versatile keyboard carries full APL and

standard ASCII character sets. Additional special function keys can be programmed to speed data entry and reduce opportunities for error. Optional minicartridge tape transports provide storage that allows the user to prepare data *off-line*, transmit it rapidly to the computer in batch, and keep or transfer development programs on the pocket-sized cartridges.

APL/3000 software and firmware can be purchased outright for $15,000. The 2641 terminal with 4 Kbytes of memory costs $4100; with tape transports it costs $5700. The 3000 Series II computer system, which makes it all possible, costs from $2350 to $7500 per month on a five-year payout lease, or $110,000 to $350,000 by direct purchase (domestic U.S. prices only, maintenance not included). We'd like to tell you more.

Mail to: Hewlett-Packard, 1503 Page Mill Road, Palo Alto, CA 94304.
Please send me further information on

( ) HP 1611A logic state analyzer
( ) HP 3000 Series II computer system and APL/3000

Name ____________

Company ____________

Address ____________

City ________ State ______ Zip ______

00650

HEWLETT *hp* PACKARD

1503 Page Mill Road, Palo Alto, California 94304

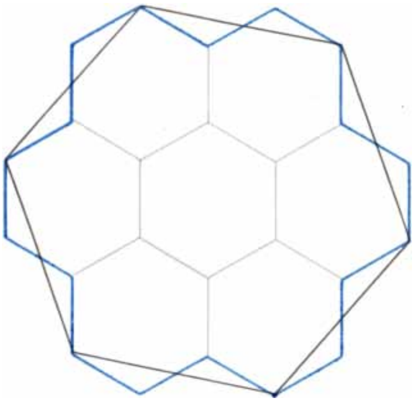
For assistance call: Washington (301) 948-6370, Chicago (312) 677-0400, Atlanta (404) 434-4000, Los Angeles (213) 877-1282.

Never buy a Bordeaux by the bottle. Buy a great, velvety Bordeaux by the label.

The B&G label.
B&G. 250 years of winemaking history from all the great wine regions of France: Bordeaux, the Loire Valley, Beaujolais, Côtes du Rhône and Burgundy.
Now B&G brings you Pontet-Latour, a great velvety wine from Bordeaux.
Superb. But affordable.
Just buy it by the label. **B&G.**

Letters of recommendation from France.

IMPORTED BY BROWNE VINTNERS, NEW YORK AND SAN FRANCISCO. AD PREPARED BY TT&P, INC.



Making the flowsnake's boundary

limit because it completely fills a plane region.

As we saw above, an enormous variety of fractal curves that do not self-intersect can be produced by simple recursive procedures. So can fractal surfaces with dimensions greater than 2, fractal solids with dimensions greater than 3 and so on. The illustration below, reproducing one that is in Mandelbrot's book, is made by replacing lines of four units with lines of eight units (as shown at the top) to produce a squarish, asymmetric snowflake with dimension log 8/log 4 = 1.5. Since each alteration of a segment adds the same amount of area as it subtracts, the limit curve bounds the same area as the original square.

Note that the Hausdorff dimension is a measure of complexity. The square snowflake is more complex than von Koch's snowflake because its dimension is higher. Mandelbrot has been working with such curves for so long that he has acquired an uncanny ability to look at a new fractal curve and, by an intuitive estimate of its complexity, guess its dimension with high accuracy.

In the light of these crazy curves, how do mathematicians currently define a curve? The scene is so crowded with monsters that no single definition covers all the objects to which the word "curve" is commonly applied. The topologist defines a curve as a set of points that are compact, connected and form a one-dimensional continuum. To make the definition clear, however, a lengthy discourse on point-set topology would be required. The definition catches well-behaved curves that diagram functions with derivatives, but it misses some of the nondifferentiable monsters we have been considering.

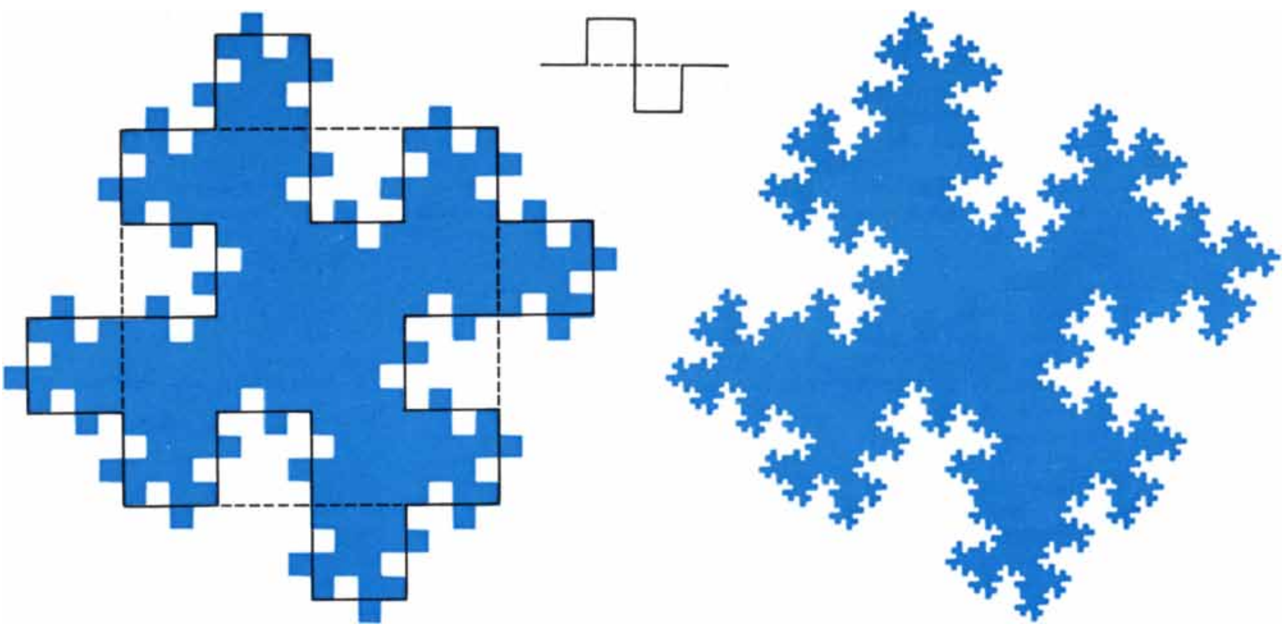
When attempts are made to define surface and volume, monsters more terrifying than flowsnakes crawl onto the landscape. This is a topic that must be postponed for a future column.

One of last month's problems was to guess how Dr. Matrix made a goblet of wine vanish inside a cylinder. The goblet was made of ice and was kept in its mold inside a freezer until Dr. Matrix was ready to perform. At the center of the table a small hole at the base of a slightly conical metal disk led into a hollow central leg. The disk was wired so that it became hot when Dr. Matrix pushed a concealed button on his desk. When the goblet melted, the water and the wine drained into the table leg.

It is easy to show how Dr. Matrix' formula of two variables, $(b + 1)/a$, generates a sequence of numbers that loops with period 5. Let a be the first number of the sequence and b the second. When the formula is applied recursively, the third number is $(b + 1)/a$, the fourth is $(a + b + 1)/ab$, the fifth is $(a + 1)/b$, the sixth is a and the seventh is b. The formula apparently was first noted by R. C. Lyness in *The Mathematical Gazette* (Vol. 26, 1942, page 62).

Many readers will be interested in two mathematically beautiful objects just placed on sale in the U.S. One is a splendid book on anamorphic art, profusely illustrated and with many color plates. Inserted in the book is a sheet of mirror paper for making a cylinder in which the distorted art can be viewed. The book, with the text by Fred Leeman, is titled *Hidden Images: Games of Perception, Anamorphic Art, Illusion—from the Renaissance to the Present.* Harry N. Abrams, Inc., is the publisher. It is a translation of a Dutch book that grew out of a major exhibit of anamorphic art earlier this year in Amsterdam. The exhibit is now touring the U.S.

The other item is an elegant but eminently playable chess set designed by Cy Endfield, heretofore available only in England. The silver- and gold-plated chess pieces are ingeniously cut at the top and the bottom so that they interlock to fit snugly over two cylinders. The leather-covered board, which folds into a box, has M. C. Escher's horse-and-rider tessellation on the squares. Around the sides is Escher's famous *Metamorphose,* a sequence of pictures that interlock much as the chessmen do.



The first three orders of Benoît Mandelbrot's square snowflake