Number Systems With a Complex Base: a Fractal Tool for Teaching Topology

Daniel Goffinet

The American Mathematical Monthly, Vol. 98, No. 3. (Mar., 1991), pp. 249-255.

Stable URL:
http://links.jstor.org/sici?sici=0002-9890%28199103%2998%3A3%3C249%3ANSWACB%3E2.0.CO%3B2-7

The American Mathematical Monthly is currently published by Mathematical Association of America.

---

courses use about the same number of teaching assistants per student. In the lecture-discussion section course we employ readers (upper-division students and some graduate students) to grade the homework that all students are required to turn in. In the self-paced course we employ tutors (upper-division students and some graduate students) on an hourly basis to help students, as described earlier.

There is also the problem of the efficient use of space. For the lecture course we use four large lecture halls three days a week for one hour and dozens of small rooms twice a week for one hour. In the self-paced course, two rooms are used all day at least three days a week. There is also a storage room. Several people here have tried to analyze this information with respect to cost and have decided that self-pacing is slightly more expensive but worth the outlay. Each institution would have to examine its own data to decide which method is more costly.

7. Other courses. Every beginning student at Berkeley who wishes to elect a mathematics course must take a diagnostic test in elementary algebra and trigonometry. Those who pass the test can choose the standard calculus course, Math 1AB (lecture-discussion section), Math 1S (self-paced), or Math 16AB, the weaker course described earlier. Those who fail the test must take a one-semester precalculus course, called Math P. After the success of the self-paced Math 1S was fully recognized, we introduced self-paced versions of Math P and Math 16AB. These courses have been operating successfully for several years, so that all our beginning mathematics students have the option of taking a lecture-discussion section course or a self-paced course. I believe that all lower division courses with large enrollments should give each student the option of attending large lectures and small discussion sections or of learning the subject at his own speed. I anticipate that the second-year mathematics course at Berkeley, which has large enrollments every term, will soon offer students these two options. It may develop that the self-paced second-year course will be even more popular than the first-year, self-paced courses are.

Number Systems With a Complex Base: a Fractal Tool for Teaching Topology

DANIEL GOFFINET
9 Rue de la Republique, St. Etienne, 42100 France

It so happens that I teach a first course in topology and a course in programming to the same students. Number systems other than the decimal one have proven to be a useful tool in the two areas. The number system with a complex base and digits 0 and 1 that is described in this note illustrates many topological notions and enthralled most of my students. An historical survey of the idea of using nonreal bases can be found in [**1**, pp. 188–190].

I gave them the program to play with, without any specific questions, in the hope of helping them to formulate and answer their own questions. Because they were only beginners in topology they mostly asked about programming details rather than the questions on connectivity that I had in mind. Still, I think it would make students with some topological background understand the necessity of the so-called "abstract" definitions.

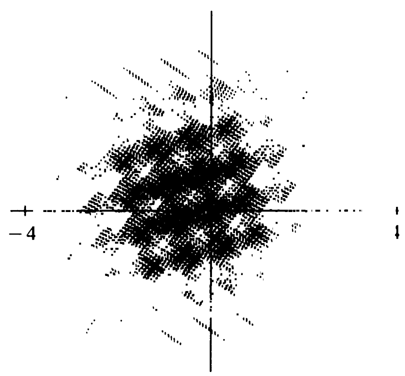
1. Definition of the set $\mathfrak{B}$. Let $b = 1/2$. Each $x \in [0,1]$ can be written as $\Sigma\{b^n: n \in A\}$ for some subset of A of $\mathbb{N}^+$, the set of positive integers. Occasionally a number has two representations (that is, different sets A). With $b = 1/3$, the set of the Σb^n, with A once more ranging over all subsets of $\mathbb{N}^+$, is a Cantor set. Representations are then unique, but there are holes in the set.

More generally, I define $\mathfrak{B}$ as the set of all the $\Sigma\{b^n: n \in A\}$. It is then easy to prove that

- If $0 < b < 1/2$, $\mathfrak{B}$ is totally disconnected and representations are unique.
- If $1/2 \leqslant b < 1$, $\mathfrak{B}$ is an interval, but different A may produce the same x.

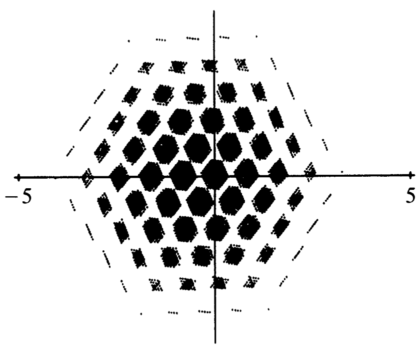
What does one get with a complex b? If b is either i or a complex cube root of 1, $\mathfrak{B}$ is easily seen to be a lattice. For other values of b inside the unit ball, see Figures 1–10. In what follows, b denotes a complex number with $|b| < 1$.

2. A simple proof that $\mathfrak{B}$ is compact. Let $\mathscr{P} = \mathscr{P}(\mathbb{N}^+)$ be the set of subsets of $\mathbb{N}^+$ with the topology induced by the metric $d(A,B) = 2^{-\min(A \Delta B)}$, where Δ denotes symmetric difference. The compactness of $\mathscr{P}$ then follows from the fact that $\mathscr{P}$ is complete and totally bounded.



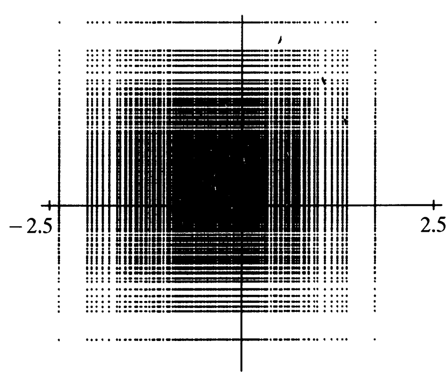
$b = 0.97e^{i\pi/1.95}$

FIG. 1.



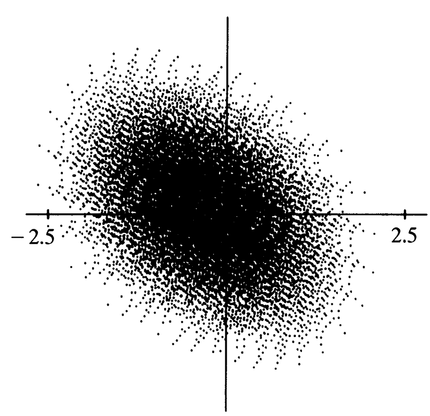
$b = -0.49 + 0.97\frac{\sqrt{3}}{2}$

FIG. 2.



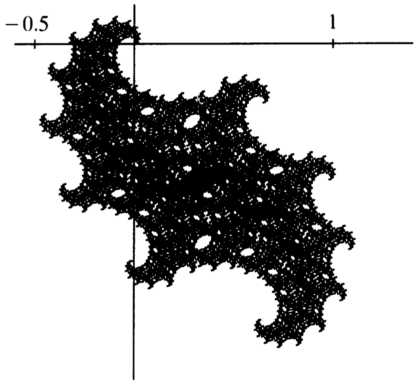
$b = 0.931i$

FIG. 3.



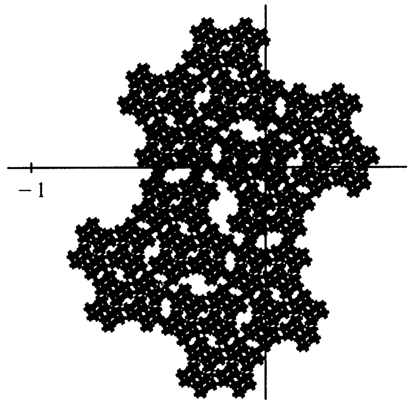
$b = 0.9\left(-\cos\frac{\pi}{13} + i\sin\frac{\pi}{13}\right)$

FIG. 4.



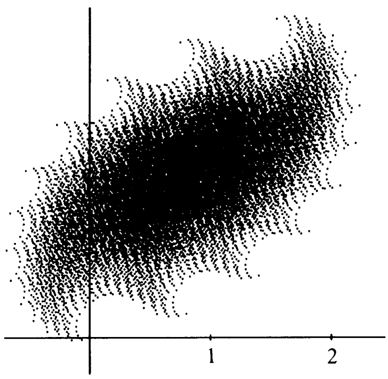
$b = 0.65 - 0.3i$

FIG. 5.



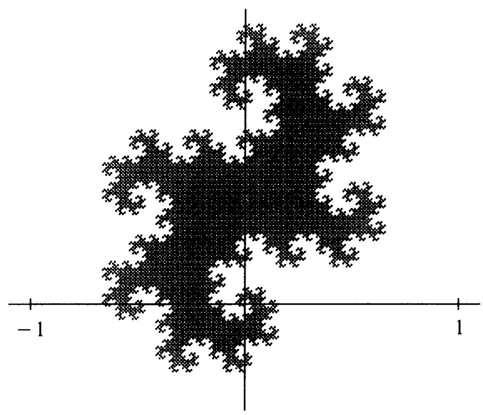
$b = -0.697e^{2\pi i/5}$

FIG. 6.



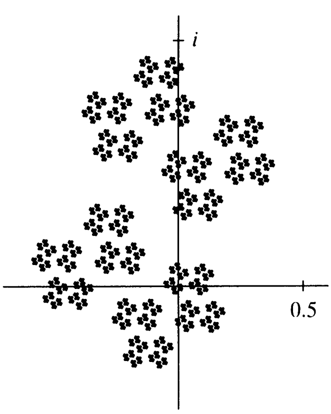
$b = 0.8 + 0.2i$

FIG. 7.



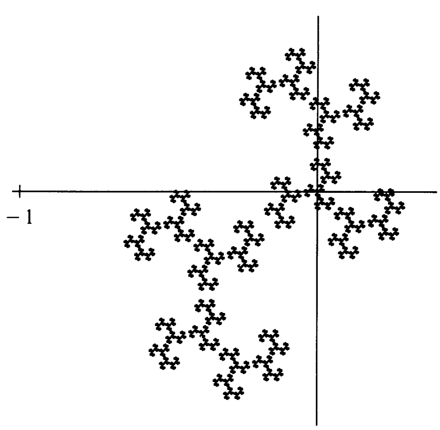
$0.5 + 0.5i$

FIG. 8.



$b = 0.2 + 0.6i$

FIG. 9.



$b = -0.62e^{2\pi i/7}$

FIG. 1–10. Some examples of $\mathfrak{B}$, where the images are produced by considering only sets A arising as subsets of $\{1, \ldots, 15\}$ (except for FIG. 9 which uses subsets of $\{1, \ldots, 13\}$).

$\mathscr{P}$ is complete. If A_n is a Cauchy sequence in $\mathscr{P}$ then for every $p \in \mathbb{N}^+$ the sets $A_n \cap [1, p]$ are eventually constant, which yields the limit.

$\mathscr{P}$ is totally bounded. For every p, $\mathscr{P}$ is a finite union of sets with diameter smaller than 2^{-p}, namely the 2^p sets obtained by considering, for each subset of $[1,p]$, all its supersets in $\mathscr{P}$. Alternatively, compactness follows from the fact that the symmetric-difference topology coincides with the product topology obtained by considering $\mathscr{P}$ as $\{0,1\}^{\mathbb{N}^+}$ (where $\{0,1\}$ is given the discrete topology).

If we define φ_b: $\mathscr{P} \to \mathbb{C}$ by $\varphi_b(A) = \Sigma\{b^n: n \in A\}$, then φ_b is continuous; hence, for each b, $\mathfrak{B} = \varphi_b(\mathscr{P})$ is compact. Another proof of compactness appears at the beginning of §3.

We will use the compactness of $\mathfrak{B}$ to prove that $\mathbb{C}$ is the union of copies of the set $\mathfrak{B}$ in the case that $b = (1+i)/2$ (see FIG. 8). The easy way to see this is to look at [**1**, §4.1]. But you can do it yourself with a pencil and graph paper. Because the set is in some ways a counterpart to $[0,1]$ (which is $\mathfrak{B}$ for $b = 1/2$), let us first look at the counterpart of $\mathbb{N}$ (which is obtained from binary expansions in $[0,1]$ by replacing each $1/2$ by 2 and considering only finite expansions). Because b is now $(1+i)/2$, $b = 1/B$ with $B = 1 - i$. Taking this B as base, and observing that $B^8 = 16$, leads us to look at $\mathscr{C}_7 = \{\Sigma\{B^k: k \in C\}: C \subseteq \{0,1,2,3,4,5,6,7\}\}$.

We start with the baby dragon $\mathscr{C}_3 = \{\Sigma\{b^k: k \in C\}: C \subseteq \{0,1,2,3\}\}$. When 4 is added to the set of admissible exponents, $\mathscr{C}_3$ becomes $\mathscr{C}_4$, which consists of two copies of $\mathscr{C}_3$. Continue until reaching $\mathscr{C}_7$, which consists of 16 $\mathscr{C}_3$s (see Figs. 11–12). Now it is easy to check that another copy of $\mathscr{C}_7$ translated 16 units rightward will fit with the first copy as in a jigsaw puzzle. Repeating yields a strip of $\mathscr{C}_7$s (see Fig. 13). Again, inspection shows that translating this doubly infinite strip by $16i$ yields a strip that fits with the first one. Thus every Gaussian integer (i.e., element of $\mathbb{G} = \mathbb{Z} + \mathbb{Z}i$) may be written uniquely in the form $c + 16(p + qi)$ with $c \in \mathscr{C}_7$ and $p, q \in \mathbb{Z}$. In short: $\mathbb{G} = 16\mathbb{G} + \mathscr{C}_7$.

Now, let us return to proving that $\mathbb{C}$ can be covered by $\mathbb{G}$-translations of $\mathfrak{B} = \{\Sigma_{k \geqslant 1} a_k b^k: a_k \in \{0,1\}\}$. Let $\mathfrak{B}_n = \{\Sigma_{1 \leqslant k \leqslant 8n} a_k b^k: a_k \in \{0,1\}\}$, with $\mathfrak{B}_0 = \{0\}$. Then $\mathfrak{B}_1 = b^8 \mathscr{C}_7 = \mathscr{C}_7/16$. Because $\mathfrak{B}_1/16 = \{\Sigma_{9 \leqslant k \leqslant 16} a_k b^k: a_k \in \{0,1\}\}$, we have that $\mathfrak{B}_2 = \mathfrak{B}_1 + (1/16)\mathfrak{B}_1 = \mathfrak{B}_1 + (1/16^2)\mathscr{C}_7$. Similarly $\mathfrak{B}_n = \mathfrak{B}_{n-1} + (1/16^n)\mathscr{C}_7$. Assume inductively that $\mathfrak{B}_{n-1} + \mathbb{G} = \mathbb{G}/16^{n-1}$ (clearly, $\mathfrak{B}_0 + \mathbb{G} = \mathbb{G}$). Then $\mathfrak{B}_n + \mathbb{G} = (\mathfrak{B}_{n-1} + \mathbb{G}) + (1/16^n)\mathscr{C}_7 = (1/16^{n-1})\mathbb{G} + (1/16^n)\mathscr{C}_7 = (1/16^n)(16\mathbb{G} + \mathscr{C}_7) = \mathbb{G}/16^n$. Since $\mathfrak{B} + \mathbb{G} \supseteq \mathfrak{B}_n + \mathbb{G}$, $\mathfrak{B} + \mathbb{G}$ contains all points in a grid of mesh $(1/16)^n$. It follows that for any square there is a finite union of $\mathbb{G}$-translates of $\mathfrak{B}$ (a compact set) which is dense in the square. Hence, $\mathfrak{B} + \mathbb{G}$ contains the square, and so $\mathfrak{B} + \mathbb{G} = \mathbb{C}$.

As with the usual base-10 representations, we lose uniqueness when we pass to the limit; for example: $b = 1 + (b^2 + b^3 + \cdots)$.

3. Going deeper into the topology of $\mathfrak{B}$. I owe the following fixed-point idea to [**5**]. If $\mathscr{K}$ is the set of nonempty, compact subsets of the plane, the Hausdorff distance is defined as $\delta(X,Y) = \max\{\sup d(x,Y), \sup d(y,X)\}$; then $(\mathscr{K}, \delta)$ is a complete space [**2**, p. 56]. I now define T_0 by $T_0(z) = bz$ and T_1 by $T_1(z) = b + bz$. The functions T_0 and T_1 are contractions of $\mathbb{C}$, and also clearly induce contractions of $(\mathscr{K}, \delta)$. If $T = T_0 \cup T_1$ is defined by $T(X) = T_0(X) \cup T_1(X)$, then T is also a contraction of $(\mathscr{K}, \delta)$, because we can easily prove that

$$\delta(X_1 \cup X_2, Y_1 \cup Y_2) \leqslant \max\{\delta(X_1, Y_1), \delta(X_2, Y_2)\}.$$

Therefore T has a unique fixed point in $\mathscr{K}$, and it is the limit of the sequence of the $T^n(K_0)$ for any K_0. Taking $K_0 = \{0\}$ we have that $T^n(K_0)$ equals those points

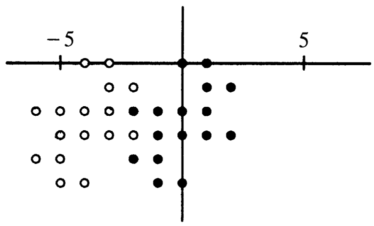


FIG. 11. Two copies of the baby dragon, $\mathscr{C}_3$, yields $\mathscr{C}_4$.

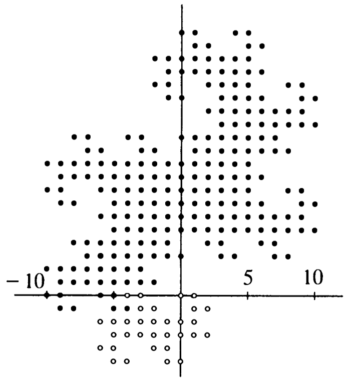


FIG. 12. The set $\mathscr{C}_7$, which is built up from 8 copies of $\mathscr{C}_4$.

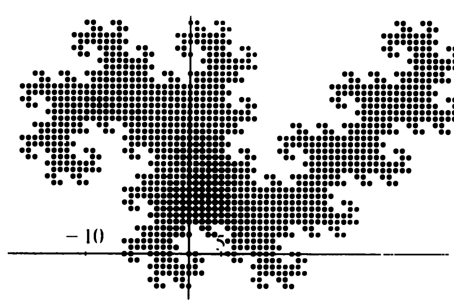


FIG. 13. The Gaussian integers can be tiled with copies of the set $\mathscr{C}_7$.

in $\mathfrak{B}$ with at most n "digits" in their representation. So $\mathfrak{B}$, the limit of the sequence, is in $\mathscr{K}$ (i.e., is compact) and is a fixed point of T.

4. Connectivity. I now claim that $\mathfrak{B}$ is either connected or totally disconnected. Set $\mathfrak{B}^0 = T_0(\mathfrak{B})$ and $\mathfrak{B}^1 = T_1(\mathfrak{B})$; then $\mathfrak{B} = T(\mathfrak{B}) = \mathfrak{B}^0 \cup \mathfrak{B}^1$.

Case 1. $\mathfrak{B}^0 \cap \mathfrak{B}^1 = \varnothing$. If $m \in \mathfrak{B}^i$, then $\mathscr{M}$, its connected component in $\mathfrak{B}$, is contained in $\mathfrak{B}^i$ ($\mathfrak{B}^0$ and $\mathfrak{B}^1$ are compact as images of $\mathfrak{B}$ by the continuous functions T_0 and T_1, respectively). But $\mathfrak{B}^i$ itself is a disjoint union of compact sets so $\mathscr{M}$ is included in some $T_jT_i(\mathfrak{B})$. By induction we can prove that $\mathscr{M}$ is a subset of the intersection of a decreasing sequence of compact sets whose diameters tend to zero. It follows that $\mathscr{M}$ is simply $\{m\}$.

Case 2. $\mathfrak{B}^0 \cap \mathfrak{B}^1 \neq \varnothing$. In this case $\mathfrak{B}$ is connected. Choose $a \in \mathfrak{B}^0 \cap \mathfrak{B}^1$, say, $a = ba_0 = b + ba_1$. Because $\mathfrak{B}$ is compact, to prove that it is connected it is enough to check that for every $z \in \mathfrak{B}$ and $\varepsilon > 0$, there is an ε-chain in $\mathfrak{B}$ from z to a. If ε is greater than the diameter of $\mathfrak{B}$ then (z, a) is such a chain. Let us show that from α-chains we can build $\lambda\alpha$-chains where $\lambda < 1$ is the modulus of b. Since $z \in \mathfrak{B}$ we have two cases:

If $z \in \mathfrak{B}^0$ then $z = bd$, and we can "α-link" d to a, then a to a_0. Now take the images by T_0 of these chains (i.e., simply multiply by b) to $\lambda\alpha$-link bd to ba to $ba_0 = a$.

If $z \in \mathfrak{B}^1$ then $z = b + bd$, and we can α-link d to a to a_1. Then take the images of these chains by T_1.

If the modulus r of b is less than 1/2, then $\mathfrak{B}^0 = \{\Sigma_{k \geqslant 2} a_k b^k\}$ and $\mathfrak{B}^1 = \mathfrak{B}^0 + b$ are disjoint. For if $\Sigma_{k \geqslant 2} a_k b^k = b + \Sigma_{k \geqslant 2} a'_k b^k$, then $b = \Sigma_{k \geqslant 2} \alpha_k b^k$ with $\alpha \in \{-1, 0, 1\}$ and, taking the modulus, $r < \Sigma_{k \geqslant 2} r^k = r^2/(1-r)$ which implies $r \geqslant 1/2$. Therefore, when $r < 1/2$ we are sure that $\mathfrak{B}$ is totally disconnected.

On the other hand, when, for a fixed argument of b, r grows, $\mathfrak{B}^0$ and $\mathfrak{B}^1$ are bigger, and may or may not meet. By looking at Figures 1–10 you can try to guess which among those $\mathfrak{B}$s are connected, checking your guesses with §7.

5. How to get nice pictures. With color of course! The idea is to take b big enough for φ not to be injective, but not too big, and to associate to each point of the screen a color according to the number of times it is generated. An example of the sort of effect this yields appears in FIGURE 5, where there are dark gray $\mathfrak{B}$-like regions inside $\mathfrak{B}$. Another example: With $b = -0.1 - 0.8i$ and the range of colors taken to be light blue, blue, green, light green, yellow, light magenta, red, and white, one gets a silky cushion with sun rays shining from the bottom left corner. If you like hairy monsters (see FIGS. 5, 6, 8, and 10) choose b with an argument very close to a root of 1. To get soft and friendly-looking $\mathfrak{B}$s, use a bigger modulus, and an arbitrary argument.

6. Programming. In order to quickly generate the set of finite parts of $\mathfrak{B}$, I used a Boolean version of the binary number system: Starting with an array of Falses—standing for 0s—I repeatedly added 1 (that is, TRUE) and then converted that array into a complex number, which was then displayed on the screen.

In the procedure given below *max* is the number of binary digits to be used (which I took to be 16 for my images); *b*[*i*] is an array element containing b^i; *add* is a complex number addition routine defined elsewhere in the program; *bin_dig* is the name of the binary-digit array.

```
procedure add_one_and_make_a_complex_number;
    var carry: Boolean
        k: 1..max
    begin
       carry := true; k := max;
       while carry do begin
             bin_dig[k] := not (bin_dig[k]);
             if not (bin_dig[k]) then k := k - 1
                                    else carry := false
                       end;
       z := 0;
        for k := 1 to max do if bin_dig[k] then add (z,
             b[k], z)
    end;
```

7. Which images are connected? Let $\mathfrak{B}_i$ be the set of Figure i ($i \leqslant 10$). However disconnected $\mathfrak{B}_2$ and $\mathfrak{B}_3$ may seem, the truth is the opposite.

The connectedness of $\mathfrak{B}_3$ is easily proved by looking at real parts: they form another $\mathfrak{B}$-like set with $b = (-0.931)^2$. Because $(0.931)^2 > 1/2$, the real parts form an interval; similarly, so do the imaginary parts. Thus $\mathfrak{B}_3$ is simply a full rectangle. If b_2 denotes the base used to construct $\mathfrak{B}_2$, then the cube of b_2 is close to 1. Thus $\mathfrak{B}_2$ contains the point $\Sigma_{n=1}^{\infty} b_2^{3n} = 1/(1 - b_2^3)$, whose actual value is about $12 + 2i$. Thus the image in FIGURE 2 is only a small part of the true image of $\mathfrak{B}_2$, which seems to be hexagonal.

Why do $\mathfrak{B}_2$ and $\mathfrak{B}_3$ look broken? Because the images were made with subsets of $\{1, \ldots, 15\}$ instead of all subsets of $\mathbb{N}^+$; when b has a modulus close to 1, all the "digits" are needed to make $\mathfrak{B}$ look whole. Another example: $\mathfrak{B}_9$ is disconnected, although one cannot really tell from the image—a proof is needed. The four upper right blobs can be separated from the four bottom left ones by a strip that can be physically measured and seen to be greater than $1/50$. The modulus of b is smaller than 0.64, so with the powers greater than 13 a gap wider than $0.64^{14}/(1-0.64)$ (which is well under $1/50$) cannot be bridged. Therefore $\mathfrak{B}_9$ is indeed disconnected.

The set of those b for which $\mathfrak{B}$ is connected is described in [**3**], where, among other things, it is proven that $\mathfrak{B}$ is connected if $|b| > \sqrt{2}/2$. Among the $\mathfrak{B}$s shown in FIGURES 1–10, only $\mathfrak{B}_9$ and $\mathfrak{B}_{10}$ are disconnected.

8. Miscellaneous. A lot of things are left to be explored. When $\mathfrak{B}$ is connected and simply connected (see FIGS. 7 and 8), is the boundary of fractal dimension strictly greater than 1 (as it seems)? Is that boundary the image of a continuous map from $[0, 1]$ in $\mathbb{R}^2$ (as it seems)? Why does $\mathfrak{B}_5$ have $\mathfrak{B}_5$-like eggs inside it (dark gray)?

I tried to replace $\mathbb{C}$ by the algebra with $(a, b) \cdot (c, d) = (ac + bd, ad + bc)$, but got nothing worth showing. Why not?

Acknowledgement. I am grateful to the referees for many suggestions.

REFERENCES

1. D. E. Knuth, The Art of Computer Programming, volume 2, Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1981.
2. J. Dieudonné, Fondements de l'Analyse Moderne, Vol. 1, Gauthier-Villars, Paris, 1965.
3. M. Barnsley, Fractals Everywhere, San Diego, Academic Press, 1988
4. W. J. Gilbert, Fractal geometry derived from complex bases, *Math. Intelligencer*, 4 (1982) 78–86
5. J. E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.*, 30 (1981) 713–747.

A Characterization of the Cantor Function

DONALD R. CHALICE
Department of Mathematics, Western Washington University, Bellingham, WA 98225

This note presents a characterization of the Cantor function that might be used as an alternative or addition to the development usually presented in real analysis courses. It may also be utilized to give a short program in *Mathematica* that easily generates the Cantor function and other similar functions which we call "devil's staircases" (see [**4**]).

THEOREM. *Any real-valued function* $F(x)$ *on* $[0, 1]$ *that is monotone increasing and satisfies* (*a*) $F(0) = 0$, (*b*) $F(x/3) = F(x)/2$, *and* (*c*) $F(1-x) = 1 - F(x)$, *is the Cantor function.*

Before presenting the proof, recall (see [**1**]) that if we consider the closed interval $[0, 1]$ and remove the open middle third, $(1/3, 2/3)$, and next remove the open middle thirds $(1/9, 2/9)$ and $(7/9, 8/9)$ of the two remaining intervals, and then remove the open middle thirds of the remaining four intervals and so on, indefinitely, what remains is the *Cantor set* C. Alternatively, x is in C iff x has a base-3 expansion consisting only of the digits 0 and 2.