Chance and Fractals

Radu V. Craiu and Xiao-Li Meng [1]

Department of Statistics, The University of Chicago

October 18, 2000

1 What are Fractals?

The name *fractal*, coined in 1975 by the mathematician Benoit Mandelbrot, known as the godfather of fractals, comes from the Latin adjective *fractus*. The corresponding Latin verb *frangere* means "to break" or "to create irregular fragments," as Mandelbrot put it. To our best knowledge, there is no unitary mathematical definition of a fractal, although many attempts have been made. In layman's terms, a fractal is a "picture" with an incredible level of detail. No matter how deep one zooms in to it, one will find irregular details as well as miniatures of parts of the original picture. Since 1975, the subject has received a great deal of research as well as public attention. Many articles and books have been written, for experts as well as for general public, among which Mandelbrot's 1977 book, *The Fractal Geometry of Nature*, is a must for anyone who is interested in the subject. A selection of further reading is given in the sidebar. For general audiences, fractals are often presented as a kind of "computer art", as they are computer-generated, colorful, with fascinating geometric shapes. The most captivating aspect of a fractal is that, at first sight, it may appear to have a highly irregular geometric shape, but with a closer look one will find that it is in fact exceedingly regular in the sense that at any detailed level the same pattern repeats. Perhaps this is best summarized by the title of Lauwerier's book, *Fractals: Endlessly Repeated Geometrical Figures*.

[1] We thank D. Murdoch for inviting us to write this article for the Statistical & Computing Column of *Chance* and for comments, and S. Lalley, M. Stein and J. Servidea for helpful exchanges. The research is supported in part by NSF grant DMS 96-26691 and in part by NSA grant MDA 904-9610007.

2 Fractals and Monte Carlo

The title of this article was inspired by a chapter title of the aforementioned Lauwerier's book, "Chance in Fractals", by which he means that in order to create fractals that can resemble reality (e.g., the west coastline of Britain), one needs to introduce a random component, that is *chance*, into fractals, and hence *stochastic fractals* (the term *stochastic* comes from a Greek verb meaning "guessing"). Or as Mandelbrot wrote in the first chapter of his 1977 book, "The most useful fractals involve *chance* and both their regularities and their irregularities are statistical." (The emphasis of *chance* was Mandelbrot's.) The reason is that deterministic fractals are too "regular", and thus cannot realistically describe certain patterns and shapes created by nature, which are at the mercy of chance. Instead of requiring that different parts of a fractal have the exact same geometrical shape, we can require them to have the same statistical properties. For example, instead of having a line segment always oriented to a particular direction, we can require its orientation angle (with respect to a well-defined axis) to follow a particular statistical distribution at any level.

The theoretical study of stochastic fractals is typically rather complicated, as one may expect. However, such fractals are usually easy to simulate on a computer using the so-called *Monte Carlo* method. To quote a key inventor of the method, Stanislaw Ulam, "Laplace asserted that the theory of probability is nothing but calculus applied to common sense. Monte Carlo is common sense applied to mathematical formulations of physical laws and processes." It is common sense because it allows us to actually follow a mathematically formulated physical process via computer simulation. As a result we are able to observe the dynamics and outcome of the process, and therefore to analyze various properties of the process even if we are unable to write down any precise formulae for these properties.

At the core of any Monte Carlo method is the generation of random numbers; by that we typically mean a sequence of independent numbers, each of them following the uniform distribution on the unit interval (0,1). It is not surprising that fractals are connected to the generation of such a sequence because an infinite set of random numbers is an extreme sort of stochastic fractal: at a first glance the sequence looks completely "chaotic" in the sense that there is no apparent pattern to speak of, yet it is extremely regular in the statistical sense because any random subsequence of it follows the exact same uniform distribution. Such sequences are usually generated by a *dynamical system*, which typically builds upon a single deterministic (or stochastic) map, say $f : \mathcal{X} \to \mathcal{X}$ ($\mathcal{X}$ is some set, for example the (0,1) interval). We take an element X_0 from $\mathcal{X}$ and by recursively applying the map f, we will obtain $X_1 = f(X_0), X_2 = f(X_1) = f(f(X_0)), ...$, and so on. The set $\{X_0, X_1, X_2, \ldots, ...\}$ is called the *orbit* of X_0 under the map f. Many fractals are just graphical representations of orbits and quite a few of the pseudorandom generators in use are based on a particular dynamical system, that is, the numbers produced by the generator are just elements from an orbit. Note that these generators are called *pseudorandom generators* because their output is not really random since the entire orbit is determined by the *seed* X_0. However, with appropriately constructed dynamical systems, the resulting pseudorandom numbers are "random" enough to be useful in many applications.

3 Antithetic Variates and Latin Hypercube Sampling

We stumbled upon fractals when we were looking for effective ways to generate *antithetic variates*, which are useful for reducing simulation errors in Monte Carlo estimation. The word *antithetic* refers to the main objective of the method, that is, to produce random numbers that are *negatively correlated*; the idea was introduced by Hammersley and Morton in a paper published in 1956. The reason for us to seek such negative correlation is clear from the following simple example. Suppose we have a draw X from a distribution symmetric about zero. Then we know $\tilde{X} = -X$ is also a draw from the same distribution because of the symmetry. Consequently, if we can make sure that for every draw X in our Monte Carlo sample there is the "opposite" (and hence antithetic) draw $-X$ in the sample, then the sample average will be a perfect estimate of the actual mean of the distribution, that is, zero. In other words, in this extreme case the Monte Carlo mean estimator will have no error because X and $\tilde{X}$ are perfectly negatively correlated, that is, $\mathrm{Corr}(X, \tilde{X}) = -1$.

In reality, we do not have such perfect estimators when we need Monte Carlo methods (e.g., if we know a distribution is symmetric about zero, then we know the mean of the distribution has to be zero as long as the distribution has a mean), but the idea of using antithetic variates to "balance out" noise and thus improve Monte Carlo efficiency is generally quite useful.

Generating a pair of antithetic variates is typically straightforward. For a pair of antithetic uniform variates on the unit interval (0, 1), we only need to take $X_1 = u$ and $X_2 = 1 - u$, where u is uniformly distributed on (0,1) (which we will denote henceforth by $u \sim U(0,1)$). Clearly, if $u \sim U(0,1)$, then $1 - u \sim U(0,1)$. Furthermore, since mathematically the correlation is the same as the cosine of the angle between the two directions defined by u and $1-u$, which are opposite to one another with respect to the center of the unit interval (0,1), the correlation between X_1 and X_2 is $\mathrm{Corr}(X_1, X_2) = \cos(180^0) = -1$. Thus the pair $\{X_1, X_2\}$ achieves *extreme antithesis* (EA), that is, the two components are as "opposite" as they can possibly be. However, it is a harder problem to generate antithetic uniform variates $\{X_1, \ldots, X_k\}$ when $k > 2$ such that they achieve EA, namely, such that the correlation between any pair of X_i and $X_{i'}$ is $-1/(k-1)$, where $i \neq i'$. Note that for $k > 2$, it is impossible to have all pair-wise correlations to be -1, because it is impossible to have, say, three directions that are pair-wise opposites. The best one can do is to have the three directions in the same plane, at a 120^0 angle apart from each other, where the angle between two directions is defined as the one not exceeding 180^0. This gives the pair-wise correlation $\cos(120^0) = -1/2$. Note that in order for $\{X_1, \ldots, X_k\}$ to achieve EA, the sum $S_k = X_1 + \cdots + X_k$ must be a constant (which is its mean, $k/2$), because EA is equivalent to the variance of S_k being zero. Recall for $k = 2$, $X_1 + X_2 = u + (1 - u) = 1$.

Among the techniques that we have investigated, we find that the *Latin hypercube sampling* method of McKay, Beckman and Conover is quite appealing for a variety of practical and theoretical reasons. The method has two steps. In the first step, we generate k independent random numbers, denoted by $u_1, \ldots, u_k$, from $U(0,1)$. In the second step, we randomly permute (i.e., "stir") the set $\{0, 1, ..., k-1\}$ into $\{j_1, j_2, \ldots, j_k\}$ and then let $X_i = (u_i + j_i)/k, \quad i = 1, \ldots, k$. It is clear that because $0 \leq j_i \leq k - 1$, all X_i's are in (0,1), and thus $\{X_1, \ldots, X_k\}$ represents a point in the

k-dimensional hypercube $\mathcal{X}_k = (0,1)^k$. In fact, it can be shown that $X_i \sim U(0,1)$ for all $i = 1, \ldots, k$ and $\text{Corr}(X_i, X_{i'}) = -(k+1)/k^2 < 0$ for any $i \neq i'$.

4 Iterative Latin Hypercube Sampling and Fractals

Although the Latin hypercube sampling method produces negatively correlated random variates, it does not achieve EA, namely, $\text{Corr}(X_i, X_{i'})$ does not achieve the minimal possible value $-1/(k-1)$. It turns out that we can get arbitrarily close to this goal by iterating the Latin hypercube sampling method. That is, we can treat the $\{X_1, \ldots, X_k\}$ obtained by the Latin hypercube sampling as the $\{u_1, \ldots, u_k\}$ from its first step, and then repeat the second step. More precisely, we define a stochastic map from $\mathcal{X}_k$ to $\mathcal{X}_k$ via

$$X_i^{(t+1)} = \frac{X_i^{(t)} + j_i^{(t)}}{k}, \quad i = 1, \ldots k, \tag{1}$$

where t indexes the iteration (with $t = 1$ corresponding to the Latin hypercube sampling method), and $\{j_1^{(t)}, \ldots, j_k^{(t)}\}$ is a new permutation of $\{0, 1, ..., k-1\}$ independent of all previous permutations. It can be shown (by induction, for example) that any $X_i^{(t)}$ follows the $U(0,1)$ distribution, and that for any $i \neq i', \text{Corr}(X_i^{(t)}, X_{i'}^{(t)}) = -(1-k^{-2t})/(k-1)$, which approaches $-1/(k-1)$ rapidly as t gets large, i.e., as the iteration goes on.

The iterative scheme (1) defines a stochastic dynamical system, whose *attractor*, namely the collection of all possible limiting points as t approaches ∞, is a so-called *self-similar* fractal, whose shape depends on the value of k. We will call such a fractal *the antihype fractal*, highlighting its use (*anti*thetic coupling) and its origin (*hype*rcube sampling). Figure 1 displays the antihype fractal with $k = 3$ on its sitting plane $x_1 + x_2 + x_3 = 3/2$. Figure 2 plots the projection of the fractal on to the (X_1, X_2)-plane, where the effect of projection is seen in the elongation at the 135^0 direction. It is also the *support* of the joint distribution of $\{X_1^{(t)}, X_2^{(t)}\}$ as t approaches ∞, namely, the set of points with positive probability density. Because it is impossible to use the ideal $t = \infty$ iterations, the two plots were actually based on $t = 10$, which is still too fine a resolution for visualization as we can only see 4 or 5 levels of detail out of the 10 possible. Nevertheless, the self-similarity is evident. The fractal is hexagon-shaped and consists of six smaller identical shapes, each of which consists of yet again six smaller but identical shapes, and the pattern repeats an infinite number of times as the scale approaches zero (at least in our minds).

The reason that the dynamical system (1) achieves EA at the limit can be understood by observing how at each iteration it forbids the pair (X_1, X_2) to take certain values. In Figure 3 we use colors to illustrate this process. Suppose we start by having the entire unit square colored in black, which represents that the initial $(X_1^{(0)}, X_2^{(0)})$ is uniformly distributed on the square $(0,1)^2$. However, after the first iteration, $(X_1^{(1)}, X_2^{(1)})$ can only take values outside the three light green squares along the main diagonal; it is actually uniformly distributed on the region outside the light green squares. This makes sense as the elimination of the light green squares clearly does not alter the marginal uniformity of $X_1^{(1)}$ (and thus any margin given the symmetry) since for any value of $X_1^{(1)}$ inside (0,1), the corresponding cumulative "height" of the allowable $X_2^{(1)}$ is the same 2/3. However, the elimination of the light green areas makes $X_1^{(1)}$ and $X_2^{(1)}$ negatively correlated because these areas are responsible for the positive correlation between $X_1^{(1)}$ and $X_2^{(1)}$. Viewing this way, it is obvious that we can further reduce the correlation by repeating the same elimination process on the remaining six blocks (which again will not alter the marginal uniformity). This is achieved by the second iteration, which takes out the $6 \times 3 = 18$ blue blocks. We can continue this process by then taking out the $6^2 \times 3 = 108$ green blocks ($t = 3$), and then the $6^3 \times 3 = 648$ red blocks ($t = 4$), and then the $6^4 \times 3 = 3,888$ yellow blocks ($t = 5$). In our mind, we can continue this elimination process indefinitely, but the graphical resolution does not allow us to go beyond five iterations. Nevertheless, we can see the black "dust specks" left in Figure 3 form the same geometric shape as the ones in Figure 2, which was based on 10 iterations. The term *dust* is used purposely here because any antihype fractal obtained as above is in fact a case of the so called "Cantor dust".

5 More Fun Fractals . . .

We conclude by mentioning that once we obtain one fractal, a well-known technique for creating more fascinating looking fractals is to mix various *similar transformations*, that is, transformations that preserve the geometric shape but not necessarily the size or the orientation of a set. Figure 4 plots a 2-D fractal, which is the support of the equal-weight mixture of the uniform distribution on the 2-D fractal in Figure 2 and its 90^0 rotation with respect to the center (0.5, 0.5). The mixture distribution is the uniform distribution on this new fractal, because the original fractal (grey) and the rotated fractal (black) have no overlap. This is easiest to prove by observing that the two fractals do not overlap at the four corner blocks in Figure 4 and thus, by self-similarity, they do not overlap anywhere because the remaining four blocks (excluding the middle empty block) are just smaller versions of the same mixture fractal. Consequently, this beautiful table-cloth-design looking fractal has the same "marginal uniformity" property as the one in Figure 2, namely, the uniform distribution on it has both margins distributed as $U(0, 1)$. We leave it as an exercise to reader to figure out the new correlation between X_1 and X_2 (Hint: no calculation is needed!), and to consider whether it is possible to have a different rotation (i.e., other than 90^0 and its equivalences such as 270^0) that preserves the marginal uniformity after the same mixture transformation.

Sidebar – References and Further Readings

Chatterjee, S. and Yilmaz, M. R. (1992), "Chaos, Fractals and Statistics," *Statistical Science*, 7, 49-121.

> For statistically oriented readers, this overview article may be of particular interest as it discusses the role of statistics in a wide variety of applications of dynamical systems in different branches of sciences.

Craiu, R. and Meng, X. L. (2000), "Multi-process Parallel Antithetic Coupling for Forward and Backward Markov chain Monte Carlo," *Technical Report*, Department of Statistics, The University of Chicago.

> This paper documents the use of antithetic variates, particularly those generated by the iterative Latin hypercube sampling, in Markov chain Monte Carlo, a very powerful class of methods for scientific and statistical computation.

Hammersley, J. M. and Handscomb, D. C. (1965), *Monte Carlo methods,* London: Methuen & Co., Ltd; New York: Barnes & Noble, Inc.

> A classic and basic reference book for anyone who wants to learn basic Monte Carlo methods. It also offers good descriptions and explanations of the antithetic principle.

Falconer, K. (1990), *Fractal Geometry – Mathematical Foundations and Applications*, New York: Wiley. & Falconer, K. (1997), *Techniques in Fractal Geometry*, New York: Wiley.

> This pair of books are for those who want to learn fractal geometry as a research subject, for they are known for their rigorous treatment of notions like fractal dimensions, attractors and possible applications of the fractal geometry in other fields.

Lauwerier, H. (1991), *Fractals: Endlessly repeated Geometrical Figures.* Princeton: Princeton University Press.

> A good book for those who want to experiment with their computer; the mathematical parts can be skipped and one would still enjoy the thrills of "computer art".

Mandelbrot, B.B. (1977),*The Fractal Geometry of Nature*, New York: Freeman and Company.

> The updated and augmented version of Mandelbrot's 1975 book in French (as well as of its 1977 English translation), the historical essay where the name "fractal" was coined. The author's main proposal is to use fractals as models of many natural objects and systems. (Except for its Chapter 39, it is not a good book to read if you have trouble to fall sleep!)

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979), "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, 21,239-245.

> This is the paper that introduced the Latin hypercube sampling method.

Palmore, J. (1995), "Chaos, Entropy and Integrals for Discrete Dynamical Systems on Lattices," *Chaos, Solitons & Fractals*, 8, 1397-1418.

> A technical paper where one can see the connections between random number generators and fractals via dynamical systems.

Ulam, S.M (1991), *Adventures of a Mathematician*, Berkeley: University of California Press.

> An autobiography of one of the great mathematicians of the twentieth century and a key inventor of the Monte Carlo method. It also contains interesting portraits of other mathematics "giants", such as Banach, von Neumann, and Erdös.

Figure 1: The 2-D Perspective View of the Antihype Fractal with $k = 3$ on Its Sitting Plane $X_1 + X_2 + X_3 - 3/2 = 0$. (The new coordinates (X_1', X_2') are related to old coordinates (X_1, X_2, X_3) via $X_1' = (X_1 + X_2 - 2X_3)/\sqrt{6}$ and $X_2' = (X_1 - X_2)/\sqrt{2}$.)

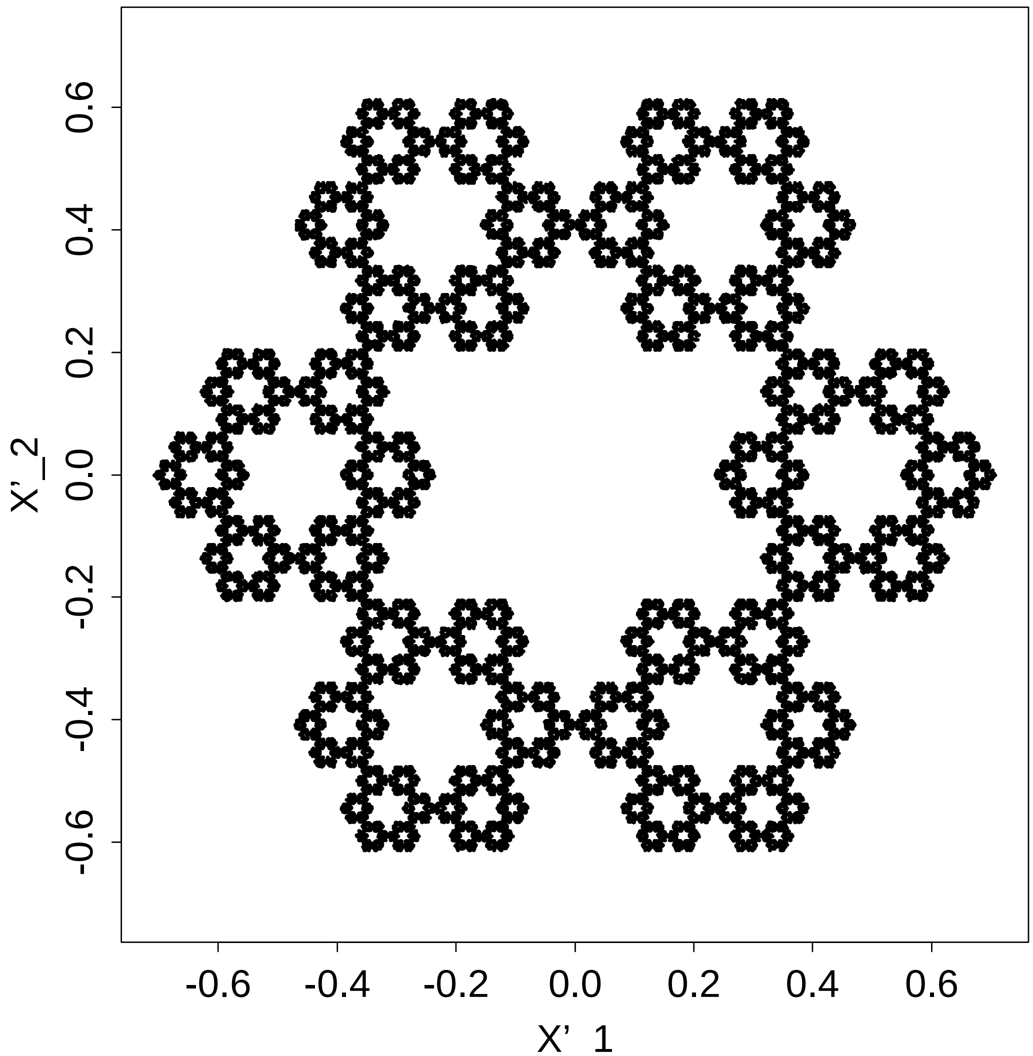

Figure 2: The 2-D Projection of the Antihype Fractal with $k = 3$.

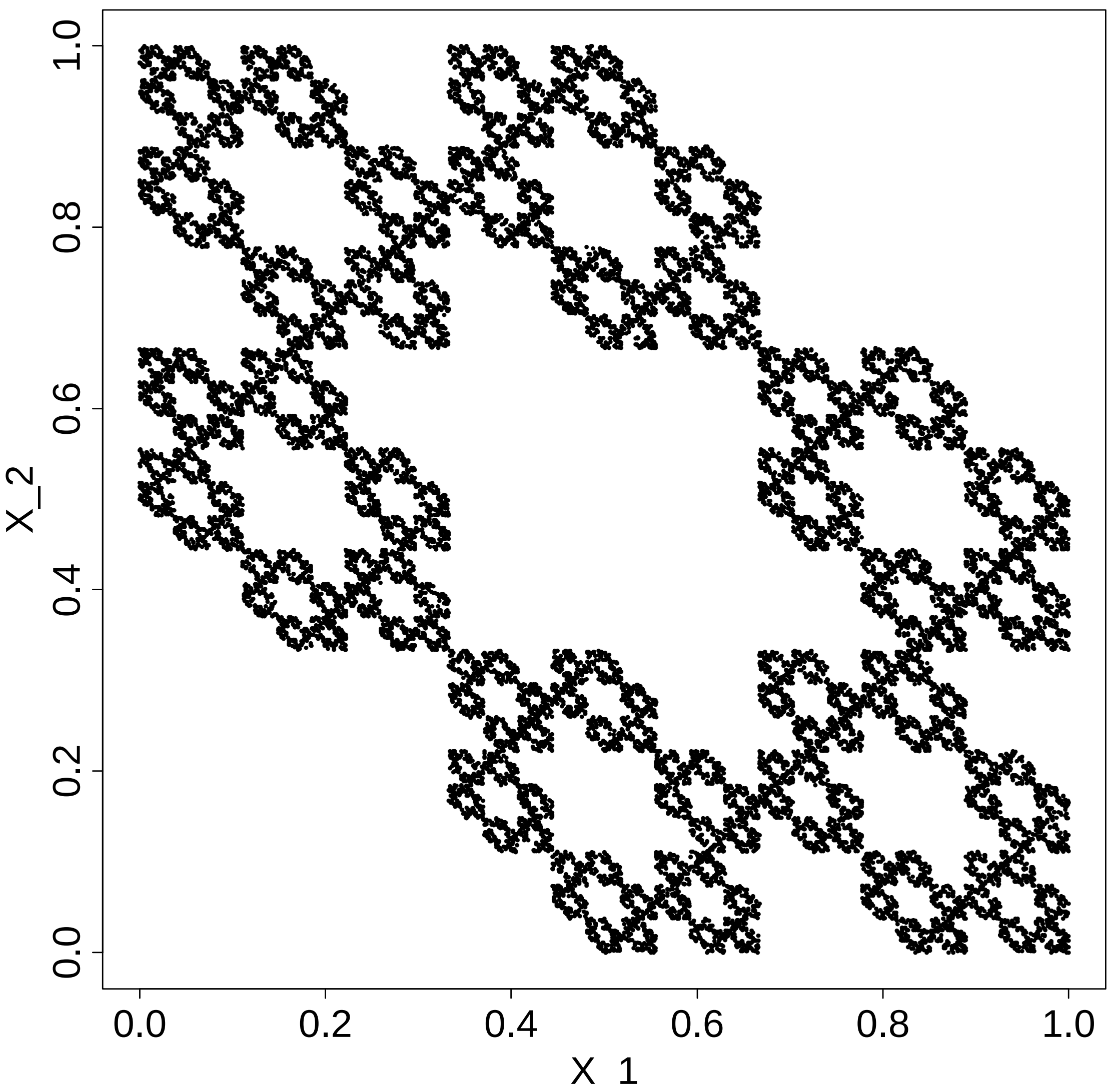

Figure 3: Sequential Elimination of Regions By the Iterative Latin Hypercube Sampling with $k = 3$ and $t = 5$.

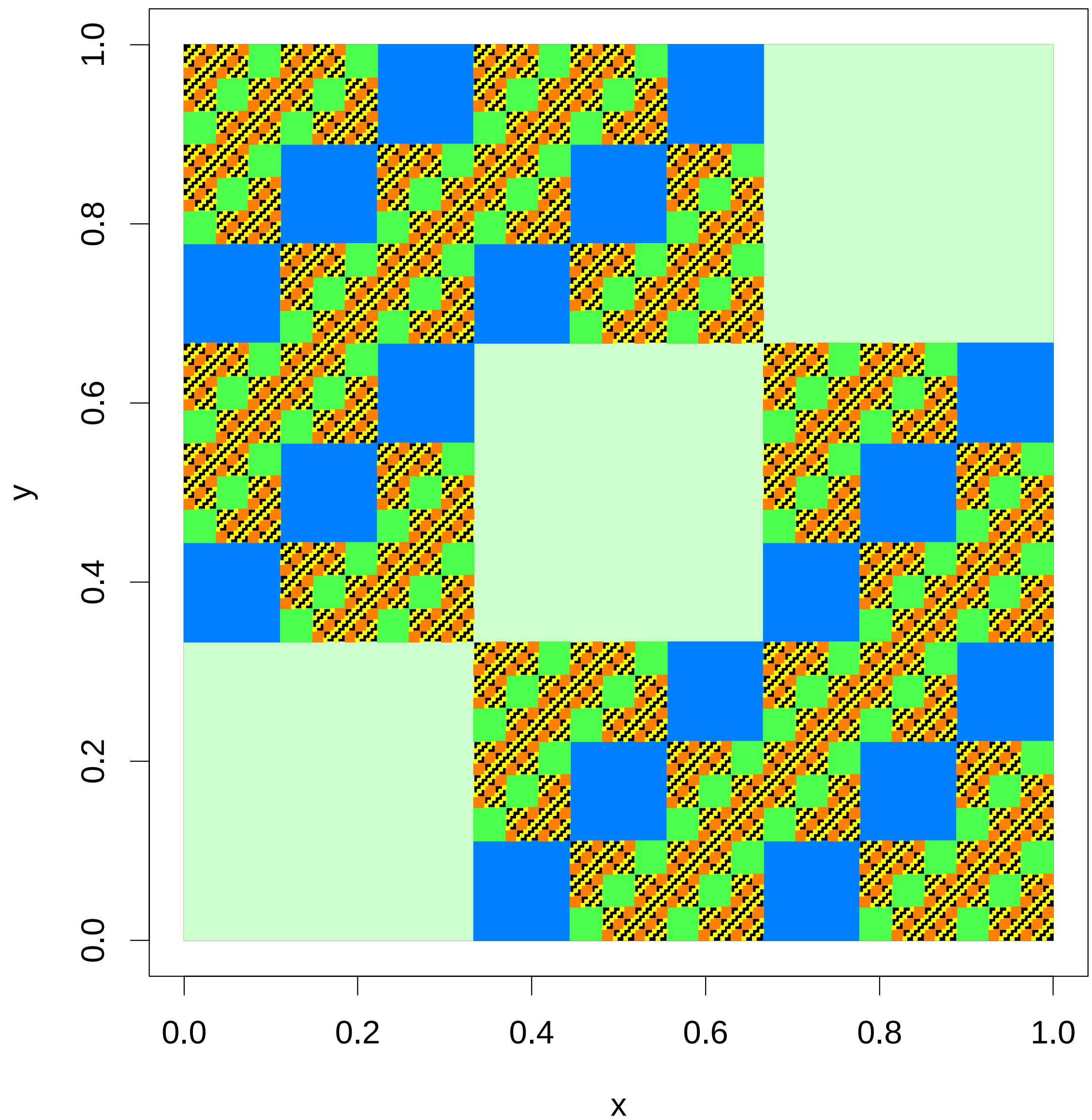

Figure 4: A New Fractal Via Rotation and Mixture.