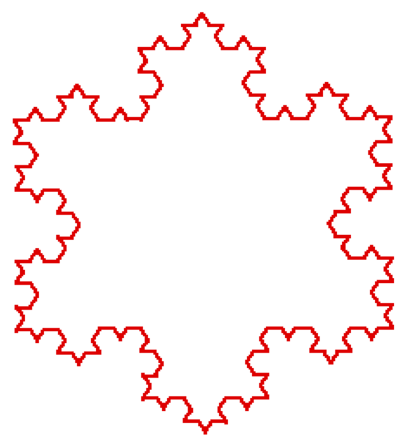
Summer 2004

When Thue-Morse Meets Koch

Jun Ma

When Thue-Morse Meets Koch

Jun Ma, Class of 2005, Kenyon College, Dr. Judy Holdener, Mathematics Department, Kenyon College



Left: The third iteration of the Koch Snowflake.

Acknowledgements: Thanks to the Kenyon College Summer Science Scholars program for providing me the research opportunity and also to Dr Judy Holdener for her guidance and support along the line of the project.

Abstract

In the project, we revealed a remarkable connection between Thue-Morse sequence and the Koch Snowflake. Using turtle geometry and polygon maps, we realized the Thue-Morse sequence as the limit of polygonal curves in the plane. We also proved that a sequence of such curves converges to the Koch snowflake in the Hausdorff metric. In the final section we considered generalized Thue-Morse sequences and provided a characterization of those that encode curves converging to the Koch snowflake

Introduction

The Thue-Morse sequence and the Koch snowflake have much in common. Both are defined iteratively. Both exhibit properties of self-similarity. Both first appeared in the early 1900's (the Koch snowflake in 1906 and the Thue-Morse sequence in 1912). And both continue to appear frequently - yet independently - in popular mathematical writing today.

The Thue-Morse sequence is a two symbol sequence typically defined by iterating a substitution map σ. Given the alphabet $A = \{a, b\}$, define the morphism $\sigma : A^* \to A^*$ by setting $\sigma(a) = ab$ and $\sigma(b) = ba$. If $\sigma^0 = a$, we see that σ generates the sequence of words:

$$\{\sigma^n(a)\}_{n\geq 0} = a,\ ab,\ abba,\ abbabaab,\ abbabaabbaababba,\ \ldots$$

This sequence converges to what is commonly known as the Thue-Morse sequence:

$$t = \lim_{n\to\infty} \sigma^n(a) = abbabaabbaababbabaababbaabbabaab\ldots.$$

A classical fractal object, the Koch snowflake was first introduced by Helge von Koch in 1906. It is constructed by starting with a line segment of unit length, extracting the middle third and replacing it with two line segments of length 1/3 (see Figure 1). The process is continued infinitely, with the middle third of any line segment at each stage being replaced with two line segments of length equal to 1/3 of the line segment.

Figure 1. Edge replacement in the Koch snowflake

The Koch snowflake is commonly defined by way of a Lindenmayer system with initial string F and rewriting rule $F \to F - F + +F - F$; "+" denotes a counterclockwise rotation of $\pi/3$ rad and "-" a clockwise rotation by the same amount. This rewriting rule can be easily translated into Turtle Program.

Turtle Geometry

A *turtle program* is defined to be any word over the alphabet $\Sigma = \{L, F\}$ where F denotes a forward motion of the turtle by one unit and L a counterclockwise rotation by some fixed angle θ. If $\theta = 2\pi/N$ (we use $\theta = \pi/3$ in our project), then the set of all words over Σ subject to the relation $R = \{L^N = \varepsilon\}$ is denoted by Σ_R^*.

A turtle state is an ordered pair $(\bar{r}, \bar{v})$ consisting of a position vector $\bar{r} \in \mathbf{R}^2$ and a unit vector $\bar{v}$ describing the turtle's heading. The command F represents the basic transformation T_F mapping the state $(\bar{r}, \bar{v})$ to the state $(\bar{r} + \bar{v}, \bar{v})$, and L represents the transformation T_L mapping $(\bar{r}, \bar{v})$ to $(\bar{r}, R_\theta \bar{v})$, where R_θ is a rotation matrix. A string w of F's and L's then describes the general turtle transformation Tw consisting of compositions of these two basic transformations in the form $Tw(\bar{r}, \bar{v}) = (\bar{r} + M\bar{v}, R\bar{v})$, where M is a matrix of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and $R = R_\theta^k$ for some positive integer k. As described in [5], the set of pairs (M, R) forms a group under the binary operation $(M_1, R_1)(M_2, R_2) = (M_1 + R_1M_2, R_1R_2)$ and there is a homomorphism $\psi : \Sigma_R^* \to G$. With this homomorphism and an initial state $(\bar{r}_0, \bar{v}_0)$, we can define a position homomorphism $g : \Sigma_R^* \xrightarrow{\psi} M_2(\mathbb{R}) \times \langle R_\theta \rangle \xrightarrow{\pi_1} M_2(\mathbb{R}) \xrightarrow{\phi_0} \mathbb{R}^2$ where $\Phi_0 : M_2(\mathbf{R}) \to \mathbf{R_2}$ is defined by $\Phi_0(M) = M\bar{v}_0$.

Thue-Morse Turtle Programs

Thue-Morse turtle programs of degree n, denoted by TM_n and $\overline{TM}_n$, are defined to be the words in Σ_R^* : $TM_n = \sigma^n(F)$ and $TM_n = \sigma^n(L)$. Their trajectories turn out to be very interesting.

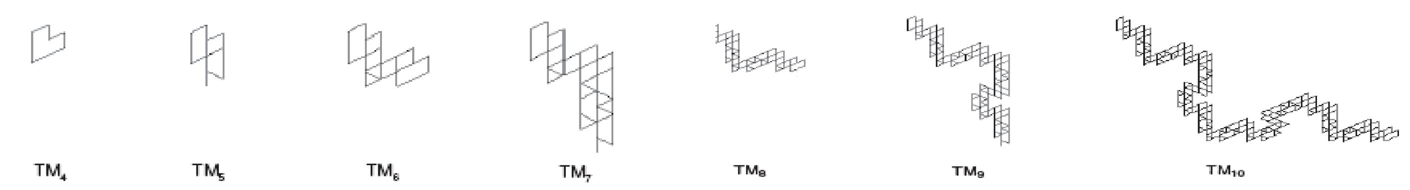


Figure 2. Trajectories of Thue-Morse turtle programs of degrees 4 through 10

Indeed, the trajectories corresponding to the even terms of the Thue-Morse sequence are starting to resemble the familiar Koch snowflake! Skeptical? Consider TM_{14}.

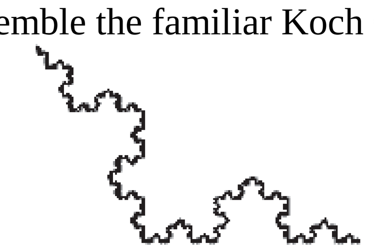


Figure 3. Trajectory of the Thue-Morse turtle program of degree 14

The Polygon Map

Our next goal is to realize turtle programs (in particular Thue-Morse turtle programs) as polygonal curves in the plane. Let S be a subset of Σ_R^* and let $H(\mathbf{R}^2)$ be the set of nonempty compact subsets of $\mathbf{R}^2$. Equipped with the position homomorphism $g : \Sigma_R^* \to \mathbf{R}^2$, we can then define a map $K[\bullet] : S \to H(\mathbf{R}^2)$ that assigns a polygon to each word over S. As F. M. Dekking presents in [[4], p. 80], define the polygon map $K[\bullet]$ on $s \in S$ to be $K[s] = \{\alpha\, g(s) : 0 \leq \alpha \leq 1\}$, and extend the map to all words over S by requiring that $K[VW] = K[V]\ U\ (g(V) \oplus K[W])$ for any $V, W \in S^*$ where $g(V) \oplus K[W] = g(V) + \pi_2 \circ \psi(V) \times K[W]$.

As indicated above, there can be many different ways of realizing a turtle program in the plane. The most general polygon map is obtained by defining $S = \Sigma^*$. In this case, $\mathrm{Kturt} : \Sigma^* \to \mathbf{R}^2$ is the map one typically associates with turtle geometry, assigning to each word $w = a_1a_2a_3\ldots a_k \in \Sigma^*$ the trajectory traversed by a turtle that follows each command $a_i \in \{F, L\}$, in turn, starting with a_1 and finishing with a_k. For all integers $k \geq 0$, if $\Omega_{2k} = \{TM_{2k}, \overline{TM}_{2k}\}$ then the polygon maps $\{K_{2k} : \Omega_{2k}^* \to H(\mathbf{R}^2)\}$ produce polygonal curves constructed out of the basic component edges $K_{2k}[TM_{2k}]$ and $K_{2k}[TM_{2k}]$. As illustrated in Figure 4, these polygon maps play a critical role in the proofs of the convergence theorems.

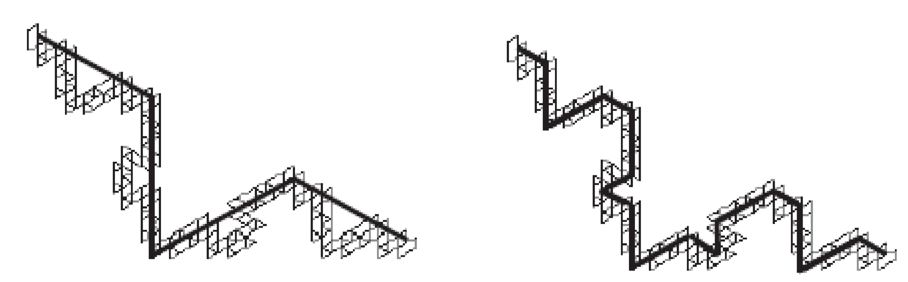


Figure 4. Left: $K_6[TM_{10}]$ overlaying $K_0[TM10] = Kturt[TM_{10}]$
Right: $K_8[TM_{10}]$ overlaying $K_0[TM_{10}] = Kturt[TM_{10}]$

Hausdorff Metric

The distance between two subsets of a metric space is defined using the Hausdorff metric (see [2]). Given the complete metric space $\mathbf{R}^2$ under the Euclidean metric d and $H(\mathbf{R}^2)$, the space of nonempty compact subsets of $\mathbf{R}^2$, the Hausdorff distance between two points $A, B \in H(\mathbf{R}^2)$ is defined by

$$h(A, B) = d(A, B) \vee d(B, A)$$

where $d(A, B)$ is the Euclidean distance between two sets:

$$d(A, B) = max\{d(x, B) : x \in A\},$$

and $d(A, B) \vee d(B, A)$ denotes $max\{d(A, B), d(B, A)\}$.

A particularly simple situation is the case where the two sets are parallel line segments $\overline{AB}$ and $\overline{CD}$. By considering two simple cases, the Hausdorff distance $h(\overline{AB}, \overline{CD})$ is computed easily using the (Euclidean) distances between the endpoints of the parallel line segments.

Figure 5. Case i (left): The projection of one line segment is contained in the other. Case ii (right): The projection of one line segment is not contained in the other. In either cases, we have
$h(\overline{AB}, \overline{CD}) = d(A, C) \vee d(B, D)$.

Main Results

Theorem 5.0.13, 5.0.14 (Convergence Theorem I, II) For positive integer $n \geq 5$, let k_n be ½ n if n is even, and be (n+1)/2 if n is odd. Then the sequence of compact sets $\{S_nK_{2kn}[TM_{2n}]\}_{n\geq 5}$ where S_n is the corresponding scaling factor converges to the Koch Snowflake in the Hausdorff metric.

Generalization

If $w, w' \in \Sigma_R^*$ satisfying the following two properties(**Theorem 6.0.18**):

(1) $|w|_L = |w'|_L = \pm 2 \bmod 6$ ($|w|_L$ denotes the number of L's in w)

(2) $\overline{g(w)}$ and $\overline{g(w')}$ lay along the same line but in the opposite direction,

and σ is the same substitution map defined in the introduction, a generalized Thue-Morse sequence is defined to be the limit $\lim_{n\to\infty} \sigma^{2n}(w)$. Following figures show two interesting examples:

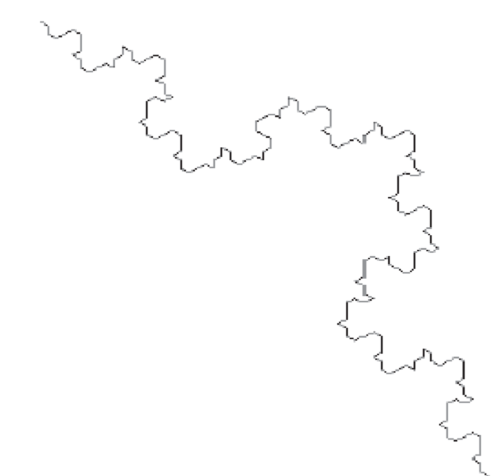


Figure 6. The polygon $Kturt\ [\sigma^6(w)]$, where $w = LFL^5FLFL$ and $w' = L^3FLF^2L^4$

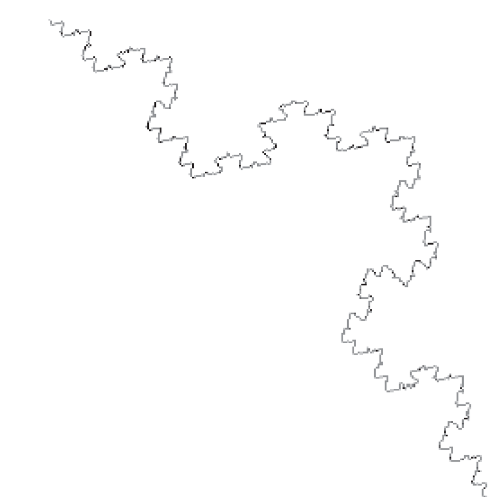


Figure 7. The polygon $Kturt\ [\sigma^8(w)]$, where $w = LFL^5FLFL$ and $w' = L^3FLF^2L^5FLF^2L$

Conclusion and Further Research

The Convergence Theorems prove that with the use of sufficiently coarse polygon maps K_{2k} the Thue-Morse sequence does indeed encode the Koch snowflake. Generalization shows that there are, in fact, many pairs $\{w, w'\}$ that encode the Koch snowflake under iteration of the substitution map σ^2. One simply needs to define w and w' to be such that they satisfy properties (1) and (2) of **Theorem 6.0.18**.

It is worth noting, however, that in the original Thue-Morse sequence, w' is closely linked to w. In particular, $w' = \bar{w}$ is obtained from w by changing all F's to L's and all L's to F's. This leads us to further research opportunity: Is it possible to find a $w \in \Sigma_R^*$ that is different from TM_{2n} or $\overline{TM}_{2n}$ such that the pair $\{w, w'\} = \{w, \bar{w}\}$ generates turtle programs $\{\sigma^{2n}(w)\}_{n\geq 0}$ that encode turtle trajectories converging to the Koch snowflake?

This, it appears, is a much more difficult question to answer. We have not been able to find such a w, nor have we proved that one does not exist. What is clear, however, is that such a w would have to meet a much more stringent set of criteria. If the Thue-Morse turtle programs TM_{2n} and $\overline{TM}_{2n}$ were the only words of the form $\{w, \bar{w}\}$ encoding the Koch snowflake, then this would establish an even tighter link between the Thue-Morse sequence and the Koch snowflake.

References

1. H. Abelson and A.A. diSessa, *Turtle geometry*, MIT Press Series in Artificial Intelligence, MIT Press, Massachusetts, 1981.
2. M. Barnsley, *Fractals everywhere*, Academic Press, Inc., 1988.
3. J. Berstel, *Axel Thue's papers on repetitions in words: a translation*, Publications du LACIM, Universite du Quebec a Montreal, No. 20,1995.
4. F. M. Dekking, *Recurrent sets*, Advances in Math. 44 (1982), 78-104.
5. J. Holdener and A. Wagaman, *A classification of periodic turtle sequences*, International Journal of Mathematics and Mathematical Sciences 2003 (2003), no. 34, 2193-2201.
6. M. Lothaire, *Combinatorics on words*, Addison-Wesley, 1983.
7. A. Luca and S. Varricchio, *Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups*, Theoretical Computer Science 63 (1989), 333-348.
8. M. Morse, *Recurrent geodesics on a surface of negative curvature*, Transactions Amer. Math. Soc. 22 (1921), 84-100.
9. M. Morse and G. Hedlund, *Unending chess, symbolic dynamics and a problem in semigroups*, Duke Math. J. 11 (1944), 1-7.
10. J.J. Pansiot, *The Morse sequence and iterated morphisms*, Inform. Process Lett. 12 (1981), 68-70.
11. P. Prusinkiewicz and A. Lindenmayer, *The algorithmic beauty of plants*, Springer-Verlag, 1990.
12. A. Thue, *Ä Uber die gegenseitige lage gleicher teile gewisser zeichenreihen*, Kra. Vidensk. Selsk. Skrifter. I. Mat.-Nat. Kl., Christiana 1912, Nr. 10.
13. H. von Koch, *Une methode geometrique elementaire pour l'etude de certaines questions de la theorie des courbes planes*, Acta Math. 30 (1906), 145-174.