Application of Gray code to the cryptanalysis of chaotic cryptosystems

GONZALO ALVAREZ, DAVID ARROYO, JUANA NUNEZ

***Abstract*— Gray codes have found many applications in engineering and mathematics. In this work we explain how to apply Gray codes to the determination of the control parameter and initial point of a chaotic orbit generated by the Mandelbrot map when the kneading sequence is known. Possible applications to the cryptanalysis of a certain type of cryptosystems based on Baptista's algorithm are also discussed.**

I. Introduction

Gray codes (or reflected binary codes) have been extensively used in engineering, telecommunications, genetics, and even mathematical puzzles. In short, the Gray code is a binary numeral system where two successive values differ in only one digit. In [1], the kneading theory on symbolic sequences, which manipulates symbols, was translated to number theory by a transformation into Gray codes. It was showed how the symbolic sequences of a 1-D quadratic map are ordered according to the Gray code for a given parameter value and different initial points as well as for a given initial point and different parameter values. The Gray code was then generalised to introduce the Gray Ordering Number (GON) -in the interval (0, 1) - allowing the simultaneous ordering of different size symbolic sequences. The introduction of Gray codes and Gray numbers in 1D quadratic maps is highly beneficial from the computing viewpoint, since the handling of symbolic sequences is substituted by the use of numbers.

From other viewpoint, in [2] the Gray codes were applied to the cryptanalysis of the Baptista cipher [3]. It was shown that given a symbolic sequence and the initial condition, the map's parameter value can be obtained. Likewise, given a symbolic sequence and the map's parameter value the initial condition can also be obtained. This result has a great importance in cryptanalysis, since many cryptosystems use as secret key these two values: initial condition and parameter.

In this presentation and starting from the preliminary work by [4], we explain how these results can be taken further and provide different ways to estimate the secret key from just the symbolic sequence. We show the limitations inherent to this method and in which cases it is best applied. These results show that in the design of a cryptosystem based on 1D-quadratic maps, it is all important to conceal the underlying symbolic sequence. We also show the quality of initial condition and parameter estimates when only partial information about the symbolic sequence is available.

Gonzalo Alvarez, David Arroyo and Juana Nunez are with the Instituto de Fisica Aplicada, Consejo Superior de Investigaciones Cientificas, Serrano 144, 28006 Madrid, Spain; gonzalo@iec.csic.es

II. Gray codes and the Mandelbrot map

The Mandelbrot map is given by

$$x_{n+1} = f_c(x_n) = x_n^2 + c. \tag{1}$$

If $x_0 = x$ is a certain initial value from which a set of N values are calculated using the previous equation and a certain value for c, the pattern or kneading sequence associated to that sequence of numbers is

$$P_{f_c}^N(x) = p_0 p_1 \dots p_{N-1}, \tag{2}$$

where

$$p_i = \begin{cases} L & \text{if} \quad x_i < 0, \\ C & \text{if} \quad x_i = 0, \\ R & \text{if} \quad x_i > 0, \end{cases} \tag{3}$$

for $i = 0, 1, \dots, N-1$. If the first symbol of $P_{f_c}^{N+1}(x)$ is not considered, we have

$$\zeta_{f_c}^N(x) = p_1 p_2 \dots p_N. \tag{4}$$

Similarly, from $P_{f_c}^{N+m}(x)$:

$$\zeta_{f_c}^{N,m}(x) = p_m p_{m+1} \dots p_{N+m-1}. \tag{5}$$

The Gray code derived from the pattern $P_{f_c}^N(x)$ is

$$G(P_{f_c}^N(x)) = g_0 g_1 \dots g_{N-1}, \tag{6}$$

where

$$g_i = \begin{cases} 0 & \text{if} \quad p_i = R, \\ 1 & \text{if} \quad p_i \in \{L, C\}, \end{cases} \tag{7}$$

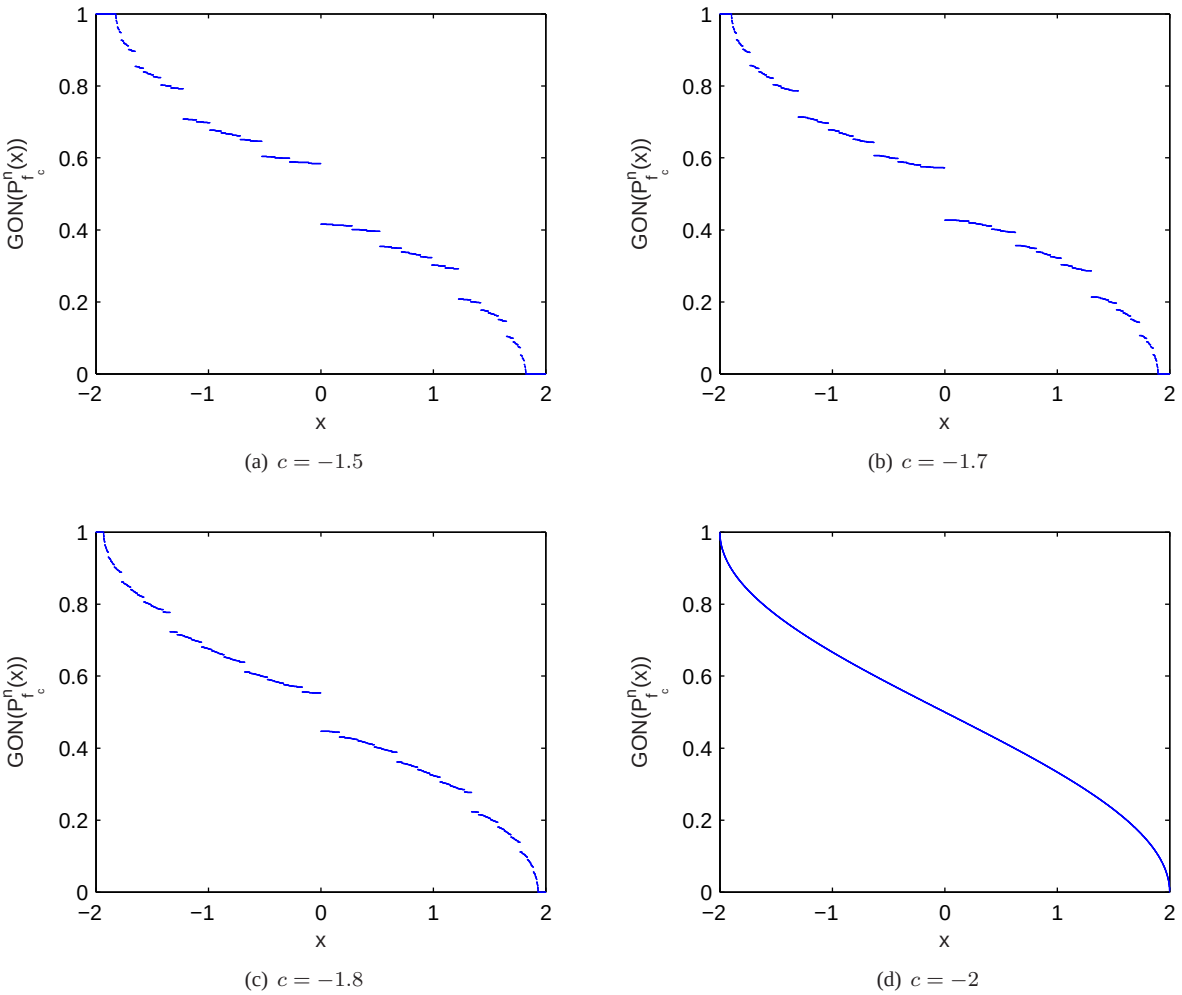
for $i = 0, 1, \dots, N-1$. The binary number corresponding to $G(P_{f_c}^N(x))$ is $U(P_{f_c}^N(x)) = u_0 u_1 \dots u_{N-1}$, and it is calculated as

$$u_i = \begin{cases} g_i & \text{if} \quad i = 0, \\ u_{i-1} \oplus g_i & \text{if} \quad i > 0, \end{cases} \tag{8}$$

Finally, the Gray Ordering Number or GON is defined as follows:

$$GON(P_{f_c}^N(x)) = 2^{-1} \cdot u_0 + 2^{-2} \cdot u_1 + \dots + 2^{-N+1} \cdot u_{N-1}. \tag{9}$$

In [5] it was proved that for an unimodal, convex function $f(x)$, the kneading sequences derived from the function iteration and different initial conditions are sorted. Moreover, in [5] and [6] it was shown the existence of an order such that if one kneading sequence is greater than other, it means that it has been generated from a greater initial condition. In the case of the Mandelbrot map, the iteration function is $f_c(x)$. This function is unimodal but concave and, consequently, a kneading sequence greater than other is a kneading sequence



(a) $c = -1.5$ (b) $c = -1.7$

(c) $c = -1.8$ (d) $c = -2$

Fig. 1. GON versus initial condition for different c values and kneading sequences of 16 symbols.

created iterating $f_c(x)$ from an smaller initial condition. In [1] it was remarked that the mentioned order is analogue to the one corresponding to the Gray codes. Therefore, it is possible to consider the kneading sequences as Gray codes and assign to each pattern a numerical value, which is its GON. Having in mind the previous considerations, the GON for the Mandelbrot map is expected to be an increasing function with respect to the initial condition. In order to verify this inference, the GON of kneading sequences of length $N = 16$, with different values for c and increasing initial condition was calculated. Figure 1 depicts these values and it confirms the expected behavior of the GON.

III. Gray codes and cryptanalysis

This relationship between the GON number and the chaotic map's control parameter and initial condition suggests that there might exist weaknesses in encryption algorithms based upon the iteration of chaotic maps, if and only if the GON number can be computed from the ciphertext or in some other way. Most cryptosystems based on the iteration of a discrete chaotic map use the initial condition and parameter as part of the secret key. In this section, the method to compute these two values from the knowledge of the underlying kneading sequence is explained. First, the control parameter is estimated with no other information than the sequence generated by the iterates of the chaotic map. Next, the initial condition is estimated. If both values are estimated within the same precision used by the computer during the computations, then the secret key is obtained.

A. Parameter estimation

So far we know that the kneading sequences generated by $f_c(x)$ are sorted in the Gray meaning. For kneading sequences of length N, the definition interval, i.e., $[c, -c]$ can be divided in $2N$ subintervals. All the values in one of those subintervals generate the same kneading sequence, and the kneading sequences associated to consecutive subintervals only differ in one symbol. Finally, the kneading sequences of the different subintervals increase from right to left. Since the minimum value of $f_c(x)$ is reached when $x = 0$, the maximum GON value for a given c is the one generated from $x = c$. In other words, $GON(\zeta_{f_c}^N(0))$ is the maximum of the function $GON(P_{f_c}^N(x))$ for $x \in \{c, -c\}$. This circumstance is illustrated in Figs. 2 and 3. Figure 2 was created according

to Algorithm 1 for $\Delta c = 10^{-4}$, $\Delta x = 10^{-4}$ and $N = 16$.

Algorithm 1: GON bifurcation diagram for the Mandelbrot map

Input: Δc
Increment value for the parameter c
Input: Δx
Increment value for the initial condition from which the kneading sequences will be generated
Input: m
Number of values that are not going to be considered in the GON evaluation process
Input: N
Kneading sequences length
Do $c = -2$
while $c \leq -1$ **do**
    $x = c + \Delta x$
    while $x < -c$ **do**
        Plot $GON(\zeta_{f_c}^{N,m}(x))$
        $x = x + \Delta x$
    end
end

The upper bound of the bifurcation diagram referred by this graphic is the function depicted in Fig. 3, i.e., $GON(\zeta_{f_c}^{N}(0))$. In [4] this fact is exploited in order to get the c value responsible of a given kneading sequence (see Algorithm 2). First of all, it is necessary to approximate $GON(\zeta_{f_c}^{N}(0))$. To do so, M subsequences are generated from the original one. If $S = s_0 s_1 \dots s_K$ is the original kneading sequence, the first subsequence is $S_1 = s_0 s_1 \dots s_{N-1}$, the second one $S_2 = s_1 s_2 \dots s_N$, the third one $S_3 = s_2 s_3 \dots s_{N+1}$, ..., the Mth subsequence is $S_M = s_{M-1} s_M \dots s_{M+N-2}$. $GON(\zeta_{f_c}^{N}(0))$ is going to be considered as the maximum value of $GON(S_i)$ for $1 \leq i \leq M$. The rest of the searching procedure appears in Algorithm 2.

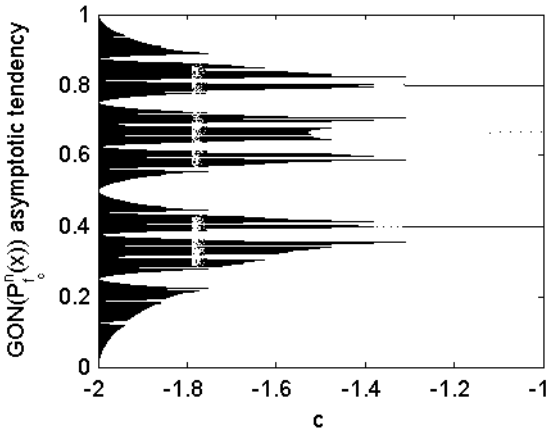


Fig. 2. Asymptotic tendency for the GON of the Mandelbrot map.

However, this algorithm presents some limitations. First

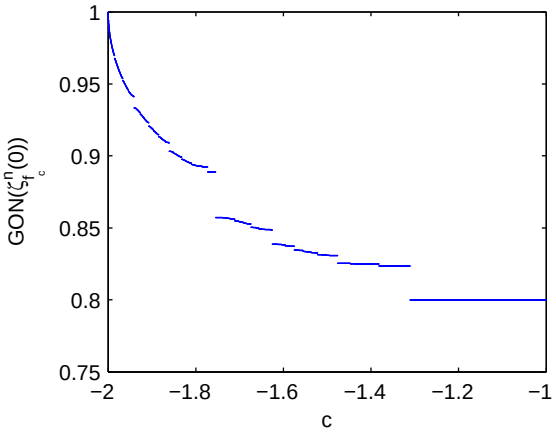


Fig. 3. GON for initial condition equal to 0 and different values for c. The kneading sequences considered are 16-symbol long.

Algorithm 2: Estimation of the parameter c for the Mandelbrot map from a given kneading sequence

Input: S
Symbolic sequence generated being c the value we want to get. S has to be $M + N$ symbol-length, at least
Input: M
Number of subsequences used in the searching for the maximum GON value
Input: N
Subsequences length
Calculate $GON\left(\zeta^{N,m}(S)\right)$ for $m = 1, \dots, M$
Determine $G_{\max} = \max\{GON\left(\zeta^{N,m}(S)\right)\}$, $m = 1, \dots, M$
Do $c_L = -2$, $c_R = -0.8$
Do $c = \frac{c_R + c_L}{2}$ and $G = GON(\zeta_{f_c}^{N}(0)) = GON(P_{f_c}^{N}(c))$
if $G > G_{\max}$ **then**
    Do $c_L = c$ and go to 7
else if $G = G_{\max}$ **then**
    Return c
else
    Do $c_R = c$ and go to 7
end

of all, when dealing with the GON and a finite precision machine, it is not possible to recover the exact value of c even if the exact value of $GON(\zeta_{f_c}^{N}(0))$ is known. For example, let us consider a kneading sequence generated from $x = 0$ and $c = -1.94798231759013328$. It was verified that the error estimation is never zero. When working with the Gray codes instead of the GON it is possible to recover the exact value of c, if the kneading sequences length, N, is big enough (see Fig. 4). Therefore, when working in a finite precision environment, it is better to deal with Gray codes instead of their numerical version. On the other hand, a perfect estimation of the sequence $\zeta_{f_c}^{N}(0)$ is not possible just looking through a certain kneading sequence S. In [4] it

is said that a better estimation of the parameter c is possible by increasing the parameters M and N. However, it only would be possible if the given input kneading sequence S contains the maximum kneading sequence, i.e., $\zeta_{f_c}^N(0)$. This happens only if the orbit generated by iterating $f_c(x)$ from a certain initial condition and using a certain value for c, contains the value 0. The probability of this fact occurring is zero. Hence, Algorithm 2 gives an estimation of the parameter c and it is not possible to improve it unless a new way to get $\zeta_{f_c}^N(0)$ from the input kneading sequence S is found. Figure 5 shows the result of the application of Algorithm 2 when the kneading sequence is generated from $c = -1.94798231759013328$ and initial condition $x_0 = -1.3807041990568$. In this simulation the length of subsequences is $N = 60$. The number of subsequences used in the maximum subsequence searching (M) was increased from 10^3 to 10^6. It was observed that the estimation error decreases as the parameter M increases: v.g. for $M \approx 10^7$ an estimation error around 10^{-15} is expected. From a cryptanalysis practical point of view, in those situations where it is not possible to get such a long kneading sequence, a new way to approximate the maximum subsequence from the given kneading sequence would be needed. This will be one of the future works to be done.

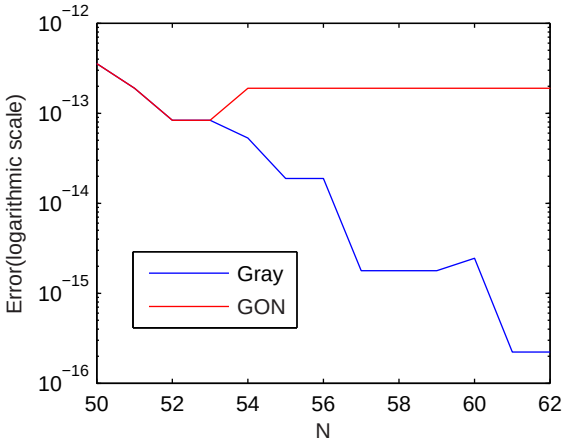


Fig. 4. Error in the estimation of c working with the GON and Gray codes.

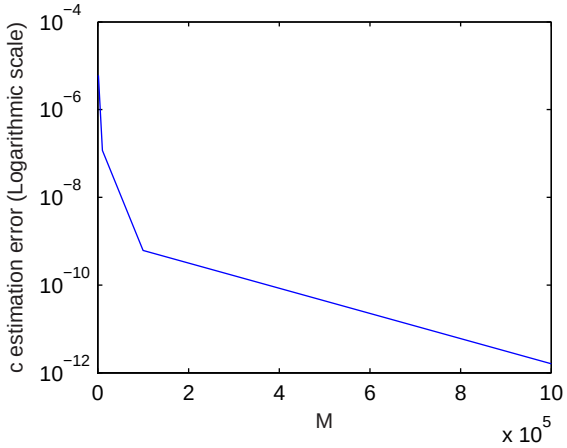


Fig. 5. Estimation error of c for $c = -1.94798231759013328$, $x_0 = -1.3807041990568$ and a kneading sequence length $N = 60$.

B. Initial condition determination

When knowing the parameter c that has been used in the generation of a certain kneading sequence, which is also known, then it is possible to recover the initial condition from which this kneading sequence was created. In fact, Fig. 1 shows how the GON increases as the initial condition decreases. It means that it is possible again to build a bisearching algorithm as the one introduced in the previous subsection. We are going to try a certain value x as the one used in the generation of the given kneading sequence. We calculate the kneading sequence for this candidate initial condition and the input c value. If the resulting kneading sequence is greater than the given one, it means we have to increase x. On the other hand, if the resulting kneading sequence is smaller than the given one, we have to decrease the candidate initial condition. The detailed method is described by Algorithm 3.

In order to verify this algorithm, several simulations were done. In Figs. 6, 7 and 8 the error in the initial condition recovering process is represented. We can verify how the error decreases as the length of the kneading sequence (N) increases. In this way, for all the different configurations, a length of the input kneading sequence greater than 80 leads to an error equal to zero or around 10^{-16}. Since all the mathematical operations were done dealing with 16 decimal values, we can say that the method described in this section allows to get the initial condition corresponding to a certain kneading sequence and a c value.

Algorithm 3: Estimation of the initial condition corresponding to the input value c and a given kneading sequence for the Mandelbrot map

Input: c

Control parameter that leads to a certain kneading sequence S from the initial condition we are looking for

Input: GON_{input}

GON of the kneading sequence S resulting of the iteration of $f_c(x)$ using the input value c and the initial condition we want to estimate

Do $x_R = -c$, $x_L = c$

Do $x = \frac{x_R + x_L}{2}$

Calculate $GON(P^n_{f_c}(x))$

if $GON(P^n_{f_c}(x)) < GON_{input}$ **then**

    Do $x_R = x$ and go to 4

else if $GON(P^n_{f_c}(x)) = GON_{input}$ **then**

    Return x

else

    Do $x_L = x$ and go to 4

end

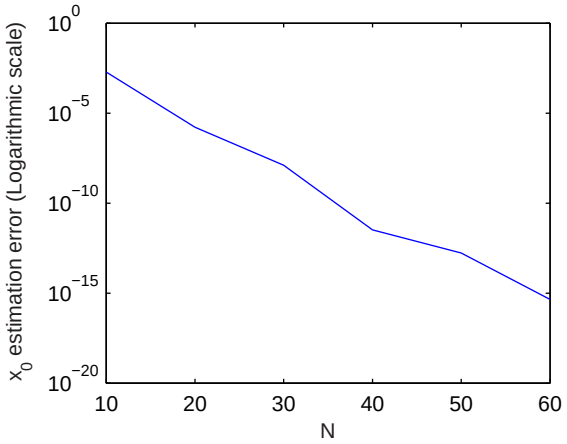


Fig. 6. Error of the initial condition x_0 used in the generation of a given kneading sequence. $c = -1.93824712239432$, $x_0 = -1.25495242770985$. N is the length of the input kneading sequence.

IV. Conclusions

In this presentation we have shown how the symbolic dynamics of the unimodal maps is a very useful tool to estimate the parameter and the initial condition that lead to a certain kneading sequence. We have remarked that given a kneading sequence it is possible to estimate the control parameter from which it was derived. However, it was mentioned that it is necessary to look for a new way to initialize the parameter estimation algorithm, since a perfect estimation requires a very long kneading sequence. The higher performance of the estimation method using Gray codes instead of GON numbers has also been shown. We have also shown how it is possible to recover the initial condition used in the generation of a kneading sequence just knowing the kneading sequence and the control parameter value.

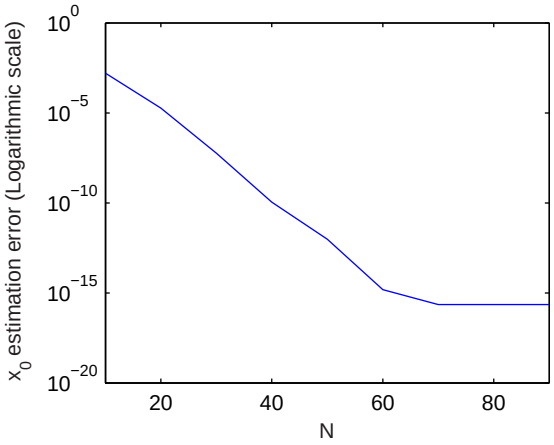


Fig. 7. Error of the initial condition x_0 used in the generation of a given kneading sequence. $c = -1.93824712239432$, $x_0 = 0.346670564996533$. N is the length of the input kneading sequence.

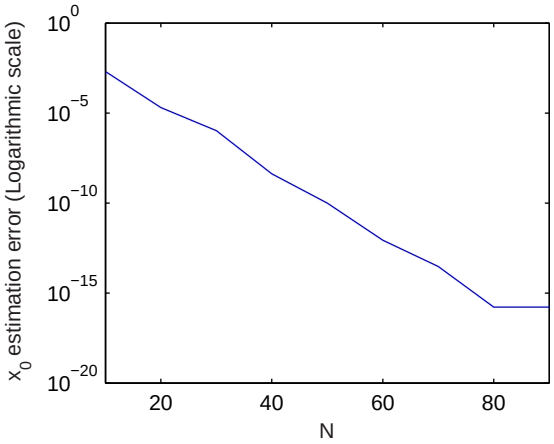


Fig. 8. Error of the initial condition x_0 used in the generation of a given kneading sequence. $c = -1.83824712239432$, $x_0 = 0.346670564996533$. N is the length of the input kneading sequence.

References

[1] G. Alvarez, M. Romera, G. Pastor, and F. Montoya, "Gray codes and 1d quadratic maps," *Electronic Letters*, vol. 34, no. 13, pp. 1304–1306, 1998.

[2] G. Alvarez, F. Montoya, M. Romera, and G. Pastor, "Cryptanalysis of an ergodic chaotic cipher," *Physics Letters A*, vol. 311, pp. 172–179, 2003.

[3] M. S. Baptista, "Cryptography with chaos," *Physics Letters A*, vol. 240, no. 1-2, pp. 50–54, 1998.

[4] X. Wu, H. Hu, and B. Zhang, "Parameter estimation only from the symbolic sequences generated by chaos system," *Chaos, solitons and Fractals*, vol. 22, pp. 359–366, 2004.

[5] N. Metropolis, M. Stein, and P. Stein, "On the limit sets for transformations on the unit interval," *Journal of Combinatorial Theory (A)*, vol. 15, pp. 25–44, 1973.

[6] L. Wang and N. D. Kazarinoff, "On the universal sequence generated by a class of unimodal functions," *Journal of Combinatorial Theory, Series A*, vol. 46, pp. 39–49, 1987.