A topological chaos framework for hash functions

Abstract

This paper presents a new procedure of generating hash functions which can be evaluated using some mathematical tools. This procedure is based on discrete chaotic iterations.

First, it is mathematically proven, that these discrete chaotic iterations can be considered as a particular case of topological chaos. Then, the process of generating hash function based on the topological chaos is detailed. Finally it is shown how some tools coming from the domain of topological chaos can be used to measure quantitatively and qualitatively some desirable properties for hash functions. An illustration example is detailed in order to show how one can create hash functions using our theoretical study.

Key-words : Discrete chaotic iterations. Topological chaos. Hash function

I. Introduction

Hash functions, such as MD5 or SHA-256, can be described by discrete iterations on a finite set. In this paper, the elements of this finite set are called cells. These cells represent the blocks of the text to which the hash function will be applied. The origin of this study goes up with the idea of using the concept of discrete *chaotic* iterations for generating new hash functions. This idea gave then quickly rise to the question of knowing if discrete chaotic iterations really generate chaos.

This article presents the research results related to this question.

First, we prove that under some conditions, discrete chaotic iterations produce chaos, precisely, they produce topological chaos in the sense of Devaney. This topological chaos is a rigorous framework well studied in the field of mathematical theory of chaos. Thanks to this result we give a process of generating hash functions.

Behind the theoretical interest connecting the field of the chaotic discrete iterations and the one of topological chaos, our study gives a framework making it possible to create hash functions that can be mathematically evaluated and compared.

Indeed, some required qualities for hash functions such as strong sensitivity to the original text, resistance to collisions and unpredictability can be mathematically described by notions from the theory of topological chaos, namely, sensitivity, transitivity, entropy and expansivity. These concepts are approached but non deepened in this article. More detailed studies will be carried out in forthcoming articles.

This study is the first of a series we intend to carry out. We think that the mathematical framework in which we are placed offers interesting new tools allowing the conception, the comparison and the evaluation of new methods of encryption in general, not only hash functions.

The rest of the paper is organized as follows.
The first next section is devoted to some recalls on two distinct domains, the domain of topological chaos and the domain of discrete chaotic iterations.
Third and fourth sections constitute the theoretical study of the present paper. Section III defines the

1

topological framework in which we are placed while section IV shows that the chaotic iterations produce a topological chaos.

The following section details, using an illustration example, the procedure to build hash functions based on our theoretical results. Section VI explains how quantitative measures could be obtained for hash functions. The paper ends by some discussions and future work.

II. Basic recalls

This section is devoted to basic definitions and terminologies in the field of topological chaos and in the one of chaotic iterations.

1 Devaney's chaotic dynamical systems

Consider a metric space $(\mathcal{X}, d)$, and a continuous function $f : \mathcal{X} \longrightarrow \mathcal{X}$.

Definition 1 f is said to be *topologically transitive* if, for any pair of open sets $U, V \subset \mathcal{X}$, there exists $k > 0$ such that $f^k(U) \cap V \neq \varnothing$.

Definition 2 An element (a point) x is a periodic element (point) for f of period $n \in \mathbb{N}$, if $f^n(x) = x$. The set of periodic points of f is denoted $Per(f)$.

Definition 3 $(\mathcal{X}, f)$ is said to be *regular* if the set of periodic points is dense in $\mathcal{X}$,

$$\forall x \in \mathcal{X}, \forall \varepsilon > 0, \exists p \in Per(f), d(x, p) \leqslant \varepsilon.$$

Definition 4 f has *sensitive dependence on initial conditions* if there exists $\delta > 0$ such that, for any $x \in \mathcal{X}$ and any neighborhood V of x, there exists $y \in V$ and $n \geqslant 0$ such that $|f^n(x) - f^n(y)| > \delta$.

δ is called the *constant of sensitivity* of f.

Let us now recall the definition of a chaotic topological system, in the sense of Devaney [4] :

Definition 5 $f : \mathcal{X} \longrightarrow \mathcal{X}$ is said to be *chaotic* on $\mathcal{X}$ if,

1. f has sensitive dependence on initial conditions,

2. f is topologically transitive,

3. $(\mathcal{X}, f)$ is regular.

Therefore, quoting Robert Devaney: "A chaotic map possesses three ingredients: unpredictability, indecomposability, and an element of regularity. A chaotic system is unpredictable because of the sensitive dependence on initial conditions. It cannot be broken down or decomposed into two subsystems, because of topological transitivity. And, in the midst of this random behavior, we nevertheless have an element of regularity, namely the periodic points which are dense."

Banks *et al.* proved in [2] that sensitive dependence is a consequence of being regular and topologically transitive.

2   Chaotic iterations

In the sequel $s[n]$ denotes the $n-$th term of a sequence s, V_i denotes the $i-$th component of a vector V, and f^k denotes the $k-$th composition of a function f. Finally, the following notation is used: $[\![1; N]\!] = \{1, 2, \ldots, N\}$.

Let us consider a *system* of a finite number N of *cells* so that each cell has a boolean *state*. Then a sequence of length N of boolean states of the cells corresponds to a particular *state of the system*.

A *strategy* corresponds to a sequence of $[\![1; \mathsf{N}]\!]$. The set of all strategies is denoted by $\mathcal{S}$.

Definition 6   Let $S \in \mathcal{S}$. The *shift* function is defined by

$$
\sigma : \quad \begin{array}{ccc} \mathcal{S} & \longrightarrow & \mathcal{S} \\ (S[n])_{n \in \mathbb{N}} & \longmapsto & (S[n+1])_{n \in \mathbb{N}} \end{array}
$$

and the *initial function* is the map which associates to a sequence, its first term

$$
i : \quad \begin{array}{ccc} \mathcal{S} & \longrightarrow & [\![1; \mathsf{N}]\!] \\ (S[n])_{n \in \mathbb{N}} & \longmapsto & S[0]. \end{array}
$$

$\mathbb{B}$ denoting $\{0, 1\}$, let $f : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$ and $S \in \mathcal{S}$ be a strategy. Let us consider the following so called *chaotic iterations* (see [6] for the general definition of such iterations).

$$
\begin{cases} x[0] \in \mathbb{B}^{\mathsf{N}} \\ \forall n \in \mathbb{N}^*, \forall i \in [\![1; \mathsf{N}]\!], x[n]_i = \begin{cases} x[n-1]_i & \text{if } S[n] \neq i \\ f(x[n])_{S[n]} & \text{if } S[n] = i. \end{cases} \end{cases} \tag{1}
$$

In other words, at the $n-$th iteration, only the $S[n]-$th cell is "iterated". Note that in a more general formulation, $f(x[n])_{S[n]}$ can be replaced by $f(x[k])_{S[n]}$, where $k \leqslant n$, modelizing for example delay transmission (see *e.g.* [1]).

III.   A topological approach of chaotic iterations

1   The new topological space

In this section we will put our study in a topological context by defining a suitable set and a suitable distance.

1.1   Defining the iteration function and the phase space

Let us denote by δ the *discrete boolean metric*, $\delta(x, y) = 0 \Leftrightarrow x = y$, and define the function

$$
F_f : \quad \begin{array}{ccc} [\![1; \mathsf{N}]\!] \times \mathbb{B}^{\mathsf{N}} & \longrightarrow & \mathbb{B}^{\mathsf{N}} \\ (k, E) & \longmapsto & \left(E_j . \delta(k, j) + f(E)_k . \overline{\delta(k, j)} \right)_{j \in [\![1; \mathsf{N}]\!]}, \end{array}
$$

where + and . are boolean operations.

Consider the phase space

$$
\mathcal{X} = [\![1; \mathsf{N}]\!]^{\mathbb{N}} \times \mathbb{B}^{\mathsf{N}},
$$

and the map

$$G_f(S, E) = (\sigma(S), F_f(i(S), E)) \tag{2}$$

Then one can remark that the chaotic iterations defined in (1) can be described by the following iterations

$$\begin{cases} X[0] \in \mathcal{X} \\ X[k+1] = G_f(X[k]). \end{cases}$$

The following result can be easily proven, by comparing $\mathcal{S}$ and $\mathbb{R}$ that,

Theorem 1 *The phase space $\mathcal{X}$ has the cardinality of the continuum.*

Note that this result is independent on the number of cells.

1.2 A new distance

We define a new distance between two points $(S, E), (\check{S}, \check{E}) \in \mathcal{X}$ by

$$d((S, E); (\check{S}, \check{E})) = d_e(E, \check{E}) + d_s(S, \check{S}),$$

where

$$\begin{cases} d_e(E, \check{E}) & = & \displaystyle\sum_{k=1}^{\mathsf{N}} \delta(E_k, \check{E}_k), \\ d_s(S, \check{S}) & = & \dfrac{9}{\mathsf{N}} \displaystyle\sum_{k=1}^{\infty} \dfrac{|S[k] - \check{S}[k]|}{10^k}. \end{cases}$$

It should be noticed that if the floor function $\lfloor d(X, Y) \rfloor = n$, then the strategies X and Y differs in n cells and that $d(X, Y) - \lfloor d(X, Y) \rfloor$ gives a measure on how the strategies S and $\check{S}$ diverge. More precisely,

- This floating part is less than 10^{-k} if and only if the first k terms of the two strategies are equal.

- If the $k-$th digit is nonzero, then the $k-$th terms of the two strategies are different.

2 Continuity of the iteration function

To prove that chaotic iterations are an example of topological chaos in the sense of Devaney [4], G_f should be continuous on the metric space $(\mathcal{X}, d)$.

Theorem 2 G_f *is a continuous function.*

PROOF We use the sequential continuity.

Let $(S[n], E[n])_{n \in \mathbb{N}}$ be a sequence of the phase space $\mathcal{X}$, which converges to (S, E). We will prove that $(G_f(S[n], E[n]))_{n \in \mathbb{N}}$ converges to $(G_f(S, E))$. Let us recall that for all n, $S[n]$ is a strategy, thus, we consider a sequence of strategy (*i.e.* a sequence of sequences).
As

$$d((S[n], E[n]); (S, E))$$

4

converges to 0, each distance $d_e(E[n], E)$ and $d_s(S[n], S)$ converges to 0. But $d_e(E[n], E)$ is an integer, so $\exists n_0 \in \mathbb{N}, d_e(E[n], E) = 0$ for any $n \geqslant n_0$.

In other words, there exists threshold $n_0 \in \mathbb{N}$ after which no cell will change its state:

$$\exists n_0 \in \mathbb{N}, n \geqslant n_0 \implies E[n] = E.$$

In addition, $d_s(S[n], S) \longrightarrow 0$, so $\exists n_1 \in \mathbb{N}, d_s(S[n], S) < 10^{-1}$ for all indices greater than or equal to n_1. This means that for $n \geqslant n_1$, all the $S[n]$ have the same first term, which is $S[0]$:

$$\forall n \geqslant n_1, S[n][0] = S[0].$$

Thus, after the $max(n_0, n_1)-$th term, states of $E[n]$ and E are the same, and strategies $S[n]$ and S start with the same first term.

Consequently, states of $G_f(S[n], E[n])$ and $G_f(S, E)$ are equal, then distance d between this two points is strictly less than 1 (after the rank $max(n_0, n_1)$).

We now prove that the distance between $(G_f(S[n], E[n]))$ and $(G_f(S, E))$ is convergent to 0. Let $\varepsilon > 0$.

- If $\varepsilon \geqslant 1$, then we have seen that the distance between $(G_f(S[n], E[n]))$ and $(G_f(S, E))$ is strictly less than 1 after the $max(n_0, n_1)$-th term (same state).

- If $\varepsilon < 1$, then $\exists k \in \mathbb{N}, 10^{-k} \geqslant \varepsilon \geqslant 10^{-(k+1)}$. But $d_s(S[n], S)$ converges to 0, so

$$\exists n_2 \in \mathbb{N}, \forall n \geqslant n_2, d_s(S[n], S) < 10^{-(k+2)},$$

   after n_2, the $k + 2$-th first terms of $S[n]$ and S are equal.

As a consequence, the $k + 1$ first entries of the strategies of $G_f(S[n], E[n])$ and $G_f(S, E)$ are the same (because G_f is a shift of strategies), and due to the definition of d_s, the floating part of the distance between $(S[n], E[n])$ and (S, E) is strictly less than $10^{-(k+1)} \leqslant \varepsilon$.

In conclusion, G_f is continuous,

$$\forall \varepsilon > 0, \exists N_0 = max(n_0, n_1, n_2) \in \mathbb{N}, \forall n \geqslant N_0, d\left(G_f(S[n], E[n]); G_f(S, E)\right) \leqslant \varepsilon.$$

In this section, we proved that chaotic iterations can be modelized as a dynamical system in a topological space. In the next section, we show that chaotic iterations are a case of topological chaos, in the sense of Devaney.

IV.   Discrete chaotic iterations are topological chaos

To prove that we are in the framework of Devaney's topological chaos, we will check the regularity and transitivity conditions.

1 Regularity

Theorem 3 *Periodic points of G_f are dense in $\mathcal{X}$.*

PROOF Let $(S, E) \in \mathcal{X}$, and $\varepsilon > 0$. We are looking for a periodic point (S', E') satisfying

$$d((S, E); (S', E')) < \varepsilon.$$

We choose $E' = E$, and we "copy" enough entries from S to S' so that the distance between (S', E) and (S, E) is strictly less than ε: a number $k = \lfloor log_{10}(\varepsilon) \rfloor + 1$ of terms is sufficient.
After this k-th iterations, the new common state is $\mathcal{E}$, and strategy S is shifted of k positions: $\sigma^k(S)$.
Then we have to complete strategy S' in order to make (E', S') periodic (at least for sufficiently large indices). To do so, we put an infinite number of 1 to the strategy S'. Then, either:

1. The first state is conserved after one iteration, so $\mathcal{E}$ is unchanged and we obtain a fixed point. Or

2. The first state is not conserved, then:

   - If the first state is not conserved after a second iteration, then we will be again in the first case above (due to the negation function).
   - Otherwise the first state is conserved, and we have indeed a fixed (periodic) point.

Thus, there exists a periodic point into every neighborhood of any point, so $(\mathcal{X}, G)$ is regular.

2 Transitivity

Contrary to the regularity, the topological transitivity condition is not automatically satisfied by any function ($f = Identity$ is not topologically transitive).
    Let us denote by $\mathcal{T}$ the set of maps f such that $(\mathcal{X}, G_f)$ is topologically transitive. Then.

Theorem 4 $\mathcal{T}$ *is a nonempty set.*

PROOF We will prove that the vectorial logical negation function f_0

$$f_0 : \quad \begin{array}{ccc} \mathbb{B}^{\mathsf{N}} & \longrightarrow & \mathbb{B}^{\mathsf{N}} \\ (x_1, \ldots, x_{\mathsf{N}}) & \longmapsto & (\overline{x_1}, \ldots, \overline{x_{\mathsf{N}}}) \end{array} \tag{3}$$

is topologically transitive.
Let $A = \mathcal{B}(X_A, r_A)$ and $B = \mathcal{B}(X_B, r_B)$ be two open balls of $\mathcal{X}$. Our goal is to start from a point of A (*i.e.* a point close to X_A) and to arrive in B (a point close to X_B).
We have to be close to X_A, then the starting state is E_A, it remains to determine the strategy S. We start by filling S with the n_0 first terms of strategy S_A of X_A, so that $(S, E_A) \in B_A$.
Let E be the image of the state E_A by mapping the n_0-th first terms of the strategy S. This new state E differs from E_B by a finite number of states, we put these cells to our strategy S (this adds n_1 integers to S). In short, starting from (S, E_A), we are in X_B after $n_0 + n_1$ iterations, and the strategy S was shifted of $n_0 + n_1$ terms (there is no more term in S).
In order to be sufficiently close to (S_B, E_B) (at a distance less than ε from (S_B, E_B)), we add as much as necessarily terms of S_B to S and we complete S with an infinity of terms equal to 1.

Remark 1 In fact, we can prove that $(\mathcal{X}, G_{f_0})$ is *highly topologically transitive* in the following sense: for every (S_A, E_A) and (S_B, E_B) of $\mathcal{X}$, there exists a point sufficiently close to (S_A, E_A) and $n_0 \in \mathbb{N}$ such that $G_{f_0}^{n_0}(S_A, E_A) = (S_B, E_B)$.

In conclusion, if $f \in \mathcal{T} \neq \varnothing$, then $(\mathcal{X}, G_f)$ is topologically transitive and regular, and then we have the result.

Theorem 5 $\forall f \in \mathcal{T}, (\mathcal{X}, G_f)$ *is chaotic, in the sense of Devaney.*

V. Hash functions based on topological chaos

1 Objective

As an application of the previous theory, we define in this section a new way to generate hash functions based on topological chaos. Our approach guarantees to obtain various desired properties in the domain of encryption. For example, the avalanche criterion is closely linked to the expansivity property (see the next section below).

The following hash function is based on the vectorial boolean negation f_0 defined in (3). Nevertheless, our procedure remains general, and can be applied with any transitive function f.

2 Application of the new hash function

Our initial condition $X0 = (S, E)$ is composed by:

- A 256 bits sequence that we call E, obtained from the original text.

- A chaotic strategy S.

In the sequel, we describe how to obtain this initial condition (S, E).

2.1 How to obtain E

The first step of our algorithm is to transform the message in a normalized 256 bits sequence E. To illustrate this step, we take an example, our original text is: *The original text*

Each character of this string is replaced by its ASCII code (on 7 bits). Then, we add a 1 to this string.

```
10101001 10100011 00101010 00001101 11111100 10110100
11100111 11010011 10111011 00001110 11000100 00011101
00110010 11111000 11101001
```

Then, we add the binary value of the length of this string, and we add 1 one more time:

```
10101001 10100011 00101010 00001101 11111100 10110100
11100111 11010011 10111011 00001110 11000100 00011101
00110010 11111000 11101001 11110001
```

Then, the whole string is copied, but in the opposite direction, this gives:

```
10101001 10100011 00101010 00001101 11111100 10110100
11100111 11010011 10111011 00001110 11000100 00011101
00110010 11111000 11101001 11110001 00011111 00101110
00111110 10011001 01110000 01000110 11100001 10111011
10010111 11001110 01011010 01111111 01100000 10101001
10001011 0010101
```

So, we obtain a multiple of 512, by duplicating enough this string and truncating at a multiple of 512. This string, in which contains the whole original text is denoted by D.

Finally, we split our obtained string into blocks of 256 bits, and apply to them the exclusive-or function, obtaining a 256 bits sequence.

```
11111010 11100101 01111110 00010110 00000101 11011101
00101000 01110100 11001101 00010011 01001100 00100111
01010111 00001001 00111010 00010011 00100001 01110010
01000011 10101011 10010000 11001011 00100010 11001100
10111000 01010010 11101110 10000001 10100001 11111010
10011101 01111101
```

So, in the context of subsection (1) , $N = 256$, and E is the above obtained sequence of 256 bits. Let us now build the strategy S.

We now have the definitive length of our digest. Note that a lot of texts have the same string. This is not a problem because the strategy we will build will depends on the whole text.

2.2   How to choose S

In order to forge our strategy, *i.e.* the sequence S of $X[0] = (S, E)$, we use the previously obtained string D, and then we start by constructing an intermediate sequence as follows:

1. We split this string into blocks of 8 bits, and we add to our sequence the corresponding decimal value of each octet.

2. We take then the first bit of this string, and put it on the end. Then we split the new string into blocks of 8 bits, and we add in the sequence decimal value associated.

3. We repeat this operation 6 times.

The general term of this sequence will be denoted by $(u[n])_n$.

Now, we are able to build our strategy S. The first term of S is the initial term of the preceding sequence. The $n-$th term is the sum (modulo 256) of the three following terms:

- the $n-$th term of the intermediate sequence (the strategy depends on the original text),

8

- the double of the $n-1$-th term of the strategy (introduction of sensitivity, with the analogy with the well known chaotic map $\theta \longmapsto 2\theta \ (mod \ 1)$),

- n (to prevent periodic behaviour).

So, the general term $S[n]$ of S is defined by

$$S[n] = (u[n] + 2 \times S[n-1] + n) \ (mod \ 256).$$

Strategy S is strong sensitive to the modification of the original text, because the map $\theta \longmapsto 2\theta \ (mod \ 1)$ is known to be chaotic in the sense of Devaney.

2.3 How to construct the digest

We apply the logical negation function to the $S[k]$-th term of E, (modulo 256). Indeed, the function f of equation (3) is defined by

$$f : \quad \begin{matrix} [\![1,256]\!] & \longrightarrow & [\![1,256]\!] \\ (E[1],\ldots,E[256]) & \longmapsto & (\overline{E[1]},\ldots,\overline{E[256]}). \end{matrix}$$

It is possible to apply the logical negation function several times the same bit.

We finally split these 256 bits into blocks of 4 bits, this will returns the hexadecimal value:

```
63A88CB6AF0B18E3BE828F9BDA4596A6A13DFE38440AB9557DA1C0C6B1EDBDBD
```

As a comparison if instead of considering the text "*The original text*" we took "*the original text*", the hash function returns:

```
33E0DFB5BB1D88C924D2AF80B14FF5A7B1A3DEF9D0E831194BD814C8A3B948B3
```

3 Example

Consider the following message (a *E. A. Poe*'s poem):

> Wanderers in that happy valley,
>   Through two luminous windows, saw
> Spirits moving musically,
>   To a lute's well−tuned law,
> Round about a throne where, sitting
>   (Porphyrogene !)
> In state his glory well befitting,
>   The ruler of the realm was seen.
>
> And all with pearl and ruby glowing
>   Was the fair palace door,
> Through which came flowing, flowing,

And sparkling evermore,
A troop of Echoes, whose sweet duty
Was but to sing,
In voices of surpassing beauty,
The wit and wisdom of their king.

Our hash function returns :

```
FF51DA4E7E50FBA7A8DC6858E9EC3353BDE2E465E1A6A1B03BEAA12A4AD694FB
```

If we put an additional space before " Was the fair palace door," the hash function returns:

```
03ABFA49B834D529669CFC1AEEC13E14EA5FFD2349582380BCBDBF8400017445
```

If we replace "Echoes" by "echoes" in the original text, the hash function returns:

```
FE54777C52D373B7AED2EA5ACAD422B5B563BB3B91E8FCB48AAE9331DAC54A9B
```

VI.   Quantitative measures

1   General definitions

In the previous section we proved that discrete iterations produce a topological chaos by checking two qualitative properties, namely transitivity and regularity. This mathematical framework offers tools to measure this chaos quantitatively.
The first of this measures is the constant of sensitivity defined in definition 4.
Intuitively, a function f has a constant of sensitivity equals to δ implies that there exist points aribitrarily close to any point x which eventually separate from x by at least δ under some iterations of f.
This induces that an arbitrarily small error on a the initial condition may become magnified upon iterations of f. (This is related to the famous *butterfly effect*).
Other important tools are defined below.

Definition 7 A function f is said to have the property of *expansivity* if

$$\exists \varepsilon > 0, \forall x \neq y, \exists n \in \mathbb{N}, d(f^n(x), f^n(y)) \geqslant \varepsilon.$$

Then, ε is the *constant of expansivity* of f. We also say f is ε-expansive.

Remark 2 A function f has a constant of expansivity equals to ε if an arbitrary small error on any initial condition is amplified till ε.

There exist other important quantitative tools such as topological entropy, which quantifies the information contained at each iteration. But this is not in the objective of this paper.
We will reconsider this quantitative measures in the next subsection, in relation with hash functions.

10

2 Quantitative evaluation of our hash function

Let f_0 be the vectorial logical negation previously used in our algorithm. In this section, sensitivity and expansivity constants of G_{f_0} will be calculated.

2.1 Sensitivity

We know that $(\mathcal{X}, G_{f_0})$ has sensitive dependence on initial conditions. Moreover, we have the following result.

Theorem 6 *The constant of sensitivity of $(\mathcal{X}, G_{f_0})$ is equal to* N.

Recall that $\mathsf{N} = 256$ in our hash function.

PROOF We have seen that sensitivity is a consequence of having Devaney's chaos property. Let us determine its constant.
Let (S, E) be a point of $\mathcal{X}$, and $\delta > 0$. Then, let us define another point (S', E') by:

- $E' = E$,

- The k-th first terms of S' are the same as those of S, where $k = \lfloor log_{10}(\varepsilon) \rfloor + 1$ such that
$$d((S, E); (S', E')) < \delta.$$

- Then, we put the terms $1, 2, 3, \ldots, \mathsf{N}$ to S'.

- S' can be completed by any terms.

Then it can be found a point (S', E') closed to (S, E) $(d((S, E); (S', E')) < \delta)$, such that states of $G_{f_0}^{k+\mathsf{N}}(S, E)$ and $G_{f_0}^{k+\mathsf{N}}(S', E')$ differ for each cell, so that the distance between this two points is greater or equal to N. This proves that we have sensitive dependence on the original text and that the constant of sensitivity is N.

2.2 Expansivity

Theorem 7 $(\mathcal{X}, G_{f_0})$ *is an expansive chaotic system. Its constant of expansivity is equal to 1.*

PROOF If $(S, E) \neq (\check{S}; \check{E})$, then:

- Either $E \neq \check{E}$, and then at least one cell is not in the same state in E and $\check{E}$. Then the distance between (S, E) and $(\check{S}; \check{E})$ is greater or equal to 1.

- Or $E = \check{E}$. Then the strategies S and $\check{S}$ are not equal. Let n_0 be the first index in which the terms S and $\check{S}$ differ. Then
$$\forall k < n_0, G_{f_0}^{n_0}(S, E) = G_{f_0}^k(\check{S}, \check{E}),$$
and $G_{f_0}^{n_0}(S, E) \neq G_{f_0}^{n_0}(\check{S}, \check{E})$, then as $E = \check{E}$, the cell which has changed in E at the n_0-th iterate is not the same than the cell which has changed in $\check{E}$, so the distance between $G_{f_0}^{n_0}(S, E)$ and $G_{f_0}^{n_0}(\check{S}, \check{E})$ is greater or equal to 2.

The property of expansivity is a kind of avalanche effect.
Remark that it can be easily proved that $(\mathcal{X}, G_{f_0})$ is not A-expansive, for any $A > 1$.

11

VII.   Discussion and future work

We proved that discrete chaotic iterations are a particular case of Devaney's topological chaos if the iteration function is topologically transitive and that the set of topologically transitive functions is non void.

We applied our results to the generation of new hash functions. Even if we used the vectorial boolean negation function, our procedure remains general and other transitive functions can be used.

By considering hash functions as an application of our theory, we have shown how some desirable aspects in encryption such as unpredictability, sensitivity to initial conditions, mixture and disorder can be mathematically guaranteed and even quantified by mathematical tools.

Theory of chaos recalls us that simple functions can have, when iterated, a very complex behaviour, while some complicated functions could have foreseeable iterations. This is why it is important to have tools for evaluating desired properties.

In our example, we used a simple topologically transitive iteration function, but it can be proved that there exist a lot of functions of this kind. Our simple function may be replaced by other "chaotic" functions which can be evaluated with the above described quantitative tools. Another important parameter is the choice of the strategy S. We proposed a particular strategy that can be easily improved by multiple ways.

We do not claim to have proposed a hash function replacing well known ones, we simply wished to show how our mathematical context allows to build such functions and especially how important properties can be measured.

Much work remains to be made, for example we are convinced that the good comprehension of the transitivity property, enables to study the problem of collisions in hash functions.

In future work we plan to investigate other forms of chaos such as Li-York chaos [5] and to explore other quantitative and qualitative tools such as entropy (see e.g. [3]) and to enlarge the domain of applications of our theoretical concepts.

References

[1] J. M. Bahi. Boolean totally asynchronous iterations. *Int. Journal of Mathematical Algorithms*, 1:331–346, 2000.

[2] J. Banks, J. Brooks, G. Cairns, and P. Stacey. On devaney's definition of chaos. *Amer. Math. Monthly*, 99:332–334, 1992.

[3] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.*, 153:401–414, 1971.

[4] R. L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Redwood City: Addison-Wesley, 2. edition, 1989.

[5] T. Y. Li and J. A. Yorke. Period three implies chaos. *Am. Math. Monthly*, 82(10):985–992, 1975.

[6] F. Robert. *Discrete Iterations: A Metric Study*, volume 6 of *Springer Series in Computational Mathematics*. 1986.