Discrete Chaos and Cryptography

J.M. Amigó(1), J. Szczepanski(2) and L. Kocarev(3)

(1) Centro de Investigación Operativa, Universidad Miguel Hernández
Avda. de la Universidad, 03202 Elche, Spain,
(2) Institute of Fundamental Technological Research, Polish Academy of Science
Swietokrzyska 21, 00-049 Warsaw, Poland,
(3) Institute for Nonlinear Science, University of California San Diego
9500 Gilman Drive, La Jolla CA 92093-0402, USA,
Email: jm.amigo@umh.es, jszczepa@ippt.gov.pl, lkocarev@ucsd.edu.

Abstract—We propose definitions of discrete Lyapunov exponent and discrete entropy for permutations on a finite set. We justify our definitions by proving that in the ‘infinite limit’, i.e., when the cardinality \( M \) of the set goes to infinity, the discrete concepts converge to their continuous counterparts for a large class of chaotic maps. Consequently, we say that a discrete-time dynamical system on a finite-state phase space is discretely chaotic if its discrete Lyapunov exponent tends to a positive number (or to \( \infty \)) when \( M \to \infty \). Possible applications of discrete chaos to cryptography are also discussed.

1. Introduction

What makes chaotic systems so attractive both for theoreticians and practitioners is their random-like behavior — in spite of being deterministic. As way of illustration, let us mention that, already in 1949, C. Shannon [1] proposed this kind of transformations to construct secure cryptosystems. It is thus no surprise that, when chaos theory flourished in the nineteen-eighties and -nineties, several cryptosystems were proposed based on the discretization of chaotic maps. Viewing how the resulting permutations mix the pixels of digital pictures [2], one cannot but admit that their ‘confusion’ strategy is implemented, roughly speaking, by means of bit permutations with strong nonlinearity (S-boxes) on subblocks of the input block and permutations with fast spreading factor on whole blocks, respectively. This being the case, the security of all these ciphers relies ultimately on such permutations delivering the right mixing and propagation properties. Here is where discrete chaos comes in: it provides tools like the discrete Lyapunov exponent and entropy to quantify the said properties. The design and certification of special-purpose permutations is just an example of possible and interesting applications of discrete chaos to cryptography. Others include the design of cryptologic algorithms, hash functions and the like.

2. The Tools of Discrete Chaos

2.1. The discrete Lyapunov Exponent

Let \( S = \{s_1, ..., s_M\} \) be a linearly ordered finite set endowed with a metric \( d(\cdot, \cdot) \), and \( F : S \to S \) be a bijection or equivalently, an \( M \)-permutation. We define the discrete Lyapunov exponent (DLE) of \( F \) as

\[
\lambda_F = \frac{1}{M-1} \sum_{j=1}^{M-1} \log \frac{d(F(s_{j+1}), F(s_j))}{d(s_{j+1}, s_j)}
\]

Following the tradition, we will use natural logarithms to calculate \( \lambda_F \). Observe that \( \lambda_F \) depends on the order and on the metric \( d \) but is invariant under rescaling and, furthermore, has the same invariances as \( d \).

In the examples and applications we will consider below, \( F \) will be a permutation on a subset \( S \) of \( \mathbb{R} \) (for instance, \( S = [0, ..., M-1] \equiv \mathbb{Z}_M \) endowed with the Euclidean distance \( d(s_i, s_j) = |s_i - s_j| \)). The case \( S = \mathbb{Z}_2^t \), with \( \mathbb{Z}_2^t \equiv \mathbb{Z}_2 \times ... \times \mathbb{Z}_2 \) (\( t \) times) being the set of binary strings of length \( k \) endowed with Hamming distance, has also interest for cryptographic applications, but it will not be considered here. Observe that if \( S = \mathbb{Z}_M \) or \( S = \mathbb{Z}_2^t \) lexicographically ordered, then \( \lambda_F \geq 0 \).

Example 1. Suppose that \( M = 2m \) and define

\[
F_{\text{max}}(s) = \begin{cases} 
    m + k & \text{if } s = 2k, \quad 0 \leq k \leq m - 1 \\
    k & \text{if } s = 2k + 1, \quad 0 \leq k \leq m - 1
\end{cases}
\]
on $\mathbb{Z}_M$. The DLE of $F_{\text{max}}$ is

$$\lambda_{F_{\text{max}}} = \frac{m}{2m - 1} \ln m + \frac{m - 1}{2m - 1} \ln (m + 1).$$

(1)

It can be shown [3] that $F_{\text{max}}$ has the largest DLE among all permutations of the set $\{0, ..., M - 1\}$.

Let $z_{j+1} = f(z_j)$, $j = 0, 1, ..., M - 1$, be a typical trajectory of length $M$ of a one-dimensional chaotic map $f : [0, 1] \rightarrow [0, 1]$, such that $z_{j+1} \neq z_j$ for all $j$ and $|z_{M-1} - z_0| < \varepsilon$. We define $f(z_{M-1}) = z_0$ and order $z_j$ according to the metric to obtain $x_j$, that is, $x_0 < x_1 < ... < x_{M-1}$, so that $x_j$ and $x_{j+1}$ are neighbors in the metric sense. Define $m_i = \{x_iN\}$, where $N$ is chosen such that $m_i \neq m_j$ for all $i$ and $j$. The map $f$ induces then the obvious permutation $F_M : \{m_0, ..., m_{M-1}\} \rightarrow \{m_0, ..., m_{M-1}\}$ with $F(m_i) = m_j$ when $f(x_i) = x_j$. The following theorem justifies calling $\lambda_{F_M}$ a discrete Lyapunov exponent.

**Theorem 1** [3]: In the above setting, $\lim_{M \rightarrow \infty} \lambda_{F_M} = \lambda_f$, where $\lambda_f$ is the Lyapunov exponent of $f$.

**Example 2.** For the general tent map,

$$f(x) = \begin{cases} \frac{x}{a}, & 0 \leq x < a \\ \frac{2 - x}{1 - a}, & a < x \leq 1 \end{cases},$$

the induced family of $M$-permutations can be given in closed form, namely,

$$F_M(s) = \begin{cases} 1, & s = 0 \\ \text{Cl} \left( \frac{M}{M - A}(M - s) \right), & 0 < s \leq A \\ \text{Fl} \left( \frac{M}{M - A}(M - s) \right) + 1, & A < s \leq M - 1 \end{cases}$$

where $\text{Cl}(x)$ and $\text{Fl}(x)$ denote, as usual, the ceiling and floor of $x$, $a = A/M$ and $s \in \{0, 1, ..., M - 1\}$. The following table shows the values of $\lambda_{F_M}$ and $\lambda_f$ for the symmetric tent map ($a = 0.5, \lambda_f = \ln 2 = 0.693147$).

<table>
<thead>
<tr>
<th>$M$</th>
<th>$\lambda_f$</th>
<th>$\lambda_{F_M} - \lambda_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>256</td>
<td>0.690440</td>
<td>2.7076 $\times 10^{-3}$</td>
</tr>
<tr>
<td>1024</td>
<td>0.692470</td>
<td>6.7690 $\times 10^{-4}$</td>
</tr>
<tr>
<td>34012</td>
<td>0.693127</td>
<td>2.0379 $\times 10^{-5}$</td>
</tr>
</tbody>
</table>

**Example 3.** For the logistic map $f(x) = 4x(1 - x)$, $0 \leq x \leq 1$, we get the following table ($\lambda_f = \ln 2$).

<table>
<thead>
<tr>
<th>$M$</th>
<th>$\lambda_f$</th>
<th>$\lambda_{F_M} - \lambda_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0.681690</td>
<td>0.011457</td>
</tr>
<tr>
<td>128</td>
<td>0.687300</td>
<td>0.005710</td>
</tr>
<tr>
<td>512</td>
<td>0.691484</td>
<td>0.001663</td>
</tr>
</tbody>
</table>

**Definition:** We say that an $M$-permutation $F_M$ on $\{0, 1, ..., M - 1\}$ has a perfect nonlinearity if the differences $|F(i + 1) - F(i)|, i = 0, 1, ..., M - 2$, take all possible values $1, 2, ..., M - 1$.

Note that this definition is weaker than the usual one for Boolean functions on binary blocks. The following example shows the existence of maps with perfect nonlinearity.

**Example 4.** Let $M = 2m$ and define

$$F_{\text{non}}(s) = \begin{cases} k, & \text{if } s = 2k \quad 0 \leq k \leq m - 1 \\ M - 1 - k, & \text{if } s = 2k + 1 \quad 0 \leq k \leq m - 1 \end{cases}$$

on $\mathbb{Z}_M$. The DLE of $F_{\text{non}}$ is

$$\lambda_{F_{\text{non}}} = \frac{1}{M - 1} \ln (M - 1)!.$$  

(2)

**2.2. The Discrete Entropy**

Our definition of discrete entropy is inspired by the so-called permutation entropy, that is a concept introduced by Bandt and Pompe in [4] as a practical complexity measure for experimental time series. The basic idea calls for comparing the values $x_1 = f(x_0), x_2 = f(x_1), ..., f$ on a typical orbit of length $M$ of a one-dimensional dynamical system and keeping track of the order patterns appearing with increasing $M$. Bandt et al. proved in [5] that permutation metric and topological entropy of piecewise monotone interval maps coincide, respectively, with Kolmogorov-Sinai (KS) and topological entropy. The result concerning permutation metric entropy and KS entropy was generalized to ergodic information sources and higher dimensional ergodic interval maps in [6].

Let $F$ be a bijection on $S = \{s_1, ..., s_M\}$ with the linear ordering $< \text{(possibly induced by a metric). Note that seemingly more general situations, like bijections on product sets endowed with the product (or lexicographical) ordering, can be reduced to the case we are envisaging here. Thus, } F \text{ is an } M \text{-permutation on a linearly ordered set.}

For $2 \leq r \leq M$ and $\pi$ the $r$-permutation $[\pi(0), \pi(1), ..., \pi(r - 1)]$ (shorthand for $0 \mapsto \pi(1), 1 \mapsto \pi(1), ...,)$, we define (i)

$$P_{\pi} = \{s \in S : F^{\pi(0)}(s) < ... < F^{\pi(r - 1)}(s)\},$$

(3)

i.e., $P_{\pi}$ is the set of points $s \in S$ such that the orbit segment $\{s, F(s), ..., F^{\pi(r - 1)}(s)\}$ is ordered under $<$ as $\pi$ (alternatively, we say that $s$ ‘defines’ the $r$-permutation $\pi$), and (ii)

$$p_{\pi} = |P_{\pi}| / \sum_{\pi \in \mathcal{P} r} |P_{\pi}|,$$

where $|\cdot|$ stands here for cardinality. We will refer to $F$ also as a substitution, reserving the word permutation for $\pi$. We define the discrete entropy rate of $F$ of rank $r$, $2 \leq r \leq M$, (based on the corresponding rate of the permutation entropy) as

$$\tilde{H}_r^{\pi}(F) = - \frac{1}{r - 1} \sum_{\pi \in \mathcal{P} r} p_{\pi} \log p_{\pi}.$$  

(4)

For entropy and entropy rate, logarithms to base 2 (and corresponding units in bits and bits per symbol, respectively) are in general preferred to other bases.

**Example 5.** For the right shift modulo $M$,

$$F_1 = [1, 2, ..., M - 1, 0],$$
we get
\[ \hat{h}_{\Pi}^{(i)}(F_1) = \frac{M - r + 1}{M(r - 1)} \log_2 \frac{M}{M - r + 1} + \frac{1}{M} \log_2 M \]
for \( 2 \leq r \leq M \). In particular, for \( M = 4 \) we have
\[ \hat{h}_{\Pi}^{(2)}(F_1) = 0.81; \quad \hat{h}_{\Pi}^{(3)}(F_1) = 0.75; \quad \hat{h}_{\Pi}^{(4)}(F_1) = 0.67. \]

As for
\[ F_2 = \{M, 0, M + 1, 1, M + 2, 2, \ldots, 2M - 1, M - 1\}, \]
the substitution on \([0, 1, \ldots, 2M - 1]\) with maximum discrete Lyapunov exponent (Example 1), we get for \( M = 2, \)
\[ \hat{h}_{\Pi}^{(2)}(F_2) = 1; \quad \hat{h}_{\Pi}^{(3)}(F_2) = 1; \quad \hat{h}_{\Pi}^{(4)}(F_2) = 0.67. \]

We see that \( \hat{h}_{\Pi}^{(3)}(F_1) \leq \hat{h}_{\Pi}^{(3)}(F_2) \) for \( r = 2, 3, 4 \). In particular, the smaller ranks \( r = 2, 3 \) show that \( F_2 \) is more random than \( F_1 \).

Suppose finally that \( F \) has non-zero entropy rates of ranks \( r = 2, 3, \ldots, q \). The possibly simplest way to encapsulate in a single number the information contained in the whole hierarchy \( \hat{h}_{\Pi}^{(2)}(F), \ldots, \hat{h}_{\Pi}^{(q)}(F) \) (without having to dissect \( F \) into cycles) consists of taking the arithmetic mean of it:
\[ h_{\Pi}(F) = \frac{1}{q - 1} \sum_{r=2}^{q} \hat{h}_{\Pi}^{(r)}(F). \]

In this way \( h_{\Pi}(F) \) takes into account both high and, most importantly, low and middle ranks on an equal footing; indeed, although the number of summands in \( \hat{h}_{\Pi}^{(r)}(F) \) grows as \( r! \) (see (4)), the sum of the non-zero terms (before getting multiplied by \( 1/(r - 1) \)) actually scales linearly in \( r \), rendering the different entropy rates comparable sizes. Moreover, if we let formally \( q \to \infty \), we recover the usual definition \( h_{\Pi}(F) = \lim_{r \to \infty} \hat{h}_{\Pi}^{(r)}(F) \), since a convergent sequence and the arithmetic mean of their successive terms have the same limit. We call \( h_{\Pi}(F) \) the (metric) discrete entropy (or just the entropy) of \( F \).

**Theorem 2** [7]: Let \( I \) be an \( n \)-dimensional interval and \( f : I \to I \) an ergodic map. Let \( F \) be a substitution on \( M \) elements obtained from \( f \), after discretizing \( I \), in a way similar to Sect. 2.1. Then \( \lim_{r \to \infty} h_{\Pi}(F_M) = h_{KS}(f) \), where \( h_{KS}(f) \) is the KS entropy of \( f \).

### 3. Discrete Chaos

For \( S = \{s_1, \ldots, s_M\} \) let us consider an arbitrary map \( F : S \to S \) (not necessarily bijective). We say that the fixed point \( s_j \) (i.e., \( F(s_j) = s_j \)) is an eventually fixed point for \( s_j \) if there exists \( n \geq 1 \) such that \( F^n(s_j) = s_j \).

**Definition**: We say that \( s_j \) is a stable fixed point for the map \( F \) if \( F(s_j) = s_j \) and \( s_j \) is an eventually fixed point for at least one of its neighbor points \( s_{i \pm 1} \). In a similar way, one can define stable periodic orbits of period \( p \) using \( F^p \) instead of \( F \).

We say that a periodic orbit is unstable if it is not stable. In particular, all periodic orbits (cycles) of permutations are unstable.

A set \( L_M = \{l_i \in \mathbb{R} : i = 1, 2, \ldots, M\} \) is said to be a (one-dimensional) finite lattice if \( l_{i+1} = l_i + \Delta = l_i + i\Delta \) for \( 1 \leq i \leq M - 1 \), where \( \Delta > 0 \) can be chosen (by rescaling the Euclidean metric) to be 1. For example, \( L_M = \{0, 1, \ldots, M - 1\} \) is a finite lattice. Given \( \mathcal{A} = \{a_1, \ldots, a_m\} \subset L_M \), for every \( M \geq M_0 \), we define \( \partial \mathcal{A} = \{a_1 \pm 1, a_2 \pm 1, \ldots, a_m \pm 1\} \) to be the neighboring set of \( \mathcal{A} \) in \( L_M \) (if \( a_1 = l_1 \) or \( a_m = l_M \), then the neighboring points are \( l_1 + 1 \) and \( l_M - 1 \), respectively). Given now a map \( F_M \) on \( L_M \), \( M \geq M_0 \), we say that \( \mathcal{A} \) is an invariant set of \( F_M \) if \( F_M(\mathcal{A}) = \mathcal{A} \).

**Definition**: We say \( \mathcal{A} \) is an attractor of \( F_M \), if \( \mathcal{A} \) is invariant under \( F_M \) and there exists \( l \in \partial \mathcal{A} \) such that \( F_M(l) \in \mathcal{A} \).

Let \( \mathcal{A}_M \subset L_M \) be an invariant set under the action of the map \( F_M : L_M \to L_M \), \( M \geq M_0 \), such that \( F_M \) restricted to \( \mathcal{A}_M \) is a bijection and write \( G_M = F_M|_{\mathcal{A}_M} \).

**Definition**: (i) We say that the map \( G_M \) is discretely chaotic on the set \( \mathcal{A}_M \) if \( \lim_{M \to \infty} \lambda_{G_M} > 0 \). (ii) We say that \( \mathcal{A}_M \) is a discretely chaotic attractor for \( F_M \) if \( \mathcal{A}_M \) is an attractor of \( F_M \) and \( \lim_{M \to \infty} \lambda_{G_M} > 0 \).

In particular, if \( F_M \) is an \( M \)-permutation, then \( F_M \) is discretely-chaotic if \( \lim_{M \to \infty} \lambda_{G_M} > 0 \). The perturbations \( F_{\text{max}} \) and \( F_{\text{non}} \) from Examples 1 and 4 are discretely chaotic.

Observe that, in strict sense, the concepts of discretely chaotic map and attractor refer to a family of maps rather than to a single map. In most applications, \( F_M \) is certainly obtained via phase space discretization and truncation of the orbits of a continuous map \( f \), as in the proofs of Theorems 1 and 2, and therefore it belongs to a family of maps (generated by \( f \)) by construction. Otherwise, if \( S = \mathbb{Z}_M \) or a translate, one can always compare \( \lambda_{F_{\text{max}}} \) to \( \lambda_{F_{\text{non}}} \), and gauge in this way the ‘distance’ from \( F_M \) to \( F_{\text{max}} \)—the most discretely chaotic permutation on \( \mathbb{Z}_M \).

### 4. Applications to Cryptography

We will now focus on the cryptographic applications of discrete chaos and, more concretely, on the quality assessment and performance comparison of S-boxes.

#### 4.1. Discrete Lyapunov exponents

As of this writing, we have analyzed the S-boxes of Rijndael cipher (the winner of the Advanced Encryption Standard –AES– competition) by means of the discrete Lyapunov exponent. The cipher is designed for 128, 192 and 256 bit block lengths but, for simplicity, we consider here the first implementation only. Rijndael applies the following transformations:
i) The ByteSub transformation $S(x)$ is a byte-level S-box (thus, $S : \mathbb{Z}^8 \rightarrow \mathbb{Z}^8$) defined as

$$S(x) = Bx^{-1} + b,$$

where $x^{-1} \in GL(2^8)$ is the multiplicative inverse of $x$ if $x \neq 0$ or $0$ if $x = 0$, $B$ is an $8 \times 8$ binary matrix $A$ obtained by successively rotating the bits of its first row $B_{ij} = (1, 0, 0, 0, 1, 1, 1, 1)$ to the right, and $b = (1, 1, 0, 0, 0, 1, 1, 0)\text{transpose}$. The ByteSub transformation defines a permutation $F$ on $[0, \ldots, 255]$ with $\lambda_F = 4.01$, while $\lambda_{\text{max}} = 4.86$ (see (1)) and $\lambda_{\text{non}} = 4.55$ (see (2)). The role of the ByteSub transformation is to mix in a strong nonlinear way the input information.

ii) Let $b_{0,0}, \ldots, b_{0,3}, b_{1,0}, \ldots, b_{1,3}$ be the 16 bytes (128 bits) of the input block. The ShiftRow transformation takes the words

$$w_i = (b_{i,0}, b_{i,1}, b_{i,2}, b_{i,3})$$

and returns $w_i \ggg C_i$, $i = 0, 1, 2, 3$, where $w \ggg C$ is the rotation of the sequence of bytes to the right by $C$ bytes. The values of $C$ are $C_i = i, i = 0, 1, 2, 3$. The role of the ShiftRow permutation is just to permute all 16 bytes of the input block, thus it is a permutation on $[0, 1, \ldots, 15]$. Its DLE turns out to be 0.93, which is substantially smaller than the maximum one (for $M = 16$) 2.13.

iii) Given an input block in the form (5), the MixColumn transformation can be viewed as a linear transformation in $GF(2^8)^4$. In fact, if $c_j = (b_{j,0}, b_{j,1}, b_{j,2}, b_{j,3})$, $0 \leq j \leq 3$, is the $j$th column of (5), then MixColumn is

$$c_j \rightarrow \begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix} c_j,$$

where the matrix entries are pair of hexadecimal numbers representing bytes in the usual way. Therefore, MixColumn induces a permutation on $[0, 1, \ldots, 2^{32} - 1]$. We have found the Lyapunov exponent of MixColumn to be 21.49 ($\lambda_{\text{max}} \approx 22.18$).

In addition to the analysis of the single transformations, the behavior of their composition (i.e., of the Rijndael cipher) has been evaluated. To this aim, we assign to each 128 bit block an integer in $[0, 1, \ldots, 2^{128} - 1]$ via its binary representation. The computation of the DLE has been performed on 7000 iterations of the Rijndael map obtaining 87.04, to be compared to $\lambda_{\text{max}} = 88.72$ (see (1) with $M = 2^{128}$).

### 4.2. Discrete entropy

As for the applications of discrete entropy, let us illustrate them with an example taken from [8]. The $4 \times 4$ S-boxes

$$S_1 = [15, 12, 2, 1, 9, 7, 10, 4, 6, 8, 5, 11, 0, 3, 13, 14]$$

S_2 = [8, 2, 4, 13, 7, 14, 11, 1, 9, 15, 6, 3, 5, 0, 10, 12]

(4-bit number $b_1b_2b_3b_4$ being identified, as usual, with the decimal number $b_12^3 + b_22^2 + b_32^1 + b_42^0$) are 0/1 balanced, nonlinear and fulfill the maximum entropy criterion. But from the discrete entropy point of view, they are quite different. $S_1$ consists of two cycles of length 7 and two fixed points. Its discrete entropies are:

$$\bar{H}_1(S_1) = 0.99; \quad \bar{H}_2(S_1) = 1.04; \quad \bar{H}_4(S_1) = 0.96; \quad \bar{H}_8(S_1) = 1.04; \quad \bar{H}_16(S_1) = 0.98; \quad \bar{H}_32(S_1) = 0.95;$$

and $h_0(S_1) = 0.85$. $S_2$ consists of two cycles of lengths 12 and 4, with

$$\bar{H}_1(S_2) = 0.99; \quad \bar{H}_2(S_2) = 1.08; \quad \bar{H}_4(S_2) = 1.17; \quad \bar{H}_8(S_2) = 3.59/(r - 1)$$

thus $h_0(S_2) = 0.68$. As expected, the discrete entropy of $S_1$ is higher and, consequently, it generates more pseudo-randomness that $S_2$.

### Conclusion

The basic conceptual framework of discrete chaos has been presented and potential applications to cryptography have been illustrated with the analysis of some S-boxes — mainly, those of AES.

### References


