Cryptographically Secure Substitutions Based on the Approximation of Mixing Maps

Janusz Szczepanski, José M. Amigó, Tomasz Michalek, and Ljupco Kocarev

***Abstract*—In this paper, we explore, following Shannon's suggestion that diffusion should be one of the ingredients of resistant block ciphers, the feasibility of designing cryptographically secure substitutions (think of S-boxes, say) via approximation of mixing maps by periodic transformations. The expectation behind this approach is, of course, that the nice diffusion properties of such maps will be inherited by their approximations, at least if the convergence rate is appropriate and the associated partitions are sufficiently fine. Our results show that this is indeed the case and that, in principle, block ciphers with close-to-optimal immunity to linear and differential cryptanalysis (as measured by the linear and differential approximation probabilities) can be designed along these guidelines. We provide also practical examples and numerical evidence for this approximation philosophy.**

***Index Terms*—Black cipher, differential cryptanalysis, linear cryptanalysis, mixing dynamical system, periodic approximation, S box.**

I. Introduction

CRYPTOGRAPHY has come to be understood as the science of secure communication. The publication in 1949 by C. E. Shannon of the paper "*Communication Theory of Secrecy Systems*" ushered in the era of *scientific secret-key cryptography* [1] by providing a theory of secrecy systems almost as comprehensive as the theory of communication he had published one year before. In his masterpiece, Shannon laid the mathematical foundations of secrecy systems, proved under what conditions a perfect security may exist and proposed design principles for practical encryption systems. Two methods were suggested as basic principles: diffusion and confusion. In the method of diffusion, "the statistical structure of the message which leads to its redundancy is dissipated into long range statistics —i.e., into statistical structure involving long combinations of letters in the cryptogram". The method of confusion seeks "to make the relation between the simple statistics of the ciphertext and the simple description of the key a very complex and involved one". Shannon also suggested mixing transformations to be used for practical encryption systems. He wrote:

Manuscript received March 26, 2004; revised August 25, 2004. This paper was recommended by Associate Editor Y. Nishio.

J. Szczepanski is with the Institute for Fundamental Technological Research, Polish Academy of Sciences, PL-00-049 Warsaw, Poland and also with the Trust and Certification Centre "CENTRAST" Co., PL-00-049 Warsaw, Poland (e-mail: jszczepa@ippt.gov.pl).

J. M. Amigó is with Centro de Investigación Operativa, Universidad Miguel Hernández, 03202 Elche, Spain (e-mail: jm.amigo@umh.es).

T. Michalek is with the Institute for Fundamental Technological Research, Polish Academy of Sciences, PL-00-049 Warsaw, Poland.

L. Kocarev is with the Institute for Nonlinear Science, University of California, San Diego, La Jolla, CA 92093-0402 USA (e-mail :lkocarev@ucsd.edu).

Digital Object Identifier 10.1109/TCSI.2004.841602

"Good mixing transformations are often formed by repeated products of two simple noncommuting operations. Hopf has shown, for example, that pastry dough can be mixed by such a sequence of operations. The dough is first rolled out into a thin slab, then folded over, then rolled, and then folded again, etc. In a good mixing transformation of a space with natural coordinates $x_1, x_2, \ldots, x_N$, the point x_i is carried by the transformation into a point x_i', with

$$x_i' = f_i(x_1, x_2, \ldots, x_N), \qquad i = 1, 2, \ldots, N,$$

and the functions f_i are complicated, involving all variables in a "sensitive" way. A small variation of any one, x_3, say, changes all the x_i' considerably. If x_3 passes through its range of possible variation, the point x_i' traces a long winding path around the space."

Suppose we have a probability or measure space X and a measure-preserving transformation $T : X \to X$. The transformation T possesses the *mixing property* (or simply, is mixing) if, for any two measurable sets $A_1, A_2 \subset X$,

$$\lim_{n\to\infty} \mu(A_1 \cap T^{-n} A_2) = \mu(A_1)\mu(A_2). \tag{1}$$

Therefore, given any set A_2 of positive measure, any set A_1 of positive measure will always intersect the set A_2 as it evolves with n from some initial 'time', and the measure of that part of A_1 which is contained in A_2 at the moment n is asymptotically proportional (for $n \to \infty$) to the measure of A_2. At the end, A_1 dilutes, so to speak, in the whole space X, becoming there an homogeneous stationary distribution. It is precisely this property which explains the origin of the following expression: "A set A_1 of positive measure in its evolution mixes uniformly in the phase space." If we think now of the set of possible messages as the phase space X and the set of sensible (high probability) messages as an initial region in X, then the mixing property of the dynamics implies the *spreading out of the influence of a single message digit over many ciphertext digits* and, by the same token, the scattering of the high probability messages (initial region) through the entire phase space as time goes on.

Cryptography is generally acknowledged as the best method of data protection against passive and active fraud [2]. Block ciphers transform a relatively short string (typically 64, 128, or 256 bits) to a string of the same length under control of a secret key. Several block encryption ciphers based on chaotic maps have been proposed in literature, in which a discretization (process that describes the way a chaotic map is implemented in the computer) is not realized by rounding the chaotic map according to the computer arithmetic, but rather is constructed

explicitly. Pichler and Scharinger [3] proposed cryptographic systems based on chaotic permutations constructed by explicitly discretizing the two dimensional baker's map. Fridrich [4] extended their ideas to chaotic permutations on any size of two dimensional lattices. Her permutations benefit from the expanding property along one axis, technically avoiding the contracting property along the other axis. The authors of [5] used two well-known chaotic maps, exponential and logistic, to construct a class of block encryption algorithms. In a recent paper [6], they analytically derived the lower bound of a number of active S-boxes in their algorithms, computed upper bounds for differential and linear probabilities, and therefore, proved the resistance of the algorithms proposed in [5] to differential and linear attacks. Masuda and Aihara [7] considered a discrete version of the skew-tent map, which exploits important chaotic properties such as the sensitive dependence on initial conditions and the exponential information decay. They discussed the difference between the discretized map and the original map, explaining the ergodic-like and chaotic-like properties of the discretized map.

In this paper, we explore the feasibility of designing cryptographically secure substitutions via approximation of mixing maps by periodic transformations. In order to adapt the continuous dynamics of a mixing map to the block structure of a cryptosystem and, most importantly, to assure that the good statistical properties of the first are transferred to the second, we use periodic approximations of dynamical systems. The approximation of automorphisms of measure spaces by periodic automorphisms first appeared in the works of Halmos and Rohlin. Clearly, one expects that the better the mixing properties of the approximated dynamical system, the better the cryptographic properties of the discrete maps obtained in the process of approximation. In general, security evaluation of block ciphers consists of three steps: first, one should prove the resistance to differential and linear attacks; second, one should check for the extensions and generalizations of differential and linear attacks; and third, one should take into account several dedicated attacks applicable to cipher with a small number of rounds. However, one should keep in mind that provable security against one or two important attacks *does not imply* that the cipher is secure: other attacks may exist. On the other hand, provable security against certain attacks is certainly a first step in the right direction. The subject of the present paper is precisely the relation between the dynamical system used for encryption and the quality of the resulting cryptosystem as quantified by its immunity to linear and differential cryptanalysis, which are currently the benchmarks for measuring the safety of the standard block ciphers.

Let F be a permutation of n-bit blocks and, as usual, denote by LP_F and DP_F the linear approximation probability and differential approximation probability of F, respectively (see Sections III and IV for precise definitions of these "probabilities"). LP_F and DP_F measure the immunity of the block cipher F to attacks mounted on the corresponding cryptanalysis, immunity being higher the smaller their values. We show in this paper that if F is a cyclic periodic approximation of a mixing automorphism and some assumptions are fulfilled, then LP_F and DP_F get asymptotically close to their greatest lower bounds $1/2^n$ and $1/2^{n-1}$, respectively, thus obtaining an arbitrarily close-to-optimal immunity to both cryptanalyzes. Therefore, we prove, as suggested by Shannon, that mixing transformations may indeed be used in encryption systems, providing an alternative to the traditional algebraic methods.

This is the outline of the paper. The notation and the necessary framework on approximation of dynamical systems are set in Section II. Although this section is slightly technical because of the generality with which it is formulated, we are going however to apply its results in rather familiar settings where technicalities can be dispensed with. In Section III we briefly review (following [8]) the properties of the substitution ciphers constructed approximating ergodic automorphisms from the point of view of linear cryptanalysis. Section III also paves the way to Section IV, devoted to differential cryptanalysis, which contains the main theoretical results. All these sections have been provided with concrete calculations in order to clarify the basic ideas. Finally, in Section V, we discuss some implementations of our approximation-based approach.

II. Notations and Preliminaries

In ergodic theory, one may study the relationship between the properties of dynamical systems and the speed of their approximations by dynamical systems of some particular fixed class, such as periodic dynamical systems. The following definitions and theorems related to this topic can be found in [9].

Let X be a set, $\mathcal{A}$ a σ-algebra of subsets of X and μ a positive measure on $(X, \mathcal{A})$. A finite measure space $(X, \mathcal{A}, \mu)$ is a *Lebesgue space* if it is isomorphic (in the sense of measure theory) to an interval of $\mathbb{R}^n$ together with countably many point masses. Suppose T is an automorphism of the Lebesgue space $(X, \mathcal{A}, \mu)$, i.e., T is a one-to-one map of X onto itself such that, for all $A \in \mathcal{A}$, we have $TA, T^{-1}A \in \mathcal{A}$ and $\mu(A) = \mu(TA) = \mu(T^{-1}A)$. We shall consider sequences of finite partitions $\{\mathcal{P}_n\}$ of the space X and sequences of automorphisms $\{T_n\}$ such that T_n preserves $\mathcal{P}_n$. The automorphism T_n *preserves the partition* $\mathcal{P}_n$, if it sends every element of $\mathcal{P}_n$ into an element of the same partition. The elements of $\mathcal{P}_n$ will be denoted by $P_k^{(n)}$, $1 \leq k \leq q_n$ (we will eventually drop the upper index of the partition elements to ease the notation). By $\mathcal{A}(\mathcal{P}_n)$ we denote the σ-algebra of subsets of the space X consisting of elements of $\mathcal{P}_n$ (except, possibly, for zero-measure sets). The notation $\mathcal{P}_n \to \mathcal{E}$ (when $n \to \infty$), where $\mathcal{E}$ is the partition of X into separate points, means that, for each $A \in \mathcal{A}$, there is a sequence of sets $A_n \in \mathcal{A}(\mathcal{P}_n)$ such that $\lim_{n\to\infty} \mu(A_n \triangle A) \to 0$, where $\triangle$ stands for symmetric set difference. Since the number of elements of the partition $\mathcal{P}_n$ is finite, the trajectory of each $P_k^{(n)}$ is finite, i.e., for some r_k, $1 \leq k \leq q_n$, we have $T_n^{r_k} P_k^{(n)} = P_k^{(n)}$. For the sequel it is not important how $T_n^{r_k}$ interchanges the points within $P_k^{(n)}$, but it is convenient to assume that $T_n^{r_k} x = x$ for any point $x \in P_k^{(n)}$. By t_n we denote the order of T_n, i.e., the smallest natural number such that $T_n^{t_n}$ is the identity.

Finally, $L^2(X, \mathcal{A}, \mu)$ will denote, as usual, the Hilbert space of the square-integrable functions $g : X \to \mathbb{C}$ with scalar product $\langle g_1, g_2 \rangle = \int_X g_1(x)\overline{g_2(x)}d\mu(x)$ and norm $\|g\| = \langle g, g \rangle^{1/2}$. Given an automorphism T of $(X, \mathcal{A}, \mu)$, the

unitary operator $U_T : L^2(X, \mathcal{A}, \mu) \to L^2(X, \mathcal{A}, \mu)$ defined by $U_T(g) = g \circ T$ is called the *Koopman operator induced by* T.

Definition 1: Suppose f is a function on the integers such that $f(n) \to 0$ monotonically.

1) An automorphism T of the space $(X, \mathcal{A}, \mu)$ possesses an approximation of the first type by periodic transformations (a.p.t. I) with speed $f(n)$, if there exists a sequence of partitions $\mathcal{P}_n \to \mathcal{E}$ and a sequence of automorphisms T_n preserving $\mathcal{P}_n$ such that

$$\sum_{k=1}^{q_n} \mu\left(TP_k^{(n)} \triangle T_n P_k^{(n)}\right) < f(q_n), \qquad n = 1, 2, \ldots.$$

2) If for the sequences $\{\mathcal{P}_n\}$, $\{T_n\}$, where T_n is a periodic automorphism of order t_n, we have the inequality

$$\sum_{k=1}^{q_n} \mu\left(TP_k^{(n)} \triangle T_n P_k^{(n)}\right) < f(t_n), \qquad n = 1, 2, \ldots,$$

and $U_{T_n} \to U_T$ in the strong topology of operators in $L^2(X, \mathcal{A}, \mu)$, then T possesses an approximation of the second type by periodic transformations (a.p.t. I) with speed $f(n)$.

3) If the automorphism T possesses an a.p.t. I and T_n cyclically permutes the elements of $\mathcal{P}_n$, then T is said to possess a cyclic approximation with speed $f(n)$.

Equivalently, we will also say that $(T_n, \mathcal{P}_n)$ is a periodic approximation (of the corresponding type) of the automorphism T. Let us illustrate these concepts with a classical example.

Example: Given an irrational number α, let T be the rotation of the circle S^1 by the angle α, i.e., $Tx = x + \alpha \pmod 1$, $x \in [0, 1)$. Thus T is an aperiodic automorphism of $X = [0, 1)$. It is known from the theory of continuous fractions [10] that for every irrational α, there exists a sequence of irreducible fractions $\{p_n/q_n\}_{n=1}^{\infty} \to \alpha$ such that

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{\sqrt{5} q_n^2}, \qquad n = 1, 2, \ldots.$$

(The number $\sqrt{5}$ is the best possible one: any larger value renders the proposition false.) Taking $\mathcal{P}_n = \{P_k^{(n)} : 1 \le k \le q_n\}$ with

$$P_k^{(n)} = \left[\frac{k-1}{q_n}, \frac{k}{q_n}\right) \tag{2}$$

and the sequence of rotations of the circle T_n by the angles p_n/q_n (so that $P_k^{(n)} \mapsto P_l^{(n)}$ with $l \equiv k + p_n \bmod q_n$ and $T_n^{q_n}$ is the identity on $\mathcal{P}_n$), it follows that

$$\sum_{k=1}^{q_n} \mu\left(TP_k^{(n)} \triangle T_n P_k^{(n)}\right) = \sum_{k=1}^{q_n} 2\left|\alpha - \frac{p_n}{q_n}\right| < 2q_n \frac{1}{\sqrt{5} q_n^2} = \frac{2}{\sqrt{5} q_n}.$$

This proves that $(T_n, \mathcal{P}_n)$ is a cyclic approximation of T with speed $f(n) = 2/(\sqrt{5}n)$.

The Rohlin–Halmos Lemma [9] states that, if T is an automorphism of the Lebesgue space $(X, \mathcal{A}, \mu)$, then for all $\varepsilon > 0$ and every positive integer n there exists a set $A \in \mathcal{A}$ (depending on ε and n) such that (*i*) the sets $A, TA, \ldots, T^{n-1}A$ are disjoint, and (*ii*) $\mu\left(\bigcup_{i=0}^{n-1} T^i A\right) > 1 - \varepsilon$. It follows that any automorphism has periodic approximations. In fact, one defines a periodic approximation $(T_n, \mathcal{P}_n)$ of T by putting

$$T_n(x) = \begin{cases} T(x), & \text{if } x \in \bigcup_{i=0}^{n-2} T^i A_n \\ T^{-n+1}(x), & \text{if } x \in T^{n-1} A_n \\ x, & \text{if } x \in X \setminus \bigcup_{i=0}^{n-1} T^i A_n \end{cases}$$

where $A_n \in \mathcal{A}$ is the set whose existence is guaranteed by the Rohlin–Halmos Lemma for $\varepsilon = 1/n$; the corresponding partition $\mathcal{P}_n$ is also defined by means of A_n. Unfortunately, the proof of the Rohlin–Halmos Lemma is not constructive, so that periodic approximations of aperiodic automorphisms are only known in simple cases, like for rigid motions in the circle (see previous example), disc, sphere and torus, while their construction in more general cases remains a challenging task.

A different question is the relation between the properties of the approximating periodic transformations and the approximated one. Intuition says that the faster the approximation speed, the worse the statistical properties of the approximated automorphism, e.g., ergodicity and mixing. On the other hand, a "good" cyclic approximation of an automorphism should guarantee its ergodicity. In fact, the following results can be proved [9].

Theorem 1:

1) Any automorphism T possesses a.p.t. I with speed $f(n) = a_n/\log n$, where a_n is an arbitrary monotonic sequence of real numbers tending to infinity.
2) If the automorphism T possesses a cyclic a.p.t. I with speed θ/n and $\theta < 4$, then T is ergodic.
3) If the automorphism T possesses an a.p.t. II with speed θ/n and $\theta < 2$, then T is not mixing.

The periodic approximation $(T_n, \mathcal{P}_n)$ can be viewed as a coarse-grained approximation of the actual dynamic T. Once the approximation speed has been correctly tuned, T_n is asymptotically the best approximation of T as far as its diffusion properties are concerned.

Example (continued): For the rotation of the circle by an irrational angle, we have shown before that $\theta = 2/\sqrt{5} < 1$, so that we recover from *Theorem 1* 2) the otherwise easy-to-prove result that such an automorphism is ergodic. It is also straightforward to show that T is not mixing directly from the definition, (1) (take, for instance, $A_1 = [1/4, 1/2)$, $A_2 = [3/4, 1)$ and see what happens when $n \to \infty$ and α is irrational). Let us show, however, how this follows from *Theorem 1* 3) in a rather simple fashion too.

In fact, since the order of T_n equals the cardinality of $\mathcal{P}_n$ (i.e., $t_n = q_n$), we only need to check that $U_{T_n} \to U_T$ in the strong topology of operators in $L^2([0, 1)^2)$ ($\equiv L^2([0, 1)^2, \mathcal{B}, \lambda)$, where $\mathcal{B}$ is the Borel σ-algebra of $[0, 1)^2$ and λ the corresponding Lebesgue measure) in order to prove that T possesses an a.p.t. II with speed $f(n) = 2/(\sqrt{5}n)$ [see *Definition 1* 2)]. According to *Theorem* 1 3), it will follow then that T is not mixing.

The Koopman operators $U_T,\ U_{T_n} : L^2([0,1)^2) \to L^2([0,1)^2)$ in question are given by

$$U_T(g)(x,y) := (g \circ T)(x,y) = g((x+\alpha, y+x) \bmod 1)$$

and

$$\begin{aligned} U_{T_n}(g)(x,y) &:= (g \circ T_n)(x,y) \\ &= g\left(\left(x + \frac{p_n}{q_n}, y + x\right) \bmod 1\right) \end{aligned}$$

for all $g \in L^2([0,1)^2)$. Thus,

$$\begin{aligned} \|U_T - U_{T_n}\| &= \sup_{\|g\| \le 1} \|U_T(g) - U_{T_n}(g)\| \\ &= \sup_{\|g\| \le 1} \|g \circ T - g \circ T_n\|. \end{aligned}$$

Now, since $C_0^\infty([0,1)^2) := \{h : [0,1)^2 \to \mathbb{C},\ h \text{ smooth with support in } (0,1)^2\}$ is dense in $L^2([0,1)^2)$ [11], there exists $g_\varepsilon \in C_0^\infty([0,1)^2)$ such that $\|g_\varepsilon - g\| < \varepsilon$ for every $\varepsilon > 0$. Therefore, for all $g \in L^2([0,1)^2)$, we have

$$\begin{aligned} \|g \circ T - g \circ T_n\| &\le \|(g - g_\varepsilon) \circ T\| + \|g_\varepsilon \circ (T - T_n)\| \\ &\quad + \|(g - g_\varepsilon) \circ T_n\| \\ &= 2\|g - g_\varepsilon\| + \|g_\varepsilon \circ (T - T_n)\| \end{aligned}$$

where

$$\begin{aligned} \|g_\varepsilon \circ (T - T_n)\|^2 &= \int_{[0,1)^2} \Big| g_\varepsilon(x+\alpha, y+x) \\ &\qquad - g_\varepsilon\left(x + \frac{p_n}{q_n}, y + x\right)\Big|^2 dxdy \\ &= O\left(\left|\alpha - \frac{p_n}{q_n}\right|^2\right). \end{aligned}$$

Since both $\|g - g_\varepsilon\|$ and $\|g_\varepsilon \circ (T - T_n)\|$ can be done arbitrarily small by choosing ε and n in convenient ways, it follows $\lim_{n\to\infty} \|U_T - U_{T_n}\| = 0$.

III. Linear Approximation Probability

Linear cryptanalysis was proposed by Matsui [12]. This attack on block ciphers exploits the statistical inhomogeneities of certain expressions called linear approximations involving different digits of the plaintext, ciphertext and the key of the encryption round considered. We go first over the basics we need. Given any pair of n-bit blocks $\alpha = (\alpha_1, \ldots, \alpha_n) \neq 0$, $\beta = (\beta_1, \ldots, \beta_n) \neq 0$, and a map $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$, define $\Phi_{\alpha,\beta}^F : \mathbb{Z}_2^n \to \mathbb{Z}_2$ as

$$\xi = (\xi_1, \ldots, \xi_n) \mapsto \Phi_{\alpha,\beta}^F(\xi) = 1 \oplus \xi \circ \alpha \oplus F(\xi) \circ \beta \quad (3)$$

where, as usual, $\oplus$ denotes addition in $\mathbb{Z}_2$ (also called the XOR operation), and $\xi \circ \alpha := \xi_1\alpha_1 \oplus \cdots \oplus \xi_n\alpha_n$ is the parity of the bitwise product of ξ and α (and analogously for $F(\xi) \circ \beta$). Next we define the linear approximation probability of F (LP_F for short) as

$$LP_F = \max_{\alpha,\beta \neq 0} LP_F(\alpha, \beta) \quad (4)$$

where

$$LP_F(\alpha, \beta) = (2p-1)^2 = 4\left(p - \frac{1}{2}\right)^2 \quad (5)$$

and

$$\begin{aligned} p &= \frac{\#\{\xi \in \mathbb{Z}_2^n : \xi \circ \alpha = F(\xi) \circ \beta\}}{2^n} \\ &= \frac{\#\{\xi \in \mathbb{Z}_2^n : \Phi_{\alpha,\beta}^F(\xi) = 1\}}{2^n}. \end{aligned}$$

Therefore, the linear approximation probability LP_F is the square of the maximal imbalance of the following event: the parity of the input bits selected by the mask α is equal to the parity of the output bits selected by the mask β. Since [12]

$$\sum_{\alpha \in \mathbb{Z}_2^n} LP_F(\alpha, \beta) = 1(\beta \in \mathbb{Z}_2^n)$$

one gets $LP_F \ge 1/2^n$. Immunity of F to linear cryptanalysis means that $LP_F(\alpha, \beta)$ should be uniformly distributed in α (resp. β) for fixed β (resp. α) so that

$$LP_F(\alpha, \beta) \simeq \frac{1}{2^n}\ (\forall\, \alpha, \beta \in \mathbb{Z}_2^n)$$

or, equivalently, $LP_F \simeq 1/2^n$. Observe for further references that, if F is a *cyclic* permutation (i.e., $\mathrm{orb}(\xi) := \{\xi, F\xi, F^2\xi, \ldots, F^{2^n-1}\xi\} = \mathbb{Z}_2^n$ for all $\xi \in \mathbb{Z}_2^n$), then (5) can be written as

$$LP_F(\alpha, \beta) = 4\left(\frac{1}{2^n}\sum_{i=0}^{2^n-1} \Phi_{\alpha,\beta}^F(F^i(\xi)) - \frac{1}{2}\right)^2 \quad (6)$$

independently of $\xi \in \mathbb{Z}_2^n$.

Let $(X, \mathcal{A}, \mu)$ be a Lebesgue space and $T : X \to X$ an automorphism. Let T_n be a periodic approximation of T and $\mathcal{P}_n = \{P_k : 1 \le k \le q_n\}$ the corresponding partition of X. For definiteness and without restriction, we choose $q_n = 2^n$ for the time being, this choice being dictated by the applications to the design of n-bit S-boxes we have in mind; other choices will be made in Section V. With this proviso, let $\Lambda : \mathcal{P}_n \to \mathbb{Z}_2^n$ be a one-to-one map that associates to each element $P_k \in \mathcal{P}_n$ an n-bit block $\overline{P_k} \in \mathbb{Z}_2^n$, $\Lambda(P_k) = \overline{P_k}$. Hence, Λ labels the elements $P_k \in \mathcal{P}_n$ with binary words of length n like, for instance, the assignment $P_k \mapsto \kappa = (\kappa_1, \ldots, \kappa_n)$, where $k = \kappa_1 \cdot 2^{n-1} + \cdots + \kappa_{n-1} \cdot 2 + \kappa_n$, does. In this setup, the action of T_n on the elements of $\mathcal{P}_n$, $P_k \mapsto P_l$, induces the obvious permutation $\overline{T}_n = \Lambda T_n \Lambda^{-1} : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$, $\overline{T}_n(\overline{P_k}) = \overline{P_l} = \overline{T_n(P_k)}$ on their n-bit Λ-labels. Needless to say, we expect that the diffusion properties the automorphism T has if it is mixing, will not be completely lost on the way from T to our substitution candidate $\overline{T}_n$ via T_n and Λ.

For $x \in X$, let $P_{k(x)} \equiv P_x$ be the element of the partition $\mathcal{P}_n$ such that $x \in P_x$ and, mimicking (3) and (6) with $\xi = \overline{P_x}$, $F(\xi) = \overline{P_{T(x)}}$, define the binary function $\Phi_{\alpha,\beta}^T : X \to \mathbb{Z}_2$ (subordinated to the partition $\mathcal{P}_n$) as

$$\Phi_{\alpha,\beta}^T(x) := 1 \oplus \overline{P_x} \circ \alpha \oplus \overline{P_{T(x)}} \circ \beta$$

and also

$$LP_T(\alpha,\beta)(x) := 4\left(\frac{1}{2^n}\sum_{i=0}^{2^n-1}\Phi^T_{\alpha,\beta}(T^i(x)) - \frac{1}{2}\right)^2. \quad (7)$$

According to these definitions

$$\begin{aligned}\Phi^{T_n}_{\alpha,\beta}(x) &= 1 \oplus \overline{P_x} \circ \alpha \oplus \overline{P_{T_n(x)}} \circ \beta \\ &= 1 \oplus \overline{P_x} \circ \alpha \oplus \overline{T_n(P_x)} \circ \beta \\ &= 1 \oplus \overline{P_x} \circ \alpha \oplus \overline{T}_n(\overline{P_x}) \circ \beta = \Phi^{\overline{T}_n}_{\alpha,\beta}(\overline{P_x})\end{aligned}$$

which shows [see (7) with T_n instead of T, and (6) with $F(\xi)$ replaced by $\overline{T}_n(\overline{P_x})$] that, for cyclic approximations $(T_n, \mathcal{P}_n)$, the equality

$$\begin{aligned}LP_{T_n}(\alpha,\beta) &= 4\left(\frac{1}{2^n}\sum_{i=0}^{2^n-1}\Phi^{T_n}_{\alpha,\beta}(T^i_n(x)) - \frac{1}{2}\right)^2 \\ &= LP_{\overline{T}_n}(\alpha,\beta) \quad (8)\end{aligned}$$

holds independently of x. The expression (8) relates a cryptographic quantity (a measure of immunity to linear cryptanalysis) to a dynamical quantity (average of $\Phi^{T_n}_{\alpha,\beta}$ along orbits), allowing to study the first one with the help of the second. Moreover, it gives us a handle for calculating the LP-value of the substitution $\overline{T}_n$ by means of the properties of T and T_n. In fact, the following theorem is proved in [8].

Theorem 2: Let $T : X \to X$ be a uniquely ergodic automorphism with invariant measure μ. Furthermore, suppose $(T_n, \mathcal{P}_n)$, $\mathcal{P}_n = \{P_k : 1 \le k \le 2^n\}$, is a cyclic approximation of T with speed $\theta/2^n$ $(0 < \theta < 4)$, such that

$$\mu(\Phi^T_{\alpha,\beta}) := \int_X \Phi^T_{\alpha,\beta}(x)d\mu(x) = \frac{1}{2} \quad (9)$$

and

$$\max_{1\le i\le 2^n-1}\sum_{k=1}^{2^n}\mu\left(T^{-i}P_k \triangle T_n^{-i}P_k\right) < \frac{\theta}{2^n}. \quad (10)$$

Then

$$LP_{\overline{T}_n} - \frac{1}{2^n} \le \frac{8+4\theta}{2^{3n/2}}.$$

The ergodic automorphism $T : X \to X$ is said to be uniquely ergodic if there exists on X only one T-invariant measure (up to normalization). Cyclic approximations with speed $f(n)$ were defined in *Definition* 1 3). If $\mu(P_k) = 1/2^n$ for $1 \le k \le 2^n$ (or, more generally, all partition elements P_k have the same measure), the condition (9) is redundant since then it follows from the ergodicity of T [8]. The proof of the following corollary can be also found in [8].

Corollary 3: If $LP_{\overline{T}_n}(\alpha,\beta)$ is asymptotically uniformly distributed with respect to α and β, then the approximation speed of $(T_n, \mathcal{P}_n)$ to the ergodic automorphism T is $\theta/2^n$ with $2 \le \theta < 4$.

According to *Theorem* 1, if $(T_n, \mathcal{P}_n)$ is a periodic approximation of T of the second type with $\theta < 2$, then T is not mixing. Therefore, *Corollary* 3 suggests that to get by this method a cryptosystem immune to linear cryptanalysis, ergodicity and approximations of the first type might not be enough, rather one should use a mixing automorphism and approximations of the second type.

Example (continued): Suppose now $q_n = 2^n$, $n \ge 3$, and enumerate lexicographically the partition elements (2) by means of n-bit blocks. Let $\pi = (\pi_1, \ldots, \pi_n) \in \mathbb{Z}_2^n$. Then, the rotation of the circle T_n by the rational angle p_n/q_n, with $p_n \pmod{q_n} = \pi_1 \cdot 2^{n-1} + \cdots + \pi_{n-1} \cdot 2 + \pi_n$, generates the block substitution

$$\overline{T}_n(\xi) = \xi + \pi(\xi \in \mathbb{Z}_2^n)$$

where $+$ denotes the sum modulo 2 of ξ and π with carry and deletion of the $(n+1)$-th leftmost digit if it is 1 (to account for the modulo 2 congruence). In order to keep track of the carries without introducing too much notational machinery, let us take a concrete π, say $\pi = (1, 0, \ldots, 0, 1)$ for simplicity, so that

$$\overline{T}_n(\xi) = (\overline{\xi_1}, \overline{\overline{\xi_2}}, \ldots, \overline{\overline{\xi_{n-1}}}, \overline{\xi_n}) \quad (11)$$

where, as usual, $\overline{\xi_i}$, $i = 1, n$, denotes the complementary digit of ξ_i (i.e., $\overline{\xi_i} = 1$ if $\xi_i = 0$ and $\overline{\xi_i} = 0$ if $\xi_i = 1$) and we use the shorthand

$$\overline{\overline{\xi_{n-k}}} := \begin{cases} \overline{\xi_{n-k}}, & \text{if } \xi_{n-k+1} = \cdots = \xi_n = 1 \\ \xi_{n-k}, & \text{otherwise} \end{cases} \quad (12)$$

for $k = 1, \ldots, n-2$. Therefore, $\xi \circ \alpha = \overline{T}_n(\xi) \circ \beta$ if and only if

$$\begin{aligned}&\xi_1\alpha_1 \oplus \xi_2\alpha_2 \oplus \cdots \oplus \xi_{n-1}\alpha_{n-1} \oplus \xi_n\alpha_n \\ &\qquad = \overline{\xi_1}\beta_1 \oplus \overline{\overline{\xi_2}}\beta_2 \oplus \cdots \oplus \overline{\overline{\xi_{n-1}}}\beta_{n-1} \oplus \overline{\xi_n}\beta_n. \quad (13)\end{aligned}$$

Thus, for masks $\alpha = \beta \ne 0$ with $\alpha_1 = \alpha_n = 0$, the condition (13) holds at least for all 2^{n-1} blocks ξ with $\xi_n = 0$, since then $\overline{\overline{\xi_2}} = \xi_2, \ldots, \overline{\overline{\xi_{n-1}}} = \xi_{n-1}$. Hence, for such masks

$$LP(\alpha,\beta) = \frac{\#\{\xi \in \mathbb{Z}_2^n : \xi \circ \alpha = \overline{T}_n(\xi) \circ \beta\}}{2^n} \ge \frac{1}{2}$$

and, consequently

$$LP_{\overline{T}_n} \ge \frac{1}{2}.$$

It should be clear that other choices of $\pi \ne 0$ are dealt with in a similar manner and lead to the same result. We conclude that the ergodicity of the aperiodic rotations of the circle does not suffice to obtain block substitutions with good properties from the standpoint of linear cryptanalysis.

IV. Differential Approximation Probability

Resistance against differential cryptanalysis of a block cipher with key K, E_K, means that, for every fixed nonzero input difference to E_K, none of the output differences occurs with high probability [13]. In order to formulate this concept mathematically, let $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ be a permutation of n-bit blocks. The immunity of F to differential cryptanalysis is measured by the

differential approximation probability of F (DP_F for short) as defined by

$$DP_F = \max_{\alpha \neq 0, \beta} DP_F(\alpha, \beta) \tag{14}$$

where α is the input difference, β the output difference and

$$DP_F(\alpha, \beta) = \frac{\#\{\xi \in \mathbb{Z}_2^n : F(\xi) \oplus F(\xi \oplus \alpha) = \beta\}}{2^n}.$$

Here $\xi \oplus \alpha = (\xi_1 \oplus \alpha_1, \ldots, \xi_n \oplus \alpha_n)$ denotes the componentwise XOR (or vector addition modulo 2) of the n-bit blocks ξ and α (and analogously for $F(\xi) \oplus F(\xi \oplus \alpha)$). It follows

$$\sum_{\beta \in \mathbb{Z}_2^n} DP_F(\alpha, \beta) = 1, \qquad \alpha \neq 0.$$

The smaller DP_F, the better the immunity of the block cipher F to differential cryptanalysis. As compared to $LP_F \geq 1/2^n$, the worse lower bound $DP_F \geq 1/2^{n-1}$ holds now because with $\eta \in H := \{\xi \in \mathbb{Z}_2^n : F(\xi) \oplus F(\xi \oplus \alpha) = \beta\}$ trivially $\eta \oplus \alpha \in H$ as well. With the approach followed in Section III in mind, we notice that, if F is a *cyclic* permutation, then $DP_F(\alpha, \beta)$ can also be written as an average on orbits of F, namely

$$DP_F(\alpha, \beta) = \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \chi_H(F^i(\xi)) \tag{15}$$

independently of $\xi \in \mathbb{Z}_2^n$, where χ_H is the characteristic function of the set $H \subset \mathbb{Z}_2^n$.

Let $(X, \mathcal{A}, \mu)$ be a Lebesgue space and $T : X \to X$ a μ-invariant automorphism. Let T_n be a periodic approximation of T and $\mathcal{P}_n = \{P_k : 1 \leq k \leq 2^n\}$ the corresponding partition. As before, let $\Lambda : \mathcal{P}_n \to \mathbb{Z}_2^n$ be a one-to-one labeling of the subsets P_k with n-bit blocks and let $\overline{T}_n$ denote the substitution $\Lambda(P_k) \equiv \overline{P}_k \mapsto \Lambda(T_n(P_k)) \equiv \overline{T_n(P_k)}$ induced on $\mathbb{Z}_2^n$ by the periodic approximation T_n. For $x \in X$, let $P_{k(x)} \equiv P_x$ be the element of the partition $\mathcal{P}_n$ such that $x \in P_x$ and set

$$\begin{aligned} H_{\alpha,\beta}^n &= \{x \in X : \overline{T_n}(\overline{P_x}) \oplus \overline{T}_n(\overline{P}_x \oplus \alpha) = \beta\} \\ &= \{x \in X : \overline{T_n(P_x)} \oplus \overline{T}_n(\overline{P}_x \oplus \alpha) = \beta\} \\ &= \{x \in X : \overline{P_{T_n(x)}} \oplus \overline{T}_n(\overline{P}_x \oplus \alpha) = \beta\}. \end{aligned} \tag{16}$$

Let us notice that, if the periodic approximation T_n is cyclic, the quantity

$$DP_{T_n}(\alpha, \beta)(x) := \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \chi_{H_{\alpha,\beta}^n}(T_n^i(x)) \tag{17}$$

is independent of x; set $DP_{T_n}(\alpha, \beta) := DP_{T_n}(\alpha, \beta)(x)$. Furthermore, since $x \in H_{\alpha,\beta}^n$ if and only if $\overline{P}_x \in \{\xi \in \mathbb{Z}_2^n : \overline{T}_n(\xi) \oplus \overline{T}_n(\xi \oplus \alpha) = \beta\}$, the equality

$$DP_{T_n}(\alpha, \beta) = DP_{\overline{T}_n}(\alpha, \beta) \tag{18}$$

trivially holds (replace F by $\overline{T}_n$ in (15)). This equation is the equivalent of (8) for differential cryptanalysis and, given the automorphism T and its cyclic approximation (T_n, P_n), it permits in a similar way to calculate the DP-value of the ensuing substitution $\overline{T}_n$.

Analogously, define

$$DP_T(\alpha, \beta)(x) := \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \chi_{H_{\alpha,\beta}^n}(T^i(x)). \tag{19}$$

Before stating precisely our main result concerning the resistance of the block cipher $\overline{T}_n$ against differential cryptanalysis, we need the following lemma.

Lemma 4:

1) Let $T = S^m$ be the product of a μ-preserving automorphism $S : X \to X$ with itself $m \geq 1$ times. Suppose $(T_n, \mathcal{P}_n)$, $\mathcal{P}_n = \{P_k : 1 \leq k \leq 2^n\}$ is a cyclic approximation of T with speed $\theta/2^n$. If T fulfills the condition (10), then

$$\left| DP_{\overline{T}_n}(\alpha, \beta) - \int_X DP_T(\alpha, \beta)(x) d\mu(x) \right| \leq \frac{\theta}{2^n}, \qquad \alpha, \beta \in \mathbb{Z}_2^n.$$

2) If, furthermore, S is mixing and $2 < \theta < 4$, then

$$\int_X DP_T(\alpha, \beta)(x) d\mu(x) \leq \frac{1 + \varepsilon_m}{2^n}$$

where ε_m can be made arbitrarily small by taking m large enough.

Proof:

1) The proof of this part of the lemma is illustrated in Fig. 1. From (17)–(19) it follows the inequality

$$\left| DP_{\overline{T}_n}(\alpha, \beta) - DP_T(\alpha, \beta)(x) \right| \leq \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \left| \chi_{H_{\alpha,\beta}^n}(T_n^i(x)) - \chi_{H_{\alpha,\beta}^n}(T^i(x)) \right|$$

where $H_{\alpha,\beta}^n$ is the set (16). On integrating this inequality over X with the normalized measure μ, we get

$$\begin{aligned} &\left| DP_{\overline{T}_n}(\alpha, \beta) - \int_X DP_T(\alpha, \beta)(x) d\mu(x) \right| \\ &\leq \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \int_X \left| \chi_{H_{\alpha,\beta}^n}(T_n^i(x)) - \chi_{H_{\alpha,\beta}^n}(T^i(x)) \right| d\mu(x) \\ &= \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \int_X \left| \chi_{T_n^{-i}(H_{\alpha,\beta}^n)}(x) - \chi_{T^{-i}(H_{\alpha,\beta}^n)}(x) \right| d\mu(x) \\ &= \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \int_X \chi_{T_n^{-i}(H_{\alpha,\beta}^n) \Delta T^{-i}(H_{\alpha,\beta}^n)}(x) d\mu(x) \\ &= \frac{1}{2^n} \sum_{i=0}^{2^n - 1} \mu\left(T_n^{-i}(H_{\alpha,\beta}^n) \Delta T^{-i}(H_{\alpha,\beta}^n)\right) \end{aligned}$$

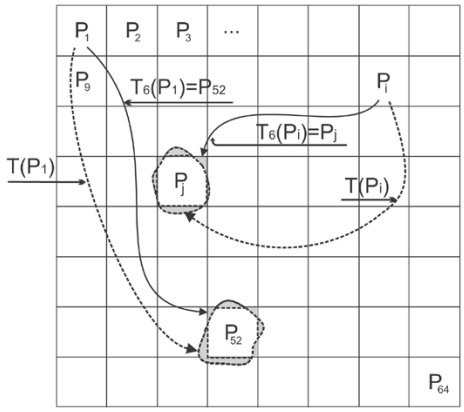

Fig. 1. Graphical illustration of the essence of periodic approximations. Here, the partition $\mathcal{P}_6$ (with $q_6 = 2^6 = 64$ elements) is depicted as a regular mesh for simplicity. The action of T_6 is exemplified on the partition elements $P = P_1$ and P_i only. The shadowed areas show the corresponding set difference $T(P) \triangle T_n(P)$ for $n = 6$. According to the approximation property this area tends to zero when n tends to infinity (i.e., when the partition becomes finer) with the rate $f(2^n) = \theta/2^n$. This estimation was used in the proof of point 1 at Lemma 4.

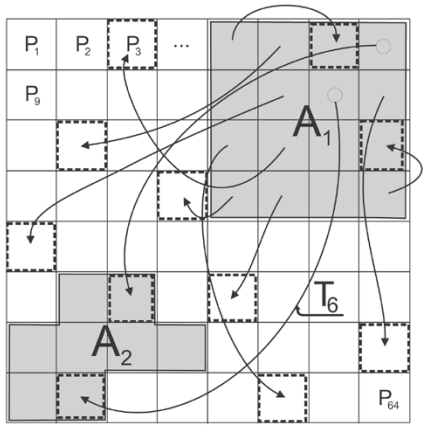

Fig. 2. Illustration of mixing. The mixing property of T implies that any region A_1 consisting of some number of blocks spreads uniformly over the whole space of blocks also under the action of the approximation T_n because of its shadowing property (see text). The number of blocks belonging to A_1 that after transformation by T_n reaches another region A_2 is proportional to the area of A_2 (see definition of mixing in the Introduction). This property was exploited in the proof of the second part of Lemma 4.

where, by the assumption (10)

$$\begin{aligned}
&\sum_{i=0}^{2^n-1} \mu\left(T_n^{-i}\left(H_{\alpha,\beta}^n\right) \Delta T^{-i}\left(H_{\alpha,\beta}^n\right)\right) \\
&\quad = \sum_{i=0}^{2^n-1}\sum_{k=1}^{2^n} \mu\left(T_n^{-i}(H_{\alpha,\beta}^n \cap P_k) \Delta T^{-i}(H_{\alpha,\beta}^n \cap P_k)\right) \\
&\quad \le \sum_{i=0}^{2^n-1}\sum_{k=1}^{2^n} \mu\left(T_n^{-i}\left(P_k\right) \Delta T^{-i}\left(P_k\right)\right) \\
&\quad < 2^n \frac{\theta}{2^n} = \theta.
\end{aligned}$$

2) The proof of this part of the lemma is illustrated in Fig. 2. Fix $n \in \mathbb{N}$, $\alpha, \beta \in \mathbb{Z}_2^n$ and consider the expression

$$\begin{aligned}
&\int_X DP_T(\alpha,\beta)(x) d\mu(x) \\
&\quad = \frac{1}{2^n} \sum_{i=0}^{2^n-1} \int_X \chi_{H_{\alpha,\beta}^n}(T_n^i(x)) d\mu(x) \\
&\quad = \frac{1}{2^n} \sum_{i=0}^{2^n-1}\sum_{k=1}^{2^n} \int_{P_k} \chi_{H_{\alpha,\beta}^n}(T_n^i(x)) d\mu(x).
\end{aligned}$$

By the definition (16) of $H_{\alpha,\beta}^n$, we have

$$\begin{aligned}
&\int_{P_k} \chi_{H_{\alpha,\beta}^n}(T_n^i(x)) d\mu(x) \\
&\quad = \mu\left(\left\{x \in P_k : \overline{P_{T_n^{i+1}(x)}} \oplus \overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) = \beta\right\}\right).
\end{aligned}$$

We introduce now the following auxiliary sets. Given $\gamma \in \mathbb{Z}_2^n$ such that $\bigcup_{x \in P_k} \left(\overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) \oplus \beta\right) \cap \gamma \neq \emptyset$, define

$$\begin{aligned}
Q_{k,i}^\gamma &= \{x \in P_k : \overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) \oplus \beta = \gamma\} \subset P_k \\
X^\gamma &= \{x \in X : \overline{P_x} = \gamma\} \subset X.
\end{aligned}$$

Then

$$\begin{aligned}
&T^{i+1}\left\{x \in Q_{k,i}^\gamma : \overline{P_{T_n^{i+1}(x)}} \oplus \overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) = \beta\right\} \\
&\quad = T^{i+1}\left\{x \in Q_{k,i}^\gamma : \overline{P_{T_n^{i+1}(x)}} = \gamma\right\} \subset T^{i+1}(Q_{k,i}^\gamma) \cap X^\gamma.
\end{aligned}$$

Moreover, since $P_k = \bigcup_{\gamma \in \mathbb{Z}_2^n} Q_{k,i}^\gamma$ (disjoint union) and the automorphism T is μ-preserving

$$\begin{aligned}
&\mu\left(\left\{x \in P_k : \overline{P_{T_n^{i+1}(x)}} \oplus \overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) = \beta\right\}\right) \\
&\quad = \sum_{\gamma \in \mathbb{Z}_2^n} \mu\left(\left\{x \in Q_{k,i}^\gamma : \overline{P_{T_n^{i+1}(x)}} \right.\right. \\
&\qquad \left.\left. \oplus \overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) = \beta\right\}\right) \\
&\quad \le \sum_{\gamma \in \mathbb{Z}_2^n} \mu\left(T^{i+1}(Q_{k,i}^\gamma) \cap X^\gamma\right)
\end{aligned}$$

holds. Due to the mixing property of T, for each $\gamma \in \mathbb{Z}_2^n$, we have

$$\mu\left(T^{i+1}(Q_{k,i}^\gamma) \cap X^\gamma\right) \le \mu(Q_{k,i}^\gamma)\mu(X^\gamma) + \frac{\varepsilon_{k,i}^\gamma}{2^{2n}}.$$

Adding over all $\gamma \in \mathbb{Z}_2^n$, the bound

$$\sum_{\gamma\in\mathbb{Z}_2^n} \mu\left(T^{i+1}(Q_{k,i}^{\gamma}) \cap X^{\gamma}\right) \leq \frac{1}{2^{2n}} + \frac{1}{2^{2n}} \sum_{\gamma\in\mathbb{Z}_2^n} \varepsilon_{k,i}^{\gamma}$$

follows, since $\sum_{\gamma\in\mathbb{Z}_2^n} \mu(Q_{k,i}^{\gamma}) = \mu(P_k) = 1/2^n$ and $\mu(X^{\gamma}) = 1/2^n$ for all $\gamma \in \mathbb{Z}_2^n$. Hence

$$\mu\left(\left\{x \in P_k : \overline{P_{T_n^{i+1}(x)}} \oplus \overline{T}_n(\overline{P_{T_n^i(x)}} \oplus \alpha) = \beta\right\}\right) \leq \frac{1}{2^{2n}}(1+\varepsilon_{k,i})$$

with $\varepsilon_{k,i} := \sum_{\gamma\in\mathbb{Z}_2^n} \varepsilon_{k,i}^{\gamma}$. Finally, $\varepsilon_m := \sum_{i=0}^{2^n-1}\sum_{k=1}^{2^n} \varepsilon_{k,i}$ can be made arbitrarily small by taking m in $T = S^m$ large enough. ∎

The main theorem of this paper is now a straightforward consequence of the previous *Lemma*.

Theorem 5: Let $T = S^m$ be the product of a mixing automorphism $S : X \to X$ with itself $m \geq 1$ times. Suppose $(T_n, \mathcal{P}_n)$, $\mathcal{P}_n = \{P_k : 1 \leq k \leq 2^n\}$ is a cyclic approximation of T with speed $\theta/2^n$ $(2 < \theta < 4)$ satisfying (10). Then

$$DP_{\overline{T}_n} \leq \frac{\theta + 1 + \varepsilon_m}{2^n}$$

where ε_m can be made arbitrarily small by taking m large enough.

Example (continued): If $\alpha = (1, 0, \ldots, 0)$, then we get from (11)

$$\begin{aligned}\overline{T}_n(\xi \oplus \alpha) &= (\overline{\xi_1 \oplus 1}, \overline{\overline{\xi_2}}, \ldots, \overline{\overline{\xi_{n-1}}}, \overline{\xi_n}) \\ &= (\xi_1, \overline{\overline{\xi_2}}, \ldots, \overline{\overline{\xi_{n-1}}}, \overline{\xi_n})\end{aligned}$$

so that

$$\begin{aligned}&\overline{T}_n(\xi) \oplus \overline{T}_n(\xi \oplus \alpha) \\ &\quad= \left(\overline{\xi_1}, \overline{\overline{\xi_2}}, \ldots, \overline{\overline{\xi_{n-1}}}, \overline{\xi_n}\right) \\ &\qquad\oplus \left(\xi_1, \overline{\overline{\xi_2}}, \ldots, \overline{\overline{\xi_{n-1}}}, \overline{\xi_n}\right) \\ &\quad= (1, 0, \ldots, 0).\end{aligned}$$

If we choose now $\beta = (1, 0, \ldots, 0) = \alpha$, it follows that

$$DP_{\overline{T}_n}(\alpha, \beta) = \frac{\#\{\xi \in \mathbb{Z}_2^n : \overline{T}_n(\xi) \oplus \overline{T}_n(\xi \oplus \alpha) = \beta\}}{2^n} = 1$$

$$DP_{\overline{T}_n} = 1.$$

We see again that ergodicity without mixing is not enough to derive cryptographically secure substitutions.

V. Examples

As already mentioned in Section II, given a Lebesgue space $(X, \mathcal{A}, \mu)$ and a mixing, μ-invariant automorphism S in X, it is in general not known how to construct periodic approximations $(S_n, \mathcal{P}_n)$ of S, where $\mathcal{P}_n = \{P_k : 1 \leq k \leq q_n\}$. So, in implementations of the approximation philosophy explained above, one has to circumvent this problem and this calls for using only the fact that $(S_n, \mathcal{P}_n)$ is a periodic approximation of S. In our case the situation is somewhat simpler, since all we need for our purposes is to come up with the reshuffling $\overline{S}_n$ of the n-bit blocks $\overline{P_k}$.

One way we propose to make the trick is the following. First observe that, if $(S_n, \mathcal{P}_n)$ is a cyclic approximation of S and $x \in P_x$, then $\mathcal{P}_n = \{P_{S^i(x)} : 0 \leq i \leq q_n - 1\}$ holds with high probability (in the sense of the measure μ), in fact, with arbitrarily high probability as q_n increases. If S is sufficiently mixing, this will be even the case for moderate q_n. Therefore, for a 'typical' $x \in X$ one may reasonably expect, $(S_n, \mathcal{P}_n)$ being a cyclic approximation of S, that the finite sequence of iterates $\{S^i x : 0 \leq i \leq q_n - 1\}$ will visit all sets $P_k \in \mathcal{P}_n$ once (and only once). Moreover, from

$$\begin{aligned}\mu\left(S^i(P_k) \triangle S_n^i(P_k)\right) &\leq \sum_{j=0}^{i-1} \mu\left(S^{i-j} S_n^j(P_k) \right. \\ &\qquad \left. \times \triangle S^{i-j-1} S_n^{j+1}(P_k)\right) \\ &= \sum_{j=0}^{i-1} \mu\left(S(S_n^j P_k) \triangle S_n(S_n^j P_k)\right)\end{aligned}$$

for $P_k \in \mathcal{P}_n$ and $1 \leq i \leq q_n$, and the cyclicity of $(S_n, \mathcal{P}_n)$, it follows that

$$\mu(S^i(P_k) \triangle S_n^i(P_k)) \leq \sum_{k=1}^{q_n} \mu(S(P_k) \triangle S_n(P_k)) \leq f(q_n)$$

where $f(q_n) \to 0$ is the speed of the approximation. This spells out that, up to the precision set by the partition $\mathcal{P}_n$, the sequence $\{S_n^i x : 1 \leq i \leq q_n\}$ "shadows" the mixing dynamics of the approximated automorphism T for most (asymptotically, almost all) $x \in X$. Remember that $\mathcal{P}_n \to \mathcal{E}$, where $\mathcal{E}$ is the partition of X into separate points, so that the partitions $\mathcal{P}_n$ gets finer (i.e., their diameters get smaller) with increasing n and thus $P_x \in \mathcal{P}_n$ locates $x \in X$ within its vanishing diameter. Let us stress again that for us to define $\overline{S}_n$ is not important how the sets P_k look like but only the order in which they are visited by the iterates of x.

It remains to enumerate the (unknown) partition elements $P_{S^i(x)}$ in order to be done; clearly, the cryptographic performance of $\overline{S}_n$ will depend in general on this enumeration. For this reason, we propose to enhance the diffusion properties of S_n by numbering the partition elements in such a way that close elements (in the sense of the metric, assuming X is a metric space) get close n-blocks (in the sense of Hamming distance, assuming $q_n = 2^l$, $l = l(n)$[1]). To make sure that the iterates $S^i(x)$, $0 \leq i \leq q_n - 1$, have scattered evenly over X, one might want to set on X a uniform mesh of q_n subintervals and check that all iterates belong to different subintervals. (For good mixing maps, it should not take many attempts to find such an x.)

To be more specific, we may assume that the S_n-dynamic on the (unknown) partition $\mathcal{P}_n$ is given by $P_{S^i(x)} \mapsto P_{S^{i+1}(x)}$, $0 \leq i \leq 2^l - 1$, for a typical $x \in X$. This dynamic is then carried over by some diffusion-enhancing labeling $\Lambda : P_{S^i(x)} \to \overline{P_{S^i(x)}}$ to $\mathbb{Z}_2^l$ in the obvious way: $\overline{P_{S^i(x)}} \mapsto \overline{P_{S^{i+1}(x)}} =: \overline{S}_n(\overline{P_{S^i(x)}})$. We

[1]We stress here that a small number of points/blocks does not satisfy this property. However, the measure of such points/blocks tends to 0 when the partition becomes finer and the quality of the designed S-boxes becomes better. This fact can be observed in the the numerical example presented in Table I.

TABLE I
VALUES OF LP AND DP FOR DIFFERENT BLOCK LENGTHS l

l	$LP_{\overline{T}}$	$LP_{\overline{T}} - 2^{-l}$	$DP_{\overline{T}}$	$DP_{\overline{T}} - 2^{-l+1}$
8	0.02995	2.60×10^{-2}	0.03125	2.34×10^{-2}
10	0.003906	2.93×10^{-3}	0.01172	9.77×10^{-3}
12	0.001682	1.44×10^{-3}	0.002440	1.95×10^{-3}
14	0.0001907	1.30×10^{-4}	0.0007324	6.10×10^{-4}
16	0.00009299	7.77×10^{-5}	0.0001526	1.22×10^{-4}
18	0.00002629	2.25×10^{-5}	0.00003052	2.29×10^{-5}

will presently explain a simple realization of such a Λ for $X = [0,1)\times[0,1) \equiv [0,1)^2$ and $l = 2n$ that, as a matter of fact, easily generalizes to other intervals and dimensions. Furthermore, in the light of *Theorem* 5, it may be convenient to "accelerate" this dynamic by using $T := S^m$, $m > 1$, instead.

In this spirit, let us consider the first 2^{2n} points $\{(x_i, y_i)\}_{i=1}^{2^{2n}}$ of the orbit of $(x_1, y_1) \in [0,1)^2$ under a mixing map $S : [0,1)^2 \to [0,1)^2$, i.e.,

$$(x_{i+1}, y_{i+1}) = S(x_i, y_i), \qquad i = 1, 2, \ldots, 2^{2n} - 1$$

where the initial point (x_1, y_1) has to be chosen according to the previous recommendations. After ordering the points (x_i, y_i) first by the size of their y-coordinates and then by the size of the corresponding x-coordinates in consecutive, disjoint groups of 2^n points, we shall associate to each of them a block $\zeta_i \in \mathbb{Z}_2^{2n}$. This amounts to labeling the partition elements as follows: if $(x_i, y_i) \in P_{(x_i,y_i)}$, then $\Lambda : P_{(x_i,y_i)} \mapsto \zeta_i =: \overline{P_{(x_i,y_i)}}$. The resulting substitution on $\mathbb{Z}_2^{2n}$ is $\overline{S}_n : \zeta_i \mapsto \zeta_{i+1}$. Thus, start by putting the y-coordinates in order of size

$$y_{k_1} < y_{k_2} < \cdots < y_{k_{2^n}} < y_{k_{2^n+1}} < \cdots y_{k_{2^{n+1}}} < y_{k_{2^{n+1}+1}} < \cdots < y_{k_{2^{2n}}}. \quad (20)$$

Next, group them in consecutive, disjoint segments of 2^n elements, $\{y_{k_{s2^n+1}}, \ldots, y_{k_{(s+1)2^n}}\}$, $0 \le s \le 2^n - 1$, and label the s-th segment by the n-bit representation ζ_s^y of s. Order, in turn, the corresponding x-coordinates, $\{x_{k_{s2^n+1}}, \ldots, x_{k_{(s+1)2^n}}\}$, by their size

$$x_{m_{s2^n+1}} < x_{m_{(s-1)2^n+2}} < \cdots < x_{m_{(s+1)2^n}}$$

and label $x_{m_{s2^n+r+1}}$, $0 \le r \le 2^n - 1$, by the n-bit representation ζ_r^x of r.

Given now (x_i, y_i), $1 \le i \le 2^{2n}$, suppose s is the rank (starting from 0) of the said segment of (20) in which y_i lies and, among the corresponding abscissas, r is the rank (starting from 0) of x_i. We assign then to (x_i, y_i) the $2n$-block $\zeta_i = (\zeta_r^x, \zeta_s^y)$ which, as above, codifies the rank parameters (r, s) in binary form. This completes the 'encoding' $\Lambda : \mathcal{P}_n \mapsto \mathbb{Z}_2^{2n}$, see also Fig. 3. Generalizations to other intervals and dimensions need no further elucidation.

For S we have considered algebraic automorphisms on $[0,1)^2$ of the form

$$\begin{cases} x_{i+1} = ax_i + by_i \mod 1 \\ y_{i+1} = cx_i + dy_i \mod 1 \end{cases}$$

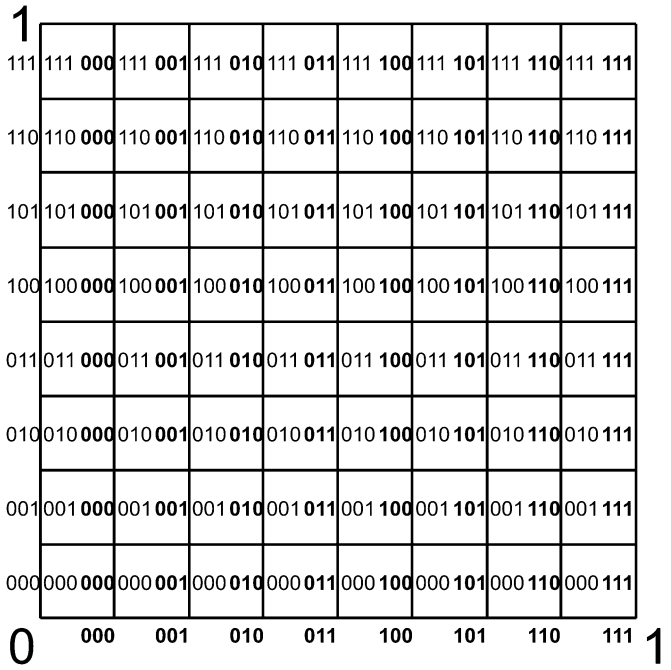

Fig. 3. The encoding method applied in the example, here $n = 64$.

with a, b, c, d integers and $ad - bc = 1$, because they are invariant with respect to Lebesgue measure and computationally efficient. Table I summarizes the results obtained in the case

$$\begin{aligned} a &= 587943273 \\ b &= 185921552200509715 \\ c &= 2 \\ d &= 632\,447\,247 \end{aligned}$$

with the accelerated dynamic $T = S^{30}$: In particular, for $n = 8$ one finds $LP_{\overline{T}} = 2^{-5.06}$ and $DP_{\overline{T}} = 2^{-5.00}$, which compare very favorably with the performances of other methods proposed in the literature for designing S-boxes of size 8, for instance in [6]. Finally, let us make a few remarks concerning the previous results.

1) For blocks of lengths 8 and 10 bits, the search was exhaustive both for LP and DP.
2) In the case of linear cryptanalysis of blocks of lengths 12 to 18 bits, the number of (randomly chosen) linear approximations considered was 2^{19}.
3) Similarly, in the case of differential cryptanalysis of blocks of lengths 12 and 14 bits, the number of (randomly chosen) pairs (input difference, output difference) checked was also 2^{18}–2^{20}. For blocks of lengths 16 and 18 bits, this number was 2^{17}.

In the next example, we use Baker's map (without acceleration) to construct cryptographically secure substitutions directly from two seed binary vectors. Let $T : [0,1]^2 \to [0,1]^2$ be Baker's map

$$T(x,y) = \begin{cases} \left(2x, \frac{1}{2}y\right), & \text{if } 0 \le x \le \frac{1}{2} \\ \left(2x-1, \frac{1}{2}(y+1)\right), & \text{if } \frac{1}{2} \le x \le 1 \end{cases}$$

and let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_r)$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_r)$ be two random binary vectors with $r \geq 2^{2n+1}$. First of all, we associate to $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ the point $(x_1, y_1) \in [0,1]^2$ by using the dyadic representation

$$\boldsymbol{\xi} \mapsto x_1 := \xi_1 \cdot 2^{-1} + \cdots + \xi_r \cdot 2^{-r} \equiv [\xi_1, \ldots, \xi_r]$$
$$\boldsymbol{\eta} \mapsto y_1 := \eta_1 \cdot 2^{-1} + \cdots + \eta_r \cdot 2^{-r} \equiv [\eta_1, \ldots, \eta_r].$$

Then, the action of T on (x_1, y_1), $T(x_1, y_1) = (x_2, y_2)$, translates in the dyadic representation (with fixed length r) into the following left and right shifts:

$$x_2 = [\xi_2, \ldots, \xi_r, 0]$$
$$y_2 = [\xi_1, \eta_1, \ldots, \eta_{r-1}].$$

In general, $(x_{k+1}, y_{k+1}) = T(x_k, y_k) = T^k(x_1, y_1)$, where

$$x_k = [\xi_k, \ldots, \xi_r, \mathbf{0_{k-1}}],$$
$$y_k = [\xi_{k-1}, \xi_{k-2}, \ldots, \xi_1, \eta_1, \ldots, \eta_{r-k+1}]$$

$1 \leq k \leq r$, and $\mathbf{0}_k$ stands for a string of k zeros. We will consider hereafter only the first 2^{2n} iterates $\{(x_k, y_k)\}_{k=1}^{2^{2n}}$ (so that $r - k \geq r - 2^{2n} \geq 2^{2n}$).

Observe next that if $\prec$ denotes the lexicographical order of r-blocks, viz.,

$$(0, \ldots, 0, 0) \prec (0, \ldots, 0, 1) \prec \cdots \prec (1, \ldots, 1, 1)$$

then

$$[\zeta_1, \ldots, \zeta_r] < [\zeta'_1, \ldots, \zeta'_r] \text{ iff } (\zeta_1, \ldots, \zeta_r) \prec (\zeta'_1, \ldots, \zeta'_r).$$

Therefore, rather than ordering the coordinates x_k and y_k according to their sizes to find out $\overline{T}_n$, we can instead order the r-blocks

$$\boldsymbol{\xi}_k := (\xi_k, \ldots, \xi_r, \mathbf{0}_{k-1})$$
$$\boldsymbol{\eta}_k := (\xi_{k-1}, \xi_{k-2}, \ldots, \xi_1, \eta_1, \ldots, \eta_{r-k+1}) \qquad (21)$$

$1 \leq k \leq 2^{2n}$, lexicographically (first the $\boldsymbol{\eta}_k$ to get $\boldsymbol{\zeta}_b^y$ and then the corresponding $\boldsymbol{\xi}_k$ in consecutive, disjoint groups of 2^n vectors to get $\boldsymbol{\zeta}_a^x$, when implementing the general procedure).

As a result, to any pair of random blocks $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{Z}_2^r$, $r \geq 2^{2n+1}$, corresponds a unique substitution $\overline{T}_n : (\boldsymbol{\zeta}_a^x, \boldsymbol{\zeta}_b^y) \mapsto (\boldsymbol{\zeta}_c^x, \boldsymbol{\zeta}_d^y)$ on $2n$-blocks, where now $(\boldsymbol{\zeta}_a^x, \boldsymbol{\zeta}_b^y)$ and $(\boldsymbol{\zeta}_c^x, \boldsymbol{\zeta}_d^y)$ are given by the lexicograpical order of the 2^n r-vector pairs $(\boldsymbol{\xi}_k, \boldsymbol{\eta}_k)$ and $(\boldsymbol{\xi}_{k+1}, \boldsymbol{\eta}_{k+1}) = T(\boldsymbol{\xi}_k, \boldsymbol{\eta}_k)$ in (21), respectively. Note that $(\boldsymbol{\xi}_{k+1}, \boldsymbol{\eta}_{k+1})$ is obtained recursively from $(\boldsymbol{\xi}_k, \boldsymbol{\eta}_k)$ by shifting the leftmost bit of $\boldsymbol{\xi}_k$ into the leftmost bit of $\boldsymbol{\eta}_k$ (with $\boldsymbol{\xi}_0 = \boldsymbol{\xi}$ and $\boldsymbol{\eta}_0 = \boldsymbol{\eta}$).

In the numerical simulation, we chose $r = 1024$, $1 \leq k \leq 256$ and $n = 4$ (8-bit blocks). For a sample of 3000 seed random vector pairs, the best performance was $LP_{\overline{T}_n} \simeq 0.08$ and $DP_{\overline{T}_n} \simeq 0.05$.

VI. Conclusions

In this paper, we proposed periodic approximations of mixing dynamical systems as a practicable way one can go to construct cryptographically secure substitutions. Indeed, although the theoretical results stated in Sections III and IV are of abstract nature, the leading principle behind can be materialized in a simple form. More importantly, numerical implementations, two of which were reported in Section V, support this approach in quite satisfactory terms.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, 1949.

[2] A. Menezes, P. van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*. Boca Raton, FL: CRC, 1997.

[3] F. Pichler and J. Scharinger, "Finite dimensional generalized baker dynamical systems for cryptographic applications," *Lect. Notes in Comput. Sci.*, vol. 1030, pp. 465–476, 1996.

[4] J. Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps," *Int. J. Bifurc. Chaos*, vol. 8, no. 6, pp. 1259–1284, 1998.

[5] G. Jakimoski and L. Kocarev, "Chaos and cryptography: Block encryption ciphers based on chaotic maps," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 2, pp. 163–169, Feb. 2001.

[6] ——, "Differential and linear probabilities of a block-encryption cipher," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 1, pp. 121–123, Jan. 2003.

[7] N. Masuda and K. Aihara, "Cryptosystems with discretized chaotic maps," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 1, pp. 28–40, Jan. 2002.

[8] J. M. Amigó and J. Szczepański, "Approximations of dynamical systems and their applications to cryptography," *Int. J. Bifurc.Chaos*, vol. 13, pp. 1937–1948, 2003.

[9] I. P. Cornfeld, S. V. Fomin, and G. S. Ya, *Ergodic Theory*. New York: Springer Verlag, 1982.

[10] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*. Oxford, U.K.: Oxford Univ. Press, 1979.

[11] C. L. DeVito, *Functional Analysis*. New York: Academic, 1987.

[12] M. Matsui, "Linear cryptanalysis method for DES cipher," in *Proc. Eurocrypt '93 Advances in Cryptology*, 1994, pp. 386–397.

[13] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *J. Crypt.*, vol. 4, pp. 3–72, 1991.

Janusz Szczepanski received the M.Sc. degree in mathematics and the Ph.D. degree in applied mathematics from Warsaw University, Warsaw, Poland, and the Polish Academy of Sciences, Warsaw, Poland, in 1979 and 1985, respectively.

He is a Researcher at the Institute of Fundamental Technological Research, Polish Academy of Sciences, and also a Consultant on cryptography with the Polish Certification Authority (Root) for Public Key Infrastructure (Trust and Certification Centre "CENTRAST" Co.), Warsaw, Poland. In 2000 and 2003, he was Visiting Scientist at the Miguel Hernández University, Elche, Spain, and in 2004 he was Visiting Scholar at the University of California, San Diego. His research interests include cryptography, information theory and application of dynamical systems and stochastic processes to biological systems.

Dr. Szczepanski received the Polish Academy of Sciences Award in 1989 and in 1992 the Kosciuszko Foundation Fellowship (NY) for a research visit to the Snowbird Research Center.

José M. Amigó received the Ph.D. degree in theoretical physics from the University of Göttingen, Germany, in 1987.

He was a Postdoctoral Fellow at the National Aerospace Laboratory, Tokyo, Japan, from 1989 to 1990, and a System Analyst with the aerospace company Construcciones Aeronáuticas S.A., Madrid, Spain from 1991 to 1997. Currently, he is Associate Professor of Applied Mathematics at Miguel Hernández University, Elche (Alicante), Spain, and also affiliated with the Operations Research Centre of this university. His scientific interests include secure communications, mathematical physics and computational neurosciences.

Tomasz Michalek received the M.Sc. degree in computer science from Jagiellonian University, Cracow, Poland, in 1999. He is working toward the Ph.D. degree in mechanics and physics of fluids at the Polish Academy of Sciences, Warsaw, Poland.

He worked as a Software Engineer in the Research and Development Division of ComArch S.A., Cracow, Poland, from 1999 to 2001. In 2003, he received training at the Nova Gorica Polytechnic, Ljubljana, Slovenia. His research interests include numerical methods, computational fluid dynamics, validation and verification in computational fluid dynamics, experimental and numerical benchmarking, parallel and distributed computing, object oriented programming and bioinformatics.

Ljupco Kocarev is a Research Scientist at the Institute for Nonlinear Science, University of California San Diego and Professor at the Graduate School of Electrical Engineering, University "Kiril i Metodij," Skopje, Macedonia. His scientific interests include nonlinear systems and circuits; coding and information theory; networks and networks on chip; and cryptography. He has coauthored more than 80 journal papers in 18 different international peer-review journals, ranging from mathematics to physics, and from electrical engineering to computer sciences.

Dr. Kocarev's work has been cited more than 2000 times according to Science Citation Index. He is a foreign member of Macedonian Academy of Sciences and Arts.