

Homogeneous Permutations. Random Feistel schemes are never homogeneous

Jacques Patarin

Abstract

In this paper we introduce two new concepts that are stronger than strong pseudorandomness: “very strong pseudorandomness” and “homogeneous permutations”. We explain why we think that those concepts are natural, and we study the values k for which the Luby-Rackoff construction with k rounds satisfy these notions.

Note: At FSE’98, I have published a paper called “About Feistel Schemes with Six (or more) Rounds”. This paper has been split in two extended version: the results on homogenous permutations are in this paper, and the results of security when $m \ll 2^{3n/4}$ are now in the extended paper called “Random Feistel Schemes: security in $m \ll 2^{3n/4}$ for ≥ 6 rounds”.

1 Introduction

In their famous paper [2], M. Luby and C. Rackoff provided a construction of pseudorandom permutations and strong pseudorandom permutations. (“Strong pseudorandom permutations” are also called “super pseudorandom permutations”: here the distinguisher can access the permutation *and* the inverse permutation at points of its choice.) The basic building block of the Luby-Rackoff construction (L-R construction) is the so called Feistel permutation based on a pseudorandom function defined by the key. Their construction consists of four rounds of Feistel scheme (for strong pseudorandom permutations) or three rounds of Feistel permutations (for pseudorandom permutations). Each round involves an application of a different pseudorandom function. This L-R construction is very attractive for various reasons: it is elegant, the proof does not involve any unproven hypothesis, almost all (secret key) block ciphers in use today are based on Feistel schemes, and the number of rounds is very small (so that their result may suggest ways of designing faster block ciphers).

In this paper, we introduce two new concepts about permutation generators: “very strong pseudorandomness” and “homogeneous permutations”. These concepts both imply that the generator is a strong pseudorandom generator. We explain why we feel that it is natural to introduce these notions, and we characterize the values k such that the L-R constructions with k rounds satisfy (or not) these notions.

Finally we formulate a few open problems and we conclude.

2 Definitions

- Let G be a permutation generator, such that G involves ℓ different pseudorandom functions of F_n to compute a permutation of B_{2n} . We denote by K the set of all ℓ -uples of functions (f_1, \dots, f_ℓ) of F_n (i.e. $K = F_n^\ell$). Thus G associates to each $k \in K$ a permutation $G(k)$ of B_{2n} . K can be seen as the set of the keys of G , and $k \in K$ as a secret key.
- Let $\alpha_1, \dots, \alpha_m$ be m distinct elements of I_{2n} , and let β_1, \dots, β_m be also m distinct elements of I_{2n} . We denote by $H(\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_m)$ the number of keys k of K such that:

$$\forall i, 1 \leq i \leq m, G(k)(\alpha_i) = \beta_i.$$

Definition 1: We say that G is a “homogeneous” permutation generator if there exist a function $\varepsilon(m, n) : \mathbf{N}^2 \rightarrow \mathbf{R}$ such that, for any integer m :

1. For all $\alpha_1, \dots, \alpha_m$ being m distinct elements of I_{2n} , and for all β_1, \dots, β_m being m distinct elements of I_{2n} , we have:

$$\left| H(\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_m) - \frac{|K|}{2^{2nm}} \right| \leq \varepsilon(m, n) \frac{|K|}{2^{2nm}}.$$

2. For any polynomial $P(n)$ and any $\alpha > 0$, an integer n_0 exists such that:

$$\forall n \geq n_0, \forall m \leq P(n), \varepsilon(m, n) \leq \alpha.$$

Remark: This definition might look a bit complex but in fact this notion of “homogeneous” permutations is a very natural notion: roughly speaking, a permutation generator is homogeneous when for all set of m cleartext/ciphertext pairs, there are always **about the same number** of possible keys that send all the cleartexts on the ciphertexts.

Definition 2: We say that G is a “very strong” permutation generator if – with the same notations as above – the function $\varepsilon(m, n)$ satisfies condition 2, and the following condition 1’ (instead of condition 1):

- 1’. For all $\alpha_1, \dots, \alpha_m$ being m distinct elements of I_{2n} , and for all β_1, \dots, β_m being m distinct elements of I_{2n} , we have:

$$H(\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_m) \geq \frac{|K|}{2^{2nm}} (1 - \varepsilon(m, n)).$$

Remark: Roughly speaking, a permutation generator is “very strong” when for all set of m cleartext/ciphertext pairs, the number of possible keys (that send all the cleartexts on the ciphertexts) is always **at least** about the average number.

Theorem 2.1 *If G is a “homogeneous permutation generator”, then G is a “very strong permutation generator”.*

Theorem 2.2 *If G is a “very strong permutation generator”, then G is a “strong permutation generator”.*

Proof: Theorem 2.1 is an obvious consequence of the definitions. Theorem 2.2 corresponds to the technique of proof we used in part I. (This way of proving strong pseudorandomness was first explicitly used in [6].)

As a result, for permutation generators, we have:

$$\text{Homogeneous} \Rightarrow \text{Very Strong} \Rightarrow \text{Strong} \Rightarrow \text{Pseudorandom}.$$

Interpretations:

In order to distinguish (with a non-negligible probability) permutations generated by a homogeneous permutation generator, from truly random permutations of B_{2n} , an enemy must know a large number of cleartext/ciphertext pairs. (More precisely, this number must increase faster than any polynomial in n , **whatever** the cleartext/ciphertext pairs may be.)

Remarks:

1. Related (but not equivalent) notions can be found in [11] (“multipermutations”) and in [5].
2. In some very special cases, this property of “homogeneity” may be useful and “strong pseudorandomness” is not enough. For example, let us assume that the enemy has a spy inside the encryption team. Let us also assume that the aim of the enemy is to distinguish the encryption algorithm from a truly random permutation, and that his spy has access to the whole database of cleartext/ciphertext pairs, but can only send very few such pairs to help distinguishing. In such a case, “homogeneity” may be a more natural property than strong pseudorandomness. However, we introduced the concepts of “homogeneity” and “very strong pseudorandomness” because they are very natural in the proofs, and not with applications in mind.

3 Examples

3.1 Ψ^4 is not homogeneous

Example 1 (with $m = 2$):

As shown in [7] p. 309 (or in [1] p. 314), if $\Psi^4[L_1, R_1] = [S_1, T_1]$ and $\Psi^4[L_2, R_2] = [S_2, T_2]$, and $R_1 = R_2, L_1 \neq L_2$, then the probability that $S_1 \oplus S_2 = L_1 \oplus L_2$ is about twice what it would be with a truly random permutation of B_{2n} (instead of Ψ^4). In [7] (and [1]), this result was used to show that the security bound given by Luby and Rackoff for Ψ^4 in a chosen-cleartext attack is tight (the attack requires $\simeq \sqrt{2^n}$ messages to ensure $S_i \oplus S_j = L_i \oplus L_j$).

Here, we use this result to show that Ψ^4 is not homogeneous, and the non-homogeneity property appears with only two (very special) messages.

Remark: However, Ψ^4 is a very strong permutation generator (and for Ψ^4 , we can take $\varepsilon(m, n) = \frac{m^2}{2^n}$). (As mentioned above, the proof of strong pseudorandomness of Ψ^4 given in [6] is also a proof of very strong pseudorandomness.)

Example 2 (with $m = 4$):

Let $R_1 = R_3$, $R_2 = R_4 = R_1 \oplus \alpha$, $S_1 = S_2$, $S_3 = S_4 = S_1 \oplus \alpha$, $L_1 = L_2$, $L_3 = L_4 = L_1 \oplus \alpha$, $T_1 = T_3$, $T_2 = T_4 = T_1 \oplus \alpha$.

Then the value H for Ψ^4 with these R, L, S, T is at least about $\frac{|F_n|^4}{2^{6n}}$ (instead of about $\frac{|F_n|^4}{2^{8n}}$ as expected if it was homogeneous). The proof of a similar property will be done in details for Ψ^6 below.

3.2 Ψ^5 is not homogeneous

If $\Psi^5[L_1, R_1] = [S_1, T_1]$ and $\Psi^5[L_2, R_2] = [S_2, T_2]$, and if $R_1 = R_2$ and $L_1 \neq L_2$, then the probability that $S_1 = S_2$ and $L_1 \oplus L_2 = T_1 \oplus T_2$ is about twice what it would be with a truly random permutation of B_{2n} (instead of Ψ^5). Therefore Ψ^5 is not homogeneous, and the non-homogeneity property appears with only two (very special) messages.

Remark: However, since here we have two equations and two indices ($S_i = S_j$ and $L_i \oplus L_j = T_i \oplus T_j$), this non-homogeneity property would require about $m = 2^n$ messages in a chosen-cleartext attack (instead of the $\sqrt{2^n}$ messages above for Ψ^4).

3.3 Ψ^6 is not homogeneous

Example 1 (with $m = 4$):

Let $\Psi^6[L_i, R_i] = [S_i, T_i]$ for $i = 1, 2, 3, 4$.

If $R_1 = R_3$, $R_2 = R_4 \neq R_1$, $S_1 = S_2$, $S_3 = S_4 \neq S_1$, $L_1 \oplus L_3 = L_2 \oplus L_4 = S_1 \oplus S_3 \neq 0$ and $T_1 \oplus T_2 = T_3 \oplus T_4 = R_1 \oplus R_2 \neq 0$, then we will see that H is at least about $2 \cdot \frac{|F_n|^6}{2^{8n}}$, instead of $\frac{|F_n|^6}{2^{8n}}$ as expected if it was homogenous. Therefore, Ψ^6 is not homogenous.

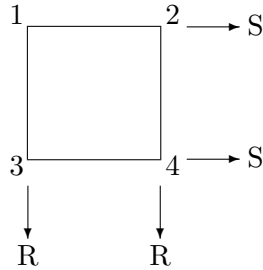


Figure 2: A representation of the equations $S_1 = S_2$, $S_3 = S_4$, $R_1 = R_3$ and $R_2 = R_4$.

Proof: We know (see [9] p.145 or [8] p.134) that the exact value of H is:

$$H = \sum_{(X,Y,P,Q) \text{ satisfying } (C)} \frac{|F_n|^6}{2^{6mn}} \cdot 2^{n(r+s+x+y+p+q)},$$

with (C) being the following set of conditions:

$$\forall i, j, 1 \leq i \leq m, 1 \leq j \leq m, i \neq j \left\{ \begin{array}{l} R_i = R_j \Rightarrow X_i \oplus L_i = X_j \oplus L_j \\ S_i = S_j \Rightarrow Y_i \oplus T_i = Y_j \oplus T_j \\ X_i = X_j \Rightarrow P_i \oplus R_i = P_j \oplus R_j \\ Y_i = Y_j \Rightarrow Q_i \oplus S_i = Q_j \oplus S_j \\ P_i = P_j \Rightarrow X_i \oplus Q_i = X_j \oplus Q_j \\ Q_i = Q_j \Rightarrow P_i \oplus Y_i = P_j \oplus Y_j. \end{array} \right.$$

and with m being the number of independent equations $R_i = R_j$, $i \neq j$, s is the number of independent equations $S_i = S_j$, $i \neq j$, etc.. up to q being the number of independent equations $Q_i = Q_j$, $i \neq j$.

We will consider two special sets of values for (X, Y, P, Q) .

First possible set.

Let X_1, Y_1, Q_1, P_1 have any value (thus we have 2^{4n} possible values here), and let $X_1 = X_2$, $X_3 = X_4 = X_1 \oplus L_1 \oplus L_3$, $Y_1 = Y_3$, $Y_2 = Y_4 = Y_1 \oplus T_1 \oplus T_2$, $Q_1 = Q_2$, $Q_3 = Q_4 = Q_1 \oplus S_1 \oplus S_3$, $P_1 \oplus P_3$ and $P_2 \oplus P_4 = P_1 \oplus R_1 \oplus R_2$.

It is easy to see that for these values all the conditions (C) are satisfied:

$$\begin{array}{ll} R_1 = R_3 \Rightarrow X_1 \oplus L_1 = X_3 \oplus L_3 & \text{(by definition of } X_3) \\ R_1 = R_3 \Rightarrow X_1 \oplus L_1 = X_3 \oplus L_3 & \text{(by definition of } X_3) \\ R_2 = R_4 \Rightarrow X_2 \oplus L_2 = X_4 \oplus L_4 & \text{(since } L_2 \oplus L_4 = L_1 \oplus L_3) \\ S_1 = S_2 \Rightarrow Y_1 \oplus T_1 = Y_2 \oplus T_2 & \text{(by definition of } Y_2) \\ S_3 = S_4 \Rightarrow Y_3 \oplus T_3 = Y_4 \oplus T_4 & \text{(since } T_1 \oplus T_2 = T_3 \oplus T_4) \\ X_1 = X_2 \Rightarrow P_1 \oplus R_1 = P_2 \oplus R_2 & \text{(by definition of } P_2) \\ X_3 = X_4 \Rightarrow P_3 \oplus R_3 = P_4 \oplus R_4 & \text{(since } R_1 \oplus R_2 = R_3 \oplus R_4) \\ Y_1 = Y_3 \Rightarrow Q_1 \oplus S_1 = Q_3 \oplus S_3 & \text{(by definition of } Q_3) \\ Y_2 = Y_4 \Rightarrow Q_2 \oplus S_2 = Q_4 \oplus S_4 & \text{(since } S_2 \oplus S_4 = S_1 \oplus S_3) \\ P_1 = P_3 \Rightarrow X_1 \oplus Q_1 = X_3 \oplus Q_3 & \text{(since } L_1 \oplus L_3 = S_1 \oplus S_3) \\ P_2 = P_4 \Rightarrow X_2 \oplus Q_2 = X_4 \oplus Q_4 & \text{(since } L_1 \oplus L_3 = S_1 \oplus S_3) \\ Q_1 = Q_2 \Rightarrow P_1 \oplus Y_1 = P_2 \oplus Y_2 & \text{(since } R_1 \oplus R_2 = T_1 \oplus T_2) \\ Q_3 = Q_4 \Rightarrow P_3 \oplus Y_3 = P_4 \oplus Y_4 & \text{(since } R_1 \oplus R_2 = T_1 \oplus T_2). \end{array}$$

Only from these X, Y, P, Q we see that:

$$H \geq 2^{4n} \cdot \frac{|F_n|^6}{2^{24n}} \cdot 2^{n(2+2+2+2+2)} = \frac{|F_n|^6}{2^{8n}}.$$

Note. Here we have $r = 2$ equalities in R and $s = 2$ equalities in S , and we have found variables X, Y, P, Q that satisfy all the equations (C) by introducing only $\mu = 4$ equations with non-zero constants (i.e. $X_3 = X_1 \oplus L_1 \oplus L_3$, $Y_2 = Y_1 \oplus T_1 \oplus T_2$, $Q_3 = Q_1 \oplus S_1 \oplus S_3$ and $P_2 = P_1 \oplus R_1 \oplus R_2$). Since all the equations of (C) are satisfied with $\mu \leq r + s$ it will give a proof of non homogeneity.

Second possible set.

There is also the ‘‘usual’’ set, i.e. the values X, Y, P, Q that we have used in the proof that Ψ^6 is super-pseudo-random (these values introduce no equalities in the X, Y, P, Q variables, so this second set is entirely disjoint from the first set).

Here we have:

- X_1 has 2^n possibilities,
- X_2 has $(2^n - 2)$ possibilities (because $X_2 \neq X_1$ and $X_2 \neq X_1 \oplus L_1 \oplus L_3$ and since here $L_1 \oplus L_3 = L_2 \oplus L_4$ these two inequalities will imply $X_2 \oplus L_2 \oplus L_4 \neq X_1$ and $X_2 \oplus L_2 \oplus L_4 \neq X_1 \oplus L_1 \oplus L_3$),
- $X_3 = X_1 \oplus L_1 \oplus L_3$, $X_4 = X_2 \oplus L_2 \oplus L_4$,
- Y_1 has 2^n possibilities, $Y_2 = Y_1 \oplus T_1 \oplus T_2$,
- Y_3 has $(2^n - 2)$ possibilities (because $Y_3 \neq Y_1$ and $Y_3 \neq Y_1 \oplus T_1 \oplus T_2$ and since here we have $T_1 \oplus T_2 = T_3 \oplus T_4$ these two inequalities will imply $Y_3 \oplus T_3 \oplus T_4 \neq Y_1$ and $Y_3 \oplus T_3 \oplus T_4 \neq Y_1 \oplus T_1 \oplus T_2$).
- P_1 has 2^n possibilities, and P_2 has $2^n - 1$ possibilities (because $P_2 \neq P_1$). Similarly P_3 has $2^n - 2$ possibilities (because $P_3 \neq P_1$ and $P_3 \neq P_2$) and P_4 has $2^n - 3$ possibilities (because $P_4 \neq P_1$, and $P_4 \neq P_2$ and $P_4 \neq P_3$),
- For the same reason Q_1, Q_2, Q_3 and Q_4 have respectively $2^n, 2^n - 1, 2^n - 2$ and $2^n - 3$ possibilities.

Only from these X, Y, P, Q we see that

$$H \geq 2^{4n} \cdot (2^n - 1)^2 \cdot (2^n - 2)^4 \cdot (2^n - 3)^2 \cdot \frac{|F_n|^6}{2^{24n}} \cdot 2^{n(2+2)} \approx \frac{|F_n|^6}{2^{8n}}.$$

Therefore by combining the first and the second set, we have $H \geq$ about $2 \frac{|F_n|^6}{2^{8n}}$, as claimed (instead of $H \approx \frac{|F_n|^6}{2^{8n}}$ if Ψ^6 was homogenous).

Remark 1.

Since $L_4 = L_1 \oplus L_2 \oplus L_3$ and $R_4 = R_2$, the index 4 is fixed the indices 1, 2 and 3 are fixed.

In fact we have here 3 indices 1, 2 and 3 and at least 4 equations on these indices that we cannot impose with a cleartext/ciphertext attack: $T_1 \oplus T_2 = R_1 \oplus R_2$, $L_1 \oplus L_3 = S_1 \oplus S_3$, $S_{4(1,2,3)} = S_3$, $T_{4(1,2,3)} = T_1 \oplus T_2 \oplus T_3$.

Thus, this example shows that Ψ^6 is not homogenous, but it does not give a cryptographic attack when $m < 2^n$.

Remark 2.

It is sometimes interesting to see if there is an attack when $2^n < m \ll 2^{2n}$, when this attack requires $\ll 2^{2n}$ computations.

However here when the index 1 is fixed (we have m possibilities for it), the index 2 is also “in a way” fixed, since $S_2 = S_1$ and $T_2 \oplus R_2 = T_1 \oplus R_1$ (because on average when 1 is fixed there will be about only one index 2 such that these two equations are satisfied). Similarly, when the index 1 is fixed, the index 3 is “in a way” fixed, since $R_3 = R_1$ and $S_3 \oplus L_3 = S_1 \oplus L_1$. So in fact, when 1 is fixed, 2, 3 and 4 are fixed. But there are still two exceptional equations: $S_{4(1)} = S_3$ and $T_{4(1)} = T_1 \oplus T_{2(1)} \oplus T_{3(1)}$, and when $m \ll 2^{2n}$ the probability that these equations occur is negligible. Therefore this example 1 does not give an attack even when $2^n < m \ll 2^{2n}$.

Example 2 (with $m = 9$):

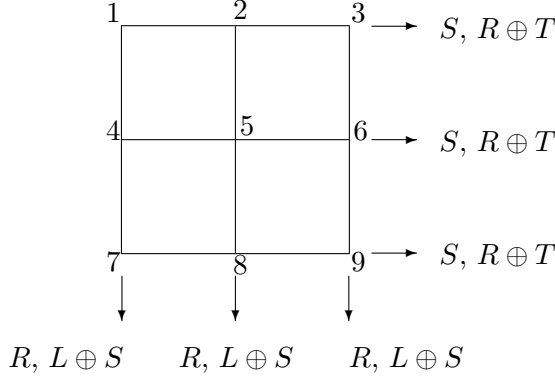


Figure 3: A representation of the 24 equations in S, L, R, T

Let $\Psi^6[L_i, R_i] = [S_i, T_i]$ for $1 \leq i \leq 9$. We study the values of H when

$$\left\{ \begin{array}{l} R_1 = R_4 = R_7 \\ R_2 = R_5 = R_8 \\ R_3 = R_6 = R_9 \\ L_1 \oplus S_1 = L_4 \oplus S_4 = L_7 \oplus S_7 \\ L_2 \oplus S_2 = L_5 \oplus S_5 = L_8 \oplus S_8 \\ L_3 \oplus S_3 = L_6 \oplus S_6 = L_9 \oplus S_9 \end{array} \right. \text{ and } \left\{ \begin{array}{l} S_1 = S_2 = S_3 \\ S_4 = S_5 = S_6 \\ S_7 = S_8 = S_9 \\ R_1 \oplus T_1 = R_2 \oplus T_2 = R_3 \oplus T_3 \\ R_4 \oplus T_4 = R_5 \oplus T_5 = R_6 \oplus T_6 \\ R_7 \oplus T_7 = R_8 \oplus T_8 = R_9 \oplus T_9 \end{array} \right.$$

All these relations are represented on Figure 3. We also assume that $R_1 \neq R_2$, $R_1 \neq R_3$, $R_2 \neq R_3$, $S_1 \neq S_4$, $S_1 \neq S_7$ and $S_4 \neq S_7$.

Then – as we will see below – for such L, R, S, T values, the value of H is at least $\frac{|F_n|^6}{2^{14n}}$, instead of $\frac{|F_n|^6}{2^{18n}}$ as expected if it was homogeneous. Therefore, Ψ^6 is not homogeneous.

Proof: Let $\alpha = R_1 \oplus R_2$, $\beta = R_1 \oplus R_3$, $\alpha' = S_1 \oplus S_4$ and $\beta' = S_1 \oplus S_7$ (by definition we have $\alpha \neq 0$, $\beta \neq 0$, $\alpha' \neq 0$ and $\beta' \neq 0$) We consider (X, Y, P, Q) values such that:

$$\left\{ \begin{array}{l} X_1 = X_2 = X_3 \\ X_4 = X_5 = X_6 = X_1 \oplus \alpha' \\ X_7 = X_8 = X_9 = X_1 \oplus \beta' \\ Y_1 = Y_4 = Y_7 \\ Y_2 = Y_5 = Y_8 = Y_1 \oplus \alpha \\ Y_3 = Y_6 = Y_9 = Y_1 \oplus \beta \end{array} \right. \text{ and } \left\{ \begin{array}{l} Q_1 = Q_2 = Q_3 \\ Q_4 = Q_5 = Q_6 = Q_1 \oplus \alpha' \\ Q_7 = Q_8 = Q_9 = Q_1 \oplus \beta' \\ P_1 = P_4 = P_7 \\ P_2 = P_5 = P_8 = P_1 \oplus \alpha \\ P_3 = P_6 = P_9 = P_1 \oplus \beta. \end{array} \right.$$

It is easy to verify that for these values all the conditions (C) are satisfied (these conditions were explicitly written for Ψ^6 in the section 3.3 above, Example 1):

$$\begin{aligned}
R_1 = R_4 &\Rightarrow X_1 \oplus L_1 = X_4 \oplus L_4 && (\text{because } \alpha' = L_1 \oplus L_4 = S_1 \oplus S_4) \\
R_1 = R_7 &\Rightarrow X_1 \oplus L_1 = X_7 \oplus L_7 && (\text{because } \beta' = S_1 \oplus S_7 = L_1 \oplus L_7) \\
R_2 = R_5 &\Rightarrow X_2 \oplus L_2 = X_5 \oplus L_5 && (\text{because } \alpha' = S_1 \oplus S_4 = S_2 \oplus S_5 = L_2 \oplus L_5) \\
R_2 = R_8 &\Rightarrow X_2 \oplus L_2 = X_8 \oplus L_8 && (\text{because } \beta' = S_1 \oplus S_7 = S_2 \oplus S_8 = L_2 \oplus L_8) \\
R_3 = R_6 &\Rightarrow X_3 \oplus L_3 = X_6 \oplus L_6 && (\text{because } \alpha' = S_1 \oplus S_4 = S_3 \oplus S_6 = L_3 \oplus L_6) \\
R_3 = R_9 &\Rightarrow X_3 \oplus L_3 = X_9 \oplus L_9 && (\text{because } \beta' = S_1 \oplus S_7 = S_3 \oplus S_9 = L_3 \oplus L_9) \\
S_1 = S_2 &\Rightarrow Y_1 \oplus T_1 = Y_2 \oplus T_2 && (\text{because } \alpha = R_1 \oplus R_2 = T_1 \oplus T_2) \\
S_1 = S_3 &\Rightarrow Y_1 \oplus T_1 = Y_3 \oplus T_3 && (\text{because } \beta = R_1 \oplus R_3 = T_1 \oplus T_3) \\
S_4 = S_5 &\Rightarrow Y_4 \oplus T_4 = Y_5 \oplus T_5 && (\text{because } \alpha = R_1 \oplus R_2 = R_4 \oplus R_5 = T_4 \oplus T_5) \\
S_4 = S_6 &\Rightarrow Y_4 \oplus T_4 = Y_6 \oplus T_6 && (\text{because } \beta = R_1 \oplus R_3 = R_4 \oplus R_6 = T_4 \oplus T_6) \\
S_7 = S_8 &\Rightarrow Y_7 \oplus T_7 = Y_8 \oplus T_8 && (\text{because } \alpha = R_1 \oplus R_2 = R_7 \oplus R_8 = T_7 \oplus T_8) \\
S_7 = S_9 &\Rightarrow Y_7 \oplus T_7 = Y_9 \oplus T_9 && (\text{because } \beta = R_1 \oplus R_3 = R_7 \oplus R_9 = T_7 \oplus T_9) \\
X_1 = X_2 &\Rightarrow P_1 \oplus R_1 = P_2 \oplus R_2 && (\text{because } \alpha = R_1 \oplus R_2) \\
X_1 = X_3 &\Rightarrow P_1 \oplus R_1 = P_3 \oplus R_3 && (\text{because } \beta = R_1 \oplus R_3) \\
X_4 = X_5 &\Rightarrow P_4 \oplus R_4 = P_5 \oplus R_5 && (\text{because } \alpha = R_1 \oplus R_2 = R_4 \oplus R_5) \\
X_4 = X_6 &\Rightarrow P_4 \oplus R_4 = P_6 \oplus R_6 && (\text{because } \beta = R_1 \oplus R_3 = R_4 \oplus R_6) \\
X_7 = X_8 &\Rightarrow P_7 \oplus R_7 = P_8 \oplus R_8 && (\text{because } \alpha = R_1 \oplus R_2 = R_7 \oplus R_8) \\
X_7 = X_9 &\Rightarrow P_7 \oplus R_7 = P_9 \oplus R_9 && (\text{because } \beta = R_1 \oplus R_3 = R_7 \oplus R_9) \\
Y_1 = Y_4 &\Rightarrow Q_1 \oplus S_1 = Q_4 \oplus S_4 && (\text{because } \alpha' = S_1 \oplus S_4) \\
Y_1 = Y_7 &\Rightarrow Q_1 \oplus S_1 = Q_7 \oplus S_7 && (\text{because } \beta' = S_1 \oplus S_7) \\
Y_2 = Y_5 &\Rightarrow Q_2 \oplus S_2 = Q_5 \oplus S_5 && (\text{because } \alpha' = S_1 \oplus S_4 = S_2 \oplus S_5) \\
Y_2 = Y_8 &\Rightarrow Q_2 \oplus S_2 = Q_8 \oplus S_8 && (\text{because } \beta' = S_1 \oplus S_7 = S_2 \oplus S_8) \\
Y_3 = Y_6 &\Rightarrow Q_3 \oplus S_3 = Q_6 \oplus S_6 && (\text{because } \alpha' = S_1 \oplus S_4 = S_3 \oplus S_6) \\
Y_3 = Y_9 &\Rightarrow Q_3 \oplus S_3 = Q_9 \oplus S_9 && (\text{because } \beta' = S_1 \oplus S_7 = S_3 \oplus S_9) \\
Q_1 = Q_2 &\Rightarrow P_1 \oplus Y_1 = P_2 \oplus Y_2 && (\text{because } P_1 \oplus P_2 = \alpha = Y_1 \oplus Y_2) \\
Q_1 = Q_3 &\Rightarrow P_1 \oplus Y_1 = P_3 \oplus Y_3 && (\text{because } P_1 \oplus P_3 = \beta = Y_1 \oplus Y_3) \\
Q_4 = Q_5 &\Rightarrow P_4 \oplus Y_4 = P_5 \oplus Y_5 && (\text{because } P_4 \oplus P_5 = \alpha = Y_4 \oplus Y_5) \\
Q_4 = Q_6 &\Rightarrow P_4 \oplus Y_4 = P_6 \oplus Y_6 && (\text{because } P_4 \oplus P_6 = \beta = Y_4 \oplus Y_6) \\
Q_7 = Q_8 &\Rightarrow P_7 \oplus Y_7 = P_8 \oplus Y_8 && (\text{because } P_7 \oplus P_8 = \alpha = Y_7 \oplus Y_8) \\
Q_7 = Q_9 &\Rightarrow P_7 \oplus Y_7 = P_9 \oplus Y_9 && (\text{because } P_7 \oplus P_9 = \beta = Y_7 \oplus Y_9) \\
P_1 = P_4 &\Rightarrow X_1 \oplus Q_1 = X_4 \oplus Q_4 && (\text{because } Q_1 \oplus Q_4 = \alpha' = X_1 \oplus X_4) \\
P_1 = P_7 &\Rightarrow X_1 \oplus Q_1 = X_7 \oplus Q_7 && (\text{because } Q_1 \oplus Q_7 = \beta' = X_1 \oplus X_7) \\
P_2 = P_5 &\Rightarrow X_2 \oplus Q_2 = X_5 \oplus Q_5 && (\text{because } Q_2 \oplus Q_5 = \alpha' = X_2 \oplus X_5) \\
P_2 = P_8 &\Rightarrow X_2 \oplus Q_2 = X_8 \oplus Q_8 && (\text{because } Q_2 \oplus Q_8 = \beta' = X_2 \oplus X_8) \\
P_3 = P_6 &\Rightarrow X_3 \oplus Q_3 = X_6 \oplus Q_6 && (\text{because } Q_3 \oplus Q_6 = \alpha' = X_3 \oplus X_6) \\
P_3 = P_9 &\Rightarrow X_3 \oplus Q_3 = X_9 \oplus Q_9 && (\text{because } Q_3 \oplus Q_9 = \beta' = X_3 \oplus X_9)
\end{aligned}$$

Therefore, from the exact value of H (given in section 3.3), and by considering only such (X, Y, P, Q) , we have:

$$H \geq 2^{4n} \cdot \frac{|F_n|^6}{2^{54n}} \cdot 2^{n(6+6+6+6+6)} = \frac{|F_n|^6}{2^{14n}},$$

as claimed (instead of $H \simeq \frac{|F_n|^6}{2^{18n}}$ if Ψ^6 was homogeneous).

Note: Here we have $6 + 6 = 12$ equalities $R_i = R_j$ or $S_i = S_k$, and only 8 equalities with nonzero constants have been used to satisfy all the (C) conditions.

So the deviation from the average is about $2^{(12-8)n}$.

Remark:

This attack shows that Ψ^6 is not homogenous as a generator of permutations. However for a specific permutation generated as a Ψ^6 this attack will not generally exist, since it requires more than 18 equations in only 9 indices and there are only 2^{2n} possible inputs.

3.4 $\forall k \in \mathbb{N}^*, \Psi^k$ is not homogeneous

Example 1 (with $m = (k/2)^2$)

For simplicity, we assume that k is even (the proof is very similar when k is odd). Let $k = 2\lambda$. Let $\Psi^k[L_i, R_i] = [S_i, T_i]$ for $1 \leq i \leq m$. We essentially generalize to Ψ^k the construction given in example 2 for Ψ^6 .

The exact value of H is:

$$H = \sum_{(X^{(1)}, \dots, X^{(k-2)}) \text{ satisfying } (C)} \frac{|F_n|^k}{2^{knm}} \cdot 2^{n(r+s+x^{(1)}+\dots+x^{(k-2)})},$$

where the $X^{(1)}, \dots, X^{(k-2)}$ variables are the intermediate round variables, and where (C) denotes the conditions on the equalities (i.e. $R_i = R_j \Rightarrow X_i^{(1)} \oplus L_i = X_j^{(1)} \oplus L_j$, etc). The proof of this formula is not difficult and is given in [8], p. 134.

We take $m = \lambda^2 (= \frac{k^2}{4})$.

We study the value H when $L_i, R_i, S_i, T_i, 1 \leq i \leq m$, satisfy the equalities illustrated by the figure 4. (For simplicity, we do not write these equalities explicitly).

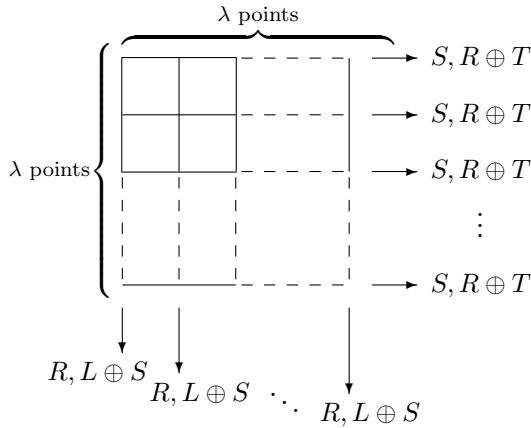


Figure 4: Modelling the $4 \cdot \lambda(\lambda - 1)$ equations in S, L, R, T .

We will consider values $X^{(1)}, \dots, X^{(k-2)}$ such that:

1. In Figure 4 the \oplus of two elements on the same line = 0 for $X^{(1)}, X^{(3)}, \dots, X^{(k-3)}$.
2. In Figure 4 the \oplus of two elements on the same column = 0 for $X^{(2)}, X^{(4)}, \dots, X^{(k-2)}$.
3. We have the $(k-2) \cdot (\lambda-1)$ equalities with non zero constant needed to satisfy all the (C) conditions.

Then, in the exact formula given above for H , for these $X^{(1)}, \dots, X^{(k-2)}$ we have: $r = \lambda(\lambda - 1)$, $s = \lambda(\lambda - 1)$, $X^{(1)} = \lambda(\lambda - 1)$, ... , $X^{(k-2)} = \lambda(\lambda - 1)$. And we have $2^{(k-2)n}$ possibilities for $X^{(1)}, \dots, X^{(k-2)}$. Then:

$$H \geq 2^{(k-2)n} \cdot \frac{|F_n|^k}{2^{knm}} \cdot 2^{nk\lambda(\lambda-1)},$$

so that, with $m = \lambda^2 = \frac{k^2}{4}$,

$$H \geq 2^{(k-2)n} \cdot \frac{|F_n|^k}{2^{2mn}}$$

(instead of $\frac{|F_n|^k}{2^{2nm}}$ if Ψ^k was homogeneous). Therefore, Ψ^k is not homogeneous, as claimed.

Note: Here we have $2\lambda(\lambda - 1)$ equalities $R_i = R_j$ or $S_i = S_k$, and only $(k - 2)(\lambda - 1)$ equalities with nonzero constants have been used to satisfy all the conditions (C). So the deviation from the average is about $2^{(k-2)n}$.

Example 2 (with $m = (k/2 - 1)^2$)

If we take $\lambda = \frac{k}{2} - 1$ (instead of $\lambda = \frac{k}{2}$), then we will have still $2\lambda(\lambda - 1)$ equalities in $S_i = S_j$ or $R_i = R_j$, $i \neq j$, and only $(k - 2)(\lambda - 1)$ equalities with nonzero constants to satisfy all the conditions (C). Here the obtained value of H will be about twice the average value. (This attack needs less points: $(\frac{k}{2} - 1)^2$ instead of $(\frac{k}{2})^2$, but the deviation from the average is less important).

Remark 1: The fact that Ψ^k is never homogeneous may explain why the proofs about the quality of pseudorandomness of the Ψ^k construction (such as theorem ??4 of section ??) are so difficult.

Remark 2: Here, in order to give an explicit construction with a non homogeneous property, we have taken $m = \mathcal{O}(k^2)$, where k is the number of rounds of the L-R construction, so m increases when k increases. It is possible to prove that this increase was a necessity: when m is fixed, then all the values of H are converging to the same value when k tends to infinity. (This property can be proved with ‘‘Markov chain’’ theory for example).

Remark 3: These attacks show that Ψ^k is not homogenous as a generator of permutations. However on a specific permutation generated as a Ψ^k this attack will not generally exist when $k \geq 7$ since there are only 2^{2n} inputs and the probability that on some of these inputs all the needed equations appear, is then negligible (when the permutation Ψ^k is fixed).

In conclusion:

Ψ^k is very strong pseudorandom $\Leftrightarrow k \geq 4$.

Ψ^k is never homogeneous (this was a surprise for me).

4 Open problem

An open problem that we mentioned is the following:

Is it possible to design homogeneous permutation generators ?

5 Conclusion

In this paper, we have defined two new natural notions about the quality of strong pseudorandom permutations: the concept of “very strong pseudorandomness” and the concept of “homogeneous permutations”. We have seen that no L-R construction gives homogeneous permutations. This result may be surprising, since it shows that – whatever the number of rounds of the L-R construction may be – there are still some “non-random places” in the resulting permutations (however, after a few rounds, the enemy is not able to choose the cleartexts or ciphertexts of his attack in order to be in one of these places: the scheme is pseudorandom).

References

- [1] W. Aiello, R. Venkatesan, *Foiling birthday attacks in length-doubling transformations*, EUROCRYPT’96, Springer, pp. 307-320.
- [2] M. Luby, C. Rackoff, *How to construct pseudorandom permutations from pseudorandom functions*, SIAM Journal on Computing, vol. 17, n. 2, pp. 373-386, April 1988.
- [3] M. Naor, O. Reingold, *On the Construction of Pseudo-Random Permutations: Luby-Rackoff revisited*, Electronic Colloquium on Computational Complexity (ECCC), Report TR 97-005. Preliminary version in: Proc. 29th Ann. ACM Symp. on Theory of Computing, 1997, pp. 189-199. To appear in the Journal of Cryptology.
- [4] U. Maurer, *A simplified and generalized treatment of Luby-Rackoff pseudorandom permutation generators*, Eurocrypt’92, Springer, pp. 239-255.
- [5] U. Maurer, J. Massey, *Local randomness in pseudorandom sequences*, Journal of Cryptology, vol. 4, pp. 135-149, 1991.
- [6] J. Patarin, *Pseudorandom Permutations based on the DES Scheme*, EUROCODE’90, LNCS 514, Springer, pp. 193-204.
- [7] J. Patarin, *New results on pseudorandom permutation generators based on the DES scheme*, CRYPTO’91, Springer, pp. 301-312.
- [8] J. Patarin, *Étude des Générateurs de Permutations Pseudo-aléatoires basés sur le schéma du DES*, Ph.D. Thesis, Université Paris VI, November 1991.
- [9] J. Patarin, *Improved security bounds for pseudorandom permutations*, 4th ACM Conference on Computer and Communications Security, April 1-4, 1997, pp. 142-150.

- [10] J. Pieprzyk, *How to construct pseudorandom permutations from single pseudorandom functions*, EUROCRYPT'90, Springer, pp. 140-150.
- [11] S. Vaudenay, *La Sécurité des Primitives Cryptographiques*, Ph.D. Thesis, École Normale Supérieure, April 1995, section II.8: "Les multipermutations".