# The "coefficients H" Technique

Jacques Patarin

Université de Versailles
45 avenue des Etats-Unis, 78035 Versailles Cedex, France
jacques.patarin@prism.uvsq.fr

**Abstract.** The "coefficient H technique" is a tool introduced in 1991 and used to prove various pseudo-random properties from the distribution of the number of keys that sends cleartext on some ciphertext. It can also be used to find attacks on cryptographic designs. We can like this unify a lot of various pseudo-random results obtained by different authors. In this paper we will present this technique and we will give some examples of results obtained.
(A more complete version of this paper will be given in the final proceedings of SAC'2008).

## 1 Introduction

The "coefficient H technique" was introduced in 1990 and 1991 in [16], [17]. Since then, it has been used many times (by myself , Henri Gilbert, Gilles Piret, Serge Vaudenay, etc.) to prove various results on pseudo-random functions and pseudo-random permutations. In this paper we will present in a self content way the "coefficient H technique", with different formulations when we study different cryptographic attacks (known plaintext attacks, chosen plaintext attacks, etc.). We will give proofs of some of these theorems and we will give some simple examples.

## 2 Notation - Definition of H

In all this paper, we will use these notations.

- KPA: Known Plaintext Attack
- CPA-1: Non-adaptive Chosen Plaintext Attack
- CPA-2: Adaptive Chosen Plaintext Attack
- CPCA-1: Non-adaptive Chosen Plaintext and Chosen Ciphertext Attack
- CPCA-2: Adaptive Chosen Plaintext and Chosen Ciphertext Attack
- $I_N = \{0,1\}^N$ ($N$ is any integer)
- $F_N$ will be the set of all applications from $I_N$ to $I_N$
- $B_N$ will be the set of all permutations from $I_N$ to $I_N$
- $\psi^k$ will denote the Feistel scheme of $F_{2n}$ with $k$ rounds with $k$ random round functions randomly chosen in $F_n$ ($n$ is any integer). $\psi^k$ is also called a random Feistel scheme or a Luby-Rackoff construction.

- $a \in_R A$ means that $a$ is randomly chosen in $A$ with a uniform distribution
- $K$ will denote a set of values that we will sometimes call "keys". In this paper we will consider that $K$ is a set of k-uples of functions $(f_1, \ldots, f_k)$ of $F_n$. (However generally only $|K|$ will be important, not the nature of the elements of $K$).
- $G$ is an application of $K \to F_N$. (Therefore, $G$ is a way to design a function of $F_N$ from k-uples $(f_1, \ldots, f_k)$ of functions of $F_n$ of $K$).

Let $m$ be an integer ($m$ will be the number of queries). Let $a = (a_i)_{1 \leq i \leq m}$ be a sequence of pairwise distinct elements of $I_N$. Let $b = (b_i)_{1 \leq i \leq m}$ be a sequence of elements of $I_N$. By definition, we will denote by $H(a, b)$ or simply by $H$ if the context of the $a_i$ and $b_i$ is clear, the number of $(f_1, \ldots, f_k) \in K$ such that:

$$\forall i,\ 1 \leq i \leq m,\ G(f_1, \ldots, f_k)(a_i) = b_i$$

Therefore, $H$ is the number of "keys" (i.e. elements of $K$) that send all the $a_i$ inputs to the exact values $b_i$.

## 3 Five basic "coefficient H" Theorems

In this section we will formulate five theorems. These theorems are the basis of a general proof technique called the "coefficient H technique", that allows to prove security results for function generators and permutation generators (and thus applies for random and pseudo-random Feistel ciphers)

These theorems were mentioned in [17] (with proofs in french) and in [20]. Since no proof in english was easily available so far we will present in this paper a proof of some of these theorems.

**Theorem 1** *(Coefficient H technique, sufficient condition for security against KPA)*

*Let $\alpha$ and $\beta$ be real numbers, $\alpha > 0$ and $\beta > 0$*

*If:*

*(1) For random values $a_i$, $b_i$, $1 \leq i \leq m$ of $I_N$ such that the $a_i$ are pairwise distinct, with probability $\geq 1 - \beta$ we have:*

$$H \geq \frac{|K|}{2^{Nm}}(1 - \alpha)$$

*Then*

*(2) For every KPA with $m$ (random) known plaintexts we have: $Adv^{KPA} \leq \alpha + \beta$, where $Adv^{KPA}$ denotes the advantage to distinguish $G(f_1, \ldots, f_k)$ when $(f_1, \ldots, f_k) \in_R K$ from a function $f \in_R F_N$*

(By "advantage" we mean here, as usual, for a distinguisher the absolute value of the difference of the two probability to output 1).

**Theorem 2** *(Coefficient H technique, sufficient condition for security against CPA-1)*

*Let $\alpha$ and $\beta$ be real numbers, $\alpha > 0$ and $\beta > 0$*

*If:*

*(1) For all sequences $a = (a_i)$, $1 \leq i \leq m$ of $m$ pairwise distinct elements of $I_N$ there exists a subset $E(a)$ of $I_N^m$ such that $|E(a)| \geq (1-\beta) \cdot 2^{Nm}$ and such that for all sequences $b = (b_i)$, $1 \leq i \leq m$ of $E(a)$ we have:*

$$H \geq \frac{|K|}{2^{Nm}}(1-\alpha)$$

*Then*

*(2) For every CPA-1 with $m$ chosen plaintexts we have: $Adv^{PRF} \leq \alpha + \beta$ where $Adv^{PRF}$ denotes the advantage to distinguish $G(f_1, \ldots, f_k)$ when $(f_1, \ldots, f_k) \in_R K$ from a function $f \in_R F_N$.*

**Theorem 3** *(Coefficient H technique, sufficient condition for security against CPA-2)*

*Let $\alpha$ and $\beta$ be real numbers, $\alpha > 0$ and $\beta > 0$. Let $E$ be a subset of $I_N^m$ such that $|E| \geq (1-\beta) \cdot 2^{Nm}$.*

*If:*

*(1) For all sequences $a_i$, $1 \leq i \leq m$, of pairwise distinct elements of $I_N$ and for all sequences $b_i$, $1 \leq i \leq m$, of $E$ we have:*

$$|H| \geq \frac{|K|}{2^{Nm}}(1-\alpha)$$

*Then*

*(2) For every CPA-2 with $m$ chosen plaintexts we have: $Adv^{PRF} \leq \alpha + \beta$ where $Adv^{PRF}$ denotes the probability to distinguish $G(f_1, \ldots, f_k)$ when $(f_1, \ldots, f_k) \in_R K$ from a function $f \in_R F_N$ (2).*

**Theorem 4** *(Coefficient H technique, sufficient condition for security against CPCA-2)*

*Let $\alpha$ be a real number, $\alpha > 0$.*

*If:*

*(1) For all sequences of pairwise distinct elements $a_i$, $1 \leq i \leq m$, and for all sequences pairwise distinct elements $b_i$, $1 \leq i \leq m$, we have:*

$$|H| \geq \frac{|K|}{2^{Nm}}(1-\alpha)$$

*Then*

*(2) For every CPA-2 with $m$ chosen plaintexts we have: $Adv^{PRF} \leq \alpha + \frac{m(m-1)}{2 \cdot 2^N}$ where $Adv^{PRF}$ denotes the probability to distinguish $G(f_1, \ldots, f_k)$ when $(f_1, \ldots, f_k) \in_R K$ from a function $f \in_R B_N$.*

**Theorem 5** *(Coefficient H technique, a more general sufficient condition for security against CPCA-2)*

*Let $\alpha$ and $\beta$ be real numbers, $\alpha > 0$ and $\beta > 0$*

*If:*

*(1a) There exists a subset $E$ of $(I_N^m)2$ such that for all $(a, b) \in E$, we have:*

$$|H| \geq \frac{|K|}{2^{Nm}}(1 - \alpha)$$

*(1b) For all CPCA-2 acting on a random permutation $f$ of $B_N$, the probability that $(a, b) \in E$ is $\geq 1 - \beta$ where $(a, b)$ denotes here the successive $b_i = f(a_i)$ or $a_i = f^{-1}(b_i)$, $1 \leq i \leq m$ that will appear.*

*Then*

*(2) For every CPCA-2 with $m$ chosen plaintexts we have: $Adv^{PRF} \leq \alpha + \frac{m(m-1)}{2 \cdot 2^N}$ where $Adv^{PRF}$ denotes the probability to distinguish $G(f_1, \ldots, f_k)$ when $(f_1, \ldots, f_k) \in_R K$ from a function $f \in_R B_N$.*

## 4 Proof of Theorem 1

Let $\phi$ be an algorithm (with no limitations in the number of computations) that takes the $(a_i, b_i)$, $1 \leq i \leq m$ in input and outputs 0 or 1. let $P_1$ be the probability that $\phi$ outputs 1 when $\forall i$, $1 \leq i \leq m$ $b_i = G(f_1, \ldots, f_k)(a_i)$ when $(f_1, \ldots, f_k) \in_R K$. Let $P_1^*$ be the probability that $\phi$ outputs 1 when $b_i = F(a_i)$ when $F \in_R F_N$. We want to prove that $|E(P_1 - P_1^*)| \alpha + \beta$. Let $D$ be the set of all pairwise distinct $a_i$, $1 \leq i \leq m$ (so $|D| \simeq 2^{Nm}(1 - \frac{m(m-1)}{2 \cdot 2^N})$). When the $a_i$, $1 \leq i \leq m$ are fixed, let $W(a)$ be the set of all $b_1, \ldots, b_m$ such that the algorithm $\phi$ outputs 1 on the input $(a_i, b_i)$, $1 \leq i \leq m$. When the $a_i$, $1 \leq i \leq m$ are fixed in $D$, then we have:

$$P_1^* = \frac{|W(a)|}{2^{Nm}} \quad (1)$$

and

$$P_1 = \frac{1}{|K|} \sum_{b \in W(a)} [Numbers\ of\ (f_1, \ldots, f_k) \in K / \forall i,\ 1 \leq i \leq m,\ G(f_1, \ldots, f_k)(a_i) = b_i]$$

so

$$P_1 = \frac{1}{|K|} \sum_{b \in W(a)} H(a, b) \quad (2)$$

Moreover, by hypothesis we have that the number $\mathcal{N}$ of $(a, b)$ such that

$$H(a, b) \geq \frac{|K|}{2^{Nm}}(1 - \alpha) \text{ satifies} : \mathcal{N} \geq |D| \cdot 2^{Nm}(1 - \beta) \quad (3)$$

When the $(a_i)$, $1 \leq i \leq m$ are fixed, let $\mathcal{N}(a)$ be the set of all $b$ such that:

$$H(a, b) \geq \frac{|K|}{2^{Nm}}(1 - \alpha)$$

From (3) we have:

$$\sum_{a \in D} |\mathcal{N}(a)| \geq |D| \cdot 2^{Nm}(1-\beta) \quad (4)$$

From (2) we have:

$$P_1 \geq \frac{1}{|K|} \sum_{b \in W(a) \cap \mathcal{N}(a)} H(a,b)$$

so

$$P_1 \geq \frac{(1-\alpha)}{2^{Nm}} |W(a) \cap \mathcal{N}(a)|$$

so

$$P_1 \geq \frac{(1-\alpha)}{2^{Nm}} (|W(a)| - |\mathcal{N}'(a)|) \quad (5)$$

where $\mathcal{N}'(a)$ is the set of all $b$ such that $b \notin \mathcal{N}(a)$. $|\mathcal{N}'(a)| = 2^{Nm} - |\mathcal{N}(a)|$, so

$$\sum_{a \in D} |\mathcal{N}'(a)| = |D| 2^{Nm} - \sum_{a \in D} |\mathcal{N}(a)|$$

so from (4) we have:

$$\sum_{a \in D} |\mathcal{N}'(a)| \leq \beta \cdot |D| \cdot 2^{Nm}, \text{ so } E(|\mathcal{N}'(a)|) \leq \beta \cdot 2^{Nm} \quad (6)$$

(where the expectation is computed when the $(a_i)$, $1 \leq i \leq m$ are randomly chosen in $D$). From (5) and (1) we have:

$$P_1 \geq (1-\alpha)(P_1^* - \frac{|\mathcal{N}'(a)|}{2^{Nm}})$$

$$P_1 \geq (1-\alpha)P_1^* - \frac{|\mathcal{N}'(a)|}{2^{Nm}}$$

so from (6) we get:

$$E(P_1) \geq (1-\alpha)E(P_1^*) - \beta$$

so

$$E(P_1) \geq E(P_1^*) - \alpha - \beta \quad (7)$$

Now if we consider the algorithm $\phi'$ that outputs 1 if and only if $\phi$ outputs 0, we have $P_1' = 1 - P_1$ and $P_1^{'*} = 1 - P_1^*$ and from (7) we get: $E(P_1') \geq E(P_1^{'*}) - \alpha - \beta$ (because (7) is true for all algorithm $\phi$, so it is true for $\phi'$). So

$$E(1 - P_1) \geq E(1 - P_1^*) - \alpha - \beta$$

so

$$E(P_1) - E(P_1^*) \leq \alpha + \beta \quad (8)$$

From (7) and (8) we get $|E(P_1 - P_1^*)| \leq \alpha + \beta$ as claimed.

## 5 Proof of Theorem 3

(We follow here a proof, in French, of this Theorem in J.Patarin, PhD Thesis, 1991, Page 27).

Let $\phi$ be a (deterministic) algorithm which is used to test a function $f$ of $F_n$. ($\phi$ can test any function $f$ from $I_N \rightarrow I_N$). $\phi$ can use $f$ at most $m$ times, that is to say that $\phi$ can ask for the values of some $f(C_i)$, $C_i \in I_N$, $1 \leq i \leq m$. (The value $C_1$ is chosen by $\phi$, then $\phi$ receive $f(C_1)$, then $\phi$ can choose any $C_2 \neq C_1$, then $\phi$ receive $f(C_2)$ etc). (Here we have adaptive chosen plaintexts). (If $i \neq j$, $C_i$ is always different from $C_j$). After a finite but unbounded amount of time, $\phi$ gives an output of "1" or "0". This output (1 or 0) is noted $\phi(f)$.

We will denote by $P_1^*$, the probability that $\phi$ gives the output 1 when $f$ is chosen randomly in $F_n$. Therefore

$$P_1^* = \frac{\text{Number of functions } f \text{ such that } \phi(f) = 1}{|F_N|}$$

where $|F_N| = 2^{N \cdot 2^N}$.

We will denote by $P_1$, the probability that $\phi$ gives the output 1 when $(f_1, \ldots, f_k) \in_R K$ and $f = G(f_1, \ldots, f_k)$. Therefore

$$P_1 = \frac{\text{Number of } (f_1, \ldots, f_k) \in K \text{ such that } \phi(G(f_1, \ldots, f_k)) = 1}{|K|}$$

We will prove:

(**"Main Lemma"**): For all such algorithms $\phi$,

$$|P_1 - P_1^*| \leq \alpha + \beta$$

Then Theorem 1 will be an immediate corollary of this "Main Lemma" since $Adv^{PRF}$ is the best $|P_1 - P_1^*|$ that we can get with such $\phi$ algorithms.

**Proof of the "Main Lemma"**

**Evaluation of $P_1^*$**

Let $f$ be a fixed function, and let $C_1, \ldots, C_m$ be the successive values that the program $\phi$ will ask for the values of $f$ (when $\phi$ tests the function $f$). We will note $\sigma_1 = f(C_1), \ldots, \sigma_m = f(C_m)$. $\phi(f)$ depends **only** of the outputs $\sigma_1, \ldots, \sigma_m$. That is to say that if $f'$ is another function of $F_n$ such that $\forall i$, $1 \leq i \leq m$, $f'(C_i) = \sigma_i$, then $\phi(f) = \phi(f')$. (Since for $i < m$, the choice of $C_{i+1}$ depends only of $\sigma_1, \ldots, \sigma_i$. Also the algorithm $\phi$ cannot distinguish $f$ from $f'$, because $\phi$ will ask for $f$ and $f'$ exactly the same inputs, and will obtain exactly the same outputs). Conversely, let $\sigma_1, \ldots, \sigma_n$ be $m$ elements of $I_N$. Let $C_1$ be the first value that $\phi$ choose to know $f(C_1)$, $C_2$ the value that $\phi$ choose when $\phi$ has obtained the answer $\sigma_1$ for $f(C_1)$,..., and $C_m$ the $m^{th}$ value that $\phi$ presents to $f$, when $\phi$ has obtained $\sigma_1, \ldots, \sigma_{m-1}$ for $f(C_1), \ldots, f(C_{m-1})$. Let $\phi(\sigma_1, \ldots, \sigma_m)$ be the output of $\phi$ (0 or 1). Then

$$P_1^* = \sum_{\substack{\sigma_1, \ldots, \sigma_n \\ \phi(\sigma_1, \ldots \sigma_m) = 1}} \frac{\text{Number of functions } f \text{ such that } \forall i,\ 1 \leq i \leq m,\ f(C_i) = \sigma_i}{2^{N \cdot 2^N}}$$

Since the $C_i$ are all distinct the number of functions $f$ such that $\forall i,\ 1 \leq i \leq m,\ f(C_i) = \sigma_i$ is exactly $|F_n|/2^{nm}$. Therefore

$$P_1^* = \frac{\text{Number of outputs } (\sigma_1, \ldots, \sigma_m) \text{ such that } \phi(\sigma_1, \ldots \sigma_m) = 1}{2^{Nm}}$$

Let $\mathcal{N}$ be the number of outputs $\sigma_1, \ldots, \sigma_m$ such that $\phi(\sigma_1, \ldots \sigma_m) = 1$. Then $P_1^* = \frac{\mathcal{N}}{2^{Nm}}$.

**Evaluation of $P_1$**

With the same notation $\sigma_1, \ldots, \sigma_n$, and $C_1, \ldots C_m$:

$$P_1 = \frac{1}{|K|} \sum_{\substack{\sigma_1, \ldots, \sigma_n \\ \phi(\sigma_1, \ldots \sigma_m)=1}} [\text{Number of } (f_1, \ldots, f_k) \in K \text{ such that}$$

$$\forall i,\ 1 \leq i \leq m,\ G(f_1, \ldots, f_k)(C_i) = \sigma_i] \quad (3)$$

Now (by definition of $\beta$) we have at most $\beta \cdot 2^{nm}$ sequences $(\sigma_1, \ldots, \sigma_m)$ such that $(\sigma_1, \ldots, \sigma_m) \notin E$. Therefore, we have at least $\mathcal{N} - \beta \cdot 2^{Nm}$ sequences $(\sigma_1, \ldots, \sigma_m)$ such that $\phi(\sigma_1, \ldots \sigma_m) = 1$ and $(\sigma_1, \ldots, \sigma_m) \in E$ (4). Therefore, from (1), (3) and (4), we have

$$P_1 \geq \frac{(\mathcal{N} - \beta \cdot 2^{Nm}) \cdot \frac{|K|}{2^{Nm}}(1-\alpha)}{|K|}$$

Therefore

$$P_1 \geq \big(\frac{\mathcal{N}}{2^{Nm}} - \beta\big)(1-\alpha)$$

$$P_1 \geq (P_1^* - \beta)(1-\alpha)$$

Thus $P_1 \geq P_1^* - \alpha - \beta$ (5), as claimed.

We now have to prove the inequality in the other side. For this, let $P_0^*$ be the probability that $\phi(f) = 0$ when $f \in_R F_N$. $P_0'^* = 1 - P_1^*$. Similarly, let $P_0$ be the probability that $\phi(f) = 0$ when $(f_1, \ldots, f_k) \in_R K$ and $f = G(f_1, \ldots, f_k)$. $P_0 = 1 - P_1$. We will have $P_0 \geq P_0^* - \alpha - \beta$ (since the outputs 0 and 1 have symmetrical hypothesis. Or, alternatively since we can always consider an algorithm $\phi'$ such that $\phi'(f) = 0 \Leftrightarrow \phi(f) = 1$ and apply (5) to this algorithm $\phi'$).

Therefore, $1 - P_1 \geq 1 - P_1^* - \alpha - \beta$, i.e. $P_1^* \geq P_1 - \alpha - \beta$ (6). Finally, from (5) and (6), we have: $|P_1 - P_1^*| \leq \alpha + \beta$, as claimed.

## 6 Examples

### 6.1 $\psi^2$

For $\psi^2$ (Feistel scheme with the round functions $(f_1, f_2) \in_R F_n^2$) let $[L_i, R_i]$, $1 \leq i \leq m$ denotes the inputs, and $[S_i, T_i]$, $1 \leq i \leq m$ denotes the outputs. We have: $S_i = L_i \oplus f_1(R_i)$ and $T_i = R_i \oplus f_2(S_i)$ (*)

For random values $[L_i, R_i], [S_i, T_i]$, $1 \leq i \leq m$ (such that $i \neq j \rightarrow L_i \neq L_j$ or $R_i \neq R_j$ with probability $> 1 - \frac{m2}{2^n}$ we have that all the $R_i$ values are pairwise

distinct and all the $S_i$ values are pairwise distinct. Moreover, if this occurs, we have exactly $H = \frac{|F_n|^2}{2^{nm}}$ (since (*) then fix $f_1$ exactly on $m$ points and $f_2$ exactly on $m$ points).

So from Theorem 1 (with $\alpha = 0$ and $\beta = \frac{m2}{2^n}$) we get:

**Theorem 6** *For two rounds, and for all algorithm A that takes m cleartext/ciphertext pairs input, and outputs 0 or 1, we have:*

$$|E(P_1 - P_1^*)| \leq \frac{m2}{2^n}$$

*when the cleartexts are randomly chosen. So when $m \leq 2^{n/2}$, $\psi^2$ will resit all known plaintext attacks.*

**Remark**. This result is tight, since when $m2$ becomes not negligible compared with $2^n$n then by counting the number $\mathcal{N}$ of $(i,j)/S_i \oplus L_i = S_j \oplus L_j$ we will be able to distinguish $\psi^2$ from a random permutation with a known plaintext attack.

With $k = 2$, $K = |B_n|^2$ and $G(f_1, \ldots, f_k) = f_1 \oplus f_2$ we obtain immediately:

**Theorem 7** *Let $\alpha$ and $\beta$ be real numbers, $\alpha > 0$ and $\beta > 0$. Let $E$ be a subset of $I_n^m$ such that $|E| \geq (1-\beta) \cdot 2^{nm}$.*

*If:*

*1) For all sequences $a_i$, $1 \leq i \leq m$, of pairwise distinct elements of $I_n$ and for all sequences $b_i$, $1 \leq i \leq m$, of $E$ we have:*

$$|H| \geq \frac{|B_n|^2}{2^{nm}}(1-\alpha)$$

*where $H$ denotes the number of $(f,g) \in B_n^2$ such that*

$$\forall i,\ 1 \leq i \leq m,\ f \oplus g(a_i) = b_i$$

*Then*

*2) For every CPA-2 with $m$ chosen plaintexts we have: $p \leq \alpha + \beta$ where $p = Adv_\phi^{PRF}$ denotes the probability to distinguish $f \oplus g$ when $(f,g) \in_R B_n^2$ from a function $h \in_R F_n$.*

### 6.2 Involutive permutations

Let assume that $G$ is a generator of permutations that generates involutive permutations $f$ (i.e. $f = f^{-1}$). then we can distinguish such $f$ from random permutations of $B_N$ with $m = 2$ queries in CPA-2 and $m = 2$ queries in CPCA-1.

**CPA-2**

In CPA-2 we ask $f(a_1) = b_1$ and $f(b_1) = b_2$, and we test if $b_2 = a_1$ This gives a CPA-2 with $m = 2$ queries. It is not in contradiction with Theorem 3 since in Theorem 3, we need property (1) on **all** sequences $a_i$, $1 \leq i \leq m$ (and

not necessary on **all** sequences $b_i$). Here if we have $a = (a_1, a_2)$, $b = (b_1, b_2)$ with $a_2 = b_1$ and $b_2 \neq a_1$, we will have $H = 0$. Therefore we will not be able to prove from Theorem 3 that $G$ is secure in CPA-2 (in fact $G$ is not secure in CPA-2) since for most $(b_1, b_2)$ there exists $(a_1, a_3)$ (take $a_2 = b_1$) such that $H = 0$.

**CPCA-1**

In CPCA-1 we ask $f(a_1) = b_1$ and $f^{-1}(a_1) = a_2$ and we test if $a_2 = b_1$. This gives a CPCA-1 distinguisher with $m = 2$ queries. We will not be able to prove from Theorem 2 that $G$ is secure in CPCA-1 (in fact $G$ is not secure in CPCA-1) since in a non-adaptive chosen plaintext/ciphertext attack we can impose that $b_2 = a_1$ of we have $a = (a_1, a_1)$, $b = (b_1, b_2)$ with $b_2 = a_1$ and $a_2 \neq b_1$ we will have $H = 0$.

### 6.3 Secret key security hierarchy

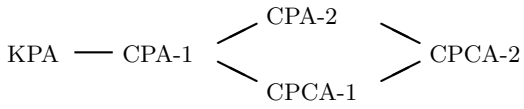


**Fig. 1.** Hierarchy of the attacks in secret key cryptography

In Figure 1, we have the well known hierarchy of attacks in secret key cryptography (cf [2], [4], [5]). With coefficients H technique we can easily prove on small examples this hierarchy, i.e. for example that there are some scheme secure in CPA-2 and not in CPCA-1, that some schemes are secure in CPA-1 and not in KPA etc. For example, we can easily prove that for a random involutive permutation of $B_N$ we will have KPA and CPA-1 security in $O(\sqrt{2^N})$. Therefore the example of Section 6.2 shows that CPA-1 $<$ CPA-2 and that CPA-1 $<$ CPCA-1.

With $f$ such that $f(0) = 0$ we will have that KPA $<$ CPA-1.

With $\psi^2$ we will have KPA $<$ CPA-1.

With $\psi^3$ we will have CPA-2 $<$ CPCA-2 and CPCA-1 $<$ CPCA-2.

With a random permutation such that $f3 = Id$ we see that sometimes CPA-2 $>$ CPCA-1

With a random permutation, such that $f^{-1}(x) = f(x) \oplus k$ where $k$ is a secret constant we see that sometimes CPCA-1 $>$ CPA-2.

## 7 Proofs with coefficient H

### 7.1 Feistel schemes $\psi^k$

I have proved many security results on $\psi^k$ generators with coefficient H. For example, in [15], the security of $\psi^5$ when $m \ll 2^n$ was proved.

### 7.2 Xor of two random permutations

With $k = 2$, $K = |B_n|^2$ and $G(f_1, \ldots, f_k) = f_1 \oplus f_2$ we obtain immediately:

**Theorem 8** *Let $\alpha$ and $\beta$ be real numbers, $\alpha > 0$ and $\beta > 0$. Let $E$ be a subset of $I_n^m$ such that $|E| \geq (1-\beta) \cdot 2^{nm}$.*

*If:*

*1) For all sequences $a_i$, $1 \leq i \leq m$, of pairwise distinct elements of $I_n$ and for all sequences $b_i$, $1 \leq i \leq m$, of $E$ we have:*

$$|H| \geq \frac{|B_n|^2}{2^{nm}}(1-\alpha)$$

*where $H$ denotes the number of $(f,g) \in B_n^2$ such that*

$$\forall i,\ 1 \leq i \leq m,\ f \oplus g(a_i) = b_i$$

*Then*

*2) For every CPA-2 with $m$ chosen plaintexts we have: $p \leq \alpha + \beta$ where $p = Adv_{\phi}^{PRF}$ denotes the probability to distinguish $f \oplus g$ when $(f,g) \in_R B_n^2$ from a function $h \in_R F_n$.*

I have conjectured this property:

$$\forall f \in F_n,\ \text{if} \bigoplus_{x \in I_n} f(x) = 0,\ \text{then } \exists (g,h) \in B_n 2,\ \text{such that } f = g \oplus h.$$

Just one day after this paper was put on eprint, J.F. Dillon pointed to us that in fact this was proved in 1952 in [3]. We thank him a lot for this information. (This property was proved again independently in 1979 in [24]).

**A new conjecture.** However I conjecture a stronger property. Conjecture:

$$\forall f \in F_n,\ \text{if} \bigoplus_{x \in I_n} f(x) = 0,\ \text{then the number } H \text{ of } (g,h) \in B_n 2,$$

$$\text{such that } f = g \oplus h \text{ satisfies } H \geq \frac{|B_n|^2}{2^{n2^n}}.$$

Variant: I also conjecture that this property is true in any group, not only with Xor.

**Remark:** in this paper, I have proved weaker results involving $m$ equations with $m \ll O(2^n)$ instead of all the $2^n$ equations. These weaker results were sufficient for the cryptographic security wanted.

### 7.3 Benes schemes

In [11] the security of Benes schemes when $m \ll 2^n$ was finally obtained (after the beginning of some proof ideas in [1]).

## 8 Attacks with coefficient H

By using the coefficient values we were able to find many generic attacks. We give here some examples.

### 8.1 For Feistel schemes $\psi^k$

From [14] we have the results.

| | KPA | CPA-1 | CPA-2 | CPCA-1 | CPCA-2 |
|---|---|---|---|---|---|
| $\Psi$ | 1 | 1 | 1 | 1 | 1 |
| $\Psi 2$ | $2^{n/2}$ | 2 | 2 | 2 | 2 |
| $\Psi 3$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ | 3 |
| $\Psi 4$ | $2^n$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ |
| $\Psi 5$ | $\leq 2^{3n/2}$ | $2^n$ | $2^n$ | $2^n$ | $2^n$ |
| $\Psi 6$ | $\leq 2^{2n}$ | $\leq 2^{2n}$ | $\leq 2^{2n}$ | $\leq 2^{2n}$ | $\leq 2^{2n}$ |
| $\Psi 7$ | $\leq 2^{3n}$ | $\leq 2^{3n}$ | $\leq 2^{3n}$ | $\leq 2^{3n}$ | $\leq 2^{3n}$ |
| $\Psi 8$ | $\leq 2^{4n}$ | $\leq 2^{4n}$ | $\leq 2^{4n}$ | $\leq 2^{4n}$ | $\leq 2^{4n}$ |
| $\Psi^k, k \geq 6$ * | $\leq 2^{(k-4)n}$ | $\leq 2^{(k-4)n}$ | $\leq 2^{(k-4)n}$ | $\leq 2^{(k-4)n}$ | $\leq 2^{(k-4)n}$ |

**Table 1.** Minimum number $\lambda$ of computations needed to distinguish a generator $\Psi^k$ (with one or many such permutations available) from random permutations with an even signature of $I_n \to I_n$. For simplicity we denote $\alpha$ for $O(\alpha)$. $\leq$ means best known attack.

* If $k \geq 7$ these attacks analyze about $2^{(k-6)n}$ permutations of the generator and if $k \leq 6$ only one permutation is needed.

### 8.2 For Feistel schemes $\psi'^k$ with $k$ random permutations for the rounds functions (instead of round functions)

From [26] we have these results:

| number $k$ of rounds | KPA | CPA-1 | CPA-2 | CPCA-1 | CPCA-2 |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | $2^{n/2}$ | 2 | 2 | 2 | 2 |
| 3 | $2^n(+)$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ | 3 |
| 4 | $2^n$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ | $2^{n/2}$ |
| 5 | $2^{3n/2}$ | $2^n$ | $2^n$ | $2^n$ | $2^n$ |
| 6 | $2^{3n}(+)$ | $2^{3n}(+)$ | $2^{3n}(+)$ | $2^{3n}(+)$ | $2^{3n}(+)$ |
| 7 | $2^{3n}$ | $2^{3n}$ | $2^{3n}$ | $2^{3n}$ | $2^{3n}$ |
| 8 | $2^{4n}$ | $2^{4n}$ | $2^{4n}$ | $2^{4n}$ | $2^{4n}$ |
| 9 | $2^{6n}(+)$ | $2^{6n}(+)$ | $2^{6n}(+)$ | $2^{6n}(+)$ | $2^{6n}(+)$ |
| 10 | $2^{6n}$ | $2^{6n}$ | $2^{6n}$ | $2^{6n}$ | $2^{6n}$ |
| 11 | $2^{7n}$ | $2^{7n}$ | $2^{7n}$ | $2^{7n}$ | $2^{7n}$ |
| 12 | $2^{9n}(+)$ | $2^{9n}(+)$ | $2^{9n}(+)$ | $2^{9n}(+)$ | $2^{9n}(+)$ |
| $k \geq 6$, $k=0 \ mod\, 3$ | $2^{(k-3)n}(+)$ | $2^{(k-3)n}(+)$ | $2^{(k-3)n}(+)$ | $2^{(k-3)n}(+)$ | $2^{(k-3)n}(+)$ |
| $k \geq 6$, $k=1$ or $2 \ mod\, 3$ | $2^{(k-4)n}$ | $2^{(k-4)n}$ | $2^{(k-4)n}$ | $2^{(k-4)n}$ | $2^{(k-4)n}$ |

**Table 2.** Maximum number of computations needed to get an attack on a $k$-round Feistel network with internal *permutations* $(+)$ is shown when the values are larger than the corresponding values with internal functions .

### 8.3 For unbalanced Feistel schemes with contracting functions

From [21] we have these results.

**Table 3.** Results on $G_k^d$ for any $k \geq 4$. For more than $2k$ rounds more that one permutation is needed or more than $2^{(2k-4)n}$ computations are needed in the best known attacks to distinguish from a random permutation with an even signature.

| | KPA | CPA-1[a] |
|---|---|---|
| $G_k^d$, $1 \leq d \leq k-1$ | 1 | 1 |
| $G_k^k$ | $2^{\frac{n(k-1)}{2}}$ | 2 |
| $G_k^{k+1}$ | $2^{\frac{n(k-1)}{2}}$ | $2^{\frac{n}{2}}$ |
| $G_k^{k+2}$ | $2^{\frac{k}{2}n}$ | $2^{\frac{3}{2}n}$ |
| $G_k^{k+3}$ | $2^{(\frac{k+1}{2})n}$ | $2^{\frac{5}{2}n}$ |
| $G_k^{k+i}$, $1 \leq i < k$ | $2^{(\frac{k+i-2}{2})n}$ | $2^{(\frac{2i-1}{2})n}$ |
| $G_k^{2k}$ | $2^{(2k-4)n}$ | $2^{(2k-4)n}$ |
| $G_k^d$, $d \geq 2k$ | $2^{(d+(k-2)\lfloor\frac{d}{k}\rfloor-2k)n}$ | $2^{(d+(k-2)\lfloor\frac{d}{k}\rfloor-2k)n}$ |

[a] Here we do not show CPA-2, CPCA-1 and CPCA-2 since for $G_k^d$, no better attacks are found compared with CPA-1.

### 8.4 For unbalanced Feistel schemes with expanding functions

From [22] we have these results.

**Table 4.** Best known attacks on $F_k^d$ for $k \geq 3$.

| | KPA | CPA-1 |
|---|---|---|
| $F1_k$ | $1$ | $1$ |
| $F2_k$ | $2^{\frac{n}{2}}$ | $2$ |
| $F3_k$ | $2^n$ | $2$ |
| $F_k^d$, $2 \leq d \leq k$ | $2^{\frac{d-1}{2}n}$ | $2$ |
| $F_k^{k+1}$ | $2^{\frac{k}{2}n}$ | $2^{\frac{n}{2}}$ |
| $F_k^{k+2}$ | $2^{\frac{k+1}{2}n}$ | $2^n$ |
| $F_k^{k+3}$ | $2^{\frac{2k+3}{4}n}$ | $2^{2n}$ or $2^{\frac{k+2}{3}n}$ |
| $F_k^d$, $k+2 \leq d \leq 2k$ | $2^{\frac{d+k}{4}n}$ | $2^{(d-k-1)n}$ or $2^{\frac{d-1}{3}n}$ |
| $F_k^{2k}$ | $2^{\frac{3k}{4}n}$ | $2^{\frac{2k-1}{3}n}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $F_k^{3k-1}$ | $2^{(k-\frac{1}{8})n}$ | $2^{(k-\frac{1}{2})n}$ |
| $F_k^{3k}$ | $2^{kn}$ | $2^{kn}$ |
| $F_k^d$, $3k \leq d \leq k2$ | $2^{(d-2k)n}$ | $2^{(d-2k)n}$ |
| $F_k^{k2}$ | $2^{(k2-2k)n}$ | $2^{(k2-2k)n}$ |
| ..... | ..... | ..... |
| $F_k^{k2+1}$ | $2^{(2k2-3k-2)n}$ | $2^{(2k2-3k-2)n}$ |
| $F_k^d$, $d \geq k2+1$ | $2^{(\lfloor 2d(1-\frac{1}{k})\rfloor -k-3)n}$ | $2^{(\lfloor 2d(1-\frac{1}{k})\rfloor -k-3)n}$ |

## 9 New Designs

### 9.1 Russian doll design

See [23] in this conference SAC' 2008.

### 9.2 Design from Random Unbalanced Feistel Schemes

This design comes directly from Table 3. (More details in the final version of this paper).

### 9.3 Hash Function Design

From 9.1 and 9.2 we are analysing a Hash function design (by Xoring two independent pseudorandom permutations, or by Xoring the input and the output of a pseudorandom permutation).

## 10 Conclusion

With the "coefficient H technique" we were able to prove many security results and to get many generic attacks. Moreover, it was a source of inspiration for the design of new schemes.

## References


1. W. Aiello and R. Venkatesan. Foiling Birthday Attacks in Length-Doubling Transformations - Benes: A Non-Reversible Alternative to Feistel. In Ueli M. Maurer, editor, *Advances in Cryptology – EUROCRYPT '96*, volume 1070 of *Lecture Notes in Computer Science*, pages 307–320. Springer-Verlag, 1996.
2. M. Bellare, A. Desai, E. Jokipii, and P. Rogaway. A Concrete Security Treatment of Symmetric Encryption: Analysis of the DES Modes of Operation. The preliminary version of this paper was entitled "A Concrete Security Treatment of Symmetric Encryption" and appeared in the Proceedings of 38th Annual Symposium of Computer Science, IEEE . 1997.
3. M. Hall Jr. A Combinatorial Problem on Abelian Groups. *Proceedings of the Americal Mathematical Society*, 3(4):584–587, 1952.
4. J. Katz and M. Yung. Characterization of Security Notions for Probabilistic. In *Private-Key Encription – STOC '2000*.
5. J. Katz and M. Yung. Unforgeable Encryption and Chosen-Ciphertext-Secure Modes of Operation. In *Fast Software Encryption 2000*.
6. M. Luby and C. Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM J. Comput.*, 17(2):373–386, 1988.
7. U. Maurer. A simplified and generalized treatment of Luby-Rackoff pseudorandom permutation generators. In *Advances in Cryptology – EUROCRYPT '92*, Lecture Notes in Computer Science, pages 239–255. Springer-Verlag, 1992.
8. U. Maurer. Indistinguishability of Random Systems. In *Advances in Cryptology – EUROCRYPT '02*, volume 2332 of *Lecture Notes in Computer Science*, pages 100–132. Springer-Verlag, 2002.
9. U. Maurer and K. Pietrzak. The Security of Many-Round Luby-Rackoff Pseudo-Random Permutations. In Eli Biham, editor, *Advances in Cryptology – EUROCRYPT 2003*, volume 2656 of *Lecture Notes in Computer Science*, pages 544–561. Springer-Verlag, 2003.
10. M. Naor and O. Reingold. On the Construction of Pseudorandom Permutations: Luby-Rackoff Revisited. *J. Cryptology*, 12(1):29–66, 1999.
11. J. Patarin. A Proof of Security in $O(2^n)$ for the Benes Scheme. In *AFRICACRYPT '2008*.
12. J. Patarin. A Proof of Security in $O(2^n)$ for the Xor of Two Random Permutations. In *ICITS '2008*.
13. J. Patarin. Generic Attacks for the Xor of k Random Permutations. eprint 2008.
14. J. Patarin. Generic Attacks on Feistel Schemes. In Colin Boyd, editor, *Advances in Cryptology – ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 222–238. Springer-Verlag. A more complete version of this paper is on eprint archives 2008.
15. J. Patarin. On Linear Systems of Equations with Distinct Variables and Small Block Size. In *ICISC 2005*.

16. J. Patarin. Pseudorandom Permutations based on the DES Scheme. In *EUROCODE '90*, volume 514 of *Lecture Notes in Computer Science*, pages 193–204. Springer-Verlag, 1990.
17. J. Patarin. Etude de Générateurs de Permutations Basés sur les Schémas du DES. In *Ph. Thesis*. Inria, Domaine de Voluceau, France, 1991.
18. J. Patarin. New Results on Pseudorandom Permutation Generarors based on the DES Scheme. In *CRYPTO '91*, Lecture Notes in Computer Science, pages 310–312. Springer-Verlag, 1991.
19. J. Patarin. How to Construct Pseudorandom and Super Pseudorandom Permutations from a Single Permutation. In *EUROCRYPTt '92*, Lecture Notes in Computer Science, pages 256–266. Springer-Verlag, 1992.
20. J. Patarin. Luby-Rackoff: 7 Rounds are Enough for $2^{n(1-\epsilon)}$ Security. In Dan Boneh, editor, *Advances in Cryptology – CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 513–529. Springer-Verlag, 2003.
21. J. Patarin, V. Nachef, and C. Berbain. Generic Attacks on Unbalanced Feistel Schemes with Contracting Functions. In Xuejia Lai and Kefei Chen, editors, *Advances in Cryptology – ASIACRYPT 2006*, volume 4284 of *Lecture Notes in Computer Science*, pages 396–411. Springer-Verlag, 2006.
22. J. Patarin, V. Nachef, and C. Berbain. Generic Attacks on Unbalanced Feistel Schemes with Expanding Functions. In Kaoru Kurosawa, editor, *Advances in Cryptology – ASIACRYPT 2007*, volume 4833 of *Lecture Notes in Computer Science*, pages 325–341. Springer-Verlag, 2007.
23. J. Patarin and Y. Seurin. Building Secure Block Ciphers on Generic Attacks Assumptions. In *SAC '2008*.
24. F. Salzborn and G. Szekeres. A Problem in Combinatorial Group Theory. *Ars Combinatoria*, 7:3–5, 1979.
25. B. Schneier and J. Kelsey. Unbalanced Feistel Networks and Block Cipher Design. In Dieter Gollmann, editor, *Fast Software Encrytion – FSE '96*, volume 1039 of *Lecture Notes in Computer Science*, pages 121–144. Springer-Verlag, 1996.
26. J. Treger and J. Patarin. Generic Attacks On Feistel Schemes with Internal Permutations. Paper in preparation.