SAPIENZA
UNIVERSITÀ DI ROMA

"SAPIENZA" UNIVERSITY OF ROME
FACULTY OF INFORMATION ENGINEERING,
INFORMATICS AND STATISTICS
DEPARTMENT OF COMPUTER SCIENCE

Cryptography

Author
Alessio Bandiera

October 5, 2025

Contents

Information and Contacts

Personal notes and summaries collected as part of the *Cryptography* course offered by the degree in Computer Science of the University of Rome "La Sapienza".

Further information and notes can be found at the following link:
https://github.com/aflaag-notes. Anyone can feel free to report inaccuracies, improvements or requests through the Issue system provided by GitHub itself or by contacting the author privately:

- Email: **alessio.bandiera02@gmail.com**
- LinkedIn: **Alessio Bandiera**

The notes are constantly being updated, so please check if the changes have already been made in the most recent version.

Suggested prerequisites:

TODO

Licence:

1

TODO

TODO

missing introduction

1.1 TODO

In this section, we will discuss **symmetric cryptography**, i.e. cryptosystems where a shared secret key is used for both encryption and decryption. Such encryption method is fast and efficient, but generally considered less secure. The core model of this approach is the **Secret Key Encryption (SKE)**, composed of:

- a shared *secret* key $K \in_R \mathcal{K}$ chosen uniformly at random
- an *encryption function* $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ that transforms plaintext into cyphertext — where $\mathcal{K}$ and $\mathcal{C}$ are the spaces of the keys and the cyphertexts, respectively
- a *decryption function* $\mathrm{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$ that transforms a cyphertext into a plaintext

Clearly, to be useful at all SKEs must be **correct**, which means that if a message $m \in \mathcal{M}$ is encrypted through $K \in_R \mathcal{K}$ getting $c \in \mathcal{C}$, the decription process on c through K must produce the original message m

Definition 1.1: Correctness of SKEs

An SKE $\Pi = (\mathrm{Enc}, \mathrm{Dec})$ is said to be **correct** if it holds that

$$\forall m \in \mathcal{M}, K \in_R \mathcal{K} \quad \mathrm{Dec}(\mathrm{Enc}(K, m)) = m$$

TODO

diagram

During the 19th century, the Dutch cryptographer Kerckhoff postulated his homonym principle, which is stated below.

Principle 1.1: Kerckhoff's principle

The security of a cryptographic system should depend solely on the secrecy of the key.

In other words, a system should be secure *even if everything is public but the key*. During his work, Shannon proposed a formal definition of **perfect secrecy** which fully follows the concept of Kerckhoff's principle.

Definition 1.2: Perfect secrecy

Given an SKE $\Pi = (\mathrm{Enc}, \mathrm{Dec})$, a random variable M over $\mathcal{M}$ and a random variable $C = \mathrm{Enc}(K, M)$ for some $K \in_R \mathcal{K}$, we say that Π has **perfect secrecy** if

$$\forall m \in \mathcal{M}, c \in \mathcal{C} \quad \Pr[M = m] = \Pr[M = m \mid C = c]$$

In other words, as Shannon originally formulated it, this definition states that the probability of m being exactly the communicated message must not depend on the cyphertext c, which implies that c can be known by everyone. Hence, this definition requires the encrypted text c to *not reveal* anything about the plaintext m. The following lemma shows some properties about perfect secrecy.

Shannon proved that perfect secrecy is achievable by some cryptosystems, but with some limitations. For instance, consider the **One Time Pad (OTP)** SKE, in which we assume that $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}^n$ for some fixed length $n \in \mathbb{N}$, and the encryption and decryption functions are defined as follows:

$$\mathrm{Enc}(K, m) = K \oplus m \qquad \mathrm{Dec}(K, c) = K \oplus c$$

(which is always possible by invertibility of the XOR function). We will prove that OTP has perfect secrecy by first providing an alternative definition to the latter.

Lemma 1.1

Let $\Pi = (\mathrm{Enc}, \mathrm{Dec})$ be an SKE, M be a random variable over $\mathcal{M}$ and C be a random variable defined as $C = \mathrm{Enc}(K, M)$ for some $K \in_R \mathcal{K}$. The following three conditions are equivalent:

1. Π has perfect secrecy
2. M and C are independent
3. $\forall m, m' \in \mathcal{M}, c \in C \quad \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] = \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m') = c]$

Proof. We will prove the statements cyclically.

- $1 \implies 2$. By perfect secrecy of Π we have that

$$\Pr[M = m] = \Pr[M = m \mid C = c] = \frac{\Pr[M = m \wedge C = c]}{\Pr[C = c]}$$

therefore, by rearranging the terms we get that

$$\Pr[M = m \wedge C = c] = \Pr[M = m] \cdot \Pr[C = c]$$

- $2 \implies 3$. Fix $m \in \mathcal{M}$ and $c \in \mathcal{C}$; we have that

$$\begin{aligned}\Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] &= \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, M) \mid M = m] \\ &= \Pr_{K \in_R \mathcal{K}}[C = c \mid M = m] && \text{(by definition)} \\ &= \Pr[C = c] && \text{(by independence of } M \text{ and } C\text{)}\end{aligned}$$

  Now fix another message $m' \in \mathcal{M}$; we can repeat the same steps and obtain that $\Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m') = c] = \Pr[C = c]$ which concludes the proof.

- $3 \implies 1$. Fix $c \in \mathcal{C}$.

  Claim: $\Pr[C = c] = \Pr[C = c \mid M = m]$

  Proof of the Claim. By assuming property 3, we get that

$$\begin{aligned}\Pr[C = c] &= \sum_{m' \in \mathcal{M}} \Pr[C = c \mid M = m'] \cdot \Pr[M = m'] && \text{(by the L.T.P.)} \\ &= \sum_{m' \in \mathcal{M}} \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, M) = c \mid M = m'] \cdot \Pr[M = m'] \\ &= \sum_{m' \in \mathcal{M}} \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m') = c] \cdot \Pr[M = m'] \\ &= \sum_{m' \in \mathcal{M}} \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] \cdot \Pr[M = m'] && \text{(by property 3)} \\ &= \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] \cdot \sum_{m' \in \mathcal{M}} \Pr[m = m'] \\ &= \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] \\ &= \Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, M) = c \mid M = m] \\ &= \Pr_{K \in_R \mathcal{K}}[C = c \mid M = m]\end{aligned}$$

  □

  Finally, by Bayes' theorem we have that

$$\begin{aligned}\Pr[M = m] &= \frac{\Pr[M = m \mid C = c] \cdot \Pr[C = c]}{\Pr[C = c \mid M = m]} \\ &= \Pr[M = m \mid C = c] && \text{(by the claim)}\end{aligned}$$

  which is precisely perfect secrecy.

□

Thanks to this lemma, we can prove the following result through the third alternative definition of perfect secrecy.

Theorem 1.1

OTP has perfect secrecy.

Proof. Fix two messages $m, m' \in \mathcal{M}$, and a cyphertext $c \in \mathcal{C}$; by the definition of the OTP system and the properties of the XOR function, we have that

$$\Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m) = c] = \Pr_{K \in_R \mathcal{K}}[K \oplus m = c] = \Pr_{K \in_R \mathcal{K}}[K = m \oplus c] = 2^{-n}$$

and through the same reasoning we can analogously show that $\Pr_{K \in_R \mathcal{K}}[\mathrm{Enc}(K, m') = c] = 2^{-n}$ which proves that the third condition of the alternative definition of perfect secrecy provided in the previous lemma holds for the OTP SKE. □

To conclude this section, we will show **Shannon's theorem about perfect secrecy**, in whic he proved an inherent limitation of perfect secrecy.

Theorem 1.2: Shannon's perfect secrecy theorem

Let $\Pi = (\mathrm{Enc}, \mathrm{Dec})$ be a non-trivial perfectly secret SKE; then, it holds that $|\mathcal{K}| \geq |\mathcal{M}|$.

Proof. Fix any cyphertext $c \in \mathcal{C}$ such that $\Pr[C = c] > 0$ — since Π is non-trivial there will always exist at least once such c. Moreover, let $\mathcal{M}'$ be the set of possible decryptions over c i.e.

$$\mathcal{M}' := \{\mathrm{Dec}(K, c) \mid K \in \mathcal{K}\}$$

We observe that $\mathcal{M}'$ contains *at most* one decryption per key, i.e. $|\mathcal{M}'| \leq |\mathcal{K}|$, since some keys may yield the same decryption.

Now, by way of contradiction, suppose that $|\mathcal{K}| < |\mathcal{M}|$, which sirectly implies that $|\mathcal{M}'| \leq |\mathcal{K}| < |\mathcal{M}| \implies |\mathcal{M}'| < |\mathcal{M}|$, implying that there must be some $m \in \mathcal{M} - \mathcal{M}'$. In particular, such message m cannot be the result of the decryption process applied on c, which means that $\Pr[M = m \mid C = c] = 0$. However, when no additional information is given every message is uniform, i.e. $\Pr[M = m] = \frac{1}{|\mathcal{M}|}$, which implies that

$$\exists m \in \mathcal{M} \quad \frac{1}{|\mathcal{M}|} \neq \Pr[M = m \mid C = c] = 0$$

contradicting the fact that Π had perfect secrecy $\lightning$. □

1.2 Message Authentication Codes (MACs)

We will now focus on the second goal of cryptography: *message integrity*, which is usually achieved through **Message Authentication Codes (MACs)**, which allows the receiver to determine if the message has been tempered with.

First, let's start with a simple model, that ignores secrecy and only cares about message integrity. This type of MACs use a deterministic **tagging function**, usually implemented though *hash functions*, but in general it is a function of the form

$$\mathrm{Tag} : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$$

where $\mathcal{T}$ is the tag space, i.e. the set of all tag strings.

TODO disegno

The idea behind tagging functions is the following:

- Alice send the message-tag pair to Bob
- Bob re-computes the tag — assuming that we do not care about secrecy, hence the key can be shared — through the same key
- if the two tags are equal, Bob can be usre that the message has not been altered

However, this simple idea only works under the assumption of **unforgeability**, which says that

1. it should be *hard* to forge a valid tag τ for a message m when the key K is not known
2. it should be *hard* to forge a valid tag τ for a message m even when a pair (m', τ') is known

In other words, the first condition requires that the tag should not reveal any information about the key that forged it, and the second condition states that no message-tag pair should reveal any information about how the tags are computed. Without this property, an adversary may be able to infer information about the key and/or the tagging function, using them to forge valid tags, which may allow them to fool the receiver. We can formalize unforgeability with the following concept.

Definition 1.3: t-time ε-statistical security

A MAC $\Pi = (\mathrm{Tag})$ is said to have ***t*-time ε-statistical security** if $\forall m, m_1, \ldots, m_t \in \mathcal{M}$ pairwise distinct, and $\forall \tau, \tau_1, \ldots, \tau_t \in \mathcal{T}$ it holds that

$$\Pr_{K \in_R \mathcal{K}}[\mathrm{Tag}(K, m) = \tau \mid \mathrm{Tag}(K, m_1) = \tau_1, \ldots, \mathrm{Tag}(K, m_t) = \tau_t] \leq \varepsilon$$

In other words, this property states that even when t message-tag pairs $(m_1, \tau_1), \ldots, (m_t, \tau_t)$ obtained through the same key K are known, the probability of any message-tag pair (m, τ) being obtained through the same key K is at most ε. Clearly, we would like ε to be as small as possible, and t to be as large as possible. However it is easy to see that $\forall t \in \mathbb{N}$ it is impossible to get $\varepsilon = 0$, since any random $\tau \in \mathcal{T}$ has always a $\frac{1}{|\mathcal{T}|}$ probability of being correct by random chance.

Furthermore, as we proved for perfect secrecy, we are going to show that the notion of "good" statistical security is achievable, even though it's highly inefficient in terms of key size.

Theorem 1.3

Any t-time $2^{-\lambda}$-statistically secure MAC must have a key of size $(t+1)\cdot\lambda$.

A good enough 1-time statisticalllhy secure MAC can be achieved through **pairwise independent hash functions**, a family of hash functions where each pair of functions forms a pari of independent random variables.

Definition 1.4: Pairwise independence

Let $\mathcal{H} = \{h_K : \mathcal{K} \to \mathcal{T}\}_{K \in_R \mathcal{K}}$ be a family of hash functions; we say that $\mathcal{H}$ is **pairwise independent** if $\forall m, m' \in \mathcal{M}$ such that $m \neq m'$ it holds that the distribution $(h_K(m), h_K(m'))$ is uniform over $\mathcal{T} \times \mathcal{T}$ when $K \in_R \mathcal{K}$.

Note that, in this definition, $h_K(m)$ and $h_K(m')$ are two random variables .

what the f does this def mean

Theorem 1.4

Any $\mathcal{H} = \{h_K : \mathcal{M} \to \mathcal{T}\}_{K \in_R \mathcal{K}}$ family of pairwise independent hash functions induces a 1-time $\frac{1}{|\mathcal{T}|}$-statistically secure MAC.

Proof. TODO □

todo

TODO

buco

1.2.1 Randomness extraction

TODO

buco

Let's first address the problem of *extracting* randomness from an unpredictable secure source (i.e. random variable) X. The first **extractor** has been introduced by Von Neumann, which yields a fair random coin from an unpredictable unfair one. Let $B \in \{0,1\}$ be the random variable describing the unpredictable unfrair coin, such that for instance $\Pr[B=0] = p < \frac{1}{2}$. Let $Y \in \{0,1\}$ be the random variable describing the coin we want to extract; its vaule will be determined by the following procedure:

- sample two values b_1, b_2 from B independently
- if $b_1 = b_2$, Y is assigned no value — noted with $Y = ?$ and the process is repeated from the beginning
- if $b_1 \neq b_2$, $Y = 1$ if and only if $b_1 = 0$ and $b_2 = 1$

First, we observe that this process can go on indefinitely, never halting on any value for Y. But assuming the procedure does halt, we have that

$$\Pr[Y=0] = \Pr[b_1 = 1 \wedge b_2 = 0] = \Pr[b_1 = 1]\cdot\Pr[b_2 = 0] = (1-p)p$$

$$\Pr[Y=1] = \Pr[b_1 = 0 \wedge b_2 = 1] = \Pr[b_1 = 0]\cdot\Pr[b_2 = 1] = p(1-p)$$

Moreover, we see that the probability of $Y = ?$ happening for m consecutive tries is at most

$$\Pr[Y = ?\text{for } m \text{ tries}] = (\Pr[Y = ?])^m = (1 - \Pr[Y = 0 \lor Y = 1])^m \leq (1 - 2p(1-p))^m$$

and as m goes to infinity, the latter probabijlity goes to 0, which implies that $\Pr[Y = 0]$ and $\Pr[Y = 1]$ will tend towards $\frac{1}{2}$ due to them having the same probability and them being the only two possible outcomes. Thus, we have achieved a fair enough coin Y.

Can we generalize this concept? That is, is it possible to design an extractor Ext that uses a random variable X to output any desired uniform distribution $\text{Ext}(X)$? Well, if the source X is unpredictable this is clearly not possible, not matter how effort — if the source is truly unpredictable, we would only be able to achieve an unpredictable extractor in any case. We can measure how "good" oure source is based on the following concept.

Definition 1.5: Min-entropy

Given a random variable X, the **min-entropy** of X is defined as follows:

$$H_{\text{min}}(X) := -\log \max_x \Pr[X = x]$$

We observe that min-entropy represents the largest value m such that each observation of X provides *as least* m bits of information. But why is it called "min-entropy"? The standard definition of entropy is the following

$$H(P) := \sum_i p_i \log\left(\frac{1}{p_i}\right)$$

and in the min-entropy, instead, we are taking the minimum value of $\log\left(\frac{1}{p_i}\right)$, in fact

$$\min_i \log\left(\frac{1}{p_i}\right) = \log\left(\min_i \frac{1}{p_i}\right) = -\log \max_i p_i$$

How does the min-entropy assign values? Consider a random variable $X \sim U_n$, where U_n denotes the uniform distribution over $\{0,1\}^n$. Clearly, for any x it holds that $\Pr[X = x] = 2^{-n}$, therefore we get that the min-entropy of X is

$$H_{\text{min}}(X) = -\log \max_x \Pr[X = x] = -\log(2^{-n}) = n$$

This result intuitively makes sense, since we assumed X to be a uniform source, which implies that each observation of X provides at least n bits of information. Instead, if we consider a constant random variable Y, i.e. such that $\Pr[Y = \bar{x}] = 1$ for some value $\bar{x}$ and $\Pr[Y \neq \bar{x}] = 0$, we get that

$$H_{\text{min}}(Y) = -\log \max_x \Pr[Y = x] = -\log 1 = 0$$

which intuitively makes sense too, since Y was assumed to be constant, and indeed each observation of Y provides at least 0 bits of information.

Is there an extractor Ext that for any random variable X it outputs a uniform distribution $Y = \text{Ext}(X)$ such that $H_{\min}(X) \geq k$ for some $k > 0$? Even under this constraint, the answer is still negative, in fact it is impossible to define such extractor even if we restrict the question to extracting only *one bit* — i.e. $\text{Ext}(X) = b \in \{0,1\}$ — and $k = n-1$.

Proposition 1.1

There is no extractor Ext such that for every random variable X over $\{0,1\}^n$ with $H_{\min}(X) \geq n-1$ it holds that $\text{Ext}(X)$ is uniform over $\{0,1\}$.

Proof. Let $\text{Ext} : \{0,1\}^n \to \{0,1\}$ be any extractor, and let $b \in \{0,1\}$ be the output minimizing the cardinality of the preimage of the extractor, i.e. the set of inputs for which the extractor output b

$$b = \arg\min_{b' \in \{0,1\}} \left|\text{Ext}^{-1}(b)\right|$$

By the pigeonhole principle, we have that

$$\left|\text{Ext}^{-1}(b)\right| \leq \frac{|\{0,1\}^n|}{2} = \frac{2^n}{2} = 2^{n-1}$$

Let X be a random variable uniform over $\text{Ext}^{-1}(b)$; since X is uniform, we have that $H_{\min}(X) \geq n-1$, however $\text{Ext}(X)$ cannot be uniform since the output is constant (in fact $\text{Ext}(X) \equiv b$).

This proves that any extractor Ext has always a bad input uniform random variable X that returns a non-uniform distribution $\text{Ext}(X)$, concluding that there is no extractor such that for all X it holds that $\text{Ext}(X)$ is uniform. □

Hence, since we cannot define an extractor that yields a uniform distribution, the best we can achieve is a distribution that is *close enough* to a uniform one. But what does it mean to be “close enough” mathematically? First, let us provide the definition of **statistical distance**.

Definition 1.6: Statistical distance

Given two random variables X and X' defined over the same set, the **statistical distance** between X and X' is defined as follows:

$$\text{SD}(X; X') = \frac{1}{2} \sum_x |\Pr[X = x] - \Pr[X' = x]|$$

We can now introduce the concept of *closeness* between distributions.

Definition 1.7: ε-closeness

Given two random variables X and X' defined over the same set, we say that X and X' are **ε-close**, written as $X \sim_\varepsilon X'$, if it holds that

$$\mathrm{SD}(X; X') \leq \varepsilon$$

We can describe the concept of ε-closeness between random variables as saying that no *unbounded adversary* A can distinguish whether a value x has been sampled from X or X', or in symbols

$$|\Pr[A(x) = 1 : x \leftarrow X] - \Pr[A(x) = 1 : x \leftarrow X']| \leq \varepsilon$$

Definition 1.8: Deterministic extractor

Given a random variable S, called *seed*, we say that $\mathrm{Ext} : \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^\ell$ is a (k, ε)**-extractor** if for every variable X such that $H_{\min}(X) \geq k$ it holds that

$$(S, \mathrm{Ext}(S, X)) \sim_\varepsilon (S, U_\ell)$$

when $S \sim U_d$.

We observe that the condition $(S, \mathrm{Ext}(S, X)) \sim_\varepsilon (S, U_\ell)$ implies that S must be *public* . This requirement is forced in order to avoid trivial extractors, such as $\mathrm{Ext}(S, X) = S$.

why?

non ho capito la fine

The goodness of the hash funcitnos is measured in terms of **collision probability**, which is defined as follows.

Definition 1.9: Collision probability

Given a random variable Y defined over a domain $\mathcal{Y}$, the **collision probability** of Y is defined as follows

$$\mathrm{Col}(Y) = \sum_{y \in \mathcal{Y}} \Pr[Y = y]^2$$

The definition can be explained as follows: the probability collision measures how likely it is that any two possible values of Y *collide*, i.e. given a copy Y' of Y we measure

$$\mathrm{Col}(Y) = \Pr[Y = Y']$$

and since Y and Y' are i.i.d. (identical and independent distribution) we get that

$$\mathrm{Col}(Y) = \Pr[Y = Y'] = \sum_{y \in \mathcal{Y}} \Pr[Y = y \wedge Y' = y] = \sum_{y \in \mathcal{Y}} \Pr[y = y]^2$$

Before proving the leftover hash lemma, consider the following property.

Proposition 1.2

Given Y random variable over some domain $\mathcal{Y}$ such that $\operatorname{Col}(Y) \leq \frac{1}{|\mathcal{Y}|}(1+4\varepsilon^2)$ for some $\varepsilon \in \mathbb{R}_{>0}$, it holds that $\operatorname{SD}(Y;U) \leq \varepsilon$ where U is the uniform distribution over $\mathcal{Y}$.

Proof. By definition, we have that

$$\operatorname{SD}(Y;U) = \frac{1}{2}\sum_{y\in\mathcal{Y}} |\Pr[Y=y]| - \Pr[U=y] = \frac{1}{2}\sum_{y\in\mathcal{Y}} |\Pr[Y=y]| - \frac{1}{|\mathcal{Y}|}$$

Now, for each $y \in \mathcal{Y}$ let

$$q_y := \Pr[Y=y] - \frac{1}{|\mathcal{Y}|} \qquad s_y := \begin{cases} 1 & q_y \geq 0 \\ -1 & q_y < 0 \end{cases} = \operatorname{sign}(q_y)$$

and consider the following two vectors

$$\begin{aligned} \overrightarrow{q} &= \begin{bmatrix} q_{y_1} & \cdots & q_{y_{|\mathcal{Y}|}} \end{bmatrix} \\ \overrightarrow{s} &= \begin{bmatrix} s_{y_1} & \cdots & s_{y_{|\mathcal{Y}|}} \end{bmatrix} \end{aligned}$$

Then, we can rewrite the previous equation as follows

$$\frac{1}{2}\sum_{y\in\mathcal{Y}} \left|\Pr[Y=y] - \frac{1}{|\mathcal{Y}|}\right| = \frac{1}{2}\sum_{y\in\mathcal{Y}} q_y s_y = \frac{1}{2}\langle \overrightarrow{q}, \overrightarrow{s}\rangle$$

By the Cauchy-Schwarz inequality we have that

$$\frac{1}{2}\langle \overrightarrow{q}, \overrightarrow{s}\rangle \leq \frac{1}{2}\sqrt{\langle \overrightarrow{q}, \overrightarrow{q}\rangle, \langle \overrightarrow{s}, \overrightarrow{s}\rangle} = \frac{1}{2}\sqrt{\left(\sum_{y\in\mathcal{Y}} q_y^2\right)|\mathcal{Y}|}$$

where $\langle \overrightarrow{s}, \overrightarrow{s}\rangle = |\mathcal{Y}|$ since every product of $\langle \overrightarrow{s}, \overrightarrow{s}\rangle$ will be either $1 \cdot 1 = 1$ or $-1 \cdot (-1) = 1$, which implies that $\langle \overrightarrow{s}, \overrightarrow{s}\rangle = \sum_{y\in\mathcal{Y}} 1 = |\mathcal{Y}|$.

Claim: $\displaystyle\sum_{y\in\mathcal{Y}} q_y^2 \leq \frac{4\varepsilon^2}{|\mathcal{Y}|}$

Proof of the Claim. We observe that

$$\begin{aligned} \sum_{y\in\mathcal{Y}} q_y^2 &= \sum_{y\in\mathcal{Y}} \left(\Pr[Y=y] - \frac{1}{|\mathcal{Y}|}\right)^2 \\ &= \sum_{y\in\mathcal{Y}} \left(\Pr[Y=y]^2 - \frac{2}{|\mathcal{Y}|}\Pr[Y=y] + \frac{1}{|\mathcal{Y}|^2}\right) \\ &= \sum_{y\in\mathcal{Y}} \Pr[Y=y] - \frac{2}{|\mathcal{Y}|}\sum_{y\in\mathcal{Y}} \Pr[Y=y] + \sum_{y\in\mathcal{Y}} \frac{1}{|\mathcal{Y}|^2} \\ &= \sum_{y\in\mathcal{Y}} \Pr[Y=y]^2 - \frac{2}{|\mathcal{Y}|} + \frac{1}{|\mathcal{Y}|} \\ &= \operatorname{Col}(Y) - \frac{1}{|\mathcal{Y}|} \end{aligned}$$

Finally, recall that $\mathrm{Col}(Y) \leq \frac{1}{|\mathcal{Y}|}(1+4\varepsilon^2)$ by hypothesis, which implies that

$$\sum_{y \in \mathcal{Y}} q_y^2 = \mathrm{Col}(Y) - \frac{1}{|\mathcal{Y}|} \leq \frac{1}{|\mathcal{Y}|}(1+4\varepsilon^2) - \frac{1}{|\mathcal{Y}|} = \frac{4\varepsilon^2}{|\mathcal{Y}|}$$

□

Lastly, by the claim we conclude that

$$\mathrm{SD}(Y;U) = \frac{1}{2}\sqrt{\left(\sum_{y \in \mathcal{Y}} q_y^2\right)|\mathcal{Y}|} \leq \frac{1}{2}\sqrt{\frac{4\varepsilon^2}{|\mathcal{Y}|}|\mathcal{Y}|} = \varepsilon$$

□

Lemma 1.2: Leftover hash lemma

Given a pairwise independent function $\mathcal{H} = \{h_S : \{0,1\}^n \to \{0,1\}^\ell\}_{S \in \{0,1\}^d}$, and a random variable X, for each $S \in \{0,1\}^d$ it holds that $\mathrm{Ext}(S,X) = h_S(X)$ is a (k,ε)-extractor for $k \geq \ell + 2\log\left(\frac{1}{\varepsilon}\right) - 2$.

Proof. Fix $S \in \{0,1\}^d$ and two random variables X, X'; moreover, let $Y = (S, \mathrm{Ext}(S,X)) = (S, h_S(X))$ and let Y' be a copy of Y defined as $Y' = (S', \mathrm{Ext}(S',X')) = (S, h_{S'}(X'))$. Note that

$$\begin{aligned}
\mathrm{Col}(Y) &= \Pr[Y = Y'] \\
&= \Pr[S = S' \wedge \mathrm{Ext}(S,X) = \mathrm{Ext}(S',X')] \\
&= \Pr[S = S' \wedge h_S(X) = h_{S'}(X')]
\end{aligned}$$

Now, since S and S' are independent from X and X', we get that

$$\begin{aligned}
\mathrm{Col}(Y) &= \Pr[S = S' \wedge h_S(X) = h_{S'}(X')] \\
&= \Pr[S = S' \wedge h_S(X) = h_S(X')] \\
&= \Pr[S = S'] \cdot \Pr[h_S(X) = h_S(X')] \\
&= \frac{1}{|\{0,1\}^d|} \cdot \Pr[h_S(X) = h_S(X')] \\
&= 2^{-d} \cdot (\Pr[X = X', h_S(X) = h_S(X')] + \Pr[X \neq X', h_S(X) = h_S(X')]) \\
&= 2^{-d} \cdot (\Pr[X = X'] + \Pr[X \neq X', h_S(X) = h_S(X')]) \\
&\leq 2^{-d} \cdot (TODO + 1^{-\ell}) \\
&= \frac{1}{2^{d+\ell}}\left(2^{\ell-k} + 1\right)
\end{aligned}$$

TODO Finally, we have that

come continua poi il calcolo sopra?

$$\begin{aligned}
\mathrm{Col}(Y) &\leq \frac{1}{2^{d+\ell}}\left(2^{\ell-k} + 1\right) \\
&\leq \frac{1}{2^{d+\ell}}\left(2^{2-2\log\left(\frac{1}{\varepsilon}\right)}\right) && \text{(by hypothesis)} \\
&= \frac{4\varepsilon^2 + 1}{|\mathcal{Y}|} && \text{(through algebraic manipulation)}
\end{aligned}$$

and thanks to the previous proposition we know that $\mathrm{Col}(Y) \leq \frac{1}{|\mathcal{Y}|}(4\varepsilon^2 + 1)$ implies that $\mathrm{SD}(Y;U) \leq \varepsilon \iff Y \sim_\varepsilon U \iff (S, \mathrm{Ext}(S,X)) \sim_\varepsilon (S, U_\ell)$ concluding that Ext is indeed a (k,ε)-extractor. □

1.3 Exercises

Problem 1.1

Given a prime $p \in \mathbb{P}$, let $\mathcal{M} = \mathcal{T} = \mathbb{Z}_p$ and $\mathcal{K} = \mathbb{Z}_p^2$. Prove that the hash family $\mathcal{H} = \{h_{(a,b)}\}_{(a,b)\in\mathbb{Z}_p^2}$, where $h_{(a,b)}(m) \equiv am + b \pmod{p}$, cannot be a 2-time statistically secure MAC.

Solution. TODO □ todo

Problem 1.2

Construct a 3-wise independent hash function family and prove its correctness.

Solution. TODO □ todo

2
TODO

TODO impagliazzo