Geometrical Cryptography

Extended abstract

L. Csirmaz* G. O. H. Katona†

Abstract

Methods using physical objects as cryptographic devices is a subject of several ongoing projects, [2, 4]; results were reported even in the Science Magazine [5]. We investigate the simplest case, when the physical object merely supplies a cryptographic key which should be extracted from a single (or a set of) measurements. The "randomness" of the physical object is guaranteed by physical laws. The object behaves as a real key, but duplication and forgery is prevented by the actual physical properties of the object. Subsequent measurements of the same object yield different results; the extracted cryptographic key, however, must remain the same with high probability. Whenever the key is needed, it is extracted freshly from the physical object.

The total data given by a measurement is grouped into a single point of an appropriately chosen *phase space*. Physical objects are identified with the result of an ideal measurement. Real measurements yield other points of the phase space "sufficiently close" to the ideal one. Cryptographic keys are assigned to objects; the key should be recovered unambiguously from any measurement of the object. For each possible cryptographic key we consider the subset of the phase space whose points correspond to physical objects with that key. The collection of these sets form a *geometric code*.

We identify three important parameters. The *advantage* is an upper bound on the expected number of random objects one has to generate until a "valid" one is found, i. e. the associated point yields a cryptographic key. The *error tolerance* tells us how punctuate we must be in the measurements so that the key could be recovered. Finally, the *security* bounds the probability of each particular key. Under quite natural assumptions on the phase space, we prove a necessary condition these parameters must satisfy; and give examples which show that our condition is tight up to constants. We also take a look at our assumptions on the phase space.

1 Introduction

Geometrical codes arise naturally when cryptographic keys are generated by observing an appropriately generated physical object. Such methods were investigated intensively, see, e. g., [2, 4, 5]. We consider the scenario when these objects, also referred to as cryptographic tokens, are used for identification, or to guarantee originality or uniqueness. Cryptographic

*CEU, Budapest

†Rényi Institute, Budapest

tokens can be attached to DVD's, whose contents are then encoded by keys extracted from the tokens. As token duplication, by assumption, is prohibitively expensive, this would prevent DVD duplications too. Originality of a paper based document can be proved off-line when a token is attached to the paper, and the issuing authority digitally signs the content of the document together with the key extracted from the token. Attaching tokens to plastic cards can prevent unauthorized duplications when the digital signature of the extracted key and relevant card information is also stored on the card.

A cryptographic token can be any physical object which can be measured repeatedly, and it can be an intrinsic part of the whole object or may be externally introduced. There are several patents for cryptographical tokens. The system in [6] uses magnetic fibers randomly sprinkled and embedded in a thin substrate. To read the identity token, a magnetic read head is passed along the substrate. [7] uses variable translucency when a sheet of paper is illuminated with a light source, and the data is read by an optical reader. Another patent [8] uses small conducting particles embedded in an insulating substrate and uses microwaves to read information. For a throughout overview of patents, see [4], where a token made of micron-scale glass spheres cured into optical-grade epoxy is investigated. For paper-based documents the token can be the collection of special fibers embedded into the paper substance; the measurement is scanning the paper under ultraviolet light.

The measurement results in a collection of digitized bulk data, coming from a scanner, from one, or several sampling devices. The data goes through filtering, smoothing, thresholding, maybe further sophisticated evaluation techniques are used as well. The preprocessing stage reduces the amount of data significantly, but – supposedly – retains all important physical parameters of the measured object. Finally, based on the preprocessed data, a cryptographic key, or extract, is generated.

In this paper we concentrate on a single aspect of this process, namely how the measurement error affects the extracted key. In this respect the preprocessing stage can be considered to be part of the measurement, or can be thought as part of the key-extracting procedure. The physical object is identified by the result of an ideally performed measurement. All other measurements yield results which differ from the ideal one, but are "sufficiently close," whatever this expression means. As we expect, the same cryptographic key should be generated from different measurements of the same object, thus sufficiently close measurements must yield the same extract.

In one of the possible scenarios the result of a measurement is a digitized, gray-scale picture consisting of, say, 1 million pixels. Each pixel has intensity which is a real number between 0 (black) and 1 (white); the whole picture thus can be considered as a single vector (point) in the 1 million dimensional Euclidean space. Two pictures are "close enough," if the average pixelwise discrepancy is smaller than, say, 0.01. In other words, the distance between two points (pictures) is measured in L_1-norm, and we tolerate error up to distance 0.01 times 1 million = 10,000. Instead of working with the pixels, we can transform the picture into the frequency domain. Similarity of pictures are often measured by how close they are in the frequency domain. Small scale noise or large scale systematic error can be filtered out quite easily simply weighing high and low frequencies with small coefficients. Using discrete frequency values, the data can also be represented as a sequence of real numbers – that is, as a point in a high dimensional Euclidean space; and closeness is measured again by a certain norm in that space.

In our model we assume that such a measurement yields a single point in a certain *phase space* T. "Closeness" is measured by some distance function, thus, in particular, T is a metric space. Physical objects (tokens) are identified by the result of the ideal measurement, that is, by a point in T. As tokens are random objects, we also have a probability distribution μ on the phase space, where for a subset A of T, $\mu(A)$ tells us the probability that a randomly produced object lies in A.

Given the result of a measurement, that is, a point p in T, some procedure, whose exact details are irrelevant at the moment, tells us what the extracted cryptographic key is. For each possible key k we denote by A_k the set of objects (tokens) to which the key k has been assigned. When the object $p \in T$ is measured, the result is some $p' \in T$ sufficiently close to p, that is the distance between p and p' is bounded by some positive ε. The key extracted from p' should be the same as the one extracted from p.

How cryptographic keys are assigned to measurements is determined uniquely by the collection of the sets A_k as k runs over all possible keys. Properties of this system have an intricate connection to the geometrical properties of the sets A_k, thus we called this collection a *geometrical code.*

Geometrical codes have three important parameters: *advantage*, *security*, and *error tolerance.* The *advantage* α is an upper bound on how many random tokens should be generated on the average until a "valid" one is found, i. e. until the associated point in the phase space T is in one of the sets A_k. If the advantage is large then, with high probability, many tokens should be discarded until we get a usable one. If the advantage is near to 1 then almost all tokens are good. It would be ideal to have advantage exactly one, however, under our assumptions it is impossible to achieve. The *error tolerance* ε tells us that the error we make in a measurement is at most ε. That is, if p' is the result of measuring a token $p \in A_k$, then the distance between p and p' is less than ε. As the extract k should also be recovered from p', the ε-neighborhoods of the A_k's must be disjoint. Finally, the *security parameter* σ bounds the probability of the sets A_k. This parameter ensures that a particular key can be generated, or replicated, with small enough probability only.

In Section 2 we give formal definitions, and state our main result about parameters of geometrical codes. In Section 3 we give examples, showing that our bound is tight up to a constant in several important cases. We take a look on our assumptions on the phase space; finally the last section contains conclusions.

2 Definition

The phase space M is a collection of point which has a metric and a measure as well. The distance between points x and y of M will be denoted by $d(x, y)$. The (open) *ball of radius r around* $p \in M$ is the set of all points whose distance from p is less than ϱ. We denote this ball by $p + \varrho$ as follows:

$$p + \varrho \stackrel{\text{def}}{=} \{x \in M \,:\, d(x,p) < \varrho\}.$$

Given any subset A of the phase space M, the (open) *ϱ-neighborhood of* A is the union of the balls $p + \varrho$ where p runs over the elements of A:

$$A + \varrho \stackrel{\text{def}}{=} \bigcup\{p + \varrho \,:\, p \in A\}.$$

We say that $A + \varrho$ is the result of *fattening* A by ϱ. From the triangle inequality it follows immediately that fattening $A + \varrho_1$ by ϱ_2, the result is included in $A + (\varrho_1 + \varrho_2)$. We call the metric *flat* whenever this inclusion is never proper for positive ϱ_1 and ϱ_2, that is if for all A and positive ϱ_1, ϱ_2 we have

$$(A+\varrho_1)+\varrho_2 = A+(\varrho_1+\varrho_2). \tag{1}$$

Apart from being a metric space, a measure μ is also given on M so that all balls are measurable. We assume also that μ is *homogeneous* in the sense that all balls of equal radius have the same measure. The *volume* (i. .e. the measure) of a ball of radius ϱ is denoted by $V(\varrho)$. We assume also that this V is continuous, strictly increasing, and takes all positive real numbers. Given a measurable set $A \subseteq M$, let $r(A)$ be the radius of the ball which has the same measure as A, namely

$$\mu(A) = V(r(A)).$$

We fix furthermore a measurable subset T of the phase space M with $\mu(T) = 1$ whose elements correspond to the possible outcomes of the measurements. For $A \subseteq T$, we interpret the measure $\mu(A)$ as the probability that the outcome is an element of A.

As an example, M can be the n-dimensional Euclidean space $\mathbb{R}^n$ with the usual (Euclidean) distance, T be the n-dimensional unit cube, and μ be the Lebesgue measure. This corresponds to the case when points of the unit cube (the phase space) are generated uniformly.

Definition 1 A *geometrical code* of size m is a collection $A_1, \ldots, A_m$ of disjoint measurable subsets of T. The code has

> *security* σ, if $\mu(A_k) \le \sigma$ for all k;
>
> *advantage* α, if $\mu(\bigcup A_k) \ge 1/\alpha$; and
>
> *error tolerance* ε if the ε-neighborhood of A_k is in T, and $(A_i+\varepsilon)\cap(A_j+\varepsilon) = \emptyset$ whenever $i \ne j$.

The *security* parameter tells us that no particular key can be obtained with probability exceeding σ. Typical values are 2^{-50} or smaller. The *advantage* says how many tokens we must produce on the average until we get a usable one; typically this value should be in the range between 1.5 and 100. The *error* indicates how punctuate we should be in measuring: making an error up to ε will not yield unrelated keys. Our goal is to get estimates on ε given security and advantage.

As keys are derived from the sets A_k, we might require all keys to have the same probability, i. e. all A_k should have the same measure, or at least they should not differ too much. Also, the sets A_k must be "simple" enough so that deriving the key k should not pose computational problems.

Definition 2 The phase space M has the *Brunn-Minkowski property*, if for all measurable sets A we have

$$\mu(A+\varepsilon) \ge \mu((p+r(A))+\varepsilon).$$

In other words, fattening any measurable set it grows at least as much as it would if it were a ball. If the metric is flat, then the inequality can be written in the more succinct form of

$$r(A+\varepsilon) \ge r(A)+\varepsilon.$$

From this form it is easy to see that rescaling the metric does not affect whether the space has the Brunn-Minkowski property.

The celebrated *Brunn-Minkowski inequality* says [1] that the n-dimensional Euclidean space with the usual metric and Lebesgue measure has the Brunn-Minkowski property. There are several generalizations of this theorem, we shall briefly recall some of them later.

We have all definitions at our disposal to state and proof our theorem, which, of course, is not in the most general form.

Theorem 1 *Suppose the phase space M is flat, has the Brunn-Minkowski property, and the function $V(\varrho)$, giving the volume of the ball of radius ϱ, is log-concave. Then any geometrical code with advantage α, security σ and error tolerance ε satisfies the inequality*

$$\varepsilon \leq V^{-1}(\alpha\sigma) - V^{-1}(\sigma). \tag{1}$$

Proof Suppose the geometrical code consists of the measurable sets $A_1, \ldots, A_m$. Let $r(A_k) = a_k$, that is a_k is the radius of the ball which has the same measure than the set A_k. As M is flat, we can apply the Brunn-Minkowski property in the second form yielding $r(A_k + \varepsilon) \geq r(A_k) + \varepsilon = a_k + \varepsilon$. In other words, $A_k + \varepsilon$ has measure at least as large as the ball with radius $a_k + \varepsilon$:

$$\mu(A_k + \varepsilon) \geq V(a_k + \varepsilon). \tag{2}$$

Let moreover a be the radius of the ball which has the measure σ, i. e. $a = V^{-1}(\sigma)$. As the code has security σ, $\mu(A_k) \leq \sigma$ for all key k, which gives $a_k \leq a$ for all k.

Both $a_k + \varepsilon$ and a are inside the interval $(a_k, a + \varepsilon)$, and if $a_k + \varepsilon$ divides this interval in the ratio λ to $1 - \lambda$, then a divides it in the ratio $1 - \lambda$ to λ. By assumption the function $\log V$ is concave, thus

$$\begin{aligned} \log V(a_k + \varepsilon) &\geq (1-\lambda)\log V(a_k) + \lambda \log V(a + \varepsilon), \\ \log V(a) &\geq \lambda \log V(a_k) + (1-\lambda)\log V(a + \varepsilon). \end{aligned}$$

Adding them up and rearranging we get

$$V(a_k + \varepsilon) \geq \frac{V(a + \varepsilon)}{V(a)} \cdot V(a_k).$$

By (2) the left hand side is at most $\mu(A_k + \varepsilon)$, and the sets $A_k + \varepsilon$ are pairwise disjoint subsets of T, consequently their sum cannot exceed the measure of T:

$$1 = \mu(T) \geq \sum_k \frac{V(a+\varepsilon)}{V(a)} \cdot V(a_k) = \frac{V(a+\varepsilon)}{V(a)} \sum_k V(a_k) \geq \frac{V(a+\varepsilon)}{V(a)} \cdot \frac{1}{\alpha}$$

since $V(a_k)$ is the measure of A_k, thus $\sum_k V(a_k) = \mu(\bigcup A_k) \geq 1/\alpha$. Noting that $V(a) = \sigma$, we get (1). ■

3 Examples

In our first example M is the n-dimensional Euclidean space $\mathbb{R}^n$ with the usual distance and measure. The domain T of the codes can either be the unit cube, or the n-dimensional ball of volume 1. The volume of the n-dimensional ball of radius ϱ is

$$V(\varrho) = \gamma_n \varrho^n \qquad \text{where} \qquad \gamma_n = \frac{\pi^{n/2}}{(n/2)!}.$$

Clearly $V(\varrho)$ is log-concave, and $\mathbb{R}^n$ has the Brunn-Minkowski property (see [1]), thus we can use Theorem 1. Estimating $k!$ by $(k/e)^k$, (1) becomes

$$\varepsilon \leq \sqrt{\frac{2\pi e}{n}}\, \sigma^{1/n} \left(\alpha^{1/n} - 1\right). \tag{3}$$

If the dimension n is at least ten times as large as $\log \alpha$ then the last factor can be replaced by $(\log \alpha)/n$. Keeping α and σ fixed, the right hand side takes its maximal value at $n = 0.66 \log(1/\sigma)$ independently of α. Then (3) simplifies to

$$\varepsilon \leq \frac{2}{3} e^{2/3} \sqrt{2\pi e}\, \frac{\log \alpha}{\log(1/\sigma)} \approx 5.37\, \frac{\log \alpha}{\log(1/\sigma)}. \tag{4}$$

For $\sigma = 2^{-50}$ and $\alpha = 1.5$ this gives the bound $\varepsilon \leq 0.01$ with n around 23.

We can also construct geometric codes with parameters quite close to the bound in (3). Cut out small cubes of edge length $\sigma^{1/n} + 2\varepsilon$ from T, and then shrink each small cube from its center until its edge becomes $\sigma^{1/n}$ (and then its volume will be σ). These sets will form a geometric code with security σ and error ε. The advantage depends on how many small cubes we can cut out of T. For simplifying the calculations, suppose that T is the unit cube. As the edge length of the small cubes is $\sigma^{1/n} + 2\varepsilon$, one can cut out

$$\left[\frac{1}{\sigma^{1/n} + 2\varepsilon}\right]^n \geq \left(\frac{1}{\sigma^{1/n} + 2\varepsilon} - 1\right)^n,$$

small cubes out of T, where $[x]$ is the integer part of x. Thus the code has advantage α whenever

$$\left(\frac{1}{\sigma^{1/n} + 2\varepsilon} - 1\right)^n \sigma \geq \frac{1}{\alpha},$$

which holds if

$$\varepsilon \leq 0.5\, \sigma^{1/n} \left(\alpha^{1/n} \frac{1}{1 + (\alpha\sigma)^{1/n}} - 1\right). \tag{5}$$

If $(\alpha\sigma)^{1/n}$ is small enough, then this is within a factor of $10\sqrt{n}$ of the theoretical.

In general, if ∂T denotes the measure of the boundary of T, then we can cut out of T at least

$$\frac{1 - \eta \sqrt{n}\, \partial T}{\eta^n}$$

small cubes with edge length η. If we assume $x = \eta \sqrt{n}\, \partial T$ to be smaller than one half, then we can use $1 - x/n$ to approximate $(1-x)^{1/n}$, which shows that the code will have advantage α whenever

$$\varepsilon \leq 0.5\, \sigma^{1/n} \left(\alpha^{1/n} \frac{1}{1 + \dfrac{\partial T}{\sqrt{n}} (\alpha\sigma)^{1/n}} - 1\right).$$

The constant times $(\alpha\sigma)^{1/n}$ error term is the *boundary error*, which comes from the fact that parts close to the boundary of T are wasted. This is proportional to the surface of T and diminishes as σ tends to zero. The discrepancy between the main constants 0.5 here and $\sqrt{2\pi e/n}$ in (3) comes from the *packing error*, namely that n-dimensional balls cannot be packed tightly.

Keeping the Lebesgue measure, we can use other metric in $\mathbb{R}^n$. One general type of metric comes from a norm $||x||$. The distance of the points (vectors) x and y is the norm of their difference:

$$d(x,y) \stackrel{\text{def}}{=} ||x-y||.$$

This metric is always flat, and "balls" are convex sets. Consequently these spaces also have the Brunn-Minkowski property ([1]). The volume of the ball of radius ϱ is $V(\varrho) = c\cdot\varrho^n$, where c is the volume of the unit ball $B_1 = \{x \in \mathbb{R}^n : ||x|| < 1\}$. As the function $V(\varrho)$ is log-concave, we can apply Theorem 1 which gives the theoretical bound

$$\varepsilon \le \frac{1}{c^{1/n}}\,\sigma^{1/n}\left(\alpha^{1/n}-1\right). \tag{6}$$

We also do have constructions in this general case, too. In fact, we have a geometric code inside T consisting of congruent parallelepipeds whenever

$$\varepsilon \le \frac{1}{(n!)^{1/n}c^{1/n}}\,\sigma^{1/n}\left(\alpha^{1/n}\frac{1}{1+\dfrac{\partial\varphi T}{\sqrt{n}}(\alpha\sigma)^{1/n}}-1\right).$$

Apart from the boundary error the two bounds agree up to the constant $(n!)^{1/n} \approx n/e$.

We have seen examples where the phase space was the n-dimensional Euclidean space and the metric was generated from an arbitrary norm. These phase spaces share several natural properties; we note particularly the following ones:

(i) the metric is translation invariant, that is the distance between a and b is the same as the distance between $a+x$ and $b+x$;

(ii) the metric is flat;

(iii) it generates the standard Euclidean topology;

(iv) the measure is homogeneous.

Now we show that our estimate in Theorem 1 applies for *all* phase spaces defined on the n-dimensional vectors, requiring only these properties.

Theorem 2 *Suppose the phase space M is defined on the points of the n-dimensional Euclidean space, and it satisfies properties* (i)–(iv) *above. Then any geometrical code in this space satisfies the inequality of Theorem 1.*

We have an example that condition (ii) is essential for Theorem 2 to hold. In that example the space is the two-dimension Euclidean space with Lebesgue measure, and a strange distance function. As Theorem 1 fails, this space cannot have the Brunn-Minkowski property either.

4 Conclusion

We have defined an abstract mathematical model of a practical problem, namely extracting reliably a cryptographic key from a physical object. The model accounted for three important parameters: *security*, *advantage*, and *error tolerance*. We have established a basic inequality between these values, based on geometrical properties of the underlying phase space.

We investigated in detail cases when this phase space is the Euclidean n-dimensional space with the usual Lebesgue measure, but not necessarily the standard metric. We have shown constructions which achieved the theoretical bounds within certain error terms. Two types of error were identified. The *boundary error* comes from the fact that the code cannot use points near to the boundary of the phase space T. This error becomes negligible when the size of individual codes tends to zero. The *packing error* measures how compactly can the space be packed with "balls." This error term is unaffected by the choice of the parameters.

Constructions also had further nice properties. All code sets were of the same measure, and they were cubes or parallelepipeds, thus had trivial structure and admit effective encoding.

We took a closer look at our assumptions on the phase space. One of them is a generalization of the celebrated Brunn-Minkowski theorem; another one is a strange restriction on the metric, namely that the metric space must be flat. We proved that flatness has interesting consequences: the volume of the ball with radius r can grow at most exponentially; furthermore on phase spaces defined on n-dimensional vectors flatness (together with natural homogeneity assumptions) ensures our estimate to hold. We also gave an example for an otherwise nice phase space where this flatness falls short. Constructions in the hyperbolic space close to the bound given by Theorem 1 would also be of interest, as well as investigating other esoteric geometries.

Problems discussed here originated from investigating off-line authentication using cryptographic tokens. Tokens are produced by a random process, they are cheap, but duplicating and copying is prohibitively expensive. The token is measured, a key is generated, and the key plus the relevant information is digitally signed. When checking authenticity, the token is measured again, the extracted key is regenerated, and the digital signature is checked. After measuring the token, signal processing techniques are used to reduce the data into several hundred (or maybe a few thousand) real numbers. Subsequent measures are expected to yield small variation in these numbers, thus "closeness" might be defined by some distance function. Being this the case, our results give a theoretical bound on how punctuate we must be on the measurements (error tolerance) given the amount of independent bits we want to extract from the tokens (security) and the percentage of the waste in generating tokens (advantage).

References

[1] Yu. D. Burago, V. A. Zalgaller: *Geometric Inequalities*, Springer, 1988

[2] B. Gassend, D. Clarke, M. van Dijk, S. Devadas, *Controlled Physical Unknown Functions*, MIT LCS TR-845, 2002

[3] A. J. Menezes, P. C. van Oorshot, S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press 1996

[4] Ravi Pappu, *Physical one-Way Functions*, PhD Thesis, March 2001

[5] Ravi Pappu, Ben Recht, Jason Taylor, Neil Gershenfeld, *Physical One-Way Functions*, Science 2002 September 20, 297, pp 2026–2030

[6] J. Brosow: Method and system for verifying authenticity safe against forgery, US patent no 4218674, 1980

[7] R. Goldman: Verification system for document substance and content, US patent no 4568936, 1986

[8] J. Samyn, Method and apparatus for checking the authenticity of documents, US patent no 4820912, 1989