Recognition of Bounded Convex Bodies.

Baptiste Huguet
Student of École Normale Superieure de RENNES
second year

SUPERVISOR:
Victor K. Ohanyan
Head of Chair of "Theory of Probability and Mathematical statistics"

Department of Mathematics and Mechanics.
YEREVAN STATE UNIVERSITY

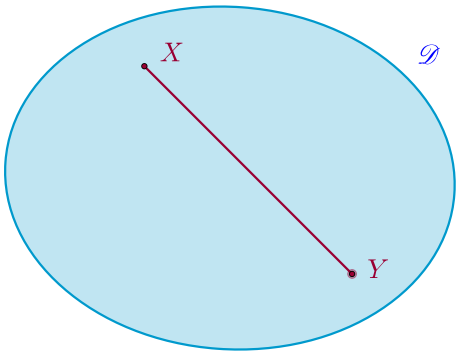
YEREVAN
June-July 2015

Contents

Introduction.

Many branches of sciences need to study and recognize geometrical object. In medical imaging, for example, you have to figure out the shape of an organ or a tumour (a three-dimensional object) with image of sections (two-dimensional view). The geometrical objects studied, can have various shapes and are often non-regular. There are many method to study them. For example a triangle can be uniquely recognize by the coordinate of its vertices. The first problem is that this method can be applied only on polygons. The second problem is more important : in this method we do not considerer the set independently from its place in the space. Yet, two triangles which differ from a translation have the same geometrical properties. We want a method who describe these properties.

This report deals with recognition of sets with an approach of stochastic geometry. The aim of our study is to discover some tools of stochastic geometry and conjectures linked to these tools. We restrain our study on compact convex sets in the Euclidean space $\mathbb{R}^n$. A set $\mathscr{D}$ is said convex when for all X and Y in $\mathscr{D}$, the segment $[X, Y] = \{tX + (1-t)Y/t \in [0,1]\}$ is included in $\mathscr{D}$.



In the first part, we study a problem of probability theory not linked to stochastic geometry : the distribution of the maximum of random variables. In the second section, we deal with the covariogram of a set and the Mathéron 's conjecture about recognition. In the third section we study the chord-length distribution of a set and show how it can characterize some class of sets. In the fourth section, we see the distance between two random points uniformly and independently distributed in a set. These three last notions are linked but we don't study the links between them.

1 Maximum of independent exponential random variables.

By $\mathbb{N}^*$ we denote the set of naturals, without 0. An interval of integers is denoted by double-brackets: $[\![\ldots]\!]$.

Let $\eta_1, \ldots, \eta_n$ be n independent random variables with exponential distribution $\mathcal{E}(\lambda_i)$ for i in $[\![1, n]\!]$. We want to study

$$\xi_n = \max_{1 \le i \le n} \eta_i \quad .$$

1.1 Distribution and Expectation.

The first characterization of a random variable is its distribution function. Let $F_n(t)$ be the distribution function of ξ_n.

Proposition 1 (Explicit form of the distribution function of ξ_n)

$$F_n(t) = \prod_{i=1}^{n} (1 - e^{-\lambda_i t}) \mathbb{1}_{t \ge 0} \,, \quad t \in \mathbb{R},\ n \in \mathbb{N}^*$$

$\mathbb{1}_A(t)$ denote the indicator function of a set A, that is $\mathbb{1}_A(t)$ equals to 1 if $t \in A$ and 0 if $t \notin A$.

Proof : For all $n \in \mathbb{N}^*$ and $t \in \mathbb{R}$, we have

$$\begin{aligned}
F_n(t) &= \mathbb{P}(\xi_n \le t) \\
&= \mathbb{P}(\eta_1 \le t \ \cap \ldots \cap \ \eta_n \le t) \\
&= \prod_{i=1}^{n} \mathbb{P}(\eta_i \le t) \quad \textit{(by independence)} \\
&= \prod_{i=1}^{n} (1 - e^{-\lambda_i t}) \mathbb{1}_{t \ge 0} \\
&= \begin{cases} 0 & \text{if} \quad t \le 0 \\ \prod_{i=1}^{n} (1 - e^{-\lambda_i t}) & \text{if} \quad t \ge 0 \end{cases}
\end{aligned}$$

□

The product can be written in the following form:

$$\prod_{i=1}^{n} (1 - e^{-\lambda_i t}) = \sum_{k=0}^{n} (-1)^k \sum_{E \subset [\![1,n]\!], |E| = k} e^{-t \sum_{l \in E} \lambda_l}$$

$$\prod_{i=1}^{n}(1-e^{-\lambda_i t}) = 1 + \sum_{k=1}^{n}(-1)^k \sum_{E\subset[\![1,n]\!],|E|=k} e^{-t\sum_{l\in E}\lambda_l}\ , \quad n\in\mathbb{N}^*,\ t\geq 0 \quad . \qquad (\star)$$

The sum $\sum_{E\subset[\![1,n]\!],|E|=k}$ is a sum for all subsets of $[\![1,n]\!]$ with k elements. It means

$$\prod_{i=1}^{n}(1-e^{-\lambda_i t}) = 1 - \sum_{i=1}^{n} e^{-t\lambda_i} + \sum_{1\leq i<j\leq n} e^{-t(\lambda_i+\lambda_j)} - \ldots + (-1)^n e^{-t\sum_{i=1}^{n}\lambda_i}.$$

Now, we are able to calculate the mathematical expectation of ξ_n.

Proposition 2 (Explicit form of $\mathbb{E}[\xi_n]$)

$$\mathbb{E}[\xi_n] = \sum_{k=1}^{n}(-1)^{k+1}\sum_{E\subset[\![1,n]\!],|E|=k}\frac{1}{\sum_{l\in E}\lambda_l}\ , \quad n\in\mathbb{N}^*$$

Proof : For all $n\in\mathbb{N}^*$, we have

$$\begin{aligned}
\mathbb{E}[\xi_n] &= \int_0^{+\infty}(1-F(t))\mathrm{d}t \\
&= \int_0^{+\infty}(1-\prod_{i=1}^{n}(1-e^{-\lambda_i t}))\mathrm{d}t \\
&= -\int_0^{+\infty}\sum_{k=1}^{n}(-1)^k\sum_{E\subset[\![1,n]\!],|E|=k} e^{-t\sum_{l\in E}\lambda_l}\mathrm{d}t \\
&= \sum_{k=1}^{n}(-1)^{k+1}\sum_{E\subset[\![1,n]\!],|E|=k}\int_0^{+\infty} e^{-t\sum_{l\in E}\lambda_l}\mathrm{d}t \\
&= \sum_{k=1}^{n}(-1)^{k+1}\sum_{E\subset[\![1,n]\!],|E|=k}\frac{1}{\sum_{l\in E}\lambda_l}
\end{aligned}$$

$\square$

In another way, we have

$$\mathbb{E}[\xi_n] = \sum_{i=1}^{n}\frac{1}{\lambda_i} - \sum_{1\leq i<j\leq n}\frac{1}{\lambda_i+\lambda_j} + \sum_{1\leq i<j<k\leq n}\frac{1}{\lambda_i+\lambda_j+\lambda_k} - \ldots + (-1)^{n-1}\frac{1}{\sum_{i=1}^{n}\lambda_i} \quad .$$

1.2 The particular case of *i.i.d* random variables.

Now, we assume that the random variables $(\eta_i)_{i\in\mathbb{N}^*}$ are *i.i.d* with distribution $\mathcal{E}(\lambda)$. In this particular case, the mathematical expectation can be greatly simplified. For each k, there are $\binom{n}{k}$ subsets of $[\![1,n]\!]$ with k elements. So we have

$$\mathbb{E}[\xi_n] = \frac{1}{\lambda}\sum_{k=1}^{n}\binom{n}{k}(-1)^{k+1}\frac{1}{k}$$

This sum is the partial sum of the harmonic series.

Proposition 3 (Explicit form of $\mathbb{E}[\xi_n]$ for *i.i.d* random variables)

$$\mathbb{E}[\xi_n] = \frac{1}{\lambda} \sum_{i=1}^{n} \frac{1}{i} \quad , \forall n \in \mathbb{N}^*$$

We prove the proposition 3 by two ways : one, classical, with mathematical induction and one, more astute, with an auxiliary function.

1.2.1 Proof by mathematical induction.

Let $\mathcal{P}_n$ the predicate define for $n \in \mathbb{N}^*$:

$$\mathcal{P}_n \ : \mathbb{E}[\xi_n] = \frac{1}{\lambda} H_n$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i} \quad .$

Basis step. For $n = 1$, the result is obvious.

Inductive step. We assume that $\mathcal{P}_n$ is true and we want to prove $\mathcal{P}_{n+1}$. We have

$$\begin{aligned}
\mathbb{E}[\xi_{n+1}] &= \frac{-1}{\lambda} \sum_{k=1}^{n+1} \binom{n+1}{k} (-1)^k \frac{1}{k} \\
&= \frac{-1}{\lambda} \sum_{k=1}^{n+1} \left[\binom{n}{k} + \binom{n}{k-1} \right] (-1)^k \frac{1}{k} \\
&= \frac{-1}{\lambda} \sum_{k=1}^{n+1} \binom{n}{k} (-1)^k \frac{1}{k} + \frac{-1}{\lambda} \sum_{k=1}^{n+1} \binom{n}{k-1} (-1)^k \frac{1}{k} \\
&= \mathbb{E}[\xi_n] + \frac{1}{\lambda} \sum_{k=0}^{n} \binom{n}{k} (-1)^k \frac{1}{k+1}
\end{aligned}$$

By hypothesis, $\mathbb{E}[\xi_n] = \frac{1}{\lambda} H_n$. Since $\binom{n}{k} \frac{1}{k+1} = \binom{n+1}{k+1} \frac{1}{n+1}$, we have

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n}{k} (-1)^k \frac{1}{k+1} &= \sum_{k=0}^{n} \binom{n+1}{k+1} (-1)^k \frac{1}{n+1} \\
&= \frac{-1}{n+1} \sum_{i=1}^{n+1} \binom{n+1}{i} (-1)^i \\
&= \frac{-1}{n+1} [(1-1)^{n+1} - 1] \\
&= \frac{1}{n+1}
\end{aligned}$$

Hence $\mathbb{E}[\xi_{n+1}] = \frac{1}{\lambda}\sum_{k=1}^{n+1}\frac{1}{k}$ and $\mathcal{P}_{n+1}$ is true.

1.2.2 The second proof.

The aim of this proof is to find an auxiliary function whom derivative or primitive integral is linked to the sum we want to calculate.

We introduce $g_n(x) = \frac{(1-x)^n - 1}{x}$ and $G_n(x) = \int_0^x g_n(t)\mathrm{d}t$. The domain of g_n can be extended by continuity in zero and we have

$$g_n(x) = \sum_{k=1}^{n}\binom{n}{k}(-1)^k x^{k-1} \quad \text{and} \quad G_n(x) = \sum_{k=1}^{n}\binom{n}{k}(-1)^k\frac{x^k}{k} \quad .$$

With those notations, we have

$$\mathbb{E}[\xi_n] = \frac{-1}{\lambda}G_n(1).$$

As we know how to factorize $a^n - b^n$, we have

$$\begin{aligned} g_n(x) &= \frac{(1-x)-1}{x}\sum_{k=0}^{n-1}(1-x)^k \\ &= -\sum_{k=0}^{n-1}(1-x)^k \end{aligned}$$

and

$$\int_0^1 (1-t)^k\mathrm{d}t = \frac{1}{k+1} \quad .$$

So we have

$$\begin{aligned} G(1) &= -\sum_{k=0}^{n-1}\frac{1}{k+1} \\ &= -\sum_{k=1}^{n}\frac{1}{k} \end{aligned}$$

This is exactly what we wan to prove.

$$\mathbb{E}[\xi_n] = \frac{1}{\lambda}\sum_{i=1}^{n}\frac{1}{i}\,, \quad n \in \mathbb{N}^*$$

1.3 The Laplace transform of the maximum.

An other interesting characterization of ξ_n is its Laplace transform.

Definition 1 For a non negative random variable η, we define its *Laplace transform* by

$$\begin{aligned} \mathcal{L}_\eta : \mathbb{R}_+ &\to \mathbb{R}_+ \\ s &\mapsto \mathbb{E}[e^{-s\eta}] \end{aligned}$$

If η has a density function f_η, we have

$$\mathcal{L}_\eta(s) = \int_0^{+\infty} e^{-st} f_\eta(t)\mathrm{d}t\,, \quad s \in \mathbb{R}_+$$

In the case of ξ_n, we have the following result :

Proposition 4 (Explicit form of $\mathcal{L}_{\xi_n}$)

$$\mathcal{L}_{\xi_n}(s) = 1 + s\sum_{k=1}^{n}(-1)^{k+1}\sum_{E\subset[\![1,n]\!],|E|=k}\frac{1}{s+\sum_{l\in E}\lambda_l}\,, \quad s\in\mathbb{R}_+,\ n\in\mathbb{N}^*$$

Proof : For all $n \in \mathbb{N}^*$ and $s \in \mathbb{R}_+$, we have

$$\begin{aligned} \mathcal{L}_{\xi_n}(s) &= \int_0^{+\infty} e^{-ts} f_{\xi_n}(t)\mathrm{d}t \\ &= [e^{-ts}(F_{\xi_n}(t)-1)]_0^{+\infty} + s\int_0^{+\infty} e^{-ts}(1-F_{\xi_n}(t))\mathrm{d}t \end{aligned}$$

Using ($\star$), we obtain

$$\begin{aligned} \mathcal{L}_{\xi_n}(s) &= 1 - s\int_0^{+\infty}\left(e^{-ts}\sum_{k=1}^{n}(-1)^k\sum_{E\subset[\![1,n]\!],|E|=k} e^{-t\sum_{l\in E}\lambda_l}\right)\mathrm{d}t \\ &= 1 + s\sum_{k=1}^{n}(-1)^{k+1}\sum_{E\subset[\![1,n]\!],|E|=k}\frac{1}{s+\sum_{l\in E}\lambda_l} \end{aligned}$$

$\square$

In another way, we have

$$\mathcal{L}_{\xi_n}(s) = 1 + s\left(\sum_{i=1}^{n}\frac{1}{s+\lambda_i} - \sum_{1\le i<j\le n}\frac{1}{s+\lambda_i+\lambda_j} + \ldots + (-1)^{n-1}\frac{1}{s+\sum_{i=1}^{n}\lambda_i}\right)$$

In the particular case of *i.i.d* random variables, we have

$$\mathcal{L}_{\xi_n}(s) = 1 - s\sum_{k=1}^{n}\binom{n}{k}(-1)^k\frac{1}{s+k\lambda}\,, \quad s\in\mathbb{R}_+,\ n\in\mathbb{N}^*.$$

2 Covariogram.

2.1 Definition.

Let V_n be the n-dimensional Lebesgue measure in $\mathbb{R}^n$. A body is a compact convex set in $\mathbb{R}^n$ with inner points.

Definition 2 The *covariogram* of a body $\mathscr{D}$ is the function :

$$\begin{array}{rl} C_{\mathscr{D}} : & \mathbb{R}^n \longrightarrow \mathbb{R}_+ \\ & X \quad \mapsto V_n(\mathscr{D} \cap (\mathscr{D} + X)) \end{array} \quad ,$$

where $(\mathscr{D} + X)$ is the translation of $\mathscr{D}$ by vector X.

The covariogram of a body $\mathscr{D}$ is invariant by translations and reflections. G. Mathéron conjectured in 1986 that the covariogram of a convex body determines this body among all convex bodies (up to translations and reflections). It had been proved that for $n \geq 4$ this conjecture is false and for $n = 2$ it is true. For $n = 3$, it is still an open problem.

The covariogram could be defined for other set than bodies, but as all sets without inner points have a covariogram identically zero, we should restrain our study to body so as to avoid non-interesting counter-examples of Mathéron conjecture.

This is a first elementary result about covariogram.

Proposition 5
For any body $\mathscr{D}$, $C_{\mathscr{D}}(0)$ is equal to $V_n(\mathscr{D})$, the volume of $\mathscr{D}$.

Each vector X in $\mathbb{R}^n$ can be described by its polar coordinates $X = (r, \omega)$ where $r \geq 0$ and $\omega \in S^{n-1}$, the n-dimensional unit sphere with centre the origin, seen a the space of direction in $\mathbb{R}^n$. So we use to study the covariogram as a function of two variables : $C_{\mathscr{D}}(r, \omega)$.

2.2 Calculation of the covariogram for a planar disc.

A disc is one of the most simple body to calculate its covariogram because of its symmetry. Obviously, the covariogram of a disc does not depend on the coordinate $\varphi \in S^1$, where (r, φ) are the planar polar coordinates of the vector $X \in \mathbb{R}^2$. That is why we forget the variable φ for discs.

Theorem 6 (Explicit form for the covariogram of a disc.)
The covariogram of the disc $\mathscr{D}(R)$, with radius R is

$$C_{\mathscr{D}}(r) = \begin{cases} 2R^2 \arccos(\frac{r}{2R}) - \frac{r}{2}\sqrt{4R^2 - r^2} & \text{if} \quad 0 \leq r \leq 2R \\ 0 & \text{if} \quad\quad r \geq 2R \end{cases}$$

Proof : Let $\mathscr{D}$ be the disc whom centre is O, the origin, and with radius R. Obviously, for any $r \geq 2R$, $C_{\mathscr{D}}(r) = 0$. We can assume that $0 \leq r \leq 2R$. $C_{\mathscr{D}}(r)$ is the area of the intersection, so it is two times the area of the doted domain in Figure 1.

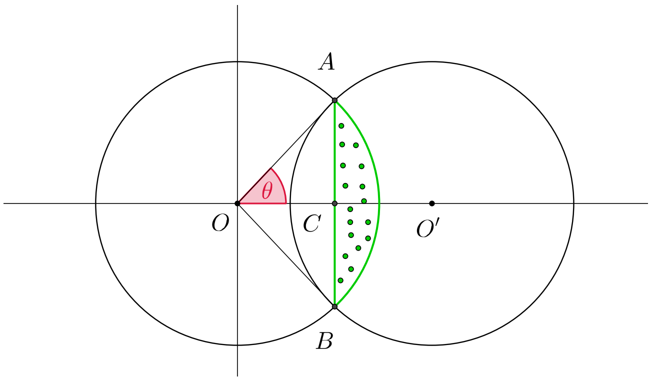


Figure 1: Intersection of two discs.

The triangle OAO' is isosceles, so

$$OC = CO' = \frac{r}{2}.$$

The triangle OAC is right in C, so

$$\theta = \arccos(\frac{r}{2R}).$$

The area of the circular sector of centre O delimited by A and B is

$$\mathcal{A}_1 = \theta R^2.$$

The area of the triangle OAB is

$$\mathcal{A}_2 = 2\,\frac{1}{2}\,\frac{r}{2} R \sin\theta.$$

Hence the doted area is

$$\mathcal{A} = \theta R^2 - \frac{rR}{2} \sin\theta.$$

□

The symmetries of discs make also easier to study the Mathéron conjecture for the class of discs.

Proposition 7
Covariogram characterizes discs in $\mathbb{R}^2$ amongst of all bodies in the plane.

Proof : Let $\mathscr{D}'$ be a body with the same covariogram as the disc $\mathscr{D}$ of centre O and radius R. We want to prove that $\mathscr{D}'$ and $\mathscr{D}$ are equal, up to translations.

Let X and Y be two points distinct in $\mathring{\mathscr{D}}'$. There is an $\varepsilon > 0$ so that $\mathcal{B}(X,\varepsilon) \subset \mathscr{D}'$ and $\mathcal{B}(Y,\varepsilon) \subset \mathscr{D}'$. Let $\vec{u}_x$ and $\vec{u}_y$ the unit direction vectors of the Cartesian frame.

$$X = Y + (X - Y) \quad \text{so} \quad X \in \mathscr{D}' \cap (\mathscr{D}' + (X - Y))$$

$$X + \varepsilon\vec{u}_x = Y\varepsilon\vec{u}_x + (X - Y) \quad \text{so} \quad X + \varepsilon\vec{u}_x \in \mathscr{D}' \cap (\mathscr{D}' + (X - Y))$$
$$X\varepsilon\vec{u}_y = Y\varepsilon\vec{u}_y + (X - Y) \quad \text{so} \quad X + \varepsilon\vec{u}_y \in \mathscr{D}' \cap (\mathscr{D}' + (X - Y))$$

As $\mathscr{D}' \cap (\mathscr{D}' + (X - Y))$ is convex and contains a triangle, its area is not zero. So $C_{\mathscr{D}'}(X - Y) \neq 0$. Hence $|X - Y| \leq 2R$. $\mathscr{D}'$ is contains in a disc of radius R. Up to a translations, we can assume that $\mathscr{D}'$ is contains in the disc $\mathscr{D}$.

Let X be a point of $\mathring{\mathscr{D}} \setminus \mathscr{D}'$. As $\mathring{\mathscr{D}} \setminus \mathscr{D}'$ is an open set, there is an $\varepsilon > 0$ so that $\mathcal{B}(X,\varepsilon) \subset (\mathring{\mathscr{D}} \setminus \mathscr{D}')$. Using the convexity of $\mathscr{D}'$, no point in the area in red can be in $\mathscr{D}'$ (*Cf* Figure 2).

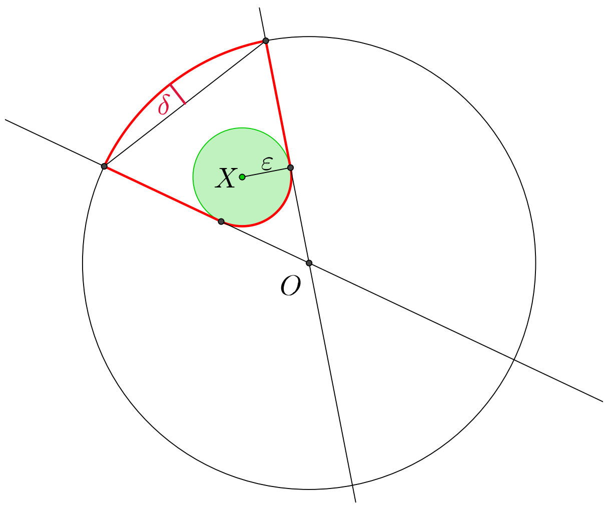


Figure 2:

Hence $\mathscr{D}' \cap (\mathscr{D}' + (2R - \delta)\frac{X}{\|X\|}) = \emptyset$ and this is in contradiction with the covariogram of $\mathscr{D}'$. So every point of $\mathscr{D}$ is in $\mathscr{D}'$. □

In Figure 3 , we can see the graph of the covariogram for a disc with radius $R = 1$.

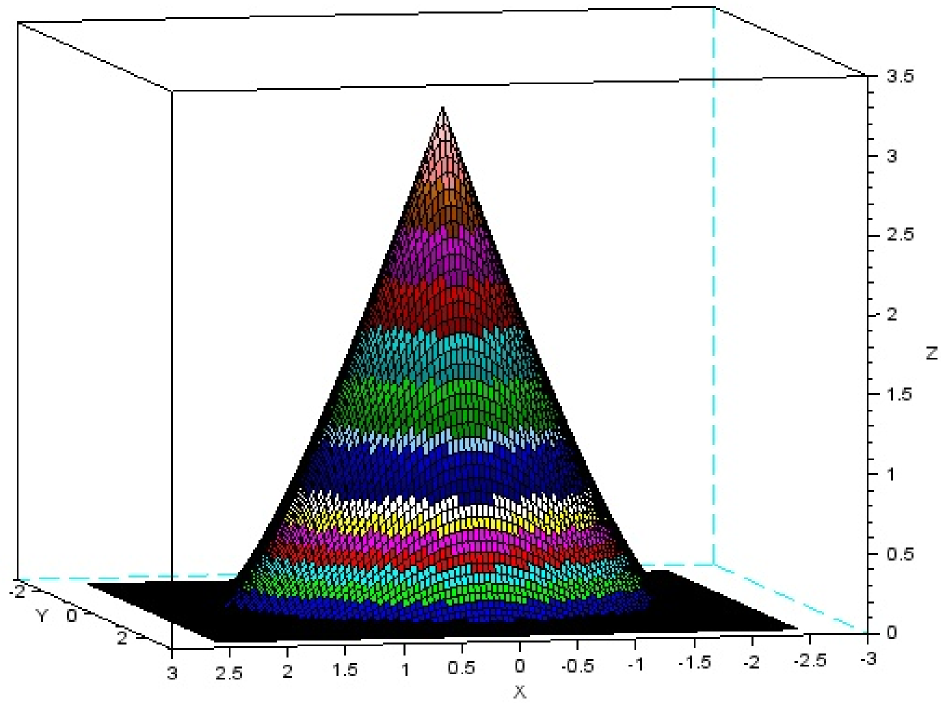


Figure 3: Covariogram for a disc with radius $R = 1$.

2.3 Links with probability.

The definition of the covariogram is purely geometric. Yet, it is linked to stochastic problems. Let $\mathscr{D}$ be a body in $\mathbb{R}^n$. The covariogram of $\mathscr{D}$ can be expressed in terms of integral. For all X in $\mathbb{R}^n$, we have :

$$\begin{aligned} C_{\mathscr{D}}(X) &= \int_{\mathbb{R}^n} \mathbb{1}_{\mathscr{D}}(Y)\mathbb{1}_{\mathscr{D}+X}(Y)\mathrm{d}Y \\ &= \int_{\mathbb{R}^n} \mathbb{1}_{\mathscr{D}}(Y)\mathbb{1}_{\mathscr{D}}(Y-X)\mathrm{d}Y \end{aligned}$$

Now, let η_1 and η_2 be two independent random variables uniformly distributed in $\mathscr{D}$. They have the same density function

$$f_\eta = \frac{1}{V_n(\mathscr{D})}\mathbb{1}_{\mathscr{D}} \quad .$$

Using the two following propositions 8 and 9, we can make a link between $C_{\mathscr{D}}$ and the random variable $\eta_1 - \eta_2$.

Proposition 8
Let η be a random variable with a density function $f_\eta(t)$. Then $-\eta$ is also a random variable with a density function $f_{-\eta}(t)$ and we have the relation

$$f_{-\eta}(t) = f_\eta(-t)\,, \quad t \in \mathbb{R} \quad .$$

Proof : For all t in $\mathbb{R}$, we have :

$$\begin{aligned} F_{-\eta}(t) &= \mathbb{P}(-\eta \leq t) \\ &= \mathbb{P}(\eta > -t) \\ &= 1 - F_\eta(-t-0) \end{aligned}$$

where $F_\eta(-t-0) = \lim\limits_{\substack{x \to -t \\ x < -t}} F_\eta(x)$.

When the density exists, the distribution function is continuous so $F_\eta(-t-0) = F_\eta(-t)$. The density function is the derivative of the distribution function. Hence, we have the result. □

Proposition 9 (Sum of random variables)
If η_1 and η_2 are two independent random variable with density, then $\eta_1 + \eta_2$ is a random variable with density and its density is the convolution of η_1 and η_2 's densities.

$$f_{\eta_1+\eta_2} = f_{\eta_1} * f_{\eta_2} \quad .$$

Proposition 10
Let $\mathscr{D}$ be a body, η_1 and η_2 two independent random variables uniformly distributed in $\mathscr{D}$. Then the density of the random vector $\eta_1 - \eta_2$ is

$$f_{\eta_1-\eta_2}(X) = \frac{1}{V_n^2(\mathscr{D})} C_{\mathscr{D}}(X)\,, \quad \forall X \in \mathbb{R}^n \quad .$$

The calculation of a covariogram is equivalent to the calculation a density of a random variable. We can reformulate the Mathéron conjecture in terms of probability : *"does the density function of the random vector between two random points dropped uniformly in a body $\mathscr{D}$, characterize its body amongst all bodies, up to translations and reflections ?"*

3 Chord-length distribution.

3.1 Invariant measures.

Let $(X, \mathcal{A})$ be a measurable space. A measure μ, on $(X, \mathcal{A})$ is a function

$$\mu : \mathcal{A} \rightarrow [0, +\infty]$$

so that : $\mu(\emptyset) = 0$ and which is σ-additive *i.e* $\mu(\bigcup_{i=1}^{+\infty} A_i) = \sum_{i=1}^{+\infty} \mu(A_i)$ for all $(A_i)_{i \in \mathbb{N}^*}$ pairwise disjoint sets in $\mathcal{A}$.

The condition $\mu(\emptyset) = 0$ avoid the measure identically equal to $+\infty$. For any measure, if it exists a measurable set so that $\mu(A) < +\infty$, the σ-additivity induce $\mu(\emptyset) = 0$.

For a measurable set as $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, there are many measures but only few have the properties that make them interesting in geometrical problems.

Definition 3 (Invariance)
Let H be a group which acts on X to the left. μ is said *left-invariant* with respect to H, if we have

$$\mu(h \cdot A) = \mu(A)\,, \quad \forall h \in H,\, \forall A \in \mathcal{A}\,.$$

A similar definition could be given for right-invariance. If μ is both left-invariant and right-invariant, we say that μ is invariant.

Definition 4 (Locally finite)
μ, a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, is *locally finite* if for any compact K, $\mu(K) < +\infty$.

Let $\mathbb{T}_n$ be the group of translations of $\mathbb{R}^n$. As $\mathbb{T}_n$ can be identified to $\mathbb{R}^n$, the action of $t \in \mathbb{T}_n$ on $x \in \mathbb{R}^n$ can be noted $t + x$. The following theorem is a very strong result about the measures on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Theorem 11
Let μ be a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, locally finite and invariant with respect to $\mathbb{T}_n$.

$$\text{Then,} \quad \exists c \geq 0,\; \mu = c V_n\,.$$

In a certain way, it is a theorem of uniqueness. For a given scale, there is one and only one locally finite invariant measure. The scale is given by $\mu([0,1]^n) = c$. It is the same matter as conversion from international system of measure to imperial system.

3.2 Lines in the plane.

Let $\mathbb{G}$ be the set of all lines of $\mathbb{R}^2$ (*Gerade* means *line* in German). For more results about $\mathbb{G}$, see [8]. There are many way to parametrize a line in the plane. We use polar coordinates. Each $g \in \mathbb{G}$ can be described by its distance from the origin and the angle that the perpendicular direction to g makes with a fixed direction (*see* Figure 4) :

$$g = (p, \varphi) \ , \quad \text{with } p \in \mathbb{R}_+ \ , \text{ and } \varphi \in S^1 \ .$$

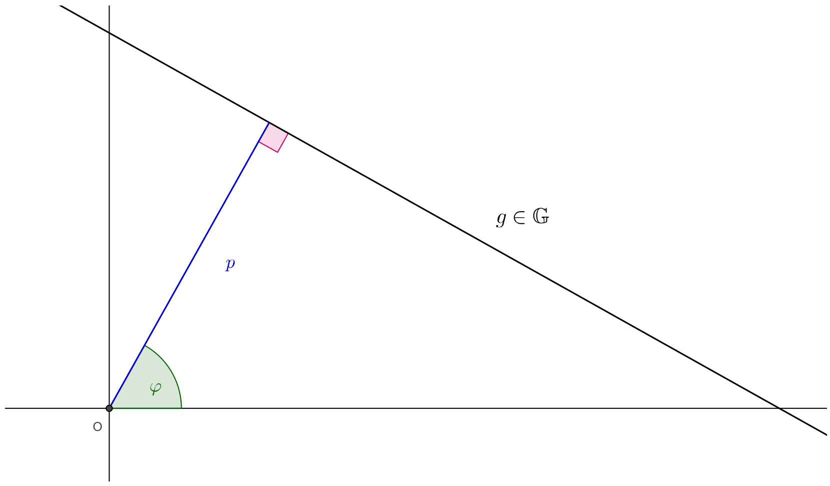


Figure 4:

We can note that topologically speaking, $\mathbb{G}$ is not equivalent to the cylinder $S^1 \times \mathbb{R}_+$ because the lines $(0,0)$ and $(0,\pi)$ are the same. $\mathbb{G}$ is a topological space homeomorphic to a Möbius strip. We note $\mathcal{B}(\mathbb{G})$ its Borel σ-algebra. The group $\mathbb{T}_2$ acts on $\mathbb{G}$ but not as translations.

Proposition 12 (Action of $\mathbb{T}_2$ on $\mathbb{G}$)
For all $t \in \mathbb{T}_2$ and $g = (p, \varphi) \in \mathbb{G}$, we have :

$$t.g = (p + \|t\| \cos(\widehat{t, \varphi}), \varphi) \quad .$$

where $\widehat{(t, \varphi)}$ is the angle between the directions of t and the perpendicular direction to g.

We have a similar result as with $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, but less strong.

Theorem 13
Let μ be a measure on $(\mathbb{G}, \mathcal{B}(\mathbb{G}))$, locally finite and invariant with respect to $\mathbb{T}_2$. Then

$$\mathrm{d}\mu = \mathrm{d}p\, m(\mathrm{d}\varphi)$$

where $\mathrm{d}p$ is the 1-dimensional Lebesgue measure and m is a measure on S^1.

There is a an infinity of measures on S^1, all the δ_{φ_0} (Dirac measure on $\varphi_0 \in S^1$) for example. If we want an uniqueness result, we need to ask a more constraigning invariance. Let $\mathfrak{M}_2$ be the group of rigid motions in the plane. It is the group of translations and rotations.

Theorem 14
Let μ be a measure on $(\mathbb{G}, \mathcal{B}(\mathbb{G}))$, locally finite and invariant with respect to $\mathfrak{M}_2$. Then

$$\mathrm{d}\mu = c\, \mathrm{d}p\, \mathrm{d}\varphi$$

where $\mathrm{d}\varphi$ is the uniform measure on S^1 and $c \geq 0$.

From now, we note $\mathrm{d}g = \mathrm{d}p\, \mathrm{d}\varphi$. It is called the density for a set of lines. A line in $\mathbb{G}$ can be determined by others coordinates than p and φ. In those cases, the form of $\mathrm{d}g$ can be obtained by changes of coordinates.

Let Ox and Oy two axes inclined with angle C ($C \neq 0 \mod \pi$). The set of lines parallel to (Ox) or (Oy) is a negligible set :

$$\mu(\{g \in \mathbb{G} / g \mathbin{/\!/} (Ox) \quad \text{or} \quad g \mathbin{/\!/} (Oy)\}) = 0 \quad .$$

Then we can study the parametrization with the coordinates (x, y) of intersections between $g \in \mathbb{G}$ and the axes and the corresponding expression of the density $\mathrm{d}g$.

Proposition 15
If lines $g \in \mathbb{G}$ (non-parallel to any axes) are parametrized by their intersections with the axes, then the density of lines is :

$$\mathrm{d}g = \left| \frac{xy \sin^2(C)}{\sigma^3} \right| \mathrm{d}x \mathrm{d}y \quad ,$$

where σ is the length of the chord between the two axes.

Proof : We introduce $f : (x, y) \mapsto (p, \varphi)$, and Jf its Jacobian. A result of integral calculus give us :

$$\mathrm{d}p\, \mathrm{d}\varphi = |\det(Jf)| \mathrm{d}x\, \mathrm{d}y \quad .$$

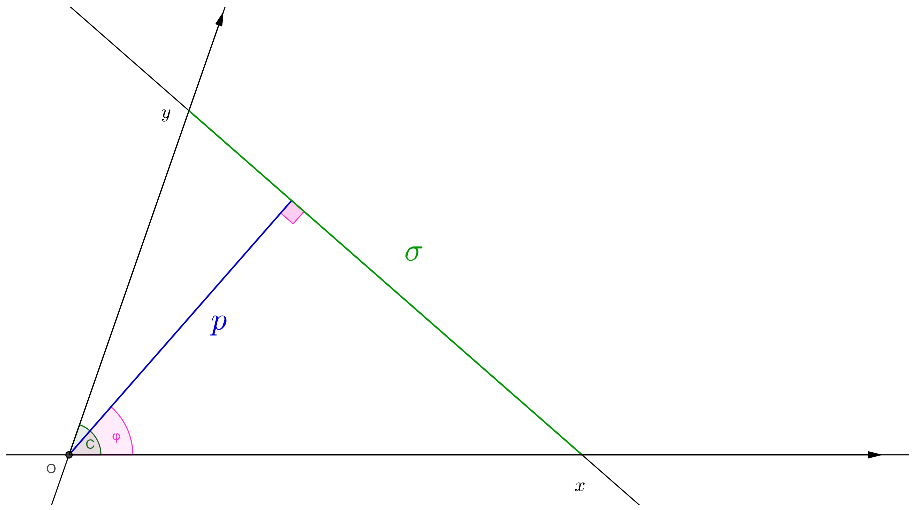


Figure 5: Change of coordinates.

We have the following relations :

$$p = x\cos(\varphi) = y\cos(C-\varphi) \quad \text{and} \quad \sigma = \sqrt{x^2+y^2-2xy\cos(C)} \quad .$$

The area of the big triangle is :

$$\frac{1}{2}xy\sin(C) = \frac{1}{2}\sigma p \quad .$$

Hence, we have

$$p = \frac{xy\sin(C)}{\sigma} \quad \text{and} \quad \varphi = \arccos(\frac{p}{x}) \quad .$$

$$Jf(x,y) = \begin{pmatrix} \frac{\partial p}{\partial x} & \frac{\partial \varphi}{\partial x} \\ \frac{\partial p}{\partial y} & \frac{\partial \varphi}{\partial y} \end{pmatrix}$$

We have :

$$\frac{\partial p}{\partial x} = y\sin(C)\frac{y^2-xy\cos(C)}{\sigma^3}$$

$$\frac{\partial p}{\partial y} = x\sin(C)\frac{x^2-xy\cos(C)}{\sigma^3}$$

$$\frac{\partial \varphi}{\partial x} = \frac{y\sin(C)}{\sigma^3}\frac{-1}{\sqrt{x^2-p^2}}$$

$$\frac{\partial \varphi}{\partial y} = \frac{-1}{\sqrt{x^2-p^2}}\frac{\partial p}{\partial y}$$

$$\det(Jf) = \frac{\partial p}{\partial x}\frac{\partial \varphi}{\partial y} - \frac{\partial p}{\partial y}\frac{\partial \varphi}{\partial x}$$

After a little calculus, we find the result announced. □

3.3 Intersection of a line with a body in the plane.

Let $\mathscr{D}$ be a body in the plane. We note $[\mathscr{D}]$ the subset of $\mathbb{G}$ of lines which intersect $\mathscr{D}$

$$[\mathscr{D}] = \{g \in \mathbb{G} / g \cap \mathscr{D} \neq 0\} \quad .$$

A classical result (*see* [1]) state that the measure of $[\mathscr{D}]$ is equal to the length of $\partial\mathscr{D}$, the boundary of $\mathscr{D}$:

$$L = \int_{[\mathscr{D}]} \mathrm{d}g \quad .$$

So as to understand this result, let's play a little game. K_1 and K_2 are two convex bonded bodies so that K_1 contains K_2. You don't know where is K_2 and you have to draw a chord of K_1 which cut K_2. What is the probability of your success?

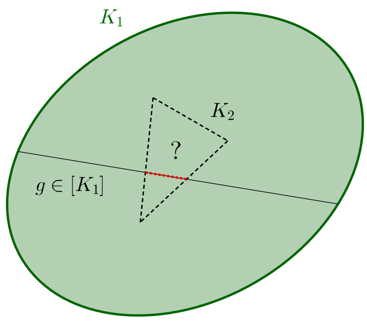


Figure 6: Intersection of K_1's chord with K_2.

According to the previous result, the probability that a random chord of K_1 intersects K_2, is equal to $\dfrac{L_2}{L_1}$, where L_1 and L_2 are the perimeters of K_1 and K_2 respectively. It is interesting to see that the probability is independent of the bodies' areas.

A line $g \in [\mathscr{D}]$ produce a chord $\chi(g)$ of length $|\chi(g)|$.

Definition 5 (Chord-length distribution.)
The *chord-length distribution* of a convex body $\mathscr{D}$ is the function :

$$\begin{aligned} F_{\mathscr{D}} : \mathbb{R} &\to [0,1] \\ x &\mapsto \frac{1}{L} \int_{\{g \in [\mathscr{D}] / |\chi(g)| \leq x\}} \mathrm{d}g \end{aligned}$$

This function is continuous (proved in 1965 by Sulanke) and there is no counter-example of non absolutely continuous chord-length distributions. So we usually consider that a density function exists.

The explicit form of chord-length distribution is known for only few class of convex bodies : ellipses, triangles ([5]), parallelograms ([6]) and regular polygons ([7]).

Mallows and Clark have proved in 1970 that the chord-length distribution doesn't characterize a convex body (up to rigid motion) by exhibiting two non-congruent convex polygons with the same chord-length distribution. Yet, for some class of polygon, there is a one-to-one correspondence between the polygon and its chord-length distribution.

It is obvious that there is less informations in the chord-length distribution, which is a function of $\mathbb{R}$, than in the covariogram, which is a function of $\mathbb{R}^2$. That is why we cannot recognize all bodies by their chord-length distribution. If we want more informations, we should study oriented chord-length distribution. The idea is to study the chord-length for lines with a fixed direction $\varphi \in S^1$, for example, for lines parallel to (Ox) only. The chord-length distribution is a kind of mean value of oriented chord-length distributions, over all the directions. A mean value is always with less informations. The link between oriented chord-length distributions and covariogram are discussed in [4].

3.4 Recognition of a triangle by its chord-length distribution.

A triangle can be recognize among the class of triangles, by studying the moments of its chord-length distribution (*see* [2]). For all convex bodies, the first and the third moments are well known. If σ denotes the length of a random chord to the body, we have :

$$\mathbb{E}[\sigma] = \frac{\pi F}{L} \quad \text{and} \quad \mathbb{E}[\sigma^3] = \frac{3F^2}{L} \quad ,$$

where L is the perimeter and F the area of the body. If we dropped a sample of lines $(g_1, \ldots, g_n)$ in a body, we obtain a sample of chord-length $(\sigma_1, \ldots, \sigma_n)$. Using the *strong law of large numbers*, we can approximate the first moments :

$$\mathbb{E}[\sigma] \approx \frac{1}{n}\sum_{i=1}^{n} \sigma_i \quad \text{and} \quad \mathbb{E}[\sigma^3] \approx \frac{1}{n}\sum_{i=1}^{n} \sigma_i^3 \quad .$$

Then it is possible to approximate the perimeter and the area. Unfortunately, a body is not characterize by its perimeter and area only. For example, the rhombus of diagonals' lengths $D = 2$ and $d = 1$, and the rectangle with sides's lengths $L = \dfrac{\sqrt{5}+1}{2}$ and $\ell = \dfrac{\sqrt{5}-1}{2}$ have the same perimeters and the same areas (*see* Figure 7).

A triangle is characterized by its sides's lengths only. We need another data than perimeter and area for deducing the three lengths. We could use the fact that the longest chord is a side, and so it is the supremum of the chord-length density support. We develop here an other way, based an the fifth moment.

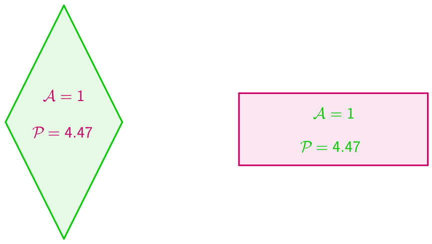


Figure 7: Two different polygons with the same perimeters and areas.

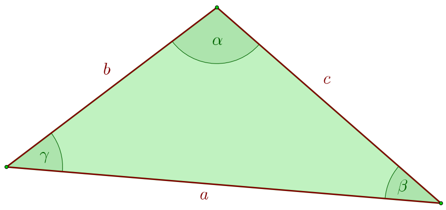


Figure 8: Triangle Δ.

For a triangle, the fifth moment is easy to calculate. Let Δ be a triangle with side a, b and c:

Proposition 16
The fifth moment of the chord-length for Δ is given by the following formula :

$$\mathbb{E}[\sigma^5] = \frac{5F^2T}{9L}$$

where $T = a^2 + b^2 + c^2$.

Proof : First we consider chords across a pair of line segments. Suppose two line segments along two axes inclined at angle C and that segments range over $a_1 \geq x \geq a_2$ and $b_1 \geq y \geq b_2$, as in figure 9.

The (unnormalised) odd moments of σ are :

$$I_{2n+3} = \int \sigma^{2n+3} \mathrm{d}G = \sin^2 C \int_{a_1}^{a_2} \int_{b_1}^{b_2} [x^2 + y^2 - 2xy\cos C]^n xy \mathrm{d}x \mathrm{d}y \quad .$$

Now, for the chords across the sides a and b of Δ, we have :

$$\begin{aligned} I_{5,ab} &= \sin^2\gamma \int_0^a \int_0^b [x^2 + y^2 - 2xy\cos\gamma] xy \mathrm{d}x \mathrm{d}y \\ &= \sin^2\gamma [\frac{a^4b^2}{8} + \frac{a^2b^4}{8} - 2\frac{a^3b^3}{9}\cos\gamma] \end{aligned}$$

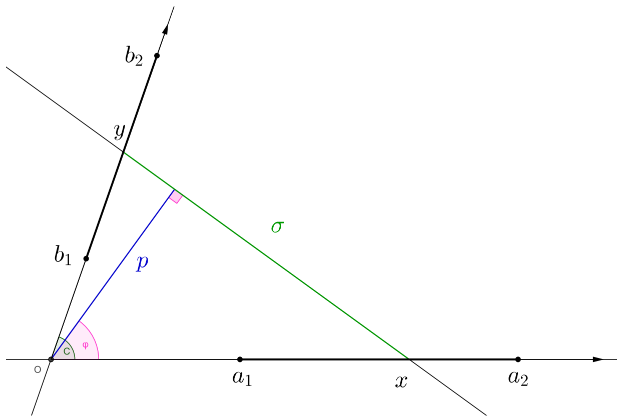


Figure 9:

Using *Al-Kashi* 's formula, we can simplify the sin and the cos :

$$I_{5,ab} = \frac{4a^2b^2 - (a^2+b^2-c^2)^2}{288}[a^2+b^2+8c^2] \quad ;$$

Using *Heron*'s formula, we obtain the result :

$$I_{5,ab} = \frac{F^2}{18}[a^2+b^2+8c^2] \quad .$$

The set of chords across Δ is the disjoint union of the sets of chords across each pairs of sides. So the fifth unnormalised moment of σ for Δ is the sum of each moments for pairs of sides. Normalising, we obtain the result announced. □

We remind the *Heron*'s formula:

Proposition 17 (*Heron*'s formula)
For a trianle of sides a, b and c, we have :

$$16F^2 = L(L-2a)(L-2b)(L-2c)$$

where L is the perimeter and F the area.

The sides of Δ are the roots of the cubic polynomial equation

$$(X-a)(X-b)(X-c) = X^3 - AX^2 + BX - C = 0\,,$$

with

$$A = a+b+c\,, \quad B = -(ab+ac+bc) \quad \text{and} \quad C = abc\,.$$

These coefficients can be expressed with the moments. We note $\mu = \mathbb{E}[\sigma]$, $\mu_3 = \mathbb{E}[\sigma^3]$ and $\mu_5 = \mathbb{E}[\sigma^5]$.

$$A = \frac{\pi^2\mu_3}{3\mu^2}\,, \quad B = \frac{\pi^2\mu_3^2}{18\mu^4} - \frac{27\mu_5}{10\mu_3} \quad \text{and} \quad C = \frac{\pi^6\mu_3^3}{216\mu^6} - \frac{9\pi^2\mu_5}{20\mu^2} - \frac{2}{3}\mu_3$$

If we have the chord-length distribution of a triangle, then we can calculate its moments, the coefficients A, B and C, and resolve the polynomial equation. Hence, a triangle is completely determined by its chord-length distribution.

3.4.1 Example of the regular triangle.

Let's see an application of this method with the case of a regular triangle. In [5], we can find the following result.

Proposition 18 (Chord-length distribution for a regular triangle.)
For a regular triangle of sides a, we have :

$$F_{\Delta}(t) = \begin{cases} 0 & \text{if} \quad t \leq 0 \\ (\frac{\pi}{3\sqrt{3}} + \frac{1}{2})\frac{t}{a} & \text{if} \quad 0 \leq t \leq \frac{a\sqrt{3}}{2} \\ \frac{2t}{a\sqrt{3}} \arcsin(\frac{a\sqrt{3}}{2t}) - \frac{2\pi t}{a3\sqrt{3}} + \frac{t}{2a} + \frac{\sqrt{4t^2-3a^2}}{2t} & \text{if} \quad \frac{a\sqrt{3}}{2} \leq t \leq a \\ 1 & \text{if} \quad t \geq a \end{cases}$$

(*see* Figure 10)

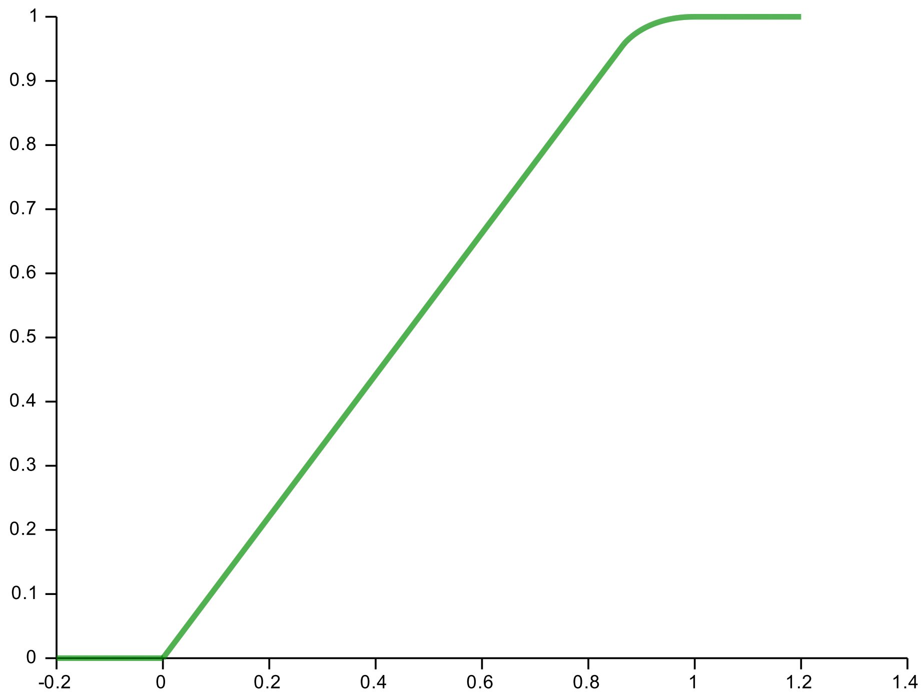


Figure 10: Chord-length distribution for a regular triangle with side $a = 1$.

With a software of numerical calculus, we obtain the moments and the constants (*see* appendix A.1). The roots of the polynomial equation are :

$$x_1 = 0.9647609\,, \quad x_2 = 1.0175865 + 0.0293943i \quad \text{and} \quad x_3 = 1.0175865 - 0.0293943i$$

This result is very near of what we were waiting for (which is $x_1 = x_2 = x_3 = 1$). We could make the error decrease by changing the method of numerical integration, for the moments, or changing the resolution of the cubic equation. Yet, it seems impossible to find the exact sides' length.

4 Distance between two random points.

For $\mathscr{P}_1$ and $\mathscr{P}_2$, two points in $\mathbb{R}^n$, we note $\rho(\mathscr{P}_1, \mathscr{P}_2)$ the distance between them. If $\mathscr{P}_1$ and $\mathscr{P}_2$ have the coordinates $(x_i^{(1)})_{1\leq i\leq n}$ and $(x_i^{(2)})_{1\leq i\leq n}$ respectively, we have the formula :

$$\rho(\mathscr{P}_1, \mathscr{P}_2) = \sqrt{\sum_{i=1}^{n} (x_i^{(1)} - x_i^{(2)})^2} \quad .$$

Let $\mathscr{D}$ be a body in $\mathbb{R}^n$ and $\mathscr{P}_1$ and $\mathscr{P}_2$ two random points uniformly dropped in $\mathscr{D}$. In this case $\rho(\mathscr{P}_1, \mathscr{P}_2)$ is a random variable. The goal of this section is to study its distribution, *i.e* the function

$$F_\rho(x) = \mathbb{P}(\rho(\mathscr{P}_1, \mathscr{P}_2) \leq x) = \frac{1}{V_n^2(\mathscr{D})} \int_{\{(\mathscr{P}_1,\mathscr{P}_2)\in\mathscr{D}^2/\rho(\mathscr{P}_1,\mathscr{P}_2)\leq x\}} \mathrm{d}\mathscr{P}_1 \mathrm{d}\mathscr{P}_2 \quad .$$

We focus our study in rectangular bodies : $[a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$.

4.1 Some preliminary results.

Let's begin with some little results of probability that we need to study the density of distances.

Proposition 19 (Square of a random variable)
If η is a random variable with density function f_η, then η^2 is a positive random variable with density function f_{η^2} and we have :

$$f_{\eta^2}(t) = \begin{cases} 0 & \text{if } \ t \leq 0 \\ \frac{1}{2\sqrt{t}}(f_\eta(\sqrt{t}) + f_\eta(-\sqrt{t})) & \text{if } \ t > 0 \end{cases}$$

Proof : For all t in $\mathbb{R}$, we have :

$$\begin{aligned} F_{\eta^2}(t) &= \mathbb{P}(\eta^2 \leq t) \\ &= \begin{cases} 0 & \text{if } \ t \leq 0 \\ \mathbb{P}(-\sqrt{t} \leq \eta \leq \sqrt{t}) & \text{if } \ t > 0 \end{cases} \\ &= \begin{cases} 0 & \text{if } \ t \leq 0 \\ F_\eta(\sqrt{t}) - F_\eta(-\sqrt{t}) & \text{if } \ t > 0 \end{cases} \end{aligned}$$

By differentiation, we obtain the result. □

Proposition 20 (Square root of a positive random variable)
If η is a positive random variable with density function f_η, then $\sqrt{\eta}$ is a positive random variable with density function $f_{\sqrt{\eta}}$ and we have :

$$f_{\sqrt{\eta}}(t) = \begin{cases} 0 & \text{if} \quad t \leq 0 \\ 2tf_\eta(t^2) & \text{if} \quad t > 0 \end{cases}$$

Proof : For all t in $\mathbb{R}$, we have :

$$F_{\sqrt{\eta}}(t) = \mathbb{P}(\sqrt{\eta} \leq t)$$

$$= \begin{cases} 0 & \text{if} \quad t \leq 0 \\ \mathbb{P}(\eta \leq t^2) & \text{if} \quad t > 0 \end{cases}$$

$$= \begin{cases} 0 & \text{if} \quad t \leq 0 \\ F_\eta(t^2) & \text{if} \quad t > 0 \end{cases}$$

By differentiation, we obtain the result. □

4.2 A rectangle in the plane.

Let $\mathscr{R}$ be a rectangle of length L and width ℓ ($l \leq L$). As the distribution of distance is invariant with respect to rigid motions, we assume that

$$(x, y) \in \mathscr{R} \Leftrightarrow (x \in [0, L] \quad \text{and} \quad y \in [0, \ell]) \quad .$$

If $\mathscr{P}_1 = (X_1, Y_1)$ is uniformly distributed in $\mathscr{R}$, then X_1 and Y_1 are independently uniformly distributed in $[0, L]$ and $[0, l]$ respectively. As $\mathscr{P}_1$ and $\mathscr{P}_2$ are independent, then X_1, X_2, Y_1 and Y_2 are independent.
We can see a way to calculate the distribution of ρ:

- calculate the density of $X_1 - X_2$;
- calculate the density of $(X_1 - X_2)^2$;
- do the same for Y;
- calculate the density of $\rho^2 = (X_1 - X_2)^2 + (Y_1 - Y_2)^2$;
- calculate the density of ρ.

This method is very general and could be applied in $\mathbb{R}^n$. The thinks that make the calculus possible (in theory) is the independence between the the coordinates of each point in a rectangular. If we want to study the distribution of distance in the disc with raduis R and centre O, the Cartesian coordinates would not be well adapted because there is a relation of dependence between x and y :

$$x^2 + y^2 \leq R^2 \quad .$$

Theorem 21
The density of the distance in a rectangle in the plane, with length L and width ℓ, is given by the formula :

$$f_\rho(t) = \begin{cases} \frac{2t}{F^2}[t^2 - 2(L+\ell)t + \pi L\ell] & \text{if} \quad 0 \leq t \leq \ell \\ \frac{2t}{F^2}[2L\sqrt{t^2-\ell^2} - 2Lt \\ +2L\ell \arcsin(\frac{\ell}{t}) - \ell^2] & \text{if} \quad \ell \leq t \leq L \\ \frac{2t}{F^2}[2L\sqrt{t^2-\ell^2} + 2\ell\sqrt{t^2-L^2} \\ +2L\ell \arcsin(\frac{L}{t}) + 2L\ell \arcsin(\frac{\ell}{t}) \\ -L^2 - \ell^2 - t^2 - \pi L\ell] & \text{if} \quad L \leq t \leq \sqrt{L^2+\ell^2} \\ 0 & \text{otherwise} \end{cases}$$

where $F = L\ell$ is the area of the rectangle.

(*see* Figure 11)

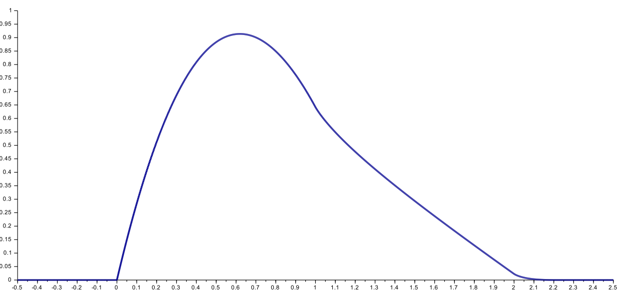


Figure 11: The density for a rectangle with $L = 2$ and $\ell = 1$.

Proof : For $i = 1, 2$, $f_{X_i}(t) = \dfrac{1}{L}\mathbb{1}_{[0,L]}(t)$ and $f_{Y_i}(t) = \dfrac{1}{\ell}\mathbb{1}_{[0,\ell]}(t)$.

Using propositions 8 and 9, for all t in $\mathbb{R}$ we have

$$f_{X_1-X_2}(t) = \frac{1}{L^2}\int_{\mathbb{R}} \mathbb{1}_{[0,L]}(s)\mathbb{1}_{[0,L]}(s-t)\mathrm{d}s$$

$$= \begin{cases} 0 & \text{if} \quad |t| > L \\ \frac{L-|t|}{L^2} & \text{otherwise} \end{cases}$$

(*see* Figure 12)

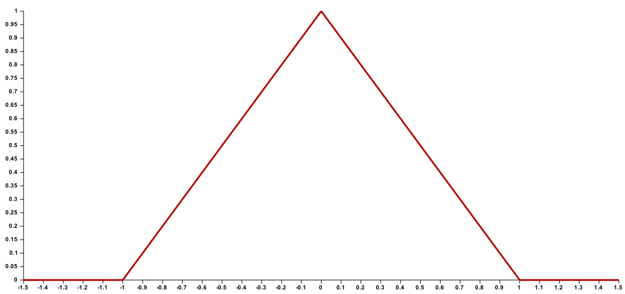


Figure 12: The density of $X_1 - X_2$, for $L = 1$.

We denote $\eta_X = X_1 - X_2$ and $\eta_Y = Y_1 - Y_2$. From proposition 19, we have

$$f_{\eta_X^2}(t) = \begin{cases} \frac{L-\sqrt{t}}{L^2\sqrt{t}} & \text{if} \quad 0 \leq t \leq L^2 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f_{\eta_Y^2}(t) = \begin{cases} \frac{\ell-\sqrt{t}}{\ell^2\sqrt{t}} & \text{if} \quad 0 \leq t \leq \ell^2 \\ 0 & \text{otherwise} \end{cases}$$

(*see* Figure 13)

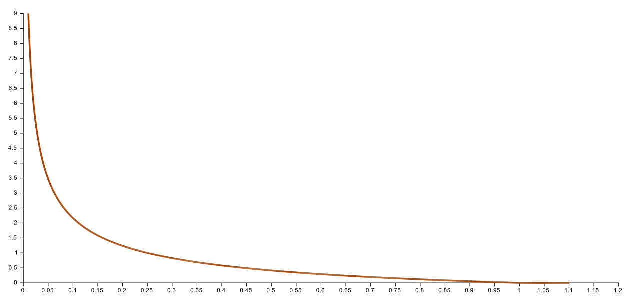


Figure 13: The density of $(X_1 - X_2)^2$, for $L = 1$.

As $\rho^2 = \eta_X^2 + \eta_Y^2$, we have :$f_{\rho^2}(t) = \int_{\mathbb{R}} f_{\eta_Y^2}(s) f_{\eta_Y^2}(t-s)\mathrm{d}s$.

$$f_{\rho^2}(t) = \begin{cases} \int_0^t \frac{L-\sqrt{s}}{L^2\sqrt{s}} \frac{\ell-\sqrt{t-s}}{\ell^2\sqrt{t-s}} \mathrm{d}s & \text{if} \quad 0 \le t \le \ell^2 \\ \int_{t-\ell^2}^t \frac{L-\sqrt{s}}{L^2\sqrt{s}} \frac{\ell-\sqrt{t-s}}{\ell^2\sqrt{t-s}} \mathrm{d}s & \text{if} \quad \ell^2 \le t \le L^2 \\ \int_{t-\ell^2}^{L^2} \frac{L-\sqrt{s}}{L^2\sqrt{s}} \frac{\ell-\sqrt{t-s}}{\ell^2\sqrt{t-s}} \mathrm{d}s & \text{if} \quad L^2 \le t \le \ell^2 + L^2 \end{cases}$$

All integrals are elementary, except it :

$$\begin{aligned} \int_a^b \frac{\mathrm{d}s}{\sqrt{s}\sqrt{t-s}} &= \int_{a-\frac{t}{2}}^{b\frac{t}{2}} \frac{\mathrm{d}u}{\sqrt{\frac{t}{2}+u}\sqrt{\frac{t}{2}-u}} \\ &= \frac{2}{t}\int_{a-\frac{t}{2}}^{b\frac{t}{2}} \frac{\mathrm{d}u}{\sqrt{1-(\frac{2u}{t})^2}} \\ &= \int_{\frac{2a}{t}-1}^{\frac{2b}{t}-1} \frac{\mathrm{d}s}{\sqrt{1-s^2}} \\ &= [\arcsin(s)]_{\frac{2a}{t}-1}^{\frac{2b}{t}-1} \end{aligned}$$

We notice that :

$$2\arcsin\left(\frac{a}{\sqrt{x}}\right) = \arcsin\left(\frac{2a^2}{x}-1\right) + \frac{\pi}{2} \quad .$$

Using the proposition 20 we find the result. □

The study of the density can reveal geometric properties of the body. For example, it is interesting to see that the density of distance in a rectangle is continuous, derivable and that the derivative function is also continuous (except at the boundary of the support). For a triangle, the derivative is discontinuous on sides' length.

In the particuliar case of a square, we have the following result.

Proposition 22
The density of the distance in a square, with length ℓ, is given by the formula :

$$f_{\rho_2}(t) = \begin{cases} \frac{2t}{F^2}[t^2 - 4\ell t + \pi\ell^2] & \text{if} \quad 0 \le t \le \ell \\ \frac{2t}{F^2}[4\ell\sqrt{t^2-\ell^2} + 4\ell^2\arcsin(\frac{\ell}{t}) - 2\ell^2 - t^2 - \pi\ell^2] & \text{if} \quad \ell \le t \le \sqrt{2\ell^2} \\ 0 & \text{otherwise} \end{cases}$$

where $F = \ell^2$ is the area of the square.

4.3 A cube in the space.

If the calculus for the rectangular in the plane were not to difficult, it becomes more complicated in higher dimensions. Whereas calculating the density for any parallelepiped, we restrain ourself to the case of the cube. We note ρ_1 the distance between two random points in a line of length ℓ, ρ_2 the distance between two random points in a square of side's length ℓ and ρ_3 the distance between two random points in a cube of side's length ℓ.

Using the same method as in the last part, we have

$$f_{\rho_3}(t) = f_{\eta_X^2} * f_{\eta_Y^2} * f_{\eta_Z^2}(t)\,, \quad \forall t_i n \mathbb{R} \quad .$$

As we work on a cube, η_X, η_X and η_X are *i.i.d.* Using the associativity of the convolution, we have

$$f_{\rho_3}(t) = f_{\rho_2^2} * f_{\rho_1^2}(t)\,, \quad \forall t \in \mathbb{R} \quad .$$

Three intervals appear : $0 \leq t \leq \ell^2, \quad \ell^2 \leq t \leq 2\ell^2 \quad$ and $\quad 2\ell^2 \leq t \leq 3\ell^2$.

If $0 \leq t \leq \ell^2$, then we have

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\int_0^t \left(s - 4\ell\sqrt{s} + \pi\ell^2\right)\left(\frac{\ell}{\sqrt{t-s}} - 1\right) \mathrm{d}s$$

$$= \frac{1}{\ell^6}\left[\int_0^t \left(\frac{\ell s}{\sqrt{t-s}} - \frac{4\ell\sqrt{s}}{\sqrt{t-s}} + \frac{\pi\ell^2}{\sqrt{t-s}}\right)\mathrm{d}s - \frac{t^2}{2} + \frac{8}{3}\ell t\sqrt{t} - \pi\ell^2 t\right]$$

For $0 \leq t \leq \ell^2$, we have

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\left[-\frac{t^2}{2} + 4\ell t\sqrt{t} - 3\pi\ell^2 t + 2\pi\ell^3\sqrt{t}\right] \quad .$$

If $\ell^2 \leq t \leq 2\ell^2$, we must separate the integral in two, because the expression of f_{ρ_2} change in ℓ^2. We have

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\int_{t-\ell^2}^{\ell^2} f_{\rho_2^2}(s) f_{\rho_1^2}(t-s)\mathrm{d}s + \frac{1}{\ell^6}\int_{\ell^2}^{t} f_{\rho_2^2}(s) f_{\rho_1^2}(t-s)\mathrm{d}s$$

We denote by $I_1(t)$ and $I_2(t)$ the two integrals.

$$
I_1(t) = \int_{t-\ell^2}^{\ell^2} \left(s - 4\ell\sqrt{s} + \pi\ell^2\right)\left(\frac{\ell}{\sqrt{t-s}} - 1\right) \mathrm{d}s
$$

$$
= \int_{t-\ell^2}^{\ell^2} \left(\frac{\ell s}{\sqrt{t-s}} - \frac{4\ell\sqrt{s}}{\sqrt{t-s}} + \frac{\pi\ell^2}{\sqrt{t-s}}\right) \mathrm{d}s - \frac{\ell^4 - (t-\ell^2)^2}{2} + \frac{8}{3}\ell\left(\ell^3 - (t-\ell^2)^{3/2}\right)
$$

$$
- \pi\ell^2(\ell^2 - t\ell^2)
$$

$$
= \ell \int_{t-\ell^2}^{\ell^2} \frac{t-u}{\sqrt{u}} \mathrm{d}u - 4\ell^2 \left[\sqrt{s}\sqrt{t-s} - t\arctan\left(\frac{\sqrt{t-s}}{\sqrt{s}}\right)\right]_{t-\ell^2}^{\ell^2} + \pi\ell^3 \int_{t-\ell^2}^{\ell^2} \frac{\mathrm{d}u}{\sqrt{u}}
$$

$$
- \frac{\ell^4 - (t-\ell^2)^2}{2} + \frac{8}{3}\ell\left(\ell^3 - (t-\ell^2)^{3/2}\right) - \pi\ell^2(\ell^2 - t\ell^2)
$$

$$
= \frac{1}{2}t^2 - 2\ell(t-\ell^2)^{3/2} - 2\ell t\sqrt{t-\ell^2} + t\ell^2(1-\pi) - 2\pi\ell^3\sqrt{t-\ell^2}
$$

$$
+ 8\ell^2 t\arctan\left(\frac{\sqrt{t-\ell^2}}{\ell}\right) + 2\ell^4
$$

$$I_2(t) = \int_{\ell^2}^{t} \left(4\ell\sqrt{t-\ell^2} - s + 4\ell^2 \arcsin\left(\frac{\ell}{\sqrt{s}}\right) - \ell^2(\pi+2)\right)\left(\frac{\ell}{\sqrt{t-s}} - 1\right) \mathrm{d}s$$

$$= \int_{\ell^2}^{t} \left(4\ell^2 \frac{\sqrt{s-\ell^2}}{\sqrt{t-s}} - \ell\frac{s}{\sqrt{t-s}} + 4\ell^3 \frac{\arcsin\left(\frac{\ell}{\sqrt{s}}\right)}{\sqrt{t-s}} - \frac{\ell^3(\pi+2)}{\sqrt{t-s}}\right) \mathrm{d}s$$

$$- \frac{8}{3}\ell(t-\ell^2)^{3/2} + \frac{t^2-\ell^4}{2} - 4\ell^2 \int_{\ell^2}^{t} \arcsin\left(\frac{\ell}{\sqrt{s}}\right) \mathrm{d}s + \ell^2(\pi+2)(t-\ell^2)$$

$$= 4\ell^2(t-\ell^2)\left[\arctan\left(\frac{\sqrt{x-\ell^2}}{t-x}\right)\right]_{\ell^2}^{t} - \ell \int_0^{t-\ell^2} \frac{t-u}{\sqrt{u}} \mathrm{d}u$$

$$+ 4\ell^3\left(\pi\ell + \pi\sqrt{t-\ell^2} - \pi\sqrt{t}\right) - \ell^3(\pi+2)\int_0^{t-\ell^2} \frac{\mathrm{d}u}{\sqrt{u}} - \frac{8}{3}\ell(t-\ell^2)^{3/2}$$

$$+ \frac{t^2}{2} - \frac{\ell^4}{2} - 4\ell^2\left[\ell\sqrt{s-\ell^2} + s\arcsin\left(\frac{\ell}{\sqrt{s}}\right)\right]_{\ell^2}^{t} + \ell^2(\pi+2)t - \ell^2(\pi+2)\ell^2$$

We notice that

$$\arcsin\left(\frac{\ell}{\sqrt{s}}\right) = \frac{\pi}{2} - \arctan\left(\frac{\sqrt{s-\ell^2}}{\ell}\right) \ , \ell^2 \leq s \leq t.$$

So we have

$$I_2(t) = \frac{t^2}{2} - 2\ell(t-\ell^2)^{3/2} - 2\ell t\sqrt{t-\ell^2} + (2\pi-8)\ell^3\sqrt{t-\ell^2} + t\ell^2(\pi+2) - 4\pi\ell^3\sqrt{t}$$

$$+ \ell^4(3\pi - \frac{5}{2}) + 4\ell^2 t \arctan\left(\frac{\sqrt{t-\ell^2}}{\ell}\right)$$

For $\ell^2 \leq t \leq 2\ell^2$, we have

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\left[t^2 - 8\ell t\sqrt{t-\ell^2} - 4\ell^3\sqrt{t-\ell^2} + 3\ell^2 t - 4\pi\ell\sqrt{t} + \ell^4(3\pi - \frac{1}{2}) + 12\ell^2 t \arctan\left(\frac{\sqrt{t-\ell^2}}{\ell}\right)\right] \quad .$$

For $2\ell^2 \leq t \leq 3\ell^2$, we have

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\int_{t-\ell^2}^{2\ell^2}\left(4\ell\sqrt{t-\ell^2} - s + 4\ell^2\arcsin\left(\frac{\ell}{\sqrt{s}}\right) - \ell^2(\pi+2)\right)\left(\frac{\ell}{\sqrt{t-s}} - 1\right)\mathrm{d}s \quad .$$

As is is the same kernel as in I_2, we give the result without more calculus.

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\left[-\frac{1}{2}t^2 - t\ell^2(\pi-3) + 4\ell t\sqrt{t-2\ell^2} + \ell^3(2\pi+6)\sqrt{t-2\ell^2} - \ell^4(2\pi + \frac{13}{2}) + 4\ell^2(t-\ell^2)\arcsin\left(\frac{\ell}{\sqrt{t-\ell^2}}\right) - 8\ell^2(t-\ell^2)\arctan\left(\frac{\sqrt{t-2\ell^2}}{\ell}\right) + 4\ell^3\int_{t-\ell^2}^{2\ell^2}\frac{\arcsin\left(\frac{\ell}{\sqrt{s}}\right)}{\sqrt{t-s}}\mathrm{d}s\right]$$

The explicit expression of the integral

$$\int \frac{\arcsin(\frac{\ell}{\sqrt{s}})}{\sqrt{t-s}}\mathrm{d}s \quad ,$$

is difficult to simplify. But we know from [9] that the result is

For $2\ell^2 \leq t \leq 3\ell^2$, we have

$$f_{\rho_3^2}(t) = \frac{1}{\ell^6}\left[-\frac{1}{2}t^2 + 4\ell t\sqrt{t-2\ell^2} + 4\ell^3\sqrt{t-2\ell^2} + 3(\pi-1)\ell^2 t - 4\pi\ell^3\sqrt{t} + \ell^4(3\pi - \frac{5}{2}) - 12\ell^2(t+\ell^2)\arctan\left(\frac{\sqrt{t-2\ell^2}}{\ell}\right) + 12\ell^3\sqrt{t}\arctan\left(\frac{\sqrt{t}\sqrt{t-2\ell^2}}{\ell^2}\right)\right] \quad .$$

4.4 Numerical approximations for cubes and hypercubes.

In [9] and [10], Johan Philip give the result for a box, a cube and an hypercube in four dimensions. It is interesting to notice that Philip use the same method to make his calculus. In the case of four dimensions, he faces integrals which cannot be expressed with usual function.

In this part, we compare the theoretical results obtained by Philip with numerical results. The idea is to implement a function `convolution(f,g,t)` where `f` and `g` are two functions and `t` a float, which calculate an approximation of

$$f * g(t) = \int_{\mathbb{R}} f(s)g(t-s)\mathrm{d}s \quad .$$

We define a function `eta_sq`, the density of η^2 for a unit segment and `sqare_sq` the density of the square of the distance for a unit square. This is not a matter as we obtained explicit expressions. We calculate the functions `app_cube_sq` and `app_hypercube_sq`, which are approximations of the densities of ρ^2 for a unit cube and hypercube. You can see the functions in appendix A.2.

Let us begin with the unit cube. As we have explicit expression of the density, it allow to compare and to see if our numericals results are faithful. In the Figure 14, we draw in green the density given by Philip and in red the approximation. The integration step is `p=0.000001` and we calculate the values for `X=[0.0001:0.01:1.8]`. It took 2 hours 42 minutes for the computer to calculate these values.

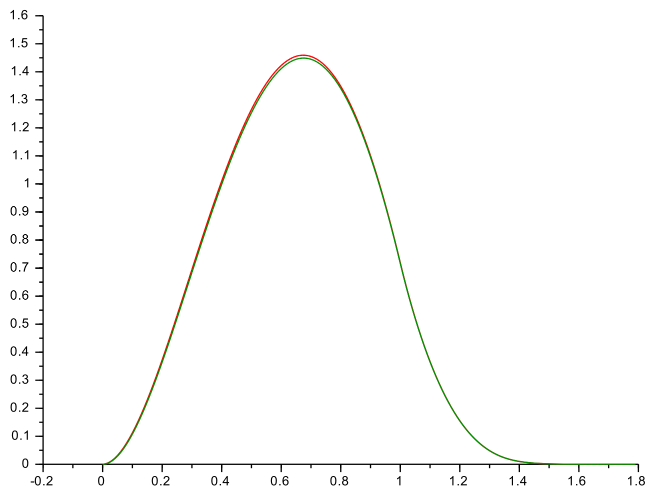


Figure 14: Exact density of the distance in an unit cube and an approximation.

The numerical approximation fit the theoretical graph. We have only 180 points, which is very few. Yet, the time of calculation is already important. If we increase the number of point, we face a big problems. For a reason of time, we must decrease the integration step, and this decrease the precision of the calculus, irregularities appear. In Figure 15, we have a illustration of this phenomenon. The integration step used isp=`0.000005` and we print the values for `X=[0.0001:0.005:1.8]`. The calculus took 1 hour 12 minutes.

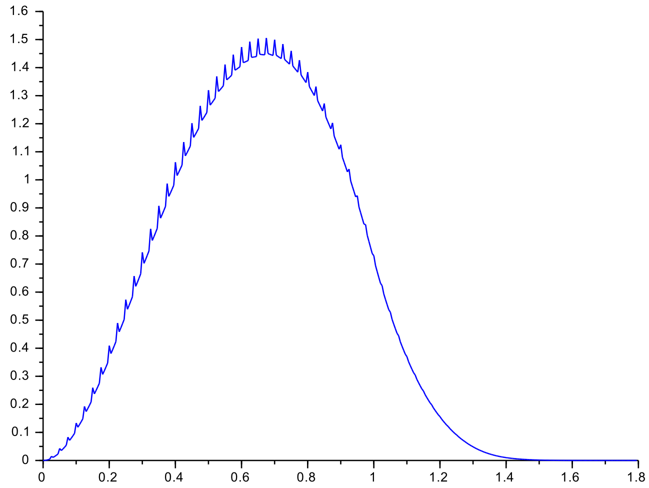


Figure 15: Peaks on the approximation.

In the case of the hypercube, the formula given by Philip is not explicit. Hence, we cannot compare theoretical graph and approximation. In Figure 16, we find an approximation of the density f distance for an unit hypercube in 4 dimensions with an integration step p=`0.000001` for the range of values `X=[0.0001:0.05:1.8]`. The computer took 51 minutes to calculate.

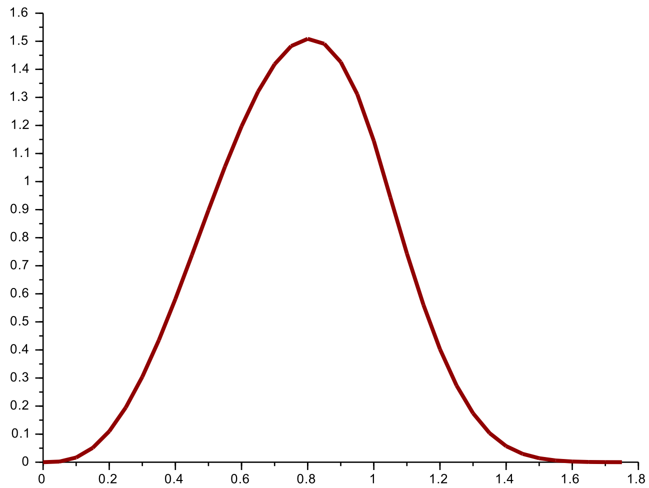


Figure 16: Approximation of the density for an unit hypercube in 4D.

Conclusion.

This rapport accustom us with three tools of stochastic geometry : covariogram, chord-length distribution and distance between random points in a body. Even if we do not go into their links in depth, they are bound by very close relation. If these relations are well-known in $\mathbb{R}^2$, they should be extended to other dimensions. The examples we choose to illustrate this tools are particular cases which make calculus easier, but we can see that, even with very simple body, it can become very complicated.

For the distance between two points in a cube and hypercube, I am disappointed of not being able to succeed in calculating the density. Our technique "step by step" cannot be used to find a formula for any n. We should search for another way. Perhaps with Fourier transform, which prevent from convolution products.

Acknowledgement.

I would like to thank V. Ohanyan for accepting me in internship. He made me discover stochastic geometry with very interesting problems. I was very glad to work with him and I hope we can work together another time. I also want to thank all members of the mathematics department with who I have work or just talk and who have welcome me in their university. I think particularly to Ara with who I had a visit "VIP" of the Matenadaran.

I wan to thank Manvel, who give me the idea of going to Armenia and who give me the email of V. Ohanyan. I want to thank Fieroussa, Gago and Sussu for welcome me in their house during all my internship and take care of me with kindness and generosity. Thanks to Samvel who made me discover his home town, and thanks to all their family for the good time I spent with them.

My trip in Armenia was a great experiment, for mathematics and for human relations. I will keep very good remembrances. I hope I will come back there.

References

[1] Luis A. SANTALÓ. *Integral Geometry and Geometric Probability*. Addison-Wesley Publishing Compagny, 2^{nd} edition, 2004.

[2] J. Gates. *"Recognition of Triangles and Quadrilaterals by Chord Length Distribution"*. Journal of Applied Probability, vol.19, pages 873-879. 1982.

[3] J. Gates. *"Some properties of chord-length distributions"*. Journal of Applied Probability, vol.24, pages 863-874. 1987.

[4] W. Nagel. *"Orientation-Dependent Chord Length Distributions Characterize Convex Polygons"*. Journal of Applied Probability, vol.30, pages 730-736. 1993.

[5] A. Gasparyan and V. K. Ohanyan. *"Recognition of Triangles by Covariogram"*. Journal of Contemporary Mathematical Analysis (Armenian Academia of sciences), vol.48, pages 110-122. 2013.

[6] A. G. Gasparyan and V. K. Ohanyan. *"Covariogram of a Parallelogram"*. Journal of Contemporary Mathematical Analysis (Armenian Academia of sciences), vol.49, pages 194-206. 2014.

[7] H. S. Harutyunyan and V. K. Ohanyan. *"Chord-length distribution function for regular polygons"*. Journal of Applied Probability, 2009

[8] R. V. AMBARTZUMYAN. *Factorization Calculus and Geometric Probability*. Cambridge University Press, 1990.

[9] Johan PHILIP. *"The probability distribution between two random points in a box"*. `https://people.kth.se/ johanph/habc.pdf`, 1991.

[10] Johan PHILIP. *"The distance between two random points in a 4- and 5- cube"*. `https://people.kth.se/ johanph/h45.pdf`, 1991.

[11] `http://www.wolframalpha.com`

A Algorithms with Scilab.

A.1 Calculation of chord-length distribution 's moments for a triangle.

These are the functions used in part 3.4.1. First we need to define the distribution function for a regular triangle with side's length $a = 1$.

```
function y=F(t)
    if t<=0 then y=0
    elseif t<=sqrt(3)/2 then y=(%pi/(3*sqrt(3))+1/2)*t
    elseif t<=1 then y=2*t*asin(sqrt(3)/(2*t))/sqrt(3)-2*%pi*t/(3*sqrt(3))+t/2
   +(sqrt(4*t*t-3))/(2*t)
    elseif t>1 then y=1
    end
endfunction
```

The function `moment(n,p)` calculate the $\mathtt{n}^{th}$ moment of the distribution `F`.

```
function m=moment(n,p)
    x=[0:p:1]
    m=0
    N=length(x)
    for i=1:N
        m=m+p*n*x(i)^(n-1)*(1-F(x(i)))
    end
endfunction
```

It is based on the simplest methods of numerical integration : the rectangle method. The integration's step is `p`. This function could be easily improve by changing the method of numerical integration, but this does not concern us.

A.2 Disance between two random points.

The function `convolution(f,g,t)` is based on the rectangular method. The float `p` is the step of integration.

```
function h=convolution(f,g,t)
    p=0.000005
    x=[0.0001:p:5]
    s=0
    for i=1:length(x)
        s=s+p*f(x(i))*g(t-x(i))
    end
    h=s
endfunction
```

The interval of integration is `x=[0.00001:p:2.1]` because we know that in our problem, all functions have their support in $[0, 2.1]$ and we need to exclude 0 for numerical reasons.

The density of η^2 for a of length $\ell = 1$ is given by the function `eta_sq`,

```
function y=eta_sq(t)
        if (t<=0)|(t>1) then y=0
          else y=(1-sqrt(t))/sqrt(t)
     end
endfunction
```

and the density of ρ_2 for the unit square is given by the function `square`,

```
function y=square_sq(t)
         if t <=0 then y=0
     elseif t<=1 then y=t-4*sqrt(t)+%pi
     elseif t<=2 then y=4*sqrt(t-1)+4*asin(1/sqrt(t))-t-2-%pi
          else y=0
     end
endfunction


function y=square(t)
     y=2*t*square_sq(t*t*sign(t))
endfunction
```

Then the approximations of densities of the square of the distance for an unit cube and an unit hypercube are given by the function `app_cube_sq` and `app_hyper_sq`

```
function y=app_cube_sq(t)
     y=convolution(square_sq,eta_sq,t)
endfunction


function y=app_hyper_sq(t)
     y=convolution(square_sq,square_sq,t)
endfunction
```

To plot the density of the distance (and not the square of the distance), we just use

```
X=[0.0001:0.005:1.8]
for j=1:length(X)
     Z(j)=2*X(j)*app_cube_sq(X(j)*X(j))
end
plot(X,Z)
```

The exact density of the distance for an unit cube is given by `cube`

```
function y=cube(t)
    if  t<=0 then y=0
    elseif  t<=1 then y=t*t*(4*%pi-6*%pi*t+8*t*t-t*t*t)
    elseif  t<=sqrt(2) then y=(6*%pi-1)*t-8*%pi*t*t+6*t*t*t+2*t*t*t*t*t+24*t*t*t
   *atan(sqrt(t*t-1))-8*t*(1+2*t*t)*sqrt(t*t-1)
    elseif  t<=sqrt(3) then y=(6*%pi-5)*t-8*%pi*t*t+6*(%pi-1)*t*t*t-t*t*t*t*t+8
   *t*(1+t*t)*sqrt(t*t-2)-24*t*(1+t*t)*atan(sqrt(t*t-2))+24*t*t*atan(t*sqrt(t*
   t-2))
    else  y=0
    end
endfunction
```

B List of useful integrals.

These are non-elementary integrals we use in the part 4.3. These expressions fit for $\ell^2 \leq x \leq t$. We found these expressions in [11].

$$\int \frac{\sqrt{x}}{\sqrt{t-x}} \mathrm{d}x = t \arctan\left(\frac{\sqrt{t}}{\sqrt{t-x}}\right) - \sqrt{x}\sqrt{t-x} \tag{1}$$

$$\int \arcsin\left(\frac{l}{\sqrt{x}}\right) \mathrm{d}x = \ell\sqrt{x-\ell^2} + x \arcsin\left(\frac{\ell}{\sqrt{x}}\right) \tag{2}$$

$$\int \frac{\arcsin\left(\frac{l}{\sqrt{x}}\right)}{\sqrt{t-x}} \mathrm{d}x = -2\sqrt{t-x} \arcsin\left(\frac{l}{\sqrt{x}}\right) - \ell \arctan\left(\frac{\ell^2+t-2x}{2\sqrt{x-\ell^2}\sqrt{t-x}}\right)$$

$$+ \mathbf{i}\sqrt{t} \log\left(\frac{\mathbf{i}\ell^2(x-2t) + 2\ell\sqrt{t}\sqrt{x-\ell^2}\sqrt{t-x} + \mathbf{i}tx}{\ell\sqrt{t}x}\right)$$

We deduce this integral from (1)

$$\int \frac{\sqrt{x-\ell^2}}{\sqrt{t-x}} \mathrm{d}x = (\ell^2 - t) \arctan\left(\frac{\sqrt{x-\ell^2}}{\sqrt{t-x}}\right) - \sqrt{x-\ell^2}\sqrt{t-x}$$