Extremal configurations of solids and the Morse theory of distance functions

Oleg Ogievetsky, joint work with S. Shlosman

arXiv: 1805.09833, 1812.09543, 1902.08995, 1904.02043, 1907.01896

Centre de Physique Théorique, Marseille, France

Yerevan, August 2019

I will talk about a circle of questions inspired by an old discussion (1694) between Isaac Newton and David Gregory about the number of non-intersecting unit spheres touching the central unit sphere.

I will talk about a circle of questions inspired by an old discussion (1694) between Isaac Newton and David Gregory about the number of non-intersecting unit spheres touching the central unit sphere.

Newton (draft for a second edition of the *Principia*) :

There are 13 stars of the first magnitude and roughly the same number of equal spheres can be arranged about a central sphere equal to them.

(The right answer is 12, K. Schutte and B. L. van der Waerden, 1953)

We will discuss moduli spaces $\mathcal{C}$ of configurations (clusters) of collections of solid bodies $\bar{\Lambda}_1, \ldots, \bar{\Lambda}_k \subset \mathbb{R}^3$, touching the central unit ball $B \subset \mathbb{R}^3$. That is, the point G of our manifold $\mathcal{C}$ is a configuration of non-intersecting solid bodies, $G = \{\Lambda_1, \ldots, \Lambda_k\}$, where each Λ_i is congruent to the corresponding shape $\bar{\Lambda}_i$ and is touching the unit ball B. It is allowed that some distances between bodies of G are zero.

The group $SO(3)$ acts on each $\mathcal{C}$. So it is natural to study these manifolds mod $SO(3)$.

A simplest (Newton–Gregory) type : arrangements of balls Λ_i of equal radius r, around B.

Wlodek Kuperberg (1990) asked a question, similar to the Newton–Gregory one, but about infinite objects :

How many non-intersecting unit cylinders can touch the unit ball ?

The examples and illustrations in this talk will mostly concern this type : arrangements of (infinite, right, circular) congruent cylinders around B.

Six cylinders can touch a unit ball :

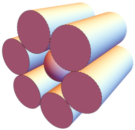


FIGURE – Cluster C_6

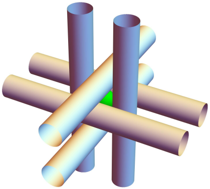


FIGURE – Cluster O_6 of cylinders

It is known that 8 cylinders can not do it (P. Brass and C. Wenk, 2000). For 7 cylinders this is an open question.

Balls :

Let $\mathcal{P}_n$ be the manifold of n-tuples of points on $\mathbb{S}^2$. To every cluster of n balls we associate a point in $\mathcal{P}_n$: it is the cluster of n points at which the balls of the cluster are touching the central ball. Consider the function D on $\mathcal{P}_n$:

$$D(p) = \min_{i \neq j} \operatorname{dist}(x_i, x_j) \quad \text{where } p = \{x_1, \ldots, x_n\} \in \mathcal{P}_n.$$

A cluster of balls, some of which touch each other, can be reconstructed from the corresponding cluster of points (giving the maximal common diameter $\tilde{D}(p)$ of balls touching B at p).

The Newton–Gregory question can be reformulated in the form : for $n = 13$, is the maximal value of the function $\tilde{D}$ (on $\mathcal{P}_{13}$) smaller than 2 ?

Cylinders :

A cylinder touching the unit sphere B has a unique generator tangent to B. The manifold of tangent lines to B we denote by M. So to every cluster of n cylinders we associate a point in M^n.

Again, a cluster of cylinders, some of which touch each other, can be reconstructed from the corresponding cluster of tangent to B lines.

As for balls, the Kuperberg question can be reformulated as the question about the maximal value of the function $D(\mathbf{m}) = \min_{1 \leq i < j \leq n} d_{u_i u_j}$ on M^n. Here d_{uv} is the distance between the tangent lines u and v. The difference with the clusters of balls is that the function d_{uv} does not have a limit when u and v become parallel.

By a deformation of the cluster G we mean a continuous curve $G(t)$ in the space of clusters, with $G(0) = G$. That means that each Λ_i touches the central ball B during the process of the deformation.

We call a cluster G *rigid,* if any deformation $G(t)$ of G has a form

$$G(t) = g(t)\, G \ ,$$

where $g(t) \in SO(3)$ is a curve of rotations of $\mathbb{R}^3$, $g(0) = e$. In other words, the only deformation of G available is the global rotation of G as a solid body.

We say that G can be unlocked, if there exists a continuous deformation $G(t)$ of G, such that for any $t > 0$ all the distances between the members Λ_i in the cluster $G(t)$ are positive (while each Λ_i always touches the central ball during the move).

Finally, we call a cluster G *critical,* if for any smooth deformation $G(t)$ of G all the distances $\mathrm{dist}(\Lambda_i(t), \Lambda_j(t))$ between the solids Λ_i which were zero at $t = 0$ – i.e. $\mathrm{dist}(\Lambda_i(0), \Lambda_j(0)) = 0$ – obey the estimate

$$\mathrm{dist}(\Lambda_i(t), \Lambda_j(t)) = o(t) \text{ as } t \to 0 . \tag{1}$$

If a critical cluster G can be unlocked, then it is called a *saddle* cluster. Other critical clusters are called *(local) maxima*, for obvious reasons. If a cluster G is a point of a sharp local maximum of the function D, that is, for any point $\mathbf{m}$ in a vicinity of G we have

$$D(\mathbf{m}) < D(G) ,$$

then G is rigid and thus not unlockable.

Example. The icosahedral cluster I_{12} of 12 equal balls gives an example of a maximal cluster. In 1943 Fejes-Tóth has shown that

(1) The maximum radius of 12 equal spheres touching a central sphere of radius 1 is

$$r_{max}(12) = \frac{1}{\sqrt{\frac{5+\sqrt{5}}{2}} - 1} \approx 1.1085085 .$$

(2) An extremal cluster achieving this radius is formed by the balls centered at 12 vertices of a regular icosahedron.

Examples of saddle clusters of balls

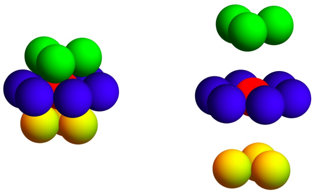


FIGURE – FCC cluster (left) and its layers (right)

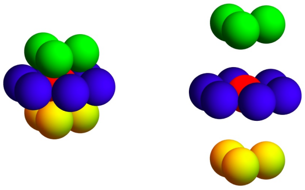


FIGURE – HCP cluster (left) and its layers (right)

The clusters FCC and HCP are saddle clusters, both of them can be unlocked and be deformed to the icosahedral cluster. For FCC the unlocking move was constructed by Coxeter. About HCP Fejes-Tóth writes : "Dagegen ist die andere doppelwabenartige Anordnung stabil" but it can be unlocked.

Questions about 6 cylinders

- Can 6 non-intersecting cylinders of radius $r > 1$ touch the unit ball ? The maximal value of r is unknown.
- Can the configurations C_6 and O_6 be unlocked ?
- Is the configuration space of six unit cylinders touching the unit ball connected ?

We have shown that the configuration C_6 is indeed unlockable. Namely, we have constructed its continuous deformation $C_6(t)$, along which quite a spacing opens between the cylinders, and at some value of t it becomes possible to arrange 6 non-intersecting cylinders of radius

$$r_{\mathfrak{m}} = \frac{1}{8}\left(3+\sqrt{33}\right) \approx 1.093070331.$$

The intermediate formulas are humongous and the fact that our minimax problem can be solved explicitly hints at a kind of integrability.

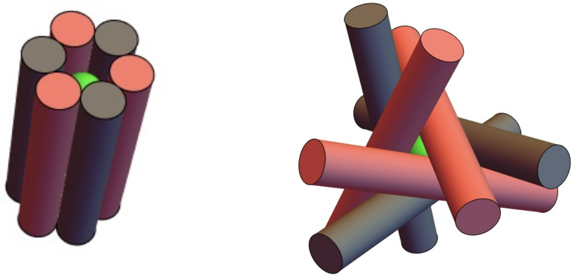
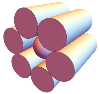
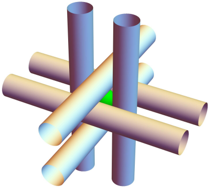


FIGURE – Two configurations of cylinders : the configuration C_6 of six parallel cylinders of radius 1 (on the left) and the configuration $C_{\mathfrak{m}}$ of six cylinders of radius $\approx$1.0931 (on the right)

We believe that our configuration of 6 cylinders with radius $r_{\mathfrak{m}}$ is in fact optimal. Here is the supporting argument.

Theorem. [O–S] Our cylinder arrangement, corresponding to the value $r_{\mathfrak{m}}$, is locally maximal, i.e. any small perturbation of our configuration decreases the corresponding radius.





Theorem. [O–S] The configuration O_6 is locally maximal : for any small perturbation of O_6 the corresponding radius is smaller than 1.

Corollary. The configuration space of six unit cylinders touching a unit ball is not connected.

In both situations, C_6 and O_6 one can check that the function D decreases along any direction. However this does not imply the maximality claim, as the following example shows.

In both situations, C_6 and O_6 one can check that the function D decreases along any direction. However this does not imply the maximality claim, as the following example shows.

Let f be a function of two variables defined by

$$f := \min\{u_1, u_2\} \text{ where } u_1 = -y + 3x^2, u_2 = y - x^2 .$$

The function f equals 0 at the origin. Consider an arbitrary ray l starting at the origin. Clearly, for some time this ray evades the 'horns' – the region between the parabolas $y = 3x^2$ and $y = x^2$. But outside the horns the function f is negative. Indeed, inside the the narrow parabola $y = 3x^2$ we have $u_1 < 0, u_2 > 0$ so f is negative there ; outside the wide parabola $y = x^2$ we have $u_1 > 0, u_2 < 0$ so f is negative there as well. Therefore the origin is a local maximum of f restricted to l, for any l. Yet the origin is not a local maximum of the function f on the plane, because inside the horns the functions u_1 and u_2 are positive so f there is positive as well.

Let $F_1(x), \ldots, F_m(x)$ be analytic functions in a neighborhood of $\mathbf{0} \in \mathbb{R}^n$ s.t. $F_u(\mathbf{0}) = 0$, $u = 1, \ldots, m$, and let $\mathrm{F}(x) := \min\{F_1(x), \ldots, F_m(x)\}$. Let us consider the differentials l_u and second differentials q_u of the functions F_u at $\mathbf{0} \in \mathbb{R}^n : F_u(x) = l_u(x) + q_u(x) + o(2)$; $o(2)$ stands for higher order terms.

We call the function $\Delta(x) := \min\{l_1(x), \ldots, l_m(x)\}$ the *PL*-differential of F. If some differentials do not vanish then the range of the differential Δ can be either a whole line $\mathbb{R}^1$, or the negative half-line. In the second case we say that $\mathbf{0} \in \mathbb{R}^n$ is a critical point of F, and that $0 \in \mathbb{R}^1$ is a critical value.

Lemma. Let $l_1, \ldots, l_m$ be linear functionals on $\mathbb{R}^n$. The two conditions are equivalent :

1. The function $\Delta(x) = \min_i l_i(x)$ is non-positive on $\mathbb{R}^n$.

2. There is a convex linear relation between l_i, i.e. for some $\lambda_1, \ldots, \lambda_r > 0$ and some $1 \leq i_1 < i_2 < \ldots < i_r \leq m$ we have $\lambda_1 l_{i_1} + \ldots + \lambda_r l_{i_r} = 0$.

We assume that the family $\{F_1(x), \ldots, F_m(x)\}$ of m analytic functions in n variables, $m \leq n$, possesses the following properties.

(A) The linear space, generated by the linear forms $l_1, \ldots, l_m$, is $(m-k)$ dimensional, with k positive.

(B) The collection $\{l_1, \ldots, l_m\}$ of linear forms can be split into k subcollections $\{l_1, \ldots, l_{m_1}\}$, $\{l_{m_1+1}, \ldots, l_{m_2}\}$, $\ldots$, $\{l_{m_{k-1}+1}, \ldots, l_m\}$ with non-intersecting spans, with **exactly one** linear relation between the functionals in each subcollection.

(C) For each $p = 1, \ldots, k$ the linear relation, from the property (B), between the functionals $\{l_{m_{p-1}+1}, \ldots, l_{m_{p+1}}\}$ is strictly convex :

$$\lambda_p^1 l_{m_{p-1}+1} + \ldots + \lambda_p^{m_p} l_{m_p} = 0 \ , \tag{2}$$

with $\lambda_p^s > 0$, $m_{p-1} + 1 \leq s \leq m_p$, $1 \leq p \leq k$.

(D) For

$$E_p = \bigcap_{u=m_{p-1}+1}^{m_p} \ker l_u \ , \ E = \bigcap_{p=1}^{k} E_p \ ,$$

and k quadratic forms Q_p, $1 \leq p \leq k$, defined by

$$Q_p = \lambda_p^1 q_{m_{p-1}+1} + \ldots + \lambda_p^{m_p} q_{m_p} \ , \tag{3}$$

the inequality

$$\min \{Q_1(\xi), ..., Q_k(\xi)\} \,|_{\xi \in E} \geq 0 \tag{4}$$

admits only the trivial solution $\xi = 0$.

Question : Classification, properties ?

Theorem. [O–S] Under the conditions (A) – (D), the origin $\mathbf{0} \in \mathbb{R}^n$ is a strict local maximum of the function $\mathsf{F}(x)$.

Note. If $m = 1$ we have the situation of a Morse function F_1. Indeed, k must be equal to 1 so the linear functional l_1 vanishes.

Both configurations, $C_{\mathfrak{m}}$ and O_6 fall into the above scenario and thus provide local maxima of the distance function.

Platonic configurations

Initially the configuration O_6 was called octahedral (wherefrom the capital O in the name), since the tangency points lie at the vertices of the regular octahedron :

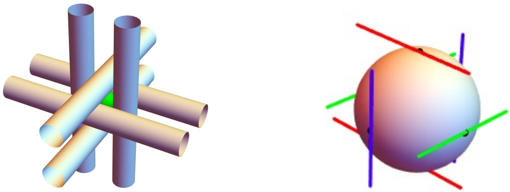


FIGURE – Octahedral configuration of tangent lines

But it should be rather called the tetrahedral configuration !

We start with the configuration of the tangent to the unit sphere lines which are continuations of the edges of the regular tetrahedron. The points of the sphere at which tangent lines pass are the edge middles of the regular tetrahedron. The initial position of the edges of the tetrahedron (in blue) :

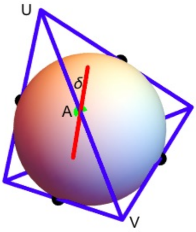


FIGURE – Sphere tangent to tetrahedron edges

Then each edge is rotated around the diameter of the unit sphere, passing through the middle of the edge, by an angle δ. When δ changes from 0 to $\pi/2$, the value of D^2 increases, achieves the maximal value 1 for $\delta_0 = \pi/4$, and afterwards decreases to 0. The point $\delta_0 = \pi/4$ is exactly the configuration O_6.

A similar construction can be performed for each pair of dual Platonic bodies. Namely, let a unit sphere touch the edge middles of a Platonic body $\mathcal{P}$. We can rotate all the edges of $\mathcal{P}$ around the axes passing through the center of the sphere and tangency points by the angle δ. When δ reaches the value $\pi/2$, the edges form the Platonic body dual to $\mathcal{P}$.

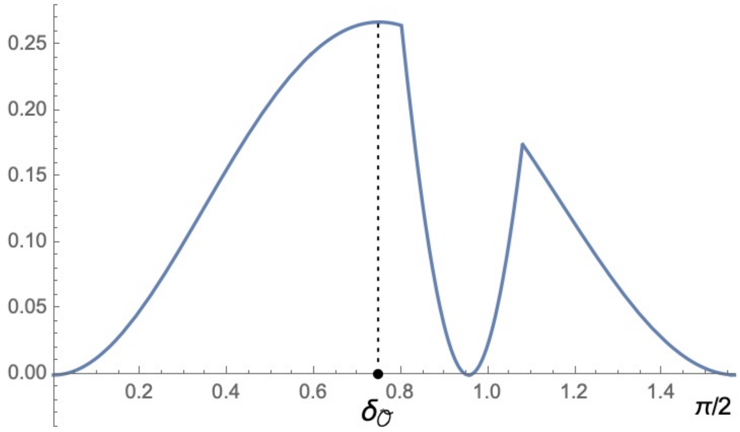


FIGURE – Graph of the square of the minimal distance for the pair octahedron-cube

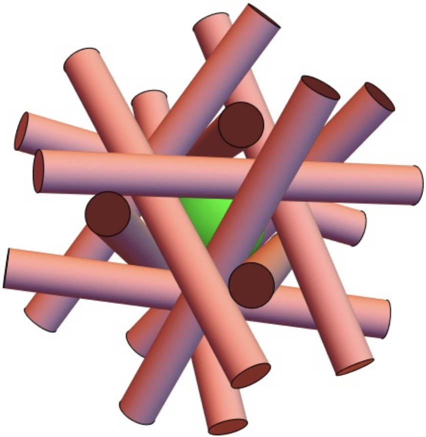


FIGURE – Octahedron/cube maximal configuration of cylinders, view from a vertex of the cube

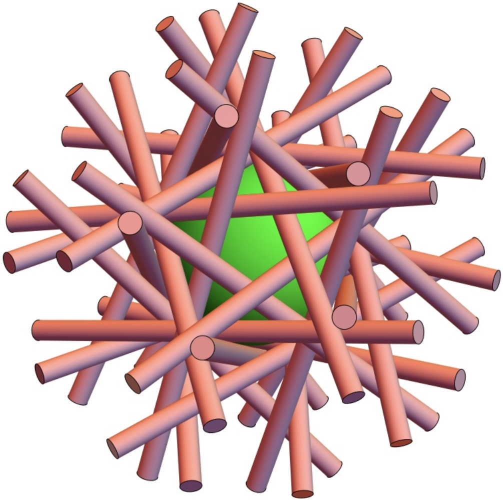


FIGURE – Maximal cluster, view from the tip of a 5-fold axis

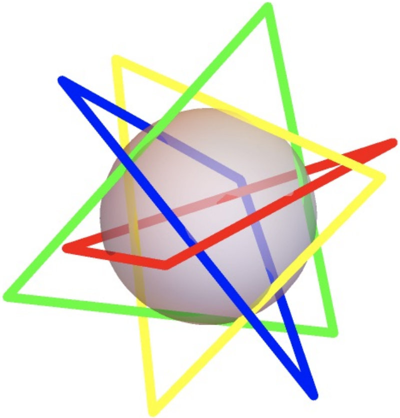


Figure – Octahedron/cube minimum

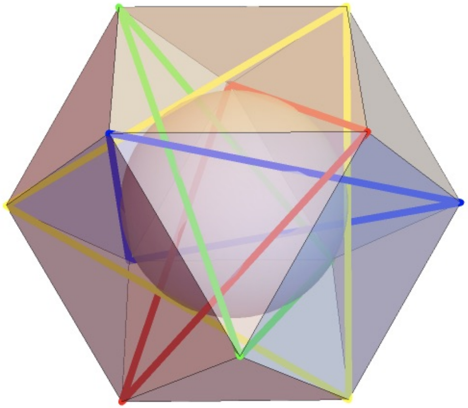


FIGURE – Octahedron/cube minimum, convex hull

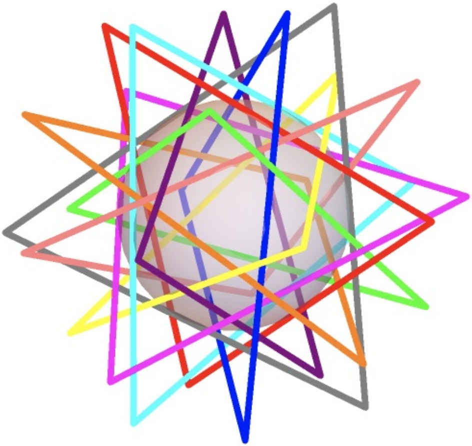


Figure – First minimum

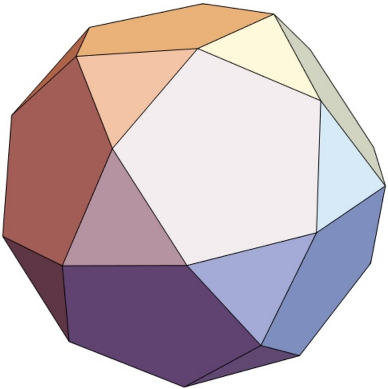


FIGURE – Icosidodecahedron

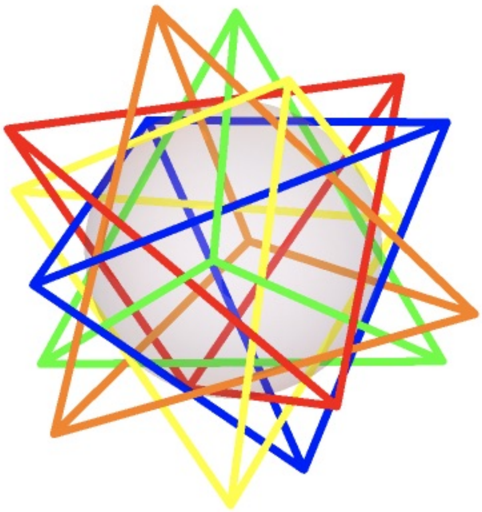


FIGURE – Second minimum

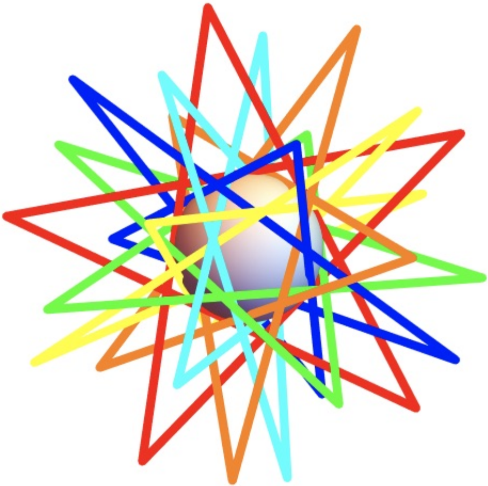


Figure – Third minimum

Let $\mathfrak{C}$ be a cluster of solid bodies $\Gamma_1, \ldots, \Gamma_L$, touching the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$. Then the solid bodies $\Gamma_j \times \mathbb{R}$, $j = 1, \ldots, L$, touch the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. We denote the so defined cluster by $\mathrm{K}(\mathfrak{C})$; this construction is due to Kuperberg, hence our notation. For example, if $\mathfrak{C}$ is a cluster of six unit discs touching the central unit disc, then $\mathrm{K}(\mathfrak{C})$ is the cluster C_6 of unit cylinders touching the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$.

Twelve unit spheres $\mathbb{S}^2$ can touch the central unit sphere in $\mathbb{R}^3$. Motivated by the fact that the configuration C_6 can be unlocked we believe that more and more space opens when we iterate the operation K. Therefore the following question arises : what is the minimal j such that thirteen bodies $\mathbb{S}^2 \times \mathbb{R}^j$ can touch the central unit sphere $\mathbb{S}^{2+j}$? Plausibly, $j = 1$.

We call a cluster G *flexible*, if it is not rigid, but during each deformation $G(t)$ some distances $\mathrm{dist}(\Lambda_i, \Lambda_j)$ which were zero at $t = 0$ remain zero at later moments t (at least up to a moment $t_0 > 0$ which might depend on the deformation $G(*)$).

Example of a flexible cluster : the configuration of four parallel non-intersecting cylinders of radius $r = 1 + \sqrt{2}$ touching the unit ball is not rigid.

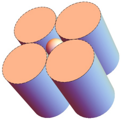


FIGURE – Initial position

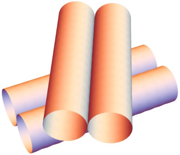


FIGURE – Motion of four cylinders

Open question. We believe that these are all possible positions of four cylinders of radius $r = 1 + \sqrt{2}$ touching the unit ball.