STUDENT SEMINAR - LECTURE ON SPHERE PACKING

OR LANDESBERG

1. Introduction

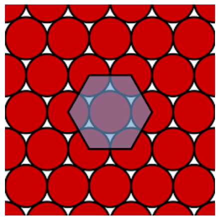
This talk will closely follow lecture notes by Henry Cohn [1] with some additions on coding [2] and Thue's theorem.

The sphere packing problem asks for the densest packing of congruent spheres (balls) in $\mathbb{R}^n$. A sphere packing in $\mathbb{R}^n$ is a collection $\mathcal{P}$ of balls of equal (unit) radius with non-intersecting interiors. We define the density of a packing (around 0) to be:

$$\Delta(\mathcal{P}) = \limsup_{R\to\infty} \frac{Vol\left(B_R(0) \cap \bigcup \mathcal{P}\right)}{Vol\left(B_R(0)\right)}$$

We define the optimal packing density of $\mathbb{R}^n$ as $\Delta_n = \sup_{\mathcal{P}} \Delta(\mathcal{P})$ where the supremum is taken over all sphere packings in $\mathbb{R}^n$.

It can easily be shown that a packing attaining this supremum always exists, but it is a theorem by Groemer that there exists an optimal packing for which the limit convergence is uniform around all points in $\mathbb{R}^n$.



What are the optimal sphere packings in low dimensions? For $n = 1$ one can quickly see that $\Delta_1 = 1$ where $\mathcal{P}$ is made of unit spheres with centers at $a + 2\mathbb{Z}$ for some $a \in \mathbb{R}$. For $n = 2$:

Theorem 1.1 (Thue 1910)**.** *The hexagonal sphere packing is the densest sphere packing in* $\mathbb{R}^2$*, with density* $\Delta_2 = \frac{\pi}{\sqrt{12}}$.

This short proof is due to Hales [3]. Let $\mathcal{H}$ be the hexagonal packing. Notice the plane can be decomposed w.r.t this packing into a collection of equilateral triangles of side length 2 connecting the centers of tangent balls. Let A be such a triangle, the area of A is equal $\sqrt{3}$ and around each of its vertices there is sixth of a circle contained inside A. Therefore the density of $\mathcal{H}$ is $\frac{\pi}{\sqrt{12}}$.

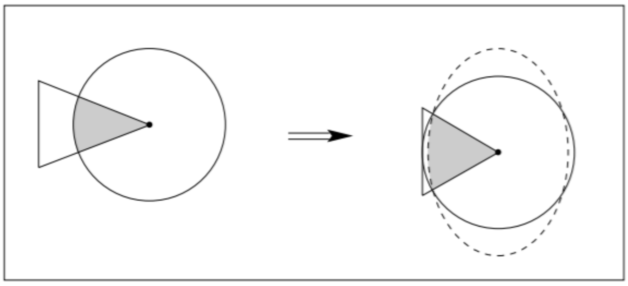
Now let $\mathcal{P}$ be any sphere packing in $\mathbb{R}^2$ with ball centers at $\{x_n\}_{n=1}^{\infty}$. Let $\tilde{\mathcal{P}}$ be the collection of balls with centers at $\{x_n\}$ and enlarged radius of $\frac{2}{\sqrt{3}}$. Notice that no three balls have intersecting interiors since the closest any three unit spheres can be is when forming an equilateral triangle of side length 2 between the vertices at their centers. In this case the enlarged balls would meet at exactly one intersection point (the circumcenter of the triangle). For any two intersecting balls in $\tilde{\mathcal{P}}$ draw a segment between the points of intersection of

Date: 22.3.17.

their boundaries. Then connect that segment to the centers of each ball to form two isosceles triangles. Partition the space into three regions:

- A Points outside all enlarged balls
- B Points inside enlarged balls and outside any triangle
- C Points inside a triangle

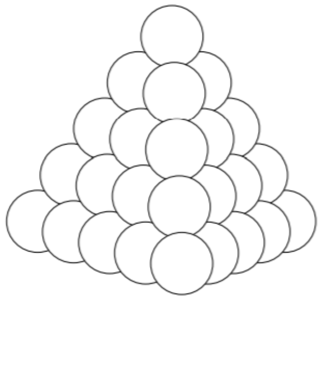
We will show the density of $\mathcal{P}$ in all three regions in less than $\frac{\pi}{\sqrt{12}}$. In region A, the density is clearly 0. In region B, the density is equal to the ratio of the area of a ball of radius 1 inside a ball of radius $\frac{2}{\sqrt{3}}$, this is $\frac{3}{4}$ which is smaller than $\frac{\pi}{\sqrt{12}}$. Let's focus on a triangle in region C, this triangle as two sides of length $\frac{2}{\sqrt{3}}$ with an angle of at most 60° between them. We are interested in the ratio between the area of the unit ball inside the triangle and the are of the triangle Notice that applying a linear transformation would not change the ratio between the resulting areas (just a constant multiple to both). By applying an appropriate transformation one can leave the side lengths fixed while changing the angle to exactly 60°, as in the picture below:



Since all radial lines between the two sides have been contracted we deduce the oval area inside the triangle (the image of the circle) is smaller than the area of a corresponding circle of unit radius, hence the density is smaller than $\frac{\pi}{\sqrt{12}}$ as required. $\square$

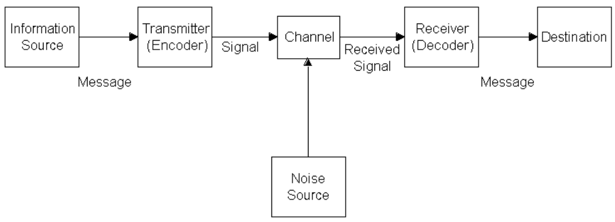
In dimension 3 things get a lot more complicated. It was Kepler who conjectured in 1611 that no packing packing of balls in $\mathbb{R}^3$ is denser than the cannonball packing, or the face-centered cubic packing, having density $\frac{\pi}{\sqrt{18}}$.

It was only around four hundred years later that this conjecture was proven by Thomas C. Hales with the help of a graduate student of his Samuel P. Ferguson and about two years of computation! In their proof the problem was reduced to calculating the maximal density of around 5,000 different configurations of spheres, each of which was analyzed. The outcome included more than 250 pages of notes and 3 gigabytes of code and data which have only been formally verified in 2014 after an additional 11 year long computation.



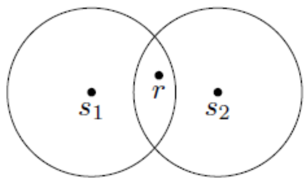
2. Error Correcting Codes

Asking for optimal packings in two and three dimensions is both natural and clearly useful (e.g. for material science) but what about higher dimensions? Is it useful to compute $\Delta_{10^{10}}$? Or understand $\liminf_{n\to\infty} \Delta_n$? Apparently so! These questions are closely related to the theory of communication over noisy channels, a discovery of Claude Shannon from 1948 (in his paper 'A Mathematical Theory of Communication'). Consider an information source feeding an encoder which transmits discrete code words through a noisy channel. At the other end of the channel there is a decoder which translates the received signal (including additional noise):



Every code word s is a point in $\mathbb{R}^n$ representing a set of distinct parameters of the signal being transmitted over a fixed time window T. Denote the sample rate of the code by $W = \frac{n}{T}$. The communication channel has one natural limitation which is an upper bound on the average energy of a code word, this being translated to a restriction that all code words be contained in a ball of fixed radius P around 0. We shall think of the noise as a random vector $u \in \mathbb{R}^n$ (usually normally distributed) added to the code word being transmitted. Hence the decoder receives at each time window the random vector $r = s + u$. Assume the noise has variance$= \sigma^2 \ll P$.

An error correcting code is a discrete set of code words $\mathcal{S} \subset B_P(0) \subset \mathbb{R}^n$ for which a decoder receiving the vector $r = s + u$ can uniquely identify s with high probability. Since $\|r - s\| \leqslant \varepsilon$ with high probability for any $\sigma^2 \leqslant \varepsilon$, such a code can be achieved when all code words in $\mathcal{S}$ are 2ε-separated, thus decreasing the probability of two words being received the same. Define the *bit-rate* of a code $\mathcal{S}$ as:



$$\rho(\mathcal{S}) = \frac{1}{T} \log_2 |\mathcal{S}|$$

Hence ε-sphere packings in $B_{P+\varepsilon}(0)$ lead to error correcting codes. And the denser the packing the higher the bit-rate. We shall now show the existence of "good" codes.

Definition 2.1. A saturated packing $\mathcal{P}$ in $\mathbb{R}^n$ is called *saturated* if every ball of unit radius in $\mathbb{R}^n$ intersects an element of $\mathcal{P}$. In other words, if no other unit ball can be added to the packing.

Clearly any sphere packing is contained in a saturated packing. In particular, there exist saturated sphere packings.

Proposition 2.1. *Every saturated packing in $\mathbb{R}^n$ has density at least 2^{-n}.*

Proof. Let $\mathcal{P}$ be a saturated sphere packing with ball centers $\{x_i\}$. Notice that every point in $\mathbb{R}^n$ is at distance strictly less than 2 from some x_i (otherwise $\mathcal{P}$ wouldn't be saturated). Hence enlarging the balls in $\mathcal{P}$ to radius 2 would lead to $\mathbb{R}^n = \bigcup_{i=1}^{\infty} B_2(x_i)$. Since $Vol(B_2(x)) = 2^n \cdot Vol(B_1(x))$, for every $R > 0$ we have:

$$Vol\left(B_R(0)\right) = Vol\left(B_R(0) \cap \bigcup_{i=1}^{\infty} B_2(x_i)\right) \leqslant \sum_{i=1}^{\infty} Vol\left(B_R(0) \cap B_2(x_i)\right) \leqslant$$

$$\leqslant \sum_{i=1}^{\infty} 2^n \cdot Vol\left(B_R(0) \cap B_1(x_i)\right) = 2^n \cdot Vol\left(B_R(0) \cap \bigcup \mathcal{P}\right)$$

Divide by $2^n \cdot Vol\left(B_R(0)\right)$ on both sides to receive desired estimate. $\square$

Corollary 2.2. *For any, P, ε, n, T there exists an error correcting code $\mathcal{S}$ in $\mathbb{R}^n$ with bit-rate greater than $W \cdot \log_2(1 + \frac{P}{\varepsilon}) - W$.*

Proof. Employing proposition 2.1 (and rescaling appropriately) one can construct an error correcting code $\mathcal{S}$ of size:

$$|\mathcal{S}| \geqslant 2^{-n} \cdot \frac{Vol(B_{P+\varepsilon}(0)))}{Vol(B_{\varepsilon}(x))} \geqslant 2^{-n} \cdot \left(\frac{P+\varepsilon}{\varepsilon}\right)^n$$

We receive an estimate of the bit-rate of $\mathcal{S}$:

$$\rho(\mathcal{S}) \geqslant \frac{1}{T}\log_2\left(2^{-n} \cdot \left(\frac{P+\varepsilon}{\varepsilon}\right)^n\right) = \frac{n}{T}\log_2\left(1 + \frac{P}{\varepsilon}\right) + \frac{-n}{T} =$$

$$= W \cdot \log_2\left(1 + \frac{P}{\varepsilon}\right) - W$$

$\square$

Remark.

(1) This is an over simplified and quite inaccurate model. Shannon and others achieve much stronger and more precise results.
(2) While an exponentially decaying lower bound $\Delta_n \geqslant 2^{-n}$ might seem very weak it is apparently not too far from reality. The best known upper bounds are of the order of $2^{-(0.5990+o(n))\cdot n}$ due to Kabatiansky and Levenshtein. One key to our lack of intuition with regards to high-dimensional geometry is the unintuitive fact that in high dimensions volume tends to accumulate at the boundary. Taking some region and shrinking it by a factor of $1-\varepsilon$ would yield a volume shrinkage of $(1-\varepsilon)^n$. Hence "holes" between spheres tend to be more significant and gluing of configurations along boundaries tends to be more subtle.
(3) Error correcting codes are often defined nowadays in a vector field over a finite field, e.g. $\mathbb{F}_2^n$. Such codes correspond in a similar manner to sphere packing problems defined using the Hamming metric (which counts the number coordinates two vectors disagree on).

3. Lattice Packings

The two optimal packings discussed in $\mathbb{R}^2, \mathbb{R}^3$ enjoy a great deal of symmetry. Is this always the case for optimal packings?
A sphere packing is called *periodic* if it consists of translated copies of some (non-trivial) finite configuration. It can easily be seen that periodic packings can get arbitrarily close to the optimal density (by copying $\Delta_n - \varepsilon$ dense bounded configurations), but there seems to be no reason to believe all optimal packings are periodic. Lattice packings have even more structure:

Definition 3.1. A *lattice* in $\mathbb{R}^n$ is a discrete subgroup $\Lambda \leqslant \mathbb{R}^n$ with finite co-volume, i.e. with $\mathbb{R}^n/\Lambda$ having finite volume.

In $\mathbb{R}^n$ this definition is equivalent to Λ being the integer span of n linearly independent vectors, $\Lambda = \{a_1 v_1 + ... + a_n v_n \mid a_1, ..., a_n \in \mathbb{Z}\}$. All currently proven optimal packings ($n = 1, 2, 3, 8, 24$) are lattice packings. This is not believed to be the case in general.

Define $l(\Lambda) = \min_{0 \neq v \in \Lambda} \|v\|$, the length of the smallest non-zero vector in the lattice Λ. Placing balls of radius $\frac{1}{2} l(\Lambda)$ at every lattice point gives a sphere packing, $\mathcal{P}_\Lambda$. Such packings are called lattice packings.

For any lattice $\Lambda = span_{\mathbb{Z}}\{v_1, ..., v_n\} \leqslant \mathbb{R}^n$ there exists a *fundamental parallelotope*:

$$\mathcal{F} = \{x_1 v_1 + ... + x_n v_n \mid x_i \in [0, 1)\}$$

with the property that any point in $\mathbb{R}^n$ is contained in $v + \mathcal{F}$ for exactly one $v \in \Lambda$, i.e.:

$$\mathbb{R}^n = \bigsqcup_{v \in \Lambda} v + \mathcal{F}$$

Notice that each vector in Λ is at the boundary of 2^n translates of $\mathcal{F}$. Accordingly, any ball in $\mathcal{P}_\Lambda$ is partitioned into 2^n regions each contained in a different translate of $\mathcal{F}$. Since $\mathcal{P}_\Lambda$ is invariant under translations by elements of Λ one can deduce that each $v + \mathcal{F}$ intersects 2^n complimenting parts of a single ball in the packing. Thus the density of a lattice packing can be easily computed to be:

$$\Delta(\mathcal{P}_\Lambda) = \frac{Vol\left(B_{\frac{1}{2}l(\Lambda)}(x)\right)}{Vol\,(\mathcal{F})}$$

Example 3.1. Taking the cubic lattice $\mathbb{Z}^n \leqslant \mathbb{R}^n$ yields a pretty sparse packing. $l(\mathbb{Z}^n) = \frac{1}{2}$ and $Vol(\mathcal{F}_{\mathbb{Z}^n}) = Vol([0, 1]^n) = 1$ and hence:

$$\Delta(\mathcal{P}_{\mathbb{Z}^n}) = \frac{\pi^{\frac{n}{2}}}{(\frac{n}{2})! \cdot 2^n}$$

where for odd dimensions $(\frac{n}{2})!$ means $\Gamma(1 + \frac{n}{2})$. This packing is asymptoticly far inferior to the exponential lower bound of proposition 2.1.

Example 3.2. For $n \geqslant 3$ the "checkerboard" lattice gives better densities:

$$D_n = \{(x_1, x_2, ..., x_n) \in \mathbb{Z}^n \mid x_1 + x_2 + ... + x_n \in 2\mathbb{Z}\}$$

The fundamental parallelotope of D_n has volume:

$$Vol\left(\mathcal{F}_{D_n}\right) = \det \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & (-1)^{n+1} \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{bmatrix} = 2$$

while $l(D_n) = \sqrt{2}$. Therefore:

$$\Delta\left(\mathcal{P}_{D_n}\right) = \frac{Vol\left(B_{\frac{\sqrt{2}}{2}}(0)\right)}{Vol\left(\mathcal{F}_{D_n}\right)} = \frac{\left(\pi/2\right)^{\frac{n}{2}}}{2 \cdot \left(\frac{n}{2}\right)!}$$

Actually D_3, D_4, D_5 are the best packings known in their dimensions.

Let's explore the gaps or holes between the spheres in the D_n lattice packing. A point $x \in \mathbb{R}^n$ is considered a hole in a packing if it is a local maximum for the distance function to the nearest sphere center. Notice that in the case of lattice packings the holes are invariant under translations by elements of the lattice. For D_n there are two kinds of holes:

- 'shallow' holes - translates of $(1, 0, ..., 0)$ which have a distance of 1 from the nearest lattice point.
- 'deep' holes - translates of $(\frac{1}{2}, ..., \frac{1}{2}, \pm\frac{1}{2})$ which have a distance of $\sqrt{\sum\left(\frac{1}{2}\right)^2} = \sqrt{\frac{n}{4}}$ (these are actually 'deeper' only when $n > 4$).

As one can see the deep holes become deeper as n increases. An interesting transition happens at $n = 8$ where the distance of the deep holes from the nearest lattice point is $\sqrt{2}$ which is exactly twice the radius of the balls in the packing! For $n = 8$ one can add another translate of D_8 at the deep holes and instantaneously double the density! This new packing is coincidently a lattice itself and is called the E_8 root lattice. In 2016, a Ukrainian mathematician named Maryna Viazovska proved E_8 is the densest packing in $\mathbb{R}^8$. Her proof was relatively short (23 pages).

Remark.

(1) It is amusing to try and imagine the 2-cube $[-1, 1]^n$ in high dimensions. Although all edges of the cube have length 2, the distance of the vertices to the origin is huge, e.g. $d(0, (1, 1, ..., 1)) = \sqrt{n}$. In some sense this is why it's very easy to pack cubes rather than spheres.

(2) A quick remark about the Leech lattice Λ_{24}, the optimal lattice in $\mathbb{R}^{24}$. Its construction uses an error correcting code called the binary Golay code which was discovered in 1949. It allows a 12 bit string to be encoded as a 24 bit string with error of up to 3 bits being correctable. This code was used for communications with the Voyager 1&2 spacecrafts for sending color pictures of Jupiter and Saturn. The lattice is also used to give an explicit construction of the Monster Group the largest sporadic simple group of size $\approx 8 \cdot 10^{53}$. The density of Λ_{24}

is $\frac{\pi^{12}}{12!}$ and was shown to be optimal in 2016 (by Cohn, Kumar, Miller et al.).

The Space of Unimdular Lattices. A lattice $\Lambda \leqslant \mathbb{R}^n$ is called *unimodular* if ${}^{\mathbb{R}^n}\!/\!_{\Lambda}$ has volume 1, or equivalently if the fundamental parallelotope has volume 1. Notice that any lattice can be mapped to a unimodular lattice using scalar multiplication, a mapping which does not effect the density of the corresponding sphere packing. Any unimodular lattice $\Lambda = \{a_1v_1 + ... + a_nv_n \mid a_i \in \mathbb{Z}\}$ can be presented as $A \cdot \mathbb{Z}^n$ where $A \in \mathrm{SL}_n(\mathbb{R})$ is the matrix with column vectors $(v_1, ..., v_n)$. Hence we have a transitive action of the group $\mathrm{SL}_n(\mathbb{R})$ on the set of all unimodular lattices. One can easily see that the stabilizer of the cubic lattice $\mathbb{Z}^n$ is the subgroup $\mathrm{SL}_n(\mathbb{Z})$ (verify this is actually a group).
The space of n-dimensional unimodular lattices X_n can thus be identified with the quotient space:

$$X_n \cong {}^{\mathrm{SL}_n(\mathbb{R})}\!/\!_{\mathrm{SL}_n(\mathbb{Z})}$$

One can show that there exists an $\mathrm{SL}_n(\mathbb{R})$ invariant probability measure m on X_n (or equivalently that $\mathrm{SL}_n(\mathbb{Z})$ is a lattice inside $\mathrm{SL}_n(\mathbb{R})$). The invariance of m means that for any $A \in \mathrm{SL}_n(\mathbb{R})$ and any measurable function f on X_n:

$$\int_{X_n} f(A \cdot \Lambda) dm(\Lambda) = \int_{X_n} f(\Lambda) dm(\Lambda)$$

The probability measure m allows us to talk about random lattices in $\mathbb{R}^n$. This is apparently useful for achieving a lower bound for the densest lattice packing.

Theorem 3.3 (Siegel Mean Value Theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a bounded Borel measurable function of compact support, then:*

$$\int_{X_n} \sum_{v \in \Lambda \backslash \{0\}} f(v) dm(\Lambda) = \int_{\mathbb{R}^n} f(x) dx$$

In other words, averaging over all unimodular lattices is the same as averaging over $\mathbb{R}^n$ itself. The idea behind the proof of this theorem is that the measure defined by the LHS of the equation above is a locally finite (Radon) $\mathrm{SL}_n(\mathbb{R})$-invariant measure on $\mathbb{R}^n$. The only possible measures are constant multiples of Lebesgue measure and this constant happens to be 1.

Corollary 3.4. *There exists a* ***lattice*** *sphere packing in $\mathbb{R}^n$ with density at least $2 \cdot 2^{-n}$.*

Proof. Let B be a ball of volume 2 in $\mathbb{R}^n$ centered at the origin. By applying the Siegel mean value theorem to the characteristic function of B one can deduce that the average number of non-zero lattice points contained in B is equal $Vol(B) = 2$. Note that since B is symmetric (w.r.t to negation $x \mapsto -x$) lattice points in B always come in pairs. A positive measure of lattices have vectors shorter than $\frac{1}{4} \cdot diam(B)$ and thus have more than two non-zero lattice point contained in B. Since the average number is 2 that means there exist lattices with 0 such points. Attaching translates of $\frac{1}{2}B$ at each vertex of such a lattice gives a packing with a copy of one ball per unit volume and density:

$$Vol\left(\frac{1}{2}B\right) = 2^{-n} \cdot Vol(B) = 2 \cdot 2^{-n}$$

as required. $\square$

This corollary improves the previous result both by a factor of 2 and by assuring it is achieved by a lattice packing.

References

[1] H. Cohn, "Packing, coding, and ground states." https://arxiv.org/abs/1603.05202, 2016.

[2] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, vol. 290 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, third ed., 1999. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.

[3] T. C. Hales, "Cannonballs and honeycombs," *Notices Amer. Math. Soc.*, vol. 47, no. 4, pp. 440–449, 2000.