### **Overview and motivation**

#### **Danylo Radchenko**

March 3, 2021

#### Problem (The sphere packing problem)

What is the best way to pack equal spheres in  $\mathbb{R}^d$ ?

#### Problem (The sphere packing problem)

What is the best way to pack equal spheres in  $\mathbb{R}^d$ ?

More precisely, we are interested in covering as large a portion of the space as possible by non-overlapping spheres (or balls) of radius R.

#### Problem (The sphere packing problem)

What is the best way to pack equal spheres in  $\mathbb{R}^d$ ?

More precisely, we are interested in covering as large a portion of the space as possible by non-overlapping spheres (or balls) of radius R.

A precise formulation of this problem requires some care (for example in defining what exactly is the portion of the space covered by spheres).

The first known formulation of the sphere packing problem (in 3 dimensions) goes back to Johannes Kepler's manuscript "The Six-Cornered Snowflake" from 1611. His interest in arrangements of spheres came from his correspondence with an English astronomer Thomas Harriot. Harriot himself was studying the problem of how to best stack cannonballs at the behest of Sir Walter Raleigh. The first known formulation of the sphere packing problem (in 3 dimensions) goes back to Johannes Kepler's manuscript "The Six-Cornered Snowflake" from 1611. His interest in arrangements of spheres came from his correspondence with an English astronomer Thomas Harriot. Harriot himself was studying the problem of how to best stack cannonballs at the behest of Sir Walter Raleigh.









Kepler has conjectured that the best way to pack spheres in three dimensions is achieved by putting layers of hexagonal lattices on top of each other.

### The Kepler Conjecture



Kepler has conjectured that the best way to pack spheres in three dimensions is achieved by putting layers of hexagonal lattices on top of each other.

The Kepler conjecture was proved in 1998 by Thomas Hales in a computer-assisted tour de force proof by exhaustion following some earlier ideas due to Fejes Tóth.

### The Kepler Conjecture



Kepler has conjectured that the best way to pack spheres in three dimensions is achieved by putting layers of hexagonal lattices on top of each other.

The Kepler conjecture was proved in 1998 by Thomas Hales in a computer-assisted tour de force proof by exhaustion following some earlier ideas due to Fejes Tóth.

A formal proof verification using automated proof checking software was completed in 2014.











hexagonal close-packed



face-centered cubic

d	Configuration	Proof
1	$\mathbb{Z}$	trivial

d	Configuration	Proof
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)

d	Configuration	Proof
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)
3	fcc, hcp,	Hales (1998/2014)

-		
d	Configuration	Proof
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)
3	fcc, hcp,	Hales (1998/2014)
8	$E_8$	Viazovska (2016)

-

d	Configuration	Proof
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)
3	fcc, hcp,	Hales (1998/2014)
8	$E_8$	Viazovska (2016)
24	$\Lambda_{24}$	Cohn-Kumar-Miller-RViazovska (2016)

d	Configuration	Proof
	comparation	11001
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)
3	<i>fcc</i> , <i>hcp</i> ,	Hales (1998/2014)
8	$E_8$	Viazovska (2016)
24	$\Lambda_{24}$	Cohn-Kumar-Miller-RViazovska (2016)

There are also reasonable conjectures for what the answer should be for d = 4, 5, 6, 7.

d	Configuration	Proof
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)
3	fcc, hcp,	Hales (1998/2014)
8	$E_8$	Viazovska (2016)
24	$\Lambda_{24}$	Cohn-Kumar-Miller-RViazovska (2016)

There are also reasonable conjectures for what the answer should be for d = 4, 5, 6, 7.

For  $d \le 8$  the best packing should be achieved by a certain root lattice. These arise in the classification of finite-dimensional simple Lie algebras.

d	Configuration	Proof
1	$\mathbb{Z}$	trivial
2	$A_2$	Thue (1890), Fejes Tóth (1940)
3	fcc, hcp,	Hales (1998/2014)
8	$E_8$	Viazovska (2016)
24	$\Lambda_{24}$	Cohn-Kumar-Miller-RViazovska (2016)

There are also reasonable conjectures for what the answer should be for d = 4, 5, 6, 7.

For  $d \le 8$  the best packing should be achieved by a certain root lattice. These arise in the classification of finite-dimensional simple Lie algebras.

The Leech lattice is of a different nature and it was first constructed from a certain exceptional error-correcting code.

In 1694 Isaac Newton and David Gregory have discussed the following question.

#### Question

Can a sphere touch 13 spheres of the same size?

In 1694 Isaac Newton and David Gregory have discussed the following question.

#### Question

Can a sphere touch 13 spheres of the same size?

Newton believed that the answer is "no", while Gregory believed that it is "yes".

In 1694 Isaac Newton and David Gregory have discussed the following question.

#### Question

Can a sphere touch 13 spheres of the same size?

Newton believed that the answer is "no", while Gregory believed that it is "yes".



In 1694 Isaac Newton and David Gregory have discussed the following question.

#### Question

Can a sphere touch 13 spheres of the same size?

Newton believed that the answer is "no", while Gregory believed that it is "yes".



One can easily arrange 12 spheres to touch a given sphere, but there is enough room to move them around.

### A configuration of 12 spheres touching another one



In general, one can ask the same question in d dimensions.

Problem (The kissing number problem)

What is the maximal number of unit spheres that can touch a given unit sphere in  $\mathbb{R}^d$ ?

In general, one can ask the same question in d dimensions.

Problem (The kissing number problem)

What is the maximal number of unit spheres that can touch a given unit sphere in  $\mathbb{R}^d$ ?

This is a spherical analogue of the sphere packing problem.

In general, one can ask the same question in d dimensions.

Problem (The kissing number problem)

What is the maximal number of unit spheres that can touch a given unit sphere in  $\mathbb{R}^d$ ?

This is a spherical analogue of the sphere packing problem.

A slightly different formulation is that we are interested in finding the maximal value of N for which there exists points  $x_1, \ldots, x_N \in \mathbb{S}^{d-1}$  satisfying

 $\langle x_i, x_j \rangle \leq 1/2$  for all  $i \neq j$ 

Any such set  $\{x_1, \ldots, x_N\}$  is called a **kissing configuration** and the maximal value of N (for given d) is called the **kissing number**.

d	kissing number	Proof
1	2	trivial

d	kissing number	Proof
1	2	trivial
2	6	easy

d	kissing number	Proof
1	2	trivial
2	6	easy
3	12	Schütte and van der Waerden (1953), Leech (1956)

d	kissing number	Proof
1	2	trivial
2	6	easy
3	12	Schütte and van der Waerden (1953), Leech (1956)
4	24	Musin (2003)

d	kissing number	Proof
1	2	trivial
2	6	easy
3	12	Schütte and van der Waerden (1953), Leech (1956)
4	24	Musin (2003)
8	240	Levenshtein (1979)

d	kissing number	Proof
1	2	trivial
2	6	easy
3	12	Schütte and van der Waerden (1953), Leech (1956)
4	24	Musin (2003)
8	240	Levenshtein (1979)
24	196560	Odlyzko and Sloane (1979)
d	kissing number	Proof
----	----------------	--
1	2	trivial
2	6	easy
3	12	Schütte and van der Waerden (1953), Leech (1956)
4	24	Musin (2003)
8	240	Levenshtein (1979)
24	196560	Odlyzko and Sloane (1979)

With the exception of d = 4 this table mirrors that for the sphere packing problem.

d	kissing number	Proof
1	2	trivial
2	6	easy
3	12	Schütte and van der Waerden (1953), Leech (1956)
4	24	Musin (2003)
8	240	Levenshtein (1979)
24	196560	Odlyzko and Sloane (1979)

With the exception of d = 4 this table mirrors that for the sphere packing problem.

For d = 8 and d = 24 the best kissing configuration is unique and comes from the exceptional lattices ( $E_8$  and the Leech lattice).

#### Definition (Spherical codes)

Given a subset  $S \subseteq [-1,1)$ , we call a set of points  $\{x_1, \ldots, x_N\} \subseteq \mathbb{S}^{d-1}$  a spherical *S*-code, if one has

 $\langle x_i, x_j \rangle \in S$  for all  $i \neq j$ 

#### Definition (Spherical codes)

Given a subset  $S \subseteq [-1,1)$ , we call a set of points  $\{x_1, \ldots, x_N\} \subseteq \mathbb{S}^{d-1}$  a spherical S-code, if one has

 $\langle x_i, x_j \rangle \in S$  for all  $i \neq j$ 

If we take S = [-1, 1/2], we recover the notion of a kissing configuration.

#### Definition (Spherical codes)

Given a subset  $S \subseteq [-1,1)$ , we call a set of points  $\{x_1, \ldots, x_N\} \subseteq \mathbb{S}^{d-1}$  a spherical S-code, if one has

$$\langle x_i, x_j \rangle \in S$$
 for all  $i \neq j$ 

If we take S = [-1, 1/2], we recover the notion of a kissing configuration.

Similarly to the kissing number problem one can ask what is the maximal size of a spherical S-code.

An interesting case arises already when |S| = k is finite. Then spherical codes are k-distance sets, i.e. there are only k distinct distances that occur between  $x_i$ 's.

## Spherical codes and spherical designs

An interesting case arises already when |S| = k is finite. Then spherical codes are k-distance sets, i.e. there are only k distinct distances that occur between  $x_i$ 's.

Delsarte, Goethals and Seidel have found a bound on the size of spherical S-codes that depends only on k = |S|. They also found that the same bound serves as a lower bound on the size of spherical designs.

#### Definition (Spherical designs)

A spherical t-design is a set  $\{x_1,\ldots,x_N\}\subseteq\mathbb{S}^{d-1}$  such that

$$\frac{1}{N}\sum_{i=1}^{N}p(x_i)=\int_{\mathbb{S}^{d-1}}p(x)d\mu(x)$$

for all polynomials p of degree  $\leq t$ .

As it turns out, a spherical *S*-code that achieves the Delsarte-Goethals-Seidel bound must also be a strong spherical design. Such configurations that achieve these bounds are called **tight spherical designs**.

As it turns out, a spherical *S*-code that achieves the Delsarte-Goethals-Seidel bound must also be a strong spherical design. Such configurations that achieve these bounds are called **tight spherical designs**.

It turns out that the shortest vectors of the  $E_8$  lattice and the Leech lattice are particularly nice tight designs. This fact is also related to their optimality as kissing configurations.

There is a discrete side to the story, which surprisingly has to do with the mathematical theory of communication as laid out by Claude Shannon in 1948.

There is a discrete side to the story, which surprisingly has to do with the mathematical theory of communication as laid out by Claude Shannon in 1948. Very roughly, the basic problem is how well can one communicate a message over a channel with noise that can distort the signal.

There is a discrete side to the story, which surprisingly has to do with the mathematical theory of communication as laid out by Claude Shannon in 1948. Very roughly, the basic problem is how well can one communicate a message over a channel with noise that can distort the signal. For simplicity, one can think of a signal as a sequence of bits, and noise can, independently and with small probability, flip some of the bits.

There is a discrete side to the story, which surprisingly has to do with the mathematical theory of communication as laid out by Claude Shannon in 1948. Very roughly, the basic problem is how well can one communicate a message over a channel with noise that can distort the signal.

For simplicity, one can think of a signal as a sequence of bits, and noise can, independently and with small probability, flip some of the bits.





### Error-correcting codes

A very simple way in which one can control the errors over a noisy channel is to group the bits into blocks, and add to each block an additional bit equal to the parity of the sum in the block.

### Error-correcting codes

A very simple way in which one can control the errors over a noisy channel is to group the bits into blocks, and add to each block an additional bit equal to the parity of the sum in the block.

This allows one to detect if a single error has occurred in a block, but does not allow one to find which bit was flipped.

### Error-correcting codes

A very simple way in which one can control the errors over a noisy channel is to group the bits into blocks, and add to each block an additional bit equal to the parity of the sum in the block.

This allows one to detect if a single error has occurred in a block, but does not allow one to find which bit was flipped.

It turns out, that one can add more control bits to also be able to detect where the error has occurred and automatically fix it. That this is possible was realized by Richard Hamming, and he also constructed the first error-correcting code: the Hamming (7,4) code.



$$d(x,y) = \#\{1 \le i \le n \colon x_i \neq y_i\}$$

$$d(x,y) = \#\{1 \le i \le n \colon x_i \neq y_i\}$$

#### Definition (Linear codes)

An [n, k, d]-binary linear code is a linear subspace  $V \subseteq \mathbb{F}_2^n$  of dimension k such that any two distinct vectors in V have Hamming distance  $\geq d$ .

$$d(x,y) = \#\{1 \le i \le n \colon x_i \neq y_i\}$$

#### Definition (Linear codes)

An [n, k, d]-binary linear code is a linear subspace  $V \subseteq \mathbb{F}_2^n$  of dimension k such that any two distinct vectors in V have Hamming distance  $\geq d$ .

With an [n, k, d]-code one can send blocks of n bits encoding k bits of information in a way that can detect up to d - 1 errors and correct up to (d - 1)/2 errors.

$$d(x,y) = \#\{1 \le i \le n \colon x_i \neq y_i\}$$

#### Definition (Linear codes)

An [n, k, d]-binary linear code is a linear subspace  $V \subseteq \mathbb{F}_2^n$  of dimension k such that any two distinct vectors in V have Hamming distance  $\geq d$ .

With an [n, k, d]-code one can send blocks of n bits encoding k bits of information in a way that can detect up to d - 1 errors and correct up to (d - 1)/2 errors.

Hamming's (7,4)-code is a binary linear [7,4,3]-code.

One can easily recognize that error-correcting codes are analogues of sphere packings for the Hamming distance. Moreover, one can construct (good) sphere packings from (good) codes.

One can easily recognize that error-correcting codes are analogues of sphere packings for the Hamming distance. Moreover, one can construct (good) sphere packings from (good) codes.

There are several ways of doing this, the easiest is to embed the codewords as elements of the set  $\{0,1\}^n \subset \mathbb{R}^n$  and extend to a  $(2\mathbb{Z})^n$ -periodic set.

There is a particularly interesting binary linear code that was discovered by Marcel J. E. Golay in 1949. This is the so-called **extended binary Golay code** with parameters [24, 12, 8].

There is a particularly interesting binary linear code that was discovered by Marcel J. E. Golay in 1949. This is the so-called **extended binary Golay code** with parameters [24, 12, 8].

It has several remarkable features:

- It is the unique binary code with parameters [24, 12, 8]
- Its automorphism group is a sporadic simple group (the Mathieu group  $M_{24}$ )
- If one removes the last coordinate one obtains a perfect [23, 12, 7] binary code (Hamming balls of radius 3 cover 𝔽<sup>23</sup><sub>2</sub>)

Linear codes are analogues to lattices in the Euclidean case. Steiner systems could be though of as analogues of spherical codes and spherical designs.

Linear codes are analogues to lattices in the Euclidean case. Steiner systems could be though of as analogues of spherical codes and spherical designs.

#### Definition (Steiner system)

A Steiner system S(n, k, l) is a set S of k-subsets of  $X = \{1, ..., n\}$  such that any *l*-subset of X is contained in exactly one element of S.

Linear codes are analogues to lattices in the Euclidean case. Steiner systems could be though of as analogues of spherical codes and spherical designs.

#### Definition (Steiner system)

A Steiner system S(n, k, l) is a set S of k-subsets of  $X = \{1, ..., n\}$  such that any *l*-subset of X is contained in exactly one element of S.

As a (rather exceptional) example: a Steiner system S(24, 8, 5) is given by the codewords of the extended binary Golay code with eight 1's.

In 1965 John Leech has constructed a remarkable lattice in 24 dimensions using the [24, 12, 8]-Golay code. Leech was interested in constructing good sphere packings in small dimensions, and it is now known that his lattice in fact gives an optimal packing in 24 dimensions.

In 1965 John Leech has constructed a remarkable lattice in 24 dimensions using the [24, 12, 8]-Golay code. Leech was interested in constructing good sphere packings in small dimensions, and it is now known that his lattice in fact gives an optimal packing in 24 dimensions.



In 1965 John Leech has constructed a remarkable lattice in 24 dimensions using the [24, 12, 8]-Golay code. Leech was interested in constructing good sphere packings in small dimensions, and it is now known that his lattice in fact gives an optimal packing in 24 dimensions.



Curiously, the construction of this lattice was not given much special prominence, and its description occupied less than a page in a supplement to his 1964 paper.

# Conway groups

Leech was interested in finding the symmetries of his lattice, but was unable to answer this question.

## Conway groups

Leech was interested in finding the symmetries of his lattice, but was unable to answer this question.

The group of symmetries of the Leech lattice was analyzed by J. H. Conway in 1968, and as a result he found three new sporadic simple groups that now bear his name.



### Conway groups

Leech was interested in finding the symmetries of his lattice, but was unable to answer this question.

The group of symmetries of the Leech lattice was analyzed by J. H. Conway in 1968, and as a result he found three new sporadic simple groups that now bear his name.



Sporadic simple groups are a finite list of 26 (or 27) groups that occupy an exceptional place in the classification of finite simple groups (among cyclic groups, alternating groups, and 16 families of groups of Lie type).

The final topic we will discuss in the seminar is modular forms. Modular forms are certain special analytic functions on  $\mathbb{H} = \{z \in \mathbb{C} : \text{Im } z > 0\}.$ 

The final topic we will discuss in the seminar is modular forms. Modular forms are certain special analytic functions on  $\mathbb{H} = \{z \in \mathbb{C} : \text{ Im } z > 0\}$ . We will be interested in modular forms for  $SL_2(\mathbb{Z})$ ; these are analytic functions  $f : \mathbb{H} \to \mathbb{C}$  that satisfy

$$f\left(rac{a au+b}{c au+d}
ight)=(c au+d)^kf( au)\,,\qquad egin{pmatrix} a&b\c&d\end{pmatrix}\in\mathsf{SL}_2(\mathbb{Z})$$

plus some growth conditions. Here  $k \in \mathbb{Z}$  is called the **weight** of the modular form f.
The final topic we will discuss in the seminar is modular forms. Modular forms are certain special analytic functions on  $\mathbb{H} = \{z \in \mathbb{C} : \text{Im } z > 0\}$ . We will be interested in modular forms for  $SL_2(\mathbb{Z})$ ; these are analytic functions  $f : \mathbb{H} \to \mathbb{C}$  that satisfy

$$f\left(rac{a au+b}{c au+d}
ight)=(c au+d)^kf( au)\,,\qquad egin{pmatrix} a&b\c&d\end{pmatrix}\in\mathsf{SL}_2(\mathbb{Z})$$

plus some growth conditions. Here  $k \in \mathbb{Z}$  is called the **weight** of the modular form f. Modular forms are a very deep subject with many interesting number-theoretic connections. The final topic we will discuss in the seminar is modular forms. Modular forms are certain special analytic functions on  $\mathbb{H} = \{z \in \mathbb{C} : \text{Im } z > 0\}$ . We will be interested in modular forms for  $SL_2(\mathbb{Z})$ ; these are analytic functions  $f : \mathbb{H} \to \mathbb{C}$  that satisfy

$$f\left(rac{a au+b}{c au+d}
ight)=(c au+d)^kf( au)\,,\qquad egin{pmatrix} a&b\c&d\end{pmatrix}\in\mathsf{SL}_2(\mathbb{Z})$$

plus some growth conditions. Here  $k \in \mathbb{Z}$  is called the **weight** of the modular form f.

Modular forms are a very deep subject with many interesting number-theoretic connections. A reason for this is a combination of two facts

- many "interesting" arithmetic sequences are encoded by modular forms
- spaces of modular forms are finite-dimensional and "computable"

To give an example, let us consider the sum of k squares function

$$r_k(n) = \#\{(a_1, \ldots a_k) \in \mathbb{Z}^k : a_1^2 + \cdots + a_k^2 = n\}$$

To give an example, let us consider the sum of k squares function

$$r_k(n) = \#\{(a_1, \ldots a_k) \in \mathbb{Z}^k : a_1^2 + \cdots + a_k^2 = n\}$$

Then the generating series

$$\theta^k(\tau) = \sum_{n\geq 0} r_k(n) q^n$$

becomes a modular form of weight k/2 if we set  $q = e^{2\pi i \tau}$  (though not for the full group  $SL_2(\mathbb{Z})$ ). This is a simplest example of a theta function.

To give an example, let us consider the sum of k squares function

$$r_k(n) = \#\{(a_1, \ldots a_k) \in \mathbb{Z}^k : a_1^2 + \cdots + a_k^2 = n\}$$

Then the generating series

$$heta^k( au) = \sum_{n\geq 0} r_k(n) q^n$$

becomes a modular form of weight k/2 if we set  $q = e^{2\pi i \tau}$  (though not for the full group  $SL_2(\mathbb{Z})$ ). This is a simplest example of a theta function.

As an application: knowing a basis for the corresponding space of modular forms for k = 4 allows one to deduce the following identity (due to Jacobi)

$$r_4(n) = 8 \sum_{4 \nmid d \mid n,} d, \quad n \geq 1.$$

To give an example, let us consider the sum of k squares function

$$r_k(n) = \#\{(a_1, \ldots a_k) \in \mathbb{Z}^k : a_1^2 + \cdots + a_k^2 = n\}$$

Then the generating series

$$heta^k( au) = \sum_{n\geq 0} r_k(n) q^n$$

becomes a modular form of weight k/2 if we set  $q = e^{2\pi i \tau}$  (though not for the full group  $SL_2(\mathbb{Z})$ ). This is a simplest example of a theta function.

As an application: knowing a basis for the corresponding space of modular forms for k = 4 allows one to deduce the following identity (due to Jacobi)

$$r_4(n) = 8 \sum_{4 \nmid d \mid n,} d, \quad n \geq 1.$$

This implies Lagrange's four-square theorem  $(r_4(n) > 0)$ .

In this seminar they will appear in two quite different roles:

In this seminar they will appear in two quite different roles:

- through classical theory of theta functions, as generating functions encoding the distance distribution of a lattice
- in construction of optimal functions for the linear programming bound as they have appeared in the work of Viazovska

In this seminar they will appear in two quite different roles:

- through classical theory of theta functions, as generating functions encoding the distance distribution of a lattice
- in construction of optimal functions for the linear programming bound as they have appeared in the work of Viazovska

These topics will be the subject of the last four talks of the seminar.