Sphere packing and the magic dimensions 8 and 24

Henry Cohn
Microsoft Research and MIT

Thurston Lectures
UC Davis
May 12, 2025

The sphere packing problem

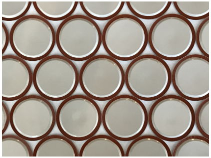
How densely can we pack identical spheres into space?

Not allowed to overlap (but can be tangent).

Density = fraction of space filled by the spheres.

One-dimensional sphere packing is boring. (density $= 1$, trivial)

Two-dimensional sphere packing is prettier and more interesting:



(density $= \pi/\sqrt{12} \approx 90.69\%$, Thue 1892)

The three-dimensional case strains human ability to prove:

(density $= \pi/\sqrt{18} \approx 74.04\%$, Hales 1998, Hales et al. 2014)

What about higher dimensions?

Sphere packing in $\mathbb{R}^n$

Solved for only two cases with $n > 3$. Viazovska's breakthrough:

Theorem (Viazovska 2017).
The E_8 root lattice achieves the greatest possible sphere packing density in $\mathbb{R}^8$, namely $\pi^4/384 \approx 25.37\%$.

Twenty-four dimensions builds on her techniques:

Theorem (Cohn, Kumar, Miller, Radchenko, and Viazovska 2017).
The Leech lattice Λ_{24} achieves the greatest possible sphere packing density in $\mathbb{R}^{24}$, namely $\pi^{12}/12! \approx 0.1930\%$.

What's so special about these dimensions? How can we understand 8 dimensions without understanding 4 through 7?

Why should we care about sphere packing?

Natural geometric problem in its own right.

Toy model for granular materials.

Error-correcting codes for continuous communication channels.
High dimensions arise naturally in practice.

Instead of aspects of the problem or applications, we'll justify sphere packing by its solutions:

A question is good if it has good answers.

What is known?

Each dimension has its own idiosyncrasies.

Good constructions are known for low dimensions.

Iteratively stacking layers from the previous dimension is not a general solution (it fails by $\mathbb{R}^{10}$).

No idea what the best high-dimensional packings look like. They may even be disordered.

Upper/lower density bounds in general.

Bounds are very far apart:

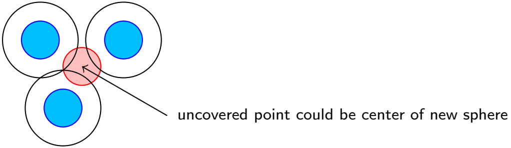
For $n = 36$, differ by a multiplicative factor of 52.

This factor grows exponentially as $n \to \infty$.

Packing in high dimensions

Easy theorem: every saturated packing (i.e., one in which no more spheres can be added) in n dimensions has density at least 2^{-n}.

Proof: Double the radius of the spheres. All of space must be then covered, since any uncovered point could have been the center of an additional sphere in the original packing.



Doubling the radius multiplies the volume by 2^n. Q.E.D.

This bound is very nearly all we know!

Asymptotics for packing in high dimensions

Greedy argument: density at least 2^{-n}.

Minkowski (1905): at least $2 \cdot 2^{-n}$.

$\vdots$

Ball (1992): at least $2n \cdot 2^{-n}$.

Vance (2011): at least $\frac{6}{e} n \cdot 2^{-n}$ when n is a multiple of 4.

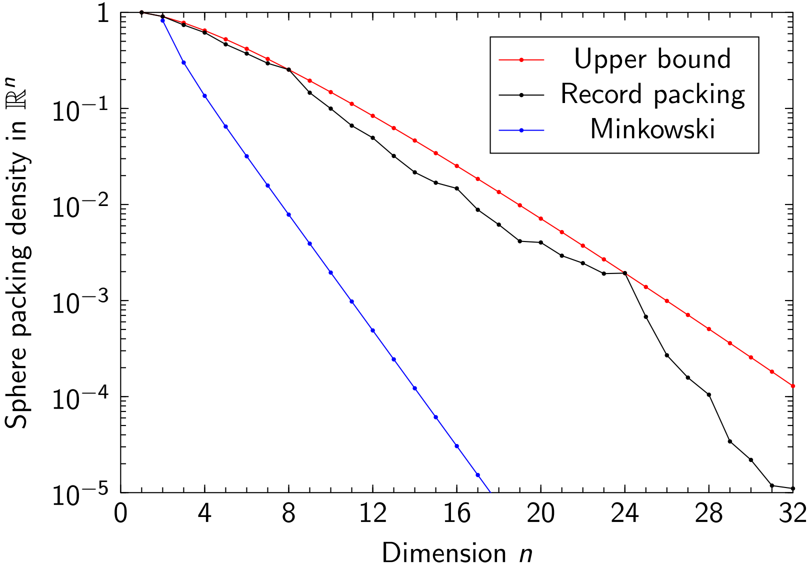
Venkatesh (2013): at least $\frac{e^{-\gamma}}{2} n \log\log n \cdot 2^{-n}$ for a certain sparse sequence of dimensions.

Campos, Jenssen, Michelen, and Sahasrabudhe (2023): at least $\frac{1}{2} n \log n \cdot 2^{-n}$.

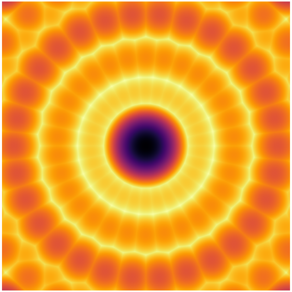
Klartag (2025): at least $cn^2 \cdot 2^{-n}$ with $c > 0$.

For comparison, the best upper bound known is $2^{-(0.599\ldots+o(1))n}$, due to Kabatiansky and Levenshtein (1978).

Record packings and bounds



The most remarkable packings



Amazing lattices in certain dimensions:

E_8 root lattice in $\mathbb{R}^8$
Leech lattice Λ_{24} in $\mathbb{R}^{24}$

Extremely symmetrical and dense packings of spheres.

Connected with many areas in mathematics and physics (e.g., string theory, modular forms, hyperbolic geometry, finite simple groups, error-correcting codes).

What makes these cases work out so well?

Notice that they rise to meet the upper bound, not vice versa.

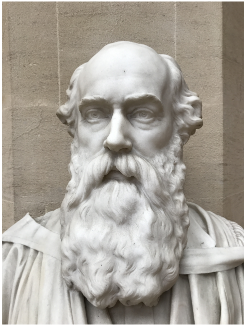
What is E_8?

Existence proved by H. J. S. Smith in 1867 via nonconstructive mass formula.

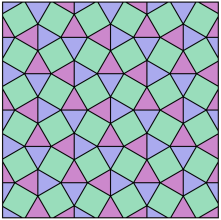
Constructed explicitly by A. Korkine and G. Zolotareff in 1873.

Appeared as the exceptional root lattice E_8 in W. Killing's 1890 classification of semisimple Lie algebras.

Constructed by T. Gosset in 1900 as the vertex set of a semiregular tessellation of $\mathbb{R}^8$ by simplices and cross polytopes (8-dimensional analogues of tetrahedra and octahedra).



H. J. S. Smith

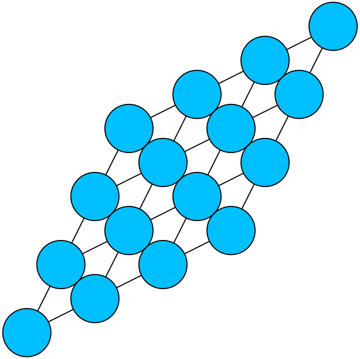


(image by R. A. Nonenmacher)

How can we pack spheres?

Lattice: integer span of n linearly independent vectors. I.e., for basis $v_1, \ldots, v_n$, center spheres at

$$\{a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \mid a_1, \ldots, a_n \in \mathbb{Z}\}.$$



Packing radius = half the shortest non-zero vector length.

Lattice density

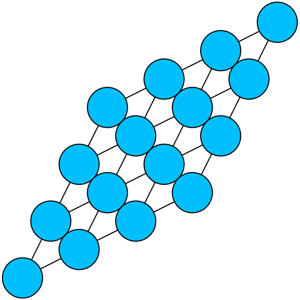
If Λ is a lattice in $\mathbb{R}^n$ with minimal vector length r, then its density is

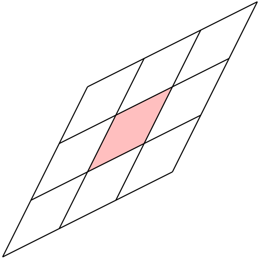
$$\mathrm{vol}(\mathrm{ball}) \cdot (\#\text{ balls per unit volume in space}),$$

which equals

$$\mathrm{vol}(B^n_{r/2}) \cdot \frac{1}{\mathrm{vol}(\mathbb{R}^n/\Lambda)}.$$

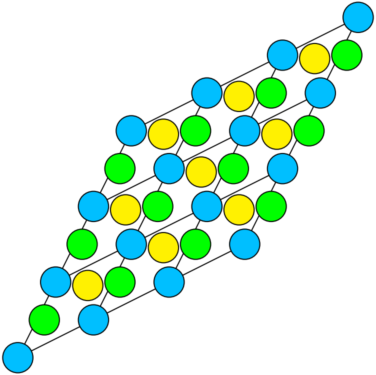
Here $\mathrm{vol}(\mathbb{R}^n/\Lambda)$ is the volume of a fundamental cell, i.e., the absolute value of the determinant of a basis matrix.





Periodic packings

In a periodic packing, spheres are not restricted to just the corners of a fundamental cell.

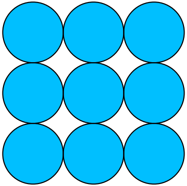


No reason to believe densest packing must be periodic, but periodic packings come arbitrarily close to the maximum density.

By contrast, lattices probably do not.

The best sphere packings currently known are not always lattice packings, but many good packings are.

Simplest lattice: $\mathbb{Z}^n$, lousy packing.



Better: for $n \geq 3$, "checkerboard" packing

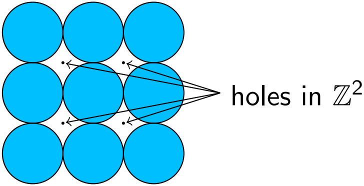
$$D_n = \{(x_1, \ldots, x_n) \in \mathbb{Z}^n \mid x_1 + \cdots + x_n \text{ is even}\}.$$

D_3, D_4, D_5 are best known packings in those dimensions, and provably best lattice packings.

What happens for larger n?

The holes in D_n grow larger and larger.

A hole is a local maximum for distance from nearest lattice point.



Where are the holes in D_n?

Two classes of holes in D_n (for $n \geq 3$):

- $(1, 0, \ldots, 0)$ at distance 1 from lattice.
- $(1/2, 1/2, \ldots, 1/2)$ at distance

$$\sqrt{\left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{n}{4}}.$$

For comparison, nearest lattice points are

$$(0, 0, \ldots, 0) \qquad \text{and} \qquad (1, 1, 0, \ldots, 0)$$

at distance $\sqrt{2}$.

Wonderful properties of dimension 8

When $n = 8$, radius $\sqrt{n/4}$ of deep hole equals distance $\sqrt{2}$ between lattice points.

We can slip another copy of D_8 into the holes! This doubles the packing density.

Result called E_8 lattice.

Leech lattice ($n = 24$) is similar in spirit, but more complicated.

Key properties of E_8

Distances between distinct lattices points are $\sqrt{2k}$ with $k = 1, 2, \ldots,$ and $\mathrm{vol}(\mathbb{R}^8/E_8) = 1$.

E_8 is a *self-dual* lattice: $E_8^* = E_8$. To obtain the *dual lattice* Λ^* of a lattice Λ,

let $v_1, \ldots, v_n$ be any basis of Λ, and

let $v_1^*, \ldots, v_n^*$ be the dual basis satisfying $\langle v_i, v_j^* \rangle = \delta_{i,j}$.

Then $v_1^*, \ldots, v_n^*$ is a basis of Λ^*.

Net result: E_8 is an *even unimodular lattice*.

Leech lattice works the same, but without points at distance $\sqrt{2}$.

So how can we prove upper bounds for density?

Linear programming bounds for sphere packing

Developed by H. Cohn and N. Elkies, based on a line of research going back to P. Delsarte in 1972.

Uses harmonic analysis, in particular the Fourier transform, to analyze pair correlations in packings.

Auxiliary functions with certain properties yield upper bounds for packing density.

No reference to special dimensions such as eight and twenty-four, yet the bounds end up being sharp in these dimensions.

Fourier transform

Define the Fourier transform $\widehat{f}$ of an integrable function $f : \mathbb{R}^n \to \mathbb{R}$ by

$$\widehat{f}(y) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \langle x,y \rangle} \, dx.$$

If $\widehat{f}$ is integrable as well, then Fourier inversion tells us that

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(y) e^{2\pi i \langle x,y \rangle} \, dy.$$

In other words, $\widehat{f}$ tells how to decompose f into complex exponentials, and vice versa.

Significance of the Fourier transform in discrete geometry

It diagonalizes the operation of translation by any vector.

Specifically,

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(y) e^{2\pi i \langle x,y\rangle}\, dy.$$

implies

$$f(x+t) = \int_{\mathbb{R}^n} \widehat{f}(y) e^{2\pi i \langle t,y\rangle} e^{2\pi i \langle x,y\rangle}\, dy.$$

I.e., translating the input to the function f by t amounts to multiplying its Fourier transform $\widehat{f}(y)$ by $e^{2\pi i \langle t,y\rangle}$.

Simultaneously diagonalizing all these translation operators makes the Fourier transform an ideal tool for studying periodic structures.

Poisson summation

Our key technical tool is the *Poisson summation formula*:

If $f\colon \mathbb{R}^n \to \mathbb{R}$ is sufficiently nice and Λ is a lattice in $\mathbb{R}^n$, then

$$\sum_{x \in \Lambda} f(x) = \frac{1}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} \sum_{y \in \Lambda^*} \widehat{f}(y).$$

Here $\mathrm{vol}(\mathbb{R}^n/\Lambda)$ is the determinant of Λ, and Λ^* is the dual lattice.

I.e., up to a scaling factor,
    summing a function over a lattice
is the same as
    summing its Fourier transform over the dual lattice.

Proof of Poisson summation

Poisson summation is the special case $t = 0$ of

$$\sum_{x \in \Lambda} f(x+t) = \frac{1}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} \sum_{y \in \Lambda^*} \widehat{f}(y) e^{2\pi i \langle y,t \rangle}.$$

Left side is periodic modulo Λ, and right side is its Fourier series.

The function $t \mapsto e^{2\pi i \langle y,t \rangle}$ is periodic modulo Λ iff $y \in \Lambda^*$.

Given left side, it's not hard to compute the Fourier coefficients by orthogonality.

Linear programming bounds

Theorem (Cohn and Elkies 2003).
Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a radial Schwartz function (i.e., smooth and rapidly decaying) and let $r > 0$, with

$$\begin{aligned} &f(0) = \widehat{f}(0) > 0, \\ &f(x) \le 0 \text{ for } |x| \ge r, \text{ and} \\ &\widehat{f}(y) \ge 0 \text{ for all } y. \end{aligned}$$

Then the sphere packing density in $\mathbb{R}^n$ is at most $\mathrm{vol}\big(B^n_{r/2}\big)$.

The volume $\mathrm{vol}\big(B^n_{r/2}\big)$ of a ball of radius $r/2$ in $\mathbb{R}^n$ is $\frac{\pi^{n/2}}{(n/2)!}\left(\frac{r}{2}\right)^n$, where $(n/2)!$ means $\Gamma(n/2+1)$ when n is odd.

We can radially symmetrize f, since all the constraints are invariant under rotation. Thus, f is a function of one (radial) variable.

Proof for lattices (general case is similar)

Suppose Λ is a lattice packing with spheres of radius $r/2$. I.e., the minimal vector length is at least r. Then

$$f(0) \geq \sum_{x \in \Lambda} f(x)$$

because $f(x) \leq 0$ for $|x| \geq r$, while

$$\frac{1}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} \sum_{y \in \Lambda^*} \widehat{f}(y) \geq \frac{\widehat{f}(0)}{\mathrm{vol}(\mathbb{R}^n/\Lambda)}$$

because $\widehat{f}(y) \geq 0$ for all y. Thus, by Poisson summation,

$$f(0) \geq \frac{\widehat{f}(0)}{\mathrm{vol}(\mathbb{R}^n/\Lambda)},$$

so $\mathrm{vol}(\mathbb{R}^n/\Lambda) \geq 1$ and thus there is at most one ball of radius $r/2$ per unit volume in space. Q.E.D.

How do we choose f?

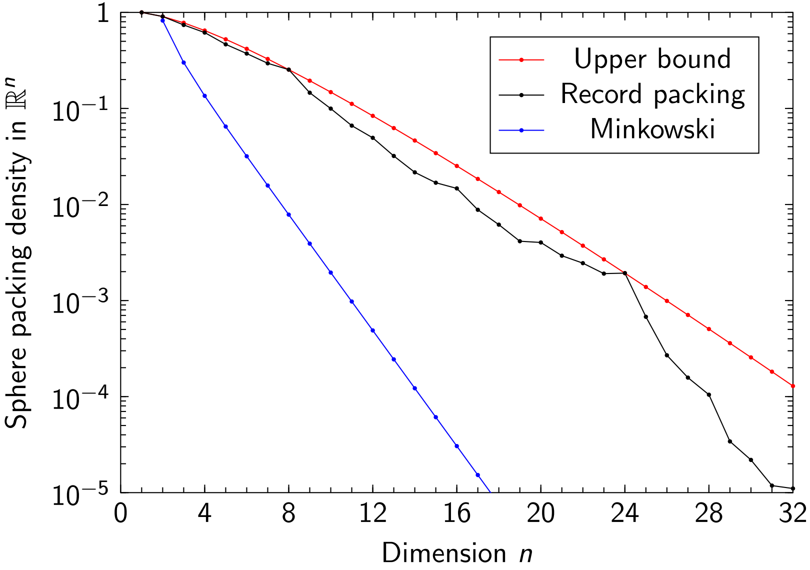
Nobody knows in general.

Get a trivial density bound of 1 via the convolution $f = \chi_B * \chi_B$, where χ_B is the characteristic function of a ball of volume 1.

More sophisticated choices give the best density bounds known in high dimensions.

In an arbitrary dimension, fall back on numerical optimization. Get good bounds in general, and seemingly sharp bounds when $n = 1$, 2, 8, or 24.

Record packings and bounds



Equality?

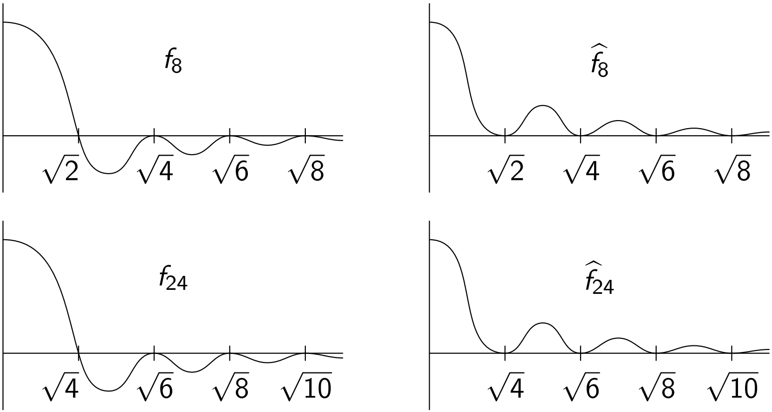
Get a sharp bound for a lattice Λ iff $f(x) = 0$ for all $x \in \Lambda \setminus \{0\}$ and $\widehat{f}(y) = 0$ for all $y \in \Lambda^* \setminus \{0\}$.

A sharp bound is not difficult for $n = 1$, is conjectured for $n = 2$, and holds for $n = 8$ and 24 (which solves the sphere packing problem).

Need *magic auxiliary functions* to get a sharp bound.

For $n = 8$ and 24, we know exactly where the roots should be: radius $\sqrt{2k}$ with $k \geq 1$ for E_8 and $k \geq 2$ for Λ_{24}.

Approximate plots (not to scale)



Numerical optimization using radial polynomials times Gaussians yields 60+ decimal digits, but cannot get an exact bound since polynomials have only finitely many roots.

Fundamental question

Do there exist radial functions f on $\mathbb{R}^8$ and $\mathbb{R}^{24}$ with these prescribed roots for f and $\widehat{f}$?

Yes, to a close numerical approximation, but that is not a proof.

Unfortunately, it's difficult to control a function and its Fourier transform simultaneously. This is the Heisenberg uncertainty principle.

Can we identify enough patterns to pin down these functions?

Quadratic coefficients

There are some numerical patterns, but not enough to determine everything:

Conjecture (Cohn and Miller). The quadratic Taylor coefficients of the magic functions f and $\widehat{f}$ (normalized with $f(0) = \widehat{f}(0) = 1$) are rational numbers when $n = 8$ or $n = 24$.

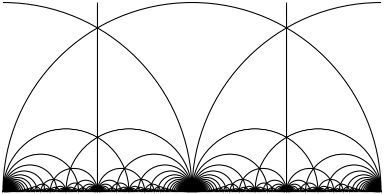
n	function	order	coefficient	conjecture
8	f	2	$-2.7000000000000000000000000000\ldots$	$-27/10$
8	$\widehat{f}$	2	$-1.5000000000000000000000000000\ldots$	$-3/2$
24	f	2	$-2.6276556776556776556776556776\ldots$	$-14347/5460$
24	$\widehat{f}$	2	$-1.3141025641025641025641025641\ldots$	$-205/156$
8	f	4	$4.2167501240968298210998965628\ldots$	?
8	$\widehat{f}$	4	$-1.2397969070295980026220596589\ldots$	?
24	f	4	$3.8619903167183007758184168473\ldots$	?
24	$\widehat{f}$	4	$-0.7376727789015322303799539712\ldots$	?

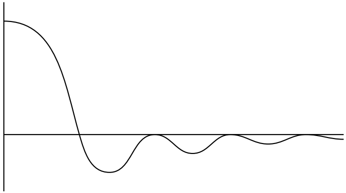
(Now proved using explicit formulas.)

Modular forms

Many people suspected the answer must involve modular forms, because they are deep special functions related to lattices.

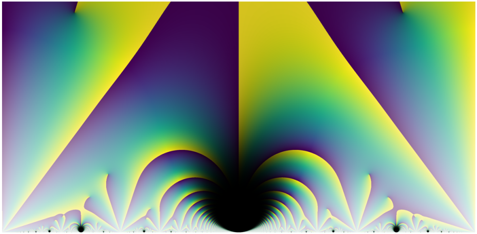
However, nobody could figure out how. Modular forms look nothing like radial functions.

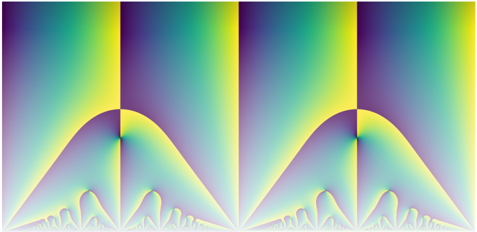




Viazovska's quasimodular forms

Viazovska found an extraordinary integral transform that turns these into the magic functions.





Eigenfunctions of the Fourier transform

Goal: radial functions such that f and $\widehat{f}$ have prescribed roots.

We'll split f into eigenfunctions of the Fourier transform via $f = f_+ + f_-$ with $\widehat{f_+} = f_+$ and $\widehat{f_-} = -f_-$.

This amounts to Fourier inversion: $\widehat{\widehat{f}} = f$ for radial f, so we can take $f_+ = \big(f + \widehat{f}\big)/2$ and $f_- = \big(f - \widehat{f}\big)/2$.

Because f and $\widehat{f}$ vanish at the same points, they share these roots with f_+ and f_-.

New goal: construct radial eigenfunctions of Fourier transform with prescribed roots. What are the magic eigenfunctions?

Modular forms

Viazovska uses modular forms to obtain the magic eigenfunctions.

Let $\mathfrak{h}$ be the upper half-plane $\{z \in \mathbb{C} : \operatorname{Im} z > 0\}$.

A *modular form* of weight k for $\mathrm{SL}_2(\mathbb{Z})$ is a holomorphic function $\varphi\colon \mathfrak{h} \to \mathbb{C}$ such that

$$\varphi(z+1) = \varphi(z) \quad \text{and} \quad \varphi(-1/z) = z^k \varphi(z)$$

for all $z \in \mathfrak{h}$, and $\varphi(z)$ remains bounded as $\operatorname{Im} z \to \infty$ ("holomorphic at infinity").

Why are these functional equations important? They turn out to come up surprisingly often in number theory.

What do they have to do with the magic functions? Not at all obvious!

Examples

Eisenstein series E_k (not to be confused with E_8!) defined by

$$E_k(z) = \frac{1}{2\zeta(k)} \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \frac{1}{(mz+n)^k}$$

is a modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$ whenever k is an even integer greater than 2 (while it vanishes when k is odd).

Proof is just rearrangement of series using absolute convergence.

What about $k = 2$? Conditional convergence leads to

$$E_2(-1/z) = z^2 E_2(z) - \frac{6i}{\pi} z,$$

so only a "quasimodular form." Nevertheless plays an important role in sphere packing.

Examples

If Λ is an even unimodular lattice in $\mathbb{R}^n$ with N_{2k} vectors of norm $2k$ for $k = 0, 1, 2, \ldots$, then

$$\Theta_\Lambda(z) = \sum_{k=0}^{\infty} N_{2k} e^{2\pi ikz}$$

is a modular form of weight $n/2$ for $\mathrm{SL}_2(\mathbb{Z})$.

The identity $\Theta_\Lambda(z+1) = \Theta_\Lambda(z)$ is trivial, while $\Theta_\Lambda(-1/z) = z^{n/2}\Theta_\Lambda(z)$ follows from Poisson summation.

Laplace transform

What do modular forms have to do with the magic functions?

Define f in terms of Gaussians by

$$f(x) = \int_0^\infty e^{-t\pi|x|^2} g(t)\, dt.$$

The Fourier transform of a wide Gaussian is narrow and vice versa:

$$\begin{aligned}\widehat{f}(y) &= \int_0^\infty t^{-n/2} e^{-\pi|y|^2/t} g(t)\, dt \\ &= \int_0^\infty e^{-t\pi|y|^2} t^{n/2-2} g(1/t)\, dt.\end{aligned}$$

In other words, taking the Fourier transform of f amounts to replacing g with $t \mapsto t^{n/2-2} g(1/t)$.

Connection with modular forms

If $g(1/t) = \varepsilon t^{2-n/2} g(t)$ with $\varepsilon = \pm 1$, then $\widehat{f} = \varepsilon f$.

This looks like a modular form on the imaginary axis.

If $g(t) = \varphi(it)$ for a modular form φ of weight k for $\mathrm{SL}_2(\mathbb{Z})$, then $\varphi(-1/z) = z^k \varphi(z)$ corresponds to $g(1/t) = i^k t^k g(t)$.

If φ is a meromorphic modular form of weight $k = 2 - n/2$, then f is a radial Fourier eigenfunction in $\mathbb{R}^n$, with eigenvalue i^k.

But can we control the roots? No obvious way to specify them.

(meromorphic modular form = ratio of holomorphic ones)

Toy version (with single roots)

Let G be the group generated by $z \mapsto z+2$ and $z \mapsto -1/z$, which has index 3 in $\mathrm{SL}_2(\mathbb{Z})$.

Let φ be a weakly holomorphic modular form of weight $2 - n/2$ for G, with its only pole at $i\infty$ and such that it vanishes at 1. Then it has a Fourier expansion

$$\varphi(z) = \sum_{k=k_0}^{\infty} c_k e^{\pi i k z},$$

where $k_0 < 0$.

Define a radial function $f\colon \mathbb{R}^n \to \mathbb{R}$ by

$$f(x) = \int_{-1}^{1} \varphi(z) e^{\pi i |x|^2 z}\, dz,$$

where the contour is the upper half of the unit circle.

Then $\widehat{f} = i^{2-n/2} f$ and $f(\sqrt{k}) = c_{-k}$ for $k \geq 0$ (i.e., $f(x) = c_{-k}$ when $|x| = \sqrt{k}$). These are simple calculations:

$$\begin{aligned}
\widehat{f}(x) &= \int_{-1}^{1} \varphi(z)(z/i)^{-n/2} e^{\pi i |x|^2(-1/z)}\, dz \\
&= i^{n/2} \int_{1}^{-1} \varphi(-1/z) z^{n/2} e^{\pi i |x|^2 z}\, dz/z^2 \\
&= i^{2-n/2} f(x)
\end{aligned}$$

and Fourier orthogonality

$$f(\sqrt{k}) = \int_{-1}^{1} \varphi(z) e^{\pi i k z}\, dz = c_{-k}.$$

So f is a radial Fourier eigenfunction that vanishes at all but finitely many square roots of integers. But getting double roots at $\sqrt{2k}$ instead is quite a bit more subtle...

Viazovska's trick

Multiply by $\sin^2(\pi|x|^2/2)$, which vanishes to second order at $|x| = \sqrt{2k}$ for $k = 1, 2, 3, \dots$.

Magic eigenfunctions have the form

$$\sin^2(\pi|x|^2/2) \int_0^\infty g(t) e^{-\pi|x|^2 t}\, dt.$$

Obvious issue: $\sin^2(\pi|x|^2/2)$ vanishes too much at $|x| = 0$ and $|x| = \sqrt{2}$. Integral must have poles to cancel unwanted roots. It converges only for $|x| > \sqrt{2}$, but we can analytically continue.

The sine factors intefere with the Laplace transform. When do we get an eigenfunction? Not obvious, but Viazovska gives an argument by shifting contours of integration.

Sufficient conditions for an eigenfunction

For a $+1$ eigenfunction in $\mathbb{R}^8$,

$$g(t) = t^2 \varphi(i/t),$$

where φ is a weakly holomorphic quasimodular form of weight 0 and depth 2 for $\mathsf{SL}_2(\mathbb{Z})$.

For a -1 eigenfunction in $\mathbb{R}^8$,

$$g(t) = \psi(it),$$

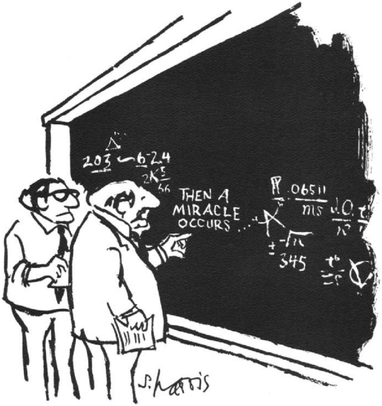
where ψ is a weakly holomorphic modular form of weight -2 for the subgroup $\Gamma(2)$ of $\mathsf{SL}_2(\mathbb{Z})$ and

$$\psi(z) = \psi(z+1) + z^2 \psi(-1/z).$$

Very strange conditions! Not at all obvious.

Obtaining the magic eigenfunctions

Look for quasimodular forms and hope for luck.



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

By now we have more insight (e.g., interpolation formulas, Dan Romik's proof), but it still feels like a miracle.

Higher dimensions

The linear programming bounds are probably not sharp in any other dimensions greater than two.

Building on work of Schrijver, Bachoc and Vallentin, and others, de Laat and Vallentin have generalized linear programming bounds to a hierarchy of semidefinite programming bounds.

Are any other dimensions accessible at low levels of the hierarchy? Perhaps optimality of D_4 in $\mathbb{R}^4$?

For more information

Henry Cohn
A conceptual breakthrough in sphere packing
Notices of the AMS **64** (2017), 102–115
arXiv:1611.01685

Henry Cohn
The work of Maryna Viazovska
Fields medal laudatio, ICM 2022
arXiv:2207.06913