# Numerical exploration in sphere packing, Fourier analysis, and physics

Henry Cohn

Lecture 3
PCMI 2022

Joint work with Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska.




*Universal optimality of the $E_8$ and Leech lattices and interpolation formulas*, arXiv:1902.05438

Viazovska's solution of the 8-dimensional sphere packing problem introduced an amazing special function, but it left many questions unanswered:

Where did this function come from? What does it mean?

Is there a broader theory?

What else can one prove using these techniques?

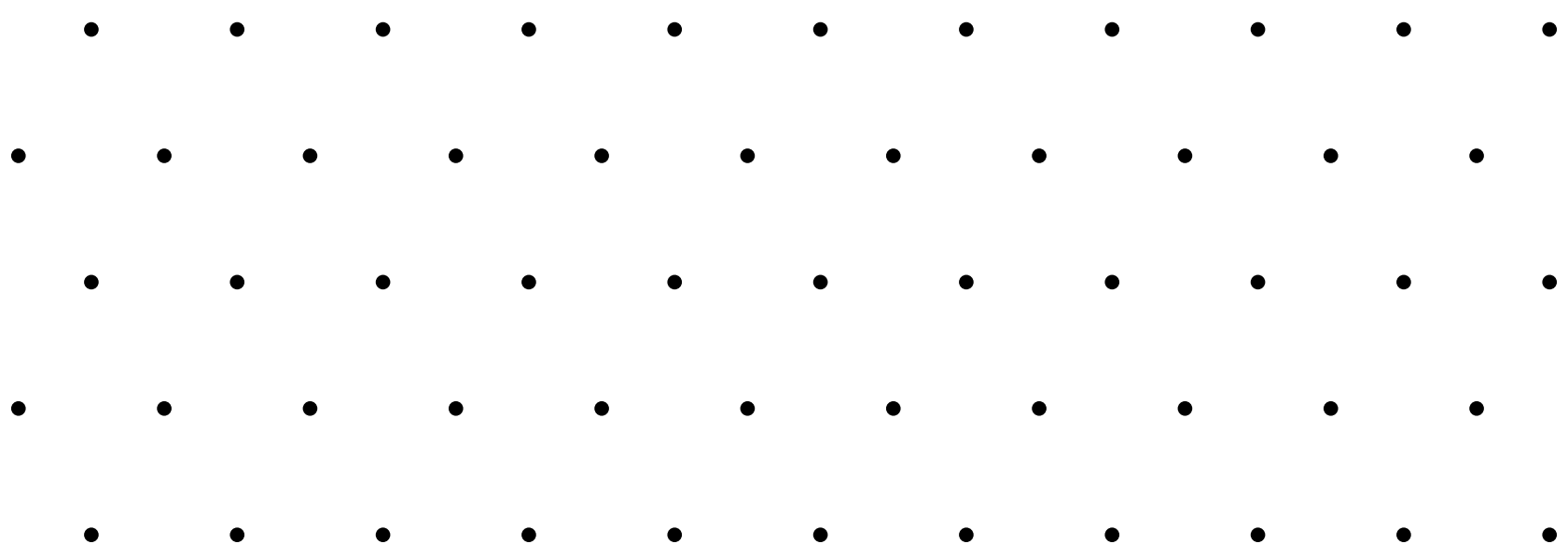
Today we'll place this special function in the context of an interpolation theorem, and see how that theorem implies a broad generalization of sphere packing optimality.

# Key question

How do particles arrange themselves under a repulsive force?

Infinitely many particles in Euclidean space.

Easy to guess the answer in $\mathbb{R}^1$ (equally spaced) or $\mathbb{R}^2$ (hexagonal).



What about higher dimensions, or proofs?

# Setting

Classical point particles in $\mathbb{R}^d$.

Locations specified by discrete, closed subset $\mathcal{C}$ of $\mathbb{R}^d$.

Pair potential function $p$, such as $p(r) = 1/r^s$ with $s > 0$ or $p(r) = e^{-\alpha r^2}$ with $\alpha > 0$. (Decreasing $\Leftrightarrow$ repulsive.)

If $|\mathcal{C}| < \infty$, potential energy is

$$\sum_{\substack{x,y \in \mathcal{C} \\ x \neq y}} p(|x - y|).$$

We care mainly about $|\mathcal{C}| = \infty$, so we must normalize to consider the average energy per particle.

# Energy

$\mathcal{C}$ has *density* $\rho$ if

$$\lim_{r\to\infty} \frac{\left|\mathcal{C} \cap B_r^d(0)\right|}{\mathrm{vol}\left(B_r^d(0)\right)} = \rho,$$

where $B_r^d(x)$ is a ball of radius $r$ about $x$.

The *lower p-energy* of $\mathcal{C}$ is

$$E_p(\mathcal{C}) := \liminf_{r\to\infty} \frac{1}{\left|\mathcal{C} \cap B_r^d(0)\right|} \sum_{\substack{x,y\in\mathcal{C}\cap B_r^d(0)\\ x\neq y}} p\big(|x-y|\big).$$

Call it the *energy* if the limit exists.

A lattice $\Lambda$ has density

$$\frac{1}{\mathrm{vol}(\mathbb{R}^d/\Lambda)}$$

and $p$-energy

$$\sum_{x \in \Lambda \setminus \{0\}} p(|x|)$$

(if absolutely convergent).

Lattices are among the simplest ways to arrange particles, but not necessarily optimal.

More generally, a *periodic configuration* $\mathcal{C} = \bigcup\!\!\!\cdot_{j=1}^{N}(\Lambda + v_j)$ has density

$$\frac{N}{\mathrm{vol}(\mathbb{R}^d/\Lambda)}$$

and $p$-energy

$$\frac{1}{N}\sum_{j,k=1}^{N}\sum_{x\in\Lambda\setminus\{v_k-v_j\}} p(|x+v_j-v_k|)$$

(again if absolutely convergent).

The non-periodic case is also important, but periodic configurations are more familiar.

# Ground state energy

Definition. A configuration $\mathcal{C}$ in $\mathbb{R}^d$ of density $\rho$ is a *ground state*, or *minimizes energy*, for potential $p$ if

1. its $p$-energy exists, and
2. all other configurations in $\mathbb{R}^d$ with density $\rho$ have lower $p$-energy at least $E_p(\mathcal{C})$.

Sphere packing is a limiting case. It amounts to maximizing the minimal distance between particles at a fixed particle density, and that is energy minimization for a steep potential.

*Crystallization problem* in mathematical physics (for classical, mesoscale materials):

Why do particles often arrange themselves
periodically at zero temperature?

# Difficulty

Ground states are a mystery.

We can't predict ground states in most cases.

We can't even predict qualitative features, such as whether there should be a periodic ground state.

It's easy to make false conjectures.

Even in low dimensions, such as $\mathbb{R}^2$, we generally can't explain what is seen in simulations.

# Number theory

A lattice $\Lambda$ has an *Epstein zeta function*

$$\zeta_\Lambda(s) = \sum_{x \in \Lambda \setminus \{0\}} \frac{1}{|x|^{2s}}$$

(for $\mathrm{Re}(s) > d/2$) and *theta series*

$$\Theta_\Lambda(z) = \sum_{x \in \Lambda} e^{\pi i z |x|^2}$$

(for $\mathrm{Im}(z) > 0$).

The energy of $\Lambda$ under $r \mapsto 1/r^s$ is $\zeta_\Lambda(s/2)$, and under $r \mapsto e^{-\alpha r^2}$ is $\Theta_\Lambda(i\alpha/\pi) - 1$.

Thus, minimizing lattice energy amounts to finding extreme values of number-theoretic special functions.

# Universal optimality

When is a ground state independent of the potential function?

Which potential functions are reasonable to consider?

A function $p\colon (0,\infty) \to \mathbb{R}$ is *completely monotonic* if it is $C^\infty$ and for all $k$, $(-1)^k p^{(k)} \geq 0$. Nonnegative, decreasing, convex, etc.

Definition. A configuration $\mathcal{C}$ is *universally optimal* if it is a ground state for all completely monotonic functions of squared distance. E.g., inverse power laws or Gaussians.

In fact, Gaussians span the cone of completely monotonic functions of squared distance (Bernstein's theorem), so

$\mathcal{C}$ is universally optimal iff it is a ground state for all Gaussians.

Equivalently, we can fix a Gaussian and vary the particle density.

# Low dimensions

Theorem (Ventavogel and Nijboer, 1979). $\mathbb{Z}$ is universally optimal in $\mathbb{R}$.

Conjecture. The hexagonal lattice $A_2$ is universally optimal in $\mathbb{R}^2$.

Proved optimal among lattices by Montgomery in 1988, but not known in general.

Previously, no ground state was known for any nice, decreasing potential function in dimension greater than 1. (No inverse power law, no Gaussian, etc.)

Dealing with long-range interactions is tough.

# Three dimensions

Consider the potential function $r \mapsto e^{-\pi r^2}$. What happens at density $\rho$? Universal optimality fails.

Conjecture. Among lattices, the face-centered cubic lattice $A_3$ is optimal for $\rho \leq 1$, and the body-centered cubic $A_3^*$ is optimal for $\rho \geq 1$.

Same energy when $\rho = 1$ by Poisson summation.
How does the phase transition near $\rho = 1$ behave?

Stillinger (1976): phase coexistence, with lower energy when

$$0.99899854\ldots < \rho < 1.00100312\ldots.$$

At $\rho = 1$, improve energy by 0.0004%. Not periodic.

Is this the full answer? No idea. It deserves further exploration.

# Main theorem: universal optimality in $\mathbb{R}^8$ and $\mathbb{R}^{24}$

Theorem. The $E_8$ root lattice in $\mathbb{R}^8$ and the Leech lattice in $\mathbb{R}^{24}$ are universally optimal, and unique among periodic packings for potentials under which they have finite energy.

Also seems to be true for the hexagonal lattice in $\mathbb{R}^2$, but we don't know how to prove it.

Simulations suggest universal optimality generally fails.
Dimensions 1, 2, 8, and 24 seem very special.

# Harmonic analysis

Recall that a *Schwartz function* $f : \mathbb{R}^d \to \mathbb{R}$ is a smooth function whose partial derivatives (of all orders) decay faster than 1/any polynomial. Think "nice function."

As always, we normalize the *Fourier transform* by

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

Schwartz functions are closed under the Fourier transform, as are radial functions (i.e., functions where $f(x)$ depends only on $|x|$).

The proof of universal optimality will require a new understanding of radial Schwartz functions on $\mathbb{R}^8$ and $\mathbb{R}^{24}$.

# Linear programming bound

Proposition (Cohn and Kumar, 2007). Let $p\colon (0,\infty) \to \mathbb{R}$ be any function, and $\rho > 0$. If $f : \mathbb{R}^d \to \mathbb{R}$ is a Schwartz function such that

1. $f(x) \leq p(|x|)$ for all $x \in \mathbb{R}^d \setminus \{0\}$ and
2. $\widehat{f}(y) \geq 0$ for all $y \in \mathbb{R}^d$,

then every subset of $\mathbb{R}^d$ of density $\rho$ has lower $p$-energy at least $\rho\widehat{f}(0) - f(0)$.

In other words, $f$ satisfying inequalities (1) and (2) certifies a lower bound for energy.

Without loss of generality, we can take $f$ to be radial: average all of its rotations about the origin.

How can we choose the best $f$ for a given $d$, $p$, and $\rho$?
Nobody knows, except for $d \in \{1, 8, 24\}$.

Numerics for potential $p(r) = e^{-\pi r^2}$ and density $\rho = 1$ in $\mathbb{R}^d$:

| $d$ | LP bound | Current record | |
|---|---|---|---|
| 1 | 0.08643481... | 0.08643481... | (equal) |
| 2 | 0.15959526... | 0.15959526... | (conj. equal) |
| 3 | 0.22321782... | 0.23153532... | |
| 4 | 0.27956960... | 0.28576449... | |
| 5 | 0.33011740... | 0.34868410... | |
| 6 | 0.37587226... | 0.38874675... | |
| 7 | 0.41756856... | 0.42445404... | |
| 8 | 0.45576289... | 0.45576289... | (equal) |
| 24 | 0.79965280... | 0.79965280... | (equal) |

Based on numerical search to optimize $f$ approximately.

Theorem (Cohn and Kumar, 2007). The LP bound proves universal optimality for $\mathbb{Z}$ in $\mathbb{R}^1$.

Ground states in one dimension are not exciting,
but they are harder to analyze than they sound.

Conjecture (Cohn and Kumar, 2007). The LP bound proves universal optimality for the hexagonal lattice $A_2$ in $\mathbb{R}^2$.

Still not known!

Theorem. The LP bound proves universal optimality for $E_8$ in $\mathbb{R}^8$ and the Leech lattice in $\mathbb{R}^{24}$.

Why is this easier for $\mathbb{R}^8$ and $\mathbb{R}^{24}$ than $\mathbb{R}^2$?

# Proof of LP bound for lattice $\Lambda$ in $\mathbb{R}^d$

Poisson summation says

$$\sum_{x \in \Lambda} f(x) = \rho \sum_{y \in \Lambda^*} \widehat{f}(y).$$

Thus,

$$\begin{aligned}
E_p(\Lambda) &= \sum_{x \in \Lambda \setminus \{0\}} p(|x|) \\
&\geq \sum_{x \in \Lambda \setminus \{0\}} f(x) && \text{because } f(x) \leq p(|x|) \\
&= -f(0) + \rho \sum_{y \in \Lambda^*} \widehat{f}(y) && \text{by Poisson summation} \\
&\geq \rho \widehat{f}(0) - f(0). && \text{because } \widehat{f}(y) \geq 0
\end{aligned}$$

The proof for non-lattices is similar in spirit.

# When does $f$ prove a sharp bound for energy?

To avoid any loss in the inequalities, we need

1. $f(x) = p(|x|)$ for all $x \in \Lambda \setminus \{0\}$, and
2. $\widehat{f}(y) = 0$ for all $y \in \Lambda^* \setminus \{0\}$.

Furthermore, these inequalities must hold to order two. I.e., the radial derivatives satisfy $f'(x) = p'(|x|)$ and $\widehat{f}\,'(y) = 0$.

For $E_8$ and the Leech lattice, we have $\Lambda^* = \Lambda$,
with vector lengths $\sqrt{2n}$ for $n \geq n_0$,
where $n_0 = 1$ for $d = 8$ and $n_0 = 2$ for $d = 24$.

In other words, for $n \geq n_0$ we need

$$\begin{aligned} f\left(\sqrt{2n}\right) &= p\left(\sqrt{2n}\right), \\ f'\left(\sqrt{2n}\right) &= p'\left(\sqrt{2n}\right), \\ \widehat{f}\left(\sqrt{2n}\right) &= 0, \\ \widehat{f}\,'\left(\sqrt{2n}\right) &= 0. \end{aligned}$$

Can we reconstruct a radial $f$ from this information?

I.e., from knowing $f\left(\sqrt{2n}\right)$, $f'\left(\sqrt{2n}\right)$, $\widehat{f}\left(\sqrt{2n}\right)$, and $\widehat{f}\,'\left(\sqrt{2n}\right)$ for all integers $n \geq n_0$.

My intuition said no, but Viazovska conjectured yes.

We prove this conjecture.

# Interpolation theorem

Theorem. Let $(d, n_0)$ be $(8, 1)$ or $(24, 2)$. Then every radial Schwartz function $f$ on $\mathbb{R}^d$ is uniquely determined by the values $f\left(\sqrt{2n}\right)$, $f'\left(\sqrt{2n}\right)$, $\widehat{f}\left(\sqrt{2n}\right)$, and $\widehat{f}\,'\left(\sqrt{2n}\right)$ for integers $n \geq n_0$.

Specifically, there exists an *interpolation basis* $a_n, b_n, \widetilde{a}_n, \widetilde{b}_n$ for $n \geq n_0$ such that for every radial Schwartz function $f$ and $x \in \mathbb{R}^d$,

$$f(x) = \sum_{n=n_0}^{\infty} f\left(\sqrt{2n}\right) a_n(x) + \sum_{n=n_0}^{\infty} f'\left(\sqrt{2n}\right) b_n(x) + \sum_{n=n_0}^{\infty} \widehat{f}\left(\sqrt{2n}\right) \widetilde{a}_n(x) + \sum_{n=n_0}^{\infty} \widehat{f}\,'\left(\sqrt{2n}\right) \widetilde{b}_n(x).$$

We construct the interpolation basis explicitly. This gives the optimal auxiliary function $f$ via

$$f(x) = \sum_{n \geq n_0} \left( p\left(\sqrt{2n}\right) a_n(x) + p'\left(\sqrt{2n}\right) b_n(x) \right).$$

To describe the interpolation basis, we use the *generating functions*

$$F(\tau, x) = \sum_{n \geq n_0} a_n(x)\, e^{2\pi i n \tau} + 2\pi i \tau \sum_{n \geq n_0} \sqrt{2n}\, b_n(x)\, e^{2\pi i n \tau}$$

and

$$\widetilde{F}(\tau, x) = \sum_{n \geq n_0} \widetilde{a}_n(x)\, e^{2\pi i n \tau} + 2\pi i \tau \sum_{n \geq n_0} \sqrt{2n}\, \widetilde{b}_n(x)\, e^{2\pi i n \tau}$$

for $x \in \mathbb{R}^d$ and $\mathrm{Im}(\tau) > 0$.

Note that $f(x) = F(\tau, x)$ for the complex Gaussian $p(r) = e^{\pi i \tau r^2}$, since $p\left(\sqrt{2n}\right) = e^{2\pi i n \tau}$ and $p'\left(\sqrt{2n}\right) = 2\pi i \tau \sqrt{2n} e^{2\pi i n \tau}$.

The Fourier transform of $x \mapsto e^{\pi i \tau |x|^2}$ is $x \mapsto (i/\tau)^{d/2} e^{\pi i (-1/\tau)|x|^2}$. In particular, the interpolation formula for complex Gaussians amounts to

$$F(\tau, x) + (i/\tau)^{d/2} \widetilde{F}(-1/\tau, x) = e^{\pi i \tau |x|^2}.$$

(They span a dense subspace of radial Schwartz functions.)

These generating functions are not quite Fourier series, because they involve both $e^{2\pi i n \tau}$ and $\tau e^{2\pi i n \tau}$, but they satisfy

$$F(\tau + 2, x) - 2F(\tau + 1, x) + F(\tau, x) = 0,$$
$$\widetilde{F}(\tau + 2, x) - 2\widetilde{F}(\tau + 1, x) + \widetilde{F}(\tau, x) = 0.$$

If $F$ and $\widetilde{F}$ satisfy these three functional equations and suitable smoothness and growth conditions,

then the interpolation theorem follows.

Recall the weight $k$ *slash operator* from modular forms: for

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

we define the action $f|_k\gamma$ of $\gamma$ on a function $f$ by

$$(f|_k\gamma)(z) = (cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right).$$

We often write $|_k^z$ or $|_k^\tau$ to indicate which of several variables is being used.

This action of $\mathrm{SL}_2(\mathbb{Z})$ will play a key role in analyzing the functional equations for $F$ and $\widetilde{F}$.

For even weight, $-I$ acts trivially and we really have an action of $\mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{\pm I\}$.

The group $\mathrm{SL}_2(\mathbb{Z})$ is generated by

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which act by linear fractional transformations as $z \mapsto z + 1$ and $z \mapsto -1/z$.

Note that

$$\begin{aligned} (i/\tau)^{d/2} \widetilde{F}(-1/\tau, x) &= \widetilde{F}(\tau, x)|^{\tau}_{d/2} S, \\ F(\tau + 1, x) &= F(\tau, x)|^{\tau}_{d/2} T, \\ F(\tau + 2, x) &= F(\tau, x)|^{\tau}_{d/2} T^2. \end{aligned}$$

Extend the slash action linearly to the group algebra

$$R = \mathbb{C}[\mathrm{PSL}_2(\mathbb{Z})]$$

consisting of formal linear combinations of elements of $\mathrm{PSL}_2(\mathbb{Z})$.

Then the function equations for $F$ and $\widetilde{F}$ amount to

$$F(\tau, x) + \widetilde{F}(\tau, x)|^{\tau}_{d/2} S = e^{\pi i \tau |x|^2},$$

$$F(\tau, x)|^{\tau}_{d/2} (T - I)^2 = 0,$$

$$\widetilde{F}(\tau, x)|^{\tau}_{d/2} (T - I)^2 = 0.$$

We use the first equation to eliminate $\widetilde{F}$, so all we need to do is solve the remaining equations for $F$.

After we eliminate $\widetilde{F}$, the functional equations amount to

$$F|^{\tau}_{d/2}(T-I)^2 = 0,$$

$$F|^{\tau}_{d/2}S(T-I)^2 = e^{\pi i\tau|x|^2}|^{\tau}_{d/2}S(T-I)^2.$$

In other words, they specify the action of the right ideal $\mathcal{I} := (T-I)^2R + S(T-I)^2R$ of the ring $R$.

One can show that $\dim_{\mathbb{C}}(R/\mathcal{I}) = 6$, so the $\mathcal{I}$ action provides substantial information about $F$ (but not quite everything).

To determine $F$ explicitly, we use a Laplace transform.
We write

$$F(\tau, x) = e^{\pi i \tau |x|^2} + 4\sin^2(\pi |x|^2/2) \int_0^\infty K(\tau, it) e^{-\pi |x|^2 t}\, dt$$

for some kernel $K(\tau, z)$. (Motivation: force $F$ to agree with complex Gaussian at desired points.)

The functional equations then amount to saying that

$K(\tau, z)$ is annihilated by $\mathcal{I}$ under $|^\tau_{d/2}$,

and has poles with certain residues when $z \in \mathrm{SL}_2(\mathbb{Z}) \cdot \tau$.

Why is this helpful? The only remnant of the inhomogeneity of the functional equations is the specified residues.

We can now write down the kernel $K$ explicitly using

1. modular forms for $\mathrm{SL}_2(\mathbb{Z})$ and $\Gamma(2)$,
2. the quasimodular form $E_2$,
3. a holomorphic logarithm of the Hauptmodul $\lambda$ for $\Gamma(2)$, and
4. lots of algebra in $R = \mathbb{C}[\mathrm{PSL}_2(\mathbb{Z})]$.

Once we have the kernel, we can prove the interpolation theorem by verifying analytic continuation and growth bounds.

Proving universal optimality requires some additional inequalities (namely $f(x) \leq p(|x|)$ and $\widehat{f}(y) \geq 0$), which can be verified by unpleasant calculations.

# Open questions

What happens in $\mathbb{R}^2$? This case matters for real-world physics.

The analogous interpolation theorem does not seem to be true in $\mathbb{R}^2$. Can it be salvaged? (For comparison, $\mathbb{R}^1$ is very different from $\mathbb{R}^8$ or $\mathbb{R}^{24}$.)

Can one give a simpler proof of the interpolation theorem, if one doesn't care about writing down an explicit interpolation basis?

How does the interpolation theorem generalize? Which values/derivatives of $f$ and $\widehat{f}$ suffice to reconstruct a radial Schwartz function?

# For more information

Henry Cohn, Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska

*Universal optimality of the $E_8$ and Leech lattices and interpolation formulas*

arXiv:1902.05438