Gaussian energy in high dimensions

Henry Cohn
Microsoft Research and MIT

Thurston Lectures
UC Davis
May 14, 2025

Where does symmetry come from?

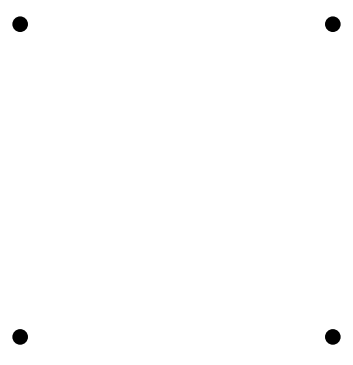
Genetics vs. systematology of the regular figures (L. Fejes Tóth):

> *"Regular arrangements are generated from unarranged, chaotic sets by the ordering effect of an economy principle, in the widest sense of the word."*

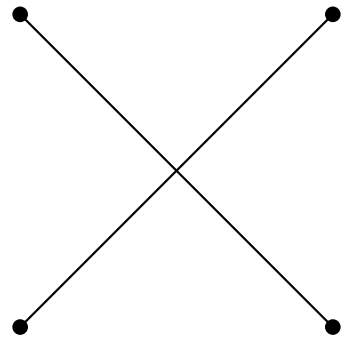
Framework: optimization problems

How symmetrical should we expect the solution of an optimization problem to be?

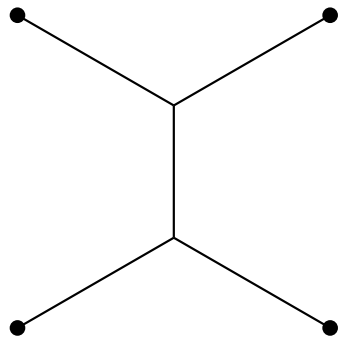
Toy example: a Steiner tree connecting the vertices of a square. What is the minimal-length connecting path?

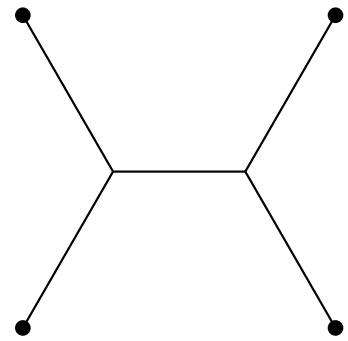


An X is suboptimal:



Of course, these are the optimal solutions (2-d soap films):





The symmetry of the problem is partially broken in each individual solution, but the set of all solutions retains the full symmetry.

How much symmetry should we expect?

It is amazing that crystals have so much symmetry, yet they break rotational symmetry (and even most translational symmetry). How does long-range order develop?

Symmetry makes it easier to solve problems, or at least guess the answers.

Which optimization problems have remarkably symmetrical answers? How can we understand why?

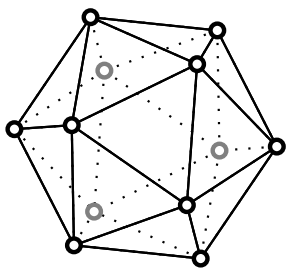
Examples

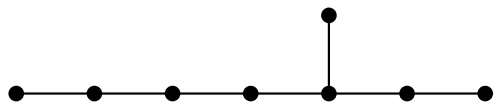
Regular polytopes

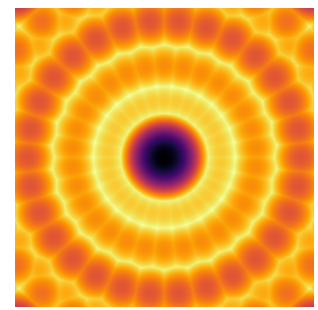
Finite simple groups

ADE classification

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







Setting: energy minimization

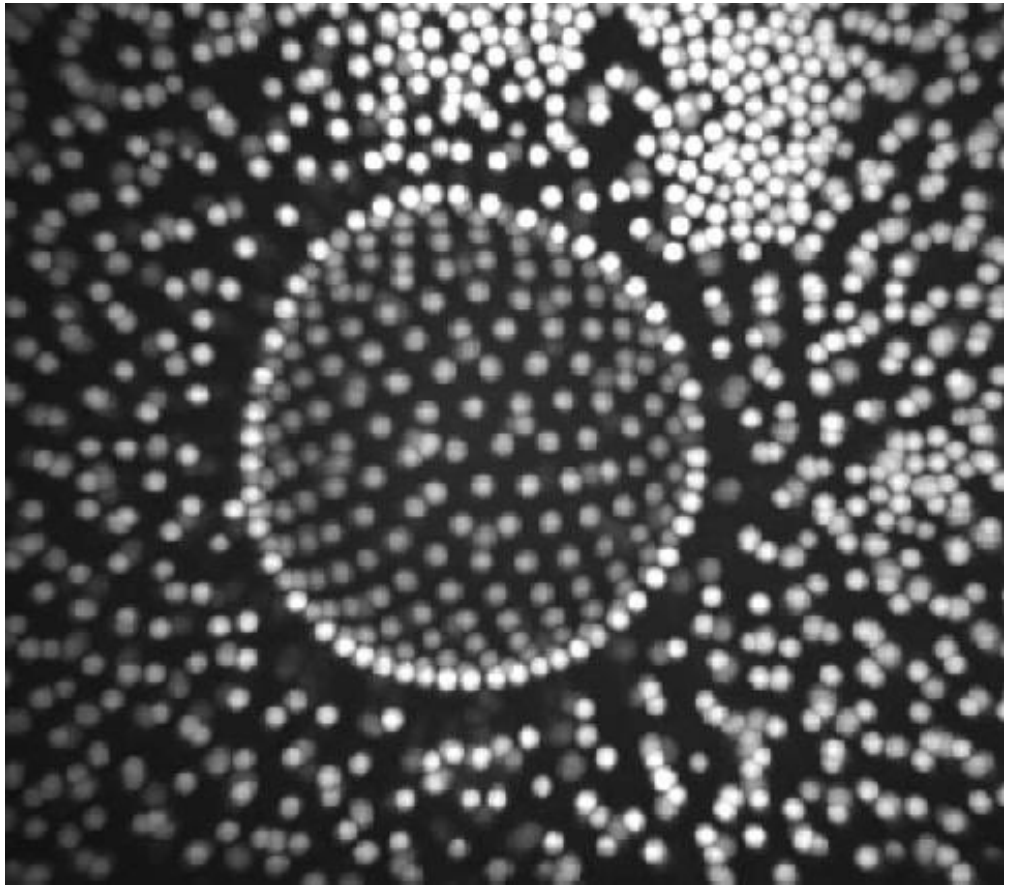
Given a collection of particles interacting according to some potential function, what do they do?

For example, identical charged particles with Coulomb interactions (potential energy $1/r$ for a pair of particles at distance r).
"Thompson problem" on S^2.

Simplest question: what is the ground state? I.e., minimal energy configuration, behavior at zero temperature.

Real physical systems

PMMA beads suspended in mixture of water and cyclohexyl bromide



(W. Irvine and P. Chaikin)

Today's story

Ground states in 8 and 24 dimensions

Bounds based on pair correlation functions

Connections with Fourier interpolation

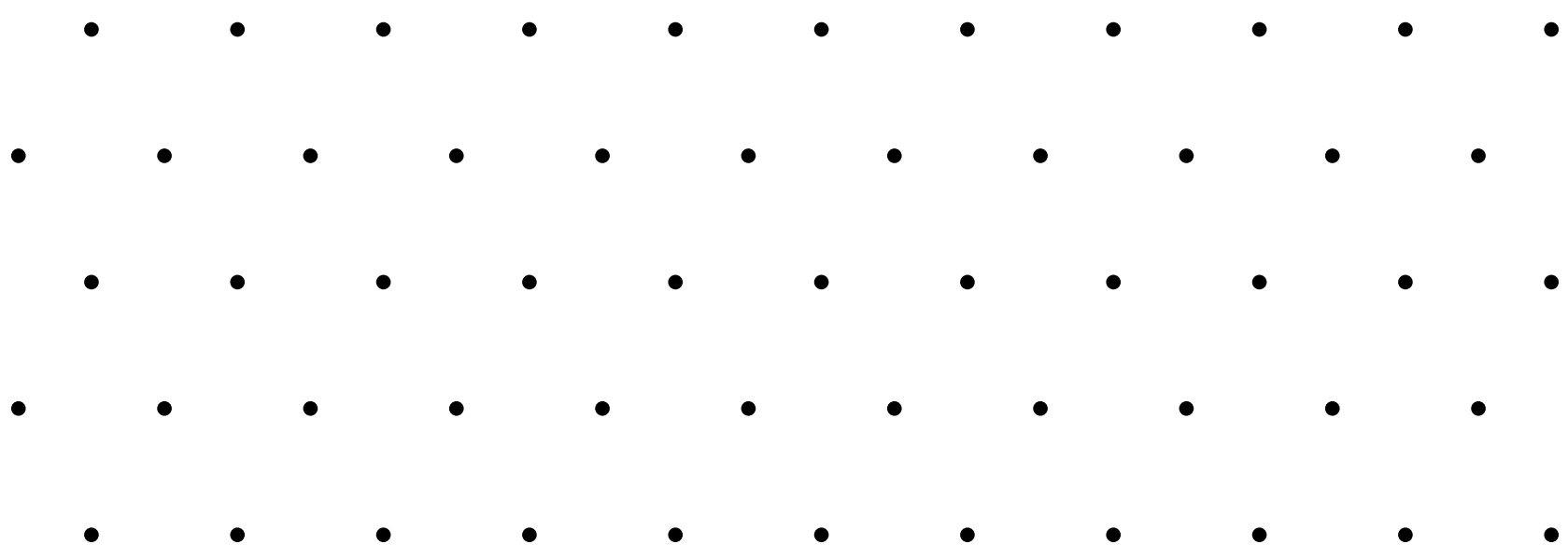
High-dimensional behavior

Broader context

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 subset $\mathcal{C}$.

Pair potential function p, such as

$$p(r) = 1/r^s \text{ with } s > 0 \text{ or}$$
$$p(r) = e^{-\alpha r^2} \text{ with } \alpha > 0.$$

Decreasing $\Leftrightarrow$ repulsive. Sphere packing is limit as s or $\alpha \to \infty$.

We'll fix density of δ particles per unit volume in space.

Energy of particle $x \in \mathcal{C}$ is

$$E_{p,x}(\mathcal{C}) = \sum_{y \in \mathcal{C} \setminus \{x\}} p(|x - y|).$$

Energy $E_p(\mathcal{C})$ is average of $E_{p,x}(\mathcal{C})$ over $x \in \mathcal{C}$. (Define via limit.)

Recall that a *lattice* in $\mathbb{R}^d$ is the $\mathbb{Z}$-span of a basis of $\mathbb{R}^d$. I.e., it is a regular grid, but not necessarily right-angled.

A lattice Λ has density

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

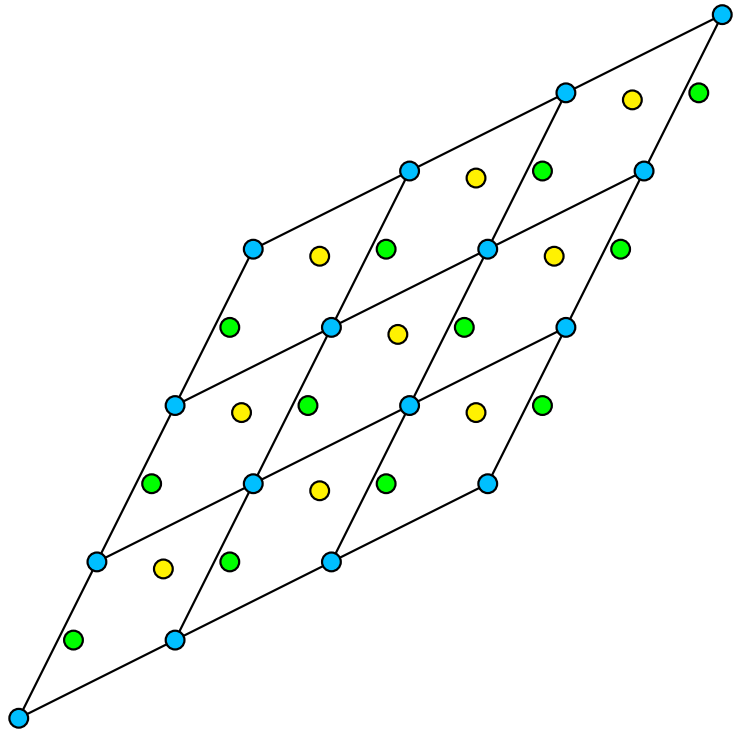
and energy

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

since every point looks the same.

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

In a *periodic configuration*, particles are not restricted to just the vertices of a fundamental cell of a lattice.



Not all important arrangements are periodic. Other possibilities: disordered, quasicrystal, phase coexistence, ...

Ground states

$\mathcal{C}$ in $\mathbb{R}^d$ of density δ is a *ground state*, or *minimizes energy*, for potential p if no other such configuration has lower energy under p.

Ground states describe behavior at zero temperature.

Sphere packing is a limiting case. Maximizing the minimal distance between particles at a fixed particle density 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 two dimensions we generally can't explain what is seen in simulations.

Number theory

A lattice Λ 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 Λ 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^∞ and for all k, $(-1)^k p^{(k)} \geq 0$. Nonnegative, decreasing, convex, etc.

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 (Ventevogel 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 δ? Universal optimality fails.

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

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

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

$$0.99899854\ldots < \delta < 1.00100312\ldots.$$

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

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

Universal optimality in $\mathbb{R}^8$ and $\mathbb{R}^{24}$

Theorem (Cohn, Kumar, Miller, Radchenko, and Viazovska, 2022). 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.

Pair correlation function

To prove this theorem, we use linear programming bounds based on pair correlations.

Let $\mathcal{C}$ be a point configuration in a metric space X with metric d.

The *pair correlation function* $p_{\mathcal{C}}$ for $\mathcal{C}$ counts how many times each distance occurs between points in $\mathcal{C}$.

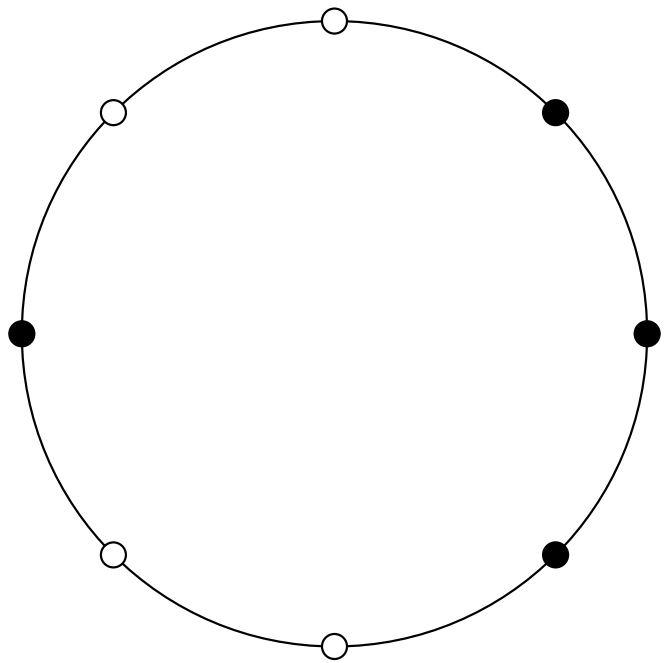
When $\mathcal{C}$ is finite, $p_{\mathcal{C}} \colon [0, \infty) \to \mathbb{N}$ is defined by

$$p_{\mathcal{C}}(r) = \frac{1}{|\mathcal{C}|} \#\{(x, y) \in \mathcal{C}^2 : d(x, y) = r\}.$$

When $\mathcal{C}$ is infinite, we should take a suitable limit to get a distribution, rather than a function.

The pair correlation function for $\mathcal{C}$ does not necessarily suffice to reconstruct $\mathcal{C}$.

In Euclidean space, it would be enough if we knew which distances correspond to which pairs of points, but we don't. For example, the black and white configurations are not isometric, despite having the same pair correlation functions:



(Fun exercise: how does this example generalize?)

This is the "phase problem" in X-ray crystallography. If $\mathcal{C}$ is the locations of the atoms in a crystal sample, then X-ray diffraction amounts to taking the Fourier transform $\widehat{P}$ of the particle distribution

$$P = \sum_{x \in \mathcal{C}} \delta_x.$$

We have

$$\widehat{P}(t) = \sum_{x \in \mathcal{C}} e^{2\pi i \langle x,t \rangle}.$$

The scattering intensity $|\widehat{P}|^2$ yields the pair correlations via

$$|\widehat{P}(t)|^2 = \sum_{x,y \in \mathcal{C}} e^{2\pi i \langle x-y,t \rangle}.$$

In fact, it gives enough information to reconstruct the set $\{x - y : x, y \in \mathcal{C}\}$, but reconstructing $\mathcal{C}$ requires the phase of $\widehat{P}$.

The identity

$$\sum_{x,y \in \mathcal{C}} e^{2\pi i \langle x-y,t \rangle} = |\widehat{P}(t)|^2 \geq 0$$

is a key positivity property for the difference set $\{x - y : x, y \in \mathcal{C}\}$ and therefore the pair correlation function. It occurs in various fields:

In physics, the nonnegativity of the structure factor.

In statistics, the autocorrelation matrix is positive semidefinite.

In quantum field theory, it underlies the conformal bootstrap.

Our proof is based on this inequality. It's noteworthy that this pair correlation inequality is enough information to settle 8 and 24 dimensions, but nowhere near enough information for 3 dimensions.

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."

We will 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 $\delta > 0$. 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 δ has p-energy at least $\delta\widehat{f}(0) - f(0)$.

In other words, f satisfying inequalities (1) and (2) certifies a lower bound for energy. It turns out (2) is related to pair correlation positivity.

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 δ?
Nobody knows, except for $d \in \{1, 8, 24\}$.

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

d	Lower 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)

Both columns are based on numerical optimization.

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 trickier 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 resolved!

Theorem (Cohn, Kumar, Miller, Radchenko, and Viazovska, 2022). 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 Λ in $\mathbb{R}^d$

Poisson summation says

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

Thus,

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

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 (Cohn, Kumar, Miller, Radchenko, and Viazovska, 2022). 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 characterize the basis, we use generating functions. The interpolation theorem is then equivalent to constructing generating functions that satisfy certain functional equations, as well as smoothness and growth conditions, and these functional equations can be solved using integral transforms of modular forms.

High dimensions

What happens in $\mathbb{R}^d$ with particle density δ and potential function $r \mapsto e^{-\alpha r^2}$, with δ and α fixed and $d \to \infty$?

Theorem (Cohn and de Courcy-Ireland, 2018).
If $\alpha < 4\pi/e$, then the minimal energy is $(\delta + o(1))(\pi/\alpha)^{d/2}$ as $d \to \infty$, which is the average energy of a random lattice.
If $\alpha > \pi e$, then the minimal energy is exponentially lower than the average energy of a random lattice.

We do not know where this transition takes place, other than somewhere between $4\pi/e$ and πe, or whether there are further transitions.

The Siegel mean value theorem

What does it mean to average over all lattices? Recall lattice = discrete subgroup of $\mathbb{R}^d$ with rank d.

It's not quite obvious, but there is a canonical probability measure on lattices with fixed determinant (i.e., fundamental cell volume). Key concept: $\mathrm{SL}_d(\mathbb{R})$-invariance.

What does the average pair correlation function look like? It measures the average number of neighbors at each distance.

Siegel mean value theorem: it is exactly the same as for a Poisson distribution (uniformly scattered points).

More precisely: given nice $f : \mathbb{R}^d \to \mathbb{R}$, the average of

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

over all lattices Λ of particle density δ equals

$$\delta \int_{\mathbb{R}^d} f(x)\, dx.$$

Why is this true? There is enough symmetry to rule out any other possible answer.

Specifically, by linearity the answer must be $\int f \, d\mu$ for some regular Borel measure μ on $\mathbb{R}^d \setminus \{0\}$ that is invariant under $\mathrm{SL}_d(\mathbb{R})$. This is only one such measure, up to scaling, and some simple consistency checks determine the constant of proportionality.

Meta principle: averaging over all possible structures is the same as having no structure at all.

This is certainly not always true. It generally depends on having a big enough symmetry group.

Breaks down a little for higher correlation functions, but get Poisson statistics modulo obvious restrictions (e.g., three points in a row for three-point correlations).

Upper bound

The average energy under $r \mapsto e^{-\alpha r^2}$ of a random lattice in $\mathbb{R}^d$ with δ points per unit volume in space is

$$\delta \int_{\mathbb{R}^d} e^{-\alpha |x|^2} \, dx = \delta (\pi/\alpha)^{d/2}.$$

Thus, there must be a lattice of at most this energy.

When $\alpha > \pi e$, we can get lower energy by conditioning on not having any especially short lattice vectors.

Lower bound

We get a lower bound by applying the Shannon sampling theorem to approximate the potential function from below by a bandlimited function.

This works up to $\alpha < 4\pi/e$, but for $\alpha < \pi$ there is a simpler proof:

In the linear programming bound, use the auxiliary function $f(r) = e^{-\alpha r^2}$. It is equal to the potential function, and it is positive definite, with a resulting lower bound of

$$\delta \widehat{f}(0) - f(0) = \delta(\pi/\alpha)^{d/2} - 1.$$

If $\alpha < \pi$, then the -1 is negligible as $d \to \infty$ and we get a lower bound of $(\delta + o(1))(\pi/\alpha)^{d/2}$.

For $\pi \leq \alpha < 4\pi/e$, getting such a bound is not nearly as simple (as far as we know), and of course this lower bound is false for $\alpha > \pi e$.

Open questions

What happens as the dimension tends to infinity?

How can one prove universal optimality in $\mathbb{R}^2$? Lots of real-world materials involve two-dimensional interfaces.

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?

How does all of this extend to higher-order correlations?

Henry Cohn, Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska, *Universal optimality of the E_8 and Leech lattices and interpolation formulas*, Annals of Mathematics **196** (2022), 983–1082, arXiv:1902.05438.

Henry Cohn and Matthew de Courcy-Ireland, *The Gaussian core model in high dimensions*, Duke Mathematical Journal **167** (2018), 2417–2455, arXiv:1603.09684.