## A macro-meso-microscopic view of disk packing

Jim Propp, UMass Lowell Brown University Mathematics Colloquium March 7, 2018 (revised Feb. 12, 2020)

(based on work with Henry Cohn and Omer Tamuz, aided by helpful conversations with Tibor Beke, Ilya Chernykh, David Feldman, Boris Hasselblatt, Alex Iosevich, Sinai Robins, and MathOverflow)

> These slides can be found at http://jamespropp.org/brown18a.pdf.

In 2016 there were breakthroughs in sphere-packing in 8 dimensions and 24 dimensions.

In 2016 there were breakthroughs in sphere-packing in 8 dimensions and 24 dimensions.

This work came on the heels of earlier work on sphere packing in 3 dimensions, from Johannes Kepler to Thomas Hales, as well as work on sphere packing in 2 dimensions, commenced by Axel Thue and László Fejes-Tóth. In 2016 there were breakthroughs in sphere-packing in 8 dimensions and 24 dimensions.

This work came on the heels of earlier work on sphere packing in 3 dimensions, from Johannes Kepler to Thomas Hales, as well as work on sphere packing in 2 dimensions, commenced by Axel Thue and László Fejes-Tóth.

But for other values of n, we know very little, and I suspect that we still aren't asking exactly the right question.

It's "obvious" to any child that in two dimensions, the optimal packing is the six-around-one hexagonal close packing.

It's "obvious" to any child that in two dimensions, the optimal packing is the six-around-one hexagonal close packing.

I want to take a fresh look at n = 2.

It's "obvious" to any child that in two dimensions, the optimal packing is the six-around-one hexagonal close packing.

I want to take a fresh look at n = 2.

It's my hope that by sharpening our notion of what it means for something to be an *optimal* sphere-packing, we'll obtain a coherent (if potentially unattainable) notion of what it might mean to classify the optimal packings in *n* dimensions.

## Part I: Three ways of measuring deviations from optimality

A sphere-packing of  $\mathbb{R}^n$  is a collection of balls in  $\mathbb{R}^n$  with disjoint interiors.

A sphere-packing of  $\mathbb{R}^n$  is a collection of balls in  $\mathbb{R}^n$  with disjoint interiors.

The density of a packing is

$$\lim_{r\to\infty}\lambda(B_r\cap S)/\lambda(B_r),$$

where  $B_r$  is the ball of radius r centered at 0, S is the union of the balls in the packing, and  $\lambda$  is Lebesgue measure (area).

Define  $\Delta_n$  as the supremum of all densities of sphere-packings in  $\mathbb{R}^n$ .

Define  $\Delta_n$  as the supremum of all densities of sphere-packings in  $\mathbb{R}^n$ .

It's not hard to show that some packing achieves this supremum.

Define  $\Delta_n$  as the supremum of all densities of sphere-packings in  $\mathbb{R}^n$ .

It's not hard to show that some packing achieves this supremum.

Trivially,  $\Delta_1 = 1$ .

The density of the 4-around-1 disk-packing is  $\pi/4 \approx 0.79$ .



The density of the 6-around-1 disk-packing is  $\pi/\sqrt{12} \approx 0.91$ .



Axel Thue and László Fejes-Tóth's theorem:  $\Delta_2 = \pi/\sqrt{12}$ .

Partial converse: Let P be a *periodic* packing of the plane by unit disks. If P achieves density  $\Delta_2$ , then P is a hexagonal close packing.

Partial converse: Let P be a *periodic* packing of the plane by unit disks. If P achieves density  $\Delta_2$ , then P is a hexagonal close packing.

Can we delete the word "periodic"?

Partial converse: Let P be a *periodic* packing of the plane by unit disks. If P achieves density  $\Delta_2$ , then P is a hexagonal close packing.

Can we delete the word "periodic"?

No.

Counterexample 1:



## Counterexample 2:



## Counterexample 3:



Counterexample 4:

There exists a disk-packing of density  $\Delta_2$  in which **no** two disks touch.

Morally, these four counterexamples seem like "cheats".

How can we sharpen our notion of optimal packing to rule them out?

I'm going to show you three possible approaches to doing this.

In the first approach, we replace disk packings by point packings P (using the centers of the disks, no two of which are at distance < 2).

The size of the infinite set *P* can be represented as the divergent infinite sum  $\sum_{(x,y)\in P} 1$ .

We'll regularize this sum by imposing a smooth cutoff at distance s from the origin, and then letting s go to infinity.

Let 
$$g_s(x, y) := \exp -(x^2 + y^2)/s^2$$
, so that  
(1) for all  $(x, y) \in \mathbb{R}^2$ ,  $g_s(x, y) \to 1$  as  $s \to \infty$ , and  
(2) for all  $s > 0$ ,  $\int_{\mathbb{R}^2} g_s < \infty$ .

Let 
$$g_s(x, y) := \exp -(x^2 + y^2)/s^2$$
, so that  
(1) for all  $(x, y) \in \mathbb{R}^2$ ,  $g_s(x, y) \to 1$  as  $s \to \infty$ , and  
(2) for all  $s > 0$ ,  $\int_{\mathbb{R}^2} g_s < \infty$ .

For P a discrete point-set in  $\mathbb{R}^2$  containing no two points at distance  $\leq 2$  (so that the unit disks centered at points in P have disjoint interiors), let

$$|P|_s := \sum_{(x,y)\in P} g_s(x,y) < \infty.$$

Let 
$$g_s(x, y) := \exp -(x^2 + y^2)/s^2$$
, so that  
(1) for all  $(x, y) \in \mathbb{R}^2$ ,  $g_s(x, y) \to 1$  as  $s \to \infty$ , and  
(2) for all  $s > 0$ ,  $\int_{\mathbb{R}^2} g_s < \infty$ .

For P a discrete point-set in  $\mathbb{R}^2$  containing no two points at distance  $\leq 2$  (so that the unit disks centered at points in P have disjoint interiors), let

$$|P|_s := \sum_{(x,y)\in P} g_s(x,y) < \infty.$$

If P is finite,  $|P|_s \rightarrow |P|$  as  $s \rightarrow \infty$ .

**Main idea**: When *P* is infinite,  $|P|_s$  diverges as  $s \to \infty$ , but the precise way it diverges gives information about the point set *P*.

**Main idea**: When *P* is infinite,  $|P|_s$  diverges as  $s \to \infty$ , but the precise way it diverges gives information about the point set *P*.

In particular, for many non-optimal packings P, we can expand  $|P|_s$  as  $\alpha s^2 + \beta s + \gamma + o(1)$ , where

- $\blacktriangleright \alpha$  tells us about density,
- $\blacktriangleright \beta$  tells us about line defects, and
- $\blacktriangleright$   $\gamma$  tells us about point defects.

Conjectural converse: Let *P* be a point-packing of the plane. If  $|P|_s = \Delta_2 s^2 + o(1)$ , then *P* is a hexagonal close packing.

Conjectural converse: Let *P* be a point-packing of the plane. If  $|P|_s = \Delta_2 s^2 + o(1)$ , then *P* is a hexagonal close packing.

"Evidence": The four counterexamples I discussed above are not counterexamples to this conjecture.

Conjectural converse: Let *P* be a point-packing of the plane. If  $|P|_s = \Delta_2 s^2 + o(1)$ , then *P* is a hexagonal close packing.

"Evidence": The four counterexamples I discussed above are not counterexamples to this conjecture.

But I've had trouble making progress with this definition.

Second approach: Go back to using disks instead of points. Let  $\mu_P$  be Lebesgue measure restricted to the union of the disks.


Second approach: Go back to using disks instead of points. Let  $\mu_P$  be Lebesgue measure restricted to the union of the disks.



Let

$$(P)_s = \int_{\mathbb{R}^2} g_s \, d\mu_P.$$

Compare:  $|P|_s$  can be defined as the integral of  $g_s$  with respect to a "Dirac comb": a measure on  $\mathbb{R}^2$  in which each point in P is assigned mass 1.

Compare:  $|P|_s$  can be defined as the integral of  $g_s$  with respect to a "Dirac comb": a measure on  $\mathbb{R}^2$  in which each point in P is assigned mass 1.

I will sometimes call  $|P|_s$  the Dirac regularization of the divergent sum  $\sum_{(x,y)\in S} 1$ , in contrast to the disk regularization  $(P)_s$  (and in contrast to the Delaunay regularization  $[P]_s$  to be described next).

Third approach ("gerrymandering"): Redistribute the mass of the disks, so that the mass associated with a sector of a disk gets reapportioned uniformly throughout its Delaunay cell.

(Note: "Delaunay" = "Delone".)

In a point-packing, a Delaunay cell is an inscriptible polygon whose vertices are points of the packing and whose circumcircle encircles no points of the packing.



Let  $\bar{\mu}_P$  be the reapportioned measure, and let

$$[P]_s = \int_{\mathbb{R}^2} g_s \ d\bar{\mu}_P.$$

**Main Theorem**: Let *P* be a distance-2 point-packing. Then  $[P]_s = \pi \Delta_2 s^2 + o(1)$  (i.e.,  $[P]_s - \pi \Delta_2 s^2 \rightarrow 0$ ) if and only if *P* is a hexagonal close packing.

## Part II: A non-Archimedean valuation on polyhedral sets

A valuation is a finitely-additive measure from a set-algebra into an abelian group:  $v(A \cup B) = v(A) + v(B) - v(A \cap B)$ , inclusion-exclusion, etc.

A valuation is a finitely-additive measure from a set-algebra into an abelian group:  $v(A \cup B) = v(A) + v(B) - v(A \cap B)$ , inclusion-exclusion, etc.

A **polyhedral set** in  $\mathbb{R}^n$  is a set specified by some Boolean formula in *n* variables involving finitely many linear equations and inequalities.

Equivalently, it's a set that belongs to the set-algebra generated by open and closed half-planes.

**Theorem** ("Throwing Gaussian darts at a noncompact target"; new?): If S is a polyhedral subset of  $\mathbb{R}^n$ , the probability that a Gaussian  $N(0, \sigma)$  random variable lies in S is given by  $p(\sigma)/\sigma^n$  plus an error term that goes to zero, where  $p(\cdot)$  is a polynomial of degree n.

Proving this theorem is essentially equivalent to setting up a valuation  $v_n$  on polyhedral sets that assigns to each (not necessarily compact) polyhedral subset of  $\mathbb{R}^n$  a generalized *n*-dimensional volume in a non-Archimedean ordered ring extension of  $\mathbb{R}$ .

If S is a compact polyhedral subset of  $\mathbb{R}^n$ ,  $v_n(S)$  will be the ordinary *n*-dimensional Lebesgue measure of S, but if S is a noncompact polyhedral set (of full dimension),  $v_n$  will be an "infinite" element of the non-Archimedean ordered ring  $\mathbb{R}[p]$ , where p is a formal infinite element satisfying  $1 \ll p \ll p^2 \ll \cdots$ .

I'll sometimes informally call the valuation  $v_n$  a "measure" even though it's not countably additive or real-valued.

I'll focus on the cases of  $\mathbb{R}$  and  $\mathbb{R}^2$ . I'll sometimes call  $v_1$  "length" and  $v_2$  "area", omitting the modifier "generalized".

Warning: Often a translation T carries a polyhedral set S into a proper subset or proper superset of itself.

E.g., 
$$S = [0, +\infty)$$
 or  $(-\infty, 0]$  in  $\mathbb{R}$ ,  $T : x \mapsto x + 1$ .

In cases like this, we can't expect S and T(S) to have the same measure.

But as we'll see there are compensations for this lack of symmetry.

The polyhedral subsets of  $\mathbb{R}^1$  are unions of isolated points, finite open intervals and infinite open rays.

Isolated points have measure 0.

Define  $v_1(I) := \text{length}(I)$  if I is a finite interval,  $v_1([x, +\infty)) = \mathfrak{p} - x$ , and  $v_1((-\infty, x)) = \mathfrak{p} + x$ .

In particular, 
$$v_1([0, +\infty)) = v_1((-\infty, 0]) = \mathfrak{p}$$
,  
 $v_1((-\infty, +\infty)) = 2\mathfrak{p} > \mathfrak{p}$ .

Note that  $v_1$  is invariant under rotation, aka negation:  $v_1(-A) = v_1(A)$ . For A, B polyhedral subsets of  $\mathbb{R}$ , define  $v_2(A \times B)$  to be  $v_1(A)v_1(B)$  (Fubini formula).

E.g., 
$$v_2([0,\infty) \times [0,\infty)) = \mathfrak{p}^2$$
, and more broadly,  
 $v_2([x,\infty) \times [x',\infty)) = (\mathfrak{p}-x)(\mathfrak{p}-x') = \mathfrak{p}^2 - (x+x')\mathfrak{p} + xx'.$ 

Special case: x' = -x.  $v_2([x, \infty) \times [-x, \infty)) = \mathfrak{p}^2 - (x + x')\mathfrak{p} + xx' = \mathfrak{p}^2 - x^2$ .

Special case: 
$$x' = -x$$
.  
 $v_2([x, \infty) \times [-x, \infty)) = \mathfrak{p}^2 - (x + x')\mathfrak{p} + xx' = \mathfrak{p}^2 - x^2$ .  
Picture proof  $(x \ge 0)$ :



This picture-proof hinges on the easily checked fact that  $v_2$  is invariant under 90-degree rotations:

$$v_2((-B) \times A) = v_1(-B)v_1(A) = v_1(A)v_1(B) = v_2(A \times B).$$

This picture-proof hinges on the easily checked fact that  $v_2$  is invariant under 90-degree rotations:

$$v_2((-B) \times A) = v_1(-B)v_1(A) = v_1(A)v_1(B) = v_2(A \times B).$$

But how do we define/compute  $v_2(S)$  when S is not a rectangle or a finite union of rectangles?

Example:  $S = \{(x, y) : 0 \le y \le x\}.$ 



We might say that this set "should" have generalized area  $(1/2)p^2$ , since two copies of it (one obtained from the other by a 45 degree rotation about the origin) form a quadrant whose generalized area is  $p^2$ .

Theorem: There is a unique valuation on polyhedral sets taking values in the ordered ring  $\mathbb{R}[p]$  satisfying the following four properties:

(1) Monotonicity: If S is a subset of S',  $v_2(S) \leq v_2(S')$ .

(2) Consistency with Lebesgue measure: If S is compact,  $v_2(S)$  is the Lebesgue measure of S.

(3) Fubini: If  $S = A \times B$ ,  $v_2(S) = v_1(A)v_1(B)$ .

(4) Rotational invariance: If S and S' are related by rotation about the origin in  $\mathbb{R}^2$ ,  $v_2(S) = v_2(S')$ .

What do these properties have to do with asymptotics of the integral of  $e^{-(x^2+y^2)/s^2}$  over a set *S*?

(1) Monotonicity: The integrand is nonnegative.

(2) Consistency with Lebesgue measure: When S is compact, the integrand goes to 1 uniformly.

(3) Fubini: The integrand factors as  $e^{-x^2/s^2}e^{-y^2/s^2}$ .

(4) Rotational invariance: The integrand is invariant under rotation.

Why is  $v_2$  is uniquely determined by properties (1)-(4)? Proof sketch: Why is  $v_2$  is uniquely determined by properties (1)-(4)?

Proof sketch:

Triangulate S, using a mix of ordinary triangles, ideal triangles with 1 point at infinity, and ideal triangles with 2 points at infinity.

Show that for each kind of triangle T, properties (1)-(4) uniquely determine the area of T.





 $\int_{[0,\mathfrak{p})} x \, dx$ 



$$\int_{[0,p)} x \, dx = x^2/2 \mid_0^p$$



$$\int_{[0,\mathfrak{p})} x \, dx = x^2/2 \mid_0^{\mathfrak{p}} = \mathfrak{p}^2/2.$$





$$\int_{[0,\mathfrak{p})}(\mathfrak{p}-y)\,dy$$



$$\int_{[0,\mathfrak{p})}(\mathfrak{p}-y)\,dy=(\mathfrak{p}y-y^2/2)\,|_0^\mathfrak{p}$$



$$\int_{[0,\mathfrak{p})}(\mathfrak{p}-y)\,dy=(\mathfrak{p}y-y^2/2)\mid_0^\mathfrak{p}=\mathfrak{p}^2-\mathfrak{p}^2/2$$



$$\int_{[0,\mathfrak{p})}(\mathfrak{p}-y)\,dy=(\mathfrak{p}y-y^2/2)\mid_0^\mathfrak{p}=\mathfrak{p}^2-\mathfrak{p}^2/2=\mathfrak{p}^2/2.$$



 $\int_{[0,\mathfrak{p})}(\mathfrak{p}-y) \, dy = (\mathfrak{p}y - y^2/2) \mid_0^{\mathfrak{p}} = \mathfrak{p}^2 - \mathfrak{p}^2/2 = \mathfrak{p}^2/2.$ We can develop, in tandem, a way of measuring polyhedral sets and a way of integrating polyhedral functions.

## Part III: Hallways
The vertical strip  $[-1, +1] \times (-\infty, +\infty)$  has generalized area  $(2)(2\mathfrak{p}) = 4\mathfrak{p}$ . So does every translate  $[x - 1, x + 1] \times (-\infty, +\infty)$ . So does every horizontal strip  $(-\infty, +\infty) \times [y - 1, y + 1]$ . The intersection of the two strips has area 4.

By inclusion-excusion, the union of the two strips has measure 4p + 4p - 4 = 8p - 4, and so does every translate of this set.



Putting it differently, the generalized measure assigned to the union of a vertical stripe of width 2 and a horizontal stripe of width 2 is independent of the choice of origin.

(You may have thought that  $\nu_1$  and  $\nu_2$  are peculiar because the origin plays a privileged role. For many shapes, this peculiarity persists; but we now see that for some shapes, it goes away.)

This kind of translation-invariance holds for a class of subsets of  $\mathbb{R}^2$ , consisting of compact polygonal **junctions** (like  $[x - 1, x + 1] \times [y - 1, y + 1]$ ) joined by **halls**, which may be finite or infinite in length; some of these **hallway networks** model simple defects in disk packings.







More generally, consider a hallway with one triangular junction: three singly-infinite strips projecting orthogonally from the sides of triangle *ABC*.



Claim: The sum of the  $v_2$  areas of the three half-strips projecting from the triangle equals p times the perimeter of triangle *ABC*, minus twice the area of triangle *ABC*.



Here's the proof in the case where the origin lies inside ABC. (Replace area by algebraic area to get this to work in the general case.)



For more general hallway networks consisting of finite and singly-infinite hallways joining up finitely many junctions, the sum of the areas of the hallways (not including the junctions) equals p times the sum of the widths of the singly-infinite hallways minus twice the sum of the areas of the junctions.

In the illustrative case of a 2-junction network, we can prove this by inclusion-exclusion. The  $v_2$  area of the halls network equals the  $v_2$  area of a single-junction network plus the  $v_2$  area of another single-junction network minus the  $v_2$  area of a single doubly-infinite strip.





plus



minus

Suppose that all the Delaunay cells of *P* are triangles and squares with side-length 2 forming a network of hallways and junctions, with  $N < \infty$  infinite hallways. (E.g., the last two examples had N = 3 and N = 4.)

## Claim:

$$[P]_{s} = (\Delta_{2}) s^{2} - (N)(2\delta\pi^{1/2}) s + (A)(2\delta) + o(1),$$
  
where  $\Delta_{2} = \pi/\sqrt{12}$ ,  
 $\delta = \pi/\sqrt{12} - \pi/4$ ,  
 $N =$  the number of infinite hallways, and  
 $A =$  the total area of the triangular junctions.

Note that the choice of origin makes no difference.

Application to crystal defects: In a packing that's close to a hexagonal close packing,

**point** defects lead to networks with area on the order of 1, **line** defects lead to networks with area on the order of  $\mathfrak{p}$ , and **density** defects lead to networks with area on the order of  $\mathfrak{p}^2$ (and the location of the origin turns out not to matter).

I'd like to try to apply this "defect calculus" to other sorts of defects in 2D and 3D lattices, but solid state physicists and materials scientists are annoyingly topological; they draw suggestive pictures, but they don't say precisely where the atoms are!

I'm not claiming that the ring  $\mathbb{R}[\mathfrak{p}]$  will be adequate for the study of all defects; there are packings with  $|P|_s$  (or  $[P]_s$ ) growing like  $\Delta_2 p^2 + \Theta(p^c)$  with 1 < c < 2, for instance (see Counterexample 4).

To include such packings in our theory, we look at truncated germs. That is, we look at the behavior of  $|P|_s$  as  $s \to \infty$  modulo  $\equiv$ , where  $P \equiv P'$  iff  $|P|_s - |P'|_s$  goes to 0.

We don't expect trichotomy  $(|P|_s - |P'|_s \text{ could exhibit}$ oscillations with amplitude that don't go to zero), but we don't need trichotomy if what we're after is a sharpened notion of optimality; to say that  $P^*$  is optimal, all we need to know is that for any P,  $\liminf(|P^*|_s - |P|_s) \ge 0$ .

## Part IV: The characterization of maximal disk-packings

Only one ingredient is missing for the proof of the main theorem:

Lemma (original source?): In a disk packing of the plane, no Delaunay cell can have local packing density exceeding  $\pi/\sqrt{12}$ . That is, the fraction of a Delaunay cell that is covered by disks of the disk-packing cannot exceed  $\pi/\sqrt{12}$ . Furthermore, equality holds if and only if the cell is an equilateral triangle.

Recall that  $\bar{\mu}_P$  is Lebesgue measure of intensity 1 on the disks reapportioned uniformly over the Delaunay regions, and that  $[P]_s = \int_{\mathbb{R}^2} g_s \ d\bar{\mu}_P$ .

If  $P^*$  is a hexagonal-close packing,  $\bar{\mu}_P$  is the uniform measure  $\Delta_2 \lambda$  (where  $\lambda$  is Lebesgue measure) on  $\mathbb{R}^2$ .

Otherwise,  $\bar{\mu}_P \leq \Delta_2 \lambda$  everywhere (by the Lemma) but there exists at least one Delaunay cell *C* whose  $\bar{\mu}_P$  measure falls short of  $\Delta_2 \lambda(C)$  by some positive amount, say *c*; then  $\liminf([P^*]_s - [P]_s) \geq c$ , and since  $[P^*]_s = \pi \Delta_2 s^2 + o(1)$ , we cannot have  $[P]_s = \pi \Delta_2 s^2 + o(1)$ .

This completes the proof.

Conjecture: The same is true for  $|\cdot|_s$  (which, unlike  $[\cdot]_s$ , extends naturally to higher dimensions).

Conjecture: the optimal 3D sphere packings are the Barlow packings (the uncountably many packings formed by layers of hexagonal close-packed spheres).

## Part V: Odds and ends

Charles Radin, Lewis Bowen and their collaborators have an interesting approach to optimal packings whose philosophical motivations are similar to ours.

However, the theory I've outlined today applies in situations where theirs doesn't.

Here are two disk-packings of a quadrant, P and P'. P is better at filling the quadrant than P' by exactly 1/4 of a disk, in the sense that  $|P|_s - |P'|_s$ ,  $(P)_s - (P')_s$ , and  $[P]_s - [P']_s$  all converge to 1/4.



Is *P* the best packing of disks in a quadrant?

Do we even know there exists a best packing?

It's not clear to me how to employ compactness principles or contraction arguments or any tools at all to prove existence of an optimum. Going back to disk-packing of the whole plane, it's not obvious that there exists a best *non-optimal* disk-packing. But I claim there is one.

Gap conjecture: The most efficient non-optimal disk-packings are hexagonal close packings with one disk missing.

Note that Counterexample 3 from earlier does not disprove this; no matter how small  $\epsilon$  is, if you displace a half-plane's worth of disks by  $\epsilon$ , the amount of deficiency introduced corresponds to removal of infinitely many disks.

As a side-project, I've studied optimal packing in 1 dimension, with various simplifications:

- 1) discrete, rather than continuous;
- 2) one-sided, rather than two-sided; and
- 3) exponential, rather than Gaussian, regularization.

Despite these simplifications, I still do not see robust methods of proving that optimal packings exist.

See my preprint One-Dimensional Packing: Maximality Implies Rationality.

See also MathOverflow post Comparing sizes of sets of natural numbers.

I'm hoping that other people, with various analytic and geometric insights, will help me figure out how to advance this point of view of packings.

Thank you!

These slides can be found at http://jamespropp.org/brown18a.pdf.