Uncertainty principles and the modular bootstrap

Henry Cohn
Microsoft Research and MIT

Thurston Lectures
UC Davis
May 16, 2025

What's an uncertainty principle?

Tradeoff between properties of $f: \mathbb{R}^d \to \mathbb{C}$ and Fourier transform $\widehat{f}$, defined for integrable f by

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

Here $\langle \cdot, \cdot \rangle$ is the inner product on $\mathbb{R}^d$. Fourier inversion says

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

when $\widehat{f}$ is integrable. I.e., $\widehat{f}$ decomposes f into complex exponentials (pure frequencies), and vice versa.

As a reminder of something we'll use later, Fourier inversion implies

$$||f||_\infty = \sup_x |f(x)| \le \int_{\mathbb{R}^d} |\widehat{f}(y) e^{2\pi i \langle x, y \rangle}|\, dy = \int_{\mathbb{R}^d} |\widehat{f}(y)|\, dy = ||\widehat{f}||_1.$$

Heisenberg's uncertainty principle

A quantum particle's position and momentum cannot simultaneously be pinned down to high precision.

You cannot confine a particle in a small box and keep it still.

This is not a statement about human knowledge or measurement, but rather about quantum reality itself.

Let's use one spatial dimension for notational convenience and choose units with Planck's constant $h = 1$. Then a particle has a wave function $f \in L^2(\mathbb{R})$, normalized with $||f||_2 = 1$.

$|f(x)|^2\,dx$ gives a probability distribution on positions x

$|\widehat{f}(y)|^2\,dy$ gives a probability distribution on momenta y

This is part of the basic setup of quantum mechanics. Now the question becomes how tightly concentrated f and $\widehat{f}$ can be.

$$\text{expected position} = \int_{\mathbb{R}} x|f(x)|^2\,dx$$
$$\text{expected momentum} = \int_{\mathbb{R}} y|\widehat{f}(y)|^2\,dy$$
$$\text{variance of position} = \int_{\mathbb{R}} x^2|f(x)|^2\,dx - \left(\int_{\mathbb{R}} x|f(x)|^2\,dx\right)^2$$
$$\text{variance of momentum} = \int_{\mathbb{R}} y^2|\widehat{f}(y)|^2\,dy - \left(\int_{\mathbb{R}} y|\widehat{f}(y)|^2\,dy\right)^2$$

Heisenberg's uncertainty principle:

$$\text{product of variances} \geq \frac{1}{16\pi^2}$$

If the wave function is concentrated about one location in position space, then its momentum must be spread out, and vice versa.

This inequality is sharp for Gaussians. Let's prove it.

WLOG translate and phase shift so expected position and momentum are zero, and assume differentiable. Want to prove

$$\int_{\mathbb{R}} x^2 |f(x)|^2 \, dx \int_{\mathbb{R}} y^2 |\widehat{f}(y)|^2 \, dy \geq \frac{||f||_2^4}{16\pi^2}.$$

Since differentiating f multiplies $\widehat{f}(y)$ by $2\pi i y$,

$$\int_{\mathbb{R}} y^2 |\widehat{f}(y)|^2 \, dy = \frac{1}{4\pi^2} \int_{\mathbb{R}} |f'(x)|^2 \, dx.$$

Now by Cauchy-Schwarz,

$$\begin{aligned}
\int_{\mathbb{R}} x^2 |f(x)|^2 \, dx \int_{\mathbb{R}} |f'(x)|^2 \, dx &\geq \left(\int_{\mathbb{R}} |x f(x) f'(x)| \, dx \right)^2 \\
&\geq \left(\int_{\mathbb{R}} x \, \mathrm{Re}(f(x) f'(x)) \, dx \right)^2 \\
&= \frac{1}{4} \left(\int_{\mathbb{R}} x \frac{d}{dx} |f(x)|^2 \, dx \right)^2 \\
&= \frac{||f||_2^4}{4}. \qquad \text{[integration by parts]}
\end{aligned}$$

Musical uncertainty

Uncertainty is by no means restricted to quantum mechanics. Instead, it's a basic phenomenon in Fourier analysis.

For example, a musical note cannot be sharply concentrated in both time and frequency.

This sounds weird at first: if a note is very short, then it must be diffuse in frequency.

But musical uncertainty is not so unfamiliar. How do you tune a musical instrument?

Tuning

Compare with a tuning fork and listen for beats as the two waves move in and out of constructive interference.

This basically amounts to the sum-to-product formula

$$\sin 2\pi\alpha x + \sin 2\pi(\alpha+\varepsilon)x = \big(2\cos \pi\varepsilon x\big)\sin 2\pi(\alpha+\varepsilon/2)x.$$

I.e., overlaying pure sine waves of frequencies α and $\alpha+\varepsilon$ sounds like a sine wave of frequency $\alpha+\varepsilon/2$, except modulated with beats at frequency $\varepsilon/2$. Your ears may not detect a frequency difference of order ε, but you can hear the amplitude variation from $\cos \pi\varepsilon x$.

To tune the instrument to within ε, you'll have to listen for a time on the order of $1/\varepsilon$ to detect the beats.

The uncertainty principle says this is optimal to within a constant factor.

Variations

There are many different uncertainty principles, such as:

$$\boxed{f \text{ and } \widehat{f} \text{ cannot both have compact support unless } f = 0.}$$

Proof: if $\widehat{f}$ has compact support, then Fourier inversion shows that f is holomorphic.

Landau, Pollak, and Slepian asked in the 1960s how much of the energy (i.e., L^2 norm) can be concentrated in a given time interval for a band-limited signal (i.e., one with a limited frequency range). Prolate spheroidal wave functions optimize this quantity.

Most uncertainty principles measure disperson or concentration, but in 2010 Bourgain, Clozel, and Kahane developed an uncertainty principle for signs of functions. That will be our topic today.

Uncertainty for signs

$f: \mathbb{R}^d \to \mathbb{R}$ is *eventually nonnegative* if $f(x) \geq 0$ for all sufficiently large $|x|$. Let the *last sign change radius* be

$$r(f) = \inf \{R \geq 0 : f(x) \text{ has the same sign for } |x| \geq R\},$$

and let $\mathcal{A}_+(d)$ be the set of functions $f: \mathbb{R}^d \to \mathbb{R}$ such that

1. $f \in L^1(\mathbb{R}^d)$, $\widehat{f} \in L^1(\mathbb{R}^d)$, and $\widehat{f}$ is real-valued (i.e., f is even),
2. f is eventually nonnegative while $\widehat{f}(0) \leq 0$, and
3. $\widehat{f}$ is eventually nonnegative while $f(0) \leq 0$.

The uncertainty principle of Bourgain, Clozel, and Kahane says

$$\mathrm{A}_+(d) := \inf_{f \in \mathcal{A}_+(d) \setminus \{0\}} \sqrt{r(f) r(\widehat{f})} > 0.$$

I.e., $r(f)$ and $r(\widehat{f})$ cannot both be small. Note that $r(f)r(\widehat{f})$ is invariant under rescaling the input to f.

Reductions

Suppose $f \in \mathcal{A}_+(d) \setminus \{0\}$. I.e., f and $\widehat{f}$ are eventually nonnegative, while $f(0) \le 0$ and $\widehat{f}(0) \le 0$.

1. By rescaling input, assume $r(f) = r(\widehat{f})$ and preserve $r(f)r(\widehat{f})$.
2. Now let $g = f + \widehat{f}$. Then $g(0) \le 0$ and g is eventually nonnegative, with $r(g) \le r(f)$. Furthermore, g is not identically zero (else f and $\widehat{f}$ would have compact support).
3. Now we have a Fourier eigenfunction with eigenvalue $+1$: $\widehat{g} = g$.
4. We can assume $g(0) = 0$. Otherwise, add a positive multiple of $x \mapsto e^{-\pi|x|^2}$ to g.

Problem ($+1$ eigenfunction uncertainty principle)

Minimize $r(g)$ over all $g\colon \mathbb{R}^d \to \mathbb{R}$ such that

1. $g \in L^1(\mathbb{R}^d) \setminus \{0\}$ *and* $\widehat{g} = g$, *and*
2. $g(0) = 0$ *and* g *is eventually nonnegative.*

Proof of BCK uncertainty principle

Normalize $\|g\|_1 = 1$, and let $g^+ = \max\{g, 0\}$ and $g^- = \max\{-g, 0\}$. Since $\int_{\mathbb{R}^d} g = \widehat{g}(0) = 0$,

$$\int_{\mathbb{R}^d} g^+ = \int_{\mathbb{R}^d} g^-.$$

Furthermore,

$$\int_{\mathbb{R}^d} g^- = \int_{B^d_{r(g)}} g^-,$$

where $B^d_{r(g)}$ is a d-dimensional ball of radius $r(g)$ and centered at the origin, because $\{x \in \mathbb{R}^d : g(x) < 0\} \subseteq B^d_{r(g)}$. It follows that

$$\int_{B^d_{r(g)}} g^- = 1/2,$$

because $\|g\|_1 = 1$.

The equation

$$\int_{B^d_{r(g)}} g^- = 1/2,$$

implies

$$\begin{aligned} 1/2 &\leq \operatorname{vol}\big(B_1^d\big)\, r(g)^d \|g\|_\infty \\ &\leq \operatorname{vol}\big(B_1^d\big)\, r(g)^d \|\widehat{g}\|_1 \\ &= \operatorname{vol}\big(B_1^d\big)\, r(g)^d \|g\|_1 \\ &= \operatorname{vol}\big(B_1^d\big)\, r(g)^d, \end{aligned}$$

and we conclude that

$$r(g) \geq \left(\frac{1}{2\operatorname{vol}\big(B_1^d\big)}\right)^{1/d} > \sqrt{\frac{d}{2\pi e}}.$$

Using a quadratic polynomial in $|x|^2$ times the Gaussian $e^{-\pi|x|^2}$ shows that $\mathsf{A}_+(d) \leq \sqrt{\frac{d+2}{2\pi}} = O\big(\sqrt{d}\big)$.

Exact values

But what is $\mathrm{A}_+(d)$ exactly? Can we obtain a sharp inequality?

So far, only when $d = 12$. It's a mystery why one seemingly arbitrary dimension has unexpected arithmetic structure.

Theorem (Cohn and Gonçalves 2019)

We have $\mathrm{A}_+(12) = \sqrt{2}$. In particular, there exists a radial Schwartz function $f : \mathbb{R}^{12} \to \mathbb{R}$ that is eventually nonnegative and satisfies $\widehat{f} = f$, $f(0) = 0$, and

$$r(f) = \sqrt{2}.$$

Moreover, as a radial function f has a double root at $|x| = 0$, a single root at $|x| = \sqrt{2}$, and double roots at $|x| = \sqrt{2j}$ for integers $j \geq 2$.

Numerics: upper bounds for $A_+(d)$

d	$A_+(d)$	d	$A_+(d)$	d	$A_+(d)$
1	0.572990	13	1.458239	25	1.894060
2	0.756207	14	1.500647	26	1.925084
3	0.887864	15	1.541603	27	1.955522
4	0.965953	16	1.581246	28	1.985407
5	1.036454	17	1.619692	29	2.014769
6	1.101116	18	1.657044	30	2.043633
7	1.161109	19	1.693390	31	2.072024
8	1.217275	20	1.728806	32	2.099965
9	1.270241	21	1.763360	33	2.127476
10	1.320483	22	1.797112	34	2.154577
11	1.368375	23	1.830115	35	2.181286
12	1.414214	24	1.862417	36	2.207618

Obtained using polynomials times Gaussians. Aside from rounding the last digit up, all digits are probably optimal for $d \geq 3$.

Open question: matching lower bounds

How can we obtain (nearly) matching lower bounds numerically?

This is a more subtle question than upper bounds, which just require producing a single function.

It ought to be possible via a suitable summation formula, but we don't know how to carry this out explicitly.

Sphere packing

The exact 12-dimensional function can be constructed using Viazovska's techniques, developed to solve the sphere packing problem in 8 and 24 dimensions.

The upper bound comes from a certain integral transform of a modular form.

The lower bound comes from the Eisenstein series E_6.

What does this problem have to do with sphere packing?

Linear programming bound

Converts an auxiliary function $f : \mathbb{R}^d \to \mathbb{R}$ into a bound for the sphere packing density Δ_d in $\mathbb{R}^d$.

Theorem (Cohn and Elkies 2003)

Suppose f is integrable, $\widehat{f}$ is also integrable, $\widehat{f}$ is real-valued (i.e., f is even), $f(0) = \widehat{f}(0) = 1$, $\widehat{f} \geq 0$ everywhere, and f is eventually nonpositive. Then

$$\Delta_d \leq \operatorname{vol}\left(B^d_{r(f)/2}\right).$$

Optimizing this bound amounts to minimizing $r(f)$. The exact optimum is not known except for

$d = 1$ (easy), $d = 8$ (Viazovska 2017), and

$d = 24$ (Cohn, Kumar, Miller, Radchenko, and Viazovska 2017),

where the optima are 1, $\sqrt{2}$, and 2, respectively. For $d = 2$, it is conjectured that the optimum is $(4/3)^{1/4}$, but no other exact values are even conjectured.

Theorem (Cohn and Elkies 2003)

Suppose f is integrable, $\widehat{f}$ is also integrable, $\widehat{f}$ is real-valued (i.e., f is even), $f(0) = \widehat{f}(0) = 1$, $\widehat{f} \geq 0$ everywhere, and f is eventually nonpositive. Then

$$\Delta_d \leq \mathrm{vol}\big(B^d_{r(f)/2}\big).$$

Given such an f, let $g = \widehat{f} - f$. Then g is not identically zero, or else f and $\widehat{f}$ would both have compact support. Furthermore, $r(g) \leq r(f)$.

Problem (-1 eigenfunction uncertainty principle)

Minimize $r(g)$ over all $g \colon \mathbb{R}^d \to \mathbb{R}$ such that

1. $g \in L^1(\mathbb{R}^d) \setminus \{0\}$ *and* $\widehat{g} = -g$, *and*
2. $g(0) = 0$ *and g is eventually nonnegative.*

Elkies and I conjectured that these problems have exactly the same answer.

Specifically, that every g from the -1 eigenfunction uncertainty principle can be lifted to an f proving the same bound for sphere packings, with $g = \widehat{f} - f$.

This is true numerically for every function anyone has ever constructed. How hard can it be to prove in general?

By far the biggest technicality in Viazovska's proof is piecing together the $+1$ and -1 eigenfunctions to get the right inequality. If we could prove this conjecture, we could dispense with the $+1$ eigenfunction entirely and substantially simplify the proof.

Comparison

These uncertainty principles (Bourgain-Clozel-Kahane 2010 and Cohn-Elkies 2003) are almost the same:

Problem ($+1$ eigenfunction uncertainty principle)

Minimize $r(g)$ over all $g\colon \mathbb{R}^d \to \mathbb{R}$ such that

1. $g \in L^1(\mathbb{R}^d) \setminus \{0\}$ *and* $\widehat{g} = g$, *and*
2. $g(0) = 0$ *and* g *is eventually nonnegative.*

Problem (-1 eigenfunction uncertainty principle)

Minimize $r(g)$ over all $g\colon \mathbb{R}^d \to \mathbb{R}$ such that

1. $g \in L^1(\mathbb{R}^d) \setminus \{0\}$ *and* $\widehat{g} = -g$, *and*
2. $g(0) = 0$ *and* g *is eventually nonnegative.*

Let $\mathcal{A}_-(d)$ be the set of functions such that

1. $f \in L^1(\mathbb{R}^d)$, $\widehat{f} \in L^1(\mathbb{R}^d)$, and $\widehat{f}$ is real-valued (i.e., f is even),
2. f is eventually nonnegative while $\widehat{f}(0) \leq 0$, and
3. $\widehat{f}$ is eventually nonpositive while $f(0) \geq 0$.

Then

$$\mathrm{A}_-(d) := \inf_{f \in \mathcal{A}_-(d) \setminus \{0\}} \sqrt{r(f) r(\widehat{f}\,)}$$

is also the optimal constant from the -1 eigenfunction uncertainty principle. Shown by similar reductions to the $+1$ case.

Again, $\mathrm{A}_-(d)$ is within a constant factor of $\sqrt{d}$.

This gives a complementary uncertainty principle to that of Bourgain, Clozel, and Kahane, for functions with opposite signs.

Meaning of linear programming bounds

What do the linear programming bounds mean in general dimensions?

When they aren't sharp, are they just some arbitrary bounds, of no particular interest once they have been superseded by improved bounds?

We believe these bounds always have an independent meaning, as optimal constants in an uncertainty principle.

But we still can't prove the underlying conjecture.

How much do $A_+(d)$ and $A_-(d)$ differ?

Pretty close in high dimensions, but almost certainly not equal.

Gonçalves and I conjecture that

$$\lim_{d \to \infty} \frac{A_+(d)}{\sqrt{d}} = \lim_{d \to \infty} \frac{A_-(d)}{\sqrt{d}},$$

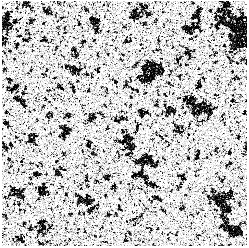
and Afkhami-Jeddi, Hartman, de Laat, Tajdini, and I conjectured that the limit is $1/\pi$.

But we can only estimate four digits reliably, so it's unclear how believable this conjecture is.

Conformal field theories (CFTs)

These uncertainty principles turn out to be connected to conformal field theory in two dimensions.

There are many different 2d CFTs, governing the behavior of various statistical mechanics models at critical points. Think scale invariance plus ε. (The ε is conceptually important but automatic in some cases.)



(Ising model image by David Wilson)

Hierarchy of exceptional structures

binary codes, Mathieu groups

$\downarrow$

lattices, Conway groups

$\downarrow$

vertex operator algebras and CFTs, monster group

Each level used to construct the next. This accounts for 20 of the 26 sporadic finite simple groups.

Conformal bootstrap

What can we learn about the space of possible CFTs from self-consistency? For example, from conformal invariance of partition functions.

Much of quantum field theory is not rigorous, but here we can formulate clean mathematics, even if some physical consequences are heuristic.

The objects may or may not be well defined, but we can specify some constraints.

Quantum gravity

The (conjectural) AdS/CFT correspondence says

3d quantum gravity in anti-de Sitter space

$$\leftrightarrow$$

2d CFT on conformal boundary

What can we learn about possibilities for quantum gravity by studying the space of CFTs ("theory space")?

This is a toy model, since AdS means negative cosmological constant (doesn't match astronomical data).

Let's treat the physics as a black box and look at the torus partition function.

Every 2d torus is conformally equivalent to a flat torus (uniformization), so it's $\mathbb{C}/\Lambda$ for some lattice Λ, which we can rescale or rotate. Pick a basis $\tau, 1$ with τ in the upper half plane.

Change of basis is action of $\mathrm{SL}_2(\mathbb{Z})$, and space of tori is quotient.

What is the partition function? Here it's a generating function for "scaling dimension" and "spin" of states.

A conformal algebra acts on our CFT. It's at least the Virasoro algebra, and possibly larger. We'll focus on $\mathrm{U}(1)^c$, the theory of c free bosons.

Hartman, Maźač, and Rastelli (2019) discovered that the spinless modular bootstrap is equivalent to linear programming bounds for sphere packings!

Write the partition function in terms of characters of $\mathrm{U}(1)^c$. If n_Δ is the multiplicity of states with scaling dimension Δ, conformal invariance of the partition function tells us that

$$\sum_{\Delta} n_\Delta e^{2\pi i \tau \Delta} = \sum_{\Delta} n_\Delta (i/\tau)^{d/2} e^{2\pi i(-1/\tau)\Delta},$$

where d is the sum of the holomorphic and antiholomorphic conformal charges ($d = 2c$ for c free bosons).

$$\sum_\Delta n_\Delta e^{2\pi i\tau\Delta} = \sum_\Delta n_\Delta (i/\tau)^{d/2} e^{2\pi i(-1/\tau)\Delta}$$

For $\mathrm{Im}(\tau) > 0$, $x \mapsto e^{2\pi i\tau|x|^2/2}$ is a complex Gaussian on $\mathbb{R}^d$ with Fourier transform $y \mapsto (i/\tau)^{d/2} e^{2\pi i(-1/\tau)(|y|^2/2)}$. I.e., setting $\Delta = |x|^2/2$ makes this look like Poisson summation:

$$\sum_\Delta n_\Delta f(\sqrt{2\Delta}) = \sum_\Delta n_\Delta \widehat{f}(\sqrt{2\Delta}).$$

This must hold for all radial Schwartz functions f on $\mathbb{R}^d$, since complex Gaussians span a dense subspace.

But Poisson summation was the only input we needed for linear programming bounds! It applies just as well regardless of whether the formula comes from a sphere packing or a CFT.

So what does this mean?

The sphere packing bounds were secretly also upper bounds for the *spectral gap* of a CFT, the smallest nonzero scaling dimension Δ_1 of a primary field (analogous to packing radius).

"Pure quantum gravity" is expected to have just Virasoro symmetry and $\Delta_1 \sim c/12$ as $c \to \infty$. A key question for the conformal bootstrap is whether such a large spectral gap is even possible, or whether gravity requires additional primary fields.

Nobody knows.

U(1) gravity

Afkhami-Jeddi, Hartman, Tajdini, and I, and independently Maloney and Witten, proposed a holographic dual to pure 3d gravity for free bosons, called U(1) gravity.

It is obtained by using results of Siegel to average over Narain lattices, a more sphisticated version of the Siegel mean value theorem from Wednesday.

This is definitely not real-world physics. Maybe it's to actual quantum gravity as the simple harmonic oscillator is to statistical mechanics. However, it gives a concrete way to obtain CFTs with large spectral gaps and carry out explicit calculations for an unusual form of pure gravity.

Let's return to Bourgain-Clozel-Kahane uncertainty.

Does it have CFT implications? These sign patterns do arise in theories with discrete anomalies or fermions, but we don't have a concrete proposal. Maybe the $d = 12$ bound is related to K3 sigma models?

Does it have packing or other discrete geometry implications? So far no one has found any.

But it has to mean something: every thread seems to lead somewhere...

For more information

J. Bourgain, L. Clozel, and J.-P. Kahane *Principe d'Heisenberg et fonctions positives*, Ann. Inst. Fourier (Grenoble) **60** (2010), no. 4, 1215–1232, arXiv:0811.4360

H. Cohn and F. Gonçalves, *An optimal uncertainty principle in twelve dimensions via modular forms*, Invent. Math. **217** (2019), no. 3, 799–831, arXiv:1712.04438

N. Afkhami-Jeddi, H. Cohn, T. Hartman, D. de Laat, and A. Tajdini, *High-dimensional sphere packing and the modular bootstrap*, J. High Energy Phys. (2020), no. 12, Paper No. 66, 44 pages, arXiv:2006.02560

N. Afkhami-Jeddi, H. Cohn, T. Hartman, and A. Tajdini, *Free partition functions and an averaged holographic duality*, J. High Energy Phys. (2021), no. 1, Paper No. 130, 42 pages, arXiv:2006.04839