Scratchwork: Theta Functions

William Duke's proof that the solutions to $n = a^2 + b^2 + c^2$ become equidistributed as $n \to \infty$ takes a quarter of a page:

- Goro Shimura shows there are many theta functions, each invariant under $\Gamma_0(4)$

$$\theta(z; u) = \sum_{m \in \mathbb{Z}^3} u(m)\, e(z|m|^2) = \sum_{n>0} n^{\ell/2}\, r_3(n) \left[\frac{1}{r_3(n)} \sum_{\xi \in V_3(n)} u(\xi) \right] e(nz)$$

  one for each spherical Harmonic $u \in L^2(S^2)$. Here $|m|^2 = m_1^2 + m_2^2 + m_3^2$ and $V_3(n) = \#\{(a,b,c) : a^2 + b^2 + c^2 = m\}$.

- Henryk Iwaniec offers a bound for the Fourier coffients of cusp forms.

$$a_n \ll_{k,\epsilon} n^{k/2 - 2/7 + \epsilon}$$

  These are tending to zero if we fix a tolerance (ϵ) and a "weight" of modular form (k).

- Combining Iwaniec and Shimura's result[1] we obtain an estimate for the sphere averages

$$\frac{1}{r_3(n)} \sum_{\xi \in V_3(n)} u(\xi) \ll_{u,\epsilon} n^{-1/28 + \epsilon}$$

  This bound depends on the spherical harmonic (u) and the tolerance (ϵ). And we need $n \not\equiv 7 \pmod 8$.

There have been many surprises along the way learning this topic. And I have problems with a lot of this discussion, because these are professors talking to other professors. Slowly turning these objections into contributions of my own!

- $\theta(z)$ is $\Gamma_0(4)$ invariant not $SL(2,\mathbb{Z})$ invariant. As subgroups the index $[SL(2,\mathbb{Z}) : \Gamma_0(4)] = 6$. For the record $\Gamma_0(4) \simeq \langle z \mapsto z+1, z \mapsto -\frac{1}{4z} \rangle$ while $SL(2,\mathbb{Z}) \simeq \langle z \mapsto z+1, z \mapsto -\frac{1}{z} \rangle$. These are like continued fractions $a = [a_0; a_1, \dots, a_n]$ all of whose digits are multiples of 4.

- $\theta(z)$ is *not* a cusp form, but if $\deg u > 0$ then $\theta(z; u)$ **is** a cusp form. The Fourier series *is* the q-series as $q = e^{2\pi i t}$ is the change of variables.

$$\theta(z; u) = 0 + a_1\, q + a_2\, q^2 + \dots$$

- Even though we are studying equidistribution of solutions to an equation $f(x) = n$ we don't necessarily study the existence of one solution. After reading Serre's **Course on Arithemtic** (GTM #85) we learn that the existence of one solution is a rather deep problem. Lastly, I don't think it's "done" by the time we discuss equidistribution.

- I didn't know what the symbol "$\ll$" meant. $1 \ll 100$ and also $100 \ll 1$. Here's one more: $10^{100} \ll x$ since *eventually* $x > 10^{100}$. These are things we learn in beginning math classes, but sometimes a brand new symbol is introduced and we have to do it again.

---

[1] and an estimate of Siegel, which I haven't even looked at, $r_3(n) \gg_\epsilon \sqrt{n^{1-\epsilon}}$.

Then I find out all of this is slightly out of date. This one pricked my last nerve. I believe, in the process of verifying Duke's claim (leaning me through Iwaniec and Shimura and Walspurger and even more) we have made new lemmas. One of them is definitnely new.

OK. Here's a problem statement: let's try to find the constant to go with the $\ll$ sign:

$$\left[a_n \ll n^{k/2-2/7+\epsilon}\right] \to \left[a_n < C\, n^{k/2-2/7+\epsilon}\right]$$

The constant C is not known (probably because nobody cares) if we fix (ϵ) and (k).

Shimura's constructions are the start of the **theta correspondence**, and I'm choosing to use the out-of-date version that doesn't use any of Waldspurger's technology. Duke makes no use of the adeles, $\mathbb{A}$; any strategy involving them will be new. Iwaniec makes heavy use of Eisenstein series and it's rather mysterious:

$$[\text{theta functions}] \to [\text{Eisenstein series}]$$

Iwaniec says it works, making such a kind of map, risk-free, but do we really understand it? I don't immediately have a problem that is unkown about it. The more we unpack, the scarier it gets.

$$[\text{Kloosterman sums}] \quad || \quad [\text{Bessel functions}]$$

Do we really understand this? In the case of 4-squares, the proof involves the Weil conjectures, yet in 3-squares this just falls out of some very tricky averaging procedures, that I do not wish to replicate.[2]

This discussion has created some deep and lingering doubts: do I understand Weyl's equidistrubiton criterion *for spheres* or requirements for Poisson summation? Do I know what it was so important that $\theta(z;u)$ was a **holomorphic** cusp form, or how to get number-theoretic information out of that? Not really.

Waldspurger Formula

Another, equally vague, proof of equidistribution comes from the study of torus integrals, but what kind of torus? As a collection of points on the sphere $G_d = \{(a,b,c) : a^2+b^2+c^2 = d\} \subseteq X = S^2$ we could assign a probability measure for each number $d \in \mathbb{Z}$, and compute averages with respect to this probability measure:

$$\int_{S^2} \phi\, \mu_d = \frac{1}{r_3(n)} \sum_{(a,b,c)\in G_d} \phi\left(\frac{a}{\sqrt{d}}, \frac{b}{\sqrt{d}}, \frac{c}{\sqrt{d}}\right)$$

There's some measure μ_d that magically when you integrate against it returns the sphere averages. The equidstirbution statement now looks very terse:

$$\lim_{d\to\infty} \mu_d = \mu$$

where μ is the Lebesgue measure on the 2-sphere $X = S^2$. To me, Lebesgue measure means that despite equidistribution, interesting sets may occur on the way as $d \to \infty$. Weak-* convergence measns:

$$W(\phi, d) := \int_X \phi\, \mu_d \to 0 \text{ as } |d| \to \infty$$

as we let ϕ range over an fixed orthogonal basis of continuous functions[3] $\phi \in L^2_0(X,\mu)$. In our case, these are **spherical harmonics**.

[2] at this time

[3] no step functions!!

The Weyl averages could be written as an "integral" over a two-sided quotient of Torus.

$$W(\phi, d) = \int_{T_d(\mathbb{Q})\backslash T_d(\mathbb{A})/K_{T_d}} \phi(z_d.t)\, dt$$

The exericse would be to understand what this torus could be; The paper I was reading had a typo. The torus is described using "restriction of scalars". Let $K = \mathbb{Q}(\sqrt{d})$:

$$T_d := \mathrm{Res}_{K/\mathbb{Q}}(\mathbb{G}_m)/\mathbb{G}_m$$

which measures how much bigger $\mathbb{G}_m$ is over K than over Q. And it would be great if we knew a definition of $\mathbb{G}_m$. In fact, it's not just a group it's a **functor** from the category from (the oppososite category) of $\mathbb{Q}$-schemes to the category of groups:

$$\begin{aligned} \mathrm{Res}_{K/\mathbb{Q}} : (\mathbb{Q}\text{-}\textbf{scheme})^{\mathsf{op}} &\to \textbf{Group} \\ \mathrm{Res}_{K/\mathbb{Q}}\, X(S) &= X(S \times_{\mathbb{Q}} K) \end{aligned}$$

where $\mathbb{G}_m$ is the multiplicative group. And then I need a definition of K_{T_d}.

...

If we keep going, there is the Waldspurger formula finally:

$$|W(\varphi, d)|^2 = c_{\varphi,d}\, \frac{L(\pi, \frac{1}{2})\; L(\pi \times \chi_d, \frac{1}{2})}{L(\chi_d, 1)^2 \sqrt{d}}$$

where φ is a "new cuspform" - the L^2 normalized newvector in some automorphic representation π. And where π' [sic] is the GL_2 automorphic respresentation corresponding to π by the Jacquet-Langlands correspondence, and $c_{\varphi,d}$ is a number.

The Waldspurger formula is used in conjunction with a **subconvexity bound**

$$\left[L(\pi \times \chi_d, \frac{1}{2}) \ll_\pi |d|^{1/2-\delta} \right] \text{ therefore } \left[W(\phi, d) \to 0 \text{ as } d \to \infty \right]$$

My problem with this reductionist point of view is that someone has to be responsible for placing these abstract results back into a context.[4] Fortunately, there is one and Akshay Venkatesh and his colleagues have developed the sparse equidistribution framework for solving this type of problem.

Unfortunately we have left out a lot of details. As long as we focus on the basics, there are several textbooks that have emerged since 2005. I like representation theory and geometry so there are at least two textbooks[5]

- Anton Deitmar **Automorphic Forms** Universitext, 2013.
- Franoise Dal'Bo **Geodesic and Horocyclic Trajectories** Universitext, 2011.
- Akshay Venkstesh **Sparse equidistribution problems, period bounds, and subconvexity** `arXiv:math/0506224` Annals of Mathematics, 172 (2010), 989-1094

[4] `https://en.wikipedia.org/wiki/Reductionism`

[5] Thus as research topics trickle down into graduate-level teaching, first textbooks written for new, cutting-edge courses may make their way into Universitext.

Comprehension Check Is $\theta(z,u) = \sum_{m\in\mathbb{Z}^3} u(m)e(z|m|^2)$ an automorphic form? What is the representation?

We could set u to be some spherical harmonic. There $u(x,y,z) \equiv 1$. Skip. There is $u(x,y,z) = \frac{z}{r}$, so let's try a schematic there:

$$\frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} z = 0$$

The answer is always zero because for every solution (x,y,z) I can find $2^3 = 8$ other solutions $(\pm x, \pm y, \pm z)$. OK so we really just want the first "octant" here,and we could have a sign $S^2 \to \{1,-1\}$ that exactly matches the sign of (x,y,z), so that $u(x,y,z) \times \text{sgn} \geq 0$.

Moving on, second degree offers our first chance at a non-trivial sum:

$$\frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} (x+iy)^2 = \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} \left(x^2-y^2\right) + \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} 2i\,xy = 0$$

There are a few more symmetries we hadn't considered. We could map a solution $(x,y,z) \mapsto (y,x,z)$. We could write this as a 3×3 matrix:

$$\left(\begin{array}{rrr} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right) \left(\begin{array}{c} x \\ y \\ z \end{array}\right) = \left(\begin{array}{r} -y \\ x \\ z \end{array}\right)$$

This group has a name $SO(3,\mathbb{Z}/2\mathbb{Z})$ sometimes written as $SO(3,\mathbb{Z}_2)$ or $SO(3,\mathbb{F}_2)$. Mentally, we are considering these symmetries as part of $SO(3,\mathbb{R})$, the group of all rotations in physical space.[6] and we need to take a moment to specify how these discrete group of symmetries embed to a continuous group of symmetries.

In any case, that average is also zero and we could multiply by $\text{sign}(x^2-y^2) = \text{sign}(x-y)\,\text{sign}(x+y)$. Likewise:

$$\begin{array}{rcrcl} \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} (x+iy)\,z & = & \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} (x+iy)\,(-z) & = & 0 \\ \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} (2z^2-x^2-y^2) & = & \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} (2z^2-z^2-z^2) & = & 0 \end{array}$$

and basically we have done all five degree-2 spherical harmonics $u \in \left\{Y_2^{-2}, Y_2^{-1}, Y_2^0, Y_2^1, Y_2^2\right\}$, and all of the averages vanish $\mathbb{E}(u) = 0$. Let's try degree 4:

$$\frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} (x+iy)^4 = \frac{1}{r_3(n)} \sum_{x^2+y^2+z^2=n} \left(x^4 - 6\,x^2y^2 + y^4\right) \overset{?}{\neq} 0$$

I don't know this one off the top of my head. Each step of the way we have to continue to refine what we mean by "evaluate". In any case, Duke does not really discuss any of this, and what could we do with the normalization of $\frac{3}{16}\sqrt{\frac{35}{2\pi}}$?.

The rotation group is beginning to feature in our discussion (but not really modular forms). Different kinds of SO_3 are being used:

$$\begin{array}{rcl} SO_3(\mathbb{Z}) & \subseteq & SO_3(\mathbb{R}) \\ SO_3(\mathbb{Z}) & \subseteq & SO_3(\mathbb{Z}_2) \end{array}$$

[6]It's almost tautolitical: if you rotate the thing once and then again you get another rotation. You should prove in 3D space that two rotations become a *single* rotation with a particular axis.

where $\mathbb{Z}_2$ could be the 2-adic numbers? And so SO_3 becomes a **functor** as well.

Back to 3-squares, overall. Is this such a fundamental problem? All of this is about a growing sphere in a cubic lattice:

$$\sqrt{n} \times \big\{ x^2 + y^2 + z^2 = 1 \big\} \subset \mathbb{Z}^3$$

and looking at the various angles. What could be more basic than that?

The "representation" of SO_3 (at least over $\mathbb{R}$) is not terribly exicing:

$$SO_3 \circlearrowright \Big\{ (x-iy)^2,\ (x-iy)z,\ (2z^2 - x^2 - y^2),\ (x-iy)z,\ (x+iy)^2 \Big\}$$

and it's highly, endlessly constructive to verify this that SO_3 is rotating these things. The discussion over $\mathbb{Z}_p$ needs discussion as well, eventually.

There's an extensive literature on modular forms, but I can never get the discussion I am looking for. And I'm still working on that, but this is why I'm proposing this more direct approach that this moment.

Getting Ahead of Ourselves Are solutions to $n = a^2 + b^2 + c^2$ points on a quaternion Shimura variety?

How do we construct the complex numbers? We have this problematic equation

$$\big\{ x^2 + 1 = 0 \big\} = \varnothing$$

and so we create this imaginary number $x = \sqrt{-1}$ and build an arithmetic out of that. This could have been done with matrices:

$$x + i\,y \mapsto \left(\begin{array}{rr} x & -y \\ y & x \end{array} \right)$$

and the 2×2 matrix arithmetic behaves the same way as the complex numbers. How about 4×4? we have the **quaternions**:

$$x + \mathbf{i}\,y + \mathbf{j}\,z + \mathbf{k}\,w \mapsto \left(\begin{array}{rrrr} x & -y & z & -w \\ y & x & w & z \\ -z & w & -x & y \\ -w & -z & -y & -x \end{array} \right)$$

Despite being 4-dimensional, this was a way of discussion rotations in 3D space and calculating with them, until the early 20th century.

There's no multiplication for 3×3 numbers. We'd like to say:

$$\mathbf{i}\,x + \mathbf{i}\,y + \mathbf{i}\,z \mapsto ?$$

but there's no right hand side. This quaternion space is called, $(B_2^\infty)^\times$ or something.

A quick google serach for quaternionic shimura varities returns many extensive resources at an undesirable level of abstraction. Or just google Shimura varities. Same. They feel that "formulas" or polynomial sections are no longer very meaningful so they resort to the language of varities and divisors, in order to express the information they want, leaving me totally lost.

There could be our lemma: the definition of Shimura variety is very complicated but we name an instance where the calculation can be done quite easily. Blah Blah Blah.

Comprehension Check #2 (Weyl Equidstribution on S^2). The discussion of spherical harmonics on Wikipedia is rather extensive.[7]

Let's propose a defintin of "sign" for points on the sphere. Umm... let's pick the x-axis:

$$\begin{array}{rcl} S^2 & \to & \{-1,1\} \\ (x,y,z) & \mapsto & \left\{ \begin{array}{rl} x & x>0 \\ -x & x<0 \end{array} \right. \end{array}$$

This is "absolute value" rather than sign. Just some notion of size. Unfortunately this function is **discontinuous** at $x=0$ which is now a circle instead of a point:

$$\big\{x=0\big\} \cap \big\{x^2+y^2+z^2=1\big\} = \Big[\bigcirc \hookrightarrow \mathbb{R}^3 \Big]$$

and we get a hoop embedded in some kind of 3-space.

How do we account for this "jumping" behavior of functions on the sphere? This set up can become very rich and quite varied.

Comprehension Check #3 (Automorphic Forms) Hopefully, getting started.

Comprehension Check #4 (Number Fields) I didn't regozize but the Gauss circle problem could be phrased in a very general way.

> Let K be a number field and let $N_{K,n}(X)$ count the number of number fields L with $[L:K]=2$ and $N^K_{\mathbb{Q}} \mathcal{D}_{L/K} < X$. Then
> $$N(X) \leq cX$$

This is not the circle problem. In the case that $K=\mathbb{Q}$, we are counting $\#\{(a,b) : b^2 - 4a \leq X\}$ and very schematically:

$$b \approx \sqrt{X} \text{ and } 1+2+3+\dots+\sqrt{X} \approx \frac{1}{2}X$$

and this result could be extended to all number fields K and another interesting case could be $[L:K]=3$.

Why did I confuse this for the Gauss circle problem? In the field $\mathbb{Q}(i)$ the ring of integers is $\mathbb{Z}[i]$ and it has a norm, $N(a+bi) = a^2+b^2$ and we could have that:

$$N(X) = \#\{a+bi \in \mathbb{Z}[i] : N(a+bi) < X\} \leq \pi X$$

I use the same letter twice. And it's reasonable to believe this should hold for any number field. Perhaps we ought to focus on primes:

$$\frac{\#\{\mathfrak{p} \in \mathbb{Z}[i] : N(\mathfrak{p}) < X \text{ and } \mathfrak{p} = e^{i(\theta \pm \epsilon)}\}}{\#\{\mathfrak{p} \in \mathbb{Z}[i] : N(\mathfrak{p}) < X\}} = 2\epsilon$$

so that primes in $\mathbb{Z}[i]$ are evenly distributed in the complex plane. If you draw the picture this is self-evident, but that's what number theory is: elaborate proofs of self-evident statements.

I am trying to get at the Sato-Tate equidistribution formula. That formula is very general. Sato-Tate can be proven over number fields, and for various classes of modular forms, in situations were we don't even know what these number sequences are.[8] perhaps I am looking for a sequence of numbers defined over $\mathbb{Z}$ which is the norm or real part of some counting problem over a number field. That could make sense to me.

[7]

[8] These "what is"-type questions are dangerious at this level. It means I get to define an invariant of some kind - probably something really dull - that captures what the definition of **is** is.

If we try to use modular forms, let $u(x) = x_1^4 - 6x_1^2x_2^2 + x_2^4$ be a spherical harmonic. Then

$$\theta(z;u) = \sum_{m\in\mathbb{Z}^3} u(m)e(z|m|^2) = \sum_{n\geq 0}\left[\sum_{m_1^2+m_2^2+m_3^2=n} u(m)\right] e(n\,z)$$

is a holomorphic cusp form of weight $k = \frac{3}{2}+4 = \frac{11}{4}$. There are two bounds we can use:

- Hecke $a_n \leq C\,n^{k/2}$ in our case $C\,n^{11/8}$.
- Trivial $a_n \ll_{k,\epsilon} n^{k/2-\frac{1}{4}+\epsilon}$ which plays out to $n^{9/8}$
- Iwaniec $a_n \ll_{k,\epsilon} n^{k/2-\frac{2}{7}+\epsilon}$ which plays out to $n^{61/56}$.

Surely these are the fundamental constants of nature. In the general terms we could rescale so that:

$$\frac{a_n}{n^{\ell/2}\,r_3(n)} = n^{\left(\frac{1}{2}(\frac{3}{2}+\ell)-\frac{1}{4}+\epsilon\right)-\left(\ell/2+\frac{1}{2}\right)} = n^{\epsilon}$$

here I'm putting that $r_3(n) \approx \sqrt{n}$ a formula which I pulled out of hat. So it's a tie. And with the Iwaniec formula – if you believe any of that – brings you just over the edge.

Really? Let's use our example choice of $u(x)$ the Weyl fraction says:

$$\frac{1}{n^2\sqrt{n}}\sum_{m_1^2+m_2^2+m_3^2=n}\left(m_1^4 - 6m_1^2\,m_2^2 + m_2^4\right) \ll n^{\epsilon} \text{ or } n^{-\frac{1}{28}+\epsilon}$$

we do have powerful means at our disposal, do you believe this is the best we can do? Even for a single $u \in L^2(S^2)$ this could be the record. This is stated out modular forms; we will use modular forms and even adeles $\mathbb{A}$ but I wanted a version were all possible approaches have equal footing

- modular forms / harmonic analysis
- dynamical systems
- additive combinatorics / sumsets
- probabilistic method

Deep recent theory shows all three approaches are more or less analogous or equivalent. I don't know if anybody does the "footwork" of passing between one approach or another, but I am certainly willing to waste my time finding such equivalences.[9]

To me, when I was a beginner, all of these look the same. And that intuition basically still stands.

[9] Odds are one approach is currently slightly better than the other and by passing between them we push them towards equilibrium.