The Unbounded Denominators Conjecture

Joint work with Frank Calegari and Yunqing Tang

Vesselin Dimitrov

University of Toronto

International Seminar on Automorphic Forms

1 February, 2022

Modular Forms

Definition

A **vector-valued modular form** on $\mathrm{SL}_2(\mathbb{Z})$ of weight k and dimension n is a pair (F, ρ) comprised of:

- A holomorphic mapping $F = (F_1, \dots, F_n) : \mathbf{H} \to \mathbb{C}^n$ on the upper half plane
- A representation $\rho : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{GL}_n(\mathbb{C})$
- Linked by $(c\tau + d)^{-k} F^{\mathrm{t}}\left(\frac{a\tau+b}{c\tau+d}\right) = \rho \begin{pmatrix} a & b \\ c & d \end{pmatrix} F^{\mathrm{t}}(\tau)$
- Such that the matrix $\rho \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_n(\mathbb{C})$ is semisimple
- All components $F_j : \mathbf{H} \to \mathbb{C}$ have *moderate growth in vertical strips*: $\forall a < b, \forall C > 0, \quad \exists A, B > 0$ such that

$$\forall \tau \in \mathbf{H}, \quad a \leq \mathrm{Re}\,\tau \leq b, \quad \mathrm{Im}\,\tau \geq C \quad \Longrightarrow \quad |F_j(\tau)| \leq A e^{B\,\mathrm{Im}\,\tau}.$$

Fourier expansions

Taken together, the latter two conditions are equivalent to the existence of Fourier expansions:

$$F_1(\tau), \dots, F_n(\tau) \in \bigoplus_{i=1}^{s} q^{\beta_i}\, \mathbb{C}((q)), \quad q := e^{\pi i \tau} \quad \text{(no } \log q \text{ terms)}$$

Theorem (Chudnovsky, Bombieri, André)

If the F_j are in $\bigoplus_{i=1}^{s} q^{\beta_i}\, R((q))$ for some finitely generated subring $R \subset \mathbb{C}$, then all the exponents are rational: $\beta_i \in \mathbb{Q}$.

That is: $F_1(\tau), \dots, F_n(\tau) \in R((q^{1/m}))$ for some $m \in \mathbb{N}$.

p-curvature conjecture (a special case)

One expects a characterization of the "S-integral denominators condition" in terms of the kernel subgroup $\ker(\rho) \lhd \mathrm{SL}_2(\mathbb{Z})$:

Grothendieck's conjecture

The following are equivalent:

1. $F_1(\tau), \ldots, F_n(\tau) \in R((q^{1/m}))$ for some finitely generated subring $R \subset \mathbb{C}$
2. $\ker \rho$ has a finite index in $\mathrm{SL}_2(\mathbb{Z})$

We will answer this question—in a greater precision—for the case that R is a finitely generated $\mathbb{Z}$-module instead of a finitely generated ring.

The Unbounded Denominators property

Theorem (Calegari, D., Tang)

The following are equivalent:

1. $F_1(\tau), \ldots, F_n(\tau) \in \mathbb{Z}((q^{1/m})) \otimes \mathbb{C}$
2. $\ker \rho$ contains $\Gamma(N)$ for some $N \in \mathbb{N}$.

$$\Gamma(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

The special case of Atkin and Swinnerton-Dyer

Theorem

Let $f(\tau) \in \overline{\mathbb{Z}}[[q^{1/N}]] \otimes \mathbb{C}$ be a holomorphic function on the upper half plane $\mathbf{H}$, expanded out in $q = e^{\pi i \tau}$. Suppose there exists a finite index subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ and an integer k such that f is a modular form of weight k and level Γ:

$$\forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma, \quad f\left(\frac{a\tau + b}{c\tau + d}\right) = (a\tau + b)^k f(\tau).$$

(And $f(\tau)$ is meromorphic locally near every cusp of the compactification of $\mathbf{H}/\Gamma$.)
Then $f(\tau)$ is modular under a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$: there exists an $N \in \mathbb{N}$ such that

$$\forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N, \quad f\left(\frac{a\tau + b}{c\tau + d}\right) = (a\tau + b)^k f(\tau).$$

A classical formula

$$
{}_2F_1\left[\begin{matrix}1/2 \;\; 1/2\\ 1\end{matrix};\frac{\left(\sum_{n\in\mathbb{Z}} q^{(n+1/2)^2}\right)^4}{\left(\sum_{n\in\mathbb{Z}} q^{n^2}\right)^4}\right]=\left(\sum_{n\in\mathbb{Z}} q^{n^2}\right)^2
$$

(An equation of the form holonomic in λ $=$ a modular form)

Jacobi's *Thetanullwerte*:

$\left(\sum_{n\in\mathbb{Z}} q^{n^2}\right)^4=\left(\sum_{n\in\mathbb{Z}} q^{(n+1/2)^2}\right)^4+\left(\sum_{n\in\mathbb{Z}}(-1)^n q^{n^2}\right)^4$

Gauss's *hypergeometric equation*

$(\lambda^2-\lambda)\frac{d^2f}{d\lambda^2}+((a+b+1)\lambda-c)\frac{df}{d\lambda}+abf=0$, with the parameters $a=b=1/2,\ c=1$

The Legendre–Picard modular lambda function (with $q=\exp(\pi i\tau)$):

$$
\lambda(\tau)=\frac{\left(\sum_{n\in\mathbb{Z}} q^{(n+1/2)^2}\right)^4}{\left(\sum_{n\in\mathbb{Z}} q^{n^2}\right)^4}=16q\prod_{n=1}^{\infty}\left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^8
$$

Modular lambda

By a mild and harmless notational abuse, let me write it with $q = \exp(\pi i \tau)$ as the argument:

$$\lambda(q) = \frac{\left(\sum_{n\in\mathbb{Z}} q^{(n+1/2)^2}\right)^4}{\left(\sum_{n\in\mathbb{Z}} q^{n^2}\right)^4} = 16q \prod_{n=1}^{\infty} \left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^8$$

- Hauptmodul of $\Gamma(2)$
- Relation $j = 256\frac{(1-\lambda+\lambda^2)^3}{\lambda^2(1-\lambda)^2}$ to Klein's modular invariant
- $(H, i) \to (\mathbb{C} \setminus \{0, 1\}, 0)$, $\tau \mapsto \lambda(q)$ is the (analytic) universal covering map.
- In the q coordinate on the unit disc, the fiber $\lambda^{-1}(0) = \{0\}$ and the *conformal size* $|\lambda'(0)| = 16$;
- The Hauptmodul $x := \lambda/16 = q - 8q^2 + \cdots \in q + q^2\mathbb{Z}[[q]]$ expresses the formal power series rings equality $\mathbb{Z}[[q]] = \mathbb{Z}[[x]]$.

General features of the classical formula

$$ {}_2F_1\left(\frac{1}{2},\frac{1}{2};1;16x\right) = \sum_{n=0}^{\infty} \binom{2n}{n}^2 x^n = \Big(\sum_{n\in\mathbb{Z}} q^{n^2}\Big)^2 \in \mathbb{Z}[[q]] = \mathbb{Z}[[x]] $$

- Replace the parameters $(\{1/2, 1/2\};\{1\})$ of this particular balanced hypergeometric function by any parameters (of any rank) $(\{a_i\};\{b_j\})$, and we still get a convergent power series on $|q|<1$: because the function $16x = \lambda$ avoids the all singularities $\{0,1,\infty\}$ of the hypergeometric ODE, and Cauchy's analyticity theorem on linear ODEs with analytic coefficients and no singularities in a disc
- Replace $\Big(\sum_{n\in\mathbb{Z}} q^{n^2}\Big)^2$ by any modular form of any weight k, then expand formally in $x=\lambda/16$ using $\mathbb{Z}[[q]] = \mathbb{Z}[[x]]$, and we always get a holonomic function satisfying a rank $k+1$ linear ODE over $\overline{\mathbb{C}(x)}$ with monodromy group commensurable with $\mathrm{Sym}^k\,\mathrm{SL}_2(\mathbb{Z})$.

Getting the linear ODE out of the modular form of weight k

If $f(\tau)$ is modular of weight k under some finite index subgroup $\Gamma < \Gamma(2)$, and $\mathbb{C}_\Gamma \subset \overline{\mathbb{C}(x)}$ is the field of Γ-automorphic functions, then the $\mathbb{C}_\Gamma$-linear span of the $k+1$ functions

$$f(\tau), \tau f(\tau), \tau^2 f(\tau), \ldots, \tau^k f(\tau)$$

is closed under d/dx.

Hecke theory gives (one, conceptual) proof of the integrality of coefficients for the *congruence* modular forms

Since the Hecke eigenvalues are manifestly algebraic integers, and since all the Fourier coefficients of an eigenform are determined as integer polynomials in the Hecke eigenvalues, it follows:

Theorem

Every holomorphic cusp form on a **congruence** subgroup of $\mathrm{SL}_2(\mathbb{Z})$ has (at every cusp) a Fourier expansion in $\overline{\mathbb{Z}}[[q^{1/N}]] \otimes \mathbb{C}$ for some N.

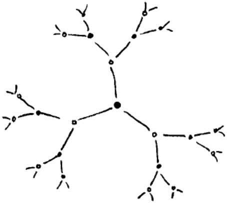
Here, $\overline{\mathbb{Z}}$ is the ring of algebraic integers.

In particular, the automorphy group of $\lambda^{1/3}$ is a certain noncongruence finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. (More precisely, this recovers a classical theorem of Fricke and Klein: $\lambda^{1/n}$—or equivalently, the Fuchsian group of the affine Fermat curve $x^n + y^n = 1$ — is congruence modular if and only if $n \mid 8$.)

Triviality of the Hecke operators in the noncongruence case

Conclusion: the Hecke operators do not give anything new on noncongruence groups. Too bad!

Yours,



J-P. Serre.

The ending of Serre's letter to Thompson (1989), in which he proved (an important case of) Atkin's conjecture of the triviality of the Hecke operators on the noncongruence modular forms.

The conjecture of Atkin and Swinnerton-Dyer (reprise)

This is the converse to the preceding:

Theorem (Calegari, D., Tang, 2021)

If $f(\tau) \in \mathbb{Q}[[q]]$ is holomorphic on $\mathbf{H}$, meromorphic on $\mathbb{Q} \cup \{\infty\}$, and modular under some finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$, but is not modular under $\Gamma(N)$ for any N, then the Fourier expansion at any cusp has (exponentially) growing denominators.

On reducing to weight-0 $\longrightarrow$ replacing $\mathbb{Z}[[q]] = \mathbb{Z}[[x]]$ in $x := \lambda/16 \longrightarrow$ restricting to $\Gamma(2)$ with $Y(2) = \mathbf{H}/\Gamma(2) \cong \mathbb{P}^1 \setminus \{0, 1/16, \infty\}$, we end up with an algebraic power series $\in \mathbb{Z}[[x]]$ with branching only over $x = 0, 1/16$ and ∞.

The basic question in the language of linear ODEs

Problem

Is there an exact description of all the $\mathbb{Z}[[x]]$ formal solutions to linear ODEs $L(f)=0$ where L ranges over all linear differential operators with coefficients in $\mathbb{C}(x)$ and no singularities outside of $0, 1/16$ and ∞?

A transcendental example is the complete elliptic integral

$$ {}_2F_1\left[\begin{matrix}1/2 & 1/2\\ & 1\end{matrix}; 16x\right] = \sum_{n=0}^{\infty} \binom{2n}{n}^2 x^n \in \mathbb{Z}[[x]] $$

$$ = \frac{1}{\sqrt{1-4x}} * \frac{1}{\sqrt{1-4x}} \quad \text{(Hadamard's product)} $$

André's algebraicity criterion

Theorem (André)

Let $f(x) \in \mathbb{Z}[[x]]$*, and consider a holomorphic mapping* $\varphi : D(0,1) \to \mathbb{C}$ *taking* $\varphi(0) = 0$ *with derivative* $|\varphi'(0)| > 1$*, and such that the germ* $f(\varphi(z)) \in \mathbb{C}[[z]]$ *is analytic on* $|z| < 1$*.*
Then $f(x)$ *is algebraic.*
(If furthermore $\varphi : D(0,1) \hookrightarrow \mathbb{C}$ *is injective, then in fact* $f(x)$ *is rational: this was Pólya's theorem.)*

A basic example: $\sqrt[4]{1-8x} \in \mathbb{Z}[[x]]$ and it meets the criterion with $\varphi(z) := \lambda(z)/8$, of conformal size $|\varphi'(0)| = 2 > 1$.

An interesting boundary case

Since $\lambda(q)$ has conformal size $\lambda'(0) = 16$, we shall scale our branch values by that factor to keep up with $\mathbb{Z}[[x]]$ expansions: $\{0, 1, \infty\} \mapsto \{0, 1/16, \infty\}$.
Let $x := \lambda(q)/16 = q - 8q^2 + \cdots \in q + q^2\mathbb{Z}[[q]]$. We may formally invert that expansion, using the equality of completed rings $\mathbb{Z}[[q]] = \mathbb{Z}[[x]]$, and write

$$q = x + 8x^2 + 91x^3 + \cdots \in x + x^2\mathbb{Z}[[x]].$$

There are infinitely many $\mathbb{Q}(x)$-linearly independent algebraic functions $f(x) \in \mathbb{Z}[[x]]$ such that $f\left(\lambda(q)/16\right) \in \mathbb{Z}[[q]]$ is convergent on the open unit q-disc $|q| < 1$.

They come from congruence modular functions! For each $N = 1, 2, 3, \ldots$, take $\lambda(q^N) \in \mathbb{Z}[[q]] = \mathbb{Z}[[x]]$ written out in terms of $x = \lambda(q)/16$.

A holonomy rank bound

The following will be applied with $t = q^{1/N}$, $p(x) := x^N$,
$x = x(t) := \sqrt[N]{\lambda(t^N)/16} : D(0,1) \to U := \mathbb{C} \setminus 16^{-1/N}\mu_N$, *and*
$\varphi : D(0,1) \to \mathbb{C} \setminus (16)^{-1/N}\mu_N$ *the universal covering map restricted disc.*

Theorem

Let $p(x) \in \mathbb{Q}[x] \setminus \mathbb{Q}$ and $x(t) = t + \cdots \in \mathbb{Q}[[t]]$ be such that $p(x(t)) \in \mathbb{Z}[[t]]$. Fix the holomorphic mapping $\varphi : \overline{D(0,1)} \to U$ with $\varphi(0) = 0$ and $|\varphi'(0)| > 1$. Then, the totality of formal functions $f(x) \in \mathbb{Q}[[x]]$ that

- fulfill a linear ODE over $\mathbb{Q}(x)$ without singularities on U, and
- have a t-expansion $f(x(t)) \in \mathbb{Z}[[t]]$,

span over $\mathbb{Q}(p(x))$ a finite-dimensional vector space of dimension at most

$$e \cdot \frac{\int_{|z|=1} \log^+ |p \circ \varphi| \, \mu_{\mathrm{Haar}}}{\log |\varphi'(0)|}.$$

($e = 2.71\ldots$ is Euler's constant)

The proof of the holonomy theorem

It follows a method of André, itself going back to D. & G. Chudnovsky in their Diophantine approximations proof of the Faltings isogeny theorem for elliptic curves over $\mathbb{Q}$. A crucial new twist (obviously inspired by Thue–Siegel–Schneider–Roth) is to let the number of auxiliary variables $\mathbf{x} := (x_1, \ldots, x_d)$ to $d \to \infty$.

Lemma (Siegel's lemma)

Let A be an $L \times M$-matrix whose entries are rational integers bounded in absolute value by B. Then, if $L > M$, the linear system $A \cdot \mathbf{x} = \mathbf{0}$ of M equations in L variables $x_1, \ldots, x_L$ has a nontrivial integral solution $\mathbf{x} \in \mathbb{Z}^L \smallsetminus \{\mathbf{0}\}$ with

$$\max_{1 \leq i \leq L} |x_i| \leq (LB)^{\frac{M}{L-M}}.$$

A Minkowski argument: pigeonholing a solution.

The proof of the holonomy theorem

Suppose there are m such functions $f_1(x), \ldots, f_m(x) \in \mathbb{Q}[[x]]$ linearly independent over $\mathbb{Q}(p(x))$. We use the m^d split variables univariate products $\prod_{s=1}^d f_{i_s}(x_s)$ and Siegel's lemma to create an auxiliary function of the form:

$$F(\mathbf{x}) = \sum_{\substack{\mathbf{i} \in \{1,\ldots,m\}^d \\ \mathbf{k} \in \{0,\ldots,D-1\}^d}} a_{\mathbf{i},\mathbf{k}}\, p(\mathbf{x})^{\mathbf{k}} \prod_{s=1}^{d} f_{i_s}(x_s) \in (\mathbf{x})^{\alpha}\, \mathbb{Q}[[\mathbf{x}]] \setminus \{0\},$$

with sub-exponentially small coefficients $a_{\mathbf{i},\mathbf{k}} = \exp(o(\alpha))$ as firstly $\alpha \to \infty$ and secondly $d \to \infty$. *With a degree D as low as possible.*

Siegel's lemma: the parameter count

$$F(\mathbf{x}) = \sum_{\substack{\mathbf{i}\in\{1,\dots,m\}^d \\ \mathbf{k}\in\{0,\dots,D-1\}^d}} a_{\mathbf{i},\mathbf{k}}\, p(\mathbf{x})^{\mathbf{k}} \prod_{s=1}^{d} f_{i_s}(x_s) \in (\mathbf{x})^{\alpha}\,\mathbb{Q}[[\mathbf{x}]] \setminus \{0\},$$

- $(mD)^d$ free parameters $a_{\mathbf{i},\mathbf{k}}$
- $\binom{\alpha+d}{d} \sim \alpha^d/d! \approx (e\alpha/d)^d$ equations to solve
- $\#\mathtt{parameters} > \#\mathtt{equations}$ if $dD > e(1+o(1))\frac{\alpha}{m}$ asymptotically
- by letting also $d \to \infty$, we can also make sure the Dirichlet exponent $\to 0$, and the coefficients a are $\exp(o(\alpha))$
- then we can asymptotically take the degree parameter $dD \sim e\,\frac{\alpha}{m}$.

The extrapolation

The idea is that the function $G(\mathbf{z}) := F(\varphi(\mathbf{z})) \in \mathbb{C}[[\mathbf{z}]]$ is analytic on $\overline{D(0,1)}$ (by Cauchy's theorem), and yet since $\varphi(z) = \varphi'(0)z + \cdots$ and $x(t) = t + \cdots$, it also inherits from $f_i(x(t)) \in \mathbb{Z}[[t]]$ and $p(x(t)) \in \mathbb{Z}[[t]]$ an integrality property of its *lexicographically lowest* term $c\,\mathbf{z}^{\beta}$:

- $c \in \varphi'(0)^{|\beta|}\mathbb{Z} \setminus \{0\}$, with total degree $|\beta| \geq \alpha$
- hence the Liouville lower bound for that coefficient: $\log|c| \geq \alpha \log|\varphi'(0)|$
- (A simplification step pointed out to us by André) We can use the plurisubharmonic property of $\log|\mathrm{holomorphic\ function}|$ together with an easy induction scheme on d to prove that, *for our lexicographically lowest monomial* $c\,\mathbf{z}^{\beta}$, we have a bound in the other direction:

$$\log|c| \leq \int_{\mathbf{T}^d} \log|G|\, \mu_{\mathrm{Haar}}.$$

The base case $d = 1$ is simply the subharmonic property of $\log|z^{-\beta} G(z)|$.

The holonomy rank bound: proof completion

- $\alpha \log|\varphi'(0)| \leq \int_{\mathbf{T}^d} \log|G| \, \mu_{\mathrm{Haar}}$
- the RHS is upper estimated by our arithmetic information from the shape of F and the asymptotically subexponential coefficients bound in Siegel's lemma:

$$\alpha \log|\varphi'(0)| \leq \int_{\mathbf{T}^d} \log|G| \, \mu_{\mathrm{Haar}} \leq dD \int_{\mathbf{T}} \log^+|p \circ \varphi| \, \mu_{\mathrm{Haar}} + o(\alpha)$$

- With the degree parameter asymptotic estimate $dD \sim e\alpha/m$, the last inequality amounts in the $\alpha \to \infty$, $d \to \infty$ limit to

$$m \leq e \frac{\int_{\mathbf{T}} \log^+|p \circ \varphi| \, \mu_{\mathrm{Haar}}}{\log|\varphi'(0)|},$$

that is precisely what we aimed to prove.

A reduction to weight 0 (modular functions)

The special case $k = 0$ of modular *functions* is no loss of generality, *viz.* multiplying by a power of Dedekind η, *and it will be assumed later on in our proof.*

To be more precise, consider the Ramanujan modular form of weight 12 on $\mathrm{SL}_2(\mathbb{Z})$:

$$\Delta(\tau/2) := q \prod_{n=1}^{\infty} (1 - q^n)^{24} \in \mathbb{Z}[[q]], \quad q = \exp(\pi i \tau).$$

WLOG, we may assume that $\Gamma \subset \Gamma(2)$. Now

$$f(\tau)^{12} j(\tau/2)^{-k} \Delta(q)^{-k} \in \mathbb{Z}[[q]]$$

is modular of weight 0 on Γ and has integer coefficients.

We pass from $\mathrm{SL}_2(\mathbb{Z}) = \Gamma(1)$ to the mod 2 level $\Gamma(2)$, since the latter acts ***freely*** *on* **H** *while the former does not.*

Fuchsian Uniformization

We need a notion of conformal size for a pointed (connected, open) Riemann surface (U, P). The Riemann–Fuchs–Koebe *uniformization theorem for Riemann surfaces* gives the correct such notion.

Let $D(0,1) = \{|z| < 1\} \subset \mathbb{C}$ be the open unit disc in the complex plane. *Then there exists an essentially* unique analytic universal covering map*

$$F : (D(0,1), 0) \to (U, P), \quad F(0) = P.$$

**Up-to precomposing by a rotation* $z \mapsto e^{i\theta} z$

Definition

We define the **uniformization radius** of the pointed Riemann surface (U, P) to be $|F'(0)|$.

The analytic universal covering of $(\mathbb{C} \smallsetminus \{\pm 1\}, 0)$

$$\tau \mapsto 2\lambda(\tau) - 1, \quad (\mathbf{H}, i) \to (\mathbb{C} \smallsetminus \{\pm 1\}, 0)$$

Due to the accidental isomorphism $Y(2) \cong \mathbb{C} \smallsetminus \mu_2$ mentioned before. The Riemann uniformization radius is thus computed explicitly:

$$\frac{\Gamma(1/4)^2}{4\pi^2} \approx 4.376879\ldots.$$

Compare: If we view $\lambda(q) = 16q + \cdots : \{|q| < 1\} \to \mathbb{C} \smallsetminus \{1\}$ instead as a holomorphic function with singleton fiber $\lambda^{-1}(0) = \{0\}$ on the unit q-disc, then the function $\sqrt{\lambda(q^2)} = 4q + \cdots : D(0,1) \to \mathbb{C} \smallsetminus \{\pm 1\}$ is still holomorphic, and materializes the *strict* lower estimate > 4 on the above uniformization radius, by virtue of factorizing *properly* via the universal covering map: *overconvergence* $>$ *a priori radius*

The universal covering $F_N : (D(0,1), 0) \to (\mathbb{C} \smallsetminus \mu_N, 0)$

The last argument with the analyticity of $\sqrt[N]{\lambda(q^N)} = 16^{1/N} q + \cdots : D(0,1) \to \mathbb{C} \smallsetminus \mu_N$ demonstrates immediately the *strict* (but asymptotically sharp "to zeroth order") lower bound $> 16^{1/N}$ on the uniformization radius $|F_N'(0)|$ at the origin of $\mathbb{C} \smallsetminus \mu_N$.

We need to be much more precise: the above classical theory generalizes, with Poincaré's ODE approach to the uniformization of Riemann surfaces, to describe the multivalued inverse "fairly explicitly" in terms of hypergeometric functions, and derive an exact formula for the uniformization radius (Kraus and Roth, 2016):

$$|F_N'(0)| = \frac{\Gamma\left(\frac{N-1}{2N}\right)^2 \Gamma\left(1+\frac{1}{N}\right)}{\Gamma\left(\frac{N+1}{2N}\right)^2 \Gamma\left(1-\frac{1}{N}\right)}$$

$$= 16^{1/N}\left(1 + \frac{\zeta(3)}{2N^3} + \frac{3\zeta(5)}{8N^5} + \ldots\right) > 16^{1/N}\Big(1 + \zeta(3)/(2N^3)\Big).$$

The universal covering of $\mathbb{C} \setminus \mu_N$ is our path to resolving a $\mathbb{Z}/N$ local monodromy at $x = 0$

- Firstly, we have seen that the unbounded denominators conjecture is secretly an arithmetic property about local systems on $Y(2) \cong \mathbb{P}^1 \setminus \{0, 1, \infty\}$.
- Secondly, if our local system (as is the case in our $k = 0$ situation with UBD, but not in the hypergeometric equation discussed above) has a finite $\mathbb{Z}/N$ local monodromy at the point $\lambda = 0$, then the N-isogeny $\lambda \mapsto \lambda^N$ trades our $\mathbb{C} \setminus \{0, 1\}$ local system into a $\mathbb{C} \setminus \mu_N$ local system.
- We want to exploit an "overconvergence boost" from the fact that the $F_N : D(0,1) \to \mathbb{C} \setminus \mu_N$ pullback of the latter local system is a trivial local system (*no singularities throughout!*) on the disc $D(0,1)$.

"Overconvergence," since that — with or without a finite $\lambda = 0$ local monodromy — were true by fiat *for the pullback under the holomorphic map $\sqrt[N]{\lambda(q^N)} : D(0,1) \to \mathbb{C} \setminus \mu_N$.*

A comparison: a Γ-automorphic function $g(\lambda(q)/16)$ versus the hypergeometric example $G(x) = \sum_{n=0}^{\infty} \binom{2n}{n}^2 x^n$

The former has a $\mathbb{Z}/N$ local monodromy (Puiseux branching) for some N. Then $g(F_N^N(z)/16)$ converges on the full unit disc $|z| < 1$ by Cauchy's analyticity theorem: for it satisfies a linear ODE with analytic coefficients and no singularities on that complex disc. *Overconvergence comes by resolving the $x = 0$ singularity via a suitable $x \mapsto x^N$ isogeny.*

The latter has an infinite (and unipotent) local monodromy at $x = 0$. Now $G(F_N^N(z)/16)$ converges only up to the “first” nonzero fiber point $F_N^{-1}\{0\} \setminus \{0\}$, giving a certain radius rather smaller than 1.

What have we got so far

We use the preceding with the choices

$$t := q^{1/N} = e^{\pi i \tau/N}, \quad x(t) := \sqrt[N]{\lambda(t^N)/16}, \quad U := \mathbb{C} \setminus 16^{-1/N}\mu_N,$$
$$\varphi(z) := 16^{-1/N} F_N(rz) \quad : \quad \overline{D(0,1)} \to U$$

for $r := 1 - 1/(2N^3)$.

Conclusion: the modular functions (weight $k = 0$, both congruence and noncongruence) that have

- $\mathbb{Z}[[q^{1/N}]]$ *Fourier expansions at the one cusp $i\infty$, and*
- *cusp widths dividing N at all the cusps,*

span over $\mathbb{C}(\lambda) = \mathbb{C}(x^N)$ a vector space of dimension at most

$$e \cdot \frac{\int_{|z|=1-1/(2N^3)} \log^+ |F_N^N| \, \mu_{\text{Haar}}}{\log |16^{-1/N} F_N'(0)| + \log r} \ll N^3 \int_{|z|=1-1/(2N^3)} \log^+ |F_N^N| \, \mu_{\text{Haar}}.$$

The $\mathbb{Z}[[q^{1/N}]]$ modular functions with Wohlfahrt level N

By Hecke theory, these include all the $\Gamma(N)$-automorphic functions. Their $\mathbb{C}(\lambda)$-linear span dimension (say N is even) equates to the index formula

$$\frac{1}{2}[\Gamma(2):\Gamma(N)] = \frac{N^3}{2[\mathrm{SL}_2(\mathbb{Z}):\Gamma(2)]}\prod_{p|N}\left(1-\frac{1}{p^2}\right) > \frac{N^3}{12\zeta(2)}.$$

Thus we have these true solutions as an $\gg N^3$ lower bound against the

$$\ll N^3 \int_{|z|=1-1/(2N^3)} \log^+ |F_N^N|\, \mu_{\mathrm{Haar}}$$

upper bound that we just proved.

Extrapolation from a single counterexample $f(q)$

But suppose we have even a single noncongruence counterexample $f(q) \in \mathbb{Z}[[q]]$, of Wohlfahrt (LCM of cusp widths) level N. *Then* $f(q^p) \in \mathbb{Z}[[q]]$ *is another counterexample at a Wohlfahrt level* Np.

An idea going back to Serre from his proof of the triviality of the Hecke operators over noncongruence subgroups — based on an amalgamated sum presentation of $\mathrm{SL}_2(\mathbb{Z}[1/p])$, and on the congruence subgroup property of that S-arithmetic group — proves that this construction is independent over the *congruence* modular forms. And thus by this construction out of a single counterexample at Wohlfahrt level N we reach as many as $2^{\pi(X)}$ *independent* counterexamples at Wohlfahrt level $N\prod_{p<X} p$.

Matching up

Hence, at the Wohlfahrt level $M := N\prod_{p<X} p \asymp_N e^{X+o(X)}$, *we have*

$$\gg M^3 2^{\pi(X)} \gg M^3 2^{X/(2\log X)}$$

examples against our upper bound of

$$\ll M^3 \int_{|z|=1-1/(2M^3)} \log^+ |F_M^M|\, \mu_{\mathrm{Haar}} \ll M^3 X$$

Now $X \sim \log M$ as we let the parameter $X \to \infty$, and so to get the desired constradiction out of a single counterexample $f(q)$, it remains to prove that the integral is sub-exponentially small in $\log M / \log\log M$.
In fact we prove that the integral is $O(\log M)$. *(But an* $\ll_\varepsilon M^\varepsilon$ *bound would not have sufficed.)*

The mean growth of the universal covering map F_N

And this was how in our *arithmetico-analytic continuation* (Gel'fond) argument with Diophantine approximations we came to require a precise — doubly uniform in both $N \geq 2$ and $r < 1$ — upper estimate on the *Nevanlinna mean proximity function at* ∞:

$$m(r, F_N) := \int_{|z|=r} \log^+ |F_N| \, \mu_{\mathrm{Haar}} \ll \frac{1}{N-1} \log \frac{N}{1-r}.$$

The general fact of the matter is: for *any* universal covering map $F : D(0,1) \to \mathbb{C} \setminus \{a_1, \ldots, a_N\}$, we have (Tsuji, 1952) the precise asymptotics as $r \to 1^-$:

$$m(r, F) = \frac{1}{N-1} \log \frac{1}{1-r} + O_{a_1, \ldots, a_N; F(0)}(1).$$

To be contrasted with the crude supremum growth asymptotic formula:

$$\sup_{|z|=r} \log |F| \asymp \frac{1}{1-r}.$$

$m(r, F_N^N) := \int_{|z|=r} \log^+ |F_N^N| \, \mu_{\mathrm{Haar}} \ll \log \frac{N}{1-r}$

- Heuristically this is plausible upon comparing to the renormalized function $F_N(q^{1/N})^N \to \lambda(q)$, the convergence taking place as q-expansions under $N \to \infty$ on any fixed disc of radius $r < 1$. But this convergence is not uniform as $r \to 1$, whereas we will need to take $r = 1 - 1/(2N^3)$.
- The growth of the map F_N is governed by the growth of the cusps of a bounded height of the (N, ∞, ∞) triangle (Fuchsian) group. The above convergence tempts us to compare these cusps (which are to some extent explicit, but of course they vary with N) to the cusps of the limit (∞, ∞, ∞) triangle group $\Gamma(2)$.
- Turns out quite hard! Although in this way we could compute a precise upper estimate on the conformal radii of the sup-level sets $|F_N| < e^M$, that turns out not sufficiently precise for our arithmetico-analytic continuation argument. The requisite mean (integrated) bound above can be translated in terms of a uniform cusp count, but the latter goes beneath what we could directly prove.

Instead, we were surprised to find that inner workings of the (general, abstract) *second main theorem* of Nevanlinna's value distribution theory exactly sufficed to give the double uniformity that we needed.

Enter Nevanlinna's lemma on the logarithmic derivative

Theorem

$m(r, F_N^N) := \int_{|z|=r} \log^+ |F_N^N| \mu_{\mathrm{Haar}} \ll \log \frac{N}{1-r}$ (an absolute and computable explicit coefficient)

Idea: By construction, the function $f := 1 - F_N^N$ is a *functional unit*, and we have a factorization in terms of logarithmic derivatives:

$$\frac{F_N^N}{F_N^N - 1} = \frac{F_N}{NF_N'} \frac{f'}{f} \text{ “} \approx \text{''} N^{-1} \frac{1 - F_N}{(1 - F_N)'} \frac{f'}{f}$$

This formula is what gets specially used about the target set $\mu_N \cup \{\infty\}$ of omitted values. (The first equality is exact, and the second is approximate on the level of the Nevanlinna characteristic function, but it now involves the two functional units $f = 1 - F_N^N$ and $1 - F_N$).

The LHS has mean proximity function $= N\, m(r, F_N) + O(1)$, so it is what we want to bound. The RHS is $m(r, F_N) + \texttt{small}$, essentially because logarithmic derivatives are small on average over circles. (Lemma on the logarithmic derivative.)

A simplified, explicit case of the lemma of the logarithmic derivative suffices for our purposes

Apply the lemma on the logarithmetic derivative to the two functional units $g := 1 - F_N^N$ and $g := 1 - F_N$, followed by a standard transformation sequence from the lemma on the logarithmic derivative $\to$ Nevanlinna's second main theorem, and a trivial supremum estimation of the emerging *double* log growth term in this bound:

Lemma

Let $g : \overline{D(0,R)} \to \mathbb{C}^\times$ *be a nowhere vanishing holomorphic function on some open neighborhood of the closed disc* $|z| \le R$*. Assume that* $g(0) = 1$*. Then, for all* $0 < r < R$*,*

$$m\Big(r, \frac{g'}{g}\Big) < \log^+ \Big\{ \frac{m(R,g)}{r} \frac{R}{R-r} \Big\} + \log 2 + 1/e.$$

Quintessence of the proof of this form of the lemma of the logarithmic derivative

It is based on Poisson's kernel, which by a simple differentiation in term yields the reproducing formula for our logarithmic derivative:

$$\frac{g'(z)}{g(z)} = \int_{|w|=R} \frac{2w}{(w-z)^2} \log|g(w)|\, \mu_{\mathrm{Haar}}(w), \quad \forall z \in D(0,R).$$

Follow it by easy estimations based on Jensen's inequality and the concavity of $\log^+|x|$ on $[1,\infty)$.

To complete the general picture: For an arbitrary meromorphic mapping $g : D(0,1) \to \mathbb{P}^1$ with $g(0) = 1$ (now possibly having a nonempty divisor of zeros and poles), if we use the full *Nevanlinna characteristic* $T(r,g) = m(r,g) + N(r,g)$, Gol'dberg and Grinshtein proved that the same type of bound persists:

$$m\Big(r, \frac{g'}{g}\Big) < \log^+\Big\{\frac{T(R,g)}{r}\frac{R}{R-r}\Big\} + 5.8501.$$

Essentially best-posible in form, and comparable to Lang's conjecture on the error term in Roth's theorem under the Osgood-Vojta dictionary to Diophantine approximation.

Conclusion of the doubly uniform mean growth bound on F_N

By the factorization identity and the lemma on the logarithmic derivative we got, uniformly $\forall 0 < r < R$ and $\forall N \geq 2$, to

$$N\, m(r, F_N) \leq m(r, F_N) + O\Big(\log^+ \Big\{ N \frac{m(R, F_N)}{r} \frac{R}{R-r} \Big\}\Big).$$

Choose $R = (1+r)/2$.
Since the $\log^+ m(R, F_N)$ erro term is *logarithmic*, it is alright to estimate it crudely by the supremum

$$\log^+ \sup_{|z|=R} \log |F_N| \ll \log \frac{N}{1-R},$$

by a simple geometric estimate based on Shimizu's lemma.

A (naive) question of the Lehmer variety

Let

$$f(q) = \sum_{r \in \mathbb{Q}^{\geq 0}} \frac{a(r)}{b(r)} q^r \in \mathbb{Q}[[q^{1/N}]], \quad \gcd\bigl(a(n), b(n)\bigr) = 1$$

be the q-expansion of a noncongruence modular form. Having proved the unboundedness $\limsup_{r\to\infty} b(r) \to \infty$, it becomes a natural question how slowly may these denominators grow in terms of the rate

$$\delta(f) := \limsup_{r\to\infty} \frac{1}{r} \log|b(r)|.$$

A (naive) question of the Lehmer variety

$$\delta(f) := \limsup_{r\to\infty} \frac{1}{r} \log|b(r)|.$$

In our discussion, it is natural to measure the growth with respect to the parameter $N = N(f) :=$ the LCM of the cusp widths of f in the $\Gamma(2)$-orbit of the infinite cusp $i\infty$.

Observe that the product $N(f) \cdot \delta(f)$ remains invariant under changing $q \mapsto q^m$, $\forall m \in \mathbb{N}$.

Hence the theoretically best-possible lower growth lower boun would look like $\delta(f) \gg N(f)^{-1}$ (unless f is congruence). *Is that bound possibly true?*

To compare: tracking our $f(q) \mapsto f(q^p)$ extrapolation argument, one can easily make everything quantitative (how large the parameter $X \to \infty$ really needed to be in terms of the initial N, etc.), and prove the following weaker lower bound on denominators growth of noncongruence modular forms: $\delta(f) \gg N(f)^{-3-o(1)}$.

Thank you for your invitation and attention, everyone stay safe and well!