2009/09/07. <u>From sums of squares to modular forms</u> Bas. <u>1</u>.

(2 × 45 minutes.) (has to do with Ila Varma's master thesis subject).

<u>References</u>: • Exercise 4.8.7 of Diamond & Shurman, A first course in modular forms.
• The end of Serre's Cours d'arithmétique

1. <u>Sums of squares</u>. A classical subject.

For $n \in \mathbb{Z}_{\geq 0}$, $m \in \mathbb{Z}_{\geq 0}$: $r_{\mathbb{Z}^n}(m) := \#\{x \in \mathbb{Z}^n \mid x_1^2 + \dots + x_n^2 = m\}$

Very rough estimates:

$$\left(\frac{2\sqrt{m}}{n}\right)^n \leqslant \sum_{k=0}^{m} r_{\mathbb{Z}^n}(k) \leqslant \left(2\sqrt{m}+1\right)^n$$

Take log: $\log(r_{\mathbb{Z}^n}(m)+1) = O(n \cdot \log m)$, and we have a similar lower bound for $\max\{r_{\mathbb{Z}^n}(k) \mid 0 \leq k \leq m\}$.

Results for small n.

$n=1$: $r_{\mathbb{Z}}(m) =$ 0 if m not square, 1 if $m=0$, 2 if $m>0$ is square.

$n=2$. (Fermat ...) Has to do with $\mathbb{Z}[i]$, splitting of primes.

Let $\chi : \mathbb{Z}/4\mathbb{Z} \longrightarrow \mathbb{C}$, $0,2 \longmapsto 0$, $1 \longmapsto 1$, $-1 \longmapsto -1$; the Dirichlet char. associated to $\mathbb{Q} \to \mathbb{Q}(i)$, $\mathrm{Gal}(\mathbb{Q}(i)/\mathbb{Q}) = (\mathbb{Z}/4\mathbb{Z})^\times$, extended by 0.

$\forall m \in \mathbb{Z}_{>0}$:

Then: $r_{\mathbb{Z}^2}(m) = 4 \cdot \sum_{d|m} \chi(d)$.

$n=3$: (Gauss) for $m \in \mathbb{Z}_{\geq 0}$: $r_{\mathbb{Z}^3}(m) = 0 \iff \left\{\begin{array}{l}\exists\, a \in \mathbb{Z}_{\geq 0},\ b \in \mathbb{Z} \\ \text{s.t. } m = 4^a \cdot (8b-1)\end{array}\right.$

The cases with n odd are too complicated for us; I'll not discuss them further.

$n=4$: (Jacobi): $r_{\mathbb{Z}^4}(m) = 8 \cdot \sum_{2 \nmid d | m} d + 16 \cdot \sum_{2 \nmid d | (m/2)} d$.

$n=6$: $\}$ $\exists$ similar formulas, certainly known, for Ila to find them.

$n=8$: (<u>Marmeth</u>: $r_{\mathbb{Z}^8}(m) = 16\,\sigma_3(m) - 32 \cdot \sigma_3(m/2) + 256\,\sigma_3(m/4)$)

$n=10$: $r_{\mathbb{Z}^{10}}(m) = \frac{4}{5} \cdot \sum_{d|m} \chi(d) \cdot d^4 + \frac{64}{5} \cdot \sum_{d|m} \chi(m/d) \cdot d^4 + \frac{8}{5} \cdot \sum_{\substack{d \in \mathbb{Z}[i],\\ |d|^2 = m}} d^4$

$n=12$: $\nexists$ similar formula.

Subject of the master thesis: find <u>all</u> such formulas.

The proofs of all these results use the theory of <u>modular forms</u>. ($n \neq 2, 3$)

$r_{\mathbb{Z}^8}(m) = 16 \cdot \sum_{d|m} d^3 - 32 \cdot \sum_{d|m/2} d^3 + 256 \sum_{d|m/4} d^3$.

2. <u>Theta functions</u>. These are generating series for the sequences of integers $r_{\mathbb{Z}^n}(m)$, where n is fixed and m varies. The motivation to look at this comes from the identity:

Let $n = n_1 + n_2$, $\{x \in \mathbb{Z}^n \mid x_1^2 + \cdots + x_n^2 = m\} = \coprod_{m = m_1 + m_2} \{(x_1, \ldots, x_{n_1}) \mid x_1^2 + \cdots + x_{n_1}^2 = m_1\} \times \{(x_{n_1+1}, \ldots x_n) \mid x_{n_1+1}^2 + \cdots + x_n^2 = m_2\}$

Identity in $\mathbb{Z}[[q]]$ (formal power series):

$$\sum_{x_1 \in \mathbb{Z}^{n_1}} q^{\|x_1\|^2} \cdot \sum_{x_2 \in \mathbb{Z}^{n_2}} q^{\|x_2\|^2} = \sum_{(x_1, x_2) \in \mathbb{Z}^{n_1+n_2}} q^{\|x_1\|^2 + \|x_2\|^2} = \underbrace{\sum_{x \in \mathbb{Z}^n} q^{\|x\|^2}}_{\sum_{m \geq 0} r_{\mathbb{Z}^n}(m) \cdot q^{m/2}}.$$

So we define: $\Theta_{\mathbb{Z}^n} := \sum_{x \in \mathbb{Z}^n} q^{\|x\|^2/2}$, then $\Theta_{\mathbb{Z}^n} = \Theta_{\mathbb{Z}}^n$.

These are called theta functions. More generally, for L a free $\mathbb{Z}$-module of finite rank with $b: L \times L \to \mathbb{Z}$ symmetric, bilinear and positive definite, one defines $\Theta_{L,b} := \sum_{x \in L} q^{b(x,x)/2}$.

These theta functions have an unexpected symmetry; a functional equation. Let $\mathbb{H} := \{z \in \mathbb{C} \mid \operatorname{Im}(z) > 0\}$, $q^{1/2}: \mathbb{H} \to \mathbb{C}$, $z \mapsto e^{\pi i z}$.

<u>Thm (whose?)</u> $\forall n \in \mathbb{Z}_{\geq 0}$, $\forall z \in \mathbb{H}$: $\Theta_{\mathbb{Z}^n}(-1/z) = (-iz)^{n/2} \cdot \Theta_{\mathbb{Z}^n}(z)$, where $z \mapsto (-iz)^{1/2}$ is holomorphic, and in $\mathbb{R}_{>0}$ for $z \in i \cdot \mathbb{R}_{>0}$.

Reference for the proof (in a more general case): Serre's Cours d'arithmétique, Ch. <u>VII</u>, §6, Prop 16. It is based on Poisson's summation formula.

As $|r_{\mathbb{Z}^n}(m)| = O(m^{n/2})$, the sum $\sum_{m \geq 0} r_{\mathbb{Z}^n}(m) \cdot q^{m/2}$ defines a holomorphic function on $\mathbb{H}$.

Let me give the main ideas of the proof. Of course the case $n=1$ suffices for us.

For $y \in \mathbb{R}_{>0}$ we have: $\Theta_{\mathbb{Z}}(iy) = \sum_{m\in\mathbb{Z}} \left(e^{\pi i \cdot iy}\right)^{m^2} = \sum_{m\in\mathbb{Z}} e^{-\pi m^2 y} =$

$$= \sum_{m\in\mathbb{Z}} f_y(m), \quad \text{where } f_y : \mathbb{R} \to \mathbb{R},\ x \mapsto e^{-\pi x^2 \cdot y}$$

Poisson's summation formula says that $\sum_{m\in\mathbb{Z}} f_y(m) = \sum_{m\in\mathbb{Z}} \hat{f}_y(m)$,

where $\hat{f}_y : \mathbb{R} \to \mathbb{R}$, $x \mapsto \int_{t\in\mathbb{R}} e^{-2\pi i x t} \cdot f_y(t)\, dt$ (Fourier transform).

Now f_y is, up to a constant factor, the Gaussian normal distribution on $\mathbb{R}$ with expected value 0 and variance $\frac{1}{\sqrt{y}}$. Therefore, $\hat{f}_y$ is, up to a constant factor (the same one) the Gaussian normal distribution with exp. value 0 and variance $\sqrt{y}$ (this is one of the very standard Fourier transforms that one learns).

So: $\Theta_{\mathbb{Z}}(iy) = \sum_{m\in\mathbb{Z}} f_y(m) = \sum_{m\in\mathbb{Z}} \hat{f}_y(m) = \frac{1}{\sqrt{y}} \sum_{m\in\mathbb{Z}} e^{-\pi m^2/y} =$

$$= \frac{1}{\sqrt{y}} \Theta_{\mathbb{Z}}(i/y), \quad i/y = -1/iy. \quad \boxtimes$$

to be precise: $\hat{f}_y(x) = \frac{1}{\sqrt{y}} \cdot e^{-\pi x^2/y}$

Intermezzo: functional equation for Riemann ζ-function.

We have now: $\Theta_{\mathbb{Z}} = \sum_{m \in \mathbb{Z}} q^{m^2/2}$, $\Theta_{\mathbb{Z}}(-1/z) = (-iz)^{1/2} \cdot \Theta_{\mathbb{Z}}(z)$.

Applying the Mellin transformation to this gives ~~merom. ext. to $\mathbb{C}$~~ + the funct. eq. of ζ, where $\zeta : \mathbb{C}_{\mathrm{Re}>1} \to \mathbb{C} : s \mapsto \sum_{n=1}^{\infty} n^{-s}$.

For details, see ~~[illegible]~~, § 3.2;
Miyake, Modular forms
Diamond-Shurman, § 4.9.

For $f : \mathbb{R}_{>0} \to \mathbb{C}$ one tries to define

$$M(f) : \mathbb{C} \to \mathbb{C}, \quad s \mapsto \int_{t=0}^{\infty} f(t) \cdot t^s \cdot \frac{dt}{t}$$

Then one has

$$\Bigg(M\Big(t \mapsto \underbrace{\frac{\Theta_{\mathbb{Z}}(it)-1}{2}}_{=\sum_{n=1}^{\infty} e^{-\pi n^2 t}}\Big)\Bigg)\left(\frac{s}{2}\right) = \pi^{-s/2}\,\Gamma\!\left(\frac{s}{2}\right) \cdot \zeta(s) =: \xi(s),$$

for $\mathrm{Re}(s) > 1$
(one uses the funct. eq. for $\Theta_{\mathbb{Z}}$ to see what happens near $t=0$).

Then one gets that:
for all $s \in \mathbb{C}$, $s \neq 0, 1$:

$$\xi(s) = \underbrace{\frac{1}{2}\int_{t=1}^{\infty} (\Theta_{\mathbb{Z}}(it) - 1)\left(t^{s/2} + t^{(1-s)/2}\right) \frac{dt}{t}}_{\text{holom. on } \mathbb{C}} - \frac{1}{s} - \frac{1}{1-s}$$

and: $\xi(1-s) = \xi(s)$.

[This method was ~~used~~ ~~in more~~ extended by Hecke to prove analytic continuation and functional equation for L-functions associated to Hecke characters.]

3. Modular forms.

We have seen that $\forall n \in \mathbb{Z}_{\geq 0}$ $\forall z \in \mathbb{H}$:

$$\begin{cases} \Theta_{\mathbb{Z}^n}(z+2) = \Theta_{\mathbb{Z}^n}(z) \\ \Theta_{\mathbb{Z}^n}(-1/z) = (-iz)^{n/2} \cdot \Theta_{\mathbb{Z}^n}(z). \end{cases}$$

In order to understand the consequences of this, we will study the action $Gl_2(\mathbb{R})^+ \times \mathbb{H} \to \mathbb{H}$, $\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, z\right) \mapsto \dfrac{az+b}{cz+d}$, and functions ~~on~~ $f: \mathbb{H} \to \mathbb{C}$ that transform ~~as~~ as: for some $k \in \mathbb{Z}$ --

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k \cdot f(z) \text{ for } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ in certain subgroups of } Sl_2(\mathbb{Z}).$$

(↖ I want to explain this....)

Recall: $\mathbb{P}^1(\mathbb{C}) = \mathbb{C}^\times \backslash (\mathbb{C}^2 - \{0\}) = \{1\text{-dim. sub-}\mathbb{C}\text{-vect. spaces of } \mathbb{C}^2\}$

$$\mathbb{P}^1(\mathbb{C}) \circlearrowleft Gl_2(\mathbb{C}) \qquad \mathbb{P}^1(\mathbb{C}) \xrightarrow{\sim} (\mathbb{C} \amalg \{\infty\})$$

$$\left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right] \longmapsto \begin{cases} z_1/z_2 & \text{if } z_2 \neq 0 \\ \infty & \text{if } z_2 = 0. \end{cases}$$

$$\left[\begin{pmatrix} z \\ 1 \end{pmatrix}\right] \longleftarrow\!\shortmid z \in \mathbb{C}$$

This explains the action by fractional linear transformations:

$$\begin{array}{ccc} \left[\begin{pmatrix} z \\ 1 \end{pmatrix}\right] & \longleftarrow\!\shortmid & z \\ \downarrow & & \\ \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} z \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} az+b \\ cz+d \end{pmatrix}\right] & \longmapsto & \dfrac{az+b}{cz+d}. \end{array}$$

so: $Gl_2(\mathbb{R})$ acts on $\mathbb{H}^\pm$

Note: $\mathbb{P}^1(\mathbb{C}) = \mathbb{H}^\pm \amalg \mathbb{P}^1(\mathbb{R})$, where $\mathbb{H}^\pm = \mathbb{C} - \mathbb{R}$.

$Gl_2(\mathbb{R})^+ := \{g \in Gl_2(\mathbb{R}) \mid \det(g) > 0\}$ preserves $\mathbb{H}^+ := \mathbb{H}^+$ and $\mathbb{H}^-$.

For $z \in \mathbb{H}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Gl_2(\mathbb{R})^+$: $\operatorname{Im}\left(\dfrac{az+b}{cz+d}\right) = \left(\dfrac{az+b}{cz+d} - \dfrac{a\bar{z}+b}{c\bar{z}+d}\right)/2i =$

$$= \ldots = \frac{\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \operatorname{Im}(z)}{|cz+d|^2}, \text{ a useful formula.}$$

We follow (more or less) Serre, Cours d'arithm. Ch. VII §2.2.

Consider:

$$\begin{tikzcd} \mathbb{P}^1(\mathbb{C}) & \mathbb{C}^2 - \{0\} \arrow[l] \text{ for "basis"} \\ \mathbb{H} \arrow[u, phantom, "\cup"] & B \arrow[l, "\text{quot}"'] \arrow[u, phantom, "\cup"] \end{tikzcd}$$

$B = \{ (z_1, z_2) \in \mathbb{C}^2 - \{0\} \mid \mathbb{R}z_1 + \mathbb{R}z_2 = \mathbb{C},\ \mathrm{Im}(z_1/z_2) > 0 \}$

$z_1/z_2 \longleftarrow\!\shortmid (z_1, z_2)$

$z \longmapsto (z, 1)$ (sec)

$\circlearrowleft\ \mathbb{C}^\times \times \mathrm{Gl}_2(\mathbb{R})^+$: $\lambda \cdot (z_1, z_2) = (\lambda z_1, \lambda z_2)$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}$

quot is ~~the~~ a quotient for the $\mathbb{C}^\times$-action, and sec is a section of quot.

We view B as {lattices ~~L~~ in $\mathbb{C}$ with oriented $\mathbb{Z}$-basis}

$$(z_1, z_2) \longmapsto (\mathbb{Z}z_1 + \mathbb{Z}z_2 \subset \mathbb{C}, (z_1, z_2)).$$

~~Let~~ $L := \{\text{lattices in } \mathbb{C}\}$, ~~then~~ $B \longrightarrow L$, $(z_1, z_2) \mapsto \mathbb{Z}z_1 + \mathbb{Z}z_2$ is the quotient for the $\mathrm{Sl}_2(\mathbb{Z})$-action on B ($\mathrm{Sl}_2(\mathbb{Z}) \subset \mathrm{Gl}_2(\mathbb{R})^+$).

For $k \in \mathbb{Z}$ we define: $M_k(B) := \{ f : B \to \mathbb{C} \mid \forall \lambda \in \mathbb{C}^\times, \forall b \in B : f(\lambda \cdot b) = \lambda^{-k} \cdot f(b) \}$

(think of decomposing $\{f : B \to \mathbb{C}\}$ for the action by $\mathbb{C}^\times$ on B)

As the $\mathbb{C}^\times$-orbits in B are precisely the ~~sets of all~~ fibers of quot, and the $\mathbb{C}^\times$ action is free, we have:

$$M_k(B) \underset{}{\overset{\text{sec}^*}{\rightleftarrows}} M_k(\mathbb{H}) = \{ g : \mathbb{H} \to \mathbb{C} \} \quad \text{(inverse maps)}$$

$$(z_1, z_2) \mapsto z_2^{-k} \cdot h(z_1/z_2) \longleftarrow\!\shortmid h$$

$$(z_1, z_2) = z_2 \cdot (z_1/z_2, 1)$$

Consequence: the right-action of $\mathrm{Gl}_2(\mathbb{R})^+$: $(f \cdot g)(b) := f(g \cdot b)$ ($f \in M_k(B)$, $g \in \mathrm{Gl}_2(\mathbb{R})^+$) on $M_k(B)$ gives a right action on $M_k(\mathbb{H})$:

$$(h \cdot g)(z) = \left((z_1, z_2) \mapsto z_2^{-k} \cdot h(z_1/z_2) \right) \left(g \cdot \begin{pmatrix} z \\ 1 \end{pmatrix} \right) =$$

$$\left(\text{write: } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) \qquad = (cz + d)^{-k} \cdot h\left(\frac{az + b}{cz + d} \right).$$

Here is the explanation that I wanted to give.