二次型的表示数问题

、Aug 2025

1 平方和函数

1.1 定义与范畴

研究整数的加法性质是数论中的一个核心问题。平方和问题即考察一个正整数 n 可以以多少种表示形式写作 k 个正整数的平方和，是该领域的一个经典课题。于是我们定义了平方和函数 $r_k(n)$，该函数记录了整数 n 表示为 k 个整数平方和的表示法数量，其中改变项的顺序或符号均会被视作不同的表示法。其形式化的定义为：

$$r_k(n) = \left|\left\{(x_1, x_2, \cdots, x_k) \in \mathbb{Z}^k | n = x_1^2 + x_2^2 + \cdots + x_k^2\right\}\right| \tag{1}$$

例如 $r_2(1) = 4$，因为 $1 = (\pm 1)^2 + 0^2 = 0^2 + (\pm 1)^2$。这一定义将其与幂级数直接联系起来。

1.2 经典的存在性定理

在模形式理论被用于解决平方和问题之前，对于一些特殊情况已经有了少许进展，这些工作的完成这今天来看都是非常著名的数学家。虽然大多数成果都仅局限于存在性的证明，而没有定量计算有多少种可能的表达方式，但也无疑是数论中较为精彩的结果。

经典的平方和定理

- **Fermat 平方和定理**（$k = 2$）：一个奇素数 p 可以写作两个平方数之和，当且仅当 $p \equiv 1 \mod 4$。该定理为 $r_2(p) > 0$ 提供了判别准则，但并未给出具体数值。学习过经典数论的同学可以使用 Gauss 整环给出一个证明。
- **Lagrange 四平方和定理**（$k = 4$）：每个正整数都可以表为四个整数的平方和。这一定理保证了对于所有 $n > 0$，均有 $r_4(n) \geq 1$，这是一个普遍性的结论。
- **Legendre 三平方和定理**（$k = 3$）：一个正整数 n 可以表为三个整数的平方和，当且仅当 n 不具有 $4^a(8b + 7)$ 的形式，其中 a, b 为非负整数。同样地，这是一个条件性的存在定理。

1.3 计数问题

模形式的核心贡献在于，其不仅解决了是否可以表为平方和问题的存在性问题，同时将其推进到了有多少种表示形式的计数问题。Jacobi 使用椭圆函数和 Theta 函数为 $k = 2, 4, 6, 8$ 等情况下的 $r_k(n)$ 函数给出了精确的解析表达式。例如 Jacobi 四平方和定理断言

$$r_4(n) = 8 \sum_{d \mid n, 4 \nmid d} d \tag{2}$$

这些精确公式是我们后续利用模形式理论推导的目标。

2 Thata 函数

我们将建立 Theta 函数，它联系了算术问题和分析问题。

2.1 Jacobi Theta 函数的定义

Jacobi Theta 函数是一个定义在上半平面 $\mathbb{H} = \tau \in \mathbb{C} \mid \mathrm{Im}(\tau) > 0$ 上的复解析函数，它由一个 q-级数定义：

$$\theta(\tau) = \sum_{n \in \mathbb{Z}} q^{n^2} \tag{3}$$

其中 $q = \mathrm{e}^{2\pi\sqrt{-1}\tau}$。由于 $\tau \in \mathbb{H}$，我们有 $|q| < 1$，确保了该级数绝对收敛，并在 $\mathbb{H}$ 上定义了一个全纯函数。

2.2 生成函数恒等式

Theta 函数与平方和问题的联系通过一个恒等式实现。通过展开 Theta 函数的 k 次幂，我们有：

$$
\begin{aligned}
\theta(\tau)^k &= \left(\sum_{n_1=-\infty}^{\infty} q^{n_1^2}\right)\cdots\left(\sum_{n_k=-\infty}^{\infty} q^{n_k^2}\right) && (4)\\
&= \sum_{n_1,\ldots,n_k\in\mathbb{Z}} q^{n_1^2+\ldots+n_k^2} && (5)\\
&= \sum_{n=0}^{\infty}\left(\sum_{\substack{n_1,\ldots,n_k\in\mathbb{Z}\\ n_1^2+\ldots+n_k^2=n}} 1\right) q^n && (6)\\
&= \sum_{n=0}^{\infty} r_k(n) q^n && (7)
\end{aligned}
$$

这个恒等式成为整个方法的基石，将一个纯粹的算术问题转化为分析问题，这种思考在使用模形式与自守形式解决数论问题的过程中具有重要的地位，并将反复出现。

3 模群及其在上半平面的作用

模形式的对称性由模群描述，我们先回顾一下模形式的基本概念：

3.1 模群 $\mathbf{SL}(2,\mathbb{Z})$

模群 $\mathbf{SL}(2,\mathbb{Z})$ 定义为所有行列式为 1 的 2×2 整数矩阵构成的群：

$$\mathbf{SL}(2,\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \, a,b,c,d\in\mathbb{Z}, ad-bc=1 \right\} \tag{8}$$

3.2 在 $\mathbb{H}$ 上的作用

模群通过分式线性变换作用于上半平面 $\mathbb{H}$ 。对于任意 $\gamma = \begin{pmatrix} a & b\ c & d \end{pmatrix} \in \mathbf{SL}(2, \mathbb{Z})$ 和 $\tau \in \mathbb{H}$，其作用定义为：

$$\gamma(\tau) = \frac{a\tau + b}{c\tau + d} \tag{9}$$

可以证明，这一作用保持 $\mathbb{H}$ 不变，因为 $\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}$，当 $\mathrm{Im}(\tau) > 0$ 时，$\mathrm{Im}(\gamma\tau)$ 也大于 0。

3.3 生成元 S 和 T

一个关键的计算事实是，$\mathbf{SL}(2, \mathbb{Z})$ 由两个基本矩阵生成：平移矩阵 $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ 和反演矩阵 $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ 。这意味着，要验证一个函数是否满足对所有模群元素的变换性质，只需验证它对 S 和 T 成立。

3.4 完全模群 $\mathbf{SL}(2, \mathbb{Z})$ 的模形式

一个函数 $f : \mathbb{H} \to \mathbb{C}$ 被称为权为 k 的 $\mathbf{SL}(2, \mathbb{Z})$ 模形式，需满足以下三个条件：

- f 在 $\mathbb{H}$ 上是全纯函数。
- 对所有 $\begin{pmatrix} a & b\ c & d \end{pmatrix} \in \mathbf{SL}(2, \mathbb{Z})$ $\quad f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau + d)^k f(\tau)$。
- f 在“无穷远点”是全纯的。这一条件通过 q-展开式来精确刻画。由变换矩阵 T 可知 $f(+1) = f()$，这意味着 f 可以表示为 $q = e^{2\pi i\tau}$ 的函数 $g(q)$。在无穷远点全纯意为 $g(q)$ 在 $q = 0$ 处没有负幂次项，即具有泰勒级数展开 $g(q) = \sum_{n\geq 0} a_n q^n$。

4 模形式的向量空间

对于给定的权 k，所有权为 k 的模形式构成的集合 $M_k(\mathbf{SL}(2, \mathbb{Z}))$ 是一个有限维复向量空间。其中常数项 $a_0 = 0$ 的模形式称为尖点形式，它们构成 M_k 的一个子空间 $S_k(\mathbf{SL}(2, \mathbb{Z}))$。

M_k 的有限维性是整个方法论的核心，它将分析问题转化为代数问题。通常，要证明两个全纯函数

f 和 g 相等，需要在一个开集上验证它们的取值。然而，如果 f 和 g 都是权为 k 的模形式，它们同属于一个维度为 d 的有限维向量空间。线性代数理论告诉我们，这样的函数由 d 个参数唯一确定。一个方便的选择是它们 q-展开式的前 d 个系数 $a_0, a_1, \cdots, a_{d-1}$。因此，要证明 $f = g$，我们无需比较无穷多个系数，只需验证它们的前 d 个系数是否一致即可。这个原理将一个无穷问题简化为一个有限的、可计算的代数验证过程，通常被称为 "Sturm 界"。

4.1 Eisenstein 级数

对于偶数 $k > 2$，我们将构造出模形式，即 Eisenstein 级数：

$$G_k(\tau) = \sum_{(m,n)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m\tau+n)^k} \tag{10}$$

这些级数是权为 k 的 $\mathbf{SL}(2,\mathbb{Z})$ 模形式。更重要的是，它们的 q-展开式可以被精确计算，并且与数论中的除数和函数 $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$ 直接相关。经过归一化处理的 Eisenstein 级数 $E_k(\tau)$ 具有如下展开式：

$$E_k(\tau) = 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n \tag{11}$$

其中 B_k 是伯努利数。

4.2 $M(\mathbf{SL}(2,\mathbb{Z}))$ 的环结构

$\mathbf{SL}(2,\mathbb{Z})$ 的模形式环具有一个经典结构，即它是由 E_4 和 E_6 生成的多项式代数，即 $M(\mathbf{SL}(2,\mathbb{Z})) = \mathbb{C}[E_4, E_6]$。这意味着任何权为 k 的模形式 $f \in M_k(\mathbf{SL}(2,\mathbb{Z}))$ 都可以唯一表示成 E_4 和 E_6 的多项式 $\sum c_{a,b}E_4^a E_6^b$，其中 $4a + 6b = k$。

5 同余子群的模形式

正如所见，$\theta(\tau)^k$ 通常情况下并非对完全模群 $\mathbf{SL}(2,\mathbb{Z})$ 是一个模形式，此时我们将引入同余子群。

5.1 同余子群的定义

我们主要关注 $\Gamma_0(N)$ 型同余子群，其定义为

$$\Gamma_0(N) = \left\{ \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \in SL(2,\mathbb{Z}) \middle| c \equiv 0 \pmod N \right\} \tag{12}$$

5.2 $\Gamma_0(N)$ 的模形式

模形式的定义可以推广至这些同余子群，其中的变换律只需要对于 $\Gamma_0(N)$ 成立，并且全纯性条件只需在 $\Gamma_0(N)$ 的所有尖点（而不只是无穷远点）满足。

下表罗列了一些权 k 下所对应的 $\theta(\tau)^k$，下面我们将依次进行探讨。

权为 $4, 8, 24$ 时的 Theta 函数

k	生成函数	权	水平(Γ)	$\dim(M_k(\Gamma))$	基元
4	$\theta(\tau)^4$	2	$\Gamma_0(4)$	2	$\Gamma_0(4)$ 的 Eisenstein 级数
8	$\theta(\tau)^8$	4	$SL(2,\mathbb{Z})$	1	E_4
24	$\theta(\tau)^{24}$	12	$SL(2,\mathbb{Z})$	2	E_{12}, Δ

6 Theta 函数的模性

6.1 $\theta(\tau)$ 的变换律

基础 Theta 函数 $\theta(\tau)$ 具有以下变换性质：

1. $\theta(\tau+1) = \theta(\tau)$
2. $\theta\left(-\frac{1}{\tau}\right) = (-\sqrt{-1}\tau)^{1/2}\theta(\tau)$

第一个性质由定义显而易见。第二个更为深刻的性质可以通过泊 Poisson 求和公式证明

6.2 证明 $\theta(\tau)^k$ 的模性

利用上述基本变换律，我们可以推导出 $\theta(\tau)^k$ 的模性质。

- 对于 $\theta(\tau)^8$，在 S 变换下

$$\theta\left(-\frac{1}{\tau}\right)^8 = \left((-\sqrt{-1}\tau)^{1/2}\right)^8 \theta(\tau)^8 = (-\sqrt{-1}\tau)^4\theta(\tau)^8 = \tau^4\theta(\tau)^8 \tag{13}$$

  这正是权为 4 的模形式的变换要求。因此，$\theta(\tau)^8$ 是 $\mathbf{SL}(2,\mathbb{Z})$ 上权为 4 的模形式。

- 同理，$\theta(\tau)^{24}$ 是 $\mathbf{SL}(2,\mathbb{Z})$ 上权为 12 的模形式。
- 对于 $\theta(\tau)^4$，情况更为复杂。它不满足对完全模群的变换律，但可以证明，它满足对同余子群 $\Gamma_0(4)$ 的变换要求，是 $\Gamma_0(4)$ 上权为 2 的模形式。

7 八平方和问题

根据之前的分析，我们知道 $\theta(\tau)^8 \in M_4(\mathbf{SL}(2,\mathbb{Z}))$，并且对于权 $k=4$，我们有 $\dim(M_4(\mathbf{SL}(2,\mathbb{Z}))) = \lfloor 4/12 \rfloor + 1 = 1$，这是一个一维向量空间。

Eisenstein 级数 $E_4(\tau)$ 是 $M_4(\mathbf{SL}(2,\mathbb{Z}))$ 中的一个非零元。由于该空间是一维的，任何非零元都可以作为基。因此，$E_4(\tau)$ 是 $M_4(\mathbf{SL}(2,\mathbb{Z}))$ 的一组基。

既然 $\theta(\tau)^8$ $E_4(\tau)$ 都属于同一个一维空间，它们之间必然只相差一个常数因子，即 $\theta(\tau)^8 = c \cdot E_4(\tau)$。为了确定常数 c，我们比较它们 q-展开式的常数项：

- $\theta(\tau)^8 = \left(\sum_{n\in\mathbb{Z}} q^{n^2}\right)^8 = \left(1 + 2q + 2q^4 + \ldots\right)^8 = 1 + 16q + \ldots$。其常数项为 $r_8(0) = 1$。
- $E_4(\tau) = 1 + 240\sum_{n=1}^{\infty} \sigma_3(n)q^n = 1 + 240q + \ldots$。其常数项为 1。

比较两者的常数项，我们得到 $1 = c \cdot 1$，因此 $c = 1$。

我们得到了函数恒等式 $\theta(\tau)^8 = E_4(\tau)$。通过逐项比较 $q^n (n \geq 1)$ 的系数，我们直接得到 $r_8(n)$ 的公式：

八平方和公式

$$\sum_{n=0}^{\infty} r_8(n) q^n = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n \tag{14}$$

因此，对于 $n \geq 1$，y 有 $r_8(n) = 240\sigma_3(n)$。(注：E_4 的不同归一化会导致系数不同，例如 $16\sigma_3(n)$ 也是常见形式，这取决于 E_4 的具体定义)。

8 四平方和问题

这个例子将展示如何处理同余子群。

我们已知 $\theta(\tau)^4$ 是 $\Gamma_0(4)$ 上权为 2 的模形式，即 $\theta(\tau)^4 \in M_2(\Gamma_0(4))$。同样利用维数公式，我们知道 $\dim(M_2(\Gamma_0(4))) = 2$。这是一个二维空间。

$M_2(\mathbf{SL}(2,\mathbb{Z}))$ 是零空间，因此我们不能使用 E_2。然而，可以通过对 E_2 进行修改来构造 $M_2(\Gamma_0(4))$ 的基。一组基由以下两个函数给出：

$$f_1(\tau) = E_2(\tau) - 2E_2(2\tau) \tag{15}$$

$$f_2(\tau) = E_2(2\tau) - 2E_2(4\tau) \tag{16}$$

由于 $\theta(\tau)^4$ 属于这个二维空间，它必然可以表示为基的线性组合：$\theta(\tau)^4 = c_1 f_1(\tau) + c_2 f_2(\tau)$。为了求解常数 c_1 c_2，我们计算并比较 q-展开式的前两项系数：

- $\theta(\tau)^4 = \left(1 + 2q + 2q^4 + \dots\right)^4 = 1 + 8q + 24q^2 + \dots$
- $f_1(\tau) = (1 - 24\sum \sigma_1(n) q^n) - 2\left(1 - 24\sum \sigma_1(n) q^{2n}\right) = -1 + 24q - 24q^2 + \dots$
- $f_2(\tau) = \left(1 - 24\sum \sigma_1(n) q^{2n}\right) - 2\left(1 - 24\sum \sigma_1(n) q^{4n}\right) = -1 - 24q^2 - \dots$

比较 q^0 和 q^1 的系数，我们得到 $c_1 = \frac{1}{3}, c_2 = -\frac{4}{3}$。

将常数和 Eisenstein 级数的展开式代回，我们得到：

$$\theta(\tau)^4 = \frac{1}{3}(E_2(\tau) - 2E_2(2\tau)) - \frac{4}{3}(E_2(2\tau) - 2E_2(4\tau)) = \frac{1}{3}E_2(\tau) - 2E_2(2\tau) + \frac{8}{3}E_2(4\tau) \tag{17}$$

将 $E_2(\tau) = 1 - 24\sum_{n=1}^{\infty}\sigma_1(n)q^n$ 代入，并比较 q^n 的系数，经过整理即可得到雅可比的四平方和公式：

四平方和公式

$$r_4(n) = 8\sum_{d|n, 4\nmid d} d \tag{18}$$

9　二十四平方和问题

这个例子揭示了模形式理论更深层次的结构，特别是尖点形式的重要性。

我们已知 $\theta(\tau)^{24} \in M_{12}(\mathbf{SL}(2,\mathbb{Z}))$。该空间的维数为 $\dim(M_{12}(\mathbf{SL}(2,\mathbb{Z}))) = \lfloor 12/12 \rfloor + 1 = 2$ 。这个二维空间可以分解为 Eisenstein 空间和尖点形式空间的直和：

$$M_{12} = \mathbb{C} \cdot E_2 \oplus S_{12} \tag{19}$$

其中 S_{12} 是一维的尖点形式空间。

一维空间 $S_{12}(\mathbf{SL}(2,\mathbb{Z}))$ 由一个非常特殊的函数——拉马努金判别式函数 $\Delta(\tau)$——所张成。它具有优美的无穷乘积表示：

$$\Delta(\tau) = q \prod_{n \geq 1} (1 - q^n)^{24} \tag{20}$$

其傅里叶系数定义了著名的拉马努金 τ-函数：

$$\Delta(\tau) = \sum_{n \geq 1} \tau(n) q^n \tag{21}$$

M_{12} 的一组基是 $\{E_{12}(\tau), \Delta(\tau)\}$ 。因此，$\theta(\tau)^{24}$ 可以表示为它们的线性组合：

$$\theta(\tau)^{24} = c_1 E_{12}(\tau) + c_2 \Delta(\tau) \tag{22}$$

同样，通过比较 q^0 和 q^1 的系数可以解出 c_1 和 c_2。

比较系数最终得到 $r_{24}(n)$ 的表达式：

二十四平方和公式

$$r_{24}(n) = \frac{16}{691} \sigma_{11}(n) + \frac{128}{691} \tau(n) \tag{23}$$

这个公式揭示了深刻的内涵。$r_{24}(n)$ 的表达式并非纯粹的除数和函数，它包含两个部分，这直接反映了模形式空间的结构分解 $M_k = E_k \oplus S_k$。Eisenstein 级数的系数是“简单”的算术函数（如除数和），而尖点形式的系数（如 $\tau(n)$）则更为复杂。$r_{24}(n)$ 的公式表明，函数空间的抽象结构分解在丢番图方程解的计数问题中有着直接而具体的算术对应物。表示数 $r_{24}(n)$ 可以看作是一个平稳增长的“主项”（来自 Eisenstein 级数）和一个剧烈波动的、量级较小的“误差项”（来自尖点形式）之和。这揭示了模曲线的几何性质（它决定了 M_k 和 S_k 的维数）与丢番图方程解的结构之间的深刻联系。