

DAVENPORT · 圆法 · 高中详解版

Waring 问题：奇异级数

Waring's problem: the singular series

本章要解决什么

上一章（圆法）已经把“把 N 写成 s 个 k 次方之和有多少种方法”的个数 $r(N)$ 拆成了一个渐近公式：

$$r(N) \approx \mathfrak{S}(N) \cdot \mathfrak{J}(N) \cdot N^{s/k-1}.$$

其中 $\mathfrak{S}(N)$ 叫**奇异级数** (singular series), $\mathfrak{J}(N)$ 叫**奇异积分**。这个公式只有在“前面那个系数 $\mathfrak{S}(N)$ 不是 0”的时候才真正有用——否则主项消失，公式什么都没说。

本章唯一的目标：证明只要 s 足够大，就有一个与 N 无关的正常数 $C_1 > 0$ ，使得

$$\mathfrak{S}(N) \geq C_1 > 0 \quad (\text{对一切 } N).$$

读完你将掌握：(1) 奇异级数到底是什么、为什么这样定义；(2) 它怎么**乘性地**拆成每个素数 p 的“局部因子” $\chi(p)$ (欧拉乘积)；(3) 每个 $\chi(p)$ 其实就是“同余方程解数的极限密度”；(4) 怎么用**原根**、**循环群**、**提升解**这些工具证明每个 $\chi(p) > 0$ ，并把它们乘起来得到 $\mathfrak{S}(N) \geq C_1 > 0$ 。

本页是**逐段精读**：黑色引用块是 Davenport 原书译文的句子，彩色框（前置知识 / 符号 / 分步推演 / 配图）是面向高中生的详解。凡原文写“显然”“容易看出”的地方，这里都拆成一步一步补全。

第 0 节 读这一章之前必须先补的“工具”

这一章用到一批超出高中课本的记号。它们其实都不难，但你必须先认得它们，后面才看得懂。下面我把每一个都从零讲清楚。等会儿正文里再次出现时，你回头看这一节即可。

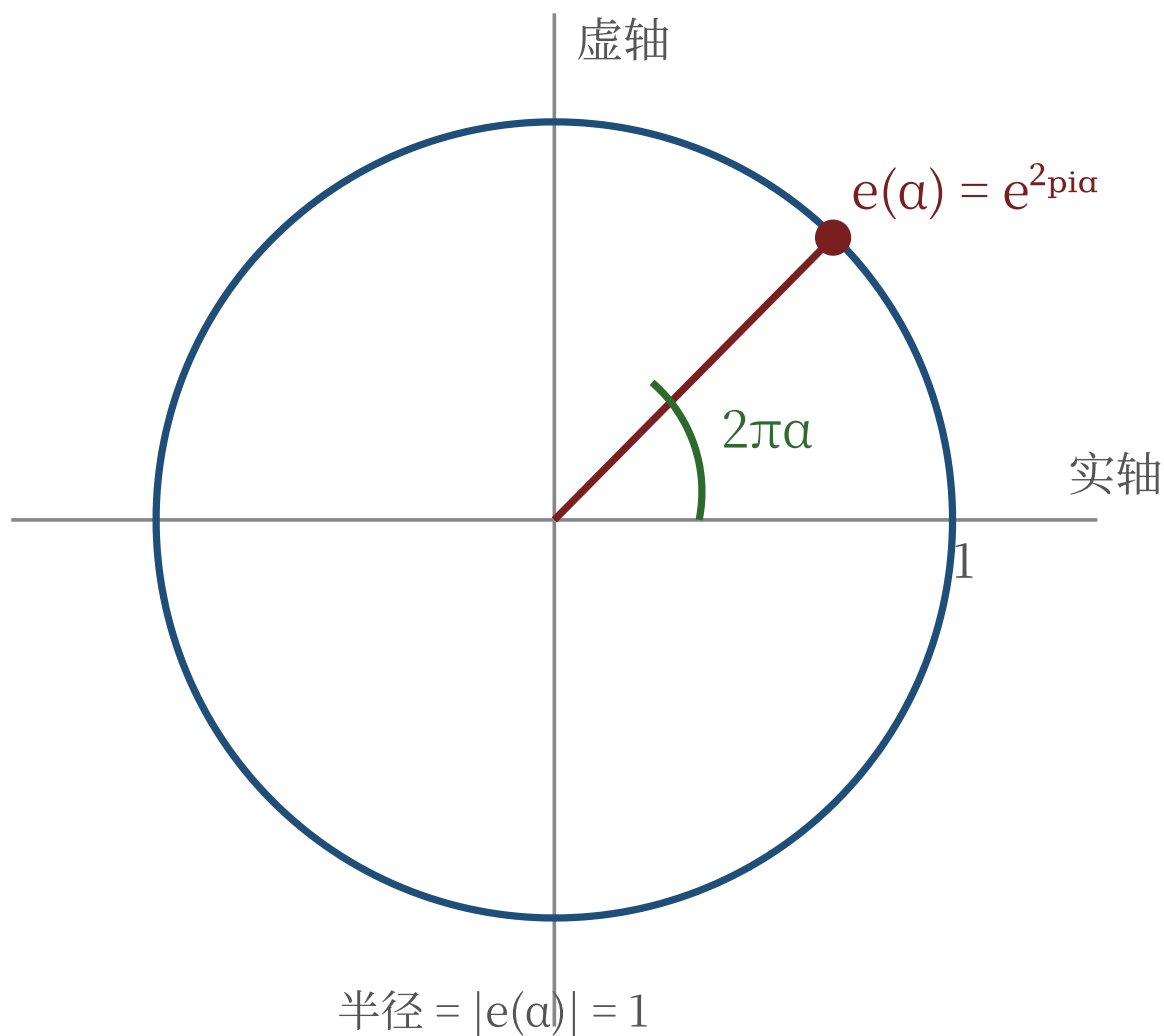
符号 $e(\alpha)$ ：复指数

记号 $e(\alpha)$ 是下面这个东西的简写：

$$e(\alpha) := e^{2\pi i \alpha} = \cos(2\pi \alpha) + i \sin(2\pi \alpha).$$

读作“e 二派 i alpha”。这里 i 是虚数单位 ($i^2 = -1$)， $\pi \approx 3.14159$ 。为什么要造这个记号？因为后面满纸都是 $e^{2\pi i(\cdots)}$ ，每次写指数太累，于是把“ $2\pi i \times$ ”这一坨打包进字母 $e()$ 里。它有三条你必须背下来的性质：

- 1 模长恒为 1。由欧拉公式 $e^{i\theta} = \cos \theta + i \sin \theta$ ，其模长 $|e^{i\theta}| = \sqrt{\cos^2 \theta + \sin^2 \theta} = 1$ 。所以 $|e(\alpha)| = 1$ 永远成立—— $e(\alpha)$ 始终是复平面单位圆上的一个点。
- 2 周期为 1。 $e(\alpha + 1) = e^{2\pi i \alpha + 2\pi i} = e^{2\pi i \alpha} \cdot e^{2\pi i} = e(\alpha) \cdot 1 = e(\alpha)$ （因为绕一整圈 2π 回到原处）。所以 $e(\alpha)$ 只看 α 的小数部分，即只依赖 $\alpha \bmod 1$ 。这正是后面“ $(\bmod 1)$ ”记号的来历。
- 3 指数相加 = 底数相乘。 $e(\alpha + \beta) = e^{2\pi i(\alpha + \beta)} = e^{2\pi i \alpha} e^{2\pi i \beta} = e(\alpha) e(\beta)$ 。



$e(\alpha)$ 就是单位圆上、从正实轴起逆时针转过角度 $2\pi\alpha$ 的那个点。 α 每增加 1，正好转满一圈回到原处——这就是“周期 1”。

符号 $\sum$ (求和) 与 $\prod$ (连乘)

$\sum_{x=1}^q f(x)$ 读作“对 x 从 1 到 q 求和”，意思是 $f(1) + f(2) + \cdots + f(q)$ 。 $\prod_p g(p)$

读作“对所有 p 连乘”，意思是把每个 $g(p)$ 乘起来 $g(p_1)g(p_2)\cdots$ 。求和号下面写条件（如 $(a, q) = 1$ ）表示只把**满足条件**的项加进来。

符号 $a \equiv b \pmod{q}$ (同余) 与 (a, q) (最大公因数)

“ $a \equiv b \pmod{q}$ ”读作“ a 与 b 模 q 同余”，意思是 $q \mid (a - b)$ （ q 整除 $a - b$ ），也就是 a, b 除以 q 余数相同。例如 $17 \equiv 2 \pmod{5}$ 。

“ (a, q) ”是 a 与 q 的最大公因数（gcd）。 $(a, q) = 1$ 表示 a, q 互素（没有大于1的公因数）。在 $1, 2, \dots, q$ 里和 q 互素的数，叫 q 的既约剩余系，它们的个数记作 $\varphi(q)$ （见下）。

符号 $\varphi(q)$ ：欧拉函数

$\varphi(q)$ （欧拉函数，phi）= $1 \sim q$ 中与 q 互素的整数个数。关键公式：对素数幂 p^ν ,

$$\varphi(p^\nu) = p^\nu - p^{\nu-1} = p^{\nu-1}(p - 1).$$

道理： $1 \sim p^\nu$ 里只有 p 的倍数（共 $p^{\nu-1}$ 个）才和 p^ν 不互素，把它们去掉就剩 $p^\nu - p^{\nu-1}$ 个。

符号 $O(\cdot)$ 与 $\ll$ ：估大小的语言

这两个记号是“不计较常数倍、只看量级”的速记。

- $f(x) \ll g(x)$ （也写成 $f = O(g)$ ）读作“ f 被 g 控制”，意思是：存在一个与 x 无关的正常数 C ，使得 $|f(x)| \leq C g(x)$ 对所有讨论范围内的 x 成立。
- 例如 $3x + 5 \ll x$ （取 $C = 8$ 即可，对 $x \geq 1$ ）。 $q^{-1-\delta} \ll q^{-1-\delta}$ 自然成立。

它的好处是：我们经常不需要算出精确常数，只要知道“某量比 $q^{-1-\delta}$ 小”，就足以判断一个级数收敛。

符号 $S_{a,q}$ ：完全指数和（上一章的主角）

这一章直接拿来用、但定义在上一章的核心对象是

$$S_{a,q} = \sum_{x=1}^q e\left(\frac{ax^k}{q}\right) = \sum_{x=1}^q e^{2\pi i ax^k/q}.$$

它把 $x = 1, 2, \dots, q$ 的 k 次方 x^k , 各自变成单位圆上一个点 $e(ax^k/q)$, 再把这 q 个点 (看成复数向量) 加起来。它度量了 “ $x^k \bmod q$ 分布得有多均匀”。上一章证明了一个关键的上界估计 (Weyl 估计的推论):

$$|S_{a,q}| \ll q^{1-1/K}, \quad K = 2^{k-1}.$$

直观上: 如果 q 个单位向量是 “随机方向”, 和的长度约为 $\sqrt{q}$; 这里得到的 $q^{1-1/K}$ 比 q 小, 说明确实有相当程度的相消。这条估计是本章一切收敛性的根基, 请记住它。

符号 $r(N)$: 表示个数

$r(N)$ 表示把 N 写成 $N = x_1^k + \dots + x_s^k$ (各 x_j 为正整数) 的解的组数。这是 Waring 问题真正关心的量。本章的全部努力, 是为了保证 $r(N)$ 的渐近主项不为零。

关键事实 正交关系 (求和挑出同余)

下面这条恒等式后面要用两次, 先证清楚。对任意整数 n :

$$\sum_{t=1}^q e\left(\frac{tn}{q}\right) = \begin{cases} q, & \text{若 } q \mid n, \\ 0, & \text{若 } q \nmid n. \end{cases}$$

1 若 $q \mid n$: 则 n/q 是整数, $e(tn/q) = e(\text{整数}) = 1$, 于是求和
 $= \underbrace{1 + 1 + \dots + 1}_{q \uparrow} = q$ 。

2 若 $q \nmid n$: 记 $\omega = e(n/q)$ 。则 $\omega \neq 1$ (因为 n/q 不是整数), 且各项是等比数列 $\omega^1, \omega^2, \dots, \omega^q$ 。用等比数列求和 $\sum_{t=1}^q \omega^t = \omega \frac{\omega^q - 1}{\omega - 1}$ 。而 $\omega^q = e(n) = 1$ (n 整数), 分子 $\omega^q - 1 = 0$, 故整个和 $= 0$ 。

这条恒等式是一台“检测器”：只有当 $q \mid n$ （即某同余式成立）时它才给出 q ，否则归零。后面用它把“同余方程的解数”写成指数和。

第 1 节 奇异级数的定义（式 5.1）

我们现在来研究奇异级数（singular series）：

$$\mathfrak{S}(N) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^q (q^{-1} S_{a,q})^s e(-aN/q). \quad (5.1)$$

这一段在讲什么：给出本章主角 $\mathfrak{S}(N)$ （花体 S，读作“奇异级数”）的定义式。它是一个二重求和：外层对模数 $q = 1, 2, 3, \dots$ 一直加到无穷，内层对 $1 \sim q$ 中与 q 互素的 a 求和。

逐块拆解式 (5.1)

$\mathfrak{S}(N)$

花体 S，奇异级数。它是一个只依赖 N （以及固定的 k, s ）的数。

$$\sum_{q=1}^{\infty}$$

外层求和：对每个正整数模数 q 各取一项 $A(q)$ （见下一节），全部加起来。这是“无穷级数”，所以叫“级数”。

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q$$

内层求和： a 跑遍 $1 \sim q$ 中和 q 互素的那些数（共 $\varphi(q)$ 个）。分数 a/q 就是上一章“主弧”的中心位置——奇异级数正是把所有这些有理点 a/q 的贡献收集起来。

$$q^{-1} S_{a,q}$$

就是 $\frac{S_{a,q}}{q}$ ：把完全指数和除以 q 做“归一化”。由 $|S_{a,q}| \ll q^{1-1/K}$ 得 $|q^{-1}S_{a,q}| \ll q^{-1/K}$ ，是个小于 1 的小量。

$(\cdot)^s$

整体 s 次方。这里的 s 就是 Waring 问题里 k 次方的个数 ($N = x_1^k + \cdots + x_s^k$ 里的 s)。

$e(-aN/q)$

第 0 节的复指数，参数是 $-aN/q$ 。这个因子带着我们要表示的目标数 N 。

为什么会冒出这么个怪东西？（动机）

别把 (5.1) 当成天上掉下来的定义。它是上一章圆法计算的**残留物**。圆法把 $r(N)$ 写成 $[0, 1]$ 上一个积分 $\int_0^1 T(\alpha)^s e(-\alpha N) d\alpha$ ，其中 $T(\alpha) = \sum_x e(\alpha x^k)$ 。这个积分被切成“主弧”（ α 靠近某个有理点 a/q ）和“次弧”两部分。把所有主弧上的贡献加总、整理后，恰好分离出两个因子：一个连续的积分 $\mathfrak{J}(N)$ （奇异积分），和一个离散的、对所有 a/q 求和的级数——那就是 (5.1) 的 $\mathfrak{S}(N)$ 。所以 $\mathfrak{S}(N)$ 是“主弧贡献里与算术（同余）有关的那一半”。它的名字“奇异”是历史叫法（singular = 来自奇点/主弧附近的主要贡献），不必纠结字面。

第 2 节 奇异级数与“同余解数”的联系

我们将会发现， $\mathfrak{S}(N)$ 的值与下列同余方程

$$x_1^k + \cdots + x_s^k \equiv N \pmod{q}$$

对一切正整数 q 的解数密切相关；并且确实地，若任何这样的同余方程无解，则 $\mathfrak{S}(N) = 0$ 。这一点从渐近公式的形式上是可以预料到的，因为此时 $r(N) = 0$ 。

这一段在讲什么：预告本章的中心思想——奇异级数 $\mathfrak{S}(N)$ 其实在度量“同余方程 $x_1^k + \cdots + x_s^k \equiv N \pmod{q}$ 好不好解”。

1 什么叫同余方程有解？就是能找到整数 $x_1, \dots, x_s$ ，使得 $x_1^k + \cdots + x_s^k$ 除以 q 的余数恰好等于 N 除以 q 的余数。这是“在 $\text{mod } q$ 的世界里”能不能凑出 N 。

2 为什么二者相关？直觉很简单：要想真的把 N 写成 s 个 k 次方之和（即 $r(N) > 0$ ），那么必要条件是：在每个模 q 下都至少凑得出来。如果存在某个 q ，连 $\text{mod } q$ 都凑不出 N ，那真正的等式当然更不可能成立，于是 $r(N) = 0$ 。

3 这与 $\mathfrak{S}(N)$ 有什么关系？既然渐近公式是 $r(N) \approx \mathfrak{S}(N) \mathfrak{J}(N) N^{s/k-1}$ ，而上面分析出“某个 q 无解 $\Rightarrow r(N) = 0$ ”，那么公式要自洽，就必须在这种情形下让 $\mathfrak{S}(N) = 0$ 来把主项压成 0。本章后面（引理 5.3）会精确证明： $\mathfrak{S}(N)$ 恰好等于各模 q 解数的某种“密度极限”的乘积，于是“某个 q 无解”确实导致 $\mathfrak{S}(N) = 0$ 。

4 本章真正要证的是反方向：只要 s 足够大，每个同余方程都有解，而且解多到使 $\mathfrak{S}(N)$ 有正的下界。这就保证主项不消失。

一个具体感受

取 $k = 2$ （平方和）、 $s = 3$ 、 $q = 8$ 。问： $x_1^2 + x_2^2 + x_3^2 \equiv 7 \pmod{8}$ 有解吗？平方数 $\text{mod } 8$ 只能是 0, 1, 4。三个这种数相加， $\text{mod } 8$ 怎么都凑不出 7（可以逐一验证）。所以 $N \equiv 7 \pmod{8}$ 的数不能写成 3 个平方和——这正是著名的“三平方和定理”里被排除的情形。此时 $\mathfrak{S}(N) = 0$ ，渐近公式的主项归零，与事实吻合。这就是“某个 q 无解 $\Rightarrow$ 表示不出来”的活例子。

第3节 把级数按 q 分项： $A(q)$ (式 5.2)

我们记

$$\mathfrak{S}(N) = \sum_{q=1}^{\infty} A(q), \quad A(q) = \sum_{\substack{a=1 \\ (a,q)=1}}^q (q^{-1}S_{a,q})^s e(-aN/q). \quad (5.2)$$

这一段在讲什么：给二重求和的内层起个名字 $A(q)$ ，于是奇异级数就写成最简单的形式“ $\mathfrak{S}(N) = \sum_q A(q)$ ”。这只是记号上的打包，没有任何新内容，但极其有用：接下来所有引理都是在研究单个 $A(q)$ 的性质。

符号 $A(q)$

$A(q)$ = 奇异级数中“模数恰为 q ”的那一整块贡献。它本身是个复数，但我们会发现实际算出来都是实数。本章策略：先研究 $A(q)$ 的乘性（引理 5.1），再把级数 $\sum_q A(q)$ 拆成对每个素数 p 的乘积（引理 5.2）。

第4节 引理 5.1： $A(q)$ 是“乘性”的

引理 5.1. 若 $(q_1, q_2) = 1$ ，则 $A(q_1 q_2) = A(q_1)A(q_2)$ 。

这条引理在讲什么、为什么重要：它说，当 q_1, q_2 互素时， A 在乘积 $q_1 q_2$ 上的值，等于在 q_1 与 q_2 上的值相乘。具备这种性质的函数叫乘性函数（multiplicative function）。这是本章的引擎：有了乘性，任意 q 都能按素因数分解 $q = p_1^{\nu_1} p_2^{\nu_2} \cdots$ 把 $A(q)$ 拆成各素数幂上的因子，从而把奇异级数变成“对每个素数各管一段”的欧拉乘积。一个全局的难题就此被切成无数个独立的、只关乎单个素数 p 的小问题。

4.1 证明的总体思路

记 $f(a, q) = (S_{a,q})^s e(-aN/q)$. 我们将证明: 若 $(a_1, q_1) = (a_2, q_2) = 1$ 且

$$\frac{a}{q} \equiv \frac{a_1}{q_1} + \frac{a_2}{q_2} \pmod{1}, \quad q = q_1 q_2, \quad (5.4)$$

则 $f(a, q) = f(a_1, q_1) f(a_2, q_2)$.

这一段在讲什么: 先把 $A(q) = \sum q^{-s} f(a, q)$ 里那一坨“与 a 有关的核心”单独命名为 $f(a, q) = (S_{a,q})^s e(-aN/q)$ (注意: 这里没有除以 q^s , 那个 q^{-s} 因子在 $A(q_1 q_2) = q^{-s} \sum f$ 里会自动配平, 因为 $q^{-s} = q_1^{-s} q_2^{-s}$). 然后宣布证明分两步走:

1 第一步 (这一小段): 建立一一对应 (5.4), 并证明对应项满足 (5.5).

2 第二步 (下一小段): 真正动手算 $f(a, q) = f(a_1, q_1) f(a_2, q_2)$.

读懂式 (5.4): 中国剩余定理的“坐标变换”

(5.4) 说: 当 q_1, q_2 互素、 $q = q_1 q_2$ 时, 把分数 $\frac{a}{q}$ 拆成 $\frac{a_1}{q_1} + \frac{a_2}{q_2}$ (在 $\text{mod } 1$ 意义下, 即只看小数部分)。这其实是**中国剩余定理 (CRT)** 的分数版。CRT 说: q_1, q_2 互素时, 一个 $\text{mod } q_1 q_2$ 的数 a , 与一对 $(\text{mod } q_1, \text{mod } q_2)$ 的数 (a_1, a_2) 是一一对应的。这里把这条对应写成了“通分”的样子。

这就足以给出所要的结果, 因为关系式 (5.4) 在既约剩余类 $a \pmod{q}$ 与既约剩余类对 $a_1 \pmod{q_1}$ 、 $a_2 \pmod{q_2}$ 之间建立了一一对应, 从而

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q f(a, q) = \left(\sum_{\substack{a_1=1 \\ (a_1,q_1)=1}}^{q_1} f(a_1, q_1) \right) \left(\sum_{\substack{a_2=1 \\ (a_2,q_2)=1}}^{q_2} f(a_2, q_2) \right).$$

这一段在讲什么、把跳步补全: 原文说“这就足以给出结果”, 跳过了为什么。我们补全:

1 对应是双射，且保持“互素”。CRT 给出 $a \leftrightarrow (a_1, a_2)$ 的一一对应。还要确认： a 与 $q = q_1 q_2$ 互素 $\iff a_1$ 与 q_1 互素且 a_2 与 q_2 互素。这成立，因为 $a \bmod q_1 \equiv a_1$ （由 (5.4) 通分知 $a \equiv a_1 q_2 \cdot (q_2^{-1}) \dots$ ，简单说 a 模 q_1 就是 a_1 那一类），所以 $(a, q_1) = (a_1, q_1)$ ，类似 $(a, q_2) = (a_2, q_2)$ ；而 $(a, q_1 q_2) = 1$ 当且仅当 a 同时与 q_1 、 q_2 互素。于是“ a 跑遍 q 的既约剩余系”正好对应“ (a_1, a_2) 跑遍 q_1, q_2 的既约剩余系的所有配对”。

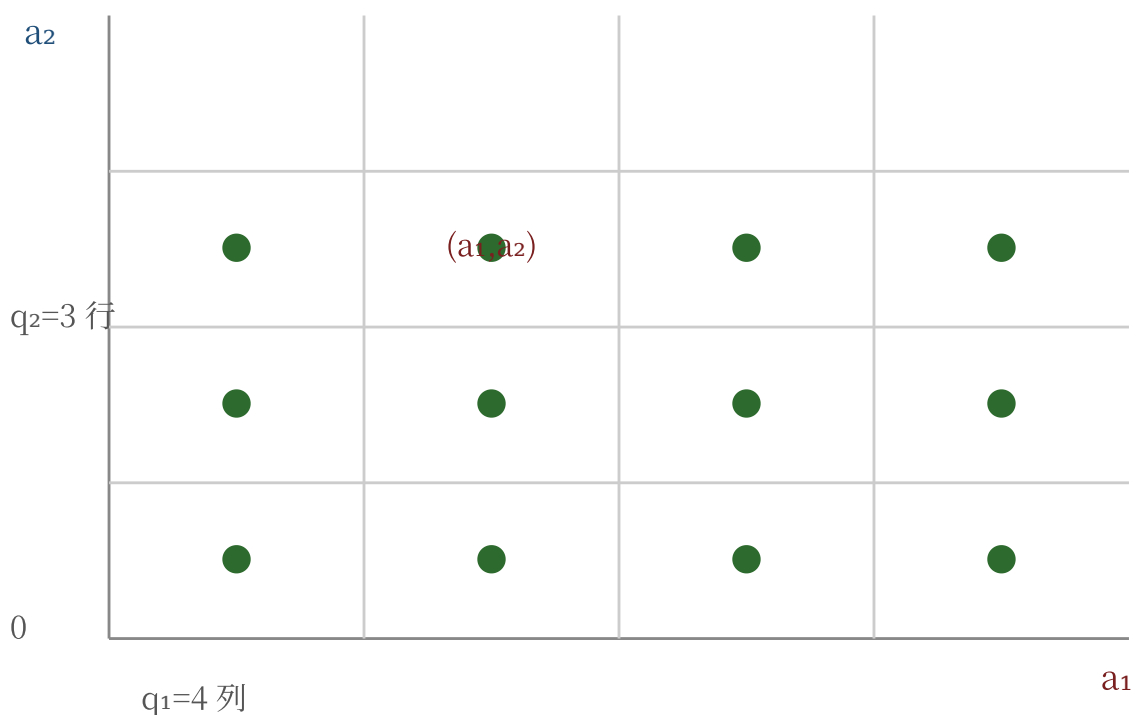
2 把求和拆成乘积。既然 $f(a, q) = f(a_1, q_1) f(a_2, q_2)$ （这是第二步要证的 (5.5)），那么

$$\sum_a f(a, q) = \sum_{a_1} \sum_{a_2} f(a_1, q_1) f(a_2, q_2) = \left(\sum_{a_1} f(a_1, q_1) \right) \left(\sum_{a_2} f(a_2, q_2) \right)$$

最后一步是“和的乘积展开”的逆用：把不含 a_2 的因子提到内层求和外即可。

3 配回 q^{-s} 。 $A(q) = q^{-s} \sum_a f(a, q)$ 。由 $q^{-s} = q_1^{-s} q_2^{-s}$ ，上式两边各乘相应因子，立即得到 $A(q_1 q_2) = A(q_1) A(q_2)$ ，即 (5.3)。

CRT 一一对应: $a \pmod{q_1 q_2} \leftrightarrow (a_1 \pmod{q_1}, a_2 \pmod{q_2})$



当 $q_1 = 4, q_2 = 3$ (互素) 时, $\pmod{12}$ 的每个 a 恰对应格点上唯一一个 (a_1, a_2) , 反之亦然。求和 $\sum_a$ 就等于先沿行、再沿列地把所有格点加一遍, 于是分裂成两个独立求和的乘积。

4.2 证明 (5.5): 用完全剩余系直接算

为了证明 (5.5), 我们使用一个相当类似的论证, 但这次用的是完全剩余系。

令 $\frac{z}{q} \equiv \frac{z_1}{q_1} + \frac{z_2}{q_2} \pmod{1}$, 则有

$$S_{a,q} = \sum_{z=1}^q e(az^k/q) = \sum_{z_1=1}^{q_1} \sum_{z_2=1}^{q_2} e\left(\frac{a}{q} q^k \left(\frac{z_1}{q_1} + \frac{z_2}{q_2}\right)^k\right).$$

这一段在讲什么: 现在要证关键的 (5.5)。核心是先把 $S_{a,q}$ 也按 CRT “拆坐标”。注意这里用的是**完全剩余系** ($z = 1, \dots, q$ 全取, 不要求互素), 因为 $S_{a,q}$ 的定义就是对所有 z 求和。

- 1 换元。** 因为 q_1, q_2 互素, CRT 告诉我们: 当 z_1 跑 $1 \sim q_1$ 、 z_2 跑 $1 \sim q_2$ 时, 由 $\frac{z}{q} \equiv \frac{z_1}{q_1} + \frac{z_2}{q_2} \pmod{1}$ 定出的 z 恰好跑遍 $1 \sim q$ 的一个完全剩余系。所以把对 z 的单重和换成对 (z_1, z_2) 的双重和, 每一项不重不漏。

- 2 代入指数。由 $\frac{z}{q} \equiv \frac{z_1}{q_1} + \frac{z_2}{q_2}$, 两边乘 q : $z \equiv q\left(\frac{z_1}{q_1} + \frac{z_2}{q_2}\right) \pmod{q}$ 。
 于是 $z^k \equiv q^k\left(\frac{z_1}{q_1} + \frac{z_2}{q_2}\right)^k$, 代进 $e(az^k/q)$ 得到右边那一长串。($e(\cdot)$ 只看 $\pmod{1}$, 所以把 z 换成同余的表达式不改变值。)

并且由于

$$\frac{a}{q} q^k \left(\frac{z_1}{q_1} + \frac{z_2}{q_2} \right)^k \equiv \frac{a_1}{q_1} (q_2 z_1)^k + \frac{a_2}{q_2} (q_1 z_2)^k \pmod{1},$$

这一段是全证明最“跳”的一步，原文一句带过，我们彻底补全。目标：证明上面这个 $\pmod{1}$ 同余式。

- 1 先通分。把括号里两项通分 ($q = q_1 q_2$):

$$\frac{z_1}{q_1} + \frac{z_2}{q_2} = \frac{q_2 z_1 + q_1 z_2}{q_1 q_2} = \frac{q_2 z_1 + q_1 z_2}{q}.$$

$$\text{于是 } \frac{a}{q} q^k \left(\frac{z_1}{q_1} + \frac{z_2}{q_2} \right)^k = \frac{a}{q} q^k \cdot \frac{(q_2 z_1 + q_1 z_2)^k}{q^k} = \frac{a}{q} (q_2 z_1 + q_1 z_2)^k.$$

记 $w = q_2 z_1 + q_1 z_2$, 要算 $\frac{a}{q} w^k \pmod{1}$ 。

- 2 用二项式定理展开 w^k 。

$$w^k = (q_2 z_1 + q_1 z_2)^k = \sum_{j=0}^k \binom{k}{j} (q_2 z_1)^{k-j} (q_1 z_2)^j.$$

- 3 看哪些项在 $\pmod{1}$ 下还活着。把 $\frac{a}{q} = \frac{a}{q_1 q_2}$ 乘进去。除了首项 ($j = 0$, 只含 q_2) 和末项 ($j = k$, 只含 q_1) 之外，**中间每一项**都同时含有 q_1 的因子和 q_2 的因子 (j 和 $k - j$ 都 ≥ 1)，即被 $q_1 q_2 = q$ 整除。中间项形如 $\frac{a}{q} \cdot (q \cdot \text{整数}) = a \cdot \text{整数}$ ，是整数，在 $\pmod{1}$ 下等于 0，可以丢掉。

- 4 只剩首末两项。

$$\frac{a}{q} w^k \equiv \frac{a}{q} (q_2 z_1)^k + \frac{a}{q} (q_1 z_2)^k \pmod{1}.$$

- 5 把 a/q 换成 $a_1/q_1, a_2/q_2$ 。由 (5.4), $\frac{a}{q} \equiv \frac{a_1}{q_1} + \frac{a_2}{q_2} \pmod{1}$ 。对第一项 $\frac{a}{q} (q_2 z_1)^k$: 替换后变成 $\left(\frac{a_1}{q_1} + \frac{a_2}{q_2}\right) (q_2 z_1)^k$ 。其中 $\frac{a_2}{q_2} (q_2 z_1)^k = a_2 q_2^{k-1} z_1^k$ 是整数 ($k \geq 1$ 故 q_2^{k-1} 是整数倍), $\pmod{1}$ 为 0, 丢掉, 只剩 $\frac{a_1}{q_1} (q_2 z_1)^k$ 。对第二项同理, 只剩 $\frac{a_2}{q_2} (q_1 z_2)^k$ 。

- 6 合并, 得证。

$$\frac{a}{q} q^k \left(\frac{z_1}{q_1} + \frac{z_2}{q_2} \right)^k \equiv \frac{a_1}{q_1} (q_2 z_1)^k + \frac{a_2}{q_2} (q_1 z_2)^k \pmod{1}.$$

我们得到

$$S_{a,q} = \sum_{z_1=1}^{q_1} e\left(\frac{a_1}{q_1} (q_2 z_1)^k\right) \sum_{z_2=1}^{q_2} e\left(\frac{a_2}{q_2} (q_1 z_2)^k\right) = S_{a_1,q_1} S_{a_2,q_2}.$$

补全这一步:

- 1 指数变乘积。由第 0 节性质 “ $e(\alpha + \beta) = e(\alpha)e(\beta)$ ”, 把上面那个和拆成两个因子:

$$e\left(\frac{a_1}{q_1} (q_2 z_1)^k + \frac{a_2}{q_2} (q_1 z_2)^k\right) = e\left(\frac{a_1}{q_1} (q_2 z_1)^k\right) e\left(\frac{a_2}{q_2} (q_1 z_2)^k\right).$$

- 2 双重和分裂。第一个因子只含 z_1 , 第二个只含 z_2 , 于是 $\sum_{z_1} \sum_{z_2}$ 分裂成两个独立求和的乘积 (同 4.1 的拆法)。

- 3 认出这就是 $S_{a_1,q_1} S_{a_2,q_2}$ 。需要说明 $\sum_{z_1=1}^{q_1} e\left(\frac{a_1}{q_1} (q_2 z_1)^k\right) = S_{a_1,q_1}$ 。理由: 当 z_1 跑遍 $\pmod{q_1}$ 的完全剩余系, 而 $(q_2, q_1) = 1$, 所以 $q_2 z_1$ 也跑遍

$\bmod q_1$ 的完全剩余系（乘一个互素的数是一个“洗牌”，不重不漏）。把求和变量从 z_1 改名为 $y = q_2 z_1$ ，就得到 $\sum_{y=1}^{q_1} e\left(\frac{a_1}{q_1} y^k\right) = S_{a_1, q_1}$ 。第二个因子同理 $= S_{a_2, q_2}$ 。

此外，由于 $e\left(-\frac{a}{q} N\right) = e\left(-\frac{a_1}{q_1} N\right) e\left(-\frac{a_2}{q_2} N\right)$ ，我们便得到 (5.5)。

1 这一步用 (5.4)： $\frac{a}{q} \equiv \frac{a_1}{q_1} + \frac{a_2}{q_2} \pmod{1}$ ，两边乘 $-N$ 再套 $e(\cdot)$ ，由“指数相加=底数相乘”立得分裂。

2 合成 (5.5)。把刚得到的 $S_{a, q} = S_{a_1, q_1} S_{a_2, q_2}$ （取 s 次方）与 $e(-aN/q) = e(-a_1 N/q_1) e(-a_2 N/q_2)$ 相乘：

$$f(a, q) = (S_{a, q})^s e(-aN/q) = (S_{a_1, q_1})^s (S_{a_2, q_2})^s e(-a_1 N/q_1) e(-a_2 N/q_2)$$

这正是 (5.5)，引理 5.1 证毕。■

注（原文的“注”）。

这段注在讲什么：提醒我们这个乘性结论跟 z^k 这个特殊形式毫无关系。证明里真正用到的只有“CRT 拆坐标 + 二项式中间项被 q 整除”。把 z^k 换成任意整系数多项式 $f(z)$ ，甚至换成多元指数和 $\sum e\left(\frac{a}{q} f(z_1, \dots, z_n)\right)$ （每个变量各跑一组完全剩余系），证明一字不差地照搬。这说明乘性是“CRT + 指数和”这套机制的普遍特征，不是 Waring 问题专属。记住这一点，将来处理别的丢番图方程时同样适用。

第 5 节 引理 5.2：欧拉乘积与收敛（式 5.6–5.8）

引理 5.2. 若 $s \geq 2^k + 1$, 则 $\mathfrak{S}(N) = \prod_p \chi(p)$, (5.6) 其中

$$\chi(p) = 1 + \sum_{\nu=1}^{\infty} A(p^\nu). \quad (5.7) \text{ 又有 } |\chi(p) - 1| \ll p^{-1-\delta} \quad (5.8) \text{ 对某个固}$$

定的 $\delta > 0$ 成立。

这条引理在讲什么、为什么是关键转折：它把“对所有 q 求和”的奇异级数，重写成“对所有素数 p 连乘”的欧拉乘积。每个因子 $\chi(p)$ （读作“chi of p”）只跟单个素数 p 有关，叫做局部因子（local factor）。这样一来，证明 $\mathfrak{S}(N) > 0$ 的全局任务，就化简为“逐个证明每个 $\chi(p) > 0$ ，并保证乘积收敛、不退化为 0”的局部任务。(5.8) 则保证“大素数的因子 $\chi(p)$ 几乎等于 1”，从而无穷乘积乖乖收敛。

符号 $\chi(p)$ 与“欧拉乘积”

$\chi(p)$

素数 p 的局部因子，定义见 (5.7)：把所有以 p 的方幂 p^ν ($\nu = 0, 1, 2, \dots$) 为模数的项 $A(p^\nu)$ 加起来 ($A(p^0) = A(1) = 1$ ，这就是 (5.7) 里那个孤零零的“1+”)。

$\prod_p$

对所有素数 $2, 3, 5, 7, 11, \dots$ 连乘。把求和写成这种连乘，是数论里反复出现的招式，因为整数的素因数分解唯一，所以“对所有 q 求和”天然能整理成“对每个素数独立处理再相乘”。

5.1 为什么级数能变成乘积（证明 5.6）

由引理 5.1 可知，若 $q = p_1^{\nu_1} p_2^{\nu_2} \cdots$ ，则 $A(q) = A(p_1^{\nu_1}) A(p_2^{\nu_2}) \cdots$ 。于是

$$\mathfrak{S}(N) = \sum_{q=1}^{\infty} A(q) = \prod_p \left\{ \sum_{\nu=0}^{\infty} A(p^\nu) \right\} = \prod_p \chi(p),$$

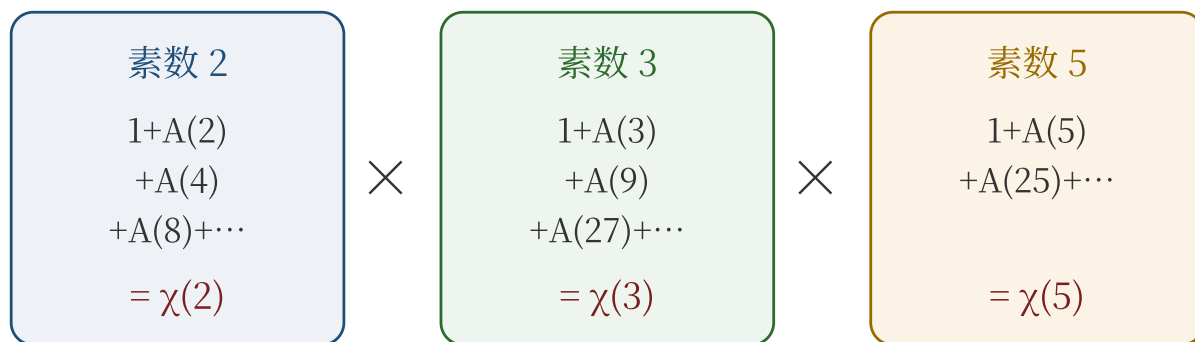
这一点由 $\sum |A(q)|$ 的收敛性所保证，而后者已在前一章中证明。

1 把乘性用满。任意 q 都唯一分解成 $q = p_1^{\nu_1} p_2^{\nu_2} \cdots$ ，各素数幂两两互素。反复用引理 5.1（互素时 A 相乘），得 $A(q) = A(p_1^{\nu_1}) A(p_2^{\nu_2}) \cdots$ 。

2 展开乘积 = 收集所有 q 。看右边 $\prod_p \left(\sum_{\nu \geq 0} A(p^\nu) \right)$ 。把这个无穷乘积“硬乘开”：从每个素数 p 的括号里各挑一项 $A(p^{\nu_p})$ （绝大多数素数挑 $\nu_p = 0$ 即 $A(1) = 1$ ），相乘得到 $\prod_p A(p^{\nu_p}) = A(q)$ ，其中 $q = \prod_p p^{\nu_p}$ 。由于素因数分解唯一，每个 q 恰好被对应一次。所以展开后正好是 $\sum_q A(q) = \mathfrak{S}(N)$ 。这就是“分配律的无穷版”。

3 为什么允许这样无限展开、重排？无穷级数/乘积的展开、重排不是无条件成立的，必须绝对收敛才行。这里的合法性来自上一章证明的 $\sum_q |A(q)| < \infty$ （绝对收敛）。绝对收敛保证“无论怎么加、怎么乘开、怎么重排，结果都一样”，于是上面的等式链成立。

奇异级数 = 各素数局部因子的连乘



$$\mathfrak{S}(N) = \chi(2) \cdot \chi(3) \cdot \chi(5) \cdot \chi(7) \cdots$$

乘性 (引理 5.1) 把无穷求和拆成“每个素数管一个括号”的无穷乘积。证明 $\mathfrak{S}(N) > 0$ 于是变成逐个证明 $\chi(p) > 0$ 。

5.2 大素数的因子几乎是 1（证明 5.8）

我们已经有了估计 $|A(q)| \ll q^{-1-1/K+\varepsilon} \ll q^{-1-\delta}$, 这蕴含 $|\chi(p) - 1| \ll \sum_{\nu=1}^{\infty} p^{-\nu(1+\delta)} \ll p^{-1-\delta}$, 这就是 (5.8)。

这一段在讲什么、把估计的来历讲清：要让无穷乘积 $\prod_p \chi(p)$ 收敛、且不会“无端塌成 0”，必须保证 $\chi(p)$ 与 1 的差随 p 增大而**快速**变小。(5.8) 正是此意。

那个 $q^{-1-1/K+\varepsilon}$ 是怎么来的（为什么要 $s \geq 2^k + 1$ ）

1 由第 0 节, $|S_{a,q}| \ll q^{1-1/K}$, $K = 2^{k-1}$ 。所以 $|q^{-1}S_{a,q}| \ll q^{-1/K}$, 取 s 次方 $|(q^{-1}S_{a,q})^s| \ll q^{-s/K}$ 。

2 $A(q)$ 是 $\varphi(q) \leq q$ 项这种东西之和, 每项还乘了一个模长为 1 的 $e(-aN/q)$ (不改变大小)。粗暴地用三角不等式:
 $|A(q)| \leq \sum_a |(q^{-1}S_{a,q})^s| \ll q \cdot q^{-s/K} = q^{1-s/K}$ 。

3 要它 $\ll q^{-1-\delta}$, 需要指数 $1 - s/K \leq -1 - \delta$, 即
 $s \geq 2K + \delta K = 2^k + (\text{一点})$ 。取 $s \geq 2^k + 1$ 就够。这正是引理 5.2 假设 $s \geq 2^k + 1$ 的原因——它恰好把每一项压到“ q^{-1} 还要再小一点”的水平, 从而保证收敛。(更精细的 Weyl 估计能把粗暴的“ q 项”改进出一个 q^ε , 这就是式中那个无伤大雅的 $+\varepsilon$ 。) 令 $\delta = 1/K = 1/2^{k-1}$ (再略减一点吸收 ε), 即得 $|A(q)| \ll q^{-1-\delta}$ 。

1 从单项估计到 $\chi(p)$ 。由 (5.7), $\chi(p) - 1 = \sum_{\nu \geq 1} A(p^\nu)$ 。三角不等式:
 $|\chi(p) - 1| \leq \sum_{\nu \geq 1} |A(p^\nu)| \ll \sum_{\nu \geq 1} (p^\nu)^{-1-\delta} = \sum_{\nu \geq 1} p^{-\nu(1+\delta)}$ 。

2 等比数列求和。这是公比 $r = p^{-(1+\delta)} < 1$ 的等比级数, 首项 r :
 $\sum_{\nu \geq 1} r^\nu = \frac{r}{1-r} = \frac{p^{-1-\delta}}{1-p^{-1-\delta}}$ 。当 $p \geq 2$ 时分母 $\geq 1 - 2^{-1-\delta} \geq$ 某正常数, 所以整体 $\ll p^{-1-\delta}$ 。这就是 (5.8)。■

第6节 推论：大素数那一截的乘积夹在 $\frac{1}{2}$ 与 $\frac{3}{2}$ 之间

推论. 若 $s \geq 2^k + 1$, 则存在 $p_0 = p_0(k)$ 使得 $\frac{1}{2} \leq \prod_{p > p_0} \chi(p) \leq \frac{3}{2}$.

这立即由 (5.8) 得出, 因为我们可以取 δ 仅依赖于 k 。同样地, 我们将会看到这个结果在 $s \geq 2k + 1$ 时也成立。

这条推论在讲什么、为什么需要它：无穷乘积 $\prod_p \chi(p)$ 是不是正数, 麻烦的地方在于“无穷多个因子相乘会不会乘成 0”。这条推论把素数分两段处理：

- **大素数段** ($p > p_0$)：因子都极接近 1, 它们的无穷乘积被牢牢夹在 $[\frac{1}{2}, \frac{3}{2}]$ 内, 绝不会塌成 0, 也不会爆掉——这一段“安全”, 本节就解决。
- **小素数段** ($p \leq p_0$)：只有有限个因子, 只要逐个证明它们都 > 0 (这是后面引理 5.5、5.6 的任务), 有限个正数相乘自然 > 0 。

两段相乘, 整个 $\mathfrak{S}(N) > 0$ 。这就是把全局问题“分而治之”的总框架。

补全“立即得出”：为什么乘积夹在 $[\frac{1}{2}, \frac{3}{2}]$

1 由 (5.8), $|\chi(p) - 1| \leq Cp^{-1-\delta}$ (C, δ 只依赖 k)。考虑 $\sum_{p > p_0} |\chi(p) - 1| \leq C \sum_{p > p_0} p^{-1-\delta}$ 。因为 $\delta > 0$, 级数 $\sum p^{-1-\delta}$ 收敛, 所以当 p_0 取得足够大时, 尾部和 $\sum_{p > p_0} p^{-1-\delta}$ 可以小于任意给定值, 比如让 $\sum_{p > p_0} |\chi(p) - 1| \leq \frac{1}{4}$ 。

2 对数控制法。写 $\chi(p) = 1 + \epsilon_p$, $\sum_{p > p_0} |\epsilon_p| \leq \frac{1}{4}$ 。用不等式 $|\log(1+x)| \leq 2|x|$ (当 $|x| \leq \frac{1}{2}$) 得 $|\sum_{p > p_0} \log \chi(p)| \leq 2 \sum |\epsilon_p| \leq \frac{1}{2}$ 。于是乘积的对数落在 $[-\frac{1}{2}, \frac{1}{2}]$, 乘积本身落在 $[e^{-1/2}, e^{1/2}] \approx [0.61, 1.65] \subset [\frac{1}{2}, \frac{3}{2}]$ 。(书中用 $\frac{1}{2}, \frac{3}{2}$ 是为了好记的安全界。)

3 关键： p_0 只跟 k 有关（因为 C, δ 只跟 k 有关），不跟 N 也不跟具体的 s 强相关，这点对最后得到“与 N 无关的正下界”至关重要。

第 7 节 $M(q)$ ：同余方程的解数（定义）

定义. 设 $M(q)$ 表示同余方程 $x_1^k + \cdots + x_s^k \equiv N \pmod{q}$ 的解数，其中 $0 < x_1, \dots, x_s \leq q$ 。

这段在讲什么、为什么现在引入它：到目前为止 $A(q), \chi(p)$ 都是“指数和”，抽象、难直接看出正负。这里引入一个看得见摸得着的算术量 $M(q)$ ：在 $\text{mod } q$ 的世界里，把 N 写成 s 个 k 次方之和，到底有几组解。下一节（引理 5.3）将证明这两者其实是一回事—— $\chi(p)$ 就是 $M(p^n)$ 的归一化极限。于是“证明 $\chi(p) > 0$ ”变成“证明同余方程有解、而且解够多”，这是高中生能动手数的事。

符号 $M(q)$

解数：满足 $1 \leq x_j \leq q$ （一个完全剩余系）且 $x_1^k + \cdots + x_s^k \equiv N \pmod{q}$ 的有序组 $(x_1, \dots, x_s)$ 的个数。注意是“有序”的： $(1, 2)$ 和 $(2, 1)$ 算两组。总的候选组数是 q^s （每个 x_j 有 q 个选择）， $M(q)$ 是其中“凑中 N ”的比例乘 q^s 。如果各 k 次方完全均匀分布，平均每个余数被命中 $q^s/q = q^{s-1}$ 次——记住这个 q^{s-1} ，它就是下面归一化的分母。

第 8 节 引理 5.3： $\chi(p)$ 就是解数的密度极限（式 5.9–5.10）

引理 5.3. 我们有 $1 + \sum_{\nu=1}^n A(p^\nu) = M(p^n)/p^{n(s-1)}$, (5.9) 从而

$$\chi(p) = \lim_{n \rightarrow \infty} M(p^n)/p^{n(s-1)}. \quad (5.10)$$

这条引理在讲什么、为什么是“桥梁”：它在抽象的指数和（左边 $A(p^\nu)$ 之和，即 $\chi(p)$ 的部分和）与具体的解数（右边 $M(p^n)$ ）之间架了一座等号桥。(5.10) 说： $\chi(p)$ 等于“解数 $M(p^n)$ 除以它的‘期望值’ $p^{n(s-1)}$ ”当 $n \rightarrow \infty$ 的极限。这个比值就是解的相对密度。从此 $\chi(p) > 0 \iff$ 解数不会比期望值小太多 $\iff$ 同余方程在越来越高的 p 次幂下持续有“足够多”的解。这就把分析问题彻底翻译成了算术问题。

8.1 把解数写成指数和

我们可以用一个程序把 $M(q)$ 表示为指数和……我们有

$$M(q) = q^{-1} \sum_{t=1}^q \sum_{x_1=1}^q \cdots \sum_{x_s=1}^q e\left(\frac{t}{q}(x_1^k + \cdots + x_s^k - N)\right),$$

因为对 t 求和当同余方程成立时给出 q ，否则给出 0。

这一步是“正交关系”的活用，把跳步补全：

- 1 对每一组 $(x_1, \dots, x_s)$ ，用第 0 节的检测器。令 $n = x_1^k + \cdots + x_s^k - N$ 。那条恒等式说 $\sum_{t=1}^q e\left(\frac{t}{q}n\right) = q$ （若 $q \mid n$ ，即同余方程成立）或 0（否则）。所以 $\frac{1}{q} \sum_{t=1}^q e\left(\frac{t}{q}n\right)$ 是一个“指示器”：方程成立时为 1，否则为 0。

- 2 把指示器对所有 $(x_1, \dots, x_s)$ 求和，就数出了解的个数。

$$M(q) = \sum_{x_1=1}^q \cdots \sum_{x_s=1}^q \left[\frac{1}{q} \sum_{t=1}^q e\left(\frac{t}{q}n\right) \right],$$

把 $\frac{1}{q}$ 和 $\sum_t$ 提到最前面，就得到原文那个四重和式。这正是上一章把 $r(N)$ 写成积分 $\int_0^1$ 的离散算术版：积分 $\int_0^1$ 换成有限和 $\frac{1}{q} \sum_{t=1}^q$ ，连续的 α 换成离散的 t/q 。

8.2 按 (t, q) 的公因数归并

我们把那些与 q 有相同最大公因子的 t 值汇集在一起。若这个最大公因子记为 q/q_1 ，则相应的 t 值为 uq/q_1 ，其中 $1 \leq u \leq q_1$ 且 $(u, q_1) = 1$ 。于是

$$M(q) = q^{-1} \sum_{q_1|q} \sum_{\substack{u=1 \\ (u, q_1)=1}}^{q_1} \sum_{x_1=1}^q \cdots \sum_{x_s=1}^q e\left(\frac{u}{q_1}(x_1^k + \cdots + x_s^k - N)\right).$$

把这个换元讲透：

1 每个 t 对应唯一的 q_1 。对 $t \in \{1, \dots, q\}$ ，设 $(t, q) = d$ （最大公因数）。记 $q_1 = q/d$ ，则 $d = q/q_1$ ，且 $q_1 \mid q$ 。所以“ t 的公因数 d ”与“因子 q_1 ”一一对应。

2 固定 q_1 ，符合 $(t, q) = q/q_1$ 的 t 长什么样。这样的 t 必是 $d = q/q_1$ 的倍数，写 $t = u \cdot \frac{q}{q_1}$ 。要让 (t, q) 恰好等于 $\frac{q}{q_1}$ （不能更大），需要 $(u, q_1) = 1$ 。 u 的范围是 $1 \leq u \leq q_1$ 。于是这些 t 就被 u （与 q_1 互素）参数化了。

3 替换分数。 $\frac{t}{q} = \frac{uq/q_1}{q} = \frac{u}{q_1}$ 。代入指数，外层 $\sum_t$ 变成“先对 $q_1 \mid q$ ，再对 $(u, q_1) = 1$ 的 u ”双重求和。得到原文式子。这一步的目的：把分母从 q 化简到它的因子 q_1 ，好让后面认出 S_{u, q_1} 。

8.3 认出 S_{u, q_1} 并合成

$$\text{现在 } \sum_{x=1}^q e\left(\frac{u}{q_1}x^k\right) = \frac{q}{q_1} \sum_{x=1}^{q_1} e\left(\frac{u}{q_1}x^k\right) = \frac{q}{q_1} S_{u, q_1}.$$

- 1 为什么能从 $\sum_{x=1}^q$ 缩到 $\sum_{x=1}^{q_1}$? 被加项 $e\left(\frac{u}{q_1}x^k\right)$ 只依赖 $x \bmod q_1$ (因为 $e(\cdot)$ 看 $\bmod 1$, 而把 x 加上 q_1 , x^k 改变量是 q_1 的倍数, $\frac{u}{q_1} \cdot (q_1 \text{倍})$ 是整数, e 不变)。所以 x 在 $1 \sim q$ 里跑, 每个 $\bmod q_1$ 的余数类恰好出现 q/q_1 次。于是 $\sum_{x=1}^q = \frac{q}{q_1} \sum_{x=1}^{q_1}$ 。后者按定义就是 S_{u,q_1} 。

- 2 每个变量 x_j 都这样处理。一共 s 个变量, 各贡献一个 $\frac{q}{q_1} S_{u,q_1}$, 相乘得到 $\left(\frac{q}{q_1}\right)^s (S_{u,q_1})^s$; 而 $-N$ 那部分给出 $e\left(-\frac{uN}{q_1}\right)$ 。

$$M(q) = q^{-1} \sum_{q_1|q} \sum_{\substack{u=1 \\ (u,q_1)=1}}^{q_1} \left(\frac{q}{q_1}\right)^s (S_{u,q_1})^s e\left(-\frac{uN}{q_1}\right) = q^{s-1} \sum_{q_1|q} A(q_1).$$

- 1 提出 q 的方幂。 $q^{-1} \cdot \left(\frac{q}{q_1}\right)^s = q^{-1} \cdot \frac{q^s}{q_1^s} = q^{s-1} \cdot q_1^{-s}$ 。把 q^{s-1} 提到最外面 (它与 q_1 无关)。

- 2 认出 $A(q_1)$ 。剩下 $q_1^{-s} \sum_{(u,q_1)=1} (S_{u,q_1})^s e(-uN/q_1)$, 对照 (5.2) 正是 $A(q_1)$ (其中 $q_1^{-s} = (q_1^{-1})^s$ 配进每个 S)。于是

$$M(q) = q^{s-1} \sum_{q_1|q} A(q_1).$$

直观意义: 解数 $M(q) =$ 期望值 q^{s-1} 乘上 “所有因子 q_1 的 $A(q_1)$ 之和”。

- 3 取 $q = p^n$ 得 (5.9)。 p^n 的因子恰是 $1, p, p^2, \dots, p^n$ 。所以 $\sum_{q_1|p^n} A(q_1) = \sum_{\nu=0}^n A(p^\nu) = 1 + \sum_{\nu=1}^n A(p^\nu)$ 。代入并把 $p^{n(s-1)}$ 除过去:

$$1 + \sum_{\nu=1}^n A(p^\nu) = \frac{M(p^n)}{p^{n(s-1)}}.$$

这就是 (5.9)。

4 取极限得 (5.10)。左边当 $n \rightarrow \infty$ 就是 $\chi(p)$ 的定义 (5.7)。所以 $\chi(p) = \lim_{n \rightarrow \infty} M(p^n)/p^{n(s-1)}$ ，即 (5.10)。■

注（原文的“注”）。

这段注在讲什么：提醒一个微妙之处——对固定的 N ，(5.9) 左边的级数其实会在某个 ν 之后停下来（后续 $A(p^\nu) = 0$ ），所以 (5.10) 不必真的取极限， n 一旦足够大，比值就稳定不动了。但“足够大”是从哪开始，既依赖 k, p ，也依赖 N 。这点提醒我们：后面证 $\chi(p) > 0$ 时，给出的下界常数 C_p 可能依赖 p ，必须小心保证最终的乘积下界与 N 无关。

第 9 节 定义 τ, k_0, γ ：刻画 p 与 k 的“纠缠程度”（式 5.11）

定义. 对每个素数 p ，设 p^τ 为整除 k 的 p 的最高幂，并令 $k = p^\tau k_0$ 。定义 γ 为

$$\gamma = \begin{cases} \tau + 1 & p > 2, \\ \tau + 2 & p = 2. \end{cases} \quad (5.11)$$

这段在讲什么、为什么需要这些记号：解同余方程 $x^k \equiv m \pmod{p^\nu}$ 的难易，关键看“指数 k 里藏了多少个 p ”。如果 k 与 p 无关 ($p \nmid k$)， k 次方映射在 $\text{mod } p^\nu$ 上很规整；一旦 $p \mid k$ ，就会出现“岔路/退化”。 τ 正是度量这种纠缠的指标， γ 则是“一旦在 $\text{mod } p^\gamma$ 这个临界精度上有解，就能自动提升到任意高精度 $\text{mod } p^\nu$ ”的临界指数（下面引理 5.4 就证这件事）。

符号 τ, k_0, γ

p^τ

整除 k 的 p 的最高幂。即 τ 是 “ k 里有几个因子 p ”。例：

$$k = 12, p = 2 \Rightarrow 12 = 4 \cdot 3, \tau = 2; p = 3 \Rightarrow \tau = 1; p = 5 \Rightarrow \tau = 0.$$

k_0

把 p 抽干净后剩下的部分： $k = p^\tau k_0$ ，且 $p \nmid k_0$ 。例 $k = 12, p = 2: k_0 = 3$ 。

γ

临界指数：奇素数取 $\tau + 1$ ， $p = 2$ 取 $\tau + 2$ 。为什么 $p = 2$ 要多加 1？因为 2 的乘法群结构特殊（ $\text{mod } 2^\nu$ 的单位群不是循环群，而是“几乎循环”，要多一层），后面引理 5.4 的 $p = 2$ 证明会看到这个 +2 的必要性。

第 10 节 引理 5.4：解能从 $\text{mod } p^\gamma$ 提升到 $\text{mod } p^\nu$

引理 5.4. 若同余方程 $y^k \equiv m \pmod{p^\gamma}$ 有解，其中 $m \not\equiv 0 \pmod{p}$ ，则同余方程 $x^k \equiv m \pmod{p^\nu}$ 对每个 $\nu > \gamma$ 都有解。

这条引理在讲什么、为什么关键：它是一种“提升 (lifting)”定理——只要在临界精度 $\text{mod } p^\gamma$ 上把 m （且 m 不被 p 整除）开出 k 次方根，就能在任意更高的精度 $\text{mod } p^\nu$ 上继续开出。这保证了“低精度有解 $\Rightarrow$ 高精度永远有解”，是后面让解数 $M(p^\nu)$ 持续不消失、从而 $\chi(p) > 0$ 的技术核心。这相当于数论里著名的 Hensel 引理在 k 次方根这个特例上的版本。

前置：原根与循环群（证明的工具）

对奇素数 p ， $\text{mod } p^\nu$ 的既约剩余系（与 p^ν 互素的剩余类，共 $\varphi(p^\nu) = p^{\nu-1}(p-1)$ 个）在乘法下构成一个循环群：存在一个“原根” g ，使得每个既约剩余类都能唯一写成 $g^{\text{某指数}}$ 。这把“乘法”变成了“指数相加”——是把开 k 次方根这件难事，化成解一次同余方程 $k\eta \equiv \mu$ 的关键。 $p = 2$ 时结构稍异，证明里单独处理。

10.1 奇素数 $p > 2$ 的情形

设 $p > 2$ 。模 p^ν 的互素既约剩余类组成一个阶为 $\phi(p^\nu) = p^{\nu-1}(p-1)$ 的循环群，可用模 p^ν 的原根 g 的各次幂来表示。若 $\nu > \gamma$ ，则 g 必然也是模 p^γ 的原根。

1 取定原根 g 。既然 $\text{mod } p^\nu$ 的单位群循环，取它的一个生成元（原根） g 。任何与 p 互素的数都 $\equiv g^{\text{指数}} \pmod{p^\nu}$ 。

2 同一个 g 也是 $\text{mod } p^\gamma$ 的原根。因为 $\nu > \gamma$ ，把 $\text{mod } p^\nu$ 的等式“降精度”到 $\text{mod } p^\gamma$ 后， g 仍生成 $\text{mod } p^\gamma$ 的单位群（这是原根的标准性质：高次幂模下的原根，约化到低次幂模仍是原根）。这样我们可以用同一个 g 同时表示两个模下的所有数，指数能直接对应。

记 $m \equiv g^\mu$, $y \equiv g^\eta$, $x \equiv g^\xi \pmod{p^\nu}$ 。那么假设 $y^k \equiv m \pmod{p^\gamma}$ 等价于 $k\eta \equiv \mu \pmod{p^{\gamma-1}(p-1)}$ 。

1 取对数（离散对数）。把 m, y, x 都写成 g 的幂，指数分别是 μ, η, ξ 。这是“离散对数”，把乘法 $\rightarrow$ 指数加法。

2 翻译假设。 $y^k \equiv m \pmod{p^\gamma}$ 即 $g^{k\eta} \equiv g^\mu \pmod{p^\gamma}$ 。两个 g 的幂相等，等价于指数模群的阶相等。 $\text{mod } p^\gamma$ 单位群的阶是 $\phi(p^\gamma) = p^{\gamma-1}(p-1)$ 。所以 $k\eta \equiv \mu \pmod{p^{\gamma-1}(p-1)}$ 。

由于 $k = p^\tau k_0$ 且 $\tau = \gamma - 1$ ，可知 μ 能被 $p^{\gamma-1}$ 整除，也能被 $(k_0, p-1)$ 整除。但现在我们可以找到 ξ 满足 $k\xi \equiv \mu \pmod{p^{\gamma-1}(p-1)}$ ，因为 μ 能被 k 与 $p^{\gamma-1}(p-1)$ 的最大公因子整除。最后一个同余方程等价于 $x^k \equiv m \pmod{p^\nu}$ 。

这是最关键、原文最跳的一步，逐句补全。我们要用“低精度可解”推出“ μ 有某些整除性”，再用这些整除性推出“高精度可解”。先回顾一个事实：

引理（一次同余方程可解判据）

同余方程 $k\xi \equiv \mu \pmod{m}$ 有解 $\iff \gcd(k, m) \mid \mu$ 。理由： $k\xi - \mu = mt$ 即 $k\xi - mt = \mu$ ，由裴蜀定理， k, m 的整数线性组合恰好能取到的值都是 $\gcd(k, m)$ 的倍数。

- 1 从假设提取 μ 的整除性。既然 $k\eta \equiv \mu \pmod{p^{\gamma-1}(p-1)}$ 有解（ η 存在），由上面判据， $\gcd(k, p^{\gamma-1}(p-1)) \mid \mu$ 。现在算这个 gcd：

$k = p^\tau k_0 = p^{\gamma-1} k_0$ （用 $\tau = \gamma - 1$ ），而 $p \nmid k_0$ 。

◦ p 的部分： $\gcd(p^{\gamma-1}, p^{\gamma-1}(p-1)) = p^{\gamma-1}$ （因 $p \nmid (p-1)$ ）。

◦ 与 p 互素的部分： $\gcd(k_0, p-1)$ 。

合起来 $\gcd(k, p^{\gamma-1}(p-1)) = p^{\gamma-1} (k_0, p-1)$ 。所以 $p^{\gamma-1} \mid \mu$ 且 $(k_0, p-1) \mid \mu$ 。这正是原文“ μ 能被 $p^{\gamma-1}$ 整除，也能被 $(k_0, p-1)$ 整除”。

- 2 用 μ 的整除性证明高精度可解。我们要解 $k\xi \equiv \mu \pmod{p^{\nu-1}(p-1)}$ 。由判据，只需 $\gcd(k, p^{\nu-1}(p-1)) \mid \mu$ 。同样算这个 gcd： $k = p^{\gamma-1} k_0$ ，而 $\nu > \gamma$ 即 $\nu - 1 \geq \gamma - 1$ ，所以 p 的部分仍是 $\gcd(p^{\gamma-1}, p^{\nu-1}) = p^{\gamma-1}$ （取较小指数）；互素部分还是 $(k_0, p-1)$ 。于是

$$\gcd(k, p^{\nu-1}(p-1)) = p^{\gamma-1} (k_0, p-1).$$

这和上一步那个 gcd 完全一样！而我们已知它整除 μ 。所以判据满足， ξ 存在，高精度方程可解。

- 3 翻译回 x 。 $k\xi \equiv \mu \pmod{p^{\nu-1}(p-1)}$ 等价于 $g^{k\xi} \equiv g^\mu \pmod{p^\nu}$ ，即 $x^k \equiv m \pmod{p^\nu}$ （其中 $x \equiv g^\xi$ ）。提升完成。

一句话抓住要点 这一步的全部精髓是：提升精度（ $\gamma \rightarrow \nu$ ）时， k 与模数的最大公因数没有变大（增加的都是 p 的更高次幂，但 k 里 p 的部分早已被 $p^{\gamma-1}$ 封

顶)。可解判据只关心这个 $\gcd$ ， $\gcd$ 不变，所以低精度能解就一定高精度能解。 $\gamma = \tau + 1$ 这个取法，正是为了让 k 里 p 的指数 $\tau = \gamma - 1$ 恰好被模数里 p 的指数“追平”。

10.2 $p = 2$ 的情形

设 $p = 2$ 。首先，若 $\tau = 0$ ，即 k 为奇数，则不成问题。因为当 x 跑遍模 2^ν 的一组既约剩余系时 x^k 也跑遍这组既约剩余系，从而同余方程 $x^k \equiv m \pmod{2^\nu}$ 对任何奇数 m 都无需任何假设即可解。

- 1 k 奇数时， k 次方是个“洗牌”。若 $\gcd(k, \varphi(2^\nu)) = 1$ ，则映射 $x \mapsto x^k$ 在单位群上是双射（“幂运算可逆”）。 $\varphi(2^\nu) = 2^{\nu-1}$ 是 2 的幂， k 奇数与之互素，所以 $x \mapsto x^k$ 把既约剩余系一一映满。于是任何奇数 m 都有 k 次方根，无需任何前提。

现在设 $k = 2^\tau k_0$ 为偶数。我们有 $x^k \equiv 1 \pmod{4}$ 对一切奇 x 成立。那些模 2^ν 且 $\equiv 1 \pmod{4}$ 的剩余类组成一个阶为 $2^{\nu-2}$ 的循环群，而 5 是一个原根。记 $m \equiv 5^\mu$ ， $y \equiv 5^\eta$ ， $x \equiv 5^\xi \pmod{2^\nu}$ 。那么假设等价于 $k\eta \equiv \mu \pmod{2^{\gamma-2}}$ 。

- 1 为什么用 $\equiv 1 \pmod{4}$ 的子群、生成元取 5。 $\text{mod } 2^\nu$ 的单位群 ($\nu \geq 3$) 不是循环群，而是 $\{\pm 1\} \times \langle 5 \rangle$ 这种结构。但当 k 偶数时，对任何奇数 x ， $x^k = (x^2)^{k/2} \equiv 1 \pmod{4}$ （奇数的平方 $\equiv 1 \pmod{8}$ ，更 $\equiv 1 \pmod{4}$ ）。所以 x^k 永远落在“ $\equiv 1 \pmod{4}$ ”这个子群里。这个子群是循环的，阶为 $2^{\nu-2}$ ，5 是它的原根（已知事实）。在子群内部就能照搬奇素数那套“取对数”的手法。

- 2 取对数翻译假设。把 m, y （必要时 m 须 $\equiv 1 \pmod{4}$ ，由有解假设保证）写成 5 的幂，假设 $y^k \equiv m \pmod{2^\gamma}$ 变成指数方程，模子群的阶。
 $\text{mod } 2^\gamma$ 对应子群阶 $2^{\gamma-2}$ 。故得 $k\eta \equiv \mu \pmod{2^{\gamma-2}}$ 。

由于 $k = 2^\tau k_0$ 且 $\tau = \gamma - 2$ ，可知 μ 能被 2^τ 整除。因此存在 ξ 使得 $k\xi \equiv \mu \pmod{2^{\nu-2}}$ ，这蕴含 $x^k \equiv m \pmod{2^\nu}$ 。

1 提取整除性。 此时 $\gamma = \tau + 2$ 即 $\tau = \gamma - 2$ 。 $k\xi \equiv \mu \pmod{2^{\gamma-2}}$ 有解，由判据 $\gcd(k, 2^{\gamma-2}) \mid \mu$ 。因 $k = 2^\tau k_0 = 2^{\gamma-2} k_0$ (k_0 奇)， $\gcd(k, 2^{\gamma-2}) = 2^{\gamma-2} = 2^\tau$ 。故 $2^\tau \mid \mu$ 。

2 提升。 要解 $k\xi \equiv \mu \pmod{2^{\nu-2}}$ ，需 $\gcd(k, 2^{\nu-2}) \mid \mu$ 。因 $\nu > \gamma$ 即 $\nu - 2 \geq \gamma - 2 = \tau$ ， $\gcd(k, 2^{\nu-2}) = 2^{\min(\tau, \nu-2)} = 2^\tau$ ，恰好仍整除 μ 。故 ξ 存在。翻译回得 $x^k \equiv m \pmod{2^\nu}$ 。引理 5.4 全部证完。■

这里 $\gamma = \tau + 2$ 中那个“+2”正是为了配平 $\equiv 1 \pmod{4}$ 子群比整个单位群“少了 2^2 这一层”，使得 k 里 2 的指数 τ 与子群模的指数恰好对齐。Davenport 在脚注里指出，Vinogradov 的原书漏掉了 $p = 2$ 这个情形，本书补上了它。

第 11 节 引理 5.5：临界精度有“非平凡解” $\Rightarrow \chi(p) > 0$

引理 5.5. 若同余方程 $x_1^k + \cdots + x_s^k \equiv N \pmod{p^\gamma}$ 有解 $x_1, \dots, x_s$ 满足它们不全被 p 整除，则 $\chi(p) > 0$ 。

这条引理在讲什么、它的地位：把前两节的成果拼起来。前面引理 5.3 说 $\chi(p) = \lim M(p^n)/p^{n(s-1)}$ ，引理 5.4 说“低精度的非退化解能提升到任意高精度”。本引理就用提升术，从“ $\pmod{p^\gamma}$ 有一个不全被 p 整除的解”造出“ $\pmod{p^\nu}$ 的大量解”，逼出 $M(p^\nu)$ 的正下界，从而 $\chi(p) > 0$ 。注意“不全被 p 整除”这个条件至关重要——它保证至少有一个变量是单位（与 p 互素），才能套用引理 5.4 开 k 次方根。

为什么要“不全被 p 整除”（非平凡解）

若所有 x_j 都被 p 整除, 那 x_j^k 被 p^k 整除, 这个解“没有信息量”, 无法提升 (引理 5.4 要求 $m \not\equiv 0$)。要求“不全被 p 整除”就是要至少留一个单位变量当“可开方的支点”。这种解叫**非平凡解** (primitive/non-trivial solution)。

设 $a_1^k + \cdots + a_s^k \equiv N \pmod{p^\gamma}$ 且 $a_1 \not\equiv 0 \pmod{p}$ 。我们任意地选取 $x_2, \dots, x_s$, 只须满足 $x_j \equiv a_j \pmod{p^\gamma}$, $0 < x_j \leq p^\nu$ 。这些选择可以有 $p^{(\nu-\gamma)(s-1)}$ 种方式。然后选取 x_1 满足 $x_1^k \equiv N - x_2^k - \cdots - x_s^k \pmod{p^\nu}$ 。

1 安排支点。不妨设非平凡解里 $a_1 \not\equiv 0 \pmod{p}$ (重新编号即可)。固定 $\nu > \gamma$ 。

2 自由选其余 $s-1$ 个变量。让 $x_2, \dots, x_s$ 各自只须与 a_j 模 p^γ 同余, 但可在 $\setminus(0$

3 用支点把第一个变量解出来。剩下要解

$$x_1^k \equiv N - x_2^k - \cdots - x_s^k \pmod{p^\nu}。记右边为 m 。$$

由引理 5.4 这是可能的, 因为右边 $\equiv a_1^k \pmod{p^\gamma}$ 且 $a_1 \not\equiv 0 \pmod{p}$ 。于是 $M(p^\nu) \geq p^{(\nu-\gamma)(s-1)} = C_p p^{\nu(s-1)}$, 其中 $C_p = p^{-\gamma(s-1)} > 0$ 。由 (5.10), 这蕴含 $\chi(p) > 0$ 。

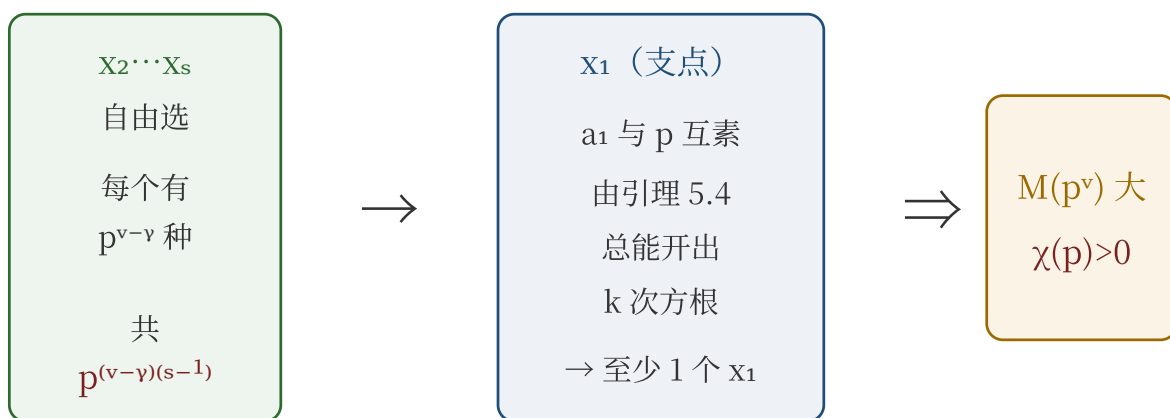
1 验证 m 满足引理 5.4 的前提。因为各 $x_j \equiv a_j \pmod{p^\gamma}$, 所以 $m = N - \sum_{j \geq 2} x_j^k \equiv N - \sum_{j \geq 2} a_j^k \equiv a_1^k \pmod{p^\gamma}$ (用原始解的关系)。又 $a_1 \not\equiv 0 \pmod{p}$ 故 $m \equiv a_1^k \not\equiv 0 \pmod{p}$ 。这恰是引理 5.4 要的“ $y^k \equiv m \pmod{p^\gamma}$ 有解 (取 $y = a_1$) 且 $m \not\equiv 0$ ”。于是 $x_1^k \equiv m \pmod{p^\nu}$ 有解。

2 数解的个数。每一种 $(x_2, \dots, x_s)$ 的选法, 都至少配得出一个 x_1 。所以 $M(p^\nu) \geq p^{(\nu-\gamma)(s-1)}$ (只数“至少一个 x_1 ”, 是下界)。

3 归一化看密度。 $\frac{M(p^\nu)}{p^{\nu(s-1)}} \geq \frac{p^{(\nu-\gamma)(s-1)}}{p^{\nu(s-1)}} = p^{-\gamma(s-1)} =: C_p > 0$ 。注意这个下界 C_p 不依赖 ν (ν 约掉了)，是个固定正数。

4 取极限。由 (5.10), $\chi(p) = \lim_{\nu \rightarrow \infty} M(p^\nu)/p^{\nu(s-1)} \geq C_p > 0$ 。证毕。■

从 $\text{mod } p^\nu$ 的一个非平凡解，造出 $\text{mod } p^\nu$ 的大量解



策略：放掉 $s-1$ 个变量自由跑（贡献 $p^{(\nu-\gamma)(s-1)}$ 种），用一个与 p 互素的“支点变量”靠引理 5.4 把方程配平。解数的下界与 ν 无关，取极限即得 $\chi(p) > 0$ 。

第 12 节 引理 5.6： s 够大时，每个同余方程都有非平凡解

引理 5.6. 若 $s \geq 2k$ (k 为奇数) 或 $s \geq 4k$ (k 为偶数)，则对一切素数 p 和一切 N 都有 $\chi(p) > 0$ 。

这条引理在讲什么、为什么是“最后一块拼图”：引理 5.5 把“ $\chi(p) > 0$ ”归约成“ $\text{mod } p^\gamma$ 有非平凡解”。本引理就无条件地证明：只要 s 够大（奇 k 要 $2k$ 个、偶 k 要 $4k$ 个），这种非平凡解必定存在，对所有 p 和所有 N 。这是把“局部 $\chi(p) > 0$ ”从

“某些好的 N, p ” 升级到 “全部” 的关键。它的证明是本章最精彩、最初等（纯属数数）的部分。

12.1 预处理：把“非平凡”这个附加条件甩掉

由引理 5.5，只须证明 (5.12) 有解不全被 p 整除。若 $N \not\equiv 0 \pmod{p}$ ，则后一要求自动满足。若 $N \equiv 0 \pmod{p}$ ，则只须解 $x_1^k + \cdots + x_{s-1}^k + 1^k \equiv N \pmod{p^\gamma}$ 。因此（把 $s-1$ 换成 s ）只须证明当 $N \not\equiv 0 \pmod{p}$ 时 (5.12) 在 $s \geq 2k-1$ (k 奇) 或 $s \geq 4k-1$ (k 偶) 下可解。

1 当 $N \not\equiv 0 \pmod{p}$ ：非平凡自动满足。若所有 x_j 都被 p 整除，则左边 $\equiv 0 \pmod{p}$ ，不可能等于 $N \not\equiv 0$ 。所以任何解都自动“不全被 p 整除”，附加条件免费成立。

2 当 $N \equiv 0 \pmod{p}$ ：手动塞一个单位变量。强行令最后一个变量 $x_s = 1$ （它 $\not\equiv 0 \pmod{p}$ ，保证非平凡），于是只须解 $x_1^k + \cdots + x_{s-1}^k \equiv N - 1 \pmod{p^\gamma}$ 。新右端 $N - 1 \not\equiv 0 \pmod{p}$ （因 $N \equiv 0$ ），回到第 1 种好情形，但变量少了一个（ $s-1$ 个）。

3 统一表述。两种情形归结为：只须证“当右端 $\not\equiv 0 \pmod{p}$ 时方程可解”，且变量个数门槛降 1。所以把目标设为： $N \not\equiv 0 \pmod{p}$ 时， $s \geq 2k-1$ （奇）或 $4k-1$ （偶）即可解。（证完后把 s 加回 1，就得到引理声称的 $2k$ 与 $4k$ 。）

12.2 奇素数 $p > 2$ ：一条漂亮的“链式”计数

设 $p > 2$ 。我们考虑一切满足 $0 < N < p^\gamma$ ， $N \not\equiv 0 \pmod{p}$ 的 N ，它们的个数为 $\phi(p^\gamma) = p^{\gamma-1}(p-1)$ 。设 $s(N)$ 表示使 (5.12) 可解的最小的 s 。若 $N \equiv z^k N' \pmod{p^\gamma}$ ，则显然 $s(N) = s(N')$ 。

符号 $s(N)$

$s(N)$ = 让 $x_1^k + \cdots + x_s^k \equiv N \pmod{p^\gamma}$ 可解所需的最少项数 s 。我们的目标就是证明：所有 N （与 p 互素）的 $s(N)$ 都 $\leq 2k - 1$ 。这样取 $s = 2k - 1$ 就对所有 N 都够用。

- 1 为什么 $s(N) = s(N')$ （当 $N \equiv z^k N'$ ）。若 $\sum x_j^k \equiv N'$ 可解，把每个 x_j 乘以 z ，则 $\sum (zx_j)^k = z^k \sum x_j^k \equiv z^k N' \equiv N$ 。所以 N' 能用 s 项凑出 $\Rightarrow N$ 也能用 s 项凑出（ $z \neq 0$ 时此变换可逆），故二者所需最少项数相同。这说明： $s(N)$ 只依赖 N 落在哪个“ k 次方剩余陪集”里。

……同余方程 $z^k \equiv a \pmod{p^\gamma}$ 可解当且仅当 α 能被 $p^{\gamma-1}\delta$ 整除，其中 $\delta = (k, p-1)$ 。…… a （模 p^γ ）的不同值的个数为 $\frac{p^{\gamma-1}(p-1)}{p^{\gamma-1}\delta} = \frac{p-1}{\delta} = r$ 。因此 N 的每个取值类都至少包含 r 个元素。

补全“ k 次方一共能取多少个不同值”这个计数：

- 1 取对数。设 g 是 $\text{mod } p^\gamma$ 的原根， $z \equiv g^\zeta$ ， $a \equiv g^\alpha$ 。则 $z^k \equiv a$ 变成 $k\zeta \equiv \alpha \pmod{\varphi(p^\gamma)}$ ，模数 $\varphi(p^\gamma) = p^{\gamma-1}(p-1)$ 。

- 2 可解判据。有解 $\iff \gcd(k, p^{\gamma-1}(p-1)) \mid \alpha$ 。前面算过这个 $\gcd = p^{\gamma-1}(k_0, p-1)$ 。记 $\delta := (k, p-1)$ 。注意 $(k, p-1) = (p^\gamma k_0, p-1) = (k_0, p-1)$ （因 $\gcd(p, p-1) = 1$ ），所以 $\gcd = p^{\gamma-1}\delta$ 。故 $z^k \equiv a$ 可解 $\iff p^{\gamma-1}\delta \mid \alpha$ 。

- 3 数 k 次方剩余的个数。 α 在 $\text{mod } p^{\gamma-1}(p-1)$ 下，能被 $p^{\gamma-1}\delta$ 整除的剩余类共有 $\frac{p^{\gamma-1}(p-1)}{p^{\gamma-1}\delta} = \frac{p-1}{\delta} =: r$ 个。这就是“ z^k 当 $z \neq 0$ 时所取不同值的个数”。

- 4 翻成 N 的语言。由步骤 1 ($s(N)$ 只看 k -次方陪集), 把所有 N 按 $s(N)$ 分类, 每一类至少含 r 个元素 (一类是若干个完整的陪集之并, 而每个陪集恰好 r 大小)。这点稍后用来做计数压制。

我们先枚举一切使 $s(N) = 1$ 的 N : ……然后枚举 $s(N) = 2$ 的……依此类推。这些集合中有些可能为空, 但我们将证明任意两个相邻的集合不能都为空。

考虑不在前 $j - 1$ 个集合中的最小的 $N' \not\equiv 0 \pmod{p}$ 。那么 $N' - 1$ 或 $N' - 2$ 之一 $\not\equiv 0 \pmod{p}$, 并且由于它小于 N' , 必定在前 $j - 1$ 个集合之中。把 N' 表示成 $(N' - 1) + 1^k$ 或 $(N' - 2) + 1^k + 1^k$, 我们推出 $s(N') \leq j + 1$ 。因此 $s(N) = j$ 、 $s(N) = j + 1$ 的两个集合不能都为空。

这是本章最妙的“链式”论证, 逐句补全。目标: 证明随着 s 增加, $s(N)$ 的可能取值不会“跳着出现留下两个连续空挡”, 从而 $s(N)$ 的最大值不会太大。

- 1 把 N 按 $s(N)$ 装进盒子。第 j 个盒子装所有满足 $s(N) = j$ 的 N 。有些盒子可能空 (某个 j 没有 N 恰好需要 j 项)。

- 2 关键引理: 相邻两盒不能同时空。反证不需要, 直接构造。取“还没被前 $j - 1$ 个盒子收纳、且与 p 互素的最小的 N' ”。

- 3 找一个比它小、已被收纳的邻居。 $N' - 1$ 和 $N' - 2$ 这两个数里, 至少有一个 $\not\equiv 0 \pmod{p}$ (因为 $p > 2$, 连续两个数不可能都是 p 的倍数, 三个数 $N', N' - 1, N' - 2$ 中至多一个被 p 整除)。设这个不被 p 整除的邻居是 $N' - 1$ (或 $N' - 2$)。它比 N' 小, 而 N' 是“最小的没被前 $j - 1$ 盒收纳者”, 所以这个更小的邻居必已被前 $j - 1$ 个盒子收纳, 即 $s(N' - 1) \leq j - 1$ (或 $s(N' - 2) \leq j - 1$)。

- 4 给 N' 现造一个表示。若邻居是 $N' - 1$: 把它的 $\leq j - 1$ 项表示后面再补一项 1^k , 得 $N' = (N' - 1) + 1^k$, 用了 $\leq j - 1 + 1 = j$ 项; 若邻居是

$N' - 2$: 补两项 $1^k + 1^k$, 得 $N' = (N' - 2) + 1^k + 1^k$, 用了 $\leq (j - 1) + 2 = j + 1$ 项。无论哪种, $s(N') \leq j + 1$ 。

- 5 **结论。** 这说明 N' 落进第 j 或第 $j + 1$ 个盒子。也就是说, 第 j 与第 $j + 1$ 盒不可能同时是空的 (否则 N' 无处可去, 矛盾)。即 “连续两个盒子不能都空”。

按 $s(N)$ 分盒, 相邻两盒不能同时为空



不会出现两个相邻的空盒 $\rightarrow$ 非空盒至少占一半 $\rightarrow$ 链不会太长

$$\frac{1}{2}(m+1) \cdot r \leq \phi(p^\gamma) \Rightarrow m \leq 2k-1$$

“相邻两盒不能都空” 逼着非空盒至少占总数的一半。每个非空盒又至少含 r 个 N 。把元素总数 $\leq \phi(p^\gamma)$ 一卡, 链长 m 就被压到 $\leq 2k - 1$ 。

设最后一个非空集合是 $s(N) = m$ 的那个。那么前 $m - 1$ 个集合中至少有 $\frac{1}{2}(m - 1)$ 个非空, 加上第 m 个, 至少有 $\frac{1}{2}(m + 1)$ 个非空集合。由于每个集合至少含 r 个数, $\frac{1}{2}(m + 1)r \leq \phi(p^\gamma) = p^{\gamma-1}(p - 1)$, 由此 $(m + 1) \leq \frac{2p^{\gamma-1}(p - 1)}{r} = 2p^{\gamma-1}\delta = 2p^\tau(k_0, p - 1) \leq 2k$. 因此 $m \leq 2k - 1 \dots\dots$

- 1 **非空盒至少占一半。** 在前 $m - 1$ 个盒子里, 由 “相邻两盒不能都空”, 每两个连续盒至少有一个非空, 所以非空的至少 $\frac{1}{2}(m - 1)$ 个。再加上第 m 盒 (按定义非空), 总非空盒数 $\geq \frac{1}{2}(m - 1) + 1 = \frac{1}{2}(m + 1)$ 。

- 2 **元素计数卡上界。** 每个非空盒至少 r 个 N , 所有 N (与 p 互素、\(\backslash\)

3 解出 m 。 $m + 1 \leq \frac{2p^{\gamma-1}(p-1)}{r}$ 。代入 $r = \frac{p-1}{\delta}$ ：
 $\frac{2p^{\gamma-1}(p-1)}{(p-1)/\delta} = 2p^{\gamma-1}\delta$ 。再用 $\gamma - 1 = \tau$ 、 $\delta = (k_0, p-1)$ ：
 $= 2p^\tau(k_0, p-1)$ 。

4 放成 $\leq 2k$ 。 $(k_0, p-1) \leq k_0$ ，故 $2p^\tau(k_0, p-1) \leq 2p^\tau k_0 = 2k$ 。于是
 $m + 1 \leq 2k$ ，即 $m \leq 2k - 1$ 。

5 收束。 m 是出现过的最大 $s(N)$ ，所以所有 N （与 p 互素）的
 $s(N) \leq 2k - 1$ 。也就是说 $s \geq 2k - 1$ 时，(5.12) 对一切这样的 N 可解。
这就完成了 $p > 2$ 的情形。

12.3 $p = 2$ 的情形

设 $p = 2$ 。若 $\tau = 0$ ，即 k 为奇数，则 (5.12) 在 $s = 1$ 时即可解……

现在设 $\tau \geq 1$ ，即 k 为偶数。我们不妨设 $0 < N < 2^\gamma$ ，因为现在 N 是奇数。在 (5.12) 中取所有 x_j 为 0 或 1，我们当然可以在 $s \geq 2^\gamma - 1$ 时解出。

现在 $2^\gamma - 1 = 2^{\tau+2} - 1 \leq 4k - 1$ 。因此当 $s \geq 4k - 1$ 时即足够……

1 k 奇数 ($\tau = 0$)。引理 5.4 证明里已说明： k 奇时 $x \mapsto x^k$ 是单位群上的双射，所以任意奇数 N （即 $N \not\equiv 0 \pmod{2}$ ）单用 $s = 1$ 项 $x_1^k \equiv N$ 就可解。此时 2 这个素数对 s 没有额外要求，唯一的约束来自奇素数那边的 $2k$ 。引理在 k 奇时成立。

2 k 偶数 ($\tau \geq 1$)。此时 N 须为奇数（否则前面已用塞 1^k 的技巧化归）。用最朴素的办法：让每个 x_j 只取 0 或 1，则 x_j^k 只取 0 或 1。要凑出 $N \bmod 2^\gamma$ （一个 $< 2^\gamma$ 的数），最多需要把若干个 1 加起来，即至多 $2^\gamma - 1$ 个 1（因为 $N \leq 2^\gamma - 1$ ）。所以 $s \geq 2^\gamma - 1$ 必能解。

- 3 把门槛翻成 $4k - 1$ 。 $\gamma = \tau + 2$ ，故 $2^\gamma - 1 = 2^{\tau+2} - 1 = 4 \cdot 2^\tau - 1$ 。又 $2^\tau \leq 2^\tau k_0 = k$ ，所以 $4 \cdot 2^\tau - 1 \leq 4k - 1$ 。于是 $s \geq 4k - 1$ 就够。把预处理时甩掉的那 1 个变量加回，得引理声称的 $s \geq 4k$ 。 k 偶情形证毕。 ■

注（原文的“注”）。

这段注在讲什么：替上面那个“取 0/1”的朴素办法辩护——它看着粗糙，其实在 $k = 2^\tau$ （纯 2 的幂）且 $\tau \geq 2$ 时一点没浪费。原因：此时对奇数 x ， $x^{2^\tau} \equiv 1 \pmod{2^{\tau+2}}$ ；对偶数 x ， $x^{2^\tau} \equiv 0 \pmod{2^{\tau+2}}$ 。也就是说 $x^k \pmod{2^\gamma}$ 本来就只能取 0 和 1 两个值，所以“只用 0、1 当 x_j ”根本没丢任何可能性，门槛 $2^\gamma - 1$ 是紧的。

为什么奇 x 的 2^τ 次幂 $\equiv 1 \pmod{2^{\tau+2}}$

这是“升幂引理 (LTE)”的一个特例：奇数 x ， $x^2 \equiv 1 \pmod{8}$ ，再反复平方，每平方一次模数的 2 指数升 1，到 2^τ 次幂时模数升到 $2^{\tau+2}$ 。偶数 x 则 x^{2^τ} 含 2^{2^τ} 这么高的 2 幂，当 $2^\tau \geq \tau + 2$ （即 $\tau \geq 2$ ）时 $\equiv 0 \pmod{2^{\tau+2}}$ 。

12.4 历史插曲： $\Gamma(k)$ 与一张表

Hardy 与 Littlewood 定义 $\Gamma(k)$ 为使下述成立的最小的 s 值：同余方程 (5.12) 对一切 p 和一切 N 都有解 $x_1, \dots, x_s$ 不全被 p 整除。在此记号下，引理 5.6 表明当 k 为奇数时 $\Gamma(k) \leq 2k$ ，当 k 为偶数时 $\Gamma(k) \leq 4k$ 。……

这段在讲什么、背景：G. H. Hardy 与 J. E. Littlewood 是 1920 年代用圆法系统研究 Waring 问题的奠基人。他们专门定义了 $\Gamma(k)$ ：保证所有同余条件都能非平凡满足所需的最小项数。这是 Waring 问题里“局部障碍”的精确度量—— $\Gamma(k)$ 个 k 次方就足以扫清所有 $\pmod{p^\gamma}$ 的限制。我们这一章的引理 5.6 给出了上界 $\Gamma(k) \leq 2k$ （奇）或

$\leq 4k$ (偶); Hardy–Littlewood 在论文 P.N.VIII 中做了更细致的研究, 确定了所有“反常” k (即 $\Gamma(k) > k$ 的 k)。下表是 $\Gamma(k)$ 的前几项:

k	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\Gamma(k)$	4	16	5	9	4	32	13	12	11	16	6	14	15	64

读这张表 (看出规律)

- $k = 2^r$ (纯 2 的幂: 4, 8, 16) 时 $\Gamma(k)$ 暴涨:
 $\Gamma(4) = 16 = 4k$, $\Gamma(8) = 32 = 4k$, $\Gamma(16) = 64 = 4k$ 。这正对应 12.3 那个“ x^k 只取 0、1”的极端退化情形, 最难, 需要满满 $4k$ 项。我们的上界 $4k$ 在这些点是精确的。
- 很多奇 k (如 3, 5, 7, 13) 的 $\Gamma(k)$ 很小, 甚至 $\leq k$: 如 $\Gamma(7) = 4$ 。这些是“好” k , 同余障碍轻微。
- 本章引理 5.6 给的是普适上界, 能覆盖最坏情形, 对证明定理 5.1 已足够, 但并非每个 k 都最优。

第 13 节 定理 5.1: 奇异级数有正下界 (全章高潮)

定理 5.1. 若 $s \geq 2^k + 1$, 则 $\mathfrak{S}(N) \geq C_1(k, s) > 0$ 对一切 N 成立。

证. 这个结果由引理 5.6 和引理 5.2 的推论得出, 因为 $2^k + 1 \geq 2k$ (k 为奇数) 以及 $2^k + 1 \geq 4k$ (k 为偶数, $k > 2$)。

把这条“一句话证明”展开成完整逻辑:

- 1 条件够强。假设 $s \geq 2^k + 1$ 。先核对它满足前面所有引理的要求:
 - k 奇: 要 $s \geq 2k$ 。而 $2^k + 1 \geq 2k$ 对 $k \geq 1$ 恒成立 (指数增长压倒线性)。✓

- k 偶且 $k > 2$: 要 $s \geq 4k$ 。验证 $2^k + 1 \geq 4k$: $k = 4$ 时 $17 \geq 16\sqrt{}$; $k = 6$ 时 $65 \geq 24\sqrt{}$; k 再大只会更宽松。 $\sqrt{}$ ($k = 2$ 是平方和的特殊情形, 需单独的经典处理, 不在此通用定理内。)
- 引理 5.2 与其推论本身就要求 $s \geq 2^k + 1$ 。 $\sqrt{}$

2 分两段制服欧拉乘积。由引理 5.2,

$\mathfrak{S}(N) = \prod_p \chi(p) = \left(\prod_{p \leq p_0} \chi(p) \right) \left(\prod_{p > p_0} \chi(p) \right)$, $p_0 = p_0(k)$ 由推论给定。

3 大素数段: 下界 $\frac{1}{2}$ 。由第 6 节推论, $\prod_{p > p_0} \chi(p) \geq \frac{1}{2} > 0$ 。

4 小素数段: 有限个正因子。由引理 5.6, 每个 $\chi(p) > 0$ (对一切 p, N)。 $p \leq p_0$ 只有有限个素数, 有限个正数相乘仍为正: $\prod_{p \leq p_0} \chi(p) > 0$ 。

5 合成正下界。两段相乘 $\mathfrak{S}(N) \geq \frac{1}{2} \prod_{p \leq p_0} \chi(p) > 0$ 。

6 下界与 N 无关——这是最后也最重要的一关。要得到统一常数 C_1 , 必须说明 $\prod_{p \leq p_0} \chi(p)$ 不会随 N 变化而趋于 0。每个 $\chi(p)$ 由引理 5.5 有下界 $C_p = p^{-\gamma(s-1)}$, 它只依赖 p, k, s , 不依赖 N (引理 5.6 保证非平凡解对一切 N 存在, 所以这个下界对一切 N 都适用)。于是

$$\mathfrak{S}(N) \geq \frac{1}{2} \prod_{p \leq p_0} C_p =: C_1(k, s) > 0,$$

这个 C_1 只跟 k, s 有关, 对一切 N 成立。定理证毕。■

第 14 节 这定理意味着什么 + 文献来历

定理 5.1 是对定理 4.1 的一个必要补充, 它表明渐近公式中的主项是 $\gg N^{s/k-1}$, 从而当 $N \rightarrow \infty$ 时 $r(N) \rightarrow \infty$ 。

1 把主项的大小定下来。上一章定理 4.1 给出

$r(N) = \mathfrak{S}(N)\mathfrak{J}(N)N^{s/k-1} + (\text{更小的误差})$ 。本章证明 $\mathfrak{S}(N) \geq C_1 > 0$ ；而奇异积分 $\mathfrak{J}(N)$ 也（在别处证明）有正下界。于是主项 $\geq C_1 \cdot (\text{常数}) \cdot N^{s/k-1} \gg N^{s/k-1}$ 。

2 结论：表示数趋于无穷。因为 $s > k$ （项数多于次数），指数 $\frac{s}{k} - 1 > 0$ ，所以 $N^{s/k-1} \rightarrow \infty$ ，从而 $r(N) \rightarrow \infty$ 。这正是 Waring 问题想要的：不仅每个大数都能表示，而且表示方法越来越多。

3 为什么说“必要补充”。定理 4.1 只给出了渐近公式的“骨架”（主项的形状和误差控制），但若主项系数可能为 0，骨架就是空的。定理 5.1 填上“系数确实为正”这块，才让整套渐近公式有实质意义。两者缺一不可。

在本章中我大体上遵循了 Vinogradov 的论述 [93, 第 2 章]。这比 Hardy 与 Littlewood 原来的论述稍微简单一些。

这段在讲历史脉络：Waring 问题（1770 年 E. Waring 提出“每个数是 $g(k)$ 个 k 次方之和”）的现代解析处理，由 Hardy–Littlewood 在 1920 年代用圆法开创。后来 I. M. Vinogradov（维诺格拉多夫）对方法做了重要简化与强化。Davenport 这本讲义在奇异级数这一章采用了 Vinogradov《数论中的三角和方法》第 2 章的处理，比 Hardy–Littlewood 的原始论证更简洁。脚注还诚实地指出：Vinogradov 的版本里漏了 $p = 2$ 这种情形（即我们引理 5.4 的后半），Davenport 在此补全。这体现了数学写作中“补全前人疏漏、把每个情形都照顾到”的严谨精神——也正是本讲解逐段补全跳步的理由。

原书脚注

1. 脚注 1（关于 $\Gamma(k)$ ）：另见 Chowla [14]。——即 $\Gamma(k)$ 的更精细研究除 Hardy–Littlewood 的 P.N.VIII 外，还可参考印度数论家 S. Chowla 的工作。

2. 脚注 2 (关于 $p = 2$): 我们的引理 5.4 中 $p = 2$ 的情形被 Vinogradov 无意中遗漏了。
——这就是本章 $p = 2$ 处理需要额外小心、并单独补证的来历。

全章逻辑回顾 (一张地图)

- 1 定义 奇异级数 $\mathfrak{S}(N) = \sum_q A(q)$ (式 5.1-5.2)。
- 2 引理 5.1: $A(q)$ 乘性 $\rightarrow$ 引理 5.2: 欧拉乘积 $\mathfrak{S} = \prod_p \chi(p)$, 大素数因子 ≈ 1 (式 5.8)。
- 3 推论: 大素数段乘积 $\in [\frac{1}{2}, \frac{3}{2}]$, 安全。剩下只需管有限个小素数。
- 4 引理 5.3: $\chi(p) = \lim M(p^n)/p^{n(s-1)}$, 把指数和翻译成解数密度。
- 5 引理 5.4: 解能从 $\bmod p^\gamma$ 提升到任意 $\bmod p^\nu$ (原根 + 一次同余可解判据)。
- 6 引理 5.5: 临界精度有非平凡解 $\Rightarrow \chi(p) > 0$ (造大量高精度解)。
- 7 引理 5.6: s 够大时 (奇 $2k$ /偶 $4k$), 非平凡解对一切 p, N 存在 (链式计数 + 0/1 朴素法)。
- 8 定理 5.1: 合成 $\Rightarrow \mathfrak{S}(N) \geq C_1 > 0$, 与 N 无关 $\Rightarrow r(N) \rightarrow \infty$ 。

