

DAVENPORT · 圆法 · 高中详解版

方程 $c_1x_1^k + \cdots + c_sx_s^k = 0$ 的整数解

The equation $c_1x_1^k + \cdots + c_sx_s^k = 0$

本章要解决的问题 给定固定整数系数 $c_1, \dots, c_s$ (都不为 0)，我们要问：方程

$$c_1x_1^k + \cdots + c_sx_s^k = 0$$

有没有非平凡的整数解（即不全为 0 的解）？甚至——有没有**无穷多组**这样的解？

本章用前几章发展出来的**圆法**（Hardy–Littlewood circle method，又叫 Hardy–Littlewood 圆法 / 解析方法）给出回答：只要变量个数 s 足够大（具体地 $s \geq 2^k + 1$ 且 $s \geq k^2(2k - 1) + 1$ ），方程就有**无穷多组**每个分量都非零的整数解。

读完你将掌握：(1) 为什么“等号成立的方程”可以用一个**积分**来数解的个数；(2) 这个积分如何拆成“奇异积分 + 奇异级数”两个因子；(3) 为什么只要证明这两个因子都是**正数**，就能断定解有无穷多个；(4) 证明奇异级数为正所需的、关于“同余方程有解”的全部细节推导。

阅读方式：下文把**译文的每一段**都先原样给出（标注「原文」），紧接着用「高中详解」逐字、逐符号、逐步骤地讲清楚。本章是第 4–7 章方法的“收口”，会反复用到那几章的工具，所有用到的工具我们都会从零讲一遍，不要求你已经读过前面。

预备：你需要先认识的几样“分析数论道具”

本章是整本书第二部分的高潮之一，几乎每一句都建立在前几章的记号之上。为了让只有高中知识的你能独立读懂，我们先把会反复出现的几个“道具”集中讲清楚。读不顺的

话可以先扫一眼，正文用到时再回头查。

道具 1 复指数 $e(\alpha)$ 我们定义

$$e(\alpha) = e^{2\pi i \alpha}.$$

这里 i 是虚数单位 ($i^2 = -1$), $e^{i\theta} = \cos \theta + i \sin \theta$ 是欧拉公式。所以

$$e(\alpha) = \cos(2\pi\alpha) + i \sin(2\pi\alpha).$$

读作"e 括号 alpha"。把 α 看成"圈数": 当 α 从 0 增到 1, 点 $e(\alpha)$ 在复平面的单位圆上正好逆时针转一整圈。它的**模长** (到原点的距离) 恒等于 1:

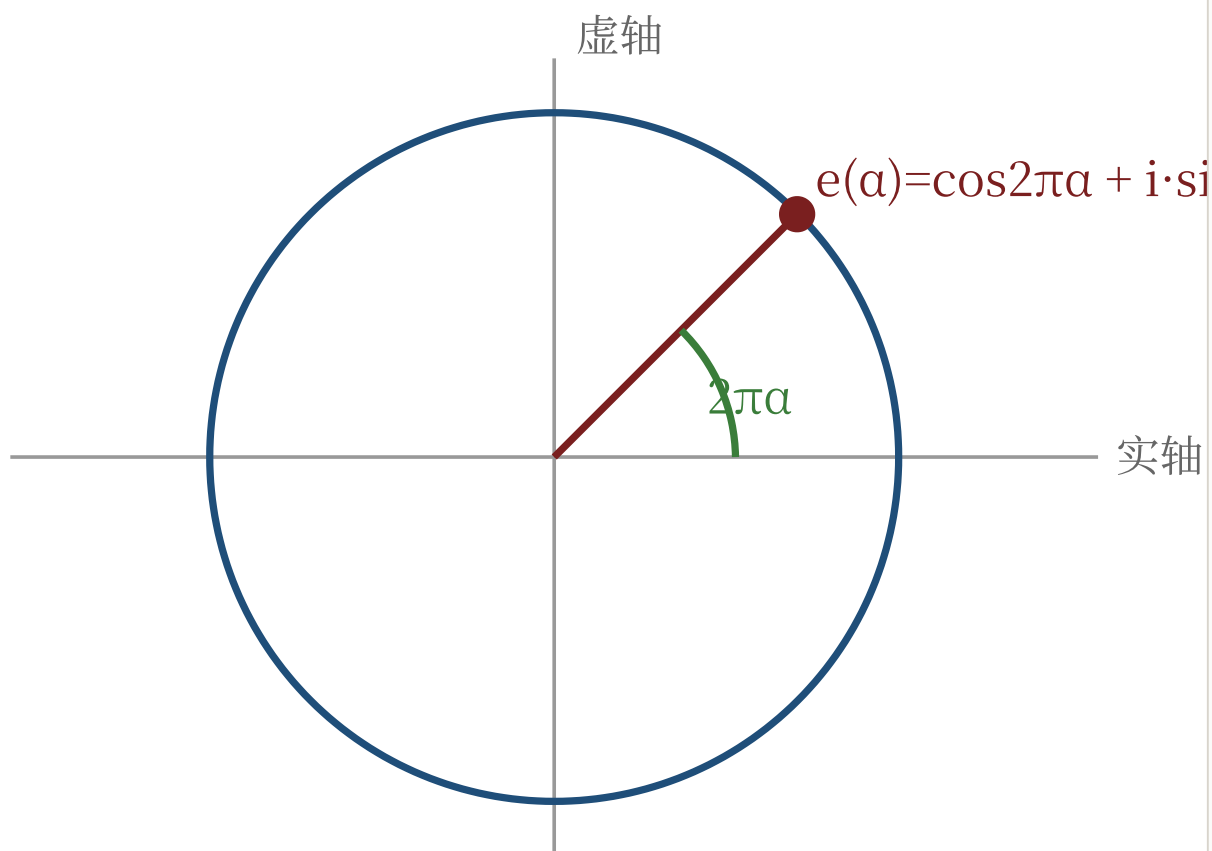
$$|e(\alpha)| = \sqrt{\cos^2 + \sin^2} = 1.$$

最重要的两条性质:

(a) **整数不变性**: 若 n 是整数, 则 $e(\alpha + n) = e(\alpha)$ (转整圈回到原地)。所以 $e(\alpha)$ 只依赖 α 的小数部分。

(b) **"挑零"性质 (正交关系)**: 对整数 m ,

$$\int_0^1 e(m\alpha) d\alpha = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0. \end{cases}$$



α 从 0 到 1, 点绕圈一周; 模长恒为 1

复指数 $e(\alpha)$ 把实数 α 缠到单位圆上。"圆法"的名字就来自这个圆。

为什么 (b) 成立 (亲手算一遍) 若 $m = 0$: $e(0 \cdot \alpha) = e(0) = 1$, 于是

$$\int_0^1 1 d\alpha = 1.$$

若 $m \neq 0$: 直接积分复指数, $\int_0^1 e^{2\pi i m \alpha} d\alpha = \left[\frac{e^{2\pi i m \alpha}}{2\pi i m} \right]_0^1 = \frac{e^{2\pi i m} - e^0}{2\pi i m}$ 。而 $e^{2\pi i m} = e(m) = 1$ (m 是整数, 转整圈), $e^0 = 1$, 分子 $= 1 - 1 = 0$ 。所以积分为 0。

这条性质是圆法的**心脏**: 它能从一大堆指数里"挑出"指数恰好为 0 的那一项——而"指数为 0"正对应"方程等号成立"。

道具 2 求和号 $\sum$ 与指数和 $\sum_{x=1}^n f(x)$ 读作"对 x 从 1 到 n 求和", 就是 $f(1) + f(2) + \cdots + f(n)$ 。本章的核心对象是**指数和**

$$T(\alpha) = \sum_{x=1}^P e(\alpha x^k),$$

即把许多个单位圆上的点 $e(\alpha x^k)$ ($x = 1, 2, \dots, P$) 当作向量首尾相接加起来。它是一个随 α 变化的复数。把若干个这样的和相乘再积分，就能数出方程的解数——这是下文的主线。

道具 3 记号 $O(\cdot)$ 与 $\ll$ $f(P) = O(g(P))$ 读作" f 是 g 的大 O "，意思是：存在一个不依赖 P 的常数 C ，使得当 P 充分大时 $|f(P)| \leq C g(P)$ 。它表示" f 的大小被 g 控制住，至多差一个固定倍数"。

$f \ll g$ 与 $f = O(g)$ 同义。 $A \asymp B$ (或 $A \gg B$ 且 $A \ll B$) 表示两者同阶。

直观：在 $\mathcal{N}(P) = (\text{主项}) P^{s-k} + O(P^{s-k-\delta})$ 里，主项是 P^{s-k} 这么大，误差项 $P^{s-k-\delta}$ 因为指数小了 δ ，当 $P \rightarrow \infty$ 时相对主项可忽略。"主项 $>$ 误差上界"正是我们要赢的那场赛跑。

道具 4 同余 $a \equiv b \pmod{m}$ 读作" a 与 b 模 m 同余"，意思是 $m \mid (a - b)$ ，即 $a - b$ 能被 m 整除，也就是" a, b 除以 m 余数相同"。本章后半段反复研究"方程取模 p^ν (p 为素数) 有多少组解"，记这个解数为 $M(p^\nu)$ 。 $p \nmid d$ 表示" p 不整除 d "， (a, q) 表示 a, q 的最大公约数， $(a, q) = 1$ 表示互素。

道具 5 奇异积分与奇异级数 (先认个脸) 圆法把解数 $\mathcal{N}(P)$ 拆成两个相乘的因子：

- **奇异积分 (singular integral)**：一个实变量的积分，记 $C'_{k,s}$ ，它来自把"方程在实数范围内"的解的"密度"积起来，反映的是实数 (无穷处) 的几何信息；
- **奇异级数 (singular series) $\mathfrak{S}$** ：一个无穷级数，反映的是每个素数 p 处 (同余) 的算术信息。

最终的渐近公式形如 " $\mathcal{N}(P) \approx C'_{k,s} \cdot \mathfrak{S} \cdot P^{s-k}$ "。要证明解有无穷多个，只需证 $C'_{k,s} > 0$ 且 $\mathfrak{S} > 0$ 。本章的全部技术工作，就是分别验证这两件事。"奇异" (singular) 只是历史沿用的名字，不含"奇怪"之意。

一、把"求整数解"翻译成积分

原文：我们现在研究上述方程在整数（正的或负的）范围内的解，其中 $c_1, \dots, c_s$ 是固定的整数，没有一个为 0。如果 k 是偶数，我们显然必须假设这些系数并非全部同号。如果 k 是奇数，则我们可以在必要时把 x_i 换成 $-x_i$ 来保证这一点。于是，稍微改变记号后，我们可以把方程写成

$$c_1 x_1^k + \dots + c_r x_r^k - c_{r+1} x_{r+1}^k - \dots - c_s x_s^k = 0. \quad (8.1)$$

这段在讲什么：给方程"摆好姿势"。我们要让方程里既有正项又有负项，否则根本不可能等于 0。下面把每句话的逻辑补全。

第一句确定了研究对象：系数 $c_1, \dots, c_s$ 是**给定的、固定的**整数（不是未知数），而且都不为 0（否则那一项消失，相当于变量更少，没意思）。未知数 $x_1, \dots, x_s$ 取**整数**，可正可负。

"如果 k 是偶数，显然必须假设系数并非全部同号"——这句话的**"显然"**要补全：

- 1 k 偶数时，任何整数的 k 次方都 ≥ 0 （因为偶数次方非负，例如 $x^2 \geq 0, x^4 \geq 0$ ）。
- 2 若所有系数 c_j 同号，比如全为正，那么每一项 $c_j x_j^k \geq 0$ ，整个左边是若干非负数之和，**只有当每一项都为 0 时总和才为 0**。
- 3 而每一项为 0 又要求每个 $x_j = 0$ （因 $c_j \neq 0$ ）。于是唯一的解是平凡解 $x_1 = \dots = x_s = 0$ ，没有非平凡解。
- 4 所以 k 偶数时若想要非平凡解，**必须有正有负**——这就是"不全同号"的来由。它不是人为限制，而是问题有意义的**必要条件**。

"如果 k 是奇数，可以把 x_i 换成 $-x_i$ 来保证这一点"——补全理由：

- 1 k 奇数时 $(-x)^k = -x^k$ （例如 $(-x)^3 = -x^3$ ）。所以把变量 x_i 换成 $-x_i$ ，等价于把系数 c_i 换成 $-c_i$ 。

2 因此当 k 奇数时，我们**总能**通过对某些变量改名 ($x_i \rightarrow -x_i$) 来调整任意一项的符号，凑出"有正有负"。即 k 奇数时"不全同号"的要求可以自动满足，不构成限制。

"稍微改变记号后写成 (8.1)": 把正系数项排在前面 (共 r 项)，负系数项排在后面 (共 $s - r$ 项)，并把负号显式写出来。这样做**不损失一般性**——只是重新编号和把符号提出来。注意 (8.1) 里写在式子中的 $c_1, \dots, c_s$ 已经是它们的绝对值。

原文：其中 $c_1, \dots, c_s$ 现在都是正整数，且 $1 \leq r \leq s - 1$ 。我们研究 (8.1) 在正整数范围内的解。

这段把约定收紧：

$c_1, \dots, c_s$ 都是正整数

负号已经被显式提到式子里，所以剩下的系数都取正值。

$1 \leq r \leq s - 1$

r 是正项的个数。 $r \geq 1$ 保证至少有一个正项， $r \leq s - 1$ 保证至少有一个负项——这正是上面分析的"有正有负"。

研究正整数解

关键的简化！由于可正可负的整数解和"全取正整数"的解一一对应（一个非平凡整数解，通过 k 奇时改符号、或本就分布在正负两侧，都能归到"在某盒子里数正解"这个标准问题），我们只在 $x_j > 0$ 的范围内计数。这样指数和的求和下标从 1 起，干净利落。"非平凡"在正整数语境下自动满足（正整数 $\geq 1 \neq 0$ ）。

原文：与第7章所处理的方程相比，第一个不同之处在于：这里不存在一个对未知数大小施加限制的大数 N 。因此我们必须自行给变量规定取值范围，而最自然的做法是选取一个大数 P ，对每个 $1 \leq j \leq s$ 定义 $P_j = \lfloor P/c_j^{1/k} \rfloor$ ，并考虑 (8.1) 满足下述条件的解的个数：

$$1 \leq x_j \leq P_j, \quad (1 \leq j \leq s). \quad (8.2)$$

这段在讲什么：第7章解的是 $c_1 x_1^k + \dots + c_s x_s^k = N$ （右边一个给定大数 N ）。那里 N 自动限制了变量大小：每个 $x_j^k \leq N/c_j$ ，所以 x_j 不会太大，"解的个数"天然有限。

可本章右边是 0，没有 N 来当"天花板"。如果不加限制，因为方程是齐次的（见下），只要有一个解，把它整体放大就有无穷多解，"数个数"就失去意义。所以我们必须自己造一个天花板。

为什么方程是"齐次的"、放大解还是解 方程 $c_1x_1^k + \cdots + c_sx_s^k = 0$ 是 k 次齐次方程：把所有 x_j 同时乘以同一个正整数 t ，左边变成 $\sum c_j(tx_j)^k = t^k \sum c_jx_j^k = t^k \cdot 0 = 0$ ，仍然是 0。所以 $(x_1, \dots, x_s)$ 是解 $\Rightarrow (tx_1, \dots, tx_s)$ 也是解。这正是"放大无穷多解"的来源——也正因如此，证明"无穷多解"其实只需找到一个非平凡解；但圆法给的更多：它给出在盒子里解数的精确阶。

造天花板的做法（把每一步说清）：

1 选一个大数 P （之后令 $P \rightarrow \infty$ ）。 P 就是我们人为设的"尺度"。

2 对第 j 个变量，定义上界 $P_j = [P/c_j^{1/k}]$ 。这里 $c_j^{1/k}$ 是 c_j 的 k 次方根， $[\cdot]$ 是取整函数（高斯括号）， $[y]$ 表示不超过 y 的最大整数（例如 $[3.7] = 3$ ）。所以 P_j 是不超过 $P/c_j^{1/k}$ 的最大整数。

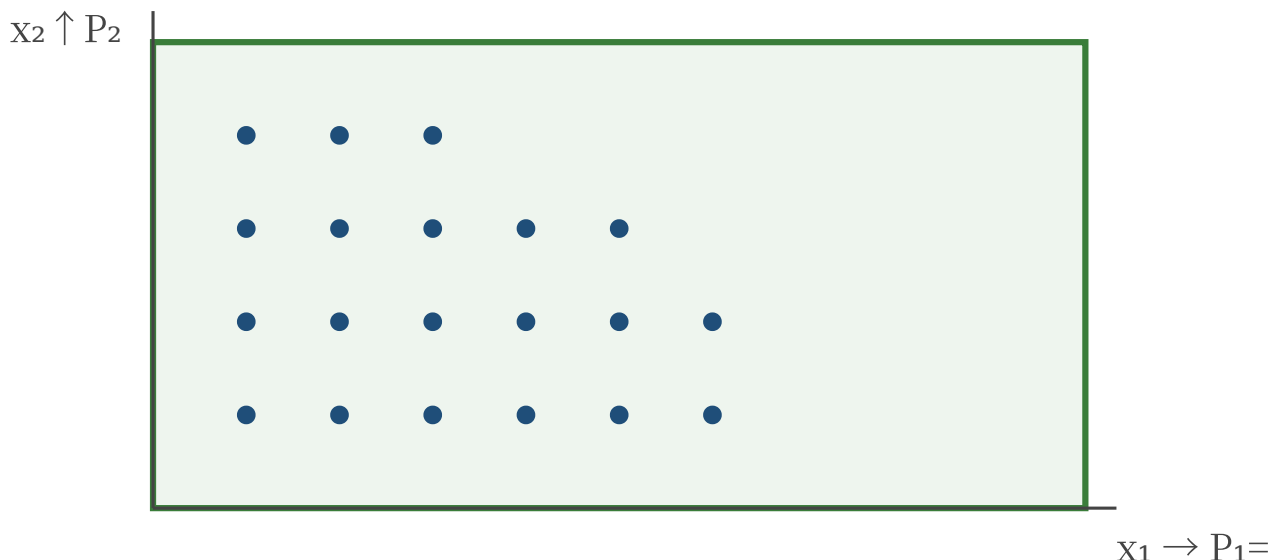
3 为什么这样取 P_j ？目的是让每一项 $c_jx_j^k$ 的大小都被同一个数 P^k 控制住，从而各项"量级齐平"。验证：当 $1 \leq x_j \leq P_j \leq P/c_j^{1/k}$ 时，

$$c_jx_j^k \leq c_jP_j^k \leq c_j\left(\frac{P}{c_j^{1/k}}\right)^k = c_j \cdot \frac{P^k}{c_j} = P^k.$$

于是无论 j 是谁， $c_jx_j^k$ 都不超过 P^k 。这保证正项之和与负项之和量级相当，方程才"配得上"等于 0。若不除以 $c_j^{1/k}$ ，系数大的那一项就会偏大，盒子形状失衡。

4 于是我们改问一个有限的问题：满足 (8.2)（即每个变量都在 1 到 P_j 之间）的解有多少组？把这个数记作 $\mathcal{N}(P)$ ，研究 $P \rightarrow \infty$ 时它的增长。

人为设定的"盒子" $1 \leq x_j \leq P_j$



没有 N 当天花板，我们就用大数 P 自己搭一个盒子，数盒子里的整点解。每个 P_j 按 $c_j^{-1/k}$ 缩放，使各项量级齐平。

原文：我们像之前在 (7.1) 中那样定义指数和 $T_j(\alpha)$ 。于是 (8.1) 在条件 (8.2) 下的解数 $\mathcal{N}(P)$ 由下式给出

$$\mathcal{N}(P) = \int_0^1 T_1(\alpha) \cdots T_r(\alpha) T_{r+1}(-\alpha) \cdots T_s(-\alpha) d\alpha.$$

这段是全章的"翻译总开关"：把"数方程的解"变成"算一个积分"。下面把指数和的定义、以及这个积分公式的来历，从零讲清。

指数和 $T_j(\alpha)$ 的定义（沿用 (7.1)）

$$T_j(\alpha) = \sum_{x=1}^{P_j} e(c_j \alpha x^k).$$

读作"对第 j 个变量的指数和"。它把 $c_j x^k$ (x 跑遍 1 到 P_j) 放到单位圆上、再加起来，是一个依赖 α 的复数。把系数 c_j 放进指数里，是为了让乘起来后恰好出现 $\sum_j c_j x_j^k$ 这个组合。

现在解释那个积分公式。回忆"道具 1"的**挑零性质**： $\int_0^1 e(m\alpha) d\alpha$ 在 $m = 0$ 时为 1，否则为 0。把 s 个指数和按"前 r 个用 α 、后 $s - r$ 个用 $-\alpha$ "乘起来再积分，一步步看会发生什么：

1 把乘积展开。 一个和乘一个和，等于"逐项配对相乘再全部相加"。所以

$$T_1(\alpha) \cdots T_r(\alpha) T_{r+1}(-\alpha) \cdots T_s(-\alpha) = \sum_{x_1=1}^{P_1} \cdots \sum_{x_s=1}^{P_s} \prod_{j=1}^r e(c_j \alpha x_j^k) \prod_{j=r+1}^s e$$

也就是：遍历盒子里每一组 $(x_1, \dots, x_s)$ ，给它配一个复指数。

2 合并指数。 同底数 e 相乘 = 指数相加（因为

$e(a)e(b) = e^{2\pi i a} e^{2\pi i b} = e^{2\pi i (a+b)} = e(a+b)$ ）。于是每一组 $(x_1, \dots, x_s)$ 对应的复指数是

$$e\left(\alpha(c_1 x_1^k + \cdots + c_r x_r^k - c_{r+1} x_{r+1}^k - \cdots - c_s x_s^k)\right) = e(\alpha m),$$

其中 $m = c_1 x_1^k + \cdots - c_s x_s^k$ 正是方程 (8.1) 的**左边**。

3 逐项积分。 把求和与积分交换次序（有限和，可以放心交换）：

$$\int_0^1 (\cdots) d\alpha = \sum_{x_1=1}^{P_1} \cdots \sum_{x_s=1}^{P_s} \int_0^1 e(\alpha m) d\alpha.$$

4 挑零。 由道具 1，内层积分 $\int_0^1 e(\alpha m) d\alpha$ 在 $m = 0$ 时给 1，在 $m \neq 0$ 时给 0。也就是说：恰好是 (8.1) 左边等于 0 的那些 $(x_1, \dots, x_s)$ 各贡献 1，其余贡献 0。

5 结论。 把这些 1 加起来，正是"盒子里使方程成立的解的组数"，即 $\mathcal{N}(P)$ 。这就证明了

$$\mathcal{N}(P) = \int_0^1 T_1(\alpha) \cdots T_r(\alpha) T_{r+1}(-\alpha) \cdots T_s(-\alpha) d\alpha.$$

每组 $(x_1, \dots, x_s) \rightarrow$ 指数 $e(am)$, $m =$ 方程左边

$m=0$

$m \neq 0$

$\rightarrow$ 只数到"方程成立"的解
总和 $= N(P)$

$$\int_0^1 e(am) d\alpha = 1 \quad \int_0^1 e(am) d\alpha = 0$$

积分 $\int_0^1$ 像一台"筛子": 把方程左边非零的项全部抹成 0, 只留下等号成立的解, 各计 1。

二、套用华林问题的处理：主项的形成

原文：我们再次沿用对华林问题的处理方式，作出与前一节相同的细微改动。唯一进一步的改动来自奇异级数与奇异积分中 N 的缺失。在 (4.10) 中，我们须把 $J(P^\delta)$ 替换为

$$(c_1 \cdots c_s)^{-1/k} \int_{|\gamma| < P^\delta} \left(\prod_{j=1}^s \int_0^1 e(\pm \gamma \xi_j^k) d\xi_j \right) d\gamma.$$

这段在讲什么："剩下的活前面都干过了"。整个圆法的骨架——把 $[0, 1]$ 切成主弧 (major arcs) 和次弧 (minor arcs)、在次弧上用 Weyl/华罗庚不等式证明贡献很小、在主弧上把指数和近似成"奇异积分 $\times$ 奇异级数"——在第 4-7 章对华林问题 ($N = x_1^k + \cdots$) 已经搭好了。本章方程几乎一模一样，只有一处**实质区别**：右边的 N 没了。所以作者说"只复述区别"。

为什么要分主弧/次弧 (动机) 积分 $\int_0^1$ 里的被积函数 $T_1 \cdots T_s$ 在某些 α 处特别大 (当 α 接近"分母小的分数" a/q 时, 许多 $e(c_j \alpha x^k)$ 几乎同相位、叠加得很大), 在其余 α 处则因相位杂乱而相互抵消、很小。把"分母小的分数附近的小区间"叫主弧 $\mathfrak{M}$

(信号)，其余叫**次弧** m (噪声)。主弧给出主项，次弧给出可忽略的误差。这就是"圆法"的战术核心。本章直接引用这套战术的结论。

$J(P^\delta)$

第 4 章 (4.10) 里出现的"奇异积分的截断形式"。它本是华林问题里把主弧上的近似积分攒起来的产物，原来含有参数 N 。

替换式中的各记号

$(c_1 \cdots c_s)^{-1/k}$: 把每个变量从 x_j 缩放成连续变量 ξ_j 时，因为 $P_j \approx P c_j^{-1/k}$ ，每个 j 贡献一个 $c_j^{-1/k}$ 因子，乘起来就是 $(c_1 \cdots c_s)^{-1/k}$ 。

γ : 主弧上把 α 写成 $\alpha = \frac{a}{q} + \frac{\beta}{P^k}$ 后，再做尺度变换得到的"局部连续变量"，相当

于在每段主弧内部的精细坐标。 $|\gamma| \xi_j \in [0, 1]$: 把离散的

$x_j \in \{1, \dots, P_j\}$ 归一化成连续变量 $\xi_j = x_j/P_j$ 后的积分变量。

$\pm$: 见下一段，前 r 个取 $+$ 、后 $s-r$ 个取 $-$ ，对应正负项。

"唯一的改动来自 N 的缺失"具体指什么？在华林问题里被积的复指数还带着 $e(-N\alpha)$ 这种"目标项"，对应到连续积分里会出现 $e(\mp\gamma)$ 之类的因子，使奇异积分依赖 N 。这里方程右边是 0，那个因子变成 $e(0) = 1$ ，于是奇异积分**不再依赖任何目标值**，纯粹是系数与 k 的函数——这一点稍后会再次强调 ($\mathfrak{S}$ 也同理)。

原文：其中符号在 $j \leq r$ 时取 $+$ ，在 $j > r$ 时取 $-$ 。如同 (4.16) 那样，我们被引向对下述积分的求值

$$J = \int_{-\infty}^{\infty} k^{-s} \left(\prod_{j=1}^s \int_0^1 \zeta_j^{-1+1/k} e(\pm\gamma\zeta_j) d\zeta_j \right) d\gamma.$$

这段把上一段的截断积分推到极限（截断范围 $P^\delta \rightarrow \infty$ 、内层积分换元），得到一个干净的、与 P 无关的常数积分 J 。逐项解释：

- 符号 $\pm$ 的规则：** $j \leq r$ (正项) 取 $+$ ， $j > r$ (负项) 取 $-$ 。它直接来自 (8.1) 中各项的符号。

2 内层换元 $\xi_j^k = \zeta_j$ 。这是关键一步，把" k 次方"拉直成线性。令 $\zeta_j = \xi_j^k$ ，当 ξ_j 从 0 到 1， ζ_j 也从 0 到 1。微分： $d\zeta_j = k\xi_j^{k-1}d\xi_j$ ，所以
 $d\xi_j = \frac{1}{k}\zeta_j^{(1-k)/k}d\zeta_j = \frac{1}{k}\zeta_j^{-1+1/k}d\zeta_j$ 。这正是积分里冒出 $\zeta_j^{-1+1/k}$ 和每项一个 $\frac{1}{k}$ 的原因； s 个 $\frac{1}{k}$ 乘起来就是前面的 k^{-s} 。

3 截断推到无穷：把 γ 纯粹的常数（只依赖 k, s 和符号分布），不再含 P 。

为什么要换元 $\xi^k = \zeta$ (动机) 原来内层积分 $\int_0^1 e(\pm\gamma\xi^k)d\xi$ 里 ξ 以 k 次方出现，难处理。换成 $\zeta = \xi^k$ 后，指数里变成 $\pm\gamma\zeta$ （线性！），代价是冒出一个权重 $\zeta^{-1+1/k}$ 。线性指数的积分我们会算，这笔交易划算。这正是第 4 章定理 4.1 处理奇异积分的标准手法。

原文：正如定理 4.1 的证明中那样，我们把变量从 ζ_s 换成 u ，其中

$$\zeta_1 + \cdots + \zeta_r - \zeta_{r+1} - \cdots - \zeta_s = u,$$

于是我们求得

$$J = k^{-s} \int_0^1 \cdots \int_0^1 \{\zeta_1 \cdots \zeta_s (\zeta_1 + \cdots \pm \zeta_{s-1})\}^{-1+1/k} d\zeta_1 \cdots d\zeta_{s-1},$$

其中 $0 < \zeta_1 + \cdots \pm \zeta_{s-1} < 1$ 。

这段在算 J 这个常数积分，目标只是看出它是正数。思路：先对 γ 积分，用到的是"傅里叶反演 / 挑零"的连续版本。把每一步补全：

1 换变量 $\zeta_s \rightarrow u$ 。固定 $\zeta_1, \dots, \zeta_{s-1}$ ，令

$$u = \zeta_1 + \cdots + \zeta_r - \zeta_{r+1} - \cdots - \zeta_s.$$

这是 u 与 ζ_s 的一次关系（系数为 -1 ），所以 $d\zeta_s = \mp du$ （雅可比行列式绝对值为 1），换元不带额外因子。 u 正是"把所有 ζ 按方程符号组合起来"的那个量——它对应方程左边。

- 2 对 γ 积分先做。把所有含 γ 的因子收集起来：每个内层积分给出 $e(\pm\gamma\zeta_j)$ ，乘起来再对 ζ_j 之外先不动、对 γ 积分，得到 $\int_{-\infty}^{\infty} e(\gamma u) d\gamma$ 形式的因子（把各 $\pm\gamma\zeta_j$ 合并，指数里 γ 的系数恰为 u ）。这是狄拉克 δ 的连续挑零：

$\int_{-\infty}^{\infty} e(\gamma u) d\gamma$ 形式上"挑出 $u = 0$ "，即强制 $\zeta_1 + \cdots \pm \zeta_s = 0$ ，也就是把 ζ_s 解出为 $\zeta_s = \zeta_1 + \cdots \pm \zeta_{s-1}$ （前 $r-?$ 的符号见原式），从而把 s 重积分降为 $s-1$ 重。

- 3 代回 ζ_s 。被积的权重原来是 $\prod_{j=1}^s \zeta_j^{-1+1/k}$ 。把 ζ_s 替换成 $\zeta_1 + \cdots \pm \zeta_{s-1}$ ，就得到

$$\{\zeta_1 \cdots \zeta_{s-1} \cdot (\zeta_1 + \cdots \pm \zeta_{s-1})\}^{-1+1/k} = \{\zeta_1 \cdots \zeta_s (\zeta_1 + \cdots \pm \zeta_{s-1})\}^{-1+1/k}$$

（最后一个写法把 ζ_s 也写出来便于对称地看）。积分范围由 $0 < \zeta_j < 1$ （每个连续变量）以及 $\zeta_s \in (0, 1)$ 即 $0 < \zeta_1 + \cdots \pm \zeta_{s-1} < 1$ 共同决定。

- 4 整理得到所写的 $(s-1)$ 重积分表达式，前面带常数 k^{-s} 。

为什么作者只关心"是否对 s 个 ζ 中某点 > 0 "，而不去精确算 J 回忆本章的真正目标：证明解有无穷多个，只需主项系数为正。精确的 J 值并不影响"无穷多解"这一结论，只影响主项的具体大小。所以作者明确说"我们所需知道的只是 $J > 0$ "。这与第 3 章导读中"我们要的是赛跑胜负、不是终点读数"的精神一致。

原文：我们所需知道的只是 $J > 0$ ，而这一点成立，是因为存在某个含于 $0 < \zeta_j < 1$ （对所有 j ）内的开集，在其上有

$$0 < \zeta_1 + \cdots \pm \zeta_{s-1} < 1.$$

把 " $J > 0$ " 补成完整论证：

- 1 被积函数 $\{\zeta_1 \cdots \zeta_s (\zeta_1 + \cdots \pm \zeta_{s-1})\}^{-1+1/k}$ 在积分区域内部处处为正：因为每个因子 $\zeta_j > 0$ 、组合 $\zeta_1 + \cdots \pm \zeta_{s-1}$ 在 $(0, 1)$ 内也为正，正数的任意实数次幂仍为正。

2 一个处处为正的连续函数，在一个**非空开集**上的积分必然 > 0 （正函数积出来不可能是 0）。

3 因此只要"满足约束 $0 < \zeta_1 + \cdots \pm \zeta_{s-1} < 1$ 且各 $0 < \zeta_j < 1$ "的区域**非空且有内点（开集）**，就有 $J > 0$ 。

4 这样的**开集确实存在**：例如取所有 ζ_j 都接近 $\frac{1}{2s}$ 这样的小正数，则正项和减负项和是一个小的正数，落在 $(0, 1)$ 内；这是一个开条件，附近一整块都满足。于是开集非空， $J > 0$ 成立。

这就完成了**奇异积分为正**（第一个因子）的验证。注意它本质上反映"方程在实数范围有一片连续解（一个 $(s-1)$ 维曲面），且我们的盒子切到了这片解"。

原文：这样，对 $s \geq 2^k + 1$ 已证明的 $\mathcal{N}(P)$ 的渐近公式取下述形式

$$\mathcal{N}(P) = \frac{C'_{k,s}}{(c_1 \cdots c_s)^{1/k}} P^{s-k} \mathfrak{S} + O(P^{s-k-\delta}), \quad (8.3)$$

其中

$$C'_{k,s} = k^{-s} \int_0^1 \cdots \int_0^1 \{\eta_1 \cdots \eta_s (\eta_1 + \cdots - \eta_{s-1})\}^{-1+1/k} d\eta_1 \cdots d\eta_{s-1}, \quad (8.4)$$

且

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^q q^{-s} S_{c_1 a, q} \cdots S_{-c_s a, q}. \quad (8.5)$$

这是本章的核心公式，把整个圆法的产出"装订"成一行。逐件拆解：

$\mathcal{N}(P)$

盒子 (8.2) 里方程 (8.1) 的解数（我们真正想知道的量）。

P^{s-k}

主项的阶。直觉： s 个变量各有约 P 种取值，共约 P^s 组；方程是一个 k 次约束，"命中 0" 的概率约 P^{-k} ；相乘得 P^{s-k} 。只要 $s > k$ ，这个指数为正， $P \rightarrow \infty$ 时主项 $\rightarrow \infty$ 。

$C'_{k,s}$ (奇异积分, (8.4))

就是上一节算出的常数 J (这里把哑变量改记为 η ，并且原文 (8.4) 把符号写成具体的 $+\cdots-$ 形态)。我们已证 $C'_{k,s} = J > 0$ 。它编码"实数 (无穷远) 处的几何"。

$$(c_1 \cdots c_s)^{-1/k}$$

缩放盒子带来的常数因子 (前面解释过)。它是正的。

$\mathfrak{S}$ (奇异级数, (8.5))

编码"每个素数处的算术"。它是否为正，是本章下半场要攻克的难点。 $\mathfrak{S}$ 这个花体 S 读作"奇异级数 S "。

$$O(P^{s-k-\delta})$$

误差项，阶比主项小 $P^{-\delta}$ ($\delta > 0$ 固定)。当 $P \rightarrow \infty$ ，它相对主项可忽略。

条件 $s \geq 2^k + 1$

这是次弧估计 (Weyl/华罗庚不等式那一套，第 3-4 章) 要求的变量数下限，用来保证误差项确实小于主项。本章把它当作已证前提引用。

奇异级数 (8.5) 里的每个符号 $S_{a,q}$ (高斯型和 / 完全指数和) 定义为

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

即把 x 跑遍模 q 的一组完全剩余系、把 $e(ax^k/q)$ 加起来。它度量" k 次方在模 q 下的分布"。

所以 $S_{c_j a, q} = \sum_{x=1}^q e(c_j a x^k / q)$ ，把系数 c_j 放进去；负项写成 $S_{-c_j a, q}$ 。

外层双重求和： q 从 1 到 ∞ (所有模)，内层 a 跑遍 $1 \leq a \leq q$ 中与 q 互素的那些 ($(a, q) = 1$ ，对应"最简分数 a/q "即各条主弧的中心)。 q^{-s} 是归一化权重。

直观： $\mathfrak{S}$ 把"每个分母 q 下、方程同余可解的程度"加权累加，得到一个衡量"方程在所有素数处算术障碍"的总指标。

$$\boxed{\begin{array}{c} N(P) \text{ 解数} \\ \text{(想求)} \end{array}} \approx \boxed{\begin{array}{c} C'_{k,s} > 0 \\ \text{奇异积分} \end{array}} \times \boxed{\begin{array}{c} \mathfrak{S} > 0 \\ \text{奇异级数} \end{array}} \times \boxed{P^{s-k}}$$

两个因子都为正 $\implies N(P) \rightarrow \infty \implies$ 无穷多解

主项 = 奇异积分 $\times$ 奇异级数 $\times P^{s-k}$ 。前者已证为正，下半章专攻后者。

原文：我们注意到，现在 $\mathfrak{S}$ 的值是一个仅依赖于系数 $c_1, \dots, -c_s$ 以及 k 的数。如前所述，定义 $\mathfrak{S}$ 的级数在 $s \geq 2k + 1$ 时绝对收敛，并可因子分解为 $\prod \chi(p)$ 。同样存在一个 p_0 使得

$$\prod_{p > p_0} \chi(p) \geq \frac{1}{2}.$$

这段讲奇异级数的三条性质，每条都补全。

(1) " $\mathfrak{S}$ 只依赖系数与 k ，不依赖别的"。这正是" N 缺失"的红利：第 7 章里奇异级数还含 N （要写成 $\mathfrak{S}(N)$ ，对不同 N 不同），这里右边为 0，求和式里没有任何依赖 N 的因子，所以 $\mathfrak{S}$ 就是一个固定的常数。这让我们的目标从"对所有 N 一致地有 $\mathfrak{S}(N) \geq C_1 > 0$ "简化为"这一个常数 $\mathfrak{S} > 0$ "。

(2) " $s \geq 2k + 1$ 时绝对收敛"。绝对收敛指 $\sum |\text{项}|$ 也有限，从而可以随意重排、拆分。这依赖一个估计（Weyl 估计的推论，第 4 章）： $|S_{a,q}| \leq C q^{1-1/k}$ 。于是每个 a 项的模 $\leq q^{-s} \prod_j C q^{1-1/k} = C^s q^{-s+s(1-1/k)} = C^s q^{-s/k}$ ；对 a 求和（至多 q 个）再对 q 求和，量级是 $\sum_q q \cdot q^{-s/k} = \sum_q q^{1-s/k}$ 。当 $1 - s/k < -1$ 即 $s > 2k$ （即 $s \geq 2k + 1$ ）时，这是收敛的 p -级数，故绝对收敛。

"因子分解为 $\prod \chi(p)$ "是什么意思（动机） $S_{a,q}$ 关于 q 具有积性（当 $q = q_1 q_2$ 且 q_1, q_2 互素时可拆成乘积，源于中国剩余定理）。绝对收敛的级数因此能按素数 p 独立地分解：

$$\mathfrak{S} = \prod_{p \text{ 素数}} \chi(p), \quad \chi(p) = \sum_{t=0}^{\infty} A(p^t),$$

其中 $A(q)$ 是 (8.5) 中固定模 q 的那一层（对互素的 a 求和）。每个 $\chi(p)$ 叫做" **p -adic 局部密度（局部因子）**"，它只关心方程在模 p 的各次幂下的解的多寡。

为什么这是好事？因为**无穷级数 > 0 难证，但"每个局部因子 > 0 且整体不退化"可以逐素数地验证**——这把一个分析问题拆成了无穷多个独立的、初等的同余计数问题。这正是下半章要做的事。

(3) "**存在 p_0 使 $\prod_{p>p_0} \chi(p) \geq \frac{1}{2}$** "。补全：由绝对收敛， $\prod_p \chi(p)$ 收敛，意味着对大的素数 p ， $\chi(p)$ 极接近 1（ $|\chi(p) - 1|$ 之和有限）。于是"尾巴"乘积 $\prod_{p>p_0} \chi(p)$ 随 p_0 增大趋于 1；取 p_0 足够大，就能让这个尾巴乘积 $\geq \frac{1}{2}$ （任何 < 1 的正下界都行， $\frac{1}{2}$ 只是方便）。**结论：大素数处的因子整体上至少贡献 $\frac{1}{2}$ ，绝不会把 $\mathfrak{S}$ 拉到 0。**剩下要担心的只有有限个小素数 $p \leq p_0$ 。

原文：为了确保 $\mathfrak{S} > 0$ （现在无需再写成 $\mathfrak{S} \geq C_1(k, s) > 0$ ，因为不再有参数 N ），只需对每个单独的 p 证明 $\chi_p > 0$ 即可。如前所述，只要满足下式即足够

$$M(p^\nu) \geq C_p p^{\nu(s-1)} \quad (8.6)$$

对所有充分大的 ν 成立，这里 $M(p^\nu)$ 表示同余式

$$c_1 x_1^k + \cdots - c_s x_s^k \equiv 0 \pmod{p^\nu}, \quad 0 \leq x < p^\nu \quad (8.7)$$

的解的总数。

这段把"证 $\mathfrak{S} > 0$ "归结为一个干净的同余计数下界。逐步说明这个归结链条：

1 从乘积到逐因子。 $\mathfrak{S} = \prod_p \chi(p)$ 。一个无穷乘积若每个因子 > 0 、且尾巴乘积有正下界（上一段的 $\frac{1}{2}$ ），则整体 > 0 。所以只需对每个素数 p 单独证 $\chi(p) > 0$ （原文记作 χ_p ）。

2 $\chi(p)$ 是什么。 它等于"方程在 p -adic 整数里的解密度"，有个具体公式：

$$\chi(p) = \lim_{\nu \rightarrow \infty} \frac{M(p^\nu)}{p^{\nu(s-1)}},$$

其中 $M(p^\nu)$ 是同余式 (8.7) 在 $\{0 \leq x_j\}$ 组数。

3 为什么分母是 $p^{\nu(s-1)}$ 。共有 s 个变量，每个有 p^ν 种取值，全空间 $p^{\nu s}$ 组；方程是一个模 p^ν 的约束，"随机命中 0" 的概率约 $p^{-\nu}$ ，期望解数约 $p^{\nu s} \cdot p^{-\nu} = p^{\nu(s-1)}$ 。所以 $M(p^\nu)/p^{\nu(s-1)}$ 是"实际解数相对随机期望的比值"，它的极限就是局部密度 $\chi(p)$ 。

4 (8.6) 的含义。若能证 $M(p^\nu) \geq C_p p^{\nu(s-1)}$ ($C_p > 0$ 是只依赖 p 的正常数) 对所有充分大的 ν 成立，那么取极限得 $\chi(p) \geq C_p > 0$ 。这就是要证 $\chi(p) > 0$ 的充分条件——把"密度为正"翻译成"解数有一个正比例下界"。

5 "无需写 $\mathfrak{S} \geq C_1(k, s) > 0$ "。第 7 章因为有 N ，需要对所有 N 一致的正下界 C_1 ；这里 $\mathfrak{S}$ 是单个常数，只需它 > 0 即可，门槛更低。

另外解释 (8.7)：把方程系数恢复成"有正有负"无所谓——下一段会说符号不影响计数；约束 $\{0 \leq x_j\}$

三、把同余问题层层化简，找出 $s_1(k)$

原文：我们现在的目标是求出 k 的某个明确的函数 $s_1(k)$ ，使得对每个 p ，只要 $s \geq s_1(k)$ ，(8.6) 即成立。于是渐近公式 (8.3) 将是有意义的，因为其中主项的阶为 P^{s-k} ，从而当 $P \rightarrow \infty$ 时 $\mathcal{N}(P) \rightarrow \infty$ 。在证明这一结果时，(8.7) 中系数的符号不起任何作用，因此我们回到原始记号，即系数前不带负号的记号。

这段宣布下半章的任务：找一个显式的变量数门槛 $s_1(k)$ （只依赖 k ），使得只要 $s \geq s_1(k)$ ，对每个素数 p 同余下界 (8.6) 都成立。一旦如此， $\chi(p) > 0$ （每个 p ） $\implies \mathfrak{S} > 0 \implies$ 主项 $\neq 0 \implies \mathcal{N}(P) \sim (\text{正常数}) P^{s-k} \rightarrow \infty \implies$ 解有无穷多个。

"符号不起作用"补全：在 (8.7) 的同余计数里，把某个 $-c_j$ 换回 $+c_j$ ，只是把变量 x_j 替换为另一套取值（模 p^ν 下 -1 是可逆的），解的**总数不变**。所以计数时可以把所有负号丢掉，统一写成 $c_1 x_1^k + \cdots + c_s x_s^k \equiv 0$ ，系数都为正。下面就用这套"无负号"的记号。

原文：第一步是导出一个未知数个数较少、且其中没有任何系数能被 p 整除的同余式，使得若 (8.6) 对这个新同余式成立，则它对原同余式也成立。我们记

$$c_j = d_j p^{h_j k + l_j}, \quad (1 \leq j \leq s),$$

其中

$$p \nmid d_j, \quad 0 \leq l_j < k.$$

这段开始"清洗系数"。问题在于：有些 c_j 可能被 p 整除（含 p 的因子），这会干扰同余计数。第一步是把每个 c_j 里的 p 因子剥离出来，化到一个"系数都不被 p 整除、变量更少"的同余式，并保证下界 (8.6) 能从新式子传回原式子。

把 c_j 按 p 的幂分解 任何正整数 c_j 都能唯一写成

$$c_j = d_j p^{e_j}, \quad p \nmid d_j,$$

其中 $e_j \geq 0$ 是 c_j 含 p 的精确次数， d_j 是去掉 p 后剩下的、与 p 互素的部分。

再把指数 e_j 按 k 做带余除法： $e_j = h_j k + l_j$ ，其中 $h_j = [e_j/k]$ 是商， l_j 是余数， $0 \leq l_j < k$ 。
 动机： $p^{h_j k} = (p^{h_j})^k$ 是一个完全 k 次方，可以"吸进变量里"（见下一段）；剩下的 p^{l_j} （指数 $\leq k-1$ ）

原文：于是 (8.7) 变为

$$\sum_{j=1}^s d_j p^{l_j} (p^{h_j} x_j)^k \equiv 0 \pmod{p^\nu}.$$

把分解代入即可，逐步看：

1 原项 $c_j x_j^k = d_j p^{h_j k + l_j} x_j^k = d_j p^{l_j} \cdot p^{h_j k} x_j^k$ 。

2 把 $p^{h_j k}$ 写成 $(p^{h_j})^k$ ，并与 x_j^k 合并： $p^{h_j k} x_j^k = (p^{h_j} x_j)^k$ 。

3 于是 $c_j x_j^k = d_j p^{l_j} (p^{h_j} x_j)^k$ ，代入即得所写同余式。注意：现在 k 次方里的"变量"变成了 $p^{h_j} x_j$ ，系数前缀只剩 $d_j p^{l_j}$ （其中 d_j 与 p 互素、 $0 \leq l_j < k$ ）。

原文：令 $h = \max h_j$ 。我们将自己限于下述形式的解

$$x_j = p^{h-h_j} y_j.$$

这样，对于大的 ν ，我们可以从同余式中消去 p^{hk} ，于是它变成

$$\sum_{j=1}^s d_j p^{l_j} y_j^k \equiv 0 \pmod{p^{\nu-hk}}, \quad (8.8)$$

受下述约束

$$0 \leq y < p^{\nu-h+h_j}.$$

这段把各项的 $p^{h_j k}$ 因子统一拉平，再约掉。策略：只数一部分解（把 x_j 限制成特定形式），这只会让计数变小，所以得到的下界更保守、但仍然够用（因为我们只要"下界"）。每步补全：

1 取 $h = \max_j h_j$ （所有 h_j 里最大的）。这是为了让接下来每一项都能凑出 p^{hk} 。

2 限制解的形式 $x_j = p^{h-h_j} y_j$ 。因为 $h \geq h_j$ ，指数 $h - h_j \geq 0$ ，这是合法的（ x_j 取 p^{h-h_j} 的倍数）。这是"只看子集解"——故计数变小，但传出去仍是下界。

3 代入并约去 p^{hk} 。把 $x_j = p^{h-h_j} y_j$ 代入第 j 项的 $(p^{h_j} x_j)^k$ ：

$$(p^{h_j} \cdot p^{h-h_j} y_j)^k = (p^h y_j)^k = p^{hk} y_j^k.$$

于是每一项都带着公共因子 p^{hk} ：第 j 项 = $d_j p^{l_j} \cdot p^{hk} y_j^k$ 。整个同余式两边都能被 p^{hk} 整除（当 ν 足够大， $\nu \geq hk$ ），可以同除 p^{hk} （模数也相应降为 $p^{\nu-hk}$ ）。结果就是 (8.8)： $\sum_j d_j p^{l_j} y_j^k \equiv 0 \pmod{p^{\nu-hk}}$ 。

4 新变量 y_j 的范围。由 $x_j = p^{h-h_j} y_j$ 且 $0 \leq x_j$

为什么允许"只数子集解"（取舍说明） 我们追求的是 (8.6) 那种下界

$M(p^\nu) \geq C_p p^{\nu(s-1)}$ 。把解限制成 $x_j = p^{h-h_j} y_j$ 这种特殊形式，只会数到原来解的

一部分，得到的计数 $\leq M(p^\nu)$ 。但如果连这个子集都已经 $\geq C_p p^{\nu(s-1)}$ 那么多，原来的 $M(p^\nu)$ 当然更多。所以"少数一点"不影响证下界，反而把系数清洗干净，是划算的取舍。

原文：如果我们用 $M'(p^{\nu-hk})$ 表示 (8.8) 满足

$$0 \leq y < p^{\nu-hk}$$

的解的个数，那么（因为 $h - h_j < hk$ ）我们有

$$M(p^\nu) \geq M'(p^{\nu-hk}).$$

因此只需对 $M'(p^\nu)$ 证明 (8.6) 的类比即可。

这段把变量范围也统一，建立 M 与 M' 的不等式。补全" $(h-h_j$

1 定义 $M'(p^{\nu-hk})$ 为 (8.8) 在更整齐的范围 $(0 \leq y_j$

2 真正从 x_j 翻译过来的范围是 $(0 \leq y_j \leq \nu - hk)$ ，所以 $p^{\nu-hk} \leq p^{\nu-h+h_j}$ ：即 M' 的范围更小或相等，被 M 实际允许的范围包含。

3 因此 M' 数的解是 M 数的解的子集（" y 范围更窄、形式更特殊"），故

$$M(p^\nu) \geq M'(p^{\nu-hk}).$$

4 结论：只要对 M' （用整齐范围 $p^{\nu-hk}$ ，再把哑变量 $\nu - hk$ 重新记成 ν ）证出 (8.6) 型下界，就传回给 $M(p^\nu)$ ，从而 $\chi(p) > 0$ 。问题成功简化为研究系数都不被 p 整除的同余式 (8.8)。

(8.7) 原同余式
系数可被 p 整除

→

(8.8) 约掉 p^{hk}
系数 $d_j p^{l_j}$, $p \nmid d_j$

→

(8.9)/(8.10) 聚成一组
 v 项系数都不整除 p

每一步只数"子集解", 得到的下界更保守但够用

三步化简: 剥 p 的幂 → 拉平 → 按 l_j 归并。最终只需对一个"干净"的同余式找到一个不平凡解。

原文: 令 $l = \max l_j$ 。在新的同余式 (8.8) 中 (但取模 p^ν)，我们按 l_j 的值把各项归并到一起。共有 k 组，其中至少有一组必含有 v 项，这里 $v \geq s/k$ 。我们在其余项中令 $y_j = p y'_j$ ，并在除去一个因子 p^l 之后，得到下述形式的同余式

$$d_1 y_1^k + \cdots + d_v y_v^k + p(d_{v+1} y_{v+1}^k + \cdots) + \cdots \equiv 0 \pmod{p^{\nu-l}}. \quad (8.9)$$

这段用抽屉原理 (鸽笼原理) 找出一大组"同型项"。各项的前缀里还残留着 p^{l_j} ($0 \leq l_j$

1 分组。把 (8.8) 的 s 个项按其 l_j 值放进 k 个"抽屉" ($l_j = 0, 1, \dots, k-1$)。

2 抽屉原理。 s 个项放进 k 个抽屉，必有一个抽屉里至少有 $\lceil s/k \rceil$ 项。把这一组 (共 v 项, $v \geq s/k$) 的公共 l_j 值记为 l (这与 $l = \max l_j$ 略有出入，但作者取这一最大、最满的组；关键是这组 p^l 是公共因子，组内系数都形如 $d_j p^l$)。

3 把这 v 项的公共 p^l 提出来。对其余项 (l_j 系数与 p 互素的 $d_1 y_1^k + \cdots + d_v y_v^k$ ；其余项至少带一个公因子 p ，写成 $p(d_{v+1} y_{v+1}^k + \cdots)$ 。模数从 p^ν 降为 $p^{\nu-l}$ 。这就是 (8.9)。

为什么这一步是关键 (动机) (8.9) 的妙处在于：前 v 项的系数 $d_1, \dots, d_v$ 全部与 p 互素，而后面所有项都至少带一个因子 p 。这意味着，模 p 时后面的项全部消失，方程的"主导部分"就是前 v 项。只要能让前 v 项非平凡地凑成 p 的高次幂的倍数 (即解

(8.10)), 后面带 p 的项就可以用"逐次提升 (Hensel 引理式的逐层修正)"补齐。于是问题彻底归结为: **系数都与 p 互素的 v 元 k 次同余式, 能否非平凡求解?**

原文: 我们仍可在右边把 $\nu - l$ 换成 ν , 因为这只是把 (8.6) 这类结果中的 C_p 改变一下。在这最后一个同余式中, 我们有

$$d_1 d_2 \cdots d_v \not\equiv 0 \pmod{p}.$$

两句都补全:

- "把 $\nu - l$ 换成 ν ". 模数差一个固定的 p^l (对"是否存在正下界"毫无影响, 所以可以放心把 $\nu - l$ 写成 ν , 记号更简洁。
- " $d_1 d_2 \cdots d_v \not\equiv 0 \pmod{p}$ ". 因为每个 d_j 都与 p 互素 ($p \nmid d_j$), 它们的乘积也与 p 互素 (素数 p 整除乘积必整除某个因子, 但它一个都不整除), 故乘积 $\not\equiv 0$ 。这再次强调前 v 项系数"干净"。

原文: 像通常那样定义 γ (见 (5.11))。利用引理 5.5 证明中所用的论证, 所求结果 (8.6) 对同余式 (8.9) 将成立, 只要下述同余式

$$d_1 y_1^k + \cdots + d_v y_v^k \equiv 0 \pmod{p^\gamma} \quad (8.10)$$

有一个解, 其中 $y_1, \dots, y_v$ 并非全部能被 p 整除。

这段把"无穷多个 ν 的下界"一举归结为"一个固定模 p^γ 的同余式有非平凡解"。这是 Hensel 提升思想的体现, 需要先讲清 γ 。

γ 的定义 (来自 (5.11)) 与提升原理 先写出 p 在 k 中的精确幂次: 设 $k = p^\tau k_0$, 其中 $p \nmid k_0$, $\tau \geq 0$ 。于是 $\tau = 0$ 当且仅当 $p \nmid k$ 。定义

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

γ 叫做 k 次幂关于 p 的"**提升门槛**"。它的作用是 Hensel (亨泽尔) 引理的精确版本:

提升原理: 若 x_0 满足 $f(x) \equiv 0 \pmod{p^\gamma}$ 且 x_0 关于该方程是"非退化的 ($p \nmid x_0$ 这一类条件)", 则它可以唯一地**逐层修正**, 得到 $\pmod{p^{\gamma+1}, p^{\gamma+2}, \dots}$ 直到任意 p^ν 的解。

直观: γ 是"看清一个 k 次方解所需的最小精度"。一旦在 p^γ 这个精度上看到了一个非

平凡解，更高精度的解就能一路自动接上去，每升一层解数乘以 p^{s-1} ，正好喂给 (8.6) 那个 $p^{\nu(s-1)}$ 的下界。

1 引理 5.5 的论证（要点复述）。它说：如果“主导部分”

$d_1 y_1^k + \cdots + d_v y_v^k \equiv 0 \pmod{p^\gamma}$ 有一个 y 不全被 p 整除的解，那么这个解（连同后面带 p 因子的项一起）可以逐层提升到模 p^ν ，并且每升一层，自由变量带来 p^{s-1} 倍的解，从而 $M'(p^\nu) \geq C_p p^{\nu(s-1)}$ ——正是 (8.6)。

2 “非平凡”为什么必要。全为 0（或全被 p 整除）的解是退化的，提升原理对它失效（导数条件不满足）。所以必须要求 $y_1, \dots, y_v$ 不全被 p 整除，即存在某个 y_j 与 p 互素，这样它对应的偏导 $k d_j y_j^{k-1}$ 在去掉 p^τ 后非零，提升才能进行。

3 归结完成。于是整章的全部困难，浓缩成一句话：对系数都与 p 互素的 v 元 k 次同余式 (8.10)，能否找到一个不全被 p 整除的解？接下来分 $p > 2$ 与 $p = 2$ 两种情况解决，并由此定出需要多少变量 v （进而多少 s ）。

四、求解 (8.10)：分 $p > 2$ 与 $p = 2$

原文：设 $p > 2$ 。我们像定理 7.2 证明中那样论证，把 (8.10) 中各项按系数 d_j 所属的 k 次幂剩余类或非剩余类归并成组。只要满足下式即足够

$$\frac{v}{p^{\gamma-1}\delta} > 2k - 1,$$

又因为 $p^{\gamma-1}\delta = p^\tau \delta \leq k$ ，所以只要满足下式即足够

$$v > k(2k - 1).$$

因此只要满足下式即足够

$$s > k^2(2k - 1).$$

这段处理奇素数 p 。核心是" k 次幂剩余类"的概念和一个抽屉/相消论证。先把概念讲清。

k 次幂剩余、剩余类、 $\delta = (k, p-1)$ 模 p 时, 非零元素 $1, 2, \dots, p-1$ 在乘法下构成一个有 $p-1$ 个元素的循环群。一个非零元 a 叫 k 次幂剩余, 如果存在 x 使 $x^k \equiv a \pmod{p}$ 。

全体 k 次幂剩余构成一个子群, 其元素个数为 $(p-1)/\delta$, 其中

$$\delta = (k, p-1) \quad (\text{即 } k \text{ 与 } p-1 \text{ 的最大公约数}).$$

于是非零元被分成 δ 个" k 次幂剩余类 (陪集)": 同一类里的元素彼此相差一个 k 次方。更一般地, 模 p^γ 时, 单位被分成 $p^{\gamma-1}\delta$ 个 k 次幂剩余类 (因为单位群有 $\varphi(p^\gamma) = p^{\gamma-1}(p-1)$ 个元, k 次方子群指数为 $(k, p^{\gamma-1}(p-1)) = p^{\gamma-1}\delta$, 这里用到 τ 的分析)。

关键不等式 $p^{\gamma-1}\delta = p^\tau \delta \leq k$: 因为 $\gamma-1 = \tau$ (p 奇), 而

$\delta = (k, p-1) = (p^\tau k_0, p-1) = (k_0, p-1) \leq k_0$ (因 $(p^\tau, p-1) = 1$)。于是 $p^\tau \delta \leq p^\tau k_0 = k$ 。这说明剩余类的个数 $p^{\gamma-1}\delta$ 至多 k 个。

现在把 " $v/(p^{\gamma-1}\delta) > 2k-1$ 即足够" 的论证补全 (定理 7.2 的思路):

- 1 把 v 个系数按剩余类分组。系数 $d_1, \dots, d_v$ (都与 p 互素) 落入至多 $p^{\gamma-1}\delta$ 个 k 次幂剩余类。由抽屉原理, 必有一类里至少有

$$\frac{v}{p^{\gamma-1}\delta} \text{ 个系数.}$$

- 2 同一类内可造相消。若 d_i, d_j 同属一个 k 次幂剩余类, 则 d_i/d_j 是一个 k 次方, 存在 t 使 $d_i \equiv d_j t^k$; 于是 $d_i \cdot 1^k + d_j \cdot (\text{适当})^k$ 可以配出该类的一个"标准代表", 多个同类项叠加可凑成 p 的更高次幂的倍数。把每一类内部的项两两/成串地相消、提升, 需要每类大约 $2k-1$ 个项才能保证造出一个非平凡的、模 p^γ 为零的组合 (这正是定理 7.2 中关于 k 次型在 p^γ 处可解性的计数结论)。

- 3 阈值。因此"某一类项数 $> 2k-1$ " 即足以保证 (8.10) 有非平凡解, 即

$$\frac{v}{p^{\gamma-1}\delta} > 2k - 1.$$

- 4 用 $p^{\gamma-1}\delta \leq k$ 放大分母。既然分母 $p^{\gamma-1}\delta \leq k$ ，那么 " $v/k > 2k - 1$ " 是比 " $v/(p^{\gamma-1}\delta) > 2k - 1$ " 更强（更难满足）的条件——满足强的就满足弱的。于是只要

$$v > k(2k - 1)$$

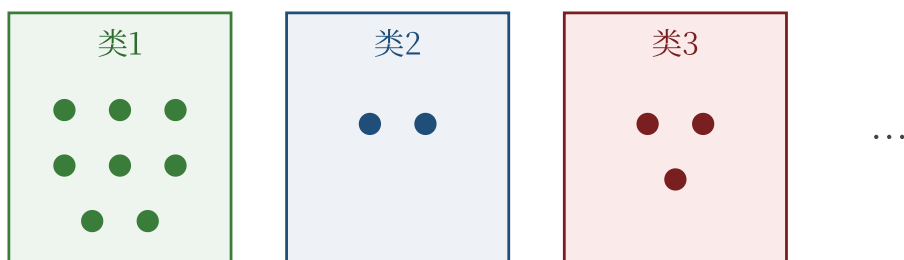
就足够（对一切奇素数 p 一致成立，因为它已经把最坏的 $p^{\gamma-1}\delta = k$ 考虑进去）。

- 5 从 v 回到 s 。回忆抽屉原理给出 $v \geq s/k$ 。所以只要 $s/k > k(2k - 1)$ ，即

$$s > k^2(2k - 1),$$

就保证 $v > k(2k - 1)$ ，进而 (8.10) 可解。

v 个系数 d_j 落入 $\leq p^{\gamma-1}\delta \leq k$ 个剩余类



必有一类装 $\geq v/k$ 个，超过 $2k-1$ 个就能凑出非平凡零解

抽屉原理两次出场：先把项分到 k 个 l_j -组（得 $v \geq s/k$ ），再把系数分到 $\leq k$ 个剩余类（得 $v/k > 2k - 1$ 即可）。

原文： 设 $p = 2$ 。再一次，若 $\tau = 0$ ，即 k 为奇数，则不成问题。若 $\tau > 0$ ，我们可以像定理 7.2 证明中那样论证，但有一个更有效的论证相当简单。我们将证明

$$d_1 y_1^k + \cdots + d_v y_v^k \equiv 0 \pmod{2^\gamma}$$

有一个解，其中 $y_1, \dots, y_v$ 并非全为偶数，只要 $v \geq 2^\gamma$ 。我们通过取 $y_j = 0$ 或 1 来找到这些解（如同与引理 5.6 相关联时所指出的，当 k 是 2 的幂时这不损失一般性）。

这段处理 $p = 2$ ，给一个比一般论证更省的特殊办法。逐点补全：

- 1 " $\tau = 0$ (k 奇) 不成问题"。此时 $p = 2 \nmid k$ ，提升门槛 $\gamma = \tau + 2 = 2$ 很小，且 k 次方映射在模 2 的单位上是双射（奇数的奇次方还是奇数，且覆盖所有奇剩余），随便取就有非平凡解，无需特殊处理。
- 2 $\tau > 0$ (即 k 偶) 才需要技巧。此时 k 含因子 2 ， k 次方在模 2^γ 下的像很受限。作者给一个干净的论证。
- 3 只取 $y_j \in \{0, 1\}$ 。当 k 是 2 的幂时（这是最坏情形，引理 5.6 处指出可归到此），注意 $0^k = 0$ ， $1^k = 1$ 。于是把 y_j 限制成 0 或 1 ，方程 $\sum d_j y_j^k$ 就退化成“选一部分系数 d_j 相加”——一个**纯粹的子集和问题**！要求“不全为偶（即不全 0 ）”就是“至少选一个 d_j ”。这是巨大的简化。
- 4 目标变为：能否从 v 个奇数 $d_1, \dots, d_v$ 中选出非空一组，使它们的和 $\equiv 0 \pmod{2^\gamma}$ ？下面用归纳证明：只要 $v \geq 2^\gamma$ 就一定能办到。

原文：首先，若 $\gamma = 1$ ，我们可以求解

$$d_1 t_1 + d_2 t_2 \equiv 0 \pmod{2}$$

方法是取 $t_1 = t_2 = 1$ （因为 d_1, d_2 为奇数）。其次，我们可以求解

$$d_1 t_1 + d_2 t_2 + d_3 t_3 + d_4 t_4 \equiv 0 \pmod{4}$$

方法是取 $t_1 = t_2 = 1, t_3 = t_4 = 0$ （若 $d_1 + d_2 \equiv 0 \pmod{4}$ ），或者取 $t_1 = t_2 = 0, t_3 = t_4 = 1$ （若 $d_3 + d_4 \equiv 0 \pmod{4}$ ），或者取 $t_1 = t_2 = t_3 = t_4 = 1$ 。这个过程继续下去，证明可以容易地通过对 γ 作归纳来完成。

把这段写成完整的归纳证明（这里 t_j 就是上面的 $y_j \in \{0, 1\}$ ）：

1 基础 $\gamma = 1$ (模 2)。有 $v \geq 2^1 = 2$ 个奇数。取 $t_1 = t_2 = 1$: 和 $= d_1 + d_2 = \text{奇} + \text{奇} = \text{偶} \equiv 0 \pmod{2}$ 。成功, 且非全 0。

2 下一层 $\gamma = 2$ (模 4)。有 $v \geq 2^2 = 4$ 个奇数 d_1, d_2, d_3, d_4 。每个奇数模 4 是 1 或 3。考察三对里的和:

○ 若 $d_1 + d_2 \equiv 0 \pmod{4}$: 取 $t_1 = t_2 = 1, t_3 = t_4 = 0$, 得证。

○ 若 $d_3 + d_4 \equiv 0 \pmod{4}$: 取 $t_3 = t_4 = 1, t_1 = t_2 = 0$, 得证。

○ 否则 $d_1 + d_2 \equiv 2$ 且 $d_3 + d_4 \equiv 2 \pmod{4}$: 取全 1, 则

$$d_1 + d_2 + d_3 + d_4 \equiv 2 + 2 = 4 \equiv 0 \pmod{4}, \text{得证。}$$

三种情况覆盖所有可能, 故模 4 总能用 4 个奇数非平凡求解。

3 归纳步骤 (从 γ 到 $\gamma + 1$)。设命题对 γ 成立: 任意 2^γ 个奇数中可选非空子集, 其和 $\equiv 0 \pmod{2^\gamma}$ 。现给 $2^{\gamma+1}$ 个奇数, 要凑出和 $\equiv 0 \pmod{2^{\gamma+1}}$ 。

1. 把这 $2^{\gamma+1}$ 个奇数分成两批, 每批 2^γ 个。对每批用归纳假设, 各得一个非空子集, 其和 $\equiv 0 \pmod{2^\gamma}$, 记两个和为 A, B , 都是 2^γ 的倍数。

2. 写 $A = 2^\gamma a, B = 2^\gamma b$ 。若 a 是偶数, 则 $A \equiv 0 \pmod{2^{\gamma+1}}$, 直接取第一批那个子集即可; b 偶数时同理取第二批。

3. 若 a, b 都是奇数, 则 $a + b$ 是偶数, 于是

$$A + B = 2^\gamma(a + b) \equiv 0 \pmod{2^{\gamma+1}}; \text{取两批子集的并 (仍非空), 其和} \\ = A + B \equiv 0 \pmod{2^{\gamma+1}}。$$

无论哪种情况都成功。归纳完成。

4 结论: 只要 $v \geq 2^\gamma$, 同余式 $\sum d_j y_j^k \equiv 0 \pmod{2^\gamma}$ 就有 $y_j \in \{0, 1\}$ 、不全为 0 的解。

为什么这个论证"更有效" (取舍说明) 对 $p = 2$ 如果照搬 $p > 2$ 的剩余类论证, 会要求 $v > k(2k - 1)$ 同样量级; 而这里用"0/1 子集和 + 归纳"只要 $v \geq 2^\gamma$, 而 $2^\gamma = 2^{\tau+2} = 4 \cdot 2^\tau \leq 4k$ (见下), 门槛 $4k$ 远小于 $k(2k - 1)$ 。所以专门为 $p = 2$ 设计的论证更省变量——这是"针对具体素数选最省工具"的取舍。

原文：条件 $v \geq 2^\gamma$ 在 $v \geq 4k$ 时即满足，因此当 $s > k(4k - 1)$ 时即满足。由于这个数小于 $k^2(2k - 1)$ ，它对结果没有影响。

把这三句的计算补全：

1 " $v \geq 4k \Rightarrow v \geq 2^\gamma$ "。对 $p = 2$, $\gamma = \tau + 2$, 故 $2^\gamma = 2^{\tau+2} = 4 \cdot 2^\tau$ 。又 $k = 2^\tau k_0$ (k_0 奇 ≥ 1), 所以 $2^\tau = k/k_0 \leq k$, 于是 $2^\gamma = 4 \cdot 2^\tau \leq 4k$ 。因此 $v \geq 4k$ 就保证 $v \geq 2^\gamma$ 。

2 " $s > k(4k - 1) \Rightarrow v \geq 4k$ "。由 $v \geq s/k$, 若 $s > k(4k - 1)$ 则 $v \geq s/k > 4k - 1$, 而 v 是整数, 故 $v \geq 4k$ 。

3 "这个数小于 $k^2(2k - 1)$, 无影响"。比较 $k(4k - 1)$ 与 $k^2(2k - 1)$: 两边同除 k , 比 $4k - 1$ 与 $k(2k - 1) = 2k^2 - k$ 。它们的差 $= (2k^2 - k) - (4k - 1) = 2k^2 - 5k + 1$, 当 $k \geq 3$ 时 > 0 。所以对我们关心的 k , $p = 2$ 的门槛 $k(4k - 1)$ 比 $p > 2$ 的门槛 $k^2(2k - 1)$ 更小, 于是把两者取最大时 $p = 2$ 不构成约束——最终的门槛由奇素数情形 $s > k^2(2k - 1)$ 决定。

五、主定理与它的"最优性"讨论

原文：综合我们的结果，我们已经证明：

定理 8.1 设 $c_1, \dots, c_s$ 是给定的整数，没有一个为 0，且当 k 为偶数时不全同号。那么，只要 $s \geq 2^k + 1$, 且

$$s \geq k^2(2k - 1) + 1, \quad (8.11)$$

方程

$$c_1 x_1^k + \dots + c_s x_s^k = 0$$

就有无穷多组整数解 $x_1, \dots, x_s$, 其中没有一个为 0。

这是本章的主结果。把两个条件的来源讲清楚, 并说明结论怎么推出来:

条件 $s \geq 2^k + 1$

来自次弧估计 (保证误差项 $O(P^{s-k-\delta})$ 真的小于主项), 即让渐近公式 (8.3) 本身成立所需。它由第 3-4 章的 Weyl/华罗庚不等式决定。

条件 (8.11): $s \geq k^2(2k-1) + 1$

来自本章下半段对奇异级数 $\mathfrak{S} > 0$ 的研究: 取整后即 $s > k^2(2k-1)$ 。它保证主项系数非零。

两条件取"且"

必须同时满足, 渐近公式才既成立又有正主项。

1 两个条件都满足时, (8.3) 成立且 $C'_{k,s} > 0$, $\mathfrak{S} > 0$, $(c_1 \cdots c_s)^{-1/k} > 0$, 于是主项 = (正常数) P^{s-k} 。

2 因 $s \geq 2^k + 1 > k$ (指数 $s-k > 0$), 当 $P \rightarrow \infty$ 主项 $\rightarrow \infty$, 而误差 $O(P^{s-k-\delta})$ 相对可忽略, 故 $\mathcal{N}(P) \rightarrow \infty$ 。

3 $\mathcal{N}(P)$ 是盒子里正整数解的个数, 它随 P 无限增大, 意味着正整数解有无穷多组; 正整数解每个分量 $\geq 1 \neq 0$, 故得到无穷多组每个分量都非零的整数解。 k 偶数时"不全同号"是前面已说明的必要前提。证毕。

原文: 条件 (8.11) 来自我们对奇异级数的研究, 它并非最佳可能。在 [27] 中, Davenport 与 Lewis 证明了: 为确保 $\mathfrak{S} > 0$, 只要

$$s \geq k^2 + 1$$

即足够。当 $k+1$ 是素数 p 时这个条件是最佳可能的。

这段是历史与最优性讨论。我们上面证出的门槛 $k^2(2k-1) + 1$ (关于 k 是三次多项式) 其实偏大, 能改进。

历史背景 H. Davenport (达文波特, 1907–1969, 英国数论大家, 本书作者) 与 D. J. Lewis 在文献 [27] 中证明: 只要变量数

$$s \geq k^2 + 1$$

就能保证奇异级数 $\mathfrak{S} > 0$ 。这是关于 k 的二次门槛, 比本章的三次门槛 $k^2(2k-1)+1$ 小得多 (例如 $k=3$: 本章给 $3^2 \cdot 5 + 1 = 46$, 而 $k^2 + 1 = 10$)。本章为了让证明自足、初等、易讲, 采用了较粗但够用的论证, 故门槛偏大——这是“证明的简洁”与“结果的最优”之间的取舍。

原文: 因为这时 $x^k \equiv 1 \pmod{p}$ (若 $x \not\equiv 0 \pmod{p}$), 于是容易推出: 含 k^2 个变量的同余式

$$p^k \mid ((x_1^k + \cdots + x_k^k) + p(x_{k+1}^k + \cdots + x_{2k}^k) + \cdots + p^{k-1}(x_{k^2-k+1}^k + \cdots + x_{k^2}^k))$$

除非所有变量都能被 p 整除, 否则无解。然而, 对 k 的大多数值, 较小的值 $k^2 + 1$ 即足够。

这段给出“ $k^2 + 1$ 不能再减”的反例, 即当 $k+1=p$ 为素数时, k^2 个变量真的不够。把“容易推出”补全:

1 **前提:** $k+1=p$ 为素数。则 $k=p-1$ 。由费马小定理: 若 $p \nmid x$, 则 $x^{p-1} \equiv 1 \pmod{p}$, 即 $x^k \equiv 1 \pmod{p}$ 。也就是说, 模 p 时, 每个不被 p 整除的数的 k 次方都恰好 $\equiv 1$ (而被 p 整除的数 k 次方 $\equiv 0$)。 k 次方在模 p 下只有 0 和 1 两个值!

2 **构造的同余式。** 考虑 k^2 个变量, 分成 k 组、每组 k 个, 第 m 组 ($m=0, 1, \dots, k-1$) 前面乘以权重 p^m :

$$F = \sum_{m=0}^{k-1} p^m (x_{mk+1}^k + \cdots + x_{m(k+k)}^k),$$

要求 $p^k \mid F$ 。

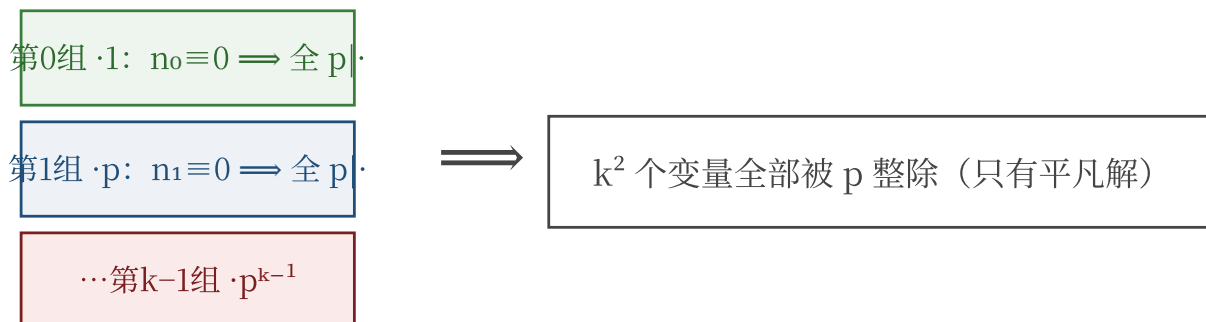
3 逐层（按 p 进制）分析，说明只有全被 p 整除才可能。假设并非所有变量都被 p 整除，取出"含 p 因子最少"的那一层来导出矛盾。

- 看模 p （最低层）。 $F \equiv x_1^k + \cdots + x_k^k \pmod{p}$ （其余项都带因子 p ）。
每个 x_i^k 是 0 或 1，所以这组里"非 p 倍数"的变量个数记为 n_0 ，则
 $F \equiv n_0 \pmod{p}$ 。要 $p^k \mid F$ 首先要 $p \mid F$ ，即 $p \mid n_0$ 。但第 0 组只有 $k = p - 1$ 个变量， $0 \leq n_0 \leq p-1$ 即第 0 组所有变量都被 p 整除。
- 于是第 0 组每项 x_i^k 实际被 p^k 整除（ $x_i = p \cdot (\cdots) \Rightarrow x_i^k$ 含 p^k ），在模 p^k 意义下第 0 组可整体忽略。把 F/p 再看模 p ，主导层变成第 1 组
 $x_{k+1}^k + \cdots + x_{2k}^k$ ，同样推出该组非 p -倍数变量个数 n_1 满足 $p \mid n_1$ ，故 $n_1 = 0$ ，第 1 组也全被 p 整除。
- 如此逐层（ $m = 0, 1, \dots, k-1$ ）推进，每一层都被迫"该组全部被 p 整除"。最终所有 k^2 个变量都被 p 整除——这就是平凡解（可以同除 p 后重复，得到唯一平凡解）。

4 结论。这个 k^2 变量的同余式没有非平凡解，于是它对应的局部因子 $\chi(p)$ 在该 p 处会退化， $\mathfrak{S}$ 可能为 0。这说明：当 $k+1$ 是素数时， $s = k^2$ 个变量确实不足以保证 $\mathfrak{S} > 0$ ；必须 $s \geq k^2 + 1$ 。所以 $k^2 + 1$ 这个门槛在这种 k 下不能再小，是"最佳可能"的。

5 "对大多数 k 更小的值即足够"。上述反例只在 $k+1$ 恰为素数时出现；对绝大多数 k ，连 $k^2 + 1$ 都不必要，更少的变量就够了。但作为对一切系数都成立的一般定理， $k^2 + 1$ 是无法再普遍降低的下限。

$k+1=p$ 素数： $x^k \equiv 0$ 或 $1 \pmod{p}$ ，逐层被逼为 0



故 $s=k^2$ 不够，门槛 k^2+1 最优（当 $k+1$ 为素数）

当 $k+1$ 是素数， k 次方在模 p 下只取 $0/1$ ，迫使 k^2 个变量逐层归零——这就是 k^2+1 最优性的反例。

六、方法的两面：盒子里的渐近 vs. 解的分布

原文：在前面对方程 $c_1x_1^k + \cdots + c_sx_s^k = 0$ 的处理中，我们已得到了：当 $P \rightarrow \infty$ 时，在 s 维盒子 $0 < x_j \leq P_j$ 中整数解个数的渐近公式。但这个盒子与方程之间并无任何独特方式的联系，结果之所以有意义，主要在于它确立了无穷多解的存在性。然而，为证明这一点，并不必须得到这样一个盒子中所有解的渐近公式；只需考虑某个特殊子集就足够了。这样，我们就可以使用类似于为估计华林问题中 $G(k)$ 所发展出来的方法。在第9章中，我们将研究维诺格拉多夫（Vinogradov）的方法，它对于大的 k 非常有效；在随后的一章中，我们将把这个方法应用于我们一直在研究的方程。

这段是承上启下的方法论反思。要点：

- 1 圆法给的"很多"。我们不仅证明了"有无穷多解"，还给出了盒子里解数的**精确渐近公式** $\mathcal{N}(P) \sim (\text{正常数})P^{s-k}$ 。这是很强的定量结果。
- 2 但盒子是"人造"的。盒子 $\setminus (0$
- 3 因此可以省力。证无穷多解，不必数清盒子里**所有解**，只要在某个**特殊子集**上证明解数 $\rightarrow \infty$ 即可。放弃"数全部"换取"门槛更低"，正是估计华林问题 $G(k)$ （表示成 k 次方和所需的最少项数）时发展出来的省力思路。
- 4 预告。第9章讲**维诺格拉多夫（I. M. Vinogradov，苏联数论大家）**的方法，对大的 k 特别有效（能把所需变量数从指数级 2^k 大幅压低）；下一章把它用到本章的方程上。这解释了为什么本章的 2^k+1 条件后面会被改进。

原文：不过，不应忽视的是：我们一直在使用的方法特别适合于研究方程之解的分布。设 $\lambda_1, \dots, \lambda_s$ 是任意实数，没有一个为 0，并满足方程

$$c_1\lambda_1^k + \cdots + c_s\lambda_s^k = 0.$$

那么本节的方法使人能够求得：当 $P \rightarrow \infty$ 时，我们的方程在盒子

$$1 - \delta < \frac{x_j}{\lambda_j P} < 1 + \delta$$

中整数解的渐近公式，这里 δ 是任意小的固定正数。

这段指出圆法的独特优势：不仅数解，还能控制解落在哪里。 详解：

$$\lambda_1, \dots, \lambda_s$$

任意一组满足方程的**实数解**（不要求整数），每个非零。它给出方程的实轨迹（一个曲面/锥面）上的一个方向。

$$\text{盒子 } 1 - \delta < \frac{x_j}{\lambda_j P} < 1 + \delta$$

等价于 x_j 落在 $\lambda_j P$ 的 $(1 \pm \delta)$ 倍范围内，即整数解 $(x_1, \dots, x_s)$ 在方向上**逼近**给定实方向 $(\lambda_1, \dots, \lambda_s)$ 、长度约为 P 。 δ 越小，方向越精确。

圆法之所以能做这件事：把积分中的"主弧"近似换成围绕指定方向 λ 的盒子，奇异积分 $C'_{k,s}$ 会相应变成**这个特定盒子上的局部积分**，仍然为正（因为实解曲面在 λ 处是光滑非退化的）。于是在任何一个对准实方向的细盒子里，整数解数仍 $\sim (\text{正常数})P^{s-k} \rightarrow \infty$ 。

原文：用几何语言表述，这一结果意味着：从原点到锥面

$$c_1x_1^k + \cdots + c_sx_s^k = 0$$

上整点的那些"射线"，若把该锥面视为 s 维空间中的实轨迹，则它们在此锥面上处处稠密。因此，尽管现在要阐述的方法在确立无穷多解方面更为有效，但它并不能完全取代前一种方法。

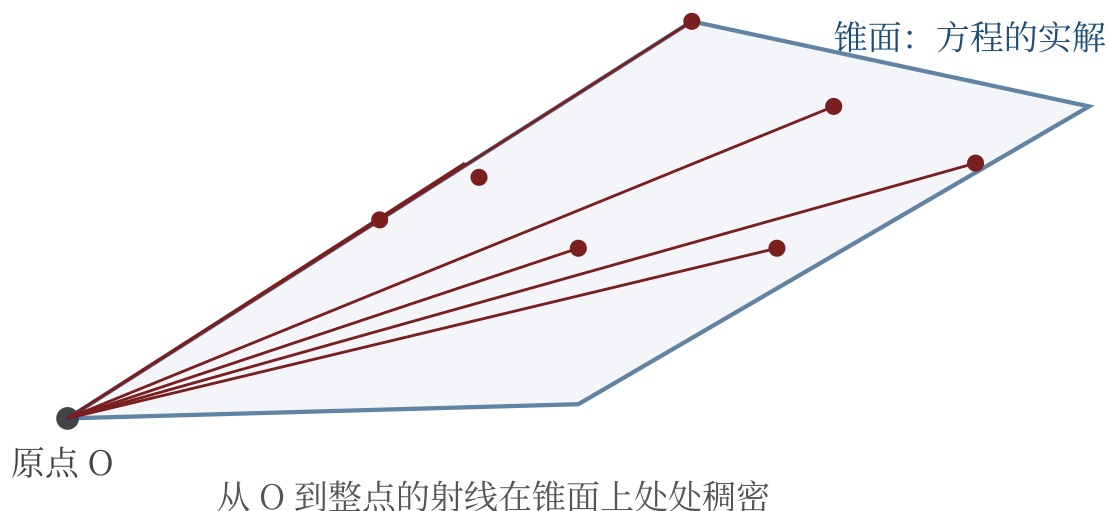
几何收尾。 把抽象结论翻译成图像：

- 1 锥面是什么。** 方程 $c_1x_1^k + \cdots + c_sx_s^k = 0$ 在 s 维实空间里的解集是一个曲面；因为方程齐次，解集对"过原点的放缩"封闭（前面证过 x 是解 $\Rightarrow tx$ 是

解)，所以这个曲面由**过原点的射线**铺成——这就是"锥面" (cone)。

2 "射线处处稠密"。上一段说：在任何对准某个实方向 λ 的细盒子里都有整点解。换句话说，**锥面上任何一个方向附近，都能找到整点**；从原点连到这些整点的射线，密密麻麻地铺满整个锥面，没有空隙——这就是"处处稠密 (everywhere dense)"。

3 两种方法的关系（呼应上一段）。"更有效的新方法"（维诺格拉多夫式，下章）擅长用更少变量证"无穷多解"，但它通常只给存在性、不给分布。而本章的圆法虽然要求变量更多，却能额外给出"解在锥面上稠密"这种**分布信息**。所以新方法"更有效"却**不能完全取代**旧方法——各有所长，这是作者特意强调的方法论结论。



齐次方程的实解是一张过原点的锥面；圆法证明：连向整点解的射线密布整个锥面。这是"分布信息"，新方法给不出。

全章小结

一句话脉络 "数整数解" $\xrightarrow{\text{挑零积分}} \mathcal{N}(P) = \int_0^1 \prod T_j d\alpha \xrightarrow{\text{主弧/次弧}}$

$$\mathcal{N}(P) \approx C'_{k,s} \cdot \mathfrak{S} \cdot P^{s-k}.$$

其中**奇异积分** $C'_{k,s} > 0$ (实数几何, 靠一个正函数在开集上的积分) 已直接验证;

奇异级数 $\mathfrak{S} > 0$ (素数算术) 经 " $\mathfrak{S} = \prod_p \chi(p)$, 逐素数证 $\chi(p) > 0$, 再化为

(8.10) 有非平凡解"层层归约, 靠 $p > 2$ 的剩余类抽屉论证 (要 $s > k^2(2k-1)$)

与 $p = 2$ 的 0/1 子集和归纳 (要 $s > k(4k-1)$, 更松) 完成。

两因子皆正 $\Rightarrow$ 主项 $\sim (\text{正常数})P^{s-k} \rightarrow \infty \Rightarrow$ **定理 8.1:** $s \geq 2^k + 1$ 且

$s \geq k^2(2k-1) + 1$ 时方程有无穷多组非零整数解。后续 Davenport-Lewis 把奇异级数门槛改进到最优的 $k^2 + 1$ (当 $k+1$ 素数时不可再降)。最后强调: 圆法不仅给存在性, 还给出解在锥面上稠密的分布信息。

返回 [全书目录](#)