

DAVENPORT · 圆法 · 高中详解版

方程 $c_1x_1^k + \cdots + c_sx_s^k = N$

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

本章要解决什么

前面几章（Waring 问题）研究的是把大整数 N 写成 s 个 k 次方之和： $N = x_1^k + \cdots + x_s^k$ 。本章把它推广为带系数的版本：

$$N = c_1x_1^k + c_2x_2^k + \cdots + c_sx_s^k,$$

这里 $c_1, \dots, c_s$ 是事先给定的正整数（"权重"）， $x_1, \dots, x_s$ 是要找的正整数。问题是：当变量个数 s 足够大时，是不是每个大的 N 都能这样表示？表示的方法有多少种？

读完本章你将掌握：① 为什么仅靠" s 大"还不够，必须加一个"同余可解"的条件；② 圆法（Hardy-Littlewood circle method）这台机器在加了系数后只需做哪些微调；③ 主结论**定理 7.1**给出表示个数 $r(N)$ 的精确渐近公式；④ **定理 7.2、7.3**说明在什么条件下这个公式的"主项"恒正，从而 $r(N) \rightarrow \infty$ 。本章不重证圆法的全部细节（那是前几章的事），而是讲清"加系数后每一处该怎么改、为什么这么改"。

阅读提示：本页把 Davenport 原书第 7 章的中文译文**逐段**展开。黑色叙述贴着原文走，彩色框（目标 / 背景 / 分步推演 / 定理 / 证明）与配图是面向高中生的零省略详解。本章大量引用前几章的"引理 3.1、4.1、5.6"等——这些是圆法的零件，我们会在用到时把它们**含义和作用**从零讲清，你不必先读完前六章也能看懂这一章在做什么。

预备知识：先把"圆法"这台机器看懂

为什么需要这一节 本章每一句话都建立在前几章的圆法框架上。高中没学过复指数、定积分、求和号这些工具，但它们其实都能从零讲明白。下面这一节把本章会反复出现的所有符号和思想**一次性**讲清；后面逐段讲解时遇到它们，你就能直接回看这里。

记号 1：求和号 $\sum$

$\sum_{x=1}^P f(x)$ 读作"西格玛，从 $x=1$ 到 P 求 $f(x)$ 的和"。它只是"把一串数加起来"的缩写：

$$\sum_{x=1}^P f(x) = f(1) + f(2) + f(3) + \cdots + f(P).$$

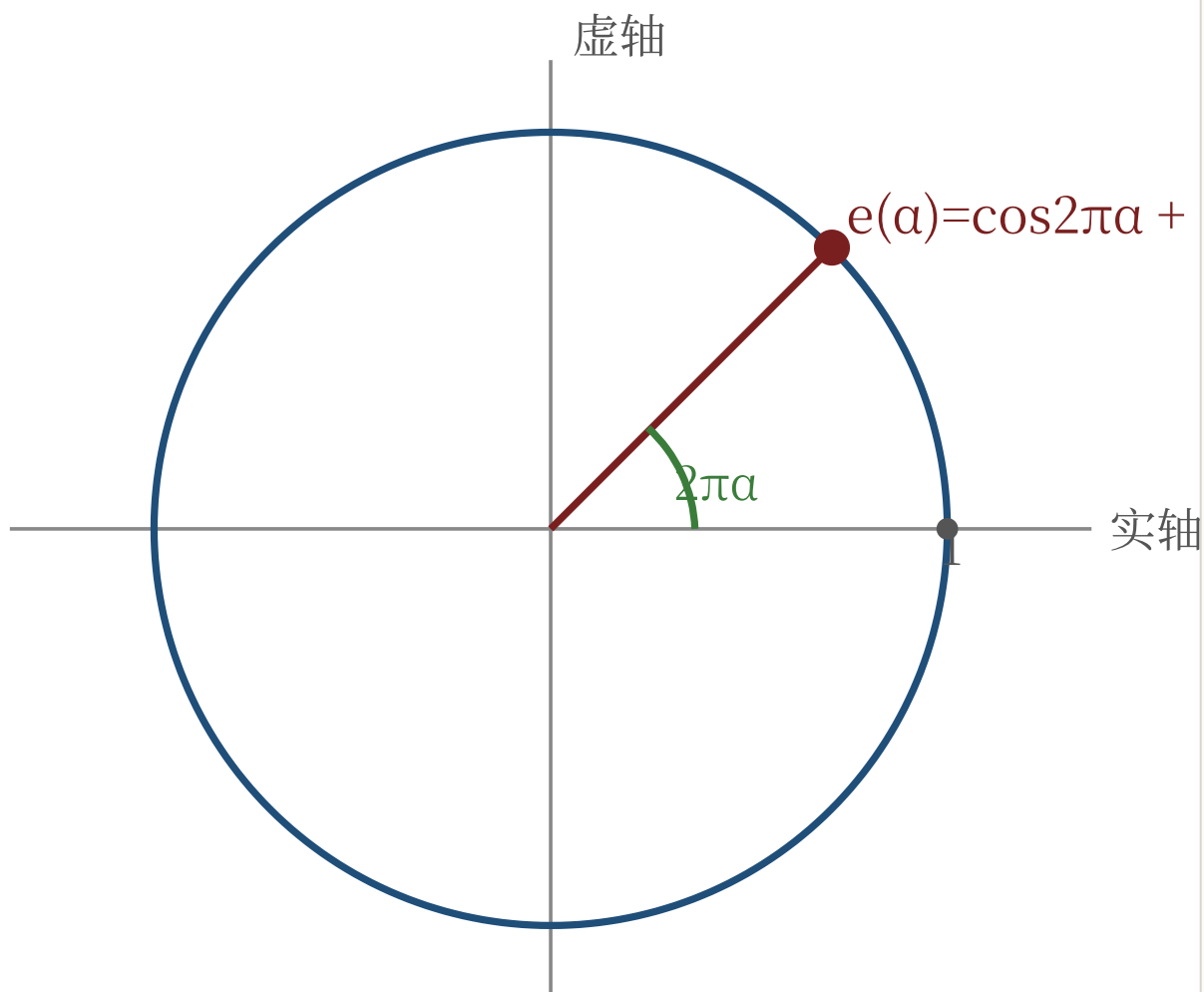
例如 $\sum_{x=1}^3 x^2 = 1^2 + 2^2 + 3^2 = 14$ 。下标 x 叫求和变量，它从 1 跑到 P ，每跑一个值就把对应的 $f(x)$ 加进来。

记号 2：复指数 $e(\alpha) = e^{2\pi i \alpha}$ （圆法的心脏）

$e(\alpha)$ 定义为

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

这里 i 是虚数单位 ($i^2 = -1$)， $\pi \approx 3.14159$ 。它是一个**复数**。复数 $z = a + bi$ 可以画成平面上的点 (a, b) 。由于 $\cos^2 + \sin^2 = 1$ ， $e(\alpha)$ 这个点永远落在**单位圆**（半径为 1 的圆）上，它与正半轴的夹角是 $2\pi\alpha$ 弧度，也就是 α 圈。



$e(\alpha)$ 是单位圆上的点，转过的角度是 α 整圈。 $\alpha = 0$ 时在 1； $\alpha = \frac{1}{4}$ 时在 i （最上）； $\alpha = \frac{1}{2}$ 时在 -1 。"圆法"的名字正来自这个圆。

它的三个关键性质（本章反复用到）：

1 模长恒为 1: $|e(\alpha)| = \sqrt{\cos^2 + \sin^2} = 1$ 。复数 $z = a + bi$ 的"模长" $|z| = \sqrt{a^2 + b^2}$ 就是它到原点的距离。

2 指数相加: $e(\alpha)e(\beta) = e(\alpha + \beta)$, 因为 $e^{2\pi i\alpha}e^{2\pi i\beta} = e^{2\pi i(\alpha+\beta)}$ 。

- 3 周期为 1: $e(\alpha + 1) = e(\alpha)$, 因为转一整圈 (角度多 2π) 回到原地。所以 $e(\alpha)$ 只关心 α 的小数部分。这就是为什么后面积分只需在区间 $[0, 1]$ 上做。

记号 3: 定积分 $\int_0^1$ 与它的"筛子"作用

$\int_0^1 g(\alpha) d\alpha$ 读作" g 在 0 到 1 上的定积分"。直观上它表示函数曲线 $y = g(\alpha)$ 在 $\alpha \in [0, 1]$ 这段下方围出的"有向面积" (曲线在横轴上方算正、下方算负)。对复值函数, 就把实部、虚部分别积分。圆法里我们只需要它的一条神奇性质:

$$\int_0^1 e(m\alpha) d\alpha = \begin{cases} 1, & m = 0, \\ 0, & m \text{ 为非零整数.} \end{cases} \quad (\text{筛子})$$

为什么 这是本章一切恒等式的根。证明只需高中的"转一圈净位移为零":

- 1 当 $m = 0$: 被积函数 $e(0) = 1$ 恒等于常数 1, 所以 $\int_0^1 1 d\alpha = 1$ 。

- 2 当 m 是非零整数: $e(m\alpha)$ 是单位圆上的点, 当 α 从 0 走到 1 时它正好匀速转了 $|m|$ 整圈。每转一整圈, 点在圆上均匀地走遍各方向, 正方向和反方向互相抵消, "平均位置"回到圆心 0。用微积分写出来:

$$\int_0^1 e(m\alpha) d\alpha = \frac{e(m\alpha)}{2\pi im} \Big|_0^1 = \frac{e(m) - e(0)}{2\pi im} = \frac{1 - 1}{2\pi im} = 0 \quad (\text{用到 } e(m) = e^0 = 1, \text{ 因为 } m \text{ 是整数、周期为 } 1)。$$

这把"筛子"如何抓表示个数? 假设我们想知道方程 $c_1 x_1^k + \cdots + c_s x_s^k = N$ 有多少组解。令 $M = c_1 x_1^k + \cdots + c_s x_s^k - N$, 它是个整数。由筛子, $\int_0^1 e(M\alpha) d\alpha$ 在 $M = 0$ (即正好凑成 N) 时给 1, 否则给 0。于是对所有变量求和再积分, 就把"恰好等于 N 的解"一个个数出来——这正是圆法把"解的个数"变成"一个积分"的核心招式。

记号 4: 生成函数 $T_j(\alpha)$ 、表示个数 $r(N)$

生成函数 把"筛子"用到我们的方程上: 对每个 j 定义一个和式

$T_j(\alpha) = \sum_x e(\alpha c_j x^k)$ (精确定义见 (7.1))。把这些和式相乘并展开, 每一项形如 $e(\alpha(c_1 x_1^k + \cdots + c_s x_s^k))$ 。再乘上 $e(-N\alpha)$ 并对 α 从 0 到 1 积分, 由筛子, 只有 $c_1 x_1^k + \cdots + c_s x_s^k = N$ 的那些项贡献 1, 其余贡献 0。于是

$$r(N) = \int_0^1 T_1(\alpha) T_2(\alpha) \cdots T_s(\alpha) e(-N\alpha) d\alpha,$$

其中 $r(N)$ 正是 N 的表示个数。整个圆法就是想办法把这个积分算出来 (算个近似值)。

记号 5: $O(\cdot)$ 与 $\ll$ ("误差不超过"的速记)

大 O 与 $\ll$ " $f(N) = O(g(N))$ " 以及等价写法 " $f(N) \ll g(N)$ " 都表示: 存在一个与 N 无关的常数 $C > 0$, 使得当 N 足够大时 $|f(N)| \leq C g(N)$ 。一句话: f 的大小最多是 g 的常数倍。例如 $3N + 100 = O(N)$, 因为当 N 大时 $3N + 100 \leq 4N$ 。它让我们专注于"增长的量级"而不纠缠具体常数。本章里 ε 表示"任意小的正数", δ, δ' 表示"某个固定的正数"。

记号 6: 同余 $\equiv \pmod{m}$

$a \equiv b \pmod{m}$ 读作" a 与 b 模 m 同余", 意思是 $a - b$ 能被 m 整除, 即除以 m 余数相同。例如 $17 \equiv 2 \pmod{5}$ 。本章里 p 总表示素数, p^ν 是它的方幂。"同余式可解"指: 能找到整数 x 让等式在模 m 意义下成立。

记号 7: 圆法的总流程 (主弧、次弧、奇异级数、奇异积分)

把整张地图先看一遍 圆法把积分区间 $[0, 1]$ 切成两类小区间:

- **主弧 $\mathfrak{M}$ (major arcs)**: 围绕"分母小的分数" a/q (q 小) 的一些很短的区间。在这些点附近 $T_j(\alpha)$ 很大, 是积分的主要贡献来源。算出来给一个漂亮的主项。
- **次弧 $\mathfrak{m}$ (minor arcs)**: 剩下的所有点。这里 $T_j(\alpha)$ 比较小, 积分贡献只是误差项, 要证明它确实小 (这要用 Weyl、Hua 不等式)。

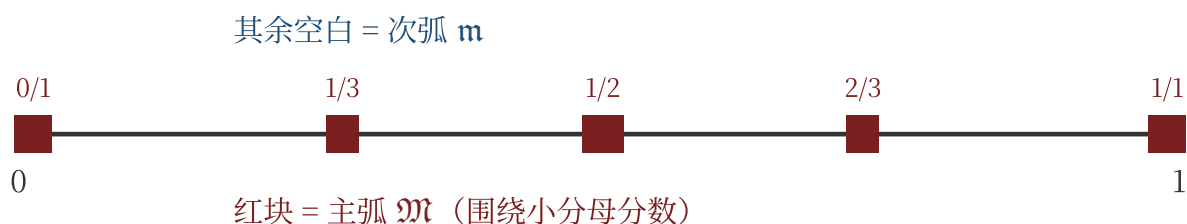
主弧上的计算最后拆成两个因子的乘积:

- **奇异积分 (singular integral)**: 来自"实数层面"的连续逼近, 本章里它给出 $N^{s/k-1}$ 这个增长量级和常数 $C_{k,s}/(c_1 \cdots c_s)^{1/k}$ 。
- **奇异级数 $\mathfrak{S}(N)$ (singular series)**: 来自"各个素数 p 的同余层面", 它把所有"模 p^v 能不能解"的局部信息打包成一个无穷级数。它是否有正下界, 决定了 N 到底能不能表示。

所以最终长这样 (即本章定理 7.1):

$$r(N) = \underbrace{\frac{C_{k,s}}{(c_1 \cdots c_s)^{1/k}} N^{s/k-1}}_{\text{奇异积分 (实数信息)}} \times \underbrace{\mathfrak{S}(N)}_{\text{奇异级数 (同余信息)}} + \underbrace{O(N^{s/k-1-\delta})}_{\text{次弧误差}}.$$

本章的全部工作, 就是把无系数版本的这套结论, **逐个零件**改成带系数 c_j 的版本。下面逐段开始。



圆法把 $[0, 1]$ (其实是单位圆, 因为 e 周期为 1) 切成红色"主弧"和剩下的"次弧"。主弧给主项, 次弧给误差。

逐段精讲

第 1 段：问题是什么，以及"为什么仅有 s 大还不够"

原文译文：我们接下来考虑这样一个问题：把一个大正整数 N 表示成 $c_1x_1^k + \cdots + c_sx_s^k$ 的形式，其中 $c_1, \dots, c_s$ 是给定的正整数，而 $x_1, \dots, x_s$ 是任意的正整数。如果不加任何进一步的假设，那么"当 $s \geq s_0(k)$ 时每个大的 N 都可表示"这一结论是不成立的。因为假设 $c_1, \dots, c_{s-1}$ 都能被某个素数 p 整除，而 c_s 不能；那么一个不能被 p 整除的整数 N ，如果它对模 p 不具有与 c_s 相同的 k 次幂特征，就必然不可表示。

这一段在讲什么 先抛出本章的研究对象（带系数的 k 次幂表示），然后立刻警告：**不能指望"只要变量够多就万事大吉"**，并造了一个反例来吓你。下面把这个反例彻底讲透。

N

要被表示的"大正整数"，是目标。

k

幂次，固定的正整数（ $k = 2$ 是平方和， $k = 3$ 是立方和……）。

$c_1, \dots, c_s$

给定的正整数系数（权重），事先固定不动。

$x_1, \dots, x_s$

要寻找的正整数变量。

s

变量（项）的个数。

$s_0(k)$

一个"只依赖 k 的阈值"。" $s \geq s_0(k)$ "是想说"项数足够多"。无系数的 Waring 问题里，项数一多就能表示所有大 N ；本段说明加了系数后**不行**。

" k 次幂特征 (kth power character)" 是什么意思 在模 p 的世界里 (即只看除以 p 的余数 $0, 1, \dots, p-1$), 把一个数能否写成"某数的 k 次方"叫它的" k 次幂特征"。能写成 $y^k \pmod{p}$ 的余数叫 k 次剩余, 不能的叫 k 次非剩余。例如模 $p=7$ 、 $k=2$ 时, 平方剩余是 $1, 2, 4$ (因为 $1^2=1, 3^2=2, 2^2=4 \pmod{7}$), 非剩余是 $3, 5, 6$ 。"具有相同的 k 次幂特征"就是落在同一类里。

把反例补成完整推导 取系数让前 $s-1$ 个都是 p 的倍数、只有最后一个 c_s 不是。看方程在模 p 下变成什么:

1 因为 $c_1, \dots, c_{s-1}$ 都能被 p 整除, 所以

$c_1 x_1^k \equiv 0, \dots, c_{s-1} x_{s-1}^k \equiv 0 \pmod{p}$ (p 的倍数模 p 余 0)。前 $s-1$ 项整体"消失"了。

2 于是原方程模 p 后只剩最后一项: $c_s x_s^k \equiv N \pmod{p}$ 。

3 这要求 N 必须能写成 c_s 乘上某个 k 次方 (模 p)。也就是说 N/c_s (模 p 下的"除法", 即乘 c_s 的逆元) 必须是个 k 次剩余——这正是" N 与 c_s 有相同的 k 次幂特征"。

4 但 k 次剩余只占余数的一部分 (不是全部)。只要取一个 N , 它不被 p 整除、且不满足上述特征要求, 那么无论 s 多大、变量怎么选, 方程模 p 都无解, 从而原方程在整数上也无解。

结论: 项数多救不了这种系数。所以" s 大 $\Rightarrow$ 可表示"在带系数情形不成立。这就逼着我们在下一段加一个新的、关于"模 p^ν 可解"的假设。 ■

第 2 段: 可以假设无公因子; 并提出"同余可解"这一必要假设

原文译文: 我们当然可以假设——在处理标题中的方程时—— $c_1, \dots, c_s$ 没有公因子。为了保证可解性, 我们将发现有必要假设: 同余式 $c_1 x_1^k + \dots + c_s x_s^k \equiv N \pmod{p^\nu}$ 对每个

素数 p 及一切足够大的 ν 都可解，并且其中各项 $c_1x_1^k, \dots, c_sx_s^k$ 不全被 p 整除（或者作出某个可由此推出的其他假设）。

这一段在讲什么 做两件事：① 说明可以不失一般性地假设系数**没有公因子**；② 把上一段反例的教训正式提炼成本章的核心**假设**——必须保证"每个素数方幂模下都能解，且不是平凡地全被 p 吃掉"。

"没有公因子"为什么可以假设 如果 $c_1, \dots, c_s$ 有公因子 $d > 1$ ，那么左边 $c_1x_1^k + \dots + c_sx_s^k$ 永远是 d 的倍数。于是只有当 N 也是 d 的倍数时才可能有解；这时两边同除以 d ，问题就化为系数 c_j/d （已无公因子）、目标 N/d 的同类问题。所以研究"无公因子"的情形不损失一般性。注意：原文说的是"不全有公因子"（do not all have a common factor），即不存在大于 1 的数同时整除全部系数。

新假设（本章的"通行证"）

$$c_1x_1^k + \dots + c_sx_s^k \equiv N \pmod{p^\nu} \quad \text{可解，且各项不全被 } p \text{ 整除.}$$

两个要点逐一解释：

1 "对每个素数 p 及足够大的 ν 可解"：这是上一段反例的"解药"。反例之所以坏，就是因为它在某个 p 下根本无解。现在我们**明文要求**每个素数层面都解得动。 ν 是 p 的方幂指数（ p^ν ）；要求对足够大的 ν 成立，是因为同余信息要"看得足够细"。

2 "各项不全被 p 整除"：这排除一种"假可解"。如果允许所有 x_j 都是 p 的倍数，那 $c_jx_j^k$ 全是 p^k 的倍数，是平凡情形，给不出有用信息。要求至少有一项不被 p 整除，才是"真正非平凡的解"。这个条件在后面定理 7.2、7.3 里是核心。

括号里"或作出某个可由此推出的其他假设"是说：也可以换一个等价或更强、能推出它的条件，本质不变。

第 3 段：只需微调；定义 P_j 与生成函数 $T_j(\alpha)$ （式 7.1）

原文译文：只需对前面的工作作出微小的改动，就能使之适用于这个更一般的方程。我们定义 $P_j = [(N/c_j)^{1/k}]$ 为 $(N/c_j)^{1/k}$ 的整数部分，并定义

$$T_j(\alpha) = \sum_{x=1}^{P_j} e(\alpha c_j x^k). \quad (7.1)$$

这一段在讲什么 宣布主旨——圆法这台机器几乎不用重造，只要把每个零件里出现 x^k 的地方换成 $c_j x^k$ 。然后给出第一个、也是最关键的改动：每个变量有自己的求和上限 P_j ，以及自己的生成函数 $T_j(\alpha)$ 。

$[\cdot]$ 取整符号 $[y]$ 表示不超过 y 的最大整数（向下取整）。例如 $[3.7] = 3$, $[5] = 5$ 。

为什么求和上限是 $P_j = [(N/c_j)^{1/k}]$

- 1 变量 x_j 是正整数，且 $c_j x_j^k$ 这一项不能超过整个目标 N （其余各项都 ≥ 0 ，最多让本项等于 N ）。即 $c_j x_j^k \leq N$ 。
- 2 解出 x_j ： $x_j^k \leq N/c_j$ ，即 $x_j \leq (N/c_j)^{1/k}$ 。（ $(\cdot)^{1/k}$ 是开 k 次方。）
- 3 因为 x_j 是整数，所以最大只能取到 $P_j = [(N/c_j)^{1/k}]$ 。这就是求和上限——把所有"有可能用到"的 x_j 都扫一遍。系数 c_j 越大，本项越"重"，能用的 x_j 范围 P_j 就越小，这很合理。

对比无系数版本 ($c_j = 1$)，那里上限是 $P = [N^{1/k}]$ ，对所有变量一样。加了系数后每个变量的上限 P_j 各不相同，这是**第一处改动**。

生成函数 $T_j(\alpha) = \sum_{x=1}^{P_j} e(\alpha c_j x^k)$ (式 7.1) 这是"记号 4"里那个生成函数的精确写法。它把第 j 个变量的所有可能取值 $x = 1, 2, \dots, P_j$ 对应的复指数 $e(\alpha c_j x^k)$ 加起来。注意指数里是 $c_j x^k$ (带系数)，不再是裸的 x^k 。把 s 个这样的和式相乘、乘 $e(-N\alpha)$ 、对 α 从 0 到 1 积分，由"筛子"就得到表示个数：

$$r(N) = \int_0^1 T_1(\alpha) \cdots T_s(\alpha) e(-N\alpha) d\alpha.$$

本章余下所有工作，都是估计这个积分。

第 4 段：Weyl 不等式仍然适用（次弧上" T_j 不会太大"）

原文译文：Weyl 不等式（引理 3.1）适用于和式 $T_j(\alpha)$ ；若 $\alpha = a/q + \beta$ 且 $|\beta| < q^{-2}$ ，则 $c_j \alpha = c_j a/q + c_j \beta$ ，且 $|c_j \beta| \ll q^{-2}$ 。这对于引理 3.1 的证明而言已经足够，因为关于 β 我们所用到的全部仅是 $|\beta| \ll q^{-2}$ 。

这一段在讲什么 检查圆法的第一个零件——Weyl 不等式——在加了系数后**照样能用**，只需把 α 换成 $c_j \alpha$ 走一遍。

Weyl 不等式（引理 3.1）是干什么的 它是次弧分析的主力。粗略地说，当 α 能被一个分母不太大的分数 a/q 很好地逼近时，指数和 $\sum_x e(\alpha x^k)$ 不会接近它的最大可能值 P ，而是被压低（出现一个 P 的负幂次的节省）。直观原因： $e(\alpha x^k)$ 这些单位圆上的点，当 α 不是"非常接近小分母分数"时，方向四面八方、互相抵消，和就小。这正是用来证明"次弧贡献只是误差"的关键。

$\alpha = a/q + \beta$ 是什么 这是**Dirichlet 逼近**：任何实数 α 都可以写成一个分数 a/q (q 不太大、 $(a, q) = 1$) 加上一个很小的余项 β ，且 $|\beta|$

为什么 Weyl 不等式对 T_j 仍成立

- 1 $T_j(\alpha) = \sum_x e(\alpha c_j x^k) = \sum_x e((c_j \alpha) x^k)$ 。也就是说，把系数 c_j 并进 α 里， $T_j(\alpha)$ 就是无系数的 Weyl 和，只不过自变量从 α 变成 $c_j \alpha$ 。
- 2 引理 3.1 的证明只需要"自变量能被某个小分母分数逼近、误差 $\ll q^{-2}$ "。现在把 $\alpha = a/q + \beta$ 乘以 c_j ： $c_j \alpha = c_j \cdot \frac{a}{q} + c_j \beta = \frac{c_j a}{q} + c_j \beta$ 。新的"分数部分"是 $c_j a/q$ （分母仍不超过 q ），新的"误差"是 $c_j \beta$ 。
- 3 误差的大小： $|(c_j \alpha) - \frac{c_j a}{q}| = |c_j \beta|$ 固定常数（与 N 、与 q 无关），所以 $c_j |\beta| \ll q^{-2}$ （被吸进 $\ll$ 的常数里）。这正是引理 3.1 唯一需要的关于误差的条件。
- 4 所以引理 3.1 原样适用于 $c_j \alpha$ ，即适用于 $T_j(\alpha)$ 。第二个零件，验收通过。

第 5 段：Hua 不等式仍然成立（关键的均值估计）

原文译文：Hua 不等式（引理 3.2）对任何一个这样的和式依然成立，因为

$$\int_0^1 |T_j(\alpha)|^{2^k} d\alpha = \int_0^1 \left| \sum_{x=1}^{P_j} e(\alpha c_j x^k) \right|^{2^k} d\alpha = \frac{1}{c_j} \int_0^{c_j} \left| \sum_{x=1}^{P_j} e(\alpha x^k) \right|^{2^k} d\alpha = \int_0^1 \left| \sum_{x=1}^{P_j} e(\alpha x^k) \right|^{2^k} d\alpha$$

这一段在讲什么 验收第三个零件——Hua 不等式。它是一个"平均意义"的估计：把 $|T_j|$ 升到 2^k 次方再在 $[0, 1]$ 上积分（求平均），结果不会太大。这一段用一串变量替换，把带系数的积分化回无系数的、已经证过的情形。下面把这串等号每一步都讲清。

Hua 不等式（引理 3.2）是什么、为什么重要 它断言

$\int_0^1 \left| \sum_{x=1}^P e(\alpha x^k) \right|^{2^k} d\alpha \ll P^{2^k - k + \varepsilon}$ 。左边的积分有个组合意义：它正好数"方程组 $x_1^k + \cdots - y_{2^{k-1}}^k = 0$ 的解数"（把 2^k 个因子配对成正负）。Hua

证明这个解数比"平凡上界"小了一个 P^k 的因子，这个节省是控制次弧、保证圆法跑通的命脉。这里我们只需知道：**它对带系数的 T_j 还成立。**

2^k 这个指数是哪来的 它是 Hua 不等式天然出现的指数（Hua 用反复的平方/差分技术，每做一次幂次翻倍，做 k 次得到 2^k ）。这也是为什么本章主定理 7.1 的门槛是 $s \geq 2^k + 1$ 。不必深究证明，只需把 2^k 当作"Hua 方法配套的项数"。

逐个等号讲清那串积分

1 第一个等号（写开定义）： $|T_j(\alpha)|^{2^k} = \left| \sum_{x=1}^{P_j} e(\alpha c_j x^k) \right|^{2^k}$ 。只是把 T_j 的定义代回去，没做任何操作。

2 第二个等号（变量替换 $\beta = c_j \alpha$ ）：令 $\beta = c_j \alpha$ ，则 $d\beta = c_j d\alpha$ ，即 $d\alpha = \frac{1}{c_j} d\beta$ 。当 α 从 0 走到 1 时， β 从 0 走到 c_j 。代入得

$$\int_0^1 \left| \sum_x e(c_j \alpha x^k) \right|^{2^k} d\alpha = \frac{1}{c_j} \int_0^{c_j} \left| \sum_x e(\beta x^k) \right|^{2^k} d\beta.$$

（把哑变量 β 再叫回 α 就是原文写法。）注意指数里现在是裸的 x^k ，系数被替换"吸收"掉了——代价是积分上限从 1 变成 c_j 、前面多了个 $\frac{1}{c_j}$ 。

3 第三个等号（周期性把上限拉回 1）：被积函数 $g(\alpha) = \left| \sum_x e(\alpha x^k) \right|^{2^k}$ 是**周期为 1** 的函数（因为 e 周期为 1，指数里 x^k 是整数， α 加 1 整个不变）。一个周期为 1 的函数，在任意长度为 1 的区间上积分都相等；在长度为 c_j （ c_j 是整数）的区间 $[0, c_j]$ 上积分，正好等于 c_j 个周期，于是

$$\frac{1}{c_j} \int_0^{c_j} g = \frac{1}{c_j} \cdot c_j \int_0^1 g = \int_0^1 g.$$

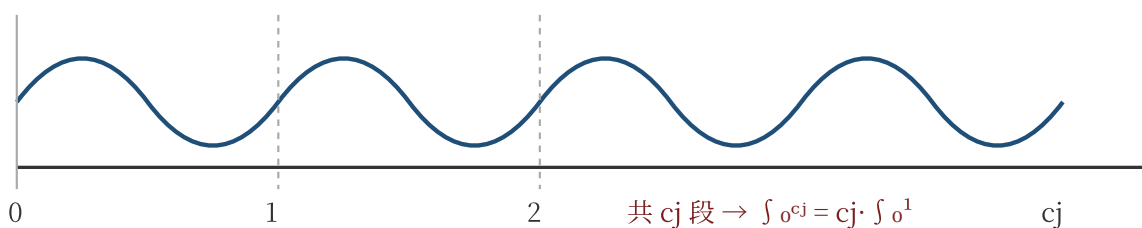
$\frac{1}{c_j}$ 和 c_j 抵消，干净地回到了 $[0, 1]$ 。

4 最后一步（用无系数的 Hua 不等式）：现在积分

$\int_0^1 \left| \sum_{x=1}^{P_j} e(\alpha x^k) \right|^{2^k} d\alpha$ 是纯无系数的、上限为 P_j 的标准 Hua 积分，引理 3.2 直接给出 $\ll P_j^{2^k-k+\varepsilon}$ 。

小结：通过"替换 $\beta = c_j \alpha$ + 利用周期性"两步，带系数的均值估计完全化归为已证的无系数情形，且误差量级一字不差（ c_j 是常数，被 $\ll$ 吸收）。这正是"只需微小改动"的具体体现。 ■

每段长度 1 的积分都一样（周期性）



周期为 1 的函数：在 $[0, c_j]$ 上的积分就是 c_j 个相同的"单位块"，所以 $\frac{1}{c_j} \int_0^{c_j} f(\alpha) d\alpha = \int_0^1 f(\alpha) d\alpha$ 。

第 6 段：用 Hölder 不等式推广到乘积（式 7.2）

原文译文：这个不等式同样可经 Hölder 不等式推广到任意 2^k 个和式之积；我们得到

$$\int_0^1 |T_1(\alpha) \cdots T_{2^k}(\alpha)| d\alpha \ll (P_1 \cdots P_{2^k})^{1-k/2^k+\varepsilon}. \quad (7.2)$$

这一段在讲什么 上一段是"同一个 T_j 自乘 2^k 次"的均值；但圆法里要估计的是不同的 $T_1, \dots, T_{2^k}$ 相乘的积分。这一段用 Hölder 不等式把"自乘"的结论搬到"互乘"，得到次弧分析直接要用的式 (7.2)。

Hölder 不等式是什么 它是高中柯西不等式的推广。柯西不等式说

$\int |fg| \leq (\int |f|^2)^{1/2} (\int |g|^2)^{1/2}$ 。Hölder 把"2 个、平方"推广到" n 个、各自适当的幂"：对 n 个函数 $f_1, \dots, f_n$,

$$\int |f_1 \cdots f_n| \leq \prod_{i=1}^n \left(\int |f_i|^n \right)^{1/n}.$$

直观就是"乘积的平均 $\leq$ 各自 n 次方平均的几何平均"。

从 (Hua) 推出 (7.2)

1 取 $n = 2^k$, 函数为 $T_1, \dots, T_{2^k}$ 。Hölder 给出

$$\int_0^1 |T_1 \cdots T_{2^k}| d\alpha \leq \prod_{j=1}^{2^k} \left(\int_0^1 |T_j|^{2^k} d\alpha \right)^{1/2^k}.$$

这一步把"互乘的积分"上界换成"每个 T_j 自乘 2^k 次的积分"的几何平均——而后者正是上一段算过的!

2 代入上一段的 Hua 估计 $\int_0^1 |T_j|^{2^k} d\alpha \ll P_j^{2^k-k+\varepsilon}$:

$$\leq \prod_{j=1}^{2^k} (P_j^{2^k-k+\varepsilon})^{1/2^k} = \prod_{j=1}^{2^k} P_j^{(2^k-k+\varepsilon)/2^k}.$$

3 化简指数: $\frac{2^k - k + \varepsilon}{2^k} = 1 - \frac{k}{2^k} + \frac{\varepsilon}{2^k}$ 。把 $\varepsilon/2^k$ 仍记为 ε (任意小的正数除以常数还是任意小), 得每个因子是 $P_j^{1-k/2^k+\varepsilon}$ 。

4 把 2^k 个因子的幂合并 (同指数相乘 = 底数相乘后取该指数):

$$\prod_{j=1}^{2^k} P_j^{1-k/2^k+\varepsilon} = (P_1 P_2 \cdots P_{2^k})^{1-k/2^k+\varepsilon}.$$

这正是式 (7.2)。 ■

为什么这条结论有用: 指数 $1 - k/2^k < 1$ 说明乘积积分比"平凡上界 $P_1 \cdots P_{2^k}$ "小了一个正幂次。这个"节省"就是后面证明"次弧只是误差"的弹药。

第 7 段: 主弧次弧的划分; 引理 4.1 (次弧贡献是误差)

原文译文：我们像之前一样定义主弧与次弧。引理 4.1 现在表明：若 $s \geq 2^k + 1$ ，则 $\int_{\mathfrak{m}} |T_1(\alpha) \cdots T_s(\alpha)| d\alpha \ll P^{s-k-\delta'}$ ，其证明与之前相同，需用到 (7.2)。引理 4.2 保持不变，只是其中的 $I(\beta)$ 要替换为 $I_j(\beta) = \int_0^{P_j} e(\beta c_j \xi^k) d\xi$ 。

这一段在讲什么 宣布主弧/次弧按老办法划分；引用引理 4.1 说明次弧上的积分确实小（是误差项），其证明只是把式 (7.2) 接进去；并指出引理 4.2 里只需把单个积分 $I(\beta)$ 换成带系数的 $I_j(\beta)$ 。

$\mathfrak{M}$ （主弧）

围绕小分母分数 a/q 的短区间之并，贡献主项。

$\mathfrak{m}$ （次弧）

$[0, 1]$ 去掉主弧后剩下的部分，贡献误差。手写体 $\mathfrak{m}$ 读 "m"。

P

$P = N^{1/k}$ ，是无系数版的"标准尺度"。所有 $P_j \leq P$ （因为 $c_j \geq 1$ ），且 P_j 与 P 只差一个常数倍，量级相同。

δ'

某个固定的小正数。指数 $s - k - \delta'$ 比"主项量级 $s - k$ " 小了 δ' ，所以次弧确实更小。

引理 4.1 的逻辑（为什么次弧是误差）

1 目标是说明 $\int_{\mathfrak{m}} |T_1 \cdots T_s| d\alpha$ 比主项 P^{s-k} 小一个正幂次。

2 分离出一个最大值因子：把 s 个因子拆成 " 2^k 个放进积分" 和 "剩下 $s - 2^k$ 个取它们在次弧上的最大值"。具体地，

$$\int_{\mathfrak{m}} |T_1 \cdots T_s| \leq \left(\max_{\mathfrak{m}} |T_{2^k+1} \cdots T_s| \right) \int_{\mathfrak{m}} |T_1 \cdots T_{2^k}|.$$

3 对最大值因子用 Weyl 不等式（第 4 段验收过它对 T_j 成立）：在次弧上每个 $|T_j| \ll P^{1-\sigma}$ 有一点节省（ $\sigma > 0$ 与 k 有关）。 $s - 2^k$ 个相乘得

$\max \ll P^{(s-2^k)(1-\sigma)}$, 关键是带一个 P 的负幂节省。

4 对积分因子用 (7.2):

$$\int_m |T_1 \cdots T_{2^k}| \leq \int_0^1 |T_1 \cdots T_{2^k}| \ll (P_1 \cdots P_{2^k})^{1-k/2^k+\varepsilon} \ll P^{2^k-k+\varepsilon}$$

(每个 $P_j \ll P$)。

5 两块相乘, 把次弧上的节省凑起来, 正好得到 $\ll P^{s-k-\delta'}$, 比主项小 δ' 个幂。条件 $s \geq 2^k + 1$ 保证"放进积分的 2^k 个之外, 至少还剩 1 个因子可供 Weyl 挤出节省"。 ■

带系数后改了什么: 什么都没改实质——只是 Weyl 和 (7.2) 都已验收对 T_j 成立, 证明逐字照搬。这就是"其证明与之前相同"。

引理 4.2 与 $I_j(\beta)$ 引理 4.2 负责主弧上的逼近: 在主弧上 (α 接近 a/q), 把离散和 $T_j(\alpha)$ 近似成"一个高斯型指数和 S 乘上一个积分 I "。带系数后, 那个连续积分由

$$I_j(\beta) = \int_0^{P_j} e(\beta c_j \xi^k) d\xi$$

代替 (ξ 是连续变量, 对应离散的 x ; β 是 α 与 a/q 的差)。和 (7.1) 一样, 指数里多了系数 c_j 、上限是各自的 P_j 。"引理 4.2 保持不变"意思是逼近的形式和误差都一样, 只需把符号换成带 j 的版本。

第 8 段: 把上限统一替换成 $Pc_j^{-1/k}$ 以简化计算

原文译文: 不过, 若把上限替换为 $Pc_j^{-1/k}$ (其中 $P = N^{1/k}$), 后面的计算会稍微简化一些; 二者之差可忽略不计。

这一段在讲什么 一个技术上的小取舍: 把 $I_j(\beta)$ 的积分上限 $P_j = [(N/c_j)^{1/k}]$ 改成更整齐的 $Pc_j^{-1/k}$, 让后面奇异积分的计算更顺手; 并解释这样替换不影响结果。

- 1 原上限 $P_j = [(N/c_j)^{1/k}]$ 。先不取整：
 $(N/c_j)^{1/k} = N^{1/k} \cdot c_j^{-1/k} = P c_j^{-1/k}$ (用 $P = N^{1/k}$)。所以 $P c_j^{-1/k}$ 就是"没取整的精确实数上限"，而 P_j 是它的整数部分。
- 2 二者之差不超过 1: $0 \leq P c_j^{-1/k} - P_j < 1$ 。把积分上限挪动不到 1 个单位，被积函数 $|e(\cdots)| = 1$ 有界，所以 $I_j(\beta)$ 的改变量不超过这段长度乘上界，是 $O(1)$ 级——相对于 I_j 本身的量级 P 来说微乎其微，可以忽略。
- 3 为什么选 $P c_j^{-1/k}$ 而不是别的：因为它能让接下来的换元一步到位（见第 10 段：令 $\eta = c_j^{1/k} \xi / P$ 时积分上限正好变成 1）。整齐的上限让奇异积分变成统一的标准形 $\int_0^1 e(\gamma \eta^k) d\eta$ ，不必处理零碎的取整。这是"为了算得干净"的合理取舍，且因第 2 步的误差可忽略，所以不损失精度。

第 9 段：引理 4.3 与定理 4.1；带系数的奇异级数（式 7.3）

原文译文：引理 4.3 与定理 4.1 的证明只需作出极轻微的改动即可适用于当前问题。一处差别在于：为把 $I_j(\beta)$ 表示成 $\int_0^1 e(\gamma \xi^k) d\xi$ 的形式而在其中作的变量替换，会产生一个因子 $|c_j|^{-1/k}$ 。另一处差别是：奇异级数现在具有稍微更一般的形式；它现在由下式给出

$$\mathfrak{S}(N) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^q q^{-s} S_{c_1 a, q} \cdots S_{c_s a, q} e(-Na/q). \quad (7.3)$$

这一段在讲什么 这是主弧计算的收官。指出带系数后只有两处差别：① 把 I_j 标准化时每个变量挤出一个 $c_j^{-1/k}$ 因子（它们乘起来就是定理 7.1 里的 $(c_1 \cdots c_s)^{-1/k}$ ）；② 奇异级数 $\mathfrak{S}(N)$ 里每个高斯和 $S_{a,q}$ 升级成带系数的 $S_{c_j a, q}$ 。这两处差别正是定理 7.1 公式相对无系数版的全部变化。

第一处差别：因子 $c_j^{-1/k}$ 怎么冒出来的

1 主弧逼近里出现连续积分 $I_j(\beta) = \int_0^{Pc_j^{-1/k}} e(\beta c_j \xi^k) d\xi$ (用上一段的整齐上限)。我们想把它化成统一的标准形 $\int_0^1 e(\gamma \eta^k) d\eta$ 。

2 换元：令 $\eta = \frac{c_j^{1/k} \xi}{P}$ ，则 $\xi = \frac{P}{c_j^{1/k}} \eta$ ， $d\xi = \frac{P}{c_j^{1/k}} d\eta = P c_j^{-1/k} d\eta$ 。
当 ξ 从 0 到 $Pc_j^{-1/k}$ 时， η 从 0 到 1。指数里
 $c_j \xi^k = c_j \cdot \frac{P^k}{c_j} \eta^k = P^k \eta^k$ ，系数 c_j 恰好被消掉，指数变成 $\beta P^k \eta^k$ 。

3 于是 $I_j(\beta) = P c_j^{-1/k} \int_0^1 e(\beta P^k \eta^k) d\eta$ 。换元的 Jacobi 因子 $d\xi = P c_j^{-1/k} d\eta$ 里那个 $c_j^{-1/k}$ ，就是原文说的"产生一个因子 $|c_j|^{-1/k}$ "。每个变量 j 贡献一个这样的因子， s 个相乘得到 $(c_1 c_2 \cdots c_s)^{-1/k}$ ——这正是定理 7.1 主项里那个分母 $(c_1 \cdots c_s)^{1/k}$ 的来源。

(绝对值 $|c_j|$ 在本章里 $c_j > 0$ 就是 c_j 自己。)

第二处差别：奇异级数 $\mathfrak{S}(N)$ (式 7.3) 逐符号拆解 先认识高斯和：

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

是一个"完整周期"上的指数和 (x 跑遍模 q 的一组代表 $1, \dots, q$)。它打包了"模 q 下 k 次方分布"的信息。带系数后把 a 换成 $c_j a$ ，得

$S_{c_j a, q} = \sum_{x=1}^q e\left(\frac{c_j a x^k}{q}\right)$ 。现在看式 (7.3) 的每一块：

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

对所有可能的分母 q 求和 (这对应主弧围绕的所有分数 a/q)。

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

对每个 q , a 跑遍与 q 互素、且在 1 到 q 之间的那些数。" $(a, q) = 1$ " 表示 a 与 q 的最大公约数是 1, 即既约分数 a/q 。这保证每个分数只数一次。

$$q^{-s}$$

归一化因子 (每个 S 大约 $q^{1-1/k}$ 量级, s 个相乘后除以 q^s 让级数有机会收敛)。

$$S_{c_1 a, q} \cdots S_{c_s a, q}$$

s 个带系数高斯和的乘积——这是相对无系数版唯一的实质变化 (无系数版是 $S_{a, q}^s$)。

$$e(-Na/q)$$

"筛子"留下的目标项, 对应主弧中心 a/q 处要凑出 N 。

它为什么叫"奇异级数"、为什么重要: 因为它把所有素数的同余信息汇总成一个数。可以证明它能写成对各素数的"局部密度"连乘 (见下一段引理 5.1 的因子分解)。如果某个素数让局部无解, 对应因子为 0, 整个 $\mathfrak{S}(N) = 0$, 于是主项消失——这与第 1 段反例完全呼应。所以 " $\mathfrak{S}(N) > 0$ " 正是 " N 可表示" 的算术判据。

第 10 段: 把 $S_{a, q}$ 的估计推广到 $S_{ca, q}$ (保证绝对收敛)

原文译文: 为了利用引理 3.1 的推论 (或当 $s \geq 2k + 1$ 时利用引理 6.4) 来确立此级数在 $s \geq 2^k + 1$ 时的绝对收敛性, 我们需要把对 $S_{a, q}$ (当 $(a, q) = 1$ 时) 的估计加以推广, 使之适用于 $S_{ca, q}$, 其中 c 是任一固定的正整数。这很容易办到, 因为若 $ca/q = a'/q'$, 则 $S_{ca, q} = \frac{q}{q'} S_{a', q'}$, 而 q/q' 作为 c 的因子是有界的。

这一段在讲什么 要证明无穷级数 (7.3) 真的收敛 (不会发散成无穷), 就得知每个 $S_{c_j a, q}$ 有多大。已有的估计是针对既约 $S_{a, q}$ ($(a, q) = 1$) 的, 但 $c_j a/q$ 不一定既约。这一段用一个约分恒等式把带系数的高斯和换算回既约的, 从而沿用已有估计。

"绝对收敛"是什么 一个无穷级数 $\sum u_n$ 绝对收敛, 指 $\sum |u_n|$ (把每项取绝对值/模长再加) 也收敛到有限值。绝对收敛比普通收敛更"稳": 可以随便交换求和次序、做因子分解, 结果不变。本章需要它来对 $\mathfrak{S}(N)$ 做后面的素数因子分解 (引理 5.1)。

为什么需要 $s \geq 2^k + 1$ (或 $s \geq 2k + 1$) 收敛靠的是 q^{-s} 这个分母"压"得够狠。已知 $|S_{a,q}| \ll q^{1-1/k}$ (既约高斯和的标准估计, 来自引理 3.1 的推论; 引理 6.4 给出更省的版本, 对 $s \geq 2k + 1$ 够用)。于是 (7.3) 的一般项绝对值 $\ll q^{-s} \cdot q \cdot (q^{1-1/k})^s = q^{1-s/k}$ (内层对 a 求和约 q 项)。当 $1 - s/k < -1$, 即 $s > 2k$, 级数 $\sum_q q^{1-s/k}$ 像 $\sum q^{-(>1)}$ 一样收敛。所以项数门槛保证收敛。

把 $S_{ca,q}$ 换算成既约和

① 分数 ca/q 约分到既约: 设 $\frac{ca}{q} = \frac{a'}{q'}$, 其中 $(a', q') = 1$ 。约掉的公因子是 $g = \gcd(ca, q)$, 于是 $q' = q/g$, 即 $q = gq'$ 。

② 恒等式 $S_{ca,q} = \frac{q}{q'} S_{a',q'}$: 理由是 $e(ca x^k/q) = e(a' x^k/q')$ (同一个数 $ca/q = a'/q'$)。当 x 跑遍模 q 的 q 个值时, $x^k \bmod q'$ 的分布是"模 q' 的分布重复了 q/q' 次", 所以

$$S_{ca,q} = \frac{q}{q'} \sum_{x=1}^{q'} e(a' x^k/q') = \frac{q}{q'} S_{a',q'}.$$

③ 这个比值有界: $\frac{q}{q'} = g = \gcd(ca, q)$ 必整除 $c \cdot \gcd(a, q)$ 。由于在 (7.3) 里 $(a, q) = 1$, 所以 $g \mid c$, 即 $\frac{q}{q'}$ 是 c 的因子, 最多等于 c 。因为 c 是固定常数, 所以 q/q' 被一个常数 (c) 一致地控制住。

④ 于是 $|S_{ca,q}| = \frac{q}{q'} |S_{a',q'}| \leq c |S_{a',q'}| \ll q'^{1-1/k} \leq q^{1-1/k}$ 。带系数高斯和的估计与既约情形同量级 (只差常数 c), 收敛证明照常成立。 ■

第 11 段：定理 7.1（本章主结果）

原文译文（定理 7.1）：设 $c_1, \dots, c_s$ 为固定的正整数。则当 $s \geq 2^k + 1$ 时，把 N 表示为 $N = c_1 x_1^k + \dots + c_s x_s^k$ ($x_1, \dots, x_s > 0$) 的表示个数 $r(N)$ 满足

$$r(N) = \frac{C_{k,s}}{(c_1 c_2 \dots c_s)^{1/k}} N^{s/k-1} \mathfrak{S}(N) + O(N^{s/k-1-\delta}) \quad (7.4)$$

其中 $\delta > 0$ 为某个固定常数， $C_{k,s}$ 同定理 4.1 中所定义，而 $\mathfrak{S}(N)$ 由 (7.3) 定义。该级数 (7.3) 在 $s \geq 2k + 1$ 时绝对收敛。

定理 7.1（带系数的渐近公式） 前面所有零件（生成函数、Weyl、Hua、(7.2)、主弧逼近、奇异级数）拼起来，就得到表示个数 $r(N)$ 的精确渐近式 (7.4)。它把 $r(N)$ 写成"主项 + 误差"。

把 (7.4) 每个因子读懂

$r(N)$

满足 $N = c_1 x_1^k + \dots + c_s x_s^k$ 的正整数解组 $(x_1, \dots, x_s)$ 的个数。这是我们最想知道的量。

$C_{k,s}$

只依赖 k, s 的正常数（与定理 4.1 即无系数情形相同），来自奇异积分 $\int e(\gamma \eta^k) d\eta$ 的标准计算，本质是 $\Gamma(1 + \frac{1}{k})^s / \Gamma(\frac{s}{k})$ 一类的体积常数。

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

分母里的系数因子，正是第 9 段每个变量挤出的 $c_j^{-1/k}$ 之积。系数越大，每项可用范围越小，表示数越少——它出现在分母上，完全合理。

$N^{s/k-1}$

主项的**增长量级**。来源：每个变量约有 $P = N^{1/k}$ 个取值， s 个变量共 $N^{s/k}$ 种组合；落在"恰好等于 N "这张超曲面上的占比约 N^{-1} （一维约束）。相乘得 $N^{s/k-1}$ 。当 $s > k$ 时这是趋于无穷的正幂，意味着表示数会变多。

$\mathfrak{S}(N)$

奇异级数（式 7.3），同余信息的总和。它是**决定主项是否非零**的关键——这是定理 7.2、7.3 要处理的。

$O(N^{s/k-1-\delta})$

误差项，比主项小 δ 个 N 的幂。它来自次弧（引理 4.1）和主弧逼近的余项。只要它真的更小，主项就主导大局。

这个公式的意义 当 $N \rightarrow \infty$ ：如果 $\mathfrak{S}(N)$ 有一个不依赖 N 的正下界，那么主项 $\asymp N^{s/k-1}$ 趋于无穷、且压过误差，于是 $r(N) \rightarrow \infty$ ，特别地 $r(N) > 0$ ，即 N 可表示。反之若 $\mathfrak{S}(N)$ 可以是 0（某素数局部无解），公式就退化、不能保证可表示。所以下面的全部任务，就是给 $\mathfrak{S}(N)$ 找正下界。注意收敛要 $s \geq 2k + 1$ ，而整个渐近式要 $s \geq 2^k + 1$ （因为次弧分析用到 Hua 的 2^k ），后者更强。

第 12 段：奇异级数与 $M(p^\nu)$ 的关系（引理 5.1–5.3）

原文译文：关于奇异级数的因子分解、以及奇异级数与 $M(p^\nu)$ 之间关系的引理 5.1、5.2、5.3 仍然适用。于是，为使 $\mathfrak{S}(N)$ 具有一个与 N 无关的正下界，只需对每个 p ，使同余式 $c_1 x_1^k + \cdots + c_s x_s^k \equiv N \pmod{p^\nu}$ 的解数 $M(p^\nu)$ 满足 $M(p^\nu) \geq C_p p^{\nu(s-1)}$ 对一切足够大的 ν 成立。

这一段在讲什么 把" $\mathfrak{S}(N)$ 有正下界"这个解析问题，**翻译成一个纯算术问题**：只要每个素数 p 的同余式解数 $M(p^\nu)$ 足够多（达到期望量级 $p^{\nu(s-1)}$ ），奇异级数就恒正。这是引理 5.1–5.3 的作用，本章说它们"仍适用"。

因子分解（引理 5.1） 奇异级数能写成对所有素数的连乘：

$\mathfrak{S}(N) = \prod_p \chi_p(N)$ ，每个 $\chi_p(N)$ 叫" p -adic 局部密度"，只依赖单个素数 p 的同余信息。因为绝对收敛（第 10 段），这个乘积有意义。**整体正 $\iff$ 每个因子正且不退化。**

$M(p^\nu)$ 是什么 它是同余式 $c_1x_1^k + \cdots + c_sx_s^k \equiv N \pmod{p^\nu}$ 的解的个数 (每个 x_j 在模 p^ν 的 p^ν 个值里取)。总组合数是 $(p^\nu)^s = p^{\nu s}$; 这些组合落在 " $\equiv N$ " 这一个同余类上的 "期望比例" 是 $1/p^\nu$, 所以 "健康" 的解数应当约为 $p^{\nu s}/p^\nu = p^{\nu(s-1)}$ 。引理 5.2/5.3 把局部密度 χ_p 与 $M(p^\nu)$ 联系起来:

$$\chi_p = \lim_{\nu \rightarrow \infty} M(p^\nu)/p^{\nu(s-1)}.$$

为什么 " $M(p^\nu) \geq C_p p^{\nu(s-1)}$ " 就够

- 1 若对每个 p 都有 $M(p^\nu) \geq C_p p^{\nu(s-1)}$, 则局部密度 $\chi_p = \lim M(p^\nu)/p^{\nu(s-1)} \geq C_p > 0$ 不退化为 0。
- 2 对所有 p 连乘: 只要这些 C_p 不会 "乘到 0" (对大 p 它们趋于 1, 乘积收敛到正数——这由收敛性 $s \geq 2k+1$ 保证), 就有 $\mathfrak{S}(N) = \prod_p \chi_p \geq$ 某个正常数。
- 3 这个下界不依赖 N , 因为脚注说明只需对有限个小素数 $p \leq p_0$ 验证 (大素数自动满足), 而每个 C_p 与 N 无关。 ■

所以问题进一步简化为: "对每个 (小) 素数, 同余式解数够多吗?" 下一段给出 "够多" 的充分条件。

第 13 段: 引理 5.4、5.5 — 把 "解数多" 化为 "有一个非平凡解" (式 7.5)

原文译文: 像之前一样定义 γ , 并利用引理 5.4, 我们如同在引理 5.5 中那样发现: 上述条件的一个充分条件是同余式 $c_1x_1^k + \cdots + c_sx_s^k \equiv N \pmod{p^\nu}$ 有一解, 其中 $c_1x_1^k, \dots, c_sx_s^k$ 不全被 p 整除。

这一段在讲什么 再把条件大幅简化: 要保证 "解数多 $M(p^\nu) \geq C_p p^{\nu(s-1)}$ ", 其实只需找到一个 "非平凡" 的解 (不全被 p 整除的那一组)。这把 "计数下

界"问题降为"存在性"问题，好验证得多。这是引理 5.4、5.5 的功劳。

γ

由 (5.11) 定义的一个临界指数（依赖 k 与 p ），它是"看清同余结构所需的最小方幂层数"。原理是：一旦在模 p^γ 找到一个非平凡解，就能用" p -adic 提升 / Hensel 型引理 5.4"把它**逐层抬到任意高的** p^ν ，而且抬升的方式多到产生足够的解。后面定理 7.3 的证明里 $\gamma = \tau + 2$ ($p = 2$) 或 $\gamma - 1 = \tau$ 等具体值就来自这里。

为什么"一个非平凡解"能撬动"很多解"（引理 5.5 思想）

- 1 假设在模 p^γ 找到一组解 $(x_1^0, \dots, x_s^0)$ ，其中至少有一项 $c_j x_j^k$ 不被 p 整除——不妨设 $x_1^0 \not\equiv 0 \pmod{p}$ 。
- 2 **提升 (lifting)**：固定 $x_2, \dots, x_s$ ，把方程看成只关于 x_1 的同余式。因为 x_1^0 处的"导数" $kc_1(x_1^0)^{k-1}$ 不被 p 整除（ $x_1^0 \not\equiv 0$ 且 $p \nmid kc_1$ 时），由引理 5.4（Hensel 引理的版本）可知：模 p^γ 的每个解都能**唯一地**提升为模 $p^{\gamma+1}, p^{\gamma+2}, \dots, p^\nu$ 的解。
- 3 **数解**：另外 $s - 1$ 个变量 $x_2, \dots, x_s$ 在模 p^ν 下可自由取（约 $p^{\nu(s-1)}$ 种），每种都通过提升配上一个 x_1 。于是解数 $M(p^\nu) \gtrsim p^{\nu(s-1)}$ ，正好达到第 12 段要的量级 $C_p p^{\nu(s-1)}$ 。
- 4 反过来，如果只有"全被 p 整除"的平凡解，提升就会失效（导数被 p 整除，Hensel 不适用），解数上不去。所以"**非平凡**"这个词是命脉——它正是第 2 段那个假设的精确化。 ■

第 14 段：定理 7.2（同余条件 $\Rightarrow$ 奇异级数有正下界）

原文译文（定理 7.2）：设 γ 由 (5.11) 定义。假设 $s \geq 2k + 1$ ，并假设对每个^(脚注)素数 p ，同余式 (7.5) 有一个解，其中 $c_1 x_1^k, \dots, c_s x_s^k$ 不全被 p 整除。则对一切满足此假

设的 N ，我们有 $\mathfrak{S}(N) \geq C(k, s) > 0$ 。

定理 7.2 (奇异级数正下界的充分条件) 把第 12、13 段串起来的正式结论：只要每个素数都有一个非平凡的局部解，奇异级数就有一个不依赖 N 的正下界。逻辑链是：非平凡局部解 (13 段) $\Rightarrow M(p^\nu) \geq C_p p^{\nu(s-1)}$ (12 段) $\Rightarrow$ 每个局部密度 $\chi_p > 0 \Rightarrow \mathfrak{S}(N) = \prod_p \chi_p \geq C(k, s) > 0$ 。

脚注的意思 (为什么只需有限个素数) 原文脚注说"只需对 $p \leq p_0(k, s)$ 假设此条件"。原因：对足够大的素数 p (大于某个只依赖 k, s 的界 p_0)，由一般的同余存在性结果 (引理 5.2 的推论)，非平凡解自动存在，不必额外假设。所以真正要检查的只有有限个小素数。这也解释了下一段为什么说"满足条件的 N 包含某个算术级数中的一切数"。

$C(k, s)$ 一个只依赖 k 和 s (不依赖 N 也不依赖具体哪组系数) 的正常数。"不依赖 N "这点至关重要：它让定理 7.1 的主项随 $N \rightarrow \infty$ 真正发散。

第 15 段：合并定理 7.1 与 7.2 — $r(N) \rightarrow \infty$ (在算术级数上)

原文译文：由定理 7.1 与 7.2，若 $s \geq 2^k + 1$ ，则当 $N \rightarrow \infty$ 时 $r(N) \rightarrow \infty$ ，前提是把 N 限制为满足定理 7.2 中同余条件的那些数。由于该同余条件只在 $p \leq p_0$ 时才需要，而 γ 与 N 无关，因此满足该同余条件的数 N 将包括某些算术级数中的一切数。

这一段在讲什么 把两大定理合并，给出本章第一个完整结论：只要项数 $s \geq 2^k + 1$ ，并且 N 满足那些有限个素数的同余条件，那么 N 足够大时表示数趋于无穷 (自然就 > 0 ，即可表示)。而满足同余条件的 N 不是零散的，它们填满一整条算术级数。

合并的逻辑 + "算术级数"的解释

1 把定理 7.2 的 $\mathfrak{S}(N) \geq C(k, s) > 0$ 代入定理 7.1 的 (7.4):

$$r(N) \geq \frac{C_{k,s} C(k, s)}{(c_1 \cdots c_s)^{1/k}} N^{s/k-1} - O(N^{s/k-1-\delta}).$$

2 主项是 $N^{s/k-1}$ 的正常数倍, 误差是 $N^{s/k-1-\delta}$ (小 δ 个幂)。当 N 足够大, 主项压过误差, 故 $r(N) \rightarrow \infty$ 。需要 $s/k - 1 > 0$, 即 $s > k$ (在 $s \geq 2^k + 1$ 下显然成立)。

3 "算术级数"从哪来: 同余条件只涉及有限个素数 $p_1, \dots, p_m \leq p_0$, 每个要求 N 落在模 $p_i^{\gamma_i}$ 的某些剩余类里。由中国剩余定理, 这些条件合起来等价于 " N 落在模 $Q = \prod p_i^{\gamma_i}$ 的某些剩余类里"。固定一个这样的剩余类 $N \equiv N_0 \pmod{Q}$, 就得到一条算术级数 $N_0, N_0 + Q, N_0 + 2Q, \dots$, 其中每个 (足够大的) 数都可表示。 γ 与 N 无关, 保证了 Q 是固定的、级数是真实存在的。 ■

第 16 段: 引入"两两互素"假设

原文译文: 如果我们作出"系数 $c_1, \dots, c_s$ 两两互素"这一假设, 那么可以证明: 只要 s 超过 k 的某个特定函数, 则该同余条件对一切 N 都满足。我们证明 (定理 7.3) ……

这一段在讲什么 上一节的结论还带个 " N 要满足同余条件"的尾巴 (只对算术级数成立)。这一段提出: 如果对系数加一个干净的条件——**两两互素**——就能去掉这个尾巴, 让结论对一切大 N 成立。下一段定理 7.3 给出具体的项数门槛。

"两两互素 (relatively prime in pairs)" 指任意两个不同系数的最大公约数都是 1: 对 $\forall (1 \leq i < j \leq s)$ 对任一素数 p , 至多有一个系数被 p 整除 (否则那两个就有公因子 p), 于是局部分析时总有 $s - 1$ 个系数与 p 互素可用。

第 17 段：定理 7.3（两两互素下的项数门槛）

原文译文（定理 7.3）：假设对 $(1 \leq i \leq s)$ 。

定理 7.3 在两两互素的前提下，只要项数 s 超过上面那个关于 k 的具体门槛，对一切 N 都有 $\mathfrak{S}(N) \geq C(k, s) > 0$ （不再需要对 N 设同余条件）。门槛按 k 的奇偶分两档：

- k 奇： $s \geq k(2k - 1) + 2$;
- k 偶： $s \geq 2k(4k - 1) + 2$ （偶数情形门槛更高，因为 $p = 2$ 的同余结构更刁）。

下面整段证明就是验证："在这个门槛下，每个素数 p 都能找到 (7.5) 的非平凡解"，从而由定理 7.2 得到正下界。

第 18 段：定理 7.3 的证明（开头：化归到 p^γ ）

原文译文：我们要证明：在所述条件下，同余式 (7.5) 有一个解，其中 $c_1 x_1^k, \dots, c_s x_s^k$ 不全被 p 整除。由于 $c_1, \dots, c_s$ 中至多有一个能被 p 整除，因此只要解出 $c_1 x_1^k + \dots + c_s x_s^k \equiv N \pmod{p^\gamma}$ （式 7.6）就足够了，其中 $x_1, \dots, x_{s-1}$ 不全被 p 整除，且假定 $c_1, \dots, c_{s-1}$ 中没有一个能被 p 整除。

这一段在讲什么 证明的开局布置：把任务从"对所有 ν "降到"只在 p^γ 找非平凡解"（由 13 段提升原理， p^γ 够了）；并利用两两互素，把那个可能被 p 整除的系数"踢到一边"，只对剩下 $s - 1$ 个与 p 互素的系数工作。

开局每一步的理由

- 1 为什么只需 p^γ ：第 13 段已说明，在 p^γ 找到一个非平凡解后，能提升到任意 p^ν 。所以把目标定为式 (7.6)： $\sum c_j x_j^k \equiv N \pmod{p^\gamma}$ 。

2 两两互素的红利：若有两个系数都被 p 整除，则它们有公因子 p ，与两两互素矛盾。所以至多一个系数被 p 整除。把它编号为 c_s （放最后），则 $c_1, \dots, c_{s-1}$ 全都不被 p 整除。

3 非平凡性只用前 $s - 1$ 个变量保证：我们将让最后一项的变量随便取（甚至取使该项被 p 整除也无妨），而要求 $x_1, \dots, x_{s-1}$ 不全被 p 整除。因为 $c_1, \dots, c_{s-1}$ 不被 p 整除，只要某个 x_i ($i \leq s - 1$) 不被 p 整除，对应的 $c_i x_i^k$ 就不被 p 整除，非平凡性即满足。

4 于是问题归结为：在 $c_1, \dots, c_{s-1}$ 都与 p 互素的前提下，解 (7.6) 且前 $s - 1$ 个变量不全被 p 整除。下面分 $p > 2$ 与 $p = 2$ 两种情形。

第 19 段：证明（情形 $p > 2$ ，奇素数）

原文译文：设 $p > 2$ 。我们在引理 5.6 的证明中已看到：当 $z \not\equiv 0 \pmod{p}$ 时，模 p^γ 下由 z^k 所取的值的个数是 $(p - 1)/\delta$ ，其中 $\delta = (k_0, p - 1)$ 。因此模 p^γ 下 k 次幂剩余与非剩余的不同类的个数为 $\frac{\phi(p^\gamma)}{(p - 1)/\delta} = \frac{p^{\gamma-1}(p - 1)\delta}{p - 1} = p^{\gamma-1}\delta$ 。（这些类是在全体互素剩余类构成的群中、由 k 次幂所构成的子群的陪集。）

这一段在讲什么 奇素数情形的核心一步：数清楚“模 p^γ 下，把系数按 k 次幂特征分类，一共能分成几类”。答案是 $p^{\gamma-1}\delta$ 类。知道了类数，下一段就用抽屉原理找出“同类系数一大堆”。

欧拉函数 φ (ϕ) $\varphi(m)$ 表示 1 到 m 中与 m 互素的整数个数。对素数方幂 p^γ ：在 $1, \dots, p^\gamma$ 中，不与 p^γ 互素的就是 p 的倍数，共 $p^{\gamma-1}$ 个，于是

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

这些与 p^γ 互素的剩余类，在乘法下构成一个有 $\varphi(p^\gamma)$ 个元素的群（互素剩余类群）。

k_0, τ

把 k 写成 $k = p^\tau k_0$, 其中 $p \nmid k_0$ (k_0 是 k 去掉所有 p 因子后的部分, τ 是 k 含 p 的次数)。对奇素数情形下面会用到 $\gamma - 1 = \tau$ 。

$\delta = (k_0, p - 1)$

k_0 与 $p - 1$ 的最大公约数。它度量" k 次幂映射 $z \mapsto z^k$ 在互素剩余群里把多少个元素压成同一个像"。

逐式讲清"类数 $= p^{\gamma-1} \delta$ "

- 1 **k 次幂子群的阶**: 互素剩余群 $(\mathbb{Z}/p^\gamma)^\times$ 的阶是 $\varphi(p^\gamma) = p^{\gamma-1}(p - 1)$ 。其中全体非零 k 次幂剩余 $\{z^k : z \not\equiv 0\}$ 恰好构成一个子群 H , 引理 5.6 算出它的阶 (即 z^k 取到的不同值的个数) 正好是 $\#H = (p - 1)/\delta$ 。

- 2 **类 = 陪集**: 把"是否同一个 k 次幂特征"作为等价关系, 相当于按" k 次幂子群 H 的陪集"分类。陪集个数 = 群的阶 $\div$ 子群的阶 = (指数)。即

$$\# \text{类} = \frac{\varphi(p^\gamma)}{\#H} = \frac{p^{\gamma-1}(p - 1)}{(p - 1)/\delta}.$$

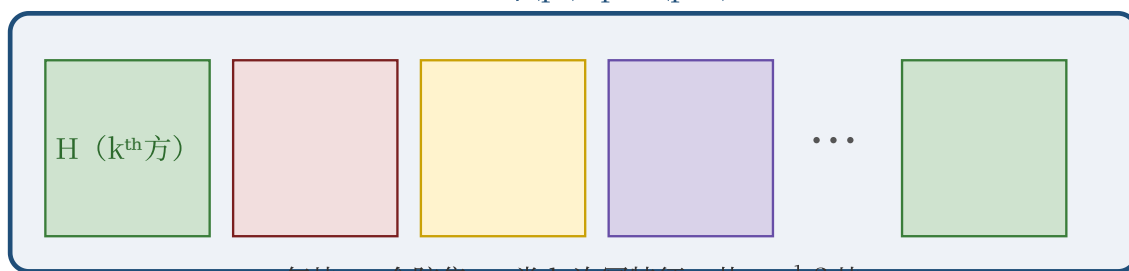
- 3 **化简**: 分子分母同时处理,

$$\frac{p^{\gamma-1}(p - 1)}{(p - 1)/\delta} = p^{\gamma-1}(p - 1) \cdot \frac{\delta}{p - 1} = p^{\gamma-1}\delta \quad ((p - 1) \text{ 约掉}).$$

所以类数 $= p^{\gamma-1}\delta$ 。

- 4 **陪集是什么 (括号那句)**: 互素剩余群里, " k 次方们" $\{z^k\}$ 构成一个子群 H ; 把整个群按 H 平移切块, 每块叫一个陪集, 块与块大小相同、互不相交, 块数就是上面算的 $p^{\gamma-1}\delta$ 。两个系数"同 k 次幂特征"当且仅当它们落在同一陪集。 ■

互素剩余群 (共 $\phi(p^\gamma)=p^{\gamma-1}(p-1)$ 个元素)



每块=一个陪集=一类 k 次幂特征, 共 $p^{\gamma-1}\delta$ 块

把 $s-1$ 个系数按"落在哪块陪集"分类, 最多 $p^{\gamma-1}\delta$ 块。系数多、块少, 必有一块装得满。

第 19 段 (续): 抽屉原理 + 化归到 Waring 同余

原文译文: 若我们按照系数 $c_1, \dots, c_{s-1}$ 所属的 k 次幂剩余或非剩余之类, 把它们分入若干集合, 则将有某一类至少包含 $(s-1)/p^{\gamma-1}\delta$ 个系数。设 t 为不小于 $(s-1)/p^{\gamma-1}\delta$ 的最小整数。我们可以把这些考虑中的系数取为前 t 个系数, 于是 $c_2 \equiv d_2^k c_1, \dots, c_t \equiv d_t^k c_1 \pmod{p^\gamma}$, 其中 d_i 不被 p 整除。令变量 $x_{t+1}, \dots$ 等于 0, 并约去 c_1 , 我们看到只需解出 $x_1^k + (d_2 x_2)^k + \dots + (d_t x_t)^k \equiv N' \pmod{p^\gamma}$, 其中诸变量不全被 p 整除。这实际上与引理 5.6 中关于 Waring 问题所考虑的同余式相同。我们在那里证明了: 只要 $t \geq 2k$, 该结论便成立。因此只需 $\frac{s-1}{p^{\gamma-1}\delta} > 2k-1$ 。由于 $\gamma-1=\tau$ 且 $p^\tau \delta \leq p^\tau k_0 = k$, 故只需 $s-1 > k(2k-1)$ 。

这一段在讲什么 把上一段数好的"类数"用**抽屉原理**转成"有一类系数特别多", 再把这一类系数 (同 k 次幂特征, 所以彼此差一个 k 次方因子) 凑成一个**标准的 Waring 同余**, 套用引理 5.6 的现成结论 $t \geq 2k$, 最后倒推出项数门槛 $s-1 > k(2k-1)$ 。这正是定理 7.3 奇数门槛的来历。

抽屉原理 把 $s-1$ 个东西放进 $p^{\gamma-1}\delta$ 个抽屉, 必有一个抽屉装了至少 $\frac{s-1}{p^{\gamma-1}\delta}$ 个。这是高中熟悉的"鸽笼原理"。

逐步推到门槛 $s-1 > k(2k-1)$

1 抽屉: $s - 1$ 个系数分进 $p^{\gamma-1}\delta$ 类 (抽屉), 必有一类含 $\geq \frac{s-1}{p^{\gamma-1}\delta}$ 个数。取 $t = \lceil \frac{s-1}{p^{\gamma-1}\delta} \rceil$ (向上取整的最小整数), 把这一类的系数重排成前 t 个 $c_1, \dots, c_t$ 。

2 同类 $\Rightarrow$ 差一个 k 次方: 同一类意味着 c_i/c_1 是 k 次剩余, 即存在 d_i (不被 p 整除) 使 $c_i \equiv d_i^k c_1 \pmod{p^\gamma}$, $i = 2, \dots, t$ 。

3 砍掉多余变量、约去 c_1 : 令 $x_{t+1}, \dots, x_{s-1}$ (以及第 s 个) 都取 0, 方程 (7.6) 只剩前 t 项: $c_1 x_1^k + c_2 x_2^k + \dots + c_t x_t^k \equiv N$ 。代入 $c_i \equiv d_i^k c_1$, 提出公因子 c_1 :

$$c_1 (x_1^k + (d_2 x_2)^k + \dots + (d_t x_t)^k) \equiv N \pmod{p^\gamma}.$$

因为 c_1 与 p 互素, 它在模 p^γ 下可逆, 两边同乘 c_1^{-1} , 记 $N' = c_1^{-1}N$, 得

$$x_1^k + (d_2 x_2)^k + \dots + (d_t x_t)^k \equiv N' \pmod{p^\gamma}.$$

4 这就是 Waring 同余: 把 $d_i x_i$ 看成新变量 y_i (因 d_i 可逆, y_i 也跑遍所有剩余、且 y_i 非零 $\iff x_i$ 非零), 方程化为 $y_1^k + y_2^k + \dots + y_t^k \equiv N'$ 。这正是无系数 Waring 问题在引理 5.6 处理的同余, 结论是: 只要 $t \geq 2k$, 就有不全被 p 整除的解。

5 翻译成对 s 的要求: 要 $t \geq 2k$ 。因 t 是 $\geq \frac{s-1}{p^{\gamma-1}\delta}$ 的最小整数, 保证 $t \geq 2k$ 的安全充分条件是 $\frac{s-1}{p^{\gamma-1}\delta} > 2k - 1$ (这样向上取整后 $t \geq 2k$)。

6 把分母 $p^{\gamma-1}\delta$ 估上界: 奇素数情形 $\gamma - 1 = \tau$, 且 $\delta = (k_0, p - 1) \leq k_0$, 于是 $p^{\gamma-1}\delta = p^\tau \delta \leq p^\tau k_0 = k$ (回忆

$k = p^\tau k_0$)。所以分母 $\leq k$ 。

7 得门槛：要 $\frac{s-1}{p^{\gamma-1}\delta} > 2k-1$ ，因分母 $\leq k$ ，只要 $\frac{s-1}{k} > 2k-1$ ，即 $s-1 > k(2k-1)$ ，亦即 $s \geq k(2k-1) + 2$ 。这正是定理 7.3 中 k 为奇数（注意奇 k 时实际门槛由此式给出）的条件。 ■

取舍说明：这里用 $\delta \leq k_0$ 、 $p^\tau k_0 = k$ 这些不等式放大得到统一门槛，是为了得到“只依赖 k 的干净条件”，代价是门槛不是最优的（见末段）。

第 20 段：证明（情形 $p = 2$ ）

原文译文：设 $p = 2$ 。首先，若 $\tau = 0$ （即 k 为奇数），则只要 N 为奇数，同余式 (7.6) 即可解，因为即使它只有一项， x^k 也取遍模 p^γ 的一切值。因此当 N 为奇或偶时，只要两项便可解，故只需 $s-1 \geq 2$ 。于是当 k 为奇数时定理结论成立。现设 $\tau \geq 1$ ，故 k 为偶数。由于每个系数 c_i ($i \leq s-1$) 都是奇数，它对模 2^γ 可取 $2^{\gamma-1}$ 个可能值。因此存在某个由 t 个相互同余的系数构成的集合，其中 $t \geq (s-1)/2^{\gamma-1}$ 。令对应于其余系数的变量等于 0，我们看到只需解出 $x_1^k + \cdots + x_t^k \equiv N' \pmod{2^\gamma}$ ，其中诸变量不全为偶数。如同在引理 5.6 的证明中那样，只要 $t \geq 4k$ 即可。因此只需 $\frac{s-1}{2^{\gamma-1}} > 4k-1$ 。由于 $k \geq 2^\tau$ 且 $\gamma = \tau + 2$ ，我们有 $2^{\gamma-1} \leq 2k$ 。因此只需 $s-1 > 2k(4k-1)$ 。这就在 k 为偶数的情形证明了定理 7.3。

这一段在讲什么 素数 2 要单独处理，因为它最“刁”（ k 次幂结构在模 2 的方幂下最复杂）。结论分两块：① 若 k 奇（此时 $\tau = 0$ ），极其简单，两项就够了；② 若 k 偶（ $\tau \geq 1$ ），用同样的“抽屉 + Waring”套路，但门槛要到 $t \geq 4k$ ，倒推出偶数门槛 $s-1 > 2k(4k-1)$ 。

k 为奇数 ($\tau = 0$)

1 k 奇时，映射 $x \mapsto x^k$ 在模 2^γ 下是双射（一一对应），即 x^k 取遍模 2^γ 的所有值。直观：奇次幂保持奇偶、且单调地遍历各剩余。所以只用一

项 x^k 就能凑出任何奇数 N (取合适的奇 x , 得非平凡解)。

2 N 为偶数时, 用两项: 先用一项凑出一个奇数 a (非平凡), 再让第二项凑 $N - a$ (也能取到, 因 x^k 遍历全部值)。所以两项足以、且能保证至少一项是非平凡的奇贡献。

3 于是无论 N 奇偶, 2 项就够, 只需 $s - 1 \geq 2$, 这被奇数门槛 $s \geq k(2k - 1) + 2$ 远远满足。 k 奇情形定理成立。 ■

k 为偶数 ($\tau \geq 1$)

1 奇系数的可能值数: 每个 c_i ($i \leq s - 1$) 是奇数 (与 $p = 2$ 互素)。奇数模 2^γ 共有 $2^{\gamma-1}$ 个 (即 $1, 3, 5, \dots, 2^\gamma - 1$, 恰好一半)。把系数按"模 2^γ 相等"分类, 至多 $2^{\gamma-1}$ 类。

2 抽屉原理: $s - 1$ 个系数分进 $\leq 2^{\gamma-1}$ 类, 必有一类含 $t \geq \frac{s-1}{2^{\gamma-1}}$ 个, 它们彼此模 2^γ 相等 (比奇素数情形更强: 直接相等, 不只是差 k 次方)。

3 化归 Waring: 把其余变量取 0, 提出这个公共系数 (它是奇数、可逆), 同余式化为 $x_1^k + \dots + x_t^k \equiv N' \pmod{2^\gamma}$, 要求变量不全为偶。引理 5.6 在 $p = 2$ 给出的门槛是 $t \geq 4k$ (比奇素数的 $2k$ 大一倍, 因为模 2 方幂下 k 次幂覆盖得更稀)。

4 翻译成对 s 的要求: 要 $t \geq 4k$, 充分条件 $\frac{s-1}{2^{\gamma-1}} > 4k - 1$ 。

5 估分母 $2^{\gamma-1}$: 这里 $\gamma = \tau + 2$, 所以 $2^{\gamma-1} = 2^{\tau+1} = 2 \cdot 2^\tau$ 。又因 k 含因子 2^τ ($k = 2^\tau k_0$, k_0 奇), 故 $k \geq 2^\tau$, 于是 $2^{\gamma-1} = 2 \cdot 2^\tau \leq 2k$ 。分母 $\leq 2k$ 。

6 得门槛：要 $\frac{s-1}{2^{\gamma-1}} > 4k-1$ ，因分母 $\leq 2k$ ，只要 $\frac{s-1}{2k} > 4k-1$ ，即 $s-1 > 2k(4k-1)$ ，亦即 $s \geq 2k(4k-1) + 2$ 。这正是定理 7.3 中 k 为偶数的条件。偶数情形证毕。 ■

为什么偶数门槛更高：根子在 $p=2$ 、 k 偶时 k 次幂在模 2^γ 下分布最不均匀，需要更多项（ $4k$ 而非 $2k$ ）才能盖住所有目标剩余，同时 γ 也更大（ $\tau+2$ ），分母更大，于是要求 s 更大。

第 21 段：结论与展望（ $s_1(k)$ 与门槛非最优）

原文译文：由定理 7.1 与 7.3 可知：我们能够指定一个数 $s_1(k)$ ，使得当 $s \geq s_1(k)$ 时，当 $N \rightarrow \infty$ 时 $r(N) \rightarrow \infty$ ；这始终以系数 c_j 两两互素为前提。定理 7.3 中给出的数绝非最优；我们只不过给出了从证明中所用简单论证路线中自然出现的那些数。原则上，可以放宽系数两两互素这一条件；对于上述结论之成立而言，真正本质的是：对任一素数 p ，都有一定数量的系数不被 p 整除。

这一段在讲什么 收尾，三件事：① 合并定理 7.1+7.3 得到最终定理——存在门槛 $s_1(k)$ ，超过它就对一切大 N 有 $r(N) \rightarrow \infty$ （只要系数两两互素）；② 坦白门槛不是最优的，只是这条简单证法的自然产物；③ 指出真正本质的条件是什么，为推广留口子。

$s_1(k)$ 把定理 7.3 的两档门槛取较大者并满足渐近式所需的 $s \geq 2^k + 1$ ，即可取

$$s_1(k) = \max \{2^k + 1, k(2k-1) + 2 \text{ (奇)}, 2k(4k-1) + 2 \text{ (偶)}\}.$$

当 $s \geq s_1(k)$ ：定理 7.3 保证 $\mathfrak{S}(N) \geq C(k, s) > 0$ 对一切 N 成立，再代入定理 7.1 的 (7.4)，主项 $\asymp N^{s/k-1}$ 发散且压过误差，得 $r(N) \rightarrow \infty$ 。前提始终是两两互素。

合并得到最终结论

1 定理 7.3: 两两互素 + $s \geq \text{门槛} \Rightarrow \mathfrak{S}(N) \geq C(k, s) > 0$ (对一切 N , 不再限于算术级数)。

2 定理 7.1: $r(N) = \frac{C_{k,s}}{(c_1 \cdots c_s)^{1/k}} N^{s/k-1} \mathfrak{S}(N) + O(N^{s/k-1-\delta})$ 。

3 代入下界: $r(N) \geq (\text{正常数}) \cdot N^{s/k-1} - O(N^{s/k-1-\delta}) \rightarrow \infty$ 。所以每个足够大的 N 都可表示, 且表示数趋于无穷。 ■

"门槛非最优"与"真正本质的条件"

- **为什么不最优:** 证明里我们做了好几处放大——用 $\delta \leq k_0$ 、 $p^{\gamma-1}\delta \leq k$ 、 $2^{\gamma-1} \leq 2k$, 又用抽屉原理只取"最满那一类"。每一步放大都让门槛变大。更精细的分析 (不在本章) 能把门槛降下来。Davenport 坦诚: 这些数只是"简单论证自然蹦出来的", 重在思路清楚, 不在最优。
- **真正本质的是什么:** 回看整条证明, 两两互素只被用来保证"对每个素数 p , 至多一个系数被 p 整除" (于是有 $s-1$ 个系数可用)。所以真正的命脉是: 对任一素数 p , 都有足够多的系数不被 p 整除。只要这点满足, 两两互素可以放宽。这与第 1 段的反例首尾呼应——反例之所以坏, 正是因为某个 p 几乎吃掉了所有系数。

脚注

原文脚注: 当然, 只需对 $p \leq p_0(k, s)$ 假设此条件即可; 参见引理 5.2 的推论。

解释 这条脚注 (对应定理 7.2 里的"对每个素数 p ") 说明: 同余条件其实不必对无穷多个素数逐一验证。存在一个只依赖 k, s 的界 $p_0(k, s)$, 对所有 $p > p_0$ 的大素数, 非平凡解由一般理论自动存在 (引理 5.2 的推论保证局部密度对大素数自动 > 0)。所以真正要人工检查的只有有限个小素数 $p \leq p_0$ 。

这也正是第 15 段"满足条件的 N 构成算术级数"的根据——有限个素数的同余条件由中国剩余定理合并成一个模 $Q = \prod_{p \leq p_0} p^\gamma$ 的条件。

全章回顾（一句话串起来）

- 1 问题： $N = c_1 x_1^k + \cdots + c_s x_s^k$ 何时可解、解多少？
- 2 警告（第1段）：光项数多不行，某素数能毁掉一切——必须加同余条件。
- 3 机器微调（3–10段）：把生成函数、Weyl、Hua、(7.2)、主弧逼近、奇异级数里所有 x^k 换成 $c_j x^k$ ；靠"换元 $\beta = c_j \alpha + \text{周期性}$ "与"约分恒等式"把带系数化归回无系数，量级不变。
- 4 主结果（定理7.1）：
$$r(N) = \frac{C_{k,s}}{(c_1 \cdots c_s)^{1/k}} N^{s/k-1} \mathfrak{S}(N) + O(N^{s/k-1-\delta}).$$
- 5 正下界 I（定理7.2）：每素数有非平凡局部解 $\Rightarrow \mathfrak{S}(N) \geq C(k, s) > 0$ （在算术级数上）。
- 6 正下界 II（定理7.3）：两两互素 + 项数过门槛 $\Rightarrow$ 对一切 N 都有 $\mathfrak{S}(N) > 0$ 。证法 = 数陪集类 + 抽屉原理 + 化归 Waring 同余。
- 7 结论： $s \geq s_1(k)$ 时 $r(N) \rightarrow \infty$ ；门槛非最优，本质条件是"每个素数都有够多系数不被它整除"。

