

DAVENPORT · 圆法 · 高中详解版

三次型的奇异级数

Cubic forms: the singular series

本页是**逐段精讲**：先把 Davenport 原书第 17 章的译文一段一段搬出来（灰色「原文译文」框），再用**普通高中生能懂的语言**把每一句、每一个符号、每一步计算都拆开讲透。凡书里写「显然」「容易看出」的地方，这里都补成一步一句的完整推导。文字偏长，但保证不跳步。

本章要解决什么 / 读完能掌握什么

整个「圆法」(circle method) 在前面几章已经把一件事做到了八九分：要数清三次方程 $C(\mathbf{x}) = 0$ 在一个大盒子里有多少组整数解，答案近似等于

$$\mathcal{N}(P) \approx P^{n-3} \mathfrak{S} J_0.$$

这里 P^{n-3} 是「主项的大小」， J_0 （奇异积分，singular integral，上一章处理过）管「实数层面有没有解、解多不多」，而本章的主角 $\mathfrak{S}$ （**奇异级数**，singular series）管「每个素数 p 的同余层面有没有解、解多不多」。

本章只干一件关键的事：**证明 $\mathfrak{S} > 0$** 。因为只要这个正数因子不为零，上式右边就会随 $P \rightarrow \infty$ 趋于无穷，于是盒子里的解必然多到无穷，从而方程一定有非平凡整数解。读完你会掌握：(1) 奇异级数怎么写成「每个素数各管一段」的**欧拉乘积**；(2) 每段 $\chi(p)$ 其实就是同余方程解数的极限密度；(3) 怎样用一个叫「性质 $\mathcal{A}(p^\ell)$ 」的条件，配合**逐次提升解**（Hensel 提升）的归纳法，保证每个 $\chi(p) > 0$ 。

开讲前：本章会反复用到的「**超纲**」记号，先一次性讲清

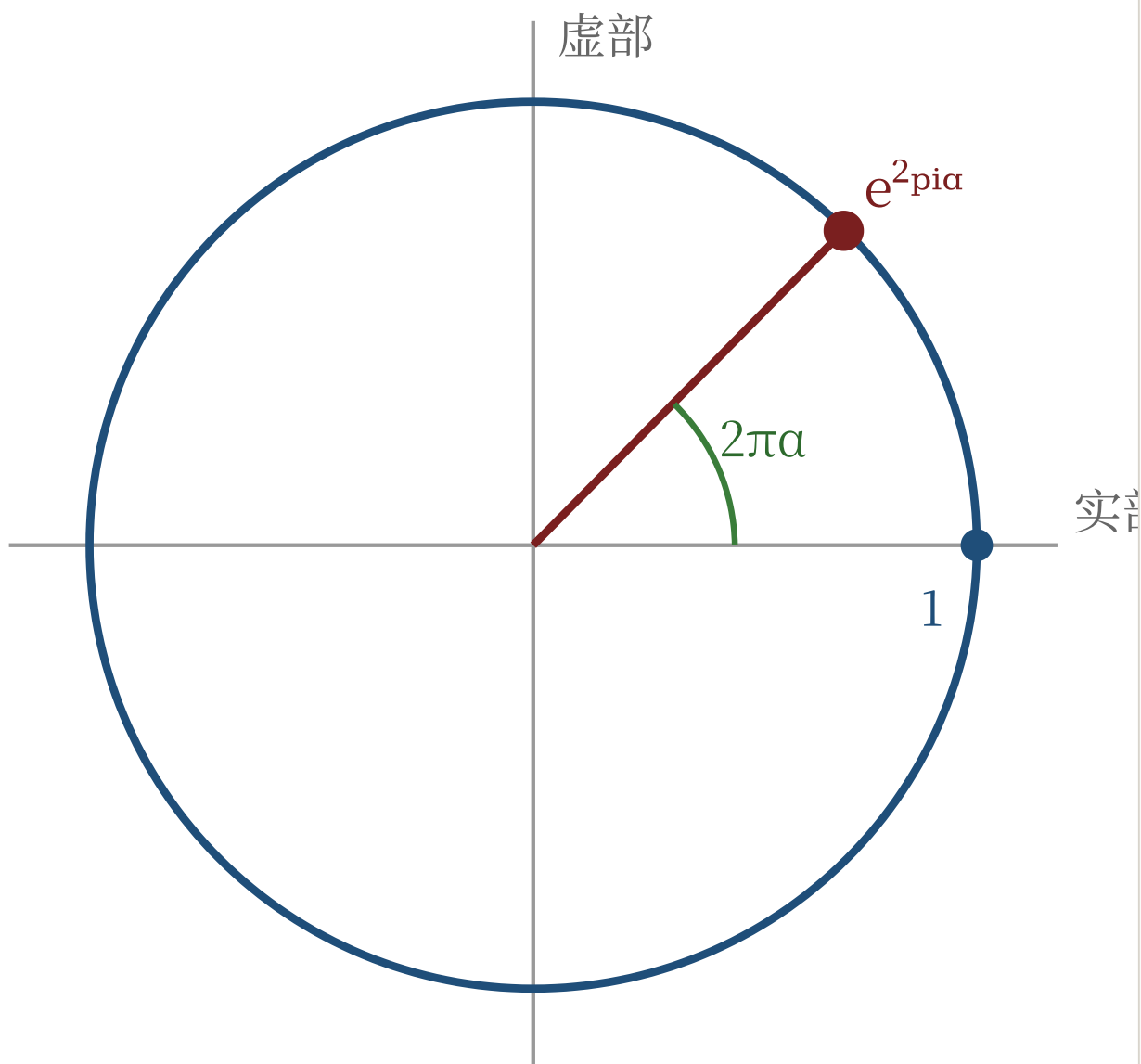
这一章的句子虽短，但每一句都压着好几个大学记号。先把它们从零讲一遍，后面遇到就不再卡壳。

记号 1: 复指数 $e(\alpha) = e^{2\pi i \alpha}$

读作「e 的 $2\pi i \alpha$ 次方」。这里 i 是虚数单位，满足 $i^2 = -1$ ； α 是一个实数。由欧拉公式 $e^{i\theta} = \cos \theta + i \sin \theta$ ，所以

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

它的几何意义：在复平面（横轴记实部、纵轴记虚部）上， $e(\alpha)$ 是单位圆上的一个点，从正东方向（点 1）出发逆时针转过 $2\pi\alpha$ 弧度（也就是转过整整 α 圈）。最有用的性质有两条：① 它的**模长**（到原点的距离）恒等于 1，记作 $|e(\alpha)| = 1$ ；② 它以 1 为**周期**： $e(\alpha + 1) = e(\alpha)$ ，因为多转一整圈回到原地。圆法里所有「指数和」都是把许多这种单位圆上的点加起来。



$e(\alpha)$ 是单位圆上转过 α 圈得到的点；不管 α 多大，它到原点的距离永远是 1。

记号 2: 求和号 $\sum$ 与同余 $\equiv \pmod{m}$

$\sum_{j=1}^N a_j$ 就是 $a_1 + a_2 + \cdots + a_N$ 的缩写，读「西格玛，j 从 1 到 N 求和」。下方还能挂条件，比如 $\sum_{(a,q)=1}$ 表示「只对那些与 q 互素的 a 求和」。

$A \equiv B \pmod{m}$ 读作「A 与 B 关于模 m 同余」，意思是 $A - B$ 能被 m 整除，也就是 A, B 除以 m 余数相同。例如 $17 \equiv 2 \pmod{5}$ （都余 2）。本章满屏的 $(\text{mod } p^\nu)$ 就是「除以素数 p 的 ν 次方，看余数」。

记号 3: $O(\cdot)$ 与 $\ll$ (数量级记号)

这两个记号都是用来表示「某个量不会比另一个量大太多」。

- $f(P) = O(g(P))$ 读「f 是 g 的大 O」，意思是存在一个与 P 无关的常数 $c > 0$ ，使得当 P 足够大时永远有 $|f(P)| \leq c g(P)$ 。直观上： f 顶多是 g 的常数倍那么大。
- $f \ll g$ 是同一件事的另一种写法 (Vinogradov 记号)，完全等价于 $f = O(g)$ ，读「f 远小于等于 g (在数量级意义下)」。

为什么数论里爱用它？因为我们常常不在乎一个误差项到底是 $3.7P^2$ 还是 $50P^2$ ，只在乎它「是 P^2 这个档次」。比如 $P^{n-3-\delta}$ 比主项 P^{n-3} 小一整个 P^δ 倍 ($\delta > 0$)，所以当 P 很大时它可以忽略。

记号 4: $o(1)$ (小 o)

$o(1)$ 表示一个「当 $P \rightarrow \infty$ 时趋于 0」的量，读「小 o 一」。所以 $J_0 + o(1)$ 的意思是「当 P 越来越大，这个括号越来越接近常数 J_0 」。和大 O 的区别：大 O 说「不超过某档」，小 o 说「相比之下可忽略 / 趋于零」。

记号 5: ε 、 δ 这种「任意小的正数」

在解析数论里， ε (epsilon) 通常代表「随便多小的一个正数」，用来给估计留一点余量。 q^ε 增长得比任何 q 的正幂都慢 (例如比 $q^{0.001}$ 还温和地可调)，它的作用是吸收掉一些「几乎不增长」的因子 (如除数函数)。 δ (delta) 则常代表一个「确实存在、固定的小正数」，用来表示「严格地多出一一点」或「严格地小一档」。

正文逐段精讲

把引理 15.1、15.4 与 16.1 的结果合在一起，我们现在已经证明：若 $n \geq 17$ ，且排除引理 13.4 的备选 A，又若盒子 $\mathfrak{B}$ 选取得当，则盒子 $P\mathfrak{B}$ 中满足 $C(\mathbf{x}) = 0$ 的整点个数 $\mathcal{N}(P)$ 满足

$$\mathcal{N}(P) = P^{n-3} \mathfrak{S}(P^\Delta) \{J_0 + o(1)\} + O(P^{n-3-\delta}),$$

其中 $\delta > 0$ 。

这段在讲什么：这是「前情提要」。前面三章（15、16 两章和它们的引理）各算出渐近公式的一块，本段把它们拼成一个完整的计数公式。下面把式子里每个符号都拆开。

$C(\mathbf{x})$

一个**三次型** (cubic form)。「型」=各项次数都相同的多项式；「三次」=每一项都是三次。 $\mathbf{x} = (x_1, \dots, x_n)$ 是 n 个整数变量打成的一组（向量）。例如 $x_1^3 + x_2^3 + \dots$ 这种。 $C(\mathbf{x}) = 0$ 就是我们想找整数解的方程。

n

变量个数。本章的核心假设是 $n \geq 17$ ：变量足够多，圆法才控制得住误差。

$\mathfrak{B}$ (哥特体 B)

一个固定的「盒子」(box)，就是 n 维空间里一块长方体区域。 $P\mathfrak{B}$ 表示把这个盒子按比例放大 P 倍。 P 是个趋于无穷的大参数。

$\mathcal{N}(P)$

放大后的盒子 $P\mathfrak{B}$ 里，**满足方程 $C(\mathbf{x}) = 0$ 的整点**（坐标全为整数的点）个数。这正是我们最终想证「趋于无穷」的量。

$\mathfrak{S}(P^\Delta)$ (哥特体 S)

奇异级数的有限截断。完整的奇异级数 $\mathfrak{S}$ 是对所有 $q = 1, 2, 3, \dots$ 求和的无穷级数；圆法里实际只能先取 q 到某个上界 P^Δ (Δ 是个小的固定正指数) 为止，得到的有限和就记 $\mathfrak{S}(P^\Delta)$ 。本章的任务之一就是说明：让 $P \rightarrow \infty$ 、上界 $P^\Delta \rightarrow \infty$ ，这个有限和会收敛到完整的 $\mathfrak{S}$ 。

J_0

奇异积分 (singular integral)，上一章（第 16 章）的产物。它是一个固定的正常数，刻画方程在「实数 / 几何」意义下解的密度。

备选 A (alternative A)

来自引理 13.4 的一个分情况讨论里的一支。整个证明分成「备选 A 成立」和「备选 A 不成立」两种局面。下一段会说明：备选 A 一旦成立，方程**直接**就有非平凡解（第 14 章已证），所以那一支不用圆法也赢了；圆法处理的是「排除备选 A」的另一支。

整条公式怎么读：盒子里解的个数 = （主项）+ （误差项）。主项是 $P^{n-3} \mathfrak{S}(P^\Delta) \{J_0 + o(1)\}$ ：

- 1 P^{n-3} 是主项的「体量」。为什么是 $n-3$ 而不是 n ？因为盒子是 n 维的，里头整点总数约 P^n 个；但加上「 $C(\mathbf{x}) = 0$ 是 1 个三次方程的约束」，会把维数压低，圆法算出净剩 P^{n-3} 这个量级（一个方程、三次，正好降 3）。
- 2 $\mathfrak{S}(P^\Delta)$ 与 J_0 是两个密度因子，分别管「同余世界」和「实数世界」。它们都为正，主项才真的有那么大。
- 3 $\{J_0 + o(1)\}$ ：括号随 P 增大越来越接近常数 J_0 ，那个 $o(1)$ 是会消失的零头。
- 4 末尾 $+O(P^{n-3-\delta})$ 是**误差项**，它的量级 $P^{n-3-\delta}$ 比主项 P^{n-3} 小了整整 P^δ 倍（ $\delta > 0$ 固定）。所以 P 一大，误差相对主项可忽略——这正是圆法千辛万苦要保证的「主项压过误差」。

原文译文

引理 15.4 中给出的级数 $\mathfrak{S}(P^\Delta)$ ，若延拓到无穷，则绝对收敛（前提是 $n \geq 17$ 且排除备选 A），因为此时由引理 15.3 有

$$q^{-n} |S_{a,q}| \ll q^{-\frac{1}{8}n + \varepsilon - 2 - \delta}.$$

这段在讲什么：把有限截断 $\mathfrak{S}(P^\Delta)$ 延长成对所有 q 的无穷级数 $\mathfrak{S}$ ，需要先确认这个无穷级数**真的收敛**（不会加出无穷大）。本段给出收敛的根据：每一项小到一定程度。先认识里头的新符号。

记号 $S_{a,q}$: 完整指数和 (complete exponential sum)

它定义为

$$S_{a,q} = \sum_{\mathbf{x} \bmod q} e\left(\frac{a C(\mathbf{x})}{q}\right),$$

即让 $\mathbf{x} = (x_1, \dots, x_n)$ 的每个坐标各跑遍模 q 的一组完全剩余 $(0, 1, \dots, q-1)$, 把对应的单位圆点 $e(aC(\mathbf{x})/q)$ 全加起来。这是圆法在「分母为 q 、分子为 a 的有理点附近」抽取出的同余信息的浓缩: 当 $C(\mathbf{x}) \equiv 0 \pmod{q}$ 的解多时, 许多项的指数接近 0、 $e(\cdot)$ 接近 1, 和就大; 解少时, 各项指数乱转、相互抵消, 和就小。所以 $S_{a,q}$ 的大小直接反映「模 q 有多少解」。

$|S_{a,q}|$ 是这个复数的模长。因为它是 q^n 个模长为 1 的复数之和, 平凡上界是 $|S_{a,q}| \leq q^n$; 引理 15.3 给的是远比这好的上界。

a, q

$q \geq 1$ 是分母, a 跑遍 $1 \leq a \leq q$ 中与 q 互素的那些数 (记 $(a, q) = 1$, 即最大公约数为 1)。分数 a/q 就是单位区间 $[0, 1]$ 上的既约有理点, 圆法围绕这些点画「主弧」。

$q^{-n}|S_{a,q}|$

把指数和除以 q^n 做归一化。因为 $S_{a,q}$ 是 q^n 项之和, 除以 q^n 后得到的是「平均每项」的大小, 也就是模 q 解的相对密度, 这是真正进入奇异级数的量。

那条不等式 $q^{-n}|S_{a,q}| \ll q^{-\frac{1}{8}n + \varepsilon - 2 - \delta}$ 怎么读:

1 引理 15.3 给出归一化指数和的上界, 指数是 $-\frac{1}{8}n + \varepsilon - 2 - \delta$ 。把 $n \geq 17$ 代进去看这个指数有多负: $-\frac{1}{8} \times 17 = -2.125$, 再 -2 得约 -4.125 , 再减 δ 、加上可忽略的 ε , 指数比 -4 还小。

2 关键是这个指数严格小于 -2 (因为 $-\frac{1}{8}n - 2 - \delta \leq -2.125 - 2 - \delta < -2$)。后面会看到, 正是「比 -2 还负」这一点, 保证级数收敛。

- 3 为什么 $n \geq 17$ 是分水岭？因为指数里 $-\frac{1}{8}n$ 这一项要足够负，才能把后面求和时产生的正幂压下去。 n 太小，这个上界就不够好，级数可能发散，整套论证就垮了。这就是「为什么是 17 而不是更小」的算术根源。

原文译文

由第 14 章的工作，备选 A 蕴含 $C(\mathbf{x}) = 0$ 有非平凡的整数解。这样我们就证明了：

这段在讲什么：收尾两支讨论中的「备选 A 这一支」。回忆前面：整个证明按引理 13.4 分成备选 A 成立 / 不成立两种局面。这里指出：如果落在备选 A，那么第 14 章已经独立地证出方程有非平凡解，于是这一支当场结束，根本不需要奇异级数。所以接下来只要专心处理「排除备选 A」的那一支即可——而那一支正好满足前面收敛性所需的条件。这样两支合起来就把定理证完，于是引出下面的定理。

什么叫「非平凡解」

方程 $C(\mathbf{x}) = 0$ 永远有一个「废解」：全取 0，即 $\mathbf{x} = (0, 0, \dots, 0)$ ，代进去当然等于 0。这个解叫**平凡解** (trivial)，它什么信息都没给。**非平凡解** (non-trivial) 指至少有一个坐标不为 0 的整数解。整本书要找的、有意义的，都是非平凡解。

定理 17.1. 若 $n \geq 17$ 且 $\mathfrak{S} > 0$ ，则方程 $C(\mathbf{x}) = 0$ 有非平凡的整数解。

定理在说什么：这是本章（乃至整套三次型论证）的「主定理半成品」。它把存在解这件大事，化简成一个纯粹关于密度因子的不等式 $\mathfrak{S} > 0$ 。换句话说：只要变量够多（ ≥ 17 ）、并且奇异级数为正，方程就一定有非平凡整数解。剩下的全部工作，就是去验证那个 $\mathfrak{S} > 0$ ——这正是本章后半段要做的。

原文译文

因为若不存在非平凡解，我们便得到

$$\mathcal{N}(P) \sim P^{n-3} \mathfrak{S} J_0 \quad \text{当 } P \rightarrow \infty,$$

这段在讲什么：用反证法证上面的定理。先认识 $\sim$ 这个符号。

记号 $\sim$ (渐近相等, asymptotically equal)

$f(P) \sim g(P)$ (当 $P \rightarrow \infty$) 读「f 渐近于 g」, 意思是两者的比值趋于 1:

$\lim_{P \rightarrow \infty} \frac{f(P)}{g(P)} = 1$ 。直观上: 当 P 很大时, f 和 g 几乎一模一样大。注意它比「等于」

弱, 但比「大 O」强。

这一步的推导 (书里跳过了, 这里补全): 为什么「没有非平凡解」会得到 $\mathcal{N}(P) \sim P^{n-3} \mathfrak{S} J_0$? 其实这只是把开头那条计数公式取极限:

1 开头公式: $\mathcal{N}(P) = P^{n-3} \mathfrak{S}(P^\Delta) \{J_0 + o(1)\} + O(P^{n-3-\delta})$ 。

2 前面已证无穷级数 $\mathfrak{S}$ 绝对收敛, 故有限截断 $\mathfrak{S}(P^\Delta) \rightarrow \mathfrak{S}$ (当 $P \rightarrow \infty$, 上界 $P^\Delta \rightarrow \infty$)。所以 $\mathfrak{S}(P^\Delta) = \mathfrak{S} + o(1)$ 。

3 括号 $\{J_0 + o(1)\} \rightarrow J_0$ 。两个「主因子」相乘:
 $\mathfrak{S}(P^\Delta) \{J_0 + o(1)\} = \mathfrak{S} J_0 + o(1)$ 。

4 于是 $\mathcal{N}(P) = P^{n-3} (\mathfrak{S} J_0 + o(1)) + O(P^{n-3-\delta}) = P^{n-3} \mathfrak{S} J_0 (1 + o(1))$ 。
最后一步把误差也并进 $o(1)$: 误差 $O(P^{n-3-\delta})$ 除以主项 $P^{n-3} \mathfrak{S} J_0$ 是 $O(P^{-\delta}) \rightarrow 0$, 故确实是 $o(1)$ (这里用到 $\mathfrak{S} J_0$ 是正常数)。

5 「比值趋于 1」正是 $\sim$ 的定义, 所以 $\mathcal{N}(P) \sim P^{n-3} \mathfrak{S} J_0$ 。

注意这一步悄悄用到了反证假设: 如果方程没有非平凡解, 那么盒子 $P\mathfrak{B}$ 里能数到的解最多只有平凡解 $\mathbf{0}$ (而且 $\mathbf{0}$ 还不一定在盒子内部), 所以 $\mathcal{N}(P)$ 只能是 0 或一个有界的小数, 绝不会趋于无穷。把这一点和上面算出的「 $\mathcal{N}(P)$ 渐近于一个趋于无穷的量」对照, 矛盾就要来了。

原文译文

从而 $\mathcal{N}(P) \rightarrow \infty$, 产生矛盾。

这段在讲什么：把矛盾点破。

- 1 由假设 $\mathfrak{S} > 0$ (定理条件) 且 $J_0 > 0$ (上一章已证为正), 故 $\mathfrak{S}J_0$ 是个正常数。
- 2 而 $P^{n-3} \rightarrow \infty$ (因为 $n \geq 17 > 3$, 指数 $n-3 \geq 14 > 0$)。正常数乘以趋于无穷的量, 仍趋于无穷。所以右边 $P^{n-3}\mathfrak{S}J_0 \rightarrow \infty$ 。
- 3 由 $\sim$, 左边 $\mathcal{N}(P)$ 也必须 $\rightarrow \infty$ 。
- 4 但「没有非平凡解」逼着 $\mathcal{N}(P)$ 有界 (上一段已说明)。「有界」和「趋于无穷」直接打架, 矛盾。
- 5 反证法收口: 假设「没有非平凡解」导致矛盾, 所以原命题成立——方程有非平凡解。定理 17.1 证毕。

原文译文

这里当然 $\mathfrak{S}$ 表示延拓到无穷的奇异级数, 即

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^q q^{-n} S_{a,q}.$$

这段在讲什么：正式给出奇异级数 $\mathfrak{S}$ 的完整定义。它是一个二重无穷和, 把上面认识的那些「归一化指数和」全加起来。逐层拆开读：

- 1 最外层 $\sum_{q=1}^{\infty}$: 分母 q 从 1 一直跑到无穷。
- 2 里层 $\sum_{\substack{a=1 \\ (a,q)=1}}^q$: 对每个固定的 q , 分子 a 跑遍 1 到 q 之间所有与 q 互素的数。这

种 a 的个数恰好是欧拉函数 $\varphi(q)$ 个（见下面记号框）。

3 被加的项 $q^{-n}S_{a,q}$ 就是前面讲过的归一化完整指数和。

为什么这个二重和叫「奇异级数」、它有什么含义

名字「singular series」是 Hardy 与 Littlewood 在 1920 年代发明圆法时起的（singular 这里不是「奇怪」，而是「特异 / 主导」的意思，因为它来自积分中那些「主弧」上的特异贡献）。它的**真实含义**是各素数同余条件的「总密度」：可以证明（本章后面就会做）它等于一串**局部密度**的乘积，每个素数 p 贡献一个因子 $\chi(p)$ ，而 $\chi(p)$ 恰好衡量「方程模 p 的各次幂有多少解」。所以 $\mathfrak{S} > 0$ 的真正含义是：方程在**每一个素数 p 处都局部可解**，并且解不会太稀。这是「整体有解」的必要前提，而圆法（在 n 够大时）让它也成为充分条件。

欧拉函数 $\varphi(q)$

$\varphi(q)$ （读「phi of q 」）表示 $1, 2, \dots, q$ 中与 q 互素的整数个数。例如 $\varphi(6) = 2$ （只有 1 和 5 与 6 互素）， $\varphi(p) = p - 1$ （素数 p 与 $1 \sim p - 1$ 全互素）。它正好数清里层求和有多少项。本章后面用乘性时会间接用到它的乘性。

原文译文

剩下需要证明的是：对每一个变元个数不少于 17 的三次型，都有 $\mathfrak{S} > 0$ 。引理 5.1 的证明适用于一般的指数和（如当时所指出的），它表明若令

$$A(q) = \sum_{\substack{a=1 \\ (a,q)=1}}^q q^{-n} S_{a,q},$$

则 $A(q)$ 是 q 的乘性函数（对互素的 q 值而言）。

这段在讲什么：明确接下来唯一的目标——证 $\mathfrak{S} > 0$ ——并迈出第一步：把奇异级数按 q 打包成 $A(q)$ ，并指出 $A(q)$ 有一个极好的性质叫「乘性」。

记号 $A(q)$ ：把奇异级数按 q 分组

$A(q)$ 就是上面二重和的内层那一整块：固定 q ，把所有合法的 a 对应的 $q^{-n} S_{a,q}$ 加起来。于是奇异级数被改写成单重和 $\mathfrak{S} = \sum_{q=1}^{\infty} A(q)$ 。引入它的动机：把「每个分母 q 的总贡献」捏成一个数，便于谈论它随 q 怎么变、级数是否收敛、能否拆成素数因子。

什么是「乘性函数」(multiplicative function)

一个定义在正整数上的函数 f 叫乘性的，如果对任意互素的 q_1, q_2 （即 $(q_1, q_2) = 1$ ）都有

$$f(q_1 q_2) = f(q_1) f(q_2).$$

注意条件是「互素」，不是任意两数都行。最常见的例子就是欧拉函数 φ 。乘性之所以宝贵：任何正整数 q 都能唯一分解成不同素数的幂相乘 $q = p_1^{\nu_1} p_2^{\nu_2} \cdots$ （这些素数幂两两互素），于是只要知道 f 在每个「素数幂」上的值，整个 f 就被完全决定，而且 $\sum_q f(q)$ 这种级数能拆成「对每个素数各管一段」的乘积（欧拉乘积）。这正是下一步要用的关键机器。

「引理 5.1 的证明适用于一般指数和」是什么意思（补全背景）：第 5 章处理 Waring 问题（ $x_1^k + \cdots$ 那种加性型）时，曾证明相应的 $A(q)$ 是乘性的。当时的证明只用到指数和 $S_{a,q}$ 的一条通用结构——中国剩余定理带来的分解。这里 Davenport 指出：那套论证对一般的（不只加性的）三次型指数和照样成立，所以结论照搬。其核心机制是：

- 1 设 $q = q_1 q_2$, $(q_1, q_2) = 1$ 。由中国剩余定理，模 q 的一个剩余 $\mathbf{x}$ 与「模 q_1 的剩余 $\mathbf{x}^{(1)}$ 加上模 q_2 的剩余 $\mathbf{x}^{(2)}$ 」一一对应，可写 $\mathbf{x} = q_2 \mathbf{x}^{(1)} + q_1 \mathbf{x}^{(2)}$ （模 q 意义下）。
- 2 把它代进三次型，并利用「分子 a 也能按 $a/q = a_1/q_1 + a_2/q_2$ 拆开」，指数 $e(aC(\mathbf{x})/q)$ 就分裂成「只含 q_1 的因子」乘「只含 q_2 的因子」。

- 3 于是整个和 $S_{a,q}$ 分裂成 $S_{a_1,q_1} \cdot S_{a_2,q_2}$ 这样的乘积；归一化、对 a 求和之后，正好得到 $A(q_1 q_2) = A(q_1) A(q_2)$ 。

细节属于第 5 章，这里只需记住结论： $A(q)$ 在互素分解下相乘。

原文译文

在 $n \geq 17$ 且 $C(\mathbf{x})$ 不表示零的前提下，由引理 15.3 我们有

$$|A(q)| \ll q^{1-\frac{1}{8}n+\varepsilon} \ll q^{-1-\delta}.$$

这段在讲什么：给出 $A(q)$ 的大小上界，为下一步「级数绝对收敛」做准备。

- 1 「 $C(\mathbf{x})$ 不表示零」就是「方程没有非平凡解」——这正是反证法里我们临时假设的局面（要推出矛盾）。在这个假设下引理 15.3 的指数和估计才适用。

- 2 第一个 $\ll$ ：把 $A(q)$ 里每一项 $|q^{-n} S_{a,q}|$ 用前面那条 $\ll q^{-\frac{1}{8}n+\varepsilon-2-\delta}$ 控制，再乘上项数（合法的 a 至多 q 个，量级 q^1 ）。粗略地：项数 q 乘以每项 $q^{-\frac{1}{8}n+\varepsilon-2-\delta}$ ，把 $-2-\delta$ 吸收进通用常数后，得到指数 $1-\frac{1}{8}n+\varepsilon$ 。这就是 $|A(q)| \ll q^{1-\frac{1}{8}n+\varepsilon}$ 的来历。

- 3 第二个 $\ll$ ：当 $n \geq 17$ 时， $1-\frac{1}{8}n \leq 1-\frac{17}{8} = 1-2.125 = -1.125$ 。再吸收掉那个可以取得很小的 ε ，指数仍然严格小于 -1 ，即可写成 $-1-\delta$ （取某个固定的小 $\delta > 0$ ，比如 $\delta = 0.1$ ）。于是 $|A(q)| \ll q^{-1-\delta}$ 。

- 4 这个 $-1-\delta$ 才是要害：它严格小于 -1 ，决定了 $\sum_q |A(q)|$ 收敛（下一段）。再次看到 17 这个门槛：若 n 不够大，指数到不了 -1 以下，级数就不收敛。

原文译文

因此 $\mathfrak{S} = \sum_q A(q)$ 绝对收敛，从而可知

$$\mathfrak{S} = \prod_p \chi(p),$$

其中

$$\chi(p) = 1 + \sum_{\nu=1}^{\infty} A(p^{\nu}).$$

这段在讲什么：这是本章结构上最漂亮的一步——把奇异级数 $\mathfrak{S}$ （对所有 q 的和）**改写成对所有素数 p 的乘积**（欧拉乘积，Euler product）。每个素数 p 贡献一个因子 $\chi(p)$ 。下面逐点补全为什么能这么做。

记号 $\prod_p$ （连乘号）与 $\chi(p)$

$\prod_p$ 读「对所有素数 p 求积」，是把每个素数对应的因子全乘起来（与 $\sum$ 把它们加起来相对）。 $\chi(p)$ （读「chi of p 」， χ 是希腊字母 chi）是素数 p 处的**局部因子**；它收集了所有「分母是 p 的方幂」的项： $q=1$ 贡献那个起头的 1（因为 $A(1)=1$ ，见下）， $q=p, p^2, p^3, \dots$ 贡献 $A(p) + A(p^2) + \dots$ 。所以 $\chi(p) = 1 + \sum_{\nu \geq 1} A(p^{\nu})$ 。

先解决一个小细节：为什么起头是 1？ 因为 $q=1$ 时只有 $a=1$ （且 $(1,1)=1$ ），而 $S_{1,1} = \sum_{\mathbf{x} \bmod 1} e(\dots)$ ，模 1 只有一个剩余、指数为 $e(\text{整数})=1$ ，故 $S_{1,1}=1$ ，归一化后 $A(1) = 1^{-n} \cdot 1 = 1$ 。这个 $q=1$ 的项不归任何素数的「方幂段」，所以单独提到每个 $\chi(p)$ 的常数项里——正是那个孤零零的 1。

为什么「绝对收敛」 $\Rightarrow$ 「能拆成欧拉乘积」（补全推导）：

1 **绝对收敛**的意思：把每一项取绝对值后再求和， $\sum_q |A(q)|$ 仍是有限数。由上一段 $|A(q)| \ll q^{-1-\delta}$ ，比较级数 $\sum_q q^{-1-\delta}$ 是收敛的 p -级数（因为指数 $1+\delta > 1$ ），所以 $\sum_q |A(q)|$ 收敛。绝对收敛极重要，因为它允许我们**任意重排、任意分组求和而不改变结果**——下面拆乘积全靠这条。

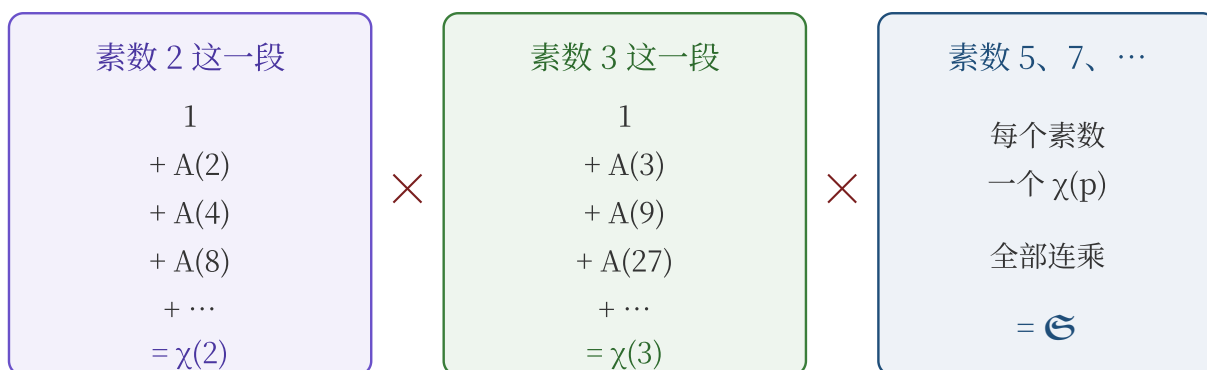
2 任何 q 唯一分解为 $q = p_1^{\nu_1} \cdots p_r^{\nu_r}$ （互素的素数幂）。由 A 的乘性， $A(q) = A(p_1^{\nu_1}) \cdots A(p_r^{\nu_r})$ 。

3 把无穷乘积 $\prod_p (1 + A(p) + A(p^2) + \cdots)$ 形式地展开：从每个素数 p 的括号里挑一项（要么挑常数 1，要么挑某个 $A(p^\nu)$ ），把挑出的相乘。由乘性，挑出的乘积恰好等于某个 $A(q)$ ，其中 q 就是「被挑中的那些素数幂的乘积」。

4 反过来，每个 q 由它的素因子分解，恰好对应一种挑法（在含 p 的因子中挑 $A(p^{\nu_p})$ ，在不含 p 的因子中挑 1）。所以把乘积展开后所有项加起来，正好不重不漏地得到 $\sum_q A(q) = \mathfrak{S}$ 。

5 「形式展开 = 真的相等」这一步需要绝对收敛保证（否则无穷重排无意义）。由第 1 步绝对收敛成立，故 $\mathfrak{S} = \prod_p \chi(p)$ 严格成立。

把「对所有 q 相加」重组为「对每个素数相乘」



乘性 + 绝对收敛 $\Rightarrow$ 奇异级数 $\mathfrak{S} = \prod_p \chi(p)$ 。从每个括号挑一项相乘，借唯一分解恰好对应一个 $A(q)$ ，不重不漏。

原文译文

在上述条件下我们还有

$$|\chi(p) - 1| \ll p^{-1-\delta},$$

所以存在 p_0 ，使得

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

正如引理 5.2 的推论所示。

这段在讲什么：处理欧拉乘积「会不会乘出 0」的问题。无穷乘积有个陷阱：即便每个因子都正，乘积也可能趋于 0（例如 $\prod(1 - \frac{1}{p}) \rightarrow 0$ ）。要保证 $\mathfrak{S} > 0$ ，必须排除这种「乘没了」。本段的办法是：证明大素数那一截的因子都非常接近 1，于是它们的乘积被 $\frac{1}{2}$ 兜底，绝不会塌到 0。

第一步： $|\chi(p) - 1| \ll p^{-1-\delta}$ 怎么来的（补全）：

1 $\chi(p) - 1 = \sum_{\nu \geq 1} A(p^\nu)$ （把那个起头的 1 减掉，剩下方幂段）。

2 取绝对值，用三角不等式（和的模 $\leq$ 模的和）： $|\chi(p) - 1| \leq \sum_{\nu \geq 1} |A(p^\nu)|$ 。

3 每项用 $|A(q)| \ll q^{-1-\delta}$ ，令 $q = p^\nu$ ： $|A(p^\nu)| \ll p^{-\nu(1+\delta)} = (p^{-(1+\delta)})^\nu$ 。

4 这是公比 $r = p^{-(1+\delta)} < 1$ 的等比级数，从 $\nu = 1$ 起求和：
 $\sum_{\nu \geq 1} r^\nu = \frac{r}{1-r}$ 。由于 $p \geq 2$ 时 $r \leq 2^{-(1+\delta)}$ 是个小于 1 的固定数，分母 $1-r$ 有正下界，所以 $\frac{r}{1-r} \ll r = p^{-1-\delta}$ 。

5 合起来 $|\chi(p) - 1| \ll p^{-1-\delta}$ ，得证。含义：素数 p 越大， $\chi(p)$ 离 1 越近，差距按 $p^{-1-\delta}$ 飞快缩小。

第二步：存在 p_0 使 $\prod_{p > p_0} \chi(p) \geq \frac{1}{2}$ （补全「引理 5.2 推论」的论证）：

1 因为 $\sum_p p^{-1-\delta}$ 收敛（指数 $1+\delta > 1$ ），它的「尾巴」可以任意小。于是能取一个足够大的 p_0 ，使得 $\sum_{p > p_0} |\chi(p) - 1| \leq \sum_{p > p_0} c p^{-1-\delta} < \frac{1}{2}$ （ c 是上面 $\ll$ 里的常数）。这一步本质是「收敛级数的余项趋于 0」。

2 写每个因子为 $\chi(p) = 1 + a_p$ ，其中 $a_p = \chi(p) - 1$ ，且 $\sum_{p > p_0} |a_p| < \frac{1}{2}$ 。

要证乘积 $\prod_{p>p_0} (1 + a_p) \geq \frac{1}{2}$ 。

- 3 用一个初等不等式：当各 a_p 满足 $\sum |a_p| < 1$ 时， $\prod (1 + a_p) \geq 1 - \sum |a_p|$ 。（直观：把乘积展开，主项是 1，其余项的绝对值之和不超过 $\prod (1 + |a_p|) - 1 \leq e^{\sum |a_p|} - 1$ ；更简单地可对因子数归纳： $(1 + a)(1 - s) \geq 1 - s - |a|$ 当 $1 - s \geq 0$ 。）

- 4 代入 $\sum_{p>p_0} |a_p| < \frac{1}{2}$ ，得 $\prod_{p>p_0} \chi(p) \geq 1 - \frac{1}{2} = \frac{1}{2}$ 。

- 5 结论的意义：所有大素数合起来的贡献至少是 $\frac{1}{2}$ ，被牢牢托住，不会塌成 0。于是判断整个 $\mathfrak{S} = (\prod_{p \leq p_0} \chi(p)) \cdot (\prod_{p > p_0} \chi(p))$ 是否为正，只剩下有限个小素数 $p \leq p_0$ 的因子 $\chi(p)$ 是否各自为正。问题从「无穷个素数」缩成了「有限个素数」。

原文译文

引理 5.3 的论证表明

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

其中 $M(p^\nu)$ 表示同余式

$$C(x_1, \dots, x_n) \equiv 0 \pmod{p^\nu}, \quad 0 \leq x_j < p^\nu$$

的解的个数。

这段在讲什么：给 $\chi(p)$ 一个看得见摸得着的几何 / 计数解释。前面 $\chi(p)$ 是一堆指数和加出来的，抽象；这里说它其实就等于「同余方程解数的密度极限」。这一步把「证 $\chi(p) > 0$ 」变成「证模 p^ν 有足够多解」，后者是纯算术、可下手。

记号 $M(p^\nu)$ ：模 p^ν 的解数

$M(p^\nu)$ 数的是：在 $0 \leq x_j < p^\nu$

为什么分母是 $p^{\nu(n-1)}$ (这是关键的「正确刻度」): 设想 $C(\mathbf{x}) \equiv 0 \pmod{p^\nu}$ 这一个同余约束「随机地」筛掉点。在 $p^{\nu n}$ 个候选点中, 约束 $C \equiv 0$ 大致每 p^ν 个里允许 1 个, 所以「期望解数」约为 $p^{\nu n}/p^\nu = p^{\nu(n-1)}$ 。换句话说 $p^{\nu(n-1)}$ 是「如果方程模 p^ν 表现得完全随机、不偏不倚」时应有的解数。于是比值

$$\frac{M(p^\nu)}{p^{\nu(n-1)}} = \frac{\text{实际解数}}{\text{随机期望解数}}$$

就是「模 p^ν 这层, 解比随机多还是少」的相对密度。取 $\nu \rightarrow \infty$ 的极限, 就抓住了素数 p 处「越来越精细」的局部解密度——这正是 $\chi(p)$ 。所以 $\chi(p) > 0$ 的含义就是: 解数 $M(p^\nu)$ 始终保持在 $p^{\nu(n-1)}$ 这个正常档次, 不会随 ν 增大而相对地萎缩到 0。

「引理 5.3 的论证」为什么给出这个极限 (补全梗概): 这里用到一条标准的「指数和 $\leftrightarrow$ 解数」对偶 (第 5 章证过, 对一般型同样成立)。核心是同余的指数和判别式:

1 对任意整数 m , 有恒等式 $\frac{1}{p^\nu} \sum_{a=0}^{p^\nu-1} e\left(\frac{am}{p^\nu}\right) = \begin{cases} 1, & p^\nu \mid m \\ 0, & \text{否则} \end{cases}$ 。(这是「若干等距单位圆点之和: 整圈则相消为 0, 原地不动则全为 1」, 是检测「 m 是否被 p^ν 整除」的开关。)

2 把 $m = C(\mathbf{x})$ 代入并对所有 $\mathbf{x} \bmod p^\nu$ 求和, 左边数出的正是解数:

$$M(p^\nu) = \frac{1}{p^\nu} \sum_{a=0}^{p^\nu-1} \sum_{\mathbf{x} \bmod p^\nu} e\left(\frac{aC(\mathbf{x})}{p^\nu}\right) = \frac{1}{p^\nu} \sum_{a=0}^{p^\nu-1} S_{a,p^\nu}.$$

3 把 a 按它与 p^ν 的最大公约数分类 (即按既约分母 p^μ , $0 \leq \mu \leq \nu$) 整理, 化简后正好得到 $\frac{M(p^\nu)}{p^{\nu(n-1)}} = 1 + A(p) + A(p^2) + \cdots + A(p^\nu)$, 即 $\chi(p)$ 的有限截断 (部分和)。

4 令 $\nu \rightarrow \infty$, 右边的部分和收敛到 $\chi(p)$ (前面已证 $\sum_\nu A(p^\nu)$ 绝对收敛), 于是得到本段的极限公式。这就把「抽象的指数和因子」与「具体的同余解密度」严丝合缝地对上了。

因此，为了证明 $\mathfrak{S} > 0$ （在当前条件下），只需证明：对每一个素数 p ，都有

$$M(p^\nu) \geq C_p p^{\nu(n-1)}, \quad C_p > 0, \quad (17.1)$$

对一切充分大的 ν 成立。

这段在讲什么：把目标彻底「落地」成一个具体的算术不等式 (17.1)。

1 回顾逻辑链： $\mathfrak{S} = \prod_p \chi(p)$ ；大素数那截 $\geq \frac{1}{2} > 0$ ；所以 $\mathfrak{S} > 0$ 当且仅当每个（有限个小素数的） $\chi(p) > 0$ 。

2 而 $\chi(p) = \lim_{\nu} M(p^\nu)/p^{\nu(n-1)} \geq 0$ （解数非负，极限非负）。要它严格大于 0，等价于这个比值不趋于 0，即存在固定正常数 C_p 使比值 $\geq C_p$ 对大 ν 都成立——这正是 (17.1)： $M(p^\nu) \geq C_p p^{\nu(n-1)}$ 。

3 注意「对充分大的 ν 」就够了：因为我们只要极限 > 0 ，前面有限个 ν 怎样无所谓。

记号 C_p ：一个依赖素数 p 、与 ν 无关的正常数下界。它具体取多少不重要，重要的是「存在某个正的它」，托住解数密度不掉到 0。

原文译文

在处理加性型的型时，我们曾发现：同余式

$$a_1 x_1^k + \cdots + a_n x_n^k \equiv 0 \pmod{p^{\gamma_1}}$$

只要存在一个解，其中 $x_1, \dots, x_n$ 并非全都被 p 整除，就足以蕴含 (17.1) 对一切充分大的 ν 成立——这里 γ_1 是一个适当的指数，依赖于 p 和 k ，以及整除各系数 a_j 的 p 的方幂。

这段在讲什么：回顾「加性型」（每项形如系数乘单变量的 k 次幂、各变量分开）当年是怎么轻松搞定 (17.1) 的，好与下面「一般三次型」的困难做对比。

加性型 $a_1 x_1^k + \cdots + a_n x_n^k$

各变量「不打架」，每项只含一个变量。Waring 问题里就是这种。 k 是次数， a_j 是系数。

γ_1

一个事先算好的、足够大的固定指数。它的取法依赖素数 p 、次数 k 、以及系数里含多少个 p 因子。

「一个非全被 p 整除的解就够」是什么意思、为什么够：当年的经验是——只要在模 p^{γ_1} 下能找到一个解 $\mathbf{x}$ ，其中至少有一个坐标（比如 x_1 ）不被 p 整除，那么就能把这个解「逐次提升」到模 $p^{\gamma_1+1}, p^{\gamma_1+2}, \dots$ 的解，而且每提升一层，解数按 p^{n-1} 倍增长，恰好维持 (17.1) 的密度。提升能进行下去的关键，是该解处导数不全为 p 的高次幂——对加性型这一点自动成立（见下一段）。

原文译文

对一般的型，情形并不那么简单。看来人们需要的不仅仅是一个 $(\text{mod } p^{\gamma_1})$ 的解；人们需要这样一个解：其偏导数 $\partial C/\partial x_1, \dots, \partial C/\partial x_n$ 并非全都能被 p 的过高方幂整除。对于加性型，存在一个明显的限制，因为关于 x_j 的导数是 $a_j k x_j^{k-1}$ ，而总有某个 j 使 x_j 不被 p 整除。

这段在讲什么：点破一般三次型与加性型的本质差别，并由此引出本章后半的核心概念。先把「偏导数」讲清。

记号：偏导数 $\partial C/\partial x_i$

读「 C 对 x_i 的偏导数」。 ∂ （读「partial」）是偏微分符号。多变量函数 $C(x_1, \dots, x_n)$ 对某个变量 x_i 求偏导，意思是把其余变量都当常数，只对 x_i 用高中的求导法则。例如若 $C = x_1^3 + 5x_1x_2^2$ ，则 $\partial C/\partial x_1 = 3x_1^2 + 5x_2^2$ （ x_2 当常数）， $\partial C/\partial x_2 = 10x_1x_2$ 。这些偏导数合起来叫梯度，衡量 C 在各方向上变化的快慢。在「提升解」时，导数告诉我们「微微挪动 $\mathbf{x}$ 会让 C 变多少」，是能否调到 0 的关键。

为什么一般型更难、为什么需要「导数不被 p 高次幂整除」的解：

- 1 提升解的算法（下面引理 17.1 会精确做）大致是：手上有模 p^k 的解 $\mathbf{x}$ ，想把它修成模 p^{k+1} 的解，办法是令 $\mathbf{y} = \mathbf{x} + p^? \mathbf{u}$ ，用泰勒展开

$C(\mathbf{y}) \approx C(\mathbf{x}) + p^?(\mathbf{u} \cdot \nabla C)$, 再解一个关于 $\mathbf{u}$ 的一次同余方程把 C 调成 0。

2 这个一次方程的系数正是各偏导数 $\partial C/\partial x_i$ 。要它「可解且解多」, 就不能让所有偏导数都被 p 整除太多次——否则一次方程退化 (系数全是 0 模 p), 调不动 C , 提升失败。

3 对加性型, $\partial C/\partial x_j = a_j k x_j^{k-1}$ (这就是把 $a_j x_j^k$ 对 x_j 求导)。既然解里「总有某个 x_j 不被 p 整除」, 那一项的导数 $a_j k x_j^{k-1}$ (在 $p \nmid a_j k$ 时) 也不被 p 整除——所以导数自动「不全被 p 整除」, 限制是白送的。

4 对一般三次型, 变量交叉耦合, 没有这种白送的保证: 一个模 p^γ 的解完全可能所有偏导数都恰好被 p 高次整除 (落在「奇点」附近), 提升就卡住。所以必须额外要求找到一个「导数不太退化」的解。这就是下面定义 $\mathcal{A}(p^\ell)$ 要精确刻画的东西。

定义. 设 p 为素数, ℓ 为正整数。我们称 $C(\mathbf{x})$ 具有性质 $\mathcal{A}(p^\ell)$, 若存在

$$C(x_1, \dots, x_n) \equiv 0 \pmod{p^{2\ell-1}} \quad (17.2)$$

的一个解, 使得

$$\partial C/\partial x_i \equiv 0 \pmod{p^{\ell-1}} \quad \text{对一切 } i \quad (17.3)$$

且

$$\partial C/\partial x_i \not\equiv 0 \pmod{p^\ell} \quad \text{对某个 } i. \quad (17.4)$$

这个定义在说什么: 它精确地描述了我们需要的「好解」长什么样。性质 $\mathcal{A}(p^\ell)$ 要求存在一个解 $\mathbf{x}$, 同时满足三条:

$$(17.2) \quad C(\mathbf{x}) \equiv 0 \pmod{p^{2\ell-1}}$$

它是个相当深的解: 不只是模 p 为 0, 而是模 $p^{2\ell-1}$ 都为 0。指数取 $2\ell - 1$ 不是随便写的, 下面引理证明里会看到它正好和泰勒展开的精度匹配。

(17.3) 所有 $\partial C / \partial x_i \equiv 0 \pmod{p^{\ell-1}}$

所有偏导数至少被 $p^{\ell-1}$ 整除。这是「导数的退化程度有个上限标尺 ℓ 」的下半句——允许导数被 p 整除，但不要太多。

(17.4) 某个 $\partial C / \partial x_i \not\equiv 0 \pmod{p^\ell}$

至少有一个偏导数不被 p^ℓ 整除。配合 (17.3)：综合起来就是「至少有一个偏导数恰好被 $p^{\ell-1}$ 整除、但不被 p^ℓ 整除」，即它含 p 的次数正好是 $\ell - 1$ 。这个「恰好 $\ell - 1$ 」的导数就是提升解时用来「撬动」 C 的杠杆。

记号 $\not\equiv$ 、 ℓ

$\not\equiv$ 是 $\equiv$ 加一道斜杠，读「不同余」，即「除以模数余数不相同 / 不被模数整除」。 ℓ （手写体小写 L）是一个正整数参数，量度「导数退化的程度 / 解的深度刻度」。本章只承诺「存在某个 ℓ 使 $\mathcal{A}(p^\ell)$ 成立」，具体哪个 ℓ 留给下一章构造。

为什么用「 $2\ell - 1$ 配 $\ell - 1, \ell$ 」这套指数（取舍理由）：这是为了让下面的提升论证刚好闭合。直觉上，泰勒展开 $C(\mathbf{x} + p^m \mathbf{u}) = C(\mathbf{x}) + p^m (\mathbf{u} \cdot \nabla C) + O(p^{2m})$ 里，一次项的「有效精度」由导数含 p 的次数（约 $\ell - 1$ ）和挪动量 p^m 共同决定，而二次以上余项是 p^{2m} 级。要让「一次项能调动 C 、而二次余项又落在更高次幂被忽略」，三个指数必须按 $2\ell - 1 = (\ell - 1) + \ell$ 这种方式咬合。下面证明的每一处模数（ $p^{2\ell-1+\nu}$ 、 $p^{\ell+\nu}$ 、 $p^{2\ell+2\nu}$ 等）都是这套配比的产物。

引理 17.1. 设 $C(\mathbf{x})$ 具有性质 $\mathcal{A}(p^\ell)$ 。则

$$M(p^{2\ell-1+\nu}) \geq p^{(n-1)\nu},$$

从而 $\chi(p) > 0$ 。

引理在说什么：这是本章的「干活引理」。它断言：只要存在一个上面那种「好解」（即性质 $\mathcal{A}(p^\ell)$ 成立），就能保证模 p 的越来越高次幂始终有足够多解——具体地，模 $p^{2\ell-1+\nu}$ 至少有 $p^{(n-1)\nu}$ 个解。把这个下界代回 $\chi(p)$ 的极限定义，立刻得到 $\chi(p) > 0$ 。下面先解释「从引理结论怎么推出 $\chi(p) > 0$ 」，再细读证明。

1 在 $\chi(p) = \lim_{\nu \rightarrow \infty} M(p^\nu)/p^{\nu(n-1)}$ 里，沿着 $\nu' = 2\ell - 1 + \nu$ 这一串取值（这是 ν 的一个子列， $\nu \rightarrow \infty$ 时 $\nu' \rightarrow \infty$ ，对收敛序列其极限相同）。

2 代入引理下界：

$$\frac{M(p^{2\ell-1+\nu})}{p^{(2\ell-1+\nu)(n-1)}} \geq \frac{p^{(n-1)\nu}}{p^{(2\ell-1+\nu)(n-1)}} = p^{(n-1)\nu - (2\ell-1+\nu)(n-1)} = p^{-(2\ell-1)(n-1)}$$

。

3 右边 $p^{-(2\ell-1)(n-1)}$ 是一个与 ν 无关的固定正数。所以比值有正的下界，取极限得 $\chi(p) \geq p^{-(2\ell-1)(n-1)} > 0$ 。这恰好就是 (17.1) 中那个 C_p 。 $\chi(p) > 0$ 证毕。

证。

证明总思路：用对 ν 的数学归纳法。每一步把「模 $p^{2\ell-1+\nu}$ 的好解」一个个提升成「模 $p^{2\ell-1+\nu+1}$ 的好解」，并且每提升一层，解数翻 p^{n-1} 倍。从 $\nu = 0$ （一个解）出发，提升 ν 次就有 $p^{(n-1)\nu}$ 个，正是要证的下界。这就是数论里著名的 **Hensel 提升**（Hensel lifting）的思想。

归纳命题（要证的东西）：同余式

$$C(x_1, \dots, x_n) \equiv 0 \pmod{p^{2\ell-1+\nu}} \quad (17.5)$$

至少有 $p^{(n-1)\nu}$ 个满足 (17.3) 与 (17.4) 的解，且这些解关于模 $p^{\ell+\nu}$ 两两不同余。

逐句解读归纳命题：

1 「满足 (17.3) 与 (17.4)」= 提升过程中始终保持「好解」的导数性质，这样下一步才提得动（导数杠杆没丢）。

2 「关于模 $p^{\ell+\nu}$ 两两不同余」是个更强、更细的要求：不仅这些解彼此不同，而且就算只看它们模 $p^{\ell+\nu}$ （较粗的尺度）也互不相同。这一条是为了**数解数时不重复**：因为 $p^{\ell+\nu} \leq p^{2\ell-1+\nu}$ （当 $\ell \geq 1$ 时 $\ell + \nu \leq 2\ell - 1 + \nu$ ），模

$p^{\ell+\nu}$ 不同 $\Rightarrow$ 模 $p^{2\ell-1+\nu}$ 当然更不同 (原文「a fortiori 更是如此」), 所以这些解作为 $M(p^{2\ell-1+\nu})$ 的解互不相同, 可以放心计数。

- 3 归纳起点 $\nu = 0$: 命题变成「模 $p^{2\ell-1}$ 至少有 $p^0 = 1$ 个满足 (17.3)(17.4) 的解」。这正是性质 $\mathcal{A}(p^\ell)$ 的假设本身——它就保证存在这样一个解。起点免费成立。

归纳步骤的核心工具——三次型的「泰勒展开」(补全为什么成立): 对任意整数 $x_1, \dots, x_n, u_1, \dots, u_n$, 有

$$C(\mathbf{x} + p^{\ell+\nu}\mathbf{u}) \equiv C(\mathbf{x}) + p^{\ell+\nu}(u_1\partial C/\partial x_1 + \dots + u_n\partial C/\partial x_n) \pmod{p^{2\ell+2\nu}}$$

- 1 这是多元泰勒展开。把 $\mathbf{x}$ 挪动 $p^{\ell+\nu}\mathbf{u}$, 按挪动量的幂次展开: 常数项 $C(\mathbf{x})$, 一次项 $p^{\ell+\nu}\sum_i u_i \partial C/\partial x_i$, 二次及更高项。
- 2 因为 C 是三次多项式, 泰勒展开到三次项就截止 (四阶以上偏导数全为 0), 所以余项只有「二次项 + 三次项」。二次项含因子 $(p^{\ell+\nu})^2 = p^{2\ell+2\nu}$, 三次项含 $(p^{\ell+\nu})^3$, 都被 $p^{2\ell+2\nu}$ 整除。
- 3 于是「模 $p^{2\ell+2\nu}$ 」下, 二次及以上余项全部消失, 只剩常数项加一次项。这就是上式——它是精确成立的同余, 不是近似。这正是前面「为什么挑指数 $2\ell - 1$ 」的回报: 挪动量的平方刚好够高, 把非线性部分扫进可忽略的模数里。

归纳推进——假设第 ν 层成立, 造第 $\nu + 1$ 层:

- 1 取第 ν 层的任意一个好解 $\mathbf{x} = (x_1, \dots, x_n)$ (共 $p^{(n-1)\nu}$ 个、模 $p^{\ell+\nu}$ 两两不同)。
- 2 由 (17.5) 知 $p^{2\ell-1+\nu} \mid C(\mathbf{x})$, 写 $C(\mathbf{x}) = ap^{2\ell-1+\nu}$ (a 为整数)。由 (17.3) 知 $p^{\ell-1} \mid \partial C/\partial x_i$, 写 $\partial C/\partial x_i = D_i p^{\ell-1}$ (D_i 为整数)。由 (17.4) 知某个

导数不被 p^ℓ 整除，即该 i 处 $D_i p^{\ell-1}$ 含 p 不到 ℓ 次，故 $p \nmid D_i$ ，即存在 i 使 $D_i \not\equiv 0 \pmod{p}$ 。这个非零的 D_i 就是接下来解一次方程的「非退化系数」。

- 3 目标：找挪动 $\mathbf{y} = \mathbf{x} + p^{\ell+\nu} \mathbf{u}$ ，使 $C(\mathbf{y}) \equiv 0 \pmod{p^{2\ell+\nu}}$ （注意模数从 $2\ell-1+\nu$ 升到 $2\ell+\nu$ ，深了一层）。把泰勒展开代入，并注意 $2\ell+\nu \leq 2\ell+2\nu$ ($\nu \geq 0$)，故那条模 $p^{2\ell+2\nu}$ 的展开在更粗的模 $p^{2\ell+\nu}$ 下也成立：

$$C(\mathbf{y}) \equiv C(\mathbf{x}) + p^{\ell+\nu} \sum_i u_i \partial C / \partial x_i \pmod{p^{2\ell+\nu}}.$$

- 4 代入 $C(\mathbf{x}) = ap^{2\ell-1+\nu}$ 与 $\partial C / \partial x_i = D_i p^{\ell-1}$ ：

$$C(\mathbf{y}) \equiv ap^{2\ell-1+\nu} + p^{\ell+\nu} \sum_i u_i D_i p^{\ell-1} = p^{2\ell-1+\nu} \left(a + \sum_i D_i u_i \right) \pmod{p^{2\ell+\nu}}$$

（合并幂次： $p^{\ell+\nu} \cdot p^{\ell-1} = p^{2\ell-1+\nu}$ ，与第一项同次，提取公因子。）

- 5 要让上式 $\equiv 0 \pmod{p^{2\ell+\nu}}$ ，等价于括号 $(a + \sum_i D_i u_i)$ 被 p 整除（因为已经有公因子 $p^{2\ell-1+\nu}$ ，再补一个 p 就到 $p^{2\ell+\nu}$ ）。即只需

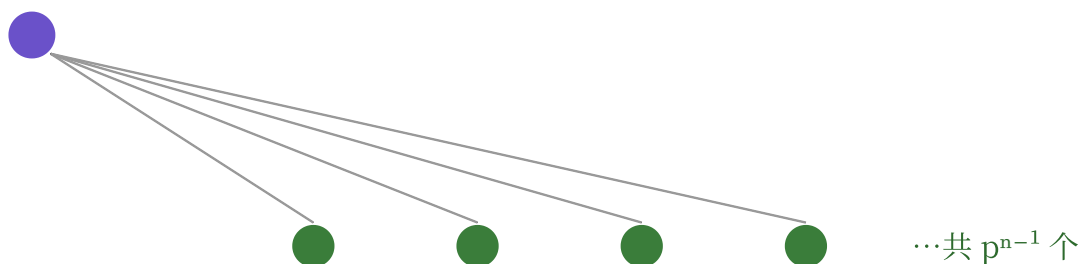
$$a + D_1 u_1 + \cdots + D_n u_n \equiv 0 \pmod{p}.$$

非线性的三次方程被化成了一个关于 $\mathbf{u}$ 的一次同余方程！这正是提升能进行的关键：泰勒展开里二次及以上项都落到了模 $p^{2\ell+\nu}$ 以外，于是每层只需解一个线性同余方程。

- 6 数这个一次方程的解：因为某个 $D_i \not\equiv 0 \pmod{p}$ （第 2 步），系数不全为 0，这是模素数 p 的一个非退化线性方程。 n 个未知数 $u_1, \dots, u_n$ （各取 $0 \sim p-1$ ），一条线性约束，解数恰为 p^{n-1} （自由地选 $n-1$ 个变量，剩下那个被非零系数 D_i 唯一定出）。这 p^{n-1} 个 $\mathbf{u}$ 关于模 p 两两不同。

Hensel 提升：每层每个好解分裂成 p^{n-1} 个更深的好解

$\text{mod } p^{2^{l-1}+\nu}$ 的一个解 $\mathbf{x}$



$$\mathbf{y} = \mathbf{x} + p^{\ell+\nu} \cdot \mathbf{u}, \text{ 其中 } \mathbf{u} \text{ 解一次同余 } a + \sum D_i u_i \equiv 0 \pmod{p}$$

每个 $\mathbf{y}$ 满足 $C(\mathbf{y}) \equiv 0 \pmod{p^{2^{l-1}+\nu}}$, 且仍是「好解」

$$\text{解数: } p^{n-1\nu} \times p^{n-1} = p^{n-1(\nu+1)}$$

归纳一步：把挪动写成 $\mathbf{y} = \mathbf{x} + p^{\ell+\nu} \mathbf{u}$, 泰勒展开后三次方程退化一次同余, 恰有 p^{n-1} 个解 $\mathbf{u}$ 。解数每层乘 p^{n-1} 。

收尾——验证新解仍是「好解」且不重复计数：

1 对每个旧解 $\mathbf{x}$, 得到 p^{n-1} 个新值 $\mathbf{y} = \mathbf{x} + p^{\ell+\nu} \mathbf{u}$, 它们满足 $C(\mathbf{y}) \equiv 0 \pmod{p^{2^{l-1}+\nu}}$, 即 (17.5) 在 $\nu+1$ 层 (模 $p^{2^{l-1}+(\nu+1)} = p^{2^{l-1}+\nu}$) 成立。✓ 深度提升到位。

2 导数性质 (17.3)(17.4) 自动遗传：由 $\mathbf{y} \equiv \mathbf{x} \pmod{p^{\ell+\nu}}$, 对任意多项式 $\partial C / \partial x_i$ (它的偏导也是多项式), 同余 $\mathbf{y} \equiv \mathbf{x}$ 蕴含 $\partial C / \partial x_i(\mathbf{y}) \equiv \partial C / \partial x_i(\mathbf{x}) \pmod{p^{\ell+\nu}}$ 。由于 $\ell + \nu \geq \ell$ (且 $\geq \ell - 1$), 模 $p^{\ell+\nu}$ 相等 $\Rightarrow$ 模 $p^{\ell-1}$ 、模 p^ℓ 也相等。所以 $\mathbf{y}$ 处各导数与 $\mathbf{x}$ 处「含 p 的次数」一致, (17.3)「全被 $p^{\ell-1}$ 整除」与 (17.4)「某个不被 p^ℓ 整除」原封不动传给 $\mathbf{y}$ 。✓ 仍是好解。

3 计数不重复：总共得到 $p^{(n-1)\nu} \times p^{n-1} = p^{(n-1)(\nu+1)}$ 个 $\mathbf{y}$ 。要确认它们模 $p^{\ell+\nu+1}$ 两两不同：① 来自不同旧解 $\mathbf{x}, \mathbf{x}'$ 的 $\mathbf{y}, \mathbf{y}'$ ——因为旧解模 $p^{\ell+\nu}$ 已不同, 而 $\mathbf{y} \equiv \mathbf{x}, \mathbf{y}' \equiv \mathbf{x}' \pmod{p^{\ell+\nu}}$, 故 $\mathbf{y}, \mathbf{y}'$ 模 $p^{\ell+\nu}$ 就不同, 模更高的 $p^{\ell+\nu+1}$ 更不同；② 来自同一旧解 $\mathbf{x}$ 的两个 $\mathbf{y} = \mathbf{x} + p^{\ell+\nu} \mathbf{u}$ 与 $\mathbf{y}' = \mathbf{x} + p^{\ell+\nu} \mathbf{u}'$ ——它们模 $p^{\ell+\nu+1}$ 相等当且仅当 $p^{\ell+\nu} \mathbf{u} \equiv p^{\ell+\nu} \mathbf{u}' \pmod{p^{\ell+\nu+1}}$, 即 $\mathbf{u} \equiv \mathbf{u}' \pmod{p}$ ；而那 p^{n-1} 个 $\mathbf{u}$

本就模 p 两两不同，故 $\mathbf{y} \neq \mathbf{y}'$ 。✓ 共 $p^{(n-1)(\nu+1)}$ 个、模 $p^{\ell+\nu+1}$ 互不同余。

- 4 于是归纳命题对 $\nu + 1$ 成立。由数学归纳法，对一切 $\nu \geq 0$ 成立，特别地解数 $\geq p^{(n-1)\nu}$ ，即 $M(p^{2\ell-1+\nu}) \geq p^{(n-1)\nu}$ 。再由引理前面那段推出 $\chi(p) > 0$ 。 ■

原文译文

关于对每个 p 都存在某个 ℓ 使得 $C(\mathbf{x})$ 具有性质 $\mathcal{A}(p^\ell)$ 的证明，构成下一章的主题。

这段在讲什么：交代「还差最后一块拼图」与下章预告。本章已经把链条接到只剩一个缺口：

- 1 本章证明了：若 C 对每个素数 p 都具有某个性质 $\mathcal{A}(p^\ell)$ ，则每个 $\chi(p) > 0$ （引理 17.1），从而 $\mathfrak{S} = \prod_p \chi(p) > 0$ ，再由定理 17.1 得方程有非平凡解。

- 2 但「每个 p 都真的存在这样的 ℓ 」本章并没有证——这需要对一般三次型在 p -进数（ p -adic）层面构造出那个「导数不太退化的好解」。这是个独立、相当技术性的命题。

- 3 这块缺口正是下一章（第 18 章「三次型： p -进问题」）的全部任务。补上它，整套「 $n \geq 17$ 的三次型必有非平凡整数解」的证明才真正完工。

本章脉络回顾（一句话串起来）

计数公式 $\mathcal{N}(P) \approx P^{n-3} \mathfrak{S} J_0 \rightarrow$ 只需 $\mathfrak{S} > 0$ （定理 17.1，反证） $\rightarrow A(q)$ 乘性 + 绝对收敛 $\Rightarrow$ 欧拉乘积 $\mathfrak{S} = \prod_p \chi(p) \rightarrow$ 大素数那截 $\geq \frac{1}{2}$ （被托住），只剩有限个小素数 $\rightarrow \chi(p) = \lim M(p^\nu)/p^{\nu(n-1)}$ ，于是只需同余解数够多 (17.1) $\rightarrow$ 引入「好解」性质 $\mathcal{A}(p^\ell)$ ，用 Hensel 逐层提升（引理 17.1）得 $\chi(p) > 0 \rightarrow$ 唯一遗留：每个 p 都有这样的 ℓ （下章证）。

返回 [全书目录](#)