

DAVENPORT · 圆法 · 高中详解版

三次型： p 进问题

Cubic forms: the p -adic problem

本章要解决什么 整本"三次型"系列（第 13–17 章）用圆法去数方程 $C(\mathbf{x}) = 0$ 的整数解，最后归结为一句话：只要"奇异级数" $\mathfrak{S} > 0$ ，方程就有非零整数解（第 17 章定理 17.1，需 $n \geq 17$ ）。而 $\mathfrak{S} > 0$ 又被拆成"对每个素数 p ，局部密度 $\chi(p) > 0$ "。第 17 章证明了：只要 C 具有"性质 $\mathcal{A}(p^\ell)$ "（某个 ℓ ），就有 $\chi(p) > 0$ 。

本章就补上最后这块拼图：证明任何 $n \geq 10$ 的非退化整系数三次型，对每个素数 p 都具有性质 $\mathcal{A}(p^\ell)$ （ ℓ 有与 p 无关的上界）。读完你将掌握：什么叫" p 进的非奇异解"、Chevalley 定理怎样用、以及 Davenport 设计的一个非常巧妙的"下降 + 不变量计数"论证——它把一个"无穷多个同余式都不可解"的假设，逼成一个关于整数 $h(C)$ 的矛盾。

读前约定 本页黑色叙述紧扣译文逐段，彩色框（目标 / 分步推演 / 符号卡 / 图）是为高中生补的细节。本章只用到：整除与同余、最大公因数、行列式、线性变换、（多元）多项式求导。这些下面都会从零讲。凡译文写"显然""容易看出"的地方，我们一律展开成一步一句。

0 预备：把本章会用到的"超纲"概念一次讲清

本章几乎每一步都依赖几个"大学才系统学、但其实高中能听懂"的概念。先集中讲透，后面就只管用。

符号卡：三次型 $C(\mathbf{x})$ 与它的系数 c_{ijk} "型"(form) 就是各项次数都相同的齐次多项式。三次型 = 每一项都是三个变量相乘（可重复）的多项式，例如 $x_1^3, x_1^2x_2, x_1x_2x_3$ 都是三次单项式。 n 个变量的三次型一般写成

$$C(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n c_{ijk} x_i x_j x_k,$$

其中 $\mathbf{x} = (x_1, \dots, x_n)$ 是变量组（粗体表示"一串数"），系数 c_{ijk} 约定关于下标 i, j, k 对称（即交换下标不改变值，例如 $c_{112} = c_{121} = c_{211}$ ）。这样写虽有重复，但好处是求导很整齐。读作" c -下标- $i j k$ "。

超纲点：齐次性 $C(p\mathbf{x}) = p^3 C(\mathbf{x})$ 每项三次，把所有变量同时乘 p ，每项就多出 $p \cdot p \cdot p = p^3$ 。于是

$$C(px_1, \dots, px_n) = p^3 C(x_1, \dots, x_n).$$

这条"提取公因子 p^3 "的小事，是本章下降法的发动机，请记牢。

超纲点：偏导数 $\partial C/\partial x_i$ 与欧拉关系 $\partial C/\partial x_i$ (读"偏 C 偏 x_i ") 表示：把除 x_i 外的变量都看成常数，只对 x_i 求导。例如对 $C = x_1^2 x_2$, $\partial C/\partial x_1 = 2x_1 x_2$, $\partial C/\partial x_2 = x_1^2$ 。对上面的对称写法，可算出一个干净公式

$$\frac{1}{3} \frac{\partial C}{\partial x_i} = \sum_j \sum_k c_{ijk} x_j x_k, \quad (0.1)$$

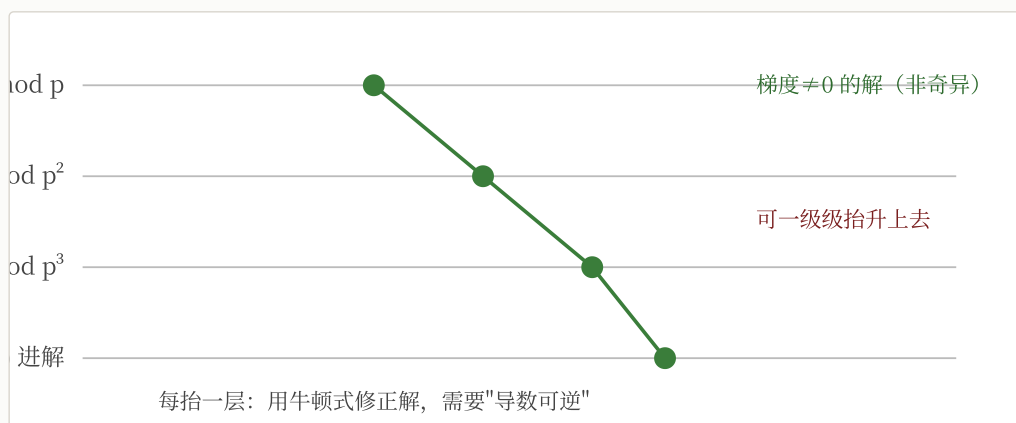
右边是个二次型 (关于 $x_1, \dots, x_n$ 次数都是 2)。它在本章是核心： n 个偏导 ($i = 1, \dots, n$) 给出 n 个二次型，它们的系数恰好就是 c_{ijk} 。还有欧拉关系： $x_1 \frac{\partial C}{\partial x_1} + \dots + x_n \frac{\partial C}{\partial x_n} = 3C$ 。

超纲点：同余 $a \equiv b \pmod{m}$ " $a \equiv b \pmod{m}$ " 读作" a 与 b 模 m 同余"，意思是 $m \mid (a - b)$ (m 整除 $a - b$)，即 a, b 除以 m 余数相同。例如 $17 \equiv 2 \pmod{5}$ 。本章的模总是素数的幂 p^k 。"方程 $C(\mathbf{x}) \equiv 0 \pmod{p^k}$ 可解"就是说能找到整数 $\mathbf{x}$ 使 $C(\mathbf{x})$ 是 p^k 的倍数。

超纲点： p 进数与"非奇异解" 高中只解"实数解"或"整数解"。数论里还有第三种世界： p 进数域 $\mathbb{Q}_p$ 。你不必懂它的构造，只需记住一个等价说法 (这是本章第一段的内容)：

- " $C(\mathbf{x}) = 0$ 在 $\mathbb{Q}_p$ 中有解" $\approx$ "对每个 k ，同余式 $C(\mathbf{x}) \equiv 0 \pmod{p^k}$ 都有解"。
- 非奇异解：一个解 $\mathbf{x}$ ，在它处梯度不全为 0 (某个 $\partial C/\partial x_i \neq 0$)。

非奇异为什么重要？因为有"亨泽尔引理(Hensel)"——它像牛顿迭代：只要在 $\text{mod } p$ (或某个 p^k) 下有一个梯度不为 0 的解，就能把它一级一级精确"抬升"到 $\text{mod } p^{k+1}, p^{k+2}, \dots$ ，直到 p 进解。性质 $\mathcal{A}(p^\ell)$ (下一卡) 就是这种"可抬升"的定量版本。



非奇异 (梯度不为零) 就像"导数不为零"，使解能沿模数塔 $p, p^2, p^3, \dots$ 一路抬升。本章要保证这种解存在。

超纲点：性质 $\mathcal{A}(p^\ell)$ (第 17 章的定义，本章天天用) 设 p 是素数、 $\ell \geq 1$ 是整数。称 $C(\mathbf{x})$ 具有性质 $\mathcal{A}(p^\ell)$ ，如果存在整数 $\mathbf{x}$ 同时满足三条：

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

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

$$\frac{\partial C}{\partial x_j} \not\equiv 0 \pmod{p^\ell} \quad (\text{对某个 } j). \quad (17.4)$$

直观读法："存在一个解，它把 C 压到很深 ($p^{2\ell-1}$ 的倍数)，同时它的梯度恰好被 $p^{\ell-1}$ 整除——不多不少"。 $\ell = 1$ 是最简单的情形：(17.2) 是 $C \equiv 0 \pmod{p}$ ，(17.3) 自动成立 ($p^0 = 1$ 整除一切)，(17.4) 是"梯度 mod p 不全为 0"。所以"具有 $\mathcal{A}(p)$ " = "有一个 mod p 的非奇异解"。性质 $\mathcal{A}(p^\ell)$ 就是"非奇异"在 p 进意义下的定量化身：它正好够喂给亨泽尔引理把解抬升上去，从而保证 $\chi(p) > 0$ 。

缩写： $p^a \parallel (g_1, \dots, g_n)$ 本章把 (17.3) + (17.4) 合写成

$$p^{\ell-1} \parallel \left(\frac{\partial C}{\partial x_1}, \dots, \frac{\partial C}{\partial x_n} \right),$$

读作" $p^{\ell-1}$ 恰好整除这组偏导"：意思是 $p^{\ell-1}$ 整除其中每一个，但 p^ℓ 至少漏掉一个（即不能整除全部）。把它想成"这组数的最大公因子中 p 的幂次恰是 $\ell - 1$ "。

超纲点：最大公因数 gcd、行列式与 n 阶子式 最大公因数 $\gcd(a_1, \dots, a_t)$ ：能同时整除全部 a_i 的最大正整数（本章记作 h.c.f. 或直接说"最大公因数"）。**行列式 det**：一个 $n \times n$ 数表算出的一个数，几何上是" n 个行向量张成的盒子的有向体积"，为 0 当且仅当这 n 行**线性相关**（即某行是其它行的线性组合）。从一个"行少列多"的矩阵里任取 n 列、与全部 n 行组成的方阵的行列式，叫一个 n 阶**子式** (minor)。本章的关键量 $h(C)$ 就是某矩阵全部 n 阶子式的最大公因数。

超纲点：线性变换、等价、幺模变换 **线性变换**：用整数把变量换一套， $x_i = \sum_r q_{ir} y_r$ ，矩阵 $Q = (q_{ir})$ 。它的行列式 $\det Q = q$ 衡量"体积放大倍数"。若两个型经一个行列式 ± 1 的**整系数线性变换**（叫**幺模变换**，unimodular）互相转化，称它们**等价**。等价的型，作为数论对象"长得一样"：能表示的整数集合、有无整数零点等等都相同。幺模变换可逆且逆也是整系数，所以是数论里"换坐标"的标准操作。

超纲点：Chevalley (–Warning) 定理 本章引理 18.3 的引擎。它说：若一个齐次多项式 $(\text{mod } p)$ 的变量个数严格大于它的次数，则它在 mod p 下必有非零解。三次型次数为 3，所以只要变量数 $n > 3$ （即 $n \geq 4$ ）， $C(\mathbf{x}) \equiv 0 \pmod{p}$ 一定有不全为 0 的解。直觉：变量太多、约束太松，"鸽子"多到必有非零解落进 0 这个"笼子"。注意临界处 $n = 3$ （三变量三次）时不保证，这正是本章下降会停在"3 个变量"的原因。

超纲点：退化 / 非退化三次型 一个型若能经某个（可逆的）线性变换写成**更少变量**的型（真的少用了某个变量），就叫**退化**；否则**非退化**。退化型本质上是"假的 n 元型"，问题可降到更少变量，所以圆法默认研究非退化型。引理 18.2 会把"非退化"翻译成关于 $h(C)$ 的一句话。

1 开篇：把" p 进解"翻译成"性质 $\mathcal{A}(p^\ell)$ "

译文第 1 段：断言"对任何 p 都存在某个 ℓ ，使 C 具有性质 $\mathcal{A}(p^\ell)$ "等价于"方程 $C(\mathbf{x}) = 0$ 在 p 进数域中有非奇异解"。我们将证明 $n \geq 10$ 时这成立。若干数学家各自独立证明了 $n \geq 10$ 时 $C = 0$ 总有非平凡 p 进解；任何一个证明都够用，因为（Lewis 教授指出）从任何非平凡解都能推出非奇异解。但作者宁用自己的证明，因为它就是为直接导向 $\mathcal{A}(p^\ell)$ 而设计的。

这段在讲什么。它交代本章的目标与策略选择。目标量是性质 $\mathcal{A}(p^\ell)$ （上面符号卡已讲透：它=" p 进非奇异解"的定量版）。作者告诉我们：这个事实早有别人证明（" $n \geq 10$ 时有非平凡 p 进解"），而且 Lewis 教授提醒——非平凡解可升级为非奇异解，所以那些证明也能用。但作者偏要重写一个证明。

为什么"另起炉灶"——取舍解读 别人的定理给的是"有解"这个定性结论，要再花一道工序把"非平凡"升成"非奇异"、再升成"性质 $\mathcal{A}(p^\ell)$ "。作者的证明则直接产出 $\mathcal{A}(p^\ell)$ 并附带 ℓ 的上界。圆法第 17 章要的恰恰是 $\mathcal{A}(p^\ell)$ （而非笼统的"有解"），所以"为目标量身定做"的证明衔接最顺、最自给自足。这就是"宁用自己证明"的真正理由：不是偏好，而是输出格式正好对接下游。

历史与人物 " p 进数"由 Hensel（亨泽尔，约 1897）引入。"齐次方程在所有 $\mathbb{Q}_p$ 与 $\mathbb{R}$ 中有解 $\implies$ 在 $\mathbb{Q}$ 中有解"这一类"局部—整体"思想，是 20 世纪数论的主线之一（Hasse 原理）。三次型恰是 Hasse 原理可能失败的最低次数（Cassels–Guy 给过反例），所以这里不能空谈"局部有解就行"，必须老老实实地对每个 p 逐一证明局部条件 $\mathcal{A}(p^\ell)$ ，再靠 n 足够大（ $n \geq 17$ ）的圆法主项压过去。文中 D. J. Lewis 是 1950–60 年代研究三次型 p 进与有理解的代表人物，作者 H. Davenport 与他同代且常通信，这条致谢是真实学术互动的痕迹。

2 搭建不变量 $h(C)$ ：把"系数"摆成一张大矩阵

译文第 2 段：设 $N = \frac{1}{2}n(n+1)$ 。设 C 为 n 行 N 列的矩阵，一般元素 c_{ijk} ， i 标行，数对 j, k （ $j \leq k$ ）标列（这些数对按固定次序排列）。设 Δ 为由 C 任意 n 列构成的典型 n 阶行列式，可能的个数为 $\binom{N}{n}$ 。必要时给 C 乘因子 6，可设诸 c_{ijk} 为整数，于是各 Δ 为整数。

这段在讲什么。它造出本章的主角矩阵 C 和量 $h(C)$ 。动机是：我们要给三次型配一个能反映"整数系数 p 倍结构"的整数不变量，好在最后一步（引理 18.7）用它"封顶"。先把记号一个个拆开。

$$N = \frac{1}{2}n(n+1)$$

满足 $1 \leq j \leq k \leq n$ 的无序数对 (j, k) 的个数。算法： $j = k$ 的有 n 个， $j < k$

C (花体 C)

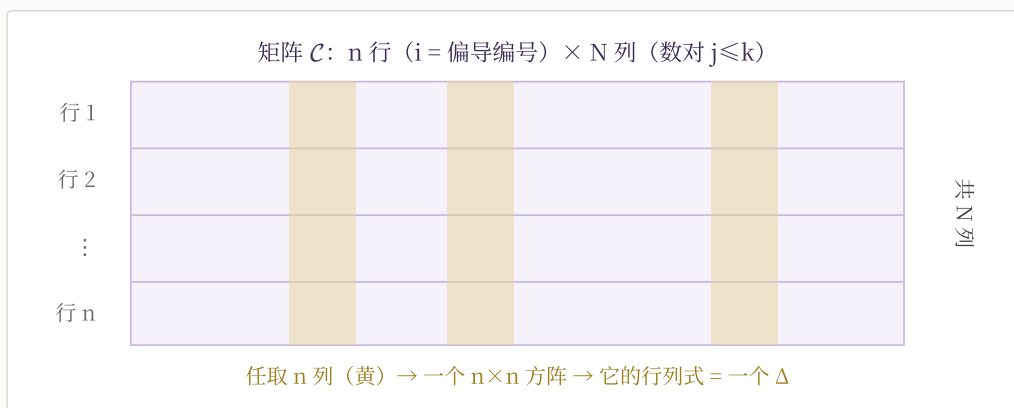
一张 $n \times N$ 的数表。第 i 行 = 第 i 个偏导 $\frac{1}{3} \frac{\partial C}{\partial x_i} = \sum_{j \leq k} (\cdots) x_j x_k$ 的全部系数排成一行；列由数对 (j, k) 标。一句话： C 的 n 行装着 C 的 n 个"偏导二次型"的系数。

Δ

从 N 列中任取 n 列、配上全部 n 行得到的 $n \times n$ 方阵的行列式（一个 n 阶子式）。这样的取法共有"从 N 个里选 n 个"种，即 $\binom{N}{n}$ 个。

1 为什么乘 6 就能让系数变整数？ 原始 C 系数可能是分数。三次型最"坏"的分母来自把对称化展开时出现的 $2, 3$ （ $x_i x_j x_k$ 三个互异下标的对称项会带 $\frac{1}{3!} = \frac{1}{6}$ 之类）。 $6 = 3!$ 是次数 3 的对称型清分母的标准公因子：给整个 C 乘 6，所有 c_{ijk} 同时变成整数。乘常数不改变方程 $C = 0$ 的解，也不改变后面要用的同余/等价结构，所以这是"无害的归一化"。

2 系数整数 $\Rightarrow$ 每个 Δ 整数：行列式是若干" n 个矩阵元相乘再带正负号求和"，整数加减乘仍是整数。



C : 行装偏导二次型的系数, 列由数对 (j, k) 标。从 N 列中挑 n 列就得到一个 n 阶子式 Δ ; 全部 $\binom{N}{n}$ 个 Δ 的最大公因数就是 $h(C)$ 。

定义 ($h(C)$) $h(C)$ = 全部 Δ 的最大公因数 (若它们不全为 0); 若全为 0, 令 $h(C) = 0$ 。

为什么要造 $h(C)$ ——动机预告 后面会证三件事, 环环相扣: (一) 做一个行列式为 q 的系数变换, 会让 h 至少"乘进"一个 q (引理 18.1); (二) 非退化 $\Leftrightarrow h(C) \neq 0$ (引理 18.2); (三) 若 C 一直没有性质 $\mathcal{A}$, 就能反复施加行列式为 p^{n-9} 的变换 (引理 18.5+18.6+18.7)。把三者合起来: $h(C)$ 必须被越来越高的 p 幂整除, 但 $h(C)$ 是个固定的正整数, 幂次不能无限大——矛盾! 于是"一直没有 $\mathcal{A}$ "不可能。 $h(C)$ 就是这个"封顶用的有限砝码"。

3 引理 18.1: 变换如何改变 $h(C)$

引理 18.1 设 $x'_i = \sum_{r=1}^n q_{ir} x_r$ ($1 \leq i \leq n$) 是整系数、行列式 $q \neq 0$ 的线性变换, 且 $C(x_1, \dots, x_n) = C'(x'_1, \dots, x'_n)$ 恒成立。则 $h(C)$ 能被 $qh(C')$ 整除。

怎么读这条引理。把 C 用一个"坐标变换"写成 C' (变换的"体积倍率"是 q)。结论: $h(C)$ 是 $qh(C')$ 的倍数。也就是说, 每做一次行列式为 q 的变换, 新型的 h 乘上 q 后仍整除旧型的 h 。这正是上面"封顶"机制要的"乘 q "。

证 (译文): 两型系数由 $c_{rst} = \sum_{i,j,k} q_{ir} q_{js} q_{kt} c'_{ijk}$ 相联系…… (随后化为矩阵式 $C = Q^T C' U$, 再用 Cauchy-Binet 公式与子式整除性收尾)。

1 系数变换公式的来历。变换是 $x'_i = \sum_r q_{ir} x_r$, 且 $C(\mathbf{x}) = C'(\mathbf{x}')$ 。把 $C'(\mathbf{x}') = \sum_{ijk} c'_{ijk} x'_i x'_j x'_k$ 里每个 x'_i 换成 $\sum_r q_{ir} x_r$:

$$C'(\mathbf{x}') = \sum_{ijk} c'_{ijk} \left(\sum_r q_{ir} x_r \right) \left(\sum_s q_{js} x_s \right) \left(\sum_t q_{kt} x_t \right).$$

把三个括号乘开、按 $x_r x_s x_t$ 收集系数，就得到 $C(\mathbf{x}) = \sum_{rst} c_{rst} x_r x_s x_t$ 中

$$c_{rst} = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n q_{ir} q_{js} q_{kt} c'_{ijk}.$$

这一步只是"代入 + 展开 + 配项"，没有任何技巧。

- 2 把数对压成单指标。造矩阵时我们已固定了一一对应："数对 (j, k) ($j \leq k$)" $\leftrightarrow$ "整数 $\mu \in \{1, \dots, N\}$ "。于是 C' 的系数 c'_{ijk} 改记 $c'_{i\mu}$ (i 是行号, μ 是列号); 同理把 (s, t) 记成 ν , C 的元素记 $c_{r\nu}$ 。这只是换记号, 方便写成矩阵。

- 3 引入 $\mathcal{U} = (u_{\mu\nu})$ 处理"无序数对"。上面 c_{rst} 的求和里, s, t 是有序跑遍 $1..n$ 的; 但列只认无序数对 $\nu = (s, t)$ ($s \leq t$)。当 $\setminus(s$

- 4 写成矩阵乘积。上式恰好是三个矩阵相乘的 (r, ν) 元: 左乘 $\mathcal{Q}^T$ (即 q_{ir} , 转置把 i 求和放到左指标)、中间 $\mathcal{C}' = (c'_{i\mu})$ 、右乘 $\mathcal{U} = (u_{\mu\nu})$:

$$\mathcal{C} = \mathcal{Q}^T \mathcal{C}' \mathcal{U},$$

其中 $\mathcal{Q} = (q_{ir})$ 是 $n \times n$, $\mathcal{U}$ 是 $N \times N$, T 是转置。这是把第 1 步那个三重和"打包"成的矩阵语言, 便于谈子式。

- 5 取一个 n 阶子式。设 Δ 是 $\mathcal{C}$ 的第 $\nu_1, \dots, \nu_n$ 列组成的行列式, 记 $\Delta = (\det \mathcal{C})_{\nu_1, \dots, \nu_n}^{1, \dots, n}$ (上标是取的行, 全取; 下标是取的列)。因为 $\mathcal{C} = \mathcal{Q}^T (\mathcal{C}' \mathcal{U})$, 取这 n 列时, 左边的 $\mathcal{Q}^T$ 是完整的 $n \times n$ 方阵 (它不动列、只搅行)。对方阵乘积取行列式:
 $\det(\mathcal{Q}^T M) = \det(\mathcal{Q}^T) \det M = q \det M$ 。所以

$$\pm \Delta = q (\det \mathcal{C}' \mathcal{U})_{\nu_1, \dots, \nu_n}^{1, \dots, n}.$$

($\det \mathcal{Q}^T = \det \mathcal{Q} = q$ 。"±"只是提醒: 换列顺序或定义里行列号约定可能差个符号, 不影响整除性。)

- 6 Cauchy-Binet 公式 ("乘积矩阵的子式 = 子式之和")。 $\mathcal{C}'$ 是 $n \times N$ 、 $\mathcal{U}$ 是 $N \times N$, 乘积 $\mathcal{C}' \mathcal{U}$ 是 $n \times N$ 。它的一个 n 阶子式可拆成

$$(\det \mathcal{C}' \mathcal{U})_{\nu_1, \dots, \nu_n}^{1, \dots, n} = \sum_{\rho_1, \dots, \rho_n} (\det \mathcal{C}')_{\rho_1, \dots, \rho_n}^{1, \dots, n} (\det \mathcal{U})_{\nu_1, \dots, \nu_n}^{\rho_1, \dots, \rho_n},$$

其中求和遍及从 $\{1, \dots, N\}$ 中无序选取 n 个指标 $\rho_1 < \dots < \rho_n$ 的全部 $\binom{N}{n}$ 种。这就是 Cauchy-Binet 公式——可以理解为"行列式版的乘法分配律"。(最熟悉的特例: 当 $N = n$ 时只剩一项, 退化成 $\det(AB) = \det A \det B$ 。)

- 7 每一项都被 $h(\mathcal{C}')$ 整除。和式每项里, 第一个因子 $(\det \mathcal{C}')_{\rho_1, \dots, \rho_n}^{1, \dots, n}$ 正是 $\mathcal{C}'$ 的某个 n 阶子式 Δ' , 按 $h(\mathcal{C}')$ 的定义它被 $h(\mathcal{C}')$ 整除; 第二个因子 $(\det \mathcal{U}) \cdots$ 是整数 (因为 q_{ir} 都是整数, $u_{\mu\nu}$ 是整数之积之和, 仍是整数)。于是每项 = $(h(\mathcal{C}')$ 的倍数) $\times$ (整数) 仍是 $h(\mathcal{C}')$ 的倍数, 整段和也是。

8 收尾。由第 5、7 步, $\Delta = \pm q \times (\text{被 } h(C') \text{ 整除的整数})$, 所以每个 Δ 都被 $qh(C')$ 整除。既然所有 Δ 都被 $qh(C')$ 整除, 它们的最大公因数 $h(C)$ 也被 $qh(C')$ 整除。■

推论 $h(C)$ 是 C 的算术不变量: 等价的型有相同的 h 值。

证 (补全)。

1 若 C, C' 等价, 变换行列式 $q = \pm 1$ 。引理 18.1 (取 $q = 1$ 那一版, 符号不影响整除) 给出: $h(C)$ 被 $h(C')$ 整除。

2 等价是对称的: 反过来用引理, $h(C')$ 被 $h(C)$ 整除。

3 两个非负整数互相整除 $\implies$ 相等。故 $h(C) = h(C')$ 。■

这条推论保证: 以后任意"换坐标" (么模变换) 都不会改变 $h(C)$, 所以 $h(C)$ 是型本身的内在量。

4 引理 18.2: 非退化 $\iff h(C) \neq 0$

引理 18.2 若 $C(\mathbf{x})$ 非退化, 则 $h(C) \neq 0$ 。

读法。这条把抽象的"非退化"翻成 $h(C)$ 这个数的"非零"。证明用逆否: 假设 $h(C) = 0$, 推出 C 退化。

1 $h(C) = 0$ 的含义。按定义, $h(C) = 0$ 即全部 n 阶子式 Δ 都为 0。一个 $n \times N$ 矩阵的所有 n 阶子式全 0, 等价于它的 n 行线性相关 (行的秩 $\leq n-1$)。

2 线性相关写成等式。存在不全为零的 $p_1, \dots, p_n$, 使第 i 行乘 p_i 再相加为零行:

$$\sum_{i=1}^n p_i c_{ijk} = 0 \quad (\text{对一切 } j, k).$$

又因为可以同除它们的最大公因数, 不妨设 $p_1, \dots, p_n$ 是最大公因数为 1 的整数 (这点后面要用来构造么模矩阵)。

3 翻成偏导的恒等式。由 (0.1): $\frac{1}{3} \frac{\partial C}{\partial x_i} = \sum_{j,k} c_{ijk} x_j x_k$ 。把它乘 p_i 对 i 求和:

$$\frac{1}{3} \sum_i p_i \frac{\partial C}{\partial x_i} = \sum_{j,k} \left(\sum_i p_i c_{ijk} \right) x_j x_k = \sum_{j,k} 0 \cdot x_j x_k = 0.$$

所以 $p_1 \frac{\partial C}{\partial x_1} + \dots + p_n \frac{\partial C}{\partial x_n} = 0$ 关于 $x_1, \dots, x_n$ 恒等成立 (不是某点, 是全体 $\mathbf{x}$)。

4 把 $(p_1, \dots, p_n)$ 补成一个么模矩阵。这是数论标准事实：若整数 $p_1, \dots, p_n$ 最大公因数为 1，则存在整数 $n \times n$ 矩阵 (p_{ir}) ，行列式 ± 1 ，且最后一列 $p_{in} = p_i$ 。（直观：以 $(p_1, \dots, p_n)$ 为一行/列，用辗转相除式的整数初等变换可补全成可逆整数矩阵。）作变量替换 $x_i = \sum_{r=1}^n p_{ir} y_r$ 。这是个么模变换，所以 C 变成的新型与 C 等价。

5 新型对 y_n 的偏导为零。链式法则：

$$\frac{\partial C}{\partial y_n} = \sum_i \frac{\partial C}{\partial x_i} \frac{\partial x_i}{\partial y_n} = \sum_i \frac{\partial C}{\partial x_i} p_{in} = \sum_i p_i \frac{\partial C}{\partial x_i} = 0$$
（用第 4 步 $p_{in} = p_i$ ，再用第 3 步那个恒等式）。所以 $\partial C / \partial y_n = 0$ 恒成立。

6 偏导恒为零 $\implies$ 不含该变量。一个多项式对 y_n 的偏导恒为 0，说明它根本不含 y_n （否则含 y_n 的项求导后不会全消）。于是新型只用到 $y_1, \dots, y_{n-1}$ ，即 C 等价于一个少一个变量的型——按定义 C 退化。

7 这与“ C 非退化”矛盾。故 $h(C) \neq 0$ 。■

译文随后说：“逆命题也成立（上述论证可逆），但我们不需要它。”——即“ $h(C) = 0 \Leftrightarrow C$ 退化”。我们只用到 $\Rightarrow$ 这一半：非退化 $\implies h(C)$ 是正整数。这正是封顶论证里“砝码大于 0”的那句保证。

5 引理 18.3：没有 $\mathcal{A}(p)$ 就能“挤出一个 p ”

引理 18.3 若 $n \geq 4$ 且 $C(\mathbf{x})$ 不具有性质 $\mathcal{A}(p)$ ，则 C 等价于一个如下类型的型

$$C'(x_1, x_2, x_3) + pC''(x_1, \dots, x_n). \quad (18.1)$$

这条引理为什么关键 回忆“ $\mathcal{A}(p)$ = 有 mod p 的非奇异解”。“没有 $\mathcal{A}(p)$ ”是很强的退化假设。引理 18.3 把这份假设兑换成**结构红利**：型能写成“3 个变量的纯三次部分 C' ”加上“ p 乘剩下一切 C'' ”。也就是说，除了三个变量，所有项都能被 p 整除。这块“ p ”正是后面下降法反复榨取的资源。

1 先有一个 mod p 的非零解（Chevalley）。三次型次数 3，变量数 $n \geq 4 > 3$ ，由 Chevalley 定理， $C(\mathbf{x}) \equiv 0 \pmod{p}$ 有 $\mathbf{x} \neq \mathbf{0}$ 的解。

2 这个解必是奇异的。“没有 $\mathcal{A}(p)$ ”恰说“没有 mod p 的非奇异解”，所以刚才那个解一定奇异： $\frac{\partial C}{\partial x_i} \equiv 0 \pmod{p}$ 对一切 i 。

3 把解搬到 $\mathbf{e}_1$ 。用一个么模变换把这非零解换成 $(1, 0, \dots, 0)$ （任何非零向量都能经整数么模变换变到标准基向量方向；若 $\gcd > 1$ 可先约掉，mod p 下不影响）。等价不改变“有无 $\mathcal{A}(p)$ ”和 $h(C)$ ，所以放心做。

4 写出 C 按 x_1 的展开。把 C 按 x_1 的幂次排：

$$C = \underbrace{apx_1^3}_{x_1^3\text{项}} + \underbrace{px_1^2(b_2x_2 + \cdots + b_nx_n)}_{x_1^2x_j\text{项}} + \underbrace{x_1B(x_2, \dots, x_n)}_{x_1\cdot\text{二次}} + \underbrace{C_{n-1}(x_2, \dots, x_n)}_{\text{不含 } x_1},$$

其中 B 是二次型、 C_{n-1} 是三次型。为什么前两块都带 p ？

- x_1^3 的系数 $= C(e_1)$ 。由第 1 步 $C(e_1) \equiv 0 \pmod{p}$ ，故该系数被 p 整除，写成 ap 。
- $x_1^2x_j$ ($j \geq 2$) 的系数正是 $\frac{\partial C}{\partial x_j}$ 在 e_1 处的值（对 x_j 求一次导、对 x_1 留二次）。
由第 2 步它 $\equiv 0 \pmod{p}$ ，故被 p 整除，写成 pb_j 。

这就解释了译文那句“诸 $x_1^2x_j$ 的系数都能被 p 整除，因为它们正是 $\partial C/\partial x_2, \dots, \partial C/\partial x_n$ 在该解处的值”。

5 反证：若 B 有系数不被 p 整除，会造出非奇异解。此时 $B(x_2, \dots, x_n)$ 不恒 $\equiv 0 \pmod{p}$ 。取一组很简单的值（形如 $1, 0, \dots, 0$ 命中某个 $b_{jj}x_j^2$ 项，或形如 $1, 1, 0, \dots, 0$ 命中某个交叉项 x_jx_k ）使 $B \not\equiv 0 \pmod{p}$ 。固定这些 $x_2, \dots, x_n$ 后，关于 x_1 的同余式

$$x_1 B(x_2, \dots, x_n) + C_{n-1}(x_2, \dots, x_n) \equiv 0 \pmod{p}$$

是关于 x_1 的一次方程，系数 $B \not\equiv 0 \pmod{p}$ 在 $\text{mod } p$ 下可逆，必有解 x_1 。对这组 $\mathbf{x}$ ：前两块 $apx_1^3, px_1^2(\cdots)$ 自动 $\equiv 0 \pmod{p}$ ，后两块刚被我们配成 0，所以 $C \equiv 0 \pmod{p}$ 。而

$$\frac{\partial C}{\partial x_1} = 3apx_1^2 + 2px_1(\cdots) + B(x_2, \dots, x_n) \equiv B \not\equiv 0 \pmod{p}.$$

这就给出一个 $\text{mod } p$ 的非奇异解 $\implies C$ 有 $\mathcal{A}(p)$ ，与假设矛盾。

6 故 B 全部系数被 p 整除。于是 $x_1B = px_1(B/p)$ 也带 p 。把带 p 的三块（ $apx_1^3, px_1^2(\cdots), x_1B$ ）合并：

$$C \text{ 等价于 } px_1Q_n(x_1, \dots, x_n) + C_{n-1}(x_2, \dots, x_n),$$

其中 Q_n 是二次型。关键观察：除了“不含 x_1 的 C_{n-1} ”，所有含 x_1 的项都带因子 p 。

7 对 C_{n-1} 递归（ $n \geq 5$ 时）。令 $x_1 = 0$ 看 $C_{n-1}(x_2, \dots, x_n)$ ：它是 $n-1 \geq 4$ 个变量的三次型。它也不能有 $\mathcal{A}(p)$ ——否则把它的非奇异解（ $x_1 = 0$ ）塞回 C 也会给 C 一个非奇异解（因为含 x_1 的项都带 p ，对 $x_1 = 0$ 不影响 $\text{mod } p$ 的值与梯度），矛盾。于是对 C_{n-1} 重复第 1-7 步：

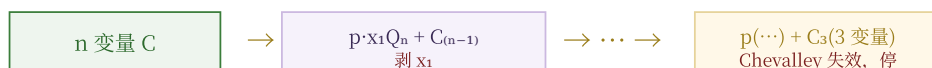
$$C_{n-1} \text{ 等价于 } px_2Q_{n-1}(x_2, \dots, x_n) + C_{n-2}(x_3, \dots, x_n).$$

8 下降到 3 个变量为止。每递归一次“剥掉一个变量、挤出一个 px_iQ ”，变量数减 1。一直降到 $C_3(x_{n-2}, x_{n-1}, x_n)$ （3 个变量）时，Chevalley 定理要求“变量数 $>$ 次数”即 > 3 ，此时 $3 \not> 3$ ，定理失效，递归停止。综合所有层：

$$C \text{ 等价于 } p(x_1Q_n + \cdots + x_{n-3}Q_4) + C_3(x_{n-2}, x_{n-1}, x_n).$$

- 9 调换变量书写顺序。把变量倒过来写 ($x_1 \leftrightarrow x_n$ 等, 这是么模变换), 那块 C_3 就落到前三个变量, 记为 $C'(x_1, x_2, x_3)$, 其余整体带 p 记为 pC'' 。得 (18.1)。■

每步剥掉一个变量、挤出一个因子 p , 直到剩 3 个变量



结论 (18.1): $C \equiv C'(x_1, x_2, x_3) + p \cdot C''(x_1, \dots, x_n)$

"除了 3 个变量的纯三次块, 其余全部带因子 p "

引理 18.3 的下降: 从 n 变量逐个剥到 3 变量, 沿途把含已剥变量的项都收进因子 p 。

6 引理 18.4: 把子型的 $\mathcal{A}(p^\lambda)$ 抬回 C 的 $\mathcal{A}(p^\ell)$

引理 18.4 若在引理 18.3 的结果中, 关于 $x_4, \dots, x_n$ 的型 $C''(0, 0, 0, x_4, \dots, x_n)$ 具有性质 $\mathcal{A}(p^\lambda)$, 则 C 对某个 $\ell \leq \lambda + 1$ 具有性质 $\mathcal{A}(p^\ell)$ 。

这条在做什么——逻辑方向 引理 18.3 把 C 写成 $C'(x_1, x_2, x_3) + pC''$ 。引理 18.4 说: 那个更小的子型 $C''(0, 0, 0, x_4, \dots, x_n)$ (把前三变量置 0) 若已经"足够好" (有 $\mathcal{A}(p^\lambda)$), 就能把这份好处抬回到原型 C (得 $\mathcal{A}(p^\ell)$, ℓ 只大一点点)。它将被反复用作"递归的归纳步"。注意我们真正要用的是它的逆否: " C 没有 $\mathcal{A}(p)$, $\mathcal{A}(p^2) \implies$ 子型 $C''(0, 0, 0, \dots)$ 也没有 $\mathcal{A}(p)$ "——一句把坏性质往下传的话。

- 1 先回顾性质 $\mathcal{A}(p^\lambda)$ 给出的"可解性" (引理 17.1 证明里的事实)。若型 C^* 有 $\mathcal{A}(p^\lambda)$, 则对每个 $\nu \geq 0$, 下面三条可同时满足:

$$C^*(\mathbf{x}) \equiv 0 \pmod{p^{2\lambda-1+\nu}}; \quad (18.2)$$

$$\frac{\partial C^*}{\partial x_i} \equiv 0 \pmod{p^{\lambda-1}} \quad (\forall i); \quad (18.3)$$

$$\frac{\partial C^*}{\partial x_j} \not\equiv 0 \pmod{p^\lambda} \quad (\exists j). \quad (18.4)$$

即简写 $p^{\lambda-1} \parallel (\partial C^*/\partial x_1, \dots, \partial C^*/\partial x_n)$ 。 $\nu = 0$ 就是 $\mathcal{A}(p^\lambda)$ 的原定义; $\nu > 0$ 是"把解亨泽尔抬升后还能保持非奇异、并让 C^* 落得更深 $p^{2\lambda-1+\nu}$ "——这是第 17 章证过的。本章只取 $\nu = 1$ 。

- 2 把假设落实到子型, 取 $\nu = 1$ 。 " $C''(0, 0, 0, x_4, \dots, x_n)$ 有 $\mathcal{A}(p^\lambda)$ ": 取 $\nu = 1$, 存在整数 $x_4, \dots, x_n$ 使

$$C''(0,0,0,x_4,\dots,x_n) \equiv 0 \pmod{p^{2\lambda}}, \quad p^{\lambda-1} \parallel \left(\frac{\partial C''}{\partial x_4}, \dots, \frac{\partial C''}{\partial x_n} \right).$$

($2\lambda - 1 + \nu = 2\lambda$ 。这里梯度只对子型的变量 $x_4, \dots, x_n$ 求。请记住指数是 $\lambda - 1$ ——下一步会因 C 多带一个 p 而升到 λ 。)

- 3 把它抬到 C : 用 $C(0,0,0,\cdot) = p C''(0,0,0,\cdot)$ 。在 (18.1) 里令 $x_1 = x_2 = x_3 = 0$: $C'(0,0,0) = 0$ (C' 是 x_1, x_2, x_3 的三次齐次型), 故

$$C(0,0,0,x_4,\dots,x_n) = p C''(0,0,0,x_4,\dots,x_n).$$

于是

- 把 (18.2) 乘 p : $C(0,0,0,x_4,\dots,x_n) \equiv 0 \pmod{p^{2\lambda+1}}$;
- 对 $i \geq 4$, $\frac{\partial C}{\partial x_i} = \frac{\partial C'}{\partial x_i} + p \frac{\partial C''}{\partial x_i} = 0 + p \frac{\partial C''}{\partial x_i}$ (C' 不含 x_i), 故梯度整体多一个 p :

$$p^\lambda \parallel \left(\frac{\partial C}{\partial x_4}, \dots, \frac{\partial C}{\partial x_n} \right).$$

(这正是要小心的一处: 子型梯度恰被 $p^{\lambda-1}$ 整除, 而 C 的对应梯度因为外层那个 p , 恰被 p^λ 整除, 相差一个 p 。译文此处把子型的指数也写作 p^λ , 按 (18.1) 的结构应是 $p^{\lambda-1}$; 我们按推导给出正确指数。)

- 4 用真实梯度估值定义 ℓ 。对这组取定的 $x_1 = x_2 = x_3 = 0, x_4, \dots, x_n$, 看全部 n 个偏导 $\partial C/\partial x_1, \dots, \partial C/\partial x_n$, 记它们公共的 p 幂次为 $\ell - 1$:

$$p^{\ell-1} \parallel \left(\frac{\partial C}{\partial x_1}, \dots, \frac{\partial C}{\partial x_n} \right).$$

我们不去管 $\partial C/\partial x_1, \partial C/\partial x_2, \partial C/\partial x_3$ 究竟多大; 但已知 $i \geq 4$ 那批里至少有一个恰被 p^λ 整除而不被 $p^{\lambda+1}$ 整除。所以这 n 个偏导的最小幂次 $\ell - 1 \leq \lambda$, 即

$$\ell \leq \lambda + 1.$$

- 5 验证这确实是 $\mathcal{A}(p^\ell)$ 。需要核对 (17.2)–(17.4):

- (17.2): 要 $C \equiv 0 \pmod{p^{2\ell-1}}$ 。由 $\ell \leq \lambda + 1$ 有 $2\ell - 1 \leq 2\lambda + 1$, 而第 3 步已给 $C \equiv 0 \pmod{p^{2\lambda+1}}$, 更高的模成立则更低的模也成立。✓
- (17.3): 要全部偏导 $\equiv 0 \pmod{p^{\ell-1}}$ 。这正是 ℓ 的定义 ($p^{\ell-1}$ 整除全部偏导)。✓
- (17.4): 要某个偏导 $\not\equiv 0 \pmod{p^\ell}$ 。这也来自 $\parallel$ 的定义 (p^ℓ 不能整除全部)。✓

故 C 具有 $\mathcal{A}(p^\ell)$, 且 $\ell \leq \lambda + 1$ 。■

我们将怎样用它 (逆否形式) 取 $\lambda = 1$: 若子型 $C''(0,0,0,\dots)$ 有 $\mathcal{A}(p)$, 则 C 有 $\mathcal{A}(p^\ell)$, $\ell \leq 2$, 即 C 有 $\mathcal{A}(p)$ 或 $\mathcal{A}(p^2)$ 。逆否: 若 C 既没有 $\mathcal{A}(p)$ 也没有 $\mathcal{A}(p^2)$, 则子型 $C''(0,0,0,\dots)$ 没有 $\mathcal{A}(p)$ 。同理取 $\lambda = 2$: C 没有 $\mathcal{A}(p), \mathcal{A}(p^2), \mathcal{A}(p^3) \implies$ 子型没有 $\mathcal{A}(p^2)$ 。"坏性质往下传"——这是引理 18.5 反复引用的台词。

7 引理 18.5: 缺 $\mathcal{A}(p), \mathcal{A}(p^2), \mathcal{A}(p^3) \Rightarrow$ 能"挤出 9 个变量的 p "

引理 18.5 若 $n \geq 10$ 且 C 不具有 $\mathcal{A}(p), \mathcal{A}(p^2), \mathcal{A}(p^3)$ 中任何一个, 则 C 等价于一个如下类型的型

$$C^*(x_1, x_2, \dots, x_9, px_{10}, \dots, px_n). \quad (18.5)$$

目标的含义与威力 (18.5) 是说: 存在一次"对角缩放"变换——前 9 个变量不动、后 $n-9$ 个变量都乘 p ——把 C 写成 C^* 。这个变换的行列式是 $1^9 \cdot p^{n-9} = p^{n-9}$ 。每施行一次, 就往 h 里塞进一个 p^{n-9} (由引理 18.1)。引理 18.3 一次只挤一个 p (行列式只含一个 p 量级), 引理 18.5 一次挤 $n-9$ 个 p ——这是后面"封顶"能跑起来的关键火力。代价是前提更强: 要同时缺 $\mathcal{A}(p), \mathcal{A}(p^2), \mathcal{A}(p^3)$ 三条。

证明思路总览: 把变量分成四组 $\{1, 2, 3\}, \{4, 5, 6\}, \{7, 8, 9\}, \{10, \dots, n\}$ 。对前三组各做一次"代入 $x_i = py_i +$ 引理 18.3 下降", 每组榨出 3 个能整体提 p 的变量; 中途若出现"某个二次型 Q 不被 p 整除", 就会反推出 C 竟有 $\mathcal{A}(p)$ 或 $\mathcal{A}(p^2)$, 与前提冲突。下面逐步。

① 第一组: 对 x_1, x_2, x_3 代入 p 。在 (18.1) (仍记作 C) 中令 $x_i = py_i$ ($i = 1, 2, 3$):

$$C(py_1, py_2, py_3, x_4, \dots, x_n) = C'(py_1, py_2, py_3) + pC''(py_1, py_2, py_3, x_4, \dots, x_n)$$

C' 是三次齐次, $C'(py_1, py_2, py_3) = p^3 C'(y_1, y_2, y_3)$ 。所以右端首项是 p^3 的倍数。

② 丢掉 p^3 的倍数, 得 (18.6)。在 $\text{mod } p^3$ 下, $p^3 C' \equiv 0$ 。再看

$pC''(py_1, py_2, py_3, x_4, \dots, x_n)$: 把 C'' 按"含几个 y_1, y_2, y_3 因子"分层——

- 不含 y_1, y_2, y_3 (即把它们当 0): 贡献 $pC''(0, 0, 0, x_4, \dots, x_n)$, 是 p^1 级;
- 恰含一个 y 、其余两因子在 $x_4, \dots, x_n$: 代入 $x_i = py_i$ 后多一个 p , 外层再乘 p , 共 p^2 级; 记成 $p^2 C_{1,2}(y_1, y_2, y_3 | x_4, \dots, x_n)$, 其中 $C_{1,2}$ 关于 y_1, y_2, y_3 一次、关于 $x_4, \dots, x_n$ 二次;
- 含两个或三个 y : 至少 p^2 (内) $\times p$ (外) = p^3 级, $\text{mod } p^3$ 归零。

所以

$$C(py_1, py_2, py_3, x_4, \dots, x_n) \equiv p^2 C_{1,2}(y_1, y_2, y_3 | x_4, \dots, x_n) + pC''(0, 0, 0, x_4, \dots, x_n) \quad (\text{m})$$

③ 第二组前的准备: 子型缺 $\mathcal{A}(p), \mathcal{A}(p^2)$ 。由引理 18.4 的逆否 (取 $\lambda = 1, 2$) 与前提" C 缺 $\mathcal{A}(p), \mathcal{A}(p^2), \mathcal{A}(p^3)$ ", 得: 子型 $C''(0, 0, 0, x_4, \dots, x_n)$ 缺 $\mathcal{A}(p)$ 和 $\mathcal{A}(p^2)$ 。它是 $n-3 \geq 7$ 个变量的三次型。

④ 第二组: 对子型用引理 18.3, 再代 p , 得 (18.7)。对 $C''(0, 0, 0, x_4, \dots, x_n)$ (缺 $\mathcal{A}(p)$) 用引理 18.3: 它等价于 " $D'(x_4, x_5, x_6) + pD''$ " 型。对 $i = 4, 5, 6$ 令 $x_i = py_i$, 仿第 1-2 步在 $\text{mod } p^2$ 下 (这里只需 $\text{mod } p^2$, 因为这块迟早还要再乘一个 p): $D'(py_4, py_5, py_6) = p^3 D' \equiv 0$, 含 y_4, y_5, y_6 的混合项 $\geq p^2$ 归零, 只剩

$$C''(0,0,0,py_4,py_5,py_6,x_7,\dots,x_n) \equiv pC^{(3)}(0,\dots,0,x_7,\dots,x_n) \pmod{p^2}, \quad (18.7)$$

其中 $C^{(3)}(0,\dots,0,x_7,\dots,x_n)$ 是新一轮的"置零子型" ($D''(0,0,0,x_7,\dots,x_n)$)。再由引理 18.4 逆否 (对子型 $C''(0,0,0,\dots)$ 缺 $\mathcal{A}(p), \mathcal{A}(p^2)$): $C^{(3)}(0,\dots,0,x_7,\dots,x_n)$ 缺 $\mathcal{A}(p)$ 。它有 $n-6 \geq 4$ 个变量。

5 把第二组代回 (18.6), 得 (18.8)。在 (18.6) 中对 $i = 4, 5, 6$ 也令 $x_i = py_i$:

- $p^2C_{1,2}(y_1, y_2, y_3 | py_4, py_5, py_6, x_7, \dots, x_n)$: $C_{1,2}$ 对 $x_4, \dots, x_n$ 是二次, 凡用到 x_4, x_5, x_6 的项代入后多带 p , 在已经有 p^2 前因子、又 $\text{mod } p^3$ 的情况下归零; 只剩"完全不碰 x_4, x_5, x_6 "的部分 $C_{1,2}(y_1, y_2, y_3 | 0, 0, 0, x_7, \dots, x_n)$ 。因 $C_{1,2}$ 关于 y_1, y_2, y_3 一次, 写成 $y_1Q_1 + y_2Q_2 + y_3Q_3$, 其中 Q_1, Q_2, Q_3 是 $x_7, \dots, x_n$ 的二次型。
- $pC''(0,0,0,py_4,py_5,py_6,x_7,\dots,x_n)$: 用 (18.7) 乘 p :
 $\equiv p \cdot pC^{(3)} = p^2C^{(3)}(0,\dots,0,x_7,\dots,x_n) \pmod{p^3}$ 。

合并:

$$C(py_1, \dots, py_6, x_7, \dots, x_n) \equiv p^2(y_1Q_1 + y_2Q_2 + y_3Q_3) + p^2C^{(3)}(0, \dots, 0, x_7, \dots, x_n) \pmod{p^3} \quad (18.8)$$

注意 y_4, y_5, y_6 已经不出现在右端 (它们的项全被 $\text{mod } p^3$ 吃掉了) ——这正是"挤出 $\{4, 5, 6\}$ 这一组 p "的体现。

6 反证: 必须 $Q_1 \equiv Q_2 \equiv Q_3 \equiv 0 \pmod{p}$ 。假设某个 (不妨 Q_1) 不恒 $\equiv 0 \pmod{p}$ 。则存在 $x_7, \dots, x_n$ 使 $Q_1 \not\equiv 0 \pmod{p}$ 。固定它们后, 关于 y_1, y_2, y_3 的一次同余式

$$y_1Q_1 + y_2Q_2 + y_3Q_3 + C^{(3)}(0, \dots, 0, x_7, \dots, x_n) \equiv 0 \pmod{p}$$

因 Q_1 可逆而可解。这给出

$$C(py_1, \dots, py_6, x_7, \dots, x_n) \equiv p^2 \cdot 0 \equiv 0 \pmod{p^3} \quad (y_4, y_5, y_6 \text{ 任取}),$$

同时

$$\frac{\partial C}{\partial y_1}(py_1, \dots, py_6, x_7, \dots, x_n) \equiv p^2Q_1 \not\equiv 0 \pmod{p^3}.$$

现在改回真实变量: 令 $x_i = py_i$ ($i = 1, \dots, 6$), 由链式法则 $\partial/\partial x_1 = p^{-1}\partial/\partial y_1$ (因 $x_1 = py_1$, 对 y_1 求导比对 x_1 求导多一个 p)。于是

$$\frac{\partial C}{\partial x_1} = p^{-1} \frac{\partial C}{\partial y_1} \equiv p^{-1} \cdot p^2Q_1 = pQ_1 \not\equiv 0 \pmod{p^2}.$$

结论: 我们找到一组 $x_1, \dots, x_n$ 使 $C \equiv 0 \pmod{p^3}$ 且 $\partial C/\partial x_1 \not\equiv 0 \pmod{p^2}$ 。设这组解的全梯度恰被 $p^{\ell-1}$ 整除; 由于 $\partial C/\partial x_1$ 不被 p^2 整除 (幂次 ≤ 1), 最小幂次 $\ell-1 \leq 1$, 故 $\ell \leq 2$; 又 $C \equiv 0 \pmod{p^3} \geq p^{2\ell-1}$ 。于是 C 具有 $\mathcal{A}(p^\ell)$ 且 $\ell \leq 2$, 即 C 有 $\mathcal{A}(p)$ 或 $\mathcal{A}(p^2)$ ——与前提冲突。所以 Q_1, Q_2, Q_3 全 $\equiv 0 \pmod{p}$ 。

7 (18.8) 简化。既然 $Q_i \equiv 0 \pmod{p}$, 那块 $p^2(y_1Q_1 + y_2Q_2 + y_3Q_3)$ 是 p^3 的倍数, $\text{mod } p^3$ 归零:

$$C(py_1, \dots, py_6, x_7, \dots, x_n) \equiv p^2 C^{(3)}(0, \dots, 0, x_7, \dots, x_n) \pmod{p^3}.$$

- 8 第三组：对 $C^{(3)}$ 再用一次引理 18.3。 $C^{(3)}(0, \dots, 0, x_7, \dots, x_n)$ (缺 $\mathcal{A}(p)$, ≥ 4 变量) 等价于 " $E'(x_7, x_8, x_9) + pE''$ " 型, 故

$$C^{(3)}(0, \dots, 0, x_7, \dots, x_n) \equiv C^{(4)}(x_7, x_8, x_9) \pmod{p}$$

($\text{mod } p$ 下带 p 的 E'' 消失, 只剩三变量纯三次块 $C^{(4)}$)。

- 9 第三组代 p , 合成 C^* 。对 $i = 7, 8, 9$ 也令 $x_i = py_i$ 。在上一步基础上:

$$C^{(3)}(0, \dots, 0, py_7, py_8, py_9, x_{10}, \dots, x_n) \equiv C^{(4)}(py_7, py_8, py_9) = p^3 C^{(4)}(y_7, y_8, y_9) \equiv 0 \pmod{p}$$

。于是 $p^2 \times (\equiv 0 \text{ mod } p) \equiv 0 \pmod{p^3}$, 即

$$C(py_1, \dots, py_9, x_{10}, \dots, x_n) \equiv 0 \pmod{p^3},$$

而且这关于 $y_1, \dots, y_9, x_{10}, \dots, x_n$ 恒成立 (不是某个解, 是全体取值都成立)。

- 10 提出 p^3 并用齐次性收尾。"左端这个型恒被 p^3 整除"意味着它每个系数都被 p^3 整除, 故可写

$$C(py_1, \dots, py_9, x_{10}, \dots, x_n) = p^3 C^*(y_1, \dots, y_9, x_{10}, \dots, x_n),$$

C^* 是整系数三次型。最后用齐次性导出干净的恒等式: 在上式里令 $y_i = x_i$ ($i \leq 9$)、 $x_j = px_j$ ($j \geq 10$), 左端变成 $C(px_1, \dots, px_9, px_{10}, \dots, px_n) = C(p\mathbf{x}) = p^3 C(\mathbf{x})$ (齐次!), 右端变成 $p^3 C^*(x_1, \dots, x_9, px_{10}, \dots, px_n)$ 。两边消去 p^3 :

$$C(x_1, \dots, x_n) = C^*(x_1, \dots, x_9, px_{10}, \dots, px_n).$$

这正是 (18.5)。■

把变量分四组, 对前三组各做一次"代 $p + 18.3$ 下降"

$\{1, 2, 3\}$ 代 $p \rightarrow$ 剥	$\{4, 5, 6\}$ 代 $p \rightarrow$ 剥	$\{7, 8, 9\}$ 代 $p \rightarrow$ 剥	$\{10, \dots, n\}$ 整体带 p
--------------------------------------	--------------------------------------	--------------------------------------	-------------------------------

途中若某 $Q_i \not\equiv 0 \pmod{p} \rightarrow$ 反推出 C 有 $\mathcal{A}(p)$ 或 $\mathcal{A}(p^2)$, 与前提矛盾

故 $Q_i \equiv 0$, 9 个变量全部"可提 p "

恒等式 (18.5): $C(\mathbf{x}) = C^*(x_1, \dots, x_9, p \cdot x_{10}, \dots, p \cdot x_n)$, 变换行列式 = p^{n-9}

引理 18.5: 缺三条性质 $\Rightarrow$ 前 9 个变量都能"代 p ", 得到行列式为 p^{n-9} 的缩放变换。这是火力最猛的一步。

8 引理 18.6: 把 C^* 的 $\mathcal{A}(p^\lambda)$ 抬回 C

引理 18.6 设 $n \geq 10$ 。若在引理 18.5 的结果中 C^* 具有 $\mathcal{A}(p^\lambda)$, 则 C 对某个 $\ell \leq \lambda + 3$ 具有 $\mathcal{A}(p^\ell)$ 。

读法。这是引理 18.4 的"加强版": 那里抬升使 $\ell \leq \lambda + 1$, 这里因为缩放变换吃进 3 个量级的 p (前 9 个变量、又涉及 p^3 的恒等式), 代价是 $\ell \leq \lambda + 3$ 。我们真正要用的还是它的逆否: " C 缺 $\mathcal{A}(p), \dots, \mathcal{A}(p^{3m}) \implies C^*$ 缺 $\mathcal{A}(p), \dots, \mathcal{A}(p^{3m-3})$ ".

- ① 用 $\mathcal{A}(p^\lambda)$ 的可解性, 取 $\nu = 3$ 。由引理 18.4 第 1 步的事实 (对 C^* , 取 $\nu = 3$, $2\lambda - 1 + 3 = 2\lambda + 2$): 存在 $y_1, \dots, y_n$ 使

$$C^*(y_1, \dots, y_n) \equiv 0 \pmod{p^{2\lambda+2}}, \quad p^{\lambda-1} \parallel \left(\frac{\partial C^*}{\partial y_1}, \dots, \frac{\partial C^*}{\partial y_n} \right).$$

(为什么挑 $\nu = 3$? 因为下一步要乘 p^3 , 得让 C^* 先落到足够深 $p^{2\lambda+2}$, 乘完正好够 $\mathcal{A}(p^\ell)$ 用。)

- ② 用恒等式 $C(py_1, \dots, py_9, y_{10}, \dots, y_n) = p^3 C^*(y_1, \dots, y_n)$ 。这正是引理 18.5 末尾恒等式的"对偶读法" (把 C^* 的前 9 个变量看成 C 缩放后的变量)。于是

$$C(py_1, \dots, py_9, y_{10}, \dots, y_n) = p^3 C^* \equiv p^3 \cdot 0 = 0 \pmod{p^{2\lambda+5}}.$$

- ③ 梯度的幂次。记复合函数 $F(\mathbf{y}) = C(py_1, \dots, py_9, y_{10}, \dots, y_n) = p^3 C^*(\mathbf{y})$ 。两边对 y_i 求导: $\frac{\partial F}{\partial y_i} = p^3 \frac{\partial C^*}{\partial y_i}$, 故 $p^{\lambda+2} \parallel (\partial F / \partial y_1, \dots, \partial F / \partial y_n)$ (在 $\lambda - 1$ 上加 3)。再把 F 的导数翻成 C 在该点的真实偏导:

- 对 $i \leq 9$ (因 $x_i = py_i$): $\frac{\partial F}{\partial y_i} = p \cdot \frac{\partial C}{\partial x_i}$, 所以 $\frac{\partial C}{\partial x_i} = p^{-1} \frac{\partial F}{\partial y_i}$, 幂次比 F 少 1;
- 对 $i \geq 10$ (因 $x_i = y_i$): $\frac{\partial C}{\partial x_i} = \frac{\partial F}{\partial y_i}$, 幂次同 F 。

设取得 $\parallel$ 的下标是 j (即 $\partial C^* / \partial y_j$ 恰被 $p^{\lambda-1}$ 整除):

- 若 $j \leq 9$: $\partial C / \partial x_j$ 幂次 = $(\lambda - 1) + 3 - 1 = \lambda + 1$;
- 若 $j \geq 10$: $\partial C / \partial x_j$ 幂次 = $(\lambda - 1) + 3 = \lambda + 2$ 。

两种情形都 $\leq \lambda + 2 < \lambda + 3$, 所以 $\partial C / \partial x_1, \dots, \partial C / \partial x_n$ 中至少有一个不被 $p^{\lambda+3}$ 整除。

- ④ 定 ℓ 并验证。对取定的点 ($x_i = py_i, i \leq 9$; $x_i = y_i, i \geq 10$), 设全梯度恰被 $p^{\ell-1}$ 整除。由第 3 步, 最小幂次 $\ell - 1 \leq \lambda + 2$, 即 $\ell \leq \lambda + 3$ 。又第 2 步给 $C \equiv 0 \pmod{p^{2\lambda+5}}$, 而 $2\ell - 1 \leq 2(\lambda + 3) - 1 = 2\lambda + 5$, 故 $C \equiv 0 \pmod{p^{2\ell-1}}$ 。三条 (17.2)-(17.4) 全满足 (同引理 18.4 第 5 步), 故 C 有 $\mathcal{A}(p^\ell)$, $\ell \leq \lambda + 3$ 。■

9 引理 18.7: 用 $h(C)$ 封顶——核心收官

引理 18.7 任何变量数 ≥ 10 的非退化整系数三次型, 对每个素数 p 都具有性质 $\mathcal{A}(p^\ell)$ (ℓ 依赖 p), 且 ℓ 有一个仅依赖该三次型的上界。

论证的精神 反证: 假设 C "病得很重"——缺掉 $\mathcal{A}(p), \mathcal{A}(p^2), \dots, \mathcal{A}(p^{3m})$ 整整 $3m$ 条。我们要证 m 不能太大。每"重病一档", 引理 18.5+18.6 就允许我们施一次行列式 p^{n-9} 的变换、并把"重病档数"降 3; 迭代 m 次后总变换行列式是 $p^{(n-9)m}$ 。引理 18.1 说 $h(C)$ 必被 $p^{(n-9)m}$ 整除; 引理 18.2 说 $h(C)$ 是个固定正整数。一个固定正整数能被多高的 p 幂整除? 有限! 于是 m 被 $\log h(C)$ 卡死。

1 设重病假设。设 C (整系数) 缺 $\mathcal{A}(p), \mathcal{A}(p^2), \dots, \mathcal{A}(p^{3m})$, m 是正整数。

2 一次迭代: 18.5 给变换、18.6 降档。因为缺 $\mathcal{A}(p), \mathcal{A}(p^2), \mathcal{A}(p^3)$ (重病假设的前三条), 由引理 18.5, C 等价于 (18.5) 型, 即存在整系数、行列式 p^{n-9} 的线性变换把 C 变成整系数型 $C^{(1)}$ (就是 C^*)。再由引理 18.6 的逆否 (C 缺 $\mathcal{A}(p), \dots, \mathcal{A}(p^{3m}) \implies C^{(1)}$ 缺 $\mathcal{A}(p), \dots, \mathcal{A}(p^{3m-3})$): 新型 $C^{(1)}$ 的"重病档数"恰好降了 3。

3 迭代 m 次。对 $C^{(1)}$ (缺 $3m-3$ 条, 仍 ≥ 3 条只要 $m \geq 1$... 逐次进行) 重复第 2 步, 每次行列式再乘 p^{n-9} 、档数再降 3。做 m 次后, 得到一个整系数、行列式 $p^{(n-9)m}$ 的总变换, 把 C 变成整系数型 $C^{(m)}$ 。(这些变换都可逆 (行列式 $\neq 0$), 所以保持非退化。)

4 用引理 18.1 封顶。这个总变换行列式 $q = p^{(n-9)m}$, 且 $C(\mathbf{x}) = C^{(m)}(\mathbf{x}')$ 。由引理 18.1, $h(C)$ 被 $q h(C^{(m)}) = p^{(n-9)m} h(C^{(m)})$ 整除。 $C^{(m)}$ 非退化, 故 $h(C^{(m)}) \geq 1$ (引理 18.2)。所以

$$p^{(n-9)m} \mid h(C), \quad \text{即} \quad p^{(n-9)m} \leq h(C).$$

5 取对数得 m 的上界。两边取对数 ($\log$ 单调增): $(n-9)m \log p \leq \log h(C)$, 故

$$(n-9)m \leq \frac{\log h(C)}{\log p} \leq \frac{\log h(C)}{\log 2},$$

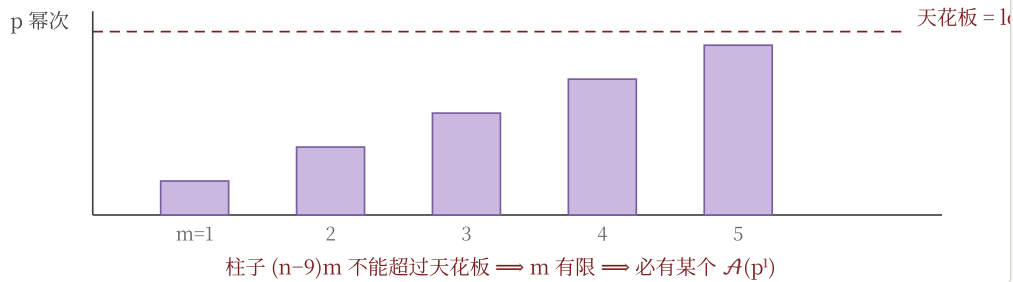
最后一步用 $p \geq 2 \Rightarrow \log p \geq \log 2$ 。这给出 m 的上界

$$m \leq \frac{\log h(C)}{(n-9) \log 2},$$

它只依赖 C (通过 $h(C)$ 与 n), 与 p 无关。

6 结论。"重病档数" $3m$ 不能无限大。所以 C 不可能缺掉所有 $\mathcal{A}(p), \mathcal{A}(p^2), \dots$ ——必在某个 $\ell \leq 3m_{\max} + 3$ 处具有 $\mathcal{A}(p^\ell)$ 。且这个 ℓ 的上界仅由 $h(C)$ 与 n 决定, 与 p 无关。引理 18.7 证毕。■

每迭代一次：行列式乘 p^{n-9} ， $h(C)$ 必含更高 p 幂——但 $h(C)$ 是固定正整数



封顶论证：每次迭代把 p^{n-9} 塞进 $h(C)$ ，而 $h(C)$ 这个固定正整数限制了能塞多少次，于是 m （从而坏档数）被卡死。

10 定理 18.1：拼回圆法主线

译文：鉴于定理 17.1 及第 17 章随后注记，引理 18.7 完成了下述结果的证明。

定理 18.1 若 $C(x_1, \dots, x_n)$ 是任意整系数三次型，且 $n \geq 17$ ，则方程 $C(x_1, \dots, x_n) = 0$ 有不全为零的整数解。

- 1 把局部条件凑齐。** 若 C 非退化且 $n \geq 10$ ，引理 18.7 给出：对每个素数 p ， C 都有某 $A(p^\ell)$ 。由第 17 章（引理 17.1）， $A(p^\ell) \Rightarrow \chi(p) > 0$ 。"对每个 p 都 $\chi(p) > 0$ "再加上必要的收敛性估计 $\Rightarrow$ 奇异级数 $\mathfrak{S} = \prod_p \chi(p) > 0$ 。
- 2 引用定理 17.1。** 第 17 章证明： $n \geq 17$ 且 $\mathfrak{S} > 0 \Rightarrow C(\mathbf{x}) = 0$ 有非平凡整数解（圆法主项 $\sim P^{n-3}\mathfrak{S}J_0$ 压过误差）。
- 3 处理退化情形。** 若 C 退化，按定义它等价于变量更少的型，可降维处理（第 17 章注记），最终仍归到非退化情形或直接得平凡可解。综上， $n \geq 17$ 时 $C = 0$ 总有非零整数解。■

全章一句话总结 本章只做一件事：保证每个素数处的局部障碍消失（ $\chi(p) > 0$ ）。手法是一个漂亮的双层结构——

- (1) 下降：**只要某个 $A(p^\ell)$ 不成立，就能用 Chevalley 定理"挤出 p "（18.3 一次挤一个、18.5 一次挤九个）；
 - (2) 封顶：**把"挤出的 p "记账到不变量 $h(C)$ ，而 $h(C)$ 是有限正整数（18.1+18.2），于是下降不能永远进行。
- 两者一夹，迫使某个 $A(p^\ell)$ 必然成立，且 ℓ 的上界与 p 无关。这种"无限下降被一个整数不变量截停"的思想，是数论里极有生命力的范式（可上溯到 Fermat 的无穷递降法）。

脚注 1 (译文): 参见文献 [23], 例如。——指 Chevalley (–Warning) 定理的标准出处。该定理 (C. Chevalley, 1935; E. Warning 同年加强) 说: 有限域 $\mathbb{F}_p$ 上变量数大于次数的齐次 (或常数项为零的) 多项式必有非零零点。它是引理 18.3"先抓到一个 mod p 非零解"的全部依据。

返回 [全书目录](#)