

DAVENPORT · 圆法 · 高中详解版

方程 $c_1x_1^k + \cdots + c_sx_s^k = 0$ (再论)

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

本章要解决什么 · 读完能掌握什么

我们要研究的方程是

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

即给定一串固定的整数系数 $c_1, \dots, c_s$, 问: 能不能找到一组**都不为 0** 的整数 $x_1, \dots, x_s$ 使它成立? 第 8 章已经证明过一个版本 (定理 8.1), 但那里要求变量个数 s 比较多。本章的目标是把第 9 章 (讨论华林问题 $G(k)$ 时发明的"稀疏集合"技巧) 搬过来, 在**更少的变量个数 s** 下仍然证明方程有无穷多组非零整数解。

读完你将理解: (1) 圆法是怎样把"方程有没有解"翻译成一个积分的; (2) 什么是**奇异级数、奇异积分、主弧 / 次弧、局部密度 $\chi(p)$** ; (3) 为什么本章要在方程里硬塞进一个因子 L^k ; (4) 最终定理 10.1 那个看起来吓人的条件 $s \geq s_0 + 3(2k \log 3k + 1)$ 是怎么来的、每一项是什么意思。

提示: 本章是第 8、9 两章的"合成与升级", 原书写得极其浓缩, 很多步骤一句"如前所述"就带过。下面我们把这些被压缩掉的背景**全部展开**, 从零讲起。

第 0 节 先把圆法的"机器"装好 (高中预备知识)

原文从第一句就默认你已经熟悉第 8、9 章的整套圆法语言。对高中生来说这些词都是新的，所以我们先用一节把所有零件讲清楚，后面逐段讲解时就能直接引用。

0.1 复指数 $e(\alpha)$ ：把"加法"变成"转圈"

记号 $e(\alpha)$

定义 $e(\alpha) = e^{2\pi i \alpha}$ ，读作"e 的 $2\pi i \alpha$ 次方"。这里 i 是虚数单位 ($i^2 = -1$)， π 是圆周率。由欧拉公式

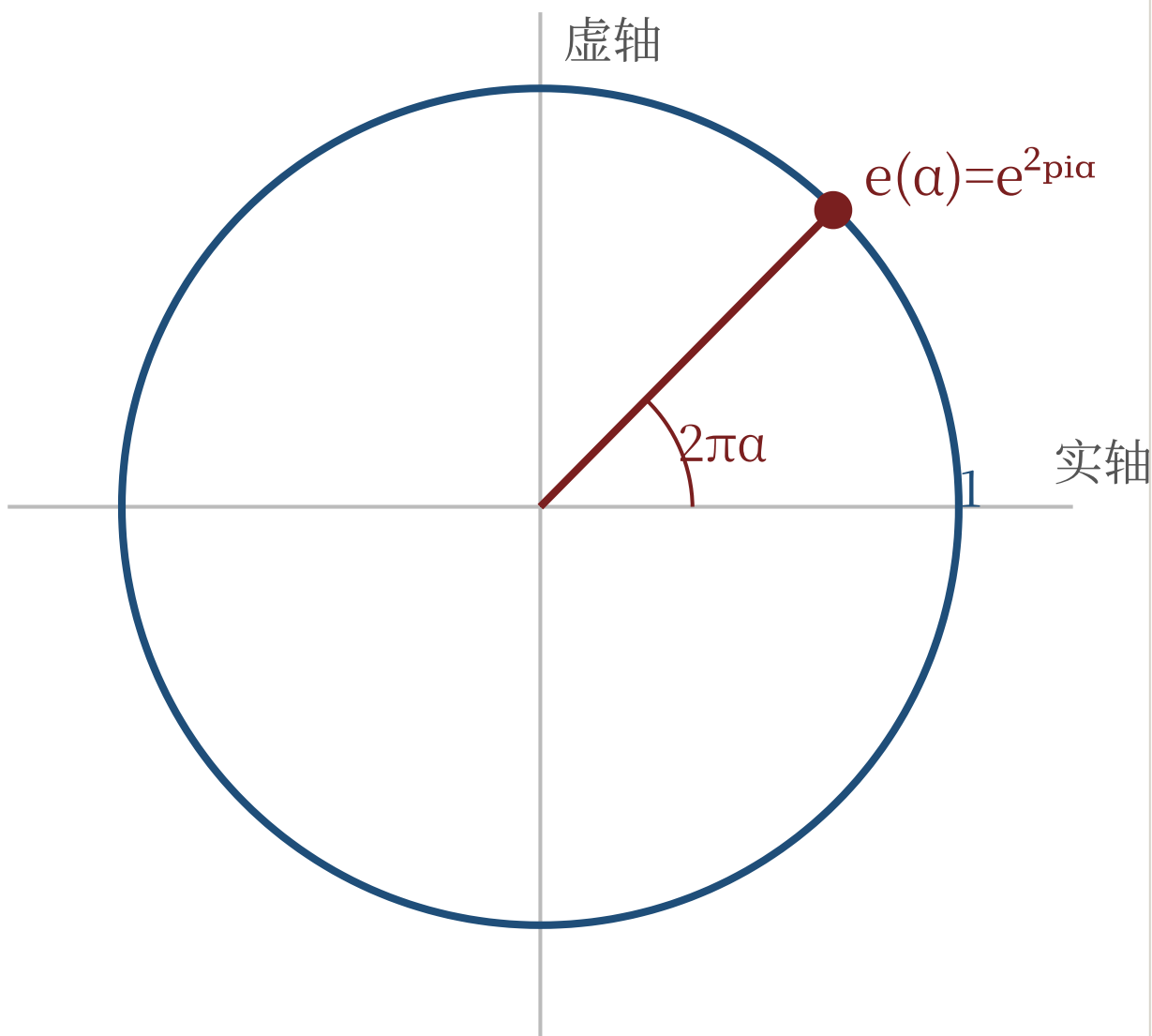
$$e^{2\pi i \alpha} = \cos(2\pi \alpha) + i \sin(2\pi \alpha).$$

所以 $e(\alpha)$ 是复平面单位圆上的一个点：把单位圆周长看成 1，从 $(1, 0)$ 出发逆时针走 α 圈所到达的点。

为什么要用它？因为它有两条神奇性质，正好是圆法的发动机：

1 把整数加法变成转角相加： $e(\alpha + \beta) = e(\alpha)e(\beta)$ 。这只是指数运算 $a^{m+n} = a^m a^n$ 的复数版。

2 整数处取值为 1：对任意整数 m ，
 $e(m) = e^{2\pi i m} = \cos(2\pi m) + i \sin(2\pi m) = 1$ （转整数圈又回到起点）。



$e(\alpha)$ 就是单位圆上转了 α 圈的那个点； α 每增加 1（转满一圈）就回到 1。

0.2 积分 $\int_0^1$ 与"正交关系"：探测器

记号 $\int_0^1 f(\alpha) d\alpha$

这是把函数 f 在区间 $[0, 1]$ 上的"平均累积量"。高中只学过有限求和；定积分可以先当成"把 $[0, 1]$ 切成无穷多个极细的小段、每段宽 $d\alpha$ 、把 $f(\alpha) d\alpha$ 全加起来"。我们只需要它的一条具体结论，不需要会算复杂积分。

正交关系（圆法的"探测器"） 对任意整数 m ,

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

证明. 分两种情况, 每一步都说清理由:

1 若 $m = 0$: 被积函数 $e(0 \cdot \alpha) = e(0) = 1$, 于是 $\int_0^1 1 d\alpha = 1$ (宽为 1、高为 1 的矩形面积)。

2 若 $m \neq 0$: 用复指数的原函数。因为 $\frac{d}{d\alpha} e^{2\pi i m \alpha} = 2\pi i m e^{2\pi i m \alpha}$, 所以 $\int_0^1 e^{2\pi i m \alpha} d\alpha = \frac{e^{2\pi i m \alpha}}{2\pi i m} \Big|_0^1 = \frac{e^{2\pi i m} - e^0}{2\pi i m}$ 。

3 由 0.1 的性质 2, $e^{2\pi i m} = e(m) = 1$ (因为 m 是整数), 又 $e^0 = 1$, 所以分子 $= 1 - 1 = 0$, 整个积分为 0。◆

这条结论为什么叫"探测器"? 因为它能判断一个整数是不是 0: 把 m 放进 $\int_0^1 e(m\alpha) d\alpha$, 得 1 就说明 $m = 0$, 得 0 就说明 $m \neq 0$ 。这正是我们求解 $c_1 x_1^k + \dots = 0$ 所需要的——我们要数的就是"使那个和恰好等于 0 的 $(x_1, \dots, x_s)$ 有多少组"。

0.3 生成函数 $T(\alpha)$ 与"解的个数 = 一个积分"

记号 $\sum$ (求和号) 与 $T_j(\alpha)$

$\sum_x$ 表示"对 x 取遍某个范围里的所有值, 把后面的式子全加起来"。对每个系数 c_j , 定义指数和 (生成函数)

$$T_j(\alpha) = \sum_x e(c_j x^k \alpha),$$

其中 x 跑遍一段事先规定好的整数范围（本章里是 $\setminus(0$

把所有 T_j 乘起来，再用上面的探测器积分，就能数解。我们一步步看为什么：

- 1 把乘积展开。根据 0.1 的性质 1（指数相乘 = 指数相加），

$$T_1(\alpha) \cdots T_s(\alpha) = \sum_{x_1} \cdots \sum_{x_s} e((c_1 x_1^k + \cdots + c_s x_s^k) \alpha).$$

也就是说，对每一组具体的 $(x_1, \dots, x_s)$ ，都贡献一项 $e(m\alpha)$ ，其中 $m = c_1 x_1^k + \cdots + c_s x_s^k$ 正是我们关心的那个和。

- 2 两边对 α 在 $[0, 1]$ 上积分，并交换"积分"与"求和"次序：

$$\int_0^1 T_1 \cdots T_s d\alpha = \sum_{x_1} \cdots \sum_{x_s} \int_0^1 e(m\alpha) d\alpha.$$

- 3 由正交关系 (0.2)，右边每一项的积分要么是 1（当 $m = 0$ ），要么是 0（当 $m \neq 0$ ）。于是求和只剩下那些真的满足 $m = 0$ 的组，每组贡献 1。

- 4 结论：

$$\int_0^1 T_1(\alpha) \cdots T_s(\alpha) d\alpha = (\text{满足 } c_1 x_1^k + \cdots + c_s x_s^k = 0 \text{ 的组数}).$$

要证"方程有解"，只要证明这个积分 > 0 即可。这就是圆法的出发点：把"解存在吗"变成"一个积分是不是正的"。

0.4 主弧与次弧：把 $[0, 1]$ 切成"信号区"和"噪声区"

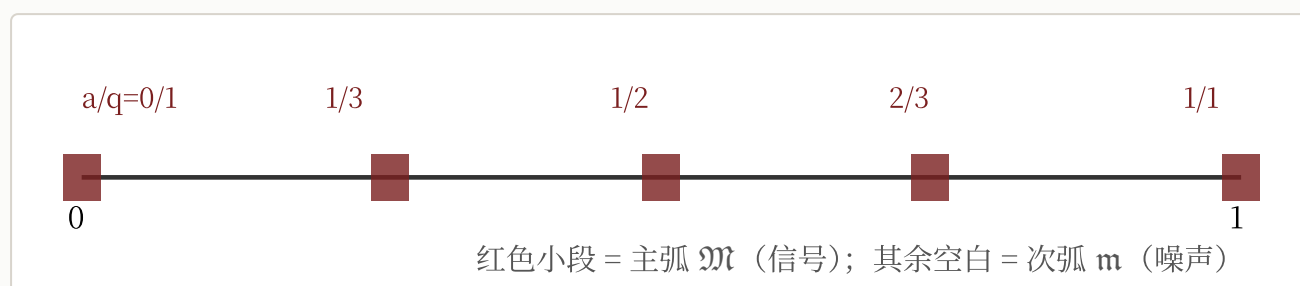
积分 $\int_0^1$ 没法直接算出来。圆法的策略是把积分区间 $[0, 1]$ 分成两类小区间：

主弧 $\mathfrak{M}$ (major arcs)

那些靠近"分母小的分数" a/q (q 较小) 的小区间。在这里指数和 $T_j(\alpha)$ 很大、行为规整，能算出一个漂亮的主项。读作"M 花体"，代表 major。

次弧 $\mathfrak{m}$ (minor arcs)

剩下的所有 α 。在这里 $T_j(\alpha)$ 因为剧烈震荡而互相抵消、整体很小，只贡献一个可控的误差（用 Weyl 不等式、van der Corput 引理等工具去压住它）。读作"m 花体"，代表 minor。



主弧是分母小的分数 a/q 周围的窄区间；其余是次弧。主弧给主项，次弧给误差。

主弧上的主项会自然地分裂成三个因子相乘：

$$\underbrace{P^{s-k}}_{\text{规模}} \times \underbrace{\mathfrak{S}}_{\text{奇异级数}} \times \underbrace{\mathfrak{J}}_{\text{奇异积分}}.$$

下面分别解释后两个，它们是本章的主角。

0.5 奇异级数 $\mathfrak{S}$ 、局部密度 $\chi(p)$ 与"模 p 可解"

奇异级数 $\mathfrak{S}$ 与欧拉乘积 $\prod_p \chi(p)$

$\mathfrak{S}$ （读作"S 花体"，singular series）是从主弧上所有分数 a/q 的贡献加起来得到的一个数。关键事实是它能写成对所有素数 p 的连乘积（叫欧拉乘积）：

$$\mathfrak{S} = \prod_p \chi(p) = \chi(2) \chi(3) \chi(5) \chi(7) \cdots$$

其中每个因子 $\chi(p)$ 只跟素数 p 有关。 $\prod$ 是连乘号，类似 $\sum$ 但把各项乘起来。

局部密度 $\chi(p)$

$\chi(p)$ （读作"chi of p"）度量方程在模 p 的世界里解有多"稠密"。精确地说，设 $N(p^t)$ 是同余方程

$$c_1 x_1^k + \cdots + c_s x_s^k \equiv 0 \pmod{p^t}$$

的解的组数（每个 x_j 在 0 到 $p^t - 1$ 里取），则

$$\chi(p) = \lim_{t \rightarrow \infty} \frac{N(p^t)}{p^{(s-1)t}}.$$

直观：随机取一组 x ，那个和落在 p^t 的某个固定余数上的"概率"约是 p^{-t} ； s 个变量、各 p^t 种取值共 p^{st} 组，落到余数 0 上的约有 $p^{st} \cdot p^{-t} = p^{(s-1)t}$ 组。 $\chi(p)$ 就是真实解数与这个"理想随机预测"的比值。

为什么 $\chi(p)$ 这么重要？因为有一条桥梁定理（圆法的灵魂）：

"局部 $\Rightarrow$ 整体"原则

如果对每一个素数 p 都有 $\chi(p) > 0$ （即方程模每个素数幂都有"非奇异"解，**局部可解**），并且奇异积分 $\mathfrak{J} > 0$ （实数范围里可解），那么当变量足够多时，原方程就真的有整数解。 $\chi(p) > 0$ 等价于那个同余方程模 p^t 一直有解且解数不退化为 0。

所以" $\mathfrak{S} > 0$ "这件事，本质上就是"方程在每个素数 p 处都过得了关"。本章前半部分全是在小心地保证 $\mathfrak{S} > 0$ ，而且要**均匀地**（下界与 M 无关）。

0.6 奇异积分 $\mathfrak{J}$ ：实数范围里的可解性

奇异积分 $\mathfrak{J}$

$\mathfrak{J}$ (singular integral) 是奇异级数在 " $p = \infty$ (实数)" 这一处的类比, 它度量方程在**实数**范围内、且各变量限制在给定盒子里时解有多稠密。 $\mathfrak{J} > 0$ 大致相当于: 在允许的实数取值范围内, 曲面 $c_1 x_1^k + \cdots = 0$ 真的穿过那个盒子。本章会看到, 正是 " c_j 当 k 为偶数时不全同号" 这个条件, 保证了 $\mathfrak{J} > 0$ 。

0.7 两个"省字"记号: $O(\cdot)$ 、 $\ll$ 与 $\gg$

大 O 、 $\ll$ 、 $\gg$

这些是估计大小时偷懒但严谨的写法:

$A = O(B)$ 或 $A \ll B$

读作" A 至多是 B 的常数倍", 意思是存在一个与变量无关的常数 $C > 0$, 使 $|A| \leq C B$ 始终成立。例如 $3n + 5 = O(n)$ 。

$A \gg B$

就是 $B \ll A$, 读作" A 至少是 B 的某个正常数倍", 即存在 $c > 0$ 使 $A \geq c B$ 。本章里 $\gg P^{s_0-k}$ 表示"这个量有一个正的、规模为 P^{s_0-k} 的下界"。

这些记号让我们能只谈"量级", 不被具体常数纠缠——因为证明解存在时, 只要主项的正下界超过误差的上界就够了, 常数多大无所谓。

0.8 同余 $\equiv$ 与 φ , 以及第 9 章的"稀疏集合"思想

同余号 $\equiv$ 与 $\bmod$

$a \equiv b \pmod{m}$ 读作" a 与 b 模 m 同余", 意思是 $m \mid (a - b)$, 即 $a - b$ 能被 m 整除 (除以 m 余数相同)。例如 $17 \equiv 2 \pmod{5}$ 。 $M \equiv 0 \pmod{L}$ 就

是" M 是 L 的整数倍"。

本章把第 9 章研究华林问题 $G(k)$ 时的核心技巧搬了过来，那技巧的精神是：不要让所有变量都自由地跑遍 k 次方，而是让一部分变量取值于一个"稀疏但够用"的集合（比如"能表成 $d_1 z_1^k + \cdots + d_\ell z_\ell^k$ 的那些数 u "）。稀疏集合里元素少，但只要个数下界够好，就能用更少的变量达到同样的覆盖效果——这正是本章能把变量个数 s 压低的原因。后面 u_1, u_2, v 就是这种"被打包好的稀疏变量"。

第 1 节 逐段精讲

下面严格按译文的段落顺序，每段先抄录原文（灰框），再做高中视角的展开，补全所有跳步。

1.1 开场：本章的目标与前提条件

原文译文

我们回到这个方程，并把上一节的方法加以改造，以便在比定理 8.1 更弱的限制条件下确立它的可解性。与第 8 章一样，我们假定 $c_1, \dots, c_s$ 是固定的整数，没有一个等于 0，并且当 k 为偶数时它们不全同号。

这段在说什么：宣布任务（用更少变量重证可解性）和两条前提。逐项拆开：

"回到这个方程"

指标题方程 $c_1 x_1^k + \cdots + c_s x_s^k = 0$ 。第 8 章已经处理过一次（结论是定理 8.1），所以叫"再论"。

"把上一节的方法加以改造"

"上一节"= 第 9 章，那里为华林问题 $G(k)$ 发明了稀疏集合 + 高效计数的技巧（见 0.8）。本章把它移植到带一般系数 c_j 的方程上。

"比定理 8.1 更弱的限制条件"

"限制条件"指的是要求变量个数 s 有多大。"更弱"= 要求更松 = 需要的变量更少。这就是本章相对第 8 章的进步：定理 10.1 的下界比定理 8.1 小。

两条前提，逐条解释为什么必须有：

1 " $c_1, \dots, c_s$ 是固定的整数，没有一个等于 0"：若某个 $c_j = 0$ ，那一项就消失了，等于减少一个变量，没意义；所以排除。"固定"是说它们是题目给定的常数，不随 P 变。我们求的解也要求每个 $x_j \neq 0$ （否则取全 0 是平凡解，不算数）。

2 "当 k 为偶数时它们不全同号"：这是保证实数可解（ $\mathfrak{J} > 0$ ）的关键。当 k 为偶数时， $x_j^k \geq 0$ 永远非负。若所有 c_j 同号（比如全正），则 $c_1 x_1^k + \dots + c_s x_s^k$ 是若干非负数之和，只有全部 $x_j = 0$ 才可能为 0，不存在非零解。所以必须让 c_j 有正有负，正负项才能互相抵消凑出 0。 k 为奇数时 x^k 可正可负，自动没问题，不需要这条。

体会一下 取 $k = 2$ （偶数）。方程 $x_1^2 + x_2^2 + x_3^2 = 0$ 的系数全是 +1（同号），非零整数解不存在（平方和为 0 必全 0）。但 $x_1^2 + x_2^2 - 2x_3^2 = 0$ 有解，例如 $x_1 = x_2 = x_3 = 1$ ： $1 + 1 - 2 = 0$ 。差别就在"不全同号"。

1.2 门槛 s_0 ：让"= 0 方程"的奇异级数为正的变量数

原文译文

设 $s_0 = s_0(k)$ 是一个整数，它具有如下性质：任何方程 $c_1 x_1^k + \dots + c_{s_0} x_{s_0}^k = 0$ 的奇异级数都是正的。由第 8 章的工作，我们可以

取

$$s_0 = k^2(2k - 1) + 1, \quad (10.1)$$

并且如那里所述，事实上可以取 $s_0 = k^2 + 1$ ，尽管我们并未证明这一点。

这段在说什么：定义一个"安全门槛" s_0 。只要变量个数达到 s_0 ，奇异级数 $\mathfrak{S}$ 就保证为正（回忆 0.5： $\mathfrak{S} > 0$ 等于"每个素数处都局部可解"）。

$$s_0 = s_0(k)$$

写成 $s_0(k)$ 是强调这个门槛依赖于次数 k ：次数越高，需要越多变量才能保证局部处处可解。它是一个只跟 k 有关、跟具体 c_j 和 M 都无关的整数。

为什么是 $s_0 = k^2(2k - 1) + 1$ ？这正是第 8 章的成果：第 8 章证明了"只要变量个数 $\geq k^2(2k - 1) + 1$ ，那个同余方程模任意素数幂都有非奇异解"，从而 $\chi(p) > 0$ 对所有 p 成立，于是 $\mathfrak{S} = \prod_p \chi(p) > 0$ 。本章直接把它当现成结论用，记成 (10.1)。我们算一下这个数的量级：展开 $k^2(2k - 1) + 1 = 2k^3 - k^2 + 1$ ，约为 $2k^3$ ，是 k 的三次量级——这正是后面要努力去"超越"的主要项。

"事实上可以取 $s_0 = k^2 + 1$ ，尽管我们并未证明"：更深的研究（戴维波特原书没在此证）表明门槛其实能降到约 k^2 （二次量级），比 $2k^3$ 小得多。作者诚实地说"我没证，但你知道有这回事"。本章的所有结论里， s_0 是一个"黑箱常数"，你换成 $k^2 + 1$ 也照样成立——本章的新意不在 s_0 本身，而在 s_0 之外只需再加很少几个变量。

1.3 换成" $= M$ "的方程：奇异级数 $\mathfrak{S}(M)$ 与尾巴自动 $\geq \frac{1}{2}$

原文译文

正如上一章那样，我们将需要关于相关方程

$$c_1 x_1^k + \cdots + c_{s_0} x_{s_0}^k = M \quad (10.2)$$

的奇异级数 $\mathfrak{S}(M)$ 的一些知识。它分解为 $\prod_p \chi(p)$ ，而我们知道

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

对某个依赖于 k 与 s 、但不依赖于 M 的 p_0 成立；只要 s_0 仅仅 $\geq 2k + 1$ ，这就成立。于是，为了确保 $\mathfrak{S}(M)$ 对某一类整数 M 有一个不依赖于 M 的正的下界，只需对每个 $p \leq p_0$ ， $\chi(p)$ 有这样一个下界即可。

这段在说什么：为什么要研究 $= M$ 而不是 $= 0$ ？因为本章的策略（见 1.6）是把方程拆成两块：前 s_0 个变量 $c_1 x_1^k + \cdots + c_{s_0} x_{s_0}^k$ 要去等于某个数 M ，剩下的变量负责凑出 $-M$ 。所以我们要研究"前 s_0 项之和 $= M$ "这个方程的奇异级数，记作 $\mathfrak{S}(M)$ ，它现在**依赖于右端的 M** 。

$\mathfrak{S}(M)$ 、 p_0

$\mathfrak{S}(M)$

方程 (10.2)（右端是 M ）的奇异级数；和 0.5 一样能写成欧拉乘积 $\prod_p \chi(p)$ ，只是现在每个 $\chi(p)$ 也跟 M 有关（因为同余方程右端变成了 M ）。

p_0

一个分界素数（cutoff）。把所有素数分成"小素数 $p \leq p_0$ "和"大素数 $p > p_0$ "两类。 p_0 只跟 k, s 有关，**不跟 M 有关**——这点很关键。

核心论断 1："大素数尾巴 $\prod_{p > p_0} \chi(p) \geq \frac{1}{2}$ 自动成立"。展开讲为什么：

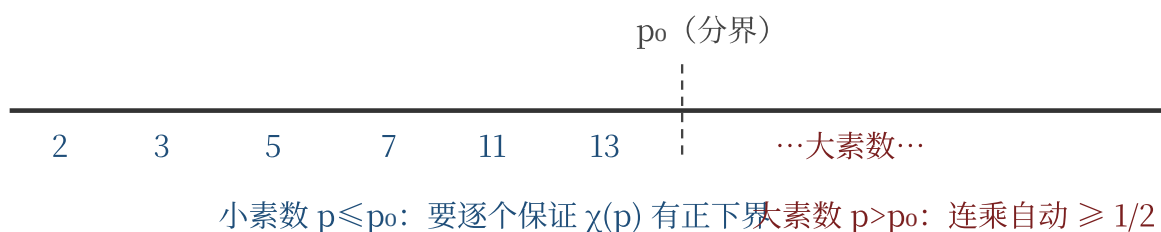
- 1 当 p 很大、不整除任何 c_j 时，每个 $\chi(p)$ 都非常接近 1，可以写成 $\chi(p) = 1 + (\text{小误差})$ 。第 5、6 章的估计给出

$|\chi(p) - 1| \leq C p^{-(s_0/k-1)}$ 这一类形式的界（指数随变量数 s_0 增大而变陡）。

2 只要 $s_0 \geq 2k + 1$ ，这个误差指数就够陡，使得级数 $\sum_{p>p_0} |\chi(p) - 1|$ 收敛（像 $\sum p^{-(1+\delta)}$ 那样的收敛级数）。

3 收敛意味着：把分界 p_0 取得足够大，就能让"所有大素数因子合起来偏离 1 的总量"任意小，特别地能保证连乘 $\prod_{p>p_0} \chi(p)$ 落在 $[\frac{1}{2}, \frac{3}{2}]$ 内，于是 $\geq \frac{1}{2}$ 。

4 由于误差界里根本没出现 M （ M 只影响小素数处的同余条件，不影响大素数的收敛速度），所以这个 p_0 和这条 $\geq \frac{1}{2}$ 都与 M 无关。



难点被压缩到有限个小素数 $p \leq p_0$ 上；无穷多个大素数的贡献自动可控。

核心论断 2（结论）：既然大素数那一半已经稳稳 $\geq \frac{1}{2}$ ，要让整个 $\mathfrak{S}(M) = \prod_{p \leq p_0} \chi(p) \cdot \prod_{p > p_0} \chi(p)$ 有一个与 M 无关的正下界，只剩下有限个小素数 $p \leq p_0$ 要操心——只要对每个 $p \leq p_0$ 把 $\chi(p)$ 也压出一个不依赖 M 的正下界，两半相乘就得到 $\mathfrak{S}(M)$ 的正下界。这是一个典型的"把无穷问题化为有限问题"的归约。

1.4 回忆第 8 章的局部处理：约化成"加性形式"

在第 8 章的工作中，我们（对每个 p ）施加了一个初步变换，化为加性形式，其依赖于整除系数 c_j 的 p 的幂次。此后我们得到一个形式 $d_1 y_1^k + \cdots + d_v y_v^k$ ，其系数不被 p 整除，并且只要同余式

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

有一个 $y_1, \dots, y_v$ 不全被 p 整除的解，这就足够了。我们证明了：当 $s \geq s_0$ 时这个条件被满足，其中 s_0 由 (10.1) 给出。

这段在说什么：复述第 8 章是怎么保证 $\chi(p) > 0$ 的，为下一段（引入 $\gamma_1(p)$ 和 L ）做铺垫。整段是一条“为什么模 p 处过关”的逻辑链。逐步看：

1 “初步变换，化为加性形式”：固定一个素数 p 。不同的 c_j 可能含有不同次数的因子 p （比如 c_1 含 p^0 ， c_2 含 $p^2 \cdots$ ）。把变量按“它的系数里含几个 p ”分组、并对变量做替换 $x_j \rightarrow p^? x_j$ ，可以把系数里的 p 因子整理干净。“加性形式”就是各项相加的纯形式 $d_1 y_1^k + \cdots$ （没有交叉项），这正是我们一直在处理的样子。

2 “系数 d_i 不被 p 整除”：变换之后留下的新系数 $d_1, \dots, d_v$ 都与 p 互素（ $p \nmid d_i$ ）。这是好事：系数不含 p ，同余方程才“干净”，可以套用标准的 Hensel 提升。

3 “只要同余式 $d_1 y_1^k + \cdots + d_v y_v^k \equiv 0 \pmod{p^\gamma}$ 有不全被 p 整除的解就足够”：这里 γ 是一个足够大的固定指数。“不全被 p 整除”= 解不是平凡的“全是 p 的倍数”（那种解可以约掉一层 p ，不算真正的本原解）。这种本原解（非奇异解）一旦在某个 p^γ 上出现，就能用 Hensel 引理逐级提升到任意高的 p^t ，从而 $N(p^t)$ 不退化、 $\chi(p) > 0$ 。

4 “当 $s \geq s_0$ 时这个条件被满足”：第 8 章正是证明了“变量够多（ $\geq s_0$ ）就一定能找到这样的本原解”。这也解释了 s_0 这个门槛的来源——它就是“保

证每个素数处都能找到本原解"所需的变量数。

d_i, y_i, v, γ

$d_1, \dots, d_v$

变换后、与 p 互素的新系数。

$y_1, \dots, y_v$

变换后的新变量。

v

变换后留下的变量个数（不超过 s ）。

γ

同余方程要解到的模的指数 (p^γ)；它是一个只跟 k （以及 p 整除 c_j 的幂次）有关的固定整数。

1.5 关键构造： $\gamma_1(p)$ 与模 L ，让 $\mathfrak{S}(M) > 0$ 均匀成立

原文译文

由此可知，对每个 p ，存在某个 $\gamma_1(p)$ （也依赖于系数 c_j ），使得当

$$M \equiv 0 \pmod{p^{\gamma_1(p)}}$$

时， $\chi(p)$ 有一个正的下界；因为这将保证最终必须求解的那个同余式与 $M = 0$ 时的同余式相同。令

$$L = \prod_{p \leq p_0} p^{\gamma_1(p)}.$$

那么，如果 $M \equiv 0 \pmod{L}$ ，就存在一个不依赖于特定 M 的、关于 $\chi(p)$ 的正下界，从而也存在关于 $\mathfrak{S}(M)$ 的正下界。

脚注：如果 p 不整除任何一个 c_j ，我们可以取 $\gamma_1(p) = \gamma$ 。

这段是本章前半的技术核心，请慢看。问题是：现在右端不是 0 而是 M ，同余方程变成 $d_1 y_1^k + \cdots \equiv M' \pmod{p^\gamma}$ (M' 是 M 经初步变换后对应的右端)。 M' 一般不是 0，它会破坏"和 $M = 0$ 时一样有本原解"的好局面。怎么办？答案：逼 M 含有足够多的 p 因子，让 M' 在模 p^γ 下变回 0。

$\gamma_1(p)$

读作"gamma-1 of p"。它是一个针对素数 p 选定的指数：只要 M 被 $p^{\gamma_1(p)}$ 整除，那么 M 经过 1.4 的初步变换后落进同余方程时恰好变成 0 (模 p^γ)，于是"要解的同余式"和 $M = 0$ 时的一模一样，从而沿用 1.4 的结论得到本原解、 $\chi(p) > 0$ 。 $\gamma_1(p)$ 比 γ 略大一点，多出来的部分用于吸收"初步变换里乘进去的 p 的幂次"，所以它也依赖于 c_j (因为变换依赖 c_j 含 p 的幂次)。

逐步说清这条推理：

- 1 固定素数 p 。1.4 告诉我们：右端为 0 时同余方程有本原解， $\chi(p) > 0$ 。
- 2 右端换成 M 后，记初步变换把右端 M 映成 M' 。 M' 含的 p 因子个数 = M 含的 p 因子个数减去 (变换里抹掉的若干个 p)。
- 3 因此只要 M 含有足够多的 p 因子 (即 $p^{\gamma_1(p)} \mid M$ ， $\gamma_1(p)$ 取得够大)，就能保证 $M' \equiv 0 \pmod{p^\gamma}$ 。
- 4 此时同余方程 $d_1 y_1^k + \cdots \equiv M' \equiv 0 \pmod{p^\gamma}$ 与 $M = 0$ 时完全相同，于是同样有本原解，同样得到 $\chi(p)$ 的一个正下界——而且这个下界跟 $M = 0$ 那次用的是同一个，与具体 M 无关。

脚注的意思：如果 p 根本不整除任何 c_j ，那么 1.4 的"初步变换"什么都不用做（系数本来就与 p 互素），于是 M 不会被变换抹掉任何 p 因子，直接取 $\gamma_1(p) = \gamma$ 即可。只有那些**整除某个 c_j** 的"坏素数"才需要把 $\gamma_1(p)$ 取得比 γ 大。这样的坏素数只有有限个（ c_j 的素因子），都 $\leq p_0$ 。

$$\text{模数 } L = \prod_{p \leq p_0} p^{\gamma_1(p)}$$

把所有小素数 $p \leq p_0$ 各自需要的 $p^{\gamma_1(p)}$ 乘起来，得到一个**固定的正整数 L** （只跟 k 、 c_j 有关，跟 M 无关）。它的妙处：" M 被 L 整除"一次性满足了所有小素数的要求。因为若 $L \mid M$ ，则对每个 $p \leq p_0$ 都有 $p^{\gamma_1(p)} \mid M$ （ L 含因子 $p^{\gamma_1(p)}$ ），于是每个小素数处都按上面的推理得到 $\chi(p)$ 的正下界。

整段的结论（把 1.3 和 1.5 接起来）：

1 大素数尾巴： $\prod_{p > p_0} \chi(p) \geq \frac{1}{2}$ (1.3, 自动, 与 M 无关)。

2 小素数：若 $L \mid M$ ，每个 $p \leq p_0$ 的 $\chi(p)$ 都有不依赖 M 的正下界 (1.5)。

3 相乘： $\mathfrak{S}(M) = \left(\prod_{p \leq p_0} \chi(p) \right) \left(\prod_{p > p_0} \chi(p) \right) \geq (\text{正数}) \times \frac{1}{2} > 0$ ，
且这个下界对所有 $L \mid M$ 的 M 统一成立。

这就是后面要反复用到的" $\mathfrak{S}(M)$ 一致正下界"。记住这句话：**只要让右端 M 是 L 的倍数，奇异级数就保证为正**。这解释了下一段为什么要在方程里塞一个 L^k 。

1.6 搭建主方程 (10.3)：把变量分组，并塞进因子 L^k

回到标题中的方程，我们取 $s = s_0 + 3\ell$ ，其中 ℓ 如第 9 章那样，并把系数分成若干组：

$$c_1, \dots, c_{s_0}; \quad d_1, \dots, d_\ell; \quad d'_1, \dots, d'_\ell; \quad e_1, \dots, e_\ell;$$

受到 $c_1, \dots, c_{s_0}$ 不全同号这一条件的约束。我们将确立方程

$$c_1 x_1^k + \dots + c_{s_0} x_{s_0}^k + L^k(u_1 + u_2 + y^k v) = 0 \quad (10.3)$$

的可解性，受到如下约束：

- i. $0 < x_j \leq P/|c_j|^{1/k}$;
- ii. u_1 是一个整数 $< \frac{1}{4}(P/L)^k$ ，它可表为 $d_1 z_1^k + \dots + d_\ell z_\ell^k$ ，对带撇系数的 u_2 也类似；
- iii. $1 < y < (P/L)^{1/2k}$;
- iv. v 是一个整数 $< \frac{1}{4}(P/L)^{k-1/2}$ ，它可表为 $e_1 t_1^k + \dots + e_\ell t_\ell^k$ 。

这段在说什么：把 $s = s_0 + 3\ell$ 个变量分成四组，各司其职，并写出本章真正要攻克方程 (10.3)。这是全章的"主战场方程"。

ℓ 、 P 、四组系数

ℓ (如第 9 章)

第 9 章里出现的一个整数参数，量级约 $2k \log 3k$ （它来自华林问题 $G(k)$ 的上界估计，本质是 $\sim k \log k$ ）。后面定理 10.1 里 $\ell = 2k \log 3k + 1$ 。它控制每个"稀疏打包变量"用掉多少个真实变量。

P

一个趋于无穷的大参数，控制所有变量的取值范围。让 $P \rightarrow \infty$ 就能得到"无穷多组解"。

四组系数 $c_j; d_i; d'_i; e_i$

原来的 s 个系数被重新命名分组：前 s_0 个保留为 c （主块），后面 3ℓ 个分成三组 d, d', e ，分别用来造出 u_1, u_2, v 这三个"打包变量"。

逐项理解方程 (10.3) 和它的约束：

1 主块 $c_1 x_1^k + \cdots + c_{s_0} x_{s_0}^k$ ：前 s_0 个变量自由地跑遍各自范围（约束 i：\ (00)\ (只要 $L \mid M$)。" $c_1, \dots, c_{s_0}$ 不全同号"正是从原前提继承来的，保证奇异积分 $\mathfrak{J} > 0$ 。

2 u_1, u_2 ：两个"被 d, d' 系数打包的稀疏数"。约束 (ii)： u_1 是能写成 $d_1 z_1^k + \cdots + d_\ell z_\ell^k$ 的整数，且 $u_1 < \frac{1}{4}(P/L)^k$ 。它用 ℓ 个真实变量 $z_1, \dots, z_\ell$ 造出，但我们把它当一个变量看。 u_2 完全类似，只是用带撇系数 d'_i 。这正是第 9 章稀疏集合思想（见 0.8）。

3 y 与 v ：再加一层结构 $y^k v$ 。约束 (iii)： \ (1

4 为什么是 $L^k(u_1 + u_2 + y^k v)$? —— L^k 是全章的画龙点睛之笔。把方程移项：

$$c_1 x_1^k + \cdots + c_{s_0} x_{s_0}^k = -L^k(u_1 + u_2 + y^k v) =: M.$$

右端 $M = -L^k(u_1 + u_2 + y^k v)$ 显然被 L^k 整除，更被 L 整除，即 $M \equiv 0 \pmod{L}$ ！于是 1.5 的结论立刻适用： $\mathfrak{S}(M) > 0$ 一致成立。塞 L^k 的唯一目的就是强行让右端是 L 的倍数，从而搭上 1.5 那台"奇异级数为正"的机器。用 L^k （ k 次方）而非 L ，是为了让这一项也是漂亮的"系数 $\times k$ 次方结构"（ $L^k \cdot u_1 = L^k(d_1 z_1^k + \cdots) = (L^k d_1) z_1^k + \cdots$ ，仍是标准 k 次型），与其它项格式统一，方便统一处理。

主块 $c_1X_1^k+\cdots+c_{s_0}X_{s_0}^k$
取到值 M , 保证 $\mathfrak{S}(M)>0$

$L^k\cdot u_1, L^k\cdot u_2$
稀疏打包(d, d')

$L^k\cdot y^{kv}$
稀疏打包(e)+ y

右端 $M = -L^k(u_1+u_2+y^{kv})$ 被 L 整除 $\implies \mathfrak{S}(M)>0$

变量总数 $s = s_0 + 3\ell$ (每个打包块用 ℓ 个真实变量)

四组变量的分工：主块负责凑 M ，三个打包块（每个 ℓ 个变量）负责凑 $-M/L^k$ ； L^k 保证 $L \mid M$ 。

1.7 变量数的会计： $s = s_0 + 3\ell$ 够用，更多自然也够

原文译文

当 $s = s_0 + 3\ell$ 时这将给出原方程的可解性，从而当 $s \geq s_0 + 3\ell$ 时更是如此。

这段在说什么：数一数变量。主块 s_0 个，加上 u_1, u_2, v 三个打包块、每块 ℓ 个真实变量，共 $s_0 + 3\ell$ 个。所以只要 $s = s_0 + 3\ell$ 就能证出解。"更是如此"（拉丁文 *a fortiori*，意为"更有力地"）：若变量更多（ $s > s_0 + 3\ell$ ），多出来的变量可以随便取定一组非零值、把它们的贡献并入常数，剩下的仍是同型方程，照样有解。所以 $s \geq s_0 + 3\ell$ 全都可解。

1.8 主弧 / 次弧的微调，与主块的主项估计 (10.4)

原文译文

主弧与次弧的定义与第 9 章相同，只是我们把 $2kqP^{k-1}$ 换成 $2kqcP^{k-1}$ ，其中 $c = \max |c_j|$ ；这是为了确保在对引理 9.2 的类比的证明中，范德科普特（van der Corput）引理的条件得到满足。代替引理 9.3，我们得到

$$\int_{\mathfrak{M}} T_1(\alpha) \cdots T_{s_0}(\alpha) e(-M\alpha) d\alpha \gg P^{s_0-k} \quad (10.4)$$

其前提是 $0 < M \leq P^k$ 且 $M \equiv 0 \pmod{L}$ 。

这段在说什么：把第 9 章的主弧分析"原样照搬"，只改一个数值参数，然后得到主块在主弧上的主项下界 (10.4)。

为什么把 $2kqP^{k-1}$ 换成 $2kqcP^{k-1}$?

主弧的**宽度**由这样一个量决定。第 9 章里系数都是 1，宽度用 $2kqP^{k-1}$ 。本章系数是一般的 c_j ，指数和 $T_j(\alpha) = \sum e(c_j x^k \alpha)$ 的相位是 $c_j x^k \alpha$ ，对 x 求导多出一个因子 c_j 。**van der Corput 引理**（一个估计指数和的工具，要求相位的导数变化被恰当控制）的适用条件因此要把这个 c_j 考虑进来。于是把宽度参数里的 " P^{k-1} " 配上 $c = \max_j |c_j|$ （最大的系数绝对值），写成 $2kqcP^{k-1}$ ，保证**对**所有 j 引理条件都满足。这是个纯技术的"放大保险系数"调整，不影响主项量级。

q 、 c 、(10.4) 各符号

q

主弧所围绕的那个分数 a/q 的分母。

$c = \max |c_j|$

所有系数绝对值里最大的那个，是一个固定常数。

$\int_{\mathfrak{M}}$

只在主弧 $\mathfrak{M}$ 上积分（不是整个 $[0, 1]$ ）。

$e(-M\alpha)$

探测器因子，挑出"前 s_0 项之和 $= M$ "的那些组（对比 0.3，这里右端是 M 而非 0，所以乘 $e(-M\alpha)$ ）。

读懂 (10.4): 左边正是"主块 $c_1x_1^k + \cdots + c_{s_0}x_{s_0}^k = M$ 在主弧上贡献的解数主项"。右边 $\gg P^{s_0-k}$ 说它有一个正下界, 量级是 P^{s_0-k} 。这个量级怎么理解? s_0 个变量各约 P 种取值, 共约 P^{s_0} 组; 要求落在某个 $= M$ 的"一维条件"上, 相当于除以 P^k (值域宽约 P^k); 得 P^{s_0-k} 。前提两条: $\setminus(00\setminus)$, 主项才真的是正的)。

1.9 (10.4) 的证明差别: 只差在奇异积分, 化成一个多重积分

原文译文

这一证明与前面的不同之处仅在于奇异积分, 它变换成下式的一个倍数:

$$\int_0^1 \cdots \int_0^1 \left\{ \zeta_1 \cdots \zeta_{s_0-1} (\pm \zeta_1 \pm \cdots \pm \zeta_{s_0-1} - \theta) \right\}^{1-1/k} d\zeta_1 \cdots d\zeta_{s_0-1},$$

积分区域为范围 $0 < \pm \zeta_1 \pm \cdots \pm \zeta_{s_0-1} - \theta < 1$, 其中那些符号是 $c_1, \dots, c_{s_0-1}$ 的符号, 而我们已假定 c_{s_0} 是负的。我们可以设那些 $\pm$ 号中至少有两个是 $+$ 。那么, 对于 $0 \leq \theta \leq 1$, 积分区域包含某个大小不依赖于 θ 的小立方体, 从而上述积分有一个正的下界。

这段在说什么: 证 (10.4) 时, 奇异级数那部分已由 1.5 解决; 剩下要保证的是**奇异积分** $\mathfrak{J} > 0$ (实数可解性)。作者把它具体写成一个多重积分, 并论证它为正。下面把这个论证彻底讲清。

ζ_i 、 θ 与这个积分

$\zeta_1, \dots, \zeta_{s_0-1}$

读作"zeta"。它们是把变量 x_j 标准化 (缩放到 $[0, 1]$ 量级) 后得到的连续积分变量, 对应前 $s_0 - 1$ 个变量。

θ

读作"theta"。它是把右端 M 标准化后得到的实参数; 因为前提 $\setminus(0$

指数 $1 - \frac{1}{k}$

来自把 x_j^k 反解成 x_j 时的换元：若 $w = x^k$ ，则 $x = w^{1/k}$ ， $dx = \frac{1}{k} w^{1/k-1} dw$ ，求导带出 $w^{-1/k}$ 这类因子；多变量累积后给出每个括号 $1 - 1/k$ 次幂。这正是 k 次方曲面的"面积元"权重，标准的奇异积分形式。

核心：为什么这个积分 > 0 ？ 关键在积分区域 $0 < \pm\zeta_1 \pm \cdots \pm \zeta_{s_0-1} - \theta < 1$ 非空，而且大到能装下一个固定大小的小立方体。逐步论证：

1 符号怎么来的： 那串 $\pm$ 正是各系数 $c_1, \dots, c_{s_0-1}$ 的正负号（正系数取 +、负系数取 -）。作者已经把第 s_0 个安排成负的（ $c_{s_0} < 0$ ），它的贡献被并进了右端的 $-\theta$ 。这一步用到了原前提" c 不全同号"——正因为不全同号，才能挑出一个负的 c_{s_0} ，并让前面的符号串里有正有负。

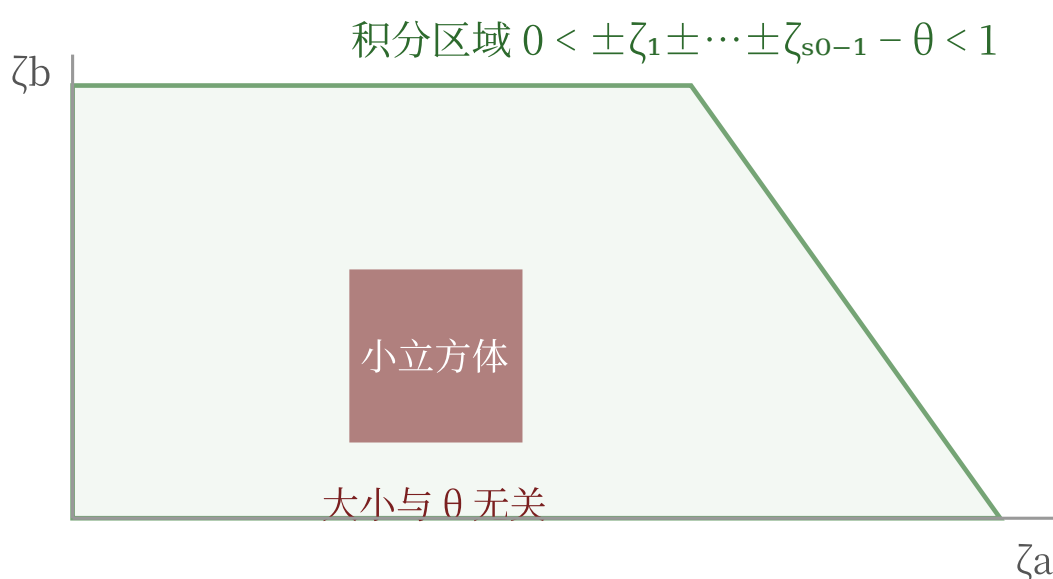
2 "至少有两个 +：" 由"不全同号"可知正系数和负系数都存在。我们可以重新编号，让 $c_1, \dots, c_{s_0-1}$ 中至少保留两个正号（ s_0 很大，正号要么本来就 ≥ 2 个，要么把负的都挪到 c_{s_0} 一侧后仍有正号剩下）。"两个 + "给了我们调节的自由度。

3 构造小立方体： 设第 a, b 两个位置是 +。固定其它 ζ 为某个适当的小值，让表达式 $\pm\zeta_1 \pm \cdots \pm \zeta_{s_0-1} - \theta$ 中的可调部分主要是 $+\zeta_a + \zeta_b$ 。当 ζ_a, ζ_b 在一个小区间（比如长度 δ 的窗口）里变动时， $+\zeta_a + \zeta_b$ 能连续地扫过一段长度 $\approx 2\delta$ 的实数区间。无论 $\theta \in [0, 1]$ 是多少，总能让整个表达式落进开区间 $(0, 1)$ 内——因为有两个可自由增减的正变量，"够灵活"地把和推到任何想要的水平。这就保证了对每个 θ ，积分区域里都含有一个边长不依赖 θ 的小立方体。

4 被积函数在该立方体上为正： 在小立方体内，所有 $\zeta_i > 0$ 且括号 $\pm\zeta_1 \pm \cdots - \theta$ 落在 $(0, 1)$ 内为正，于是花括号里整体为正，其 $1 - 1/k$

次幂也为正且有正下界。"正函数在正体积区域上积分"必为正，且下界不依赖 θ 。

- 5 结论： $\mathfrak{J} > 0$ （一致地）。这正是实数可解性：在允许的盒子里，曲面 $c_1 x_1^k + \cdots = M$ 确实穿过，且有一整块解。



有两个 $+$ 号的正变量提供自由度，无论 θ 取多少，区域里总含一个固定大小的小立方体，使奇异积分有正下界。

1.10 拼起来得到 (10.4)

原文译文

我们已经看到，出现在引理 9.3 的类比中的奇异级数，对于 $M \equiv 0 \pmod{L}$ 有一个正的下界，因此我们得到 (10.4)。

这段在说什么：收口。主弧主项 $= P^{s_0-k} \times \mathfrak{S}(M) \times \mathfrak{J}$ 。其中：

- 1 奇异级数 $\mathfrak{S}(M)$: 由 1.5, 只要 $M \equiv 0 \pmod{L}$ 就有正下界 (与 M 无关)。
- 2 奇异积分 $\mathfrak{J}$: 由 1.9, 有一致正下界。
- 3 两个正下界相乘再乘 P^{s_0-k} , 得到 $\gg P^{s_0-k}$, 正是 (10.4)。证毕。

注意 (10.4) 同时用到了前提的两半: 奇异级数为正用到了 " $L \mid M$ " (数论/局部条件), 奇异积分为正用到了 " c 不全同号" (实数/无穷处条件)。两者缺一不可, 正对应 0.5 的 "局部 $\implies$ 整体" 原则需要每个素数处和实数处都过关。

1.11 稀疏打包数的计数: 改造引理 9.4

原文译文

引理 9.4 仍然适用, 只需作微小改动, 即可给出可表为 $d_1 z_1^k + z$ 形状的数 u 或 v 的个数的下界, 其中 z 可表为 $d_2 z_2^k + \cdots + d_\ell z_\ell^k$; 并且, 如果变量 z_1 与 z 被限制在适当的范围内, 这些数全都互不相同。其范围当然依赖于 $d_1, \dots, d_\ell$ 。我们对 $U_\ell(X)$ 得到与前面相同的下界, 只差一个依赖于 $d_1, \dots, d_\ell$ 的常数。

这段在说什么: 要让前面 (10.4) 那台机器跑起来, 还得知道 "打包数 u 、 v 到底有多少个"——也就是稀疏集合不能太稀, 得有足够多元素。这由第 9 章的引理 9.4 提供, 本章只需微调。

$U_\ell(X)$: 稀疏集合的计数函数

$U_\ell(X)$ 表示"不超过 X 、且能表成 $d_1 z_1^k + \cdots + d_\ell z_\ell^k$ 形状的不同整数的个数"。我们需要它有一个好的下界（足够多），这样 u_1, u_2, v 的可取值才够丰富，能覆盖足够多的 M 。

"互不相同"这个细节为什么重要，怎么保证：

- 1 把表示拆成两层： $u = d_1 z_1^k + z$ ，其中"内层" $z = d_2 z_2^k + \cdots + d_\ell z_\ell^k$ 。这是把一个 ℓ 变量的和，看成" $d_1 z_1^k$ "加上一个 $(\ell - 1)$ 变量的内层和 z 。
- 2 **为什么要它们互不相同：**计数时如果同一个数 u 被不同的 $(z_1, \dots, z_\ell)$ 重复算到，下界就会被"灌水"虚高，进而后面用到这些计数的不等式会失真。所以要保证不同的 (z_1, z) 给出不同的 u ，下界才是"真材实料"的下界。
- 3 **怎么保证：**把 z_1 限制在一段范围、把内层 z 也限制在一段范围，使得 $d_1 z_1^k$ 的"步长"大于内层 z 的"波动范围"。这样 $u = d_1 z_1^k + z$ 里， $d_1 z_1^k$ 决定了 u 的"大档位"、 z 决定"档位内的细节"，两段不重叠 $\implies$ 不同 (z_1, z) 必给出不同的 u 。这就像十进制里十位和个位不会互相串位。范围当然要依赖系数 $d_1, \dots, d_\ell$ （系数大小决定步长和波动幅度）。
- 4 **结论：** $U_\ell(X)$ 的下界与第 9 章（系数全为 1 时）相同，只多一个依赖 $d_1, \dots, d_\ell$ 的常数因子。常数无所谓（回忆 0.7：我们只关心量级），所以稀疏集合"够大"这件事照样成立。

1.12 两个打包变量 u_1, u_2 ，用 Cauchy-Schwarz 不等式合并

我们现在需要两个指数和 $R_1(\alpha), R_2(\alpha)$ ，分别对应于两个不同的数集 u_1, u_2 ；但由于

$$\int_0^1 |R_1(\alpha)R_2(\alpha)| d\alpha \leq \left\{ \int_0^1 |R_1(\alpha)|^2 d\alpha \int_0^1 |R_2(\alpha)|^2 d\alpha \right\}^{1/2},$$

我们得到与引理 9.4 的推论中相同的节省。

这段在说什么：方程 (10.3) 里有**两个**打包变量 u_1 （用系数 d ）和 u_2 （用系数 d' ），需要分别给它们各配一个指数和 $R_1(\alpha)$ 、 $R_2(\alpha)$ 。在估计次弧误差时会冒出乘积 $|R_1 R_2|$ 的积分，要把它压住。工具就是那条不等式——它是大名鼎鼎的 **Cauchy-Schwarz 不等式** 的积分版。

$R_1(\alpha), R_2(\alpha)$ 与 $\int_0^1 |R|^2$ 的含义

$$R_1(\alpha) = \sum_{u_1} e(L^k u_1 \alpha)$$

对所有合法的打包数 u_1 （满足约束 ii）求和的指数和； R_2 同理对 u_2 。

$|R_i(\alpha)|$ 是复数 R_i 的**模长**（到原点的距离， $|a + bi| = \sqrt{a^2 + b^2}$ ）。

$$\int_0^1 |R_i(\alpha)|^2 d\alpha$$

由正交关系 (0.2) 展开， $\int_0^1 |R_i|^2 d\alpha$ 恰好等于“集合里数的个数”（每个数和自己配对贡献 1，不同的数互相抵消）。所以这两个积分就是 u_1 、 u_2 各自可取值的**数量**，正是 1.11 给出下界的那个 U_ℓ 。

Cauchy-Schwarz 不等式（积分版） 对任意（复值）函数 f, g ,

$$\int_0^1 |f(\alpha)g(\alpha)| d\alpha \leq \left(\int_0^1 |f|^2 d\alpha \right)^{1/2} \left(\int_0^1 |g|^2 d\alpha \right)^{1/2}.$$

高中里你见过它的向量版: " $\vec{a} \cdot \vec{b} \leq |\vec{a}| |\vec{b}|$ ", 即

$(\sum a_i b_i)^2 \leq (\sum a_i^2)(\sum b_i^2)$ 。把"求和 $\sum$ "换成"积分 $\int$ ", 向量换成函数, 就是上式。

这条不等式带来什么"节省", 逐步说清:

- 1 直接处理混合项 $\int_0^1 |R_1 R_2| d\alpha$ 不好办 (两个不同集合纠缠在一起)。
- 2 用 Cauchy-Schwarz 把它拆开成 $(\int |R_1|^2)^{1/2} (\int |R_2|^2)^{1/2}$, 于是只需分别估计 $\int |R_1|^2$ 和 $\int |R_2|^2$ 。
- 3 而这两个积分 (由上面的注记) 就是 u_1, u_2 各自的元素个数, 已经在 1.11 算好了下界/上界。
- 4 于是混合项的处理退化成单个集合的处理, 和第 9 章只有一个打包变量时 (引理 9.4 的推论) 得到完全相同的"节省" (即相同量级的误差节约)。
"节省"指: 误差被压低的幅度, 正好够让主项盖过误差。

为什么要用两个 u_1, u_2 而不是一个? 因为方程 (10.3) 需要 " $u_1 + u_2$ " 这种两数之和才能覆盖足够稠密的 M 取值 (两个稀疏集合相加, 和集更稠密); 这是第 9 章"加法能让稀疏集合变稠"思想的体现。而代价 (处理两个指数和) 被 Cauchy-Schwarz 一招化解, 不增加本质困难。

1.13 收尾: 引理 9.5 的推论照搬, 得到定理

原文译文

引理 9.5 的推论本质上不变, 证明也如前完成。于是我们得到如下结果。

这段在说什么：把所有零件装到第 9 章的总装配线（引理 9.5 的推论）上。那条装配线干的事是：主弧主项（正，量级 P^{s_0-k} ，由 1.8–1.10）+ 稀疏变量的足量供应（由 1.11–1.12）> 次弧误差（由微调后的引理 9.2 类比压住），于是总积分 > 0 ，方程 (10.3) 有解。由于本章对第 9 章的改动都只是“换常数、套 Cauchy-Schwarz”这类不触及结构的微调，所以“本质上不变，证明如前完成”。把 $P \rightarrow \infty$ ，解的组数趋于无穷，得到无穷多组解。

1.14 最终结论：定理 10.1

原文译文

定理 10.1. 设 $c_1, \dots, c_s$ 是整数，没有一个等于 0，并且当 k 为偶数时它们不全同号。那么，如果

$$s \geq s_0 + 3(2k \log 3k + 1), \quad (10.5)$$

其中 s_0 取 (10.1) 中给出的值，则方程 $c_1 x_1^k + \dots + c_s x_s^k = 0$ 有无穷多个整数解 $x_1, \dots, x_s$ ，没有一个等于 0。

把 (10.5) 的每一块拆开看，确认它和前面对得上：

s_0

$= k^2(2k - 1) + 1$ （来自 (10.1)，第 8 章给的“局部处处可解”门槛，量级 $\approx 2k^3$ ）。这是主块用掉的变量。

3

三个打包变量 u_1, u_2, v （见 1.6）。

$\ell = 2k \log 3k + 1$

每个打包变量用掉的真实变量数（来自第 9 章 $G(k)$ 的估计，量级 $\approx 2k \log k$ ）。这里 $\log$ 是自然对数。

$$\text{合计 } s_0 + 3\ell = s_0 + 3(2k \log 3k + 1)$$

正是 1.7 的 $s = s_0 + 3\ell$ 。

为什么说这比定理 8.1"限制更弱"? 因为新增部分 $3(2k \log 3k + 1) \approx 6k \log k$ 是**对数量级**，远小于主项 $s_0 \approx 2k^3$ 。也就是说，本章基本上做到了"只要变量数达到保证局部可解的门槛 s_0 ，再多添几个（对数级）变量就够了"。相比之下，第 8 章的定理 8.1 需要的额外变量更多。本章用第 9 章的稀疏打包技巧，把"额外开销"从多项式级压到了对数级——这正是把"上一节的方法加以改造"换来的回报。

代入感受一下量级 取 $k = 3$ （三次方程）。则

$$s_0 = k^2(2k - 1) + 1 = 9 \cdot 5 + 1 = 46; \text{ 又}$$

$$2k \log 3k + 1 = 6 \ln 9 + 1 \approx 6 \times 2.197 + 1 \approx 14.2, \text{ 于是 (10.5) 的右端}$$

$s_0 + 3(2k \log 3k + 1) \approx 46 + 3 \times 14.2 \approx 88.6$ 。由于 s 必须是整数，条件就是 $s \geq 89$ 。即：对一般三次型 $c_1 x_1^3 + \cdots + c_s x_s^3 = 0$ （系数非零，不要求不同号，因为 $k = 3$ 为奇数），只要变量数 ≥ 89 就保证有无穷多组非零整数解。其中 46 个用于"保证模每个素数都有解"，其余约 43 个用于第 9 章的稀疏打包凑数。

全章逻辑一图流

1 圆法：解数 $= \int_0^1$ （探测器把"= 0"翻译成积分，0.3）。

2 切成主弧+次弧：主项 $= P^{s-k} \mathfrak{S} \mathfrak{J}$ (0.4)。

3 要 $\mathfrak{S}(M) > 0$ 一致：大素数尾巴自动 $\geq \frac{1}{2}$ (1.3) + 小素数用 $L = \prod p^{\gamma_1(p)}$ 控制，只要 $L \mid M$ (1.4–1.5)。

4 造方程 (10.3)：塞 L^k 强行让右端 M 被 L 整除 (1.6)；用 $s_0 + 3\ell$ 个变量 (1.7)。

5 主弧主项 (10.4): 奇异级数靠 $L \mid M$ 为正、奇异积分靠"不全同号"为正 (1.8–1.10)。

6 稀疏打包够量: 改造引理 9.4 给 U_ℓ 下界 (1.11), 两个 u 用 Cauchy–Schwarz 合并 (1.12)。

7 主项 $>$ 误差 $\implies$ 积分 $> 0 \implies$ 无穷多解 (1.13–1.14, 定理 10.1)。

返回 [全书目录](#)