

DAVENPORT · 圆法 · 高中详解版

前言、目录与编者序

Foreword, Contents & Editorial Preface

本章要解决什么·读完能掌握什么 这一章是全书的「门厅」：书名页、目录、三篇专家撰写的前言（导读）、以及编者序。它本身**不证明任何定理**，但它把全书会反复出现的**所有核心词汇与记号**第一次集中端了出来——丢番图方程、华林问题、型（form）、圆法、生成函数 $T(\alpha)$ 、主弧与次弧、奇异级数、奇异积分、 $G(k)$ 、 p -进、局部到整体……。本讲解的目标是：把这些对高中生而言完全陌生的名词，**从零讲清楚**，让你读完这扇门厅后，手里握着一张「全书地图」，知道后面 20 章各自在干什么、为什么这么排。所以这一章我们会**故意慢**：每出现一个新名词、新记号，都停下来从最基础讲起。

阅读方式：黑色引用块是译文的原话（按译文段落顺序逐段抄录），其下的彩色框与配图是面向高中生的逐句详解。凡译文里一笔带过的名词，我们都补全成完整解释；凡涉及超出高中的概念（复指数、定积分、求和号、同余、 O 与 $\ll$ 记号等），都从零讲起。

读前必备：先认全本书反复出现的核心词汇A primer on the vocabulary used throughout the book

前言里专家们默认你已经懂「丢番图方程」「型」「圆法」「 $G(k)$ 」这些词。高中课本里没有它们。所以在逐段读前言之前，我们先用这一节把这些词一次性讲透。后面读到时你就能轻松跟上，不必每次卡壳。

① 丢番图方程是什么

丢番图方程 (Diophantine equation) 一个整系数的多项式方程，但我们只允许未知数取整数值（有时限定正整数），并问：它有没有整数解？有多少个？

「丢番图」是公元 3 世纪古希腊数学家丢番图 (Diophantus) 的名字。

例： $x^2 + y^2 = z^2$ 在整数范围内问解，就是丢番图方程（解如 3, 4, 5）；而 $x^n + y^n = z^n$ ($n \geq 3$ 无正整数解) 就是著名的费马大定理。关键词是「整数解」——同一个方程若允许实数解往往很容易，限定整数后立刻变难。

② 「型」是什么：齐次多项式

型 (form) · 次数 d (degree) · 齐次 (homogeneous) 「型」就是每一项次数都相同的多项式，专业叫「齐次多项式」。这个相同的次数叫它的次数 d 。

读法：degree 读「迪格瑞」，就是「次数」。

例： $x^2 + xy + y^2$ 每项都是 2 次 $\rightarrow$ 这是一个 2 次型（二次型，quadratic form）； $3x_1^3 + 4x_2^3 + 5x_3^3$ 每项都是 3 次 $\rightarrow$ 3 次型（三次型，cubic form）。

为什么数论特别爱「齐次」？因为齐次型有一个漂亮性质：若 $(x_1, \dots, x_n)$ 是 $F = 0$ 的解，则 $(tx_1, \dots, tx_n)$ 也是解（两边都乘 t^d ）。所以解会成「整条射线」出现， $(0, \dots, 0)$ 总是一个解——这个解叫平凡解 (trivial solution)。真正难的问题是：有没有非平凡整数解（不全为 0 的整数解）？这正是本书第 8 章往后的主线。

对角型 (diagonal form) 最简单的一类型：只有「纯幂」项、没有交叉项，形如 $c_1x_1^k + c_2x_2^k + \dots + c_sx_s^k$ 。因为把系数排成矩阵时只有对角线上有数，故称「对角」。本书前半部分（华林问题、第 7、8、10、20 章）打交道的几乎都是对角型；一般（非对角）型留到第 11 章以后。

③ 华林问题与 k 次幂之和

k 次幂 (k -th power) 就是某个整数的 k 次方: $1^k, 2^k, 3^k, \dots$ 。 $k = 2$ 是平方数 $1, 4, 9, 16, \dots$; $k = 3$ 是立方数 $1, 8, 27, 64, \dots$ 。

1770 年, 英国数学家**华林 (Edward Waring)** 断言: 对每个固定的 k , 都存在一个数 s , 使得每个正整数都能写成至多 s 个 k 次幂之和。例如 $k = 2$ (平方数) 时, 拉格朗日早已证明 $s = 4$ 够用——每个正整数都是 4 个平方数之和 (如 $7 = 4 + 1 + 1 + 1$, $31 = 25 + 4 + 1 + 1$)。 $k = 3$ (立方数) 时大约 $s = 9$ 够用, 等等。这就是**华林问题 (Waring's problem)**, 是本书第 1-10 章的主题。

$g(k)$ 与 $G(k)$ 这是华林问题里两个核心记号, 务必分清:

- $g(k)$: 使「**每一个**正整数都是至多 s 个 k 次幂之和」成立的最小 s 。它常被几个「捣乱的小数字」(如 **23, 239**) 拖大。
- $G(k)$: 使「**每一个充分大的**正整数都是至多 s 个 k 次幂之和」成立的最小 s 。「充分大」意思是「除了有限多个例外」。

读法: 大写 $G(k)$ 读「大 G of k」, 小写 $g(k)$ 读「小 g of k」。

$G(k)$ 才是真正反映问题深层难度的量, 因为它撇开了少数例外的偶然干扰。已知 $g(2) = G(2) = 4$; 而 $G(3) \leq 7$ 但 $g(3) = 9$ (**23** 和 **239** 这两个例外把 $g(3)$ 顶到了 9)。本书第 9 章、以及第一篇前言里那一长串数字, 讲的都是 $G(k)$ 。

④ 圆法的一句话核心: 把「数解的个数」变成「算一个积分」

这是全书的方法论灵魂, 叫**哈代-李特尔伍德圆法 (Hardy-Littlewood circle method)**, 1920 年代由英国的哈代 (G. H. Hardy) 和李特尔伍德 (J. E. Littlewood) 创立。下面三个记号是它的全部零件, 逐个讲。

求和号 $\sum$ (sigma) $\sum$ 是希腊字母大写 sigma, 读「西格玛」, 意思是「把一串东西加起来」。 $\sum_{x=1}^P a_x$ 表示 $a_1 + a_2 + \dots + a_P$ 。例:

$$\sum_{x=1}^3 x^2 = 1 + 4 + 9 = 14。$$

复指数 $e(\alpha) = e^{2\pi i \alpha}$ 这是圆法的「心脏」，高中没学过，从零讲。

先认 i ：虚数单位，满足 $i^2 = -1$ 。一个复数 $a + bi$ 可以画成平面上的点 (a, b)

。

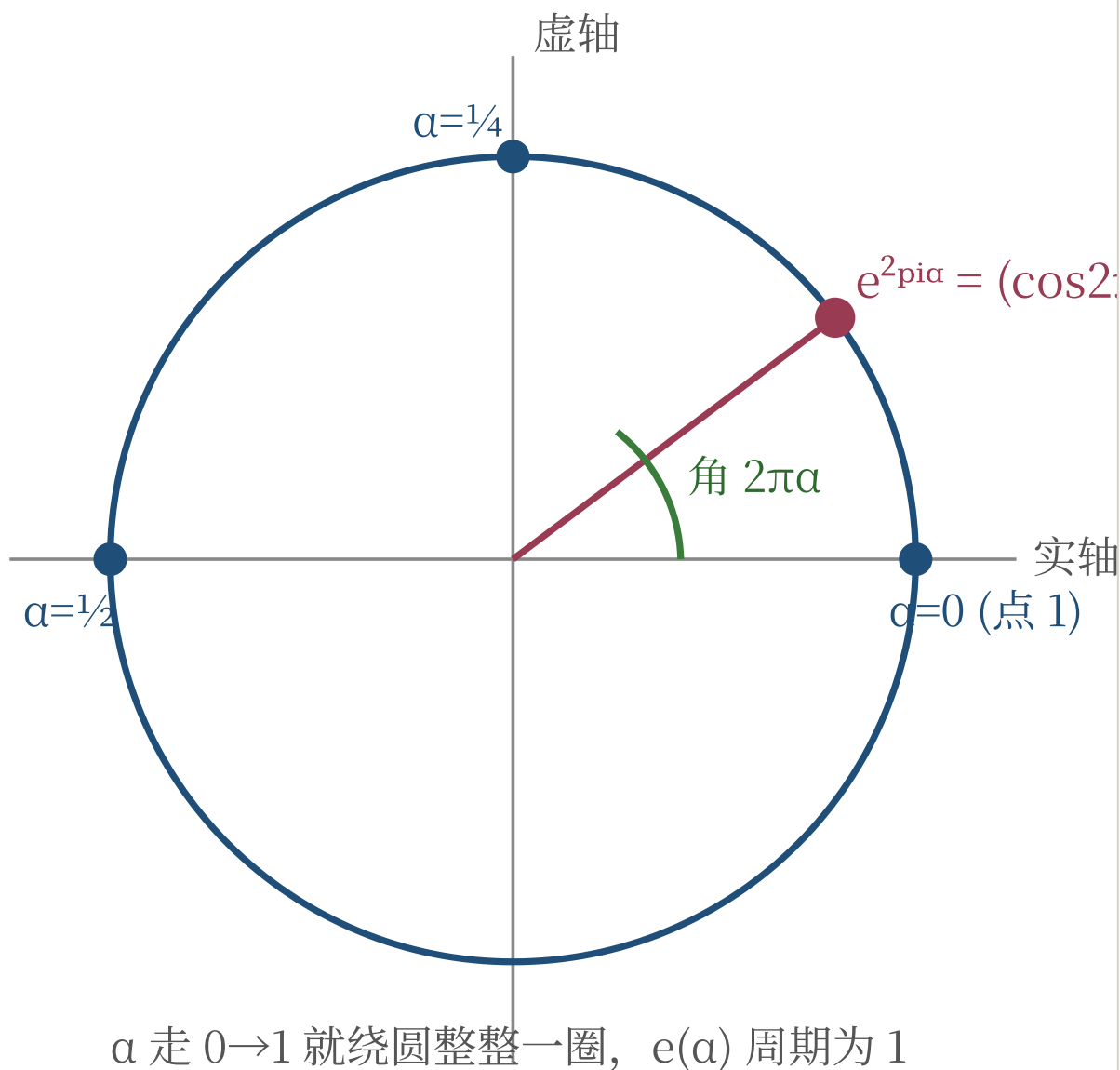
欧拉公式（大学内容，先接受）： $e^{i\theta} = \cos \theta + i \sin \theta$ 。它表示单位圆上转了角度 θ （弧度）的那个点，因为 $(\cos \theta, \sin \theta)$ 恰在半径为 1 的圆上。

本书为书写方便，定义 $e(\alpha) := e^{2\pi i \alpha}$ 。读法：「 e of α 」。它表示把整个单位圆等分后、转过「 α 整圈」的点： $\alpha = 0$ 在 $(1, 0)$ ； $\alpha = \frac{1}{4}$ 转了四分之一圈到 $(0, 1)$ ； $\alpha = \frac{1}{2}$ 到 $(-1, 0)$ ； $\alpha = 1$ 转满一圈又回到 $(1, 0)$ 。

两个一定要记住的性质：

(i) **周期 1**： $e(\alpha + 1) = e(\alpha)$ （多转一整圈回到原地）。所以 α 只需在 $[0, 1]$ 上看就够了——这就是为什么后面积分区间是 $[0, 1]$ 。

(ii) **模长恒为 1**： $|e(\alpha)| = 1$ （永远在单位圆上，长度不变）。这里 $|z|$ 是复数 z 到原点的距离，叫「模长」。



$e(\alpha) = e^{2\pi i \alpha}$ 是单位圆上「转过 α 整圈」的点：永远在圆上（模长 1），且 α 每增加 1 就回到原处（周期 1）。这就是把 $[0, 1]$ 看成「一个圆」、方法得名「圆法」的原因。

定积分 $\int_0^1 f(\alpha) d\alpha$ $\int$ 是拉长的 S (sum 的首字母)，读「积分」。 $\int_0^1 f(\alpha) d\alpha$

直观上就是：把区间 $[0, 1]$ 切成无穷多个极细的小段，每段宽 $d\alpha$ 、高 $f(\alpha)$ ，把这些细长条的「面积」全加起来。它是「连续版的求和」。本书里 f 取复数值，但运算规则一样。我们只需要它的一条核心性质，下面 $T(\alpha)$ 处会用到。

生成函数 $T(\alpha) = \sum_{x=1}^P e(\alpha x^k)$ (generating function) 为什么要造这个怪

东西？因为它能把「数解」变成「算积分」。 $T(\alpha)$ 把从 1 到 P 的每个数 x 的 k 次幂 x^k ，统统塞进复指数当指数，再加起来。它是一串单位圆上的点的和。

读法：「 T of alpha」， T 取自 trigonometric（三角）或 generating。

圆法的核心恒等式（第 4 章会正式登场，这里先认脸）是：把 N 写成 s 个 k 次幂之和的表法数 $r(N)$ 等于

$$r(N) = \int_0^1 T(\alpha)^s e(-N\alpha) d\alpha.$$

为什么这条等式成立？关键是复指数的一条「正交性」：

1 把 $T(\alpha)^s = \left(\sum_x e(\alpha x^k) \right)^s$ 展开，得到对所有 s 元组 $(x_1, \dots, x_s)$ 求和： $T(\alpha)^s = \sum_{x_1, \dots, x_s} e(\alpha(x_1^k + \dots + x_s^k))$ 。这一步只是把乘积按分配律拆开。

2 再乘上 $e(-N\alpha)$ 并对 α 在 $[0, 1]$ 上积分，把求和与积分交换次序，得到 $\sum_{x_1, \dots, x_s} \int_0^1 e(\alpha(x_1^k + \dots + x_s^k - N)) d\alpha$ 。

3 用复指数的正交性：对任意整数 m ， $\int_0^1 e(\alpha m) d\alpha = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$
(道理： $m \neq 0$ 时 $e(\alpha m)$ 在圆上均匀绕 $|m|$ 圈，正负抵消，积分为 0； $m = 0$ 时被积函数恒等于 1，积分得 1。)

4 于是只有满足 $x_1^k + \dots + x_s^k - N = 0$ 的那些组对积分贡献 1，其余贡献 0。把贡献 1 的组数一数，恰好就是 N 的表法数 $r(N)$ 。等式得证。

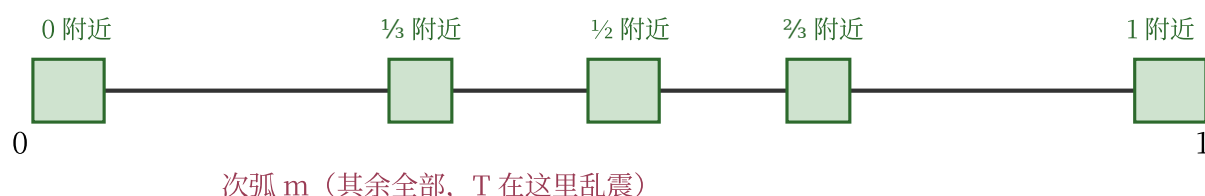
这就是圆法的全部魔法：「 N 有几种写法」这个离散的计数问题，被翻译成「算一个积分」这个分析问题。剩下的全部工作，就是想办法把这个积分算出来（或估出主项）。

主弧 $\mathfrak{M}$ 与次弧 $\mathfrak{m}$ (major arcs / minor arcs) 怎么算上面那个积分？哈代和李特尔伍德的洞见是：把积分区间 $[0, 1]$ （看成圆）切成两类小段——

- **主弧 $\mathfrak{M}$** (major arcs, 「大弧」)： α 离某个分母很小的分数 a/q （如 $0, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \dots$ ）很近的那些小区间。在这里 $T(\alpha)$ 很大、有漂亮的近似公式，贡献了积分的主项。
- **次弧 $\mathfrak{m}$** (minor arcs, 「小弧」)：剩下的所有 α 。这里 $T(\alpha)$ 剧烈震荡、忽大忽小、没有封闭公式，我们既算不出也不想算，只需证明它们加起来比主项小，淹没不了主项即可。

记号： $\mathfrak{M}$ 、 $\mathfrak{m}$ 是花体（哥特体）的大、小写 M ，读音同 M 。「弧 (arc)」这个字正来自把 $[0, 1]$ 卷成圆后，这些小区间就成了圆上的一段段「弧」。

绿块 = 主弧 $\mathfrak{M}$ （围着小分母分数， T 在这里很大、有公式）



圆法的招牌动作：把 $[0, 1]$ 切成围着小分母分数 a/q 的若干「主弧」（绿块）和剩下的「次弧」。主项几乎全来自主弧；次弧只需证明它「小到可忽略」。

奇异级数 $\mathfrak{S}$ (singular series) 与奇异积分 $\mathfrak{J}$ (singular integral) 把主弧上的贡献算到底，主项会自然分解成两个因子的乘积：

- **奇异积分 $\mathfrak{J}$** ：来自实数侧的「连续」信息，大致衡量方程在实数范围有没有解、解的「体积」多大。
- **奇异级数 $\mathfrak{S}$** ：一个无穷乘积（对每个素数 p 各一个因子），来自整数同余侧的信息，衡量方程「对每个模 p^j 有没有解」。

读法： $\mathfrak{S}$ 是花体 S (singular series), $\mathfrak{J}$ 是花体 J。「奇异 (singular)」是历史叫法，并不是「奇怪」，你就当成专有名词。

一句话直觉：表法数 $\approx$ (奇异级数 $\mathfrak{S}$) $\times$ (奇异积分 $\mathfrak{J}$) $\times P^{s-k}$ 。要让方程有大量解，就必须证明 $\mathfrak{S} > 0$ 且 $\mathfrak{J} > 0$ 。所以前言里反复出现「奇异级数有界离 0 (bounded away from 0)」这句话，意思就是「 $\mathfrak{S}$ 不会塌到 0、稳稳大于某个正数」——这正是「方程在每个素数模下都有解」的体现。

⑤ 估计的语言： $O(\cdot)$ 与 $\ll$ 记号

大 O 记号 $O(\cdot)$ 与 Vinogradov 记号 $\ll$ 做估计时我们常常不在乎精确值、只在乎量级 (大约多大)，于是用这两个记号。

· $f(P) = O(g(P))$ 读「 f 是大 O of g 」，意思是：存在一个常数 C ，使得当 P 足够大时 $|f(P)| \leq C g(P)$ 。直白说： f 的大小不超过 g 的某个常数倍。

· $f \ll g$ 与 $f = O(g)$ 意思完全相同 ($\ll$ 是 Vinogradov 引入的记号，读「 f 远小于号 g 」，但实际含义就是「至多是常数倍」，并非真的「小很多」)。

例： $3P^2 + 5P + 7 = O(P^2)$ ，因为 P 大时它被 $15P^2$ 压住；同样写成 $3P^2 + 5P + 7 \ll P^2$ 。

· $o(g)$ 读「小 o of g 」，比 O 强：表示 $f/g \rightarrow 0$ ，即 f 与 g 相比可以忽略不计。例： $P \log P = o(P^2)$ 。

· P^ε (epsilon, 读「埃普西龙」)： ε 是任意小的正数， P^ε 表示「比任何 P 的正幂都增长得慢、但允许有一点点增长」的微小因子，估计里常被当作「几乎可忽略的尾巴」。

⑥ 同余、 p -进、局部到整体

同余 $a \equiv b \pmod{m}$ 读「 a 同余 b 模 m 」，意思是 $a - b$ 能被 m 整除，即 a, b 除以 m 余数相同。例： $17 \equiv 2 \pmod{5}$ ，因为两者除以 5 都余 2。「剩余类 r 模 16」就是指所有除以 16 余 r 的整数构成的一族。

p -进 / p -adic、域 $\mathbb{Q}_p$ 、局部到整体 (local-to-global) 一个整数方程能不能有整数解，有两道「必须先过」的关卡：

关卡一 (实数 / 无穷处)：方程在实数范围内有没有解？例如 $x^2 + y^2 = -1$ 在实数里就无解，那整数里更不可能。

关卡二 (每个素数 p)：方程对每个模 $p, p^2, p^3, \dots$ 有没有解（同余意义下）？若对某个 p 连同余方程都无解，那整数解当然也不存在。

把「关卡二对固定素数 p 的全部信息」打包成的数系，就叫 **p -进数域 $\mathbb{Q}_p$** （读「 p -adic」， $\mathbb{Q}$ 是有理数集的记号，读「Q」）。「方程在 $\mathbb{Q}_p$ 里有解」 $\approx$ 「方程对所有 p 的幂模都有解」。这些是**局部条件** (local，只看一个素数或只看实数)。真正想要的整数解是**整体条件** (global)。

局部到整体原理 (local-to-global principle)：如果所有局部关卡（实数 + 每个 p ）都通过了，能否保证整体（整数 / 有理数）解一定存在？对二次型，哈塞-闵可夫斯基定理说「是！」；对更高次型一般「不是」（前言里塞尔默的例子 $3x_1^3 + 4x_2^3 + 5x_3^3 = 0$ 就是反例：它过了所有局部关卡却没有非平凡有理解）。本书后半部分的一大主题，就是：**当变量个数足够多时，局部到整体原理能不能救回来。**

把以上词汇备齐，前言里那些「奇异级数有界离 0」「 p -进解存在」「主弧上估计 $T(\alpha)$ 」就都能读懂了。下面正式逐段讲解译文。

书名页信息 Title page & publication data

《丢番图方程与丢番图不等式的解析方法》(Analytic Methods for Diophantine Equations and Diophantine Inequalities)，第二版。
作者 H. 达文波特，本版由 T. D. 布朗宁编辑整理，剑桥大学出版社，收入「剑桥数学文库」。

这段在说：书叫什么、谁写的、谁整理的、谁出版。

书名拆开看就是本书的两条主线：「丢番图方程」（求多项式方程的整数解，第 1–19 章）+「丢番图不等式」（把等号换成「绝对值小于某个数」，第 20 章）；前缀「解析方法」（analytic methods）指的就是用分析学工具（积分、复指数、级数）来研究这些本属于「整数」的问题——这正是圆法的精神：用连续的分析手段去抓离散的整数。「第二版」「剑桥数学文库」说明这是一本经典老书的重印精装版。

哈罗德·达文波特是二十世纪真正伟大的数学家之一。本书基于他 1960 年代初在密歇根大学所作的讲座……全书有三大主题：华林问题以及对角型对整数的表示；多变量型方程组的整数可解性；以及对角不等式的整数可解性。

这是封底简介，点明全书三大主题。

这三句话其实就是全书的「三幕剧」，对应我们上面备齐的词汇，可以一一对上号：

- **第一幕（华林问题 + 对角型表示，第 1–10 章）**：把正整数 N 写成若干 k 次幂之和——这是对角型 $x_1^k + \cdots + x_s^k = N$ 的「表示问题」。这是圆法最经典的舞台。
- **第二幕（多变量型方程组，第 11–19 章）**：从「对角型」升级到「一般的型」（带交叉项），问 $F(x_1, \dots, x_n) = 0$ 有没有非平凡整数解。难度陡增，达文波特对三次型的攻关是其巅峰。
- **第三幕（对角不等式，第 20 章）**：把等号 $= 0$ 松绑成 $|\cdots|$ 不等式。这是达文波特与海尔布朗 1946 年开创的方向。

「达文波特是二十世纪真正伟大的数学家之一」不是客套：他在数论多个分支都有奠基性工作，本书正是他亲授讲义的整理，所以行文「惜墨如金、几无赘述」——这也是为什么我们这份讲解要替他把省略的细节补回来。

为本书第二版新增了一篇内容详尽的前言，由三位顶尖专家分别描述了相关领域的现代背景与最新进展；同时还补充了一份完整的参考文献。

说明第二版多了什么：三篇专家前言 + 完整参考文献。

这点很重要，它决定了你接下来读到的「前言」的**性质**：它不是达文波特本人写的，而是 1963 年原书出版四十多年后，三位当代专家（Vaughan、Heath-Brown、Freeman）回头写的「导读 + 最新进展综述」。所以前言里会大量出现「现在我们已知……」「后来某某改进到……」这类**把老讲义放进现代视角**的话。读前言时要意识到：达文波特讲义本身是 1960 年代的水平，而前言告诉你这六十年里学界走到了哪。

版权页信息：剑桥大学出版社……© 哈罗德·达文波特遗产管理委员会
1963, 2005……ISBN-13 978-0-521-60583-0 平装……第一版最初由
Campus Publishers 于 1963 年出版。

这是版权页的例行信息：出版社地址、版权年份、书号、初版信息。

这些是出版法律信息，与数学无关，扫一眼即可。值得一提的两点：版权年份「1963, 2005」对应初版（1963, Campus Publishers）与这一精装重印版（2005, 剑桥）；「ISBN」是国际标准书号，是每本书的唯一身份证号。英文原版版权页还列出了电子书号（eBook, NetLibrary, 978-0-511-08086-9），中译省去了，无关紧要。

目录：全书 20 章的地图Contents

这一段是目录，列出三篇前言、编者序和 20 章的标题与页码。

目录是你这趟旅程的**地图**。与其逐行抄页码，不如把 20 章按「三幕剧」重新归类，看清它们的逻辑链条——这正是读懂全书结构的关键。

第一幕·华林问题与对角型（第 1–10 章）

第 1 章**引论**给全书定调；第 2 章讲华林问题的**历史**；第 3 章是技术工具箱——**外尔不等式（Weyl）与华罗庚不等式（Hua）**，专门用来「压次弧」；第 4 章用圆法导出表法数的**渐近公式**（主项 + 误差）；第 5、6 章专攻主项里的**奇异级**

数；第 7、8、10 章把方法推广到带系数的对角方程 $c_1x_1^k + \cdots + c_sx_s^k = N$ 或 $= 0$ ；第 9 章是华林问题的核心量 $G(k)$ 。

第二幕·多变量一般型（第 11–19 章）

第 11 章伯奇定理（一般齐次方程）；第 12 章引入数的几何这件几何化工具；第 13–18 章是全书技术巅峰——三次型的逐步攻关（双线性方程→次弧优弧→奇异积分→奇异级数→ p -进问题），最终证明「16 个变量的三次型必有非平凡整数零点」；第 19 章把视野推到更高次型（伯奇的一般结果）。

第三幕·丢番图不等式（第 20 章）

单独一章，讲达文波特–海尔布朗方法，处理 $(|\lambda_1x_1^2 + \cdots + \lambda_sx_s^2|$

附录

参考文献（第 134 页起）与索引（第 139 页起）。

注意页码用罗马数字（vii, xi, xv, xix）的部分（前言与编者序）排在用阿拉伯数字编页的正文（第 1 章起）之前——这是西方学术书的惯例：正文前的「前置材料」单独用罗马数字编页，正文才从「第 1 页」重新数起。所以前言在「第 vii 页」，引论却在「第 1 页」，并不矛盾。

前言（一）·华林问题（第 1–10 章）Foreword — Waring's problem

撰文：R. C. Vaughan（宾夕法尼亚州立大学）

撰写人沃恩（Vaughan）是华林问题领域当今最权威的专家之一（他和伍利 Wooley 联手刷新了一大批 $G(k)$ 的纪录）。这篇前言的任务是：告诉你达文波特讲义在华林问题上没讲什么、以及六十年来学界把这些问题推到了哪一步。

达文波特撰写这些讲义之时，自达文波特与维诺格拉多夫约四分之一世纪以前的重要工作以来，华林问题上的进展寥寥；其主要旨趣在于报

告后面各章所述、关于型的更晚近的工作。事实上，至少就华林问题而言，当时普遍存在一种看法：哈代–李特尔伍德方法合理可得的一切结论都已被发掘殆尽……其后可在后面各章中加以发展，以服务于零的一般整型，尤其是三次型，的表示研究之需。

这段交代写作背景：1960年代初，人们以为圆法在华林问题上「榨干了」，所以讲义只把华林问题当作圆法的入门引子，重头戏放在后面的「型」。

- 「维诺格拉多夫 (Vinogradov)」：苏联数学家，圆法的另一位巨匠，1930年代用「维诺格拉多夫方法」大幅改进了华林问题（尤其大 k 时）。「四分之一世纪以前」指 1930 年代，相对讲义写作的 1960 年代初恰好约 25 年。
- 「该方法基本上已经被用尽」：这是当时的普遍误判——前言下一段会告诉你，过去二十年（1980 年代后）圆法在华林问题上又有了「相当多的进展」。把这句话记住，它是理解整篇前言情绪的钥匙：达文波特那一代以为故事讲完了，结果远没有。
- 「零的一般整型……的表示」：「零的表示」指研究方程 $F(\mathbf{x}) = 0$ 的非平凡整数解（把 0 表示成型的值）。这点明了全书真正的野心不在华林问题，而在后面的一般型。

书中没有论及达文波特本人的基础性工作，即 $G(4) = 16$ 、 $G(5) \leq 23$ 、 $G(6) \leq 36$ ，也没有论及维诺格拉多夫关于大 k 时 $G(k) \leq 2k \log k + o(k \log k)$ 的结果，或达文波特关于几乎所有自然数都是四个正立方数之和的证明。在更技术性的层面上，书中也未尝试给出引理 4.2 与 9.2 的更精细形式，即对优弧上母函数 $T(\alpha)$ 的估计……尽管这类精细化在应用中可能极为有用。

列举讲义「故意略去」的那些更深的结果，以提醒读者：讲义是入门，不是前沿。

逐个名词拆解（这里第一次密集出现 $G(k)$ 与 $\log$ 、 $o(\cdot)$ 等记号，回顾上面「必备词汇」）：

- $G(4) = 16$ ：每个充分大的整数都是 16 个四次幂之和，且 16 不能再少。这是达文波特的招牌成就之一，唯一被精确确定（等号）的 $G(k)$ ($k \geq 3$)。

- $G(5) \leq 23, G(6) \leq 36$: 当年达文波特给出的上界（注意是「 $\leq$ 」，只是上界，未必最优）。
- $G(k) \leq 2k \log k + o(k \log k)$: 维诺格拉多夫的「大 k 渐近上界」。这里 $\log k$ 是自然对数（增长极慢的函数）； $o(k \log k)$ 表示一个与 $k \log k$ 相比可忽略的尾巴。直觉： $G(k)$ 大约像 $2k \log k$ 那样增长——比 k 略快一点点，但远慢于 2^k 。这告诉我们 $G(k)$ 其实增长得相当「温和」。
- 「几乎所有自然数都是四个正立方数之和」：注意是「几乎所有」（除掉一个稀疏的例外集），不是「所有」；而且是正立方数（不准用负数或 0）。这是个比 $G(3)$ 更精细的结果。
- 「优弧上母函数 $T(\alpha)$ 的估计」：「优弧」即我们说的主弧 $\mathfrak{M}$ (major arcs, 本书译作「优弧」), 「母函数」即生成函数 $T(\alpha)$ 。意思是：引理 4.2、9.2 给出的是 $T(\alpha)$ 在主弧上的一个够用版本，但其实有人（达文波特–海尔布朗、华罗庚）给出了更精细的版本。「华罗庚」即中国数学家华罗庚 (Hua Loo-Keng), 第 3 章的「华氏不等式」就是他的工作。

在过去二十年里，华林问题取得了相当多的进展。受哈代与李特尔伍德、达文波特、维诺格拉多夫等人研究中某些思想之启发，人们发展出了一些极为灵活的方法……这些重要思想中某些萌芽的初露端倪，在引理 9.4 与 9.5 中可见一斑。

推翻上面那句「方法已用尽」的误判：近二十年（相对 2005 年，即 1980 年代以来）圆法又活了，而且新方法的种子在讲义的引理 9.4、9.5 里已能看到。

这段在说一件励志的事：被宣判「死亡」的方法，因为几个灵活的新思想（主要来自沃恩、伍利）**起死回生**。而且这些新思想并非凭空出现，它们的「萌芽」就藏在达文波特讲义第 9 章的两条引理里——这也是为什么值得重读这本老讲义：经典里埋着未来的种子。

当 $s \geq 2^k + 1$ 时，定理 4.1 所建立的、将大自然数 n 表为至多 s 个 k 次幂之和的表法数的渐近公式，对 $3 \leq k \leq 10$ 曾是前沿水平；但对更大的 k ，维诺格拉多夫所采用的方法更为优越。目前的进展状况是：已知渐近公式在 $s \geq 2^k$ ($k = 3, 4, 5$)、 $s \geq 7 \cdot 2^{k-3}$ ($k = 6, 7, 8$)，以

及 $s \geq s_1(k)$ (其中 $s_1(k) = k^2(\log k + \log \log k + O(1))$, $k \geq 9$) 时成立。

报告「渐近公式成立需要多少个变量 s 」这条战线的最新纪录, 并把它和讲义里的定理 4.1 ($s \geq 2^k + 1$) 对比。

「**渐近公式 (asymptotic formula)**」是什么 就是「主项 + 可忽略误差」形式的精确计数公式, 长得像 $r(N) = (\text{主项}) + o(\text{主项})$ 。它比「 $r(N) > 0$ (只说有解)」强得多: 它告诉你解的**确切数量级**。要得到渐近公式, 需要变量个数 s 足够大, 让主弧的主项压过次弧的误差。所以这段全在比较「 s 至少要多大, 渐近公式才成立」。

- $s \geq 2^k + 1$ (**定理 4.1**): 达文波特讲义给的门槛。 2^k 是 2 的 k 次方, 增长很快 ($k = 10$ 时已是 1024)。这个门槛在 $3 \leq k \leq 10$ 曾是最好的。
- 新纪录把门槛压低了: 对 $k = 3, 4, 5$ 只要 $s \geq 2^k$ (少了 1 个变量, 看似微小, 实则极难); 对 $k = 6, 7, 8$ 压到 $7 \cdot 2^{k-3}$ 。
- $s_1(k) = k^2(\log k + \log \log k + O(1))$: 大 k 时的门槛。注意它是 k^2 量级 (k 的平方乘上对数因子), $O(1)$ 表示一个有界的常数项。 $\log \log k$ 是「对数的对数」, 增长慢到几乎是常数。直觉: **渐近公式所需的 s 大约是 $k^2 \log k$ 量级**。(这比后面「只要有解」所需的 $G(k) \sim k \log k$ 大一个 k 因子——要算准比只要有解更费变量, 合情合理。)

第 3 章中的注里关于 $k = 3$ 情形的讨论至今仍有现实意义。尽管八个立方数之和的渐近公式现已确立, 但在指数 \2

聚焦最经典的立方数 ($k = 3$) 情形: 八个立方数的渐近公式虽已证明, 但中间指数处的关键估计仍没本质改进, 并给出沃恩费尽心力才得到的一个次弧积分上界。

逐个零件解释这个行间公式 (这是本前言第一个真正的数学式, 慢讲):

- $\int_m |T(\alpha)|^8 d\alpha$: 在次弧 m (这里译作「劣弧」, 即 minor arcs, 与「优弧/主弧」相对) 上, 对 $|T(\alpha)|$ 的8次方积分。8次方对应「8个立方数之和」($s = 8$, 因为 T^s 展开后正是 s 个数之和)。 $|T(\alpha)|$ 是复数 $T(\alpha)$ 的模长 (到原点的距离)。这个积分衡量「次弧贡献的噪声总量」。
- $\ll P^5(\log P)^{-\gamma}$: 右边是上界。前面讲过 $\ll$ 就是「至多常数倍」。 P^5 是主项的量级 (这里 $s - k = 8 - 3 = 5$, 恰好是 $P^{s-k} = P^5$)。关键在那个 $(\log P)^{-\gamma}$: $\gamma > 0$, 所以 $(\log P)^{-\gamma}$ 是一个小于1、随 P 增大缓慢趋于0的因子。它的作用是: 让次弧积分比主项 P^5 还小那么一丁点 (小一个对数因子) ——刚好够让噪声淹不过信号, 从而证出渐近公式。
- 「凸性界 (convexity bound)」与「指数 $\backslash(2$: 这里 m 指对 $|T|$ 取的方次 ($\int |T|^m$)。在 $m = 2$ 和 $m = 4$ 两个端点处, 积分有「平凡的、靠直接展开就能得到」的精确值 (凸性来自这两个端点把中间夹住); 而 $\backslash(2$ 这一段仍是未被攻克的硬骨头。

然而 Hooley 在某个哈塞-韦伊 L -函数的 (尚未证明的) 黎曼假设之下证明了

$$\int_m |T(\alpha)|^6 d\alpha \ll P^{3+\varepsilon}$$

这进而蕴含七个立方数之和的渐近公式。遗憾的是, 目前甚至连该 L -函数是否能解析延拓到临界带内都尚不清楚。

介绍一个「有条件」的更强结果: 胡利 (Hooley) 在某个尚未证明的大假设下, 把六次方的次弧积分压到了 $P^{3+\varepsilon}$, 从而推出「七个立方数」的渐近公式; 但代价是这个假设本身还遥不可及。

- $\int |T|^6 \ll P^{3+\varepsilon}$: 现在是6次方 (对应7个立方数里留6个在次弧), $s - k = 6 - 3 = 3$, 所以 P^3 是主项量级, P^ε 是那条「微小可忽略尾巴」。能把次弧积分压到 $P^{3+\varepsilon}$ 是极强的结论。

- 「 L -函数」「黎曼假设」「解析延拓」「临界带」：这些是解析数论的高深对象，高中完全不必掌握细节，只需理解一句话——胡利的结果建立在一个至今无人能证明的大猜想之上（黎曼假设的某个推广）。所以它是「有条件的结果」：如果那个猜想对，则七个立方数的渐近公式成立。末句更悲观：连这个 L -函数能不能「延拓」到关键区域都还不知道，意味着这条路目前走不通。这教给你一个真实的数学现状：很多漂亮结论是悬在未证猜想之上的「如果……那么……」。

对于 $G(k)$ ，我们目前已知的最佳结果有： $G(3) \leq 7$ 、 $G(4) = 16$ 、 $G(5) \leq 17$ 、 $G(7) \leq 33$ 、 $G(6) \leq 21$ 、 $G(8) \leq 42$ 、 $G(9) \leq 50$ 、 $G(11) \leq 59$ 、 $G(12) \leq 76$ 、……、 $G(20) \leq 142$ ，以及一般地 $G(k) \leq s_2(k)$ ，其中 $s_2(k) = k(\log k + \log \log k + O(1))$ 。

这是华林问题的「成绩单」：把当前已知的 $G(k)$ 最佳上界一个个列出来，并给出大 k 的统一上界公式。

订正一处译文笔误 对照英文原书，这一串数字里 $k = 10, 11$ 处中译有抄写错位。原书是：

$G(9) \leq 50$, $G(10) \leq 59$, $G(11) \leq 67$, $G(12) \leq 76, \dots$ 。中译漏写了 $G(10) \leq 59$ ，并把 $G(11)$ 误写成 ≤ 59 （应为 ≤ 67 ）。请以英文原书为准。这类「成绩单」式的数字只需领会趋势，不必背。

怎么读这张成绩单（不必记数字，要记规律）：

- 只有 $G(4) = 16$ 用了等号（精确值已知），其余都是 $\leq$ （只知上界，真值可能更小）。这说明 $G(k)$ 的精确值至今几乎全是未解之谜，连 $G(3)$ 都只知道在 4 和 7 之间。
- 这些数字（7, 16, 17, 21, 33, 42, ..., 142）大致随 k 线性偏快地增长。统一上界 $s_2(k) = k(\log k + \log \log k + O(1))$ 告诉你： $G(k)$

大约是 $k \log k$ 量级——和前面渐近公式所需的 $k^2 \log k$ 相比，「只要有解」省了一个 k 因子。这是合理的：要证「有解」比要证「有几个解」容易。

- 这些纪录大多由 Vaughan 与 Wooley（本前言作者及其合作者）创造，方括号里的 [87][88][89][90] 是参考文献编号（指向书末文献表）。

设 $G^\#(4)$ 表示满足如下性质的最小正整数 s ：只要 $1 \leq r \leq s$ ，模 16 余 r 的剩余类中每个充分大的 n 都是 s 个四次幂之和。事实上，达文波特曾证明 $G^\#(4) \leq 14$ ，而我们现在能证明 $G^\#(4) \leq 12$ 。林尼克关于立方数华林问题的工作并不使用哈代–李特尔伍德方法，而是基于三元二次型理论。

引入一个 $G(k)$ 的「精细变体」 $G^\#(4)$ ，并提到林尼克 (Linnik) 用了完全不同的（非圆法）路线处理立方数。

$G^\#(4)$ （读「G sharp of 4」） $\#$ 是「井号 / 升号 sharp」。它和 $G(4)$ 的区别在于更尊重四次幂的同余障碍。背景：四次幂模 16 只能取 0 或 1（因为偶数四次方被 16 整除，奇数四次方 $\equiv 1$ ）。所以「 n 是 s 个四次幂之和」会受到「 n 模 16 是几」的强烈限制——某些剩余类天生需要更多项。 $G^\#(4)$ 把这种逐个剩余类的要求纳进定义：要求对 $1 \leq r \leq s$ 的每一个剩余类 $r \bmod 16$ 都成立。它是一个比 $G(4)$ 更贴合「真实局部障碍」的量。达文波特证到 ≤ 14 ，沃恩改进到 ≤ 12 。你不必深究，只需领会：四次幂这个例子里，模 16 的同余信息（局部信息）实实在在地影响着需要多少项——这正是「奇异级数 / 局部条件」在具体例子中的体现。

「林尼克 (Linnik) ……基于三元二次型理论」：值得记住的一点——并不是所有华林问题都靠圆法。林尼克处理 $G(3) \leq 7$ 用的是「三元二次型」（三个变量的二次型）的算术理论，是一条完全不同的路。这提醒你：圆法虽强，但不是唯一武器。

沃森给出了一个类似但更简单的证明。然而这些证明所给出的关于七个立方数之和表法数的信息相对较弱。……对于 $G(k)$ 的正确值应当是多少，达文波特未给出任何提示。最简单的猜测是

$$G(k) = \max\{k + 1, \Gamma(k)\}$$

其中 $\Gamma(k)$ 的定义见定理 5.1 之前那一段。这将蕴含：当 $k \geq 3$ 时，若 $k = 2^l$ 则 $G(k) = 4k$ ；而当 $k \neq 2^l$ 时 $k + 1 \leq G(k) \leq \frac{3}{2}k$ 。

指出非圆法证明的弱点（给不出表法数的好下界），并给出关于 $G(k)$ 真值的「最简单猜想」。

- 「信息相对较弱」：林尼克/沃森的方法能证「有解」，但说不清「有多少种写法」。圆法的优势恰在这里——它天生给出表法数 $r(N)$ 的渐近公式，所以即便结论相同，圆法版本的证明「含金量」更高。
- $\max\{a, b\}$ ：取 a, b 里较大的那个。读「max of a and b 」。
- $\Gamma(k)$ ：这里是本书自定义的一个量（注意不是大学里的「伽马函数」，别混淆！它的精确定义在第 5 章定理 5.1 前），来自同余障碍——它算的是「为了在所有素数模下都有解，理论上至少需要多少项」，本质上就是奇异级数非零所要求的下限。读「大 Gamma of k 」。
- $k = 2^l$ ： k 恰是 2 的某个幂（如 $k = 4, 8, 16$ ）。这种 k 的同余障碍特别强（像上面四次幂模 16 那样），导致猜想值 $G(k) = 4k$ 偏大；其余 k 障碍弱， $G(k)$ 落在 $k + 1$ 到 $\frac{3}{2}k$ 之间。这条猜想的精神是： $G(k)$ 的真值由「同余障碍 $\Gamma(k)$ 」和「最起码要 $k + 1$ 个变量」两者中较大的那个决定。「至今无人证明」。

关于引理 9.2 以及该证明之后的注，我们现在已知：在更弱的假设 $(q, a) = 1$ 、 $q|\beta| \leq \frac{1}{2k}P^{1-k}$ 、 $\alpha = \beta + a/q$ 之下，我们有更强的估计

$$T(\alpha) = q^{-1} S_{a,q} I(\beta) + O(q^{\frac{1}{2}+\varepsilon}).$$

而且仅在假设 $(q, a) = 1$ 之下，我们就有

$$T(\alpha) = q^{-1} S_{a,q} I(\beta) + O(q^{\frac{1}{2}+\varepsilon} (1 + P^k |\beta|)^{\frac{1}{2}}).$$

……后一结果使得我们能够对立方数情形作出处理，其中所有的弧都是优弧。

给出讲义引理 9.2 的「升级版」：在更弱的条件下得到对 $T(\alpha)$ 更精确的主弧近似公式。

这是本前言里技术性最强的一段，但它的**结构**其实就是我们反复说的「主弧上 $T(\alpha)$ 有漂亮近似」。逐个零件拆：

- $(q, a) = 1$ ：表示整数 q 与 a **互素**（最大公因数为 1）。这保证分数 a/q 是既约的。读「 q 与 a 的最大公约数是 1」。
- $\alpha = \beta + a/q$ ：把 α 写成「一个小分母分数 a/q 」加上「一个很小的偏移 β 」。这正是「 α 落在围着 a/q 的主弧里」的标准写法—— β 衡量 α 离分数中心有多远。
- $q|\beta| \leq \frac{1}{2^k} P^{1-k}$ ：限制偏移 β 多小（即主弧多窄）。「更弱的假设」指的就是：允许 β 比讲义里更大一些，主弧可以更宽，结论却仍然成立——这正是「升级」的含义。
- $q^{-1} S_{a,q} I(\beta)$ ：近似公式的主项，恰好分解成两个因子： $S_{a,q}$ 是一个**有限的指数和**（叫「高斯型和 / Gauss sum」，纯粹来自同余信息，是奇异级数的「砖块」）， $I(\beta)$ 是一个**积分**（来自实数侧，是奇异积分的「砖块」）。看，主弧上的 $T(\alpha)$ 自然分裂成「**同余因子** × **实数因子**」——这正是后面奇异级数 $\mathfrak{S}$ 与奇异积分 $\mathfrak{J}$ 的源头！
- $O(q^{1/2+\varepsilon})$ ：误差项，量级约 $\sqrt{q}$ （再带个可忽略的 P^ε 尾巴，这里写成 q^ε ）。 q 越小（主弧越「正中」），误差越小，近似越准——这正符合「主弧中心附近 T 最听话」的直觉。
- 「**所有的弧都是优弧**」：第二个更弱的公式（只要 $(q, a) = 1$ ）威力大到：对立方数（ $k = 3$ ），可以让整个 $[0, 1]$ 都当主弧来处理，根

本不需要单独对付难缠的次弧。这是一个方法论上的飞跃——把最难的次弧问题彻底绕开了。

关于哈代–李特尔伍德方法应用于华林问题的现代入门及一些较晚近进展，见 Vaughan [86]；关于华林问题的综述，见 Vaughan and Wooley [91]。

给出进一步阅读的两个权威参考。

这是「延伸阅读指引」：[86] 是沃恩的名著《The Hardy–Littlewood Method》（圆法的现代标准教材），[91] 是华林问题的综述文章。方括号数字都指向书末参考文献。对你而言，记住「想系统学圆法，去看 Vaughan [86]」即可。

第 7 章讨论的是：给定一个自然数序列 $\{c_j\}$ ，方程

$$c_1 x_1^k + \cdots + c_s x_s^k = N \quad (1)$$

对大自然数 N 的可解性。……然而 s 须取多大方有解，其障碍纯粹是局部的。……在对 s 加上相应条件后，便能得到对所数解的个数的近似公式。这将对任意大的、其奇异级数有界离 0 的 N ，给出解的个数的一个正下界。

介绍第 7 章：把华林问题推广成「带系数 c_j 的对角方程 (1)」，并指出可解的唯一障碍是「局部障碍」，由奇异级数是否离 0 来判定。

- $\{c_j\}$ ：花括号表示「一系列数」 $c_1, c_2, \dots, c_s$ ，这里是给定的正整数系数。方程 (1) 比华林问题（所有系数都是 1）更一般。
- 「障碍纯粹是局部的」：这是本书的中心思想之一，回顾「局部到整体」那一节——意思是：方程 (1) 会不会无解，**完全由局部条件（实数侧 + 各素数模侧）决定**，没有别的隐藏障碍。只要局部都通过（即奇异级数 $\mathfrak{S} > 0$ 且离 0），整数解就一定存在，而且数量可观。

- 「正下界」：不仅证明有解，还证明解的个数 $\geq$ 某个正的量级（约 $\mathfrak{S} \cdot \mathfrak{J} \cdot P^{s-k}$ ）。这比「至少有一个解」强得多。

达文波特简要勾勒了为使方法适应这一情形而须对论证作出的细微改动，本章其余部分则致力于证明：上述关于奇异级数的条件，本质上等价于所期望的局部可解性条件。

说明第 7 章后半的核心任务：证明「奇异级数离 0」与「局部可解」其实是同一件事。

这句话把「奇异级数 $\mathfrak{S}$ 」这个看起来很分析、很吓人的无穷乘积，和「每个素数模下方程有解」这个很初等、很算术的条件划上了等号。这是圆法里一个反复出现的美妙对应：分析对象（奇异级数） $\iff$ 算术条件（局部可解性）。理解了这一点，你就抓住了奇异级数存在的全部意义——它是局部信息的「分析化身」。

在第 8 章与第 10 章中，达文波特将该方法调整以处理

$$c_1 x_1^k + \cdots + c_s x_s^k = 0 \quad (2)$$

此时 c_j 可为整数，且当 k 为偶数时不必全部同号。当然此方程总有解，因此主要旨趣在于确立存在并非所有 x_j 都为 0 的整数解。……在第 8 章中……这要求 s 取相当大的值方能保证有解。在第 10 章中，借助第 9 章处理华林问题所用的维诺格拉多夫论证之变体，这一要求得到了一定程度的放宽。

介绍第 8、10 章：把右边从 N 换成 0，问题变成「求齐次对角方程 (2) 的非平凡整数解」；第 8 章用最朴素的圆法（需要很多变量），第 10 章用维诺格拉多夫技巧把变量需求降下来。

- 为什么右边是 0 反而更难？方程 (2) 显然总有 $(0, 0, \dots, 0)$ 这个解（平凡解），所以「有没有解」是废话；真问题是**有没有不全为 0 的整数解**。这就把焦点从「数解的个数」转到「排除平凡解后还剩没剩」，这正是全书第二、三幕（一般型）的核心问法。

- 「 k 为偶数时不必全部同号」：若 k 是偶数（如 $k = 2$ ）， $x^k \geq 0$ 恒成立，若所有 c_j 同号则 $\sum c_j x_j^k = 0$ 只能全取 0（无非平凡解）。所以必须允许 c_j 有正有负，正负项才能相互抵消出非平凡解。这是个很自然的「符号条件」。 k 为奇数时 x^k 本身可正可负，就不需要这个条件。
- 第 8 章 vs 第 10 章：朴素圆法（第 8 章）门槛高（需要大 s ）；引入维诺格拉多夫的精细技巧（第 10 章，源自第 9 章）后门槛降低。这就是「方法越精细，需要的变量越少」的典型权衡。

尽管第 10 章的论证相对简单，但从哲学角度看它存在缺陷：它和方程 (1) 的局部可解性一样，需要讨论当 N 非零时方程 (2) 的局部可解性——而这当然本不应当需要。……达文波特与刘易斯曾证明 $k^2 + 1$ 个变量足以使奇异级数有界离 0，而当 $k + 1$ 为素数时，存在 k^2 个变量却无非平凡解的方程。

指出第 10 章方法的一个「哲学瑕疵」，并给出「保证 (2) 有非平凡解需要多少变量」这个独立问题的精确答案： $k^2 + 1$ 够、 k^2 不一定够。

- 「哲学缺陷」：第 10 章证 $= 0$ 的方程时，技术上却被迫去讨论一个 $N \neq 0$ 的辅助方程的局部可解性。这在逻辑上「绕了远路」——按理说证 $= 0$ 不该牵扯 $\neq 0$ 。作者诚实地指出这是个不优雅之处（虽然能补救，但要更多细节）。这教你一种**数学审美**：能证出结论还不够，证明的逻辑结构是否「干净」也被讲究。
- $k^2 + 1$ 够、 k^2 不够（当 $k + 1$ 为素数）：这是一对漂亮的「上界 + 下界」夹逼。 $k^2 + 1$ 个变量足以保证奇异级数离 0（从而有解）；而当 $k + 1$ 恰是素数时，能构造出 k^2 个变量却没有非平凡解的反例。两者只差 1 个变量，说明 $k^2 + 1$ 这个门槛几乎是最优的、**卡得死死的**。这种「上界恰好等于下界加一」的结果，是数论里最令人满意的一类。「刘易斯（Lewis）」是达文波特的长期合作者。

他们还借助哈代-李特尔伍德方法证明：当 $s \geq k^2 + 1$ 且 $k \leq 6$ 或 $k \geq 18$ 时，(2) 可解。后来 Vaughan、[83]、[85]

消除了这一间隙。当将与华林问题相关的 Vaughan 与 Wooley 方法加以调整后可以表明：只要相应的奇异级数有界离 0……所需变量数远少于此即可保证 (2) 有非平凡解。

补充进展：达文波特-刘易斯的结果原本在 $7 \leq k \leq 17$ 有「空隙」，后被沃恩等人填满；而更现代的方法能把变量需求压得比 $k^2 + 1$ 还低。

这段是典型的「填空隙」叙事：一个定理最初只对某些 k ($k \leq 6$ 或 $k \geq 18$) 成立，中间一段 k (7 到 17) 悬而未决，后来由 Vaughan [81] ($11 \leq k \leq 17$)、[83] ($7 \leq k \leq 9$)、[85] ($k = 10$) 逐段补齐。最后一句再次强调主旋律：只要奇异级数离 0 (局部可解)，更精细的现代方法所需的变量远少于 $k^2 + 1$ 。「局部可解 $\Rightarrow$ 整体可解」这条线索贯穿始终。

在后面各章中，哈代-李特尔伍德方法被以各种、有时相当精巧的方式加以调整。然而，前 10 章主要结果被直接应用的唯一之处，是第 11 章伯奇定理证明中对定理 8.1 (或定理 10.1) 的使用。后来伯奇给出了一个完全初等的证明……

承上启下：前 10 章 (华林问题) 的成果，主要是为第 11 章伯奇定理铺路；而伯奇后来还有一个不依赖圆法的初等证明。

这是第一篇前言的收尾，它把「第一幕」和「第二幕」缝合起来：原来前 10 章辛辛苦苦建立的对角方程理论 (定理 8.1、10.1)，唯一的「**直接用武之地**」是第 11 章证明伯奇定理 (关于一般齐次方程)。这告诉你全书的逻辑依赖关系——华林问题是热身，一般型才是目标。末句又补一刀：伯奇本人后来用林尼克的初等方法给出了替代证明，连圆法都可以不用。这再次提醒：同一个定理往往有多条证明路径，圆法只是其中一条 (虽然常常是给信息最多的一条)。

前言（二）· 多变量型（第 11–19 章）Foreword — Forms in many variables

撰文：D. R. Heath-Brown（牛津大学）

撰写人希思–布朗（Heath-Brown）是解析数论的顶尖人物。这一篇导读对应全书「第二幕」：从对角型升级到**一般的型**（带交叉项的齐次多项式），这是达文波特讲义真正的技术巅峰所在。

设 $F(x_1, \dots, x_n)$ 为一个整系数的 d 次型。当 $d \geq 3$ 时，方程 $F(x_1, \dots, x_n) = 0$ 是否有非平凡整数解的问题极为自然、极为一般，也极为困难。然而对二次型，哈塞–闵可夫斯基定理给出了完整答案：它断言非平凡解存在，当且仅当在 $\mathbb{R}$ 中以及在每个 p -进域 $\mathbb{Q}_p$ 中都存在这样的解。

提出第二幕的总问题——一般 d 次型 $F = 0$ 有没有非平凡整数解；并指出二次型（ $d = 2$ ）已被哈塞–闵可夫斯基定理完美解决。

- $F(x_1, \dots, x_n)$: n 个变量、次数为 d 的型（齐次多项式），整系数。 $d \geq 3$ 是难的开始—— $d = 1$ （线性）小学就会， $d = 2$ （二次）已被完全解决， $d \geq 3$ 才是雷区。
- 「当且仅当」：读「iff」，表示「 $\iff$ 」（充要条件，两边等价）。
- **哈塞–闵可夫斯基定理（Hasse–Minkowski）**：二次型的完美局部到整体定理。回顾上面「局部到整体」那节——它说：二次型 $F = 0$ 有非平凡（有理）解 $\iff$ 它在实数 $\mathbb{R}$ 里有解、并且在每个 p -进域 $\mathbb{Q}_p$ 里都有解。也就是：**所有局部关卡都过 $\iff$ 整体有解**。这是数论里最干净漂亮的定理之一，是「局部到整体」原理的黄金样板。本幕的全部努力，就是问：对 $d \geq 3$ ，这个黄金样板还能不能（部分地）保住？

对于更高次的型，已知这一结论不成立，正如塞尔默的例子

$$3x_1^3 + 4x_2^3 + 5x_3^3 = 0$$

所表明的那样。尽管如此，仍有希望：若变量数不太小，我们应当依然拥有某种哈塞-闵可夫斯基定理那样的「局部到整体」原理。

给出一个著名反例：塞尔默的三次型在所有局部关卡都有解，整体却没有非平凡解——所以局部到整体对 $d = 3$ 不成立。但作者立刻给出希望：变量够多时也许能挽回。

- 塞尔默 (Selmer) 的例子 $3x_1^3 + 4x_2^3 + 5x_3^3 = 0$ ：这是一个 3 变量的三次型。它通过了所有局部关卡（在 $\mathbb{R}$ 和每个 $\mathbb{Q}_p$ 里都有非平凡解），但没有非平凡的有理数解。它一锤定音地证明：对三次型，「局部都有解」并不能保证「整体有解」，哈塞-闵可夫斯基的黄金定理彻底失效。这是数论史上极重要的反例。
- 「若变量数不太小……依然拥有」：这是整个第二幕的战略希望。塞尔默的反例只有 3 个变量（很少）。作者的赌注是：当变量个数 n 相对次数 d 足够大时，局部到整体原理会「自动恢复」。后面所有章节（伯奇定理、达文波特的三次型 16 变量结果）都在兑现这个希望。

事实证明：若变量数相对于次数足够大，则 p -进条件会自动成立。这是布劳尔所证明的……其攻击路线采用多重嵌套归纳，因而所需的变量数极大。阿廷曾猜测 $d^2 + 1$ 个变量总是足够的……然而其后人们发现了许多反例。其中第一个由特尔贾尼安给出，涉及一个 18 个变量的四次型，它没有非平凡的 2-进解。

先解决「局部关卡」本身：变量够多时， p -进解会自动存在（布劳尔），阿廷猜测门槛是 $d^2 + 1$ ，但这个猜测被反例推翻了。

- 「 p -进条件会自动成立」：好消息——你甚至不必逐个检验 p -进关卡，只要变量足够多，这些局部关卡自动全过。这是布劳尔

(Brauer) 1945 年的成果,「处理此类问题的第一个一般方法」。

- 「**多重嵌套归纳**」: 布劳尔的证明用了一层套一层的数学归纳法, 逻辑上有效但极其浪费, 导致它给出的「足够变量数」是个天文数字 (远大于实际所需)。这是「方法能行但效率极低」的典型。
- **阿廷 (Artin) 猜测 $d^2 + 1$** : 阿廷猜「 $d^2 + 1$ 个变量永远够让 p -进解存在」, 且这个 $d^2 + 1$ 几乎最优 (容易造出 d^2 变量却只有平凡 p -进解的型)。 d^2 是 d 的平方。这个猜测很优美。
- **特尔贾尼安 (Terjanian) 的反例**: 然而阿廷猜测错了——特尔贾尼安造出一个 18 变量的四次型 ($d = 4$, $d^2 + 1 = 17$, 所以 18 个变量按猜测「绰绰有余」), 它却没有非平凡的 2-进解 (连 $p = 2$ 的局部关卡都过不了)。这推翻了阿廷猜测, 展示了 $p = 2$ (最小素数) 常常是最难缠的特例。「2-进」就是 $p = 2$ 时的 p -进。

目前尚无涉及素数次型的已知反例, 在这种情形下, 阿廷猜想是否成立仍是一个未决问题。……达文波特在第 18 章给出了其中之一, 针对 $d = 3$ 这一最简单的情形, 确立了当 $n \geq 10$ 时 p -进解的存在性。对于 $d \geq 4$, 此类途径仅在 p 足够大时方才奏效。……阿克斯与科亨证明了: 当 p 相对于 d 足够大时, $d^2 + 1$ 个变量总足以解决 p -进问题。阿克斯-科亨的证明因其使用数理逻辑方法而引人注目。

交代 p -进问题的现状: 素数次型情形阿廷猜测尚未判定; 达文波特在第 18 章解决了 $d = 3$ 的最优结果 ($n \geq 10$); 大素数情形阿克斯-科亨用数理逻辑证明了阿廷猜测。

- 「**素数次型**」: 次数 d 本身是素数 (如 $d = 3, 5, 7$) 的型。特尔贾尼安的反例次数 $d = 4$ (合数)。对素数次型, 至今**没有反例**, 所以阿廷猜测对它们可能仍对——这是个公开问题。
- **达文波特第 18 章**: $d = 3$ 时 $n \geq 10$ 保证 p -进解存在, 且 10 是最优 (存在 9 变量无非平凡 p -进零点的三次型)。这是全书第 18 章的主结果。又一个「上界恰好卡在最优」的漂亮结果。

- **阿克斯-科亨 (Ax-Kochen) 用数理逻辑**：这是个传奇——他们证明阿廷猜测「对足够大的 p 成立」，用的竟是**数理逻辑 (模型论)** 而非数论本身的工具！这告诉你数学不同分支会以意想不到的方式互相帮忙。代价是「足够大的 p 」没有明确界，小素数仍需另想办法（下段）。

对于小素数，似乎需要另辟蹊径，而 Wooley 重新审视了布劳尔归纳途径，从而确立了 $d^{2^d} + 1$ 个变量对每个域 $\mathbb{Q}_p$ 都足够。在 $d = 4$ 与 $d = 5$ 的情形得到合理大小的界（比如 1000 以下），仍是一个重要的未决问题。

小素数情形由伍利改进布劳尔方法解决，但给出的界 $d^{2^d} + 1$ 大得离谱；把它降到「1000 以下」仍未做到。

- $d^{2^d} + 1$ ：注意这是「 d 的（2 的 d 次方）次方」，是**双重指数**，大得吓人—— $d = 4$ 时 $2^4 = 16$ ，于是 $4^{16} \approx 40$ 亿。这说明伍利的结果虽然「对所有素数都成立、无条件」，但界极不实用。这是数论的常态：「无条件但界巨大」与「界小但有条件」之间常需取舍。
- 「把 $d = 4, 5$ 的界降到 1000 以下仍未做到」——一个具体、诚实、悬而未决的小目标。

型在 $\mathbb{Q}$ 上（而非 $\mathbb{Q}_p$ 上）的问题则截然不同。对于偶数次的型，不存在能保证非平凡整数解存在的 n 值，正如例子

$$x_1^d + \cdots + x_n^d = 0$$

所示。……然而伯奇能够将这一归纳途径仅就奇数次的型加以调整，从而证明：对任意奇整数 $d \geq 1$ ，存在相应的 $n(d)$ ，使得 $F(x_1, \dots, x_n) = 0$ 在 $n \geq n(d)$ 时总有非平凡解。

转到真正想要的「整体（有理 / 整数）解」问题：偶数次型天生没救（举反例），但奇数次型有救——这就是第 11 章的伯奇定理。

- **为什么偶数次型没救？** 看反例 $x_1^d + \cdots + x_n^d = 0$ ：若 d 是偶数，每个 $x_i^d \geq 0$ ，要它们的和为 0 就只能全都为 0（平凡解），不论 n 多大都凑不出非平凡解。所以对偶数次型，再多变量也无法保证非平凡整数解——这是一个无法逾越的「符号障碍」（和前面第 8 章那个符号条件同源）。
- **奇数次型有救（伯奇定理）：** d 为奇数时 x^d 可正可负，能相互抵消。伯奇（Birch）证明：对每个奇数 d ，存在一个门槛 $n(d)$ ，只要变量数 $n \geq n(d)$ ，奇数次型 $F = 0$ 必有非平凡整数解。这就是第 11 章的核心定理，也兑现了第二幕开头那个「变量够多就有救」的希望（至少对奇数次）。 $n(d)$ 是「依赖 d 的一个门槛函数」。

这一工作由达文波特在第 11 章中加以描述。……尽管伯奇工作所产生的 $n(d)$ 值大到无法写出，Wooley 通过对伯奇途径的仔细调整，给出了更合理的估计。

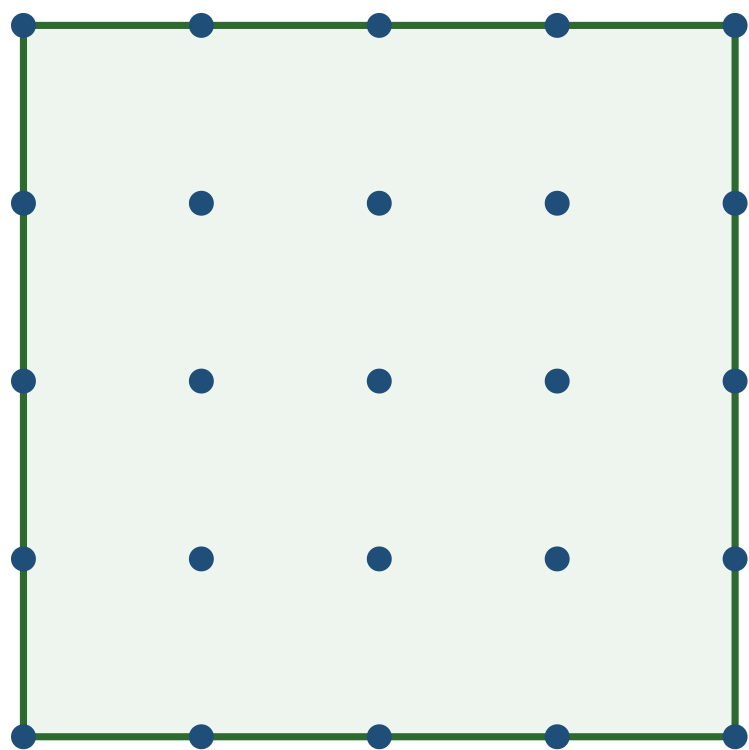
补充：伯奇定理的门槛 $n(d)$ 原本大到写不出来，后被伍利改良到较合理。

又是「定理成立但界巨大、后人改良」的故事。伯奇的归纳证明（和布劳尔一样多重归纳）给出的 $n(d)$ 大到无法写下，伍利 [99] 仔细优化后给出了可写出的合理估计。读到这里你应该已经熟悉这个反复出现的模式了：先有人证明「存在某个门槛」，再有人把门槛压到实用。

达文波特本人对该领域的一大贡献，是借助圆法对三次型的攻关。外尔方法的自然应用——如第 13 章所述——会导出一个涉及双线性型的丢番图不等式组；这一背景下的关键结果是引理 13.2。借助数的几何的技巧，达文波特得以将这些不等式转化为方程。

进入本幕高潮：达文波特用圆法攻三次型。技术链条是——外尔方法 → 双线性不等式组（第 13 章，关键是引理 13.2）→ 用「数的几何」把不等式变成方程。

- 「外尔方法 (Weyl's method)」：处理指数和 $T(\alpha)$ 的一套基本技巧（第 3 章「外尔不等式」就是它），核心动作是「平方放大、错位相减」，把高次的指数和降次。把它用到三次型上，会自然冒出「双线性型」。
- 「双线性型 (bilinear form)」：形如 $\sum a_{ij}x_iy_j$ 的式子——对 x 是线性的、对 y 也是线性的（各自一次，合起来二次）。三次型经外尔方法降次后，导数结构里会出现双线性型。
- 「数的几何 (geometry of numbers)」：闵可夫斯基开创的一门学问（第 12 章），研究格点（整数坐标的点）与凸体的关系，核心是「足够大的对称凸体里一定含有非零格点」。达文波特用它把「一组双线性不等式」翻译成「一组方程」——这是关键的桥梁，因为不等式松散难控，方程才好下手。



边长 P 的方盒 ($|x_i| \leq P$)

蓝点 = 整数坐标的格点

「数的几何」研究格点（整数坐标的点）与几何区域的关系。圆法里反复说的「边长为 P 的方盒中的解」，就是数所有坐标都满足 $|x_i| \leq P$ 的整点。格点越密、方盒越大，里面的解越多。

在他关于此主题的头两篇论文中，这些方程被用来推出 F 必表示一个形如 $a_1x_1^3 + F'(x_2, \dots, x_m)$ （对某 m ）
讲达文波特方法的演进：前两篇论文「化归到对角型」但浪费变量；第三篇用更几何的处理，不浪费变量、效率高得多（这正是第 14 章的内容）。

- 「 F 必表示形如 $a_1x_1^3 + F'(x_2, \dots, x_m)$ 的型」：意思是能把三次型 F 化简，剥出一个单独的纯立方项 $a_1x_1^3$ ，剩下一个变量更少的三次型 F' 。反复这么剥，最终把一般三次型化归到**对角**三次型（只剩纯立方项），从而能套用前 10 章的对角型理论。
- 「**浪费 $n - m$ 个变量**」：每剥一次都要扔掉一批变量（从 n 个降到 m 个）。变量是「资源」，扔得越多，最后能证的结论越弱（需要起始变量更多）。第三篇论文的革新就在于一个**变量都不浪费**——这是把「16 变量」这个惊人结果做出来的关键。这体现了一个深刻的方法论：**在圆法里，变量是稀缺资源，谁省得多谁就赢。**

这第三种方法的直接应用表明：对任意 17 个或更多变量的三次型 $F = 0$ 有非平凡整数解，这正是定理 18.1 给出的结果。然而在 [22] 中通过一处轻微的精细化表明 16 个变量即足够。……其主要目标是为边长为 P 的方盒中解的个数证明一个渐近公式。达文波特达成了这一点，前提是上述双线性方程的解的个数增长不太快。

给出达文波特的招牌成果：17 变量（定理 18.1）→ 精细化到 16 变量的三次型必有非平凡整数零点；并说明证明的目标是「方盒中解数的渐近公式」，前提是双线性方程的解不要增长太快。

- 「16 个变量的三次型必有非平凡整数零点」：这是达文波特一生最精致的成就之一。无论三次型多复杂（只要整系数、16 个变量），方程 $F = 0$ 一定有不全为 0 的整数解。注意这是**无条件的整体结论**，不需要假设局部可解——变量够多时局部条件自动满足，呼应第二幕开篇的希望。
- 「**渐近公式 + 前提**」：证明的真正目标比「有解」更强——给出边长 P 方盒里解的个数的精确渐近公式。但这有个**前提**：那些双线性方程的解不能增长太快。下一段会讲这个前提如何被「二选一」论证巧妙绕过。

处理后一问题的论证导向两种可能：要么双线性方程的解的个数确实得到适当控制，要么原三次型出于几何原因而有非平凡整数零点。无论哪种情形，三次型都有非平凡整数零点。……型

$$x_1^3 + x_2(x_3^2 + \cdots + x_n^2)$$

在 $x_1 = x_2 = 0$ 时恒为零，因此在边长为 P 的方盒中有 $\gg P^{n-2}$ 个解。这个例子表明：人们一般不能指望得到与定理 17.1 相关联的那类渐近公式。

解释证明那「略显不寻常的逻辑结构」：用「二选一」论证保证无论如何都有零点；并举一个例子说明为什么「渐近公式」不可能对所有三次型成立。

- 「**二选一 (dichotomy) 论证**」：这是一个非常聪明的逻辑设计。证明分叉成两种情况：(甲) 双线性方程的解被控

制住 $\rightarrow$ 渐近公式成立 $\rightarrow$ 有零点；(乙) 双线性方程的解失控 $\rightarrow$ 但这种「失控」本身（出于几何原因）就直接逼出了一个非平凡零点。无论落在哪个分支，结论「 F 有非平凡整数零点」都成立。这种「敌人的强反而帮了我」的论证，是高级证明的典型美感。

- 例子 $x_1^3 + x_2(x_3^2 + \cdots + x_n^2)$ ：为什么这个例子重要？它说明渐近公式（精确计数）做不到普适。看：只要 $x_1 = x_2 = 0$ ，无论 $x_3, \dots, x_n$ 取什么，这个型都等于 0。于是它有一整个「 $(n-2)$ 维的解平面」，方盒里光这些解就有 $\gg P^{n-2}$ 个——这远多于「随机」三次型该有的解数（约 P^{n-3} ）。所以对这种「退化」的型，解多到不符合标准渐近公式。结论：「有解」可以普适，但「精确渐近公式」只能在排除这类退化型后才成立。这正是落入上面「分支乙」的型。

16 个变量的结果可以说是达文波特最精致的成就之一……达文波特的途径被施密特大大推广……对于单个型 $F(x_1, \dots, x_n)$ ，结果可借助不变量 $h(F)$ 来表述， $h(F)$ 定义为使得可以写成

$$F(\mathbf{x}) = G_1(\mathbf{x})H_1(\mathbf{x}) + \cdots + G_h(\mathbf{x})H_h(\mathbf{x})$$

的最小整数 h ，其中 G_i, H_i 为系数有理的非常数型。

介绍施密特 (Schmidt) 对达文波特方法的大推广，并引入衡量型「复杂度」的关键不变量 $h(F)$ 。

不变量 $h(F)$ (读「h of F」) 这是衡量一个型 F 有多「可分解」的量。定义：把 F 写成若干个「两个型的乘积」之和 $G_1H_1 + \cdots + G_hH_h$ （每个 G_i, H_i 都是非常数的、有理系数的型）所需的最少项数 h 。

直觉： $h(F)$ 小 $\iff F$ 「高度可约 / 退化」（能用很少几个乘积拼出来，结构特殊）； $h(F)$ 大 $\iff F$ 「一般 / 难拆」。
它精确刻画了上一段那个退化例子的「退化程度」。
 $\mathbf{x}$ 是向量记号，表示 $(x_1, \dots, x_n)$ 一整组变量，黑体表示「一串数打包成一个向量」。

考察达文波特关于 16 变量三次型的论证便可确立：
对任意满足 $h(F) \geq 16$ 的三次型 F ，标准的哈代-李特尔伍德渐近公式成立。当 $h(F) \leq 15$ 且 $n \geq 16$ 时，型 F 仍有非平凡整数零点，因为人们可以取诸 $H_i(\mathbf{x})$ 为线性型，并使用 $H_1, \dots, H_h$ 的一个公共零点。

用 $h(F)$ 把「16 变量」结果重新表述得更精准： $h(F) \geq 16$ 时有渐近公式； $h(F) \leq 15$ 时虽无渐近公式但仍有零点（用另一条几何理由）。

这里把上面那个「二选一」用 $h(F)$ 量化了，非常漂亮：

- $h(F) \geq 16$ （型「足够一般」）：渐近公式成立，解数有精确主项。这对应分支甲。
- $h(F) \leq 15$ 且 $n \geq 16$ （型「足够退化」）：渐近公式可能失效，但退化本身送来了零点——因为 h 小意味着能把 F 写成 $\sum G_i H_i$ ，其中可取 H_i 都是线性型；线性方程组 $H_1 = \dots = H_h = 0$ （ $h \leq 15$ 个方程、 $n \geq 16$ 个未知数，方程比未知数少）必有非零公共解，在那点上 $F = \sum G_i H_i = 0$ 。这对应分支乙，且现在给出了构造性的零点。两个分支合起来覆盖一切三次型。

在他的推广中，施密特能够得到一个明确的函数 $n(d)$ ，使得对任意满足 $h(F) \geq n(d)$ 的 d 次型，哈代-李特尔伍德公式成立。……于是，当人们从单

个 $d = 5$ 次型出发时，便想了解三次型方程组的零点。在这方面，施密特证明：一个整系数三次型方程组若至少有 $(10r)^5$ 个变量，则有非平凡整数零点。

施密特把上面的框架推广到任意次数 d ，并指出处理 $h(F)$ 小的型会自然引出「型方程组」，进而给出三次型方程组的零点结果。

- 「从 $d = 5$ 引出三次型方程组」：这是个微妙但重要的递归结构。处理一个退化的 5 次型 ($h(F)$ 小) 时，会冒出更低次的型组成的方程组——具体地，一个 5 次型分解 $\sum G_i H_i$ 里，若 G_i, H_i 取 2 次和 3 次，研究公共零点就要研究一组三次型的公共零点。于是问题从「单个高次型」递归到「一组低次型」。这是施密特推广的核心机制。
- $(10r)^5$ 个变量： r 个三次型组成的方程组，只要变量数 $\geq (10r)^5$ 就有非平凡公共整数零点。 $(10r)^5$ 是「 $10r$ 的 5 次方」，对 $r = 1$ （单个三次型）给出 100000 个变量——远大于达文波特的 16，说明「方程组」比「单个型」难得多，代价是变量数暴涨。

达文波特的结果还被普利桑兹在另一方向上加以推广，他证明：若型 F 的系数以及解被允许位于一个代数数域中，则结果仍然成立。在这一更广的背景下，16 个变量仍然足够。

另一个推广方向：把「整数」换成「代数数域」里的数，达文波特的 16 变量结论依然成立。

「代数数域 (algebraic number field)」是把有理数 $\mathbb{Q}$ 扩大后的数系（比如允许 $\sqrt{2}$ 、 i 这类「代数数」参与）。普利桑

兹 (Pleasants) 的工作说明达文波特的方法很**稳健**：换一个更大的数系，关键的数字「16」纹丝不动。这是好定理的标志——它的核心不依赖于「恰好是整数」这件事。你不必掌握代数数域的细节，记住「16 这个数在更一般的数系里仍然有效」即可。

若假设型 F 非奇异（这是通有情形），则可以证明 10 个变量即足够。这里 10 这个数是最优的，因为存在 9 个变量、无非平凡 p -进零点的型。……这些工作使用哈代-李特尔伍德方法，但不使用外尔不等式，而是依赖泊松求和公式以及对「完整」指数和的估计。涉及非奇异型的完整指数和可借助德利涅关于有限域上簇的黎曼假设极为高效地加以估计……

引入「非奇异」这个附加假设：若 F 非奇异，则 10 个变量就够（且最优），但用的是另一套工具（泊松求和 + 完整指数和 + 德利涅的深刻定理）。

非奇异 (non-singular)·通有 (generic) 一个型 $F = 0$ 叫「非奇异」，粗略说就是它定义的曲面足够「光滑」、没有尖点 / 自交等坏点（用偏导数语言： F 与它所有一阶偏导不同时为零）。「通有 (generic)」意思是「绝大多数、随手写一个就是」——奇异型是稀有的例外。直觉：加上「非奇异」这个良好假设后，型的行为规整得多，于是门槛从 16 一下降到 **10**（且 10 最优：有 9 变量无非平凡 p -进零点的非奇异型）。代价是要换更精密的工具。这又是一次「加强假设 → 改进结论」的权衡。

- 「泊松求和公式 (Poisson summation)」：一个把「对整点求和」与「对其傅里叶变换求和」联系起来的强力公

式，是外尔不等式之外的另一条技术路线。

- 「完整指数和 (complete exponential sum)」 「德利涅 (Deligne) 的黎曼假设」：这些是代数几何 / 解析数论交汇处的深刻工具。德利涅 1974 年证明的「有限域上簇的黎曼假设」(韦伊猜想的最后一块，菲尔兹奖工作) 能**极高效地**估计非奇异型相关的指数和。你只需领会：
「非奇异」之所以能把门槛降到 10，是因为它能调用代数几何最深刻的成果。

……但素数幂模的和仍是一个相当大的难题。……

以其最简单的形式，上述方法可处理 13 个或更多变量的非奇异三次型。然而借助克卢斯特曼对圆法的精细化，这可降至 10。为处理 9 个变量的型，Hooley 采用了一种明显更精微的分析，旨在……仅节省一个 $\log P$ 的幂次。

细化上一段：朴素方法只到 13 变量，用克卢斯特曼精细化降到 10，胡利再用更精微的分析啃下 9 变量（仅靠抠出一个 $\log P$ 因子）。

- 「素数幂模 p^j (prime power moduli)」：德利涅的定理只直接管「素数模 p 」或「无平方因子模」的和；模 $p^2, p^3, \dots$ (素数的高次幂) 的和仍很难。这是技术上的一道坎。
- 门槛的逐级下降：13 $\rightarrow$ 10 $\rightarrow$ 9：这一连串数字生动展示了「方法越精，门槛越低」。克卢斯特曼 (Kloosterman) 精细化是圆法的一种加强版（更细地处理主弧 / 次弧），把 13 压到 10。胡利 (Hooley) 为了再啃下「9 变量」，付出了不成比例的努力——整段分析只为「抠出一个 $\log P$ 的幂次」这点微小优势。这让你

真切体会到：在前沿，每降低一个变量都要付出巨大代价，越往后越难。

伯奇的工作（在第 19 章中作了概述）……当 F 非奇异时，只要 $n > (d-1)2^d$ （且奇异级数与奇异积分为正），伯奇便能确立一个渐近公式。对 $d=3$ 这弱于 Hooley 的结果，但该方法对任意 d 值都奏效。事实上，其后的研究未能就任何 $d > 3$ 改进伯奇的结果。

介绍第 19 章伯奇 [6] 的一般性结果：对任意次数 d 的非奇异型，变量数 $n > (d-1)2^d$ 时有渐近公式；虽然对 $d=3$ 不如胡利，但它的普适性（任意 d ）至今无人超越。

- $n > (d-1)2^d$ ：伯奇的门槛。代入 $d=3$ 得 $n > 2 \cdot 8 = 16$ ，即 $n \geq 17$ ——比胡利的 9 弱，但伯奇的优势是这个公式对每个 d 都成立，而胡利的 9、10 只对三次型。「奇异级数与奇异积分为正」就是我们反复说的局部条件 $\mathfrak{S} > 0, \mathfrak{J} > 0$ （保证局部有解）。
- 「未能就任何 $d > 3$ 改进」：一个发人深省的事实——伯奇 1962 年的结果，对所有高于 3 次的型，六十年来无人改进分毫。这标记了当前数学的真实边界，也说明三次型之所以被达文波特、胡利反复深耕，正因为它是唯一能突破伯奇一般界的特例。

伯奇的论证基于外尔不等式，并导出一个与达文波特工作中的双线性方程类似的多重线性方程组。这些方程用一种不同于达文波特所用的技巧来处理——达文波特的技巧更简单、几何意味更明显，但需要关于 F 之奇异点的信息。

收尾第二篇前言：比较伯奇与达文波特两套技巧的异同——都从外尔不等式出发、都导出（多重）线性方程组，但处理方式不同，各有优劣。

这是对整个第二幕的技术总结，点出两条路线的取舍：达文波特的技巧「更简单、更几何」，但需要知道 F 奇异点的信息（所以更适合非奇异型）；伯奇的技巧更一般、不需要奇异点信息，但更抽象。这再次体现数学方法的权衡——没有放之四海皆优的方法，简单 / 一般 / 几何直观三者常常不可兼得，要看你手头的型是什么样子来选工具。

前言（三）· 丢番图不等式（第 20 章）

Foreword — Diophantine inequalities

撰文：D. E. Freeman（卡尔顿大学）

撰写人弗里曼（Freeman）专攻丢番图不等式。这篇导读对应全书「第三幕」（仅第 20 章）：把等号 $= 0$ 松绑成「绝对值小于一个数」，研究的不再是方程而是不等式。这是达文波特与海尔布朗 1946 年开创的方向。

在最后一章中，达文波特阐述了他与海尔布朗 1946 年合作的开创性工作。他们展示了如何调整哈代-李特尔伍德方法，使之能在丢番图不等式上产生结果。……现今通常称为达文波特-海尔布朗方法。

点题：第 20 章讲达文波特-海尔布朗方法——把圆法从「方程」搬到「不等式」。

「达文波特-海尔布朗方法（Davenport-Heilbronn method）」是圆法的一个重要变种。原版圆法处理

$F(\mathbf{x}) = N$ (精确等于一个整数), 靠的是复指数的正交性 $\int_0^1 e(\alpha m) d\alpha = [m = 0]$ 。但不等式 $\|(F(\mathbf{x}))\|$

设 s 为满足 $s \geq 5$ 的整数, $\lambda_1, \dots, \lambda_s$ 为实数, 它们不全同号, 且不全成有理数之比。本章内容是一个证明, 表明: 给定任意正实数 C , 存在丢番图不等式

$$|\lambda_1 x_1^2 + \dots + \lambda_s x_s^2|$$

陈述第 20 章的主定理: 在「不全同号、不全成有理比」两个条件下, 对角二次不等式 (3) 必有非平凡整数解。

λ_i (lambda, 读「兰布达」) 与两个关键条件

$\lambda_1, \dots, \lambda_s$ 是不等式 (3) 的实数系数 (注意: 是实数, 不必是整数! 这是不等式问题与方程问题的根本区别)。两个条件务必理解:

- 「不全同号」: λ_i 必须有正有负。为什么? 若全是正数, 则 $\lambda_1 x_1^2 + \dots \geq 0$ 且只有全取 0 才接近 0, 凑不出非平凡的「接近 0」。必须正负搭配才能让大项相互抵消、把和压进 $(-C, C)$ 这个小窗口。(k 为奇数时 x^k 本身可正可负, 故不需要此条件。)

- 「不全成有理数之比」 (not all in rational ratio): 即诸 λ_i 之间至少有一对的比值是无理数。为什么需要? 若所有比值都是有理数, 可同乘一个公分母把 λ_i 全变成整数, 于是 (3) 退化成一个普通的整数方程问题 ($C < 1$ 时就是 $= 0$), 不再是「真正的不等式问题」。无理比是不等式问题的灵魂——正是它让取值能「稠密地」逼近 0。

对比 (3) 与方程 $\sum \lambda_i x_i^2 = 0$ ：不等式版本要松得多（只要落进宽为 $2C$ 的窗口，不必精确为 0），所以能在系数为**任意实数**时谈论解——这是它比方程问题更普适、也更需要新方法的原因。末句还说明这一切可推广到 k 次幂、变量数 $\geq 2^k + 1$ 的情形。

证明是哈代–李特尔伍德方法的一个巧妙改造。人们对某个大正数 P 估计 (3) 中满足 $|x_i| \leq P$ 的整数解的个数。这里不像哈代–李特尔伍德方法那样在一个单位区间上积分，而是在整条实轴上对一个适当的衰减核积分。这里没有多个优弧，主要贡献来自以零为中心的一个区间，而最难界定的区域则由一部分中等大小的数构成。处理后一区域时，要用到诸比值之一为无理数这一假设。

概述证明思路：把圆法的「在 $[0, 1]$ 上积分」改成「在整条实轴上对衰减核积分」，并描述新版的「主弧 / 次弧」结构。

这段把「达文波特–海尔布朗方法」与原版圆法的三处**关键改造**讲清楚了，逐条对照：

1 积分区域： $[0, 1]$ （圆） $\rightarrow$ **整条实轴** $\mathbb{R}$ 。原版靠周期性把区间卷成圆；不等式问题没有这种周期性，必须在整条实轴上积。

2 被积的「探测器」：复指数 $\rightarrow$ **衰减核**（decaying kernel）。原版用 $e(-N\alpha)$ 来「挑出」精确等于 N 的解（靠正交性）。

不等式要挑出的是「落在 $(-C, C)$ 窗口内」的解，于是换成一个在 0 附近大、远处迅速衰减到 0 的「核函数」 $K(\alpha)$ （典型取 $K(\alpha) = \left(\frac{\sin \pi C \alpha}{\pi \alpha}\right)^2$ 这类）。它像一个「软窗口」，把窗口内的解算进来、窗口外的滤掉。「衰减」保证了实轴上的积分收敛。

3 **区域划分：多个优弧 $\rightarrow$ 三段（中心 / 中等 / 远端）。**原版把 $[0, 1]$ 切成围着各个分数的多个主弧。新版变成：(i) **以 0 为中心的一个区间**提供主项（取代原来的多主弧）；(ii) 远端因衰减核而可忽略；(iii) **「中等大小」的区域最难对付**——正是在这里，要动用「某对系数之比为无理数」这个假设，才能证明它的贡献不会盖过主项。

所以「无理比」假设的作用现在具体了：**它专门用来制服那块最难的「中等区域」**。有理比时这块区域会出问题（产生伪主项），无理比让它乖乖变小。

在讲义中，达文波特猜测 (3) 即便对 $s \geq 3$ 也非平凡可解；……事实上，马古利斯对这两个问题都给出了肯定回答，确立了

$$|Q(x_1, \dots, x_s)| < \varepsilon$$

……这确立了奥本海姆猜想，因为它蕴含此类型在整数点处的值在实轴上稠密。我们注意到，马古利斯的证明所用的技巧不同于哈代–李特尔伍德方法。

讲后续重大进展：达文波特猜 $s \geq 3$ 也行、并问能否推广到非对角的一般二次型；这两个问题都被马古利斯（用全新方法）解决，即著名的奥本海姆猜想。

- **达文波特的两个猜想：**(a) 把变量门槛从 $s \geq 5$ 降到 $s \geq 3$ ；(b) 把「对角型」推广到「一般（非对角）不定二次型 Q 」。达文波特只是猜，没证。
- **马古利斯（Margulis）解决奥本海姆猜想（Oppenheim conjecture）：**这是 20 世纪数论的里程碑。马古利斯证明：对一般不定二次型 Q ($s \geq 3$ 、系数不全成有理比)， $|Q(\mathbf{x})| < \varepsilon$ 对任意小的 $\varepsilon > 0$ 都有非平凡整数解。「 ε (epsilon)」这里就是任意小的正数。这等价于说 Q 在整点处的取值在实轴上稠密 (dense, 意思是「任何一个实数附近都有 Q 的取值」)。
- **「技巧不同于圆法」：**极其重要的一句——马古利斯用的是遍历理论 / 齐性动力系统（研究群在空间上作用的轨道），而非圆法！这告诉你：圆法虽是本书主角，但解决奥本海姆这种问题最终靠的是另一个完全不同的数学领域。这正是数学的迷人之处——一个问题的突破常来自意想不到的方向。马古利斯因相关工作获菲尔兹奖。

关于更高次的型，达文波特提到皮特曼就三次型给出了一个结果，但他评论说，对更高奇数次的型证明类似结果似乎涉及一种「原理上的困难」。施密特在一系列论

文中提供了解决这一困难所需的关键结果。……他证明：给定奇正整数 $d_1, \dots, d_R$ ，存在一个常数 $C(d_1, \dots, d_R)$ ……则存在非平凡整数向量 $\mathbf{x}$ 使得

$$|F_1(\mathbf{x})| < \varepsilon, \quad |F_2(\mathbf{x})| < \varepsilon, \quad \dots, \quad |F_R(\mathbf{x})| < \varepsilon$$

讲高次型不等式：达文波特指出有「原理上的困难」，施密特用「圆法 + 对角化」突破，证明了「奇数次不等式组」的普适可解性。

- 「原理上的困难 (difficulty of principle)」：达文波特坦言，从三次推广到更高奇数次型，不只是「技术更繁」，而是**方法本身有一道原理性的坎**，靠原方法硬推过不去。承认「这里有真正的困难」本身是诚实而珍贵的——它指引了后人该往哪使劲。
- 施密特的突破：圆法 + 对角化 (diagonalization)。他把达文波特–海尔布朗方法和一个「对角化程序」（把一般型先化简成对角型再处理，呼应第二幕达文波特剥变量的思想）结合，跨过了那道坎。
- 结论的含义：给定任意一组奇数次数 $d_1, \dots, d_R$ ，存在只依赖这些次数的常数 C ，使得只要变量数 $s \geq C$ ，那么**任意 R 个实型 $F_1, \dots, F_R$ 都能被同一个非平凡整点 $\mathbf{x}$ 同时压进 ε 窗口**。即「一个整点同时让一组型都接近 0」。 R 是型的个数， $\mathbf{x}$ 是 s 维整向量。这是不等式版的「伯奇定理」。

有许多结果给出诸如 $C(d_1, \dots, d_R)$ 之类的下界……布吕德恩与库克在系数满足某些条件下，对对角型方程组给出了这样一个结果；纳德萨林甘与皮特曼对 R 个对角三次型方程组给出了明确的下界。

补充：很多后续工作在给上面那个抽象常数 $C(d_1, \dots, d_R)$ 算出具体的、可写下的下界。

施密特的常数 C 是「存在但没明确算出」的（像前面布劳尔、伯奇的门槛一样）。这段提到的两组工作（布吕德恩-库克 Brüdern-Cook、纳德萨林甘-皮特曼 Nadesalingam-Pitman）就是在把这个抽象常数**具体化**，针对特定型（对角型、对角三次型组）给出能写下的数值下界。这又是「先证存在、再求明确界」的经典分工。

人们也可以问及涉及一般正定型（其系数不全成有理数之比）的不等式。我们当然不期望这些型在整数点处的值在实轴上稠密；因此相关的问题——由埃斯特曼提出——是：当变量数足够大时，随着这些值趋于无穷，相邻的值之间的间隔是否趋于零。

转向一类不同的型——「正定型」，并引出埃斯特曼 (Estermann) 的「间隔问题」。

正定型 (positive definite form) 与「间隔问题」

「正定」二次型指除原点外取值恒为正的型（如 $x_1^2 + x_2^2 + \dots$ 这类，配上无理系数）。

它和前面的「不定型 (indefinite)」相反——不

定型有正有负、取值能逼近 0，所以能问「稠密」；正定型取值恒正且趋于 $+\infty$ ，**不可能**在整条实轴上稠密。

于是埃斯特曼换了个更精细的问法（「间隔问题，gaps problem」）：把正定型在所有整点的取值从小到大排成一系列 $v_1, v_2, \dots$ ，相邻两个取值之间的间隔 $v_{n+1} - v_n$ 会不会随取值增大而趋于 0？直觉：取值虽不能填满实轴，但会不会越来越「密」？这是个很自然的精细化问题。

对于对角二次型，达文波特与刘易斯指出：当变量数 s 至少为 5 时，这可由雅尼克与瓦尔菲什的一个结果轻松推出。……他们的方法实质上表明……对任意充分大的整数点 $\mathbf{x}_0$ ，存在许多整数点 $\mathbf{x}$ ，对它们有 $|Q(\mathbf{x}) - Q(\mathbf{x}_0)| < \varepsilon$ ……1999 年，本特库斯与格策用强有力的新技巧解决了间隔问题。

报告间隔问题的进展：对角二次型 $s \geq 5$ 由旧结果即得；一般情形经达文波特-刘易斯、库克-拉加万推进，最终 1999 年本特库斯与格策（Bentkus-Götze）用新技巧彻底解决。

- $|Q(\mathbf{x}) - Q(\mathbf{x}_0)| < \varepsilon$ ：这是间隔问题的等价表述——给定任一个大整点 $\mathbf{x}_0$ ，能找到「许多」别的整点 $\mathbf{x}$ ，使 $Q(\mathbf{x})$ 与 $Q(\mathbf{x}_0)$ 相差不到 ε 。这正说明取值「扎堆」、间隔趋于 0。「许多 (many)」是可精确定义的数量级。
- **进展链条**：雅尼克-瓦尔菲什 (Jarník-Walfisz, 旧结果) $\rightarrow$ 达文波特-刘易斯 (迈出

一步) → 库克-拉加万 (Cook-Raghavan, 阐明) → 本特库斯-格策 (1999, 彻底解决, 对 $s \geq 5$ 任意正定二次型)。这是一个典型的「接力赛」, 跨越半个多世纪。本特库斯-格策的新技巧后来催生了丢番图不等式领域的大量新工作。

自达文波特与海尔布朗的工作以来, 关于保证 k 次对角丢番图不等式非平凡可解所需的 s 的下界, 已有许多改进。对每个正整数 k , 设 $G_{\text{ineq}}(k)$ 表示满足如下性质的最小正整数 s_0 : 对所有 $s \geq s_0$ ……都存在

$$|\lambda_1 x_1^k + \cdots + \lambda_s x_s^k| < \varepsilon \quad (4)$$

的非平凡整数解。……达文波特与罗斯给出了一个改进; 他们证明存在常数 $C_1 > 0$ 使得

$$G_{\text{ineq}}(k) \leq C_1 k \log k.$$

为不等式问题定义一个与 $G(k)$ 平行的量 $G_{\text{ineq}}(k)$, 并给出达文波特-罗斯的上界 $C_1 k \log k$ 。

$G_{\text{ineq}}(k)$ (下标 **ineq** = **inequality**, 「不等式版的 G 」) 完全模仿华林问题里的 $G(k)$ 定义, 只是把「方程有解」换成「不等式 (4) 有解」: 它是能保证「 k 次对角不定不等式 (4) 必有非平凡整数解」的最小变量数 s_0 。把它和 $G(k)$ 对照, 就能看出「不等式版」和「方程版」的难度关系。达文波特-罗斯 (Davenport-Roth) 的结果 $G_{\text{ineq}}(k) \leq C_1 k \log k$ 说明: 不等式版所需变量

也大约是 $k \log k$ 量级——和华林问题的 $G(k)$ 同一个量级。这是个令人欣慰的对应。罗斯 (Roth) 即因「有理逼近」获菲尔兹奖的那位 K. F. Roth。

事实上，达文波特-海尔布朗方法足够灵活，使得不等式的界大致与华林问题工作所给出的界相平行。特别地，对大 k ，人们有

$$G_{\text{ineq}}(k) \leq k(\log k + \log \log k + 2 + o(1)). \quad (5)$$

……伍利晚近的工作表明： $G(k)$ 的界一般（除少数例外）也适用作 $G_{\text{ineq}}(k)$ 的界。

强调一个漂亮的「平行现象」：不等式版的界 (5) 与华林问题版 $G(k)$ 的界几乎一模一样，伍利甚至证明二者基本通用。

把 (5) 与 第一篇前言里 $G(k) \leq s_2(k) = k(\log k + \log \log k + O(1))$ 并排看，你会发现它们几乎相同（只是 $O(1)$ 被精确成 $+2 + o(1)$ ）。这传达一个深刻信息：「方程问题」和「不等式问题」尽管表面不同、用的探测器不同，深层难度却是同一个量级。达文波特-海尔布朗方法的「灵活」正体现于此——它能把华林问题的成果几乎原样搬到不等式上。伍利 [101] 把这个「平行」做成了定理： $G(k)$ 的界基本上可直接拿来当 $G_{\text{ineq}}(k)$ 的界。

正如达文波特所指出的，第 20 章中的证明仅适用于一个由大的 P 构成的序列，该序列依赖于诸系数之比值的有理逼近性质。……我们会期望 (4) 中满足 $|x_i| \leq P$ 的整数解的个数 $N(P)$ 满足

$$N(P) = C(s, k, \lambda_1, \dots, \lambda_s) \varepsilon P^{s-k} + o(P^{s-k}). \quad (6)$$

指出第 20 章证明的一个局限：它只对「特殊的一列 P 」给出结果，而非对所有 P ；理想情况则是对所有 P 都有渐近公式 (6)。

- 「仅适用于一个由大的 P 构成的序列」：这是达文波特–海尔布朗原证明的一个**诚实的瑕疵**。它不能对每个 P 都给出解数，只能对一串特殊挑选的 P （这串 P 的选取还依赖于系数比值 λ_i/λ_j 能被有理数逼近得多好）。直觉上：无理数的有理逼近时好时坏，证明只能挑「逼近恰好」的那些 P 下手。
- **渐近公式 (6)**：理想目标。 $N(P)$ 是方盒 $|x_i| \leq P$ 内解的个数，主项是 $C \cdot \varepsilon \cdot P^{s-k}$ （解数随方盒变大像 P^{s-k} 增长，正比于窗口宽度 ε ）， $o(P^{s-k})$ 是可忽略误差。 C 是只依赖 s, k, λ_i 的正常数。这个公式形式上和华林问题的渐近公式如出一辙——再次印证「方程 / 不等式平行」。

……本特库斯与格策对系数不全成有理数之比的一般正定二次型，对所有正 P ，确立了 (6) 的相应类比……通过改造他们的工作，弗里曼得以对所有正 P 证明：对 k 次的不定对角型，存在诸如 (6) 那样的渐

近公式。伍利借助减少所需变量数的巧妙思想，对这一工作作了相当大的简化与改进。

报告「对所有 P 」这个更强目标的进展：本特库斯-格策、弗里曼（本前言作者）、伍利相继把渐近公式 (6) 做到了对所有 P 成立。

这段把上一段的「瑕疵」补好了：达文波特只能对特殊 P 序列，而本特库斯-格策、弗里曼（Freeman，本文作者本人的贡献）、伍利逐步把它升级到所有正 P 。注意伍利的贡献方式很典型——「减少所需变量数」，这是圆法里永恒的主题（变量是资源，省变量就是进步），呼应第二幕「不浪费变量」。

特别地，关于大 k 时渐近公式的存在性，……

$$G_{\text{asympt}}(k) \leq k^2 (\log k + \log \log k + O(1))$$

最后，我们指出：埃斯金、马古利斯与莫泽斯运用不同于达文波特-海尔布朗方法的技巧，事实上更早地证明了：对至少四个变量、系数不全成有理数之比、且符号差不等于 $(2, 2)$ 的一般不定二次型，期望类型的渐近公式存在。

收尾第三篇前言：定义「保证渐近公式所需变量数」 $G_{\text{asympt}}(k)$ 并给出其大 k 上界；最后提到埃斯金-马古利斯-莫泽斯用动力系统方法更早得到的相关结果。

- $G_{\text{asympt}}(k)$ (下标 **asympt** = **asymptotic**): 保证「渐近公式 (6) (对所有 P) 成立」所需的最小变量数。它的界是 $k^2(\log k + \log \log k + O(1))$, 约 $k^2 \log k$ 量级——比「只要有解」的 $G_{\text{ineq}}(k) \sim k \log k$ 大一个 k 因子。完全平行于华林问题里「渐近公式需 $k^2 \log k$ 、只要有解需 $k \log k$ 」的关系 (见第一篇前言)。三个量 $G, G_{\text{ineq}}, G_{\text{asympt}}$ 至此构成一张完整的对照表。
- 「符号差 (signature)」与 $(2, 2)$: 不定二次型可化为 p 个正平方减 q 个正平方, (p, q) 就叫它的「符号差」。「符号差 $\neq (2, 2)$ 」是埃斯金-马古利斯-莫泽斯 (Eskin-Margulis-Mozes) 结果的一个技术性排除条件 (恰好 2 正 2 负这个特殊情形要单独处理)。他们用的又是**动力系统** (非圆法) 方法, 且「更早」就得到了——再次提醒你: 圆法之外另有天地, 同一座山峰常有多条登顶路线。

编者序 Editorial preface

撰文: T. D. Browning (牛津大学数学研究所)

布朗宁 (Browning) 是本书第二版的编者。这篇短序是全章最「人情味」的部分——他讲述自己为什么、以及怎样把这本鲜为人知的老讲义重新整理出版。读它能让你理解这本书背后的人和故事。

像许多数学家一样，我第一次接触数论是通过达文波特的著作《高等算术》。……而在更高深的层面上，达文波特《乘性数论》的第三版近来也已问世。……在得知达文波特还撰写过一套鲜为人知、处理一个具有重大现实意义之领域的讲义后，我立刻被驱使着去尝试将它重新付印。在此过程中，我尽力在原有格式中保留尽可能多的材料，仅仅订正了我所遇到的错误。

编者自述与达文波特著作的渊源，以及重印这套讲义的动机与原则（保留原貌、只改错）。

- **《高等算术》(The Higher Arithmetic) 与《乘性数论》(Multiplicative Number Theory)**：达文波特的两本名著，前者是面向初学者的数论入门经典（已出到第七版），后者是研究生级别的解析数论标准教材。布朗宁借此说明达文波特至今影响巨大。
- **编辑原则：「保留原貌、只订正错误」**。这对你阅读很重要——意味着正文（第 1–20 章）**根本就是 1960 年代达文波特讲课的原样**，那种「惜墨如金、跳步很多」的讲义风格被刻意保留了。这恰恰解释了为什么需要我们这份「补全每一步」的高中详解：原书为节省篇幅省略的推导，正是初学者最需要补回的部分。

正如书名所示，本书关注的是运用解析方法研究某些多项式方程与不等式的整数解。……这一解析方法通常被称为「哈代–李特尔伍德圆法」，其威力可由能用它

处理的数论问题之多样范围而清楚地展现出来。本书前半部分用于讨论该方法在其最经典背景下的应用：华林问题以及对角型对整数的表示。在第 11–19 章中……展示了该方法有时如何能被调整以处理一般齐次多项式方程组……最后，在第 20 章中……丢番图不等式背景下所作的工作。

编者用一段话再次总览全书结构——又一次「三幕剧」，与封底简介、目录互相印证。

读到这里，这「三幕剧」你应该已经烂熟于心了：

(1) 华林问题 + 对角型（圆法的经典主场，第 1–10 章）；(2) 一般齐次方程组（圆法的进阶调整，第 11–19 章）；(3) 丢番图不等式（圆法的跨界变体，第 20 章）。编者特意强调圆法的「威力体现在能处理问题的多样性」——同一套「把计数变积分、切主弧次弧」的思想，竟能横跨方程与不等式、对角与一般、整数与实系数。这种「一招吃遍天」的统一性，正是值得你花力气学圆法的根本理由。

达文波特讲义的本版已得到大幅充实，办法是增加了一篇前言，其主要目的在于将讲义所描述的知识状况置于现代视角之下。我极为感谢弗里曼教授、希思–布朗教授与沃恩教授将他们的权威赋予这一项目。我还要感谢莉莲·皮尔斯与卢克·伍德沃德……帮助我将达文波特的原始讲义转录为 LaTeX……最后，我很荣幸向巴斯大学的詹姆斯·达文波特以及剑桥大学出版社的戴维·特拉纳表达谢意……

致谢段：感谢三位前言作者、两位转录助手，以及达文波特的后人与出版社编辑。

这是例行的致谢，但有两点值得留意：(1) 三篇前言的作者正是前面读到的弗里曼、希思-布朗、沃恩——他们「把权威赋予这个项目」，意味着前言的现代综述部分出自各自领域的最顶尖人物之手，分量十足，可放心当作可靠的「领域地图」。(2) 「**转录为 LaTeX**」：LaTeX 是数学排版的标准软件；把 1960 年代的手写 / 打字讲义逐字录入 LaTeX 是一项浩大而易出错的工作（前面 $G(k)$ 那处笔误或许正源于此），这解释了重印一本老书背后的实际工作量。「詹姆斯·达文波特 (James Davenport)」是作者哈罗德·达文波特的后人。

T. D. Browning, 数学研究所, 牛津大学,
24-29 St. Giles', Oxford, OX1
3LB, browning@maths.ox.ac.uk

编者署名与通讯地址。

这是编者的落款（姓名、单位、地址、邮箱），是学术出版的惯例，无需多解释。至此，全书的「门厅」走完。

本章小结·你现在手里有什么 读完这扇门厅，你应当带走两样东西：

① **一套词汇表**：丢番图方程 / 型 / 对角型 / 华林问题 / $g(k)$ 与 $G(k)$ / 复指数 $e(\alpha)$ / 生成函数 $T(\alpha)$ / 圆法核心恒等式

$r(N) = \int_0^1 T^s e(-N\alpha) d\alpha$ / 主弧 $\mathfrak{M}$ 与次弧 $\mathfrak{m}$ / 奇异级数 $\mathfrak{S}$ 与奇异积分 $\mathfrak{J}$ / 局部到整体与

p -进 / O 与 $\ll$ 记号。后面 20 章不会再从头解释它们，但你已经备齐了。

② **一张全书地图（三幕剧）**：第 1–10 章用圆法打华林问题与对角型（信号在主弧、噪声在次弧、主项 = 奇异级数 $\times$ 奇异积分 $\times P^{s-k}$ ）；第 11–19 章升级到一般型，巅峰是达文波特「16 变量三次型必有非平凡零点」；第 20 章跨界到不等式，是达文波特–海尔布朗方法。

③ **三条贯穿始终的主线索**：(i) *局部到整体*——局部条件（奇异级数离 0）何时能保证整体有解；(ii) *变量是稀缺资源*——方法越精，所需变量越少，每省一个都极难；(iii) *圆法不是唯一武器*——林尼克、马古利斯、阿克斯–科亨等用了完全不同的路线。带着这张地图，去读第 1 章吧。

返回 [全书目录](#)