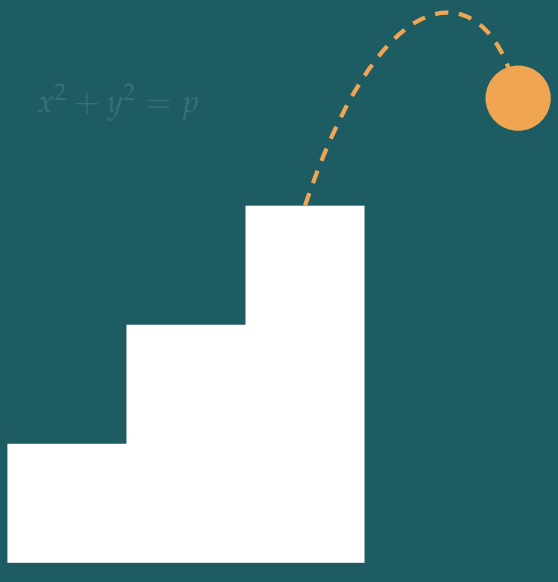
STEP 备试 · 大学数学

Beyond STEP

STEP II 及 III

2022 年详细解答



KEITH McNULTY

Beyond STEP

以历届 STEP 试题为门
走进大学数学

STEP II 及 III

2022 年详细解答

Keith McNulty

Beyond STEP：以历届 STEP 试题为门，走进大学数学

2022 年详细解答

2022 年 STEP 试卷（STEP II 及 III）全部 24 题的详细解答。自样本试卷以来每一道 STEP 题目的详细解答、*Beyond STEP* 全书以及各章的单行本，均可在 `beyondstep.keithmcnulty.org` 免费阅读及下载。

阅读解答之前，请先认真尝试解题。每个答案均已用数值方法核对，但解答数量如此之多，难免会有疏漏：如果发现有可疑之处，请相信自己的演算，并仔细检查。

由机器从英文翻译而成；尚未经审阅。

2026 年 11 月首次于网上发表。

本 PDF 内凡属作者本人创作的内容，包括正文、解答、插图、练习及答案，均按照 Creative Commons CC0 1.0 Universal 授权条款贡献予公有领域：你可以为任何目的复制、改编及分享，毋须征求许可。详见 `creativecommons.org/publicdomain/zero/1.0`。

本书讨论的 STEP 试题，版权属 Cambridge University Press & Assessment 所有。本书并无转载这些试题：每一题均附连结，通往官方来源的题目，即 STEP Support Programme 的试题资料库或 OCR 的历届试卷。

如有错漏，欢迎到 github.com/keithmcnulty/step_book_mcnulty 提交 issue 指正。

目录

STEP II

第 1 题 分部积分法 纯数学 . . . 3

第 2 题 递推关系 纯数学 . . . 8

第 3 题 斐波那契级数与小数 纯数学 . . . 12

第 4 题 绝对值函数 纯数学 . . . 16

第 5 题 三角形中的不等式 纯数学 . . . 21

第 6 题 正交轨线 纯数学 . . . 25

第 7 题 复系数多项式的根 纯数学 . . . 29

第 8 题 矩阵的不变直线 纯数学 . . . 33

第 9 题 摩擦力与力矩 力学 . . . 37

第 10 题 斜面上的抛体 力学 . . . 41

第 11 题 期望损失的最优化 概率与统计 . . . 46

第 12 题 均值、中位数与众数 概率与统计 . . . 50

STEP III

第 1 题 多项式的根 纯数学 . . . 54

第 2 题 无穷递降法 纯数学 . . . 58

第 3 题 隐函数微分法 纯数学 . . . 62

第 4 题 无穷乘积 纯数学 . . . 66

第 5 题 对称定积分 纯数学 . . . 70

第 6 题 滚动圆与切线 纯数学 . . . 74

第 7 题 向量积变换 纯数学 . . . 78

第 8 题 康美弗定理与无理性 纯数学 . . . 82

第 9 题 碰撞与恢复系数 力学 . . . 86

第 10 题 弹簧的平衡 力学 . . . 90

第 11 题项分布的平均偏差 概率与统计 94
第 12 题圆的随机弦 概率与统计 98

详细解答

STEP II 2022，第 1 题

纯数学 · 分部积分法

STEP II 2022，第 1 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

每一部分都建基于同一招：分部积分（integration by parts）会产生第二个积分，而这个积分恰好抵消（或几乎抵消）你手上已有的某一项。在 (i) 和 (ii) 中，这意味着要看出被积函数（integrand）是某个乘积的导数（derivative）。在 (iii) 中，同一招显示两个积分之差只不过是 e^x/x 在上下限之间的变化，于是 (a) 的草图便能完成其余的工作。

解答

第 (i) 部分

对第一项进行分部积分，取 $u=\sqrt{1+x^3}$ 及 $\frac{\mathrm{d}v}{\mathrm{d}x}=1$，于是 $v=x$ 及 $\frac{\mathrm{d}u}{\mathrm{d}x}=\dfrac{3x^2}{2\sqrt{1+x^3}}$：

$$\begin{aligned}\int 2\sqrt{1+x^3}\,\mathrm{d}x &= 2x\sqrt{1+x^3}-2\int x\cdot\frac{3x^2}{2\sqrt{1+x^3}}\,\mathrm{d}x\\ &= 2x\sqrt{1+x^3}-\int\frac{3x^3}{\sqrt{1+x^3}}\,\mathrm{d}x.\end{aligned}$$

最后的积分正是被积函数的第二项，所以把它加上去便会将其抵消：

$$\int\left(2\sqrt{1+x^3}+\frac{3x^3}{\sqrt{1+x^3}}\right)\mathrm{d}x=2x\sqrt{1+x^3}+c.$$

用乘积法则（product rule）对 $2x\sqrt{1+x^3}$ 求导，会得回 $2\sqrt{1+x^3}+\dfrac{3x^3}{\sqrt{1+x^3}}$，由此印证结果。

第 (ii) 部分

把被积函数拆成 $\dfrac{\sin x}{x}+\dfrac{2\sin x}{x^3}$。两部分各自都没有初等原函数（elementary antiderivative）（第一部分就是著名的 Si 被积函数），所以它们必须合并处理，正如 (i) 一样。对第二部分进行分部积分，取 $u=\sin x$ 及 $\frac{\mathrm{d}v}{\mathrm{d}x}=2x^{-3}$，于是 $v=-x^{-2}$：

$$\int \frac{2\sin x}{x^3}\,\mathrm{d}x=-\frac{\sin x}{x^2}+\int \frac{\cos x}{x^2}\,\mathrm{d}x.$$

接着对 $\dfrac{\cos x}{x^2}$ 进行分部积分，取 $u=\cos x$ 及 $v=-x^{-1}$：

$$\int \frac{\cos x}{x^2}\,\mathrm{d}x=-\frac{\cos x}{x}-\int \frac{\sin x}{x}\,\mathrm{d}x.$$

$\int \frac{\sin x}{x}\,\mathrm{d}x$ 与第一部分互相抵消，剩下

$$\int (x^2+2)\frac{\sin x}{x^3}\,\mathrm{d}x=-\frac{\sin x}{x^2}-\frac{\cos x}{x}+c=-\frac{\sin x+x\cos x}{x^2}+c.$$

验证一下：对 $-\dfrac{\sin x}{x^2}$ 求导得 $-\dfrac{\cos x}{x^2}+\dfrac{2\sin x}{x^3}$，对 $-\dfrac{\cos x}{x}$ 求导得 $\dfrac{\sin x}{x}+\dfrac{\cos x}{x^2}$；$\cos x/x^2$ 项互相抵消，总和正是被积函数。

第 (iii)(a) 部分

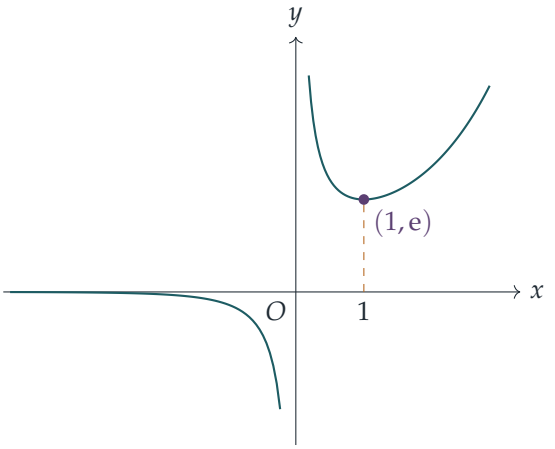
设 $y=\dfrac{\mathrm{e}^x}{x}$，定义域为 $x\neq 0$。由商法则（quotient rule），

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{x\mathrm{e}^x-\mathrm{e}^x}{x^2}=\frac{\mathrm{e}^x(x-1)}{x^2}.$$

它只在 $x=1$ 时为零，所以唯一的驻点（stationary point）是 $(1,\mathrm{e})$。当 $x<1$（$x\neq 0$）时导数为负，当 $x>1$ 时导数为正，所以 $(1,\mathrm{e})$ 是极小点（minimum）。

图像其余的形状：

- 当 $x<0$ 时，$y<0$ 且 y 递减。当 $x\to-\infty$ 时，$y\to 0^-$（指数函数占上风），所以负 x 轴是一条渐近线（asymptote）。当 $x\to 0^-$ 时，$y\to-\infty$。
- 当 $x>0$ 时，$y>0$。当 $x\to 0^+$ 时，$y\to+\infty$；y 下降至 $x=1$ 处的极小值 e，然后增加至 $+\infty$，增长速度比任何多项式都快。
- y 轴是一条垂直渐近线，而曲线与两条坐标轴都不相交。



第 (iii)(b) 部分

对 $\dfrac{\mathrm{e}^x}{x^2}$ 进行分部积分，取 $u=\mathrm{e}^x$ 及 $v=-x^{-1}$：

$$\int \frac{\mathrm{e}^x}{x^2}\,\mathrm{d}x = -\frac{\mathrm{e}^x}{x} + \int \frac{\mathrm{e}^x}{x}\,\mathrm{d}x, \qquad \text{所以} \qquad \int \left(\frac{\mathrm{e}^x}{x} - \frac{\mathrm{e}^x}{x^2}\right)\mathrm{d}x = \frac{\mathrm{e}^x}{x} + c.$$

（等价地，$\frac{\mathrm{e}^x}{x} - \frac{\mathrm{e}^x}{x^2} = \frac{\mathrm{e}^x(x-1)}{x^2}$ 正是我们在 (a) 中求得的导数。）

想一想 在 a 与 $2a$ 之间积分，究竟是否合理?

两个被积函数在 $x=0$ 处都趋向无穷，而在 0 附近 e^x/x^2 的积分发散，所以积分区间必须避开 0。这要求 $a\neq 0$；这样 a 与 $2a$ 同号，0 永远不会在两者之间，积分便是有限的。

两个积分相等，当且仅当它们的差为零：

$$\left[\frac{\mathrm{e}^x}{x}\right]_a^{2a} = \frac{\mathrm{e}^{2a}}{2a} - \frac{\mathrm{e}^a}{a} = \frac{\mathrm{e}^a}{2a}\left(\mathrm{e}^a - 2\right) = 0.$$

由于 $\mathrm{e}^a/(2a)\neq 0$，我们需要 $\mathrm{e}^a=2$，所以

$$a = \ln 2.$$

第 (iii)(c) 部分

与 (b) 完全一样，对于整数 $m\neq n$，两个积分要是有限的，0 就不能落在 m 与 n 之间的闭区间内，所以 m 和 n 均非零且同号。它们的差是

$$\int_m^n \frac{\mathrm{e}^x}{x}\,\mathrm{d}x - \int_m^n \frac{\mathrm{e}^x}{x^2}\,\mathrm{d}x = \frac{\mathrm{e}^n}{n} - \frac{\mathrm{e}^m}{m},$$

所以我们需要两个同号的相异整数，使 $f(x)=\mathrm{e}^x/x$ 在该两处取相同的值。换句话说，一条水平线必须在两个整数 x 值处与 (a) 的图像相交。

想一想 $x > 0$ 部分的曲线并非一一对应（one-to-one），因为它先下降再回升。那么，为何两个正整数不能给出相同的高度？

因为转向点（turning point）位于 $x = 1$，而 1 本身就是最小的正整数。逐一考虑各情况：

- *两者皆负*。由 (a) 可知，f 在 $x < 0$ 上严格递减（strictly decreasing），所以 $m \neq n$ 推出 $f(m) \neq f(n)$。
- *两者皆正*。这时 $m, n \geq 1$，而当 $x > 1$ 时 $f'(x) > 0$，所以 f 在 $[1, \infty)$ 上严格递增。同样地，$m \neq n$ 推出 $f(m) \neq f(n)$。

一条水平线可以与正数分支相交两次，但这时两个交点之一必定位于 $0 < x < 1$，而该范围内没有整数。所以没有任何一对相异整数符合要求，因此这是不可能的，得证。

验算

- 对于 (i)，在 $x = 1$ 处原函数的值为 $2\sqrt{2}$，而积分在 $[0, 1]$ 上的数值为 2.8284…，两者吻合。
- 对于 (ii)，积分在 $[1, 10]$ 上的数值为 1.4711…，与 $-(\sin x + x \cos x)/x^2$ 的变化相同。
- 对于 (iii)(b)，取 $a = \ln 2$ 时两个积分都约为 1.9224。另请注意 $f(\ln 2) = 2/\ln 2 = f(2 \ln 2)$：两个积分限位于极小点两侧的同一高度，正如草图所示必然如此。
- 对于 (iii)(c)，任何两个整数最接近的情况是 $f(1) = \mathrm{e} \approx 2.718$ 对 $f(2) = \frac{1}{2}\mathrm{e}^2 \approx 3.695$。

常见错误

- 在 (b) 和 (c) 中，忘记包含 0 的区间会令积分失去意义，因而看不出为何 m 和 n 必须同号。
- 在 (c) 中，不加限制地断言“f 是一一对应的”：事实并非如此（例如 $f(\ln 2) = f(2 \ln 2)$）。论证需要用到位于 $x = 1$ 的转向点。

STEP 之外

第 (i) 和 (ii) 部分是这样的例子：被积函数的两部分各自都没有初等原函数，但它们的和却有。微分代数（differential algebra）中的刘维尔定理（Liouville's theorem）精确刻画了哪些函数具有初等原函数；e^x/x、$\sin x/x$ 和 $\sqrt{1 + x^3}$ 都没有，这正是抵消之所以关键的原因。第 (iii)(c) 部分是一个单射性（injectivity）论证，而分析课程会把这类关于单调

性（monotonicity）和中值定理（mean value theorem）的推理严格化；可参阅 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP II 2022，第 2 题

纯数学 · 递推关系

> **STEP II 2022，第 2 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：
>
> www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

把关系式两边乘以 2，它就变成一个关于二阶差分（second differences）的陈述：$u_{n+2} - 2u_{n+1} + u_n$。在 (i) 中它等于零，所以一阶差分（first differences）是常数，数列（sequence）是线性的。在 (ii) 和 (iii) 中，二阶差分是一个给定的常数或一个给定的 n 的线性函数，所以技巧是减去一个恰好具有该二阶差分的多项式（polynomial），剩下的便是 (i) 中那类数列。

解答

第 (i) 部分

把关系式加倍并移项，对所有 $n \geq 1$ 得

$$u_{n+2} - u_{n+1} = u_{n+1} - u_n.$$

所以每个差分都等于前一个差分，由归纳法（induction）可知每个差分都等于第一个差分 $d = u_2 - u_1$。这就是相邻两项之间的固定差。

把 $n - 1$ 个这样的差分加起来，

$$u_n = u_1 + (n-1)d = dn + (u_1 - d),$$

若 $d \neq 0$，这是 n 的 1 次多项式；若 $d = 0$，则是 0 次多项式（即常数）。无论哪种情况，次数（degree）都不超过 1。

第 (ii) 部分

设 $v_n = t_n + pn^2$。由于

$$\tfrac{1}{2}\big(p(n+2)^2 + pn^2\big) = p(n^2+2n+2) = p(n+1)^2 + p,$$

v_n 的关系式变成

$$t_{n+1} + p(n+1)^2 = \tfrac{1}{2}(t_{n+2} + t_n) + p(n+1)^2 + p - p,$$

即对所有 $n \geq 1$, 有 $t_{n+1} = \frac{1}{2}(t_{n+2} + t_n)$。由第 (i) 部分, 存在常数 A 和 B 使得 $t_n = A + Bn$, 所以

$$v_n = pn^2 + Bn + A.$$

因为 $p \neq 0$, n^2 项不会消失, 所以 v_n 的次数恰好是 2。

由 $v_1 = v_2 = 0$:

$$p + B + A = 0, \qquad 4p + 2B + A = 0.$$

两式相减得 $3p + B = 0$, 所以 $B = -3p$, 从而 $A = 2p$。因此

$$v_n = p(n^2 - 3n + 2) = p(n-1)(n-2).$$

第 (iii) 部分

这次题目没有提示代换, 所以我们要自己找一个。把关系式加倍,

$$w_{n+2} - 2w_{n+1} + w_n = 2(an + b).$$

? **想一想** 哪一个 n 的多项式, 其二阶差分是 n 的线性函数?

每取一次差分, 多项式的次数就降一, 所以我们需要一个三次多项式 (cubic)。记 $\Delta^2 f(n) = f(n+2) - 2f(n+1) + f(n)$。直接展开得

$$\Delta^2 n^3 = (n^3 + 6n^2 + 12n + 8) - 2(n^3 + 3n^2 + 3n + 1) + n^3 = 6n + 6,$$
$$\Delta^2 n^2 = (n^2 + 4n + 4) - 2(n^2 + 2n + 1) + n^2 = 2.$$

所以 $\Delta^2(cn^3 + en^2) = 6cn + 6c + 2e$, 而当

$$c = \frac{a}{3}, \qquad e = b - 3c = b - a.$$

时, 它对每个 n 都等于 $2an + 2b$。现在设 $w_n = t_n + \frac{a}{3}n^3 + (b-a)n^2$。由于 Δ^2 是线性的, $\Delta^2 w_n = \Delta^2 t_n + 2(an + b)$, 于是关系式变成 $\Delta^2 t_n = 0$, 这正是第 (i) 部分的关系式。所以 $t_n = A + Bn$, 且

$$w_n = \frac{a}{3}n^3 + (b-a)n^2 + Bn + A.$$

因为 $a \neq 0$，三次项存在，所以 w_n 的次数是 3。

由 $w_1 = w_2 = 0$：

$$\frac{a}{3} + (b-a) + B + A = 0,$$
$$\frac{8a}{3} + 4(b-a) + 2B + A = 0.$$

第二式减去第一式得 $\frac{7a}{3} + 3(b-a) + B = 0$，所以 $B = \frac{2a}{3} - 3b$，再由第一式得 $A = 2b$。因此

$$w_n = \frac{a}{3}n^3 + (b-a)n^2 + \left(\frac{2a}{3} - 3b\right)n + 2b.$$

由于 $w_1 = w_2 = 0$，因式 $n-1$ 和 $n-2$ 必定出现，再比较 n^3 项和常数项，便得到较简洁的形式

$$w_n = \frac{(n-1)(n-2)(an+3b)}{3}.$$

（展开 $(n-1)(n-2)(an+3b)$ 得 $an^3 + (3b-3a)n^2 + (2a-9b)n + 6b$，与上式一致。）

验算

- 在 (iii) 中设 $a = 0$、$b = p$ 应该得回 (ii)，结果的确如此：$(n-1)(n-2)\cdot 3p/3 = p(n-1)(n-2)$。
- 直接运行递推关系（recurrence）。取 $a = 3$、$b = 0$，有 $w_{n+2} = 2w_{n+1} - w_n + 6n$，所以由 $w_1 = w_2 = 0$ 得 $w_3 = 6$、$w_4 = 24$、$w_5 = 60$。公式给出 $(n-1)(n-2)n$，即 $6, 24, 60$。
- 次数为 3 的数列，其三阶差分是常数。对 $0, 0, 6, 24, 60$，差分是 $0, 6, 18, 36$，然后是 $6, 12, 18$，再然后是 $6, 6$：是常数，等于 $2a$。

常见错误

- 说 (i) 中的数列次数恰好是 1：若 $u_1 = u_2$，它是常数。
- 在 (iii) 中只尝试 $w_n = t_n + kn^3$。单靠 n^3 项会产生 $6kn + 6k$，所以还需要一个 n^2 项来调整常数。
- 忘记由 $p \neq 0$（或 $a \neq 0$）去证明“次数恰好是 2”（或 3）。

STEP 之外

这是解 $y'' = f(x)$ 的离散类比（discrete analogue）：通解（general solution）等于一个特解（particular solution）加上齐次方程（homogeneous equation）的解，而这里齐次方

程的解就是线性数列 $A + Bn$。同样的“余函数（complementary function）加特解”结构支配着线性微分方程（linear differential equations），而猜测多项式特解的方法几乎可以原封不动地搬过去。Boyce 和 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）讲述了连续的版本。

详细解答

STEP II 2022，第 3 题

纯数学 · 斐波那契级数与小数

STEP II 2022，第 3 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

递推关系（recurrence）$F_{r+1} - F_r - F_{r-1} = 0$ 是整道题的核心。把它乘以 $\frac{1}{10}$ 的幂再求和，然后重新编排下标（re-indexing），每一部分都会变成 S_n 的倍数加上几个边界项，这正是生成函数（generating function）的技巧。第 (i) 部分是安全网：它说明斐波那契数（Fibonacci numbers）的增长不会快过 2 的幂，所以除以 10（或 100）的幂之后，尾项（tail）非常小，很容易估计上界。

解答

第 (i) 部分

首先，每个 $F_k \geq 0$（用归纳法（induction）证明：因为 $F_0, F_1 \geq 0$，而非负数之和仍是非负数）。其次，这个数列从不递减：$F_1 \geq F_0$，而对 $k \geq 1$，有 $F_{k+1} - F_k = F_{k-1} \geq 0$。

所以对每个 $k \geq 1$，

$$F_{k+1} = F_k + F_{k-1} \leq F_k + F_k = 2F_k.$$

现在固定 $n \geq 1$，对 $r \geq n$ 进行归纳。当 $r = n$ 时，命题变成 $F_n \leq F_n$。若对某个 $r \geq n$ 有 $F_r \leq 2^{r-n}F_n$，则由于 $r \geq 1$，

$$F_{r+1} \leq 2F_r \leq 2 \cdot 2^{r-n}F_n = 2^{(r+1)-n}F_n.$$

归纳完成，所以只要 $n \geq 1$ 且 $r \geq n$，便有 $F_r \leq 2^{r-n}F_n$。

第 (ii) 部分

我们把左边三个和各自用 S_n 表示。

第一个和。设 $k = r+1$：

$$\sum_{r=1}^{n} \frac{F_{r+1}}{10^{r-1}} = 100 \sum_{k=2}^{n+1} \frac{F_k}{10^k} = 100\Big(S_n - \frac{F_1}{10} + \frac{F_{n+1}}{10^{n+1}}\Big)$$
$$= 100S_n - 10F_1 + \frac{F_{n+1}}{10^{n-1}}.$$

第二个和。这就是 $10S_n$。

第三个和。设 $k = r-1$：

$$\sum_{r=1}^{n} \frac{F_{r-1}}{10^{r-1}} = \sum_{k=0}^{n-1} \frac{F_k}{10^k} = F_0 + S_n - \frac{F_n}{10^n}.$$

从第一个和减去第二和第三个和，得

$$(100-10-1)S_n - 10F_1 - F_0 + \frac{F_n}{10^n} + \frac{F_{n+1}}{10^{n-1}} = 89S_n - 10F_1 - F_0 + \frac{F_n}{10^n} + \frac{F_{n+1}}{10^{n-1}},$$

这就是所求的结果。

第 (iii) 部分

把 (ii) 左边的三个和合并成一个。它的第 r 项是 $(F_{r+1} - F_r - F_{r-1})/10^{r-1} = 0$，所以左边等于零。代入 $F_0 = 0$ 和 $F_1 = 1$，得

$$89S_n = 10 - \frac{F_n}{10^n} - \frac{F_{n+1}}{10^{n-1}}. \qquad (*)$$

? **想一想** 我们怎样知道当 $n \to \infty$ 时，剩下的两项会趋于零?

这正是第 (i) 部分发挥作用的地方。在那里取 $n = 1$，得对所有 $r \geq 1$ 都有 $F_r \leq 2^{r-1}$，所以

$$0 \leq \frac{F_n}{10^n} \leq \frac{1}{2}\Big(\frac{1}{5}\Big)^n, \qquad 0 \leq \frac{F_{n+1}}{10^{n-1}} \leq \frac{2^n}{10^{n-1}} = 10\Big(\frac{1}{5}\Big)^n,$$

而两个上界都趋于 0。在 $(*)$ 中令 $n \to \infty$，得 $89S_n \to 10$，即

$$\sum_{r=1}^{\infty} \frac{F_r}{10^r} = \frac{10}{89}.$$

至于尾项，在第 (i) 部分中取 $n = 7$（注意 $F_7 = 13$）：对 $r \geq 7$，有 $F_r \leq 2^{r-7} \cdot 13$，所以

$$\sum_{r=7}^{\infty} \frac{F_r}{10^r} \leq \frac{13}{10^7} \sum_{j=0}^{\infty} \frac{2^j}{10^j} = \frac{13}{10^7} \cdot \frac{1}{1-\frac{1}{5}} = \frac{65}{4} \times 10^{-7} = 1.625 \times 10^{-6} < 2 \times 10^{-6}.$$

$\frac{1}{89}$ 的各位数字。首六项之和为

$$S_6 = 0.1 + 0.01 + 0.002 + 0.0003 + 0.00005 + 0.000008 = 0.112358.$$

尾项 $T = \frac{10}{89} - S_6$ 严格大于零（因为它的各项都是正数），而且小于 2×10^{-6}，所以

$$0.112358 < \frac{10}{89} < 0.112360, \qquad \text{因此} \qquad 0.0112358 < \frac{1}{89} < 0.0112360.$$

严格介于这两个界之间的每个数都以 0.011235… 开头，所以 $\frac{1}{89}$ 小数点后的首六位数字是 $0, 1, 1, 2, 3, 5$。严格的下界很重要：它排除了展开式是 0.0112357… 的可能。

第 (iv) 部分

把目标数在开头的 00 之后两位两位地读：

$$0.00\ 01\ 01\ 02\ 03\ 05\ 08\ 13\ 21\ 34\ 55\ldots$$

这些两位数组正是 $F_1, F_2, \ldots, F_{10}$，每个都写成两位数字。所以这个数应该是 $\frac{1}{100}U$，其中

$$U = \sum_{r=1}^{\infty} \frac{F_r}{100^r}.$$

*求 U 的和。*由第 (i) 部分得 $F_r \le 2^{r-1}$，所以这个级数绝对收敛（converges absolutely）。因此我们可以像 (ii) 一样，逐项平移和合并收敛级数：

$$U - \frac{U}{100} - \frac{U}{100^2} = \frac{F_1}{100} + \sum_{r=2}^{\infty} \frac{F_r - F_{r-1} - F_{r-2}}{100^r} = \frac{1}{100}.$$

这里 $r = 1$ 的项是 $F_1/100$（因为 $F_0 = 0$），而之后每个括号都等于零。所以 $U \cdot \frac{9899}{10000} = \frac{1}{100}$，得 $U = \frac{100}{9899}$。因此候选答案是

$$\frac{U}{100} = \frac{1}{9899},$$

其中 $r = 1$、$s = 9899$，两者都是小于 10000 的正整数。

*验证各位数字。*设 U_{10} 为首十项之和。由于 $F_1, \ldots, F_{10}$ 每个都小于 100，各项不会互相重叠，而且

$$U_{10} = 0.01010203050813213455$$

恰好成立，共有 20 位小数。由第 (i) 部分取 $n = 11$（且 $F_{11} = 89$），余下各项满足

$$0 < U - U_{10} \le \frac{89}{100^{11}} \sum_{j=0}^{\infty} \Big(\frac{2}{100}\Big)^j = \frac{89}{10^{22}} \cdot \frac{100}{98} < 10^{-20}.$$

所以 $U_{10} < U < U_{10} + 10^{-20}$，$U$ 的首 20 位小数就是 U_{10} 的那些数字。除以 100 会把这些数字向右移两位：

$$\frac{1}{9899} = 0.0001010203050813213455\ldots,$$

这正是所要求的开头。

验算

- 用长除法（long division）得 $1/89 = 0.011235955\ldots$。数字 011235 吻合，而接下来的数字 9,5,5 显示斐波那契数一旦达到两位数，就会互相“进位”。
- (iii) 中 $r = 6$ 之后的尾项实际上是 $10/89 - 0.112358 \approx 1.55 \times 10^{-6}$，稳稳低于我们的上界 1.625×10^{-6}。
- 用电脑把 $1/9899$ 计算到 22 位小数，得 0.0001010203050813213455。下一组两位数是 90 而不是 $F_{11} = 89$，因为 $F_{12} = 144$ 超出了它的两位数位置，向前一组进位 1。

常见错误

- 在 (ii) 重新编排下标时把边界项弄错。把每个平移后的和的首项和末项写出来，就可以避免。
- 在 (iii) 中没有证明 $F_n/10^n$ 和 $F_{n+1}/10^{n-1}$ 趋于零。第 (i) 部分正是为此而设的。
- 忘记尾项是严格大于零的。单靠一个上界并不能确定第六位数字。

STEP 之外

函数 $G(x) = \sum F_r x^r$ 是斐波那契数的生成函数，而上面的平移论证说明，当 $|x| < \frac{1}{2}$ 时 $G(x) = x/(1 - x - x^2)$。代入 $x = \frac{1}{10}$ 和 $x = \frac{1}{100}$，便得到 $\frac{10}{89}$ 和 $\frac{100}{9899}$。生成函数把递推关系转化为代数运算，是组合数学（combinatorics）的核心工具。Herbert Wilf 的 *generatingfunctionology*（《生成函数学》）是一本生动的入门书。要严格地处理“重新排列收敛级数”，属于大学一年级分析（analysis）课程的内容：可参阅 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP II 2022，第 4 题

纯数学 · 绝对值函数

STEP II 2022，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

绝对值符号（modulus sign）只在其内部式子变号之处才起作用，因此在每两个这样的点之间的区间上，每个绝对值都可以换成其内部式子的正号或负号形式。反过来看，一项 $A|x-p|$ 会在 $x=p$ 处造成一个转角，该处的斜率（gradient）跳变 $2A$。因此，要写出单一公式，我们先配合斜率的跳变（在最后一部分还要配合 x^2 系数（coefficient）的跳变），然后再定出常数。

解答

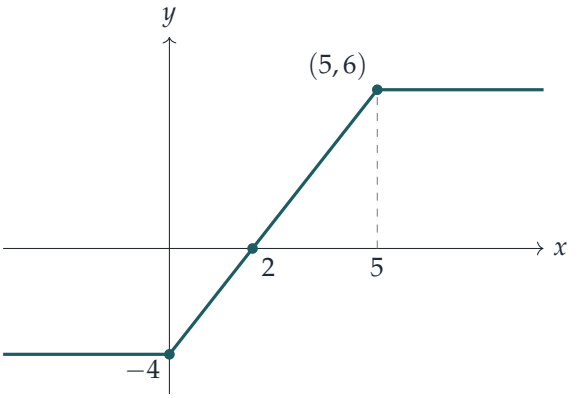
第 (i) 部分

临界点是 $x=0$（$|x|$ 在此转换）和 $x=5$（$|x-5|$ 在此转换）。

- 当 $x \le 0$: $|x| = -x$, 而由于 $x-5<0$, $|x-5| = 5-x$。所以 $f(x) = -x-(5-x)+1 = -4$。
- 当 $0 \le x \le 5$: $|x| = x$ 及 $|x-5| = 5-x$。所以 $f(x) = x-(5-x)+1 = 2x-4$。
- 当 $x \ge 5$: $|x| = x$ 及 $|x-5| = x-5$。所以 $f(x) = x-(x-5)+1 = 6$。

在 $x=0$ 和 $x=5$ 处，相关的两个表达式都成立（每个绝对值在该处要么为零，要么符号确定），所以重叠的端点是一致的：$-4 = 2(0)-4$ 及 $2(5)-4 = 6$。这就是所求的分段（piecewise）形式。

图像先是高度为 -4 的水平线，然后以斜率 2 从 $(0,-4)$ 经 $(2,0)$ 上升至 $(5,6)$，之后是高度为 6 的水平线。



第 (ii) 部分

第 (i) 部分启发我们尝试

$$\mathrm{g}(x) = A|x| + B|x-5| + Cx + D.$$

由左至右经过 $x = 0$ 时，$A|x|$ 一项由 $-Ax$ 变为 Ax，所以斜率跳变 $2A$；同理，$B|x-5|$ 令斜率在 $x = 5$ 处跳变 $2B$。已知的 g 在三段上的斜率分别是 -1、3、1，所以

$$2A = 3 - (-1) = 4, \qquad 2B = 1 - 3 = -2,$$

得 $A = 2$、$B = -1$。当 $x \geq 5$ 时，公式为 $2x - (x-5) + Cx + D = (1+C)x + 5 + D$，而这必须等于 $x + 10$，所以 $C = 0$ 及 $D = 5$。因此

$$\boxed{\mathrm{g}(x) = 2|x| - |x-5| + 5.}$$

验算其余两段：当 $x \leq 0$，$-2x - (5-x) + 5 = -x$；当 $0 \leq x \leq 5$，$2x - (5-x) + 5 = 3x$。两者都与 g 的定义相符。

? **想一想** 为甚么这种形式的公式有机会成功?

g 的每一段都是线性的，而且各段连续（continuous）地接合（在 $x = 0$ 处两侧都是 0，在 $x = 5$ 处两侧都是 15）。由线性式的绝对值之和加上一个线性函数所组成的函数，必定是连续且分段线性的，而转角只出现在各绝对值转换之处。假如 g 有跳跃，就没有这样的公式能产生它。

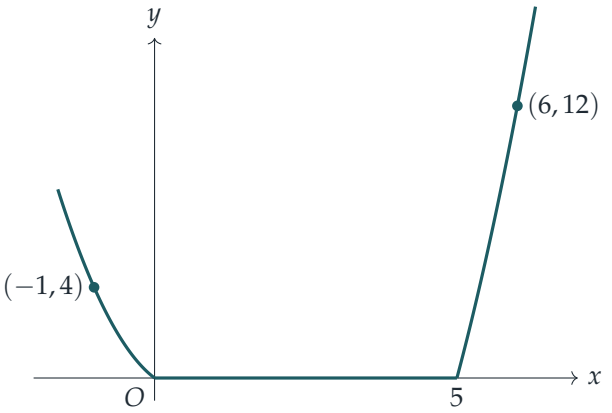
第 (iii) 部分

现在转换点是 $x = 0$（对应 $|x|$）以及 $x(x-5)$ 的根，即同样是 0 和 5。乘积 $x(x-5)$ 在 $x < 0$ 及 $x > 5$ 时为正，在两者之间为负。

▶ 当 $x \leq 0$：$\mathrm{h}(x) = x^2 - x + 4x + (x^2 - 5x) = 2x^2 - 2x = 2x(x-1)$。

- 当 $0 \le x \le 5$：$\mathrm{h}(x) = x^2 - x - 4x - (x^2 - 5x) = 0$。
- 当 $x \ge 5$：$\mathrm{h}(x) = x^2 - x - 4x + (x^2 - 5x) = 2x^2 - 10x = 2x(x-5)$。

所以 h 在 $[0,5]$ 上恒等于零。在左边，它是抛物线（parabola）$y = 2x(x-1)$ 的一部分：它以斜率 $4x - 2 = -2$ 到达原点，并在 $x \to -\infty$ 时无界地增大（例如 $\mathrm{h}(-1) = 4$）。在右边，它是 $y = 2x(x-5)$ 的一部分：它以斜率 $4x - 10 = 10$ 离开 $(5,0)$（例如 $\mathrm{h}(6) = 12$）。两臂并非互为镜像：右臂离开横轴时陡峭五倍。图像是连续的，在 $(0,0)$ 和 $(5,0)$ 处有转角，而且处处 $\mathrm{h}(x) \ge 0$。



第 (iv) 部分

第 (iii) 部分展示了我们需要的工具：$|x(x-5)|$ 在 $[0,5]$ 以外等于 $x^2 - 5x$，在其内则等于 $-(x^2 - 5x)$，所以它可以把一个 x^2 项“开启”和“关闭”。函数 k 的 x^2 系数在 $[0,5]$ 以外为 0，在其内为 2。在 $ax^2 + d|x(x-5)|$ 中，x^2 的系数在外面是 $a+d$，在里面是 $a-d$，所以我们需要 $a + d = 0$ 及 $a - d = 2$，即 $a = 1$、$d = -1$。于是 $x^2 - |x(x-5)|$ 具有正确的二次项表现，而且

$$x^2 - |x(x-5)| = \begin{cases} 5x & x \le 0 \text{ 或} x \ge 5, \\ 2x^2 - 5x & 0 \le x \le 5. \end{cases}$$

从 k 减去此式，剩下一个分段线性的余下部分：

$$\mathrm{k}(x) - \left(x^2 - |x(x-5)|\right) = \begin{cases} 5x & x \le 5, \\ 50 - 5x & x \ge 5. \end{cases}$$

它只在 $x = 5$ 处有一个转角，该处斜率由 5 跳变为 -5，即跳变 -10；所以它是 $-5|x-5|$ + (线性函数)。当 $x \le 5$ 时，$-5|x-5| = 5x - 25$，所以要加上 25 才得到 $5x$。这启发我们得出

$$\boxed{\mathrm{k}(x) = x^2 - |x(x-5)| - 5|x-5| + 25.}$$

验证。我们在每个区间上代入。

- 当 $x \le 0$: $x(x-5) \ge 0$ 且 $x-5 < 0$, 所以公式给出 $x^2 - (x^2-5x) - 5(5-x) + 25 = 5x + 5x = 10x$。
- 当 $0 \le x \le 5$: $x(x-5) \le 0$ 且 $x-5 \le 0$, 所以它给出 $x^2 + (x^2-5x) - 5(5-x) + 25 = 2x^2$。
- 当 $x \ge 5$: $x(x-5) \ge 0$ 且 $x-5 \ge 0$, 所以它给出 $x^2 - (x^2-5x) - 5(x-5) + 25 = 50$。

这些正正就是 k 的三段，所以该公式对每个实数 x 都正确。

这个单一公式并非唯一。例如 $|x(x-5)| = |x|\,|x-5|$，所以答案可以改写，而其他组合（例如用到 $x|x|$ 的）也同样可行。任何答案只要通过上述三个区间的检验，都是可以接受的。

验算

- 代入数值：由定义及 $2(2) - 7 + 5$ 均得 $\mathrm{g}(-2) = 2$；$\mathrm{g}(7) = 17 = 14 - 2 + 5$。至于 k：$\mathrm{k}(3) = 18$ 而 $9 - |-6| - 10 + 25 = 18$；$\mathrm{k}(-1) = -10$ 而 $1 - 6 - 30 + 25 = -10$。
- 连续性是一个快捷的合理性检查：任何由绝对值构成的单一公式都是连续的，而 g 和 k 在接合点确实吻合（$\mathrm{k}(5) = 50 = 2 \cdot 5^2$）。
- 四条公式都已用电脑在 $[-20, 25]$ 中 10^5 个随机点上与其分段形式比较，结果在舍入误差范围内一致。

常见错误

- 把 h 的两臂画成对称。左臂是 $2x(x-1)$，右臂是 $2x(x-5)$，在转角处的斜率分别是 -2 和 10。
- 忘记 $|x(x-5)|$ 在两个点变号，所以在 (iii) 中它在中间的区间翻转，而不只是在 0 处。
- 写出 k 的公式后没有代入全部三个区间验证；验证是必需的。

STEP 之外

这里追踪斜率跳变的做法，让我们初尝一个一般事实：$\mathbb{R}$ 上每个只有有限个转角的连续分段线性函数，都是一个线性函数加上若干 $|x-p|$ 项的组合。这正是数值分析（numerical analysis）中线性样条（linear splines）以及神经网络（neural networks）中“ReLU”单元背后的概念，因为 $\max(x, 0) = \frac{1}{2}(x + |x|)$。若想仔细了解连续性及分段定义的函数，可

参阅 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP II 2022，第 5 题

纯数学 · 三角形中的不等式

STEP II 2022，第 5 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

从 P 把三角形分成三块，便可看出无论 P 怎样移动，加权和（weighted sum）$ax+by+cz$ 始终固定为 2Δ。之后每一部分都是在这个单一线性约束（linear constraint）下的最优化（optimisation）问题。第 (i) 部分把 $x+y+z$ 与这个固定量作比较；(iii)(a) 的恒等式（identity）对 $x^2+y^2+z^2$ 做同样的事（它其实是乔装了的柯西－施瓦茨不等式（Cauchy–Schwarz inequality））；至于最大值，我们把 (i) 反过来用，以 c 代替 a。

解答

第 (i) 部分

由于 $a>b$ 且 $y\geq 0$，得 $by\leq ay$；由于 $a>c$ 且 $z\geq 0$，得 $cz\leq az$。把两式相加，再在两边加上 ax，得

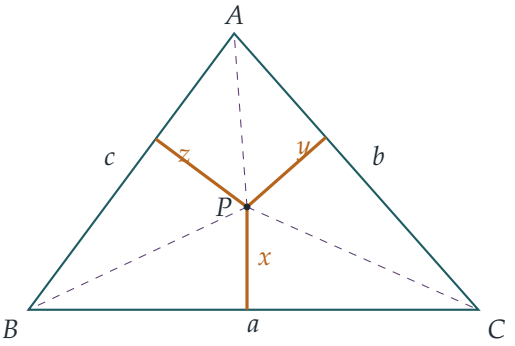
$$ax+by+cz\leq ax+ay+az=a(x+y+z),$$

此即所求。因为 $a-b>0$ 及 $a-c>0$，等号成立当且仅当 $(a-b)y+(a-c)z=0$，即 $y=z=0$。稍后我们会用到这个等号成立的情况。

第 (ii)(a) 部分

把 P 与各顶点连起来，便把 ABC 分成三角形 PBC、PCA 和 PAB（若 P 在某条边上，其中一些可能是退化（degenerate）的）。三角形 PBC 的底为 $BC=a$，高为 x，其余依此类推，所以

$$\Delta=\tfrac{1}{2}ax+\tfrac{1}{2}by+\tfrac{1}{2}cz, \qquad \text{即} \qquad ax+by+cz=2\Delta.$$



反过来，任何满足 $ax+by+cz=2\Delta$ 的 $x,y,z\geq 0$，都恰好是三角形内（或边上）唯一一点 P 到三边的距离：三个三角形的面积 $\frac{1}{2}ax$、$\frac{1}{2}by$、$\frac{1}{2}cz$ 都非负，总和为 Δ，而它们唯一确定 P（它们就是 P 的面积坐标（areal coordinates））。因此，下文可以把问题当作纯代数问题处理：在所有满足 $ax+by+cz=2\Delta$ 的非负 x,y,z 中求最优值。

第 (ii)(b) 部分

依次用第 (i) 部分和 (ii)(a)，

$$x+y+z\geq \frac{ax+by+cz}{a}=\frac{2\Delta}{a}.$$

等号成立当且仅当 $y=z=0$，此时 $ax=2\Delta$。所以总和的最小值为

$$x+y+z=\frac{2\Delta}{a},\qquad \text{取得于}\quad x=\frac{2\Delta}{a},\ y=0,\ z=0.$$

从几何角度看，$y=z=0$ 表示 P 同时在 CA 和 AB 上，所以 $P=A$，而 x 就是从 A 到最长边 BC 的高（altitude）。

第 (iii)(a) 部分

展开三个差的平方，

$$\begin{aligned}(bx-ay)^2+(cy-bz)^2+(az-cx)^2 &= b^2x^2+a^2y^2+c^2y^2+b^2z^2+a^2z^2+c^2x^2\\ &\quad -2(abxy+bcyz+cazx),\end{aligned}$$

而

$$(ax+by+cz)^2=a^2x^2+b^2y^2+c^2z^2+2(abxy+bcyz+cazx).$$

两式相加，交叉项互相抵消，而 a^2,b^2,c^2 其中之一与 x^2,y^2,z^2 其中之一的每个乘积都恰好出现一次，得 $(a^2+b^2+c^2)(x^2+y^2+z^2)$。这就证明了此恒等式对所有实数值皆成立。

第 (iii)(b) 部分

把 $ax+by+cz=2\Delta$ 代入恒等式：

$$(a^2+b^2+c^2)(x^2+y^2+z^2)=(bx-ay)^2+(cy-bz)^2+(az-cx)^2+4\Delta^2\geq 4\Delta^2,$$

因为平方数非负。因此

$$x^2+y^2+z^2\geq\frac{4\Delta^2}{a^2+b^2+c^2}.$$

? **想一想** 下界只有在真正能取到时才是最小值。三个平方项能否在一个真实存在的点 P 同时为零?

等号成立需要 $bx=ay$、$cy=bz$ 及 $az=cx$。这些条件表示 (x,y,z) 是 (a,b,c) 的倍数：若 $x=\lambda a$, 则 $y=\frac{b}{a}x=\lambda b$, $z=\frac{c}{a}x=\lambda c$, 三条方程都成立。由约束条件得 $\lambda(a^2+b^2+c^2)=2\Delta$，所以 $\lambda=\dfrac{2\Delta}{a^2+b^2+c^2}>0$。于是 x,y,z 全为正数，且满足 $ax+by+cz=2\Delta$，故根据 (ii)(a) 的说明，它们对应一个真实的点 P（严格位于三角形内部）。所以平方和的最小值为

$$\frac{4\Delta^2}{a^2+b^2+c^2},\qquad \text{取得于}\quad x=\frac{2a\Delta}{a^2+b^2+c^2},\ y=\frac{2b\Delta}{a^2+b^2+c^2},\ z=\frac{2c\Delta}{a^2+b^2+c^2}.$$

第 (iv) 部分

? **想一想** 第 (i) 部分用到 a 是最长边。如果改用最短边 c，会怎样?

由于 $a>c$、$b>c$ 且 $x,y\geq 0$, 把第 (i) 部分的论证反过来, 得 $ax+by+cz\geq c(x+y+z)$, 等号成立当且仅当 $x=y=0$。所以

$$x+y+z\leq\frac{2\Delta}{c}.$$

此外，由于 $x,y,z\geq 0$,

$$x^2+y^2+z^2=(x+y+z)^2-2(xy+yz+zx)\leq(x+y+z)^2,$$

等号成立当且仅当 $xy+yz+zx=0$。把两者结合（由于两边都非负，平方是安全的），

$$x^2+y^2+z^2\leq(x+y+z)^2\leq\frac{4\Delta^2}{c^2}.$$

当 $x=y=0$ 时，两个不等式都取等号：此时 $xy+yz+zx=0$ 且 $cz=2\Delta$。所以平方和的最大值为

$$\frac{4\Delta^2}{c^2},\qquad \text{取得于}\quad x=0,\ y=0,\ z=\frac{2\Delta}{c},$$

即 $P=C$，也就是最短边所对的顶点。（第二个不等式只在 $x=y=0$ 时取等号，所以这是取得最大值的唯一一点。）

验算

取边长为 $a = 7$、$b = 6$、$c = 5$ 的锐角三角形。海伦公式（Heron's formula）给出 $\Delta = \sqrt{9 \cdot 2 \cdot 3 \cdot 4} = 6\sqrt{6} \approx 14.70$。答案于是为：$x + y + z$ 的最小值是 $12\sqrt{6}/7 \approx 4.20$；$x^2 + y^2 + z^2$ 的最小值是 $864/110 = 432/55 \approx 7.85$；最大值是 $864/25 = 34.56$。大小次序合理：在 $P = C$ 时，平方和就是从 C 出发的高的平方，即 $(12\sqrt{6}/5)^2 = 34.56$。我们用电脑检验了大量随机锐角三角形，每个三角形取数以千计的随机点 P，证实了 $ax + by + cz = 2\Delta$ 及全部三个界，而极值分别在 A、在距离与 (a, b, c) 成比例的点，以及在 C 取得。

常见错误

- 给出下界却没有证明它能取到。在 (iii)(b) 中，必须检查等号成立的点的距离非负且满足约束条件，才能确定它是一个真实存在的 P。
- 混淆顶点：$x + y + z$ 的最小值在 A（最长边所对的顶点）取得，但 $x^2 + y^2 + z^2$ 的最大值在 C（最短边所对的顶点）取得。
- 在 (iv) 中忽略条件 $x, y, z \geq 0$；没有这个条件，$x^2 + y^2 + z^2 \leq (x + y + z)^2$ 并不成立。

STEP 之外

(iii)(a) 中的恒等式是拉格朗日恒等式（Lagrange's identity），它证明了三维空间中的柯西－施瓦茨不等式 $(\mathbf{u} \cdot \mathbf{v})^2 \leq |\mathbf{u}|^2 |\mathbf{v}|^2$：最小值点就是 (x, y, z) 与 (a, b, c) 平行之处，亦即从原点到平面 $ax + by + cz = 2\Delta$ 的垂足。像 $x^2 + y^2 + z^2$ 这样的凸函数（convex function）在三角形上的最大值总是在某个顶点取得，这让我们初窥线性最优化与凸最优化的世界。(iii)(b) 中找到的点是三角形的类似重心（symmedian point，又称勒穆瓦纳点（Lemoine point））。想进一步了解点积与投影，可参阅 Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）。

详细解答

STEP II 2022，第 6 题

纯数学 · 正交轨线

STEP II 2022，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

两条曲线在某点互相垂直，当且仅当它们在该点的斜率（gradient）之积为 -1。因此，对于一整族曲线，我们需要一条在点 (x, y) 成立的斜率公式，而且无论族中哪一条曲线经过该点都适用。这意味着要在曲线方程与其导数（derivative）之间消去参数（parameter）（a、b、c 或 k），只剩下一条只含 x、y 和 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 的微分方程（differential equation）。然后把斜率 m 换成 $-1/m$ 再求解，便得到与之垂直的曲线族。

解答

全题中 $x > 0$ 且 $y > 0$。

第 (i)(a) 部分

把 $x^2 + y^2 = 2ax$ 对 x 作隐函数微分（implicit differentiation），得

$$2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 2a.$$

这式子仍然含有 a，而 a 随曲线而变。由曲线方程得 $2a = (x^2 + y^2)/x$，代入后乘以 x，

$$2x^2 + 2xy\frac{\mathrm{d}y}{\mathrm{d}x} = x^2 + y^2 \quad \Longrightarrow \quad 2xy\frac{\mathrm{d}y}{\mathrm{d}x} = y^2 - x^2,$$

即所求。

对于第二族，把 $x^2 + y^2 = 2by$ 微分（记住右方亦通过 y 依赖于 x），得

$$2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 2b\frac{\mathrm{d}y}{\mathrm{d}x}.$$

此时 $2b = (x^2+y^2)/y$；代入后乘以 y，

$$2xy + 2y^2\frac{\mathrm{d}y}{\mathrm{d}x} = (x^2+y^2)\frac{\mathrm{d}y}{\mathrm{d}x} \quad\Longrightarrow\quad (x^2-y^2)\frac{\mathrm{d}y}{\mathrm{d}x} = 2xy.$$

第 (i)(b) 部分

假设两族中各有一条曲线经过 (x,y)，且 $y \neq x$。由于 $x, y > 0$，故 $2xy \neq 0$；又因 $y \neq x$ 且两者皆为正，故 $x^2 - y^2 \neq 0$。因此两条微分方程都可以相除：两个斜率为

$$m_1 = \frac{y^2-x^2}{2xy}, \qquad m_2 = \frac{2xy}{x^2-y^2}, \qquad \text{故} \qquad m_1 m_2 = \frac{y^2-x^2}{x^2-y^2} = -1.$$

两个斜率都是有限值，且乘积为 -1，所以两条切线（tangent）互相垂直。

> ? **想一想** 在直线 $y = x$ 上会出现甚么问题？这是否表示结果在那里不成立？

在 $y = x$ 上，第二条方程变成 $0 \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = 2x^2 \neq 0$，没有任何有限斜率能满足它：第二条曲线的切线是铅垂的。出问题的是代数，而不是几何，所以我们直接考察。

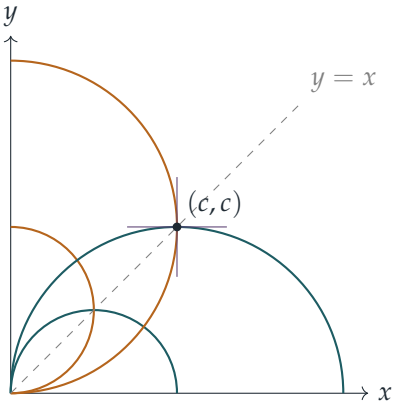
若 (c,c) 在 $x^2+y^2 = 2ax$ 上，则 $2c^2 = 2ac$，故 $a = c$；同理 $b = c$。第一条曲线是圆心 $(c,0)$、半径 c 的圆，而 (c,c) 是它的最高点，所以该点的切线是水平的（事实上当 $y = x$ 时 $m_1 = 0$）：

$$y = c.$$

第二条曲线是圆心 $(0,c)$、半径 c 的圆，而 (c,c) 是它的最右点，所以其切线是铅垂的：

$$x = c.$$

水平线与铅垂线互相垂直，所以**是的**：在每一个交点，包括 $y = x$ 上的交点，切线都互相垂直。



第一族的圆（青绿色）在原点与 y 轴相切；第二族的圆（琥珀色）在原点与 x 轴相切。

第 (ii) 部分

对于 $y = c\ln x$，有 $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{c}{x}$。为了消去 c，注意在这类曲线上位于第一象限的点，$y > 0$ 迫使 $\ln x \neq 0$，故 $c = y/\ln x$，且

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{x\ln x}.$$

这个斜率永不为零，所以经过同一点的垂直曲线的斜率必为 $-x\ln x/y$。因此第二族满足

$$y\frac{\mathrm{d}y}{\mathrm{d}x} = -x\ln x.$$

分离变量（separating variables），并以 $u = \ln x$、$\mathrm{d}v = x\,\mathrm{d}x$ 对 $x\ln x$ 作分部积分（integration by parts），

$$\int y\,\mathrm{d}y = -\int x\ln x\,\mathrm{d}x \quad\Longrightarrow\quad \tfrac{1}{2}y^2 = -\tfrac{1}{2}x^2\ln x + \tfrac{1}{4}x^2 + C.$$

乘以 4，第二族为

$$2y^2 + 2x^2\ln x - x^2 = k,$$

每个常数 k 对应一条曲线（只计 $x > 0$、$y > 0$ 的部分）。反过来，对此式作隐函数微分，得 $4y\frac{\mathrm{d}y}{\mathrm{d}x} + 4x\ln x + 2x - 2x = 0$，正是上面的方程，所以凡这些曲线与某条 $y = c\ln x$ 相交之处，斜率之积都是 $\frac{y}{x\ln x}\cdot\frac{-x\ln x}{y} = -1$。

第 (iii) 部分

把 $y^2 = 4k(x+k)$ 微分，得 $2y\frac{\mathrm{d}y}{\mathrm{d}x} = 4k$，由于 $y > 0$，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2k}{y}.$$

想一想 这里只有一族曲线。那么“第二条曲线”从何而来?

固定第一象限中一点 (x, y)，问族中哪些曲线经过它。k 的值必须满足

$$4k^2 + 4xk - y^2 = 0,$$

这是 k 的二次方程（quadratic）。其判别式（discriminant）$16x^2 + 16y^2$ 为正，所以有两个相异实根 k_1、k_2，而它们的积为

$$k_1k_2 = -\frac{y^2}{4}.$$

（此值为负，所以两根皆非零，符合曲线族的要求。）二次方程最多只有两个根，所以若族中两条不同的曲线在 (x, y) 相交，它们的参数必定正是 k_1 和 k_2。它们在该点的斜率为 $2k_1/y$ 和 $2k_2/y$，乘积为

$$\frac{4k_1k_2}{y^2} = \frac{4}{y^2}\cdot\left(-\frac{y^2}{4}\right) = -1.$$

所以在每一个这样的点，切线都互相垂直，即所求。（这些曲线是焦点（focus）在原点的抛物线（parabola）；这族曲线就是它自己的正交轨线。）

验算

- 在 (i) 中，取 $(x,y)=(1,2)$：则 $a=\frac{5}{2}$、$b=\frac{5}{4}$，斜率为 $m_1=\frac{3}{4}$ 和 $m_2=-\frac{4}{3}$，乘积为 -1。
- 在 (ii) 中，在 $(e,1)$ 处，对数曲线有 $c=1$，斜率为 $1/e$；第二族的公式给出斜率 $-e\cdot 1/1=-e$。乘积同样是 -1。
- 在 (iii) 中，在 $(0,2)$ 处，二次方程为 $4k^2-4=0$，故 $k=\pm 1$：抛物线 $y^2=4(x+1)$ 和 $y^2=-4(x-1)$，在该点的斜率分别为 1 和 -1。
- 四个结果亦已用电脑在数千个随机点上验证，所用的是曲线本身的数值导数。

常见错误

- 微分 $2by$ 时漏掉右方的因子 $\frac{\mathrm{d}y}{\mathrm{d}x}$。
- 在 (ii) 中，把垂直斜率写成 $-x/c$，然后在仍含 c 的情况下积分：c 属于第一条曲线，必须在求解前消去。
- 在 (i)(b) 中除以 x^2-y^2 而没有说明它为何非零，因而忽略了直线 $y=x$ 需要另行处理。

STEP 之外

处处以直角相交的曲线族称为**正交轨线**（orthogonal trajectories）。它们在物理学中以场线和等势线的形式出现，在复分析（complex analysis）中亦然：解析函数（analytic function）的实部与虚部的等高线总是互相垂直，这就是为甚么 (i) 中的圆来自 $1/z$，而 (iii) 中自身正交的抛物线来自 $\sqrt{z}$。Boyce 和 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）讨论了正交轨线，而 Needham 的 *Visual Complex Analysis*（《直观复分析》）则把复分析的图像解释得精彩动人。

详细解答

STEP II 2022，第 7 题

纯数学 · 复系数多项式的根

STEP II 2022，第 7 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

方程 $\mathrm{h}(w) = 0$ 可以很自然地写成 $w^5(nw+1) = -(w+n)$，两边取模（modulus）后，就变成一个只关于 $|w|$ 的命题。一旦知道所有根都在单位圆（unit circle）上，每个非实根都会与它的共轭（conjugate）配成一对，而这个共轭恰好也是它的倒数（reciprocal）。由此便得到二次因式（quadratic factor），之后只需比较系数（comparing coefficients），再作一个正负号的论证即可。

解答

第 (i) 部分

(a) 把 $\mathrm{h}(w) = 0$ 的各项分组为

$$w^5(nw+1) = -(w+n).$$

两个括号都不可能为零。若 $w = -1/n$，则 $\mathrm{h}(w) = n - \frac{1}{n} \neq 0$；若 $w = -n$，则 $\mathrm{h}(w) = n^5(n^2-1) \neq 0$，因为 $n \geq 2$。两边取模再平方，得

$$|w^5|^2\,|nw+1|^2 = |w+n|^2.$$

对任何复数 u 及实数 c，有 $|u+c|^2 = (u+c)(\bar{u}+c) = |u|^2 + 2c\,\mathrm{Re}(u) + c^2$。所以

$$|w+n|^2 = |w|^2 + 2n\,\mathrm{Re}(w) + n^2 = \mathrm{f}(w),$$
$$|nw+1|^2 = n^2|w|^2 + 2n\,\mathrm{Re}(w) + 1 = \mathrm{g}(w).$$

两者皆为正，故 $|w^5|^2 = \mathrm{f}(w)/\mathrm{g}(w)$。取正平方根，便得 $|w^5| = \sqrt{\mathrm{f}(w)/\mathrm{g}(w)}$，即为所求。

(b) 两者相减时，$\mathrm{Re}(w)$ 项互相抵消：

$$\mathrm{f}(w)-\mathrm{g}(w)=n^2+|w|^2-n^2|w|^2-1=(n^2-1)(1-|w|^2).$$

用反证法（proof by contradiction），假设 $|w|<1$。由于 $n^2-1>0$，得 $\mathrm{f}(w)>\mathrm{g}(w)>0$，所以 $\mathrm{f}(w)/\mathrm{g}(w)>1$。由 (a)，$|w|^5>1$，故 $|w|>1$，与 $|w|<1$ 矛盾。因此 $|w|\geq 1$。

(c) 同样的论证反过来亦成立。若 $|w|>1$，则 $\mathrm{f}(w)<\mathrm{g}(w)$，所以 $|w|^5<1$，即 $|w|<1$，矛盾。故 $|w|\leq 1$ 亦成立，结合 (b)，得 $|w|=1$。

想一想 有没有更简洁的理由说明 $|w|\leq 1$?

有的。系数 $n,1,0,0,0,1,n$ 倒过来读也一样，所以 $z^6\,\mathrm{h}(1/z)=\mathrm{h}(z)$。因此若 w 是根，$1/w$ 也是根；把 (b) 应用于 $1/w$，便得 $|w|\leq 1$。

第 (ii) 部分

(a) 由于 h 是实系数（real coefficients）多项式，它的非实根必成共轭对出现。六个根互不相同且都不是实数，所以它们组成三对 w_k, $\bar{w}_k$（$k=1,2,3$）。由第 (i) 部分，$|w_k|=1$，故 $w_k\bar{w}_k=1$，而

$$(z-w_k)(z-\bar{w}_k)=z^2-(w_k+\bar{w}_k)z+w_k\bar{w}_k=z^2-a_kz+1,$$

其中 $a_k=2\,\mathrm{Re}(w_k)$ 为实数。多项式 h 的次数（degree）为 6，首项系数（leading coefficient）为 n，所以

$$\mathrm{h}(z)=n(z^2-a_1z+1)(z^2-a_2z+1)(z^2-a_3z+1).$$

(b) 乘积中的 z^5 项，来自从其中一个因式取 $-a_kz$、从另外两个因式各取 z^2，所以 z^5 的系数是 $-n(a_1+a_2+a_3)$。在 h 中这个系数是 1，所以

$$a_1+a_2+a_3=-\frac{1}{n}.$$

(c) 先把其中两个因式相乘：

$$(z^2-a_1z+1)(z^2-a_2z+1)=z^4-(a_1+a_2)z^3+(2+a_1a_2)z^2-(a_1+a_2)z+1.$$

再乘以 z^2-a_3z+1，z^3 项来自 $z^3\times 1$、$z^2\times(-a_3z)$ 及 $z\times z^2$：

$$-(a_1+a_2)-a_3(2+a_1a_2)-(a_1+a_2)=-2(a_1+a_2+a_3)-a_1a_2a_3.$$

h 中没有 z^3 项，所以 $a_1a_2a_3=-2(a_1+a_2+a_3)$，再由 (b)，

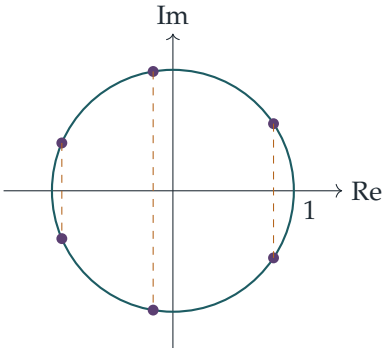
$$a_1a_2a_3=\frac{2}{n}.$$

(d)

想一想 a_k 的正负号能告诉我们甚么关于实部的信息?

每对共轭根 $w_k, \bar{w}_k$ 有相同的实部（real part）$\mathrm{Re}(w_k) = a_k/2$。所以实部为负的根的数目，是负数 a_k 的数目的两倍。

由于 $a_1a_2a_3 = 2/n > 0$，所有 a_k 都不为零，而且其中负数的个数是偶数：零个或两个。它们不可能全部为正，因为 $a_1 + a_2 + a_3 = -1/n < 0$。所以恰好有两个 a_k 为负、一个为正，六个根之中有**四个**的实部为负。



上图显示 $n = 2$ 的情况。六个根以三对共轭根的形式落在单位圆上，其中两对位于虚轴（imaginary axis）的左边。

验算

- 作为“没有实根”的合理性检查：$\mathrm{h}(1) = 2n + 2$ 和 $\mathrm{h}(-1) = 2n - 2$ 都不为零。由于所有根都在单位圆上，唯一可能的实根只有 ± 1。
- 当 $n = 2$ 时，数值求解得 $a_k \approx -1.8368,\ -0.3272,\ 1.6640$。它们之和为 $-0.5 = -\frac{1}{2}$，乘积为 $1.000 = \frac{2}{2}$，且有两个为负，与 (b)、(c)、(d) 相符。以电脑检验 $n = 2, \ldots, 29$，确认每次所有根的模都是 1，而且都有四个根的实部为负。

常见错误

- 未检查 $w = -1/n$ 不是根，便直接除以 $nw + 1$。
- 在 (c) 中漏掉构成 z^3 的三种方式之一，特别是来自 $2 + a_1a_2$ 中常数 2 的那一项 $-2a_3$。
- 在 (d) 中数了负数 a_k 的个数，却忘记每一个都对应两个根。

STEP 之外

系数倒过来读也一样的多项式称为回文多项式（palindromic polynomial，又称自倒多项式，self-reciprocal）。它的根以 $w, 1/w$ 成对出现，而代换 $u = z + 1/z$ 可以把它的次数减半。这里的 $a_k = w_k + 1/w_k$ 正正就是这个 u。探讨所有根何时都落在单位圆上的问题，会引出 Schur–Cohn 判别法，以及复分析（complex analysis）中的儒歇定理（Rouché's theorem）等结果。Needham 的 *Visual Complex Analysis*（《直观复分析》）是从几何角度欣赏这些概念的绝佳读物。

详细解答

STEP II 2022，第 8 题

纯数学 · 矩阵的不变直线

> **STEP II 2022，第 8 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：
>
> www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

一条穿过原点的直线，如果线上每一点都被映射到同一条直线上的点，这条直线就是不变直线（invariant line）；因此只需看一个方向向量（direction vector）被映射到哪里便足够。把直线写成 $y = mx$，并要求 $\mathbf{M}\binom{1}{m}$ 是 $\binom{1}{m}$ 的倍数，整个问题就化为一个关于 m 的二次方程（quadratic equation）。之后，第 (ii) 和第 (iii) 部分都是关于这个方程两根之和与两根之积的问题；而铅垂线没有斜率（gradient），每次都需要另行处理。

解答

第 (i) 部分

考虑直线 $y = mx$，其方向为 $\binom{1}{m}$。它的像是

$$\mathbf{M}\begin{pmatrix}1\\m\end{pmatrix} = \begin{pmatrix}a+bm\\c+dm\end{pmatrix}.$$

由于 $\mathbf{M}$ 是线性（linear）的，这条直线是不变直线，当且仅当这个像位于 $y = mx$ 上，即当且仅当

$$c + dm = m(a + bm) \quad \Longleftrightarrow \quad bm^2 + (a-d)m - c = 0.$$

所以每条非铅垂的不变直线，其斜率都满足这个方程；而方程的每个根都给出一条不变直线。

在把斜率称为“两个根”之前，我们必须先确认这真的是一个二次方程。y 轴的方向为 $\binom{0}{1}$，其像为 $\binom{b}{d}$；这个像位于 y 轴上，当且仅当 $b = 0$。所以如果两条不变直线都不是 y 轴，则 $b \neq 0$。这时方程是真正的二次方程，最多只有两个根。两条不变直线都不是铅垂

线，而且斜率不同，所以它们的斜率是两个相异的根。这就表示它们的斜率正是方程的两个根，得证。

如果其中一条不变直线是 y 轴，则 $b = 0$，方程退化为 $(a-d)m - c = 0$。由于 $a \neq d$（这正是用到这个条件的地方），另一条不变直线的斜率是

$$m = \frac{c}{a-d}.$$

第 (ii) 部分

情况 $b \neq 0$。设两条直线的斜率为 m_1 和 m_2。由第 (i) 部分，

$$m_1 + m_2 = -\frac{a-d}{b}, \qquad m_1 m_2 = -\frac{c}{b}.$$

若两条直线与 x 轴所成的角分别为 α_1 和 α_2，则

$$\tan(\alpha_1 - \alpha_2) = \frac{m_1 - m_2}{1 + m_1 m_2}.$$

两直线之间的夹角为 45°，即上式等于 ± 1，所以 $(m_1 - m_2)^2 = (1 + m_1 m_2)^2$。把左边写成 $(m_1 + m_2)^2 - 4m_1 m_2$，得

$$\frac{(a-d)^2}{b^2} + \frac{4c}{b} = \left(1 - \frac{c}{b}\right)^2.$$

两边乘以 b^2，得 $(a-d)^2 + 4bc = (b-c)^2$，即

$$(a-d)^2 = (b-c)^2 - 4bc.$$

情况 $b = 0$。其中一条直线是 y 轴，另一条的斜率为 $c/(a-d)$。要使它与 y 轴成 45°，它的斜率必须是 ± 1，所以 $c = \pm(a-d)$，从而 $(a-d)^2 = c^2$。当 $b = 0$ 时，所求结果的右边也等于 c^2，所以结果在这个情况下同样成立。

第 (iii) 部分

> ? **想一想** “与 $y = x$ 成相等的角”用斜率来说是甚么意思?

两条穿过原点的不同直线与 $y = x$ 成相等的角，当且仅当其中一条是另一条关于 $y = x$ 的镜像（mirror image）。关于 $y = x$ 的反射会把坐标对调，因此它把 $y = mx$ 映射到 $x = my$；当 $m \neq 0$ 时，这就是 $y = \frac{1}{m}x$。它把 x 轴映射到 y 轴，并把 y 轴映射到 x 轴。

情况 $b \neq 0$。两条直线都不是铅垂线。如果其中一条的斜率为 0，它的镜像便是 y 轴，而 y 轴并非不变直线，所以两条直线不会对称。除此以外，条件就是 $m_2 = 1/m_1$，即 $m_1 m_2 = 1$。由第 (i) 部分，这就是 $-c/b = 1$，即 $b + c = 0$。

情况 $b = 0$。其中一条直线是 y 轴，所以另一条必须是 x 轴：$c/(a-d) = 0$，即 $c = 0$。在 $b = 0$ 的前提下，这同样就是条件 $b + c = 0$。

所以在所有情况下，充要条件都是

$$b + c = 0, \quad \text{即} \quad c = -b.$$

（当矩阵是对称矩阵（symmetric matrix），即 $b = c$ 时，不变直线反而互相垂直。这里则是两个非对角元素互为相反数。）

第 (iv) 部分

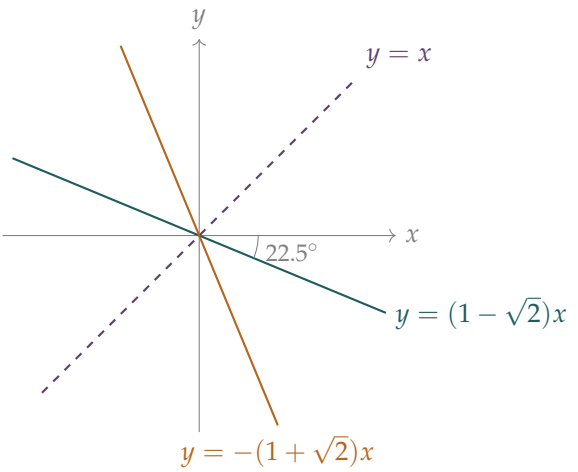
把 $c = -b$ 代入第 (ii) 部分的条件：

$$(a-d)^2 = (2b)^2 + 4b^2 = 8b^2.$$

我们需要 $b \neq 0$：若 $b = c = 0$，则 $(a-d)^2 = 0$，与 $a \neq d$ 矛盾。取 $b = 1$、$c = -1$ 及 $a - d = 2\sqrt{2}$，例如 $a = \sqrt{2}$、$d = -\sqrt{2}$：

$$\mathbf{M} = \begin{pmatrix} \sqrt{2} & 1 \\ -1 & -\sqrt{2} \end{pmatrix}.$$

第 (i) 部分的二次方程是 $m^2 + 2\sqrt{2}\,m + 1 = 0$，有相异实根 $m = -\sqrt{2} \pm 1$。所以这个矩阵恰好有两条不变直线，与题目的假设相符。两根之积为 1，所以两条直线关于 $y = x$ 互为镜像。由于 $\tan 22.5^\circ = \sqrt{2} - 1$ 及 $\tan 67.5^\circ = \sqrt{2} + 1$，它们与 x 轴所成的角分别为 -22.5° 和 -67.5°。因此两条直线相交成 45°，而每条都与 $y = x$ 成 67.5°。



验算

- 对于第 (iv) 部分的例子，$(a-d)^2 = 8$，而 $(b-c)^2 - 4bc = 4+4 = 8$，所以第 (ii) 部分的条件成立。我们也可以直接验证其中一条直线：$\mathbf{M}\begin{pmatrix}1\\1-\sqrt{2}\end{pmatrix} = \begin{pmatrix}\sqrt{2}+1-\sqrt{2}\\-1-\sqrt{2}+2\end{pmatrix} = \begin{pmatrix}1\\1-\sqrt{2}\end{pmatrix}$，所以 $y = (1-\sqrt{2})x$ 上的每一点其实都是不动点。
- 第 (i) 部分的一个简单特例：对于 $\mathbf{M} = \begin{pmatrix}2 & 0\\0 & 1\end{pmatrix}$，不变直线就是两条坐标轴。$y$ 轴情况的公式给出另一条直线的斜率为 $0/(2-1) = 0$，正确无误。
- 第 (i) 至 (iii) 部分已用电脑以数以千计的随机矩阵验证：特征向量的方向满足该二次方程，成 $45°$ 的直线对满足该恒等式，而且两条直线关于 $y = x$ 对称，当且仅当 $b + c = 0$。

常见错误

- 忘记铅垂线。二次方程只能反映斜率，所以每一部分都必须另行处理 $b = 0$ 的情况。
- 在第 (iii) 部分写成 $m_1 = -m_2$。这是关于坐标轴的对称。关于 $y = x$ 的对称需要 $m_1 m_2 = 1$。
- 使用 $\tan\theta = 1$ 而漏掉 $\pm$。把式子平方成 $(m_1 - m_2)^2 = (1 + m_1 m_2)^2$，便可同时处理两个符号。

STEP 之外

穿过原点的不变直线，就是 $\mathbf{M}$ 的特征向量（eigenvector）方向；而第 (i) 部分的二次方程，就是把特征向量方程 $\mathbf{Mv} = \lambda\mathbf{v}$ 中的 λ 消去后所得的结果。实对称矩阵的特征向量必定互相垂直（谱定理（spectral theorem））。这道题则展示了反对称的非对角元素如何令特征向量反而关于 $y = x$ 对称。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）对特征向量和谱定理有系统的讲解。

详细解答

STEP II 2022，第 9 题

力学 · 摩擦力与力矩

STEP II 2022，第 9 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

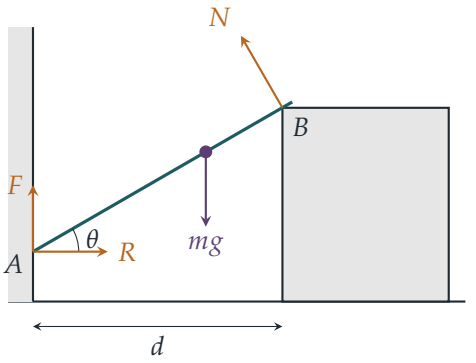
木板是轻的，所以只有三样东西作用于它：质点、棱柱边缘处的反作用力（reaction），以及墙壁。对木板与墙壁的接触点取力矩（moments），可以一次过消去墙壁的两个力，并立即得出边缘处的反作用力。然后沿铅垂方向分解（resolve）便得出摩擦力（friction），而之后每一部分都只是把摩擦不等式 $|F| \leq \mu R$ 重新整理而已。

解答

模型。木板是一根轻刚性杆，质量是一个重量为 mg 的质点，棱柱的边缘是光滑的，而墙壁是粗糙的，摩擦系数（coefficient of friction）为 μ。设 A 为木板与墙壁的接触点，B 为木板搁在边缘上的点。由于 B 与墙壁的水平距离为 d，

$$AB = d\sec\theta.$$

在 B 处木板搁在一条边缘上，所以光滑反作用力 N 垂直于木板。在 A 处，墙壁提供水平的法向反作用力 R 和铅垂的摩擦力 F，我们取向上为正（F 为负即表示它向下作用）。



对 A 取力矩。墙壁的两个力都通过 A。重量作用于沿木板距离 x 处，所以其力臂为 $x\cos\theta$；N 在距离 $d\sec\theta$ 处垂直于木板：

$$N\,d\sec\theta = mg\,x\cos\theta \quad\Longrightarrow\quad N = \frac{mg\,x\cos^2\theta}{d}.$$

分解。N 的方向与铅垂线成角 θ，并向墙壁倾斜。沿水平及铅垂方向，

$$R = N\sin\theta = \frac{mg\,x\cos^2\theta\sin\theta}{d},$$
$$F = mg - N\cos\theta = mg\left(1 - \frac{x\cos^3\theta}{d}\right).$$

注意 $R > 0$，所以木板确实压在墙壁上。

第 (i) 部分

若 $x = d\sec^3\theta$，则 $x\cos^3\theta/d = 1$，所以 $F = 0$：墙壁处没有摩擦力作用。

第 (ii) 部分

在有摩擦的情况下，平衡（equilibrium）要求 $|F| \le \mu R$。

若 $x > d\sec^3\theta$，则 $x\cos^3\theta/d > 1$，所以 $F < 0$（摩擦力向下作用；没有它的话，端点 A 会沿墙壁向上滑）。条件 $-F \le \mu R$ 即为

$$mg\left(\frac{x\cos^3\theta}{d} - 1\right) \le \mu\,\frac{mg\,x\cos^2\theta\sin\theta}{d}.$$

乘以正数 $d\sec^3\theta/(mg)$，得

$$x - d\sec^3\theta \le \mu\,x\tan\theta, \qquad \text{因此} \qquad \mu \ge \frac{x - d\sec^3\theta}{x\tan\theta},$$

这里除以了 $x\tan\theta > 0$（因为 θ 是锐角）。这就是所求的结果。

若 $x < d\sec^3\theta$，则 $F > 0$，而条件 $F \le \mu R$ 同样地变成

$$\mu \ge \frac{d\sec^3\theta - x}{x\tan\theta}.$$

想一想 摩擦力向哪个方向作用？我们怎样不靠猜测就知道？

我们不需要猜。取 F 向上为正，铅垂方程便会告诉我们它的正负号。摩擦力总可以写成 $|F| \le \mu R$，而 F 的正负号只是决定我们得到两个不等式中的哪一个。

第 (iii) 部分

(ii) 的两种情况都是说 $|x - d\sec^3\theta| \le \mu x\tan\theta$，这等同于以下两式同时成立：

$$x - d\sec^3\theta \le \mu x\tan\theta \qquad 及 \qquad d\sec^3\theta - x \le \mu x\tan\theta.$$

第二式给出 $d\sec^3\theta \le x(1+\mu\tan\theta)$。由于 $1+\mu\tan\theta > 0$，

$$\frac{x}{d} \ge \frac{\sec^3\theta}{1+\mu\tan\theta}.$$

第一式给出 $x(1-\mu\tan\theta) \le d\sec^3\theta$。若 $\mu < \cot\theta$，则 $\mu\tan\theta < 1$，所以 $1-\mu\tan\theta > 0$，我们可以相除而不必反转不等号：

$$\frac{x}{d} \le \frac{\sec^3\theta}{1-\mu\tan\theta}.$$

（若 $\mu \ge \cot\theta$，则没有上界：无论质量位于木板上多高的位置，摩擦力都能把它承住。）

第 (iv) 部分

质量位于墙壁与边缘之间，当且仅当 $x < AB = d\sec\theta$。结合 (iii) 的下界，

$$\frac{d\sec^3\theta}{1+\mu\tan\theta} \le x < d\sec\theta.$$

除以 $d\sec\theta > 0$ 并乘以 $1+\mu\tan\theta > 0$，

$$\sec^2\theta < 1+\mu\tan\theta \quad\Longleftrightarrow\quad \tan^2\theta < \mu\tan\theta,$$

这里用了 $\sec^2\theta = 1+\tan^2\theta$。由于 $\tan\theta > 0$，我们可以除以它，得 $\tan\theta < \mu$，正如所求。

验算

- 取 $\theta = 45°$、$d = 1$ 及 $\mu = 0.5$。则 $\sec^3\theta = 2\sqrt{2} \approx 2.83$，而 (iii) 容许 $1.89 \le x \le 5.66$。例如在 $x = 2$ 处，$F = mg(1-2/2\sqrt{2}) \approx 0.29mg$，$N = mg$，$R = mg/\sqrt{2}$，而

$\mu R \approx 0.35mg$，所以平衡成立。这时质量位于边缘之外（$AB = \sqrt{2}$），与 (iv) 一致，因为 $\tan 45^\circ = 1 > \mu$。

- 极限情况 $\mu = 0$（光滑墙壁）：(iii) 中的上下界收缩为 $x = d\sec^3\theta$，即 (i) 中唯一的位置。
- 以电脑对大量随机的 θ、d、μ 及 x 值进行检验，确认各力平衡（包括对其他点取力矩），并且 $|F| \le \mu R$ 成立当且仅当 (iii) 的不等式成立。

常见错误

- 把边缘处的反作用力当作铅垂。木板是搁在一条边缘上，所以光滑反作用力垂直于木板，而不是垂直于棱柱的顶面。
- 在力矩方程中，以 d 而非 $d\sec\theta$ 作为距离 AB。
- 把 $x(1 - \mu\tan\theta) \le d\sec^3\theta$ 相除时，没有检查 $1 - \mu\tan\theta$ 的正负号；这正是条件 $\mu < \cot\theta$ 出现的原因。

STEP 之外

这块木板是静定的（statically determinate）：三个未知反作用力，三条平衡方程，所以摩擦力被完全确定。若再加一个支点，方程便不再能确定各力；这时工程师会引入梁的弹性来决定负荷如何分配。(iii) 中的上下界是“平衡区域”的一个小例子，这个主题在稳定性的研究中会再次出现。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）对刚体静力学和摩擦有细致的论述。

详细解答

STEP II 2022，第 10 题

力学 · 斜面上的抛体

STEP II 2022，第 10 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

第 (i) 部分是标准的轨迹方程（trajectory equation），窍门在于利用 $\sec^2\alpha = 1 + \tan^2\alpha$，把它看成关于 $\tan\alpha$ 的二次方程（*quadratic*）。目标点能够被击中，当且仅当这个二次方程有实根，于是整道题便化为一个判别式（discriminant）条件。用水平坐标 (x, y) 写出来，这个条件原来描述的是一个圆盘（disc），而之后每一部分都只是读出一条直线与一个圆的相交位置。

解答

全题我们把抛射体（projectile）视为只受重力作用而自由运动的质点（particle），不计空气阻力（air resistance），并取 $0 < \theta < \frac{\pi}{2}$。

第 (i) 部分

以发射点为基准量度水平距离和高度。在时间 t，

$$s = ut\cos\alpha, \qquad h = ut\sin\alpha - \tfrac{1}{2}gt^2.$$

由水平运动得 $t = \dfrac{s}{u\cos\alpha}$，代入后得

$$h = s\tan\alpha - \frac{gs^2}{2u^2\cos^2\alpha},$$

即为所求。

第 (ii) 部分

设 $P=(x,y,y\tan\theta)$，并记 $s=\sqrt{x^2+y^2}$ 为它与大炮的水平距离。由于大炮可以指向任何方向，我们把它瞄准在经过 P 的铅垂平面内；运动便在该平面内进行，而由第 (i) 部分，我们需要一个角 α（可以是负数，因为大炮可以向下指），使得

$$y\tan\theta = s\tan\alpha - \frac{gs^2}{2u^2}\left(1+\tan^2\alpha\right).$$

若 $s=0$，则 P 就是原点，垂直向上发射便可到达；而由于 $\sec^2\theta>\tan^2\theta$，下面的不等式在该点亦成立。对于 $s>0$，设 $T=\tan\alpha$，它可以是任何实数。条件是

$$\frac{gs^2}{2u^2}T^2 - sT + \left(y\tan\theta+\frac{gs^2}{2u^2}\right)=0$$

有实根 T，即它的判别式非负：

$$s^2 \geq 4\cdot\frac{gs^2}{2u^2}\left(y\tan\theta+\frac{gs^2}{2u^2}\right) \iff u^4 \geq 2gu^2y\tan\theta+g^2s^2,$$

这是除以 $s^2>0$ 再乘以 u^4 的结果。除以 g^2 并利用 $s^2=x^2+y^2$，

$$x^2+y^2+\frac{2u^2\tan\theta}{g}y \leq \frac{u^4}{g^2}.$$

对 y 配方（completing the square），并利用 $1+\tan^2\theta=\sec^2\theta$，

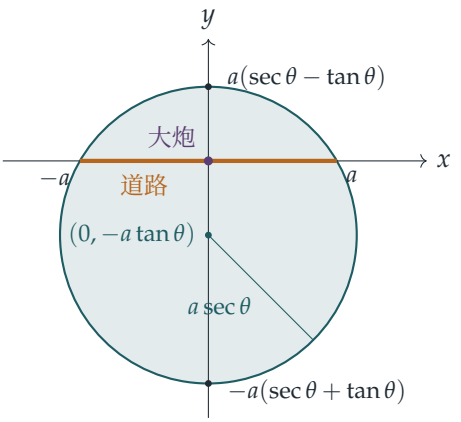
$$x^2+\left(y+\frac{u^2\tan\theta}{g}\right)^2 \leq \frac{u^4}{g^2}\left(1+\tan^2\theta\right)=\frac{u^4\sec^2\theta}{g^2},$$

这正是所求的条件。

? **想一想** 抛射体会不会在到达 P 之前就撞上斜坡地面?

不会。在发射所在的铅垂平面内，抛射体的高度减去地面的高度，是关于水平距离的二次式，首项系数为负（抛物线减去一条直线），而它在大炮处和 P 处均为零。所以它在两者之间为正：抛射体一直保持在地面之上，直至落在 P。

由此起记 $a=u^2/g$。可到达的点，就是其水平投影位于以 $(0,-a\tan\theta)$ 为圆心、$a\sec\theta$ 为半径的圆盘内的那些点。



上图（以 $\theta = 30^\circ$ 绘制）显示从上方俯视的可到达区域；“沿斜面向上”就是 y 的正方向。

第 (iii) 部分

沿斜面正上方即 $x = 0$、$y > 0$。条件变成 $(y + a\tan\theta)^2 \le a^2\sec^2\theta$, 而由于 $y + a\tan\theta > 0$, 这就是

$$y \le a(\sec\theta - \tan\theta).$$

斜面上水平坐标为 y 的点，沿斜面量度与原点的距离是 $y\sec\theta$。所以沿斜面向上最远的点，距离为

$$a\sec\theta(\sec\theta - \tan\theta) = a\,\frac{1-\sin\theta}{\cos^2\theta} = a\,\frac{1-\sin\theta}{(1-\sin\theta)(1+\sin\theta)} = \frac{u^2}{g(1+\sin\theta)},$$

即为所求。

沿斜面正下方即 $x = 0$、$y < 0$。此时条件给出 $y \ge -a(\sec\theta + \tan\theta)$，所以沿斜面向下最远的点与大炮的距离为

$$a\sec\theta(\sec\theta + \tan\theta) = a\,\frac{1+\sin\theta}{\cos^2\theta} = \frac{u^2}{g(1-\sin\theta)}$$

第 (iv) 部分

道路是直线 $y = 0$（它位于斜面上，高度为 0）。在条件中代入 $y = 0$，得

$$x^2 \le a^2\sec^2\theta - a^2\tan^2\theta = a^2,$$

所以可到达的道路是 $-a \le x \le a$，长度为 $\dfrac{2u^2}{g}$，与在水平地面上相同。

现在把大炮放在 $(0, c, c\tan\theta)$。它与道路（即 x 轴）的距离是 $\sqrt{c^2 + c^2\tan^2\theta} = |c|\sec\theta$, 所以 $r = |c|\sec\theta$。第 (ii) 部分的一切只取决于相对于大炮的位置，而从斜面上任何一点

看，斜面都是一样的，因此我们可以沿用该条件，只需把 y 换成 $y-c$。道路（$y=0$）上可到达的部分满足

$$x^2 \leq a^2\sec^2\theta - (a\tan\theta - c)^2.$$

想一想 大炮应该放在道路上方还是下方?

可到达的长度是 $2\sqrt{a^2\sec^2\theta-(a\tan\theta-c)^2}$，当被减去的平方项为零，即 $c = a\tan\theta > 0$ 时最大。所以大炮应放在道路上方的斜面上，距离恰好使道路穿过可到达圆盘的圆心，此时道路截出一条直径。最佳距离是

$$r = a\tan\theta\sec\theta = \frac{u^2\sin\theta}{g\cos^2\theta},$$

而道路的最大长度就是圆盘的直径，

$$2a\sec\theta = \frac{2u^2}{g\cos\theta}.$$

验算

- 取 $\theta = 0$（水平地面），第 (ii) 部分变成 $x^2+y^2 \leq (u^2/g)^2$，即熟悉的各方向最大射程 u^2/g；(iii) 的两个答案都化为 u^2/g，而 (iv) 中 $r=0$，长度为 $2u^2/g$。
- 沿斜面向上的答案 $u^2/\big(g(1+\sin\theta)\big)$ 正是斜面（inclined plane）上向上发射的标准最大射程，这是圆盘正确的好迹象。
- 取 $u = 12\,\mathrm{m\,s^{-1}}$、$g = 9.8\,\mathrm{m\,s^{-2}}$ 及 $\theta = 0.5\,\mathrm{rad}$，用电脑扫描所有发射角，得出沿斜面向上的最大射程为 9.932 m，向下为 28.226 m，与 $u^2/\big(g(1\pm\sin\theta)\big)$ 吻合；移动大炮后的最佳道路长度为 $33.49\,\mathrm{m} = 2u^2/(g\cos\theta)$。

常见错误

- 把 (x,y) 当作沿斜面量度的距离。它们是水平坐标，所以沿斜面向上的距离是 $y\sec\theta$，而不是 y。
- 忘记大炮可以指向水平线以下：$\tan\alpha$ 可取所有实数，正因如此，判别式条件就是问题的全部。
- 把移动后的大炮放在道路下方。圆盘的圆心在大炮的斜坡下方，所以大炮必须位于道路上方。

STEP 之外

第 (ii) 部分的圆盘，是所有可能轨迹的包络（*envelope*）——“安全抛物面”（paraboloid of safety）$z = \dfrac{u^2}{2g} - \dfrac{g(x^2+y^2)}{2u^2}$——的一个截面；用任何平面与这个抛物面相交，都会得到一条圆锥曲线（conic），在这里其投影是一个圆。曲线族的包络，在大学一年级的力学和微分方程（differential equations）课程中都是很自然的课题。Kleppner 和 Kolenkow 所著的 *An Introduction to Mechanics*（《力学导论》）是一个好去处，可以看到以这种几何上的细致手法处理抛射体问题。

详细解答

STEP II 2022，第 11 题

概率与统计 • 期望损失的最优化

STEP II 2022，第 11 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

按那 T 枝被测试的记忆棒的情况来划分事件。只要其中任何一枝受感染，就不会有任何东西放上网络，损失为零；只有当全部 T 枝都干净时，网络上的那场赌博才会出现。把期望损失（expected loss）写成 q 的函数，(ii) 就变成单变量的最大化（maximisation）问题；而比较相邻的 T 值（用差分（difference）而不是导数（derivative），因为 T 是整数）则可处理 (iii)。

解答

全文中，损失为 $-x$ 即表示获利 £x。设 L 为以英镑计的净损失。各记忆棒是否受感染互相独立（independent），每枝干净的概率（probability）为 q，所以任何指定的 m 枝记忆棒全部干净的概率是 q^m。由对称性（symmetry）可知，选哪 T 枝来测试并不重要。

第 (i) 部分

先考虑 $T > 0$。共有三种结果。

- 某枝被测试的记忆棒受感染（概率为 $1 - q^T$）：其余的全被销毁，没有任何记忆棒被使用，$L = 0$。
- 全部 T 枝被测试的记忆棒都干净，但其余 $N - T$ 枝中至少有一枝受感染（概率为 $q^T(1 - q^{N-T})$）：我们要支付消毒费用，所以 $L = D$。
- 全部 N 枝记忆棒都干净（概率为 $q^T q^{N-T} = q^N$）：被使用的 $N - T$ 枝每枝赚 £1，所以 $L = -(N - T)$。

因此

$$\begin{aligned}\mathrm{E}(L) &= D\,q^T\left(1-q^{N-T}\right)-(N-T)\,q^N\\ &= D\,q^T-(N-T+D)\,q^N.\end{aligned}$$

当 $T=0$ 时，全部 N 枝记忆棒都被使用：$L=D$ 的概率为 $1-q^N$，$L=-N$ 的概率为 q^N，得 $D-(N+D)q^N$。这正是同一公式取 $T=0$ 的结果，所以对每个 $0\le T<N$，

$$\mathrm{E}(L)=D\,q^T-(N-T+D)\,q^N.$$

第 (ii) 部分

α 的界。由于 $D>0$、$T\ge 0$ 及 $N-T>0$，分子和分母分别满足 $DT\ge 0$ 及 $N(N-T+D)>0$，所以 $\alpha\ge 0$。此外

$$N(N-T+D)-DT=N(N-T)+D(N-T)=(N-T)(N+D)>0,$$

所以 $DT<N(N-T+D)$，即 $\alpha<1$。

最大值。固定 N、D、T，对 $0\le q\le 1$ 记 $f(q)=Dq^T-(N-T+D)q^N$。

若 $T=0$，则 $\alpha=0$ 且 $f(q)=D-(N+D)q^N$，它随 q 递减，所以最大值出现在 $q=0$，即 $q^N=0=\alpha$ 之处。

若 $T\ge 1$，求导：

$$f'(q)=DTq^{T-1}-N(N-T+D)q^{N-1}=N(N-T+D)\,q^{T-1}\left(\alpha-q^{N-T}\right).$$

> **想一想** 我们怎样知道这个驻点（stationary point）是极大点，而且是整个区间上的最大值?

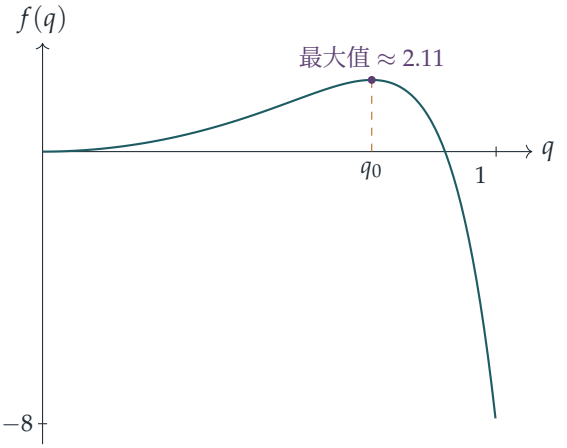
当 $0<q\le 1$ 时，因子 $N(N-T+D)q^{T-1}$ 是正数。因为 $N-T\ge 1$，q^{N-T} 由 0 严格递增（strictly increasing）至 1，而 $0<\alpha<1$，所以恰好有一个 $q_0\in(0,1)$ 使 $q_0^{N-T}=\alpha$。当 $q<q_0$ 时 $f'(q)>0$，当 $q>q_0$ 时 $f'(q)<0$。所以 f 在 q_0 之前上升，之后下降：它在 $[0,1]$ 上的最大值出现在 $q^{N-T}=\alpha$ 之处，得证。

最大值是多少。在 q_0 处有 $q_0^{N-T}=\alpha$，所以 $q_0=\alpha^{1/(N-T)}$ 及 $q_0^T=\alpha^{T/(N-T)}=\alpha^k$。于是

$$\begin{aligned}f(q_0)&=q_0^T\left(D-(N-T+D)\,q_0^{N-T}\right)=\alpha^k\left(D-(N-T+D)\cdot\frac{DT}{N(N-T+D)}\right)\\ &=\alpha^k\left(D-\frac{DT}{N}\right)=\frac{D(N-T)\,\alpha^k}{N}.\end{aligned}$$

所以最大的期望净损失是 $\pounds\, D(N-T)\alpha^k/N$。（当 $T=0$ 时，这就是 $\pounds D$；若取 $\alpha^0=1$，便与 $f(0)=D$ 吻合。）

下图显示 $N=10$、$T=2$、$D=5$ 时的 f，此时 $\alpha=\frac{1}{13}$。



第 (iii) 部分

现在固定 N、D 和 q，并记 $E_T = Dq^T - (N - T + D)q^N$。由于 T 只取整数值，“随 T 递增”的意思是对 $0 \le T \le N-2$ 有 $E_{T+1} > E_T$。相减得

$$\begin{aligned} E_{T+1} - E_T &= Dq^{T+1} - Dq^T + q^N \\ &= q^T\left(q^{N-T} - Dp\right), \end{aligned}$$

这里用了 $q - 1 = -p$。

> ? **想一想** 括号内的式子与 T 有关。怎样才能找到一个 β 同时适用于所有 T?

因为 $0 < q \le 1$ 及 $N - T \le N$，我们有 $q^{N-T} \ge q^N$，所以只须令 $q^N > Dp$ 便足够，而这个条件与 T 无关。由伯努利（Bernoulli）不等式（用简单的归纳法（induction）即可证明：若 $(1-p)^n \ge 1 - np$，则 $(1-p)^{n+1} \ge (1-np)(1-p) = 1 - (n+1)p + np^2$），

$$q^N = (1-p)^N \ge 1 - Np.$$

所以若 $1 - Np > Dp$，即 $p < \dfrac{1}{N+D}$，则 $q^{N-T} \ge q^N \ge 1 - Np > Dp$，而且 $q^T > 0$。因此，取

$$\beta = \frac{1}{N+D} > 0,$$

每个 $p < \beta$ 都使所有 T 均有 $E_{T+1} > E_T$，所以期望损失是 T 的递增函数。

对这样的 p，$E_0, E_1, \ldots, E_{N-1}$ 中最小的是 E_0。所以当 p 足够小时，最佳策略是 $T = 0$：甚么都不测试，使用全部 N 枝记忆棒。这是合理的：如果感染很罕见，测试大多只会销毁本来可以赚钱的好记忆棒。

验算

- 取 $N=1$、$T=0$，公式给出 $D-(1+D)q$：以概率 p 损失 D，以概率 q 获利 1，与预期相符。
- $q=1$（完全没有感染）给出 $D-(N-T+D)=-(N-T)$，即必定获利 $N-T$。
- 取 $N=10$、$T=2$、$D=5$：$\alpha=\frac{1}{13}$，$k=\frac{1}{4}$，最大值为 $4\cdot 13^{-1/4}\approx 2.107$，与上图的峰值吻合。我们亦用电脑列出 $N\le 7$ 时所有感染情况，验证了第 (i) 部分的公式。
- 取 $N=10$、$D=5$、$p=0.3$（远大于 $\beta=\frac{1}{15}$），期望损失其实随 T 递减，由 $T=0$ 时约 4.58 降至 $T=9$ 时的 0.03：可见对 p 确实需要加上某种条件。

常见错误

- 遗漏“某枝被测试的记忆棒受感染”这个结果：它对期望值的贡献是 0，但仍然占去 $1-q^T$ 的概率。
- 把获利算成 $£N$ 而不是 $£(N-T)$：即使被测试的记忆棒是干净的，它们也会被销毁。
- 在 (iii) 中不加说明就对整数 T 求导，或选取了与 T 有关的 β。

STEP 之外

这是统计决策理论（statistical decision theory）中的一个小问题：当某个参数（p）未知时，选择一个行动（这里是 T）以使期望损失最小。第 (ii) 部分体现了极小化极大（minimax）的思想：它为每个策略找出在所有可能的 p 之下的最坏情况，这正是谨慎的管理者可能会拿来比较的东西。测试成本与漏检缺陷成本之间的同一种取舍，也是验收抽样（acceptance sampling）和群组检测（group testing）背后的原理。Blitzstein 与 Hwang 的 *Introduction to Probability*（《概率导论》）是建立本题所用期望值工具的好起点。

详细解答

STEP II 2022，第 12 题

概率与统计 • 平均值、中位数与众数

STEP II 2022，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

一切都源自累积分布函数（cumulative distribution function）F，而这里它只是一个简单的多项式。由于 F 在 $[0,1]$ 上递增，数 a 位于中位数（median）之下，当且仅当 $\mathrm{F}(a) < \frac{1}{2}$。因此，要比较平均值（mean）或众数（mode）与中位数的大小，我们根本不需要求出中位数本身：只需计算 F 在平均值或众数处的值，再与 $\frac{1}{2}$ 比较即可。

解答

第 (i) 部分

由于总概率为 1，

$$1 = k\int_0^1 \left(x^n - x^{n+1}\right)\mathrm{d}x = k\left(\frac{1}{n+1} - \frac{1}{n+2}\right) = \frac{k}{(n+1)(n+2)},$$

所以 $k = (n+1)(n+2)$，正如所求。于是

$$\mu = k\int_0^1 \left(x^{n+1} - x^{n+2}\right)\mathrm{d}x = (n+1)(n+2)\cdot\frac{1}{(n+2)(n+3)} = \frac{n+1}{n+3}.$$

第 (ii) 部分

当 $0 \le x \le 1$ 时，

$$\mathrm{F}(x) = (n+1)(n+2)\left(\frac{x^{n+1}}{n+1} - \frac{x^{n+2}}{n+2}\right) = x^{n+1}\big((n+2) - (n+1)x\big).$$

由于在 $(0,1)$ 上 $\mathrm{f}>0$，F 在该区间上严格递增（strictly increasing），而中位数 m 是唯一满足 $\mathrm{F}(m)=\frac{1}{2}$ 的点。因此

$$\mu < m \iff \mathrm{F}(\mu) < \tfrac{1}{2}.$$

现在代入 $\mu=\frac{n+1}{n+3}$，

$$(n+2)-(n+1)\mu = \frac{(n+2)(n+3)-(n+1)^2}{n+3} = \frac{3n+5}{n+3},$$

所以

$$\mathrm{F}(\mu) = \left(\frac{n+1}{n+3}\right)^{n+1}\frac{3n+5}{n+3}.$$

因此，条件 $\mathrm{F}(\mu)<\frac{1}{2}$ 两边乘以正数 $2\left(\frac{n+3}{n+1}\right)^{n+1}$ 后，变为

$$\frac{2(3n+5)}{n+3} < \left(\frac{n+3}{n+1}\right)^{n+1}.$$

由于 $\frac{2(3n+5)}{n+3}=\frac{6(n+3)-8}{n+3}=6-\frac{8}{n+3}$ 而且 $\frac{n+3}{n+1}=1+\frac{2}{n+1}$，这正是题目给出的不等式。所以只要该不等式成立，$\mu$ 就小于中位数。

要证明这个不等式，设 $N=n+1$，于是 $N\geq 3$ 是整数。$\left(1+\frac{2}{N}\right)^N$ 的二项展开（binomial expansion）中每一项都是正数，而且共有 $N+1\geq 4$ 项，所以总和至少等于首四项之和：

$$\begin{aligned}\left(1+\frac{2}{N}\right)^N &\geq 1+N\cdot\frac{2}{N}+\frac{N(N-1)}{2}\cdot\frac{4}{N^2}+\frac{N(N-1)(N-2)}{6}\cdot\frac{8}{N^3}\\ &= 3+2\left(1-\frac{1}{N}\right)+\frac{4}{3}\left(1-\frac{3}{N}+\frac{2}{N^2}\right)\\ &= \frac{19}{3}-\frac{6}{N}+\frac{8}{3N^2}.\end{aligned}$$

想一想 这个四项之和只是一个下界。我们怎样证明它大于 $6-\frac{8}{N+2}$？

我们把两者相减，再乘以正的分母 $3N^2(N+2)$ 以消去分母：

$$\begin{aligned}&3N^2(N+2)\left[\frac{19}{3}-\frac{6}{N}+\frac{8}{3N^2}-6+\frac{8}{N+2}\right]\\ &\quad= N^2(N+2)-18N(N+2)+8(N+2)+24N^2\\ &\quad= N^3+8N^2-28N+16.\end{aligned}$$

当 $N\geq 3$ 时，我们有 $N^2+8N-28\geq 9+24-28=5>0$，所以 $N^3+8N^2-28N+16=N(N^2+8N-28)+16>0$。因此

$$6-\frac{8}{n+3} < \frac{19}{3}-\frac{6}{N}+\frac{8}{3N^2} \leq \left(1+\frac{2}{n+1}\right)^{n+1},$$

即题目给出的不等式对每个整数 $n>1$ 都成立，所以 X 的中位数大于 μ。

第 (iii) 部分

在 $(0,1)$ 上，

$$\mathrm{f}'(x) = k\left(nx^{n-1} - (n+1)x^n\right) = kx^{n-1}\big(n - (n+1)x\big),$$

当 $x < \frac{n}{n+1}$ 时它是正数，当 $x > \frac{n}{n+1}$ 时它是负数。所以 f 先升后降，众数为 $M = \frac{n}{n+1}$。

与第 (ii) 部分一样，$M > m$ 当且仅当 $\mathrm{F}(M) > \frac{1}{2}$。由于 $(n+2) - (n+1)M = (n+2) - n = 2$，

$$\mathrm{F}(M) = 2\left(\frac{n}{n+1}\right)^{n+1},$$

所以我们需要 $\left(\frac{n}{n+1}\right)^{n+1} > \frac{1}{4}$，即

$$\left(1 + \frac{1}{n}\right)^{n+1} < 4.$$

> **想一想** 既然 n 可以是 2 或以上的任何整数，题目给出关于 $\left(1+\frac{1}{x}\right)^{x+1}$ 的事实怎样帮到我们?

由于该函数在 x 为正数时递减，它在任何整数 $n \geq 2$ 处的值都不大于它在 $x = 2$ 处的值：

$$\left(1 + \frac{1}{n}\right)^{n+1} \leq \left(\frac{3}{2}\right)^3 = \frac{27}{8} < 4.$$

因此 $\mathrm{F}(M) > \frac{1}{2}$，即 X 的众数大于其中位数。

综合各部分，对每个 $n > 1$ 都有 平均值 < 中位数 < 众数。

验算

- 取 $n = 2$：平均值为 $\frac{3}{5} = 0.6$，众数为 $\frac{2}{3} \approx 0.667$。以数值方法解 $4x^3 - 3x^4 = \frac{1}{2}$，得中位数 ≈ 0.6143，恰好介乎两者之间。此外，$\mathrm{F}(0.6) = 0.4752 < \frac{1}{2}$，而 $\mathrm{F}(\frac{2}{3}) = \frac{16}{27} \approx 0.593 > \frac{1}{2}$。
- 当 $n = 2$（$N = 3$）时，展开式刚好有四项，所以下界取等号：两边都等于 $\frac{125}{27} \approx 4.63$，稳稳大于 $6 - \frac{8}{5} = 4.4$。
- 平均值 < 中位数 < 众数这个大小次序，已用电脑对 2 至 79 的每个 n 验证过。

常见错误

- 试图明确求出中位数：方程 $x^{n+1}(n+2-(n+1)x) = \frac{1}{2}$ 没有简洁的解，而你也不需要求出它。

- 弄错不等号方向：由于 F 递增，$\mu < m$ 对应的是 $\mathrm{F}(\mu) < \frac{1}{2}$。
- 忘记说明为何可以舍去后面的二项式项（因为它们全是正数）。

STEP 之外

这个密度函数是一个贝塔分布（Beta distribution）$\mathrm{Beta}(n+1, 2)$，它是 $[0,1]$ 上最重要的分布族之一，在贝叶斯统计（Bayesian statistics）中处于核心地位。平均值 < 中位数 < 众数这个次序，是左边有长尾（负偏态，negative skew）的分布的典型特征，但它并非普遍定律，所以这里必须加以证明。第 (iii) 部分中的递减函数在 $x \to \infty$ 时趋向 e，这是一个经典极限，你将会在分析（analysis）课程中正式学到。Blitzstein 与 Hwang 的 *Introduction to Probability*（《概率导论》）对贝塔分布有出色的论述。

详细解答

STEP III 2022，第 1 题

纯数学 · 多项式的根

STEP III 2022，第 1 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

把参数化（parametrisation）代入圆的方程，就能把“两条曲线在哪里相交?”转化为“一个四次方程（quartic）的根是甚么?”；此后一切都只靠根与系数的关系（relationships between roots and coefficients）完成，完全不必解这个四次方程。另一个有用的技巧是：倒数（reciprocals）$1/t_i$ 正是把同一个四次方程的系数倒序排列后所得方程的根。至于最后一部分，“相切”（touching）的意思就是有重根（repeated root）。

解答

第 (i) 部分

参数为 t 的点位于 C_2 上，当且仅当

$$(ct-a)^2+\left(\frac{c}{t}-b\right)^2=r^2.$$

展开得 $c^2t^2-2act+a^2+\dfrac{c^2}{t^2}-\dfrac{2bc}{t}+b^2=r^2$。由于 $t\neq 0$，我们可以乘以 t^2 而不会增加或失去任何解：

$$c^2t^4-2act^3+(a^2+b^2-r^2)t^2-2bct+c^2=0.$$

因此每个 t_i 都是 $(*)$ 的根。不同的 t 值给出不同的点（它们的 x 坐标 ct 各不相同），所以四个点 P_i 对应四个互不相同的参数。一个四次方程最多只有四个根，因此 $t_1,\ldots,t_4$ 恰好就是 $(*)$ 的全部根。

第 (ii) 部分

把 $(*)$ 除以 c^2，各根之和以及两两之积的和分别为

$$\sum t_i = \frac{2a}{c}, \qquad \sum_{i<j} t_i t_j = \frac{a^2+b^2-r^2}{c^2}.$$

利用 $\left(\sum t_i\right)^2 = \sum t_i^2 + 2\sum_{i<j} t_i t_j$，

$$\sum_{i=1}^{4} t_i^2 = \frac{4a^2}{c^2} - \frac{2(a^2+b^2-r^2)}{c^2} = \frac{2}{c^2}\left(a^2-b^2+r^2\right),$$

这正是所求的结果。

> **想一想** 怎样才能求得 $\sum 1/t_i^2$，而不必处理繁复的公分母?

在 $(*)$ 中令 $u = 1/t$。由于常数项 $c^2 \neq 0$，没有任何根为零，所以可以把 $(*)$ 除以 t^4：

$$c^2 - 2acu + (a^2+b^2-r^2)u^2 - 2bcu^3 + c^2u^4 = 0,$$

即 $c^2u^4 - 2bcu^3 + (a^2+b^2-r^2)u^2 - 2acu + c^2 = 0$。它的根就是各个 $1/t_i$。这正是把 $(*)$ 中的 a 与 b 互换所得的方程，所以同样的计算给出

$$\sum_{i=1}^{4} \frac{1}{t_i^2} = \frac{2}{c^2}\left(b^2-a^2+r^2\right).$$

第 (iii) 部分

点 P_i 为 $(ct_i,\, c/t_i)$，所以 $OP_i^2 = c^2t_i^2 + \dfrac{c^2}{t_i^2}$。求和并利用第 (ii) 部分，

$$\begin{aligned}\sum_{i=1}^{4} OP_i^2 &= c^2\sum t_i^2 + c^2\sum \frac{1}{t_i^2} \\ &= 2(a^2-b^2+r^2) + 2(b^2-a^2+r^2) \\ &= 4r^2,\end{aligned}$$

这就是所求的结果。留意圆心 (a,b) 已完全消失。

第 (iv) 部分

> **想一想** 以 $(*)$ 的根而言，“相切”是甚么意思?

设 $f(t) = (ct-a)^2 + (c/t-b)^2 - r^2$，即动点到圆心距离的平方减去 r^2。曲线在参数 t_0 处与圆相切，当 $f(t_0) = 0$ 且 $f'(t_0) = 0$：该点在圆上，而且在一阶近似下，它与圆心的

距离没有改变，因此曲线在该处的方向垂直于半径，也就是圆的切线方向。（这里用到速度 $(c, -c/t^2)$ 永不为零这一事实。）$(*)$ 中的四次多项式就是 $t^2 f(t)$，而由于 $t_0 \neq 0$，f 在 t_0 有二重根当且仅当 $(*)$ 在 t_0 有二重根。

所以，若 C_1 与 C_2 在两个不同的点相切，其参数为 $\alpha \neq \beta$，则 α 和 β 都是 $(*)$ 的二重根。这已涵盖全部四个根，因此

$$c^2t^4 - 2act^3 + (a^2 + b^2 - r^2)t^2 - 2bct + c^2 = c^2(t-\alpha)^2(t-\beta)^2.$$

各根之积。比较常数项，$c^2\alpha^2\beta^2 = c^2$，所以 $(\alpha\beta)^2 = 1$，即 $\alpha\beta = \pm 1$。

各根之和。比较 t^3 项，$2(\alpha+\beta) = \dfrac{2a}{c}$，所以 $\alpha + \beta = \dfrac{a}{c}$。

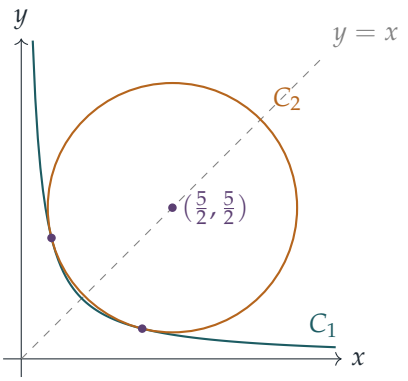
每三根之积的和。比较 t 项，$2\alpha\beta(\alpha+\beta) = \dfrac{2b}{c}$，所以 $\alpha\beta(\alpha+\beta) = \dfrac{b}{c}$。

把后两式相除（或把第二式代入第三式），得 $b = \alpha\beta\, a$。因此

$$\alpha\beta = 1 \implies b = a, \qquad \alpha\beta = -1 \implies b = -a,$$

所以圆心 (a, b) 位于 $y = x$ 或 $y = -x$ 上。（若 $a = 0$，则 $b = 0$，圆心为原点，同时位于两条直线上。）

下图展示 $c = 1$、$\alpha = 2$、$\beta = \frac{1}{2}$ 的情况。比较系数得 $a = b = \frac{5}{2}$ 及 $r^2 = \frac{17}{4}$，而圆与双曲线（hyperbola）相切于 $(2, \frac{1}{2})$ 和 $(\frac{1}{2}, 2)$。



验算

- 在第 (iii) 部分，取以原点为圆心的圆 $x^2 + y^2 = r^2$，并设 $c = 1$ 及 $r^2 > 2$。由对称性，四个交点为 $(p, q), (q, p), (-p, -q), (-q, -p)$，每点与原点的距离都是 r，所以总和确实是 $4r^2$。
- 在图中的例子里，$(\frac{5}{2}, \frac{5}{2})$ 到 $(2, \frac{1}{2})$ 距离的平方为 $\frac{1}{4} + 4 = \frac{17}{4} = 4.25$，正如所料。该处的半径方向 $(-\frac{1}{2}, -2)$ 与曲线方向 $(1, -\frac{1}{4})$ 互相垂直，证实了相切。

- 第 (ii) 部分的公式已用电脑对数千个随机的圆进行验证，方法是以数值方法解该四次方程。

常见错误

- 弄错 $\sum t_i t_j$ 的正负号。对于 $t^4 + pt^3 + qt^2 + \ldots$，各根之和是 $-p$，但两两之积的和是 $+q$。
- 在第 (iv) 部分只用了各根之积。这只给出 $\alpha\beta = \pm 1$，单凭这一点并不能说明 (a, b) 的任何性质。还需要再比较一个系数，才能把 a 和 b 引进来。
- 没有说明为何“相切”会给出二重根。理由是条件 $f(t_0) = f'(t_0) = 0$，连同 $t_0 \neq 0$。

STEP 之外

根与系数之间的关系就是初等对称函数（elementary symmetric functions）。一条经典定理指出，各根的任何对称多项式（symmetric polynomial），例如 $\sum t_i^2$ 或 $\sum t_i^{-2}$，都可以用系数表示。倒数代换 $u = 1/t$ 让我们初步窥见射影（projective）几何的思想；而一条圆锥曲线（conic）与一个圆最多相交于四点，这是关于平面曲线相交的贝祖定理（Bézout's theorem）的特例。如想进一步了解多项式及其根，可参阅 Allenby, *Rings, Fields and Groups*（《环、域与群》）；如想多做这类风格的练习，可参阅 Siklos, *Advanced Problems in Mathematics*（《高等数学问题》）。

详细解答

STEP III 2022，第 2 题

纯数学 · 无穷递降法

STEP III 2022，第 2 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

本题在 *Beyond STEP* 第 1.5 节无穷递降法中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch01/infinite-descent.html

关键思路

这是费马（Fermat）的无穷递降法（*method of infinite descent*）。假设存在一组非零整数解。通过整除性（divisibility）论证，可证明所有变量都有一个公因数。把它除去，便得到同一类（或密切相关的一类）的新解，而且各数都是绝对值严格更小的非零整数。这个过程不可能无止境地进行下去，因为正整数不可能无止境地递减。每一部分的工夫在于找出应考虑的质数（prime）（2 或 3），并验证新解确实与旧解具有相同的形式。

解答

全程我们会用到两个事实。若质数 m 整除 x^k，则 m 整除 x。另外，若 x 是 $m \geq 2$ 的非零倍数，则 x/m 是非零整数，且 $|x/m| < |x|$。

第 (i) 部分

假设 $a^3 + 2b^3 + 4c^3 = 0$，其中 a, b, c 为非零整数。那么

$$a^3 = -2\left(b^3 + 2c^3\right)$$

是偶数。奇数的立方是奇数，所以 a 是偶数。设 $a = 2p$，则 p 是非零整数，且 $|p| = \frac{1}{2}|a| < |a|$。代入得

$$8p^3 + 2b^3 + 4c^3 = 0 \quad \Longrightarrow \quad 4p^3 + b^3 + 2c^3 = 0.$$

这就是第一个结论。

现在对新方程重复同一论证，得 $b^3 = -2\left(2p^3 + c^3\right)$，所以 b 是偶数，设 $b = 2q$，于是

$$4p^3 + 8q^3 + 2c^3 = 0 \quad \Longrightarrow \quad 2p^3 + 4q^3 + c^3 = 0.$$

再做一次，$c^3 = -2\left(p^3 + 2q^3\right)$，所以 $c = 2r$，且

$$2p^3 + 4q^3 + 8r^3 = 0 \quad \Longrightarrow \quad p^3 + 2q^3 + 4r^3 = 0.$$

这里 $p = a/2$、$q = b/2$ 和 $r = c/2$ 都是非零整数，且 $|p| < |a|$、$|q| < |b|$ 及 $|r| < |c|$，正如所求。

> **想一想** 为甚么这表示根本没有解，而不只是“另一组解”？

假设存在非零整数解。在所有这些解之中，选取 $|a|$ 尽可能小的一组。这是做得到的，因为任何非空的正整数集合都有最小元素（least element）。上述论证于是给出另一组非零整数解 (p, q, r)，且 $|p| < |a|$，这与 $|a|$ 的最小性矛盾。所以这样的整数 a、b、c 并不存在。（等价地说，这个过程会给出一个由正整数组成、无穷且严格递减的数列 $|a| > |a/2| > |a/4| > \cdots$，而这是不可能的。）

第 (ii) 部分

这次要用的质数是 3。假设 $9a^3 + 10b^3 + 6c^3 = 0$，其中 a, b, c 为非零整数。

- 由于 $10b^3 = -3\left(3a^3 + 2c^3\right)$，3 整除 $10b^3$。因为 3 不整除 10，所以它整除 b^3，从而整除 b。设 $b = 3q$，则 $9a^3 + 270q^3 + 6c^3 = 0$，即
$$3a^3 + 90q^3 + 2c^3 = 0.$$
- 现在 $2c^3 = -3\left(a^3 + 30q^3\right)$，所以 $3 \mid c$。设 $c = 3r$：$3a^3 + 90q^3 + 54r^3 = 0$，即
$$a^3 + 30q^3 + 18r^3 = 0.$$
- 最后 $a^3 = -3\left(10q^3 + 6r^3\right)$，所以 $3 \mid a$。设 $a = 3p$：$27p^3 + 30q^3 + 18r^3 = 0$，即
$$9p^3 + 10q^3 + 6r^3 = 0.$$

所以 $(p, q, r) = (a/3, b/3, c/3)$ 是同一方程的非零整数解，且 $|p| < |a|$。与第 (i) 部分一样，选取 $|a|$ 最小的一组解，便得出矛盾，所以不存在非零整数解。

第 (iii) 部分

任何不是 3 的倍数的整数，都可以写成 $3n \pm 1$，其中 n 为某个整数。那么

$$(3n \pm 1)^2 = 9n^2 \pm 6n + 1 = 3\left(3n^2 \pm 2n\right) + 1,$$

即比 3 的某个倍数大 1。另一方面，3 的倍数的平方是 3 的倍数（事实上更是 9 的倍数）。

因此每个平方数除以 3 的余数是 0 或 1，而余数为 0 当且仅当该数本身是 3 的倍数。对于 x^2+y^2，可能的余数是

$$0+0=0, \qquad 0+1=1, \qquad 1+1=2.$$

所以 x^2+y^2 是 3 的倍数，仅当两个余数都是 0，亦即仅当 x 和 y 都是 3 的倍数。

现在假设 $a^2+b^2=3c^2$，其中 a,b,c 为非零整数。右方是 3 的倍数，所以由刚证明的结果，$a=3p$ 及 $b=3q$。于是

$$9p^2+9q^2=3c^2 \quad\Longrightarrow\quad c^2=3\left(p^2+q^2\right).$$

所以 c^2 是 3 的倍数。由第一个结果，c 必定是 3 的倍数，否则 c^2 会比 3 的某个倍数大 1。设 $c=3r$：

$$9r^2=3\left(p^2+q^2\right) \quad\Longrightarrow\quad p^2+q^2=3r^2.$$

这是同一方程的非零整数解，且 $|p|<|a|$。选取 $|a|$ 最小的一组解，便得出与之前相同的矛盾，所以不存在非零解。

第 (iv) 部分

这里我们以 4 为模（modulo）进行计算。偶数 $2m$ 的平方是 $4m^2\equiv 0$，而奇数 $2m+1$ 的平方是 $4m^2+4m+1\equiv 1 \pmod 4$。所以三个平方数之和的余数，等于三项中奇数的个数（0、1、2 或 3）。它是 4 的倍数，仅当三个数都是偶数。

? **想一想** 若把 $a^2+b^2+c^2=4abc$ 中的所有数减半，会得回同一条方程吗?

不会。若 $a=2p$、$b=2q$ 及 $c=2r$，方程变为 $4(p^2+q^2+r^2)=32pqr$，即 $p^2+q^2+r^2=8pqr$。系数加倍了。所以我们改为证明一个更强、而且在减半之下确实保持不变的命题：

命题。对任何整数 $k\geq 2$，都不存在非零整数使得 $a^2+b^2+c^2=2^kabc$。

假设存在。在所有这样的解之中（$k\geq 2$ 可取任何值），选取 $|a|$ 尽可能小的一组。由于 $k\geq 2$，右方 2^kabc 是 4 的倍数。所以 $a^2+b^2+c^2$ 是 4 的倍数，由上面的说明，a、b 和 c 都是偶数：$a=2p$、$b=2q$、$c=2r$。于是

$$4\left(p^2+q^2+r^2\right)=2^k\cdot 8pqr \quad\Longrightarrow\quad p^2+q^2+r^2=2^{k+1}\,pqr.$$

这是一组非零整数解，其指数为 $k+1\geq 2$，且 $|p|<|a|$，与 a 的选取矛盾。所以命题成立。取 $k=2$ 的情况，即知不存在非零整数 a,b,c 使得 $a^2+b^2+c^2=4abc$。

验算

- 以电脑搜索所有满足 $|a|,|b|,|c| \leq 40$ 的非零 a,b,c，四条方程都找不到任何解，对 $2 \leq k \leq 6$ 的 $a^2+b^2+c^2=2^k abc$ 亦然。
- 令第 (iv) 部分得以成立的，正是系数中 2 的幂。某些这种形式的方程确实有解，例如 $1^2+1^2+1^2=3\cdot 1\cdot 1\cdot 1$ 及 $3^2+3^2+3^2=1\cdot 3\cdot 3\cdot 3$。在这些例子中系数是奇数，所以模 4 的论证无从入手。
- 试把递降法套用在一个“差一点”的例子上：$a^2+b^2=3c^2$ 在 $(a,b,c)=(1,1,1)$ 时不成立，因为 $1+1=2$ 不是 3 的倍数，这正是第 (iii) 部分中“$1+1=2$”的情况。

常见错误

- 忘记说明递降为何导致矛盾。“我们得到更小的解”并不构成证明，除非你引用最小解（或正整数不可能无止境递减这一事实）。
- 在第 (iv) 部分把数减半后，声称方程不变。事实并非如此：系数加倍了，这正是命题必须加强的原因。
- 写“$3 \mid c^2$ 所以 $3 \mid c$”而不给理由。这里的理由是第 (iii) 部分的第一个结果（更一般地说，因为 3 是质数）。

STEP 之外

费马正是用无穷递降法证明 $x^4+y^4=z^2$ 没有正整数解，从而解决了他的最后定理（费马大定理，Last Theorem）中 $n=4$ 的情况。这里的每一次递降，实际上都等于证明：以某个适当的质数幂为模时，方程没有变量不全被该质数整除的解。这是迈向 p 进数（p-adic numbers）以及现代数论中局部—整体（local–global）思想的第一步。著名的方程 $a^2+b^2+c^2=3abc$（马尔可夫方程（Markov equation））有无穷多组解，它们由一个优美的树状结构生成。Davenport 的 *The Higher Arithmetic*（《高等算术》）正好以这个程度讲解递降法、同余（congruence）和平方和。

详细解答

STEP III 2022，第 3 题

纯数学 · 隐函数微分法

STEP III 2022，第 3 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

用隐函数求导（implicit differentiation）。驻点（stationary point）要求 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 的分子为零，由此可把 y 以 x 表示（或得出两者之间的一个线性关系）。把它代回曲线方程，便只剩下一条关于 x 的方程，问题于是变成：这条方程何时有两个实根（real roots）？在 (i) 中，它其实是一条变了装的二次方程（quadratic）；在 (ii) 中，它是一条四次方程（quartic），而数算其根的方法是找出它的转向点（turning point）。

解答

第 (i) 部分

将 $ax^2 + bxy + cy^2 = 1$ 对 x 作隐函数求导，得

$$2ax + by + (bx + 2cy)\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

在驻点处 $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$，所以 $2ax + by = 0$，即 $y = -\frac{2a}{b}x$（因为 $abc \neq 0$，所以可以除以 b）。代入曲线方程，

$$ax^2 - 2ax^2 + \frac{4a^2c}{b^2}x^2 = 1$$
$$\frac{a(4ac - b^2)}{b^2}x^2 = 1.$$

若 $4ac - b^2 \leq 0$，则对所有 x，左方都 ≤ 0（记得 $a > 0$），所以根本没有驻点。因此，若曲线有两个驻点，必有 $4ac - b^2 > 0$，即 $b^2 < 4ac$，得证。（反过来，当 $b^2 < 4ac$ 时，得 $x = \pm|b|/\sqrt{a(4ac - b^2)}$，恰好给出两个点。）

想一想 在这些点上，$bx+2cy$ 会不会也是零，以致导数其实没有定义？

若 $2ax+by=0$ 与 $bx+2cy=0$ 同时成立，由于行列式（determinant）$4ac-b^2$ 不为零，唯一的解是 $x=y=0$，而这点并不在曲线上。所以这两点确实满足 $\frac{\mathrm{d}y}{\mathrm{d}x}=0$。

第 (ii) 部分

将 $ay^3+bx^2y+cx=1$ 作隐函数求导，得

$$(3ay^2+bx^2)\frac{\mathrm{d}y}{\mathrm{d}x}+2bxy+c=0.$$

因此在驻点处 $2bxy+c=0$。由于 $c\neq 0$，故 $x\neq 0$，于是 $y=-\dfrac{c}{2bx}$。代入曲线方程，

$$-\frac{ac^3}{8b^3x^3}-\frac{cx}{2}+cx=1,\qquad \text{即}\qquad -\frac{ac^3}{8b^3x^3}+\frac{cx}{2}=1.$$

两边乘以 $8b^3x^3$ 并移项，

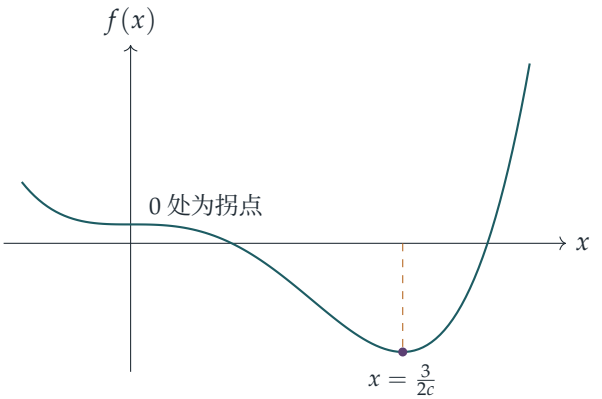
$$4cb^3x^4-8b^3x^3-ac^3=0,$$

这正是所求的方程。

把左方记作 $f(x)$。由于 $y=-c/(2bx)$ 由 x 决定，每个驻点的 x 坐标各不相同。所以有两个驻点，就表示 f 至少有两个相异实根。要数算 f 的根，可以看它的导数：

$$f'(x)=16cb^3x^3-24b^3x^2=8b^3x^2(2cx-3).$$

因子 x^2 永不变号，所以 f' 只在 $x=\frac{3}{2c}$ 处变号。设 $c>0$。那么首项系数（leading coefficient）$4cb^3$ 为正（因为 $b>0$），f 在 $x<\frac{3}{2c}$ 时严格递减，在 $x>\frac{3}{2c}$ 时严格递增；$x=0$ 只是一个水平拐点（stationary point of inflection）。因此 f 在 $\frac{3}{2c}$ 两侧各最多有一个根，而它有两个相异根当且仅当其最小值为负。若 $c<0$，情况全部倒转：f 先增后减，有两个相异根当且仅当其最大值为正。两种情况下，条件都是 $c\,f\left(\frac{3}{2c}\right)<0$。



$a=-1$，$b=c=1$ 的情况：$f(x)=4x^4-8x^3+1$ 只有一个转向点，即 $x=\frac{3}{2}$ 处的极小点，并有两个实根。

现在在转向点求值：

$$\begin{aligned} f\left(\tfrac{3}{2c}\right) &= 4cb^3\cdot\frac{81}{16c^4}-8b^3\cdot\frac{27}{8c^3}-ac^3 \\ &= \frac{81b^3}{4c^3}-\frac{108b^3}{4c^3}-ac^3=-\frac{27b^3+4ac^6}{4c^3}. \end{aligned}$$

所以

$$c\,f\left(\tfrac{3}{2c}\right)=-\frac{4ac^6+27b^3}{4c^2},$$

而这式为负当且仅当 $4ac^6+27b^3>0$。因此，若曲线有两个驻点，则 $4ac^6+27b^3>0$。

第 (iii) 部分

第二及第三条方程表示 $2bxy+c$ 以及 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 的系数 $3ay^2+bx^2$ 同时为零，所以这是曲线上一个令 (ii) 的方法失效的点。由于 $c\neq 0$，第二条方程给出 $x\neq 0$ 及 $y=-\dfrac{c}{2bx}$。前两条方程正是我们在 (ii) 中所用的，所以 x 满足

$$4cb^3x^4-8b^3x^3-ac^3=0. \qquad (*)$$

把 $y=-c/(2bx)$ 代入第三条方程，

$$\frac{3ac^2}{4b^2x^2}+bx^2=0 \quad\Longrightarrow\quad 4b^3x^4=-3ac^2.$$

乘以 c，可见 $(*)$ 的第一项为 $-3ac^3$，于是 $(*)$ 变成

$$-3ac^3-8b^3x^3-ac^3=0 \quad\Longrightarrow\quad x^3=-\frac{ac^3}{2b^3}.$$

把 $x^4=-\dfrac{3ac^2}{4b^3}$ 除以此式（因 $x\neq 0$，可以这样做），得

$$x=\frac{-3ac^2/(4b^3)}{-ac^3/(2b^3)}=\frac{3}{2c}.$$

所以这个解恰好位于 (ii) 中 f 的转向点。把 $x=\frac{3}{2c}$ 代回 $x^3=-\frac{ac^3}{2b^3}$：

$$\frac{27}{8c^3}=-\frac{ac^3}{2b^3} \quad\Longrightarrow\quad 54b^3=-8ac^6 \quad\Longrightarrow\quad 4ac^6+27b^3=0,$$

得证。（等价地，由 (ii) 的公式可知 $f\left(\frac{3}{2c}\right)=0$。）这是 (ii) 的边界情况：四次方程的两个根在转向点合并成一个重根（repeated root），而曲线在该处有一个奇点（singular point），而不是两个驻点。

验算

- 在 (i) 中取 $a = c = 1$、$b = 1$：曲线 $x^2 + xy + y^2 = 1$ 是一个椭圆（ellipse），且 $b^2 = 1 < 4$；它的驻点位于 $x = \pm\frac{1}{\sqrt{3}}$、$y = \mp\frac{2}{\sqrt{3}}$，这些点确实在曲线上。
- 在 (ii) 中取 $a = -1$、$b = c = 1$，得 $4ac^6 + 27b^3 = 23 > 0$，而 f 的最小值为 $f(\frac{3}{2}) = -5.75 = -\frac{23}{4}$，与 $-\frac{27b^3+4ac^6}{4c^3}$ 吻合。
- 在 (iii) 中取 $a = -2$、$c = 1.3$，并取 b 使 $27b^3 = -4ac^6$；那么 $x = \frac{3}{2c}$、$y = -\frac{c}{2bx}$ 在舍入误差范围内满足全部三条方程。以电脑对数千组随机的 (a, b, c) 作检验，亦证实驻点的数目与 (i) 和 (ii) 中的条件相符。

常见错误

- 忘记 $x = 0$ 是四次多项式的一个临界点：它是拐点（point of inflection）而非转向点，指出这一点才能断定最多只有两个根。
- 忽略 c 的正负号：只有当 $c > 0$ 时，转向点才是极小点。把条件写成 $c\,f(\frac{3}{2c}) < 0$，便可一次过处理两种情况。
- 在 (i) 中没有指出 $b^2 = 4ac$ 的情况会得出 $0 = 1$，所以这情况同样没有驻点。

STEP 之外

量 $4ac^6 + 27b^3$ 在相差一个因子的意义下，就是该四次方程的判别式（discriminant）：它为零当且仅当四次方程有重根，正如二次方程的 $b^2 - 4ac$（以及三次方程 $t^3 + pt + q$ 的 $4p^3 + 27q^2$）一样。第 (iii) 部分展示了它的几何意义：重根对应曲线上的一个奇点，在该点 $ay^3 + bx^2y + cx$ 的两个偏导数（partial derivatives）同时为零。奇点与判别式是代数几何（algebraic geometry）的起点；若想以浅白而直观的方式认识与多项式及曲线相关的概念，可读 Needham 的 *Visual Complex Analysis*（《直观复分析》）；若想做更多同类型的题目，可读 Siklos 的 *Advanced Problems in Mathematics*（《高等数学难题》）。

详细解答

STEP III 2022，第 4 题

纯数学 · 无穷乘积

STEP III 2022，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

本题在 *Beyond STEP* 第 2.5 节韦达乘积中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch02/viete.html

关键思路

一切都建基于倍角公式（double-angle formula）$\sinh 2\theta = 2\sinh\theta\cosh\theta$，并反复应用：每用一次，$\sinh$ 的自变量便减半，同时留下一个 $\cosh$ 因子。减半 n 次之后，剩下的 $\sinh(x/2^n)$ 非常小，而且表现得跟 $x/2^n$ 一样，所以令 $n\to\infty$ 便能把一个有限的恒等式变成一个无穷乘积（infinite product）。最后两部分就是在精心挑选的 x 值上计算这个乘积，再配合它的三角函数版本。

解答

第 (i) 部分

记 $P_n(x) = \cosh\left(\frac{x}{2}\right)\cosh\left(\frac{x}{4}\right)\cdots\cosh\left(\frac{x}{2^n}\right)$。我们证明对每个正整数 n，都有 $\sinh x = 2^n P_n(x)\sinh\left(\frac{x}{2^n}\right)$。

基础情况（*base case*）$n=1$。在倍角公式中取 $\theta = x/2$，得 $\sinh x = 2\cosh\left(\frac{x}{2}\right)\sinh\left(\frac{x}{2}\right)$，这正是 $n=1$ 时的命题。

归纳步骤（*inductive step*）。假设结果对某个 $n=k$ 成立。在倍角公式中取 $\theta = x/2^{k+1}$，

$$\sinh\left(\frac{x}{2^k}\right) = 2\cosh\left(\frac{x}{2^{k+1}}\right)\sinh\left(\frac{x}{2^{k+1}}\right),$$

所以

$$\begin{aligned}\sinh x &= 2^k P_k(x)\sinh\left(\frac{x}{2^k}\right) = 2^k P_k(x)\cdot 2\cosh\left(\frac{x}{2^{k+1}}\right)\sinh\left(\frac{x}{2^{k+1}}\right)\\ &= 2^{k+1}P_{k+1}(x)\sinh\left(\frac{x}{2^{k+1}}\right),\end{aligned}$$

这就是 $n = k+1$ 时的结果。由归纳法（induction），结果对所有正整数 n 都成立。

推论。若 $x \neq 0$，则 $x/2^n \neq 0$，所以 $\sinh(x/2^n) \neq 0$（sinh 唯一的零点在 0）。因此我们可以把恒等式除以 $x\sinh(x/2^n)/2^n$，得到

$$\frac{\sinh x}{x}\cdot\frac{x/2^n}{\sinh(x/2^n)} = P_n(x),$$

即为所求。

第 (ii) 部分

当 $y \neq 0$ 时，把麦克劳林级数（Maclaurin series）除以 y，得

$$\frac{\sinh y}{y} = 1 + \frac{y^2}{3!} + \frac{y^4}{5!} + \cdots = 1 + y^2\sum_{r=1}^{\infty}\frac{y^{2r-2}}{(2r+1)!}.$$

想一想 只说“第一项之后的每一项都趋于 0”足够吗?

不完全足够：项数有无穷多，所以我们应该为它们的和找一个界。当 $0 < |y| \leq 1$ 时，$y^{2r-2} \leq 1$，所以

$$0 \leq \frac{\sinh y}{y} - 1 \leq y^2\sum_{r=1}^{\infty}\frac{1}{(2r+1)!} = y^2(\sinh 1 - 1),$$

而当 $y \to 0$ 时，右边趋于 0。因此 $\dfrac{\sinh y}{y} \to 1$，取倒数（reciprocal）后得 $\dfrac{y}{\sinh y} \to 1$。现在固定 $x \neq 0$，并在 (i) 的恒等式中令 $n \to \infty$。设 $y = x/2^n$；那么 $y \to 0$，所以因子 $\dfrac{x/2^n}{\sinh(x/2^n)}$ 趋于 1。于是左边趋于 $\dfrac{\sinh x}{x}$，而右边是部分积（partial product）$P_n(x)$。所以部分积收敛（converge）至 $\dfrac{\sinh x}{x}$，这正是以下命题：

$$\frac{\sinh x}{x} = \cosh\left(\frac{x}{2}\right)\cosh\left(\frac{x}{4}\right)\cdots\cosh\left(\frac{x}{2^n}\right)\cdots.$$

第 (iii) 部分

取 $x = \ln 2$，则 $\mathrm{e}^{x/2^n} = 2^{1/2^n}$，所以

$$\cosh\left(\frac{x}{2^n}\right) = \frac{2^{1/2^n} + 2^{-1/2^n}}{2} = \frac{2^{2/2^n}+1}{2\times 2^{1/2^n}} = \frac{1+2^{1/2^{n-1}}}{2\times 2^{1/2^n}},$$

这里把分子和分母同乘以 $2^{1/2^n}$。当 $n=1$ 时，得 $\cosh\left(\frac{x}{2}\right)=\dfrac{1+2}{2\sqrt{2}}=\dfrac{3}{2\sqrt{2}}$；当 $n=2$ 时，得

$$\cosh\left(\frac{x}{4}\right)=\frac{1+2^{1/2}}{2\times 2^{1/4}},$$

即为所求。

在 (ii) 的左边，$\sinh(\ln 2)=\frac{1}{2}\left(2-\frac{1}{2}\right)=\frac{3}{4}$，所以 $\dfrac{\sinh x}{x}=\dfrac{3}{4\ln 2}$。

在右边，首 N 个因子的部分积为

$$P_N(\ln 2)=\frac{\prod_{n=1}^{N}\left(1+2^{1/2^{n-1}}\right)}{2^N\cdot 2^{\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^N}}}=\frac{3}{2}\cdot\prod_{k=1}^{N-1}\frac{1+2^{1/2^k}}{2}\cdot\frac{1}{2^{1-1/2^N}},$$

这里我们抽出了 $n=1$ 的因子 $(1+2)/2$，以 $k=n-1$ 重新编号，并把指数中的等比级数（geometric series）求和。当 $N\to\infty$ 时，$2^{1-1/2^N}\to 2$，而由 (ii) 得 $P_N(\ln 2)\to\dfrac{3}{4\ln 2}$。

因此部分积 $\prod_{k=1}^{N-1}\frac{1+2^{1/2^k}}{2}$ 收敛，而且

$$\frac{3}{4\ln 2}=\frac{3}{2}\cdot\frac{1}{2}\cdot\prod_{k=1}^{\infty}\frac{1+2^{1/2^k}}{2},$$

约去 $\frac{3}{4}$ 后得

$$\frac{1}{\ln 2}=\frac{1+2^{\frac{1}{2}}}{2}\times\frac{1+2^{\frac{1}{4}}}{2}\times\frac{1+2^{\frac{1}{8}}}{2}\times\cdots.$$

第 (iv) 部分

同一论证换成 sin 和 cos 后一字不改地适用，因为 $\sin 2\theta=2\sin\theta\cos\theta$ 具有相同的形式。由归纳法得

$$\sin x=2^n\cos\left(\frac{x}{2}\right)\cdots\cos\left(\frac{x}{2^n}\right)\sin\left(\frac{x}{2^n}\right).$$

当 $0<x<\pi$ 时，$0<x/2^n<\pi$，所以 $\sin(x/2^n)\neq 0$，可以像之前一样相除。麦克劳林级数 $\sin y=y-\frac{y^3}{3!}+\cdots$ 给出：当 $0<|y|\le 1$ 时，

$$\left|\frac{\sin y}{y}-1\right|\le y^2\sum_{r=1}^{\infty}\frac{1}{(2r+1)!}=y^2(\sinh 1-1)\to 0,$$

所以 $\sin y/y\to 1$，令 $n\to\infty$，得

$$\frac{\sin x}{x}=\cos\left(\frac{x}{2}\right)\cos\left(\frac{x}{4}\right)\cos\left(\frac{x}{8}\right)\cdots.$$

（另一个方法是在 (ii) 中代入 $x=\mathrm{i}t$，并利用 $\cosh(\mathrm{i}t)=\cos t$ 和 $\sinh(\mathrm{i}t)=\mathrm{i}\sin t$；但 (ii) 中的级数论证是针对实数 y 的，所以直接重做一遍会更干净。）

取 $x=\frac{\pi}{2}$：左边为 $\dfrac{1}{\pi/2}=\dfrac{2}{\pi}$，而各因子为 $\cos\left(\frac{\pi}{4}\right)$、$\cos\left(\frac{\pi}{8}\right)$、$\cos\left(\frac{\pi}{16}\right)$，如此类推。

想一想 怎样把 $\cos(\pi/8)$、$\cos(\pi/16)$、... 化成嵌套平方根?

把 $\cos 2\theta = 2\cos^2\theta - 1$ 写成 $2\cos\theta = \sqrt{2+2\cos 2\theta}$ 的形式；这里取正根，因为当 $0 < \theta < \frac{\pi}{2}$ 时 $\cos\theta > 0$。设 $c_n = 2\cos\left(\pi/2^{n+1}\right)$。那么 $c_1 = 2\cos\frac{\pi}{4} = \sqrt{2}$，而取 $\theta = \pi/2^{n+2}$，得

$$c_{n+1} = \sqrt{2+c_n}.$$

所以 $c_1 = \sqrt{2}$、$c_2 = \sqrt{2+\sqrt{2}}$、$c_3 = \sqrt{2+\sqrt{2+\sqrt{2}}}$，如此类推，而乘积的第 n 个因子是 $c_n/2$。因此

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2} \times \frac{\sqrt{2+\sqrt{2}}}{2} \times \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \times \cdots,$$

这就是 1593 年的韦达（Viète）公式。

验算

- 在 (i) 中代入 $n = 1$ 和 $x = 2\theta$：会得回倍角公式，理应如此。
- 在 (iii) 中，$\cosh(\frac{1}{2}\ln 2) = \frac{3}{2\sqrt{2}} \approx 1.0607$，而一般 n 的公式在 $n = 1$ 时与此吻合。
- 数值上，(iii) 的乘积很快便稳定下来：首 6o 个因子给出 $1.442695\ldots$，与 $1/\ln 2 = 1.4426950\ldots$ 吻合；而 (iv) 的乘积给出 $0.6366197\ldots$，与 $2/\pi$ 吻合。两者均已用电脑验证至约 15 位有效数字（significant figures）。

常见错误

- 除以 $\sinh(x/2^n)$ 时，没有指出它因 $x \neq 0$ 而不为零。
- 在 (iii) 中出现差一错误：$n = 1$ 的因子是 $3/(2\sqrt{2})$，并不是 $(1+2^{1/2^k})/2$ 的形式；而零散的幂 $2^{1/2}\cdot 2^{1/4}\cdots$ 合起来给出因子 2，而不是 1。
- 没有从级数作任何论证便假设“$y/\sinh y \to 1$”，而这正是题目要求证明的。

STEP 之外

无穷乘积 $\prod a_n$ 定义为其部分积的极限，正如无穷级数是其部分和（partial sums）的极限；取对数（logarithm）可以把两者互相转换，所以 $\prod(1+u_n)$ 的收敛与 $\sum u_n$ 的收敛密切相关。这里 $\sin x/x$ 的乘积，与欧拉（Euler）更著名的乘积 $\sin x = x\prod_{k\geq 1}\left(1 - x^2/(k^2\pi^2)\right)$ 并列，后者可推出 $\sum 1/k^2 = \pi^2/6$。若想了解如何严谨处理极限，使第 (ii) 部分这类论证变得严密，可阅读 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP III 2022，第 5 题

纯数学 · 对称定积分

STEP III 2022，第 5 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

在对称区间（symmetric interval）$[-a, a]$ 上，代换（substitution）$x \mapsto -x$ 让我们可以把积分“折叠”到 $[0, a]$ 上，把 x 处的值与 $-x$ 处的值配成一对。对于形如 $1/(1+\text{某式})$ 的被积函数（integrand），这一对 $\frac{1}{1+p} + \frac{1}{1+q}$ 恰好在 $pq = 1$ 时化简为 1，而 $\mathrm{e}^{x} \cdot \mathrm{e}^{-x} = 1$ 正是典型例子。整道题的重点，就是看出这种配对，并且知道在甚么情况下这是答案等于 a 的唯一途径。

解答

第 (i) 部分

设 $I = \displaystyle\int_{-a}^{a} \frac{1}{1+\mathrm{e}^{x}}\,\mathrm{d}x$。代入 $x = -u$（于是 $\mathrm{d}x = -\mathrm{d}u$，而积分上下限互换后又换回来），

$$I = \int_{-a}^{a} \frac{1}{1+\mathrm{e}^{-u}}\,\mathrm{d}u = \int_{-a}^{a} \frac{\mathrm{e}^{u}}{\mathrm{e}^{u}+1}\,\mathrm{d}u,$$

这里把分子和分母同乘以 e^{u}。把这式与 I 的原式相加，

$$2I = \int_{-a}^{a} \frac{1+\mathrm{e}^{x}}{1+\mathrm{e}^{x}}\,\mathrm{d}x = \int_{-a}^{a} 1\,\mathrm{d}x = 2a,$$

所以对所有 $a \geq 0$，都有 $I = a$，得证。

第 (ii) 部分

解释。对 $a \geq 0$，定义 $G(a) = \displaystyle\int_{0}^{a} g(x)\,\mathrm{d}x$。由于 g 连续（continuous），微积分基本定理（fundamental theorem of calculus）告诉我们 G 可导（differentiable），且 $G'(a) = g(a)$

（在 $a=0$ 处为单侧导数（one-sided derivative））。但 G 在 $[0,\infty)$ 上恒等于零，所以它在该处的导数为零。因此对所有 $a\geq 0$，都有 $g(a)=0$。（连续性很重要：没有它的话，g 可以在个别孤立点上不为零，而不影响任何积分。）

折叠积分。在 0 处分开，并在左半部分代入 $x=-u$：

$$\int_{-a}^{0}\frac{\mathrm{d}x}{1+\mathrm{f}(x)}=\int_{0}^{a}\frac{\mathrm{d}u}{1+\mathrm{f}(-u)}.$$

把 a 写成 $a=\int_0^a 1\,\mathrm{d}x$，便得到

$$\int_{-a}^{a}\frac{\mathrm{d}x}{1+\mathrm{f}(x)}-a=\int_{0}^{a}g(x)\,\mathrm{d}x,\qquad \text{其中}\quad g(x)=\frac{1}{1+\mathrm{f}(x)}+\frac{1}{1+\mathrm{f}(-x)}-1.$$

由于 $\mathrm{f}\geq 0$，两个分母都至少为 1，所以 g 连续。

> **想一想** 刚才解释的结果只是单向的命题。它给出“当且仅当”（if and only if）的哪一个方向？另一个方向又从何而来？

- 若对所有 $x\geq 0$ 都有 $g(x)=0$，则 $\int_0^a g=0$，于是对所有 $a\geq 0$，该积分都等于 a。这个方向是显然的。
- 反过来，若对所有 $a\geq 0$，该积分都等于 a，则对所有 $a\geq 0$ 都有 $\int_0^a g=0$，再由第一段的结果（套用于这个连续的 g），得出对所有 $x\geq 0$ 都有 $g(x)=0$。

所以积分条件成立，当且仅当对所有 $x\geq 0$ 都有 $g(x)=0$。

乘积条件。设 $p=\mathrm{f}(x)$，$q=\mathrm{f}(-x)$。两边同乘以 $(1+p)(1+q)>0$，

$$\begin{aligned}\frac{1}{1+p}+\frac{1}{1+q}=1 &\iff (1+q)+(1+p)=(1+p)(1+q)\\ &\iff 2+p+q=1+p+q+pq\\ &\iff pq=1.\end{aligned}$$

所以对所有 $x\geq 0$ 都有 $g(x)=0$，当且仅当对所有 $x\geq 0$ 都有 $\mathrm{f}(x)\mathrm{f}(-x)=1$。最后，把 x 换成 $-x$ 时，式子 $\mathrm{f}(x)\mathrm{f}(-x)$ 不变，所以它对所有 $x\geq 0$ 等于 1，恰好等同于它对所有实数 x 等于 1。这样便完成了整条等价（equivalence）链。

第 (iii) 部分

跟之前完全一样地折叠积分，这次在左半部分用上 $\mathrm{h}(-u)=\mathrm{h}(u)$：

$$\begin{aligned}\int_{-a}^{a}\frac{\mathrm{h}(x)}{1+\mathrm{f}(x)}\,\mathrm{d}x &= \int_{0}^{a}\frac{\mathrm{h}(x)}{1+\mathrm{f}(x)}\,\mathrm{d}x+\int_{0}^{a}\frac{\mathrm{h}(-u)}{1+\mathrm{f}(-u)}\,\mathrm{d}u\\ &= \int_{0}^{a}\mathrm{h}(x)\left(\frac{1}{1+\mathrm{f}(x)}+\frac{1}{1+\mathrm{f}(-x)}\right)\mathrm{d}x\\ &= \int_{0}^{a}\mathrm{h}(x)\,\mathrm{d}x,\end{aligned}$$

这是因为根据第 (ii) 部分证明的等价关系，只要 $\mathrm{f}(x)\mathrm{f}(-x)=1$，括号内的式子便等于 1。（条件 $\mathrm{f}\geq 0$ 及连续性只是确保每个被积函数都连续。）

第 (iv) 部分

> ? **想一想** 怎样把 $\dfrac{\mathrm{e}^{-x}}{\cosh x}$ 写成 $\dfrac{\text{常数}}{1+\mathrm{f}(x)}$ 的形式?

由于 $\cosh x=\frac{1}{2}(\mathrm{e}^{x}+\mathrm{e}^{-x})$，

$$\frac{\mathrm{e}^{-x}\cos x}{\cosh x}=\frac{2\mathrm{e}^{-x}\cos x}{\mathrm{e}^{x}+\mathrm{e}^{-x}}=\frac{2\cos x}{1+\mathrm{e}^{2x}},$$

这里把分子和分母同乘以 e^{x}。取 $\mathrm{f}(x)=\mathrm{e}^{2x}$，它连续、为正，并满足 $\mathrm{f}(x)\mathrm{f}(-x)=\mathrm{e}^{2x}\mathrm{e}^{-2x}=1$；再取 $\mathrm{h}(x)=2\cos x$，它连续且为偶函数（even function）。以 $a=\frac{1}{2}\pi$ 套用第 (iii) 部分，得

$$\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi}\frac{\mathrm{e}^{-x}\cos x}{\cosh x}\,\mathrm{d}x=\int_{0}^{\frac{1}{2}\pi}2\cos x\,\mathrm{d}x=\Big[2\sin x\Big]_{0}^{\frac{1}{2}\pi}=2.$$

验算

- 第 (i) 部分是第 (iii) 部分在 $\mathrm{f}(x)=\mathrm{e}^{x}$、$\mathrm{h}\equiv 1$ 时的特例，给出 $\int_{0}^{a}1\,\mathrm{d}x=a$，与预期相符。
- 一个反例：$\mathrm{f}(x)=x^{2}$ 有 $\mathrm{f}(x)\mathrm{f}(-x)=x^{4}\neq 1$，而确实 $\int_{-1}^{1}\frac{\mathrm{d}x}{1+x^{2}}=\frac{\pi}{2}\neq 1$。
- 用辛普森法则（Simpson's rule）对第 (iv) 部分的积分作数值计算，得到 2.0000000；第 (iii) 部分的恒等式亦已用电脑对多个随机选取的 $\mathrm{f}=\mathrm{e}^{(\text{奇函数})}$ 及偶函数 h 加以验证。

常见错误

- 在第 (ii) 部分只证明了“当且仅当”的其中一个方向：从积分条件推回 $g = 0$ 的那一步，正正需要用到第一个结果（以及 g 的连续性）。
- 忘记说明为何 $\mathrm{f}(x)\mathrm{f}(-x) = 1$ 对 $x \geq 0$ 成立便足以推出它对所有 x 成立：原因是这个式子在 $x \mapsto -x$ 之下对称。
- 在第 (iv) 部分，试图直接积分 $\mathrm{e}^{-x}\cos x/\cosh x$，而不是先把它改写成 $2\cos x/(1+\mathrm{e}^{2x})$。

STEP 之外

第 (ii) 部分的论证，其实是把微积分基本定理“倒过来”用；这也是分析（analysis）中经常出现连续性假设的原因：积分“看不见”个别点上发生的事，所以只有在连续的情况下，才能从一个函数的积分还原出它的函数值。折叠技巧与把函数拆成偶部分和奇部分的做法是近亲；在这里，$\frac{1}{1+\mathrm{f}(x)} - \frac{1}{2}$ 是奇函数（odd function），恰好当 $\mathrm{f}(x)\mathrm{f}(-x) = 1$。Abbott 的 *Understanding Analysis*（《理解分析》）对微积分基本定理以及为何需要连续性有仔细的论述。

详细解答

STEP III 2022，第 6 题

纯数学 · 滚动圆与切线

STEP III 2022，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

第 (i) 部分是一个二阶泰勒展开（Taylor expansion），而它真正的重点在最后一行的提醒：当一阶项消失时，比值由二阶项决定。在第 (ii) 部分中，滚动条件表明两个圆经过接触点的弧长相等，由此可知 C_1 转了多少。在 (c) 的特殊点上，弦（chord）中的一阶项恰好互相抵消，所以我们需要用到 (i) 中的二阶项。

解答

第 (i) 部分

当 α 很小时，$\cos\alpha \approx 1 - \frac{1}{2}\alpha^2$ 及 $\sin\alpha \approx \alpha$，误差的阶为 α^3。因此

$$\begin{aligned}\cos(\theta+\alpha) - \cos\theta &= \cos\theta\cos\alpha - \sin\theta\sin\alpha - \cos\theta \\ &\approx \cos\theta\left(1 - \tfrac{1}{2}\alpha^2\right) - \alpha\sin\theta - \cos\theta = -\alpha\sin\theta - \tfrac{1}{2}\alpha^2\cos\theta,\end{aligned}$$

此即所求。同理，

$$\sin(\theta+\alpha) - \sin\theta \approx \alpha\cos\theta - \tfrac{1}{2}\alpha^2\sin\theta.$$

把 $(*)$ 的分子和分母同除以 α，

$$\frac{\sin(\theta+\alpha) - \sin\theta}{\cos(\theta+\alpha) - \cos\theta} \approx \frac{\cos\theta - \frac{1}{2}\alpha\sin\theta}{-\sin\theta - \frac{1}{2}\alpha\cos\theta} \longrightarrow -\frac{\cos\theta}{\sin\theta} = -\cot\theta,$$

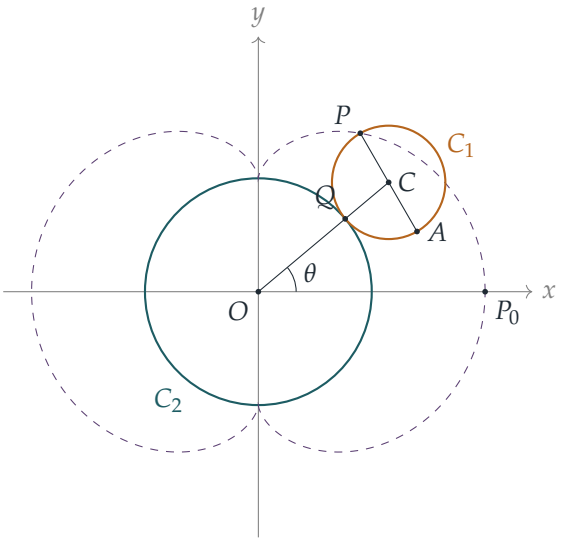
这个极限（limit）没有问题，因为分母趋向 $-\sin\theta \neq 0$。

若 $\sin\theta = 0$，则 $\cos\theta = \pm 1$，分母的一阶项便消失：

$$(*) \approx \frac{\alpha\cos\theta}{-\frac{1}{2}\alpha^2\cos\theta} = -\frac{2}{\alpha}.$$

所以 $(*)$ 没有有限的极限：它的大小无界地增长，当 $\alpha \to 0^+$ 时趋向 $-\infty$，当 $\alpha \to 0^-$ 时趋向 $+\infty$。（从几何上看，单位圆上由角 θ 到 $\theta+\alpha$ 的弦变成铅垂，其斜率（gradient）趋向无限大。）

第 (ii)(a) 部分



图中所示为 $n=3$ 的情况，虚线曲线是 P 的轨迹。设 C 为 C_1 的圆心。由于两圆在 Q 相切，点 O、Q、C 共线（collinear），且 $OC=(n-1)a+a=na$，所以

$$\overrightarrow{OC}=na(\cos\theta,\sin\theta),$$

而从 C 看，点 Q 位于方向 $\theta+\pi$。

设 A 为 C_1 上与 P 直径相对（diametrically opposite）的点；这就是开始时与 C_2 相接触的 C_1 上的点，位于 $((n-1)a,0)$。无滑动滚动（rolling without slipping）表示 C_1 上由 A 到 Q 的弧长，等于接触点在 C_2 上走过的弧长，即 $(n-1)a\theta$。由于 C_1 的半径为 a，角 ACQ 等于 $(n-1)\theta$。

? **想一想** 从 C 看，A 位于 Q 的顺时针方向还是逆时针方向?

当 C_1 沿 C_2 外侧逆时针滚动时，C_1 本身也逆时针自转，所以它已经滚过的点会被逆时针带离接触点。因此 $\overrightarrow{CA}$ 指向方向 $\theta+\pi+(n-1)\theta=n\theta+\pi$，而 $\overrightarrow{CP}=-\overrightarrow{CA}$ 指向方向 $n\theta$。（验证：当 $\theta=0$ 时，这给出 $P=((n+1)a,0)$。C_1 总共转了 $n\theta$，这就是一枚硬币绕另一枚硬币滚动时那个熟悉的“多转一圈”。）因此

$$\overrightarrow{OP}=na(\cos\theta,\sin\theta)+a(\cos n\theta,\sin n\theta),$$

所以 $x(\theta) = a(n\cos\theta + \cos n\theta)$，此即所求，而且

$$y(\theta) = a(n\sin\theta + \sin n\theta).$$

第 (ii)(b) 部分

把两式平方后相加，

$$OP^2 = a^2\big(n^2 + 1 + 2n(\cos\theta\cos n\theta + \sin\theta\sin n\theta)\big) = a^2\big(n^2 + 1 + 2n\cos(n-1)\theta\big).$$

令它等于 $(n-1)^2a^2 = (n^2 - 2n + 1)a^2$，得 $2n\cos(n-1)\theta = -2n$，即 $\cos(n-1)\theta = -1$。所以

$$\theta = \frac{(2k+1)\pi}{n-1}, \qquad k \in \mathbb{Z},$$

在绕 C_2 的每一圈中给出 $n-1$ 个不同的位置。这些正是 P 本身成为接触点的时刻：曲线的尖点（cusp）。

第 (ii)(c) 部分

设 $\theta_0 = \pi/(n-1)$，则 $n\theta_0 = \theta_0 + \pi$，所以

$$\cos n\theta_0 = -\cos\theta_0, \qquad \sin n\theta_0 = -\sin\theta_0.$$

对每一项应用第 (i) 部分，其中 $n\theta$ 项以 $n\alpha$ 作为微小增量：

$$\begin{aligned}
\frac{y(\theta_0+\alpha) - y(\theta_0)}{a} &\approx n\big(\alpha\cos\theta_0 - \tfrac{1}{2}\alpha^2\sin\theta_0\big) \\
&\quad + \big(n\alpha\cos n\theta_0 - \tfrac{1}{2}n^2\alpha^2\sin n\theta_0\big) \\
&= n\alpha\cos\theta_0 - \tfrac{1}{2}n\alpha^2\sin\theta_0 - n\alpha\cos\theta_0 + \tfrac{1}{2}n^2\alpha^2\sin\theta_0 \\
&= \tfrac{1}{2}n(n-1)\alpha^2\sin\theta_0,
\end{aligned}$$

$$\begin{aligned}
\frac{x(\theta_0+\alpha) - x(\theta_0)}{a} &\approx n\big(-\alpha\sin\theta_0 - \tfrac{1}{2}\alpha^2\cos\theta_0\big) \\
&\quad + \big(-n\alpha\sin n\theta_0 - \tfrac{1}{2}n^2\alpha^2\cos n\theta_0\big) \\
&= \tfrac{1}{2}n(n-1)\alpha^2\cos\theta_0,
\end{aligned}$$

两者的误差都是 α^3 阶。一阶项恰好互相抵消，正如 (i) 中 $\sin\theta = 0$ 的情况，所以答案由二阶项决定。把分子和分母同除以 $\frac{1}{2}n(n-1)\alpha^2$，

$$\frac{y(\theta_0+\alpha) - y(\theta_0)}{x(\theta_0+\alpha) - x(\theta_0)} = \frac{\sin\theta_0 + O(\alpha)}{\cos\theta_0 + O(\alpha)} \longrightarrow \tan\theta_0,$$

前提是 $\cos\theta_0 \neq 0$，而这在 $n \geq 4$ 时成立。当 $n = 3$ 时，$\theta_0 = \frac{1}{2}\pi$；同样的推导显示倒数比值趋向 $\cot\theta_0 = 0$，所以弦变成铅垂。

这个极限就是 θ_0 处切线（tangent）的斜率，因为它是弦斜率的极限。而

$$x(\theta_0) = a(n\cos\theta_0 - \cos\theta_0) = (n-1)a\cos\theta_0, \qquad y(\theta_0) = (n-1)a\sin\theta_0,$$

所以 OP 与 x 轴成角 θ_0，斜率为 $\tan\theta_0$（当 $n = 3$ 时为铅垂）。切线与 OP 方向相同，所以 $(x(\theta_0), y(\theta_0))$ 处的切线平行于 OP。事实上，由于切线经过 P，它就落在直线 OP 上。

验算

- 当 $\theta = 0$、$\alpha = 0.001$ 时，比值 $(*)$ 为 -1999.9998，对比 $-2/\alpha = -2000$。
- 当 $n = 5$、$\theta_0 = \frac{1}{4}\pi$ 及 $\alpha = 10^{-4}$ 时，弦的斜率为 1.0004，接近 $\tan\frac{1}{4}\pi = 1$。
- 在 θ_0 处，$OP = (n-1)a$，所以 (c) 中的点就是 (b) 中求得的 $k = 0$ 尖点。这就是该处切线沿径向的原因。

常见错误

- 把 C_1 的转角当作 $(n-1)\theta$（滚过的弧），忘记了 C 绕 O 运行所带来的额外 θ。
- 在 (c) 中只用一阶近似：这只会得到 $0/0$，甚么也说明不了。
- 忘记 $n = 3$ 时 $\theta_0 = \frac{1}{2}\pi$，此时 $\tan\theta_0$ 没有定义。

STEP 之外

这条曲线是一条有 $n-1$ 个尖点的外摆线（epicycloid）。在每个尖点处，速度（velocity）$(x'(\theta), y'(\theta))$ 为零，所以一般的公式 $\mathrm{d}y/\mathrm{d}x = y'/x'$ 不再适用。第 (c) 部分通过推到二阶来处理这个问题，这其实就是乔装了的洛必达法则（l'Hôpital's rule）。参数化曲线（parametrised curve）上速度为零的点称为奇点（singular points），它们是曲线微分几何（differential geometry）入门课程的研究对象。关于第 (i) 部分背后的极限与泰勒近似，可参阅 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP III 2022，第 7 题

纯数学 · 向量积变换

STEP III 2022，第 7 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

本题在 *Beyond STEP* 第 7.5 节三维空间中的旋转中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch07/rotations-3d.html

关键思路

一切都建基于把一个向量（vector）分解为沿 $\mathbf{n}$ 的部分和垂直（perpendicular）于 $\mathbf{n}$ 的部分。映射（map）$\mathbf{r} \mapsto \mathbf{n} \times \mathbf{r}$ 会消去平行部分，并把垂直部分绕 $\mathbf{n}$ 转过一个直角，而不改变其长度。第 (i) 部分一旦证明了连续施行这个映射两次会得到垂直部分的负值，第 (ii) 和 (iii) 部分的映射便可以逐一观察它们分别对这两个部分的作用而直接读出。

解答

第 (i) 部分

设 $\mathbf{n} = (a,b,c)^{\mathsf{T}}$ 及 $\mathbf{r} = (x,y,z)^{\mathsf{T}}$，则

$$\mathrm{f}(\mathbf{r}) = \mathbf{n} \times \mathbf{r} = \begin{pmatrix} bz - cy \\ cx - az \\ ay - bx \end{pmatrix}.$$

再施行 f 一次，$\mathbf{n} \times \mathrm{f}(\mathbf{r})$ 的 x 分量（component）等于 b 乘以 $\mathrm{f}(\mathbf{r})$ 的第三个分量，减去 c 乘以第二个分量：

$$\begin{aligned} b(ay - bx) - c(cx - az) &= aby - b^2x - c^2x + acz \\ &= -x(b^2 + c^2) + aby + acz, \end{aligned}$$

此即所求。由于 $\mathbf{n}$ 是单位向量（unit vector），$b^2+c^2=1-a^2$，所以上式变成

$$-x+a^2x+aby+acz=a(ax+by+cz)-x=(\mathbf{n}\cdot\mathbf{r})\,a-x.$$

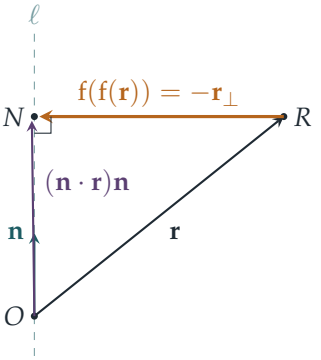
这正是 $(\mathbf{n}\cdot\mathbf{r})\mathbf{n}-\mathbf{r}$ 的 x 分量。由于叉积（cross product）在这种循环重新标记下保持不变，只要循环替换 $a\to b\to c\to a$ 及 $x\to y\to z\to x$，便可同样得出 y 分量和 z 分量。因此

$$\mathrm{f}(\mathrm{f}(\mathbf{r}))=(\mathbf{n}\cdot\mathbf{r})\mathbf{n}-\mathbf{r}.$$

几何意义。写 $\mathbf{r}=\mathbf{r}_{\parallel}+\mathbf{r}_{\perp}$，其中 $\mathbf{r}_{\parallel}=(\mathbf{n}\cdot\mathbf{r})\mathbf{n}$ 是沿 $\mathbf{n}$ 的分量，而 $\mathbf{r}_{\perp}=\mathbf{r}-(\mathbf{n}\cdot\mathbf{r})\mathbf{n}$ 垂直于 $\mathbf{n}$。于是

$$\mathrm{f}(\mathrm{f}(\mathbf{r}))=-\mathbf{r}_{\perp}.$$

设 ℓ 为通过 O 而方向为 $\mathbf{n}$ 的直线，并设 N 为由 R 到 ℓ 的垂足（foot of the perpendicular），故 $\overrightarrow{ON}=\mathbf{r}_{\parallel}$。于是 $\mathrm{f}(\mathrm{f}(\mathbf{r}))=\overrightarrow{ON}-\overrightarrow{OR}=\overrightarrow{RN}$：它是由 R 垂直地直指轴 ℓ 的向量，其长度就是 R 与 ℓ 的距离。



第 (ii) 部分

轴。由于 $\mathbf{n}\times\mathbf{n}=\mathbf{0}$，我们有 $\mathrm{f}(\mathbf{n})=\mathbf{0}$，因而 $\mathrm{f}(\mathrm{f}(\mathbf{n}))=\mathbf{0}$，得 $\mathrm{g}(\mathbf{n})=\mathbf{n}$。由线性（linearity）得 $\mathrm{g}(\lambda\mathbf{n})=\lambda\mathbf{n}$，所以直线 ℓ 上每一点都是不动点（fixed point）。

垂直向量。现在取满足 $\mathbf{n}\cdot\mathbf{r}=0$ 的 $\mathbf{r}$。由第 (i) 部分，$\mathrm{f}(\mathrm{f}(\mathbf{r}))=-\mathbf{r}$，所以

$$\mathrm{g}(\mathbf{r})=\mathbf{r}+\sin\theta\,(\mathbf{n}\times\mathbf{r})-(1-\cos\theta)\mathbf{r}=\cos\theta\,\mathbf{r}+\sin\theta\,(\mathbf{n}\times\mathbf{r}).$$

? **想一想** 当 $\mathbf{r}\perp\mathbf{n}$ 时，$\mathbf{n}\times\mathbf{r}$ 是怎样的?

它同时垂直于 $\mathbf{n}$ 和 $\mathbf{r}$，所以它位于通过 O 且垂直于 $\mathbf{n}$ 的平面上，并与 $\mathbf{r}$ 成直角。它的长度是 $|\mathbf{n}||\mathbf{r}|\sin 90^\circ=|\mathbf{r}|$。因此 $\mathbf{r}$ 与 $\mathbf{n}\times\mathbf{r}$ 是该平面上两个互相垂直、长度相等的向量，而 $\mathbf{n}\times\mathbf{r}$ 就是把 $\mathbf{r}$ 绕 $\mathbf{n}$ 按右手方向（right-handed sense）转过 90° 的结果（即从 $\mathbf{n}$ 的末端回望 O

时的逆时针方向)。正如 $\cos\theta\,\mathbf{i}+\sin\theta\,\mathbf{j}$ 是把 $\mathbf{i}$ 朝 $\mathbf{j}$ 旋转 θ 所得，向量 $\cos\theta\,\mathbf{r}+\sin\theta\,(\mathbf{n}\times\mathbf{r})$ 就是把 $\mathbf{r}$ 绕 $\mathbf{n}$ 按同一方向旋转 θ 所得。

一般向量。f 是线性的（叉积对第二个自变量是线性的），所以 g 也是线性的。把一般的 $\mathbf{r}$ 写成 $\mathbf{r}_{\parallel}+\mathbf{r}_{\perp}$，g 保持 $\mathbf{r}_{\parallel}$ 不变，并把 $\mathbf{r}_{\perp}$ 绕 $\mathbf{n}$ 旋转 θ，这正是绕 ℓ 的旋转所做的事。

因此 $R\mapsto P$ 是一个**绕通过 O、方向为 $\mathbf{n}$ 的直线旋转 θ 角的旋转（rotation）**，方向为绕 $\mathbf{n}$ 的右手方向（即从 $\mathbf{n}$ 的末端望向 O 时的逆时针方向)。

第 (iii) 部分

利用第 (i) 部分，

$$\mathrm{h}(\mathbf{r})=-\mathbf{r}-2\big((\mathbf{n}\cdot\mathbf{r})\mathbf{n}-\mathbf{r}\big)=\mathbf{r}-2(\mathbf{n}\cdot\mathbf{r})\mathbf{n}.$$

以两个分量表示，$\mathrm{h}(\mathbf{r})=\mathbf{r}_{\perp}-\mathbf{r}_{\parallel}$：

- $\mathrm{h}(\mathbf{n})=\mathbf{n}-2\mathbf{n}=-\mathbf{n}$，所以沿 $\mathbf{n}$ 的分量被反转；
- 若 $\mathbf{n}\cdot\mathbf{r}=0$，则 $\mathrm{h}(\mathbf{r})=\mathbf{r}$，所以平面 $\mathbf{n}\cdot\mathbf{r}=0$ 上每一点都是不动点。

从几何上看，$\overrightarrow{RQ}=-2(\mathbf{n}\cdot\mathbf{r})\mathbf{n}$ 垂直于平面 $\Pi:\ \mathbf{n}\cdot\mathbf{r}=0$，而 RQ 的中点的位置向量（position vector）是 $\mathbf{r}-(\mathbf{n}\cdot\mathbf{r})\mathbf{n}$，它位于 Π 上。所以 Q 是 R 的镜像：映射 $R\mapsto Q$ 是**对通过 O 且垂直于 $\mathbf{n}$ 的平面的反射（reflection）**。

验算

- 取 $\mathbf{n}=\mathbf{k}=(0,0,1)^{\mathsf{T}}$。则 $\mathbf{n}\times\mathbf{r}=(-y,x,0)^{\mathsf{T}}$，施行两次得 $(-x,-y,0)^{\mathsf{T}}$，与 $(\mathbf{n}\cdot\mathbf{r})\mathbf{n}-\mathbf{r}=(0,0,z)^{\mathsf{T}}-(x,y,z)^{\mathsf{T}}$ 吻合。于是 $\mathrm{g}(\mathbf{r})=(x\cos\theta-y\sin\theta,\ x\sin\theta+y\cos\theta,\ z)^{\mathsf{T}}$，正是我们熟悉的绕 z 轴的旋转；而 $\mathrm{h}(\mathbf{r})=(x,y,-z)^{\mathsf{T}}$，是对平面 $z=0$ 的反射。
- $\theta=0$ 给出 $\mathrm{g}(\mathbf{r})=\mathbf{r}$，而施行 h 两次得到恒等映射（identity），这正符合“转零角度的旋转”和反射应有的性质。
- 以 2000 组随机的单位向量 $\mathbf{n}$、向量 $\mathbf{r}$ 和角 θ 进行电脑验算，证实了 x 分量的公式、$\mathrm{f}(\mathrm{f}(\mathbf{r}))$ 的恒等式、g 与绕 $\mathbf{n}$ 按右手方向旋转 θ 一致，以及 h 与对平面 $\mathbf{n}\cdot\mathbf{r}=0$ 的反射一致。

常见错误

- 忘记使用 $a^2+b^2+c^2=1$：若 $\mathbf{n}$ 不是单位向量，(i) 中的恒等式并不成立。
- 在 (ii) 中只验证了 $\mathbf{r}\perp\mathbf{n}$ 的情况，却没有解释为何这一点连同 $\mathrm{g}(\mathbf{n})=\mathbf{n}$ 及线性，便足以处理所有 $\mathbf{r}$。

- 把 (iii) 说成是对直线 ℓ 的反射：ℓ 上的点是被反转，而不是保持不动；保持不动的是垂直于 $\mathbf{n}$ 的平面。

STEP 之外

(ii) 中的公式就是 Rodrigues 旋转公式（Rodrigues′ rotation formula），而 f 的矩阵是反对称（antisymmetric）的“叉积矩阵” K，所以旋转可写成 $I + \sin\theta\, K + (1 - \cos\theta)K^2$，这也等于 $\exp(\theta K)$。(iii) 中的映射是 Householder 反射 $I - 2\mathbf{n}\mathbf{n}^{\mathsf{T}}$，是数值线性代数（numerical linear algebra）中的重要工具。在大学一年级的课程中，它们分别以行列式（determinant）为 $+1$ 和 -1 的正交矩阵（orthogonal matrix）的身份出现。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）是认识它们的好起点。

详细解答

STEP III 2022，第 8 题

纯数学 · 棣美弗定理与无理性

STEP III 2022，第 8 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

本题在 *Beyond STEP* 第 5.5 节名不副实的有理角中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch05/rational-angles.html

关键思路

把 $\cos\theta$ 提出后套用棣美弗定理（De Moivre's theorem），可将 $\sin k\theta$ 和 $\cos k\theta$ 表示为 $\cos^k\theta$（或 $\sin\theta\cos^{k-1}\theta$）乘以一个关于 $\tan^2\theta = \sec^2\theta - 1$ 的多项式（polynomial）。当 $\sec^2\theta - 1$ 是偶整数时，该多项式的值是一个整数，而它的奇偶性（parity）可以从首项直接读出。若角度（以度为单位）是有理数（rational），某个 $\sin k\theta$ 便会等于零，而奇偶性论证说明这种情况永远不会发生。

解答

第 (i) 部分

假设 $\cos\theta \neq 0$，使 $\sec\theta$ 存在。由棣美弗定理，

$$\cos k\theta + \mathrm{i}\sin k\theta = (\cos\theta + \mathrm{i}\sin\theta)^k = \cos^k\theta\,(1+\mathrm{i}t)^k, \qquad t = \tan\theta.$$

按二项式展开（binomial expansion），$(1+\mathrm{i}t)^k = \sum_{j=0}^{k}\binom{k}{j}\mathrm{i}^j t^j$。$j$ 为奇数的项是虚数，且 $\mathrm{i}^{2r+1} = (-1)^r\mathrm{i}$；$j$ 为偶数的项是实数，且 $\mathrm{i}^{2r} = (-1)^r$。取虚部（imaginary part），

$$\begin{aligned}\sin k\theta &= \cos^k\theta\left(\binom{k}{1}t - \binom{k}{3}t^3 + \binom{k}{5}t^5 - \cdots\right)\\ &= \cos^k\theta\,\tan\theta\left(k - \binom{k}{3}t^2 + \binom{k}{5}t^4 - \cdots\right).\end{aligned}$$

由于 $\cos^k\theta\tan\theta = \sin\theta\cos^{k-1}\theta$ 及 $t^2 = \tan^2\theta = \sec^2\theta - 1$，上式即为

$$\sin k\theta = \sin\theta\cos^{k-1}\theta\left(k - \binom{k}{3}(\sec^2\theta-1) + \binom{k}{5}(\sec^2\theta-1)^2 - \cdots\right),$$

此即所求。括号会自然终止，因为当 $j > k$ 时 $\binom{k}{j} = 0$。同样地取实部，

$$\cos k\theta = \cos^k\theta\left(1 - \binom{k}{2}(\sec^2\theta-1) + \binom{k}{4}(\sec^2\theta-1)^2 - \cdots\right).$$

第 (ii) 部分

此处 $\cos\theta = 1/a$ 且 $a \geq 3$，故 $0° < \theta < 90°$。特别地，$\sin\theta \neq 0$ 且 $\cos\theta \neq 0$。此外 $\sec^2\theta - 1 = a^2 - 1$，由于 a 是奇数，这是一个 偶整数。记 $X = a^2 - 1$。

假设 k 是使 $\sin k\theta = 0$ 的最小正整数。那么 $k \neq 1$，因为 $\sin\theta \neq 0$，所以 $k > 1$，可应用第 (i) 部分：

$$0 = \sin k\theta = \sin\theta\cos^{k-1}\theta\left(k - \binom{k}{3}X + \binom{k}{5}X^2 - \cdots\right).$$

因子 $\sin\theta\cos^{k-1}\theta$ 不为零，所以括号等于零。首项以外的每一项都至少含有一个偶整数 X 的因子，而二项式系数（binomial coefficient）都是整数，所以

$$0 = k - \binom{k}{3}X + \binom{k}{5}X^2 - \cdots \equiv k \pmod 2.$$

因此 k 是偶数。于是

$$0 = \sin k\theta = 2\sin\left(\tfrac{1}{2}k\theta\right)\cos\left(\tfrac{1}{2}k\theta\right).$$

但 $\frac{1}{2}k$ 是整数且 $0 < \frac{1}{2}k < k$，所以由 k 的最小性（minimality），$\sin(\frac{1}{2}k\theta) \neq 0$。因此 $\cos(\frac{1}{2}k\theta) = 0$，此即所求。

想一想 “θ 是有理数”这个假设在哪里用上?

正是它令 k 得以存在。设 $\theta = p/q$，其中 p、q 为正整数（记住 θ 以度为单位）。那么 $180q\,\theta = 180p$，故 $\sin(180q\,\theta) = 0$。所以确实存在正整数 m 使 $\sin m\theta = 0$，因此也存在最小的一个，记为 k。它满足上述假设，所以 k 是偶数；记 $n = \frac{1}{2}k$，便有 $\cos n\theta = 0$。

现在证明 $\cos n\theta$ 不可能为零。若 $n = 1$，则 $\cos n\theta = 1/a \neq 0$。若 $n > 1$，第 (i) 部分的余弦公式给出

$$\cos n\theta = \frac{1}{a^n}\left(1 - \binom{n}{2}X + \binom{n}{4}X^2 - \cdots\right).$$

同样, 首项以外的每一项都是偶数, 所以括号是一个奇整数, 特别地不为零。故 $\cos n\theta \neq 0$, 矛盾。因此 θ 是无理数（irrational）。

第 (iii) 部分

现在 $\cot\phi = 1/b$，所以 $\tan\phi = b$，其中 $b \geq 2$，且 $0° < \phi < 90°$。同样 $\sin\phi \neq 0$ 且 $\cos\phi \neq 0$，而这次

$$\sec^2\phi - 1 = \tan^2\phi = b^2,$$

由于 b 是偶数，这是一个偶整数。所以只要以 $X = b^2$ 代替 $a^2 - 1$，第 (ii) 部分的论证便完全适用。完整写出如下：

设 $\phi = p/q$，其中 p、q 为正整数。那么 $\sin(180q\,\phi) = 0$，所以存在最小的正整数 k 使 $\sin k\phi = 0$，而由于 $\sin\phi \neq 0$，$k > 1$。由第 (i) 部分，因为 $\sin\phi\cos^{k-1}\phi \neq 0$，

$$0 = k - \binom{k}{3}b^2 + \binom{k}{5}b^4 - \cdots \equiv k \pmod 2,$$

所以 k 是偶数。于是 $0 = \sin k\phi = 2\sin(\tfrac{1}{2}k\phi)\cos(\tfrac{1}{2}k\phi)$，而由最小性，$\sin(\tfrac{1}{2}k\phi) \neq 0$，所以 $\cos n\phi = 0$，其中 $n = \tfrac{1}{2}k$。若 $n = 1$，这表示 $\cos\phi = 0$，但这是错的。若 $n > 1$，

$$\cos n\phi = \cos^n\phi\left(1 - \binom{n}{2}b^2 + \binom{n}{4}b^4 - \cdots\right),$$

其中 $\cos^n\phi \neq 0$，而括号是奇数，故不为零。这个矛盾说明 ϕ 是无理数。

验算

- 在第 (i) 部分代入 $k = 2$ 及 $k = 3$：公式给出 $\sin 2\theta = 2\sin\theta\cos\theta$、$\cos 2\theta = \cos^2\theta - \sin^2\theta$ 及 $\sin 3\theta = \sin\theta(3\cos^2\theta - \sin^2\theta) = 3\sin\theta - 4\sin^3\theta$。对于不超过 15 的 k，两条公式在随机角度下也与直接计算的结果吻合。
- 奇偶条件是必要的。取 $a = 2$（偶数）得 $\theta = 60°$，取 $b = 1$（奇数）得 $\phi = 45°$，两者都是有理数。在这两种情况下，$X = 3$ 或 1 都是奇数，所以论证恰好在应该失效的地方失效。
- 当 $a = 3$ 时 $\theta \approx 70.5288°$，当 $b = 2$ 时 $\phi \approx 63.4349°$。电脑验算确认，对于所有小于 30 的奇数 a 及偶数 b，以及每个 $k < 2000$，$\sin k\theta$ 和 $\sin k\phi$ 都与零保持相当距离。

常见错误

- 在约去 $\sin\theta\cos^{k-1}\theta$ 之前，没有说明为何它 $\neq 0$。
- 没有解释 k 从何而来：最小的 k 之所以存在，只是因为有理数 θ 令 $\sin(180q\,\theta) = 0$。
- 没有利用 k 的最小性排除 $\sin(\tfrac{1}{2}k\theta) = 0$，便断定 $\cos(\tfrac{1}{2}k\theta) = 0$。

STEP 之外

撇除前面的因子，这里的括号就是切比雪夫多项式（Chebyshev polynomials），而奇偶性技巧则是把整数恒等式模一个质数（reducing modulo a prime）来处理的初步体验。一个著名的相关结果是 *Niven* 定理（Niven's theorem）：在 180° 的有理数倍 θ 之中，使 $\cos\theta$ 为有理数的只有满足 $\cos\theta \in \{0, \pm\frac{1}{2}, \pm 1\}$ 的那些。其证明方法是证明 $2\cos\theta$ 是一个代数整数（algebraic integer）。Davenport 的 *The Higher Arithmetic*（《高等算术》）是建立这类论证背后的数论基础的好起点。

详细解答

STEP III 2022，第 9 题

力学 · 碰撞与恢复系数

STEP III 2022，第 9 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

这里每一次碰撞其实都是同一种碰撞：一个运动中的质点撞向一个静止、质量为其 k 倍的质点。因此第 (i) 部分开头求出的因子 α 和 β 可以在第 (ii) 部分重用，因为在那里 B（质量 km）撞向 C（质量 k^2m）。之后，每个“在同一点相遇”的条件，都只不过是把同一时刻的两个表达式设为相等。在第 (ii) 部分我们会得到两条关于 k 和 e 的方程，而其中一条原来正正就是第 (i) 部分的条件。

解答

我们把 A、B、C 视为在光滑表面上沿同一直线运动的质点，而碰撞是瞬时的。所有速度均以 A 初始运动的方向为正方向。

第 (i) 部分

第一次碰撞。动量（momentum）守恒，而牛顿恢复定律（Newton's law of restitution）指出分离速率等于接近速率的 e 倍：

$$mu = mv_1 + kmv_2, \qquad v_2 - v_1 = eu.$$

把 $v_1 = v_2 - eu$ 代入第一条方程，得 $u = v_2 - eu + kv_2$，因此

$$v_2 = \frac{1+e}{k+1}u = \beta u, \qquad v_1 = \frac{1+e}{k+1}u - eu = \frac{1+e-ek-e}{k+1}u = \frac{1-ke}{k+1}u = \alpha u,$$

即为所求。

第二次碰撞。由第一次碰撞起计时。B 用时 $D/(\beta u)$ 走完到墙壁的距离 D。它以速率 $e\beta u$ 反弹，所以再用时 $D/(2e\beta u)$ 往碰撞点方向走回 $\frac{1}{2}D$。若第二次碰撞在那里发生，则发生

时刻为

$$T = \frac{D}{\beta u}\left(1 + \frac{1}{2e}\right) = \frac{(2e+1)D}{2e\beta u}.$$

该点距离墙壁 $\frac{1}{2}D$，因此它同时位于第一次碰撞点之后 $\frac{1}{2}D$ 处。A 以恒定速度 αu 运动，所以必须有 $\alpha > 0$ 及 $\alpha u T = \frac{1}{2}D$：

$$\alpha \frac{(2e+1)D}{2e\beta} = \frac{D}{2} \quad \Longleftrightarrow \quad \alpha(2e+1) = e\beta.$$

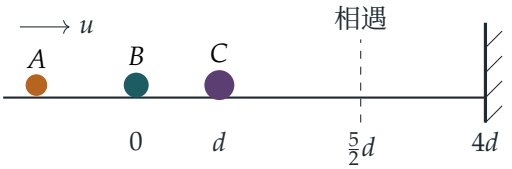
两边乘以 $k+1$，得 $(1-ke)(2e+1) = e(1+e)$，因此

$$ke(2e+1) = (2e+1) - e - e^2 = 1 + e - e^2 \quad \Longrightarrow \quad k = \frac{1+e-e^2}{e(2e+1)},$$

这就是所求的结果。这个值确实给出 $\alpha > 0$：它使 $ke = \frac{1+e-e^2}{2e+1}$，而由于当 $e > 0$ 时 $1 + e - e^2 < 1 + 2e$，此值小于 1。

第 (ii) 部分

把原点设在 B 的起点，使 C 从 $x = d$ 出发，而墙壁位于 $x = 4d$。所要求的相遇点距离墙壁 $\frac{3}{2}d$，即位于 $x = \frac{5}{2}d$。



设 A 撞上 B 的一刻为 $t = 0$。这次碰撞与第 (i) 部分的完全相同，所以碰撞后 A 的速度为 αu，B 的速度为 βu，其中 α 和 β 与之前相同。

想一想 B 撞上 C 时会发生甚么？我们需要重做碰撞的代数运算吗？

不需要。C 静止，而它的质量 $k^2 m$ 是 B 的质量 km 的 k 倍。所以这又是第 (i) 部分的碰撞，质量比同样是 k，e 相同，只是入射速率由 u 变成 βu。B 在时刻 $t_1 = d/(\beta u)$ 撞上 C。碰撞后 B 的速度为 $\alpha\beta u$，C 的速度为 $\beta^2 u$。

现在我们计算每个质点何时到达 $x = \frac{5}{2}d$，并假设在此之前没有其他碰撞发生（我们会在最后验证这一点）。

- A 从 $x = 0$ 以 αu 运动，所以在 $t = \dfrac{5d}{2\alpha u}$ 到达。
- B 在 t_1 到达 $x = d$，然后以速率 $\alpha\beta u$ 走 $\frac{3}{2}d$。它在 $t = \dfrac{d}{\beta u} + \dfrac{3d}{2\alpha\beta u}$ 到达。

▶ C 以 $\beta^2 u$ 走 $3d$ 到墙壁，再以 $e\beta^2 u$ 往回走 $\frac{3}{2}d$。它到达的时刻为

$$t = \frac{d}{\beta u} + \frac{3d}{\beta^2 u}\left(1 + \frac{1}{2e}\right) = \frac{d}{\beta u} + \frac{3(2e+1)d}{2e\beta^2 u}.$$

要发生同时碰撞，三个时刻必须相等。

B 与 C。消去公共项 $d/(\beta u)$，再乘以 $2\beta^2 u/(3d)$：

$$\frac{1}{\alpha}\beta = \frac{2e+1}{e} \quad\Longleftrightarrow\quad e\beta = \alpha(2e+1).$$

这正正是第 (i) 部分的条件，所以

$$k = \frac{1+e-e^2}{e(2e+1)}. \tag{1}$$

（这并非巧合。在两种情况下，都是一个质点离开碰撞后，以入射速率的 β 倍走 D 到墙壁，再折返 $\frac{1}{2}D$；与此同时，另一个质点以入射速率的 α 倍向前走 $\frac{1}{2}D$。这里 $D = 3d$，而入射速率为 βu。）

A 与 B。把 $\dfrac{5}{2\alpha} = \dfrac{1}{\beta} + \dfrac{3}{2\alpha\beta}$ 乘以 $2\alpha\beta$，得 $5\beta = 2\alpha + 3$。再乘以 $k+1$：

$$5(1+e) = 2(1-ke) + 3(k+1) \quad\Longrightarrow\quad 5e = k(3-2e) \quad\Longrightarrow\quad k = \frac{5e}{3-2e}. \tag{2}$$

? **想一想** 我们有两个 k 的表达式。怎样求出 e?

令 (1) 与 (2) 相等。由于 $0 < e < 1$，两个分母都是正数，所以可以交叉相乘：

$$\begin{aligned} 5e^2(2e+1) &= (3-2e)(1+e-e^2) \\ 10e^3 + 5e^2 &= 3 + e - 5e^2 + 2e^3 \\ 0 &= 8e^3 + 10e^2 - e - 3. \end{aligned}$$

试代 $e = \frac{1}{2}$，得 $1 + \frac{5}{2} - \frac{1}{2} - 3 = 0$，所以 $2e-1$ 是一个因式：

$$8e^3 + 10e^2 - e - 3 = (2e-1)(4e^2 + 7e + 3) = (2e-1)(4e+3)(e+1).$$

其余的根 $-\frac{3}{4}$ 和 -1 都是负数，所以满足 $0 < e < 1$ 的唯一根是 $e = \frac{1}{2}$，即为所求。于是 $k = \frac{5}{4}$。

验证情境是否自洽。当 $e = \frac{1}{2}$、$k = \frac{5}{4}$ 时，得 $\alpha = \frac{1}{6}$ 及 $\beta = \frac{2}{3}$。在 B 撞上 C 之前，A（速率 $\frac{1}{6}u$）比 B（速率 $\frac{2}{3}u$）慢，所以不能提早追上 B。之后，B（速率 $\frac{1}{9}u$）比 C（速率 $\frac{4}{9}u$）慢，所以 B 不能在 C 到达墙壁前追上 C。这时 B 比 A 慢，所以 A 逐渐追近 B，而它们只在我们算出的那一点相遇。因此不会有额外的碰撞，三个质点都在 $t = 15d/u$ 时到达 $x = \frac{5}{2}d$。

验算

- 在第 (i) 部分，$v_1 + kv_2 = u$（动量）及 $v_2 - v_1 = eu$（恢复系数）。若 $e = 1$ 且 $k = 1$，则 $\alpha = 0$ 及 $\beta = 1$：在完全弹性碰撞中，质量相等的两个质点互换速度，理应如此。
- 在第 (ii) 部分，取 $u = d = 1$ 及上述数值，三个到达时刻分别为：A 是 $\frac{5/2}{1/6} = 15$，B 是 $\frac{1}{2/3} + \frac{3/2}{1/9} = \frac{3}{2} + \frac{27}{2} = 15$，$C$ 是 $\frac{3}{2} + \frac{3}{4/9} + \frac{3/2}{2/9} = \frac{3}{2} + \frac{27}{4} + \frac{27}{4} = 15$。
- 公式 (1) 在 $e = \frac{1}{2}$ 时给出 $k = \frac{5/4}{1} = \frac{5}{4}$，与 (2) 一致：$\frac{5/2}{2} = \frac{5}{4}$。

常见错误

- 在第 (i) 部分，应该由碰撞点量度相遇点，却由墙壁量度。在这里两者恰巧相同，都是 $\frac{1}{2}D$；但在第 (ii) 部分，距离墙壁 $\frac{3}{2}d$ 的相遇点，距离 B 的起点是 $\frac{5}{2}d$。
- 从头重做 B–C 的碰撞而漏了一个因子。留意到质量比再次是 k，便可省去这些功夫。
- 忘记 C 从墙壁反弹后的速率是 $e\beta^2 u$，而不是 $\beta^2 u$。

STEP 之外

这道题之所以行得通，是因为问题具有*自相似性*（self-similar）：把质量乘以 k、速率乘以 β，就把一次碰撞变成下一次碰撞，所以同一套代数运算会重复出现。同样的想法也出现在质量按几何级数递减的一串球之中——这样的球串能非常有效地把能量传递下去；它亦是著名的“用碰撞方块数出 π”问题背后的道理，而该问题在速度空间中、以动量和能量作为约束来分析最为简洁。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）有系统地处理碰撞问题，包括在质心参考系（centre-of-mass frame）中进行分析，这使得像 α 和 β 这样的公式一目了然。

详细解答

STEP III 2022，第 10 题

力学 • 弹簧的平衡

STEP III 2022，第 10 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

由于 P 处的角是直角，两条弹簧的拉力互相垂直。因此沿每条弹簧方向分解（resolve），便能各自单独求出每个张力（tension），无须解联立方程。至于能量方面，弹性势能项与重力势能项合并得出奇地简洁，使 p 成为 θ 的一个简单函数（function）。第 (ii) 部分真正的重点，其实在于 θ 所受的物理限制：两条弹簧都必须处于伸长状态。

解答

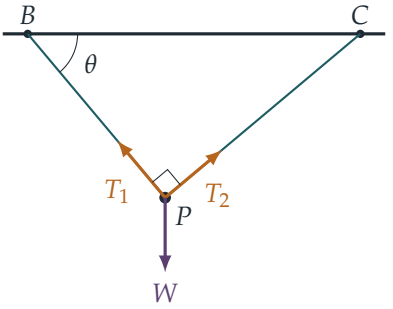
我们把 P 视为质点，并假设弹簧是轻的且遵守胡克定律（Hooke's law）。于是，一条弹性模量（modulus）为 λ、自然长度（natural length）为 a 的弹簧被拉长至长度 l 时，其张力为 $\lambda(l-a)/a$，所储存的能量为 $\lambda(l-a)^2/(2a)$。

第 (i) 部分

由于 $BC = 2a$ 且角 $BPC = 90°$，三角形 BPC 的斜边为 BC，所以

$$PB = 2a\cos\theta, \qquad PC = 2a\sin\theta, \qquad \angle PCB = 90° - \theta.$$

设 T_1 和 T_2 分别为 PB 和 PC 中的张力。



直线 PB 与水平方向成角 θ，所以它与铅垂方向成角 $90^\circ-\theta$；同理，PC 与铅垂方向成角 θ。由于 $PB\perp PC$，沿 PB 方向分解时只涉及 T_1 和重量，而重量沿 BP 方向（背离 B）的分量为 $W\cos(90^\circ-\theta)=W\sin\theta$。所以

$$T_1=W\sin\theta,\qquad \text{同理，沿}PC,\qquad T_2=W\cos\theta.$$

由胡克定律，

$$T_1=\frac{s_1W(2a\cos\theta-a)}{a}=s_1W(2\cos\theta-1),\qquad T_2=s_2W(2\sin\theta-1).$$

令两者相等，得

$$s_1=\frac{\sin\theta}{2\cos\theta-1},\qquad s_2=\frac{\cos\theta}{2\sin\theta-1},$$

这就是所求的 s_1，并一并得出 s_2。

第 (ii) 部分

重力势能。P 位于 BC 下方，距离等于三角形从 P 出发的高，即 $PB\sin\theta=2a\sin\theta\cos\theta=a\sin2\theta$。以 BC 为零势能水平，重力势能（gravitational potential energy）为 $-Wa\sin2\theta$。

弹性势能。PB 的伸长量（extension）为 $a(2\cos\theta-1)$，因此利用第 (i) 部分，

$$\frac{s_1W\,a^2(2\cos\theta-1)^2}{2a}=\tfrac12Wa\,s_1(2\cos\theta-1)^2=\tfrac12Wa\sin\theta\,(2\cos\theta-1),$$

同理，PC 中的能量为 $\frac12Wa\cos\theta\,(2\sin\theta-1)$。

总和。相加得

$$\begin{aligned}-paW&=-Wa\sin2\theta+\tfrac12Wa\big(2\sin\theta\cos\theta-\sin\theta+2\sin\theta\cos\theta-\cos\theta\big)\\&=-Wa\sin2\theta+Wa\sin2\theta-\tfrac12Wa(\sin\theta+\cos\theta),\end{aligned}$$

所以

$$p=\tfrac12(\sin\theta+\cos\theta)=\tfrac{\sqrt2}{2}\sin\big(\theta+\tfrac{\pi}{4}\big).$$

想一想 θ 可以是 0 与 $\frac{\pi}{2}$ 之间的任何值吗?

不可以。模量 s_1W 和 s_2W 均为正数，而张力 $W\sin\theta$ 和 $W\cos\theta$ 亦为正数，所以两条弹簧都必须处于伸长状态：$2\cos\theta - 1 > 0$ 且 $2\sin\theta - 1 > 0$。（若其中一项为零，该弹簧便不会产生任何力，却仍须提供非零的张力。）因此 $\cos\theta > \frac{1}{2}$ 且 $\sin\theta > \frac{1}{2}$，即

$$\tfrac{\pi}{6} < \theta < \tfrac{\pi}{3}, \qquad \text{故} \qquad \tfrac{5\pi}{12} < \theta + \tfrac{\pi}{4} < \tfrac{7\pi}{12}.$$

这个区间关于 $\frac{\pi}{2}$ 对称。在此区间上，$\sin(\theta + \frac{\pi}{4})$ 最大为 1（在 $\theta = \frac{\pi}{4}$ 时取得），并严格大于它在端点的值 $\sin\frac{5\pi}{12} = \sin\frac{7\pi}{12}$。而

$$\sin\tfrac{5\pi}{12} = \sin\left(\tfrac{\pi}{4} + \tfrac{\pi}{6}\right) = \tfrac{\sqrt{2}}{2}\cdot\tfrac{\sqrt{3}}{2} + \tfrac{\sqrt{2}}{2}\cdot\tfrac{1}{2} = \frac{\sqrt{2}(\sqrt{3}+1)}{4},$$

所以

$$\tfrac{1}{2}\sqrt{2} \;\geq\; p \;>\; \tfrac{\sqrt{2}}{2}\cdot\frac{\sqrt{2}(\sqrt{3}+1)}{4} = \tfrac{1}{4}(1+\sqrt{3}),$$

得证。

一位有效数字。我们需要 $0.65 \leq p < 0.75$，并以平方的方法来证明根式（surd）的估值。由于 $1.6^2 = 2.56 < 3$，得 $\sqrt{3} > 1.6$，所以 $p > \frac{1}{4}(2.6) = 0.65$。由于 $1.5^2 = 2.25 > 2$，得 $\sqrt{2} < 1.5$，所以 $p \leq \frac{1}{2}\sqrt{2} < 0.75$。因此，准确至一位有效数字（significant figure），$p = 0.7$。

验算

- 当 $\theta = \frac{\pi}{4}$ 时，整个装置是对称的；事实上 $s_1 = s_2 = \frac{1}{\sqrt{2}-1} = \sqrt{2}+1$，且 $p = \frac{1}{2}\sqrt{2} \approx 0.7071$。
- 当 $\theta \to \frac{\pi}{6}^+$ 时，$s_1 \to \infty$：弹簧 PB 处于自然长度，必须无限坚硬才能承受张力。此时 $p \to \frac{1}{4}(1+\sqrt{3}) \approx 0.6830$，即下界。
- 我们用电脑对 $(\frac{\pi}{6}, \frac{\pi}{3})$ 中许多随机的 θ 作验证，确认所求得的位置是总势能的驻点（stationary point）（因此确实是平衡位置（equilibrium）），而且能量等于 $-\frac{1}{2}Wa(\sin\theta + \cos\theta)$。

常见错误

- 把重力势能写成正号：P 位于零势能水平的*下方*。
- 在整个 $(0, \frac{\pi}{2})$ 上求 $\frac{1}{2}(\sin\theta + \cos\theta)$ 的范围，因而得出错误的下界 $\frac{1}{2}$；弹簧必须处于张紧状态。
- 直接引用 $\sqrt{3} \approx 1.732$，却没有说明误差不会影响舍入结果。

STEP 之外

第 (ii) 部分中的抵消并非巧合：在平衡位置，总势能取驻值；而对于遵守胡克定律的弹簧，弹性势能恰好是（张力 $\times$ 伸长量）的一半，这使它与克服重力所做的功紧密相连。把平衡位置视为势能的驻点来研究，并通过二阶导数（second derivative）判断其稳定性（stability），是大学一年级力学的核心主题之一。Kleppner 与 Kolenkow 合著的 *An Introduction to Mechanics*（《力学导论》）对这种能量观点有细致的阐述。

详细解答

STEP III 2022，第 11 题

概率与统计 · 二项分布的平均偏差

STEP III 2022，第 11 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

关键思路

硬币公平，所以 X 的分布（distribution）关于其均值（mean）对称（symmetric），即低于均值的概率与高于均值的概率一一对应。利用这种对称性，可以把绝对值（absolute value）化为单侧的和。接着，(i)(c) 中的恒等式（identity）把 $\sum r\binom{2n}{r}$ 化为一个较小的二项分布（binomial distribution）的尾部和（tail sum），而这个尾部和同样可以用对称性求出。

解答

全题中 $X \sim \mathrm{B}(N, \frac{1}{2})$，所以 $\mathrm{P}(X = r) = \binom{N}{r}2^{-N}$，且 $\mu = \frac{1}{2}N$。由于 $\binom{N}{r} = \binom{N}{N-r}$，我们有 $\mathrm{P}(X = r) = \mathrm{P}(X = N - r)$：分布关于 $\frac{1}{2}N$ 对称。

第 (i)(a) 部分

当 $N = 2n$ 时，映射 $r \mapsto 2n - r$ 把 $\{0, \ldots, n-1\}$ 映到 $\{n+1, \ldots, 2n\}$，所以由对称性得

$$\mathrm{P}(X \leq n-1) = \mathrm{P}(X \geq n+1).$$

三个事件 $X \leq n-1$、$X = n$、$X \geq n+1$ 涵盖了所有结果，所以 $2\,\mathrm{P}(X \leq n-1) + \mathrm{P}(X = n) = 1$，由此得

$$\mathrm{P}(X \leq n-1) = \tfrac{1}{2}\big(1 - \mathrm{P}(X = n)\big).$$

第 (i)(b) 部分

此时 $\mu = n$，所以

$$\delta = \sum_{r=0}^{2n} |r-n| \binom{2n}{r} \frac{1}{2^{2n}}.$$

$r = n$ 的项为零。在 $r \geq n+1$ 的各项中代入 $r = 2n - s$，得 $|r-n| = n - s$ 及 $\binom{2n}{r} = \binom{2n}{s}$，其中 $0 \leq s \leq n-1$。因此，高于均值的各项恰好重复了低于均值的各项：

$$\delta = 2\sum_{r=0}^{n-1} (n-r)\binom{2n}{r}\frac{1}{2^{2n}} = \sum_{r=0}^{n-1}(n-r)\binom{2n}{r}\frac{1}{2^{2n-1}},$$

得证。

第 (i)(c) 部分

对于 $1 \leq r \leq 2n$，

$$r\binom{2n}{r} = \frac{r\,(2n)!}{r!\,(2n-r)!} = 2n \cdot \frac{(2n-1)!}{(r-1)!\,(2n-r)!} = 2n\binom{2n-1}{r-1}.$$

现在把 (b) 中的和拆成 $2^{2n-1}\delta = A - B$，其中

$$A = n\sum_{r=0}^{n-1}\binom{2n}{r}, \qquad B = \sum_{r=1}^{n-1} r\binom{2n}{r}$$

（B 中 $r = 0$ 的项为零，所以略去）。

和 A。由 (a)，$\sum_{r=0}^{n-1}\binom{2n}{r} = 2^{2n}\,\mathrm{P}(X \leq n-1) = 2^{2n-1} - \frac{1}{2}\binom{2n}{n}$，所以

$$A = n\,2^{2n-1} - \tfrac{n}{2}\binom{2n}{n}.$$

? **想一想** 当 s 由 0 到 $n-2$ 时，怎样求 $\binom{2n-1}{s}$ 的总和?

和 B。由上述恒等式，令 $s = r - 1$，

$$B = 2n\sum_{s=0}^{n-2}\binom{2n-1}{s}.$$

这一行 $\binom{2n-1}{s}$ $(0 \leq s \leq 2n-1)$ 共有偶数个项，总和为 2^{2n-1}。它在 $s \mapsto 2n-1-s$ 下对称，而这个映射把 $\{0,\ldots,n-1\}$ 与 $\{n,\ldots,2n-1\}$ 互换，所以前一半的和是 2^{2n-2}。我们的和比这一半少了最后一项，所以

$$\sum_{s=0}^{n-2}\binom{2n-1}{s} = 2^{2n-2} - \binom{2n-1}{n-1} = 2^{2n-2} - \tfrac{1}{2}\binom{2n}{n},$$

这里用了 $\binom{2n}{n} = \binom{2n-1}{n-1} + \binom{2n-1}{n} = 2\binom{2n-1}{n-1}$。因此

$$B = n\,2^{2n-1} - n\binom{2n}{n}.$$

合并起来。

$$2^{2n-1}\delta = A - B = \left(n - \frac{n}{2}\right)\binom{2n}{n} = \frac{n}{2}\binom{2n}{n}, \qquad \text{所以} \qquad \delta = \frac{n}{2^{2n}}\binom{2n}{n}.$$

第 (ii) 部分

现在 $N = 2n+1$，$\mu = n + \frac{1}{2}$，这并不是 X 可取的值。对称映射 $r \mapsto 2n+1-r$ 把 $\{0,\dots,n\}$ 与 $\{n+1,\dots,2n+1\}$ 互换，并且保持 $|r-\mu|$ 和 $\binom{2n+1}{r}$ 不变。所以，与 (i)(b) 一样，

$$\delta = 2\sum_{r=0}^{n}\left(n + \tfrac{1}{2} - r\right)\binom{2n+1}{r}\frac{1}{2^{2n+1}} = \frac{1}{2^{2n+1}}\sum_{r=0}^{n}(2n+1-2r)\binom{2n+1}{r}.$$

- 由同样的对称性，这一行 $\binom{2n+1}{r}$ 前一半的和是 2^{2n+1} 的一半：$\sum_{r=0}^{n}\binom{2n+1}{r} = 2^{2n}$。
- 由恒等式 $r\binom{2n+1}{r} = (2n+1)\binom{2n}{r-1}$（证明与 (i)(c) 完全相同），再用 (i)(a)，

$$\sum_{r=1}^{n} r\binom{2n+1}{r} = (2n+1)\sum_{s=0}^{n-1}\binom{2n}{s} = (2n+1)\left(2^{2n-1} - \tfrac{1}{2}\binom{2n}{n}\right).$$

因此

$$\sum_{r=0}^{n}(2n+1-2r)\binom{2n+1}{r} = (2n+1)2^{2n} - (2n+1)\left(2^{2n} - \binom{2n}{n}\right) = (2n+1)\binom{2n}{n},$$

而当 $N = 2n+1$ 时，

$$\boxed{\delta = \frac{2n+1}{2^{2n+1}}\binom{2n}{n}}$$

验算

- $N = 3$（$n = 1$）：$|X - \frac{3}{2}|$ 以概率 $\frac{1}{4}$ 取值 $\frac{3}{2}$，以概率 $\frac{3}{4}$ 取值 $\frac{1}{2}$，所以 $\delta = \frac{3}{8} + \frac{3}{8} = \frac{3}{4}$。公式给出 $\frac{3\cdot 2}{8} = \frac{3}{4}$。当 $n = 0$（$N = 1$）时，公式给出 $\frac{1}{2}$，同样正确。
- $N = 4$（$n = 2$）：(i)(c) 给出 $\frac{2\cdot 6}{16} = \frac{3}{4}$，与 $N = 3$ 的结果相同。事实上，$N = 2n+1$ 时的 δ 总是等于 $N = 2n+2$ 时的 δ：由于 $\binom{2n+2}{n+1} = \frac{2(2n+1)}{n+1}\binom{2n}{n}$，把偶数情况公式中的 n 换成 $n+1$，便化简为奇数情况的公式。
- 经电脑验算，对所有 $N \le 60$，两条公式都与 δ 的定义完全吻合；例如 $N = 10$ 给出 $\delta = \frac{315}{256} \approx 1.230$。

常见错误

- 忘记对称性带来的因子 2；(i)(b) 中出现 2^{2n-1} 而不是 2^{2n}，正是因为这个因子。
- 作平移 $s = r - 1$ 后弄错 s 的范围：和止于 $n-2$，比半行少一项，而正是缺少的那一项 $\binom{2n-1}{n-1}$ 令答案不为零。
- 在 (ii) 中把 μ 当作整数：其实没有任何一项为零，而且这一行的两半各有 $n+1$ 项。

STEP 之外

斯特林公式（Stirling's formula）给出 $\binom{2n}{n} \approx 4^n/\sqrt{\pi n}$，所以当 N 很大时，$\delta \approx \sqrt{N/(2\pi)}$。由于 $\sigma = \frac{1}{2}\sqrt{N}$，这就是 $\delta \approx \sigma\sqrt{2/\pi}$，恰好是正态分布（normal distribution）的平均绝对偏差（mean absolute deviation），正如中心极限定理（central limit theorem）所预测。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）既讲解了二项分布的正态近似，也讲解了其背后的中心极限定理。

详细解答

STEP III 2022，第 12 题

概率与统计 • 圆的随机弦

STEP III 2022，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2022 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

本题在 *Beyond STEP* 第 12.3 节“随机”是甚么意思？中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch12/bertrand.html

关键思路

本题中所有的弦都源自同一个事实：在半径为 a 的圆中，与圆心的垂直距离为 d 的弦（chord），长度为 $2\sqrt{a^2-d^2}$。因此要做的，就是用随机量把 d 表示出来，然后求平均。在第 (ii) 部分，代换（substitution）$u = a^2 - r^2\sin^2 t$ 是很自然的选择，因为 R 的概率密度（probability density）恰好提供了这个代换所需要的因子 r。

解答

第 (i) 部分

设 Θ 为角 AOB，则 Θ 在 $[0, 2\pi]$ 上均匀分布（uniformly distributed），密度为 $\frac{1}{2\pi}$。三角形 AOB 是顶角为 Θ 的等腰三角形。从 O 向 AB 作垂线，这条垂线同时平分顶角和弦，所以

$$AB = 2a\sin\tfrac{1}{2}\Theta.$$

这对所有 $\Theta \in [0, 2\pi]$ 都成立，因为在这个范围内 $\sin\frac{1}{2}\Theta \geq 0$；当 $\Theta > \pi$ 时，我们量度的是优角（reflex angle），但弦是同一条。因此

$$\mathrm{E}(AB) = \frac{1}{2\pi}\int_0^{2\pi} 2a\sin\tfrac{1}{2}\theta\,\mathrm{d}\theta = \frac{a}{\pi}\Big[-2\cos\tfrac{1}{2}\theta\Big]_0^{2\pi} = \frac{a}{\pi}(2+2) = \frac{4a}{\pi}.$$

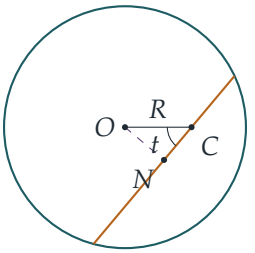
第 (ii) 部分

R 的密度。对于 $0 \le r \le a$，事件 $R \le r$ 表示 C 落在以 O 为圆心、半径为 r 的圆盘内。由于概率与面积成正比，

$$\mathrm{P}(R \le r) = \frac{\pi r^2}{\pi a^2} = \frac{r^2}{a^2}, \qquad \text{所以} \qquad f_R(r) = \frac{2r}{a^2} \quad (0 \le r \le a),$$

而在其他情况下 $f_R(r) = 0$。

弦的长度。设 N 为从 O 到弦的垂足。在直角三角形 ONC 中，斜边是 $OC = R$，而在 C 处的角是 t，所以 $ON = R\sin t$。



从圆心所作的垂线平分弦，N 就是弦的中点，所以在由 O、N 和弦的一个端点组成的三角形中应用毕氏定理（Pythagoras' theorem），得

$$\text{弦长} = 2\sqrt{a^2 - R^2\sin^2 t}.$$

由于 $R < a$，根号内的式子是正数。注意这里计算的是整条弦，包括 C 两侧的部分。

期望长度。固定 t，而 C 为随机，

$$\mathrm{L}(t) = \int_0^a 2\sqrt{a^2 - r^2\sin^2 t}\,\frac{2r}{a^2}\,\mathrm{d}r.$$

取 $0 < t < \frac{1}{2}\pi$，使得 $\sin t \ne 0$。代入 $u = a^2 - r^2\sin^2 t$，则 $\mathrm{d}u = -2r\sin^2 t\,\mathrm{d}r$；当 $r = 0$ 时 $u = a^2$，当 $r = a$ 时 $u = a^2(1 - \sin^2 t) = a^2\cos^2 t$。于是

$$\mathrm{L}(t) = \frac{2}{a^2\sin^2 t}\int_{a^2\cos^2 t}^{a^2} u^{1/2}\,\mathrm{d}u = \frac{2}{a^2\sin^2 t}\cdot\frac{2}{3}\left(a^3 - a^3\cos^3 t\right) = \frac{4a(1-\cos^3 t)}{3\sin^2 t},$$

其中用到 $(a^2\cos^2 t)^{3/2} = a^3\cos^3 t$，这是成立的，因为对于锐角，$\cos t \ge 0$。这就是所求的公式。

> **想一想** 怎样才能把分母 $\sin^2 t$ 化成可以与 $1 - \cos^3 t$ 约去的形式?

设 $c = \cos t$。那么 $\sin^2 t = 1 - c^2 = (1-c)(1+c)$，而 $1 - c^3$ 是立方差（difference of cubes），即 $(1-c)(1+c+c^2)$。由于 $c \ne 1$，可以约去 $1 - c$：

$$\mathrm{L}(t) = \frac{4a}{3}\cdot\frac{1+c+c^2}{1+c} = \frac{4a}{3}\left(c + \frac{1}{1+c}\right),$$

因为 $c(1+c)+1 = 1+c+c^2$。最后，$1+\cos t = 2\cos^2\frac{1}{2}t$，所以 $\dfrac{1}{1+\cos t} = \frac{1}{2}\sec^2\frac{1}{2}t$，从而

$$\mathrm{L}(t) = \frac{4a}{3}\Big(\cos t + \tfrac{1}{2}\sec^2\big(\tfrac{1}{2}t\big)\Big),$$

即为所求。这个形式在 $t=0$ 时没有问题：它给出 $\frac{4a}{3}\cdot\frac{3}{2} = 2a$，这是正确的，因为沿 OC 方向经过 C 的弦就是直径。

第 (iii) 部分

T 在 $[0, \frac{1}{2}\pi]$ 上的密度为 $\frac{2}{\pi}$，所以使用 L 的第二个形式（它很容易积分，因为 $\frac{\mathrm{d}}{\mathrm{d}t}\tan\frac{1}{2}t = \frac{1}{2}\sec^2\frac{1}{2}t$），

$$\begin{aligned}\mathrm{E}\big(\mathrm{L}(T)\big) &= \frac{2}{\pi}\cdot\frac{4a}{3}\int_0^{\pi/2}\Big(\cos t + \tfrac{1}{2}\sec^2\tfrac{1}{2}t\Big)\,\mathrm{d}t = \frac{8a}{3\pi}\Big[\sin t + \tan\tfrac{1}{2}t\Big]_0^{\pi/2}\\ &= \frac{8a}{3\pi}(1+1) = \frac{16a}{3\pi}.\end{aligned}$$

验算

- 当 $t = \frac{1}{2}\pi$ 时，弦垂直于 OC，而 $\mathrm{L} = \frac{4a}{3}(0+\frac{1}{2}\cdot 2) = \frac{4a}{3}$。直接计算，$\int_0^a 2\sqrt{a^2-r^2}\,\frac{2r}{a^2}\,\mathrm{d}r = \frac{4a}{3}$，结果相同。
- 当 t 由 0 增至 $\frac{1}{2}\pi$ 时，$\mathrm{L}(t)$ 应由 $2a$ 下降至 $\frac{4}{3}a$；而 (iii) 的答案 $\frac{16}{3\pi}a \approx 1.698a$ 正好介于两者之间，平均值理应如此。
- 取 $a = 1.7$ 和 $t = 0.7$，将弦长乘以 R 的密度作数值积分，得 3.01799，与 $\mathrm{L}(0.7)$ 的两条公式结果一致。(i) 的答案 $\frac{4}{\pi}a \approx 1.273a$ 亦已用数值积分核实。

常见错误

- 以为 R 在 $[0,a]$ 上均匀分布。点是按*面积*均匀分布的，而离圆心越远，面积越多，因此密度是 $2r/a^2$。
- 只计算弦在 C 一侧的部分，或者写成 $ON = R\cos t$（那是沿弦方向的距离 CN）。
- 在第 (iii) 部分积分 $\mathrm{L}(t)$ 的第一个形式，而不是第二个形式；题目特意设计了第二个形式，让你一行就能积分。

STEP 之外

这是几何概率（geometric probability）：“随机选取”只有在确定了分布之后才有意义。答案 $\frac{4a}{\pi}$ 和 $\frac{16a}{3\pi}$ 之所以不同，是因为它们描述的是选取弦的不同方式，这正是贝特朗悖论

（Bertrand's paradox）的要点。积分几何（integral geometry）利用在旋转和平移下保持不变的直线测度，把这类问题系统化。关于区域上的连续分布以及本题这类期望值的计算，可参阅 Blitzstein 和 Hwang 的 *Introduction to Probability*（《概率导论》）。

Beyond STEP

STEP 是大多数中学生一生中应考过最艰深的数学考试，也是最出色的考试之一：题目简短、陌生而深刻，当中最精彩的，其实是乔装了的大学数学小品。

本书精选最有趣的 STEP 试题，由 1987 年首次开考之前的样本试卷（specimen papers），到 2025 年的试卷，并以每一道题目作为一扇门。每道题目都会仔细解答，沿途提出优秀学生会问的问题，然后带你穿过这扇门：这个想法从何而来，到了大学会演变成甚么，以及接下来该读些甚么。

本书为正在准备 STEP、有天分又有志投身数学的学生而写。他们想要的不止是模拟试卷：他们想开始像数学系本科生那样思考，并一探前方有甚么正等着自己。

- 共 13 章，涵盖纯数学（pure mathematics）、力学（mechanics）、概率（probability）与统计（statistics）
- 136 道 STEP 历届试题，每题均附引导式解答
- 每道题目背后的大学数学
- 附答案的延伸练习，以及延伸阅读

2022 年详细解答. 全书各章及
所有详细解答均可免费阅读、下载及分享，网址为
beyondstep.keithmcnulty.org

文字内容已贡献予公有领域（CC0 1.0）。STEP 试题 © Cambridge University Press & Assessment。

