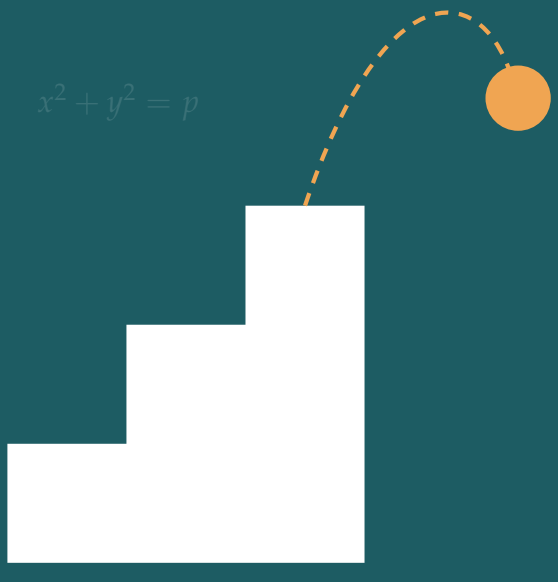
STEP 备试 · 大学数学

Beyond STEP

STEP II 及 III

2024 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

2024 年详细解答

Keith McNulty

Beyond STEP：以历届 STEP 试题为门，走进大学数学

2024 年详细解答

2024 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 连续整数之和 纯数学 . . . 3

第 2 题 级数与积分 纯数学 . . . 7

第 3 题 半角正切代换 纯数学 . . . 11

第 4 题 向量与角平分线 纯数学 . . . 15

第 5 题 值域与丢番图方程 纯数学 . . . 19

第 6 题 二项式级数的系数 纯数学 . . . 23

第 7 题 曲线描绘 纯数学 . . . 27

第 8 题 算术几何平均 纯数学 . . . 32

第 9 题 抛体运动 力学 . . . 36

第 10 题 加速楔上的摩擦力 力学 . . . 40

第 11 题 分组检测的期望值 概率与统计 . . . 44

第 12 题 离散顺序统计量 概率与统计 . . . 48

STEP III

第 1 题 裂项相消级数 纯数学 . . . 52

第 2 题 不等式与绝对值 纯数学 . . . 56

第 3 题 单调函数与导数 纯数学 . . . 61

第 4 题 抛物线的切线 纯数学 . . . 65

第 5 题 矩阵微分方程 纯数学 . . . 69

第 6 题 线性微分方程 纯数学 . . . 73

第 7 题 e^2 的无理性 纯数学 . . . 77

第 8 题 相交圆锥曲线 纯数学 . . . 81

第 9 题 斜碰撞 力学 . . . 85

第 10 题 倾倒与滑动 力学 . . . 90

第 11 题 项分布的期望值 概率与统计 95
第 12 题 何概率 概率与统计 98

详细解答

STEP II 2024，第 1 题

纯数学 · 连续整数之和

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

等式两边都是算术级数（arithmetic series），所以各自等于（项数）×（平均项）。令两者相等，便得到 n、k 与 c 之间的一条二次关系式，而题目余下部分就是针对这条关系式的数论（number theory）问题。固定 k 后，很容易解出 c，再由奇偶性（parity）或余数判断 c 何时是正整数。若改为固定 $c=1$，便得到一条佩尔（Pell）型方程，而 (iv)(b) 中的映射正是由旧解产生新解的标准方法。

解答

第 (i) 部分

由 c 开始往上的 $n+k$ 个整数是 $c, c+1, \ldots, c+n+k-1$。它们的和是

$$S_1 = (n+k)c + \frac{(n+k)(n+k-1)}{2}.$$

接下来的 n 个整数是 $c+n+k, \ldots, c+2n+k-1$，其和为

$$S_2 = n(c+n+k) + \frac{n(n-1)}{2}.$$

把两个和都乘以二并展开，

$$2S_1 = 2nc + 2kc + n^2 + 2nk + k^2 - n - k,$$
$$2S_2 = 2nc + 2n^2 + 2nk + n^2 - n.$$

相减得 $2S_1 - 2S_2 = 2kc + k^2 - k - 2n^2$。因此

$$S_1 = S_2 \iff 2kc + k^2 - k = 2n^2 \iff 2n^2 + k = 2ck + k^2.$$

每一步都是可逆的（我们只是展开、乘以二和移项），所以两个和相等当且仅当 $2n^2+k=2ck+k^2$，得证。

第 (ii) 部分

(a) 当 $k=1$ 时，条件为 $2n^2+1=2c+1$，所以 $c=n^2$。对每个 n 这都是正整数，因此每一个正整数 n 都可以，且 $c=n^2$。（例如 $n=2$：$4+5+6=7+8$。）

(b) 当 $k=2$ 时，条件为 $2n^2+2=4c+4$，即

$$n^2=2c+1, \qquad c=\frac{n^2-1}{2}.$$

要 c 是整数，n^2 必须是奇数，所以 n 必须是奇数。要 c 是正数，需要 $n^2>1$，这就排除了 $n=1$。反过来，每个奇数 $n\geq 3$ 都给出正整数 c。所以可能的值是

$$n\in\{3,5,7,\dots\}, \qquad c=\tfrac{1}{2}(n^2-1).$$

$n=3$、$c=4$ 的情况就是等式 $4+5+6+7+8=9+10+11$。

第 (iii) 部分

当 $k=4$ 时，条件为 $2n^2+4=8c+16$，即

$$n^2=4c+6=4(c+1)+2.$$

所以 n^2 除以 4 的余数将会是 2。但平方数除以 4 的余数只能是 0 或 1：若 $n=2m$，则 $n^2=4m^2$；若 $n=2m+1$，则 $n^2=4(m^2+m)+1$。所以没有解。

第 (iv) 部分

当 $c=1$ 时，条件变成 $2n^2+k=2k+k^2$，即

$$2n^2=k(k+1). \qquad (*)$$

换句话说 $n^2=\frac{1}{2}k(k+1)$：我们要找同时是完全平方数（perfect square）的三角形数（triangular number）。

(a) 尝试较小的 k：$k=1$ 给出 $2n^2=2$，所以 $n=1$；接着 $k=2,\dots,7$ 给出 $k(k+1)/2=3,6,10,15,21,28$，都不是平方数；而 $k=8$ 给出 $2n^2=72$，所以 $n=6$。因此两个可能情况是

$$k=1,\ n=1 \qquad \text{和} \qquad k=8,\ n=6.$$

（验证第二个：$1+2+\cdots+14=105$ 及 $15+16+\cdots+20=105$。）

(b) 设 $2N^2 = K^2 + K$。令 $N' = 3N + 2K + 1$ 及 $K' = 4N + 3K + 1$。则

$$\begin{aligned} 2N'^2 &= 2(9N^2 + 4K^2 + 1 + 12NK + 6N + 4K) \\ &= 18N^2 + 8K^2 + 24NK + 12N + 8K + 2, \\ K'^2 + K' &= (16N^2 + 9K^2 + 1 + 24NK + 8N + 6K) + (4N + 3K + 1) \\ &= 16N^2 + 9K^2 + 24NK + 12N + 9K + 2. \end{aligned}$$

相减，

$$2N'^2 - (K'^2 + K') = 2N^2 - K^2 - K = 0,$$

这里用了假设。所以 $2N'^2 = K'(K' + 1)$。由于 N 和 K 是正整数，N' 和 K' 也是正整数；由 $(*)$ 及第 (i) 部分，N' 是 n 的一个可能值，对应的 $k = K'$。

想一想 这些古怪的公式从何而来?

把 $(*)$ 乘以 4 再加 1：它变成 $(2k+1)^2 - 8n^2 = 1$。令 $x = 2k + 1$ 及 $y = n$，这就是佩尔方程 $x^2 - 8y^2 = 1$，其所有解可以由 $x + y\sqrt{8}$ 乘以最小解 $3 + \sqrt{8}$ 逐一产生。写出 $(x + y\sqrt{8})(3 + \sqrt{8}) = (3x + 8y) + (x + 3y)\sqrt{8}$，再转回 k 和 n，便恰好得到 $N' = 3N + 2K + 1$ 及 $2K' + 1 = 3(2K + 1) + 8N$，即 $K' = 4N + 3K + 1$。

(c) 把映射应用于 $(N, K) = (1, 1)$，只会得到 $N' = 6$、$K' = 8$，即 (a) 中的第二个解。把它应用于 $(N, K) = (6, 8)$：

$$N' = 18 + 16 + 1 = 35, \qquad K' = 24 + 24 + 1 = 49,$$

再把它应用于 $(35, 49)$：

$$N' = 105 + 98 + 1 = 204, \qquad K' = 140 + 147 + 1 = 288.$$

所以另外两个解是

$$k = 49,\ n = 35 \qquad \text{和} \qquad k = 288,\ n = 204.$$

确实 $2 \cdot 35^2 = 2450 = 49 \cdot 50$ 及 $2 \cdot 204^2 = 83232 = 288 \cdot 289$。

验算

- 开首的例子是 $n = 3$、$k = 2$、$c = 4$：$2 \cdot 9 + 2 = 20 = 2 \cdot 4 \cdot 2 + 4$，所以它满足 (i) 的关系式，并符合 (ii)(b)，其中 $c = \frac{1}{2}(9 - 1)$。
- 用电脑搜索所有 $n, k < 40$ 及 $c < 200$，证实两个和相等恰好当 $2n^2 + k = 2ck + k^2$；另以 $c = 1$ 及 $n, k < 400$ 搜索，只找到 $(k, n) = (1, 1), (8, 6), (49, 35), (288, 204)$，与 (iv) 吻合。

- 在 (iv)(b) 中，注意由 (N, K) 到 (N', K') 的一步对任何 N、K 都保持 $2N^2 - K^2 - K$ 这个量不变，这是用小数字快速检查代数运算的方法，例如 $N = K = 0$：此时 $N' = K' = 1$，两个值都是 0。

常见错误

- 在 (ii)(b) 中把 $n = 1$ 也包括在内：它给出 $c = 0$，并不是正整数。
- 在 (i) 中数错第二组由哪里开始：它由 $c + n + k$ 开始，而不是 $c + n + k + 1$。
- 在 (iv)(c) 中再次给出 $k = 8$、$n = 6$：这是映射的第一个输出，但在 (a) 已经找到。

STEP 之外

方程 $x^2 - 8y^2 = 1$ 是一条佩尔方程；对任何非平方数 d，$x^2 - dy^2 = 1$ 的所有解都来自单一基本解（fundamental solution）的幂，通过 $x + y\sqrt{d}$ 这类数的乘法得出。这是初次窥见 $\mathbb{Z}[\sqrt{d}]$ 这类环（ring）中的单位（unit），而这是代数数论（algebraic number theory）的核心课题。Davenport 的 *The Higher Arithmetic*（《高等算术》）通过连分数（continued fraction）处理佩尔方程。

详细解答

STEP II 2024，第 2 题

纯数学 · 级数与积分

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

第 (i) 部分把每个积分 $\int_0^1 x^m/(8+x^3)\,\mathrm{d}x$ 都化成一个级数（series），所以凡是由相同元素组成的级数，都是这些积分的组合。要求出第 (ii) 或 (iii) 部分的级数之和，可以把括号拆成形如 $1/(3k+m+1)$ 的项，再把对应的被积函数（integrand）合并成一个分式，然后积分。因式分解（factorisation）$8+x^3=(x+2)(x^2-2x+4)$ 正是令所得积分变得容易处理的关键。

解答

全题中记

$$I_m=\int_0^1 \frac{x^m}{8+x^3}\,\mathrm{d}x.$$

第 (i) 部分

先提出因子 8，再对 $|u|<1$ 利用 $(1+u)^{-1}=\sum_{k\ge 0}(-1)^k u^k$，并取 $u=x^3/8$，得

$$(8+x^3)^{-1}=\frac{1}{8}\Big(1+\frac{x^3}{8}\Big)^{-1}=\sum_{k=0}^{\infty}\frac{(-1)^k x^{3k}}{8^{k+1}}=\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{3(k+1)}}\,x^{3k}.$$

此式在 $|x^3/8|<1$，即 $|x|<2$ 时成立。

现在乘以 x^m，然后在 $[0,1]$ 上逐项（term by term）积分；这个区间位于成立范围之内（而且题目已说明不必理会收敛（convergence）问题）：

$$I_m=\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{3(k+1)}}\int_0^1 x^{3k+m}\,\mathrm{d}x=\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{3(k+1)}}\cdot\frac{1}{3k+m+1},$$

这是因为当 $n = 3k + m \geq 0$ 时，$\int_0^1 x^n \,\mathrm{d}x = 1/(n+1)$。这就是所求的结果。

第 (ii) 部分

三个分母 $3k+3$、$3k+2$ 和 $3k+1$ 分别是 $m = 2, 1, 0$ 时的 $3k+m+1$。因此由第 (i) 部分，该级数等于

$$I_2 - 2I_1 + 4I_0 = \int_0^1 \frac{x^2 - 2x + 4}{8 + x^3} \,\mathrm{d}x.$$

想一想 为甚么 $x^2 - 2x + 4$ 会跟某些东西约掉?

因为 $8 + x^3$ 是两个立方数之和（sum of two cubes）：在 $a^3 + b^3 = (a+b)(a^2 - ab + b^2)$ 中取 $a = x$、$b = 2$，得

$$8 + x^3 = (x+2)(x^2 - 2x + 4).$$

因此被积函数化简为 $\dfrac{1}{x+2}$，所以该级数等于 $\displaystyle\int_0^1 \frac{1}{x+2} \,\mathrm{d}x$，即为所求。最后，

$$\int_0^1 \frac{1}{x+2} \,\mathrm{d}x = \Big[\ln(x+2)\Big]_0^1 = \ln 3 - \ln 2 = \ln \tfrac{3}{2}.$$

第 (iii) 部分

首先把括号按 k 拆成部分分式（partial fractions）。设 $2k+1 = \alpha(3k+2) + \beta(3k+1)$。代入 $k = -\frac{1}{3}$ 得 $\frac{1}{3} = \alpha$；代入 $k = -\frac{2}{3}$ 得 $-\frac{1}{3} = -\beta$，所以 $\beta = \frac{1}{3}$。因此

$$\frac{72(2k+1)}{(3k+1)(3k+2)} = \frac{24}{3k+1} + \frac{24}{3k+2},$$

而由第 (i) 部分（取 $m = 0$ 及 $m = 1$），该级数等于

$$24I_0 + 24I_1 = \int_0^1 \frac{24(1+x)}{(x+2)(x^2 - 2x + 4)} \,\mathrm{d}x.$$

现在把这个被积函数拆成部分分式：

$$\frac{1+x}{(x+2)(x^2 - 2x + 4)} = \frac{A}{x+2} + \frac{Bx + C}{x^2 - 2x + 4},$$

于是

$$1 + x = A(x^2 - 2x + 4) + (Bx + C)(x + 2).$$

代入 $x = -2$ 得 $-1 = 12A$, 所以 $A = -\frac{1}{12}$。比较 x^2 的系数得 $0 = A + B$, 所以 $B = \frac{1}{12}$。比较常数项得 $1 = 4A + 2C$, 所以 $C = \frac{2}{3}$。(验算 x 的系数: $-2A + 2B + C = \frac{1}{6} + \frac{1}{6} + \frac{2}{3} = 1$。)
乘以 24，得

$$\begin{aligned}\frac{24(1+x)}{8+x^3} &= -\frac{2}{x+2} + \frac{2x+16}{x^2-2x+4} \\ &= -\frac{2}{x+2} + \frac{2x-2}{x^2-2x+4} + \frac{18}{(x-1)^2+3},\end{aligned}$$

其中我们把二次式的导数 $2x-2$ 分拆出来，并对余下部分配方（completing the square）。我们把每一部分在 $[0,1]$ 上积分。

$$\begin{aligned}
\int_0^1 \frac{-2}{x+2}\,\mathrm{d}x &= -2\ln\tfrac{3}{2} = -2\ln 3 + 2\ln 2,\\
\int_0^1 \frac{2x-2}{x^2-2x+4}\,\mathrm{d}x &= \Big[\ln(x^2-2x+4)\Big]_0^1 = \ln\tfrac{3}{4} = \ln 3 - 2\ln 2,\\
\int_0^1 \frac{18}{(x-1)^2+3}\,\mathrm{d}x &= \frac{18}{\sqrt{3}}\Big[\arctan\frac{x-1}{\sqrt{3}}\Big]_0^1\\
&= 6\sqrt{3}\Big(0-\arctan\big(-\tfrac{1}{\sqrt{3}}\big)\Big) = 6\sqrt{3}\cdot\frac{\pi}{6} = \pi\sqrt{3}.
\end{aligned}$$

最后一行用了 $\int \frac{\mathrm{d}u}{u^2+c^2} = \frac{1}{c}\arctan\frac{u}{c}$，其中 $u=x-1$、$c=\sqrt{3}$。把三部分相加，$\ln 2$ 的项互相抵消，得

$$\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{3(k+1)}}\left(\frac{72(2k+1)}{(3k+1)(3k+2)}\right) = \pi\sqrt{3}-\ln 3,$$

所以 $a=3$ 及 $b=3$。

验算

- 这些级数收敛得非常快（每一项最多只有前一项的八分之一），所以取几项已足以验算。在第 (ii) 部分，级数首三项给出 $0.41667 - 0.01198 + 0.00084 \approx 0.40553$，而 $\ln\frac{3}{2}\approx 0.40547$。
- 在第 (iii) 部分，$\pi\sqrt{3}-\ln 3 \approx 5.4414 - 1.0986 = 4.3428$；用电脑把级数首 200 项相加，结果在小数点后十二位都吻合。即使只取 $k=0$ 那一项，也得 $\frac{1}{8}\cdot 36 = 4.5$，数量级正确。

常见错误

- 展开式写成 8^{-k} 而不是 $8^{-(k+1)}$：一开始提出的因子 $\frac{1}{8}$ 不可遗漏。
- 把 $\frac{1}{3k+3}$ 对应到 I_3 而不是 I_2：分母是 $3k+m+1$，所以 m 比常数项小一。
- 忘记 $\arctan\big(-\frac{1}{\sqrt{3}}\big)$ 是负数，以致 $\pi\sqrt{3}$ 那一项的正负号出错。

STEP 之外

像第 (i) 部分那样交换求和与积分的次序，在这里是合理的，因为该幂级数（power series）在 $[0,1]$ 上一致收敛（converges uniformly）；何时可以这样交换，是大学一年级分析（analysis）课程的核心主题之一。同样的方法可以像第 (iii) 部分那样，准确求出 $1-\frac{1}{4}+$

$\frac{1}{7} - \frac{1}{10} + \cdots = \int_0^1 \frac{\mathrm{d}x}{1+x^3}$ 之和。Abbott 的 *Understanding Analysis*（《理解分析》）对一致收敛和逐项积分有细致的讲解。

详细解答

STEP II 2024，第 3 题

纯数学 · 半角正切代换

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

第 (i) 部分建立了一个对照关系：圆上角度为 ϕ 的点，对应从 N 经过该点的直线与 y 轴相交处的高度 $\tan\frac{1}{2}\phi$。把这个对应反过来用，使每个高度 q 都恰好对应圆上的一个点。这样，题目中每个关于 q 的函数都会变成一条关于 $\frac{1}{2}\theta$ 的三角恒等式，我们便可以直接看出角度 θ 的变化，从而得知点的变化。

解答

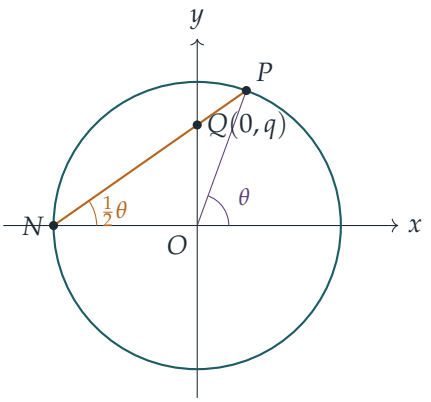
第 (i) 部分

由于 $P \neq N$，我们有 $\cos\theta \neq -1$，所以直线 NP 不是铅垂线。它的斜率（gradient）是

$$\frac{\sin\theta - 0}{\cos\theta - (-1)} = \frac{2\sin\frac{1}{2}\theta\cos\frac{1}{2}\theta}{2\cos^2\frac{1}{2}\theta} = \tan\tfrac{1}{2}\theta,$$

这里用了倍角公式（double-angle formulas）$\sin\theta = 2\sin\frac{1}{2}\theta\cos\frac{1}{2}\theta$ 和 $1+\cos\theta = 2\cos^2\frac{1}{2}\theta$（而且 $\cos\frac{1}{2}\theta \neq 0$，因为 $|\frac{1}{2}\theta| < \frac{1}{2}\pi$）。该直线通过 $(-1,0)$，所以它是 $y = (x+1)\tan\frac{1}{2}\theta$；代入 $x=0$ 便得 $q = \tan\frac{1}{2}\theta$，即所求。

从几何角度看，这就是圆周角（angle at the circumference）：由于 P 在圆心所对的角为 θ，弦 NP 在 N 处与 x 轴所成的角为 $\frac{1}{2}\theta$。



反过来：由高度得出点

我们会不断用到逆命题（converse）。给定任意实数 q，通过 N 和 $(0,q)$ 的直线是 $y = q(x+1)$。代入 $x^2+y^2=1$：

$$x^2 - 1 + q^2(x+1)^2 = 0 \iff (x+1)\left[(1+q^2)x - (1-q^2)\right] = 0.$$

根 $x=-1$ 对应 N，所以另一个交点是

$$\left(\frac{1-q^2}{1+q^2},\ \frac{2q}{1+q^2}\right).$$

若 $q = \tan\frac{1}{2}\phi$，则由 t 公式（t-formulas）可知此点为 $(\cos\phi, \sin\phi)$。由于 $\phi \mapsto \tan\frac{1}{2}\phi$ 是从 $(-\pi,\pi)$ 到 $\mathbb{R}$ 的双射（bijection），我们得到以下结果。

对照表。若对任何角 ϕ（不一定在 $(-\pi,\pi)$ 内）有 $\tan\frac{1}{2}\phi = q$，则从 N 经过 $(0,q)$ 的直线与圆的另一个交点为 $(\cos\phi, \sin\phi)$。

这对每个 ϕ 都成立，因为把 ϕ 换成 $\phi \pm 2\pi$ 既不改变 $\tan\frac{1}{2}\phi$（它作为 $\frac{1}{2}\phi$ 的函数，周期为 π），也不改变该点。

第 (ii) 部分

(a) 由于 $\tan\frac{1}{4}\pi = 1$，由和角公式（addition formula）得

$$\tan\left(\tfrac{1}{2}\theta + \tfrac{1}{4}\pi\right) = \frac{\tan\frac{1}{2}\theta + 1}{1 - \tan\frac{1}{2}\theta} = \frac{1+q}{1-q} = \mathrm{f}_1(q),$$

这就是 $\tan\frac{1}{2}\left(\theta + \frac{1}{2}\pi\right)$，即所求。条件 $q \neq 1$ 意味着 $\theta \neq \frac{1}{2}\pi$，所以分母不为零。

(b) 由 (a)，Q_1 的高度是 $\tan\frac{1}{2}\phi$，其中 $\phi = \theta + \frac{1}{2}\pi$。由对照表，$P_1 = \left(\cos(\theta+\frac{1}{2}\pi), \sin(\theta+\frac{1}{2}\pi)\right)$。所以

P_1 是P 的像，变换为绕O 逆时针旋转（rotation）$\frac{1}{2}\pi$.

（被排除的值 $q = 1$ 对应 $\theta = \frac{1}{2}\pi$，这正是唯一一个在此旋转下像为 N 本身的点。）

第 (iii) 部分

(a) P_2 是角度为 $\theta + \frac{1}{3}\pi$ 的点，所以应用第 (i) 部分（如有需要，先把角度化到 $(-\pi, \pi)$ 之内，这不会改变半角的 tan 值），并利用 $\tan\frac{1}{6}\pi = \frac{1}{\sqrt{3}}$：

$$\mathrm{f}_2(q) = \tan\left(\tfrac{1}{2}\theta + \tfrac{1}{6}\pi\right) = \frac{q + \frac{1}{\sqrt{3}}}{1 - \frac{q}{\sqrt{3}}} = \frac{\sqrt{3}q + 1}{\sqrt{3} - q}.$$

条件 $q \neq \sqrt{3}$ 即 $\theta \neq \frac{2}{3}\pi$，这恰好是 P_2 会等于 N、以致直线 NP_2 不存在的情况。

(b) 这次把 $\tan\frac{1}{4}\pi = 1$ 用在差角公式上：

$$\mathrm{f}_3(q) = \frac{1-q}{1+q} = \frac{\tan\frac{1}{4}\pi - \tan\frac{1}{2}\theta}{1 + \tan\frac{1}{4}\pi\tan\frac{1}{2}\theta} = \tan\left(\tfrac{1}{4}\pi - \tfrac{1}{2}\theta\right) = \tan\tfrac{1}{2}\left(\tfrac{1}{2}\pi - \theta\right).$$

由对照表，P_3 是角度为 $\frac{1}{2}\pi - \theta$ 的点，即 $(\sin\theta, \cos\theta)$。P 的坐标互换了，所以

$$P_3 \text{ 是}P \text{ 的反射（reflection），镜线为} y = x.$$

(c)

> **想一想** f_2^{-1} 对角度做了甚么？

由 (a)，对每个角 ϕ，f_2 都把 $\tan\frac{1}{2}\phi$ 映射到 $\tan\frac{1}{2}(\phi + \frac{1}{3}\pi)$（计算过程从未用到 $\phi = \theta$）。因此 $\mathrm{f}_2\big(\tan\frac{1}{2}(\psi - \frac{1}{3}\pi)\big) = \tan\frac{1}{2}\psi$；又由于 f_2 是一对一（one-to-one）的（解 $s = \mathrm{f}_2(q)$ 得唯一的 $q = \frac{\sqrt{3}s-1}{s+\sqrt{3}}$，其中 $s \neq -\sqrt{3}$），

$$\mathrm{f}_2^{-1}\left(\tan\tfrac{1}{2}\psi\right) = \tan\tfrac{1}{2}\left(\psi - \tfrac{1}{3}\pi\right).$$

所以 f_2^{-1} 把角度转回 $\frac{1}{3}\pi$。

现在逐步追踪角度经过这三个步骤的变化。由于 $0 < q < 1$，我们有 $0 < \theta < \frac{1}{2}\pi$。

- f_2：角度 $\theta + \frac{1}{3}\pi \in (\frac{1}{3}\pi, \frac{5}{6}\pi)$。此时 $\mathrm{f}_2(q) > 0$，所以可以应用 f_3。
- f_3：角度 $\frac{1}{2}\pi - (\theta + \frac{1}{3}\pi) = \frac{1}{6}\pi - \theta \in (-\frac{1}{3}\pi, \frac{1}{6}\pi)$。这不是 $-\frac{2}{3}\pi$，所以高度不是 $-\sqrt{3}$，可以应用 f_2^{-1}。
- f_2^{-1}：角度 $\frac{1}{6}\pi - \theta - \frac{1}{3}\pi = -\theta - \frac{1}{6}\pi$。

所以 $\mathrm{f}_4(q) = \tan\frac{1}{2}\left(-\theta - \frac{1}{6}\pi\right)$，而由对照表，$P_4$ 是角度为 $-\frac{1}{6}\pi - \theta$ 的点。

关于通过 O、角度为 α 的直线作反射，会把角度为 θ 的点映射到角度为 $2\alpha-\theta$ 的点（镜线平分这两个方向）。这里 $2\alpha=-\frac{1}{6}\pi$，所以

$$P_4 \text{ 是}P \text{ 的反射，镜线通过}O\text{，其倾角为 } -\tfrac{1}{12}\pi\text{（从正}x \text{ 轴量起），}$$

即直线 $y=-\tan\frac{1}{12}\pi\,x=-(2-\sqrt{3})\,x$。

这是合理的：先旋转 $\frac{1}{3}\pi$，再关于 $y=x$ 反射，然后转回去，等同于关于“把镜线 $y=x$ 转回 $\frac{1}{3}\pi$”后所得的直线作反射，而 $\frac{1}{4}\pi-\frac{1}{3}\pi=-\frac{1}{12}\pi$。明确地写出来，设 $t=\tan\frac{1}{12}\pi=2-\sqrt{3}$，

$$\mathrm{f}_4(q)=-\tan\left(\tfrac{1}{2}\theta+\tfrac{1}{12}\pi\right)=-\frac{q+2-\sqrt{3}}{1-(2-\sqrt{3})q}.$$

验算

- 在 (iii)(b) 中取 $\theta=\frac{1}{2}\pi$：$P=(0,1)$，$q=1$，$\mathrm{f}_3(1)=0$，所以 $Q_3=O$ 而 $P_3=(1,0)$，这确实是 $(0,1)$ 关于 $y=x$ 的反射。
- 对 (c) 取 $q=\frac{1}{2}$：$\mathrm{f}_2(\frac{1}{2})=\frac{\sqrt{3}+2}{2\sqrt{3}-1}\approx 1.5146$，接着 $\mathrm{f}_3\approx -0.2046$，然后 $\mathrm{f}_2^{-1}\approx -0.8868$，与封闭形式 $-\frac{\frac{1}{2}+2-\sqrt{3}}{1-\frac{1}{2}(2-\sqrt{3})}\approx -0.8868$ 吻合。
- 用电脑对 100 000 个随机点 P 进行检验，在舍入误差范围内确认了以上每一个关系（旋转 $\frac{1}{2}\pi$、f_2 的公式、关于 $y=x$ 的反射，以及关于倾角为 $-\frac{1}{12}\pi$ 的直线的反射）。

常见错误

- 断言 P_1 位于角度 $\theta+\frac{1}{2}\pi$，却没有留意这个角度可能落在 $(-\pi,\pi)$ 之外；这之所以无碍，只是因为当 ϕ 改变 2π 时，$\tan\frac{1}{2}\phi$ 和该点都会重复。
- 在 (c) 中，把 f_2^{-1} 的旋转方向弄反了（再加一次 $\frac{1}{3}\pi$），得出角度 $\frac{1}{2}\pi-\theta$，因而错误地令 P_4 与 P_3 相同。

STEP 之外

把角度为 θ 的点映射到 $q=\tan\frac{1}{2}\theta$ 的映射，就是把圆投影到一条直线上的球极投影（stereographic projection），而函数 f_1、f_2、f_3 都是莫比乌斯变换（Möbius transformations）$q\mapsto\frac{aq+b}{cq+d}$。圆的每一个旋转和反射都会变成直线上的这类变换，而对称变换的复合则对应 2×2 矩阵（matrices）$\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ 的乘法。Needham 的 *Visual Complex Analysis*（《可视化复分析》）把这幅图景发展得非常精彩，并将其推广到黎曼球面（Riemann sphere）。

详细解答

STEP II 2024，第 4 题

纯数学 · 向量与角平分线

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

把所有向量都缩放成单位向量（unit vector）。若 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{y}}$ 分别是沿 OX 和 OY 方向的单位向量，那么 $\hat{\mathbf{x}}+\hat{\mathbf{y}}$ 就是一个菱形（rhombus）的对角线，而菱形的对角线平分（bisect）其所在的角。题目给出的 $\mathbf{b}$ 只不过是 $\hat{\mathbf{x}}+\hat{\mathbf{y}}$ 的正数倍，所以每个平分向量（bisecting vector）都是两个单位向量之和的正数倍。这样一来，第 (ii) 部分的点积（dot product）就会化简成同一个对称的式子。

解答

全文以 θ 表示角 XOY。由于 O、X、Y 不共线（collinear），$0<\theta<\pi$，所以

$$-1<\cos\theta<1.$$

下面两个严格不等式都会用到。

第 (i)(a) 部分

设 $\hat{\mathbf{x}}=\mathbf{x}/|\mathbf{x}|$ 及 $\hat{\mathbf{y}}=\mathbf{y}/|\mathbf{y}|$，于是 $\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}=\cos\theta$。那么

$$\mathbf{b}=|\mathbf{x}|\mathbf{y}+|\mathbf{y}|\mathbf{x}=|\mathbf{x}||\mathbf{y}|\,(\hat{\mathbf{x}}+\hat{\mathbf{y}}).$$

这是 $\mathbf{x}$ 与 $\mathbf{y}$ 的线性组合，所以它位于平面 XOY 上。它不是零向量，因为若 $\hat{\mathbf{x}}+\hat{\mathbf{y}}=\mathbf{0}$，则 $\hat{\mathbf{y}}=-\hat{\mathbf{x}}$，那就会令 O、X、Y 共线。

现在比较两个角。我们有

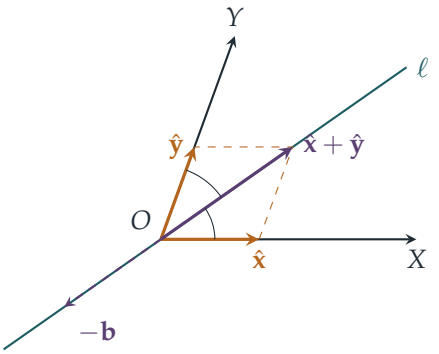
$$\mathbf{b}\cdot\hat{\mathbf{x}}=|\mathbf{x}||\mathbf{y}|\,(1+\cos\theta)=\mathbf{b}\cdot\hat{\mathbf{y}},$$

所以 $\mathbf{b}$ 与 $\overrightarrow{OX}$ 的夹角 α 以及 $\mathbf{b}$ 与 $\overrightarrow{OY}$ 的夹角 β 满足

$$\cos\alpha = \frac{\mathbf{b}\cdot\hat{\mathbf{x}}}{|\mathbf{b}|} = \frac{|\mathbf{x}||\mathbf{y}|(1+\cos\theta)}{|\mathbf{b}|} = \frac{\mathbf{b}\cdot\hat{\mathbf{y}}}{|\mathbf{b}|} = \cos\beta.$$

两个角都在 $[0,\pi]$ 之内，而余弦函数在这区间上是一一对应（one-to-one）的，所以 $\alpha=\beta$。此外 $1+\cos\theta>0$，所以 $\cos\alpha>0$，两个角都小于 90°。因此 $\mathbf{b}$ 是角 XOY 的一个平分向量。

其他平分向量。



由于 $\hat{\mathbf{x}}$ 与 $\hat{\mathbf{y}}$ 长度相等，它们所张成的平行四边形（parallelogram）是一个菱形，其对角线 $\hat{\mathbf{x}}+\hat{\mathbf{y}}$ 把角 XOY 分成相等的两半。在平面 XOY 上，与 OX 和 OY 成等角的方向，只有沿这条对角线 ℓ 的两个方向。若把一个向量从 ℓ 转开，它就会靠近其中一边而远离另一边，两个角便不再相等。沿著 ℓ，$\mathbf{b}$ 的方向与每一边所成的角为 $\frac{1}{2}\theta<90°$；相反方向与每一边所成的角为 $180°-\frac{1}{2}\theta>90°$，因此要排除。所以任何平分向量都沿 ℓ 并与 $\mathbf{b}$ 同向，即它是 $\mathbf{b}$ 的正数倍。

同一论证的代数版本：平面上任何向量都可写成 $\mathbf{v}=p\hat{\mathbf{x}}+q\hat{\mathbf{y}}$。两角相等即 $\mathbf{v}\cdot\hat{\mathbf{x}}=\mathbf{v}\cdot\hat{\mathbf{y}}$，而

$$\mathbf{v}\cdot\hat{\mathbf{x}}-\mathbf{v}\cdot\hat{\mathbf{y}} = (p+q\cos\theta)-(p\cos\theta+q) = (p-q)(1-\cos\theta).$$

由于 $\cos\theta\neq1$，这迫使 $p=q$，所以 $\mathbf{v}=p(\hat{\mathbf{x}}+\hat{\mathbf{y}})$。只有当 $p>0$ 时，两个角才是锐角。

第 (i)(b) 部分

点 B 位于直线 XY 上，当且仅当对某实数 t 有 $\overrightarrow{OB}=(1-t)\mathbf{x}+t\mathbf{y}$。我们需要

$$\lambda|\mathbf{y}|\,\mathbf{x}+\lambda|\mathbf{x}|\,\mathbf{y} = (1-t)\mathbf{x}+t\mathbf{y}.$$

由于 $\mathbf{x}$ 与 $\mathbf{y}$ 不平行，我们可以比较系数：$\lambda|\mathbf{y}|=1-t$ 及 $\lambda|\mathbf{x}|=t$。两式相加得

$$\lambda=\frac{1}{|\mathbf{x}|+|\mathbf{y}|},\qquad t=\frac{|\mathbf{x}|}{|\mathbf{x}|+|\mathbf{y}|}.$$

于是 $\overrightarrow{XB} = t(\mathbf{y}-\mathbf{x})$ 及 $\overrightarrow{BY} = (1-t)(\mathbf{y}-\mathbf{x})$，其中 $0 < t < 1$，所以 B 位于 X 与 Y 之间，且

$$XB : BY = t : (1-t) = |\mathbf{x}| : |\mathbf{y}| = OX : OY.$$

这就是角平分线定理（angle bisector theorem）：从 O 出发的角平分线按相邻两边之比分割 XY。

第 (i)(c) 部分

这里 $\overrightarrow{OB}$ 是 $\mathbf{b}$ 的非零倍数，所以 $OB \perp XY$ 即 $\mathbf{b}\cdot(\mathbf{y}-\mathbf{x}) = 0$。利用 $\mathbf{b} = |\mathbf{x}||\mathbf{y}|(\hat{\mathbf{x}}+\hat{\mathbf{y}})$ 及 $\mathbf{y}-\mathbf{x} = |\mathbf{y}|\hat{\mathbf{y}} - |\mathbf{x}|\hat{\mathbf{x}}$,

$$\begin{aligned}\mathbf{b}\cdot(\mathbf{y}-\mathbf{x}) &= |\mathbf{x}||\mathbf{y}|\big(|\mathbf{y}|\cos\theta + |\mathbf{y}| - |\mathbf{x}| - |\mathbf{x}|\cos\theta\big)\\ &= |\mathbf{x}||\mathbf{y}|\,(|\mathbf{y}|-|\mathbf{x}|)(1+\cos\theta).\end{aligned}$$

由于 $|\mathbf{x}|, |\mathbf{y}| > 0$ 且 $1+\cos\theta > 0$，这式只有在 $|\mathbf{y}| = |\mathbf{x}|$ 时才等于零。即 $OX = OY$，三角形 XOY 是等腰（isosceles）三角形。

第 (ii) 部分

设 $\mathbf{p}$、$\mathbf{q}$、$\mathbf{r}$ 分别为沿 $\overrightarrow{OP}$、$\overrightarrow{OQ}$、$\overrightarrow{OR}$ 的单位向量。四点中任何三点都不共线，特别是 O, P, Q 不共线，所以第 (i)(a) 部分适用于每一个角。因此所选的平分向量必定是

$$\mathbf{u} = \alpha(\mathbf{p}+\mathbf{q}), \qquad \mathbf{v} = \beta(\mathbf{q}+\mathbf{r}), \qquad \mathbf{w} = \gamma(\mathbf{r}+\mathbf{p}),$$

其中 $\alpha, \beta, \gamma > 0$，而且三者都不是零向量。

> ? **想一想** 三个平分向量都是任意的正数倍。我们怎样才能同时对三个角下结论?

比例因子都是正数，所以不会改变点积的正负号。而我们只需要正负号：两个非零向量的夹角是锐角、直角还是钝角，取决于它们的点积是正、零还是负。记 $S = \mathbf{p}\cdot\mathbf{q} + \mathbf{q}\cdot\mathbf{r} + \mathbf{r}\cdot\mathbf{p}$。利用 $\mathbf{q}\cdot\mathbf{q} = 1$,

$$\mathbf{u}\cdot\mathbf{v} = \alpha\beta\,(\mathbf{p}\cdot\mathbf{q} + \mathbf{p}\cdot\mathbf{r} + \mathbf{q}\cdot\mathbf{q} + \mathbf{q}\cdot\mathbf{r}) = \alpha\beta\,(1+S).$$

完全同理，利用 $\mathbf{r}\cdot\mathbf{r} = 1$ 及 $\mathbf{p}\cdot\mathbf{p} = 1$,

$$\mathbf{v}\cdot\mathbf{w} = \beta\gamma\,(1+S), \qquad \mathbf{w}\cdot\mathbf{u} = \gamma\alpha\,(1+S).$$

由于 $\alpha\beta$、$\beta\gamma$、$\gamma\alpha$ 全是正数，三个点积的正负号都与 $1+S$ 相同。所以当 $1+S > 0$ 时三个角都是锐角，当 $1+S = 0$ 时都是直角，当 $1+S < 0$ 时都是钝角。 ■

验算

- 取 $\mathbf{x} = (3,0,0)$ 及 $\mathbf{y} = (0,4,0)$。那么 $\mathbf{b} = (12,12,0)$，它平分该直角。取 $\lambda = \frac{1}{7}$ 得 $B = (\frac{12}{7}, \frac{12}{7}, 0)$，它位于 $\frac{x}{3} + \frac{y}{4} = 1$ 上，且 $XB : BY = 3 : 4$。
- 第 (ii) 部分的三种情况确实都会出现。若 P、Q、R 在同一平面上两两相隔 120°，则 $S = -\frac{3}{2}$，而各平分向量之间的夹角为 120°（钝角）。若 $\mathbf{p}, \mathbf{q}, \mathbf{r}$ 指向一个以 O 为中心的正四面体（regular tetrahedron）的三个顶点，则每个点积都是 $-\frac{1}{3}$，所以 $S = -1$，三个角都是直角。三个几乎平行的向量则给出锐角。
- 两部分的结果亦已用电脑在数以千计的随机配置上验证过。

常见错误

- 忘记必须用到 $\cos\theta \neq \pm 1$。(i)(c) 中的因子 $1 + \cos\theta$ 以及唯一性论证中的因子 $1 - \cos\theta$ 之所以不为零，全因各点不共线。
- 在第 (ii) 部分，把平分向量当作某些特定向量，例如 $|\mathbf{x}|\mathbf{y} + |\mathbf{y}|\mathbf{x}$。题目容许它们自由选取，所以你必须引用第 (i)(a) 部分，才能说它们是这些向量的正数倍。

STEP 之外

把向量写成 $|\mathbf{x}|\hat{\mathbf{x}}$ 并以单位向量运算，是一个在整个线性代数（linear algebra）中都大有裨益的习惯。(ii) 中的恒等式是 $\mathbf{p}, \mathbf{q}, \mathbf{r}$ 的格拉姆矩阵（*Gram matrix*）几何的一小部分，该矩阵的元素就是这些向量两两之间的点积。由 $|\mathbf{p} + \mathbf{q} + \mathbf{r}|^2 = 3 + 2S \geq 0$ 可知 $S \geq -\frac{3}{2}$，所以钝角的情况是有界限的。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）正是以这种精神讲解点积、投影和格拉姆矩阵。

详细解答

STEP II 2024，第 5 题

纯数学 · 值域与丢番图方程

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

把多项式的自变量平移一个整数，它在 $\mathbb{Z}$ 上的值域（range）并不会改变，所以我们通过配方（completing the square）（或“配立方”）把每个函数化成最简单的平移形式。这样，问两个值域是否相交，就变成问一条含两个整数未知数的方程是否有解；而窍门在于把方程整理成两个整数之积 = 一个固定的小数的形式。一个小数的因数分解只有寥寥几种可能，我们可以逐一检查。

解答

第 (i) 部分

配方得

$$\mathrm{f}_1(n) = (n+3)^2 + 2 = \mathrm{F}_1(n+3).$$

当 n 取遍 $\mathbb{Z}$ 时，$m = n+3$ 亦然（每个整数 m 都会出现，取 $n = m-3$ 即可）。因此 f_1 的值域是 $\{m^2+2 : m \in \mathbb{Z}\}$，正好就是 F_1 的值域。这就是所要证明的。

第 (ii) 部分

同理 $\mathrm{g}_1(n) = (n-1)^2 + 4$，故按同样的论证，$\mathrm{g}_1$ 的值域是 $\{k^2+4 : k \in \mathbb{Z}\}$。由第 (i) 部分，若有公共值，便需要整数 m、k 使得

$$m^2 + 2 = k^2 + 4 \quad \Longleftrightarrow \quad (m-k)(m+k) = 2.$$

由于 $m-k$ 与 $m+k$ 相差 $2k$，是偶数，所以两者奇偶性（parity）相同。若两者皆为奇数，其积为奇数，不可能等于 2。若两者皆为偶数，其积是 4 的倍数，同样不可能等于 2。因

此这样的 m、k 并不存在，两个值域的交集（intersection）为空。

第 (iii) 部分

配方得

$$\mathrm{f}_2(n) = (n-1)^2 - 7, \qquad \mathrm{g}_2(n) = (n-2)^2 - 2,$$

所以两个值域分别是 $\{m^2-7\}$ 和 $\{k^2-2\}$，其中 m、k 取遍 $\mathbb{Z}$。公共值需要

$$m^2 - 7 = k^2 - 2 \quad \Longleftrightarrow \quad (m-k)(m+k) = 5.$$

由于 5 是质数（prime），数对 $(m-k, m+k)$ 只能是 $(1,5)$、$(5,1)$、$(-1,-5)$、$(-5,-1)$ 其中之一。两个因数相加得 $2m = \pm 6$，相减得 $2k = \pm 4$，所以在每种情况下都有 $m^2 = 9$ 及 $k^2 = 4$。公共值为

$$m^2 - 7 = 9 - 7 = 2 = 4 - 2 = k^2 - 2.$$

所以同时属于两个值域的整数只有 $\boxed{2}$。验证一下：$\mathrm{f}_2(4) = \mathrm{f}_2(-2) = 2$，而 $\mathrm{g}_2(4) = \mathrm{g}_2(0) = 2$。

第 (iv) 部分

对 p 配方，

$$p^2 + pq + q^2 = \left(p + \tfrac{1}{2}q\right)^2 + \tfrac{3}{4}q^2 \geq 0,$$

因为两项都是平方乘以正常数。（等号只在 $q = 0$ 且 $p = 0$ 时成立。）

至于三次函数，我们把 f_3 与 $(n-1)^3 = n^3 - 3n^2 + 3n - 1$ 作比较：

$$\mathrm{f}_3(n) = (n-1)^3 + 4n + 1 = (n-1)^3 + 4(n-1) + 5.$$

所以设 $m = n - 1$（它同样取遍整个 $\mathbb{Z}$），则 f_3 的值域是 $\{m^3 + 4m + 5\}$，而 g_3 的值域是 $\{k^3 + 4k - 6\}$。公共值需要整数 m、k 使得

$$\begin{aligned} & m^3 + 4m + 5 = k^3 + 4k - 6 \\ \Longleftrightarrow \quad & (k^3 - m^3) + 4(k-m) = 11 \\ \Longleftrightarrow \quad & (k-m)\left(k^2 + km + m^2 + 4\right) = 11. \end{aligned}$$

? **想一想** 本部分开首的不等式对这条方程有甚么帮助?

在该不等式中取 $p = k$、$q = m$，便得 $k^2 + km + m^2 + 4 \geq 4$。特别地，第二个因式是正整数，所以 $k - m$ 亦必须是正数。由于 11 是质数，只有两个可能：

- $k-m=11$ 且 $k^2+km+m^2+4=1$。这不可能，因为第二个因式至少是 4。
- $k-m=1$ 且 $k^2+km+m^2+4=11$。

在第二种情况下，代入 $k=m+1$：

$$\begin{aligned}(m+1)^2+m(m+1)+m^2&=7\\ 3m^2+3m+1&=7\\ m^2+m-2=(m+2)(m-1)&=0,\end{aligned}$$

所以 $m=1$（此时 $k=2$）或 $m=-2$（此时 $k=-1$）。对应的值为

$$m=1:\ 1+4+5=10,\qquad m=-2:\ -8-8+5=-11.$$

所以同时属于两个值域的整数是 $\boxed{10\text{ 和 }-11}$。

回到原来的变量：$f_3(2)=8-12+14=10$ 而 $g_3(2)=8+8-6=10$；另外 $f_3(-1)=-1-3-7=-11$ 而 $g_3(-1)=-1-4-6=-11$。有趣的是，两个公共值都来自相同的输入 n。

验算

- 如上文般代回验证：f_2 和 g_2 在 $n=4$ 时都取值 2，而 f_3、g_3 在 $n=2$（值为 10）及 $n=-1$（值为 -11）时相等。
- 列出一些较小的值，可快速检查第 (ii) 部分是否合理：f_1 给出 $2,3,6,11,18,27,\ldots$，而 g_1 给出 $4,5,8,13,20,29,\ldots$。两个列表都没有对方的数。
- 以电脑搜索所有 $|n|\le 3000$,证实 (i) 的两个值域相同、(ii) 的两个值域不相交(disjoint)，而唯一的公共值是 (iii) 中的 2，以及 (iv) 中的 10、-11。

常见错误

- 以同一个 n 解 $f(n)=g(n)$ 来找公共值。两个输入可以不同，所以我们需要两个独立的未知数。在 (iii) 中，$f_2(n)=g_2(n)$ 碰巧给出正确答案，但方法是错的。
- 在 (ii) 中，只论证 $(m-k)(m+k)=2$ 迫使 $\{m-k,m+k\}=\{1,2\}$，却忘记了负数因数对。奇偶性论证可一次过处理所有情况。
- 在 (iv) 中，忘记排除 11×1 这个因数分解（factorisation）。正是那条不等式把它排除掉。

STEP 之外

像 $m^2 - k^2 = 5$ 和 $k^3 + 4k = m^3 + 4m + 11$ 这类只计算整数解的方程，称为丢番图方程（Diophantine equations）。把一边因式分解，再利用另一边的因数，是解这类方程最有力的初等工具之一。(ii) 中的奇偶性技巧可以推广：一个整数可表示为两个平方数之差，当且仅当它是奇数或 4 的倍数。若想深入了解，Davenport 的 *The Higher Arithmetic*（《高等算术》）是一本极佳的大学一年级数论（number theory）入门书。

详细解答

STEP II 2024，第 6 题

纯数学 · 二项式级数的系数

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

每一部分都靠相邻项之比（ratio of successive terms）推动。二项式系数（binomial coefficient）$\binom{2n}{n}$ 和二项级数（binomial series）的系数都是乘积，每走一步便多乘一个因子，所以把第 n 项与第 $(n-1)$ 项比较，就得到一条简短的递推关系（recurrence）。两个数列若服从同一条递推关系，而且首项相同，它们便相等。最后的恒等式则来自把 $(1-x)^{-3/2}$ 写成 $(1-x)^{-1/2}\times(1-x)^{-1}$ 并比较系数（comparing coefficients），因为乘以 $(1-x)^{-1}=1+x+x^2+\cdots$ 会把一列系数变成它的累计和。

解答

第 (i) 部分

设 $P(n)$ 为命题 $T_n=\dfrac{1}{2^{2n}}\binom{2n}{n}$。

基本情况（*base case*）。$P(0)$ 成立，因为 $\frac{1}{2^0}\binom{0}{0}=1=T_0$。

归纳步骤（*inductive step*）。假设对某个 $n\geq 1$，$P(n-1)$ 成立。由于 $\binom{2n}{n}=\dfrac{(2n)!}{n!\,n!}$ 及 $\binom{2n-2}{n-1}=\dfrac{(2n-2)!}{(n-1)!\,(n-1)!}$，两者相除得

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

于是，利用递推关系及 $P(n-1)$，

$$\begin{aligned} T_n = \frac{2n-1}{2n}\,T_{n-1} &= \frac{2n-1}{2n}\cdot\frac{1}{2^{2n-2}}\binom{2n-2}{n-1} \\ &= \frac{2n-1}{2n}\cdot\frac{1}{2^{2n-2}}\cdot\frac{n}{2(2n-1)}\binom{2n}{n} \\ &= \frac{1}{4\cdot 2^{2n-2}}\binom{2n}{n} = \frac{1}{2^{2n}}\binom{2n}{n}, \end{aligned}$$

这正是 $P(n)$。由归纳法（induction），$P(n)$ 对所有 $n = 0, 1, 2, \ldots$ 皆成立。

第 (ii) 部分

由一般指数的二项式定理（binomial theorem for a general index），$(1-x)^{-1/2}$ 中 x^r 的系数为

$$a_r = \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{1}{2}-r+1\right)}{r!}\,(-1)^r = \frac{\frac{1}{2}\cdot\frac{3}{2}\cdots\frac{2r-1}{2}}{r!},$$

因为分子中 r 个因子各自为负，而 $(-x)^r$ 带来的 $(-1)^r$ 正好抵消这些符号。由 a_{r-1} 到 a_r，分子多了一个因子 $\frac{2r-1}{2}$，分母多了一个因子 r，所以当 $r \geq 1$ 时

$$\frac{a_r}{a_{r-1}} = \frac{(2r-1)/2}{r} = \frac{2r-1}{2r}, \qquad \text{即} \qquad a_r = \frac{2r-1}{2r}\,a_{r-1}.$$

> ? **想一想** 为甚么这条递推关系能立即推出 $a_r = T_r$?

两个数列的首项都是 1（因为常数项 $a_0 = 1$，而 $T_0 = 1$），而且由第 $r-1$ 项到第 r 项遵从同一规则。一行归纳便能完成：若 $a_{r-1} = T_{r-1}$，则 $a_r = \frac{2r-1}{2r}a_{r-1} = \frac{2r-1}{2r}T_{r-1} = T_r$。所以对所有 $r \geq 0$ 都有 $a_r = T_r$，再由 (i)，

$$a_r = \frac{1}{2^{2r}}\binom{2r}{r}.$$

第 (iii) 部分

同样地，$(1-x)^{-3/2}$ 中 x^r 的系数为

$$b_r = \frac{\frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2r+1}{2}}{r!}.$$

a_r 与 b_r 的分子都是 r 个连续“半奇数”的乘积。b_r 的乘积少了 a_r 的第一个因子 $\frac{1}{2}$，而在末端多了一个因子 $\frac{2r+1}{2}$。$r!$ 互相约去，所以

$$\frac{b_r}{a_r} = \frac{(2r+1)/2}{1/2} = 2r+1.$$

（当 $r = 0$ 时，两个乘积都是空乘积，$b_0 = a_0 = 1$，与此相符。）因此

$$b_r = (2r+1)\,a_r = \frac{2r+1}{2^{2r}}\binom{2r}{r}, \qquad r = 0, 1, 2, \ldots.$$

第 (iv) 部分

由于 $(1-x)^{-3/2} = (1-x)^{-1/2}(1-x)^{-1}$，且 $(1-x)^{-1} = \sum_{s\geq 0} x^s$，

$$\sum_{n=0}^{\infty} b_n x^n = \left(\sum_{r=0}^{\infty} a_r x^r\right)\left(\sum_{s=0}^{\infty} x^s\right).$$

在右方，x^n 项来自从第一个括号选取 $a_r x^r$、从第二个括号选取 x^{n-r}，其中 $r = 0, 1, \ldots, n$。每一种选法的系数都是 $a_r \cdot 1$，所以 x^n 的系数为 $a_0 + a_1 + \cdots + a_n$。比较两边 x^n 的系数，

$$b_n = \sum_{r=0}^{n} a_r.$$

代入 (ii) 及 (iii) 的公式，得

$$\frac{2n+1}{2^{2n}}\binom{2n}{n} = \sum_{r=0}^{n} \frac{1}{2^{2r}}\binom{2r}{r},$$

即所求。（上述推导对 $n = 0$ 亦成立，此时两边都等于 1。）

验算

- 小数值情况：$a_0, a_1, a_2, a_3 = 1, \frac{1}{2}, \frac{3}{8}, \frac{5}{16}$，与 $\binom{2r}{r}/4^r = 1, \frac{2}{4}, \frac{6}{16}, \frac{20}{64}$ 吻合。
- 当 $n = 3$，(iv) 的左边为 $\frac{7\cdot 20}{64} = \frac{35}{16}$，右边为 $1 + \frac{1}{2} + \frac{3}{8} + \frac{5}{16} = \frac{35}{16}$。
- 直接展开，$(1-x)^{-3/2} = 1 + \frac{3}{2}x + \frac{15}{8}x^2 + \cdots$，而 $b_2 = 5 \cdot \frac{3}{8} = \frac{15}{8}$。
- 以精确分数进行的电脑验算确认 (i)–(iv) 对每个 $n \leq 29$ 皆成立。在 $x = 0.3$ 处把两个级数各自的首 30 项相加，得 1.19523 及 1.70747，分别与 $(0.7)^{-1/2}$ 及 $(0.7)^{-3/2}$ 相符。

常见错误

- 二项式系数的符号错误：$\binom{-1/2}{r}$ 的符号正负交替，但 $(-x)^r$ 带来的因子 $(-1)^r$ 抵消了这一点，所以每个 a_r 和 b_r 都是正数。
- 忘记基本情况，或单凭递推关系便断言 $a_r = T_r$：首项 $a_0 = T_0 = 1$ 也必须相符。

STEP 之外

把幂级数（power series）乘以 $(1-x)^{-1}$ 以得出部分和（partial sums），是*生成函数*（generating functions）的基本工具；生成函数把关于数列的问题转化为幂级数的代数运算。数 $\binom{2n}{n}/4^n$ 亦出现在概率（probability）之中：它是一枚公平硬币掷 $2n$ 次恰好出现

n 次正面的概率，而斯特林公式（Stirling's formula）显示当 n 很大时，它的表现约为 $1/\sqrt{\pi n}$。Siklos 的 *Advanced Problems in Mathematics*（《数学进阶难题》）收录更多同类题目，而 Grimmett 与 Welsh 的 *Probability: An Introduction*（《概率导论》）则有系统地发展生成函数。

详细解答

STEP II 2024，第 7 题

纯数学 · 曲线描绘

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

每个括号都是一个圆的方程，所以 C_1 不过是两个圆。至于 C_2，把括号展开后，所有项都会归结为 x^2+y^2 和 x^2，因此用一条水平直线去截它，就会得到一个关于 x^2 的二次方程（quadratic）。于是计算交点数目就变成计算这个二次方程有多少个正根，而这只需结合判别式（discriminant）、两根之和与两根之积便可确定。

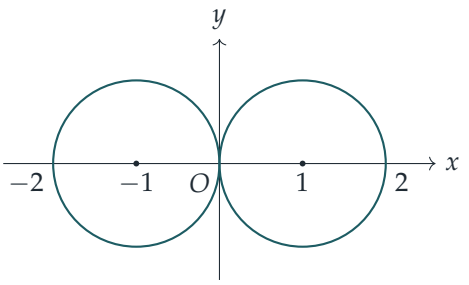
解答

全题中记

$$A = y^2+(x-1)^2-1 = x^2+y^2-2x, \qquad B = y^2+(x+1)^2-1 = x^2+y^2+2x.$$

第 (i) 部分

一个乘积等于零，当且仅当其中一个因子等于零，所以 C_1 是圆 $A=0$ 与圆 $B=0$ 的并集（union）。这两个圆的半径都是 1，圆心分别在 $(1,0)$ 和 $(-1,0)$。两圆都经过原点，并在原点相切（它们在该处的公切线（tangent）是 y 轴）。曲线的最高点是 $(\pm1,1)$，与 x 轴相交于 $x=-2,0,2$。



第 (ii)(a) 部分

由于 A 和 B 分别具有 $s-2x$ 和 $s+2x$ 的形式，其中 $s=x^2+y^2$，故

$$AB=(x^2+y^2)^2-4x^2.$$

在直线 $y=k$ 上，方程 $AB=\frac{1}{16}$ 变为

$$\begin{aligned}(x^2+k^2)^2-4x^2&=\tfrac{1}{16}\\ x^4+2k^2x^2+k^4-4x^2-\tfrac{1}{16}&=0\\ x^4+2(k^2-2)x^2+\left(k^4-\tfrac{1}{16}\right)&=0,\end{aligned}$$

此即所求。

现在设 $u=x^2$，得

$$u^2+2(k^2-2)u+\left(k^4-\tfrac{1}{16}\right)=0. \qquad (*)$$

每个根 $u>0$ 给出两个交点 $x=\pm\sqrt{u}$；根 $u=0$ 给出一个交点（$x=0$）；根 $u<0$ 则不给出任何交点。所以我们需要知道 $(*)$ 各根的正负号。以下三个量可以告诉我们：

- 四分之一判别式（quarter-discriminant）$(k^2-2)^2-\left(k^4-\frac{1}{16}\right)=\frac{65}{16}-4k^2$，当 k 小于、等于或大于 $\frac{\sqrt{65}}{8}$ 时，它分别为正、零或负；
- 两根之和 $2(2-k^2)$；
- 两根之积 $k^4-\frac{1}{16}$，当 $k>0$ 时，它与 $k-\frac{1}{2}$ 同号。

注意 $\frac{1}{2}<\frac{\sqrt{65}}{8}\approx 1.008$，而且只要两根是实数，就有 $k^2\le\frac{65}{64}<2$，所以两根之和为正。

> ? **想一想** 如果两根之积为负，我们还需要检查判别式吗？

不需要：两根之积为负，必然迫使两根为实数（此时判别式 b^2-4ac 为正）且一正一负。现在逐一考虑各种情况。

- $0<k<\frac{1}{2}$：两根之积为负，所以恰有一个正根。**2 个交点**。

- $k = \frac{1}{2}$：$(*)$ 变为 $u^2 - \frac{7}{2}u = 0$，所以 $u = 0$ 或 $u = \frac{7}{2}$，得 $x = 0$ 及 $x = \pm\sqrt{7/2}$。**3 个交点**。
- $\frac{1}{2} < k < \frac{\sqrt{65}}{8}$：判别式为正，所以有两个相异实根。它们的积与和都是正数，所以两根皆为正。**4 个交点**。
- $k = \frac{\sqrt{65}}{8}$：有一个重根（repeated root）$u = 2 - k^2 = \frac{63}{64} > 0$，得 $x = \pm\frac{\sqrt{63}}{8}$。**2 个交点**。
- $k > \frac{\sqrt{65}}{8}$：没有实根。**没有交点**。

第 (ii)(b) 部分

由 (a) 可知，当 $0 < k \leq \frac{\sqrt{65}}{8}$ 时，直线 $y = k$ 与 C_2 相交，而 k 更大时则不相交。C_2 的方程只通过 y^2 涉及 y，所以负的 k 不会给出更高的点。因此 C_2 上最大的 y 坐标是 $\frac{\sqrt{65}}{8}$，在 $x = \pm\frac{\sqrt{63}}{8}$ 处取得。

在 C_1 上，最高点是两个圆的顶点 $(\pm 1, 1)$。由于 $\frac{\sqrt{63}}{8} < \frac{\sqrt{64}}{8} = 1$，$C_2$ 的最高点比 C_1 的最高点**更接近 y 轴**（不过稍为高一点，因为 $\frac{\sqrt{65}}{8} > 1$）。

第 (ii)(c) 部分

假设 $A < 0$ 且 $B < 0$。由 $A < 0$ 得 $2x > x^2 + y^2 \geq 0$，所以 $x > 0$。由 $B < 0$ 得 $-2x > x^2 + y^2 \geq 0$，所以 $x < 0$。两者互相矛盾，所以 A 和 B 不可能同时为负。（从几何上看，两个开圆盘（open discs）互不重叠。）

在 C_2 上有 $AB = \frac{1}{16} > 0$，所以 A 和 B 都不为零，且同号。它们不可能同时为负，所以两者皆为正。$A > 0$ 表示 $(x-1)^2 + y^2 > 1$，即该点严格位于以 $(1, 0)$ 为圆心的圆之外。同理，$B > 0$ 表示该点严格位于以 $(-1, 0)$ 为圆心的圆之外。因此 C_2 上每一点都位于两个圆之外：C_2 完全位于 C_1 之外。

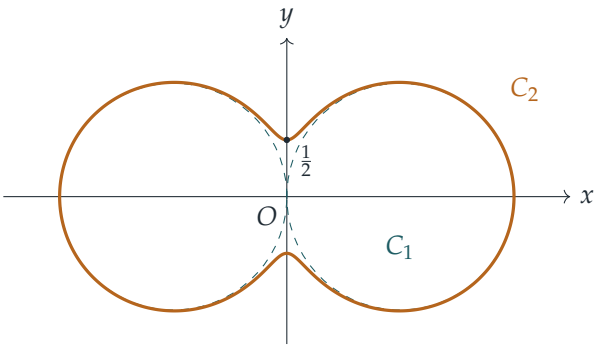
第 (ii)(d) 部分

有助描绘 C_2 的特征：

- 由于方程中只出现 x^2 和 y^2，曲线关于两条坐标轴都对称（symmetric）。
- 在 $y = 0$ 上，$(*)$ 给出 $u = 2 + \frac{\sqrt{65}}{4}$，所以 C_2 与 x 轴相交于 $x = \pm\sqrt{2 + \frac{\sqrt{65}}{4}} \approx \pm 2.004$，刚好在 ± 2 之外。
- 在 $x = 0$ 上，$k^4 = \frac{1}{16}$，所以 C_2 与 y 轴相交于 $(0, \pm\frac{1}{2})$。
- 它的最高点和最低点是 $\left(\pm\frac{\sqrt{63}}{8}, \pm\frac{\sqrt{65}}{8}\right) \approx (\pm 0.99, \pm 1.01)$。

(a) 中的交点数目揭示了曲线的形状。当 $|y| < \frac{1}{2}$ 时只有两个交点（外侧两边）。当 $\frac{1}{2} <$

$|y| < \frac{\sqrt{65}}{8}$ 时有四个交点，所以曲线在两个隆起之间向下凹至 $(0, \pm\frac{1}{2})$。因此 C_2 是一条包围两个圆的单一闭合“花生形”曲线，在 y 轴处向内收窄。



（把 C_2 写成极坐标形式（polar form），得 $r^2 = 2\cos^2\theta + \sqrt{4\cos^4\theta + \frac{1}{16}}$，上图的曲线正是这样画出来的。）

验算

- 当 $k = \frac{\sqrt{65}}{8}$ 时，判别式 $\frac{65}{16} - 4k^2$ 等于零，与曲线的顶部吻合。重根 $u = \frac{63}{64}$ 给出 $x^2 + y^2 = \frac{128}{64} = 2$ 及 $4x^2 = \frac{63}{16}$，而 $2^2 - \frac{63}{16} = \frac{1}{16}$ 确实成立。
- 点 $(0, \frac{1}{2})$：$A = B = \frac{1}{4}$，所以 $AB = \frac{1}{16}$。
- 用电脑计算 $y = k$ 与 C_2 的交点数目，对于 $k = 0.1$、0.49、0.51、1.0、1.007 和 1.01，分别得到 2、2、4、4、4 和 0 个交点；而数值搜索得出最大高度为 $1.00778 \approx \frac{\sqrt{65}}{8}$，位于 $x \approx 0.9922 \approx \frac{\sqrt{63}}{8}$。

常见错误

- 把 $u = 0$ 当作给出两个交点：$x = \pm\sqrt{0}$ 只是一个点，这就是为甚么 $k = \frac{1}{2}$ 给出 3 个交点而不是 4 个。
- 计算交点数目时遗漏了相切（tangency）的情况 $k = \frac{\sqrt{65}}{8}$。
- 在 (c) 中只得出“A 和 B 同号”，而没有排除两者同时为负的情况——而这正是前半部分的重点所在。

STEP 之外

C_2 属于曲线族 $(x^2+y^2)^2-4x^2=c$，即函数 AB 的等值曲线（level curves）。随着 c 增大，曲线的形状会改变：先是两个位于圆内的独立卵形（当 $c<0$），然后在 $c=0$ 时成为 8 字形（figure-of-eight）曲线 C_1，之后成为单一的花生形，而当 $c\geq 4$ 时其凹陷便会消失。拓扑（topology）的改变恰好发生在 AB 于原点的临界点（critical point）所对应的水平上。追踪等值曲线在经过临界点时如何变化，正是莫尔斯理论（Morse theory）的出发点；而与此密切相关的卡西尼卵形线（Cassini ovals）——其 8 字形成员就是伯努利双纽线（lemniscate of Bernoulli）——也有完全相同的表现。Siklos 的 *Advanced Problems in Mathematics*（《数学进阶问题》）收录了更多类似 STEP 风格、描绘这类隐函数曲线（implicit curves）的题目。

详细解答

STEP II 2024，第 8 题

纯数学 · 算术几何平均

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

本题在 *Beyond STEP* 第 2.6 节算术几何平均中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch02/agm.html

关键思路

两个数列互相夹近：算术平均数逐步下降，几何平均数逐步上升，而两者之差每次至少减半，所以两个数列都趋于同一个极限 M，即 x_0 与 y_0 的算术—几何平均数（arithmetic–geometric mean）。在第 (ii) 部分，代换（substitution）把数对 (p,q) 的积分变成这两个数的平均数所成数对的积分，所以在整个迭代过程中积分的值始终不变。令 $n \to \infty$，积分便化为一个我们能够计算的积分。

解答

第 (i) 部分

首先用归纳法（induction）证明对每个 n 都有 $0 < y_n < x_n$。当 $n = 0$ 时成立。假设 $0 < y_n < x_n$，则 $\sqrt{x_n} \neq \sqrt{y_n}$，所以

$$0 < \left(\sqrt{x_n} - \sqrt{y_n}\right)^2 = x_n + y_n - 2\sqrt{x_n y_n},$$

整理得 $\frac{1}{2}(x_n + y_n) > \sqrt{x_n y_n}$，即 $y_{n+1} < x_{n+1}$。此外 $y_{n+1} = \sqrt{x_n y_n} > 0$。归纳法至此完成，所以对所有 $n \geq 0$ 都有 $y_{n+1} < x_{n+1}$。

利用 $y_n < x_n$：

$$x_{n+1} = \tfrac{1}{2}(x_n + y_n) < \tfrac{1}{2}(x_n + x_n) = x_n,$$
$$y_{n+1} = \sqrt{x_n y_n} > \sqrt{y_n y_n} = y_n,$$

其中第二行用到 $x_n > y_n > 0$ 以及平方根是递增函数这一事实。

所以 (y_n) 是递增数列，而且有上界（bounded above）：$y_n < x_n < x_{n-1} < \cdots < x_0$，所以对所有 n 都有 $y_n < x_0$。根据该定理，存在某个 $M \le x_0$ 使得 $y_n \to M$。（由于数列从 y_0 开始递增，亦有 $M \ge y_0 > 0$。）

至于两者之差，左边的不等式 $x_{n+1} - y_{n+1} > 0$ 正是我们最先证明的结果。至于右边的不等式，由于 $y_{n+1} > y_n$，

$$x_{n+1} - y_{n+1} < x_{n+1} - y_n = \tfrac{1}{2}(x_n + y_n) - y_n = \tfrac{1}{2}(x_n - y_n).$$

反复应用此式，得

$$0 < x_n - y_n < \frac{x_0 - y_0}{2^n} \to 0,$$

所以由夹逼原理（squeezing）得 $x_n - y_n \to 0$。最后，$x_n = y_n + (x_n - y_n)$ 是一个趋于 M 的数列与一个趋于 0 的数列之和，所以 $x_n \to M$ 亦成立。

第 (ii) 部分

记 $a = \tfrac{1}{2}(p+q)$、$g = \sqrt{pq}$，并把关于 t 的积分记作 J。

代换。对 $x > 0$，令 $t = \tfrac{1}{2}\left(x - \tfrac{pq}{x}\right)$，则

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{1}{2}\left(1 + \frac{pq}{x^2}\right) = \frac{x^2 + pq}{2x^2} > 0,$$

所以 t 严格递增，当 $x \to 0^+$ 时趋于 $-\infty$，当 $x \to \infty$ 时趋于 $+\infty$：此代换把 $(0,\infty)$ 一一对应（one-to-one）地映射到整条实数线上。现在改写根号内的两个因子。

$$pq + t^2 = \tfrac{1}{4}\left(x^2 + 2pq + \frac{p^2q^2}{x^2}\right) = \frac{1}{4}\left(x + \frac{pq}{x}\right)^2 = \left(\frac{x^2+pq}{2x}\right)^2,$$
$$\begin{aligned}\tfrac{1}{4}(p+q)^2 + t^2 &= \tfrac{1}{4}\left(p^2 + 2pq + q^2 + x^2 - 2pq + \frac{p^2q^2}{x^2}\right)\\ &= \frac{x^4 + (p^2+q^2)x^2 + p^2q^2}{4x^2} = \frac{(x^2+p^2)(x^2+q^2)}{4x^2}.\end{aligned}$$

由于 $x > 0$ 且 $x^2 + pq > 0$，两者乘积的平方根为

$$\frac{x^2+pq}{2x} \cdot \frac{\sqrt{(p^2+x^2)(q^2+x^2)}}{2x}.$$

因此

$$J = \int_0^\infty \frac{x^2+pq}{2x^2}\cdot\frac{4x^2}{(x^2+pq)\sqrt{(p^2+x^2)(q^2+x^2)}}\,\mathrm{d}x$$
$$= 2\int_0^\infty \frac{\mathrm{d}x}{\sqrt{(p^2+x^2)(q^2+x^2)}} = 2\,\mathrm{I}(p,q).$$

直接解读同一个积分。 由于 $\frac{1}{4}(p+q)^2 = a^2$ 及 $pq = g^2$，J 的被积函数（integrand）是 $1/\sqrt{(a^2+t^2)(g^2+t^2)}$，它是 t 的偶函数（even function）。所以

$$J = 2\int_0^\infty \frac{\mathrm{d}t}{\sqrt{(a^2+t^2)(g^2+t^2)}} = 2\,\mathrm{I}(a,g),$$

这样做是合理的，因为 a、g 均为正数且 $g < a$（由于 $p \neq q$，这由第 (i) 部分的不等式得出）。比较 J 的两个表达式，得

$$\mathrm{I}(p,q) = \mathrm{I}\big(\mathrm{a}(p,q),\mathrm{g}(p,q)\big),$$

此即所求。

计算 $\mathrm{I}(x_0,y_0)$。 取 $(p,q) = (x_n,y_n)$ 应用上式（由于 $0 < y_n < x_n$，这是容许的），得 $\mathrm{I}(x_n,y_n) = \mathrm{I}(x_{n+1},y_{n+1})$，所以

$$\mathrm{I}(x_0,y_0) = \mathrm{I}(x_n,y_n) \quad \text{对每个} n.$$

> **想一想** 我们想令 $n\to\infty$，并说积分趋于 $\mathrm{I}(M,M)$。在无穷区间上的积分，怎样才能合理地把极限移到积分号之内?

我们用夹逼的方法，避开任何微妙的极限论证。对 $0 < y_n < x_n$ 及每个 x，

$$(y_n^2+x^2)^2 \le (x_n^2+x^2)(y_n^2+x^2) \le (x_n^2+x^2)^2,$$

所以取平方根再取倒数（这会令不等号反向），然后积分，得

$$\int_0^\infty \frac{\mathrm{d}x}{x_n^2+x^2} \le \mathrm{I}(x_n,y_n) \le \int_0^\infty \frac{\mathrm{d}x}{y_n^2+x^2}.$$

对 $c>0$，$\displaystyle\int_0^\infty \frac{\mathrm{d}x}{c^2+x^2} = \Big[\frac{1}{c}\arctan\frac{x}{c}\Big]_0^\infty = \frac{\pi}{2c}$，所以

$$\frac{\pi}{2x_n} \le \mathrm{I}(x_0,y_0) \le \frac{\pi}{2y_n} \quad \text{对每个} n.$$

中间一项不随 n 改变，而由于 $x_n\to M$、$y_n\to M$ 且 $M>0$，上下两界都趋于 $\pi/(2M)$。因此

$$\mathrm{I}(x_0,y_0) = \frac{\pi}{2M}.$$

验算

- 取 $x_0 = 2$、$y_0 = 1$。迭代得 $(1.5,\ 1.41421)$，然后是 $(1.45711,\ 1.45648)$，再下一步两者在五位有效数字内一致，即 $M \approx 1.45679$。所以 $\pi/(2M) \approx 1.07826$，而用数值方法计算 $\mathrm{I}(2,1)$ 亦得 1.07826。
- 若 $p = q$，则恒等式显然成立，因为此时 $\mathrm{a}(p,p) = \mathrm{g}(p,p) = p$。答案 $\pi/(2M)$ 与此吻合，因为 $\mathrm{I}(M,M) = \pi/(2M)$。
- 收敛速度极快：最初几步之后，两者之差大约每次平方一次，远比我们证明的减半要好。

常见错误

- 在归纳法中使用 $y_{n+1} < x_{n+1}$，却没有先确立 $y_n < x_n$（以及 $y_n > 0$，使平方根有意义）。
- 忘记关于 t 的积分是在整条实数线上进行，因此等于 $\mathrm{I}(a,g)$ 的*两倍*；两个因子 2 必须互相抵消。
- 毫无理据地写出 $\mathrm{I}(x_n,y_n) \to \mathrm{I}(M,M)$。上面的夹逼论证提供了一个可靠的理由。

STEP 之外

共同极限 M 就是高斯（Gauss）的算术—几何平均数，而积分 I 是一个*完全椭圆积分*（complete elliptic integral）：高斯在 1799 年发现了这个恒等式，它提供了一种快得惊人的方法来计算椭圆积分，甚至可以计算 π 本身（高斯—勒让德算法（Gauss–Legendre algorithm））。开头引用的单调收敛定理（monotone convergence theorem），以及交换极限与积分次序时所需的谨慎，都是大学第一门分析（analysis）课程的核心主题；Abbott 的 *Understanding Analysis*（《理解分析》）是正式认识这些内容的绝佳起点。

详细解答

STEP II 2024，第 9 题

力学 • 抛体运动

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

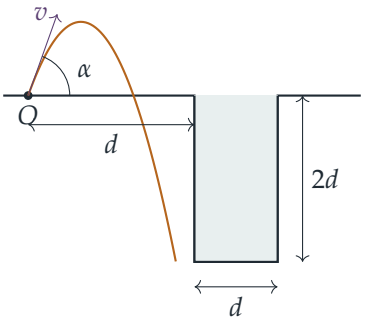
www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

一切都由轨迹方程（trajectory equation）$y = x\tan\alpha - \dfrac{gx^2}{2v^2}\sec^2\alpha$ 推出，把它写成以 $t = \tan\alpha$ 为变量的二次式（quadratic）。要干净利落地落在沟底，意味着两件事：质点到达近边时仍在地面之上；而在它本应到达对面一侧之前，已经低于沟底的水平。每个条件都是关于 t 的二次不等式（quadratic inequality），而配方（completing the square）可把它们化成所需的形式。第 (ii) 部分则只需证明 t 的两个容许范围有重叠。

解答

把质点视为只受重力作用的点状抛射体（projectile），不计空气阻力（air resistance）。取抛射点为原点，x 沿水平方向指向沟渠，y 竖直向上。于是近壁是直线 $x = d$，远壁是 $x = 2d$，而沟底是 $y = -2d$，其中 $d < x < 2d$。



在时刻 T，水平速度为 $v\cos\alpha$，竖直速度为 $v\sin\alpha - gT$，消去 T 得

$$y = x\tan\alpha - \frac{gx^2}{2v^2\cos^2\alpha} = xt - \frac{x^2(1+t^2)}{2\lambda d},$$

其中 $t = \tan\alpha$，并用了 $v^2 = \lambda dg$ 及 $\sec^2\alpha = 1 + t^2$。

第 (i)(a) 部分

要避免在沟渠前撞到地面或撞到近壁，质点在 $x = d$ 时必须仍严格高于地面。（若它在该处位于 $y = 0$ 或以下，它便已撞到地面，或将撞上近壁。）所以

$$y(d) = dt - \frac{d(1+t^2)}{2\lambda} > 0 \iff 2\lambda t - 1 - t^2 > 0 \iff t^2 - 2\lambda t + \lambda^2 < \lambda^2 - 1,$$

这里乘以了 $2\lambda/d > 0$。即

$$(\tan\alpha - \lambda)^2 < \lambda^2 - 1,$$

此即所求。左边是平方，所以 $\lambda^2 - 1 > 0$；由于 $\lambda = v^2/(dg) > 0$，得 $\lambda > 1$。

第 (i)(b) 部分

现在考虑远壁。路径是一条开口向下的抛物线（parabola），在 $x = d$ 时高于沟底水平 $-2d$。它在 $x = 2d$ 之前到达 $y = -2d$（即沟底），当且仅当它在 $x = 2d$ 时已低于该水平，即 $y(2d) < -2d$。反之若 $y(2d) \geq -2d$，它到达 $x = 2d$ 时仍未触及沟底，因此会碰到远壁（或者整条沟渠都飞越过去）。

> ? **想一想** 为甚么只看 $x = 2d$ 这一点就足够?

因为 $y(x) + 2d$ 是关于 x 的二次式，首项系数（leading coefficient）为负，且在 $x = d$ 时为正。它在 $x = d$ 之后恰有一个根（root），而这个根位于 $2d$ 之前，当且仅当该二次式在 $x = 2d$ 时已经为负。

所以我们需要

$$y(2d) = 2dt - \frac{4d^2(1+t^2)}{2\lambda d} < -2d \iff \lambda t - (1+t^2) < -\lambda \iff t^2 - \lambda t + 1 - \lambda > 0,$$

这里乘以了 $\lambda/(2d) > 0$。乘以 4 再配方，

$$4t^2 - 4\lambda t + \lambda^2 > \lambda^2 - 4 + 4\lambda, \qquad \text{即} \qquad (2\tan\alpha - \lambda)^2 > \lambda^2 + 4(\lambda - 1),$$

此即所求。

推出 $\alpha > 45°$。由 (a)，$\lambda > 1$，所以 $\lambda^2 + 4(\lambda - 1) > \lambda^2 > 0$。记 $s = \sqrt{\lambda^2 + 4\lambda - 4}$，则 $s > \lambda$。(b) 中的不等式表示

$$2t > \lambda + s \qquad \text{或} \qquad 2t < \lambda - s.$$

由于 $s > \lambda$，第二种情况给出 $t < 0$。但 (a) 迫使 $t > 0$：由 $(t-\lambda)^2 < \lambda^2 - 1 < \lambda^2$ 得 $|t-\lambda| < \lambda$，所以 $t > 0$。（从物理上看，质点必须上升才能越过近边。）因此第一种情况成立，并且

$$2t > \lambda + s > \lambda + \lambda = 2\lambda > 2,$$

所以 $\tan\alpha > \lambda > 1$。由于 $0 < \alpha < 90°$，得 $\alpha > 45°$。

第 (ii) 部分

反过来，两个条件合起来已经足够。若 $y(d) > 0$，由于 $y(0) = 0$ 且 y 是凹函数（concave），在整个区间 $0 < x \le d$ 上都有 $y > 0$，所以质点越过地面和近壁。若同时 $y(2d) < -2d$，它会在严格介乎 d 与 $2d$ 之间的某个 x 到达沟底，而在此之前一直留在沟渠之内。所以我们只需证明：对每个 $\lambda > 1$，存在某个 $t > 0$ 同时满足两个不等式。

由 (a)，容许值为 $\lambda - \sqrt{\lambda^2-1} < t < \lambda + \sqrt{\lambda^2-1}$。由 (b) 连同 $t > 0$，容许值为 $t > \frac{1}{2}(\lambda + s)$。两者对某个 t 同时成立，只要

$$\tfrac{1}{2}(\lambda + s) < \lambda + \sqrt{\lambda^2-1} \iff s < \lambda + 2\sqrt{\lambda^2-1}.$$

（(a) 的下界不会造成麻烦，因为 $\frac{1}{2}(\lambda+s) > \lambda > \lambda - \sqrt{\lambda^2-1}$。）两边皆为正，所以可以平方：

$$\begin{aligned} & \lambda^2 + 4\lambda - 4 < \lambda^2 + 4\lambda\sqrt{\lambda^2-1} + 4(\lambda^2-1) \\ \iff \quad & 4\lambda < 4\lambda\sqrt{\lambda^2-1} + 4\lambda^2 \\ \iff \quad & 1 < \sqrt{\lambda^2-1} + \lambda, \end{aligned}$$

这里除以了 $4\lambda > 0$。由于 $\lambda > 1$，最后一行成立。所以区间

$$\tfrac{1}{2}\left(\lambda + \sqrt{\lambda^2 + 4\lambda - 4}\right) < \tan\alpha < \lambda + \sqrt{\lambda^2-1}$$

是非空的（non-empty）。选取 α 使 $\tan\alpha$ 落在其中（这是可行的，因为 tan 在 $(0°, 90°)$ 上取遍所有正值），质点便会落在沟底而不碰到任何一侧。

验算

- 取 $\lambda = 2$。区间为 $1 + \sqrt{2} < \tan\alpha < 2 + \sqrt{3}$，恰好是 $67.5° < \alpha < 75°$。把 $\alpha = 70°$ 代入 $y = xt - x^2(1+t^2)/(4d)$，得 $y(d) \approx 0.61d > 0$ 及 $y(2d) \approx -3.05d < -2d$，与预期相符。
- 当 $\lambda \to 1^+$ 时，区间两端都趋于 1，所以可行范围收缩至 $\alpha = 45°$：当 $v^2 = dg$ 时，质点只能勉强到达近边，这需要最大射程（maximum range）的角度。
- 用电脑随机检验数以千计的 (λ, α) 组合，按时间逐步推进运动，结果在每一种情况下都与两个不等式吻合。

常见错误

- 把条件当成 $y(d) > -2d$（高于沟底）而非 $y(d) > 0$：近壁一直延伸到地面高度。
- 在 (b) 中忘记排除 $2\tan\alpha < \lambda - s$ 这一支，而这需要用到 $\tan\alpha > 0$。
- 在 (ii) 中未先指出两边皆为正便把不等式平方。

STEP 之外

这里一个物理条件（“落在沟渠内”）变成了发射参数空间中由不等式所围出的一个区域，而问题在于该区域是否非空。这正是最佳化（optimisation）与控制（control）的起点：描述可行（feasible）输入的集合，并找出其边界。Kleppner 与 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）以大学一年级程度讨论抛射运动及其延伸（空气阻力、旋转参考系）。

详细解答

STEP II 2024，第 10 题

力学 · 加速楔上的摩擦力

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

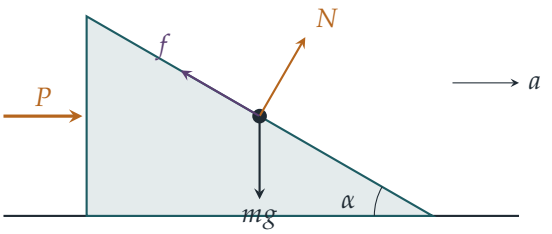
www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

整道题建基于两个观察。第一，水平面光滑（smooth），所以作用于整个系统（棱柱连同质点）的唯一水平外力是 P，由此可立即得出共同加速度。第二，由于质点相对棱柱不动，它的水平加速度与棱柱完全相同，因此只需对质点沿斜面及垂直斜面分解力，便可直接求得摩擦力和法向反作用力（normal reaction）。第 (iii) 部分的不等式其实就是“摩擦力 $\leq \mu\times$ 法向反作用力”，再用复角公式（compound-angle formulae）整理一下而已。

解答

我们把质点视为质点（particle），把棱柱（prism）视为在光滑水平面上滑动的刚体（rigid body），并设质点与斜面之间的接触是粗糙（rough）的，摩擦系数（coefficient of friction）为 μ。把竖直面放在左边，使斜面向右下倾斜；这样 P 便把棱柱推向右方。



作用于质点的力有：重量 mg、垂直于斜面的法向反作用力 N，以及摩擦力。我们把摩擦力视为一个带正负号的量 f，以沿斜面向上为正（若 f 结果为负，则摩擦力实际上沿斜面向下）。摩擦力的大小为 $F = |f|$。

第 (i) 部分

假设系统处于平衡（equilibrium），即所有物体都静止。对质点沿垂直斜面及沿斜面方向分解力，得

$$N = mg\cos\alpha, \qquad f = mg\sin\alpha.$$

质点不滑动的条件是 $|f| \le \mu N$，即 $mg\sin\alpha \le \mu\, mg\cos\alpha$，亦即 $\mu \ge \tan\alpha$（由于 α 是锐角，可以除以 $mg\cos\alpha > 0$）。因此若 $\mu < \tan\alpha$，便不可能平衡：质点必定滑动。

第 (ii) 部分

加速度。把棱柱和质点合起来考虑。水平面光滑，只施加竖直方向的力；质点与棱柱之间的力属于内力。因此唯一的水平外力是 P，而两者的加速度同为 a（水平向右），故

$$P = (M+m)a, \qquad a = \frac{P}{M+m}.$$

对质点分解力。质点的加速度是水平的 a。它沿斜面向上的分量是 $-a\cos\alpha$（斜面向左上升，而质点向右加速），沿向外法线方向的分量是 $a\sin\alpha$。

> ? **想一想** 为甚么要沿斜面及垂直斜面分解，而不是沿水平及竖直方向分解?

沿斜面方向，只有 f 和重量有分量；垂直斜面方向，只有 N 和重量有分量。所以每条方程只含一个未知量，免去解联立方程。代价是加速度也要分解，但这并不困难。

沿斜面向上：

$$f - mg\sin\alpha = -ma\cos\alpha \quad\Longrightarrow\quad f = m(g\sin\alpha - a\cos\alpha).$$

代入 $a = P/(M+m)$，得

$$f = \frac{m}{M+m}[(M+m)g\sin\alpha - P\cos\alpha],$$

因此摩擦力的大小为

$$F = \frac{m}{M+m}\,|(M+m)g\sin\alpha - P\cos\alpha|,$$

与题目所求相符。这里的绝对值确实不可少：P 较小时摩擦力沿斜面向上；但 P 较大时，斜面推向质点的力足够大，摩擦力必须沿斜面向下，才能阻止质点沿斜面上移。

垂直斜面方向（向外法线）：

$$N - mg\cos\alpha = ma\sin\alpha \quad\Longrightarrow\quad N = \frac{m}{M+m}[(M+m)g\cos\alpha + P\sin\alpha].$$

由于 $P > 0$ 且 α 是锐角，此式为正，所以不需要绝对值，而质点亦保持与斜面接触。

第 (iii) 部分

质点相对棱柱保持静止，当且仅当摩擦力足以提供所需：$F \le \mu N$。设 $G = (M+m)g$，并约去正因子 $m/(M+m)$：

$$|G\sin\alpha - P\cos\alpha| \le \tan\lambda\,(G\cos\alpha + P\sin\alpha).$$

乘以 $\cos\lambda > 0$，这等价于以下一对不等式（inequalities）：

$$-\sin\lambda\,(G\cos\alpha + P\sin\alpha) \;\le\; \cos\lambda\,(G\sin\alpha - P\cos\alpha) \;\le\; \sin\lambda\,(G\cos\alpha + P\sin\alpha).$$

? **想一想** 这些不等式如何变成 $\alpha \pm \lambda$ 的正切?

把含 G 的项移到一边，含 P 的项移到另一边；这样每个括号都是一个复角展开式。

右边的不等式（沿斜面向上的摩擦力不可超过其极限）：

$$G(\sin\alpha\cos\lambda - \cos\alpha\sin\lambda) \le P(\cos\alpha\cos\lambda + \sin\alpha\sin\lambda),$$
$$G\sin(\alpha-\lambda) \le P\cos(\alpha-\lambda).$$

由于 $0 < \alpha - \lambda < \frac{1}{4}\pi$，有 $\cos(\alpha-\lambda) > 0$，所以此式即 $P \ge G\tan(\alpha-\lambda)$。

左边的不等式（沿斜面向下的摩擦力不可超过其极限）：

$$P(\cos\alpha\cos\lambda - \sin\alpha\sin\lambda) \le G(\sin\alpha\cos\lambda + \cos\alpha\sin\lambda),$$
$$P\cos(\alpha+\lambda) \le G\sin(\alpha+\lambda).$$

这里要用到条件 $\lambda < \alpha < \frac{1}{4}\pi$：它给出 $0 < \alpha+\lambda < \frac{1}{2}\pi$，所以 $\cos(\alpha+\lambda) > 0$，可以相除得 $P \le G\tan(\alpha+\lambda)$。

两式必须同时成立，故

$$(M+m)g\tan(\alpha-\lambda) \le P \le (M+m)g\tan(\alpha+\lambda),$$

与题目所求相符。

验算

- 当 $P = 0$ 时，下界给出 $0 \ge G\tan(\alpha-\lambda)$，由于 $\lambda < \alpha$，此式不成立：这与第 (i) 部分一致，因为 $\mu = \tan\lambda < \tan\alpha$。
- 当 $P = G\tan\alpha$ 时得 $f = 0$：此时的推力恰好令光滑质点能静止不动，这正是熟悉的“加速楔子”结果 $a = g\tan\alpha$。这个值理应落在区间之内，事实亦然。

- 数值上，取 $M = 3$、$m = 1$、$g = 9.8$、$\alpha = 30°$、$\lambda = 10°$，容许范围是 $14.27 \le P \le 32.89$（牛顿）。当 $P = 20$ 时，公式给出 $F \approx 0.570$ 及 $N \approx 10.99$，所以 $F/N \approx 0.052 < \tan 10° \approx 0.176$。用电脑对数以千计的随机情况直接解质点的方程，结果与两条公式及该区间都吻合。

常见错误

- 在第 (ii) 部分取 $N = mg\cos\alpha$：斜面正在加速，所以法向反作用力并非静止时的数值。
- 漏掉绝对值，结果只求得 P 的下界；上界来自摩擦力沿斜面向下的情况。
- 除以 $\cos(\alpha + \lambda)$ 时没有指出它是正数；这正是用到 $\alpha < \frac{1}{4}\pi$ 的地方。

STEP 之外

更简洁的看法是在棱柱的参考系中处理问题，该参考系以 a 加速：在其中，质点会受到一个向后的额外“虚拟力”（fictitious force）ma，问题就变成在粗糙斜面上的普通静力学，只是等效重力倾侧了 $\arctan(a/g)$ 这个角度。此时条件就是：这个倾侧角与 α 的差必须在摩擦角（angle of friction）λ 之内，这正好就是所求得的区间。非惯性参考系（non-inertial frames）与虚拟力在 Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）中有详细阐述。

详细解答

STEP II 2024，第 11 题

概率与统计 · 分组检测的期望值

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

只有当混合样本的检测通常呈阴性时，合并检测（pooling）才划算；而每人的期望（expected）检测次数原来是 $1+\frac{1}{k}-(1-p)^k$。要问这个数何时不超过 1，就变成比较 $(1-p)$ 与 $k^{-1/k}$ 的大小，这正是第 (i) 部分要讨论 $x^{1/x}$ 的原因：当 p 较大时，最佳的组别人数就是使 $k^{1/k}$ 最大的那个整数。

解答

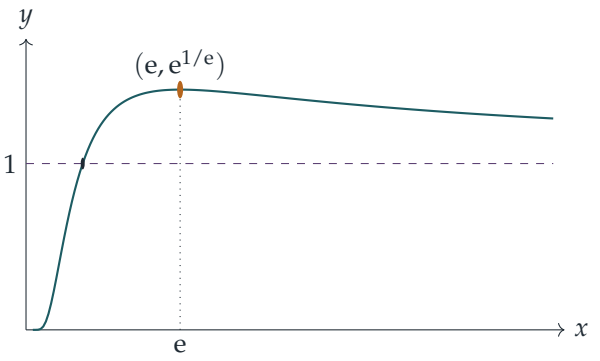
第 (i) 部分

对 $x>0$，记 $y=x^{1/x}=\mathrm{e}^{(\ln x)/x}$，于是 $\ln y=\dfrac{\ln x}{x}$。求导（differentiating）得

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1-\ln x}{x^2}.$$

由于 $y>0$，$\frac{\mathrm{d}y}{\mathrm{d}x}$ 与 $1-\ln x$ 同号：当 $0<x<\mathrm{e}$ 时为正，在 $x=\mathrm{e}$ 时为零，当 $x>\mathrm{e}$ 时为负。因此曲线只有一个转折点（turning point），即位于 $(\mathrm{e},\mathrm{e}^{1/\mathrm{e}})\approx(2.72,\,1.44)$ 的极大值（maximum）点。

至于图像两端：当 $x\to 0^+$ 时，$(\ln x)/x\to-\infty$，所以 $y\to 0$；当 $x\to\infty$ 时，$(\ln x)/x\to 0$，所以 $y\to 1$，且从上方趋近（因为对 $x>1$ 有 $\ln y>0$）。曲线经过点 $(1,1)$。



对正整数 n 而言，此函数在 $(0, \mathrm{e}]$ 上递增，在 $[\mathrm{e}, \infty)$ 上递减，所以 $1^{1/1} < 2^{1/2}$ 且 $3^{1/3} > 4^{1/4} > 5^{1/5} > \cdots$。因此最大值是 $2^{1/2}$ 或 $3^{1/3}$。把两者都升到六次方，得 $2^3 = 8$ 和 $3^2 = 9$，所以 $3^{1/3} > 2^{1/2}$，且

$$\max_n n^{1/n} = 3^{1/3}.$$

第 (ii) 部分

记 $q = 1 - p$。考虑一组 k 人。由独立性（independence），组内无人带有该酶的概率是 q^k，而只有在这种情况下混合样本才呈阴性。所以该组所用的检测次数为

$$\begin{cases} 1 \text{ 次检测} & \text{概率为} q^k, \\ 1 + k \text{ 次检测} & \text{概率为} 1 - q^k, \end{cases}$$

其期望检测次数是 $1 + k(1 - q^k)$。共有 $r = N/k$ 组，而期望值具可加性，所以期望总次数为

$$E = \frac{N}{k}\left(1 + k(1 - q^k)\right) = N\left(1 + \frac{1}{k} - (1 - p)^k\right).$$

第 (iii) 部分

由第 (ii) 部分，$E \le N$ 当且仅当

$$\frac{1}{k} \le (1 - p)^k \iff 1 - p \ge k^{-1/k} = \frac{1}{k^{1/k}},$$

这里取了正的 k 次方根（两边均为正数，而 $t \mapsto t^{1/k}$ 是递增的）。

> ? **想一想** 哪个 k 容许最大的 p?

对给定的 k，条件是 $p \le 1 - k^{-1/k}$，而当 $k^{1/k}$ 最大时，这个上界也最大。由第 (i) 部分，对每个整数 $k > 1$ 都有 $k^{1/k} \le 3^{1/3}$，所以无论选哪个 k，都必须有 $p \le 1 - 3^{-1/3}$。反过来，

N 是 3 的倍数，所以 $k = 3$ 是容许的组别人数，而取 $k = 3$ 时条件正好是 $p \le 1 - 3^{-1/3}$（等号成立时 $E = N$）。因此最大值为

$$p = 1 - 3^{-1/3} \approx 0.307.$$

要证明这个值大于 $\frac{1}{4}$：

$$1 - 3^{-1/3} > \tfrac{1}{4} \iff 3^{-1/3} < \tfrac{3}{4} \iff \tfrac{1}{3} < \tfrac{27}{64} \iff 64 < 81,$$

其中中间一步两边取立方（对正数而言，取立方保持大小次序）。最后一个式子成立，所以 $1 - 3^{-1/3} > \frac{1}{4}$。

第 (iv) 部分

由二项式展开（binomial expansion），

$$(1-p)^k = 1 - kp + \binom{k}{2}p^2 - \cdots.$$

若 pk 很小，则 $\binom{k}{2}p^2 < \frac{1}{2}(pk)^2$ 更小得多，之后各项亦然，所以 $(1-p)^k \approx 1 - pk$。代入第 (ii) 部分的结果，

$$E \approx N\left(1 + \frac{1}{k} - (1 - pk)\right) = N\left(\frac{1}{k} + pk\right).$$

（事实上，伯努利不等式（Bernoulli's inequality）$(1-p)^k \ge 1 - pk$ 表明这个近似永远不会低估 E；而把展开式多保留一项，则可证明误差最多为 $\frac{1}{2}(pk)^2 N$。）

取 $p = 0.01$ 和 $k = 10$，得 $pk = 0.1$，确实很小，而

$$E \approx N\left(\frac{1}{10} + 0.1\right) = 0.2N,$$

所以这个方案只需约 20% 的检测次数，相比之下逐一检测每个人需要 N 次。

验算

- 取 $p = 0.01$ 和 $k = 10$ 作精确计算：$E/N = 1.1 - 0.99^{10} \approx 1.1 - 0.9044 = 0.1956$，约为 19.6%，与近似值之差在误差上界 $\frac{1}{2}(0.1)^2 = 0.005$ 之内。
- 当 $p = 1 - 3^{-1/3}$ 且 $k = 3$ 时：$(1-p)^3 = \frac{1}{3}$，所以 $E = N(1 + \frac{1}{3} - \frac{1}{3}) = N$，刚好落在边界上。
- 检查第 (ii) 部分是否合理：当 $p \to 0$ 时，$E \to N/k$（每组一次检测）；当 $p \to 1$ 时，$E \to N(1 + \frac{1}{k})$（每组都要重新逐一检测）。此公式亦已用电脑逐一列举小组别的所有结果加以验证。

常见错误

- 第 (i) 部分答 $e^{1/e}$：n 必须是整数，所以应比较 $2^{1/2}$ 与 $3^{1/3}$。
- 第 (iii) 部分只验证了 $k=3$，却没有解释为何其他组别人数不可能容许更大的 p。
- 在呈阳性的组别中漏计混合检测本身，得出 k 次而非 $k+1$ 次检测。

STEP 之外

这就是 Dorfman 的群组检测（group testing）方案，它在第二次世界大战期间为筛查士兵的血液样本而设计，后来在 COVID-19 检测中再度被采用。把 $\frac{1}{k}+pk$ 最小化，得 $k\approx 1/\sqrt{p}$（所以这里 $k=10$ 是最佳选择），而所需检测次数约为 $2\sqrt{p}\,N$；信息理论（information theory）中研究的更巧妙的自适应（adaptive）方案，效果还要更好。关于背后的概率理论，可参阅 Blitzstein 与 Hwang 的 *Introduction to Probability*（《概率导论》）。

详细解答

STEP II 2024，第 12 题

概率与统计 • 离散顺序统计量

STEP II 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

在第一个游戏中，胜出者是持有唯一最小数字的玩家，而这样的玩家最多只有一个。因此“有胜出者”可以按照谁胜出、以哪个数字胜出，分拆成互斥（disjoint）的情况，而每一种情况都只是若干独立（independent）概率的简单乘积。题目给出 $\sum i^2$ 和 $\sum i^3$ 的公式，正是提示所得的和是多项式求和。第二个游戏的做法相同，只是现在要考虑一对有序的玩家（一个最低、一个最高），并要对两个数字求和。

解答

在整个解答中，k 个数字互相独立，且每个都在 $\{1,2,\ldots,n\}$ 上均匀分布（uniform），因此任何一位玩家得到任何一个特定数字的概率都是 $\frac{1}{n}$，而得到某个含 s 个值的集合中的数字的概率是 $\frac{s}{n}$。

第 (i) 部分

固定 a，其中 $1 \le a \le n-1$。Ada 必须得到 a（概率为 $\frac{1}{n}$），而其余 $k-1$ 位玩家每人都必须得到 $a+1,\ldots,n$ 这 $n-a$ 个数字之一（每人的概率为 $\frac{n-a}{n}$）。由独立性，

$$\mathrm{P}(\text{Ada 得到}a\text{, 其他人全部较大}) = \frac{1}{n}\left(\frac{n-a}{n}\right)^{k-1} = \frac{(n-a)^{k-1}}{n^k}.$$

（当 $a = n$ 时此概率为零，这就是 a 只取到 $n-1$ 的原因。）

设 W 为有胜出者的事件。胜出者所持的数字严格小于其他所有人的数字，所以不可能有两个胜出者：否则两者都必须比对方小。因此 W 是以下事件的互斥并集：对 k 位玩家及 $a = 1,\ldots,n-1$，事件“这位玩家得到 a，而其他人都得到较大的数字”。由对称性

(symmetry)，每位玩家的贡献都与 Ada 相同，所以

$$\mathrm{P}(W) = k\sum_{a=1}^{n-1}\frac{(n-a)^{k-1}}{n^k} = \frac{k}{n^k}\sum_{j=1}^{n-1} j^{k-1},$$

其中我们代入了 $j = n-a$，它同样由 1 取到 $n-1$。

当 $k=4$ 时，利用题目给出的 $\sum i^3$ 公式，并以 $n-1$ 代替 n，

$$\mathrm{P}(W) = \frac{4}{n^4}\cdot\frac{1}{4}(n-1)^2n^2 = \frac{(n-1)^2}{n^2},$$

即为所求。

第 (ii) 部分

固定 d 和 a，其中 $1 \le d \le n-2$ 且 $1 \le a \le n-d-1$；这些条件保证 $a+d+1 \le n$，并且 a 与 $a+d+1$ 之间至少严格夹着一个数字。Ada 必须得到 a，Bob 必须得到 $a+d+1$（各自的概率为 $\frac{1}{n}$），而其余 $k-2$ 位玩家每人都必须得到 $a+1,\ldots,a+d$ 这 d 个数字之一。由独立性，

$$\mathrm{P}(\text{Ada 得到}a,\ \text{Bob 得到}a+d+1,\ \text{其他人严格介乎两者之间}) = \frac{1}{n^2}\left(\frac{d}{n}\right)^{k-2} = \frac{d^{k-2}}{n^k}.$$

想一想 为甚么在对玩家以及对 a 和 d 求和时，每一种出现两个胜出者的方式都恰好出现一次?

有两个胜出者，即是一位玩家持有唯一最小的数字，而另一位玩家持有唯一最大的数字。与第 (i) 部分一样，低位胜出者是唯一的，高位胜出者亦然，因此结果决定了一对有序的不同玩家（低位胜出者，高位胜出者），连同较小的数字 a 以及到较大数字的差距 $d \ge 1$：差距必须至少为 1，因为其余 $k-2 \ge 2$ 位玩家的数字都严格介乎两者之间。这些选择不同，所得的事件就互斥。有序玩家对共有 $k(k-1)$ 对，每一对的贡献都与（Ada，Bob）相同，所以

$$\mathrm{P}(\text{两个胜出者}) = k(k-1)\sum_{d=1}^{n-2}\sum_{a=1}^{n-d-1}\frac{d^{k-2}}{n^k} = \frac{k(k-1)}{n^k}\sum_{d=1}^{n-2}(n-1-d)\,d^{k-2},$$

因为内层的和有 $n-d-1$ 个相等的项。

现在取 $k=4$，并记 $m = n-1$，于是 d 由 1 取到 $m-1$。利用题目给出的两条公式，并以 $m-1$ 代替 n，

$$\begin{aligned}\sum_{d=1}^{m-1}(m-d)\,d^2 &= m\cdot\tfrac{1}{6}(m-1)m(2m-1) - \tfrac{1}{4}(m-1)^2m^2\\ &= \tfrac{1}{12}m^2(m-1)\big[2(2m-1)-3(m-1)\big]\\ &= \tfrac{1}{12}m^2(m-1)(m+1).\end{aligned}$$

因此，由于 $m-1=n-2$ 及 $m+1=n$，

$$\mathrm{P}(\text{两个胜出者})=\frac{12}{n^4}\cdot\frac{1}{12}(n-1)^2(n-2)n=\frac{(n-2)(n-1)^2}{n^3},$$

这正是所求的结果。

当 $k=4$ 时比较一个与两个胜出者

设 L 为有玩家持有唯一最小数字的事件，H 为有玩家持有唯一最大数字的事件。第 (i) 部分给出 $\mathrm{P}(L)=\frac{(n-1)^2}{n^2}$。把每个数字 x 换成 $n+1-x$，会把均匀分布的数字变成均匀分布的数字，并把“最小”与“最大”互换，所以同样有 $\mathrm{P}(H)=\mathrm{P}(L)$。

恰好两个胜出者就是事件 $L\cap H$，因此记 $T=\mathrm{P}(L\cap H)=\frac{(n-2)(n-1)^2}{n^3}$。恰好一个胜出者，即是 L 或 H 发生但不同时发生，所以

$$\mathrm{P}(\text{恰好一个})=\mathrm{P}(L)+\mathrm{P}(H)-2\,\mathrm{P}(L\cap H)=\frac{2(n-1)^2}{n^2}-2T.$$

两个胜出者比一个胜出者更有可能，当且仅当

$$\begin{aligned}T>\frac{2(n-1)^2}{n^2}-2T &\iff 3\cdot\frac{(n-2)(n-1)^2}{n^3}>\frac{2(n-1)^2}{n^2}\\ &\iff 3(n-2)>2n \iff n>6,\end{aligned}$$

其中我们除以正数 $\frac{(n-1)^2}{n^3}$（因为 $n\geq 3$，所以它是正数）。当 $n=6$ 时两个概率相等，因此最小值是

$$n=7.$$

验算

- 当 $n=3$ 及 $k=4$ 时共有 81 个结果。以 1 胜出需要其余三个数字都在 $\{2,3\}$ 中：共 $4\cdot 2^3=32$ 个结果。以 2 胜出需要其余三个都是 3：共 4 个结果。所以 $\mathrm{P}(W)=\frac{36}{81}=\frac{4}{9}=\frac{(3-1)^2}{3^2}$。
- 当 $n=6$ 时，“恰好一个”和“恰好两个”的概率都是 $\frac{25}{54}\approx 0.463$；当 $n=7$ 时，两者分别是 $\frac{144}{343}\approx 0.420$ 和 $\frac{180}{343}\approx 0.525$。这些数值，以及两部分中在 $n\leq 8$ 和 $k=2,3,4$ 时的公式，都已经用电脑逐一枚举（enumeration）所有结果加以核实。
- 当 $n\to\infty$ 时，数字相同的情况变得罕见，所以有四位玩家时，几乎总会同时有唯一最低和唯一最高的数字：的确 $\frac{(n-2)(n-1)^2}{n^3}\to 1$。

常见错误

- 在第 (i) 部分漏掉因子 k，或在第 (ii) 部分漏掉因子 $k(k-1)$。在第 (ii) 部分中，两位胜出者是有序的，因为一位是低位、一位是高位。
- 把“恰好一个胜出者”算成 $\mathrm{P}(L)+\mathrm{P}(H)-\mathrm{P}(L\cap H)$，但这其实是“至少一个”；重叠部分必须减去两次。
- 答 $n=6$：此时两个概率相等，并非严格的大小关系。

STEP 之外

把一个事件按“谁”和“在哪里”分拆成互斥的部分，是求顺序统计量（order statistics）——即样本中最小值和最大值——的分布的离散版本。对于连续样本，同样的论证会给出最小值与最大值的联合密度（joint density），而令本题变得有趣的数字相同情况亦会消失。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）是认识这些概念的好起点。

详细解答

STEP III 2024，第 1 题

纯数学 · 裂项相消级数

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

这里每一个和都用同一方法处理：先把每项拆成部分分式（partial fractions），找出其中属于 $1/r^2$ 倍数的部分（它们加起来是 S_N 的倍数），再把其余部分整理成 $f(r)-f(r+1)$ 形式的差，这些差会裂项相消（telescope）。全程只处理到 N 为止的有限和，到最后每一部分都有已知极限时，才让 $N\to\infty$。

解答

第 (i) 部分

以 $r^2(r+1)$ 为公分母，

$$\frac{1}{r+1}-\frac{1}{r}+\frac{1}{r^2}=\frac{r^2-r(r+1)+(r+1)}{r^2(r+1)}=\frac{r^2-r^2-r+r+1}{r^2(r+1)}=\frac{1}{r^2(r+1)},$$

得证。

现在由 $r=1$ 加到 N。各 $1/r^2$ 项之和为 S_N，其余两项则裂项相消：

$$\sum_{r=1}^{N}\left(\frac{1}{r+1}-\frac{1}{r}\right)=\left(\tfrac{1}{2}-1\right)+\left(\tfrac{1}{3}-\tfrac{1}{2}\right)+\cdots+\left(\tfrac{1}{N+1}-\tfrac{1}{N}\right)=\frac{1}{N+1}-1.$$

因此

$$\sum_{r=1}^{N}\frac{1}{r^2(r+1)}=S_N-1+\frac{1}{N+1}.$$

当 $N\to\infty$ 时，$S_N\to\frac{1}{6}\pi^2$ 而 $\frac{1}{N+1}\to 0$，所以右方趋于 $\frac{1}{6}\pi^2-1$。按定义，无穷和就是部

分和（partial sums）的极限，所以

$$\sum_{r=1}^{\infty} \frac{1}{r^2(r+1)} = \tfrac{1}{6}\pi^2 - 1.$$

第 (ii) 部分

设

$$\frac{1}{r^2(r+1)(r+2)} = \frac{A}{r} + \frac{B}{r^2} + \frac{C}{r+1} + \frac{D}{r+2}.$$

两边乘以 $r^2(r+1)(r+2)$，并使用遮盖法（cover-up rule）：代入 $r=0$ 得 $B = \frac{1}{2}$；$r=-1$ 得 $C = \frac{1}{(-1)^2\cdot 1} = 1$；$r=-2$ 得 $D = \frac{1}{4\cdot(-1)} = -\frac{1}{4}$。当 r 很大时，左方的数量级为 $1/r^4$，所以右方各 $1/r$ 项必须互相抵消：$A + C + D = 0$，得 $A = -\frac{3}{4}$。所以

$$\frac{1}{r^2(r+1)(r+2)} = \frac{1}{2r^2} - \frac{3}{4r} + \frac{1}{r+1} - \frac{1}{4(r+2)}.$$

? **想一想** $-\frac{3}{4r}$、$\frac{1}{r+1}$、$-\frac{1}{4(r+2)}$ 这三项看来不像单一个可裂项相消的差。怎样才能令它们裂项相消?

把中间的系数拆成 $1 = \frac{3}{4} + \frac{1}{4}$，使每一部分都可与相邻的项配对：

$$-\frac{3}{4r} + \frac{1}{r+1} - \frac{1}{4(r+2)} = -\frac{3}{4}\left(\frac{1}{r} - \frac{1}{r+1}\right) + \frac{1}{4}\left(\frac{1}{r+1} - \frac{1}{r+2}\right).$$

现在每个括号都会裂项相消：

$$\sum_{r=1}^{N}\left(\frac{1}{r} - \frac{1}{r+1}\right) = 1 - \frac{1}{N+1}, \qquad \sum_{r=1}^{N}\left(\frac{1}{r+1} - \frac{1}{r+2}\right) = \frac{1}{2} - \frac{1}{N+2}.$$

因此

$$\begin{aligned}\sum_{r=1}^{N} \frac{1}{r^2(r+1)(r+2)} &= \tfrac{1}{2}S_N - \tfrac{3}{4}\left(1 - \frac{1}{N+1}\right) + \tfrac{1}{4}\left(\frac{1}{2} - \frac{1}{N+2}\right)\\ &= \tfrac{1}{2}S_N - \frac{5}{8} + \frac{3}{4(N+1)} - \frac{1}{4(N+2)}.\end{aligned}$$

令 $N \to \infty$，最后两项趋于 0，而 $S_N \to \frac{1}{6}\pi^2$，所以

$$\sum_{r=1}^{\infty} \frac{1}{r^2(r+1)(r+2)} = \frac{\pi^2}{12} - \frac{5}{8}.$$

第 (iii) 部分

目标式的右方有 $\frac{2}{r^2(r+1)}$，所以比较这两项。利用第 (i) 部分，

$$\frac{2}{r^2(r+1)} = \frac{2}{r+1} - \frac{2}{r} + \frac{2}{r^2}.$$

另一方面，

$$\frac{1}{r^2(r+1)^2}=\left(\frac{1}{r}-\frac{1}{r+1}\right)^2=\frac{1}{r^2}+\frac{1}{(r+1)^2}-\frac{2}{r(r+1)}=\frac{1}{r^2}+\frac{1}{(r+1)^2}-\frac{2}{r}+\frac{2}{r+1},$$

其中两步都用了 $\frac{1}{r}-\frac{1}{r+1}=\frac{1}{r(r+1)}$。相减得

$$\frac{1}{r^2(r+1)^2}-\frac{2}{r^2(r+1)}=\frac{1}{(r+1)^2}-\frac{1}{r^2},$$

即

$$\frac{1}{r^2(r+1)^2}=\frac{2}{r^2(r+1)}-\left(\frac{1}{r^2}-\frac{1}{(r+1)^2}\right).$$

想一想 为甚么不直接把无穷级数逐项相加?

只有在知道每一部分都收敛（converge）之后，重新排列无穷和才是安全的，所以我们先加到 N。括号部分裂项相消：

$$\sum_{r=1}^{N}\frac{1}{r^2(r+1)^2}=\sum_{r=1}^{N}\frac{2}{r^2(r+1)}-\left(1-\frac{1}{(N+1)^2}\right).$$

由第 (i) 部分，当 $N\to\infty$ 时右方第一个和收敛（至 $2(\frac{1}{6}\pi^2-1)$），而 $\frac{1}{(N+1)^2}\to 0$。所以左方亦收敛，并且

$$\sum_{r=1}^{\infty}\frac{1}{r^2(r+1)^2}=\sum_{r=1}^{\infty}\frac{2}{r^2(r+1)}-1,$$

得证。两边都等于 $\frac{1}{3}\pi^2-3$。

验算

- 在第 (ii) 部分取 $N=1$: 和只有一项 $\frac{1}{6}$，而公式给出 $\frac{1}{2}-\frac{5}{8}+\frac{3}{8}-\frac{1}{12}=\frac{12-15+9-2}{24}=\frac{1}{6}$。
- 数值上，$\frac{1}{6}\pi^2-1\approx 0.644934$，$\frac{1}{12}\pi^2-\frac{5}{8}\approx 0.197467$，$\frac{1}{3}\pi^2-3\approx 0.289868$。用电脑把每个级数的首一百万项相加，三个数值都至少准确到小数点后九位。
- 三个部分中的有限恒等式，已用有理数运算对 N 直至 59 的每一个值逐一精确核实。

常见错误

- 弄错裂项相消和的首尾项，尤其是 $\frac{1}{r+1}-\frac{1}{r+2}$ 的和，它由 $\frac{1}{2}$ 开始，而不是 1。
- 拆成像 $\sum\frac{1}{r}$ 这样各自发散（diverge）的无穷级数，而不是把调和型（harmonic-type）的部分放在有限和中一并处理。

STEP 之外

$\sum 1/r^2 = \pi^2/6$ 这个值是欧拉（Euler）对巴塞尔问题（Basel problem）的解答，在大学里通常用傅里叶级数（Fourier series）或 $\sin x$ 的乘积公式来证明。这里对有限和的小心处理，正是分析学（analysis）课程要严格化的内容：级数就是其部分和的极限，收敛级数的和可以合并，发散级数则不可以。Abbott 的 *Understanding Analysis*（《理解分析》）正是这样处理级数的，很适合作为下一本读物。

详细解答

STEP III 2024，第 2 题

纯数学 · 不等式与绝对值

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

本题中每个平方根（square root）内的式子都是恒为正的二次式（quadratic），而且每条不等式（inequality）的两边都是非负的，所以我们可以两边平方而不改变解集（solution set）。平方之后，每个问题都变成关于某个二次式正负号的命题（或者当 x^2 项互相抵消时，变成关于一个线性函数（linear function）的命题）。从几何角度看，$y = \sqrt{4x^2 - 8x + 64}$ 是一条双曲线（hyperbola）的上半部分，而题目中的直线则是它的弦（chord）和渐近线（asymptote）。

解答

第 (i) 部分

配方（completing the square）得 $4x^2 - 8x + 64 = 4(x-1)^2 + 60 > 0$，所以平方根对所有 x 都有定义。每条不等式的两边都是非负的，而对于 $A, B \geq 0$，我们有 $A \leq B \iff A^2 \leq B^2$。

(a) 两边平方，

$$\begin{aligned} 4x^2 - 8x + 64 \leq x^2 + 16x + 64 &\iff 3x^2 - 24x \leq 0 \\ &\iff 3x(x-8) \leq 0 \iff 0 \leq x \leq 8. \end{aligned}$$

(b) 两边平方，

$$\begin{aligned} 4x^2 - 8x + 64 \leq 9x^2 - 48x + 64 &\iff 0 \leq 5x^2 - 40x \\ &\iff 5x(x-8) \geq 0 \iff x \leq 0 \text{ 或 } x \geq 8. \end{aligned}$$

在两部分中，等号都在 $x = 0$（各边均等于 8）和 $x = 8$（各边均等于 16）时成立。

第 (ii) 部分

(a) 利用平方差（difference of two squares），

$$\left(\sqrt{4x^2-8x+64}+2(x-1)\right)\mathrm{f}(x)=(4x^2-8x+64)-4(x-1)^2=60.$$

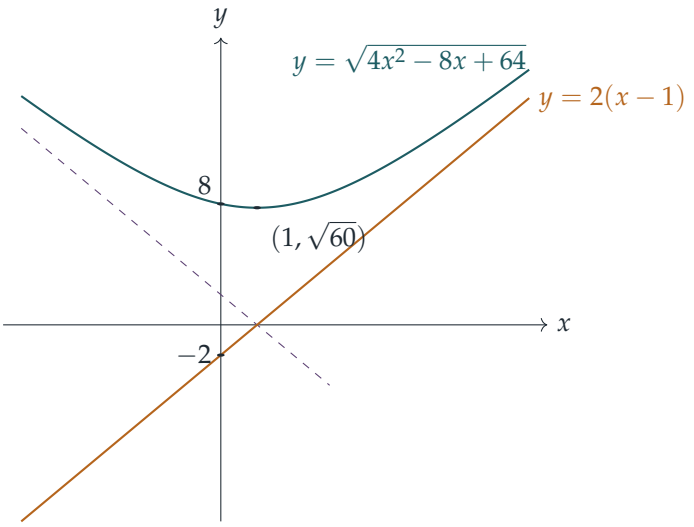
当 $x>1$ 时，括号内的式子为正，所以可以相除：

$$\mathrm{f}(x)=\frac{60}{\sqrt{4x^2-8x+64}+2(x-1)}.$$

当 $x\to\infty$ 时，分母至少为 $2(x-1)$，而它趋向无穷大，所以 $\mathrm{f}(x)\to 0$，正如所求。

(b) 设 $y=\sqrt{4(x-1)^2+60}$。那么 $y>0$ 且 $y^2-4(x-1)^2=60$，所以这条曲线是一条以 $(1,0)$ 为中心的双曲线的上支。它关于 $x=1$ 对称，在 $x=1$ 处取最小值 $\sqrt{60}=2\sqrt{15}\approx 7.75$，并与 y 轴相交于 $(0,8)$。

由于 $y>\sqrt{4(x-1)^2}=2|x-1|$，曲线严格位于两条直线 $y=\pm 2(x-1)$ 的上方。(a) 部分说明当 $x\to\infty$ 时，曲线趋近 $y=2(x-1)$。由关于 $x=1$ 的对称性，当 $x\to-\infty$ 时，曲线趋近 $y=-2(x-1)$。直线 $y=2(x-1)$ 的斜率（gradient）为 2，并经过 $(1,0)$ 和 $(0,-2)$。



虚线是另一条渐近线 $y=-2(x-1)$。

第 (iii) 部分

同样地，$4x^2-5x+4=4\left(x-\frac{5}{8}\right)^2+\frac{39}{16}>0$，所以平方根总是有定义。把（非负的）两边平方，该不等式等价于

$$(m^2-4)x^2+(2mc+5)x+(c^2-4)\geq 0. \qquad (*)$$

想一想 $(*)$ 左边是甚么类型的式子，才可能恰好在 $x \geq 3$ 时非负?

若 $m^2 - 4 > 0$, 当 $x \to -\infty$ 时左边趋向 $+\infty$, 于是非常负的 x 也会满足 $(*)$。若 $m^2 - 4 < 0$, 当 $x \to +\infty$ 时左边趋向 $-\infty$，于是很大的 x 会不满足。两者都不容许，所以 $m^2 = 4$，而 $(*)$ 变成线性不等式。线性不等式 $ax + b \geq 0$ 的解集为 $\{x : x \geq 3\}$，当且仅当 $a > 0$ 且 $3a + b = 0$。

取 $m = 2$。那么 $a = 4c + 5$，$b = c^2 - 4$，所以我们需要

$$3(4c + 5) + c^2 - 4 = c^2 + 12c + 11 = (c + 1)(c + 11) = 0, \qquad 4c + 5 > 0.$$

根 $c = -11$ 给出 $a = -39 < 0$，不符合。根 $c = -1$ 给出 $a = 1 > 0$。所以

$$m = 2, \qquad c = -1$$

可行。验证一下：$(2x - 1)^2 - (4x^2 - 5x + 4) = x - 3$，它恰好在 $x \geq 3$ 时非负。（取 $m = -2$、$c = 1$ 会得到相同的 $|mx + c|$，同样成立。）

第 (iv) 部分

设 $g(x) = x^2 + px + q$，$L(x) = mx + c$。由于 $|g| \leq L \iff -L \leq g \leq L$，该不等式成立当且仅当以下两者同时成立：

$$g(x) - L(x) \leq 0 \qquad 及 \qquad g(x) + L(x) \geq 0.$$

$g - L$ 和 $g + L$ 都是首项系数（leading coefficient）为 1 的二次式。

- 使 $g - L \leq 0$ 的集合是一个闭区间 $[a, b]$（可能是空集或单一点）。
- 使 $g + L \geq 0$ 的集合是全体实数除去一个开区间 (r, s)。若 $g + L$ 没有实根，则甚么都不用除去。

所以解集是 $[a, b]$ 除去 (r, s)。要使它等于 $[-5, 1] \cup [5, 7]$，我们需要 $a = -5$ 及 $b = 7$（外侧两端），而被除去的空隙必须恰好是 $(1, 5)$，所以 $r = 1$，$s = 5$。因此

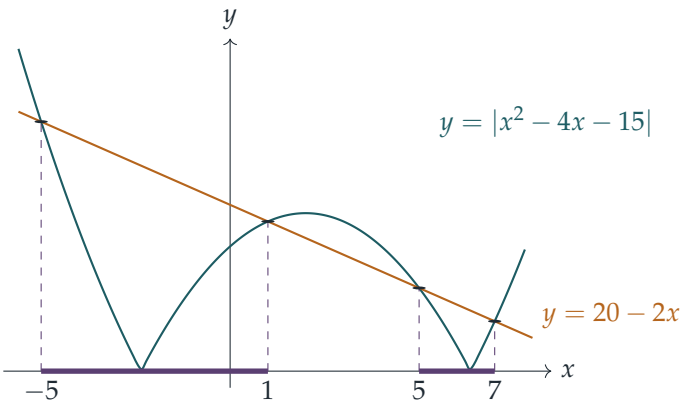
$$g(x) - L(x) = (x + 5)(x - 7) = x^2 - 2x - 35,$$
$$g(x) + L(x) = (x - 1)(x - 5) = x^2 - 6x + 5.$$

两式相加再除以二，得 $g(x) = x^2 - 4x - 15$。两式相减再除以二，得 $L(x) = -2x + 20$。因此

$$p = -4, \qquad q = -15, \qquad m = -2, \qquad c = 20.$$

以上每一步都是可逆的，所以取这些值时，解集确实是 $[-5, 1] \cup [5, 7]$：我们需要 $(x + 5)(x - 7) \leq 0$ 及 $(x - 1)(x - 5) \geq 0$。这些值也是唯一可行的值。

在图中，g 的根为 $2 \pm \sqrt{19} \approx -2.36,\ 6.36$。直线 $y = 20 - 2x$ 在 $x = -5$ 和 $x = 7$ 处与 $|g|$ 的外侧两臂相交，并在 $x = 1$ 和 $x = 5$ 处穿过被反射上来的拱形部分。



验算

- 在 (ii)(a) 中，当 $x = 100$ 时，$f(100) = \sqrt{39264} - 198 \approx 0.151$，与 $60/(\sqrt{39264} + 198) \approx 0.151$ 吻合。
- 在 (iv) 中，当 $x = 0$ 时：$|-15| = 15 \le 20$，所以 0 是解。当 $x = 3$ 时：$|9 - 12 - 15| = 18 > 14$，所以 3 不是解。当 $x = 8$ 时：$|64 - 32 - 15| = 17 > 4$，所以 8 不是解。这些都与 $[-5, 1] \cup [5, 7]$ 相符。用电脑在 x 的细密网格上检查，亦确认了全部四个解集。

常见错误

- 两边平方却没有说明为何可以这样做。这里之所以成立，只因为两边都是非负的。
- 在 (iii) 中保留真正的二次式。二次不等式的解集不可能是单侧的半直线，所以 x^2 项必须抵消。
- 在 (iv) 中把端点配对错了：外侧两端 $-5, 7$ 属于 $g - L$，而内侧两端 $1, 5$（即空隙）属于 $g + L$。

STEP 之外

曲线 $y^2 - 4(x-1)^2 = 60$ 是一条双曲线，而第 (ii) 部分展示了如何求渐近线：把曲线与候选直线之差有理化（rationalise）。同样的技巧——乘以“共轭式”（conjugate）——是大学一年级分析（analysis）中求极限的标准工具。像 (iv) 那样把 $|g| \le L$ 写成两个单侧条件，正是那里处理含绝对值（absolute value）不等式的惯常做法。Abbott 的 *Understanding*

Analysis（《理解分析》）是很好的下一步读物。

详细解答

STEP III 2024，第 3 题

纯数学 · 单调函数与导数

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

一切都取决于一个导数（derivative）。把 g 求导并通分后，分子（numerator）会化简成一个关于 x 的线性（linear）式子，因此它的正负一眼就能看出。与第 (iii) 部分的联系在于：g 恰好就是 $\ln f$ 的导数。所以 g 的正负（第 (ii) 部分的草图把它直观地呈现出来）决定了 f 是上升还是下降。

解答

第 (i) 部分

由于 $x>0$，可写成 $\ln\left(1+\frac{1}{x}\right)=\ln(x+1)-\ln x$。于是

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

而以分母（denominator）x^2+x 运用商法则（quotient rule），

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x+c}{x^2+x}=\frac{(x^2+x)-(x+c)(2x+1)}{x^2(x+1)^2}=\frac{-x^2-2cx-c}{x^2(x+1)^2}.$$

两式相减，x^2 项互相抵消：

$$\mathrm{g}'(x)=\frac{-x^2-x+x^2+2cx+c}{x^2(x+1)^2}=\frac{(2c-1)x+c}{x^2(x+1)^2}.$$

分母为正，所以 $\mathrm{g}'(x)$ 的正负与 $N(x)=(2c-1)x+c$ 的正负相同。

(a) 若 $c\geq\frac{1}{2}$，则 $(2c-1)x\geq 0$ 且 $c\geq\frac{1}{2}>0$，所以对每个 $x>0$ 都有 $N(x)>0$。因此对所有 $x>0$，斜率（gradient）均为正。

(b) 若 $0 \le c < \frac{1}{2}$，则 $1-2c>0$，而

$$N(x)<0 \iff (1-2c)x > c \iff x > \frac{c}{1-2c}.$$

所以斜率为负当且仅当 $x > \dfrac{c}{1-2c}$。（当 $c=0$ 时，这就是所有 $x>0$。）记 $x_0 = \dfrac{c}{1-2c}$。当 $c>0$ 时，g 在 $(0, x_0]$ 上递增，在 $[x_0, \infty)$ 上递减。

第 (ii) 部分

我们还需要知道 x 很大时的情况。g 的两项都趋向 0，所以 $g(x) \to 0$，但草图应该显示它从哪一侧趋近。利用 $\ln(1+t) = t - \frac{1}{2}t^2 + \cdots$，取 $t = \frac{1}{x}$，再加上 $\dfrac{x+c}{x(x+1)} = \left(\dfrac{1}{x} + \dfrac{c}{x^2}\right)\left(1 - \dfrac{1}{x} + \cdots\right) = \dfrac{1}{x} + \dfrac{c-1}{x^2} + \cdots$，得到

$$g(x) = \frac{\frac{1}{2} - c}{x^2} + \cdots \quad \text{对很大的} x.$$

> **想一想** 若 g 递增并趋向 0，它必定位于轴的哪一侧?

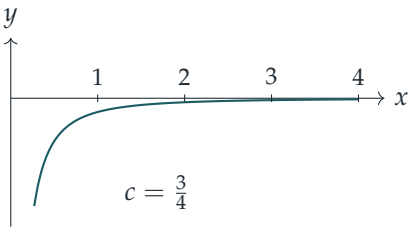
它必定在轴下方。若 g 严格递增，则对任何 x，所有 $y > x$ 都有 $g(x) < g(y)$；令 $y \to \infty$ 便得 $g(x) < 0$（严格不等，因为 $g(x) < g(x+1) \le 0$）。同理，一个趋向 0 的严格递减函数必定为正。

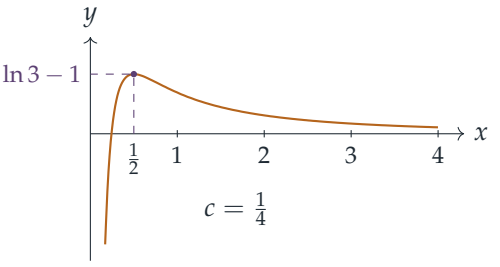
(a) $c = \frac{3}{4}$。由 (i)(a)，g 处处递增。它从 $-\infty$（当 $x \to 0$）上升趋向 0，并一直留在轴下方。它以 $-\frac{1}{4x^2}$ 的方式趋近，所以永不穿过轴，也没有转向点。例如 $g(1) = \ln 2 - \frac{7}{8} \approx -0.18$。

(b) $c = \frac{1}{4}$。这时 $x_0 = \dfrac{1/4}{1/2} = \frac{1}{2}$。曲线从 $-\infty$ 上升至 $x = \frac{1}{2}$ 处的极大值，其中

$$g\left(\tfrac{1}{2}\right) = \ln 3 - \frac{3/4}{3/4} = \ln 3 - 1 \approx 0.099 > 0.$$

之后它从上方以 $\frac{1}{4x^2}$ 的方式递减趋向 0。在 $[\frac{1}{2}, \infty)$ 上它递减至 0，所以在那里为正。在 $(0, \frac{1}{2})$ 上它从 $-\infty$ 递增至一个正值，所以恰好穿过轴一次，位置略小于 $\frac{1}{4}$（数值上 $x \approx 0.243$）。





（两图的纵向比例不同：右图经过拉伸，好让那个细小的正值隆起看得见。两图中 y 轴都是渐近线（asymptote）。）

第 (iii) 部分

由于对 $x > 0$ 有 $f(x) > 0$，我们可以取对数（logarithm）：

$$\ln f(x) = (x + c)\ln\left(1 + \frac{1}{x}\right).$$

用积法则（product rule）求导，并利用 (i) 中的导数，

$$\frac{f'(x)}{f(x)} = \ln\left(1 + \frac{1}{x}\right) + (x + c)\cdot\left(-\frac{1}{x(x+1)}\right) = g(x).$$

所以 $f'(x) = f(x)\,g(x)$，又因 $f(x) > 0$，f' 与 g 同号。另外留意，对每个 c，当 $x \to \infty$ 时 $g(x) \to 0$，因为它的两项都趋向 0。

(a) $c \geq \frac{1}{2}$。由 (i)(a)，g 在 $x > 0$ 上严格递增并趋向 0。由 (ii) 中方框之后的论证，对所有 $x > 0$ 有 $g(x) < 0$。因此 $f'(x) < 0$，f 是递减的。

(b) $0 < c < \frac{1}{2}$。由 (i)(b)，g 在 $[x_0, \infty)$ 上严格递减并趋向 0，所以在那里 $g > 0$；特别地 $g(x_0) > 0$。在 $(0, x_0]$ 上，g 连续（continuous）且严格递增，而当 $x \to 0$ 时 $g(x) \to -\infty$（题目已给出，因为 $c > 0$）。因此它恰好在一点 $x_1 \in (0, x_0)$ 取值 0，并在 $(0, x_1)$ 上 $g < 0$，在 (x_1, ∞) 上 $g > 0$。所以 $f'(x_1) = 0$，而 f' 在该处由负变正。因此 f 在 x_1 有一个转向点（turning point）（极小值）。

> ? **想一想** 为甚么单凭 $f'(x_1) = 0$ 并不足够?

驻点（stationary point）不一定是转向点（想想 $y = x^3$ 在原点的情况）。令 x_1 成为转向点的，是 f' 的变号，而这正是 g 的形状所提供的。

(c) $c = 0$。这时对所有 $x > 0$ 有 $g'(x) = \dfrac{-x}{x^2(x+1)^2} < 0$，所以 g 严格递减并趋向 0。因此对所有 $x > 0$ 有 $g(x) > 0$，于是 $f'(x) > 0$，而 $f(x) = \left(1 + \frac{1}{x}\right)^x$ 是递增的。

验算

- 在某一点以数值方法对 g 求导，例如 $x = 1$、$c = \frac{1}{4}$：公式给出 $g'(1) = \frac{-1/2+1/4}{4} = -\frac{1}{16}$，与差商（difference quotient）吻合。
- $c = \frac{1}{4}$：$g(0.2) \approx -0.083$ 而 $g(0.3) \approx 0.056$，与在 0.243 附近唯一一次穿过轴一致。对 $c = \frac{3}{4}$，$g(3) \approx -0.025$，仍然刚好在轴下方。
- 在 (iii)(b) 中取 $c = \frac{1}{4}$，f 的极小值位于同一个 $x \approx 0.243$；用电脑在细密网格上检查 f，对 $(0, \frac{1}{2})$ 中多个 c 值都恰好找到一个转向点。
- 对 $c = 0$，$f(x) = (1 + \frac{1}{x})^x$ 递增至 e；对 $c = 1$，$(1 + \frac{1}{x})^{x+1}$ 递减至 e。这正是大家熟悉、从两边夹住 e 的那对数列（sequences）。

常见错误

- 单凭“g 递增”就断定 $\mathrm{g} < 0$。你还需要当 $x \to \infty$ 时的极限 $\mathrm{g} \to 0$。
- 在 (ii)(b) 中把曲线画成一直为负。极大值 $\ln 3 - 1$ 虽小但为正，所以曲线必须穿过轴，然后从上方回落趋近轴。
- 对 $(1 + \frac{1}{x})^{x+c}$ 求导时把指数当作常数。应先取对数。

STEP 之外

$(1 + \frac{1}{x})^{x}$ 递增而 $(1 + \frac{1}{x})^{x+1}$ 递减、两者都趋向 e，这是大学一年级证明定义 e 的极限存在的经典方法。它是单调收敛定理（monotone convergence theorem）的一个例子。这道题找出了行为转变的精确临界值 $c = \frac{1}{2}$，而展开式 $\mathrm{g}(x) \approx (\frac{1}{2} - c)/x^2$ 解释了原因。想进一步了解单调数列与极限，可以读 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP III 2024，第 4 题

纯数学 · 抛物线的切线

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

一切都取决于两直线夹角的 tan 公式，而这其实只是 $\tan(A-B)$ 公式换了个面貌。一旦知道在 $x=p$ 和 $x=q$ 处的切线（tangent）相交于点 (X,Y)，而 $p+q$ 和 pq 都可以用 X 和 Y 表示，那么任何关于夹角的条件（涉及 $(p-q)^2$ 和 pq）都可以单纯用 X 和 Y 改写。改写后的条件就是所求的轨迹（locus）。

解答

第 (i) 部分

设两直线与 x 轴正方向分别成 α 角和 β 角，于是 $m_1=\tan\alpha$，$m_2=\tan\beta$（两直线都有斜率（gradient），所以都不是铅垂线）。两直线的夹角在不计正负号及加上 180° 的倍数的意义下等于 $\alpha-\beta$。若两者之间的锐角为 45°，则 $\alpha-\beta$ 等于 $\pm45^\circ$ 加上 180° 的某个倍数；由于 tan 的周期为 180°，

$$\tan(\alpha-\beta)=\tan(\pm45^\circ)=\pm1.$$

特别地，$\cos(\alpha-\beta)\neq0$，所以复角公式（compound-angle formula）适用，且 $1+m_1m_2\neq0$：

$$\tan(\alpha-\beta)=\frac{\tan\alpha-\tan\beta}{1+\tan\alpha\tan\beta}=\frac{m_1-m_2}{1+m_1m_2}.$$

因此 $\dfrac{m_1-m_2}{1+m_1m_2}=\pm1$，得证。更一般地，两直线之间的锐角 θ 满足 $\tan\theta=\left|\dfrac{m_1-m_2}{1+m_1m_2}\right|$，我们会在第 (iii) 部分用到这一点。

第 (ii) 部分

在 C 上有 $y=\dfrac{x^2}{4a}$，所以 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=\dfrac{x}{2a}$。在 $\left(p,\frac{p^2}{4a}\right)$ 处的切线为

$$y-\frac{p^2}{4a}=\frac{p}{2a}(x-p)\quad\Longleftrightarrow\quad y=\frac{p}{2a}x-\frac{p^2}{4a},$$

同样，在 $x=q$ 处的切线为 $y=\dfrac{q}{2a}x-\dfrac{q^2}{4a}$。两式相减，

$$\frac{p-q}{2a}\,x=\frac{p^2-q^2}{4a}=\frac{(p-q)(p+q)}{4a}.$$

由于 $p\neq q$，可以除以 $p-q$，得 $x=\frac{1}{2}(p+q)$，得证。代入第一条切线，

$$y=\frac{p(p+q)}{4a}-\frac{p^2}{4a}=\frac{pq}{4a}.$$

所以两切线相交于

$$(X,Y)=\left(\tfrac{1}{2}(p+q),\,\frac{pq}{4a}\right),\qquad\text{所以}\qquad p+q=2X,\quad pq=4aY.\qquad(*)$$

现在假设在 $x=p$ 和 $x=q$ 处的切线互成 45°。它们显然是不同的直线，所以 $p\neq q$，且 $(*)$ 成立。两者的斜率为 $m_1=\frac{p}{2a}$ 和 $m_2=\frac{q}{2a}$，所以

$$\frac{m_1-m_2}{1+m_1m_2}=\frac{(p-q)/(2a)}{1+pq/(4a^2)}=\frac{2a(p-q)}{4a^2+pq}.$$

由第 (i) 部分，此式等于 ±1；平方后可消去正负号：

$$4a^2(p-q)^2=(4a^2+pq)^2.$$

> **想一想** 条件涉及的是 $p-q$ 而不是 $p+q$，我们怎样消去 p 和 q?

利用 $(p-q)^2=(p+q)^2-4pq$，这是对称（symmetric）的表达式，所以可以应用 $(*)$：

$$(p-q)^2=4X^2-16aY,\qquad 4a^2+pq=4a^2+4aY=4a(a+Y).$$

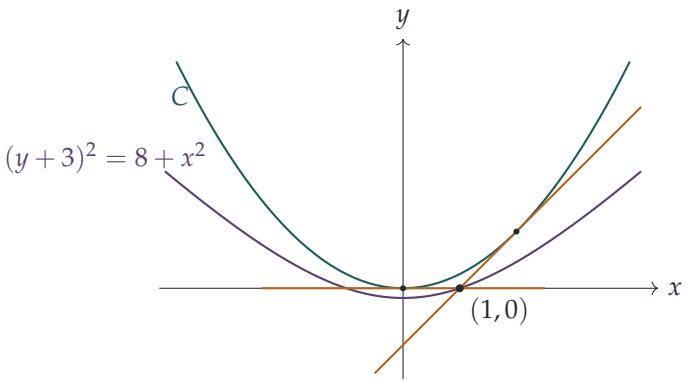
条件变成

$$4a^2\left(4X^2-16aY\right)=16a^2(a+Y)^2.$$

两边除以 $16a^2$（因为 $a\neq0$，所以可以这样做），

$$X^2-4aY=Y^2+2aY+a^2,$$
$$\text{所以}\qquad X^2+8a^2=Y^2+6aY+9a^2=(Y+3a)^2.$$

因此交点位于 $(y+3a)^2=8a^2+x^2$ 上，得证。



取 $a = 1$：在 $x = 2$ 处（斜率 1）和 $x = 0$ 处（斜率 0）的切线互成 45°，两者相交于 $(1, 0)$，这点位于双曲线（hyperbola）的上支上。

第 (iii) 部分

设 (X, Y) 是 $(y + 7a)^2 = 48a^2 + 3x^2$ 上的一点，而 C 在 $x = p$ 和 $x = q$ $(p \neq q)$ 处的两条切线相交于此点，所以 $(*)$ 成立。根据第 (i) 部分末尾的说明，两切线之间的锐角 θ 满足

$$\tan^2\theta = \left(\frac{2a(p-q)}{4a^2+pq}\right)^2 = \frac{4a^2(4X^2-16aY)}{16a^2(a+Y)^2} = \frac{X^2-4aY}{(Y+a)^2},$$

这里用了与第 (ii) 部分相同的代换。（若 $Y = -a$，两切线会互相垂直；下面会证明这在此曲线上不可能发生。）

> **想一想** 曲线的方程怎样帮助我们处理 $X^2 - 4aY$?

展开曲线的方程：$Y^2 + 14aY + 49a^2 = 48a^2 + 3X^2$，即

$$3X^2 = Y^2 + 14aY + a^2.$$

于是

$$3(X^2 - 4aY) = Y^2 + 14aY + a^2 - 12aY = Y^2 + 2aY + a^2 = (Y + a)^2.$$

由此得出两个结果。第一，$Y \neq -a$：否则 $X^2 - 4aY = 0$，即 $X^2 = -4a^2$，但 $a \neq 0$，这是不可能的。所以 $X^2 - 4aY = \frac{1}{3}(Y + a)^2 > 0$，这也确认了该点位于 C 之外：以 p 和 q 为根的二次方程 $t^2 - 2Xt + 4aY = 0$ 有两个相异实根，所以经过曲线上每一点确实都有两条切线。第二，

$$\tan^2\theta = \frac{\frac{1}{3}(Y+a)^2}{(Y+a)^2} = \frac{1}{3}.$$

所以无论该点位于曲线上何处，都有 $\tan\theta = \frac{1}{\sqrt{3}}$：两切线之间的锐角是常数，等于 30°。

验算

- 取 $a = 1$，在 $x = 2$ 和 $x = 0$ 处的切线分别是 $y = x - 1$ 和 $y = 0$，互成 45°。它们相交于 $(1, 0)$，而 $(0+3)^2 = 9 = 8 + 1^2$。
- 对于第 (iii) 部分，取 $a = 1$ 及 $X = 0$：则 $(Y+7)^2 = 48$，所以 $Y = 4\sqrt{3} - 7 \approx -0.072$。经过 $(0, Y)$ 的切线在 $x = \pm\sqrt{-4Y} \approx \pm 0.536$ 处与曲线相切，斜率为 ± 0.268。由于 $\tan 15^\circ \approx 0.268$，每条切线与水平线成 15°，两者相差 30°。
- 用电脑对一千个随机的 a 值以及每条双曲线两支上的点进行检验，确认了 45° 和 30° 的结果。

常见错误

- 试图分别解出 p 和 q。其实只需要对称组合 $p + q$ 和 pq，再配合 $(p-q)^2 = (p+q)^2 - 4pq$。
- 在第 (i) 部分忘记 $\pm$：把条件平方后，45° 角的两种方向都得以保留，正因如此才得到单一条方程。
- 在第 (iii) 部分，约去之前没有检查 $Y + a \neq 0$。

STEP 之外

从点 (X, Y) 到抛物线（parabola）的切线由二次方程 $t^2 - 2Xt + 4aY = 0$ 决定，而使两切线以固定角度相交的点的轨迹总是一条双曲线（当角度为 90° 时，它退化为准线（directrix）$y = -a$）。这些概念属于射影几何（projective geometry），其中极点与极线（pole–polar）的关系把一个点与它的切点弦（chord of contact）联系起来。如想做更多同类型的题目，并参考如何入手的评注，可参阅 Siklos 的 *Advanced Problems in Mathematics*（《高等数学难题》）。

详细解答

STEP III 2024，第 5 题

纯数学 · 矩阵微分方程

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

本题在 *Beyond STEP* 第 7.7 节永恒不变的量中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch07/lax-pairs.html

关键思路

一切都取决于迹（trace）的两个性质：它是线性（linear）的，而且 $\operatorname{tr}(\mathbf{XY}) = \operatorname{tr}(\mathbf{YX})$。第 (ii) 部分给出行列式（determinant）变化率的公式，即雅可比公式（Jacobi's formula）。它把“$\det \mathbf{M}$ 是否不变?”转化为“某个迹是否为零?”。在 (iii) 中，交换子（commutator）$\mathbf{MN} - \mathbf{NM}$ 的迹必定为零，由此得到全部三个不变量（invariants）。在 (iv) 中，同一条公式表明 $\operatorname{tr} \mathbf{N}$ 量度的是 $\ln|\det \mathbf{M}|$ 的变化快慢，于是问题变成：一个矩阵能否迹不变而行列式改变?

解答

第 (i) 部分

设 $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 及 $\mathbf{N} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$。$\mathbf{MN}$ 的对角元素是 $ap + br$ 和 $cq + ds$，而 $\mathbf{NM}$ 的对角元素是 $pa + qc$ 和 $rb + sd$。所以

$$\operatorname{tr}(\mathbf{MN}) = ap + br + cq + ds = \operatorname{tr}(\mathbf{NM}).$$

此外，$\mathbf{M} + \mathbf{N}$ 的对角元素是 $a + p$ 和 $d + s$，所以

$$\operatorname{tr}(\mathbf{M} + \mathbf{N}) = \operatorname{tr}(\mathbf{M}) + \operatorname{tr}(\mathbf{N}).$$

同理，对纯量 k 有 $\operatorname{tr}(k\mathbf{M}) = k\operatorname{tr}(\mathbf{M})$，而 $\mathrm{d}\mathbf{M}/\mathrm{d}t$ 的迹等于 $\operatorname{tr}\mathbf{M}$ 的导数。下文会随意使用这些性质。

第 (ii) 部分

元素 a, b, c, d 均随 t 而变（撇号表示 $\mathrm{d}/\mathrm{d}t$），由乘积法则（product rule）得

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det\mathbf{M}) = \frac{\mathrm{d}}{\mathrm{d}t}(ad - bc) = a'd + ad' - b'c - bc'.$$

由于 $\mathbf{M}$ 是非奇异（non-singular）矩阵，$\mathbf{M}^{-1} = \dfrac{1}{\det\mathbf{M}}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$，而

$$\mathbf{M}^{-1}\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \frac{1}{\det\mathbf{M}}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \frac{1}{\det\mathbf{M}}\begin{pmatrix} da' - bc' & * \\ * & -cb' + ad' \end{pmatrix}.$$

它的迹是 $\dfrac{a'd + ad' - b'c - bc'}{\det\mathbf{M}}$，即 $\dfrac{1}{\det\mathbf{M}}\dfrac{\mathrm{d}}{\mathrm{d}t}(\det\mathbf{M})$，得证。

第 (iii) 部分

行列式。由 (ii) 及迹的线性，

$$\frac{1}{\det\mathbf{M}}\frac{\mathrm{d}}{\mathrm{d}t}(\det\mathbf{M}) = \operatorname{tr}\big(\mathbf{M}^{-1}(\mathbf{MN} - \mathbf{NM})\big) = \operatorname{tr}(\mathbf{N}) - \operatorname{tr}(\mathbf{M}^{-1}\mathbf{NM}).$$

对两个矩阵 $\mathbf{M}^{-1}\mathbf{N}$ 和 $\mathbf{M}$ 应用 (i)，得 $\operatorname{tr}(\mathbf{M}^{-1}\mathbf{N}\,\mathbf{M}) = \operatorname{tr}(\mathbf{M}\,\mathbf{M}^{-1}\mathbf{N}) = \operatorname{tr}(\mathbf{N})$。所以右方为零。由于 $\det\mathbf{M} \neq 0$，得 $\frac{\mathrm{d}}{\mathrm{d}t}(\det\mathbf{M}) = 0$。

迹。由 (i)，$\dfrac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}(\mathbf{M}) = \operatorname{tr}\Big(\dfrac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}\Big) = \operatorname{tr}(\mathbf{MN}) - \operatorname{tr}(\mathbf{NM}) = 0$。

平方的迹。只要保持因子的次序，乘积法则对矩阵同样成立：

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M}^2) &= \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}\mathbf{M} + \mathbf{M}\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = (\mathbf{MN} - \mathbf{NM})\mathbf{M} + \mathbf{M}(\mathbf{MN} - \mathbf{NM}) \\ &= \mathbf{M}^2\mathbf{N} - \mathbf{NM}^2,\end{aligned}$$

因为两个 $\mathbf{MNM}$ 项互相抵消。这同样是交换子的形式，只是以 $\mathbf{M}^2$ 代替 $\mathbf{M}$，所以由 (i) 知其迹为零。因此 $\operatorname{tr}(\mathbf{M}^2)$ 亦是常数。三个量都与 t 无关。

> ? **想一想** 当 $\mathbf{N}$ 也随 t 而变时，怎样解 $\mathrm{d}\mathbf{M}/\mathrm{d}t = \mathbf{MN} - \mathbf{NM}$?

把 $\mathbf{N}$ 中与一切矩阵都可交换的部分分拆出来。写 $\mathbf{N} = t\mathbf{I} + t\mathbf{J}$，其中 $\mathbf{J} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$。单位

矩阵与 $\mathbf{M}$ 可交换，所以 $\mathbf{MN}-\mathbf{NM}=t(\mathbf{MJ}-\mathbf{JM})$。设 $\mathbf{M}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$，则

$$\mathbf{MJ}-\mathbf{JM}=\begin{pmatrix}0 & a\\ 0 & c\end{pmatrix}-\begin{pmatrix}c & d\\ 0 & 0\end{pmatrix}=\begin{pmatrix}-c & a-d\\ 0 & c\end{pmatrix}.$$

比较元素，得四条纯量方程：

$$a'=-tc,\qquad b'=t(a-d),\qquad c'=0,\qquad d'=tc.$$

以 $t=0$ 时的值 A,B,C,D 依次求解：

- $c'=0$，所以 $c=C$。
- 于是 $a'=-Ct$ 及 $d'=Ct$，所以 $a=A-\frac{1}{2}Ct^2$ 及 $d=D+\frac{1}{2}Ct^2$。
- 现在 $a-d=A-D-Ct^2$，所以 $b'=(A-D)t-Ct^3$，得 $b=B+\frac{1}{2}(A-D)t^2-\frac{1}{4}Ct^4$。

因此

$$\mathbf{M}=\begin{pmatrix}A-\frac{1}{2}Ct^2 & B+\frac{1}{2}(A-D)t^2-\frac{1}{4}Ct^4\\ C & D+\frac{1}{2}Ct^2\end{pmatrix}.$$

验算一下：迹是 $A+D$，而展开 $ad-bc$ 可见 t 的各次幂全部抵消，剩下 $AD-BC$。这与本部分前半的结果一致。

第 (iv) 部分

现在 $\mathrm{d}\mathbf{M}/\mathrm{d}t=\mathbf{MN}$，所以 $\mathbf{M}^{-1}\,\mathrm{d}\mathbf{M}/\mathrm{d}t=\mathbf{N}$，而 (ii) 给出

$$\operatorname{tr}(\mathbf{N})=\frac{1}{\det\mathbf{M}}\frac{\mathrm{d}}{\mathrm{d}t}(\det\mathbf{M}).$$

所以对所有 t 有 $\operatorname{tr}\mathbf{N}=0$，当且仅当 $\det\mathbf{M}$ 是常数。迹不变并不能说明行列式的任何情况，所以答案应该是否。要证明这一点，只需一个反例（counterexample）。

> ? **想一想** 题目只通过微分方程给出 $\mathbf{N}$，我们该怎样构造反例?

倒过来做。任取一个非奇异的 $\mathbf{M}(t)$，使其迹为非零常数而行列式不是常数，然后定义 $\mathbf{N}=\mathbf{M}^{-1}\,\mathrm{d}\mathbf{M}/\mathrm{d}t$。这样方程便自动成立。取

$$\mathbf{M}=\begin{pmatrix}1 & t\\ -t & 1\end{pmatrix},\qquad \operatorname{tr}\mathbf{M}=2,\qquad \det\mathbf{M}=1+t^2\neq 0.$$

于是

$$\mathbf{N}=\mathbf{M}^{-1}\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}=\frac{1}{1+t^2}\begin{pmatrix}1 & -t\\ t & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}=\frac{1}{1+t^2}\begin{pmatrix}t & 1\\ -1 & t\end{pmatrix}.$$

乘出来可确认 $\mathbf{MN} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathrm{d}\mathbf{M}/\mathrm{d}t$。此外 $\mathbf{N}$ 是非奇异的，因为 $\det \mathbf{N} = 1/(1+t^2)$。然而

$$\operatorname{tr}(\mathbf{N}) = \frac{2t}{1+t^2},$$

对每个 $t \neq 0$ 都不为零。所以 $\operatorname{tr}(\mathbf{N}) = 0$ 不一定成立。真正必然成立的是 $\operatorname{tr}(\mathbf{MN}) = \operatorname{tr}(\mathrm{d}\mathbf{M}/\mathrm{d}t) = 0$，这是另一回事。

验算

- 在 (iii) 中，取 $A = 1$、$B = 0$、$C = 2$、$D = 3$ 及 $t = 1$。公式给出 $\mathbf{M} = \begin{pmatrix} 0 & -\frac{3}{2} \\ 2 & 4 \end{pmatrix}$，其迹为 4，行列式为 3，与 $t = 0$ 时相同。它的平方的迹是 $0 - 3 - 3 + 16 = 10$，而初始矩阵的也是：$1 + 0 + 0 + 9 = 10$。
- 对随机的 A, B, C, D，(iii) 中 $\mathbf{M}$ 的公式与电脑数值积分该微分方程的结果吻合至小数点后八位。以随机且随时间变化的 $\mathbf{N}$ 进行同样的积分时，三个不变量都保持不变。
- 在 (iv) 中，当 $t = 1$ 时，$\mathbf{N} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$，其迹为 1；而 $\det \mathbf{M} = 1 + t^2$ 在该处的对数导数（logarithmic derivative）是 $2t/(1+t^2) = 1$，正如 (ii) 所预测。

常见错误

- 写成 $\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M}^2) = 2\mathbf{M}\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}$：矩阵不一定可交换，所以两种次序都必须保留。
- 以为 $\operatorname{tr}(\mathbf{MN}) = \operatorname{tr}\mathbf{M}\operatorname{tr}\mathbf{N}$。这是错的。成立的只有线性法则和循环法则（cyclic rule）。
- 在 (iv) 中，由 $\operatorname{tr}(\mathbf{MN}) = 0$ 直接推出 $\operatorname{tr}(\mathbf{N}) = 0$。答“否”需要一个明确的反例。

STEP 之外

第 (ii) 部分是雅可比公式在 2×2 的情况，该公式对 $n \times n$ 矩阵都成立。配合 $\mathbf{M}' = \mathbf{MN}$，它给出 $\det \mathbf{M}(t) = \det \mathbf{M}(0) \exp \int_0^t \operatorname{tr}\mathbf{N}$。当 $\mathbf{N}$ 是常数时，这就变成著名的恒等式 $\det e^{\mathbf{N}} = e^{\operatorname{tr}\mathbf{N}}$。形如 $\mathbf{M}' = \mathbf{MN} - \mathbf{NM}$ 的方程称为 Lax 方程（Lax equations）。$\operatorname{tr}(\mathbf{M}^k)$ 这些量总是守恒（conserved）的，这正是它们出现在可积系统（integrable systems）理论中的原因。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）是阅读迹、行列式和矩阵指数（matrix exponential）的好去处。

详细解答

STEP III 2024，第 6 题

纯数学 · 线性微分方程

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

控制量（control）u 以相同方式加到每一条方程上，所以坐标的任何一个系数总和为零的组合，都完全不受 u 影响。只要选对这个组合，它的导数（derivative）就是它本身的倍数，因此无论 u 怎样变化，它都会按指数（exponential）增长或衰减。指数函数除非一开始就是零，否则永远不会等于零，由此便得到必要条件（necessary condition）。当这些条件成立时，运动就简化为单一条一阶（first-order）方程，而我们可以用常数 u 把它解出来。

解答

第 (i) 部分

(a) 把两条方程相减，含 u 的项互相抵消：

$$\frac{\mathrm{d}x}{\mathrm{d}t} - \frac{\mathrm{d}y}{\mathrm{d}t} = (-x + 3y + u) - (x + y + u) = -2(x - y).$$

因此 $w = x - y$ 满足 $\frac{\mathrm{d}w}{\mathrm{d}t} = -2w$，于是对任何函数 u，都有

$$x - y = (x_0 - y_0)\,\mathrm{e}^{-2t}$$

若质点在某个 $t > 0$ 时位于原点，则该时刻 $x - y = 0$。由于 $\mathrm{e}^{-2t} \neq 0$，这迫使 $x_0 - y_0 = 0$，即 $x_0 = y_0$。

(b) 当 $x_0 = y_0$ 时，(a) 的公式给出对所有 t 都有 $x - y = 0$，所以质点一直留在直线 $y = x$ 上。把 $y = x$ 代入第一条方程，得

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 2x + u.$$

当 u 为常数时，这是一条常系数线性方程。其特解（particular solution）为常数 $-\frac{1}{2}u$，所以

$$x = -\tfrac{1}{2}u + \left(x_0 + \tfrac{1}{2}u\right)\mathrm{e}^{2t}.$$

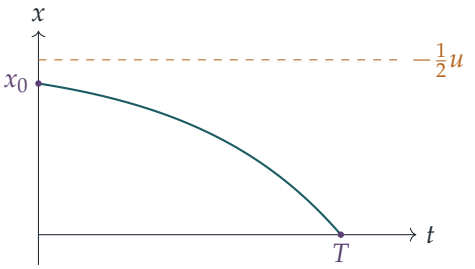
我们需要 $x(T) = 0$（这时 $y(T) = x(T) = 0$ 亦随之成立）：

$$\left(x_0 + \tfrac{1}{2}u\right)\mathrm{e}^{2T} = \tfrac{1}{2}u \quad \Longrightarrow \quad x_0\,\mathrm{e}^{2T} = \tfrac{1}{2}u\left(1 - \mathrm{e}^{2T}\right).$$

由于 $T > 0$，$\mathrm{e}^{2T} \neq 1$，所以可以相除：

$$\boxed{u = -\frac{2x_0\,\mathrm{e}^{2T}}{\mathrm{e}^{2T} - 1}}$$

这个正负号是合理的。若不加控制，$\dot{x} = 2x$ 会把质点推离原点，所以 u 必须朝与 x_0 相反的方向推。（若 $x_0 = 0$，质点在 $u = 0$ 时已静止于原点。）



图中所示为 $x_0 = 1$、$T = 1$ 的情况。质点从不稳定平衡点（unstable equilibrium）$x = -\frac{1}{2}u$ 稍下方出发，被推离该点，并恰好在时间 T 到达原点。

第 (ii) 部分

(a) 仿照第 (i) 部分，我们寻找一个 $ax + by + cz$，其中 $a + b + c = 0$（使 u 抵消），而且它的导数是它本身的倍数。“x，然后 $-2y$，然后 z”这个规律提示我们取 $w = x - 2y + z$：

$$\begin{aligned}\frac{\mathrm{d}w}{\mathrm{d}t} &= (4y - 5z + u) - 2(x - 2z + u) + (x - 2y + u)\\ &= -x + 2y - z = -w.\end{aligned}$$

因此无论 u 是甚么，都有 $x - 2y + z = (x_0 - 2y_0 + z_0)\,\mathrm{e}^{-t}$。在原点处 $w = 0$，而 $\mathrm{e}^{-t} \neq 0$，所以 $x_0 - 2y_0 + z_0 = 0$。也就是说，$y_0 = \frac{1}{2}(x_0 + z_0)$，即 x_0 与 z_0 的平均值。

> **想一想** 是否还有第二个同样不受 u 影响的组合?

最后两条方程都含有 $x + u$，所以把它们相减，可以同时消去 x 和 u。设 $v = y - z$，则

$$\frac{\mathrm{d}v}{\mathrm{d}t} = (x - 2z + u) - (x - 2y + u) = 2(y - z) = 2v,$$

所以 $y-z=(y_0-z_0)\,\mathrm{e}^{2t}$。

(b) 若质点在某个 $t>0$ 时位于原点，则该时刻 $v=0$。由于 $\mathrm{e}^{2t}\neq 0$，得 $y_0=z_0$。结合 (a)，有 $x_0=2y_0-z_0=y_0$，所以 $x_0=y_0=z_0$。

(c) 当 $x_0=y_0=z_0$ 时，w 和 v 的初始值都是 0，所以对所有 t 它们都保持为 0：$y=z$，从而 $x=2y-z=y$。质点沿直线 $x=y=z$ 运动，而每条方程都变成

$$\frac{\mathrm{d}x}{\mathrm{d}t}=4x-5x+u=-x+u$$

（第二和第三条方程给出 $x-2x+u$，结果相同）。当 u 为常数时，

$$x=u+(x_0-u)\,\mathrm{e}^{-t}.$$

令 $x(T)=0$，得 $u\left(1-\mathrm{e}^{-T}\right)=-x_0\,\mathrm{e}^{-T}$。由于 $T>0$，可以相除；再把分子分母同乘 e^{T}，得

$$\boxed{u=-\frac{x_0}{\mathrm{e}^{T}-1}}$$

这里不受控制的运动 $\dot{x}=-x$ 本身已会衰减趋向原点，但只是在 $t\to\infty$ 时才到达。只需一个朝向原点的小推动（当 $T\to\infty$ 时趋于零），就能令质点在有限时间内到达原点。

验算

- 极限：当 $T\to\infty$ 时，(ii)(c) 的答案趋于 0，与自然衰减相符；而 (i)(b) 的答案趋于 $-2x_0$，这使 $x_0=-\frac{1}{2}u$ 成为平衡点。质点于是停留在那里，永远不会到达原点，这正是有限的 T 需要绝对值稍大的 u 的原因。当 $T\to 0^{+}$ 时，两个答案都趋于无穷大，这是应有的结果。
- 数值验算：取 $x_0=1$、$T=1$，第 (i)(b) 部分给出 $u\approx -2.3130$，用数值方法逐步推进方程，在 $t=1$ 时落在 $(0,0)$，误差在 10^{-8} 以内。(ii)(c) 取 $u=-1/(\mathrm{e}-1)\approx -0.5820$ 时亦同样成立。此外，亦验证了当 u 是随时间变化的非常数函数时，不变量 (invariant) $x-y$、$x-2y+z$ 和 $y-z$ 仍然按各自的指数函数变化。

常见错误

- 在 (i)(a) 或 (ii)(a)、(b) 中假设 u 为常数。这些部分针对的是*任何* u，而 u 被抵消正是整个关键所在。
- 在 (i)(b) 和 (ii)(c) 中，忘记说明*为甚么*对所有时间都有 $y=x$（或 $x=y=z$）。这是由初始值为零的指数公式得出的。

STEP 之外

这是对控制理论（control theory）的初步体验。把系统写成 $\dot{\mathbf{r}} = A\mathbf{r} + \mathbf{b}u$，我们找到的组合就是 A 的左特征向量（left eigenvector），而且与 $\mathbf{b}$ 正交（orthogonal），所以任何控制都无法影响它们。卡尔曼（Kalman）的秩条件（rank condition）决定是否每个状态都能被引导到原点。这里该条件不成立，这正是为甚么只有特殊的起点才可行。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）涵盖特征向量方面的内容，而 Boyce 与 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）则涵盖常微分方程的线性系统。

详细解答

STEP III 2024，第 7 题

纯数学 · e^2 的无理性

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

本题在 *Beyond STEP* 第 1.4 节*e 是无理数，e^2 也是*中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch01/e-irrational.html

关键思路

两个级数都是指数级数（exponential series）乘以一个阶乘（factorial）之后的尾项（tail）。把 $\mathrm{e} = \sum 1/r!$ 乘以 $(2n)!$，则直至 $r = 2n$ 的每一项都变成整数，而其后的部分恰好就是 $\mathrm{f}(2n)$；$1/\mathrm{e}$ 与 $\mathrm{g}(2n)$ 亦有同样情况。假如 e^2 是有理数（rational），这些细小尾项的某个组合便必须是整数，但它却被严格夹在 0 与 1 之间。

解答

记

$$a_k = \frac{1}{(n+1)(n+2)\cdots(n+k)}, \qquad k \geq 1,$$

于是 $\mathrm{f}(n) = a_1 + a_2 + a_3 + \cdots$，而 $\mathrm{g}(n) = a_1 - a_2 + a_3 - \cdots$。注意 $a_{k+1} = a_k/(n+k+1)$。

第 (i) 部分

每个 a_k 都是正数，所以 $\mathrm{f}(n) > 0$。a_k 分母中的每个因子都至少是 $n+1$，所以 $a_k \leq 1/(n+1)^k$，而当 $k \geq 2$ 时不等式是严格的（因为出现了因子 $n+2$）。与公比为 $\frac{1}{n+1} < 1$ 的几何级数（geometric series）比较，

$$\mathrm{f}(n) < \sum_{k=1}^{\infty} \frac{1}{(n+1)^k} = \frac{\frac{1}{n+1}}{1 - \frac{1}{n+1}} = \frac{1}{n}.$$

因此 $0 < \mathrm{f}(n) < \frac{1}{n}$。

第 (ii) 部分

相邻各项严格递减：由于 $n+k+1 > 1$，$a_{k+1} = a_k/(n+k+1) < a_k$。从头开始两两分组，

$$\mathrm{g}(n) = (a_1 - a_2) + (a_3 - a_4) + \cdots > 0,$$

因为每个括号都是正数。改为从第一项之后开始分组，

$$\mathrm{g}(n) = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots < a_1 = \frac{1}{n+1},$$

因为每个括号同样是正数。所以 $0 < \mathrm{g}(n) < \frac{1}{n+1}$。

第 (iii) 部分

设 $m = 2n$。利用 $\mathrm{e} = \sum_{r\ge 0} 1/r!$，

$$m!\,\mathrm{e} = \sum_{r=0}^{m} \frac{m!}{r!} + \sum_{r=m+1}^{\infty} \frac{m!}{r!}.$$

当 $r \le m$ 时，$m!/r! = (r+1)(r+2)\cdots m$ 是整数，所以第一个和是整数，记之为 A。在第二个和中，$m!/(m+k)! = 1/\big((m+1)\cdots(m+k)\big)$，这正是 $\mathrm{f}(m)$ 的第 k 项。所以 $m!\,\mathrm{e} = A + \mathrm{f}(m)$，而 $(2n)!\,\mathrm{e} - \mathrm{f}(2n) = A$ 是整数。

同理，由 $1/\mathrm{e} = \sum_{r\ge 0}(-1)^r/r!$ 得

$$\frac{m!}{\mathrm{e}} = \underbrace{\sum_{r=0}^{m} (-1)^r \frac{m!}{r!}}_{B\in\mathbb{Z}} + \sum_{k=1}^{\infty} \frac{(-1)^{m+k}}{(m+1)(m+2)\cdots(m+k)}.$$

> **想一想** 为甚么题目用 $2n$ 而不用 n?

由于 $m = 2n$ 是偶数，$(-1)^{m+k} = (-1)^k$，所以尾项是 $-a_1 + a_2 - a_3 + \cdots = -\mathrm{g}(m)$（此处 a_k 改由 m 构成）。因此 $m!/\mathrm{e} = B - \mathrm{g}(m)$，而 $\dfrac{(2n)!}{\mathrm{e}} + \mathrm{g}(2n) = B$ 是整数。（若 m 为奇数，尾项会是 $+\mathrm{g}(m)$，我们便需要改用减号。）

第 (iv) 部分

利用第 (iii) 部分中的整数 A 和 B，

$$\begin{aligned} q\,\mathrm{f}(2n) + p\,\mathrm{g}(2n) &= q\big((2n)!\,\mathrm{e} - A\big) + p\Big(B - \frac{(2n)!}{\mathrm{e}}\Big) \\ &= (2n)!\Big(q\mathrm{e} - \frac{p}{\mathrm{e}}\Big) - qA + pB \\ &= pB - qA, \end{aligned}$$

因为 $q\mathrm{e} = p/\mathrm{e}$。由于 p, q, A, B 都是整数，所以对每个正整数 n，$q\,\mathrm{f}(2n) + p\,\mathrm{g}(2n)$ 亦是整数。

第 (v) 部分

用反证法（proof by contradiction），假设 e^2 是有理数。由于 $\mathrm{e}^2 > 0$，我们可写 $\mathrm{e}^2 = p/q$，其中 p, q 是正整数；两边除以 e 得 $q\mathrm{e} = p/\mathrm{e}$。由第 (iv) 部分，对每个 $n \geq 1$，$N_n = q\,\mathrm{f}(2n) + p\,\mathrm{g}(2n)$ 都是整数。

> **想一想** 怎样才能令 N_n 不是整数?

把第 (i) 及 (ii) 部分应用于 $2n$，将它夹住。由于 $p, q > 0$，

$$0 < N_n < \frac{q}{2n} + \frac{p}{2n+1} < \frac{p+q}{2n}.$$

现在取足够大的 n，使 $2n > p + q$。那么 $0 < N_n < 1$，即 N_n 是严格介乎 0 与 1 之间的整数，这是不可能的。因此 e^2 是无理数（irrational）。

验算

细小的情况可以准确求和。当 $n = 2$ 时，第 (iii) 部分指出 $24\mathrm{e} - \mathrm{f}(4)$ 是整数；事实上 $\sum_{r=0}^{4} 24/r! = 65$，而 $\mathrm{f}(4) = 24\mathrm{e} - 65 \approx 0.2388$，稳稳落在 $(0, \frac{1}{4})$ 之内。同样地，$\mathrm{f}(2) = 2\mathrm{e} - 5 \approx 0.4366 < \frac{1}{2}$，而 $\mathrm{g}(2) = 1 - 2/\mathrm{e} \approx 0.2642 < \frac{1}{3}$。电脑验算确认了：(i) 和 (ii) 的界限对 $n \leq 24$ 成立；(iii) 的整数性对 $2n \leq 22$ 成立；以及对随机的 p, q，$0 < q\,\mathrm{f}(2n) + p\,\mathrm{g}(2n) < (p+q)/(2n)$ 成立。

常见错误

- 声称 $a_k < 1/(n+1)^k$ 对所有 k 成立：在 $k = 1$ 时等号成立，所以严格不等式来自 $k \geq 2$ 的各项。
- 忘记 $1/\mathrm{e}$ 展开式中尾项的符号：它之所以是 $-\mathrm{g}(2n)$，只因为 $2n$ 是偶数。
- 停在“N_n 很小”这一步：要得出矛盾，还需要 $N_n > 0$，这正是 (i) 和 (ii) 中下界重要的原因。

STEP 之外

这是傅里叶（Fourier）证明 e 为无理数的经典方法，并通过把 $\mathrm{e}^2 = p/q$ 拆成 $q\mathrm{e} = p/\mathrm{e}$ 加以推广，使两个级数都快速收敛（converge）。同一思路经过更细致的处理，可证明对

每个非零有理数 r，e^r 都是无理数；刘维尔（Liouville）亦用它证明了 e 不是任何整系数二次方程（quadratic equation）的根。估计级数尾项的界是大学一年级分析（analysis）的核心技巧；可参阅 Abbott 的 *Understanding Analysis*（《理解分析》）。

详细解答

STEP III 2024，第 8 题

纯数学 · 相交圆锥曲线

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

形如 $y^2 - x^2 + \ldots$ 的二次式（quadratic expression），当且仅当配方（completing the square）后不留下常数项时，才会分解为两条斜率（gradient）为 ± 1 的直线。题目其余部分的窍门在于：把一条曲线方程的某个倍数加到另一条曲线的方程上，不会失去任何公共点；而选取恰当的倍数，便能把棘手的四次（quartic）问题化为“哪些直线与抛物线（parabola）相交”，而这只不过是两个二次方程。

解答

第 (i) 部分

乘积为零当且仅当其中一个因子为零，因此满足 $(y - x + 3)(y + x - 5) = 0$ 的点，就是 $y = x - 3$（斜率为 1）上的点，连同 $y = -x + 5$（斜率为 -1）上的点。

对于一般方程，就两个变量分别配方：

$$y^2 - x^2 + py + qx + r = \left(y + \tfrac{1}{2}p\right)^2 - \left(x - \tfrac{1}{2}q\right)^2 - K, \qquad \text{其中} K = \tfrac{1}{4}\left(p^2 - q^2\right) - r.$$

设 $Y = y + \frac{1}{2}p$ 及 $X = x - \frac{1}{2}q$，这是一个平移（translation），不会改变斜率。方程变为 $(Y - X)(Y + X) = K$。

若 $p^2 - q^2 = 4r$, 则 $K = 0$, 方程为 $(Y - X)(Y + X) = 0$; 由第一段可知, 这是直线 $Y = X$ 与 $Y = -X$ 组成的一对直线，斜率分别为 1 和 -1。在原坐标下，它们是

$$y - x + \tfrac{1}{2}(p + q) = 0 \quad \text{及} \quad y + x + \tfrac{1}{2}(p - q) = 0.$$

反过来，假设该方程表示一对斜率为 ± 1 的直线；我们证明 $K = 0$。其中一条直线的斜

率为 1，故具有 $Y = X + c$ 的形式，而且直线上每一点都满足该方程。在这条直线上，$(Y - X)(Y + X) = c(2X + c)$，所以对每一个实数 X 都有 $2cX + c^2 = K$。X 的线性函数只有在其系数为零时才是常数，所以 $c = 0$，于是 $K = c^2 = 0$。因此 $p^2 - q^2 = 4r$，“当且仅当”（if and only if）得证。

第 (ii) 部分

若某点位于 C_1 上，则该处有 $y^2 + 2sy + s(s+1) - x = 0$；若它位于 C_2 上，则该处有 $y - x^2 = 0$。所以在公共点处，表达式

$$y^2 + 2sy + s(s+1) - x + k(y - x^2)$$

的值为 $0 + k \cdot 0 = 0$，不论 k 取何值。

取 $k = 1$，方程为

$$y^2 - x^2 + (2s+1)y - x + s(s+1) = 0,$$

它具有第 (i) 部分的形式，其中 $p = 2s + 1$、$q = -1$、$r = s(s+1)$。于是

$$p^2 - q^2 = 4s^2 + 4s + 1 - 1 = 4s(s+1) = 4r,$$

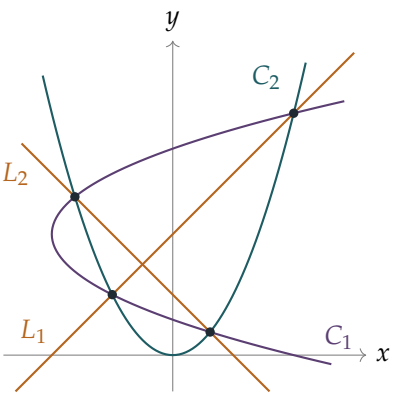
所以由第 (i) 部分，这是一对斜率为 1 和 -1 的直线。利用第 (i) 部分的因式分解（factorisation），并注意 $\frac{1}{2}(p+q) = s$ 及 $\frac{1}{2}(p-q) = s+1$，方程为

$$(y - x + s)(y + x + s + 1) = 0,$$

所以这两条直线是

$$L_1 : \ y = x - s \qquad 及 \qquad L_2 : \ y = -x - s - 1.$$

由第一段，$C_1 \cap C_2$ 的每一点都满足这个方程，所以全部四个交点都位于 $L_1 \cup L_2$ 上。（这对每个 s 都成立；我们从来不需要这四个点存在。）



图中所示为 $s = -2$ 的情况：C_1 为 $x = (y-2)^2 - 2$，而 C_1 与 C_2 的四个公共点分别有两个位于直线 $y = x + 2$ 上、两个位于直线 $y = 1 - x$ 上。

第 (iii) 部分

C_1 与 C_2 的每个公共点都位于 C_2 上，并位于 L_1 或 L_2 上。代入 $y = x^2$：

$$\begin{aligned} \text{对于} L_1: \quad & x^2 - x + s = 0, \\ \text{对于} L_2: \quad & x^2 + x + s + 1 = 0. \end{aligned}$$

C_2 上的点由其 x 坐标唯一确定，所以每条直线与 C_2 至多相交于两点。因此，要有四个不同的公共点，每个二次方程都必须有两个不同的实根，所以两个判别式（discriminant）都为正：

$$1 - 4s > 0 \quad \text{及} \quad 1 - 4(s+1) > 0.$$

第二个不等式给出 $s < -\frac{3}{4}$，即所求。（第一个不等式 $s < \frac{1}{4}$ 随之自动成立。）

> **想一想** 在第 (iv) 部分，为甚么只找出四个同时位于 C_2 和 $L_1 \cup L_2$ 上的点并不足够?

第 (ii) 部分只证明了 C_1 与 C_2 的公共点位于这两条直线上；我们还需要反方向的结论，而且必须确保这四个点不会重合。

第 (iv) 部分

设 $s < -\frac{3}{4}$。那么上述两个判别式都为正（$-3 - 4s > 0$，以及 $1 - 4s > 4 > 0$），所以 L_1 与 C_2 相交于两个不同的点，L_2 亦然。

这四个点互不相同。若某点被重复计算，它便同时位于 L_1 和 L_2 上。两直线相交于 $x - s = -x - s - 1$ 之处，即点 $\left(-\frac{1}{2}, -\frac{1}{2} - s\right)$。此点只在 $-\frac{1}{2} - s = \frac{1}{4}$，即 $s = -\frac{3}{4}$ 时才位于 C_2 上，而这已被排除。所以我们得到四个不同的点。

它们每一个都位于 C_1 上。在每一点处都有 $y - x^2 = 0$（因为它在 C_2 上），而第 (ii) 部分 $k = 1$ 的方程亦成立（因为它在 L_1 或 L_2 上）。两式相减，得

$$y^2 + 2sy + s(s+1) - x = 0,$$

这正是 C_1 的方程。

所以 C_1 与 C_2 至少有四个不同的公共点；由第 (iii) 部分的论证，它们不可能多于四个。因此两者恰好相交于四个不同的点。

验算

- 把 $y = x^2$ 代入 C_1，得四次方程 $x^4 + 2sx^2 - x + s^2 + s = 0$。展开 $(x^2 - x + s)(x^2 + x + s + 1)$ 正好得到此式，印证了这对直线。

- 当 $s=-2$ 时，四个点为 $(2,4)$、$(-1,1)$ 及 $\left(\frac{1}{2}(-1\pm\sqrt{5}),\ \frac{1}{2}(3\mp\sqrt{5})\right)$；每一点都满足 $x=(y-2)^2-2$，例如 $(2,4)$：$(4-2)^2-2=2$。
- 在边界 $s=-\frac{3}{4}$ 时，两个二次方程为 $x^2-x-\frac{3}{4}=0$（根为 $\frac{3}{2},-\frac{1}{2}$）及 $\left(x+\frac{1}{2}\right)^2=0$，所以只有两个不同的公共点，其中 $x=-\frac{1}{2}$ 为两条直线所共有。用电脑计算该四次方程的实根数目，在 $s=-0.76,-1,-2$ 时得 4 个，在 $s=-0.74,-0.5$ 时得 2 个。

常见错误

- 在第 (i) 部分只证明“若”的方向：一对直线必然给出 $p^2-q^2=4r$，这一点需要论证（例如上文的论证）。
- 在第 (iv) 部分，忘记检查两条直线没有共用 C_2 上的点，或忘记证明所找到的点确实位于 C_1 上。

STEP 之外

方程族 $F+kG=0$ 称为二次曲线束（pencil of conics）：族中每一条曲线都经过 $F=0$ 与 $G=0$ 的公共点。选取 k 使该曲线退化（degenerate）为一对直线，是求两条二次曲线（conics）交点的经典方法。在射影几何（projective geometry）中，做法是令一个 3×3 行列式（determinant）为零，从而得到一个关于 k 的三次方程。在本题中，不论 s 为何值，曲线束在 $k=1$ 时恰好都是退化的。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）讲解了二次曲线背后的对称矩阵（symmetric matrices）。Siklos 的 *Advanced Problems in Mathematics*（《高等数学难题》）收录了更多同类型的题目。

详细解答

STEP III 2024，第 9 题

力学 · 斜碰撞

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

一切都在桌面这个平面上发生，所以不妨从上方俯视：每个球都变成一个以其最低点为圆心、半径为 r 的圆。对于光滑（smooth）球体，冲量（impulse）沿连心线（line of centres）作用，因此我们把速度分解为沿这条线的分量（动量守恒和恢复系数在此适用）和垂直于这条线的分量（保持不变）。此后 B 沿直线运动，而它会否撞到 C，就变成一个关于点到直线距离的问题。

解答

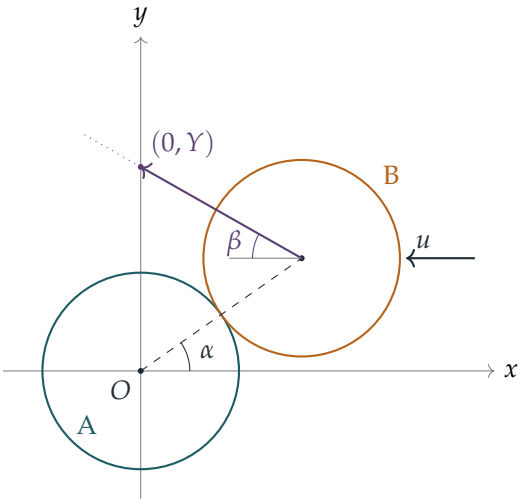
我们把球体视为光滑且质量相等，放在光滑桌面上，因此任何地方都没有摩擦，而除了碰撞的瞬间外，每个球都以恒定速率沿直线运动。我们以最低点来处理，因为最低点在水平方向的运动与球心完全相同；两个半径为 r 的球，当它们的最低点相距 $2r$ 时便互相接触。

第 (i) 部分

B 的最低点沿直线 $y = 2r\sin\alpha$ 移动。当它与 O 的距离为 $2r$ 时便接触 A，即在点 $(2r\cos\alpha,\ 2r\sin\alpha)$（取正的 x 坐标，因为 B 从 x 值很大的正方向驶来）。因此碰撞时，从 A 指向 B 的连心线的单位向量（unit vector）为

$$\mathbf{n} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad \text{而} \qquad \mathbf{t} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$$

与之垂直。



碰撞前 B 的速度为 $\begin{pmatrix} -u \\ 0 \end{pmatrix}$，沿 **n** 的分量为 $-u\cos\alpha$，沿 **t** 的分量为 $u\sin\alpha$。A 处于静止。

由于球体光滑，两者之间的冲量沿 **n** 方向。因此 B 的 **t** 分量保持为 $u\sin\alpha$（而 A 不会获得任何 **t** 分量）。设碰撞后沿 **n** 的分量分别为 v_A 和 v_B。沿 **n** 方向的动量守恒（momentum conservation）（质量相等）及牛顿（Newton）恢复定律（law of restitution）给出

$$\begin{aligned} v_A + v_B &= -u\cos\alpha, \\ v_B - v_A &= \tfrac{1}{3}\big(0 - (-u\cos\alpha)\big) = \tfrac{1}{3}u\cos\alpha. \end{aligned}$$

（沿 **n** 的接近速率为 $u\cos\alpha$，而位于 **n** 正方向一侧的 B 与 A 分离。）两式相加，得 $v_B = -\frac{1}{3}u\cos\alpha$，以及 $v_A = -\frac{2}{3}u\cos\alpha$。

所以碰撞后 B 的速度为

$$\begin{aligned} -\tfrac{1}{3}u\cos\alpha \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} + u\sin\alpha \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix} &= \begin{pmatrix} -\frac{1}{3}u(\cos^2\alpha + 3\sin^2\alpha) \\ \frac{2}{3}u\sin\alpha\cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{3}u(1 + 2\sin^2\alpha) \\ \frac{2}{3}u\sin\alpha\cos\alpha \end{pmatrix}, \end{aligned}$$

这里用了 $\cos^2\alpha = 1 - \sin^2\alpha$。这就是所求的结果。

第 (ii) 部分

两个分量都不为零，其中 x 分量为负，y 分量为正（因为 α 是锐角）。所以 B 此时向左上方运动，其路径斜率的大小为

$$\frac{\frac{2}{3}u\sin\alpha\cos\alpha}{\frac{1}{3}u(1 + 2\sin^2\alpha)} = \frac{2\sin\alpha\cos\alpha}{1 + 2\sin^2\alpha} = \tan\beta,$$

其中 β 是路径与 x 负方向所成的锐角。B 的最低点从 $(2r\cos\alpha, 2r\sin\alpha)$ 出发。要到达 y 轴，x 须减少 $2r\cos\alpha$，同时 y 增加 $2r\cos\alpha\tan\beta$。所以它在 $(0, Y)$ 穿过 y 轴，其中

$$Y = 2r\sin\alpha + 2r\cos\alpha\tan\beta = 2r(\cos\alpha\tan\beta + \sin\alpha),$$

正如所求。

第 (iii) 部分

碰撞后 B 以恒定速率沿直线运动，沿着经过 $(0, Y)$、单位方向向量为 $\mathbf{d} = (-\cos\beta, \sin\beta)$ 的直线 ℓ。C 一直静止于 $(0, h)$，直至被撞到为止。B 撞到 C 当且仅当 $(0, h)$ 到 B 的路径的距离跌到 $2r$ 以下。

$(0, h)$ 到 ℓ 的垂直距离（perpendicular distance），就是 $(0, h) - (0, Y) = (0, h - Y)$ 垂直于 $\mathbf{d}$ 的分量的大小。ℓ 的一个单位法向量为 $(\sin\beta, \cos\beta)$，所以这个距离是

$$\big|(0,\, h - Y)\cdot(\sin\beta, \cos\beta)\big| = |h - Y|\cos\beta,$$

因为 β 是锐角，所以 $\cos\beta > 0$。若 $h > Y + 2r\sec\beta$，则 $h - Y > 2r\sec\beta > 0$，而距离为 $(h - Y)\cos\beta > 2r$。B 的路径上每一点（不只是整条直线上最近的那一点）与 C 的最低点的距离都至少有这么远。所以 B 永远不会进入 C 的 $2r$ 范围内，也就不会撞到它。

（也没有其他东西阻碍：A 以速度 $-\frac{2}{3}u\cos\alpha\,\mathbf{n}$ 离开，进入 $y < 0$ 的区域，远离 B 和 C。）

> **想一想** 第 (iv) 部分要求我们证明 $Y < 2r\sec\beta$。能否把它化为单一个三角恒等事实?

可以：两边乘以 $\cos\beta > 0$，左边就变成一条复角公式（compound-angle formula）。

第 (iv) 部分

除以 $2r$ 并乘以 $\cos\beta > 0$，不等式 $Y < 2r\sec\beta$ 等价于

$$\cos\alpha\sin\beta + \sin\alpha\cos\beta < 1, \qquad \text{即} \qquad \sin(\alpha + \beta) < 1.$$

显然 $\sin(\alpha + \beta) \le 1$，且等号只在 $\alpha + \beta = \frac{\pi}{2}$ 时成立（因为 $0 < \alpha + \beta < \pi$）。这意味着 $\tan\beta = \cot\alpha$，即

$$\frac{2\sin\alpha\cos\alpha}{1 + 2\sin^2\alpha} = \frac{\cos\alpha}{\sin\alpha} \iff 2\sin^2\alpha = 1 + 2\sin^2\alpha$$

（除以 $\cos\alpha \neq 0$），这是不可能的。因此 $\sin(\alpha + \beta) < 1$，从而 $Y < 2r\sec\beta$。

由此。现在

$$Y + 2r\sec\beta < 4r\sec\beta,$$

所以我们需要 $\sec\beta$ 在所有锐角 α 下的上界，即 $\tan\beta$ 的上界。分子分母同除以 $\cos^2\alpha$，并记 $t=\tan\alpha>0$，

$$\tan\beta=\frac{2\tan\alpha}{\sec^2\alpha+2\tan^2\alpha}=\frac{2t}{1+3t^2}\le\frac{2t}{2\sqrt{3}\,t}=\frac{1}{\sqrt{3}},$$

这里用了算术—几何平均不等式（AM–GM inequality）$1+3t^2\ge 2\sqrt{3}\,t$（等号在 $t=1/\sqrt{3}$ 即 $\alpha=\frac{\pi}{6}$ 时成立）。所以 $0<\beta\le\frac{\pi}{6}$，从而 $\sec\beta\le\sec\frac{\pi}{6}=\frac{2}{\sqrt{3}}$。因此对每个锐角 α，

$$Y+2r\sec\beta<4r\sec\beta\le\frac{8r}{\sqrt{3}}.$$

若 $h>\dfrac{8r}{\sqrt{3}}$，则 $h>Y+2r\sec\beta$，由第 (iii) 部分可知，无论 α 取何值，B 都不会撞到 C。

验算

- 动量：A 的速度 $-\frac{2}{3}u\cos\alpha\,(\cos\alpha,\sin\alpha)$ 加上第 (i) 部分中 B 的速度，总和为 $(-u,0)$，符合质量相等时应有的结果。
- 当 $\alpha=0$（正面碰撞）时，公式给出 B 的速度为 $(-\frac{1}{3}u,0)$，而 A 得到 $(-\frac{2}{3}u,0)$：正是质量相等、$e=\frac{1}{3}$ 时常见的一维答案。
- 取 $\alpha=\frac{\pi}{6}$，得 $\tan\beta=1/\sqrt{3}$，所以 $\beta=\frac{\pi}{6}$，$\alpha+\beta=\frac{\pi}{3}$，而 $Y=2r(\frac{\sqrt{3}}{2}\cdot\frac{1}{\sqrt{3}}+\frac{1}{2})=2r$，明显小于 $2r\sec\beta=4r/\sqrt{3}\approx 2.31r$。于是 $Y+2r\sec\beta\approx 4.31r$，低于 $8r/\sqrt{3}\approx 4.62r$。
- 我们用电脑对 20,000 个随机的 α 值作了验证：逐步推进 B 的直线路径，得到的穿越点正是上面的 Y；在略高于 $h=Y+2r\sec\beta$ 时 B 避过 C，略低于此值时则撞到 C；而 $\sec\beta$ 的最大值为 $1.1547\ldots=2/\sqrt{3}$。

常见错误

- 把恢复系数应用于 x 分量，而不是沿连心线的分量。光滑球体之间的冲量沿连心线作用，而这里连心线与 x 轴成角 α。
- 以为碰撞时连心线沿 x 轴或 y 轴。接触点是由两球相距 $2r$ 这一条件决定的。
- 在第 (iii) 部分中，只检验 C 的最低点是否在单一点 $(0,Y)$ 的 $2r$ 范围内，而不是整条路径的 $2r$ 范围内。

STEP 之外

把速度分解为沿某个选定单位向量及垂直于它的分量，就是投影到一组标准正交基（orthonormal basis）上，这个概念贯穿大学一年级的线性代数和力学。光滑球体的斜碰

（oblique collision）是散射（scattering）的入门模型，物理学家用它来研究气体和粒子碰撞。Kleppner 与 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）就以这种向量方法处理二维碰撞。

详细解答

STEP III 2024，第 10 题

力学 • 倾倒与滑动

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

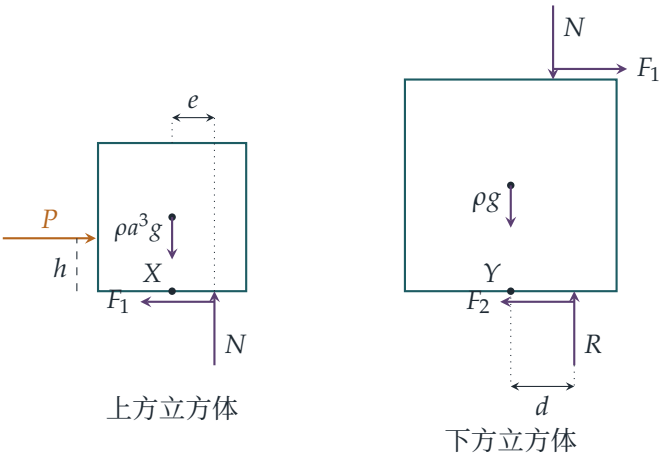
www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

每一部分都是在比较平衡（equilibrium）被打破的四种方式：上方立方体滑动、下方立方体滑动、上方立方体翻倒（topple），或两个立方体一同翻倒。每一种情况各自在某个 P 的临界值发生；当 P 由零开始增大时，哪个临界值最小，哪种情况就先发生。选取适当的点取力矩（moments），即对整个系统取关于 Y 的力矩、对上方立方体单独取关于 X 的力矩，便可以毫不费力地求得第 (i) 部分中法向反作用力（normal reaction）的位置；而当其中一个反作用力到达其作用面的边缘时，翻倒便开始。

解答

我们把两个立方体视为刚体（rigid body）。下方立方体的质量为 ρ，上方立方体的质量为 ρa^3。设 P 水平向右作用，作用点位于下方立方体顶面之上高度 h 处，即离地面高度 $1+h$。由于 P 是水平的，它关于任何一点的力矩只取决于其高度，而与它作用于该面上的哪个位置无关。



图中 N 和 F_1 是两个立方体之间的法向反作用力和摩擦力（friction）（根据牛顿第三定律（Newton's third law），在每个立方体上分别画出），作用点与 X 的距离为 e；R 和 F_2 是地面施加的法向反作用力和摩擦力，作用点与 Y 的距离为 d。

第 (i) 部分

整个系统。外力只有 P、两个重量、R 和 F_2。分解各力，得

$$F_2 = P, \qquad R = \rho g + \rho a^3 g = (1 + a^3)\rho g.$$

对 Y 取力矩。两个重量都沿着通过 Y 的铅垂线作用（每个质心（centre of mass）都在 Y 的正上方），而 F_2 沿着地面通过 Y 作用，所以这些力都没有力矩。剩下的就是高度为 $1 + h$ 的 P，与距离为 d 的 R 互相平衡：

$$R\,d = P(1 + h) \quad \Longrightarrow \quad d = \frac{P(1 + h)}{(1 + a^3)\rho g},$$

此距离由 Y 沿 P 的方向量度。这正是所求的结果。

单独考虑上方立方体。分解各力，得 $F_1 = P$ 及 $N = \rho a^3 g$。对 X 取力矩，重量和 F_1 都没有力矩，所以 $Ne = Ph$，从而

$$e = \frac{Ph}{\rho a^3 g},$$

同样由 X 沿 P 的方向量度。

第 (ii) 部分

在任何物体移动之前，摩擦力分别为两个立方体之间的 $F_1 = P$ 及地面上的 $F_2 = P$。当某接触处所需的摩擦力超过 μ 乘以该处的法向反作用力时，该处便会滑动（slip），所以

$$\text{上方立方体在下方立方体上滑动的条件为} P > \mu\rho a^3 g,$$
$$\text{下方立方体在地面上滑动的条件为} P > \mu(1+a^3)\rho g.$$

由于 $a^3 < 1 + a^3$，第一个临界值必定严格较小。当 P 由零开始增大时（假如没有先发生翻倒），上方立方体会先到达其极限并开始滑动，而此时地面的摩擦力仍远未到达极限。上方立方体一旦开始滑动，它施加于下方立方体的摩擦力为 $\mu\rho a^3 g$，仍然小于地面所能提供的 $\mu(1+a^3)\rho g$，因此下方立方体保持不动。所以打破平衡的是上方立方体的滑动，而不是下方立方体的滑动。

第 (iii) 部分

当作用于一个平面上的物体所受的法向反作用力到达该面的边缘时，物体便即将翻倒。上方立方体底面的半宽为 $\frac{1}{2}a$，下方立方体底面的半宽为 $\frac{1}{2}$。因此，以 ρg 为单位量度所有力，P 的四个临界值为

$$P_{\text{滑动}} = \mu a^3 \text{（上方在下方上）}, \qquad P_{\text{地面}} = \mu(1+a^3),$$
$$P_{\text{上翻}} = \frac{a^4}{2h} \text{（由} e = \tfrac{1}{2}a \text{ 得）}, \qquad P_{\text{同翻}} = \frac{1+a^3}{2(1+h)} \text{（由} d = \tfrac{1}{2} \text{ 得）}.$$

下方立方体翻倒时会把上方立方体一并带倒，这就是为甚么 $d = \frac{1}{2}$ 表示两个立方体一同翻倒。

现在设 $a = 1$，于是 $P_{\text{滑动}} = \mu$、$P_{\text{上翻}} = \frac{1}{2h}$ 及 $P_{\text{同翻}} = \frac{1}{1+h}$。根据第 (ii) 部分，我们可以不理会 $P_{\text{地面}}$。

> ? **想一想** 我们需要担心上方立方体单独翻倒吗?

不需要，但必须加以验证。由于 $h < a = 1$，我们有 $2h < 1 + h$，所以

$$P_{\text{上翻}} = \frac{1}{2h} > \frac{1}{1+h} = P_{\text{同翻}}.$$

两个立方体会在上方立方体能够单独翻倒之前一同翻倒。因此只需比较 $P_{\text{滑动}} = \mu$ 与 $P_{\text{同翻}} = 1/(1+h)$：

- 若 $\mu(1+h) < 1$，则 $\mu < \frac{1}{1+h}$，滑动的临界值先到达，上方立方体在下方立方体上滑动；
- 若 $\mu(1+h) > 1$，则 $\frac{1}{1+h} < \mu$，翻倒的临界值先到达，两个立方体一同翻倒。

第 (iv) 部分

若没有发生滑动，便只需比较 $P_{\text{上翻}}$ 与 $P_{\text{同翻}}$。上方立方体先翻倒，当且仅当

$$\frac{a^4}{2h} < \frac{1+a^3}{2(1+h)} \iff a^4(1+h) < h(1+a^3) \iff a^4 < h\big(1+a^3-a^4\big) = h\big(1+a^3(1-a)\big),$$

其中我们在两边乘以正数 $2h(1+h)$。这正是题目所给的条件，所以在此条件下，打破平衡的是上方立方体的翻倒。

第 (v) 部分

当 $a=\frac{1}{2}$ 时，我们需要满足三个条件。

- 第 (iv) 部分的条件：$h\big(1+\frac{1}{8}\cdot\frac{1}{2}\big) > \frac{1}{16}$，即 $\frac{17}{16}h > \frac{1}{16}$，所以 $h > \frac{1}{17}$。
- 题目设置要求 $h < a = \frac{1}{2}$。
- 翻倒之前没有滑动：$P_{\text{上翻}} < P_{\text{滑动}}$，即 $\dfrac{1/16}{2h} < \dfrac{\mu}{8}$，亦即 $\mu > \dfrac{1}{4h}$。（根据第 (ii) 部分，在地面上的滑动永远不会最先发生。）由于我们同时需要 $\mu < 1$，这要求 $h > \frac{1}{4}$。

所以选取 $h=\frac{2}{5}$，它满足 $\frac{1}{4} < h < \frac{1}{2}$ 及 $h > \frac{1}{17}$。这时我们需要 $\frac{5}{8} < \mu < 1$；取 $\mu = \frac{3}{4}$。于是各临界值（以 ρg 为单位）为

$$P_{\text{上翻}} = \tfrac{5}{64} \approx 0.078, \quad P_{\text{滑动}} = \tfrac{3}{32} \approx 0.094, \quad P_{\text{同翻}} = \tfrac{45}{112} \approx 0.402, \quad P_{\text{地面}} = \tfrac{27}{32} \approx 0.844.$$

最小的是 $P_{\text{上翻}}$，所以当 P 增大时，上方立方体会在任何滑动发生之前翻倒：第 (iv) 部分的情况确实会出现。

验算

- 当 $P \to 0$ 时，d 和 e 都趋向 0：没有推力时，每个反作用力都作用于质心的正下方，理应如此。
- 单独对下方立方体取力矩（其重量、距离为 d 的 R 和 F_2，以及方向相反、与 X 距离为 e 的 N 和 F_1），得 $Rd = Ne + F_1\cdot 1$，即 $P(1+h) = Ph + P$。所以第 (i) 部分的两个答案互相吻合。
- 对于 $a=\frac{1}{2}$、$h=\frac{2}{5}$、$\mu=\frac{3}{4}$，上述四个临界值都经过精确计算；而第 (ii)–(iv) 部分中的比较亦已用电脑就大量随机选取的 a、h 和 μ 加以验证。

常见错误

- 在第 (iii) 部分中，只比较滑动与两个立方体一同翻倒，却从未证明上方立方体不会先翻倒。
- 对上方立方体取关于地面上某点的力矩，这样会把接触面上的摩擦力 F_1 牵涉进来；对 X 取力矩则可避免。
- 在整个系统的力矩方程中，误把 h 而不是 $1+h$ 当作 P 离地面的高度。

STEP 之外

这里每一种失效方式都是关于各力的一条不等式，只要所有不等式同时成立，平衡便得以维持；最先失效的，就是临界值最小的那一条。这种“一组线性约束中有一条起决定作用（binding）”的想法，正是工程学中用于砌体及堆叠结构的极限分析（limit analysis）的核心，更广泛而言亦是线性规划（linear programming）的核心。若想深入了解摩擦力、扭矩（torque）及刚体静力学，可参阅 Kleppner 和 Kolenkow 所著的 *An Introduction to Mechanics*（《力学导论》）。

详细解答

STEP III 2024，第 11 题

概率与统计 · 二项分布的期望值

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

第 (i) 部分的恒等式（identity）把和式下半部分的每一项 $r\binom{2n}{r}$ 与上半部分的一项配对起来。这种对称性（symmetry）正是处理 X 所需要的：正面与反面次数中较大者，不是正面的次数，就是反面的次数；把正面和反面的角色互换，便能把期望值（expectation）折叠到和式的上半部分。之后，第 (iii) 部分只是一个比值检验（ratio test），而第 (iv) 部分只需一行。

解答

第 (i) 部分

当 $1 \le r \le 2n$ 时，把 r 与 $r!$ 约去，得

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

另一边，$2n-(2n+1-r) = r-1$，所以

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

同样是把分子的因子与阶乘（factorial）约去。（我们需要 $1 \le r \le 2n$，使 $2n+1-r$ 介乎 1 与 $2n$ 之间，这样两个阶乘才有意义。）两个表达式相同，故恒等式成立。

至于和式，$r = 0$ 的项为零。把其余各项分成 $1 \le r \le n$ 和 $n+1 \le r \le 2n$ 两组。在较低的一组中，设 $s = 2n+1-r$。当 r 由 1 走到 n，s 便由 $2n$ 递减至 $n+1$，而根据恒等式，

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

因此较低一组等于较高一组，从而

$$\sum_{r=0}^{2n} r\binom{2n}{r} = 2\sum_{r=n+1}^{2n} r\binom{2n}{r},$$

得证。

第 (ii) 部分

设 H 为正面的次数，则 $H \sim \mathrm{B}(2n, \frac{1}{2})$，且 $\mathrm{P}(H=r) = \binom{2n}{r}/2^{2n}$。反面的次数为 $2n-H$，所以 $X = \max(H, 2n-H)$。即是说，当 $H \geq n+1$ 时 $X = H$；当 $H = n$ 时 $X = n$；当 $H \leq n-1$ 时 $X = 2n-H$。因此

$$2^{2n}\,\mathrm{E}(X) = \sum_{r=n+1}^{2n} r\binom{2n}{r} + n\binom{2n}{n} + \sum_{r=0}^{n-1}(2n-r)\binom{2n}{r}.$$

在最后一个和式中设 $s = 2n-r$，它由 $2n$ 递减至 $n+1$。由于 $\binom{2n}{r} = \binom{2n}{2n-r}$，该和式变成 $\sum_{s=n+1}^{2n} s\binom{2n}{s}$，与第一个和式相同。（这就是正反面的对称性：得到 s 次反面的机会与得到 s 次正面的机会相同。）

> **想一想** $\sum_{r=n+1}^{2n} r\binom{2n}{r}$ 只涵盖一半的范围，该如何求它的值?

这正是第 (i) 部分的用途。题目已给出全范围的版本：以 $2n$ 代替 n，得 $\sum_{r=0}^{2n} r\binom{2n}{r} = 2n\cdot 2^{2n-1}$。所以由第 (i) 部分，

$$2\sum_{r=n+1}^{2n} r\binom{2n}{r} = 2n\cdot 2^{2n-1} = n\,2^{2n}.$$

因此

$$2^{2n}\,\mathrm{E}(X) = n\,2^{2n} + n\binom{2n}{n}, \qquad \text{所以} \qquad \mathrm{E}(X) = n\left(1 + \frac{1}{2^{2n}}\binom{2n}{n}\right),$$

得证。

第 (iii) 部分

设 $a_n = \binom{2n}{n}/2^{2n}$，它是正数。于是

$$\frac{a_{n+1}}{a_n} = \frac{(2n+2)!}{(n+1)!\,(n+1)!}\cdot\frac{n!\,n!}{(2n)!}\cdot\frac{1}{4} = \frac{(2n+2)(2n+1)}{4(n+1)^2} = \frac{2n+1}{2n+2},$$

这里用了 $2n+2 = 2(n+1)$。这个比值小于 1，而 $a_n > 0$，所以对每个 $n \geq 1$ 都有 $a_{n+1} < a_n$。此数列（sequence）严格递减（strictly decreasing）。

第 (iv) 部分

你赢得的金额正是第 (ii) 部分的 X，而你要付 £n，所以每付一英镑的期望赢得金额为

$$\frac{\mathrm{E}(X)}{n} = 1 + \frac{1}{2^{2n}}\binom{2n}{n} = 1 + a_n.$$

由第 (iii) 部分，n 最小时 a_n 最大，所以你应选 $n = 1$。这时每付一英镑可取回 $1 + \frac{1}{2} = \frac{3}{2}$ 英镑，即每英镑的期望利润为 50 便士（50p）。（若“赢得金额”指的是利润，则为 $\mathrm{E}(X)/n - 1 = a_n$，结论不变。）

验算

- 当 $n = 1$（掷两次）时，HH 或 TT 对应 $X = 2$，HT 或 TH 对应 $X = 1$，所以 $\mathrm{E}(X) = \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 1 = \frac{3}{2}$。公式给出 $1\cdot(1 + \frac{2}{4}) = \frac{3}{2}$。
- 当 $n = 2$ 时，公式给出 $2(1 + \frac{6}{16}) = \frac{11}{4}$，所以每英镑的回报为 $\frac{11}{8} = 1.375 < 1.5$。当 $n = 3$ 时，回报为 $\frac{21}{16} = 1.3125$，正如第 (iii) 部分所预测，仍在递减。
- 对于 $n \le 5$，已用电脑逐一检查全部 2^{2n} 个掷币序列，确认 $\mathrm{E}(X)$ 的公式正确。

常见错误

- 把 $H = n$ 的项计算两次，或完全漏掉。这一项不属于分割的任何一边，只贡献一次 $n\binom{2n}{n}$。
- 在第 (iv) 部分中，去最大化会无限增大的 $\mathrm{E}(X)$ 本身，而不是每付一英镑的 $\mathrm{E}(X)$。

STEP 之外

超出的部分 $\mathrm{E}(X) - n = na_n$ 是简单随机漫步（simple random walk）与起点之期望距离的一半。由史特灵公式（Stirling's formula），$a_n \approx 1/\sqrt{\pi n}$，所以这个超出量的增长与 $\sqrt{n/\pi}$ 相若：正面与反面次数的典型差距，按掷币次数的平方根增长。与此同时，每英镑的超出量 a_n 趋向 0，这正是大数定律（law of large numbers）在起作用。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（《概率导论》）讲述了随机漫步及这些极限。

详细解答

STEP III 2024，第 12 题

概率与统计 · 几何概率

STEP III 2024，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 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

若在单位正方形（unit square）内均匀（uniformly）随机选取一点，则 $\mathrm{P}(R \le r)$ 正是正方形内落在以原点为圆心、半径为 r 的圆之内那部分的面积。当 $r > 1$ 时，圆会从正方形的两条边突出，而留在正方形内的区域可以整齐地分拆为两个直角三角形和一个扇形（sector）。至于期望值（expectation），先求导得出概率密度函数（density），再对 $rf(r)$ 积分；用一次分部积分（integration by parts）便能消去麻烦的 $\cos^{-1}$ 项，最后以双曲代换（hyperbolic substitution）完成计算。

解答

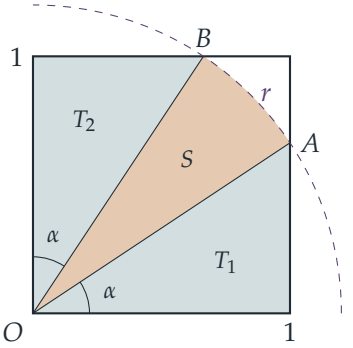
第 (i) 部分

由于概率等于面积，$\mathrm{P}(R \le r)$ 就是正方形内与 O 的距离不超过 r 的点所组成集合的面积。

当 $0 \le r \le 1$ 时，四分之一圆盘 $x, y \ge 0$，$x^2 + y^2 \le r^2$ 完全位于正方形之内，因此

$$\mathrm{P}(R \le r) = \tfrac{1}{4}\pi r^2 \qquad (0 \le r \le 1).$$

当 $1 \le r \le \sqrt{2}$ 时，圆 $x^2 + y^2 = r^2$ 与边 $x = 1$ 相交于 $A = \left(1, \sqrt{r^2 - 1}\right)$，与边 $y = 1$ 相交于 $B = \left(\sqrt{r^2 - 1}, 1\right)$；由于 $0 \le \sqrt{r^2 - 1} \le 1$，这两点都在正方形上。设 α 为 OA 与 x 轴所成的角，则 $\cos\alpha = 1/r$，即 $\alpha = \cos^{-1}(r^{-1})$。由于图形关于 $y = x$ 对称，OB 与 y 轴所成的角同样是 α。



正方形内位于圆内的部分由以下各部分组成：

- 顶点为 O、$(1,0)$、A 的三角形 T_1：直角位于 $(1,0)$，两条直角边分别为 1 和 $\sqrt{r^2-1}$，故面积为 $\frac{1}{2}\sqrt{r^2-1}$；
- 顶点为 O、$(0,1)$、B 的三角形 T_2：它是 T_1 的镜像，面积同样是 $\frac{1}{2}\sqrt{r^2-1}$；
- 介于 OA 与 OB 之间、半径为 r 的扇形 S，其圆心角为 $\frac{\pi}{2}-2\alpha$，故面积为 $\frac{1}{2}r^2\left(\frac{\pi}{2}-2\alpha\right)=\frac{1}{4}\pi r^2-r^2\alpha$。

相加得

$$\mathrm{P}(R\le r)=\sqrt{r^2-1}+\tfrac{1}{4}\pi r^2-r^2\cos^{-1}(r^{-1})\qquad(1\le r\le\sqrt{2}),$$

此即所求。（在 $r=1$ 时，两条公式都给出 $\frac{1}{4}\pi$；而在 $r=\sqrt{2}$ 时，我们得到 $1+\frac{1}{2}\pi-2\cdot\frac{\pi}{4}=1$，正如所料。）

第 (ii) 部分

记 $F(r)=\mathrm{P}(R\le r)$，并以 $f=F'$ 表示概率密度函数。在 $0<r<1$ 上，$f(r)=\frac{1}{2}\pi r$。在 $1<r<\sqrt{2}$ 上，我们需要

$$\frac{\mathrm{d}}{\mathrm{d}r}\cos^{-1}(r^{-1})=-\frac{1}{\sqrt{1-r^{-2}}}\cdot\left(-\frac{1}{r^2}\right)=\frac{1}{r\sqrt{r^2-1}},$$

这里用了当 $r>0$ 时 $\sqrt{1-r^{-2}}=\sqrt{r^2-1}/r$。于是由乘积法则（product rule），

$$\begin{aligned}f(r)&=\frac{r}{\sqrt{r^2-1}}+\tfrac{1}{2}\pi r-\left(2r\cos^{-1}(r^{-1})+\frac{r^2}{r\sqrt{r^2-1}}\right)\\&=\tfrac{1}{2}\pi r-2r\cos^{-1}(r^{-1}).\end{aligned}$$

两个 $\dfrac{r}{\sqrt{r^2-1}}$ 项互相抵消，这是个好兆头。因此

$$\begin{aligned}\mathrm{E}(R)&=\int_0^{\sqrt{2}}rf(r)\,\mathrm{d}r=\int_0^{\sqrt{2}}\tfrac{1}{2}\pi r^2\,\mathrm{d}r-2\int_1^{\sqrt{2}}r^2\cos^{-1}(r^{-1})\,\mathrm{d}r\\&=\frac{\pi(\sqrt{2})^3}{6}-2I=\frac{\sqrt{2}\,\pi}{3}-2I,\end{aligned}$$

其中 $I = \displaystyle\int_1^{\sqrt{2}} r^2 \cos^{-1}(r^{-1})\,\mathrm{d}r$。

> **想一想** 怎样才能消去 I 里面的反余弦?

进行分部积分，对 $\cos^{-1}(r^{-1})$ 求导（上面已求得它的导数是代数式），并对 r^2 积分：

$$\begin{aligned} I &= \Big[\tfrac{1}{3}r^3\cos^{-1}(r^{-1})\Big]_1^{\sqrt{2}} - \int_1^{\sqrt{2}} \frac{r^3}{3}\cdot\frac{1}{r\sqrt{r^2-1}}\,\mathrm{d}r \\ &= \tfrac{1}{3}\cdot 2\sqrt{2}\cdot\frac{\pi}{4} - 0 - \frac{1}{3}\int_1^{\sqrt{2}} \frac{r^2}{\sqrt{r^2-1}}\,\mathrm{d}r, \end{aligned}$$

这里用了 $\cos^{-1}(1/\sqrt{2}) = \frac{\pi}{4}$ 及 $\cos^{-1}1 = 0$。（被积函数在 $r=1$ 处无界，但在该处它的表现与 $1/\sqrt{2(r-1)}$ 相若，所以积分收敛（converges）。）因此

$$\mathrm{E}(R) = \frac{\sqrt{2}\,\pi}{3} - \frac{\sqrt{2}\,\pi}{3} + \frac{2}{3}\int_1^{\sqrt{2}} \frac{r^2}{\sqrt{r^2-1}}\,\mathrm{d}r = \frac{2}{3}\int_1^{\sqrt{2}} \frac{r^2}{\sqrt{r^2-1}}\,\mathrm{d}r,$$

此即所求结果。

第 (iii) 部分

$\sqrt{r^2-1}$ 提示我们设 $r = \cosh t$，则 $\mathrm{d}r = \sinh t\,\mathrm{d}t$，且 $\sqrt{r^2-1} = \sinh t$（因为 $t \geq 0$）。积分上下限变为 $t = 0$ 和 $t = T$，其中 $\cosh T = \sqrt{2}$，所以 $\sinh T = 1$，且

$$T = \cosh^{-1}\sqrt{2} = \ln\big(\sqrt{2}+\sqrt{2-1}\big) = \ln\big(\sqrt{2}+1\big).$$

于是

$$\begin{aligned} \int_1^{\sqrt{2}} \frac{r^2}{\sqrt{r^2-1}}\,\mathrm{d}r &= \int_0^T \frac{\cosh^2 t}{\sinh t}\,\sinh t\,\mathrm{d}t = \int_0^T \tfrac{1}{2}(1+\cosh 2t)\,\mathrm{d}t \\ &= \Big[\tfrac{1}{2}t + \tfrac{1}{2}\sinh t\cosh t\Big]_0^T = \tfrac{1}{2}\ln\big(\sqrt{2}+1\big) + \tfrac{1}{2}\cdot 1\cdot\sqrt{2}. \end{aligned}$$

乘以 $\frac{2}{3}$，得

$$\mathrm{E}(R) = \tfrac{1}{3}\Big(\sqrt{2} + \ln\big(\sqrt{2}+1\big)\Big),$$

此即所求。

验算

- F 在 $r=1$ 处连续（两段公式都给出 $\frac{\pi}{4} \approx 0.785$），并在 $r=\sqrt{2}$（即最远的那个角）处达到 1。
- 概率密度函数 $\frac{1}{2}\pi r - 2r\cos^{-1}(r^{-1})$ 在 $r=\sqrt{2}$ 处等于零，理应如此：只有那唯一的角点才有这么远。

- 数值上，$\mathrm{E}(R) = \frac{1}{3}(1.41421 + 0.88137) \approx 0.7652$，合理地介于 $\frac{1}{2}$ 与 $\sqrt{2}$ 之间。以电脑模拟 4×10^5 个随机点，得到的平均距离约为 0.766，而其经验分布函数（empirical distribution function）在多个 r 值上与 F 的差距都在 0.002 以内。

常见错误

- 把扇形的圆心角当成 $\frac{\pi}{2} - \alpha$ 而非 $\frac{\pi}{2} - 2\alpha$：在两条边的每一条都切出了一个三角形。
- 弄错 $\cos^{-1}(r^{-1})$ 导数的正负号：两个负号（分别来自 $\cos^{-1}$ 和 r^{-1}）互相抵消，所以导数是正的。

STEP 之外

这是几何概率（geometric probability）的初体验：由单位正方形的一个角到正方形内随机一点的平均距离。与之相关的问题——单位正方形内两个随机点之间的平均距离——有一个类似但更复杂、含有 $\ln(1+\sqrt{2})$ 的闭合形式，是二重积分（double integrals）与变量代换（change of variables）的经典练习。Grimmett 与 Welsh 所著的 *Probability: An Introduction*（《概率导论》）介绍了由几何定义的分布，以及如何由分布函数（distribution function）计算期望值。

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

2024 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。

