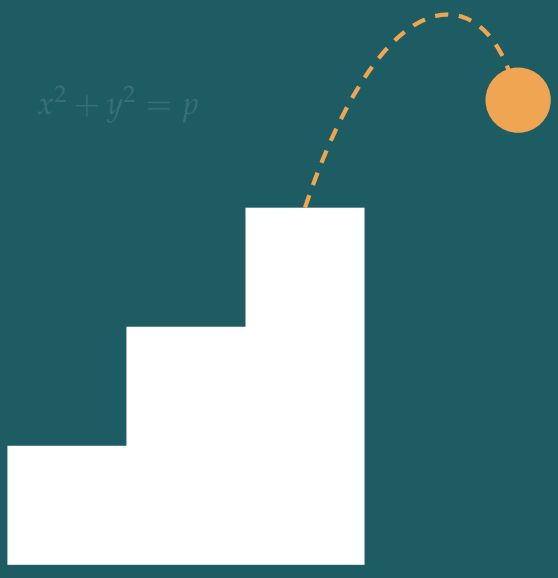
STEP 备试 · 大学数学

Beyond STEP

STEP I, II 及 III

2010 年详细解答



KEITH McNULTY

Beyond STEP

以历届 STEP 试题为门
走进大学数学

STEP I, II 及 III

2010 年详细解答

Keith McNulty

Beyond STEP：以历届 STEP 试题为门，走进大学数学

2010 年详细解答

2010 年 STEP 试卷（STEP I, II 及 III）全部 39 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 I

第 1 题 配方法 纯数学 . . . 3

第 2 题 曲线描绘 纯数学 . . . 7

第 3 题 和化积公式 纯数学 . . . 12

第 4 题 代换积分法 纯数学 . . . 16

第 5 题 二项式系数恒等式 纯数学 . . . 19

第 6 题 降阶法 纯数学 . . . 22

第 7 题 向量与相交直线 纯数学 . . . 25

第 8 题 无穷递降法 纯数学 . . . 29

第 9 题 摩擦力与平衡 力学 . . . 32

第 10 题 动量运动学 力学 . . . 36

第 11 题 圆上的碰撞 力学 . . . 40

第 12 题 尾和求期望值 概率与统计 . . . 43

第 13 题 泊松过程的等候时间 概率与统计 . . . 47

STEP II

第 1 题 曲率与密切圆 纯数学 . . . 51

第 2 题 三角函数积分法 纯数学 . . . 54

第 3 题 斐波那契数与递推关系 纯数学 . . . 58

第 4 题 代换积分法 纯数学 . . . 61

第 5 题 向量夹角与圆锥 纯数学 . . . 64

第 6 题 正四面体的几何 纯数学 . . . 68

第 7 题 三次方程 纯数学 . . . 72

第 8 题 曲线之间的面积 纯数学 . . . 76

第 9 题 相撞的抛体 力学 . . . 80

第 10 题 弹性碰撞 力学 . . . 83
第 11 题 力平衡 力学 . . . 87
第 12 题 均值与中位数 概率与统计 . . . 91
第 13 题 独立事件 概率与统计 . . . 95

STEP III

第 1 题 平均值与方差 纯数学 . . . 99
第 2 题 定积分 纯数学 . . . 103
第 3 题 单位根 纯数学 . . . 107
第 4 题 多项式的公共根 纯数学 . . . 111
第 5 题 坐标几何作图 纯数学 . . . 115
第 6 题 三维空间中的旋转 纯数学 . . . 119
第 7 题 麦克劳林级数 纯数学 . . . 122
第 8 题 利用商法则积分 纯数学 . . . 126
第 9 题 圆上的运动 力学 . . . 130
第 10 题 微小振动 力学 . . . 133
第 11 题 摩擦力与相对运动 力学 . . . 136
第 12 题 几何分布的期望值 概率与统计 . . . 140
第 13 题 线性组合的相关性 概率与统计 . . . 144

详细解答

STEP I 2010，第 1 题

纯数学 · 配方法

> **STEP I 2010，第 1 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s1-q1

关键思路

这个恒等式（identity）把一个不易处理的二次式改写成两个平方之和再加一个常数。只要留意到第二条方程也可以用同样的两个括号 $x-y+2$ 和 $2x-y$ 来表示，问题就变成两个新未知数的两条简单方程。这两条方程对新未知数的平方而言是线性的，所以可以直接求解。

解答

恒等式

展开右边，

$$\begin{aligned} &a(x-y+2)^2+b(cx+y)^2+d \\ &\quad =a(x^2+y^2-2xy+4x-4y+4)+b(c^2x^2+2cxy+y^2)+d \\ &\quad =(a+bc^2)x^2+(a+b)y^2+(2bc-2a)xy+4ax-4ay+(4a+d). \end{aligned}$$

由于这是恒等式，我们可以与 $5x^2+2y^2-6xy+4x-4y$ 比较系数（comparing coefficients）：

$$\begin{aligned}
x:&\quad 4a=4, && \text{所以}a=1;\\
y:&\quad -4a=-4, && \text{与上式吻合;}\\
y^2:&\quad a+b=2, && \text{所以}b=1;\\
xy:&\quad 2bc-2a=-6, && \text{所以}2c=-4,\ c=-2;\\
x^2:&\quad a+bc^2=5, && \text{验算：}1+4=5;\\
\text{常数项}:&\quad 4a+d=0, && \text{所以}d=-4.
\end{aligned}$$

注意，单凭 x^2 的系数只能得出 $c^2=4$；决定正负号的是 xy 的系数。所以

$$a=1,\qquad b=1,\qquad c=-2,\qquad d=-4,$$

即

$$5x^2+2y^2-6xy+4x-4y\equiv(x-y+2)^2+(2x-y)^2-4,$$

其中我们把 $(-2x+y)^2$ 写成 $(2x-y)^2$。

联立方程

根据恒等式，第一条方程变成

$$(x-y+2)^2+(2x-y)^2=13.$$

想一想 第二条方程看来不一样。它可以用同样的两个括号来表示吗?

它含 x 和 y 的一次项是 $8x-8y$，是第一条方程的两倍。这提示我们试试 $2(x-y+2)^2$，它的一次项正好相符：

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

左边余下的部分是

$$6x^2+3y^2-8xy+8x-8y-2(x-y+2)^2=4x^2+y^2-4xy-8=(2x-y)^2-8.$$

所以第二条方程是 $2(x-y+2)^2+(2x-y)^2-8=14$，即

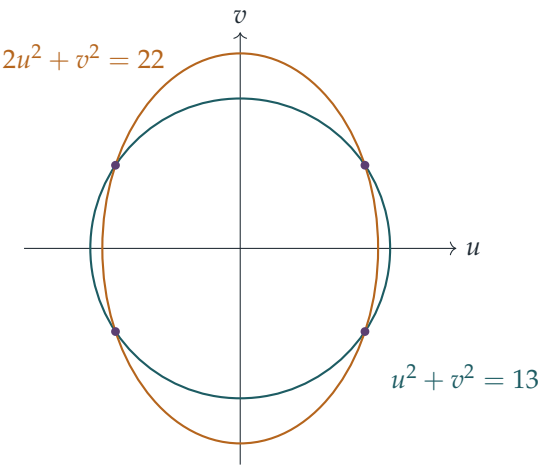
$$2(x-y+2)^2+(2x-y)^2=22.$$

（你也可以沿用求恒等式的方法，在 $\alpha(x-y+2)^2+\beta(\gamma x+y)^2+\delta$ 中比较系数而得出这个结果。）

现在设 $u = x - y + 2$ 及 $v = 2x - y$。方程组变成

$$u^2 + v^2 = 13, \qquad 2u^2 + v^2 = 22.$$

两式相减得 $u^2 = 9$，于是 $v^2 = 4$。所以 $u = \pm 3$ 及 $v = \pm 2$，两者的正负号可独立选取，共得四组 (u, v)。



要还原为 x 和 y，可以相减：$v - (u - 2) = (2x - y) - (x - y) = x$，所以

$$x = v - u + 2, \qquad y = 2x - v = v - 2u + 4.$$

每组 (u, v) 恰好对应一组 (x, y)，因为这个变量代换（change of variables）是线性而且可逆的。四个解如下：

u	v	x	y
3	2	1	0
3	−2	−3	−4
−3	2	7	12
−3	−2	3	8

所以

$$(x, y) = (1, 0), \quad (-3, -4), \quad (7, 12), \quad (3, 8).$$

验算

代入 $(1, 0)$：第一条方程得 $5 + 4 = 9$，第二条得 $6 + 8 = 14$。对于 $(7, 12)$：$245 + 288 - 504 + 28 - 48 = 9$ 及 $294 + 432 - 672 + 56 - 96 = 14$。电脑已在一千个随机有理点上验

证这两个恒等式，并确认四组解全部满足两条方程。以穷举法搜索所有满足 $|x|, |y| \leq 50$ 的整数对，找到的正好就是这四个解。两条圆锥曲线（conic）最多相交于四点，所以我们不可能有所遗漏。

常见错误

- 由 $c^2 = 4$ 直接取 $c = 2$，而没有核对 xy 的系数。
- 只取正平方根，结果只找到一个解而不是四个。
- 直接从原来的方程消去一个变量。这样会得出一条四次方程（quartic），用手计算非常费力。

STEP 之外

把二次型(quadratic form)写成平方的组合,是二次型理论的起点。把对称矩阵(symmetric matrix）对角化（diagonalise），总可以找到一个线性变量代换，把 $Ax^2 + Bxy + Cy^2$ 化为 $\lambda_1 u^2 + \lambda_2 v^2$，而圆锥曲线正是这样分类的。这一题的两条方程碰巧共用同一对平方，所以才能如此干净俐落地分离开来。Strang 的 *Introduction to Linear Algebra*（线性代数入门）清楚讲解了对称矩阵和二次型。

详细解答

STEP I 2010，第 2 题

纯数学 · 曲线描绘

STEP I 2010，第 2 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q2

关键思路

用商法则（quotient rule）求导，并抽出永不为零的因子 e^x。这样驻点（stationary point）就是一个二次式的根，所以“有两个驻点”其实是关于其判别式（discriminant）的一个条件。至于绘画略图，同一个二次式的正负号告诉我们曲线在哪里上升、在哪里下降，而曲线在渐近线（asymptote）附近以及当 $x \to \pm\infty$ 时的表现，则补足其余的细节。

解答

设 $y = \dfrac{x-a}{x-b}\mathrm{e}^x$，其定义域为 $x \neq b$。利用积法则（product rule），并由 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{x-a}{x-b}\right) = \dfrac{(x-b)-(x-a)}{(x-b)^2} = \dfrac{a-b}{(x-b)^2}$，可得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x-a}{x-b}\mathrm{e}^x + \frac{a-b}{(x-b)^2}\mathrm{e}^x = \frac{\mathrm{e}^x\left[(x-a)(x-b)+(a-b)\right]}{(x-b)^2}.$$

由于 $\mathrm{e}^x > 0$，而且当 $x \neq b$ 时 $(x-b)^2 > 0$，所以 $\mathrm{d}y/\mathrm{d}x$ 的正负号与下式的正负号相同：

$$q(x) = (x-a)(x-b) + (a-b) = x^2 - (a+b)x + ab + a - b,$$

而驻点正好就是 q 的根当中不等于 $x = b$ 的那些。

? **想一想** q 的根会否刚好落在 $x = b$，即曲线没有定义的地方？

这并不会造成问题：$q(b) = a - b$，所以只有当 $a = b$ 时 $x = b$ 才是根。但这时曲线只不过是 $y = \mathrm{e}^x$（在 $x = b$ 处有一个空洞），根本没有驻点。因此在任何情况下，驻点都恰好

是 q 的实根，而驻点有两个，当且仅当 q 有两个相异实根，即其判别式为正：

$$(a+b)^2 - 4(ab+a-b) = (a-b)^2 - 4(a-b) = (a-b)(a-b-4) > 0.$$

两个因子之积为正，当且仅当两者皆正或两者皆负。两者皆负即 $a-b<0$；两者皆正即 $a-b>4$（这已保证 $a-b>0$）。因此，若有两个驻点，必须有

$$a-b<0 \quad \text{或} \quad a-b>4,$$

这正是所要证明的。（反过来说，两个条件中任何一个成立，都会有两个驻点。）

第 (i) 部分

取 $a=0$、$b=\frac{1}{2}$，则 $a-b=-\frac{1}{2}<0$，并且

$$y = \frac{x\,\mathrm{e}^x}{x-\frac{1}{2}}, \qquad q(x) = x^2 - \tfrac{1}{2}x - \tfrac{1}{2} = (x-1)\big(x+\tfrac{1}{2}\big).$$

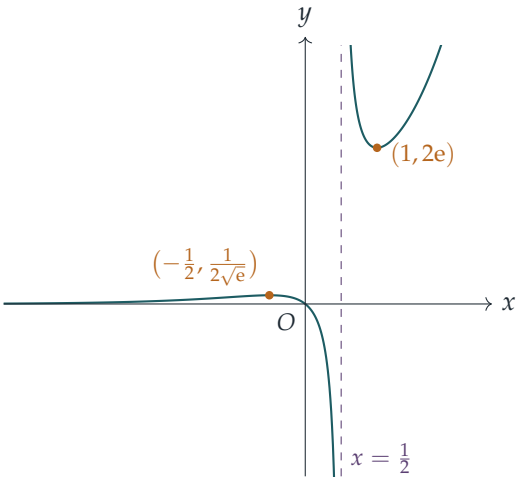
驻点位于 $x=-\frac{1}{2}$ 和 $x=1$，分别在铅垂渐近线 $x=\frac{1}{2}$ 的两侧，符合题目要求。它们的 y-值为

$$y\big(-\tfrac{1}{2}\big) = \frac{-\frac{1}{2}\,\mathrm{e}^{-1/2}}{-1} = \frac{1}{2\sqrt{\mathrm{e}}} \approx 0.30, \qquad y(1) = \frac{\mathrm{e}}{\frac{1}{2}} = 2\mathrm{e} \approx 5.44.$$

由于当 $x<-\frac{1}{2}$ 时 $q>0$，当 $-\frac{1}{2}<x<1$（其中 $x\neq\frac{1}{2}$）时 $q<0$，而当 $x>1$ 时 $q>0$，所以点 $\big(-\frac{1}{2}, \frac{1}{2\sqrt{\mathrm{e}}}\big)$ 是局部极大点（local maximum），而 $(1, 2\mathrm{e})$ 是局部极小点（local minimum）。

其他特征：

- 当 $x\to-\infty$ 时，$\frac{x}{x-1/2}\to 1$ 而 $\mathrm{e}^x\to 0$，所以 $y\to 0$。当 $x<0$ 时，x 和 $x-\frac{1}{2}$ 皆为负数，所以 $y>0$：曲线从上方趋近负 x-轴。
- 曲线只在原点与坐标轴相交，而且当 $0<x<\frac{1}{2}$ 时 $y<0$。
- 当 $x\to\frac{1}{2}^-$ 时，$y\to-\infty$；当 $x\to\frac{1}{2}^+$ 时，$y\to+\infty$（分子趋向 $\frac{1}{2}\mathrm{e}^{1/2}>0$）。
- 当 $x\to\infty$ 时，$y\to\infty$，增长的速度与 e^x 相若。



第 (ii) 部分

取 $a=\frac{9}{2}$、$b=0$，则 $a-b=\frac{9}{2}>4$，所以同样有两个驻点。这时

$$y=\frac{x-\frac{9}{2}}{x}\,\mathrm{e}^x, \qquad q(x)=x^2-\tfrac{9}{2}x+\tfrac{9}{2}=\tfrac{1}{2}(2x-3)(x-3),$$

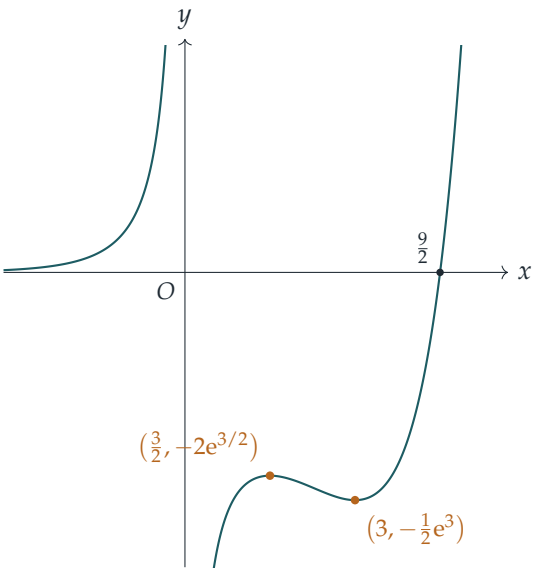
所以驻点位于 $x=\frac{3}{2}$ 和 $x=3$，两者都在渐近线 $x=0$ 的右方。它们的 y-值为

$$y\left(\tfrac{3}{2}\right)=\frac{-3}{\frac{3}{2}}\,\mathrm{e}^{3/2}=-2\mathrm{e}^{3/2}\approx -8.96, \qquad y(3)=\frac{-\frac{3}{2}}{3}\,\mathrm{e}^3=-\tfrac{1}{2}\mathrm{e}^3\approx -10.04.$$

q 的正负号如下：当 $x<\frac{3}{2}$（$x\neq 0$）时为正，在 $\left(\frac{3}{2},3\right)$ 上为负，当 $x>3$ 时为正，所以 $\left(\frac{3}{2},-2\mathrm{e}^{3/2}\right)$ 是局部极大点，而 $\left(3,-\frac{1}{2}\mathrm{e}^3\right)$ 是局部极小点。留意两点都位于 x-轴下方，而且极小点只比极大点略低，所以两点之间的凹陷很浅。

其他特征：

- 当 $x\to-\infty$ 时，$y\to 0$；当 $x<0$ 时，$x-\frac{9}{2}$ 和 x 皆为负数，所以 $y>0$。在 $x<0$ 上曲线递增（因为该处 $q>0$），由刚好高于横轴之处一直向上升往渐近线。
- 当 $x\to 0^-$ 时，$y\to+\infty$；当 $x\to 0^+$ 时，$y\to-\infty$（分子趋向 $-\frac{9}{2}$）。
- 唯一的截点（intercept）是 $\left(\frac{9}{2},0\right)$；在 $\left(0,\frac{9}{2}\right)$ 上 $y<0$，而当 $x>\frac{9}{2}$ 时 $y>0$，并且当 $x\to\infty$ 时 $y\to\infty$。



这里 y-轴本身就是铅垂渐近线。

验算

- 在随机选取的 (a,b,x) 上以数值方法计算 y 的导数，结果与 $\mathrm{d}y/\mathrm{d}x$ 的公式吻合；对随机选取的 (a,b) 数算 q 变号的次数，亦发现恰好在 $a-b<0$ 或 $a-b>4$ 时才有两个驻点。
- 在 (i) 中，$y(1)=5.4366=2\mathrm{e}$ 及 $y(-\frac{1}{2})=0.3033$；在 (ii) 中，$y(\frac{3}{2})=-8.9634$ 及 $y(3)=-10.0428$，而各点左右两旁邻近的函数值亦确认了它们是极大点还是极小点。
- 检查条件是否合理：在 (i) 中 $a-b=-\frac{1}{2}$，在 (ii) 中 $a-b=\frac{9}{2}$，两者都在容许的范围之内，理应如此。

常见错误

- 把判别式展开后，没有察觉它就是 $(a-b)^2-4(a-b)$；只要以 $a-b$ 来表示，它便可以简洁地因式分解。
- 忘记检查 q 的根不可能落在渐近线 $x=b$ 上。
- 在 (ii) 中，把极大点和极小点画在 x-轴的上下两侧：其实两者都在轴的下方，而曲线只在 $x=\frac{9}{2}$ 处穿过该轴。

STEP 之外

这里所用的方法, 即根据导数的正负号和各个极限来判断曲线的形状, 在大学第一门分析（analysis）课程中会得到严格的论证：一方面借着中值定理（Mean Value Theorem）（在某区间上导数为正，确实意味着函数递增），另一方面借着无穷远处极限的严谨定义。当 $x \to \pm\infty$ 时，e^x 胜过任何有理函数（rational function），这也是在该课程中证明的一个标准结果。Abbott 的 *Understanding Analysis* 把这一切都讲解得十分精彩。

详细解答

STEP I 2010，第 3 题

纯数学 · 和化积公式

STEP I 2010，第 3 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q3

关键思路

和差化积公式（sum-to-product formulae）的作用，正是把两个正弦（或余弦）之差化为乘积。椭圆（ellipse）上连接参数（parameter）为 u 和 v 的两点的弦（chord），其斜率（gradient）为 $\dfrac{b(\sin v - \sin u)}{a(\cos v - \cos u)}$。分子和分母都是差，所以可用这些公式把它们因式分解，约去公因式之后，斜率只取决于和 $u+v$。于是比较两条弦，就归结为比较两个余切（cotangent）。

解答

三角恒等式

利用复角公式（compound-angle formulae），

$$\begin{aligned}\sin(x+y)-\sin(x-y) &= (\sin x\cos y+\cos x\sin y)-(\sin x\cos y-\cos x\sin y)\\ &= 2\cos x\sin y,\end{aligned}$$

正如所求。现在，对任意 A 和 B，设 $x=\frac{1}{2}(A+B)$ 及 $y=\frac{1}{2}(A-B)$，则 $x+y=A$ 及 $x-y=B$。这个恒等式（identity）变成

$$\sin A-\sin B = 2\cos\tfrac{1}{2}(A+B)\,\sin\tfrac{1}{2}(A-B).$$

同样地，

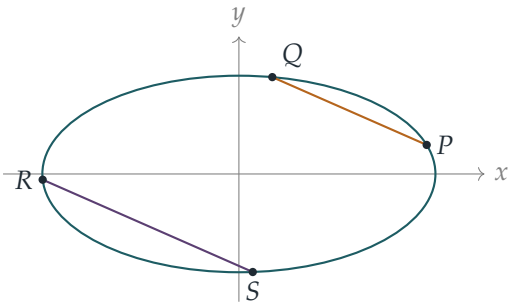
$$\begin{aligned}\cos(x+y)-\cos(x-y) &= (\cos x\cos y-\sin x\sin y)-(\cos x\cos y+\sin x\sin y)\\ &= -2\sin x\sin y,\end{aligned}$$

再作相同的代换，便得

$$\cos A - \cos B = -2\sin\tfrac{1}{2}(A+B)\,\sin\tfrac{1}{2}(A-B).$$

平行弦

四点全都位于椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 上，按逆时针方向依次为 P、Q、R、S。下图显示 $a=2$、$b=1$ 时的一对平行弦。



弦的斜率。在 $[0,2\pi)$ 内取两个参数 $u<v$，使得通过 $(a\cos u, b\sin u)$ 和 $(a\cos v, b\sin v)$ 的弦不是垂直的，即 $\cos v \neq \cos u$。在两个恒等式中取 $A=v$、$B=u$，可知其斜率为

$$\frac{b(\sin v-\sin u)}{a(\cos v-\cos u)}=\frac{2b\cos\frac{1}{2}(u+v)\,\sin\frac{1}{2}(v-u)}{-2a\sin\frac{1}{2}(u+v)\,\sin\frac{1}{2}(v-u)}=-\frac{b}{a}\cot\tfrac{1}{2}(u+v).$$

这样约分是合理的：$0<v-u<2\pi$，所以 $\frac{1}{2}(v-u)$ 严格介乎 0 与 π 之间，故 $\sin\frac{1}{2}(v-u)\neq 0$。由于原来的分母不等于零，$\sin\frac{1}{2}(u+v)\neq 0$ 亦成立，所以余切有定义。

记 $\alpha=\frac{1}{2}(p+q)$ 及 $\beta=\frac{1}{2}(r+s)$。PQ 和 SR 都不是垂直的，所以

$$\text{斜率：}\ PQ=-\frac{b}{a}\cot\alpha,\qquad \text{斜率：}\ SR=-\frac{b}{a}\cot\beta,$$

其中 $\sin\alpha\neq 0$ 及 $\sin\beta\neq 0$。两条非垂直的直线平行，当且仅当它们的斜率相等；又因为 $b/a\neq 0$，

$$PQ\parallel SR \iff \cot\alpha=\cot\beta.$$

? **想一想** 为甚么不能直接“约去 cot”而得出 $\alpha=\beta$?

因为 cot 并非一对一（one-to-one）：它以 π 为周期（period）重复取值。为了顾及这一点，把两者之差写成单一个正弦：

$$\cot\alpha-\cot\beta=\frac{\cos\alpha\sin\beta-\sin\alpha\cos\beta}{\sin\alpha\sin\beta}=\frac{\sin(\beta-\alpha)}{\sin\alpha\sin\beta}.$$

分母不等于零，所以

$$PQ \parallel SR \iff \sin(\beta-\alpha)=0 \iff \beta-\alpha=k\pi\text{，对于某整数}k.$$

确定 k 的值。我们有

$$\beta-\alpha=\tfrac{1}{2}\big[(r-p)+(s-q)\big].$$

由于 $p<r$ 及 $q<s$，两个括号内的值都是正数，所以 $\beta-\alpha>0$。由于四个参数都在 $[0,2\pi)$ 内，每个括号内的值都小于 2π，所以 $\beta-\alpha<2\pi$。严格介乎 0 与 2π 之间的 π 的倍数只有 π 本身，所以 $k=1$。因此

$$PQ \parallel SR \iff \tfrac{1}{2}(r+s-p-q)=\pi \iff r+s-p-q=2\pi.$$

每一步都是等价关系，所以两个方向都已证明，正如所求。

（这两条直线不可能重合：一条直线与椭圆最多相交于两点，而 P、Q、R、S 是四个相异的点。所以这里的“平行”是指真正平行而且互不重合。）

验算

- 取 $p=0$、$q=\frac{\pi}{2}$、$r=\pi$、$s=\frac{3\pi}{2}$，则 $r+s-p-q=2\pi$。此时 $P=(a,0)$、$Q=(0,b)$、$R=(-a,0)$、$S=(0,-b)$，而 PQ 和 SR 的斜率都是 $-b/a$。当 $a=2$、$b=1$ 时，两个斜率都等于 -0.5。
- 对称性：条件即是 $\frac{1}{2}(r+s)=\frac{1}{2}(p+q)+\pi$。变换 $(x,y)\mapsto(-x,-y)$ 会令点的参数增加 π，而把弦 PQ 旋转半周会得出一条平行的弦，与结果吻合。
- 数值检验：随机选取 a,b 以及 $p<q<r<s$，在 $r+s-p-q=2\pi$ 的情况下（约 5000 个特意构造的例子），两个斜率完全相等，其他情况下则从不相等。

常见错误

- 由 $\cot\alpha=\cot\beta$ 直接推出 $\alpha=\beta$。这样会得到 $r+s=p+q$，在这里是不可能的。cot 的周期正是整题的关键。
- 忘记说明为何可以约去 $\sin\frac{1}{2}(v-u)$，或者没有利用 $p<q<r<s$ 这个次序来排除 $k=0$ 和 $k=2$。

STEP 之外

把椭圆写成 $(a\cos t, b\sin t)$，可见它是单位圆（unit circle）在伸缩变换（stretch）$(x,y)\mapsto(ax,by)$ 之下的像。这类线性映射（linear map）把平行线变为平行线，所以这道题其实

是关于圆的问题。在圆上，由角 p 到角 q 的弦垂直于角 $\frac{1}{2}(p+q)$ 处的半径，于是结果几乎一目了然。这种利用映射下保持不变的性质来简化问题的想法，贯穿线性代数和几何。Strang 的 *Introduction to Linear Algebra*（线性代数入门）是认识线性映射如何运作的好去处。Siklos 的 *Advanced Problems in Mathematics*（高等数学难题）则有更多 STEP 风格的三角恒等式练习。

详细解答

STEP I 2010，第 4 题

纯数学 · 代换积分法

STEP I 2010，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q4

关键思路

这个代换（substitution）的用意，是令 x 和 $x+1$ 同时变成分母同为 t^2-1 的简单分数，于是两者乘积的平方根可以准确地开出来。积分随即化为附注中的标准积分。至于体积，把 y 平方后会得到两个容易处理的对数项和一个交叉项（cross term），而这个交叉项正正就是我们刚求得的积分。

解答

积分

取 $x>0$。代换 $x=\dfrac{1}{t^2-1}$（其中 $t>1$）是合理的：解出 t 得

$$t=\sqrt{1+\frac{1}{x}}=\frac{\sqrt{x+1}}{\sqrt{x}},$$

而当 t 取遍 $(1,\infty)$ 时，x 取遍 $(0,\infty)$（递减地），每个 t 对应唯一一个 x 值。

现在

$$x+1=\frac{1}{t^2-1}+1=\frac{t^2}{t^2-1},\qquad \text{所以}\qquad x(x+1)=\frac{t^2}{(t^2-1)^2}.$$

由于 $t>1$，t 和 t^2-1 都是正数，所以

$$\sqrt{x(x+1)}=\frac{t}{t^2-1}.$$

? **想一想** 为甚么开平方根时需要 $t>1$?

平方根永不为负，所以 $\sqrt{t^2/(t^2-1)^2} = |t|/|t^2-1|$。条件 $t>1$ 让我们可以把两个绝对值（modulus）符号都去掉。这也解释了为何这样选取：$t^2-1>0$ 正是令 $x>0$ 的条件。求导，得 $\dfrac{\mathrm{d}x}{\mathrm{d}t} = -\dfrac{2t}{(t^2-1)^2}$。因此

$$\begin{aligned}\int \frac{1}{\sqrt{x(x+1)}}\,\mathrm{d}x &= \int \frac{t^2-1}{t}\cdot\left(-\frac{2t}{(t^2-1)^2}\right)\mathrm{d}t = -2\int \frac{1}{t^2-1}\,\mathrm{d}t \\ &= -2\cdot\frac{1}{2}\ln\left|\frac{t-1}{t+1}\right| + c = \ln\left(\frac{t+1}{t-1}\right) + c,\end{aligned}$$

这里用了附注中 $a=1$ 的情况，以及 $t>1$ 令 $(t-1)/(t+1)$ 为正数这一事实。

现在利用 $t = \sqrt{x+1}/\sqrt{x}$ 换回 x。把分子和分母同乘以 $\sqrt{x}$，

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

其中我们把分子和分母同乘以 $\sqrt{x+1}+\sqrt{x}$，并用了平方差（difference of two squares）公式。故此

$$\int \frac{1}{\sqrt{x(x+1)}}\,\mathrm{d}x = 2\ln\left(\sqrt{x}+\sqrt{x+1}\right) + c,$$

正如所求。

体积

当 $x>0$ 时，曲线上 $y>0$，因为由 $\sqrt{x}<\sqrt{x+1}$ 可得 $1/\sqrt{x}>1/\sqrt{x+1}$。所以曲线与 x 轴之间的区域旋转所扫出的体积是 $V = \pi\int y^2\,\mathrm{d}x$。展开后，

$$y^2 = \frac{1}{x}+\frac{1}{x+1}-\frac{2}{\sqrt{x}\sqrt{x+1}} = \frac{1}{x}+\frac{1}{x+1}-\frac{2}{\sqrt{x(x+1)}},$$

而最后一项是第一部分中被积函数（integrand）的两倍。所以

$$V = \pi\Big[\ln x + \ln(x+1)\Big]_{1/8}^{9/16} - 2\pi\Big[2\ln\left(\sqrt{x}+\sqrt{x+1}\right)\Big]_{1/8}^{9/16}.$$

积分上下限是特意选取的，好让各个平方根都很整齐：

- 当 $x=\frac{9}{16}$：$\sqrt{x}=\frac{3}{4}$，$\sqrt{x+1}=\frac{5}{4}$，两者之和为 2，而 $x(x+1)=\frac{225}{256}$；
- 当 $x=\frac{1}{8}$：$\sqrt{x}=\frac{1}{2\sqrt{2}}$，$\sqrt{x+1}=\frac{3}{2\sqrt{2}}$，两者之和为 $\frac{4}{2\sqrt{2}}=\sqrt{2}$，而 $x(x+1)=\frac{9}{64}$。

第一个方括号的值是

$$\ln\frac{225}{256}-\ln\frac{9}{64} = \ln\left(\frac{225}{256}\cdot\frac{64}{9}\right) = \ln\frac{25}{4},$$

而第二个是

$$2\ln 2 - 2\ln\sqrt{2} = 2\ln 2 - \ln 2 = \ln 2.$$

因此

$$V = \pi \ln \frac{25}{4} - 2\pi \ln 2 = \pi \ln \frac{25}{4} - \pi \ln 4 = \pi \ln \frac{25}{16} = 2\pi \ln \frac{5}{4},$$

这正是我们要的结果。

验算

- 对答案求导：

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

  因为括号内等于 $\dfrac{\sqrt{x+1} + \sqrt{x}}{2\sqrt{x}\sqrt{x+1}}$。

- 数值上，用辛普森法则（Simpson's rule）可得 $\pi \int_{1/8}^{9/16} y^2 \,\mathrm{d}x \approx 1.402052$，与 $2\pi \ln \frac{5}{4} \approx 1.402052$ 吻合。辛普森法则亦印证了 $\int_{1/8}^{9/16} \frac{\mathrm{d}x}{\sqrt{x(x+1)}} = \ln 2 \approx 0.693147$。

常见错误

- 漏掉 $\mathrm{d}x/\mathrm{d}t$ 中的负号。正是这个负号把 $\ln \frac{t-1}{t+1}$ 变成 $\ln \frac{t+1}{t-1}$。
- 把 y^2 展开成 $\frac{1}{x} - \frac{1}{x+1}$：交叉项 $-2/\sqrt{x(x+1)}$ 正是第一部分的用意所在。
- 在体积公式中漏掉因子 π，或者把 π 写成 2π。

STEP 之外

答案 $2\ln(\sqrt{x} + \sqrt{x+1})$ 其实就是 $2\,\mathrm{arsinh}\,\sqrt{x}$，因为 $\mathrm{arsinh}\,u = \ln\left(u + \sqrt{u^2+1}\right)$。用代换 $x = \sinh^2\theta$，一行就可得到同样的结果。选取能把平方根有理化（rationalise）的代换，是一套有系统的技巧（欧拉代换，Euler substitutions），而它背后支撑着这样一个结果：凡是 x 与 $\sqrt{ax^2+bx+c}$ 的有理函数（rational function），其积分都可以用初等函数（elementary functions）表示。Siklos 的 *Advanced Problems in Mathematics* 中还有很多同类的积分。

详细解答

STEP I 2010，第 5 题

纯数学 · 二项式系数恒等式

STEP I 2010，第 5 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q5

关键思路

这些和式每一个都是 $(1+x)^n$ 的二项展开式（binomial expansion）的化身：先对展开式做一些处理，再代入 $x=1$。$\binom{n}{k}$ 前面的因子 k，正是对 x^k 求导（differentiate）所产生的；因子 $\frac{1}{k+1}$，则是对 x^k 积分（integrate）所产生的。因此我们对恒等式（identity）

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

求导或积分，然后代入 $x=1$。至于 k^2，我们需要求导两次，并在两次之间乘以 x。

解答

以下 n 一直是正整数，而

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n \qquad (*)$$

对所有实数 x 均成立。由于等式两边都是多项式（polynomial），我们可以逐项求导或逐项积分，而恒等式依然成立。

第 (i) 部分

在 $(*)$ 中代入 $x=1$。左边是 2^n，而每个 x^k 都变成 1，所以

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

这正是所要证明的结果。

第 (ii) 部分

把 $(*)$ 对 x 求导。常数项消失，而 $\binom{n}{k}x^k$ 变成 $k\binom{n}{k}x^{k-1}$：

$$n(1+x)^{n-1} = \binom{n}{1} + 2\binom{n}{2}x + 3\binom{n}{3}x^2 + \cdots + n\binom{n}{n}x^{n-1}. \qquad (**)$$

代入 $x = 1$ 得

$$\binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n} = n\,2^{n-1},$$

这正是所要证明的结果。

第 (iii) 部分

这次我们想要的是 $\binom{n}{k}$ 除以 $k+1$，而这正好可以由 $\int_0^1 x^k\,\mathrm{d}x = \frac{1}{k+1}$ 得到。把 $(*)$ 由 0 积分至 1，左边是

$$\int_0^1 (1+x)^n\,\mathrm{d}x = \left[\frac{(1+x)^{n+1}}{n+1}\right]_0^1 = \frac{2^{n+1}-1}{n+1},$$

而右边是

$$\sum_{k=0}^{n}\binom{n}{k}\int_0^1 x^k\,\mathrm{d}x = \sum_{k=0}^{n}\frac{1}{k+1}\binom{n}{k}.$$

因此

$$\binom{n}{0} + \frac{1}{2}\binom{n}{1} + \frac{1}{3}\binom{n}{2} + \cdots + \frac{1}{n+1}\binom{n}{n} = \frac{1}{n+1}\left(2^{n+1}-1\right),$$

这正是所要证明的结果。留意积分下限十分重要：正因为由 0 开始积分，才会出现那个 -1。

第 (iv) 部分

> **想一想** 如果我们只是把 $(**)$ 再求导一次，得到的系数会是 $k(k-1)$，而不是 k^2。怎样才能得到 k^2 呢?

在第二次求导之前，先把 $(**)$ 乘以 x。这样 $k\binom{n}{k}$ 旁边的幂便还原为 x^k，于是下一次求导带下来的因子恰好又是 k：

$$nx(1+x)^{n-1} = \sum_{k=1}^{n} k\binom{n}{k}x^k.$$

求导，左边使用积法则（product rule），得

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

（当 $n=1$ 时，左边第二项含有因子 $n-1=0$，所以这一项根本不存在；恒等式仍然成立。）现在代入 $x=1$：

$$\begin{aligned}\binom{n}{1}+2^2\binom{n}{2}+3^2\binom{n}{3}+\cdots+n^2\binom{n}{n} &= n\,2^{n-1}+n(n-1)2^{n-2}\\ &= n\,2^{n-2}\big(2+(n-1)\big)\\ &= n(n+1)2^{n-2},\end{aligned}$$

这正是所要证明的结果。

另一个做法是写成 $k^2=k(k-1)+k$。把 $(*)$ 求导两次再代入 $x=1$，得 $\sum k(k-1)\binom{n}{k}=n(n-1)2^{n-2}$，再加上第 (ii) 部分的结果，便得到相同的答案。

验算

取 $n=3$，此时各系数为 $1,3,3,1$。

- (i) $1+3+3+1=8=2^3$。
- (ii) $3+6+3=12=3\cdot 2^2$。
- (iii) $1+\frac{3}{2}+1+\frac{1}{4}=\frac{15}{4}=\frac{1}{4}(2^4-1)$。
- (iv) $3+12+9=24=3\cdot 4\cdot 2$。

我们亦已用分数作精确计算，对由 1 至 40 的每一个 n 验证了全部四个恒等式。

常见错误

- 在 (iii) 中，忘记了积分下限的贡献，得出 $\frac{2^{n+1}}{n+1}$。
- 在 (iv) 中，直接对 $(**)$ 求导，得到的是 $\sum k(k-1)\binom{n}{k}$，而不是 $\sum k^2\binom{n}{k}$。

STEP 之外

这些恒等式都有概率上的意义。设 X 为掷一枚公平硬币 n 次所得正面的次数，则 $\mathrm{P}(X=k)=\binom{n}{k}/2^n$，所以第 (ii) 部分说明 $\mathrm{E}(X)=\frac{n}{2}$，第 (iv) 部分说明 $\mathrm{E}(X^2)=\frac{n(n+1)}{4}$，由此得 $\mathrm{Var}(X)=\frac{n}{4}$。对母函数（generating function）求导以求出矩（moments），是大学一年级的标准技巧。Grimmett and Welsh 的 *Probability: An Introduction*（概率入门教科书）对此有完整的阐述。

详细解答

STEP I 2010，第 6 题

纯数学 · 降阶法

> **STEP I 2010，第 6 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s1-q6

关键思路

只要知道二阶线性微分方程（linear second-order equation）的一个解，把通解（general solution）写成这个已知解乘以一个未知函数 u，所有只含 u 本身（不带导数）的项便会互相抵消。余下的只涉及 u' 和 u''，所以它其实是关于 $v = u'$ 的一阶方程，而且这个方程是可分离变量（separable）的。这种技巧称为*降阶法*（reduction of order）。

解答

验证 $y = \mathrm{e}^x$ 是解

若 $y = \mathrm{e}^x$，则 $y' = y'' = \mathrm{e}^x$，所以

$$(x-1)\mathrm{e}^x - x\mathrm{e}^x + \mathrm{e}^x = (x - 1 - x + 1)\mathrm{e}^x = 0,$$

因此 $y = \mathrm{e}^x$ 满足 $(*)$。

关于 u 的方程

设 $y = u\mathrm{e}^x$。由积法则（product rule），

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= (u' + u)\,\mathrm{e}^x,\\ \frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= (u'' + u')\,\mathrm{e}^x + (u' + u)\,\mathrm{e}^x = (u'' + 2u' + u)\,\mathrm{e}^x.\end{aligned}$$

代入 $(*)$ 的左方，并抽出因子 e^x：

$$\begin{aligned}&\mathrm{e}^x\left[(x-1)(u''+2u'+u)-x(u'+u)+u\right]\\&\quad=\mathrm{e}^x\left[(x-1)u''+(2x-2-x)u'+(x-1-x+1)u\right]\\&\quad=\mathrm{e}^x\left[(x-1)u''+(x-2)u'\right].\end{aligned}$$

含 u 的项互相抵消，情况与我们验证 $y=\mathrm{e}^x$ 时完全一样（那次验证正是 $u=1$ 的情形）。由于 e^x 永不为零，$y=u\mathrm{e}^x$ 满足 $(*)$ 当且仅当

$$(x-1)\frac{\mathrm{d}^2u}{\mathrm{d}x^2}+(x-2)\frac{\mathrm{d}u}{\mathrm{d}x}=0,$$

这就是 $(**)$，正如题目所求。

先解 v，再解 u

设 $v=\dfrac{\mathrm{d}u}{\mathrm{d}x}$，$(**)$ 便化为一阶方程

$$(x-1)\frac{\mathrm{d}v}{\mathrm{d}x}=-(x-2)\,v.$$

其中一个解是 $v=0$。除此之外，在 $x\neq1$ 且 $v\neq0$ 的区间上，我们可以分离变量：

$$\int\frac{1}{v}\,\mathrm{d}v=-\int\frac{x-2}{x-1}\,\mathrm{d}x.$$

想一想 怎样求 $\dfrac{x-2}{x-1}$ 的积分?

把分子用分母表示：$x-2=(x-1)-1$，所以

$$\frac{x-2}{x-1}=1-\frac{1}{x-1}.$$

因此

$$\ln|v|=-x+\ln|x-1|+c,\qquad\text{所以}\qquad v=C(x-1)\mathrm{e}^{-x},$$

其中 $C=\pm\mathrm{e}^c$。若容许 $C=0$, 便连 $v=0$ 这个解也包括在内, 所以 $v=C(x-1)\mathrm{e}^{-x}$, 其中 C 为任意常数。(不难直接验证,这个解对所有 x(包括 $x=1$)都满足 $(x-1)v'+(x-2)v=0$。)

现在求积分以得出 u。运用分部积分法（integration by parts），对 $(x-1)$ 求导、对 e^{-x} 求积分，

$$\begin{aligned}\int(x-1)\mathrm{e}^{-x}\,\mathrm{d}x&=-(x-1)\mathrm{e}^{-x}+\int\mathrm{e}^{-x}\,\mathrm{d}x\\&=-(x-1)\mathrm{e}^{-x}-\mathrm{e}^{-x}=-x\mathrm{e}^{-x}.\end{aligned}$$

所以 $u = -Cx\mathrm{e}^{-x} + B$。把 $-C$ 改记为 A，

$$u = Ax\mathrm{e}^{-x} + B,$$

其中 A 和 B 为任意常数。

最后部分

上文已证明，只要 u 满足 $(**)$，$y = u\mathrm{e}^{x}$ 便满足 $(*)$。我们求得的 $u = Ax\mathrm{e}^{-x} + B$ 满足 $(**)$，因为它的导数是 $v = -A(x-1)\mathrm{e}^{-x}$，而这正是关于 v 的方程的解。因此，对任意常数 A 和 B，

$$y = u\mathrm{e}^{x} = \left(Ax\mathrm{e}^{-x} + B\right)\mathrm{e}^{x} = Ax + B\mathrm{e}^{x}$$

都满足 $(*)$，正如题目所求。

验算

直接验证 $y = x$：$y' = 1$，$y'' = 0$，所以 $(*)$ 的左方是 $0 - x + x = 0$。由于 $(*)$ 是线性的，x 和 e^{x} 这两个解的任何组合 $Ax + B\mathrm{e}^{x}$ 同样是解。再作一次数值验算：取 $A = 2$、$B = -1$ 及 $x = 0.5$，则 $y = 1 - \mathrm{e}^{0.5} \approx -0.64872$，$y' = 2 - \mathrm{e}^{0.5} \approx 0.35128$，$y'' = -\mathrm{e}^{0.5} \approx -1.64872$，得出 $(-0.5)(-1.64872) - 0.5(0.35128) - 0.64872 = 0.82436 - 0.17564 - 0.64872 = 0$。

常见错误

- 忘记积法则在 y'' 中给出的是 $2u'$（而不是 u'）。
- 没有先分拆出 1，便直接把 $\frac{x-2}{x-1}$ 当作对数来积分。
- 对 v 求积分以得出 u 时漏掉常数。正是这个常数 B 把原来的解 e^{x} 带回来。

STEP 之外

降阶法适用于任何线性方程 $y'' + p(x)y' + q(x)y = 0$：若 y_1 是其中一个解，代入 $y = uy_1$ 总会得出一个关于 u' 的一阶方程。配合关于朗斯基行列式（Wronskian）的阿贝尔（Abel）公式，可以证明所有解构成一个二维向量空间（vector space），所以在每个不包含奇点（singular point）$x = 1$ 的区间上，$Ax + B\mathrm{e}^{x}$ 其实就是本题的通解。Boyce 与 DiPrima 的 *Elementary Differential Equations*（一本常微分方程入门教科书）对这两个概念都有讲解。

详细解答

STEP I 2010，第 7 题

纯数学 · 向量与相交直线

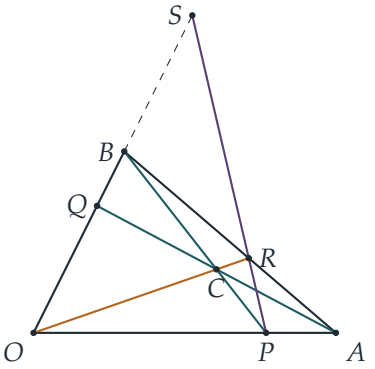
STEP I 2010，第 7 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q7

关键思路

由于 O、A 和 B 不共线（collinear），平面上每一点都可以写成 $\lambda\mathbf{a}+\mu\mathbf{b}$ 的形式，而且写法唯一。因此，要找出两条直线的交点，我们可以把每条直线上的一般点写成这个形式，然后比较 $\mathbf{a}$ 和 $\mathbf{b}$ 的系数（coefficient）。一点位于直线 OB 上，当且仅当它的 $\mathbf{a}$ 系数为零；位于 AB 上，则当且仅当它的两个系数之和为 1。这两个事实几乎包办了全部工作。



图中取 $\alpha=\frac{1}{2}$，$\beta=\frac{7}{20}$。留意 S 位于直线 OB 上，但在线段（segment）OB 之外，而 Q 则位于 O 与 B 之间。

解答

求 p 和 q

直线 BC 上的一般点为

$$\mathbf{b}+\mu(\mathbf{c}-\mathbf{b})=\mu\alpha\,\mathbf{a}+\big(1-\mu(1-\beta)\big)\mathbf{b}.$$

直线 OA 上的点具有 $\lambda\mathbf{a}$ 的形式，没有 $\mathbf{b}$ 分量。由于 $\mathbf{a}$ 与 $\mathbf{b}$ 不平行（因为 O、A、B 不共线），我们可以比较系数。$\mathbf{b}$ 的系数必须为零，所以 $\mu=\dfrac{1}{1-\beta}$；这样做是可以的，因为 $\beta<1-\alpha<1$，所以 $1-\beta\neq 0$。于是

$$\mathbf{p}=\mu\alpha\,\mathbf{a}=\frac{\alpha\mathbf{a}}{1-\beta},$$

正如所求。

整个设置在对调 $A\leftrightarrow B$ 及 $\alpha\leftrightarrow\beta$ 之下是对称的（这同时对调了 $P\leftrightarrow Q$），所以我们可以直接写出

$$\mathbf{q}=\frac{\beta\mathbf{b}}{1-\alpha}.$$

点 R

首先，$\mathbf{r}=\dfrac{1}{\alpha+\beta}\,\mathbf{c}$ 是 $\mathbf{c}$ 的纯量倍数（scalar multiple），所以 R 位于直线 OC 上。

其次，把它写成 $\mathbf{r}=t\mathbf{a}+(1-t)\mathbf{b}$，其中 $t=\dfrac{\alpha}{\alpha+\beta}$。事实上 $1-t=\dfrac{\beta}{\alpha+\beta}$，这正是 $\mathbf{b}$ 在 $\mathbf{r}$ 中的系数。形如 $t\mathbf{a}+(1-t)\mathbf{b}=\mathbf{b}+t(\mathbf{a}-\mathbf{b})$ 的点恰好就是直线 AB 上的点，所以 R 也位于 AB 上。

点 S

直线 PR 上的一般点为 $\mathbf{p}+t(\mathbf{r}-\mathbf{p})$，它等于

$$\left[\frac{\alpha}{1-\beta}+t\left(\frac{\alpha}{\alpha+\beta}-\frac{\alpha}{1-\beta}\right)\right]\mathbf{a}\;+\;\frac{t\beta}{\alpha+\beta}\,\mathbf{b}.$$

要使这一点位于 OB 上，$\mathbf{a}$ 的系数必须为零。两边除以 $\alpha\neq 0$，得

$$\frac{1}{1-\beta}=t\cdot\frac{(\alpha+\beta)-(1-\beta)}{(1-\beta)(\alpha+\beta)},\qquad\text{所以}\qquad t=\frac{\alpha+\beta}{\alpha+2\beta-1}.$$

（题目告诉我们这两条直线相交，这就迫使 $\alpha+2\beta\neq 1$。若 $\alpha+2\beta=1$，则 PR 与 OB 平行。）因此

$$\mathbf{s}=\frac{t\beta}{\alpha+\beta}\,\mathbf{b}=\frac{\beta}{\alpha+2\beta-1}\,\mathbf{b}.$$

想一想 怎样由位置向量（position vector）$k\mathbf{b}$ 得出该点与 O 的距离和与 B 的距离之比?

若 X 的位置向量为 $k\mathbf{b}$，则 $\overrightarrow{OX} = k\mathbf{b}$ 而 $\overrightarrow{BX} = (k-1)\mathbf{b}$，所以

$$\frac{OX}{BX} = \frac{|k|}{|1-k|}.$$

这里的绝对值（absolute value）很重要，因为 k 可以是负数，也可以大于 1，图中的 S 正是如此。

两个比

对于 Q, 我们有 $k = \dfrac{\beta}{1-\alpha}$, 而且 $0 < k < 1$, 因为 $0 < \beta < 1-\alpha$。于是 $1-k = \dfrac{1-\alpha-\beta}{1-\alpha}$, 所以

$$\frac{OQ}{BQ} = \frac{\beta}{1-\alpha-\beta}.$$

对于 S，我们有 $k = \dfrac{\beta}{\alpha+2\beta-1}$，所以

$$1-k = \frac{\alpha+2\beta-1-\beta}{\alpha+2\beta-1} = \frac{\alpha+\beta-1}{\alpha+2\beta-1}, \qquad \frac{k}{1-k} = \frac{\beta}{\alpha+\beta-1} = -\frac{\beta}{1-\alpha-\beta}.$$

取绝对值，并利用 $\beta > 0$ 和 $1-\alpha-\beta > 0$，得

$$\frac{OS}{BS} = \left|\frac{k}{1-k}\right| = \frac{\beta}{1-\alpha-\beta} = \frac{OQ}{BQ}.$$

证明至此完成。取绝对值之前的那个负号说明：Q 内分（divides internally）OB，而 S 则外分（divides externally）它，两者的比相同。

验算

- 试取 $\alpha = \frac{1}{2}$，$\beta = \frac{7}{20}$（即图中的情况）。这时 $\mathbf{q} = 0.7\mathbf{b}$ 而 $\mathbf{s} = 1.75\mathbf{b}$，所以 $OQ/BQ = 0.7/0.3 = \frac{7}{3}$ 及 $OS/BS = 1.75/0.75 = \frac{7}{3}$。公式给出 $\beta/(1-\alpha-\beta) = 0.35/0.15 = \frac{7}{3}$。
- 我们随机选取 $\mathbf{a}$、$\mathbf{b}$、α、β，直接解成对的直线方程以计算全部四个交点，再与公式比较，并检查两个比是否相等。在超过 8000 个个案中，结果全部吻合。
- 对称性：对调 $\alpha \leftrightarrow \beta$ 及 $\mathbf{a} \leftrightarrow \mathbf{b}$，会把 $\mathbf{p}$ 的公式变成 $\mathbf{q}$ 的公式，正该如此。

常见错误

- 忘记取绝对值。$\mathbf{s}$ 的系数可以是负数，也可以大于 1，这时不小心直接用 $k/(1-k)$，便会得出一个负的“长度之比”。

► 比较 **a** 和 **b** 的系数，却没有说明为甚么可以这样做。这个做法之所以成立，是因为 **a** 与 **b** 不平行。

STEP 之外

点 Q 和 S 以相同的比内分和外分 OB，所以它们是关于 O 和 B 的调和共轭（*harmonic conjugates*）。这里的图形，即四条直线连同它们的交点，就是经典的完全四边形（*complete quadrilateral*）。调和分割（harmonic division）在投影（projection）之下保持不变，是射影几何（projective geometry）的核心概念之一。把点用两个独立向量表示，再比较系数，这个习惯正是线性代数（linear algebra）的起点；在线性代数中，它发展成为基（basis）的概念。Strang 的 *Introduction to Linear Algebra*（线性代数入门教材）阐述了这方面的内容。

详细解答

STEP I 2010，第 8 题

纯数学 · 无穷递降法

STEP I 2010，第 8 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q8

关键思路

两部分用的是同一个技巧，称为*无穷递降法*（infinite descent）。首先考虑方程模（modulo）某个质数（prime）的情况（第 (i) 部分取 3，第 (ii) 部分取 5），会发现这个质数必定整除每一个未知数。把它除去之后，便得到*完全相同的方程*的一个更小的解，而一个非零整数不可能无止境地被 3（或 5）整除。真正要花工夫的，只是找出哪些余数（residue）可能出现，而在第 (ii) 部分，四次方（fourth power）的末位数字已替我们完成这一步。

解答

第 (i) 部分

移项整理，得

$$a^3 = 9c^3 - 3b^3 = 3\left(3c^3 - b^3\right),$$

所以 3 整除 a^3。由于 3 是质数，若它整除乘积 $a \cdot a \cdot a$，则它必整除其中一个因子，所以 $3 \mid a$。

设 $a = 3a_1$，其中 a_1 为整数。代入后除以 3：

$$27a_1^3 + 3b^3 = 9c^3 \quad \Longrightarrow \quad 9a_1^3 + b^3 = 3c^3,$$

所以 $b^3 = 3\left(c^3 - 3a_1^3\right)$ 可被 3 整除，同理可得 $3 \mid b$。设 $b = 3b_1$；则

$$9a_1^3 + 27b_1^3 = 3c^3 \quad \Longrightarrow \quad c^3 = 3\left(a_1^3 + 3b_1^3\right),$$

所以亦有 $3 \mid c$。这证明了 a、b 和 c 全都可被 3 整除。

现在再设 $c = 3c_1$。最后一条方程变成 $27c_1^3 = 3a_1^3 + 9b_1^3$，即

$$a_1^3 + 3b_1^3 = 9c_1^3.$$

? **想一想** 我们已证明 3 整除所有未知数。为甚么这样就迫使它们全部等于零?

因为 $(a_1, b_1, c_1) = (a/3,\, b/3,\, c/3)$ 同样是原来方程的一个整数解，所以刚才证明的一切对它同样适用：a_1、b_1、c_1 可被 3 整除，即 a、b、c 可被 9 整除。把这个论证重复 k 次，可知对每个 $k \geq 1$，3^k 都整除 a、b 和 c 各数。但一旦 $3^k > |n|$，非零整数 n 便不能被 3^k 整除。因此 $a = b = c = 0$，而这显然是一个解，所以它是唯一的解。

（等价的说法：假如存在 $(0,0,0)$ 以外的解，从中选取 $|a| + |b| + |c|$ 为正数而又尽可能小的一个；那么 $(a/3, b/3, c/3)$ 也是一个解，而且对应的正数值更小，矛盾。）

第 (ii) 部分

由于 $p^4 = |p|^4$，p^4 的末位数字只取决于 $|p|$ 的末位数字：

n 的末位数字	0	1	2	3	4	5	6	7	8	9
n^4 的末位数字	0	1	6	1	6	5	6	1	6	1

（例如 $2^4 = 16$、$3^4 = 81$、$7^4 = 2401$。）所以四次方的末位数字是 0、1、5 或 6，而且末位是 0 或 5 当且仅当 $5 \mid n$。

现在逐项看末位数字。

- $5r^4$ 是 5 的倍数，所以末位是 0 或 5。
- $2q^4$ 的末位是 0（若 q^4 的末位是 0 或 5）或 2（若 q^4 的末位是 1 或 6）。
- p^4 的末位是 0、1、5 或 6。

把 $\{0,1,5,6\}$ 中的一个末位数字与 $\{0,2\}$ 中的一个相加，结果的末位是 $0,1,5,6,2,3,7,8$ 其中之一。要得到 0 或 5，只有 $0+0$ 和 $5+0$ 两种方式。所以 $2q^4$ 的末位是 0，也就是说 q^4 的末位是 0 或 5，即 $5 \mid q$；而 p^4 的末位是 0 或 5，即 $5 \mid p$。这证明了 p 和 q 可被 5 整除。

（另一做法：根据费马小定理（Fermat's little theorem），当 $5 \nmid n$ 时 $n^4 \equiv 1 \pmod 5$，所以 $p^4 + 2q^4$ 模 5 同余（congruent）于 $0+0$、$1+0$、$0+2$ 或 $1+2$，当中只有第一个是 0。）

设 $p = 5p_1$ 及 $q = 5q_1$。则

$$625\left(p_1^4 + 2q_1^4\right) = 5r^4 \quad \Longrightarrow \quad r^4 = 125\left(p_1^4 + 2q_1^4\right),$$

所以 $5 \mid r^4$，又因为 5 是质数，故 $5 \mid r$。设 $r = 5r_1$，得

$$625r_1^4 = 125\left(p_1^4 + 2q_1^4\right) \quad \Longrightarrow \quad p_1^4 + 2q_1^4 = 5r_1^4.$$

所以 $(p/5, q/5, r/5)$ 同样是同一方程的整数解。

与第 (i) 部分完全一样，重复这个步骤可知对每个 k，5^k 都整除 p、q 和 r 各数，而这对非零整数来说是不可能的。因此唯一的整数解是 $p = q = r = 0$。

验算

用电脑搜索，在 $|a|, |b|, |c| \leq 60$ 的范围内，$a^3 + 3b^3 = 9c^3$ 除 $(0,0,0)$ 外没有其他解；在 $|p|, |q|, |r| \leq 150$ 的范围内，$p^4 + 2q^4 = 5r^4$ 亦然。逐一检查 p 和 q 末位数字的全部 100 种组合，可确认只有当两个末位数字都是 0 或 5 时，$p^4 + 2q^4$ 的末位才是 0 或 5。再用手算快速核对一下：取 $p = q = 1$，得 $p^4 + 2q^4 = 3$，不是 5 的倍数，与上表所示一致。

常见错误

- 证明了 3（或 5）整除每个未知数之后便停下来。单凭这一点并不能得出零；关键在于除去该因数后，得到的是同一条方程的解，所以论证可以重复下去。
- 由“$3 \mid a^3$”推出“$3 \mid a$”而不说明理由。这一步成立是因为 3 是质数；换成例如由 $9 \mid a^3$ 推出 $9 \mid a$ 就不成立（试试 $a = 3$）。
- 列出末位数字时忘记了负整数。四次方是偶数次幂，所以 $p^4 = |p|^4$，上表已涵盖所有情况。

STEP 之外

无穷递降法可追溯至费马（Fermat），他用这个方法证明了 $x^4 + y^4 = z^2$ 没有正整数解（从而证明了费马最后定理（Fermat's Last Theorem）中 $n = 4$ 的情况）。这里用到的想法是：如果一条方程模某个质数的幂没有非平凡（non-trivial）解，它就不可能有非平凡的整数解；这个想法进一步发展，便成为 p 进数（p-adic numbers）和哈塞原理（Hasse principle）的研究。达文波特（Davenport）的 *The Higher Arithmetic* 是一本介绍同余与递降法的优美入门书。

详细解答

STEP I 2010，第 9 题

力学 • 摩擦力与平衡

STEP I 2010，第 9 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q9

关键思路

先把地面上的那个顶点放在一个已知位置，再沿两条互相垂直的边走到质心（centre of mass）：这样便得出质心与墙的距离。至于平衡（equilibrium），“两个接触点均处于极限状态”意味着每个摩擦力（friction）都等于 μ 乘以相应的法向反作用力（normal reaction），因此沿两个方向分解力（resolving）后，两个反作用力都可用重量表示。之后只需一条力矩（moment）方程便大功告成，而设 $\mu = \tan\lambda$ 则可把代数运算化为复角公式（compound-angle formulae）。

解答

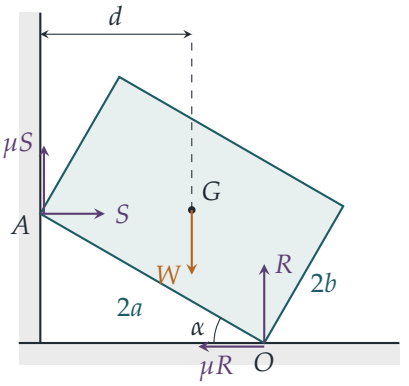
模型

薄片（lamina）是一块重量为 W 的均匀刚性长方形，所以它的质心 G 就是它的中心。设 O 为地面上的顶点，A 为靠墙的顶点，于是 OA 是一条长度为 $2a$ 的边，与地面成角 α，而由 O 出发、长度为 $2b$ 的边则向上并向远离墙的方向伸展。各力如下：

- 在 O 点：法向反作用力 R 向上，摩擦力 F 沿水平方向；
- 在 A 点：法向反作用力 S 沿水平方向指离墙壁，摩擦力 P 沿铅垂方向；
- 重量 W 作用于 G，方向向下。

薄片正处于即将滑动的状态，O 会向远离墙的方向滑动，A 则会沿墙向下滑动，而摩擦力阻碍这些运动：F 指向墙壁，P 向上。由于两个接触点的摩擦力均达极限（limiting friction），

$$F = \mu R, \qquad P = \mu S.$$



质心与墙的距离

由于 OA 的长度为 $2a$，且与水平线成角 α，所以 O 与墙的水平距离为 $2a\cos\alpha$。

由 O 出发，沿 O 点的两条边各走一半，便到达 G。沿 OA（向墙的方向）走 a，我们便向墙移近了 $a\cos\alpha$。长度为 $2b$ 的边与 OA 垂直，所以它与水平线成角 $\frac{1}{2}\pi-\alpha$，并向远离墙的方向倾斜；沿它走 b，我们便远离墙壁 $b\cos(\frac{1}{2}\pi-\alpha)=b\sin\alpha$。因此 G 与墙的距离为

$$d=2a\cos\alpha-a\cos\alpha+b\sin\alpha=a\cos\alpha+b\sin\alpha,$$

正如所求。同理，G 在地面之上的高度为 $a\sin\alpha+b\cos\alpha$，而 A 在地面之上的高度为 $2a\sin\alpha$。

平衡条件

沿水平和铅垂方向分解力，

$$S=\mu R,$$
$$R+\mu S=W.$$

把第一式代入第二式，得 $R(1+\mu^2)=W$，所以

$$R=\frac{W}{1+\mu^2},\qquad S=\frac{\mu W}{1+\mu^2}.$$

现在对 O 取力矩，这样便消去了 R 和 μR。在水平方向上，G 比 O 更接近墙壁 $2a\cos\alpha-d=a\cos\alpha-b\sin\alpha$，所以重量的力矩为 $W(a\cos\alpha-b\sin\alpha)$，倾向使薄片朝墙的方向转动。在 A 点，力 S 的作用线位于 O 之上 $2a\sin\alpha$ 的高度，而 μS 沿墙作用，与 O 的水平距离为 $2a\cos\alpha$；两者都倾向使薄片朝远离墙的方向转动。所以

$$W(a\cos\alpha-b\sin\alpha)=S\cdot 2a\sin\alpha+\mu S\cdot 2a\cos\alpha=\frac{2a\mu W(\sin\alpha+\mu\cos\alpha)}{1+\mu^2}.\qquad(*)$$

? **想一想** 设 $\mu = \tan\lambda$ 如何把右边化简?

设 $\mu = \tan\lambda$，则 $1 + \mu^2 = \sec^2\lambda$，所以

$$\frac{\mu}{1+\mu^2} = \tan\lambda\cos^2\lambda = \sin\lambda\cos\lambda,$$

$$\sin\alpha + \mu\cos\alpha = \frac{\sin\alpha\cos\lambda + \cos\alpha\sin\lambda}{\cos\lambda} = \frac{\sin(\alpha+\lambda)}{\cos\lambda}.$$

因此 $(*)$ 的右边是 $2aW\sin\lambda\sin(\alpha+\lambda)$。约去 W，并利用 $2\sin P\sin Q = \cos(P-Q) - \cos(P+Q)$，其中 $P = \alpha+\lambda$，$Q = \lambda$，

$$\begin{aligned} b\sin\alpha &= a\cos\alpha - 2a\sin\lambda\sin(\alpha+\lambda) \\ &= a\cos\alpha - a\big(\cos\alpha - \cos(\alpha+2\lambda)\big) \\ &= a\cos(2\lambda+\alpha), \end{aligned}$$

这就是所求的结果。

正方形薄片

若薄片是正方形，则 $b = a$，两边除以 $a \neq 0$，

$$\cos(2\lambda+\alpha) = \sin\alpha = \cos\left(\tfrac{1}{2}\pi - \alpha\right).$$

所以对某个整数 k，有 $2\lambda+\alpha = \pm\left(\frac{1}{2}\pi - \alpha\right) + 2k\pi$。我们需要选出正确的分支。这里 $0 < \alpha < \frac{1}{2}\pi$（薄片是倾斜的），而且由于 $\mu > 0$，有 $0 < \lambda < \frac{1}{2}\pi$，所以 $0 < 2\lambda+\alpha < \frac{3}{2}\pi$。

- 取 $+$ 号得 $2\lambda+\alpha = \frac{1}{2}\pi - \alpha + 2k\pi$；只有 $k = 0$ 落在 $(0, \frac{3}{2}\pi)$ 之内（因为 $\frac{1}{2}\pi - \alpha$ 位于 $(0, \frac{1}{2}\pi)$ 之内），由此得 $\lambda = \frac{1}{4}\pi - \alpha$。
- 取 $-$ 号得 $2\lambda = -\frac{1}{2}\pi + 2k\pi$，即 $\lambda = -\frac{1}{4}\pi + k\pi$，而这样的值没有一个位于 $(0, \frac{1}{2}\pi)$ 之内。

因此 $\lambda = \frac{1}{4}\pi - \alpha$，正如所求。（特别地，这种极限平衡需要 $\alpha < \frac{1}{4}\pi$，这与 G 必须位于 O 靠墙一侧的要求吻合：$a\cos\alpha > a\sin\alpha$。）

验算

- 若 $\mu = 0$（$\lambda = 0$），条件变为 $a\cos\alpha = b\sin\alpha$：$G$ 在 O 的正上方，这正是薄片无需墙壁帮助、仅靠顶点便能平衡的情况。
- 对于 $\alpha = 15^\circ$ 的正方形，结果给出 $\lambda = 30^\circ$，所以 $\mu = 1/\sqrt{3} \approx 0.577$。把 $a = b = 1$、$\alpha = \frac{\pi}{12}$、$\mu = 1/\sqrt{3}$ 代入 $(*)$，两边都等于 $W(\cos 15^\circ - \sin 15^\circ) \approx 0.707W$。

- 在数值方面，我们随机选取了 2000 组长方形和角度，以二分法（bisection）从力和力矩方程求出 μ，并确认 $a\cos(2\lambda+\alpha)=b\sin\alpha$ 成立，误差在 10^{-14} 以内。

常见错误

- 弄错摩擦力的方向：应先判断每个接触点会向哪个方向滑动，然后取相反的方向。
- 把与墙的距离 $a\cos\alpha+b\sin\alpha$ 当作对 O 的力臂（moment arm）；对 O 的力臂其实是 $a\cos\alpha-b\sin\alpha$。
- 忘记了 $\lambda=\frac{3}{4}\pi$ 也是该三角方程的解；它之所以被排除，是因为 $\mu>0$ 使得 $0<\lambda<\frac{1}{2}\pi$。

STEP 之外

满足 $\tan\lambda=\mu$ 的角 λ 称为摩擦角（angle of friction）：在极限摩擦时，总接触力与法线成角 λ。另有一个巧妙的方法，利用的是以下事实：刚体（rigid body）在三个力作用下平衡，这三个力必须共点（concurrent）：两个总接触力与重量必须交于一点，于是问题便化为纯粹的几何问题。Kleppner 与 Kolenkow 的 *An Introduction to Mechanics*（力学入门）在讲述牛顿力学其他内容的同时，也阐述了刚体静力学和摩擦。

详细解答

STEP I 2010，第 10 题

力学 · 向量运动学

STEP I 2010，第 10 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q10

关键思路

一切都源自对 $\mathbf{r}$ 求导再取点积(dot product)：两个向量之间的夹角可由 $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$ 求得，而垂直 (perpendicular) 即表示点积为零。至于第二个质点，关键的观察是：Q 在时刻 t 的位置，正是 P 在时刻 $t-T$ 的位置。把位置写成复数 (complex number) $z = \mathrm{e}^{(1+\mathrm{i})t}$，最后一部分便化为一行计算。

解答

速度与加速度

记 $\mathbf{r} = \mathrm{e}^t(\cos t,\ \sin t)$。由乘积法则（product rule），

$$\begin{aligned}\mathbf{v} = \dot{\mathbf{r}} &= \mathrm{e}^t\big(\cos t - \sin t,\ \sin t + \cos t\big),\\ \mathbf{a} = \dot{\mathbf{v}} &= \mathrm{e}^t\big((\cos t - \sin t) + (-\sin t - \cos t),\ (\sin t + \cos t) + (\cos t - \sin t)\big)\\ &= \mathrm{e}^t\big(-2\sin t,\ 2\cos t\big).\end{aligned}$$

速度与位移之间的夹角

两者的大小分别为 $|\mathbf{r}| = \mathrm{e}^t$ 及

$$|\mathbf{v}|^2 = \mathrm{e}^{2t}\left[(\cos t - \sin t)^2 + (\sin t + \cos t)^2\right] = \mathrm{e}^{2t}\cdot 2,$$

因为交叉项 $\mp 2\sin t\cos t$ 互相抵消。所以 $|\mathbf{v}| = \sqrt{2}\,\mathrm{e}^t$。点积为

$$\mathbf{r}\cdot\mathbf{v} = \mathrm{e}^{2t}\left[\cos t(\cos t - \sin t) + \sin t(\sin t + \cos t)\right] = \mathrm{e}^{2t}(\cos^2 t + \sin^2 t) = \mathrm{e}^{2t}.$$

因此，若 θ 为 $\mathbf{r}$ 与 $\mathbf{v}$ 之间的夹角，则

$$\cos\theta = \frac{\mathbf{r}\cdot\mathbf{v}}{|\mathbf{r}||\mathbf{v}|} = \frac{\mathrm{e}^{2t}}{\mathrm{e}^t\cdot\sqrt{2}\,\mathrm{e}^t} = \frac{1}{\sqrt{2}},$$

所以对任何 t 都有 $\theta = \frac{\pi}{4}$，正如题目所求。（由于 $\mathrm{e}^t > 0$，两个向量都永不为零，所以夹角恒有定义。）

加速度与位移垂直

$$\mathbf{r}\cdot\mathbf{a} = \mathrm{e}^{2t}\left[-2\sin t\cos t + 2\sin t\cos t\right] = 0,$$

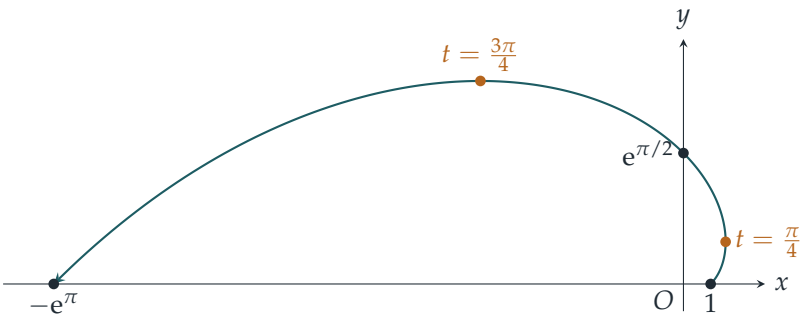
而且两个向量均非零，所以 $\mathbf{a}$ 恒与 $\mathbf{r}$ 垂直。还可留意，$\mathbf{a} = 2\mathrm{e}^t(-\sin t, \cos t)$ 正是把 $\mathbf{r}$ 旋转 $\frac{\pi}{2}$ 后再放大至两倍。

$0 \le t \le \pi$ 时的路径

在极坐标（polar coordinates）中，路径为 $r = \mathrm{e}^\theta$，其中 $\theta = t$：质点沿逆时针方向转动时，与 O 的距离按指数增长，这是一条等角螺线（equiangular spiral），又称对数螺线（logarithmic spiral）。几个有用的参考点：

- $t = 0$：位于点 $(1, 0)$，沿方向 $(1, 1)$ 运动，与 x 轴成 $\frac{\pi}{4}$ 角。
- $t = \frac{\pi}{4}$：$\dot{x} = \mathrm{e}^t(\cos t - \sin t) = 0$，所以 x 达到最大，位于点 $\left(\frac{1}{\sqrt{2}}\mathrm{e}^{\pi/4}, \frac{1}{\sqrt{2}}\mathrm{e}^{\pi/4}\right) \approx (1.55, 1.55)$。
- $t = \frac{\pi}{2}$：位于点 $(0, \mathrm{e}^{\pi/2}) \approx (0, 4.81)$。
- $t = \frac{3\pi}{4}$：$\dot{y} = \mathrm{e}^t(\sin t + \cos t) = 0$，所以 y 达到最大，位于 $\left(-\frac{1}{\sqrt{2}}\mathrm{e}^{3\pi/4}, \frac{1}{\sqrt{2}}\mathrm{e}^{3\pi/4}\right) \approx (-7.46, 7.46)$。
- $t = \pi$：位于点 $(-\mathrm{e}^\pi, 0) \approx (-23.1, 0)$，以 $\frac{\pi}{4}$ 角穿过坐标轴，朝左下方前进。

在路径上每一点，路径都与由 O 出发的径向线成 $\frac{\pi}{4}$ 角，其形状由此确定。



第二个质点

想一想 Q 在时刻 t 位于何处？试以 P 的运动来表示。

Q 到达每一点的时间都比 P 迟 T，所以在时刻 t，它正处于 P 在时刻 $t-T$ 的位置：

$$\mathbf{r}_Q(t) = \mathbf{r}(t-T) = \mathrm{e}^{t-T}\big(\cos(t-T),\, \sin(t-T)\big).$$

于是

$$\begin{aligned} |PQ|^2 &= |\mathbf{r}(t)|^2 + |\mathbf{r}(t-T)|^2 - 2\,\mathbf{r}(t)\cdot\mathbf{r}(t-T) \\ &= \mathrm{e}^{2t} + \mathrm{e}^{2t-2T} - 2\mathrm{e}^{2t-T}\big(\cos t\cos(t-T) + \sin t\sin(t-T)\big) \\ &= \mathrm{e}^{2t}\big(1 - 2\mathrm{e}^{-T}\cos T + \mathrm{e}^{-2T}\big), \end{aligned}$$

其中用到 $\cos t\cos(t-T) + \sin t\sin(t-T) = \cos\big(t-(t-T)\big) = \cos T$。因此

$$|PQ| = k\,\mathrm{e}^{t}, \qquad k = \sqrt{1 - 2\mathrm{e}^{-T}\cos T + \mathrm{e}^{-2T}},$$

而 k 只取决于 T，与 t 无关。所以 P 与 Q 之间的距离与 e^t 成正比，正如题目所求。（括号内的式子等于 $|1-\mathrm{e}^{-T}\mathrm{e}^{\mathrm{i}T}|^2$，当 $T\neq 0$ 时为正数，因为 $\mathrm{e}^{-T}\neq 1$；所以 $k>0$。）

*复数版本。*设 $z = x+\mathrm{i}y$，则 P 位于 $z_P = \mathrm{e}^t\mathrm{e}^{\mathrm{i}t} = \mathrm{e}^{(1+\mathrm{i})t}$，而 Q 位于 $z_Q = \mathrm{e}^{(1+\mathrm{i})(t-T)}$。所以

$$z_P - z_Q = \mathrm{e}^{(1+\mathrm{i})t}\big(1-\mathrm{e}^{-(1+\mathrm{i})T}\big), \qquad |z_P - z_Q| = \mathrm{e}^t\,\big|1-\mathrm{e}^{-(1+\mathrm{i})T}\big|.$$

这亦显示 $\overrightarrow{QP}$ 是 z_P 的一个固定复数倍，所以线段 PQ 与 OP 的夹角保持不变：三角形 OQP 的形状始终如一，只是不断放大和转动。

验算

- 对 $z = \mathrm{e}^{(1+\mathrm{i})t}$ 求导，得 $\dot z = (1+\mathrm{i})z$ 及 $\ddot z = (1+\mathrm{i})^2 z = 2\mathrm{i}z$。乘以 $1+\mathrm{i} = \sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}$ 相当于旋转 $\frac{\pi}{4}$，乘以 $2\mathrm{i}$ 则相当于旋转 $\frac{\pi}{2}$：最初的两个结果一目了然。
- 以数值验证：随机选取 $t, \mathbf{r}$ 与 $\mathbf{v}$ 之间的夹角算得 $0.785398\ldots = \frac{\pi}{4}$，且 $\mathbf{r}\cdot\mathbf{a} = 0$。取 $T=1$ 时，在所有测试过的 t 值上都有 $|PQ|/\mathrm{e}^t = 0.85895\ldots$，与 $\sqrt{1-2\mathrm{e}^{-1}\cos 1+\mathrm{e}^{-2}}$ 吻合。
- 当 $T\to 0$ 时，$k\approx T\sqrt{2}$，所以 $|PQ|\approx\sqrt{2}\,\mathrm{e}^t T = |\mathbf{v}|T$，即在时间 T 内走过的距离，理应如此。

常见错误

- 把 Q 的位置当作 $\mathbf{r}(t+T)$：Q 在 P 的后面，所以它位于 $\mathbf{r}(t-T)$。（答案仍然与 e^t 成正比，但常数不同。）
- 把螺线画成顺时针方向转动，或者画成路径经过原点。质点与 O 的距离为 $\mathrm{e}^t \geq 1$，并随着质点沿逆时针方向运动而增加。

STEP 之外

曲线 $r = \mathrm{e}^\theta$ 就是对数螺线：以 O 为中心把它缩放，所得的曲线都是它本身旋转后的结果，这正是相交角保持不变、以及 OQP 形状不变的原因。在大学一年级的课程中，平面上的运动常以极坐标单位向量 $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}$ 来处理；在这套表示法下，本题的径向加速度 $\ddot{r} - r\dot{\theta}^2$ 等于零，直接得出垂直的结论。Kleppner 与 Kolenkow 合著的 *An Introduction to Mechanics*（力学导论）正是为这类问题而发展极坐标的方法。

详细解答

STEP I 2010，第 11 题

力学 · 圆上的碰撞

STEP I 2010，第 11 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

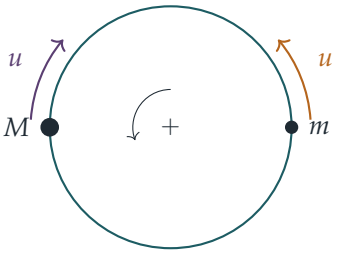
step.maths.org/questions/10-s1-q11

关键思路

虽然两个质点（particle）在圆上运动，但每次碰撞（collision）其实都是沿公切线方向的普通一维碰撞：光滑凹槽只会在垂直于运动的方向施力，所以它既不能改变速率，也不会影响切向的冲量（impulse）。因此我们采用动量守恒（conservation of momentum）和牛顿实验定律（Newton's experimental law），并以沿圆周量度的*带符号*速度来列式。此后，两个事实已包含全部信息：带符号的总动量永远不变，而每次碰撞都令相对速率乘以 e。

解答

我们把两个质点看成光滑水平凹槽内的点。取逆时针方向为正，设 m 最初以 u 逆时针运动，M 以 u 顺时针运动，所以它们的带符号速度分别是 u 和 $-u$。



在两次碰撞之间，每个质点所受的唯一水平力就是光滑凹槽的法向反作用力（normal reaction），它垂直于速度。这个力不作功，所以每个质点的速率保持不变。碰撞时，两个质点互相施加的冲量沿公切线方向，大小相等、方向相反，而凹槽并不提供任何切向冲量。因此每次碰撞中，沿圆周的带符号总动量不变，而牛顿定律亦适用于带符号速度。

第一次碰撞

设第一次碰撞后的带符号速度为 p（对于 m）和 q（对于 M）。由动量守恒和牛顿定律得

$$mp + Mq = mu - Mu,$$
$$q - p = e\big(u - (-u)\big) = 2eu.$$

把 $q = p + 2eu$ 代入第一条方程，得 $(m+M)p = (m-M)u - 2eMu$，所以

$$p = -\frac{u\,(M - m + 2eM)}{M+m}, \qquad q = p + 2eu = \frac{u\,\big(2em - (M-m)\big)}{M+m}.$$

因此第一次碰撞后，两个质点的速率分别为

$$\text{质点}m = \frac{u\,(M - m + 2eM)}{M+m}, \qquad \text{质点}M = \frac{u\,|2em - (M-m)|}{M+m}.$$

想一想 哪些正负号表示质点改变了方向?

质点 m 原本沿正方向运动，所以它恰好在 $p < 0$ 时反向。由于 $M > m$ 且 $e \geq 0$，括号内的 $M - m + 2eM$ 是正数，所以 m 必定反向。质点 M 原本沿负方向运动，所以它恰好在 $q > 0$ 时反向，即当

$$2em > M - m.$$

所以在这个条件下，两个质点都改变方向，正如题目所要求。

再经过 $2n$ 次碰撞之后

为何 $mv - MV = u(M-m)$。开始时，带符号的总动量是 $mu - Mu = -u(M-m)$。碰撞令它守恒，而在两次碰撞之间速率不变，所以它永远保持这个值。

由于两个质点在每次碰撞时都反向，第一次碰撞后 m 顺时针运动而 M 逆时针运动；其后每次碰撞都把这个情况对调。再经过 $2n$ 次碰撞之后，总共发生了 $2n+1$ 次碰撞，是奇数次，所以 m 又是顺时针运动，M 又是逆时针运动。它们的带符号速度是 $-v$ 和 V，而

$$m(-v) + MV = -u(M-m) \quad \Longrightarrow \quad mv - MV = u(M-m).$$

相对速率。每次碰撞后，两个质点以恒定速率沿圆周向相反方向分开，然后在圆的另一边再次相遇，而互相接近时的相对速率与分开时相同。牛顿定律指出，碰撞后的分离速率等于碰撞前接近速率的 e 倍。最初的接近速率是 $2u$，所以经过 $2n+1$ 次碰撞后，分离速率是 $2ue^{2n+1}$。由于两个质点向相反方向运动，这个分离速率就是 $v + V$：

$$v + V = 2ue^{2n+1}.$$

求解。把第二条方程乘以 M，再与第一条相加：

$$(m+M)v = u(M-m) + 2Mue^{2n+1},$$

又把第二条方程乘以 m，再减去第一条：

$$(m+M)V = 2mue^{2n+1} - u(M-m).$$

因此

$$v = \frac{u\big(M-m+2Me^{2n+1}\big)}{M+m}, \qquad V = \frac{u\big(2me^{2n+1}-(M-m)\big)}{M+m}.$$

验算

- 代入 $n=0$，便得回上面求得的第一次碰撞后的速率。
- 取 $m=2$、$M=3$、$e=0.8$、$u=1$，公式在 $n=0$ 时给出 $(v,V)=(1.16,\,0.44)$，在 $n=1$ 时给出 $(0.8144,\,0.2096)$。逐步模拟两个质点绕圆运动并碰撞的过程，在第 1 次和第 3 次碰撞后得到的正是这些数值。
- 动能（kinetic energy）不应增加。当 $e=1$ 时，$v = u\,\frac{3M-m}{M+m}$ 而 $V = u\,\frac{3m-M}{M+m}$，你可以验证 $\frac{1}{2}mv^2 + \frac{1}{2}MV^2 = \frac{1}{2}(m+M)u^2$。

常见错误

- 混淆速率与速度：牛顿定律和动量守恒都必须以带符号速度来写。到最后才转换为速率。
- 忘记“再经过 $2n$ 次碰撞”即总共 $2n+1$ 次，所以幂次是 e^{2n+1} 而不是 e^{2n}。
- 为凹槽提供的向心力（centripetal force）而担心。它垂直于运动方向，所以既不改变速率，也不会出现在碰撞方程之中。

STEP 之外

V 的公式显示，当 $e<1$ 时，题目给定的假设不可能永远成立：一旦 $2me^{2n+1}$ 跌至低于 $M-m$，较重的质点便不再反向。这里的动量只是在旋转的意义下守恒。凹槽的力总是指向通过圆心的方向，所以真正守恒的量是关于圆心的*角动量*（angular momentum），而圆的半径刚好约去。这类源自对称性的守恒定律，是大学一年级力学的核心主题；Kleppner 与 Kolenkow 所著的 *An Introduction to Mechanics*（力学导论）对角动量和碰撞有仔细的论述。

详细解答

STEP I 2010，第 12 题

概率与统计 · 以尾和求期望值

STEP I 2010，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q12

关键思路

尾和公式（tail-sum formula）的来源，是把 $n\,\mathrm{P}(X=n)$ 写成 $\mathrm{P}(X=n)$ 相加 n 次，然后把各项按另一种次序重新归并。就企鹅这个问题而言，事件 $X \geq n$ 比 $X = n$ 容易描述得多：开了 $n-1$ 个盒之后仍须再开，当且仅当这 $n-1$ 个盒里全是同一种企鹅。因此尾概率（tail probability）十分简单，而刚证明的公式就把 $\mathrm{E}(X)$ 化成两个等比级数（geometric series）。

解答

定义与尾和公式

由于 X 只取 $1, 2, 3, \ldots$ 这些值，

$$\mathrm{E}(X) = \sum_{m=1}^{\infty} m\,\mathrm{P}(X=m),$$

前提是这个级数收敛（converge）。

对每个 $n \geq 1$，事件 $X \geq n$ 是各事件 $X = m$（$m \geq n$）的不相交并集（disjoint union），所以 $\mathrm{P}(X \geq n) = \sum_{m \geq n} \mathrm{P}(X=m)$。因此

$$\begin{aligned}\sum_{n=1}^{\infty} \mathrm{P}(X \geq n) &= \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} \mathrm{P}(X=m) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{m} \mathrm{P}(X=m) \\ &= \sum_{m=1}^{\infty} m\,\mathrm{P}(X=m) = \mathrm{E}(X).\end{aligned}$$

想一想 为甚么可以对调两个求和的次序?

这个二重和是把 $\mathrm{P}(X=m)$ 对所有满足 $1 \le n \le m$ 的数对 (n,m) 相加。第一行是固定 n，对 m 求和；第二行是固定 m，对 $n=1,\ldots,m$ 这 m 个值求和。每一项都是非负的，而对于非负项级数，总和与求和的次序无关（两边都可能是无穷大，若然则两边同为无穷大）。不妨把各项想像成排成一个三角形，第 m 行有 m 个 $\mathrm{P}(X=m)$：逐行相加得 $\mathrm{E}(X)$，逐列相加得尾和。公式得证。

尾概率

按题目原意，假设不同盒子的内容互相独立（independent）。设 D_k 为第 k 个盒内是企鹅爸爸的事件，则 $\mathrm{P}(D_k)=p$，$\mathrm{P}(D_k')=q$。

取 $n \ge 2$。$X \ge n$ 成立，当且仅当首 $n-1$ 个盒尚未凑齐两种企鹅，即它们全是爸爸或全是妈妈。这两种情况互不相交，故由独立性得

$$\mathrm{P}(X \ge n) = p^{n-1} + q^{n-1} \qquad (n \ge 2).$$

特别地，取 $n=4$，

$$\mathrm{P}(X \ge 4) = p^3 + q^3,$$

正是所求。此外，$X \ge 2$ 必然成立，因为一个盒不可能同时装有两种企鹅，所以 $\mathrm{P}(X \ge 1) = \mathrm{P}(X \ge 2) = 1$。（上面的公式在 $n=2$ 时给出 $p+q=1$，与此相符；但在 $n=1$ 时却给出 2，所以 $n=1$ 这一项必须另行处理。）

期望值

由尾和公式，

$$\begin{aligned}\mathrm{E}(X) &= \mathrm{P}(X \ge 1) + \sum_{n=2}^{\infty} \left(p^{n-1} + q^{n-1}\right) \\ &= 1 + \sum_{k=1}^{\infty} p^k + \sum_{k=1}^{\infty} q^k \\ &= 1 + \frac{p}{1-p} + \frac{q}{1-q},\end{aligned}$$

其中两个等比级数都收敛，因为 $0<p<1$ 及 $0<q<1$（这里正是用到 $p \ne 0$、$q \ne 0$ 的地方）。由于 $1-p=q$ 及 $1-q=p$，

$$\mathrm{E}(X) = 1 + \frac{p}{q} + \frac{q}{p} = 1 + \frac{p^2+q^2}{pq}.$$

而 $p^2+q^2 = (p+q)^2 - 2pq = 1-2pq$，所以

$$\mathrm{E}(X) = 1 + \frac{1-2pq}{pq} = 1 + \frac{1}{pq} - 2 = \frac{1}{pq} - 1,$$

正是所求。

最后一部分

对正数 p 和 q 应用算术平均—几何平均不等式（AM–GM inequality），

$$\sqrt{pq} \le \frac{p+q}{2} = \frac{1}{2}, \qquad \text{所以} \qquad 0 < pq \le \frac{1}{4}.$$

想一想 取倒数之后，不等号朝哪个方向?

由于 $pq > 0$，取倒数会令不等号反向：$\dfrac{1}{pq} \ge 4$。因此

$$\mathrm{E}(X) = \frac{1}{pq} - 1 \ge 4 - 1 = 3,$$

等号成立当且仅当 $p = q = \frac{1}{2}$。（等价地，$pq = p(1-p) = \frac{1}{4} - (p - \frac{1}{2})^2 \le \frac{1}{4}$。）

验算

- 取 $p = q = \frac{1}{2}$：$X - 1$ 是等到第一个与第 1 盒不同的盒子所需的次数，是一个平均值为 2 的几何变量（geometric variable），所以 $\mathrm{E}(X) = 3$，与 $\frac{1}{1/4} - 1$ 吻合。
- 当 $p \to 0$ 时，$\mathrm{E}(X) \to \infty$：企鹅爸爸十分罕有，所以要等很久。这与 $\frac{1}{pq} - 1$ 趋向无穷大相符。
- 直接计算，对 $n \ge 2$ 有 $\mathrm{P}(X = n) = p^{n-1}q + q^{n-1}p$。取 $p = \frac{1}{3}$，以数值方法把 $n\,\mathrm{P}(X = n)$ 相加得 3.5；另外模拟 200 000 名收集者（取 $p = 0.3$），所得平均值为 3.761，而准确值是 $\frac{1}{0.21} - 1 \approx 3.762$。

常见错误

- 在 $n = 1$ 时也使用 $p^{n-1} + q^{n-1}$，结果得出一个等于 2 的“概率”，期望值亦因而多了 1。
- 写成 $\mathrm{P}(X \ge 4) = p^4 + q^4$：只要首三个盒相同，就需要开第四个盒。
- 在最后一部分忘记了正数之间的不等式取倒数后会反向。

STEP 之外

尾和公式有一个连续版本的孪生公式：对于非负随机变量（random variable），$\mathrm{E}(X) = \int_0^\infty \mathrm{P}(X > x)\,\mathrm{d}x$,证明方法同样是对调次序(在一般情况下,这一步由托内利定理(Tonelli's

theorem）保证成立）。企鹅问题是赠券收集问题（coupon collector's problem）中只有两种类型的情况；若有 N 种机会均等的类型，所需盒数的期望值是 $N(1+\frac{1}{2}+\cdots+\frac{1}{N})$，约为 $N\ln N$。这两个想法都见于 Grimmett 与 Welsh 的 *Probability: An Introduction*（概率入门），以及 Blitzstein 与 Hwang 的 *Introduction to Probability*（概率导论）。

详细解答

STEP I 2010，第 13 题

概率与统计 · 泊松过程的等候时间

STEP I 2010，第 13 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s1-q13

关键思路

“第一则短讯在第 1 至第 2 小时之间收到”等同于“第一个小时内没有短讯，但并非头两个小时内都没有短讯”。两者各自都是泊松分布（Poisson distribution）中零个事件的概率，也就是单独一个指数式。之后代入 $x = \mathrm{e}^{\lambda}$，一切便化为一条二次方程，余下的工作交给韦达定理（Vieta's formulas，即根之和与根之积）便可完成，Mildred 的部分也是如此。

解答

我们采用惯常的泊松模型：George 在任何长度为 t 小时的时段内收到的短讯数目服从平均值为 λt 的泊松分布，而互不重叠的时段内的数目互相独立（independent）。时间由早上开始时起计。由于时间是连续的，第一则短讯刚好在第 1 小时或刚好在第 2 小时收到的概率为零，所以端点包括与否并无影响。

George 的方程

设 T 为 George 收到第一则短讯前所经过的时间（以小时计）。那么 $T > t$ 当且仅当 $[0, t]$ 内没有短讯，所以

$$\mathrm{P}(T > t) = \mathrm{P}(\text{Poisson}(\lambda t) = 0) = \mathrm{e}^{-\lambda t}.$$

事件 $\{1 < T \leq 2\}$ 就是从 $\{T > 1\}$ 中剔除 $\{T > 2\}$，而 $\{T > 2\} \subseteq \{T > 1\}$，所以

$$p = \mathrm{P}(1 < T \leq 2) = \mathrm{P}(T > 1) - \mathrm{P}(T > 2) = \mathrm{e}^{-\lambda} - \mathrm{e}^{-2\lambda}.$$

全式乘以 $\mathrm{e}^{2\lambda}$，得 $p\mathrm{e}^{2\lambda} = \mathrm{e}^{\lambda} - 1$，即

$$p\mathrm{e}^{2\lambda} - \mathrm{e}^{\lambda} + 1 = 0,$$

正是所求。

λ 的两个正值

设 $x = \mathrm{e}^{\lambda}$。由于 $\lambda \mapsto \mathrm{e}^{\lambda}$ 是由 $(0,\infty)$ 到 $(1,\infty)$ 的严格递增双射（bijection），λ 的正数解恰好对应于下式中满足 $x > 1$ 的解：

$$px^2 - x + 1 = 0.$$

这里 $p > 0$（它是某事件的概率，而该事件在例如 $[1,2]$ 内刚好收到一则短讯、之前一则也没有的情况下必定发生），所以这确实是一条二次方程。

> ? **想一想** 要证明两根为实数并不困难。但我们怎样知道两根都大于 1?

*相异实根。*判别式（discriminant）为 $1 - 4p > 0$，因为 $4p < 1$，所以方程有两个相异实根 $x_1 < x_2$。

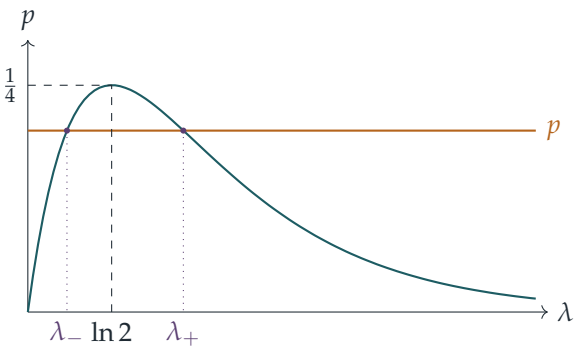
*两根均大于 1。*根据韦达定理，$x_1 + x_2 = \frac{1}{p}$ 及 $x_1 x_2 = \frac{1}{p}$。于是

$$(x_1 - 1)(x_2 - 1) = x_1 x_2 - (x_1 + x_2) + 1 = \tfrac{1}{p} - \tfrac{1}{p} + 1 = 1 > 0,$$

所以 $x_1 - 1$ 与 $x_2 - 1$ 同号。两者之和为 $x_1 + x_2 - 2 = \frac{1}{p} - 2 > 0$，因为由 $p < \frac{1}{4}$ 可得 $\frac{1}{p} > 4$。因此两者皆为正数：$x_1 > 1$ 及 $x_2 > 1$。

故此 $\lambda_{\pm} = \ln\left(\dfrac{1 \pm \sqrt{1-4p}}{2p}\right)$ 是满足该方程的两个相异的 λ 正值。

看看图像便明白个中道理。把 $p = \mathrm{e}^{-\lambda} - \mathrm{e}^{-2\lambda}$ 视为 λ 的函数，它在 $\lambda = 0$ 时为 0，当 $\lambda \to \infty$ 时趋向 0，并且（设 $u = \mathrm{e}^{-\lambda}$，则 $p = u - u^2$）在 $u = \frac{1}{2}$，即 $\lambda = \ln 2$ 时取得极大值 $\frac{1}{4}$。高度为 $p < \frac{1}{4}$ 的水平线与曲线相交两次：一部冷清的电话（根本很少收到短讯）和一部繁忙的电话（短讯通常在第一个小时内便收到）可以给出相同的 p。



Mildred：$\lambda_1+\lambda_2$

每部电话单独来看，都与 George 的情况完全相同，所以 e^{λ_1} 和 e^{λ_2} 都满足 $px^2-x+1=0$。这条二次方程只有两个根，而 $\lambda_1\neq\lambda_2$，所以 e^{λ_1} 和 e^{λ_2} 就是它的两个根 x_1 和 x_2。（这需要 $p<\frac{1}{4}$，而此条件成立，因为存在两个相异的值：当 $p=\frac{1}{4}$ 时方程有重根。）由根之积，

$$e^{\lambda_1+\lambda_2}=e^{\lambda_1}e^{\lambda_2}=x_1x_2=\frac{1}{p},\qquad \text{所以}\qquad \lambda_1+\lambda_2=\ln\frac{1}{p}=-\ln p.$$

Mildred 的第一则短讯

> **想一想** Mildred 两部电话合共收到的短讯总数服从甚么分布？

两个数目是独立的泊松变量，所以它们之和服从泊松分布，其平均值为两个平均值之和：在 t 小时内，她合共收到 $\text{Poisson}\big((\lambda_1+\lambda_2)t\big)$ 则短讯。换一个说法，设 T_1,T_2 为两部电话各自的等候时间；它们互相独立，而她的第一则短讯在 $T=\min(T_1,T_2)$ 时收到，且

$$\mathrm{P}(T>t)=\mathrm{P}(T_1>t)\,\mathrm{P}(T_2>t)=e^{-(\lambda_1+\lambda_2)t}.$$

无论用哪一种方法，设 $\mu=\lambda_1+\lambda_2$，则 $e^{-\mu}=p$，再沿用处理 George 时的同一论证，可得

$$\mathrm{P}(1<T\leq 2)=e^{-\mu}-e^{-2\mu}=p-p^2=p(1-p).$$

验算

- 取 $\lambda=0.8$，$p=e^{-0.8}-e^{-1.6}\approx 0.2474$；模拟 2×10^5 个指数分布（exponential distribution）的等候时间，所得结果为 0.2462。
- 取 $p=0.2$：两根为 $\lambda_1\approx 0.3235$ 和 $\lambda_2\approx 1.2859$，两者皆为正数，其和为 $1.6094=\ln 5=-\ln 0.2$。模拟 $\min(T_1,T_2)$ 所得的最后一个概率为 0.1601，而 $p-p^2=0.16$。
- 合理性检查：$p(1-p)<p$，理应如此，因为有两部电话时，第一则短讯往往较早收到，令“一小时内没有短讯”的可能性降低。

常见错误

- 证明了 x 的两根是正实数，却忘记 $\lambda>0$ 需要 $x>1$，而不仅是 $x>0$。
- 在最后一部分把两部电话的概率相加，或把它们相乘（p^2）。正确的事件是“任何一部电话收到的第一则短讯”，由合并后的速率决定。

STEP 之外

泊松过程（Poisson process）中第一个事件出现前的等候时间服从指数分布，而本题用到的合并性质（速率分别为 λ_1、λ_2 的独立泊松过程合并成一个速率为 $\lambda_1 + \lambda_2$ 的泊松过程）是泊松过程的基本性质之一，与之并列的还有“稀疏化”（thinning，即每个事件各自独立地以概率 q 保留）。这些想法是排队论（queueing theory）以及应用概率中许多内容的基础。Grimmett 与 Welsh 所著的 *Probability: An Introduction*（概率入门教材）对泊松过程的处理方式，正好自然地承接本题。

详细解答

STEP II 2010，第 1 题

纯数学 · 曲率与密切圆

STEP II 2010，第 1 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q1

关键思路

“相切”（touches）的意思是圆与曲线共有点 P，而且在该点的斜率（gradient）相同；“斜率的变化率”就是 $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$。因此我们要找一个圆，使它在 P 点的 y、$\frac{\mathrm{d}y}{\mathrm{d}x}$ 和 $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$ 全部与曲线的相应数值一致。把一般圆的方程作两次隐函数微分（implicit differentiation），问题便化为关于圆心的简单方程。

解答

曲线在 P 点的情况

对于 $y = 1 - x + \tan x$，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -1 + \sec^2 x = \tan^2 x, \qquad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2\tan x \sec^2 x.$$

当 $x = \frac{1}{4}\pi$ 时，$\tan x = 1$ 且 $\sec^2 x = 2$，所以

$$y = 2 - \tfrac{1}{4}\pi, \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = 1, \qquad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 4.$$

因此 $P = \left(\frac{1}{4}\pi,\ 2 - \frac{1}{4}\pi\right)$。

圆

设圆的方程为 $(x-a)^2 + (y-b)^2 = r^2$。对 x 作隐函数微分，得

$$(x-a) + (y-b)\frac{\mathrm{d}y}{\mathrm{d}x} = 0,$$

再微分一次（对第二项使用积法则（product rule）），得

$$1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+(y-b)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=0.$$

想一想 为甚么可以把曲线的数值直接代入圆的方程?

这两条方程在圆上每一点都成立，其中 y、$\mathrm{d}y/\mathrm{d}x$ 和 $\mathrm{d}^2y/\mathrm{d}x^2$ 是圆本身的数值。题目的两个条件是说：在 P 点，圆通过 P，而且斜率与 C 相同，二阶导数（second derivative）也相同。所以在 P 点我们可以取 $x=\frac{1}{4}\pi$、$y=2-\frac{1}{4}\pi$、$\mathrm{d}y/\mathrm{d}x=1$ 和 $\mathrm{d}^2y/\mathrm{d}x^2=4$。

由第二条方程得

$$1+1+4(y-b)=0 \quad\Longrightarrow\quad y-b=-\tfrac{1}{2} \quad\Longrightarrow\quad b=2-\tfrac{1}{4}\pi+\tfrac{1}{2}=\tfrac{5}{2}-\tfrac{1}{4}\pi.$$

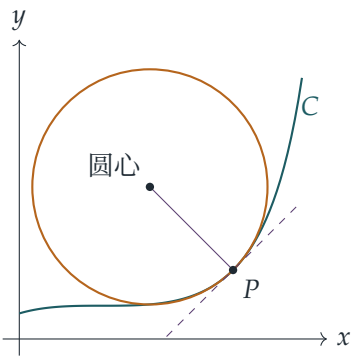
再由第一条方程得

$$x-a=-(y-b)\cdot 1=\tfrac{1}{2} \quad\Longrightarrow\quad a=\tfrac{1}{4}\pi-\tfrac{1}{2}.$$

最后，$r^2=(x-a)^2+(y-b)^2=\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$。所以密切圆（osculating circle）的

$$\text{圆心为}\left(\tfrac{1}{4}\pi-\tfrac{1}{2},\ \tfrac{5}{2}-\tfrac{1}{4}\pi\right) \quad\text{，半径为}\quad \frac{1}{\sqrt{2}}.$$

从几何上看，这个结果是合理的。圆心与 P 相距 $\frac{1}{\sqrt{2}}$，位于方向 $(-1,1)$ 上，而这个方向与斜率为 1 的切线（tangent）垂直。圆心位于 $y>2-\frac{1}{4}\pi$ 的一侧，即曲线的上方；这是对的，因为 $\mathrm{d}^2y/\mathrm{d}x^2>0$，曲线在该处向上凹（concave upwards）。



验算

- 一般的曲率半径（radius of curvature）公式 $\rho=\left(1+y'^2\right)^{3/2}/|y''|$ 给出 $2^{3/2}/4=1/\sqrt{2}$，与上面求得的结果一致。

- 以数值计算,圆心约为 $(0.2854,\ 1.7146)$,P 约为 $(0.7854,\ 1.2146)$;两点的距离为 0.70711。取圆的下半部分 $y = b - \sqrt{r^2 - (x-a)^2}$,在 $x = 0.7854$ 处作数值微分,得斜率 1.0000 和二阶导数 4.000,与曲线吻合。

常见错误

- 把圆心放在切线的错误一侧:$y - b$ 的正负号由第二条方程决定,所以让代数运算来判断就可以了。
- 忘记对 $(y-b)\,\mathrm{d}y/\mathrm{d}x$ 微分时要用积法则,而 $(\mathrm{d}y/\mathrm{d}x)^2$ 这一项正是由此而来。

STEP 之外

密切圆的半径就是曲率半径,它的倒数 $\kappa = |y''|/(1+y'^2)^{3/2}$ 就是曲率(curvature),是曲线微分几何(differential geometry)中的核心量。对于空间曲线,曲率与挠率(torsion)一同出现在弗勒内–塞雷公式(Frenet–Serret formulae)中,这组公式描述曲线如何弯曲和扭转。Siklos 的 *Advanced Problems in Mathematics*(进阶数学难题集)收录了更多同类的 STEP 题目。

详细解答

STEP II 2010，第 2 题

纯数学 · 三角函数积分法

STEP II 2010，第 2 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q2

关键思路

三次式 $7\sin x - 8\sin^3 x$ 看来不易积分，但只要留意到它其实由三倍角公式（triple-angle formula）砌成，即 $\sin 3x$ 的两倍再加上少许余项，问题便迎刃而解。接着利用余弦的三倍角公式，把答案变回 $c = \cos\alpha$ 的多项式（polynomial）。至于 Eustace 的积分，利用 $\sin^2\alpha = 1 - c^2$ 可把他那个（错误的）答案化为同一个变量，再比较两者，便得到一个 c 的四次方程（quartic），而题目已就它的根给了我们提示。

解答

三倍角公式

利用复角公式（addition formulae）及 $\sin^2 x = 1 - \cos^2 x$，

$$\begin{aligned}\cos 3x = \cos(2x + x) &= \cos 2x \cos x - \sin 2x \sin x \\ &= (2\cos^2 x - 1)\cos x - 2\sin^2 x \cos x \\ &= 2\cos^3 x - \cos x - 2(1 - \cos^2 x)\cos x \\ &= 4\cos^3 x - 3\cos x,\end{aligned}$$

正如所求。同样地，利用 $\cos^2 x = 1 - \sin^2 x$，

$$\begin{aligned}\sin 3x = \sin(2x + x) &= \sin 2x \cos x + \cos 2x \sin x \\ &= 2\sin x \cos^2 x + (1 - 2\sin^2 x)\sin x \\ &= 2\sin x(1 - \sin^2 x) + \sin x - 2\sin^3 x \\ &= 3\sin x - 4\sin^3 x.\end{aligned}$$

第 (i) 部分

由正弦的公式可得 $8\sin^3 x = 6\sin x - 2\sin 3x$，所以

$$7\sin x - 8\sin^3 x = \sin x + 2\sin 3x.$$

这样积分便容易计算了：

$$\mathrm{I}(\alpha) = \Big[-\cos x - \tfrac{2}{3}\cos 3x\Big]_0^{\alpha} = -\cos\alpha - \tfrac{2}{3}\cos 3\alpha + 1 + \tfrac{2}{3}.$$

由余弦的公式，$\cos 3\alpha = 4c^3 - 3c$，所以

$$\mathrm{I}(\alpha) = -c - \tfrac{2}{3}(4c^3 - 3c) + \tfrac{5}{3} = -\tfrac{8}{3}c^3 + c + \tfrac{5}{3},$$

这正是我们要证明的结果。

至于使 $\mathrm{I}(\alpha) = 0$ 的 c 值：取 $\alpha = 0$，积分显然等于零，所以 $c = 1$ 符合要求（验算：$-\frac{8}{3} + 1 + \frac{5}{3} = 0$）。事实上这是唯一的值，因为 $8c^3 - 3c - 5 = (c-1)(8c^2 + 8c + 5)$，而当中二次式的判别式（discriminant）为 $64 - 160 < 0$。

第 (ii) 部分

按 Eustace 的法则，$\int \sin x\,\mathrm{d}x = \frac{1}{2}\sin^2 x$ 及 $\int \sin^3 x\,\mathrm{d}x = \frac{1}{4}\sin^4 x$，所以他算出的定积分（definite integral）之值为

$$\mathrm{E}(\alpha) = \Big[\tfrac{7}{2}\sin^2 x - 2\sin^4 x\Big]_0^{\alpha} = \tfrac{7}{2}s^2 - 2s^4, \qquad s = \sin\alpha.$$

为了与第 (i) 部分比较，写成 $s^2 = 1 - c^2$：

$$\mathrm{E}(\alpha) = \tfrac{7}{2}(1 - c^2) - 2(1 - c^2)^2 = \tfrac{3}{2} + \tfrac{1}{2}c^2 - 2c^4.$$

$\cos\beta = -\frac{1}{6}$ 这个值。 当 $c = -\frac{1}{6}$ 时，

$$\mathrm{I}(\beta) = -\tfrac{8}{3}\cdot\big(-\tfrac{1}{216}\big) - \tfrac{1}{6} + \tfrac{5}{3} = \tfrac{1}{81} + \tfrac{3}{2} = \tfrac{245}{162},$$
$$\mathrm{E}(\beta) = \tfrac{3}{2} + \tfrac{1}{2}\cdot\tfrac{1}{36} - 2\cdot\tfrac{1}{1296} = \tfrac{3}{2} + \tfrac{1}{72} - \tfrac{1}{648} = \tfrac{3}{2} + \tfrac{8}{648} = \tfrac{245}{162}.$$

两者相等，所以 Eustace 只是碰巧得到 $\mathrm{I}(\beta)$ 的正确值。

所有 α 的值。 Eustace 的答案正确，当且仅当 $\mathrm{E}(\alpha) = \mathrm{I}(\alpha)$，即（乘以 6 并整理各项后）

$$12c^4 - 16c^3 - 3c^2 + 6c + 1 = 0.$$

? **想一想** 没有现成公式可用，我们怎样把这样的四次式因式分解（factorise）？

我们已经知道两个根。由刚才的计算，$c=-\frac{1}{6}$ 符合方程，所以 $6c+1$ 是一个因式。另外 $c=1$（即 $\alpha=0$）也必定符合，因为 I 和 E 在该处都等于 0。除去因式 $6c+1$（比较 c^3 和 c 的系数），

$$12c^4-16c^3-3c^2+6c+1=(6c+1)(2c^3-3c^2+1),$$

然后，由于 $c=1$ 是这个三次式的根，

$$2c^3-3c^2+1=(c-1)(2c^2-c-1)=(c-1)^2(2c+1).$$

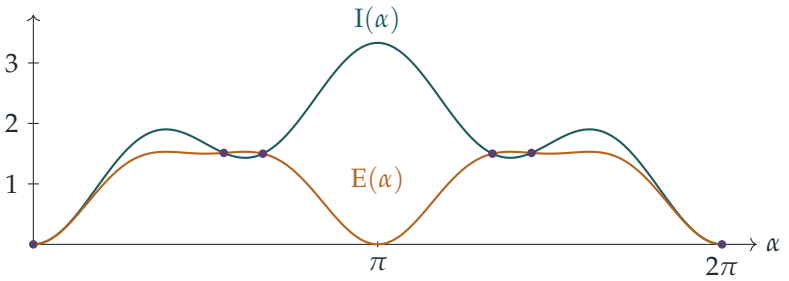
所以条件就是

$$(6c+1)(2c+1)(c-1)^2=0,\qquad \text{即}\qquad c=1,\ -\tfrac{1}{2}\ \text{或}\ -\tfrac{1}{6}.$$

由于 $c=\cos\alpha$，而 $[-1,1]$ 内每一个 c 值都对应某个 α，所以 Eustace 得到 $\mathrm{I}(\alpha)$ 的正确值，当且仅当对某个整数 n，

$$\alpha=2n\pi,\qquad \alpha=2n\pi\pm\tfrac{2\pi}{3},\qquad \text{或}\qquad \alpha=2n\pi\pm\arccos\left(-\tfrac{1}{6}\right).$$

（负的 α 也已包括在内：I 和 E 都只通过 $\cos\alpha$ 取决于 α。）下图显示在 $0\le\alpha\le2\pi$ 的范围内，真正的积分（蓝绿色）和 Eustace 的答案（琥珀色）；两条曲线相交于 $\alpha=0$、$\frac{2\pi}{3}$、$\arccos(-\frac{1}{6})\approx1.738$，以及这些点关于 $\alpha=\pi$ 的镜像。



验算

- 当 $\alpha=\pi$ ($c=-1$) 时，公式给出 $\mathrm{I}(\pi)=\frac{8}{3}-1+\frac{5}{3}=\frac{10}{3}$，与 $\int_0^\pi(\sin x+2\sin 3x)\,\mathrm{d}x=2+\frac{4}{3}$ 吻合。
- 对原来的被积函数 (integrand) 使用森逊法则 (Simpson's rule)，得 $\mathrm{I}(\beta)\approx1.5123457$，与 $\frac{245}{162}=1.5123456\ldots$ 吻合；当 $\alpha=\frac{2\pi}{3}$ 时，积分和 Eustace 的答案都等于 1.5。

常见错误

- 忘记代入下限：$-\cos0-\frac{2}{3}\cos0=-\frac{5}{3}$，那个 $+\frac{5}{3}$ 正是由此而来。

- 没有发现 $c = 1$ 是四次方程的二重根，甚或完全遗漏了它，只写出 $c = -\frac{1}{2}$ 和 $c = -\frac{1}{6}$。
- 只给出 $[0, \pi]$ 内的 α；题目要求的是所有值。

STEP 之外

把 $\sin^3 x$ 这类幂写成 $\sin x$ 与 $\sin 3x$ 的组合，正是傅里叶分析（Fourier analysis）的起点：每一个幂 $\sin^n x$ 都是有限个 $\sin kx$ 或 $\cos kx$ 之和，利用 $\sin x = (e^{ix} - e^{-ix})/2i$ 便可很快求出。以 $\cos x$ 表示 $\cos nx$ 的多项式，就是切比雪夫多项式（Chebyshev polynomial）T_n，其中 $T_3(c) = 4c^3 - 3c$。Needham 的 *Visual Complex Analysis*（《复分析：可视化方法》）说明了为何有了棣美弗定理（de Moivre's theorem），这些恒等式全都变得自然而然。

详细解答

STEP II 2010，第 3 题

纯数学 · 斐波那契数与递推关系

STEP II 2010，第 3 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q3

关键思路

已知四项，未知数也有四个：把 $n = 0, 1, 2, 3$ 代入 $(*)$ 再解方程。窍门是用一条方程除以另一条，这样可以消去 a，并把条件化为关于 $\lambda + \mu$ 和 $\lambda\mu$ 的陈述。这两个对称函数（symmetric functions）使 λ 和 μ 成为某个二次方程（quadratic equation）的根。求出公式之后，第 (iii) 部分的和就可以拆成两个几何级数（geometric series）。

解答

第 (i) 部分

把 $n = 0, 1, 2, 3$ 代入 $(*)$，得

$$a + b = 0, \qquad a\lambda + b\mu = 1,$$
$$a\lambda^2 + b\mu^2 = 1, \qquad a\lambda^3 + b\mu^3 = 2.$$

由第一条得 $b = -a$，于是其余三条变为

$$a(\lambda - \mu) = 1, \qquad a(\lambda^2 - \mu^2) = 1, \qquad a(\lambda^3 - \mu^3) = 2.$$

特别地，由 $a(\lambda - \mu) = 1$ 可知 $a \neq 0$ 且 $\lambda \neq \mu$。因式分解，

$$\lambda^2 - \mu^2 = (\lambda - \mu)(\lambda + \mu), \qquad \lambda^3 - \mu^3 = (\lambda - \mu)(\lambda^2 + \lambda\mu + \mu^2),$$

所以把第二及第三条方程除以 $a(\lambda - \mu) = 1$（因为它不为零，所以可以这样做），得

$$\lambda + \mu = 1 \qquad \text{和} \qquad \lambda^2 + \lambda\mu + \mu^2 = 2,$$

其中第二条正是所求的结果。

想一想 怎样从这两条方程求出 λ 和 μ 本身?

把第二条用和与积表示：$\lambda^2+\lambda\mu+\mu^2=(\lambda+\mu)^2-\lambda\mu$，所以 $1-\lambda\mu=2$，即 $\lambda\mu=-1$。和为 1、积为 -1 的两个数是以下方程的根：

$$t^2-t-1=0,\qquad \text{所以}\qquad t=\frac{1\pm\sqrt{5}}{2}.$$

哪一个根是 λ? 由 $a(\lambda-\mu)=1$ 及 $a>0$，需要 $\lambda-\mu>0$，所以 λ 是较大的根。于是 $\lambda-\mu=\sqrt{5}$，而

$$\lambda=\frac{1+\sqrt{5}}{2},\qquad \mu=\frac{1-\sqrt{5}}{2},\qquad a=\frac{1}{\sqrt{5}},\qquad b=-\frac{1}{\sqrt{5}}.$$

（若把两根对调，则必须 $a=-1/\sqrt{5}$，这是不容许的。）反过来，由于上面每一步都可逆，这些值满足全部四条方程。

第 (ii) 部分

我们有 $F_6=(\lambda^6-\mu^6)/\sqrt{5}$。与其直接展开 $\left(\frac{1+\sqrt{5}}{2}\right)^6$，不如利用 λ 是 $t^2=t+1$ 的根这一事实来降次：

$$\begin{aligned}\lambda^2&=\lambda+1,\\ \lambda^3&=\lambda^2+\lambda=2\lambda+1,\\ \lambda^6&=(2\lambda+1)^2=4\lambda^2+4\lambda+1=4(\lambda+1)+4\lambda+1=8\lambda+5.\end{aligned}$$

μ 满足同一条二次方程，所以完全同样的推导亦成立，得 $\mu^6=8\mu+5$。因此

$$F_6=\frac{(8\lambda+5)-(8\mu+5)}{\sqrt{5}}=\frac{8(\lambda-\mu)}{\sqrt{5}}=\frac{8\sqrt{5}}{\sqrt{5}}=8.$$

第 (iii) 部分

利用 $(*)$，

$$\sum_{n=0}^{\infty}\frac{F_n}{2^{n+1}}=\frac{1}{2\sqrt{5}}\sum_{n=0}^{\infty}\left[\left(\frac{\lambda}{2}\right)^n-\left(\frac{\mu}{2}\right)^n\right].$$

由于 $\lambda\approx 1.618$ 而 $\mu\approx -0.618$，两个公比（common ratio）$\lambda/2$ 和 $\mu/2$ 的模（modulus）都小于 1。所以两个几何级数都收敛（converge），我们可以把和拆成两者之差：

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{F_n}{2^{n+1}}&=\frac{1}{2\sqrt{5}}\left[\frac{1}{1-\lambda/2}-\frac{1}{1-\mu/2}\right]=\frac{1}{\sqrt{5}}\left[\frac{1}{2-\lambda}-\frac{1}{2-\mu}\right]\\ &=\frac{1}{\sqrt{5}}\cdot\frac{\lambda-\mu}{(2-\lambda)(2-\mu)}=\frac{1}{4-2(\lambda+\mu)+\lambda\mu},\end{aligned}$$

这里用了 $\lambda-\mu=\sqrt{5}$。由 $\lambda+\mu=1$ 及 $\lambda\mu=-1$，分母为 $4-2-1=1$，所以

$$\sum_{n=0}^{\infty}\frac{F_n}{2^{n+1}}=1.$$

验算

公式应该给出整数：$F_4=(\lambda^4-\mu^4)/\sqrt{5}$，配合 $\lambda^4=3\lambda+2$ 得 $F_4=3$；而 $F_5=5$、$F_6=8$ 符合从前四项看到的规律 $F_{n+2}=F_{n+1}+F_n$。至于第 (iii) 部分，部分和（partial sums）$0,\ \frac{1}{4},\ \frac{3}{8},\ \frac{1}{2},\ \frac{19}{32},\ldots$ 慢慢逼近 1；用数值方法把首 300 项相加，在机器精度内得 1.0000000。各项只以约 $(0.809)^n$ 的速度衰减，所以收敛得很慢。

常见错误

- 忘记利用 $a>0$ 来决定哪个根是 λ：没有这个条件，答案便不能确定。
- 除以 $\lambda-\mu$ 时没有说明它为何不为零（它不为零，因为 $a(\lambda-\mu)=1$）。
- 把无穷级数拆成两个级数时，没有检查每个级数是否收敛。

STEP 之外

这里的公式是斐波那契数（Fibonacci numbers）的比奈公式（Binet's formula），而所用的方法是处理线性递推关系（linear recurrences）的一般方法：λ 和 μ 是 $F_{n+2}=F_{n+1}+F_n$ 的特征方程（characteristic equation）$t^2=t+1$ 的根。第 (iii) 部分其实是在 $x=\frac{1}{2}$ 处求生成函数（generating function）$\sum F_n x^n=x/(1-x-x^2)$ 的值（再取一半），结果同样是 1。生成函数是大学一年级组合数学（combinatorics）和概率的重要工具；关于黄金比例（golden ratio）λ 以及斐波那契数在连分数（continued fractions）理论中的角色，可参阅 Davenport 的 *The Higher Arithmetic*（《高等算术》）。

详细解答

STEP II 2010，第 4 题

纯数学 · 代换积分法

STEP II 2010，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q4

关键思路

代换（substitution）$u = a - x$ 会把区间的两端对调，并把 $f(x)$ 变成 $f(a-x)$，反之亦然。因此这个积分等于它的“拍档”，后者的分子正是分母的另一半。把两个版本相加，分子便等于分母，被积函数（integrand）就变成 1。题目余下的部分，就是要在伪装之下认出这个结构。第 (ii) 部分用的是同一招，只是代换改为 $x \mapsto 1/x$，它会把 $[\frac{1}{2}, 2]$ 的两端对调。

解答

第 (i) 部分

我们全程假设在 $[0, a]$ 上 $f(x) + f(a-x) \neq 0$，使被积函数有定义。在 I 中代入 $u = a - x$，于是 $\mathrm{d}x = -\mathrm{d}u$。当 $x = 0$ 时，$u = a$；当 $x = a$ 时，$u = 0$。因此

$$I = \int_a^0 \frac{f(a-u)}{f(a-u)+f(u)}\,(-\mathrm{d}u) = \int_0^a \frac{f(a-u)}{f(u)+f(a-u)}\,\mathrm{d}u.$$

由于 u 是哑变量（dummy variable），我们可以把它改名为 x，这样便得到所需的 I 的第二种形式。

现在把 I 的两个表达式相加：

$$2I = \int_0^a \frac{f(x)+f(a-x)}{f(x)+f(a-x)}\,\mathrm{d}x = \int_0^a 1\,\mathrm{d}x = a, \qquad \text{所以} \qquad I = \frac{a}{2}.$$

第一个积分。

? **想一想** $f(x) + f(a-x)$ 藏在 $\ln(2 + x - x^2)$ 的哪里？

这个二次式可以因式分解：$2+x-x^2=(1+x)(2-x)$。所以

$$\ln(2+x-x^2)=\ln(1+x)+\ln(2-x),$$

这是成立的，因为两个因式在 $[0,1]$ 上都是正数。取 $a=1$ 及 $f(x)=\ln(1+x)$。那么 $f(1-x)=\ln(2-x)$，而被积函数正好是 $\dfrac{f(x)}{f(x)+f(1-x)}$。分母永不为零，因为在 $[0,1]$ 上 $2+x-x^2\geq 2$。因此

$$\int_0^1 \frac{\ln(x+1)}{\ln(2+x-x^2)}\,\mathrm{d}x=\frac{1}{2}.$$

第二个积分。 用复角公式（compound-angle formula）展开分母：

$$\sin\left(x+\tfrac{\pi}{4}\right)=\sin x\cos\tfrac{\pi}{4}+\cos x\sin\tfrac{\pi}{4}=\frac{\sin x+\cos x}{\sqrt{2}}.$$

所以被积函数是 $\sqrt{2}\,\dfrac{\sin x}{\sin x+\cos x}$。取 $a=\frac{\pi}{2}$ 及 $f=\sin$，便有 $f(\frac{\pi}{2}-x)=\cos x$。分母 $\sin x+\cos x$ 在 $[0,\frac{\pi}{2}]$ 上是正数。由上面的结果，

$$\int_0^{\pi/2}\frac{\sin x}{\sin(x+\frac{\pi}{4})}\,\mathrm{d}x=\sqrt{2}\cdot\frac{\pi/2}{2}=\frac{\sqrt{2}\,\pi}{4}.$$

第 (ii) 部分

把这个积分记为 J。它的被积函数是 $\dfrac{1}{x}\cdot\dfrac{\sin x}{\sin x+\sin(1/x)}$。分母没有问题：对于 $x\in[\frac{1}{2},2]$，x 和 $1/x$ 都落在 $[\frac{1}{2},2]\subset(0,\pi)$ 之内，所以 $\sin x>0$ 且 $\sin(1/x)>0$。

> **想一想** 哪一个代换可以把 $\sin x$ 和 $\sin\frac{1}{x}$ 对调，同时保持上下限 $\frac{1}{2}$ 和 2?

答案是 $x=1/u$。这时 $\mathrm{d}x=-\mathrm{d}u/u^2$。当 $x=\frac{1}{2}$ 时，$u=2$；当 $x=2$ 时，$u=\frac{1}{2}$。因子 $\frac{1}{x}$ 变成 u，所以

$$J=\int_2^{1/2}u\cdot\frac{\sin(1/u)}{\sin(1/u)+\sin u}\cdot\left(-\frac{\mathrm{d}u}{u^2}\right)=\int_{1/2}^{2}\frac{\sin(1/u)}{u\left(\sin u+\sin\frac{1}{u}\right)}\,\mathrm{d}u.$$

把 u 改名为 x，再与 J 原来的形式相加：

$$2J=\int_{1/2}^{2}\frac{\sin x+\sin\frac{1}{x}}{x\left(\sin x+\sin\frac{1}{x}\right)}\,\mathrm{d}x=\int_{1/2}^{2}\frac{1}{x}\,\mathrm{d}x=\ln 2-\ln\tfrac{1}{2}=2\ln 2.$$

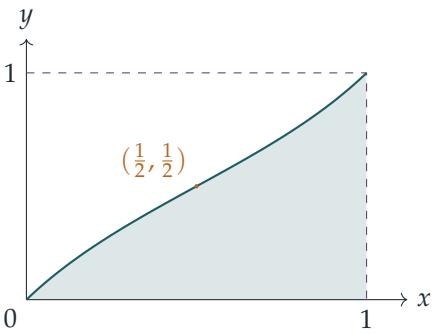
因此

$$\int_{1/2}^{2}\frac{\sin x}{x\left(\sin x+\sin\frac{1}{x}\right)}\,\mathrm{d}x=\ln 2.$$

因子 $\frac{1}{x}$ 在这里是不可或缺的。测度（measure）$\frac{\mathrm{d}x}{x}$ 在 $x\mapsto 1/x$ 之下保持不变（方向除外），正如 $\mathrm{d}x$ 在 $x\mapsto a-x$ 之下保持不变一样。

验算

- 被积函数 $g(x) = \frac{f(x)}{f(x)+f(a-x)}$ 满足 $g(x) + g(a-x) = 1$。因此它的图像绕 $(\frac{a}{2}, \frac{1}{2})$ 旋转半圈后不变，所以曲线下的面积是 $a \times 1$ 长方形面积的一半。下图以第一个积分为例说明这一点。



- 数值上，用辛普森法则（Simpson's rule）计算，第一个积分得 0.500000，第二个得 1.110721（而 $\frac{\sqrt{2}\pi}{4} \approx 1.110721$），第 (ii) 部分得 0.693147（而 $\ln 2 \approx 0.693147$）。

常见错误

- 改写 $\sin(x + \frac{\pi}{4})$ 时漏掉因子 $\sqrt{2}$，结果得出 $\frac{\pi}{4}$ 而不是 $\frac{\sqrt{2}\pi}{4}$。
- 在第 (ii) 部分，忘了跟进 $\frac{1}{x}$ 的去向：它变成 u，再与 $-\frac{1}{u^2}$ 合并，剩下 $-\frac{1}{u}$。

STEP 之外

这两招都利用了区间的一种对称性，而这种对称性令积分的“测度”保持不变：反射（reflection）$x \mapsto a - x$ 保持 $\mathrm{d}x$ 不变，而反演（inversion）$x \mapsto 1/x$ 保持 $\frac{\mathrm{d}x}{x}$ 不变。到了大学，这会发展成不变测度（invariant measure）的概念。例如，$\frac{\mathrm{d}x}{x}$ 是正实数在乘法之下的自然测度，也是梅林变换（Mellin transform）的基础。关于代换法在实分析（real analysis）上的基础，可参阅 Abbott 的 *Understanding Analysis*。想看更多 STEP 风格的积分技巧，可参阅 Siklos 的 *Advanced Problems in Mathematics*。

详细解答

STEP II 2010，第 5 题

纯数学 · 向量夹角与圆锥

> **STEP II 2010，第 5 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s2-q5

关键思路

这一题的一切都归结为一个事实：方向为 $\mathbf{d}$ 的直线与向量（vector）$\mathbf{a}$ 之间的夹角 θ 满足 $\cos\theta = \dfrac{\mathbf{d}\cdot\mathbf{a}}{|\mathbf{d}|\,|\mathbf{a}|}$。“倾角相等”和“倾角为 α”都可以直接化为关于 $\mathbf{d}$ 各分量（components）的方程。至于最后一部分，留意那个笛卡儿方程（Cartesian equation）与你刚刚求得的关系式形状相同。通过 O 而与某条固定直线成相同夹角的所有直线，合起来就是一个圆锥（cone）。

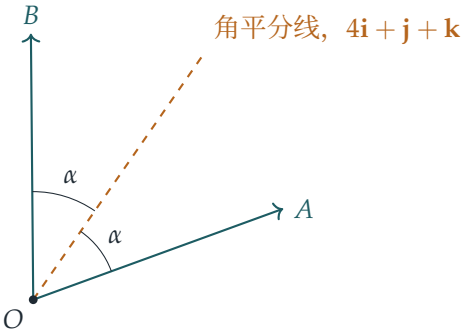
解答

设 $\mathbf{a} = \mathbf{i}+\mathbf{j}+\mathbf{k}$ 及 $\mathbf{b} = 5\mathbf{i}-\mathbf{j}-\mathbf{k}$，则 $|\mathbf{a}| = \sqrt{3}$ 而 $|\mathbf{b}| = \sqrt{27} = 3\sqrt{3}$。另外 $\mathbf{a}\cdot\mathbf{b} = 5-1-1 = 3$。因此

$$\cos 2\alpha = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{3}{\sqrt{3}\cdot 3\sqrt{3}} = \frac{1}{3}.$$

由于 $\cos 2\alpha > 0$，角 2α 是锐角，所以 α 介乎 0 与 $\frac{\pi}{4}$ 之间。稍后我们会用到

$$\cos^2\alpha = \tfrac{1}{2}(1+\cos 2\alpha) = \tfrac{2}{3}.$$



（此图只是平面 OAB 的示意图。）

第 (i) 部分

设 $\mathbf{d} = m\mathbf{i} + n\mathbf{j} + p\mathbf{k}$ 为 L_1 的方向。$\mathbf{d}$ 与 OA 及 OB 所成夹角的余弦分别是

$$\frac{\mathbf{d}\cdot\mathbf{a}}{|\mathbf{d}|\sqrt{3}} = \frac{m+n+p}{\sqrt{3}\,|\mathbf{d}|} \qquad \text{及} \qquad \frac{\mathbf{d}\cdot\mathbf{b}}{|\mathbf{d}|\,3\sqrt{3}} = \frac{5m-n-p}{3\sqrt{3}\,|\mathbf{d}|}.$$

由于 cos 在 $[0,\pi]$ 上是一对一（one-to-one）的，夹角相等即余弦相等。两边同乘 $3\sqrt{3}\,|\mathbf{d}|$，得

$$3(m+n+p) = 5m-n-p \quad \Longleftrightarrow \quad m = 2(n+p).$$

> **想一想** 直线本身并没有指定的方向。假如直线与 OA 成角 θ，但与 OB 成角 $\pi-\theta$，那又如何?

如果我们把“两条直线之间的夹角”理解为锐角的那一个，那么两个余弦只需在相差一个正负号的意义下相等。取另一个符号，得 $3(m+n+p) = -(5m-n-p)$，即 $4m+n+p = 0$。这些直线位于通过 O 且垂直于角平分线的平面上，而该平面包含 $\angle AOB$ 的外角平分线（external bisector）。题目预期的答案是 $m = 2(n+p)$，但知道还有第二族直线存在，也是值得的。

要求 $\angle AOB$ 的角平分线（bisector），我们把沿 OA 和 OB 方向的单位向量（unit vector）相加。两个单位向量之和平分它们之间的夹角，因为它们张成一个菱形（rhombus）。所以

$$\frac{\mathbf{a}}{\sqrt{3}} + \frac{\mathbf{b}}{3\sqrt{3}} = \frac{3\mathbf{a}+\mathbf{b}}{3\sqrt{3}} = \frac{8\mathbf{i}+2\mathbf{j}+2\mathbf{k}}{3\sqrt{3}},$$

于是可取 $m = 4$、$n = 1$、$p = 1$（或其任何非零倍数）。验算一下：$4 = 2(1+1)$，所以这个方向满足上面的关系式。

第 (ii) 部分

设 $\mathbf{e} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$ 为 L_2 的方向。这条直线与 OA 成角 α，当且仅当

$$\frac{u+v+w}{\sqrt{3}\sqrt{u^2+v^2+w^2}} = \pm\cos\alpha.$$

这里的 $\pm$ 顾及了沿直线的两个指向。平方后正负号消失，再利用 $\cos^2\alpha = \frac{2}{3}$，得

$$(u+v+w)^2 = 2(u^2+v^2+w^2).$$

把左边展开为 $u^2+v^2+w^2+2(vw+wu+uv)$ 并整理，便得到关系式

$$u^2+v^2+w^2 = 2(vw+wu+uv).$$

> ? **想一想** 一个关于方向的条件，怎样变成一个曲面？

任取一点 $(x,y,z) \neq (0,0,0)$。它恰好位于一条通过 O 的直线上，即方向为 $x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ 的那条直线。根据第 (ii) 部分，该点满足

$$x^2+y^2+z^2 = 2(yz+zx+xy)$$

当且仅当它所在的那条通过 O 的直线与 OA 成角 α。原点也满足这个方程。因此，这个曲面（surface）就是所有通过 O 且与直线 OA 成倾角 α 的直线的并集。

这是一个**（双叶）直立圆锥（right circular cone）**，顶点在原点，轴为直线 $x=y=z$（即直线 OA）。它的半顶角（semi-vertical angle）α 满足

$$\cos\alpha = \sqrt{\tfrac{2}{3}}, \qquad \tan\alpha = \tfrac{1}{\sqrt{2}}, \qquad \alpha \approx 35.3^\circ,$$

即 $\angle AOB$ 的一半。它有两叶（nappes），因为 μ 取遍所有实数。第 (i) 部分的角平分线位于这个圆锥上：对于 $(4,1,1)$，我们有 $16+1+1 = 18 = 2(1+4+4)$。

验算

- 点 $(1,1,0)$ 满足 $2 = 2(0+0+1)$。它与 $\mathbf{i}+\mathbf{j}+\mathbf{k}$ 的夹角的余弦为 $2/(\sqrt{2}\sqrt{3}) = \sqrt{2/3}$，与预期相符。
- 数值上，$\arccos(1/3)/2 = 0.61548\,\text{rad}$，而角平分线 $4\mathbf{i}+\mathbf{j}+\mathbf{k}$ 与 OA 和 OB 所成的角都是 $0.61548\,\text{rad}$。
- 恒等式 $x^2+y^2+z^2-2(yz+zx+xy) = (x^2+y^2+z^2)(2-3c^2)$ 成立，其中 c 是与 $\mathbf{i}+\mathbf{j}+\mathbf{k}$ 所成夹角的余弦。所以左边等于零，当且仅当 $c^2 = \frac{2}{3}$。

常见错误

- 用 $\mathbf{a}+\mathbf{b}$ 来求角平分线，而不是用单位向量之和。由于 $|\mathbf{b}| \neq |\mathbf{a}|$，$\mathbf{a}+\mathbf{b} = 6\mathbf{i}$ 并不是角平分线。
- 使用 $\cos\alpha = \frac{1}{2}\cos 2\alpha$ 或类似的式子。这里需要的是倍角公式（double-angle formula）$\cos^2\alpha = \frac{1}{2}(1+\cos 2\alpha)$。
- 只把曲面描述为“一个圆锥”，而没有指出它的顶点、轴和半顶角。

STEP 之外

这个圆锥的方程是一个二次型（quadratic form）：$\mathbf{x}^{\mathsf{T}} M\mathbf{x} = 0$，其中 M 是对称矩阵（symmetric matrix）。在大学一年级的线性代数（linear algebra）中，你会按 M 的特征值（eigenvalues）为这类曲面分类。这里 M 的对角线元素为 1，非对角线元素为 -1，所以它的特征值是 -1（沿 $\mathbf{i}+\mathbf{j}+\mathbf{k}$ 方向）和 2（二重）。一个负特征值加上两个相等的正特征值，正是圆锥的标志：它是一个圆形的圆锥，轴沿着那个与别不同的特征值所对应的特征向量（eigenvector）。Strang 的 *Introduction to Linear Algebra*（线性代数入门）在讨论对称矩阵和二次型时，对此有详细的阐述。

详细解答

STEP II 2010，第 6 题

纯数学 · 正四面体的几何

STEP II 2010，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

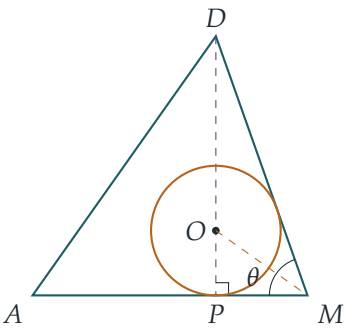
step.maths.org/questions/10-s2-q6

关键思路

正四面体（regular tetrahedron）的对称性极高，几乎所有事情都发生在同一个铅垂截面（cross-section）上：通过一条棱的中点、其对顶点和顶尖的那个平面。在这个平面内，每个立体问题都化为一个直角三角形。至于球，对称性告诉我们球心必定在哪里，而第 (ii) 部分求得的角则告诉我们球心与每个面相距多远。

解答

以下一律设 M 为 BC 的中点。每个面都是边长为 1 的等边三角形，所以每条中线（median）（例如 AM 或 DM）的长度为 $\sqrt{1-\frac{1}{4}}=\frac{\sqrt{3}}{2}$。



第 (i) 部分

由于 D 与 A、B、C 等距，而 DP 垂直于平面 ABC，由毕氏定理（Pythagoras' theorem）得 $PA^2=DA^2-DP^2=DB^2-DP^2=PB^2$，同理 $PB=PC$。所以 P 是等边三角形 ABC

的外心（circumcentre），亦即它的形心（centroid）。形心位于每条中线上由顶点起计三分之二的位置，所以

$$PA = \tfrac{2}{3}\cdot\tfrac{\sqrt{3}}{2} = \tfrac{1}{\sqrt{3}}, \qquad PM = \tfrac{1}{3}\cdot\tfrac{\sqrt{3}}{2} = \tfrac{1}{2\sqrt{3}}.$$

然后，由三角形 DPA 中位于 P 的直角，得

$$PD^2 = DA^2 - PA^2 = 1 - \tfrac{1}{3} = \tfrac{2}{3}, \qquad \text{所以} \qquad PD = \sqrt{\tfrac{2}{3}} = \frac{\sqrt{6}}{3}.$$

第 (ii) 部分

六条棱的情况完全相同（四面体的对称变换可以把任何一条棱移到任何另一条棱上），所以只需求出面 ABC 与面 DBC 之间的夹角 θ，这两个面相交于 BC。

? **想一想** 图中哪一个角才*真正*是这两个面之间的夹角?

两个平面相交于一条直线时，它们的夹角要用两条直线来量度：两条直线各在一个平面内，并且在同一点上垂直于交线。在等边三角形 ABC 中，中线 AM 垂直于 BC；在等边三角形 DBC 中，中线 DM 垂直于 BC。所以 $\theta = \angle AMD$，而由于 P 在 AM 上，介乎 A 与 M 之间，故亦有 $\theta = \angle PMD$。

三角形 DPM 在 P 处有一直角，其中 $PM = \frac{1}{2\sqrt{3}}$，斜边 $DM = \frac{\sqrt{3}}{2}$。因此

$$\cos\theta = \frac{PM}{DM} = \frac{1/(2\sqrt{3})}{\sqrt{3}/2} = \frac{1}{3},$$

正如所求。（作为验算，$PD^2 = DM^2 - PM^2 = \frac{3}{4} - \frac{1}{12} = \frac{2}{3}$，与第 (i) 部分吻合。）

第 (iii) 部分

最佳的球。设 O 为 DP 上满足 $OP = r$ 的点，其中 r 有待求出，使得 O 与三个斜面的距离也同样是 r。根据对称性，只需处理面 DBC。

平面 AMD 垂直于 BC（它包含 AM 和 DM，两者都垂直于 BC），所以它垂直于面 DBC，而由 O 到该面的垂线就在这个截面之内：它就是由 O 到直线 MD 的垂线。由于 O 与直线 MP 的距离为 OP，与直线 MD 的距离为 r，而我们想这两个距离相等，所以 O 必须在 M 处的角 θ 的角平分线（bisector）上。于是

$$r = PM\tan\tfrac{\theta}{2}.$$

由第 (ii) 部分，利用 $\tan^2\frac{\theta}{2} = \dfrac{1-\cos\theta}{1+\cos\theta}$，

$$\tan^2\tfrac{\theta}{2} = \frac{1-\frac{1}{3}}{1+\frac{1}{3}} = \frac{1}{2}, \qquad \tan\tfrac{\theta}{2} = \frac{1}{\sqrt{2}}$$

（取正值，因为 θ 是锐角）。因此

$$r = \frac{1}{2\sqrt{3}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2\sqrt{6}} = \frac{\sqrt{6}}{12}.$$

以 O 为球心、半径为 $\frac{\sqrt{6}}{12}$ 的球与四个面全部相切，而且位于四面体之内，因为 O 是一个内点，与每个面的距离都至少为 r。

? **想一想** 我们怎样知道，无论把球放在哪里，都不可能比这个球更大?

再大就放不下。在四面体内任取一点 X，设它与四个面的距离为 d_1, d_2, d_3, d_4。把 X 与各顶点连起来，便将四面体分成四个棱锥（pyramid），每个面上各有一个，每个的底面积都是 $\frac{\sqrt{3}}{4}$。把它们的体积相加，

$$V = \tfrac{1}{3} \cdot \tfrac{\sqrt{3}}{4} (d_1 + d_2 + d_3 + d_4).$$

但取 $X = D$（此时三个距离为 0，余下一个为 PD）可知 $V = \frac{1}{3} \cdot \frac{\sqrt{3}}{4} PD$。因此对每一个内点都有

$$d_1 + d_2 + d_3 + d_4 = PD = \frac{\sqrt{6}}{3}.$$

以 X 为球心、半径为 ρ 的球，只有在每个 $d_i \geq \rho$ 时才能放在四面体之内，所以 $4\rho \leq \frac{\sqrt{6}}{3}$，即 $\rho \leq \frac{\sqrt{6}}{12}$。所以最大的球的半径为

$$r = \frac{\sqrt{6}}{12} = \tfrac{1}{4} PD,$$

等号恰好在四个距离全部相等时（即在 O 处）成立。

验算

- 坐标：取 $A = (0,0,0)$、$B = (1,0,0)$、$C = (\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$、$D = (\frac{1}{2}, \frac{\sqrt{3}}{6}, \sqrt{2/3})$，则六条棱的长度全为 1，而 $PD \approx 0.8165$。
- 任意两个面的向外单位法向量（unit normal）的点积（dot product）为 $-\frac{1}{3}$，印证了 $\cos\theta = \frac{1}{3}$（约 70.5°）。
- 点 $(\frac{1}{2}, \frac{\sqrt{3}}{6}, \frac{\sqrt{6}}{12})$ 与四个面的距离全为 $0.20412\ldots = \frac{\sqrt{6}}{12}$。随机选取的内点，其距离之和总是 $0.8165\ldots$，而其中最小的距离从不超过 0.2041。
- 球心位于 PD 上由底部起计四分之一的高度。这与以下事实吻合：正四面体的内心（incentre）就是它的形心，而形心位于由一个面到其对顶点的四分之一处。

常见错误

- 把 $\angle DAM$ 或 $\angle DBA$ 当作两个面之间的夹角。所用的两条直线必须都垂直于公共棱，而 DM 和 AM 正是如此。
- 把 PA 当作中线的一半，而不是中线的三分之二。
- 求出了与四个面相切的球，却没有说明为何放在别处的球不可能更大。距离之和的恒等式只需两行便可解决这一点。

STEP 之外

到各个面的距离之和为常数，这是等边三角形的维维亚尼定理（Viviani's theorem）的三维版本。它源自重心坐标（barycentric coordinates），即把单纯形（simplex）内任何一点表示为其顶点的加权平均。角 $\arccos\frac{1}{3}$ 在化学中也会出现：甲烷中键与键之间的夹角是 $\arccos(-\frac{1}{3}) \approx 109.5°$，亦即同一个四面体的两个顶点在中心所张的角。想看更多这类立体推理，可参阅 Siklos 的 *Advanced Problems in Mathematics*；若想了解运用点积和叉积（cross product）的向量方法，可参阅 Strang 的 *Introduction to Linear Algebra*。

详细解答

STEP II 2010，第 7 题

纯数学 · 三次方程

STEP II 2010，第 7 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q7

关键思路

首项系数(leading coefficient)为正的三次曲线,当且仅当它的两个转向点(turning points)位于轴的同一侧时，才只与轴相交一次。第 (ii) 部分是解不含 x^2 项三次方程的经典技巧：选取代换 $x = u + q/u$，使 x^3 中的交叉项与 $-3qx$ 互相抵消，剩下一条关于 u^3 的二次方程。而第 (iii) 部分原来是同一条三次方程换了个样子，只是以 p 代替 x。

解答

第 (i) 部分

设 $f(x) = x^3 - 3qx - q(1+q)$。则 $f'(x) = 3x^2 - 3q$，它在 $x = \pm\sqrt{q}$ 处为零（由于 $q > 0$，两者均为实数）。由于 $f''(x) = 6x$，$x = -\sqrt{q}$ 处的点是局部极大点（local maximum），而 $x = \sqrt{q}$ 处的点是局部极小点（local minimum）。

在局部极大点，

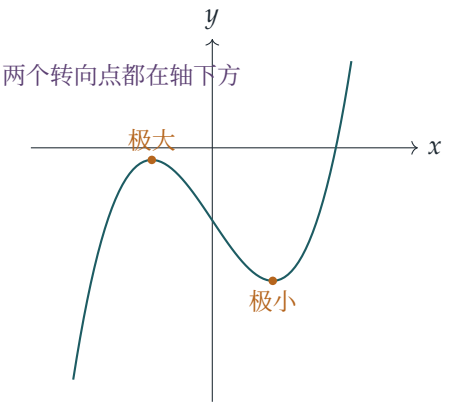
$$\begin{aligned} f(-\sqrt{q}) &= -q\sqrt{q} + 3q\sqrt{q} - q - q^2 \\ &= -q\left(q - 2\sqrt{q} + 1\right) \\ &= -q\left(\sqrt{q} - 1\right)^2. \end{aligned}$$

由于 $q > 0$ 且 $q \neq 1$（因此 $\sqrt{q} \neq 1$），此值严格小于零。局部极小值更低：

$$f(\sqrt{q}) = q\sqrt{q} - 3q\sqrt{q} - q - q^2 = -2q\sqrt{q} - q - q^2 < 0.$$

所以两个转向点都在 x 轴下方。当 $x < \sqrt{q}$ 时，曲线永远不会高于其局部极大值，而该值为负，因此曲线在此范围内不与轴相交。当 $x > \sqrt{q}$ 时，函数严格递增，由负值增至 $+\infty$，

因此恰好与轴相交一次。故此曲线只与 x 轴相交一次。



第 (ii) 部分

用二项式定理（binomial theorem）将 $x = u + q/u$ 立方，

$$\begin{aligned} x^3 &= u^3 + 3u^2 \cdot \frac{q}{u} + 3u \cdot \frac{q^2}{u^2} + \frac{q^3}{u^3} \\ &= u^3 + \frac{q^3}{u^3} + 3q\left(u + \frac{q}{u}\right) = u^3 + \frac{q^3}{u^3} + 3qx. \end{aligned}$$

所以 $x^3 - 3qx = u^3 + q^3/u^3$，三次方程变为

$$u^3 + \frac{q^3}{u^3} = q(1+q).$$

两边乘以 u^3（注意代换要有意义，须有 $u \neq 0$），得 u^3 所满足的二次方程

$$\left(u^3\right)^2 - q(1+q)\,u^3 + q^3 = 0.$$

想一想 这条二次方程怎样因式分解?

两根之和必须为 $q + q^2$，两根之积为 $q^3 = q \cdot q^2$，所以两根是 q 和 q^2：

$$\left(u^3 - q\right)\left(u^3 - q^2\right) = 0.$$

取实数立方根，得 $u = q^{1/3}$ 或 $u = q^{2/3}$。两者给出相同的 x：

$$q^{1/3} + \frac{q}{q^{1/3}} = q^{1/3} + q^{2/3}, \qquad q^{2/3} + \frac{q}{q^{2/3}} = q^{2/3} + q^{1/3}.$$

直接验证：设 $s = q^{1/3}$ 及 $x = s + s^2$，则 $x^3 = s^3 + 3s^4 + 3s^5 + s^6$ 而 $3qx = 3s^4 + 3s^5$，所以 $x^3 - 3qx = s^3 + s^6 = q(1+q)$。由第 (i) 部分，当 $q > 0$、$q \neq 1$ 时，此三次方程只有一个实根，所以该实根为

$$x = q^{1/3} + q^{2/3}.$$

第 (iii) 部分

由于 $\alpha+\beta=p$ 及 $\alpha\beta=q$，

$$\alpha^3+\beta^3=(\alpha+\beta)^3-3\alpha\beta(\alpha+\beta)=p^3-3qp,$$

此即所求。

若其中一根是另一根的平方，则 $\alpha^2=\beta$ 或 $\beta^2=\alpha$，因此无论哪种情况都有

$$(\alpha^2-\beta)(\beta^2-\alpha)=0.$$

这个式子关于 α 和 β 对称（symmetric），所以可以用 p 和 q 表示。展开得

$$\begin{aligned}(\alpha^2-\beta)(\beta^2-\alpha)&=\alpha^2\beta^2-(\alpha^3+\beta^3)+\alpha\beta\\&=q^2-(p^3-3qp)+q.\end{aligned}$$

令此式等于零，得到关系式

$$p^3-3qp-q(1+q)=0.$$

这正是第 (ii) 部分的三次方程，只是以 p 代替 x。当 $q>0$、$q\neq1$ 且 p 为实数时，p 必然是它唯一的实根，所以

$$p=q^{1/3}+q^{2/3}.$$

这是合理的：此时两根为 $\alpha=q^{1/3}$ 及 $\beta=q^{2/3}=\alpha^2$，其和为 p，其积为 q。

验算

取 $q=8$：第 (ii) 部分给出 $x=2+4=6$，而确实 $216-144-72=0$。在第 (iii) 部分，$p=6$ 给出 $t^2-6t+8=(t-2)(t-4)$，且 $4=2^2$。我们用电脑对数以千计的随机 q 验证了 (i) 中两个转向点的函数值，对随机复数 u 验证了恒等式 $x^3-3qx=u^3+q^3/u^3$，并确认 $p=q^{1/3}+q^{2/3}$ 总是给出一对其中一根为另一根平方的根。

常见错误

- 在 (i) 中只检查局部极小值为负。单凭这一点，曲线仍可能与轴相交三次；关键在于局部极大值为负，而这正是用到 $q\neq1$ 的地方。
- 在 (ii) 中担心 u^3 的两个值会给出两个不同的根。它们给出相同的 x，这是必然的，因为 u 和 q/u 的地位是对称的。

STEP 之外

(ii) 中的代换是一般三次方程的卡尔达诺公式（Cardano's formula）的核心，也正是在这里，复数（complex numbers）在数学中首次变得无可回避：当三次方程有三个实根时，关于 u^3 的二次方程反而有复数根。被排除的情况 $q = 1$ 是一个很好的例子。此时 $p = -1$ 也是允许的，而 $t^2 + t + 1 = 0$ 的根是 ω 和 ω^2，即单位的复数立方根（complex cube roots of unity），两者互为对方的平方。至于多项式方程何时可以用根式（radicals）求解，完整的故事就是伽罗瓦理论（Galois theory），Allenby 的 *Rings, Fields and Groups*（《环、域与群》）对此有很好的讲述。

详细解答

STEP II 2010，第 8 题

纯数学 · 曲线之间的面积

STEP II 2010，第 8 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q8

关键思路

由于 $\sin x \le 1$，曲线 $y = \mathrm{e}^{-x}\sin x$ 永远不会高于 $y = \mathrm{e}^{-x}$；两者恰好在 $\sin x = 1$ 之处相遇，而且在该处只是相切（touch）。因此每一块面积都只是 $\mathrm{e}^{-x}(1-\sin x)$ 在相邻两个切点之间的单一积分，毋须担心正负号改变。$\mathrm{e}^{-x}\sin x$ 的积分是标准的“分部积分两次”题型，而由于积分上下限全都是 $\sin x = 1$ 及 $\cos x = 0$ 的点，反导数（antiderivative）在这些点上会化简成非常简单的式子。

解答

草图

C_1 是大家熟悉的衰减指数函数（decaying exponential），由 $(0,1)$ 开始，递减并趋向 0。C_2 是一条阻尼正弦波（damped sine wave）：它由原点开始，在 π 的每个倍数处都等于零，而且因为 $-1 \le \sin x \le 1$，它被夹在 $y = \mathrm{e}^{-x}$ 与 $y = -\mathrm{e}^{-x}$ 之间。振荡幅度收缩得非常快；第一个负向凹陷的深度已只有约 0.014（在 $x = \frac{5\pi}{4}$ 处）。

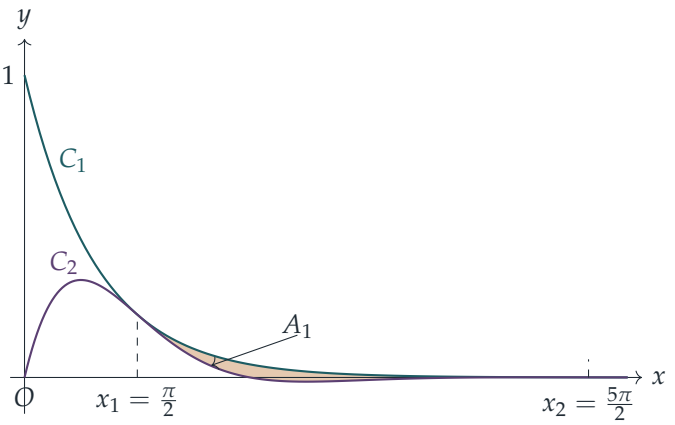
两条曲线在哪里相遇？由于 $\mathrm{e}^{-x} > 0$，

$$\mathrm{e}^{-x}\sin x = \mathrm{e}^{-x} \iff \sin x = 1 \iff x = \tfrac{\pi}{2} + 2k\pi, \quad k = 0, 1, 2, \ldots$$

在这些点上，两条曲线相切而非相交：C_2 处处位于 C_1 之上或之下。事实上，两者在该处的斜率（gradient）亦相同，因为

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x}\sin x\right) = \mathrm{e}^{-x}(\cos x - \sin x) = -\mathrm{e}^{-x} \quad \text{当 } \sin x = 1,\ \cos x = 0.$$

留意 C_2 的第一个极大值在 $x = \frac{\pi}{4}$（即 $\cos x = \sin x$ 之处），早于在 $x = \frac{\pi}{2}$ 的第一个切点：草图应显示 C_2 先到达峰顶，然后回落，与 C_1 相切。



面积 A_n

第 n 个切点是方程 $\sin x = 1$ 在 $x > 0$ 时的第 n 个解：

$$x_n = \tfrac{\pi}{2} + 2(n-1)\pi = \tfrac{(4n-3)\pi}{2}, \qquad x_{n+1} = x_n + 2\pi.$$

在 x_n 与 x_{n+1} 之间，有 $\mathrm{e}^{-x} - \mathrm{e}^{-x}\sin x = \mathrm{e}^{-x}(1-\sin x) \geq 0$，所以 C_1 在整段区间都是上边界，因此

$$A_n = \int_{x_n}^{x_{n+1}} \mathrm{e}^{-x}(1-\sin x)\,\mathrm{d}x.$$

想一想 怎样求 $\mathrm{e}^{-x}\sin x$ 的积分?

进行两次分部积分（integration by parts），便会回到原本的积分。设 $I = \int \mathrm{e}^{-x}\sin x\,\mathrm{d}x$。先取 $u = \sin x$、$\mathrm{d}v = \mathrm{e}^{-x}\,\mathrm{d}x$ 进行分部积分，再取 $u = \cos x$ 做一次，

$$\begin{aligned} I &= -\mathrm{e}^{-x}\sin x + \int \mathrm{e}^{-x}\cos x\,\mathrm{d}x \\ &= -\mathrm{e}^{-x}\sin x - \mathrm{e}^{-x}\cos x - \int \mathrm{e}^{-x}\sin x\,\mathrm{d}x, \end{aligned}$$

所以 $2I = -\mathrm{e}^{-x}(\sin x + \cos x)$，而 $I = -\frac{1}{2}\mathrm{e}^{-x}(\sin x + \cos x) + c$。（你可以通过求导来验证。）

因此 $\mathrm{e}^{-x}(1-\sin x)$ 的一个反导数是

$$F(x) = -\mathrm{e}^{-x} + \tfrac{1}{2}\mathrm{e}^{-x}(\sin x + \cos x).$$

巧妙之处来了：在每个切点上都有 $\sin x = 1$ 及 $\cos x = 0$，所以在这些点上

$$F(x) = -\mathrm{e}^{-x} + \tfrac{1}{2}\mathrm{e}^{-x} = -\tfrac{1}{2}\mathrm{e}^{-x}.$$

因此，利用 $x_{n+1} = x_n + 2\pi$，

$$\begin{aligned} A_n = F(x_{n+1}) - F(x_n) &= \tfrac{1}{2}\mathrm{e}^{-x_n} - \tfrac{1}{2}\mathrm{e}^{-x_n-2\pi} \\ &= \tfrac{1}{2}\mathrm{e}^{-x_n-2\pi}\left(\mathrm{e}^{2\pi} - 1\right). \end{aligned}$$

最后，$x_n + 2\pi = \frac{(4n-3)\pi}{2} + \frac{4\pi}{2} = \frac{(4n+1)\pi}{2}$，所以

$$A_n = \tfrac{1}{2}\left(\mathrm{e}^{2\pi} - 1\right)\mathrm{e}^{-(4n+1)\pi/2},$$

此即所求。

无穷级数之和

由公式可见 $A_{n+1} = \mathrm{e}^{-2\pi}A_n$，所以各面积组成一个几何级数（geometric series），首项为 $A_1 = \frac{1}{2}(\mathrm{e}^{2\pi} - 1)\mathrm{e}^{-5\pi/2}$，公比（common ratio）为 $r = \mathrm{e}^{-2\pi}$。由于 $0 < r < 1$，该级数收敛（converges），而且

$$\begin{aligned} \sum_{n=1}^{\infty} A_n = \frac{A_1}{1 - \mathrm{e}^{-2\pi}} &= \frac{\frac{1}{2}(\mathrm{e}^{2\pi} - 1)\mathrm{e}^{-5\pi/2}}{\mathrm{e}^{-2\pi}(\mathrm{e}^{2\pi} - 1)} \\ &= \tfrac{1}{2}\mathrm{e}^{-\pi/2}. \end{aligned}$$

验算

- 总和就是 x_1 右方两曲线之间的全部面积，所以它应等于 $\int_{\pi/2}^{\infty} \mathrm{e}^{-x}(1 - \sin x)\,\mathrm{d}x = \lim_{X\to\infty} F(X) - F(\frac{\pi}{2}) = 0 + \frac{1}{2}\mathrm{e}^{-\pi/2}$。结果的确如此，而且完全不需要用到级数。
- 数值上，$A_1 \approx 0.10375$，而 $\frac{1}{2}\mathrm{e}^{-\pi/2} \approx 0.10394$：正如草图所示，几乎全部面积都在第一个区域内，因为 $\mathrm{e}^{-2\pi} \approx 0.0019$。对 $\mathrm{e}^{-x}(1 - \sin x)$ 在 $[\frac{\pi}{2}, \frac{5\pi}{2}]$ 上使用辛普森法则（Simpson's rule），得到相同的 0.103746。

常见错误

- 误以为切点就是 C_2 的极大值点。极大值在 $x = \frac{\pi}{4} + k\pi$；两曲线则在 $x = \frac{\pi}{2} + 2k\pi$ 相切。
- 把 $\sin x = \pm 1$ 的每个解都算进去，或把 $\sin x = 1$ 的解当作相隔 π：相邻两个切点相隔 2π。
- 计算 $\int \mathrm{e}^{-x}\sin x\,\mathrm{d}x$ 时弄错正负号。把答案求导来检查一下。

STEP 之外

曲线 C_2 是阻尼振荡的形状，即有摩擦的弹簧的运动，而包络线（envelope）$\pm e^{-x}$ 决定了振荡衰减的快慢。像 $\int e^{-x} \sin x \, dx$ 这类积分，若视为 $\int e^{(-1+i)x} \, dx$ 的虚部（imaginary part）来计算，会快得多；这个技巧在大学一年级的微分方程（differential equations）及拉普拉斯变换（Laplace transforms）课程中经常用到。Boyce 与 DiPrima 合著的 *Elementary Differential Equations*（《初等微分方程》）详细讨论了阻尼振荡。

详细解答

STEP II 2010，第 9 题

力学 · 相撞的抛体

STEP II 2010，第 9 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

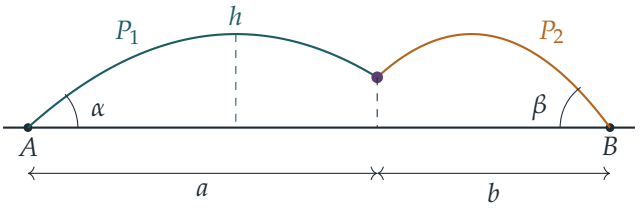
step.maths.org/questions/10-s2-q9

关键思路

两粒同时抛出的质点以相同的加速度 g 下落，所以要令两者在碰撞一刻的高度相等，唯一的可能是它们的初始竖直速度相等。这样两者的竖直运动便完全一样：最大高度同为 h，在同一时刻到达，而且（若没有碰撞）会在同一时刻落地。碰撞发生在飞行最高点与落地之间，而每粒质点走过的水平距离与时间成正比，因此这段距离被夹在两个简单的界限之间。

解答

我们把两粒质点都视为只受重力 g 作用、自由运动的抛射体（projectile），并忽略空气阻力（air resistance）。设 P_1 的速率为 u，P_2 的速率为 v，取抛出时刻为 $t = 0$，并设碰撞发生于时刻 $t_c > 0$。



相同的竖直运动

在时刻 t，P_1 与 P_2 的高度分别为

$$y_1 = ut\sin\alpha - \tfrac{1}{2}gt^2, \qquad y_2 = vt\sin\beta - \tfrac{1}{2}gt^2.$$

碰撞时 $y_1 = y_2$，故 $(u\sin\alpha - v\sin\beta)\,t_c = 0$；由于 $t_c > 0$，

$$u\sin\alpha = v\sin\beta.$$

因此两粒质点的竖直运动完全相同。特别地，两者在同一时刻 T 到达相同的最大高度 h，其中

$$T = \frac{u\sin\alpha}{g} = \frac{v\sin\beta}{g}, \qquad h = \frac{u^2\sin^2\alpha}{2g} = \frac{v^2\sin^2\beta}{2g},$$

而且两者都会在时刻 $2T$ 回到地面。

想一想 “在空中”和“在 P_1 到达其最大高度之后”这两个条件，告诉了我们关于 t_c 的甚么信息?

P_1 在时刻 T 位于最高点，所以“之后”给出 $t_c > T$。两粒质点只在 $0 < t < 2T$ 期间位于地面之上，所以要在空中碰撞，便需要 $t_c < 2T$。因此

$$T < t_c < 2T.$$

b 的界限

在水平方向上，P_2 以恒定速率 $v\cos\beta$ 运动，所以 $b = v\cos\beta\, t_c$。在时刻 T，它走过的水平距离会是

$$v\cos\beta\, T = \frac{v^2\sin\beta\cos\beta}{g} = \frac{v^2\sin^2\beta}{2g}\cdot\frac{2\cos\beta}{\sin\beta} = 2h\cot\beta.$$

因此

$$b = v\cos\beta\, T\cdot\frac{t_c}{T} = 2h\cot\beta\cdot\frac{t_c}{T},$$

又由于 $1 < t_c/T < 2$，且 $2h\cot\beta > 0$（因为 β 是锐角），

$$2h\cot\beta < b < 4h\cot\beta,$$

得证。

a 的界限

完全同理，$a = u\cos\alpha\, t_c$，而

$$u\cos\alpha\, T = \frac{u^2\sin\alpha\cos\alpha}{g} = 2h\cot\alpha,$$

所以 $a = 2h\cot\alpha\cdot t_c/T$，再由 $1 < t_c/T < 2$ 得

$$2h\cot\alpha < a < 4h\cot\alpha.$$

这里用的是同一个 h：按定义，h 是 P_1 的最大高度；而由于两者的竖直运动完全相同，它同时也是 P_2 的最大高度。

验算

- 边界情况：若碰撞恰好发生在最高点（$t_c = T$），则 $a = 2h\cot\alpha$，即熟悉的半射程（half-range）$\frac{u^2\sin 2\alpha}{2g}$；若碰撞发生在落地一刻（$t_c = 2T$），则 $a = 4h\cot\alpha$，即全射程 $\frac{u^2\sin 2\alpha}{g}$。题目中的严格不等式正好介乎两者之间。
- 数值验证：取 $g = 9.8$、$u = 20$、$\alpha = 0.7$、$\beta = 0.9$（弧度），$v = u\sin\alpha/\sin\beta$ 及 $t_c = 1.5T$，得 $h \approx 8.470$ 及 $a \approx 30.17$，介乎 $2h\cot\alpha \approx 20.11$ 与 $4h\cot\alpha \approx 40.22$ 之间（事实上它恰好等于 1.5×20.11）。我们用这个方法测试了 20 000 个随机设置，逐步模拟 P_1 的运动以确认 h，并检查两粒质点确实重合。

常见错误

- 把两粒质点到达碰撞点的时间当作不同：它们是同时抛出的，所以两者用的是同一个 t_c。
- 忘记解释上界的由来：上界来自碰撞发生在空中，即在共同落地时刻 $2T$ 之前。

STEP 之外

关键的一步是转换参考系（frame of reference）：在一个以加速度 g 自由下落的参考系中，两粒质点都以恒定速度沿直线运动，所以只有当它们在该参考系中互相瞄准时才会碰撞，这正是经典的“猴子与猎人”（monkey and hunter）问题的思路。以加速参考系思考问题，以及当加速度不均匀时随之出现的惯性力（fictitious forces），是大学一年级力学的核心主题。Kleppner 与 Kolenkow 合著的 *An Introduction to Mechanics*（《力学导论》）讨论了抛射体、相对运动及非惯性参考系（non-inertial frames），并附有大量优秀习题。

详细解答

STEP II 2010，第 10 题

力学 · 弹性碰撞

> **STEP II 2010，第 10 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s2-q10

本题在 *Beyond STEP* 第 10.2 节叠球效应中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch10/stacked-balls.html

关键思路

一切都建基于一次碰撞：一粒质点以完全弹性的方式撞向另一粒静止的质点。动量守恒（conservation of momentum）加上牛顿实验定律（Newton's experimental law，取 $e = 1$）便可求出传递出去的速率，而这个速率只取决于两个质量之比。在第 (ii) 部分，每次碰撞的质量比都是 λ，所以每次碰撞都把速率乘以同一个因子，余下的工作交给等比数列（geometric progression）就可以了。

解答

把各质点视为点质量，在光滑桌面上沿一条直线运动，所有速度均以 B 的初始运动方向为正方向量度。

第 (i) 部分

设 B 和 A 在紧接碰撞之后的速度分别为 v_B 和 v_A。由动量守恒得

$$bmu = bm\,v_B + m\,v_A,$$

又因为碰撞是完全弹性（perfectly elastic）的（$e = 1$），按牛顿实验定律，分离速率等于接近速率：

$$v_A - v_B = u.$$

把 $v_B = v_A - u$ 代入动量方程，并约去 m，

$$bu = b(v_A - u) + v_A \quad \Longrightarrow \quad v_A = \frac{2bu}{b+1}.$$

所以 A 的速率是 $\dfrac{2b}{b+1}\,u$。

这条公式只用到质量比 b，而且推导对任何 $b > 0$ 都成立。由于 $b+1 > b$ 且 $u > 0$，

$$v_A = 2u \cdot \frac{b}{b+1} < 2u,$$

不论 b 取甚么值都是如此。所以 A 永远不可能以 B 初始速率的两倍运动；不过只要把 b 取得够大，它的速率可以任意接近 $2u$。

第 (ii) 部分

沿直线依次排列的质量为

$$B_1 : \lambda^n m, \quad B_2 : \lambda^{n-1} m, \quad \ldots, \quad B_n : \lambda m, \quad A : m.$$

每粒质点的质量都是下一粒的 λ 倍。所以每当一粒以速率 v 运动的质点撞上下一粒（静止的）质点，第 (i) 部分的结果便适用，其中 $b = \lambda$，而被撞的质点会以速率 kv 离开，其中

$$k = \frac{2\lambda}{\lambda+1}.$$

由第 (i) 部分的方程 $v_B = v_A - u$ 可知，撞击的质点会继续向前运动，速率为

$$kv - v = \frac{\lambda - 1}{\lambda + 1}\, v,$$

由于 $\lambda > 1$，这个值是正数。

于是 B_1 令 B_2 得到速率 ku，然后 B_2 令 B_3 得到速率 k^2u，如此类推，直至以 $k^{n-1}u$ 运动的 B_n 令 A 得到速率

$$v_A = k^n u = \left(\frac{2\lambda}{\lambda+1}\right)^n u.$$

? **想一想** 会不会有额外的碰撞，即某个 B_i 把速率传递出去之后，又追上前面的那一粒？

全部 n 次碰撞之后，B_i 以速率 $\dfrac{\lambda-1}{\lambda+1}k^{i-1}u$ 向前运动，其中 $i = 1, \ldots, n$，而 A 则以速率 $k^n u$ 运动。由于 $\lambda > 1$，我们有 $k > 1$（因为 $2\lambda > \lambda + 1$），所以这些速率沿着直线严格递增：每粒质点都比它后面的那一粒快，彼此愈离愈远。因此不会再有碰撞，A 的最终速率就是 $k^n u$。

因为 $k > 1$，当 $n \to \infty$ 时 $k^n \to \infty$。所以对任何目标速率 V，只要把 n 取得够大，使 $k^n u > V$，便可令 A 的速率超过 V：A 的速率没有上限。

$\lambda = 4$ 的情况

这时 $k = \frac{8}{5}$，我们要找出满足 $\left(\frac{8}{5}\right)^n > 20$ 的最小 n。取以 10 为底的对数（logarithm；这是递增函数，所以不等式的方向不变），

$$n \log_{10} \tfrac{8}{5} > \log_{10} 20.$$

写成 $\frac{8}{5} = \frac{2^4}{10}$ 和 $20 = 2 \times 10$，

$$\log_{10} \tfrac{8}{5} = 4 \log_{10} 2 - 1 \approx 0.20412,$$
$$\log_{10} 20 = \log_{10} 2 + 1 \approx 1.30103.$$

由于 $\log_{10} \frac{8}{5} > 0$，我们可以把两边同除以它，得

$$n > \frac{1.30103}{0.20412} \approx 6.37.$$

满足此条件的最小整数是 $n = 7$。（这里的余裕相当充足：6.37 离整数很远，所以 $\log_{10} 2$ 的近似值不会影响答案。）

验算

- 在第 (i) 部分，$b = 1$ 给出 $v_A = u$ 和 $v_B = 0$：质量相等的质点互换速度，正该如此。当 $b \to \infty$ 时，$v_A \to 2u$。
- 动能（kinetic energy）：以 $v_A = \frac{2b}{b+1}u$ 和 $v_B = \frac{b-1}{b+1}u$ 代入，得 $bv_B^2 + v_A^2 = \frac{b(b-1)^2 + 4b^2}{(b+1)^2}u^2 = bu^2$，所以能量守恒。
- 直接计算：$1.6^6 \approx 16.78 < 20$ 而 $1.6^7 \approx 26.84 > 20$，确认 $n = 7$。用精确分数逐步模拟整条质点链，同样得出 A 的速率在刚好 n 次碰撞之后恰为 $k^n u$。

常见错误

- 把质量的次序弄反：B_1 是最重的（$\lambda^n m$），B_n 是各个 B 之中最轻的（λm），而 A 比它还要轻。
- 把碰撞次数数成 $n - 1$ 次而不是 n 次：速率沿 $B_1 \to B_2 \to \cdots \to B_n \to A$ 传递，共有 n 步。
- 忘记排除其后再发生碰撞的可能，或者答了非整数 6.37，而不是最小整数 $n = 7$。

STEP 之外

单独一次弹性碰撞最多只能把速率加倍，但一连串碰撞却没有这个限制。这正是“叠球”示范背后的原理：把一个轻球叠在重球上面一同放下，轻球反弹的高度会远远超过它落下的高度。利用动量和能量来分析碰撞（更简洁的做法是在质心参考系（centre-of-mass frame）中分析，在那里弹性碰撞只是把每个速度反转），是大学一年级力学的核心课题。Kleppner and Kolenkow, *An Introduction to Mechanics*（力学导论）对此有深入的讨论。

详细解答

STEP II 2010，第 11 题

力学 · 三力平衡

STEP II 2010，第 11 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q11

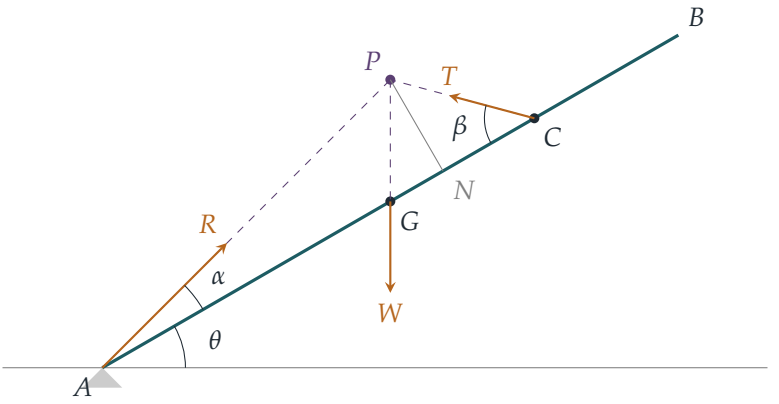
关键思路

作用于杆上的力只有三个：杆的重量、A 点的反作用力（reaction）和 C 点的张力（tension）。当三个互不平行的力使一个物体保持平衡（equilibrium）时，它们的作用线（line of action）必定通过同一点。因此，由 A 出发的反作用力所在直线、由 C 出发的绳所在直线，以及通过杆中点的铅垂线，三者相交于一点。从这个公共点出发，沿杆的方向和垂直于杆的方向量度距离，整道题便化为简单的直角三角学，完全不必求出各个力的大小。

解答

模型与图示

我们把杆视为均匀刚性杆（uniform rigid rod），所以它的重量 W 作用于中点 G，其中 $AG = 2L$。绳是轻绳，所以它在 C 点沿绳本身的方向拉着杆，其中 $AC = 3L$。支承在 A 点对杆施加单一个力 R（题目没有说明支承是光滑还是粗糙，而我们也不需要知道）。下图按比例绘出 $\theta = 30°$、$\beta = 45°$ 的情况。



证明公式

重量沿铅垂方向作用，其作用线通过 G；绳的作用线通过 C，反作用力的作用线则通过 A。这三条直线并不平行，所以三者都通过同一点，我们称这点为 P。

? **想一想** 为甚么三条作用线必定相交于一点?

设 R 和 T 的作用线相交于 P。对 P 取力矩（moment），R 和 T 的力矩都是零，所以杆若要处于转动平衡，重量对 P 的力矩也必须是零。这就表示通过 G 的铅垂线同样通过 P。

设 N 为由 P 到杆的垂足，并设 $h = PN$。在直角三角形 PNA 中，A 处的角是 α，所以 $AN = h\cot\alpha$。在直角三角形 PNC 中，C 处的角是 β，所以 $NC = h\cot\beta$。由于 $AN + NC = AC$，

$$h\cot\alpha + h\cot\beta = 3L. \tag{1}$$

（若 α 或 β 是钝角，N 便落在 A 或 C 之外，但带正负号的长度 $h\cot\alpha$ 与 $h\cot\beta$ 之和仍然等于 AC，所以 (1) 在任何情况下都成立。）

现在利用第三条作用线。杆与水平线的夹角是 θ，所以垂线 PN 与铅垂线 PG 的夹角也是 θ。在直角三角形 PNG 中，P 处的角是 θ，所以 $GN = h\tan\theta$，而且 N 位于 G 的上方一侧（离 A 较远），如图所示。因此

$$h\cot\alpha = AN = AG + GN = 2L + h\tan\theta. \tag{2}$$

由 (1) 减去 (2)，得 $h\cot\beta = L - h\tan\theta$，即

$$L = h(\cot\beta + \tan\theta).$$

把 L 的这个值代入 (2)：

$$h\cot\alpha = 2h(\cot\beta + \tan\theta) + h\tan\theta = h(3\tan\theta + 2\cot\beta).$$

由于 $h \neq 0$（否则 P 会落在杆上的 G 点，而系于 C 点的绳便会沿着杆的方向，无法抵销重量对 A 的力矩），我们可以除以 h：

$$\cot\alpha = 3\tan\theta + 2\cot\beta,$$

这正是所需的结果。

$\theta = 30°$、$\beta = 45°$ 的情况

由 $\tan 30° = \frac{1}{\sqrt{3}}$ 及 $\cot 45° = 1$，

$$\cot\alpha = \frac{3}{\sqrt{3}} + 2 = 2 + \sqrt{3}.$$

现在我们需要 $\cot 15°$。利用 $\tan(A - B)$ 的公式，取 $A = 45°$、$B = 30°$，

$$\tan 15° = \frac{1 - \frac{1}{\sqrt{3}}}{1 + \frac{1}{\sqrt{3}}} = \frac{\sqrt{3} - 1}{\sqrt{3} + 1} = \frac{(\sqrt{3} - 1)^2}{2} = 2 - \sqrt{3},$$

所以

$$\cot 15° = \frac{1}{2 - \sqrt{3}} = \frac{2 + \sqrt{3}}{(2 - \sqrt{3})(2 + \sqrt{3})} = 2 + \sqrt{3}.$$

因此 $\cot\alpha = \cot 15°$。角 α 介乎 $0°$ 与 $180°$ 之间，而 cot 在这个区间上严格递减（strictly decreasing），所以每个值它都恰好取一次。故此 $\alpha = 15°$。

验算

- *另一条独立的途径。*对 A 取力矩，得 $T \cdot 3L\sin\beta = W \cdot 2L\cos\theta$。然后沿杆的方向和垂直于杆的方向分解各力，便得到 R 的两个分量，而 $\cot\alpha$ 就是它们之比；这样可以重新得出该公式。我们用这个方法，对 10^5 组随机的 (θ, β) 作了数值检验，两者吻合至约 10^{-15}。
- *特殊值。*当 $\theta = 30°$、$\beta = 45°$ 时，以数值方法求平衡，得 $\alpha = 15.000°$，而 $\cot 15° = 3.7320\ldots = 2 + \sqrt{3}$。图是按比例绘画的：$P$ 位于 G 的正上方，比 A 高出 $\sqrt{3}L$。
- *极限情况。*若 $\theta \to 0$ 而 $\beta = 90°$（杆水平，绳铅垂），公式给出 $\cot\alpha = 0$，所以 $\alpha = 90°$：反作用力也是铅垂的，而当其余各力全是铅垂时，这是必然的。

常见错误

- 把重量放在 C 点，或取 $AG = 3L/2$：重量作用于整枝杆的中点，与 A 相距 $2L$。

- 弄错 GN 的正负号。由于杆从 A 起向上倾斜，通过 G 的铅垂线与由 P 所作的垂线，相交于 G 远离 A 的一侧，所以 $AN = 2L + h\tan\theta$，而不是 $2L - h\tan\theta$。
- 写到 $\cot\alpha = 2 + \sqrt{3}$ 便停下来，既没有证明这就是 $\cot 15^\circ$，也没有说明为甚么这能唯一确定 α。

STEP 之外

三力定理（three-force theorem）是静力学（statics）中一个普遍事实的特殊情况。作用于刚体（rigid body）上的任何力系，都可以化简为通过某一选定点的单一个力，再加上一个力偶（couple），而平衡就是指两者皆为零。到了大学，这会变成一个向量条件：合力与总力矩 $\sum \mathbf{r}_i \times \mathbf{F}_i$ 都等于零，而这个条件正是刚体动力学的基础。Kleppner 与 Kolenkow 所著的 *An Introduction to Mechanics*（力学导论）就是用这套向量语言来阐述力矩（torque）与平衡的。

详细解答

STEP II 2010，第 12 题

概率与统计 • 平均值与中位数

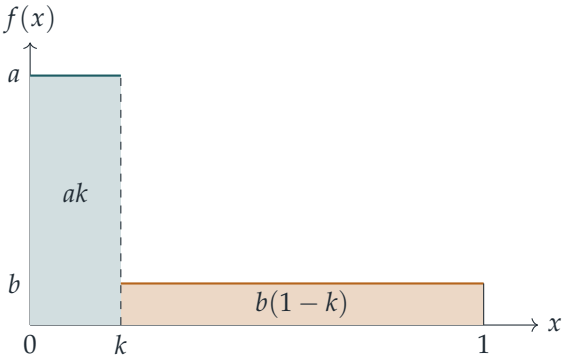
STEP II 2010，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s2-q12

关键思路

这个密度函数（density function）是只有两级的阶梯函数（step function），所以每个概率都是一个长方形的面积，而唯一真正的条件就是总面积等于 1。这个条件把 k 用 a 和 b 表示出来；一旦消去 k，余下的全是代数运算。至于中位数（median），问题只在于那个较高的第一个长方形是否已经占了一半的概率。



解答

图中两个长方形的面积分别是 ak 和 $b(1-k)$，所以密度函数的积分等于 1，当且仅当

$$ak + b(1-k) = 1, \qquad \text{即} \qquad k = \frac{1-b}{a-b}.$$

（由于 $a > b$，我们可以除以 $a - b$。）以下一直采用这个 k 值。

为何 $a > 1$ 且 $b < 1$。 方程 $ak + b(1-k) = 1$ 表示 1 是 a 和 b 的加权平均（weighted average），权重为 k 和 $1-k$，两者都严格介乎 0 与 1 之间。权重为正的加权平均必定严格介乎两个数值之间，所以 $b < 1 < a$。不用文字来说，就是：$1 - b = k(a-b) > 0$ 及 $a - 1 = (1-k)(a-b) > 0$。

第 (i) 部分

在每一段上对 $xf(x)$ 积分，

$$\mathrm{E}(X) = \int_0^k ax\,\mathrm{d}x + \int_k^1 bx\,\mathrm{d}x = \tfrac{1}{2}ak^2 + \tfrac{1}{2}b(1-k^2) = \tfrac{1}{2}\big(b + (a-b)k^2\big).$$

现在 $(a-b)k^2 = \dfrac{(1-b)^2}{a-b}$，所以

$$\mathrm{E}(X) = \frac{b(a-b) + (1-b)^2}{2(a-b)} = \frac{ab - b^2 + 1 - 2b + b^2}{2(a-b)} = \frac{1 - 2b + ab}{2(a-b)},$$

正如所求。

第 (ii) 部分

累积分布函数（cumulative distribution function）是

$$\mathrm{F}(x) = \begin{cases} ax & 0 \le x < k, \\ ak + b(x-k) & k \le x \le 1. \end{cases}$$

它在 $[0,1]$ 上连续而且严格递增，所以中位数 M 是 $\mathrm{F}(M) = \frac{1}{2}$ 的唯一解。该用哪一条公式，取决于第一个长方形是否占了至少一半的概率。

> **想一想** 怎样把“第一个长方形的面积至少是 $\frac{1}{2}$”化为关于 a 和 b 的条件?

代入 k 的值，并消去正数分母 $a-b$：

$$ak \ge \tfrac{1}{2} \iff 2a(1-b) \ge a - b \iff 2a - 2ab \ge a - b \iff a + b \ge 2ab.$$

情况 $a + b \ge 2ab$。这时 $\mathrm{F}(k) = ak \ge \frac{1}{2}$，所以中位数位于 $[0,k]$ 之内，并满足 $aM = \frac{1}{2}$：

$$M = \frac{1}{2a}.$$

情况 $a + b \le 2ab$。这时 $ak \le \frac{1}{2}$，所以中位数位于 $[k,1]$ 之内。从右端入手最为简便：M 右方的概率是 $b(1-M)$，而它必须等于 $\frac{1}{2}$。因此

$$M = 1 - \frac{1}{2b} = \frac{2b-1}{2b}.$$

当 $a + b = 2ab$ 时，两条公式都适用，而且结果一致，因为这时 $1 - \frac{1}{2b} = \frac{2ab-a}{2ab} = \frac{b}{2ab} = \frac{1}{2a}$。

第 (iii) 部分

我们在每个情况下计算 $\mathrm{E}(X) - M$，并证明它是正数。

情况 $a + b \geq 2ab$。通分，以 $2a(a-b)$ 为公分母，

$$\mathrm{E}(X) - \frac{1}{2a} = \frac{a(1-2b+ab) - (a-b)}{2a(a-b)} = \frac{b - 2ab + a^2 b}{2a(a-b)} = \frac{b(a-1)^2}{2a(a-b)}.$$

情况 $a + b \leq 2ab$。通分，以 $2b(a-b)$ 为公分母，

$$\begin{aligned}\mathrm{E}(X) - \frac{2b-1}{2b} &= \frac{b(1-2b+ab) - (2b-1)(a-b)}{2b(a-b)} \\ &= \frac{(b - 2b^2 + ab^2) - (2ab - 2b^2 - a + b)}{2b(a-b)} = \frac{a(1-b)^2}{2b(a-b)}.\end{aligned}$$

想一想 为甚么这些数值是严格为正，而不仅仅是非负？

在两个情况下，a、b 和 $a-b$ 都是正数，而平方项严格为正，因为我们一开始已证明 $a > 1$ 且 $b < 1$。所以在任何情况下都有 $\mathrm{E}(X) - M > 0$，即 $M < \mathrm{E}(X)$。

验算

- 取 $a = 3$、$b = \frac{1}{2}$：这时 $k = \frac{1}{5}$，$a + b = \frac{7}{2} \geq 3 = 2ab$，而 $\mathrm{E}(X) = \frac{1 - 1 + \frac{3}{2}}{5} = 0.3$，$M = \frac{1}{6}$。两者之差 $0.1\overline{3}$ 与 $\frac{b(a-1)^2}{2a(a-b)} = \frac{2}{15}$ 吻合。
- 取 $a = \frac{5}{4}$、$b = \frac{9}{10}$：这次 $a + b = 2.15 \leq 2.25 = 2ab$，所以 $M = 1 - \frac{5}{9} = \frac{4}{9} \approx 0.444$，而 $\mathrm{E}(X) = \frac{13}{28} \approx 0.464$。
- 当 $a \to 1^+$（因而 $b \to 1^-$）时，密度函数趋向均匀分布（uniform distribution）的密度函数，两个差值都趋向 0：平均值（mean）与中位数同为 $\frac{1}{2}$，按对称性理应如此。

我们以 300 对随机选取的 (a, b) 作了数值检验：对 F 使用二分法（bisection）求得的中位数与分情况的公式吻合，以中点求积法（midpoint quadrature）求得的平均值与第 (i) 部分吻合，而 $\mathrm{E}(X) - M$ 亦等于上面已因式分解的表达式。

常见错误

- 忘记消去 k：答案须只用 a 和 b 表示。
- 把第二个中位数写成 $b(M-k) = \frac{1}{2}$ 的解，忘记了在 $[0, k)$ 上已经累积的概率 ak。
- 只得出 $\mathrm{E}(X) \geq M$ 的结论：严格不等式需要 $a \neq 1$ 和 $b \neq 1$，而这正是题目开首部分所提供的。

STEP 之外

这个密度函数是递减的，所以它是“右偏”（skewed to the right）的：长长的尾部把平均值拉到中位数之外。这种现象很常见，但并非必然；大学的第一门概率课程会说明，对于单峰分布（unimodal distribution），平均值、中位数和众数（mode）之间有何关系，以及那些耳熟能详的经验法则在甚么情况下会失效。像这样的阶梯函数，也是直方图（histogram）的基本组件，同时是测度论（measure theory）中用来定义积分的简单函数（simple function）的基本组件。想进一步了解连续分布，可参阅 Grimmett and Welsh, *Probability: An Introduction*（概率论入门教材）。

详细解答

STEP II 2010，第 13 题

概率与统计 · 独立事件

> **STEP II 2010，第 13 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s2-q13

关键思路

Rosalind 要合格，必须对 Pardeep 至少胜一局，并且对 Quentin 至少胜一局。“至少一次”用补集（complement）处理总是最容易：n 局中至少胜一局的概率，等于 1 减去全部落败的概率。各局独立（independent）时，每种策略的成功概率便是两个这类因子的乘积，而比较策略就是把两者之差因式分解（factorise）。以 $a = 1 - p$ 和 $b = 1 - q$ 表示落败的概率，可使代数运算保持简洁。

解答

我们假设不同对局的结果互相独立，并且无论之前发生过甚么，每局对 Pardeep（或 Quentin）的胜出概率都是 p（或 q）。对局的次序并不重要：无论她取得资格后是否停止比赛，她取得资格的结果都完全相同。

设 Rosalind 与 Pardeep 对赛 m 局、与 Quentin 对赛 n 局，令 A 为她对 Pardeep 至少胜一局的事件（event），B 为她对 Quentin 至少胜一局的事件。那么 $\mathrm{P}(A) = 1 - (1-p)^m$，因为只有当她 m 局全败时 A 才不发生；同理 $\mathrm{P}(B) = 1 - (1-q)^n$。这两个事件涉及不同的对局，所以互相独立，于是

$$\mathrm{P}(\text{获接纳}) = \big(1 - (1-p)^m\big)\big(1 - (1-q)^n\big).$$

以下一律记 $a = 1 - p$ 和 $b = 1 - q$。由于 $0 < p < q < 1$，我们有 $0 < b < a < 1$。

第 (i) 部分

与 Pardeep 对赛两局、与 Quentin 对赛一局，成功概率为 $(1 - a^2)(1 - b)$；与 Quentin 对赛两局、与 Pardeep 对赛一局，成功概率为 $(1 - a)(1 - b^2)$。把各个平方差（difference

of squares）因式分解，

$$\begin{aligned}(1-a^2)(1-b)-(1-a)(1-b^2) &= (1-a)(1-b)\big[(1+a)-(1+b)\big]\\ &= (1-a)(1-b)(a-b)\\ &= pq(q-p).\end{aligned}$$

由于 $p,q>0$ 且 $q>p$，此式严格为正。所以与 Pardeep 对赛两局，获接纳的机会较大，她应该选择这个方案。

第 (ii) 部分

三种策略的成功概率分别为

$$S_1=(1-a^3)(1-b),\qquad S_2=(1-a^2)(1-b^2),\qquad S_3=(1-a)(1-b^3).$$

三者都含有因子 $(1-a)(1-b)=pq$，因为 $1-x^k=(1-x)(1+x+\cdots+x^{k-1})$。

策略 1 必定胜过策略 3。我们有

$$\begin{aligned}S_1-S_3 &= pq\big[(1+a+a^2)-(1+b+b^2)\big]\\ &= pq\,(a-b)(1+a+b)>0,\end{aligned}$$

因为 $a>b$。所以策略 3 永远不是最佳选择，只需比较策略 1 和策略 2。

策略 1 对策略 2。我们有

$$S_1-S_2=pq\big[(1+a+a^2)-(1+a)(1+b)\big]=pq\big[a^2-b(1+a)\big].$$

换回 p 和 q，

$$S_1-S_2=pq\,D,\qquad \text{其中}\qquad D=(1-p)^2-(1-q)(2-p).$$

由于 $pq>0$，策略 1 优于策略 2 当且仅当 $D>0$。

> ? **想一想** 条件 $q-p>\frac{1}{2}$ 怎样有助令 D 为正?

D 中唯一涉及 q 的项是 $-(1-q)(2-p)$，而由于 $2-p>0$，它随 q 增加而增加。所以我们把 q 换成条件所容许的最小值。若 $q-p>\frac{1}{2}$，则 $1-q<\frac{1}{2}-p$，又因为 $2-p>0$，

$$(1-q)(2-p)<\big(\tfrac{1}{2}-p\big)(2-p)=1-\tfrac{5}{2}p+p^2.$$

因此

$$D>(1-2p+p^2)-\big(1-\tfrac{5}{2}p+p^2\big)=\tfrac{1}{2}p>0.$$

所以 $S_1>S_2$，连同 $S_1>S_3$，这就证明了当 $q-p>\frac{1}{2}$ 时 Rosalind 应该选择策略 1。

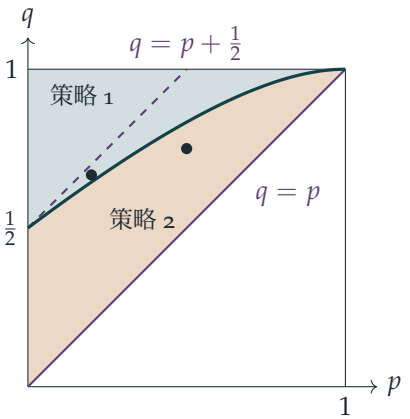
当 $q-p<\frac{1}{2}$ 时，两者都有可能较优。以下是两个例子：

- $p=\frac{1}{2}, q=\frac{3}{4}$（所以 $q-p=\frac{1}{4}$）。这时 $D=\frac{1}{4}-\frac{1}{4}\cdot\frac{3}{2}=-\frac{1}{8}<0$。事实上 $S_1=\frac{7}{8}\cdot\frac{3}{4}=\frac{42}{64}$ 而 $S_2=\frac{3}{4}\cdot\frac{15}{16}=\frac{45}{64}$，所以**策略 2 较佳**。
- $p=\frac{1}{5}$，$q=\frac{2}{3}$（所以 $q-p=\frac{7}{15}<\frac{1}{2}$）。这时 $D=\frac{16}{25}-\frac{1}{3}\cdot\frac{9}{5}=\frac{16}{25}-\frac{15}{25}=\frac{1}{25}>0$。事实上 $S_1=\frac{61}{125}\cdot\frac{2}{3}=\frac{122}{375}$ 而 $S_2=\frac{9}{25}\cdot\frac{8}{9}=\frac{120}{375}$，所以**策略 1 较佳**。

在 (p,q) 平面上，整个情况一目了然。令 $D=0$ 便得出边界曲线 $q=1-\dfrac{(1-p)^2}{2-p}$。可以验证

$$\left(p+\tfrac{1}{2}\right)-\left(1-\frac{(1-p)^2}{2-p}\right)=\frac{p}{2(2-p)}>0,$$

所以该曲线位于直线 $q=p+\frac{1}{2}$ 稍下方。在曲线之上，策略 1 最佳；在曲线与对角线 $q=p$ 之间，策略 2 最佳。



图中两点就是上面的两个例子：$(\frac{1}{2},\frac{3}{4})$ 位于策略 2 的区域内，而 $(\frac{1}{5},\frac{2}{3})$ 则刚好在曲线之上。

验算

在第 (i) 部分，令 $q\to 1$：对 Quentin 必胜，所以只有对 Pardeep 的对局才有影响，而这类对局显然愈多愈好，这与差值 $pq(q-p)\to p(1-p)>0$ 吻合。至于第 (ii) 部分，我们随机抽取了 2000 对满足 $p<q$ 的数值，以分数精确地列举全部 2^4 种结果，确认了 S_1, S_2, S_3 的公式和各个因式分解。以 $p=\frac{1}{2}$、$q=\frac{3}{4}$ 进行模拟（simulation），得出 $S_1\approx 0.657$ 和 $S_2\approx 0.703$，而精确值为 $\frac{21}{32}=0.65625$ 和 $\frac{45}{64}\approx 0.7031$。

常见错误

- 把“至少胜一局”的概率算成 mp 而不是 $1-(1-p)^m$；把互相重叠的事件的概率相加，会重复计算。
- 忘记策略 3：要证明策略 1 最佳，必须同时胜过另外两种策略。
- 举出 $q-p<\frac{1}{2}$ 的例子，却没有检查条件 $0<p<q<1$。

STEP 之外

直觉告诉我们“把力气花在自己最弱的地方”，但第二部分显示这并非永远正确：多对 Pardeep 一局的效益会递减（diminishing returns），而当 p 不太小时，多对 Quentin 一局更为划算。把固定数目的试验分配到各处，使若干个“至少成功一次”的概率之积达到最大，正是可靠性理论（reliability theory）和实验设计（design of experiments）背后那类最优化（optimisation）问题的一个小例子。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门教材）仔细讲解了独立性和补集，并有许多同类风格的题目。

详细解答

STEP III 2010，第 1 题

纯数学 · 平均值与方差

> **STEP III 2010，第 1 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s3-q1

关键思路

A 和 B 是首 n 个数的平均值（mean）和方差（variance）；C 和 D 则是再加入一个数之后的平均值和方差。方差最容易以“平方的平均值减去平均值的平方”这个形式处理，而这正是第 (ii) 部分给我们的结果。有了这个形式，D 便成为 x_{n+1} 的二次式（quadratic），而两条不等式（inequality）都是关于这个二次式的陈述。

解答

为简洁起见，全文记 $x = x_{n+1}$。

第 (i) 部分

按定义 $\sum_{k=1}^{n} x_k = nA$，所以

$$C = \frac{1}{n+1}\left(\sum_{k=1}^{n} x_k + x\right) = \frac{nA + x}{n+1}.$$

第 (ii) 部分

把平方展开，并利用 $\sum_{k=1}^{n} x_k = nA$ 及 $\sum_{k=1}^{n} 1 = n$：

$$\begin{aligned} B &= \frac{1}{n}\sum_{k=1}^{n}\left(x_k^2 - 2Ax_k + A^2\right) = \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \frac{2A}{n}\cdot nA + \frac{1}{n}\cdot nA^2 \\ &= \frac{1}{n}\sum_{k=1}^{n} x_k^2 - A^2, \end{aligned}$$

此即所求。

第 (iii) 部分

第 (ii) 部分对任何一组数都成立，所以把它应用于 $n+1$ 个数 $x_1,\ldots,x_{n+1}$（其平均值为 C），得

$$D=\frac{1}{n+1}\sum_{k=1}^{n+1}x_k^2-C^2.$$

再由第 (ii) 部分，$\sum_{k=1}^{n}x_k^2=n(B+A^2)$，故 $\sum_{k=1}^{n+1}x_k^2=n(B+A^2)+x^2$。以第 (i) 部分代入 C，

$$D=\frac{nB+nA^2+x^2}{n+1}-\frac{(nA+x)^2}{(n+1)^2}.$$

两边乘以 $(n+1)^2$ 并展开：

$$\begin{aligned}(n+1)^2D&=(n+1)\left(nB+nA^2+x^2\right)-\left(n^2A^2+2nAx+x^2\right)\\&=n(n+1)B+(n^2+n-n^2)A^2-2nAx+(n+1-1)x^2\\&=n(n+1)B+n\left(x^2-2Ax+A^2\right)\\&=n(n+1)B+n(x-A)^2.\end{aligned}$$

因此

$$\boxed{D=\frac{n}{n+1}B+\frac{n}{(n+1)^2}\left(x_{n+1}-A\right)^2.}$$

第一条不等式。 乘以 $n+1$，

$$(n+1)D=nB+\frac{n}{n+1}\left(x_{n+1}-A\right)^2\ \geq\ nB,$$

因为最后一项是一个平方的非负倍数。这对 x_{n+1} 的任何值都成立，而等号成立当且仅当 $x_{n+1}=A$。

? **想一想** 怎样才能把 $D<B$ 转化为关于 x_{n+1} 的条件，而又不必担心负数的平方根？

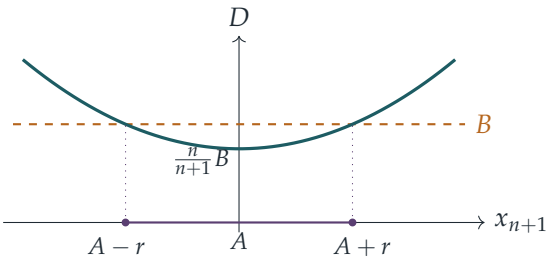
第二条不等式。 首先留意 $B\geq 0$，因为它是一些平方的平均值，所以 $\sqrt{(n+1)B/n}$ 是实数。现在利用方框中的公式，以下每一步都是可逆的（我们只是在两边减去同一个量，以

及乘以正数 $(n+1)^2/n$）：

$$\begin{aligned} D < B &\iff \frac{n}{(n+1)^2}(x_{n+1}-A)^2 < B - \frac{n}{n+1}B = \frac{B}{n+1} \\ &\iff (x_{n+1}-A)^2 < \frac{(n+1)B}{n} \\ &\iff |x_{n+1}-A| < \sqrt{\frac{(n+1)B}{n}} \\ &\iff A - \sqrt{\frac{(n+1)B}{n}} < x_{n+1} < A + \sqrt{\frac{(n+1)B}{n}}. \end{aligned}$$

第三步是在两边取非负平方根；对非负数之间的严格不等式而言，这样做会保持不等式成立。这就是所求的“当且仅当”。（若 $B = 0$，则该区间是空集，而事实上此时恒有 $D \geq 0 = B$。）

下图显示当 $n = 3$、$A = 0$、$B = 1$ 时，D 作为 x_{n+1} 的函数的图像：这是一条抛物线，在 $x_{n+1} = A$ 处取得最小值 $\frac{n}{n+1}B$，并恰好在上面求得的区间上低于 B 这个水平。



其中 $r = \sqrt{(n+1)B/n}$。

验算

- 取 $n = 1$：此时 $B = 0$，而公式给出 $D = \frac{1}{4}(x_2 - x_1)^2$；对于两个数 x_1, x_2 而言这是正确的，因为它们的平均值就是它们的中点。
- 取 $x_1, x_2, x_3 = 0, 0, 3$，则 $A = 1$、$B = 2$，再加入 $x_4 = 1 = A$。直接计算得 $D = \frac{1}{4}(1+1+4+0) = \frac{3}{2} = \frac{3}{4}B$，正是最小值。
- 我们在 20 000 组随机数列上以分数精确验证了全部三个结果，包括等价关系 $D < B \iff (x_{n+1}-A)^2 < (n+1)B/n$，并在区间两端点皆为有理数的 216 个情况中，确认了在两端点处 $D = B$。

常见错误

- 在第 (ii) 部分展开时，忘记 $\sum_{k=1}^{n} A^2 = nA^2$，而不是 A^2。

- 以 $\sum x_k^2$ 表示 D 后便停下来：题目要求以 B、A、x_{n+1} 和 n 表示，所以必须利用第 (ii) 部分把这个和替换掉。
- 对 $(x_{n+1}-A)^2 < k$ 开平方时只得出 $x_{n+1}-A < \sqrt{k}$，漏掉了区间的下端。

STEP 之外

D 的公式是一条**更新公式**（updating formula）：它让电脑可以逐个数值地维持一个数据流的平均值和方差，而无须储存全部数据；这个想法在 Welford 算法（Welford's algorithm）中得到进一步改良。$n(x_{n+1}-A)^2/(n+1)^2$ 这一项，是把方差分解为“组内”和“组间”两部分的一个小例子，而这种分解正是统计学中方差分析（analysis of variance）的基础。如欲深入了解，可参阅 Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）。

详细解答

STEP III 2010，第 2 题

纯数学 · 定积分

STEP III 2010，第 2 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q2

关键思路

这里每个分母都是一个二次式（quadratic）（以 x 或 x^2 为变量），而其根原来正是 $\pm e^{a}$ 和 $\pm e^{-a}$。把双曲函数（hyperbolic function）系数写成指数形式，因式分解（factorisation）便一目了然，因为 e^{a} 与 e^{-a} 之积为 1，之和为 $2\cosh a$（或之差为 $2\sinh a$）。之后，部分分式（partial fractions）把一切化为对数或反正切（arctangent），而真正的功夫在于把答案整理成双曲函数的形式。

解答

第 (i) 部分

按定义，

$$\cosh a = \tfrac{1}{2}\left(e^{a} + e^{-a}\right).$$

因此 $2\cosh a = e^{a} + e^{-a}$ 且 $e^{a}\cdot e^{-a} = 1$，即

$$x^2 + 2x\cosh a + 1 = (x + e^{a})(x + e^{-a}).$$

当 $x \geq 0$ 时两个因式皆为正，所以被积函数（integrand）在 $[0,1]$ 上连续。由于 $a > 0$，两根 $-e^{a}$ 与 $-e^{-a}$ 相异，部分分式为

$$\frac{1}{(x+e^{-a})(x+e^{a})} = \frac{1}{e^{a}-e^{-a}}\left(\frac{1}{x+e^{-a}} - \frac{1}{x+e^{a}}\right) = \frac{1}{2\sinh a}\left(\frac{1}{x+e^{-a}} - \frac{1}{x+e^{a}}\right),$$

你可以把它重新合并来验证：分子变成 $(x+e^a)-(x+e^{-a})=e^a-e^{-a}$。积分得

$$\int_0^1 \frac{\mathrm{d}x}{x^2+2x\cosh a+1} = \frac{1}{2\sinh a}\Big[\ln\frac{x+e^{-a}}{x+e^a}\Big]_0^1 = \frac{1}{2\sinh a}\left(\ln\frac{1+e^{-a}}{1+e^a}-\ln\frac{e^{-a}}{e^a}\right).$$

第一个对数可以漂亮地化简：$1+e^a = e^a(e^{-a}+1)$，所以 $\dfrac{1+e^{-a}}{1+e^a} = e^{-a}$。第二个是 $\ln e^{-2a} = -2a$。因此

$$\int_0^1 \frac{\mathrm{d}x}{x^2+2x\cosh a+1} = \frac{1}{2\sinh a}(-a+2a) = \frac{a}{2\sinh a},$$

这正是所求的结果。

第 (ii) 部分

第一个积分。 现在系数是 $2\sinh a = e^a - e^{-a}$，常数项是 $-1 = e^a\cdot(-e^{-a})$，所以

$$x^2+2x\sinh a-1 = (x+e^a)(x-e^{-a}).$$

> **想一想** 被积函数在整个范围 $[1,\infty)$ 上都表现良好吗?

两根为 $-e^a<0$ 和 e^{-a}，而由于 $a>0$，有 $e^{-a}<1$。所以当 $x\ge 1$ 时两个因式皆为正，被积函数连续。当 x 很大时，它的表现类似 $1/x^2$，所以这个瑕积分（improper integral）收敛（converges）。部分分式的做法同前，现在用 $e^a-(-e^{-a})=2\cosh a$：

$$\frac{1}{(x-e^{-a})(x+e^a)} = \frac{1}{2\cosh a}\left(\frac{1}{x-e^{-a}}-\frac{1}{x+e^a}\right).$$

由 1 积分至 X，再令 $X\to\infty$，

$$\int_1^X \frac{\mathrm{d}x}{x^2+2x\sinh a-1} = \frac{1}{2\cosh a}\left(\ln\frac{X-e^{-a}}{X+e^a}-\ln\frac{1-e^{-a}}{1+e^a}\right).$$

当 $X\to\infty$ 时，分数 $\frac{X-e^{-a}}{X+e^a}\to 1$，所以第一个对数趋向 0。至于第二个，从分子抽出 e^{-a}：

$$\frac{1-e^{-a}}{1+e^a} = e^{-a}\,\frac{e^a-1}{e^a+1} = e^{-a}\tanh\tfrac{a}{2},$$

这里用了 $\dfrac{e^a-1}{e^a+1} = \dfrac{e^{a/2}-e^{-a/2}}{e^{a/2}+e^{-a/2}} = \tanh\frac{a}{2}$（分子分母同除以 $e^{a/2}$）。所以第二个对数是 $-a+\ln\tanh\frac{a}{2}$，而

$$\int_1^\infty \frac{\mathrm{d}x}{x^2+2x\sinh a-1} = \frac{a-\ln\tanh\frac{a}{2}}{2\cosh a} = \frac{a+\ln\coth\frac{a}{2}}{2\cosh a}.$$

分子中两项皆为正（因为 $0<\tanh\frac{a}{2}<1$），这正符合被积函数为正的要求。

第二个积分。 这是第 (i) 部分的二次式，只是把 x 换成 x^2：

$$x^4 + 2x^2\cosh a + 1 = (x^2 + e^a)(x^2 + e^{-a}),$$

对所有实数 x，两个因式皆为正，所以被积函数连续；而它在无穷远处是 $O(1/x^4)$，所以积分收敛。把 x^2 当作变量，套用同样的部分分式，

$$\frac{1}{(x^2+e^{-a})(x^2+e^a)} = \frac{1}{2\sinh a}\left(\frac{1}{x^2+e^{-a}} - \frac{1}{x^2+e^a}\right).$$

每一部分都是标准的反正切积分：对 $c > 0$，

$$\int_0^\infty \frac{\mathrm{d}x}{x^2+c^2} = \left[\frac{1}{c}\arctan\frac{x}{c}\right]_0^\infty = \frac{\pi}{2c}.$$

分别取 $c = e^{-a/2}$ 和 $c = e^{a/2}$（每一部分各自收敛，所以可以把积分拆开），

$$\int_0^\infty \frac{\mathrm{d}x}{x^4+2x^2\cosh a+1} = \frac{1}{2\sinh a}\cdot\frac{\pi}{2}\left(e^{a/2} - e^{-a/2}\right) = \frac{\pi}{2}\cdot\frac{2\sinh\frac{a}{2}}{2\sinh a}.$$

最后，二倍角公式（double-angle formula）$\sinh a = 2\sinh\frac{a}{2}\cosh\frac{a}{2}$ 把 $\sinh\frac{a}{2}$ 约去：

$$\int_0^\infty \frac{\mathrm{d}x}{x^4+2x^2\cosh a+1} = \frac{\pi}{4\cosh\frac{a}{2}}.$$

想一想 有没有办法预先看出 $\cosh\frac{a}{2}$ 会出现?

有。设该积分为 I。代换（substitution）$x \mapsto 1/x$ 显示同样有 $I = \int_0^\infty \frac{x^2}{x^4+2x^2\cosh a+1}\,\mathrm{d}x$，所以

$$2I = \int_0^\infty \frac{1+1/x^2}{x^2+1/x^2+2\cosh a}\,\mathrm{d}x.$$

令 $u = x - 1/x$，则 $\mathrm{d}u = (1+1/x^2)\,\mathrm{d}x$ 且 $x^2+1/x^2 = u^2+2$，而 u 由 $-\infty$ 走到 ∞。分母变成 $u^2+2+2\cosh a = u^2+4\cosh^2\frac{a}{2}$，所以 $2I = \pi/(2\cosh\frac{a}{2})$，答案相同。

验算

- 当 $a \to 0$ 时，第 (i) 部分变成 $\int_0^1 (1+x)^{-2}\,\mathrm{d}x = \frac{1}{2}$，而 $a/(2\sinh a) \to \frac{1}{2}$。最后一个积分变成 $\int_0^\infty (1+x^2)^{-2}\,\mathrm{d}x = \frac{\pi}{4}$，与 $\pi/(4\cosh 0)$ 吻合。
- 数值上，取 $a = 1$，辛普森法则（Simpson's rule）给出第 (i) 部分为 0.4254590641，(ii) 的第一个积分为 0.5741556186，第二个为 0.6965059227，与 $\frac{1}{2\sinh 1}$、$\frac{1-\ln\tanh\frac{1}{2}}{2\cosh 1}$ 及 $\frac{\pi}{4\cosh\frac{1}{2}}$ 在小数点后十位一致。

常见错误

- 把部分分式的常数弄错：它是 $1/(\text{两根之差})$，在第 (i) 部分是 $1/(2\sinh a)$，但在 (ii) 的第一个积分中是 $1/(2\cosh a)$。
- 忘记验证 $e^{-a} < 1$，即分母的根位于 $[1, \infty)$ 以外。
- 把答案留在指数形式：题目要求用双曲函数表示，所以 $\frac{e^a-1}{e^a+1}$ 应写成 $\tanh\frac{a}{2}$。

STEP 之外

在大学里，(ii) 的第二个积分通常用围道积分（contour integration）来做：被积函数在 $\pm ie^{\pm a/2}$ 处有单极点（simple poles），把上半平面（upper half-plane）内的留数（residues）加起来，几行便得出 $\pi/(4\cosh\frac{a}{2})$。同一方法也能处理 $\int_0^\infty \mathrm{d}x/(1+x^n)$，以及许多用实分析方法难以攻克的积分。Needham 的 *Visual Complex Analysis*（《直观复分析》）从几何角度阐述留数及其背后的道理。

详细解答

STEP III 2010，第 3 题

纯数学 · 单位根

> **STEP III 2010，第 3 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/10-s3-q3

本题在 *Beyond STEP* 第 6.1 节以单位根为根的多项式中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch06/cyclotomic.html

关键思路

每个单位根 a 都有唯一的阶（order）：即使 $a^m = 1$ 成立的最小正整数 m。所谓本原 n 次单位根（primitive nth root of unity），就是阶恰好为 n 的单位根。因此，要求 $\mathrm{C}_n(x)$，我们先列出所有 n 次单位根 $\mathrm{e}^{2\pi \mathrm{i}k/n}$，剔除那些在较低次幂时已经回到 1 的根，再把余下的 $(x-a)$ 全部相乘。而每个数只有一个阶这一事实，也正是解决最后一部分的关键。

解答

以下各部分中，n 次单位根为 ω^k，其中 $k = 0, 1, \ldots, n-1$，$\omega = \mathrm{e}^{2\pi \mathrm{i}/n}$；并且 $(\omega^k)^m = 1$ 当且仅当 n 整除 km。由于 C_n 是首一（monic）多项式，没有重根，而且其根恰好是所有本原 n 次单位根，故

$$\mathrm{C}_n(x) = \prod_{a \text{ 为本原单位根}} (x-a).$$

4 次单位根

4 次单位根为 $1, \mathrm{i}, -1, -\mathrm{i}$。由于 $1^1 = 1$ 及 $(-1)^2 = 1$，这两个根都不是本原的。但 $\mathrm{i}^1 = \mathrm{i}$、$\mathrm{i}^2 = -1$、$\mathrm{i}^3 = -\mathrm{i}$，全部不等于 1；$-\mathrm{i}$ 的情况亦相同（它的各次幂是前者的共轭（conjugate））。所以本原 4 次单位根为 i 和 $-\mathrm{i}$，从而

$$\mathrm{C}_4(x) = (x-\mathrm{i})(x+\mathrm{i}) = x^2 + 1,$$

与题目所求相符。

第 (i) 部分

$n = 1$。唯一的 1 次单位根是 1，而不存在满足 $0 < m < 1$ 的 m，所以 1 是本原的：$C_1(x) = x - 1$。

$n = 2$。2 次单位根为 ± 1；由于 $1^1 = 1$，1 被剔除，而 $(-1)^1 \neq 1$。所以 $C_2(x) = x + 1$。

$n = 3$ **及** $n = 5$。当 n 为质数（prime）时，对于满足 $1 \leq k \leq n-1$ 的根 ω^k，$(\omega^k)^m = 1$ 只在 $n \mid km$ 时成立；又因 n 是质数且不整除 k，这迫使 $n \mid m$。所以除 1 以外的每个 n 次单位根都是本原的。它们正是

$$\frac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + \cdots + x + 1,$$

的根，而这个多项式是首一的，且各根互不相同。因此

$$C_3(x) = x^2 + x + 1, \qquad C_5(x) = x^4 + x^3 + x^2 + x + 1.$$

$n = 6$。设 $\omega = e^{\pi i/3}$，则 ω^k 的阶为 $6/\gcd(k,6)$。所以 $\omega^0 = 1$ 的阶为 1，$\omega^3 = -1$ 的阶为 2，ω^2, ω^4 的阶为 3，只有 ω 和 $\omega^5 = \overline{\omega}$ 的阶为 6。两者之和为 $2\cos\frac{\pi}{3} = 1$，之积为 $|\omega|^2 = 1$，所以

$$C_6(x) = x^2 - x + 1.$$

（作为交叉验证：每个 6 次单位根的阶为 1、2、3 或 6，所以 $x^6 - 1 = C_1C_2C_3C_6$；而事实上 $(x-1)(x+1)(x^2+x+1)(x^2-x+1) = (x^3-1)(x^3+1) = x^6 - 1$。）

第 (ii) 部分

设 a 为 $x^4 + 1 = 0$ 的任一根。则 $a^4 = -1$，所以 $a^8 = 1$。a 的阶 d 整除 8（写 $8 = qd + r$，其中 $0 \leq r < d$；则 $a^r = a^8(a^d)^{-q} = 1$，由 d 的最小性得 $r = 0$）。所以 $d \in \{1, 2, 4, 8\}$。若 d 为 1、2 或 4，则 $a^4 = 1$，与 $a^4 = -1$ 矛盾。因此 $x^4 + 1$ 的每个根都是本原 8 次单位根。

反过来，设 $\omega = e^{\pi i/4}$，则 ω^k 的阶为 8 当且仅当 k 为奇数（若 k 为偶数，则 $(\omega^k)^4 = 1$），所以本原 8 次单位根恰好有四个：$\omega, \omega^3, \omega^5, \omega^7$。四次多项式 $x^4 + 1$ 有四个互不相同的根（即 $e^{(2j+1)\pi i/4}$，$j = 0, 1, 2, 3$），全部都是本原 8 次单位根，所以它们正正就是这四个根。因此 $C_8(x) = x^4 + 1$。

其他 n 均不成立：C_n 的根的阶全部是 n，而 $x^4 + 1$ 的根的阶是 8。所以答案是 $n = 8$。

第 (iii) 部分

第 (i) 部分中用于 $n=3$ 及 $n=5$ 的论证，适用于任何质数 p：除 1 以外的每个 p 次单位根都是本原的。因此

$$\mathrm{C}_p(x)=\frac{x^p-1}{x-1}=x^{p-1}+x^{p-2}+\cdots+x+1.$$

（当 $p=2$ 时，此式给出 $x+1$，与第 (i) 部分一致。）

第 (iv) 部分

用反证法（proof by contradiction），假设对某些正整数 q,r,s，有 $\mathrm{C}_q(x)\equiv\mathrm{C}_r(x)\mathrm{C}_s(x)$。

首先，C_s 至少有一个根，因为 $\mathrm{e}^{2\pi\mathrm{i}/s}$ 是本原 s 次单位根：$(\mathrm{e}^{2\pi\mathrm{i}/s})^m=\mathrm{e}^{2\pi\mathrm{i}m/s}=1$ 只在 $s\mid m$ 时成立。C_r 亦同理。所以 C_r 和 C_s 的次数（degree）都至少为 1。

取 C_r 的任一根 a。由恒等式，$\mathrm{C}_q(a)=\mathrm{C}_r(a)\mathrm{C}_s(a)=0$，所以 a 既是本原 q 次单位根，也是本原 r 次单位根。

> ? **想一想** 同一个数能否在 $r\neq q$ 的情况下，既是本原 r 次单位根，又是本原 q 次单位根?

不能。假如 $r<q$，则 $a^r=1$ 且 $0<r<q$，这与 a 是本原 q 次单位根矛盾。若 $q<r$，则把两者的角色对调即可。所以 $r=q$；同理（利用 C_s 的一个根）可得 $s=q$。

但这样一来，恒等式就变成 $\mathrm{C}_q=\mathrm{C}_q^2$。比较次数，得 $\deg\mathrm{C}_q=2\deg\mathrm{C}_q$，所以 $\deg\mathrm{C}_q=0$，与 C_q 有根这一事实矛盾。（另一做法：C_q^2 有重根，而 C_q 没有。）因此不存在这样的 q、r、s。

验算

- C_n 的次数等于在 $1\le k\le n$ 中满足 $\gcd(k,n)=1$ 的 k 的个数：当 $n=1,\ldots,6$ 时，次数为 $1,1,2,2,4,2$；当 $n=8$ 时为 4，全部吻合。
- 每个 n 次单位根都有某个整除 n 的阶 d，所以 x^n-1 等于 C_d 对 n 的所有因数 d 之积。当 $n=4$：$(x-1)(x+1)(x^2+1)=x^4-1$；当 $n=8$：$(x^4-1)(x^4+1)=x^8-1$。
- 以数值方法，对 $n\le 40$ 把各本原根的 $(x-a)$ 相乘，可重现上述每一个多项式，只有在 $n=8$ 时得到 x^4+1，并且在 $q,r,s\le 30$ 的范围内找不到任何恒等式 $\mathrm{C}_q=\mathrm{C}_r\mathrm{C}_s$。

常见错误

- 把 1（或 -1）算作本原根：C_n 并不是 x^n-1。

- 在第 (ii) 部分，只验证 x^4+1 整除 x^8-1，却没有排除较小的阶：重点在于 $a^4=-1\neq 1$。
- 在第 (iv) 部分，单凭次数论证：$\deg \mathrm{C}_q=\deg \mathrm{C}_r+\deg \mathrm{C}_s$ 是可能成立的（例如 $4=2+2$），所以导出矛盾的是根，而不是次数。

STEP 之外

这些 C_n 就是分圆多项式（cyclotomic polynomials）。恒等式 $x^n-1=\prod_{d\mid n}\mathrm{C}_d(x)$ 说明它们的系数必定是整数；而一个更深入的定理指出，每个分圆多项式在有理数上都是不可约（irreducible）的。它们在伽罗瓦理论（Galois theory）中占核心地位，亦是高斯（Gauss）发现哪些正多边形可以用尺规作图（ruler and compass construction）作出的关键。Allenby 的 *Rings, Fields and Groups*（《环、域与群》）发展了这方面的代数，而 Needham 的 *Visual Complex Analysis*（《可视化复分析》）则为单位根提供了几何图像。

详细解答

STEP III 2010，第 4 题

纯数学 · 多项式的公共根

STEP III 2010，第 4 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q4

关键思路

若 α 同时满足两条首项系数（leading coefficient）相同的二次方程（quadratic），那么它亦满足两者之差，而 x^2 项互相抵消，只剩下一条线性（linear）方程。这样便能确定 α，再把它代回其中一条二次方程，就得到所求条件。在第 (ii) 部分，只要先把三次多项式（cubic）除以二次多项式，同一技巧便再次奏效：在二次方程的根上，三次多项式的值与其余式（remainder）完全相同，于是问题又回到两条二次方程，可直接引用第 (i) 部分的结果。

解答

全文中，“根”指复数范围内的根，因此每条二次方程都有两个根（按重数（multiplicity）计算）。

第 (i) 部分

设 α 同时满足两条方程。以第一条减去第二条，得

$$(a-c)\alpha + (b-d) = 0.$$

由于 $a \neq c$，可除以 $a-c$，得

$$\alpha = -\frac{b-d}{a-c},$$

即为所求。

条件。记 $k = a - c \neq 0$ 及 $m = b - d$，则条件为 $m^2 - amk + bk^2 = 0$。

($\Rightarrow$)若存在公共根(common root),由第一个结果,它必为 $\alpha = -m/k$。代入 $\alpha^2 + a\alpha + b = 0$,

$$\frac{m^2}{k^2} - \frac{am}{k} + b = 0,$$

再乘以 $k^2 \neq 0$，得 $m^2 - amk + bk^2 = 0$。

($\Leftarrow$) 设 $m^2 - amk + bk^2 = 0$，并定义 $\alpha = -m/k$。把条件除以 k^2，再把上一步倒过来看，便知 $\alpha^2 + a\alpha + b = 0$。另外，按 α 的取法，$(a-c)\alpha + (b-d) = k\alpha + m = 0$。因此

$$\alpha^2 + c\alpha + d = (\alpha^2 + a\alpha + b) - \big((a-c)\alpha + (b-d)\big) = 0 - 0 = 0,$$

所以 α 是公共根。这就证明了在 $a \neq c$ 时两者等价。

> ? **想一想** 为甚么单靠正方向并不足够?

正方向只说明如果存在公共根，条件便成立。要证明“当且仅当”(if and only if)，我们还必须证明条件能产生一个公共根。正因如此，我们才由条件定义 α，并验证它同时满足两条方程。

去掉 $a \neq c$。设 $a = c$。此时条件变为 $(b-d)^2 = 0$，即 $b = d$。

- 若 $b = d$，两条方程完全相同，自然有公共根（它们的两个根都是公共根），而条件亦成立。
- 若 $b \neq d$，假如有公共根 α，相减便得 $(a-c)\alpha + (b-d) = b - d = 0$，矛盾。所以没有公共根，而条件亦不成立。

两种情况下，“有公共根”与“条件成立”都一致，所以**是的，即使没有 $a \neq c$ 这个条件，结果仍然成立**。

(唯一微妙之处在于只考虑实根的情况。若 $a = c$、$b = d$ 且 $a^2 < 4b$，条件成立，但两条相同的方程却没有实根。当 $a \neq c$ 时则没有这个问题，因为对实系数而言，公共根 $-(b-d)/(a-c)$ 必然是实数。)

第 (ii) 部分

把三次多项式除以 $x^2 + ax + b$。由于 $(x+1)(x^2 + ax + b) = x^3 + (a+1)x^2 + (a+b)x + b$,

$$x^3 + (a+1)x^2 + qx + r = (x+1)(x^2 + ax + b) + (q - a - b)x + (r - b).$$

因此若在 $x = \alpha$ 处 $x^2 + ax + b = 0$，则三次多项式在该处的值为 $(q-a-b)\alpha + (r-b)$，而二次多项式

$$(x^2 + ax + b) + (q - a - b)x + (r - b) = x^2 + (q-b)x + r.$$

在该处的值亦然。所以 α 是该二次方程与三次方程的公共根，当且仅当它是

$$x^2+ax+b=0 \qquad 与 \qquad x^2+(q-b)x+r=0.$$

的公共根。现在以 $c=q-b$ 及 $d=r$ 应用第 (i) 部分。此时 $a-c=a+b-q$ 及 $b-d=b-r$，所以存在公共根当且仅当

$$(b-r)^2-a(b-r)(a+b-q)+b(a+b-q)^2=0.$$

由于无论 $a=c$ 与否，第 (i) 部分的结果都成立，这里不需要额外条件。

例子。把两条方程都除以 2，使它们成为首一（monic）方程：

$$x^2+\tfrac{5}{2}x+b=0, \qquad x^3+\tfrac{7}{2}x^2+\tfrac{5}{2}x+\tfrac{1}{2}=0.$$

它们具备所需形式, 其中 $a=\frac{5}{2}$ (而且确实 $a+1=\frac{7}{2}$)、$q=\frac{5}{2}$ 及 $r=\frac{1}{2}$。于是 $a+b-q=b$ 及 $b-r=b-\frac{1}{2}$，条件变为

$$\left(b-\tfrac{1}{2}\right)^2-\tfrac{5}{2}\left(b-\tfrac{1}{2}\right)b+b\cdot b^2=0.$$

乘以 4:

$$\begin{aligned}(2b-1)^2-5b(2b-1)+4b^3&=0\\ 4b^3-6b^2+b+1&=0.\end{aligned}$$

观察可知 $b=1$ 是一个根，而

$$4b^3-6b^2+b+1=(b-1)(4b^2-2b-1).$$

二次因式给出 $b=\dfrac{2\pm\sqrt{4+16}}{8}=\dfrac{1\pm\sqrt{5}}{4}$。所以两条方程有公共根，当且仅当

$$b=1 \qquad 或 \qquad b=\frac{1+\sqrt{5}}{4} \qquad 或 \qquad b=\frac{1-\sqrt{5}}{4}.$$

另一方法。这里的三次方程不含 b, 所以可以直接解出。试 $x=-\frac{1}{2}$, 得 $-\frac{1}{4}+\frac{7}{4}-\frac{5}{2}+1=0$, 所以

$$2x^3+7x^2+5x+1=(2x+1)(x^2+3x+1),$$

其根为 $-\frac{1}{2}$ 及 $\frac{1}{2}(-3\pm\sqrt{5})$。公共根 x 必须满足 $2b=-(2x^2+5x)$。当 $x=-\frac{1}{2}$ 时, 得 $b=1$。至于另外两个根, $x^2=-3x-1$, 所以 $2x^2+5x=-x-2$, 而 $b=\frac{1}{2}(x+2)=\frac{1}{4}(1\pm\sqrt{5})$。这与上面的答案一致。

验算

- 当 $b = 1$ 时，二次式 $2x^2 + 5x + 2 = (2x + 1)(x + 2)$ 有根 $-\frac{1}{2}$，而它是三次方程的根。
- $b = \frac{1}{4}(1 + \sqrt{5}) \approx 0.8090$：二次方程的根为 ≈ -0.3820 及 -2.1180，而 $-0.3820 \approx \frac{1}{2}(-3 + \sqrt{5})$ 是三次方程的根。
- 以电脑测试 20 000 组随机整数 a, b, c, d（以及 a, b, q, r），结果显示条件式恰好在两条方程有公共复数根时为零，包括 $a = c$ 的情况。

常见错误

- 只证明“有公共根 $\Rightarrow$ 条件成立”，忘记证明逆命题（converse）。
- 回答 $a = c$ 的问题时说“不成立，因为我们除以了 $a - c$”。证明的确失效，但结果仍然成立，因为两边都化为 $b = d$。
- 未先除以 2 便应用第 (ii) 部分：公式要求方程是首一的。

STEP 之外

第 (i) 部分的表达式是两条二次多项式的结式（resultant）：一个以系数为变量的多项式，恰好在两个多项式有公共根时为零。结式源于西尔维斯特（Sylvester）行列式，并且是用来求解多项式方程组的消元理论（elimination theory）的基础。第 (ii) 部分中“相除并保留余式”的步骤，正是多项式的欧几里得算法（Euclidean algorithm）的第一步；这个算法求两个多项式的最高公因式（highest common factor），方法与整数的情况如出一辙。Allenby 的 *Rings, Fields and Groups*（《环、域与群》）在这个框架下探讨多项式除法与最高公因式。

详细解答

STEP III 2010，第 5 题

纯数学 • 坐标几何作图

STEP III 2010，第 5 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q5

关键思路

这题全靠坐标完成：先求出 R 和 S，用截距式（intercept form）写出直线 PQ 和 RS，再看两者各自与“过 A 且垂直（perpendicular）于 AC 的直线”交于何处。它们交于同一点，而该点必然就是 T。至于作图，关键在于一个观察：已知正方形的对角线（diagonal）正好是我们需要的边长，而第一部分正好提供了一个只用直尺（straight-edge）就能作出与该对角线成直角的方法。

解答

证明 $TA \perp AC$

过 $C(a,a)$ 的直线很容易处理。直线 CP 连接 (a,a) 与 $(an,0)$，所以其斜率（gradient）为 $\dfrac{a}{a-an} = \dfrac{1}{1-n}$（由于 $n<1$，斜率有定义），方程为

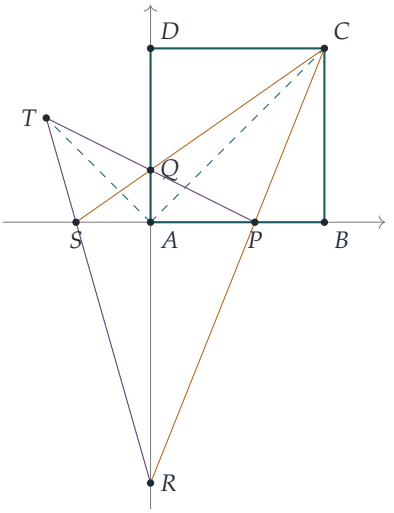
$$y = \frac{x-an}{1-n}.$$

直线 DA 是 $x=0$，所以

$$R = \left(0,\ -\frac{an}{1-n}\right).$$

同理，CQ 连接 (a,a) 与 $(0,am)$，斜率为 $1-m \neq 0$，方程为 $y = am + (1-m)x$。令 $y=0$（即直线 BA），

$$S = \left(-\frac{am}{1-m},\ 0\right).$$



现在用截距式。直线 PQ 的截距为 an 和 am；直线 RS 的截距为 $-\frac{am}{1-m}$ 和 $-\frac{an}{1-n}$。所以

$$\begin{aligned} PQ: &\quad \frac{x}{n}+\frac{y}{m}=a, \\ RS: &\quad -\frac{(1-m)x}{m}-\frac{(1-n)y}{n}=a. \end{aligned}$$

由于 AC 的斜率为 1，过 A 且垂直于 AC 的直线是 $y=-x$。把 $y=-x$ 代入两条方程：

$$\begin{aligned} PQ: &\quad x\Big(\frac{1}{n}-\frac{1}{m}\Big)=a, \\ RS: &\quad x\Big(-\frac{1}{m}+1+\frac{1}{n}-1\Big)=x\Big(\frac{1}{n}-\frac{1}{m}\Big)=a. \end{aligned}$$

两者是同一条方程。由于 $m\neq n$，括号内的值不为零，因此两条直线都与 $y=-x$ 交于唯一一点

$$\Big(\frac{amn}{m-n},-\frac{amn}{m-n}\Big).$$

想一想 我们找到了一个同时在 PQ 和 RS 上的点。为甚么它一定是 T?

两条不同的直线最多只有一个交点。直线 PQ 和 RS 是不同的直线（PQ 经过 x 轴上 $x>0$ 的点 P，而 RS 与 x 轴交于 $x<0$ 的点 S）。所以两者的任何公共点都就是它们的交点 T。因此

$$T=\Big(\frac{amn}{m-n},-\frac{amn}{m-n}\Big),$$

它位于 $y=-x$ 上。向量(vector) $\overrightarrow{AT}$ 是 $(1,-1)$ 的倍数，而 $\overrightarrow{AC}=(a,a)$，且 $(1,-1)\cdot(a,a)=0$。所以 TA 垂直于 AC，得证。

留意整个推导从未用到 $m<n$，只用到 $m\neq n$。推导也从未用到 P 和 Q 的实际位置，所以只要 $AP\neq AQ$，边 AB 上的任意一点 P 与边 AD 上的任意一点 Q 所得的 T，都会落在过 A 且垂直于 AC 的直线上。

作图

面积为 $2a^2$ 的正方形，边长为 $\sqrt{2}\,a$，正好是对角线 AC 的长度。所以我们以 AC 为一边作正方形。用坐标表示，这个正方形的顶点为

$$A(0,0),\quad E(a,-a),\quad F(2a,0),\quad C(a,a).$$

它的各边 $\overrightarrow{AE}=(a,-a)$、$\overrightarrow{EF}=(a,a)$、$\overrightarrow{FC}=(-a,a)$ 和 $\overrightarrow{CA}=(-a,-a)$ 长度都是 $\sqrt{2}\,a$，而且相邻两边互相垂直，所以其面积为 $2a^2$。

> **想一想** 用没有刻度的直尺，我们只能连接两点和延长直线。这样怎可能作出直角？

第一部分正是这样一个方法，因为其中每一步都只是连接两个已知点或求两直线的交点。只用直尺：

1. 画出对角线 AC。在 AB 上任取一点 P，在 AD 上任取一点 Q，使两者与 A 的距离明显不相等（无须量度；我们只需要 $AP\neq AQ$）。
2. 延长 CP 与直线 DA 交于 R，延长 CQ 与直线 BA 交于 S。延长 PQ 和 RS，交于 T。
3. 画出直线 TA。由上面的结果，它垂直于 AC。把 CB 延长越过 B，与直线 TA 交于 E。（直线 CB 是 $x=a$，直线 TA 是 $y=-x$，所以 $E=(a,-a)$。）
4. 在对角 C 重复第 1 至 3 步：在 CD 和 CB 上取与 C 距离明显不相等的点，进行同样的作图。绕正方形中心的半周旋转（half-turn）会互换 A 和 C，并把正方形映射到自身，所以这样得到的是过 C 且垂直于 AC 的直线，即 $x+y=2a$。
5. 把 AB 延长越过 B，与这条直线交于 $F=(2a,0)$。

于是 $AEFC$ 是边长为 $AC=\sqrt{2}\,a$ 的正方形，所以其面积为 $2a^2$。（它的对角线 AF 和 CE 分别位于直线 AB 和 CB 上，相交于新正方形的中心 B。）

验算

- 取 $a=1$、$m=0.3$、$n=0.6$。则 $R=(0,-1.5)$、$S=(-\frac{3}{7},0)$、$T=(-0.6,0.6)$。点 T 在通过 R 和 S 的直线上：$-\frac{0.7}{0.3}(-0.6)-\frac{0.4}{0.6}(0.6)=1.4-0.4=1$。它也在 PQ 上：$-1+2=1$。而且 T 在 $y=-x$ 上。
- 我们用电脑以分数精确计算了 R、S 和 T，对 a、m、n 作了 2000 组随机选择，$m<n$ 和 $m>n$ 两种次序都有涵盖。在每一种情况下，$\overrightarrow{AT}\cdot\overrightarrow{AC}=0$，而且所作出的 $AEFC$ 都有四条相等的边，边长平方为 $2a^2$，并且两两成直角相交。

常见错误

- 硬解 PQ 和 RS 的联立方程。这样做行得通，但很繁复。把一条候选直线（$y = -x$）分别代入两条方程检验，快得多。
- 忘记解释为甚么公共点真的是 T。你需要说明两条直线是不同的。
- 在作图中“作垂线”或“截取某长度”却没有说明怎样做。每一步都必须是连接两点或求两直线的交点。

STEP 之外

只用直尺的作图属于射影几何（*projective geometry*），在其中连接与相交是仅有的运算；而像隐藏在 R、S 和 T 之中的那些比例，则通过交比（cross-ratio）和调和点列（harmonic range）来研究。彭赛列（Poncelet）和施泰纳（Steiner）的一条经典定理指出：只要预先画好一个圆及其圆心，单凭直尺就能完成直尺与圆规所能做的一切作图；在本题中，正方形扮演了类似的角色。考克斯特（Coxeter）的 *Projective Geometry*（《射影几何》）是这个学科一本优美的入门书。

详细解答

STEP III 2010，第 6 题

纯数学 · 三维空间中的旋转

STEP III 2010，第 6 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q6

关键思路

绕一条轴的旋转（rotation）会令轴上每一点保持不动，而对垂直于该轴的平面内的一切，则完全如同该平面上的一个旋转般转动。因此每一步都可化为一个二维旋转：找出被转动的平面，写下该平面内一对标准正交（orthonormal）向量，然后把它们旋转。任何沿轴方向的向量都只会原封不动，而单是这一点观察，便已替 Q 的计算完成了大部分工作。

解答

以 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 表示沿各坐标轴的单位向量（unit vectors），于是 P_0、Q_0、R_0 的位置向量（position vectors）分别为 $\mathbf{i}$、$\mathbf{j}$、$\mathbf{k}$。

第 (i) 部分

绕 z 轴旋转会把 x-y 平面转过 ϕ，而 $\mathbf{i}$ 朝 $\mathbf{j}$ 的方向移动。因此

$$P_1 = (\cos\phi,\ \sin\phi,\ 0).$$

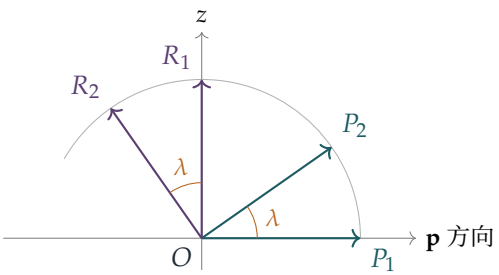
为了稍后之用：同一个旋转把 $\mathbf{j}$ 送到 $Q_1 = (-\sin\phi, \cos\phi, 0)$，因为在该平面内 Q 始终领先 P 一个直角；而 R_0 位于轴上，故保持不动：$R_1 = (0,0,1)$。

第 (ii) 部分

设 $\mathbf{p} = \overrightarrow{OP_1} = (\cos\phi, \sin\phi, 0)$。在 x-y 平面内垂直于 OP_1 的直线，其方向为

$$\mathbf{u} = (-\sin\phi,\ \cos\phi,\ 0),$$

因为 $\mathbf{u}\cdot\mathbf{p}=0$ 及 $\mathbf{u}\cdot\mathbf{k}=0$。垂直于这条轴的平面，就是由标准正交向量对 $\mathbf{p}$ 和 $\mathbf{k}$ 所张成的平面，而第二次旋转把这个平面转过 λ，其中 $\mathbf{p}$ 朝 $\mathbf{k}$ 的方向移动：



在该平面内，这个旋转把

$$\mathbf{p}\mapsto\cos\lambda\,\mathbf{p}+\sin\lambda\,\mathbf{k},\qquad \mathbf{k}\mapsto-\sin\lambda\,\mathbf{p}+\cos\lambda\,\mathbf{k},$$

其中第二式成立，是因为 $\mathbf{k}$ 始终领先 $\mathbf{p}$ 一个直角。因此

$$P_2=(\cos\lambda\cos\phi,\ \cos\lambda\sin\phi,\ \sin\lambda),\qquad R_2=(-\sin\lambda\cos\phi,\ -\sin\lambda\sin\phi,\ \cos\lambda).$$

想一想 在第二次旋转中，Q 会去到哪里?

第一次旋转后，Q 位于 $Q_1=(-\sin\phi,\cos\phi,0)$，这恰好就是 $\mathbf{u}$：它正正位于第二次旋转的轴上，所以不会移动。因此

$$Q_2=(-\sin\phi,\ \cos\phi,\ 0).$$

作为验算，P_2、Q_2、R_2 是两两互相垂直的单位向量（例如 $\overrightarrow{OP_2}\cdot\overrightarrow{OR_2}=-\sin\lambda\cos\lambda\cos^2\phi-\sin\lambda\cos\lambda\sin^2\phi+\sin\lambda\cos\lambda=0$），经过旋转后理应如此。

第 (iii) 部分

在这一部分中，我们一直假设 $P_2\neq\pm P_0$，使 O、P_0、P_2 确实能决定一个平面。

想一想 为甚么旋转角会等于角 P_2OP_0?

由于 P 一直留在平面 OP_0P_2 内，旋转轴必定是经过 O 并垂直于该平面的直线（旋转会令一点在经过该点而垂直于轴的平面内移动）。所以 P 沿着这个平面上的大圆（great circle）移动，而平面内每个向量都转过旋转角。因此该角就是 $\overrightarrow{OP_2}$ 与 $\overrightarrow{OP_0}$ 之间的夹角；又由于两者皆为单位向量，

$$\cos\theta=\overrightarrow{OP_2}\cdot\overrightarrow{OP_0}=(\cos\lambda\cos\phi)(1)+(\cos\lambda\sin\phi)(0)+(\sin\lambda)(0)=\cos\phi\cos\lambda,$$

此即所求。（若绕远路转过 $2\pi-\theta$，所得的余弦值亦相同。）

旋转轴同时垂直于 $\overrightarrow{OP_2}$ 和 $\overrightarrow{OP_0}$，所以它沿着两者的向量积（cross product）方向：

$$\overrightarrow{OP_2} \times \overrightarrow{OP_0} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos\lambda\cos\phi & \cos\lambda\sin\phi & \sin\lambda \\ 1 & 0 & 0 \end{vmatrix} = (0,\ \sin\lambda,\ -\cos\lambda\sin\phi).$$

所以沿轴方向的一个向量为

$$\begin{pmatrix} 0 \\ \sin\lambda \\ -\cos\lambda\sin\phi \end{pmatrix}.$$

取这个定向时，按右手定则（right-hand rule）转过 θ 的旋转会把 P_2 带到 P_0。留意这条轴没有 x 分量，这是理所当然的，因为它垂直于 OP_0。

验算

- 若 $\lambda = 0$，第二次旋转毫无作用，所以第三次旋转只是抵消第一次：$\theta = \phi$，而轴沿 $(0, 0, -\sin\phi)$ 方向，即 z 轴。公式与此吻合。
- 若 $\phi = 0$，唯一的旋转是绕 y 轴进行的；而轴确实沿 $(0, \sin\lambda, 0)$ 方向，且 $\theta = \lambda$。
- 以数值验证：对 20 000 组随机的 (ϕ, λ)，用 Rodrigues 公式（Rodrigues' formula）施行各旋转，均能重现 P_2、Q_2、R_2，而绕 $(0, \sin\lambda, -\cos\lambda\sin\phi)$ 转过 θ 的旋转亦会把 P_2 送回 P_0。当 $\phi = 0.7$、$\lambda = 0.4$ 时，$\cos\theta \approx 0.7045$，故 $\theta \approx 0.789$。

常见错误

- 在第二步中旋转了 Q：它位于轴上，所以应保持不动。
- 弄错 R_2 的正负号：当 P 向上朝 $\mathbf{k}$ 倾斜时，R 会向下倾并远离 OP_1，所以它的水平部分是 $-\sin\lambda\,\mathbf{p}$。

STEP 之外

这里的两个角分别是经度（longitude）ϕ 和纬度（latitude）λ，而 $\cos\theta = \cos\phi\cos\lambda$ 正是球面余弦定理（spherical cosine rule）的直角情形，在航海中用来计算大圆距离。更一般地说，两个旋转的合成（composition）仍然是绕单一条轴的旋转（欧拉（Euler）旋转定理），而以 $\overrightarrow{OP_2}$、$\overrightarrow{OQ_2}$、$\overrightarrow{OR_2}$ 为各列所组成的，正是合成旋转的正交矩阵（orthogonal matrix）。Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）讨论了旋转矩阵和正交矩阵；Needham 的 *Visual Complex Analysis*（《直观复分析》）则展示了如何用复数（complex numbers）描述球面的旋转。

详细解答

STEP III 2010，第 7 题

纯数学 · 麦克劳林级数

STEP III 2010，第 7 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q7

关键思路

整个解法的关键，在于摆脱那个麻烦的平方根：只要把 $\sqrt{1-x^2}\,y'$ 写成一个简洁的式子，再微分一次，二阶微分方程（second-order differential equation）便自然得出。之后反复微分，系数会呈现出一个规律，再用归纳法（induction）加以证明。把 $x=0$ 代入一般关系式，便得到麦克劳林系数（Maclaurin coefficients）的递推关系（recurrence）；当 m 为偶数时，这个递推关系会遇上一个等于零的因子，于是级数就此终止。

解答

微分方程

设 $\theta = \arcsin x$，则 $\sin\theta = x$，其中 $|\theta| < \frac{1}{2}\pi$；而当 $|x|<1$ 时，$\cos\theta = \sqrt{1-x^2} > 0$。于是 $\dfrac{\mathrm{d}\theta}{\mathrm{d}x} = \dfrac{1}{\sqrt{1-x^2}}$，且 $y = \cos m\theta$，由链式法则（chain rule）得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{m\sin m\theta}{\sqrt{1-x^2}}, \qquad \text{亦即,} \qquad \sqrt{1-x^2}\,\frac{\mathrm{d}y}{\mathrm{d}x} = -m\sin m\theta.$$

将此式对 x 微分，左边使用乘积法则（product rule）：

$$\sqrt{1-x^2}\,\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - \frac{x}{\sqrt{1-x^2}}\,\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{m^2\cos m\theta}{\sqrt{1-x^2}} = -\frac{m^2 y}{\sqrt{1-x^2}}.$$

两边同乘 $\sqrt{1-x^2}$（它不等于零），得

$$(1-x^2)\,y'' - x\,y' + m^2 y = 0,$$

此即所求。（此处及以下，$y^{(n)}$ 表示 $\mathrm{d}^n y/\mathrm{d}x^n$。）

三阶及四阶方程

将刚证明的方程微分。乘积 $(1-x^2)y''$ 给出 $(1-x^2)y'''-2xy''$，而 xy' 给出 $xy''+y'$，因此

$$(1-x^2)\,y'''-3x\,y''+(m^2-1)\,y'=0.$$

以同样方法再微分一次，

$$(1-x^2)\,y^{(4)}-2x\,y'''-3x\,y'''-3y''+(m^2-1)\,y''=0,$$

故

$$(1-x^2)\,y^{(4)}-5x\,y'''+(m^2-4)\,y''=0.$$

一般关系式

x 的系数依次为 $1,3,5$，常数项依次为 m^2-0^2，m^2-1^2，m^2-2^2，因此我们猜想对每个整数 $n\geq 0$，

$$(1-x^2)\,y^{(n+2)}-(2n+1)\,x\,y^{(n+1)}+(m^2-n^2)\,y^{(n)}=0. \qquad (*)$$

我们对 n 用归纳法证明。$n=0$ 的情况就是最先证明的方程。假设 $(*)$ 对某个 n 成立。将其微分，

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

合并同类项，$x\,y^{(n+2)}$ 的系数为 $-(2n+3)=-\big(2(n+1)+1\big)$，而 $y^{(n+1)}$ 的系数为 $m^2-(n^2+2n+1)=m^2-(n+1)^2$。因此

$$(1-x^2)\,y^{(n+3)}-\big(2(n+1)+1\big)x\,y^{(n+2)}+\big(m^2-(n+1)^2\big)y^{(n+1)}=0,$$

这正是以 $n+1$ 代替 n 的 $(*)$。由归纳法，$(*)$ 对所有 $n\geq 0$ 成立。

麦克劳林级数

当 $x=0$ 时 $\theta=0$，故 $y(0)=\cos 0=1$，且 $y'(0)=-m\sin 0=0$。将 $x=0$ 代入 $(*)$，中间一项便消失：

$$y^{(n+2)}(0)=(n^2-m^2)\,y^{(n)}(0).$$

由于 $y'(0)=0$，所有奇数阶导数在 0 处均为零。至于偶数阶导数，

$$\begin{aligned}y''(0)&=(0-m^2)\cdot 1=-m^2,\\ y^{(4)}(0)&=(2^2-m^2)(-m^2)=m^2(m^2-2^2),\\ y^{(6)}(0)&=(4^2-m^2)\,m^2(m^2-2^2)=-m^2(m^2-2^2)(m^2-4^2),\end{aligned}$$

一般而言 $y^{(2k)}(0) = (-1)^k m^2(m^2-2^2)(m^2-4^2)\cdots\left(m^2-(2k-2)^2\right)$。因此麦克劳林级数（Maclaurin series）为

$$y = 1 - \frac{m^2x^2}{2!} + \frac{m^2(m^2-2^2)\,x^4}{4!} - \frac{m^2(m^2-2^2)(m^2-4^2)\,x^6}{6!} + \cdots,$$

而首三个非零项为 $1 - \dfrac{m^2x^2}{2!} + \dfrac{m^2(m^2-2^2)x^4}{4!}$。（只要 $m \neq 0, \pm 2$，这些项都不为零；对于这些值，级数的项数较少，下一部分会加以解释。）

偶数 m

设 m 为偶数，并记 $|m| = 2p$。因子 $m^2 - (2p)^2$ 等于零，而一旦 $2k - 2 \geq 2p$，它便会出现在 $y^{(2k)}(0)$ 之中。所以阶数大于 $|m|$ 的每个导数在 $x = 0$ 处都等于零，麦克劳林级数在 $x^{|m|}$ 项便终止。

想一想 一个终止的级数只告诉我们函数在 $x = 0$ 这一点的各阶导数。我们怎样知道对每个 $|x| < 1$，y 真的等于这个多项式（polynomial）？

我们再次利用 $(*)$，这次是在整个区间上。取 $n = |m|$，使 $m^2 - n^2 = 0$，并设 $v = y^{(n+1)}$。于是 $(*)$ 变成

$$(1-x^2)\,v' - (2n+1)\,x\,v = 0 \qquad (|x| < 1).$$

这个一阶方程的积分因子（integrating factor）为 $(1-x^2)^{(2n-1)/2}$；事实上

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)^{(2n+1)/2}\,v\right] = (1-x^2)^{(2n-1)/2}\left[(1-x^2)\,v' - (2n+1)\,x\,v\right] = 0,$$

所以 $(1-x^2)^{(2n+1)/2}\,v$ 在 $(-1,1)$ 上是常数。它在 $x = 0$ 的值为 $v(0) = y^{(n+1)}(0) = 0$，因为 $n+1$ 是奇数。因此 $v \equiv 0$，即对所有 $|x| < 1$，$y^{(|m|+1)}(x) = 0$，故 y 是次数最多为 $|m|$ 的多项式。多项式等于其本身的麦克劳林级数，所以对 $|x| < 1$，

$$\cos(m \arcsin x) = 1 - \frac{m^2x^2}{2!} + \frac{m^2(m^2-2^2)\,x^4}{4!} - \cdots,$$

右边就是上面那个会终止的级数。

现在设 $|\theta| < \frac{1}{2}\pi$，并令 $x = \sin\theta$。则 $|x| < 1$，且 $\arcsin(\sin\theta) = \theta$（对 θ 的限制正是在这里用上），所以

$$\cos m\theta = 1 - \frac{m^2\sin^2\theta}{2!} + \frac{m^2(m^2-2^2)\sin^4\theta}{4!} - \cdots,$$

这是一个关于 $\sin\theta$ 的多项式，其开首各项与所求相符。

最后，$\sin^{|m|}\theta$ 的系数为 $y^{(|m|)}(0)/|m|!$，即 $\pm\, m^2(m^2-2^2)\cdots\left(m^2-(|m|-2)^2\right)/|m|!$。这些因子都不为零，因为当 $2j \leq |m| - 2$ 时 $(2j)^2 < m^2$。所以当 $m \neq 0$ 时，这个多项式的次数恰好是 $|m|$（即当 m 为正偶数时，次数为 m）；当 $m = 0$ 时，它是常数 1。

验算

- $m=2$：级数给出 $1-2\sin^2\theta$，这正是我们熟悉的 $\cos 2\theta$。
- $m=4$：级数给出 $1-8\sin^2\theta+\frac{16\cdot 12}{24}\sin^4\theta=1-8\sin^2\theta+8\sin^4\theta$。直接计算，$\cos 4\theta=1-2\sin^2 2\theta=1-8\sin^2\theta\cos^2\theta=1-8\sin^2\theta+8\sin^4\theta$。两者一致。
- 我们利用 $\cos(m\arcsin x)$ 在各点的泰勒展开（Taylor expansion），在 200 组随机的 (m,x) 上，以数值方法验证了 $n=0,\ldots,8$ 时的 $(*)$（最大相对残差约为 10^{-13}），并对所有不超过 12 的偶数 m 验证了 $\cos m\theta$ 与该多项式相等。例如当 $m=6$、$\theta=1$ 时，两边都等于 0.96017。

常见错误

- 微分 $(1-x^2)y^{(n+2)}$ 时，漏掉乘积法则产生的 $-2xy^{(n+2)}$ 项：正是这一项把 $2n+1$ 变成 $2n+3$。
- 以为终止的麦克劳林级数必定等于原函数。这里确实如此，但需要用 $v=y^{(|m|+1)}$ 的论证（或其他理据）来说明。
- 把次数写成 $m/2$：这个多项式只含 $\sin\theta$ 的偶数次幂，但最高次幂是 $\sin^m\theta$。

STEP 之外

若改为代入 $x=\cos\phi$，同类方程便化为切比雪夫（Chebyshev）方程 $(1-x^2)y''-xy'+n^2y=0$，其多项式解 $T_n(\cos\phi)=\cos n\phi$ 就是切比雪夫多项式（Chebyshev polynomials），在数值逼近中占核心地位。我们把 $x=0$ 代入经微分的方程以得出系数递推关系的技巧，正是线性常微分方程（linear ODEs）的*幂级数法*（power series method）的起点；在这个方法中，级数在特殊参数值下终止，正是勒让德（Legendre）多项式、埃尔米特（Hermite）多项式和切比雪夫多项式产生的方式。Boyce 与 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）有一章清晰讲解级数解。

详细解答

STEP III 2010，第 8 题

纯数学 • 利用商法则积分

STEP III 2010，第 8 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q8

关键思路

表达式 $\mathrm{QR}' - \mathrm{Q}'\mathrm{R}$ 正是对商 R/Q 求导时得到的分子（numerator）。因此，只要能找到一个 R 产生所给的 P，任何形如 P/Q^2 的被积函数（integrand）都可以一眼积出来。在两个部分中，分母都告诉我们 Q 是甚么，然后我们通过比较系数（comparing coefficients）求出 R。在第 (ii) 部分，微分方程（differential equation）的左边原来恰好是 $\mathrm{Q}y' - \mathrm{Q}'y$，即是同一个结构换了个样子出现。

解答

一般积分

由商法则（quotient rule），

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{R}(x)}{\mathrm{Q}(x)}\right) = \frac{\mathrm{Q}(x)\mathrm{R}'(x) - \mathrm{Q}'(x)\mathrm{R}(x)}{\left(\mathrm{Q}(x)\right)^2} = \frac{\mathrm{P}(x)}{\left(\mathrm{Q}(x)\right)^2},$$

所以

$$\int \frac{\mathrm{P}(x)}{\left(\mathrm{Q}(x)\right)^2}\,\mathrm{d}x = \frac{\mathrm{R}(x)}{\mathrm{Q}(x)} + C.$$

第 (i) 部分

这里 $\mathrm{Q}(x) = 1 + 2x + 3x^2$，所以 $\mathrm{Q}'(x) = 2 + 6x$。设 $\mathrm{R}(x) = a + bx + cx^2$，则 $\mathrm{R}'(x) = b + 2cx$，

$$\begin{aligned}\mathrm{QR}' - \mathrm{Q}'\mathrm{R} &= (1 + 2x + 3x^2)(b + 2cx) - (2 + 6x)(a + bx + cx^2) \\ &= (b - 2a) + (2c - 6a)x + (2c - 3b)x^2.\end{aligned}$$

（两个乘积各自产生的 x^3 项 $6cx^3$ 互相抵消。）我们需要它等于 $5x^2 - 4x - 3$，所以

$$b - 2a = -3, \qquad 2c - 6a = -4, \qquad 2c - 3b = 5.$$

由头两条得 $b = 2a - 3$ 及 $c = 3a - 2$。代入第三条，

$$2(3a - 2) - 3(2a - 3) = 6a - 4 - 6a + 9 = 5,$$

这对任何 a 都成立。所以第三条方程没有提供新信息，每个 a 值都可行：

$$\mathrm{R}(x) = a + (2a - 3)x + (3a - 2)x^2.$$

取 $a = 1$ 得到最整洁的选择 $\mathrm{R}(x) = 1 - x + x^2$，因此

$$\int \frac{5x^2 - 4x - 3}{(1 + 2x + 3x^2)^2}\,\mathrm{d}x = \frac{1 - x + x^2}{1 + 2x + 3x^2} + C.$$

分母永不为零（其判别式（discriminant）为 $4 - 12 < 0$），所以此式对所有实数 x 成立。

不唯一性。 我们刚看到有效的 R 组成一个单参数族，所以选择并不唯一：例如 $a = 0$ 给出 $\mathrm{R} = -3x - 2x^2$，$a = 1$ 给出 $\mathrm{R} = 1 - x + x^2$，两者都可行。把一般 a 的选择与 $a = 0$ 的选择比较，

$$\big(a + (2a - 3)x + (3a - 2)x^2\big) - \big(-3x - 2x^2\big) = a + 2ax + 3ax^2 = a\,\mathrm{Q}(x).$$

? **想一想** 为甚么在 R 上加上 Q 的倍数不会带来任何分别?

若以 $\mathrm{R} + k\mathrm{Q}$ 取代 R，其中 k 为常数，则

$$\mathrm{Q}(\mathrm{R} + k\mathrm{Q})' - \mathrm{Q}'(\mathrm{R} + k\mathrm{Q}) = \mathrm{QR}' - \mathrm{Q}'\mathrm{R} + k(\mathrm{QQ}' - \mathrm{Q}'\mathrm{Q}) = \mathrm{QR}' - \mathrm{Q}'\mathrm{R},$$

所以 P 不变。相应地，答案 R/Q 变成 $\mathrm{R}/\mathrm{Q} + k$：不同的 R 给出的原函数（antiderivative）只相差一个常数，而这个常数会被任意的积分常数（constant of integration）吸收。例如，

$$\frac{1 - x + x^2}{1 + 2x + 3x^2} = \frac{-3x - 2x^2}{1 + 2x + 3x^2} + 1.$$

第 (ii) 部分

设 $\mathrm{Q}(x) = 1 + \cos x + 2\sin x$。则 $\mathrm{Q}'(x) = -\sin x + 2\cos x$，所以方程中 y 的系数 $\sin x - 2\cos x$ 恰好是 $-\mathrm{Q}'(x)$。因此方程为

$$\mathrm{Q}\frac{\mathrm{d}y}{\mathrm{d}x} - \mathrm{Q}'y = 5 - 3\cos x + 4\sin x.$$

在任何 $\mathrm{Q} \neq 0$ 的区间上，我们除以 Q^2，并看出左边是一个商的导数：

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y}{\mathrm{Q}}\right) = \frac{5 - 3\cos x + 4\sin x}{\mathrm{Q}^2}.$$

（换句话说，$1/\mathrm{Q}^2$ 是一个积分因子（integrating factor）。）要积出右边，我们寻找一个 R 使得 $\mathrm{QR}' - \mathrm{Q}'\mathrm{R} = 5 - 3\cos x + 4\sin x$。仿照 Q 的形式，试设 $\mathrm{R} = a + b\cos x + c\sin x$，于是 $\mathrm{R}' = -b\sin x + c\cos x$。展开，

$$\begin{aligned}\mathrm{QR}' &= -b\sin x + c\cos x + (2c - b)\sin x\cos x + c\cos^2 x - 2b\sin^2 x,\\ \mathrm{Q}'\mathrm{R} &= -a\sin x + 2a\cos x + (2c - b)\sin x\cos x + 2b\cos^2 x - c\sin^2 x.\end{aligned}$$

相减后，$\sin x\cos x$ 项互相抵消，而平方项合并为 $(c - 2b)(\cos^2 x + \sin^2 x) = c - 2b$。所以

$$\mathrm{QR}' - \mathrm{Q}'\mathrm{R} = (c - 2b) + (c - 2a)\cos x + (a - b)\sin x.$$

与 $5 - 3\cos x + 4\sin x$ 比较：

$$c - 2b = 5, \qquad c - 2a = -3, \qquad a - b = 4.$$

第一条方程减去第二条得 $2a - 2b = 8$，又是第三条方程，所以与第 (i) 部分一样有一个自由参数，原因也相同。取 $b = 0$ 得 $c = 5$、$a = 4$，所以 $\mathrm{R} = 4 + 5\sin x$。因此

$$\frac{y}{\mathrm{Q}} = \frac{4 + 5\sin x}{1 + \cos x + 2\sin x} + C,$$

而通解（general solution）为

$$y = 4 + 5\sin x + C(1 + \cos x + 2\sin x),$$

其中 C 为任意常数。虽然我们求解时除以了 Q，但直接代入可以确认这个函数对每个 x 都满足方程。其他 b 的选择只会令 C 平移：例如 $b = -2$ 给出 $\mathrm{R} = 2 - 2\cos x + \sin x = (4 + 5\sin x) - 2\mathrm{Q}$。

验算

- 在第 (i) 部分，由我们的原函数，被积函数的 $\int_0^1$ 应为 $\frac{1}{6} - 1 = -\frac{5}{6}$；辛普森法则（Simpson's rule）给出 $-0.83333\ldots$，而且对每个 a 的选择，在随机区间上结果都吻合。

- 在第 (ii) 部分，代入 $y = 4 + 5\sin x$：左边为 $(1 + \cos x + 2\sin x)5\cos x - (2\cos x - \sin x)(4 + 5\sin x) = 5 - 3\cos x + 4\sin x$。而 $y = \mathrm{Q}$ 给出 $\mathrm{QQ}' - \mathrm{Q}'\mathrm{Q} = 0$，即余函数（complementary function）。
- 以数值方法（龙格－库塔法（Runge–Kutta））由 $y(0.5) = 1.7$ 解方程至 $x = 1.5$，得 3.910638，与取相应 C 的公式吻合。

常见错误

- 以为三个未知数的三条方程必定有唯一解。这里它们是相依的，而这正是不唯一性问题的重点所在。
- 在第 (ii) 部分弄错 Q' 的正负号：$\mathrm{Q}' = 2\cos x - \sin x$，所以 y 的系数是 $-\mathrm{Q}'$，而不是 $+\mathrm{Q}'$。
- 忘记在最终答案中积分常数是乘以 Q 的：通解是 $y = \mathrm{R} + C\mathrm{Q}$，而不是 $y = \mathrm{R} + C$。

STEP 之外

第 (ii) 部分的方程是一条一阶线性微分方程（first-order linear differential equation），其通解的形状是“特解（particular solution）加上齐次方程（homogeneous equation）之解的倍数”，这个模式贯穿所有线性问题。表达式 $\mathrm{QR}' - \mathrm{Q}'\mathrm{R}$ 是 Q 与 R 的朗斯基行列式（Wronskian），它在二阶线性方程中担当核心角色，尤其是在降阶法（reduction of order）和参数变易法（variation of parameters）之中。Boyce 与 DiPrima 的 *Elementary Differential Equations*（《初等微分方程》）同时涵盖积分因子和朗斯基行列式。

详细解答

STEP III 2010，第 9 题

力学 • 圆上的运动

STEP III 2010，第 9 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

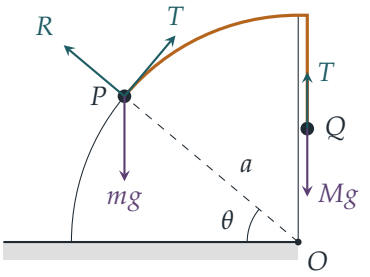
step.maths.org/questions/10-s3-q9

关键思路

由于绳子不可伸长（inextensible），而木块光滑（smooth），P 和 Q 始终以相同的速率运动，而且整个系统的能量守恒（conservation of energy）。这样便能把 P 的速率写成 θ 的函数（function），完全不用解任何微分方程（differential equation）。接着，由于 P 在圆周上运动，沿指向 O 的方向分解力，便可求出法向反作用力（normal reaction）。保持接触即表示这个反作用力始终非负，而利用凹性（concavity）论证，可知只需检查木块顶部的情况便足够。

解答

模型。 P 和 Q 是质点（particle），绳子轻而不可伸长，木块光滑，因此整条绳子上的张力（tension）T 处处相同。设 P 位于圆弧上角度 θ 的位置。绳子由 P 沿圆弧伸延至木块顶部，然后沿铅垂面旁边垂直向下连到 Q。



速率。 当 P 由 $\theta = 0$ 移动到 θ 时，它走过了长度为 $a\theta$ 的圆弧，因此 Q 下降了距离 $a\theta$。两者都以速率 $v = a\dot{\theta}$ 运动。与此同时，P 上升了高度 $a\sin\theta$。木块光滑，而绳子对整个系统不做净功（因为它不可伸长），所以能量守恒。由静止开始，

$$\tfrac{1}{2}(m+M)v^2 = Mga\theta - mga\sin\theta, \qquad \text{所以} \qquad v^2 = \frac{2ga\,(M\theta - m\sin\theta)}{m+M}.$$

由于 $M\theta > m\theta \geq m\sin\theta$，此式在 $0 < \theta \leq \frac{1}{2}\pi$ 时为正，所以系统确实会由静止开始运动。

法向反作用力。 P 绕 O 沿半径为 a 的圆周运动，因此它指向 O 的加速度为 v^2/a。作用于 P 的力如下：

- 沿 OP 方向、背离 O 的反作用力 R；
- 沿切线（tangent）方向的张力 T，它没有径向分量；
- 重力 mg，由于 OP 与水平方向成角 θ，它指向 O 的分量为 $mg\sin\theta$。

沿指向 O 的方向分解，

$$mg\sin\theta - R = \frac{mv^2}{a} = \frac{2mg\,(M\theta - m\sin\theta)}{m+M}.$$

因此

$$\begin{aligned} R &= \frac{mg\big[(m+M)\sin\theta - 2M\theta + 2m\sin\theta\big]}{m+M} \\ &= \boxed{\frac{mg\big[(3m+M)\sin\theta - 2M\theta\big]}{m+M}.} \end{aligned}$$

> **想一想** 对 $[0, \frac{1}{2}\pi]$ 内的每一个 θ，R 都必须非负。我们是否需要找出 R 的最小值?

条件。 P 保持接触，当且仅当 $R \geq 0$ 在整个过程中成立，即

$$f(\theta) = (3m+M)\sin\theta - 2M\theta \geq 0 \quad \text{对所有} 0 \leq \theta \leq \tfrac{1}{2}\pi.$$

我们不需要求最小值。由于在这个区间上 $f''(\theta) = -(3m+M)\sin\theta \leq 0$，$f$ 是凹函数，所以它的图像位于连接两个端点的弦（chord）之上或弦上。又因为 $f(0) = 0$，所以：

- 必要性（*necessity*）。若 P 在 $\theta = \frac{1}{2}\pi$ 时仍保持接触，则 $f(\frac{1}{2}\pi) = 3m + M - \pi M \geq 0$。
- 充分性（*sufficiency*）。若 $f(\frac{1}{2}\pi) \geq 0$，则由 $(0,0)$ 到 $(\frac{1}{2}\pi, f(\frac{1}{2}\pi))$ 的弦位于横轴之上或横轴上。由凹性可知 f 位于弦之上或弦上，所以在整个区间上 $f \geq 0$。

因此条件是 $3m \geq (\pi - 1)M$，即

$$\frac{m}{M} \geq \frac{\pi - 1}{3},$$

得证。注意 $(\pi - 1)/3 \approx 0.714 < 1$，所以这与 $m < M$ 并无矛盾。

验算

- 当 $\theta = 0$ 时，得 $R = 0$。这是正确的：开始时 P 位于木块底部，半径 OP 是水平的，P 的重量由桌面承托。
- 对 P 沿切线方向分解、对 Q 沿铅垂方向分解，可得 $T = \dfrac{Mmg(1 + \cos\theta)}{m + M} > 0$，所以绳子在整个过程中确实保持拉紧。
- 数值上，取 $m = 0.8$、$M = 1$ 及 $a = 1$（比值 $0.8 > 0.714$），逐步数值求解运动方程（equation of motion），并用牛顿第二定律（Newton's second law）计算 R，结果与公式吻合至 10^{-13}，而且 R 始终非负。若取 $m/M = \frac{2}{3}$，则 R 在接近顶部时为负。

常见错误

- 忘记 Q 也有动能（kinetic energy），因而写成 $\frac{1}{2}mv^2$ 而非 $\frac{1}{2}(m + M)v^2$。
- 把张力放进径向方程。在光滑表面上，绳子沿 P 点的切线方向。
- 只检查 $\theta = \frac{1}{2}\pi$ 时 $R \geq 0$，却没有说明为何这样已经足够。正是 f 的凹性令这样做足够。

STEP 之外

在这里，能量守恒代替了求解运动方程 $(m + M)a\ddot{\theta} = g(M - m\cos\theta)$。这就是首次积分（first integral）的概念，它在拉格朗日力学（Lagrangian mechanics）中处于核心地位：绳子是一个约束（constraint），而单一个坐标 θ 便能描述整个系统。本题属于“珠子离开球面”这一类问题，Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）对这类问题有详细的讨论。

详细解答

STEP III 2010，第 10 题

力学 · 微小振动

STEP III 2010，第 10 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q10

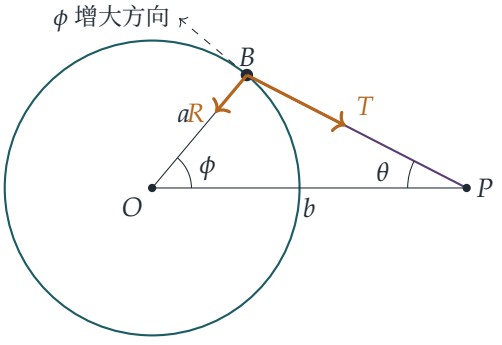
关键思路

珠子（bead）被限制在圆环上运动，所以只有沿圆环切线（tangent）方向的力分量会改变它的运动；法向反作用力（normal reaction）以及重力（由圆环的竖直反作用力抵消）完全不会出现在方程中。其余一切都是三角形 OBP 的几何。正弦定理（sine rule）把这几何关系转化为弦的长度；在角度很小时，它又给出第二部分所要求的 θ 与 ϕ 之间的线性关系。

解答

运动方程

把珠子视为光滑圆环上的质点，把弦视为轻而有弹性的弦。取 P 在正 x 轴上，并以逆时针方向量度 ϕ，使 B 位于与 OP 成角 ϕ 的位置，如图所示。在三角形 OBP 中，O 处的角为 ϕ，P 处的角为 θ，因此 B 处的角为 $\pi - \theta - \phi$。



设 $l = BP$。在三角形 OBP 中应用正弦定理，

$$\frac{l}{\sin\phi} = \frac{a}{\sin\theta} = \frac{b}{\sin(\theta+\phi)},$$

这里用了 $\sin(\pi-\theta-\phi) = \sin(\theta+\phi)$。特别地，$l = \dfrac{a\sin\phi}{\sin\theta}$。

弦在整个运动过程中都是拉紧的，因为 B 与 P 的最短距离是 $b-a$，而 $b-a > c$。所以由胡克定律（Hooke's law），张力（tension）为

$$T = \frac{\lambda(l-c)}{c} = \lambda\left(\frac{a\sin\phi}{c\sin\theta} - 1\right).$$

在水平面内的力有沿 BP 的张力 T，以及光滑圆环沿半径 OB 方向的法向反作用力 R。（在竖直方向，重量由反作用力的竖直部分抵消。）珠子在半径为 a 的圆上运动，所以它沿切线方向（取 ϕ 增大的方向）的加速度为 $a\ddot{\phi}$。只有 T 有切向分量。

弦 BP 与向内的半径 BO 成角 $\pi-\theta-\phi$，所以它与向外的半径成角 $\theta+\phi$，其沿切线的分量大小为 $T\sin(\theta+\phi)$。这个分量指向 P，即 ϕ 减小的方向。沿切线分解，

$$ma\ddot{\phi} = -\lambda\left(\frac{a\sin\phi}{c\sin\theta} - 1\right)\sin(\theta+\phi),$$

即为所求。（当 $\phi < 0$ 时，图形关于 OP 作反射；此时 $\theta < 0$ 亦成立，而以有向角计算，同一方程仍然成立。）

微小振动

正弦定理给出精确关系 $a\sin(\theta+\phi) = b\sin\theta$。当 θ 和 ϕ 都很小时，两边都有 $\sin x \approx x$，所以

$$a(\theta+\phi) \approx b\theta.$$

因此 $\theta \approx \dfrac{a\phi}{b-a}$，以及 $\theta+\phi \approx \dfrac{b\theta}{a} \approx \dfrac{b\phi}{b-a}$。

> ? **想一想** 当两个角都趋于零时，$\dfrac{a\sin\phi}{c\sin\theta}$ 会怎样？它看起来像 $0/0$。

这个商其实就是 l/c，而当 B 趋近平衡位置（equilibrium）时，长度 l 趋于 $b-a$。等价地，由上面的关系可得 $\dfrac{a\sin\phi}{\sin\theta} \approx \dfrac{a\phi}{\theta} \approx b-a$。所以在首阶近似（leading order）下，张力就是常数的平衡张力

$$T_0 = \lambda\left(\frac{b-a}{c} - 1\right) = \frac{\lambda(b-a-c)}{c},$$

由于 $c < b-a$，它是正数。因子 $\sin(\theta+\phi)$ 本身已经很小，所以 T 的任何修正项只会在二阶才有贡献。

代入运动方程并保留一阶项，

$$ma\ddot{\phi} \approx -\frac{\lambda(b-a-c)}{c}\cdot\frac{b\phi}{b-a}, \qquad \text{所以} \qquad \ddot{\phi} = -\omega^2\phi, \quad \omega^2 = \frac{\lambda b(b-a-c)}{mac(b-a)}.$$

这是简谐运动（simple harmonic motion），而在 $\theta = \phi = 0$ 附近微小振动的周期（period）为

$$\frac{2\pi}{\omega} = 2\pi\sqrt{\frac{mac(b-a)}{\lambda b(b-a-c)}}.$$

验算

- 量纲（dimensions）：λ 是力，所以 $mac(b-a)/(\lambda b(b-a-c))$ 的量纲是 $\mathrm{M\,L}/(\mathrm{M\,L\,T}^{-2}) = \mathrm{T}^2$，正是周期应有的量纲。
- 当 $c \to b-a$ 时，平衡张力消失，周期变成无限大：松弛的弦不提供恢复力（restoring force），这是合理的。
- 数值上，取 $a = 1$、$b = 2.5$、$c = 0.7$、$\lambda = 3$、$m = 0.5$，公式给出周期 1.8586。从一个微小位移出发，以数值方法逐步推进精确的运动（力直接由 B 的位置计算），得出的值在六位有效数字内相同。在随机选取的位置上，精确的运动方程亦与直接的向量分解结果吻合。

常见错误

- 符号弄错：切向拉力指向 P，即指向 $\phi = 0$，所以它必须带负号。
- 以为平衡时张力为零。在平衡位置，弦被拉长至长度 $b-a > c$，所以张力是 $\lambda(b-a-c)/c$，而正是这个张力提供恢复力。
- 使用 $\theta + \phi \approx \phi$。这里 θ 与 ϕ 同阶，把它略去会得出错误的周期。

STEP 之外

这里的方法是求微小振动频率的标准做法：找出平衡位置，把运动方程在其附近线性化（linearise），然后读出 ω^2。一个巧妙之处是，恢复力完全来自平衡张力通过弦方向的改变而产生的作用，这与绷紧的结他弦是同一机制。在大学一年级的课程中，这会发展为拉格朗日力学（Lagrangian mechanics）和简正模（normal modes）；Kleppner 与 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）是细心学习微小振动分析的好地方。

详细解答

STEP III 2010，第 11 题

力学 · 摩擦力与相对运动

STEP III 2010，第 11 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

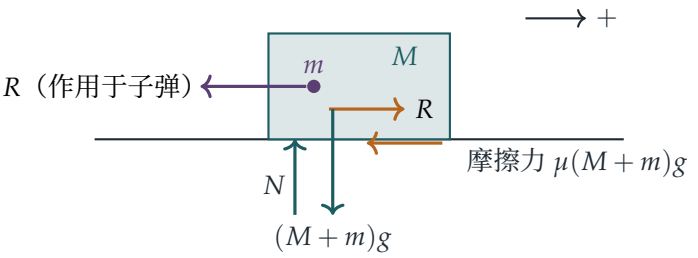
step.maths.org/questions/10-s3-q11

关键思路

这里动量并不守恒（conservation of momentum），因为地面的摩擦力自始至终都作用在木块上，所以熟悉的“子弹嵌入木块”捷径行不通。我们改为把子弹和木块视为两个独立的质点，各自有恒定的加速度，并由大小相等、方向相反的阻力 R 联系起来。既然加速度恒定，*suvat* 方程便能完成所有工作，而相对运动只不过是两者之差。

解答

我们把子弹和木块模拟为沿水平直线运动的质点，并以子弹的运动方向为正方向。当子弹在木块内时，木块以力 R 向后推子弹；根据牛顿第三定律（Newton's third law），子弹亦以力 R 向前推木块。



法向反作用力（*normal reaction*）。铅垂方向上没有任何物体加速。子弹由木块承托，所以地面承受两者的重量：$N = (M + m)g$。

摩擦力（*friction*）。由于题目给出 $R > (M + m)\mu g$，作用在木块上的向前力 R 超过可提供的最大摩擦力 $\mu N = (M + m)\mu g$。因此木块确实会滑动，而一旦木块移动，摩擦力便

取其极限值（limiting value）$\mu(M+m)g$，方向向后。

第 (i) 部分

子弹所受的唯一水平力是向后的 R，所以子弹以 R/m 减速。木块所受的向前合力为 $R-(M+m)\mu g$，所以它向前加速，加速度为

$$A=\frac{R-(M+m)\mu g}{M}>0.$$

相对于木块，子弹的加速度为 $-\frac{R}{m}-A$，所以它相对于木块的减速度（deceleration）大小为

$$a=\frac{R}{m}+\frac{R-(M+m)\mu g}{M},$$

此即所求。

第 (ii) 部分

相对于木块，子弹以速率 u 开始运动，并以 a 均匀减速，所以它在以下时间后相对于木块静止：

$$t_1=\frac{u}{a}.$$

在那一刻，木块的速率就是共同速率 v。木块由静止开始以恒定加速度 A 运动，所以

$$v=At_1=\frac{Au}{a},\qquad \text{即}\qquad av=Au=\frac{Ru-(M+m)\mu gu}{M},$$

这就是所要求的结果。

想一想 为甚么不直接用动量守恒求 v?

子弹—木块系统的动量只有在没有外加水平力作用时才守恒。这里地面的摩擦力在整段时间 t_1 内都作用在木块上，并带走动量 $(M+m)\mu g\,t_1$。（对于真实的子弹，这段时间极短，所以 $v\approx mu/(M+m)$，但题目要求的是准确值。）

第 (iii) 部分

在时间 t_1 内，木块由静止均匀加速至速率 v，所以它移动的距离等于平均速率乘以时间：

$$s_1=\frac{0+v}{2}t_1=\frac{uv}{2a}.$$

第 (iv) 部分

当子弹不再穿过木块时，两者便作为一个质量为 $M+m$ 的单一物体一起运动。其重量仍为 $(M+m)g$，所以摩擦力为向后的 $\mu(M+m)g$，减速度为 μg。由速率 v 开始，它再滑行

$$s_2=\frac{v^2}{2\mu g}$$

才停下（之后保持静止，因为已没有水平力去克服静摩擦力（static friction））。总距离为

$$s_1 + s_2 = \frac{uv}{2a} + \frac{v^2}{2\mu g} = \frac{uv}{2a} + \frac{v}{2\mu g} \cdot \frac{Au}{a} = \frac{uv}{2a\mu g}(\mu g + A),$$

其中我们对其中一个因子 v 使用了第 (ii) 部分的 $v = Au/a$。

想一想 怎样消去 a 和 A，得出简洁的答案?

把两者都化成以 mM 为分母，并寻找公因子：

$$\mu g + A = \frac{M\mu g + R - (M+m)\mu g}{M} = \frac{R - m\mu g}{M},$$
$$a = \frac{MR + mR - m(M+m)\mu g}{mM} = \frac{(M+m)(R - m\mu g)}{mM}.$$

所以 $\dfrac{\mu g + A}{a} = \dfrac{m}{M+m}$，其中因子 $R - m\mu g$ 被约去（它是正数，因为 $R > (M+m)\mu g > m\mu g$）。因此木块移动的总距离为

$$\frac{uv}{2\mu g} \cdot \frac{m}{M+m} = \frac{muv}{2(M+m)\mu g},$$

此即所求。

$R < (M+m)\mu g$ 的情况

此时作用在木块上的向前推力 R 小于极限摩擦力 $(M+m)\mu g$，所以静摩擦力足以使木块保持不动：木块始终不会移动。子弹只是在静止的木块内以 R/m 减速，在穿入距离 $\dfrac{mu^2}{2R}$ 后停下（假设木块足够厚），而木块自始至终都留在原位。

验算

- 取 $m = 0.02\,\mathrm{kg}$、$M = 2\,\mathrm{kg}$、$u = 300\,\mathrm{m\,s^{-1}}$、$\mu = 0.5$、$g = 9.8\,\mathrm{m\,s^{-2}}$ 及 $R = 1000\,\mathrm{N}$。则 $a \approx 50\,495\,\mathrm{m\,s^{-2}}$，$v \approx 2.941\,\mathrm{m\,s^{-1}}$，$s_1 \approx 0.0087\,\mathrm{m}$，$s_2 \approx 0.8827\,\mathrm{m}$，以及 $s_1 + s_2 \approx 0.8914\,\mathrm{m}$，与 $muv/\big(2(M+m)\mu g\big)$ 相符。对两个质点作逐步数值模拟，亦得出相同的 v 和总距离。
- 当 $R \to \infty$（即瞬间碰撞）时，$v = Au/a \to mu/(M+m)$，正是动量守恒给出的答案，理应如此。
- 量纲（dimensions）：$muv/\big((M+m)\mu g\big)$ 的量纲是 (速率)^2/(加速度)，即长度。

常见错误

- 对木块使用 $N = Mg$，忘记嵌入其中的子弹的重量同样由地面承受。

- 在撞击过程中使用动量守恒，忽略了摩擦力的冲量。
- 把 R 当作木块所受的合力，而不是 R 减去摩擦力。

STEP 之外

冲量（impulse）可以为整个运动提供一个快速的一致性检验：由于系统最终静止，在木块移动的整段时间内，地面摩擦力的冲量必须等于子弹的初始动量 mu。把系统分拆成几部分，用牛顿第三定律追踪内力，并仔细处理外力，是大学一年级力学的核心技巧，并会引伸至质心参考系（centre-of-mass frame）和非弹性碰撞（inelastic collisions）。Kleppner 和 Kolenkow 的 *An Introduction to Mechanics*（《力学导论》）发展了这些概念，书中有很多正是这类风格的题目。

详细解答

STEP III 2010，第 12 题

概率与统计 · 几何分布的期望值

STEP III 2010，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q12

关键思路

级数 S 是一个算术几何级数（arithmetico-geometric series），而减去 rS 这个技巧可以把它化成普通的几何级数（geometric series）。求出 S 之后，题中每一个期望值（expectation）都正好是这种形式的和：射击次数按线性增长，而射到那一步的概率则按几何方式递减。所以整道题归结为把每个期望值写成一个和，再读出 r 和 d 的正确数值。

解答

级数

由于 $|r| < 1$，级数绝对收敛（converges absolutely）（其各项的绝对值不超过某常数乘以 $n|r|^n$），所以可以逐项相减。把 r 的同次幂对齐，

$$\begin{aligned} S &= 1 + (1+d)r + (1+2d)r^2 + (1+3d)r^3 + \cdots, \\ rS &= \qquad r + (1+d)r^2 + (1+2d)r^3 + \cdots, \end{aligned}$$

对于 $n \geq 1$，r^n 的系数之差都正好是 d。因此

$$(1-r)S = 1 + d\,(r + r^2 + r^3 + \cdots) = 1 + \frac{rd}{1-r},$$

这里把几何级数求和，因为 $|r| < 1$，所以这一步成立。两边除以 $1 - r \neq 0$，

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

即为所求。

Arthur 单独射击

假设 $0 < a \leq 1$（若 $a = 0$，Arthur 永远射不中）。设 N 为 Arthur 所需的射击次数。$N = n$ 当且仅当他首 $n-1$ 次都射不中，而第 n 次射中，所以由独立性（independence）得

$$\mathrm{P}(N = n) = a'^{n-1}a, \qquad n = 1, 2, 3, \ldots$$

因此

$$\mathrm{E}(N) = \sum_{n=1}^{\infty} n\, a\, a'^{n-1} = a\big(1 + 2a' + 3a'^2 + \cdots\big).$$

括号内正是 $r = a'$、$d = 1$ 时的 S（a'^n 的系数是 $1+n$），而且 $0 \leq a' < 1$，所以

$$\mathrm{E}(N) = a\left(\frac{1}{1-a'} + \frac{a'}{(1-a')^2}\right) = a\left(\frac{1}{a} + \frac{a'}{a^2}\right) = \frac{a+a'}{a} = \frac{1}{a},$$

这里先用了 $1 - a' = a$，再用了 $a + a' = 1$。（当 $a = 1$ 时，这条公式同样给出 1。）

比赛：谁会胜出

以下取 $0 < a < 1$ 及 $0 < b < 1$，使得 $0 < a'b' < 1$。第 $1, 3, 5, \ldots$ 次射击属于 Arthur，第 $2, 4, 6, \ldots$ 次射击属于 Boadicea。把 Arthur 射一次、接着 Boadicea 射一次称为一个回合；一个回合中无人射中的概率是 $a'b'$。

Arthur 在第 $2k+1$ 次射击时胜出（$k = 0, 1, 2, \ldots$），当且仅当首 k 个回合都无人射中，然后 Arthur 射中。由独立性，这事件的概率是 $(a'b')^k a$，而这些事件互斥（disjoint），所以

$$\alpha = \sum_{k=0}^{\infty} (a'b')^k a = \frac{a}{1-a'b'}.$$

同理，Boadicea 在第 $2k+2$ 次射击时胜出，当且仅当 k 个回合都无人射中，然后 Arthur 射不中而 Boadicea 射中，其概率是 $(a'b')^k a'b$。所以

$$\beta = \sum_{k=0}^{\infty} (a'b')^k a'b = \frac{a'b}{1-a'b'} = \frac{(1-a)b}{1-(1-a)(1-b)}.$$

验证一下：$1 - a'b' = a + b - ab = a + a'b$，所以 $\alpha + \beta = 1$，即比赛以概率 1 结束。

比赛：期望长度

设 T 为比赛中的射击次数。由上一部分，$T = 2k+1$ 的概率是 $(a'b')^k a$，而 $T = 2k+2$ 的概率是 $(a'b')^k a'b$。记 $q = a'b'$，

$$\mathrm{E}(T) = a\sum_{k=0}^{\infty} (2k+1)q^k + a'b\sum_{k=0}^{\infty} (2k+2)q^k.$$

想一想 这两个和怎样套进级数 S 的形式?

第一个和是 $1+3q+5q^2+\cdots$，即 $r=q$、$d=2$ 时的 S。第二个和是 $1+2q+3q^2+\cdots$ 的两倍，而后者是 $r=q$、$d=1$ 时的 S。所以

$$\sum_{k=0}^{\infty}(2k+1)q^k=\frac{1}{1-q}+\frac{2q}{(1-q)^2}=\frac{1+q}{(1-q)^2},$$
$$\sum_{k=0}^{\infty}(2k+2)q^k=2\left(\frac{1}{1-q}+\frac{q}{(1-q)^2}\right)=\frac{2}{(1-q)^2}.$$

因此

$$\mathrm{E}(T)=\frac{a(1+q)+2a'b}{(1-q)^2}.$$

现在与目标式比较。利用 $\alpha=a/(1-q)$ 及 $\beta=a'b/(1-q)$，

$$\frac{\alpha}{a}+\frac{\beta}{b}=\frac{1}{1-q}+\frac{a'}{1-q}=\frac{(1+a')(1-q)}{(1-q)^2}.$$

所以只需证明 $a(1+q)+2a'b=(1+a')(1-q)$。利用 $1-q=a+a'b$，

$$(1+a')(1-q)=(1+a')(a+a'b)=a+aa'+a'b+a'^2b,$$
$$a(1+q)+2a'b=a+aa'(1-b)+2a'b=a+aa'-aa'b+2a'b.$$

两者之差（第一式减第二式）是

$$a'^2b+a'b-2a'b+aa'b=a'b\,(a'+a-1)=0,$$

因为 $a+a'=1$。因此

$$\mathrm{E}(T)=\frac{\alpha}{a}+\frac{\beta}{b},$$

即为所求。

验算

- 若 $a=1$，Arthur 第一次便射中：$\alpha=1$，$\beta=0$，$\mathrm{E}(T)=1$，与 $\alpha/a+\beta/b$ 吻合。
- 若 $a=b=p$，则 $\alpha/p+\beta/p=(\alpha+\beta)/p=1/p$。这是对的：此时比赛只是一连串独立的射击，每次射中的概率都是 p，而由第一部分，首次射中平均需要 $1/p$ 次射击。
- 取 $a=0.3$、$b=0.45$，得 $\alpha=0.3/0.615\approx0.4878$ 及 $\mathrm{E}(T)=1.7/0.615\approx2.764$；模拟（simulation）200 000 场比赛得出 0.486 和 2.769。

常见错误

- 忘记 Boadicea 只有在 Arthur 刚刚射不中时才可能胜出，于是写成 $\beta = b/(1 - a'b')$；漏掉的因子 a' 会令 $\alpha + \beta \neq 1$，从而暴露出来。
- 数回合而不是数射击次数：Boadicea 在第 $k+1$ 回合胜出的比赛共有 $2k+2$ 次射击，而不是 $k+1$ 次。

STEP 之外

这类结果通常可以通过以第一回合为条件（“首步分析”（first-step analysis））更简洁地证明：若第一回合无人射中，比赛在两次射击后便重新开始，所以 $\mathrm{E}(T) = a \cdot 1 + a'b' \cdot 2 + a'b'\big(2 + \mathrm{E}(T)\big)$，一行便得出 $\mathrm{E}(T) = (1 + a')/(1 - a'b')$。首步分析是研究马尔可夫链（Markov chains）的基本工具，通常在数学学位的第一或第二年学习。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）涵盖了生成函数（generating functions）和这类论证方法。

详细解答

STEP III 2010，第 13 题

概率与统计 · 线性组合的相关性

STEP III 2010，第 13 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/10-s3-q13

本题在 *Beyond STEP* 第 13.7 节制造相关性中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch13/correlation.html

关键思路

这里的一切都是平均值为 0、方差为 1 的独立（independent）变量的线性组合（linear combination），所以每个期望值（expectation）、方差（variance）和协方差（covariance）都可以直接从系数读出：若 $U = \sum u_k Z_k$ 及 $V = \sum v_k Z_k$，则 $\operatorname{Cov}(U, V) = \sum u_k v_k$。这道题其实是在逐个变量地构造三个具有任意指定相关系数（correlation）的标准化（standardised）变量，本质上就是乔装了的格拉姆–施密特（Gram–Schmidt）正交化过程。每个新条件都确定一个新的系数。

解答

全程我们使用协方差的双线性（bilinearity）、$\operatorname{Var}(Z_k) = 1$，以及当 $j \neq k$ 时 $\operatorname{Cov}(Z_j, Z_k) = 0$（这是因为各个 Z_k 互相独立）。特别地，对常数 u_k、v_k，

$$\operatorname{Cov}\Big(\sum_k u_k Z_k,\ \sum_k v_k Z_k\Big) = \sum_{j,k} u_j v_k \operatorname{Cov}(Z_j, Z_k) = \sum_k u_k v_k.$$

Z_1 与 Z_2 的相关系数

由独立性得 $\operatorname{Cov}(Z_1, Z_2) = \mathrm{E}(Z_1 Z_2) - \mathrm{E}(Z_1)\mathrm{E}(Z_2) = 0$，所以 $\operatorname{Corr}(Z_1, Z_2) = 0$。

变量 Y_2

记 $s=\sqrt{1-\rho_{12}^2}$，由于 $-1<\rho_{12}<1$，它是正实数。于是 $Y_2=\rho_{12}Z_1+sZ_2$，由期望值的线性（linearity of expectation）得

$$\mathrm{E}(Y_2)=\rho_{12}\mathrm{E}(Z_1)+s\,\mathrm{E}(Z_2)=0.$$

由上述协方差法则，

$$\mathrm{Var}(Y_2)=\rho_{12}^2+s^2=\rho_{12}^2+(1-\rho_{12}^2)=1,\qquad \mathrm{Cov}(Y_1,Y_2)=\mathrm{Cov}(Z_1,\rho_{12}Z_1+sZ_2)=\rho_{12}.$$

由于 $\mathrm{Var}(Y_1)=\mathrm{Var}(Y_2)=1$，定义中的分母为 1，因此

$$\mathrm{Corr}(Y_1,Y_2)=\rho_{12}.$$

变量 Y_3

条件 $\mathrm{E}(Y_3)=0$ 对任何 a,b,c 都自动成立，所以它没有提供任何信息。其余三个条件各给出一条方程。由于 Y_1,Y_2,Y_3 的方差都是 1，每个相关系数都等于对应的协方差。

与 Y_1 的相关系数。$\mathrm{Cov}(Z_1,\,aZ_1+bZ_2+cZ_3)=a$，所以

$$a=\rho_{13}.$$

与 Y_2 的相关系数。$\mathrm{Cov}(\rho_{12}Z_1+sZ_2,\,aZ_1+bZ_2+cZ_3)=\rho_{12}a+sb$，所以 $\rho_{12}\rho_{13}+sb=\rho_{23}$，又因 $s\neq 0$，

$$b=\frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}.$$

单位方差。$a^2+b^2+c^2=1$，由于 $c\geq 0$，我们取非负的平方根：

$$\begin{aligned}c^2&=1-\rho_{13}^2-\frac{(\rho_{23}-\rho_{12}\rho_{13})^2}{1-\rho_{12}^2}\\&=\frac{(1-\rho_{12}^2)(1-\rho_{13}^2)-(\rho_{23}-\rho_{12}\rho_{13})^2}{1-\rho_{12}^2}\\&=\frac{1-\rho_{12}^2-\rho_{13}^2-\rho_{23}^2+2\rho_{12}\rho_{13}\rho_{23}}{1-\rho_{12}^2},\end{aligned}$$

其中最后一步里 $\rho_{12}^2\rho_{13}^2$ 项互相抵消。因此

$$c=\sqrt{\frac{1-\rho_{12}^2-\rho_{13}^2-\rho_{23}^2+2\rho_{12}\rho_{13}\rho_{23}}{1-\rho_{12}^2}}.$$

想一想 如果平方根下的式子是负数，会怎样？

那么不存在实数 c，而且根本不存在具有这三个相关系数的随机变量：$\rho_{12},\rho_{13},\rho_{23}$ 这三个数彼此并不相容。例如，$\rho_{12}=\rho_{13}=0.9$ 而 $\rho_{23}=-0.9$ 是不可能的，因为如果 Y_2 和 Y_3 都与 Y_1 十分接近，它们之间就不可能强烈地反向相关。题目默认所给的相关系数是可以实现的，而这正正就是该分子非负的条件。

最后部分：X_i

对 $i=1,2,3$，取

$$X_i=\mu_i+\sigma_i Y_i,$$

其中 $\sigma_i>0$（我们可以取 σ_i 为正，因为题目只指定了 σ_i^2）。于是 $\mathrm{E}(X_i)=\mu_i+\sigma_i\cdot 0=\mu_i$，且 $\mathrm{Var}(X_i)=\sigma_i^2\mathrm{Var}(Y_i)=\sigma_i^2$。加上常数不会改变协方差，所以

$$\mathrm{Corr}(X_i,X_j)=\frac{\sigma_i\sigma_j\,\mathrm{Cov}(Y_i,Y_j)}{\sqrt{\sigma_i^2\sigma_j^2}}=\frac{\sigma_i\sigma_j}{|\sigma_i\sigma_j|}\,\rho_{ij}=\rho_{ij},$$

这里用了 $\sigma_i\sigma_j>0$。所以这些 X_i 具备所有要求的性质。

验算

- 若 $\rho_{13}=\rho_{23}=0$，则得 $a=b=0$、$c=1$，所以 $Y_3=Z_3$，与其他变量独立，理应如此。若 $\rho_{12}=0$，则 $Y_2=Z_2$，且 $a=\rho_{13}$、$b=\rho_{23}$、$c=\sqrt{1-\rho_{13}^2-\rho_{23}^2}$，正是显而易见的选择。
- 数值上，取 $\rho_{12}=0.5$、$\rho_{13}=-0.3$、$\rho_{23}=0.2$，得 $a=-0.3$、$b\approx 0.4041$、$c\approx 0.8641$，且 $a^2+b^2+c^2=1$。以 2×10^5 组三元组进行模拟（所用的 Z_k 是均匀分布而非正态分布），得到 X_i 的样本相关系数为 0.500、−0.302、0.200，平均值和标准差亦正确。

常见错误

- 忘记 $\mathrm{Var}(Y_2)=1$ 意味着相关系数就是协方差，结果除错了东西。
- 写成 $X_i=\mu_i+\sigma_i^2Y_i$：方差按乘数的*平方*缩放。
- 对其中一个变量取 $\sigma_i<0$，这会令它的相关系数变号。

STEP 之外

这三个条件表示系数矩阵 A（各行为 $(1,0,0)$、$(\rho_{12},s,0)$、(a,b,c)）满足 $AA^{\mathsf{T}}=R$，其中 R 是相关矩阵（correlation matrix）。这就是 R 的*乔列斯基分解*（Cholesky decomposition），它存在若且唯若 R 是半正定（positive semi-definite）的；而 c^2 的分子正是 $\det R$。统计学家在实际中正是用这个方法模拟相关的变量；若 Z_k 是正态分布的，所得的就是多元正态分布（multivariate normal distribution）。Blitzstein 与 Hwang 的 *Introduction to Probability*（《概率导论》）涵盖协方差、相关系数和多元正态分布；Strang 的 *Introduction to Linear Algebra*（《线性代数导论》）讲解正定矩阵（positive definite matrix）和乔列斯基分解。

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

2010 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。

