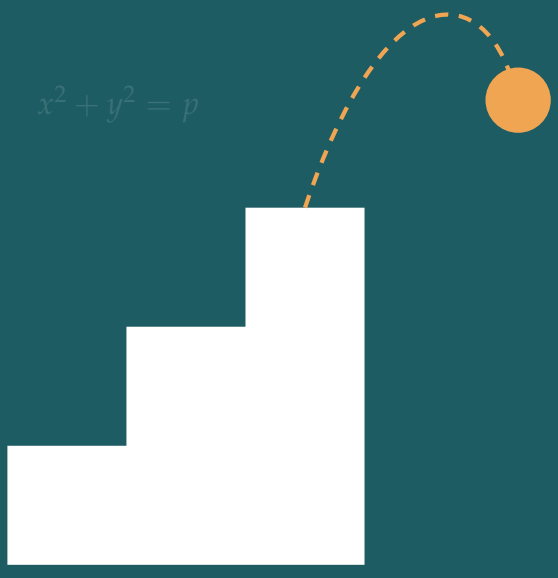
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

1991 年詳細解答



KEITH McNULTY

Beyond STEP

以歷屆 STEP 試題為門
走進大學數學

STEP I, II 及 III

1991 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

1991 年詳細解答

1991 年 STEP 試卷（STEP I, II 及 III）全部 48 題的詳細解答。自樣本試卷以來每一道 STEP 題目的詳細解答、*Beyond STEP* 全書以及各章的單行本，均可在 `beyondstep.keithmcnulty.org` 免費閱讀及下載。

閱讀解答之前，請先認真嘗試解題。每個答案均已用數值方法核對，但解答數量如此之多，難免會有疏漏：如果發現有可疑之處，請相信自己的演算，並仔細檢查。

由機器從英文翻譯而成；尚未經審閱。

2026 年 11 月首次於網上發表。

本 PDF 內凡屬作者本人創作的內容，包括正文、解答、插圖、練習及答案，均按照 Creative Commons CC0 1.0 Universal 授權條款貢獻予公有領域：你可以為任何目的複製、改編及分享，毋須徵求許可。詳見 `creativecommons.org/publicdomain/zero/1.0`。

本書討論的 STEP 試題，版權屬 Cambridge University Press & Assessment 所有。本書並無轉載這些試題：每一題均附連結，通往官方來源的題目，即 STEP Support Programme 的試題資料庫或 OCR 的歷屆試卷。

如有錯漏，歡迎到 github.com/keithmcnulty/step_book_mcnulty 提交 issue 指正。

目錄

STEP I

第 1 題 三角恆等式 純數學 . . . 3

第 2 題 變率 純數學 . . . 6

第 3 題 複數等比級數 純數學 . . . 10

第 4 題 幾何中的最優化 純數學 . . . 14

第 5 題 正則向量組 純數學 . . . 18

第 6 題 駐點與廣義積分 純數學 . . . 22

第 7 題 可分離變量微分方程 純數學 . . . 26

第 8 題 代換積分法 純數學 . . . 30

第 9 題 不等式與調和級數 純數學 . . . 34

第 10 題 懸鏈的平衡 力學 . . . 37

第 11 題 滑輪與衝量 力學 . . . 41

第 12 題 斜面及摩擦力的平衡 力學 . . . 45

第 13 題 動量守恆 力學 . . . 49

第 14 題 角形的計數 概率與統計 . . . 53

第 15 題 態近似 概率與統計 . . . 57

第 16 題 條件概率 概率與統計 . . . 61

STEP II

第 1 題 二次方程與判別式 純數學 . . . 65

第 2 題 雙曲線的漸近線 純數學 . . . 69

第 3 題 對稱聯立方程 純數學 . . . 73

第 4 題 餘弦之和 純數學 . . . 76

第 5 題 積分歸約公式 純數學 . . . 80

第 6 題 積分的界限 純數學 . . . 84

第 7 題 函數方程 純數學 88
第 8 題 非線性一階常微分方程 純數學 92
第 9 題 矩陣群 純數學 96
第 10 題 三維空間中的直線與平面 純數學 99
第 11 題 拋體與發射點的距離 力學 103
第 12 題 受迫彈性振動 力學 106
第 13 題 剛體轉動 力學 110
第 14 題 相對速度與追蹤 力學 114
第 15 題 期望值的線性性質 概率與統計 118
第 16 題 項分佈的正態近似 概率與統計 122

STEP III

第 1 題 級數與展開式 純數學 126
第 2 題 幾何中的複數 純數學 129
第 3 題 分段函數與反函數 純數學 133
第 4 題 三維空間中的直線與平面 純數學 137
第 5 題 極坐標曲線與常微分方程 純數學 141
第 6 題 雙曲線性變換 純數學 145
第 7 題 利用對稱性積分 純數學 150
第 8 題 歸約公式與求和 純數學 154
第 9 題 橢圓的弧長 純數學 158
第 10 題 牛頓恆等式 純數學 162
第 11 題 剛體平衡 力學 165
第 12 題 曲線上的運動 力學 169
第 13 題 拋體與天花板的碰撞 力學 173
第 14 題 擺及彈簧的圓周運動 力學 177
第 15 題 樣本中配對的計數 概率與統計 181
第 16 題 聯合分佈與邊際分佈 概率與統計 185

詳細解答

STEP I 1991，第 1 題

純數學 • 三角恆等式

> **STEP I 1991，第 1 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q1

關鍵思路

角度所滿足的條件讓我們可以消去其中一個角：由於 $\psi = \frac{1}{2}\pi - (\theta + \phi)$，我們有 $\sin\psi = \cos(\theta + \phi)$，於是該恆等式（identity）便成為一個關於兩個自由角的命題。其後，乘積 $\sin\theta\sin\phi$ 與 $\cos(\theta + \phi)$ 結合起來得出 $\cos(\theta - \phi)$，而乘積 $\cos(\theta + \phi)\cos(\theta - \phi)$ 的形式十分簡潔。至於第二部分，所選的特殊角令 ψ 等於 θ 的一半，有關 $\frac{1}{10}\pi$ 的資料正是由此而來。

解答

恆等式

由於 $\theta + \phi + \psi = \frac{1}{2}\pi$，

$$\sin\psi = \sin(\tfrac{1}{2}\pi - (\theta + \phi)) = \cos(\theta + \phi).$$

因此恆等式的左方是

$$L = \sin^2\theta + \sin^2\phi + \cos(\theta + \phi)[\cos(\theta + \phi) + 2\sin\theta\sin\phi].$$

在括號之內，$\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$，所以

$$\cos(\theta + \phi) + 2\sin\theta\sin\phi = \cos\theta\cos\phi + \sin\theta\sin\phi = \cos(\theta - \phi).$$

現在我們需要乘積 $\cos(\theta+\phi)\cos(\theta-\phi)$。把複角公式（compound-angle formulae）乘開，

$$\begin{aligned}\cos(\theta+\phi)\cos(\theta-\phi) &= (\cos\theta\cos\phi-\sin\theta\sin\phi)(\cos\theta\cos\phi+\sin\theta\sin\phi)\\ &= \cos^2\theta\cos^2\phi-\sin^2\theta\sin^2\phi\\ &= \cos^2\theta\,(1-\sin^2\phi)-(1-\cos^2\theta)\sin^2\phi\\ &= \cos^2\theta-\sin^2\phi.\end{aligned}$$

因此

$$L=\sin^2\theta+\sin^2\phi+\cos^2\theta-\sin^2\phi=\sin^2\theta+\cos^2\theta=1,$$

這就是所需的恆等式。

$\sin\frac{1}{10}\pi$ 所滿足的三次方程

取 $\theta=\phi=\frac{1}{5}\pi$。那麼 $\psi=\frac{1}{2}\pi-\frac{2}{5}\pi=\frac{1}{10}\pi$。設

$$x=\sin\tfrac{1}{10}\pi,\qquad s=\sin\tfrac{1}{5}\pi.$$

由恆等式得

$$2s^2+x^2+2s^2x=1,\qquad \text{即}\qquad 2s^2(1+x)=1-x^2. \tag{$*$}$$

> **想一想** 為甚麼不直接利用 $\sin\psi=\cos 2\theta=1-2s^2$ 來消去 s？

這個關係式對每一個滿足 $\psi=\frac{1}{2}\pi-2\theta$ 的 θ 都成立，所以把 $s^2=\frac{1}{2}(1-x)$ 代入 $(*)$ 只會得出 $(1-x)(1+x)=1-x^2$：這是對的，卻毫無內容。我們必須用上 $\frac{1}{5}\pi$ 的某種特殊性質。這個特殊之處就是 $\frac{1}{5}\pi=2\cdot\frac{1}{10}\pi$，所以 $\theta=2\psi$，而二倍角公式（double-angle formula）把 s 與 x 聯繫起來：

$$s=\sin 2\psi=2\sin\psi\cos\psi,\qquad \text{所以}\qquad s^2=4x^2(1-x^2).$$

把這個結果代入 $(*)$，

$$8x^2(1-x^2)(1+x)=1-x^2.$$

由於 $0<\frac{1}{10}\pi<\frac{1}{2}\pi$，我們有 $0<x<1$，所以 $1-x^2\neq 0$，可以把它約去：

$$8x^2(1+x)=1,\qquad \text{即}\qquad 8x^3+8x^2-1=0.$$

所以 $\sin\frac{1}{10}\pi$ 滿足題目給出的三次方程（cubic），證明完畢。

驗算

恆等式可以用 $\theta = \frac{1}{2}\pi$、$\phi = \psi = 0$ 來檢驗：左方是 $1+0+0+0=1$。取 $\theta = \phi = \psi = \frac{1}{6}\pi$ 時，左方是 $3\cdot\frac{1}{4} + 2\cdot\frac{1}{8} = 1$。

至於三次方程，$\sin\frac{1}{10}\pi = \sin 18^\circ \approx 0.309017$，而 $8x^3 + 8x^2 - 1$ 在該處的值在捨入誤差範圍內等於 0。事實上，這個三次式可以因式分解為 $(2x+1)(4x^2+2x-1)$，其根為 $-\frac{1}{2}$ 和 $\frac{1}{4}(-1\pm\sqrt{5})$，而 $\frac{1}{4}(\sqrt{5}-1) \approx 0.309017$ 正是大家熟悉的 $\sin 18^\circ$ 的準確值。

常見錯誤

- 只用 $\sin\psi = \cos 2\theta$ 來消去 s，結果得到恆等式 $0=0$；見上面的方框。
- 除以 $1-x^2$ 卻沒有說明它為何不等於零。這裏 $x = \sin\frac{1}{10}\pi$ 嚴格介乎 0 與 1 之間。

STEP 之外

三次方程的另外兩個根並非偶然：$-\frac{1}{2} = \sin\left(-\frac{1}{6}\pi\right)$，而 $\frac{1}{4}(-1-\sqrt{5}) = -\sin\frac{3}{10}\pi$。這類方程記錄了 sin 和 cos 在 π 的有理數倍處的取值之間的代數關係，而要理解這些關係，最好的途徑是單位根（roots of unity）：$\cos\frac{1}{5}\pi$ 與五次單位根，即 $z^5 = 1$ 的解，有密切的關係。這就是為甚麼以直尺和圓規作正五邊形，歸根究柢取決於 $\sqrt{5}$。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）是從幾何角度認識這些想法的絕佳讀物。

詳細解答

STEP I 1991，第 2 題

純數學 • 變率

> **STEP I 1991，第 2 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q2

關鍵思路

把融化規律改用半徑來表達。對球體而言，$\frac{\mathrm{d}V}{\mathrm{d}t} = 4\pi r^2 \frac{\mathrm{d}r}{\mathrm{d}t}$，而表面積同樣是 $4\pi r^2$，所以「體積減少的速率與表面積成正比」其實只是說半徑以恆定速率縮短。兩個雪球的比例常數相同，因此在任何一段時間內，兩個半徑縮短的長度都相同。於是所有量都變成單一變量的函數（function），這個變量就是每個半徑縮短了的長度 x。

解答

兩個半徑以相同的恆定速率縮短

設雪球的半徑為 r，體積為 $V = \frac{4}{3}\pi r^3$，表面積為 $S = 4\pi r^2$。融化規律是 $\frac{\mathrm{d}V}{\mathrm{d}t} = -kS$，其中 $k > 0$ 為常數，兩個雪球的常數相同。由鏈式法則（chain rule），

$$4\pi r^2 \frac{\mathrm{d}r}{\mathrm{d}t} = -k \cdot 4\pi r^2 \quad \Longrightarrow \quad \frac{\mathrm{d}r}{\mathrm{d}t} = -k,$$

這裏兩邊除以 $4\pi r^2$，只要雪球仍然存在（$r > 0$），這一步便成立。所以經過時間 t 後，兩個半徑都減少了相同的長度 $x = kt$：頭部的半徑為 $2R - x$，身體的半徑為 $3R - x$，適用於 $0 \le x \le 2R$（頭部在 $x = 2R$ 時消失）。

高度減半時的體積

由於頭部放在身體之上，Frosty 的高度是兩個直徑之和：

$$h = 2(2R - x) + 2(3R - x) = 10R - 4x.$$

最初 $h = 10R$，所以高度減半即 $10R - 4x = 5R$，也就是 $x = \frac{5}{4}R$。這個值小於 $2R$，所以頭部仍然存在，此時兩個半徑分別為 $\frac{3}{4}R$ 和 $\frac{7}{4}R$。因子 $\frac{4}{3}\pi$ 在比值中約去，所以

$$\frac{V}{V_0} = \frac{\left(\frac{3}{4}\right)^3 + \left(\frac{7}{4}\right)^3}{2^3 + 3^3} = \frac{(27 + 343)/64}{35} = \frac{370}{2240} = \frac{37}{224}.$$

所以當 Frosty 的高度減半時，他的體積只剩下原來的 $\frac{37}{224} \approx 0.165$：體積隨長度的三次方而變化。

$\mathrm{d}V/\mathrm{d}S$ 的最大值

記 $a = 2R - x$ 及 $b = 3R - x$。當 $0 \le x < 2R$ 時，

$$V = \tfrac{4}{3}\pi(a^3 + b^3), \qquad S = 4\pi(a^2 + b^2),$$

又因為 $\frac{\mathrm{d}a}{\mathrm{d}x} = \frac{\mathrm{d}b}{\mathrm{d}x} = -1$，

$$\frac{\mathrm{d}V}{\mathrm{d}x} = -4\pi(a^2 + b^2), \qquad \frac{\mathrm{d}S}{\mathrm{d}x} = -8\pi(a + b).$$

想一想 V 和 S 都是時間（或 x）的函數。那麼 $\mathrm{d}V/\mathrm{d}S$ 是甚麼意思？

當 x 增加時，S 嚴格遞減（因為 $\frac{\mathrm{d}S}{\mathrm{d}x} < 0$），所以 V 可以看成 S 的函數。由鏈式法則 $\frac{\mathrm{d}V}{\mathrm{d}S} = \frac{\mathrm{d}V}{\mathrm{d}x} \Big/ \frac{\mathrm{d}S}{\mathrm{d}x}$：

$$\frac{\mathrm{d}V}{\mathrm{d}S} = \frac{a^2 + b^2}{2(a + b)}.$$

現在利用 $b - a = R$，這個關係始終不變。設 $s = a + b = 5R - 2x$。那麼 $a^2 + b^2 = \frac{1}{2}\big((a+b)^2 + (b-a)^2\big) = \frac{1}{2}(s^2 + R^2)$，所以

$$\frac{\mathrm{d}V}{\mathrm{d}S} = \frac{s^2 + R^2}{4s} = \frac{s}{4} + \frac{R^2}{4s} = f(s), \text{ 記為此.}$$

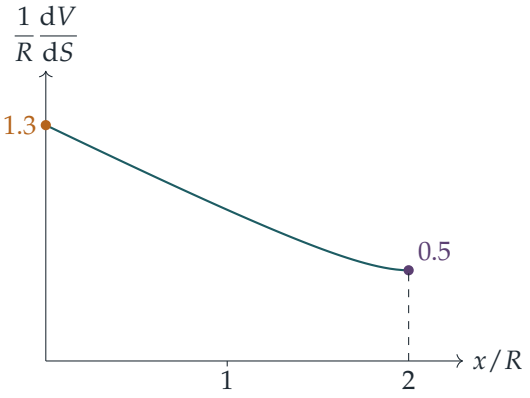
當 x 由 0 增至 $2R$ 時，s 由 $5R$ 減至 R。而

$$f'(s) = \frac{1}{4} - \frac{R^2}{4s^2} = \frac{s^2 - R^2}{4s^2} > 0 \quad \text{當} s > R,$$

所以 f 在 $[R, 5R]$ 上遞增，其最大值在 $s = 5R$ 處取得，也就是在最開始的時候（$x = 0$）：

$$\left(\frac{\mathrm{d}V}{\mathrm{d}S}\right)_{\max} = \frac{5R}{4} + \frac{R}{20} = \frac{13R}{10}.$$

（直接計算：$\frac{4R^2 + 9R^2}{2 \cdot 5R} = \frac{13}{10}R$。）其後這個值持續下降，到頭部剛好消失時降至 $f(R) = \frac{1}{2}R$。



驗算

- 對每個雪球按 $\frac{dV}{dt} = -kS$ 逐步模擬（取 $R = k = 1$），在 $t \approx 1.25$ 時高度減半，體積比為 0.16518，與 $\frac{37}{224} = 0.165178\ldots$ 吻合。
- 模擬所得的 V 和 S 的有限差分（finite differences）自始至終與 $\frac{a^2+b^2}{2(a+b)}$ 相符，而且從不超過 $1.3R$。
- 單位：dV/dS 是體積除以面積，所以是長度，$\frac{13}{10}R$ 正是如此。

常見錯誤

- 把高度當成兩個半徑之和，而不是兩個直徑之和。（在這題中碰巧不會改變 x 的值，但做法是錯的。）
- 計算了 V/S 而不是 dV/dS；對單一球體而言，$V/S = r/3$，但 $dV/dS = r/2$。
- 一味尋找駐點（stationary point）：f 的轉向點在 $s = R$，即區間的端點，所以最大值在區間的另一端。

STEP 之外

這道題的核心在於：「速率與表面積成正比」這條規律，會化為半徑以恆定速率變化，這是幾何如何令微分方程（differential equation）變得輕而易舉的第一個例子。後半部分是參數式導數（parametric derivative），即沿曲線 $(S(t), V(t))$ 的 dV/dS，以及在閉區間（closed interval）上求最大值時須有的謹慎：最大值可能位於端點，而不是駐點；Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門）解釋了為何這樣的最大值必定存在。至於以微分方程建立模型，可參考 Boyce 與 DiPrima 的 *Elementary Differential*

Equations（初等微分方程）。

詳細解答

STEP I 1991，第 3 題

純數學 • 複數等比級數

> **STEP I 1991，第 3 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q3

關鍵思路

在阿根圖（Argand diagram）上，乘以 $re^{i\phi}$ 會把一個向量放大 r 倍，並按逆時針方向轉過角 ϕ。路徑上每一條線段都是前一條線段按相同的比例縮放、按相同的角度旋轉而成，所以每個線段向量都等於前一個乘以一個固定的複數（complex number）。於是這些線段構成一個等比數列（geometric progression），把它們加起來便得出每個頂點的位置。

解答

相鄰線段之比

由 P_n 到 P_{n+1} 的向量由 $z_{n+1}-z_n$ 表示，而由 P_{n-1} 到 P_n 的向量則由 z_n-z_{n-1} 表示。兩個複數相除時，模（modulus）相除，輻角（argument）相減：

$$\left|\frac{z_{n+1}-z_n}{z_n-z_{n-1}}\right| = \frac{P_nP_{n+1}}{P_{n-1}P_n} = d, \qquad \arg\frac{z_{n+1}-z_n}{z_n-z_{n-1}} = \arg(z_{n+1}-z_n) - \arg(z_n-z_{n-1}) = \theta,$$

因為第二條線段的長度是第一條的 d 倍，而由一條有向線段的方向按逆時針轉到下一條的方向，所轉的角是 θ。（輻角只確定到相差 2π 的倍數，但這並不影響該複數本身。）因此，對所有 $n \geq 2$，

$$\frac{z_{n+1}-z_n}{z_n-z_{n-1}} = d\,\mathrm{e}^{\mathrm{i}\theta} = d(\cos\theta + \mathrm{i}\sin\theta).$$

z_{n+1} 的公式

設 $w = de^{i\theta}$。第一條線段是 $z_2 - z_1 = 1$，而其後每一條線段都是前一條的 w 倍，所以由歸納法（induction）可得

$$z_{k+1} - z_k = w^{k-1} \qquad (k \geq 1).$$

把 $k = 1, 2, \ldots, n$ 的各式相加，左邊逐項相消（telescoping），只剩下 $z_{n+1} - z_1 = z_{n+1}$：

$$z_{n+1} = 1 + w + w^2 + \cdots + w^{n-1} = \frac{w^n - 1}{w - 1} = \frac{d^n e^{in\theta} - 1}{de^{i\theta} - 1}.$$

求等比級數（geometric series）之和時需要 $w \neq 1$，而 $d \neq 1$ 這個條件正是用在這裏，因為 $|w| = d$。

$\theta = \frac{1}{3}\pi$、$d = 2$ 的情況

這時 $w = 2e^{i\pi/3} = 1 + i\sqrt{3}$，所以分母非常簡單：$w - 1 = i\sqrt{3}$。由於 $\frac{1}{i} = -i$，

$$z_{n+1} = \frac{w^n - 1}{i\sqrt{3}} = -\frac{i}{\sqrt{3}}\left(2^n \cos\tfrac{n\pi}{3} - 1 + i\,2^n \sin\tfrac{n\pi}{3}\right),$$

從而

$$\operatorname{Re} z_{n+1} = \frac{2^n \sin\frac{n\pi}{3}}{\sqrt{3}}, \qquad \operatorname{Im} z_{n+1} = \frac{1 - 2^n \cos\frac{n\pi}{3}}{\sqrt{3}}.$$

想一想 $\operatorname{Im} z_{n+1}$ 的正負號怎樣告訴我們路徑穿過實軸（real axis）？

這條路徑是連貫的：它是由線段首尾相接而成的一條連續折線。如果某個頂點嚴格位於實軸下方，而其後某個頂點嚴格位於實軸上方（或者反過來），那麼兩者之間的那段路徑必定穿過實軸。所以我們要一次又一次地找出位於實軸兩側的頂點。

$\cos\frac{n\pi}{3}$ 的值以 6 為週期重複。取任意整數 $k \geq 1$。

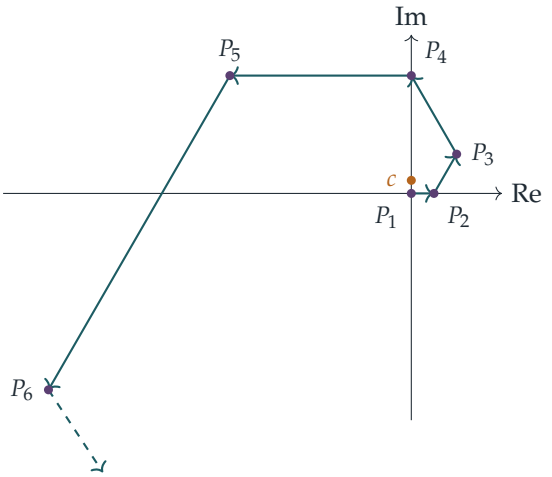
- 當 $n = 6k$ 時，$\cos\frac{n\pi}{3} = \cos 2k\pi = 1$，所以 $\operatorname{Im} z_{6k+1} = \dfrac{1 - 2^{6k}}{\sqrt{3}} < 0$，因為 $2^{6k} > 1$。
- 當 $n = 6k + 3$ 時，$\cos\frac{n\pi}{3} = \cos(2k\pi + \pi) = -1$，所以 $\operatorname{Im} z_{6k+4} = \dfrac{1 + 2^{6k+3}}{\sqrt{3}} > 0$。

因此，對所有 $k \geq 1$，P_{6k+1} 在實軸下方，而 P_{6k+4} 在實軸上方。所以對每個 $k = 1, 2, 3, \ldots$，路徑由 P_{6k+1} 到 P_{6k+4} 的那一段都會由下而上穿過實軸。這些是路徑上互不相同的部分，每個 k 對應一段，所以路徑穿過實軸無限多次。（按同樣的論證，路徑在 P_{6k+4} 與 P_{6k+7} 之間也會由上而下穿過實軸。）證明至此完成。

事實上，頂點是每三個一組地繞圈的：$\operatorname{Im} z_1, \operatorname{Im} z_2, \ldots$ 的正負號規律是 $0, 0, +, +, +, -, -, -, +, +, +, .$ 這背後有一個巧妙的幾何原因。把上式寫成

$$z_{n+1} = \frac{i}{\sqrt{3}} - \frac{i}{\sqrt{3}}\,w^n,$$

可見 P_{n+1} 與定點 $c = \mathrm{i}/\sqrt{3}$ 的距離是 $2^n/\sqrt{3}$，而每走一步，它便繞 c 轉過 $60°$。各頂點位於一條環繞 c 的螺線上，這條螺線每走一步便放大一倍，而 c 剛好在實軸上方不遠處，所以螺線會不斷掃過實軸。



圖中顯示 $P_1, \ldots, P_6$；下一個頂點是 $P_7 = -21\sqrt{3}\,\mathrm{i}$，遠在下方。路徑第一次穿過實軸是在 P_5P_6 上，其後一直留在實軸下方，直至在 P_8P_9 上由下而上穿過，即由 $P_8 = 64 - 21\sqrt{3}\,\mathrm{i}$ 到 $P_9 = 128 + 43\sqrt{3}\,\mathrm{i}$。

驗算

- 直接由幾何關係計算：$z_3 = 1 + w = 2 + \sqrt{3}\,\mathrm{i}$，而公式給出 $\operatorname{Im} z_3 = (1 - 4\cos\frac{2\pi}{3})/\sqrt{3} = 3/\sqrt{3} = \sqrt{3}$。同樣地，再加上 $w^2 = -2 + 2\sqrt{3}\,\mathrm{i}$ 和 $w^3 = -8$，得 $z_5 = -8 + 3\sqrt{3}\,\mathrm{i}$，而在公式中取 $n = 4$，得 $\operatorname{Re} z_5 = 16\sin\frac{4\pi}{3}/\sqrt{3} = -8$ 和 $\operatorname{Im} z_5 = (1+8)/\sqrt{3} = 3\sqrt{3}$。
- 若 $d = 1$，各頂點會落在一個正多邊形上，路徑會閉合；等比級數的求和公式恰好在這個情況下失效，這正是應有的結果。
- 我們按長度和轉角以數值方法建構了這條路徑（d、θ 和起點均隨機選取），並確認了比值 $d\mathrm{e}^{\mathrm{i}\theta}$、$z_{n+1}$ 的閉合式（closed form），以及在 $\theta = \frac{1}{3}\pi$、$d = 2$ 時上述的正負號規律：在 $z_1, \ldots, z_{41}$ 之中，$\operatorname{Im} z$ 共變號 13 次。

常見錯誤

- 把比值寫成 $\frac{1}{d}\mathrm{e}^{-\mathrm{i}\theta}$：分子是較後的那條線段，所以它才是長 d 倍、並轉過 $+\theta$ 的那一條。

- 求和時項數差一：由 P_1 到 P_{n+1} 共有 n 條線段，所以級數加到 w^{n-1} 為止。
- 單憑一個位於軸上的頂點便聲稱路徑穿過實軸：你需要的是嚴格位於實軸兩側的頂點（這裏 z_1 和 z_2 都在實軸上，所以要由 $k \geq 1$ 開始）。

STEP 之外

規則 $z_{n+1} - z_n = w(z_n - z_{n-1})$ 是一個線性遞歸關係（linear recurrence），而把每個頂點帶到下一個頂點的映射 $z \mapsto c + w(z - c)$，是以其不動點（fixed point）c 為中心的旋轉相似變換（spiral similarity）：即旋轉與放大的結合。這類映射是莫比烏斯變換（Möbius transformation）最簡單的例子，把它們反覆迭代，便會產生貝殼上常見的等角螺線（又稱對數螺線，equiangular (logarithmic) spiral）。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）以這種幾何眼光看待複數乘法，闡述得十分精彩。

詳細解答

STEP I 1991，第 4 題

純數學 • 幾何中的最優化

> **STEP I 1991，第 4 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q4

關鍵思路

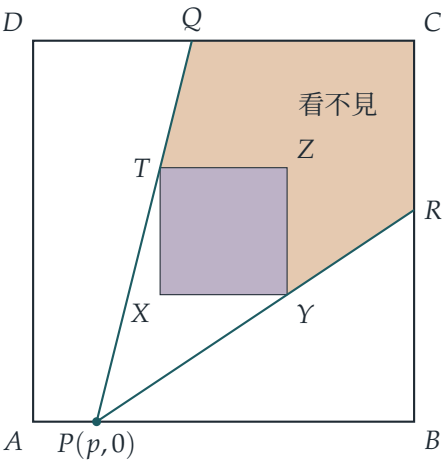
守衛能看見庭院中的每一點，只有被塔樓擋住的點除外，所以我們按照提示，求出從他所在位置 $P=(p,0)$ 望去時塔樓「陰影」的面積。陰影的邊界是由 P 出發、剛好擦過塔樓的兩條射線。求出這兩條射線與圍牆的交點，把陰影當作多邊形計算其面積，問題便化為在 $[0,4]$ 上求一個單變數函數的最大值。記得檢查端點。

解答

建立模型

庭院是大正方形減去塔樓，面積為 $144-16=128$。設守衛位於 $P=(p,0)$，其中 $0\le p\le 4$，並設 $H(p)$ 為他看不見的面積。可見面積為 $128-H(p)$，所以 H 最大時可見面積最小，反之亦然。

由於 $p\le 4$，P 位於直線 $y=4$ 之下，並在直線 $x=4$ 的左方（或在該直線上）。因此從 P 望去，塔樓露出它的底面 XY 和左面 XT，而它的輪廓由頂角 $T(4,8)$ 和 $Y(8,4)$ 決定。頂角 $Z(8,8)$ 位於角 TPY 之內，是看不見的。看不見的區域位於路徑 $T\to Z\to Y$ 之外，夾在射線 PT 與 PY 之間。



擦邊射線與圍牆的交點

由 P 出發穿過 $T(4,8)$ 的射線，每橫移 $4-p$ 便上升 8，所以它到達 $y=12$ 時

$$x = p + \tfrac{12}{8}(4-p) = 6 - \tfrac{p}{2}, \qquad \text{由此得} Q = \left(6-\tfrac{p}{2},\ 12\right).$$

由於 $4 \le 6-\frac{p}{2} \le 6$，這一點確實在上方的圍牆 DC 上。

由 P 出發穿過 $Y(8,4)$ 的射線，每橫移 $8-p$ 便上升 4，所以它到達 $x=12$ 時

$$y = \frac{4(12-p)}{8-p}, \qquad \text{由此得} R = \left(12,\ \frac{4(12-p)}{8-p}\right).$$

當 p 由 o 增至 4 時，這個值由 6 增至 8，所以 R 在右方的圍牆 BC 上。

看不見的面積

看不見的區域是多邊形 $TQCRYZ$。求其面積最簡便的方法是

$$H = [PRCQ] - [PYZT],$$

其中 $[\cdots]$ 表示面積：即兩條射線之間一直伸延至圍牆的整個楔形，減去楔形中位於塔樓遠側兩面之前的部分（塔樓前方看得見的地面，連同塔樓本身）。

對 $P(p,0)$、$R(12,q)$、$C(12,12)$、$Q(r,12)$ 使用鞋帶公式（shoelace formula），其中 $q = \frac{4(12-p)}{8-p}$ 及 $r = 6-\frac{p}{2}$：

$$\begin{aligned}[PRCQ] &= \tfrac{1}{2}\big(pq + (144-12q) + (144-12r) - 12p\big) \\ &= \tfrac{1}{2}\big(288 - 12p - 12r - (12-p)q\big) = 108 - 3p - \frac{2(12-p)^2}{8-p},\end{aligned}$$

這裏用了 $12r = 72 - 6p$。同樣地，對 $P(p,0)$、$Y(8,4)$、$Z(8,8)$、$T(4,8)$：

$$[\,PYZT\,] = \tfrac{1}{2}\big(4p + 32 + 32 - 8p\big) = 32 - 2p.$$

因此

$$H(p) = 76 - p - \frac{2(12-p)^2}{8-p}.$$

? **想一想** 有沒有一個代換（substitution）能令這個函數容易求最大、最小值？

分母不好處理，這提示我們設 $u = 8 - p$，於是 u 的取值範圍是 $[4,8]$，而 $12 - p = u + 4$。這樣

$$H = 76 - (8-u) - \frac{2(u+4)^2}{u} = 68 + u - \frac{2u^2 + 16u + 32}{u} = 52 - \left(u + \frac{32}{u}\right).$$

第 (i) 部分：最小可見面積

我們需要 H 盡可能大，即 $u + \frac{32}{u}$ 盡可能小。根據算術平均—幾何平均不等式（AM–GM inequality），對 $u > 0$，

$$u + \frac{32}{u} \;\geq\; 2\sqrt{32} = 8\sqrt{2},$$

等號恰好在 $u = \frac{32}{u}$ 時成立，即 $u = 4\sqrt{2}$。這個值在 $[4,8]$ 之內（因為 $4\sqrt{2} \approx 5.66$），所以是容許的。於是 $p = 8 - u$，而當守衛位於以下位置時，可見面積最小：

$$\boxed{p = 8 - 4\sqrt{2} \;\approx 2.34 \text{ 個單位，距離點} A}$$

在這一點，他看不見的面積是 $52 - 8\sqrt{2}$，所以他看見 $76 + 8\sqrt{2} \approx 87.3$ 平方單位。

第 (ii) 部分：最大可見面積

現在我們需要 H 盡可能小，即 $f(u) = u + \frac{32}{u}$ 在 $[4,8]$ 上盡可能大。由於 $f'(u) = 1 - \frac{32}{u^2}$，$f$ 在 $[4, 4\sqrt{2}]$ 上遞減，在 $[4\sqrt{2}, 8]$ 上遞增，所以它的最大值在端點取得。我們有

$$f(4) = 4 + 8 = 12, \qquad f(8) = 8 + 4 = 12,$$

兩者相等。所以 H 在 $u = 8$ 和 $u = 4$ 兩處，即 $p = 0$ 和 $p = 4$ 兩處，都取得其最小值 $52 - 12 = 40$。可見面積最大，為 $128 - 40 = 88$ 平方單位，此時守衛站在

$$\boxed{\text{位於} A\ (p = 0) \text{ 或} E\ (p = 4)}$$

而 AE 上再沒有其他點能做到。

驗算

- 當 $p = 0$ 時，我們可以直接手算陰影：兩條射線與圍牆相交於 $(6, 12)$ 和 $(12, 6)$，而看不見的六邊形面積為 40，與 $H(0)$ 吻合。
- 當 $p = 4$ 時，穿過 T 的射線沿直線 $x = 4$ 筆直向上，而 $H(4) = 40$ 再次成立。兩個端點的值相等是個好兆頭：公式沒有偏向一邊。
- 我們在 600×600 的網格上點算看不見的點，逐一檢驗由 P 出發的視線是否穿過塔樓。當 $p = 8 - 4\sqrt{2}$ 時，結果是 40.687，而 $52 - 8\sqrt{2} = 40.686$。當 $p = 1$ 時，結果是 40.429，而 $H(1) = \frac{283}{7} \approx 40.429$。

常見錯誤

- 用 Z 而不是 T 或 Y 作為擦邊點。從左下方望去，塔樓的輪廓由 T 和 Y 這兩個頂角決定，而不是由遠處的那個頂角決定。
- 做第 (ii) 部分時只求出駐點（stationary point）便停下來。駐點給出的是最小可見面積，所以最大值必定來自端點，而這裏有兩個。

STEP 之外

這是可見性（visibility）問題的一個小例子，這類問題屬於計算幾何（computational geometry）的研究範圍（「美術館問題」（art gallery problem）問的是：要看守一個多邊形房間的每一點，需要多少名守衛）。當中求最大、最小值的一步，展示了分析學中的一個基本事實：閉區間上的連續函數必定取得其最大值和最小值，取得之處不是駐點便是端點。這就是極值定理（extreme value theorem），其嚴謹證明可見於 Abbott, *Understanding Analysis*。

詳細解答

STEP I 1991，第 5 題

純數學 • 正則向量組

STEP I 1991，第 5 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s1-q5

關鍵思路

條件 (ii) 說這些向量（vector）之和為 $\mathbf{0}$，而要把一條向量等式化為一個數，最省力的方法就是與其中一個向量取點積（dot product），條件 (i) 正好告訴我們怎樣計算這些點積。此後，整道題目其實是在點算數目：若一個向量與某幾個固定向量的點積全部已被指定，它可處的位置就寥寥無幾，所以集合內只容得下少數幾個向量。

解答

恆等式 $(n-1)\beta = -\alpha^2$

由 (ii)，$\mathbf{a}_1 + \mathbf{a}_2 + \cdots + \mathbf{a}_n = \mathbf{0}$。將等式兩邊與 $\mathbf{a}_1$ 取點積。由 (i)，$\mathbf{a}_1 \cdot \mathbf{a}_1$ 一項等於 α^2，而其餘 $n-1$ 項 $\mathbf{a}_1 \cdot \mathbf{a}_j$（$j \neq 1$）每項均為 β，所以

$$\alpha^2 + (n-1)\beta = 0, \qquad \text{即} \qquad (n-1)\beta = -\alpha^2,$$

這正是所要證明的。特別地，$\beta = -\alpha^2/(n-1) < 0$。

二維情況

現在 $\alpha^2 = \mathbf{a}_1 \cdot \mathbf{a}_1 = 1$，所以每個 $\mathbf{a}_i$ 都是單位向量（unit vector），並且

$$\beta = -\frac{1}{n-1}.$$

對 $i \geq 2$，設 $\mathbf{a}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$。條件 $\mathbf{a}_1 \cdot \mathbf{a}_i = \beta$ 即是 $x_i = \beta$，於是 $|\mathbf{a}_i| = 1$ 迫使 $y_i = \pm\sqrt{1-\beta^2}$。因此，除 $\mathbf{a}_1$ 以外的每個向量都必定是以下兩者之一：

$$\begin{pmatrix} \beta \\ \sqrt{1-\beta^2} \end{pmatrix}, \qquad \begin{pmatrix} \beta \\ -\sqrt{1-\beta^2} \end{pmatrix}.$$

想一想 知道了候選向量之後，怎樣限制 n 的大小？

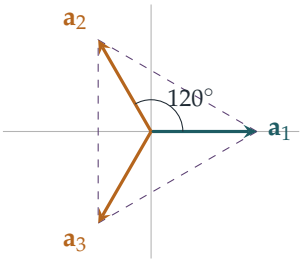
集合中的向量互不相同，所以 $\mathbf{a}_2, \ldots, \mathbf{a}_n$ 是 $n-1$ 個不同的候選向量。候選向量最多只有兩個，所以 $n-1 \leq 2$，即 $n \leq 3$。由於 $n \geq 2$，故 $n = 2$ 或 $n = 3$。

$n = 2$。此時 $\beta = -1$，兩個候選向量重合，$\mathbf{a}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$。這確實可行：$\mathbf{a}_1 \cdot \mathbf{a}_2 = -1 = \beta$，而且 $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{0}$。

$n = 3$。此時 $\beta = -\frac{1}{2}$，$\sqrt{1-\beta^2} = \frac{\sqrt{3}}{2}$，而 $\mathbf{a}_2, \mathbf{a}_3$ 必須是那兩個不同的候選向量：

$$\begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \quad \text{和} \quad \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}$$

（次序不拘）。驗算：$\mathbf{a}_2 \cdot \mathbf{a}_3 = \frac{1}{4} - \frac{3}{4} = -\frac{1}{2} = \beta$，而三個向量之和為 $\mathbf{0}$。它們指向一個以原點為中心的等邊三角形（equilateral triangle）的頂點。



（確定 β 的另一個方法：當 $n \geq 3$ 時，我們有 $-1 < \beta < 0$，所以兩個候選向量確實不同，而 $\mathbf{a}_2, \mathbf{a}_3$ 必須分別是這兩個向量。於是 $\mathbf{a}_2 \cdot \mathbf{a}_3 = \beta^2 - (1-\beta^2) = 2\beta^2 - 1$，令它等於 β 便得 $(2\beta+1)(\beta-1) = 0$。所以 $\beta = -\frac{1}{2}$，這同樣表示 $n = 3$。）

三維情況

同樣有 $\alpha = 1$ 及 $\beta = -1/(n-1)$。$n = 2$ 的情況與之前一樣：$\beta = -1$，而滿足 $\mathbf{a}_1 \cdot \mathbf{a}_2 = -1$ 的單位向量只有 $\mathbf{a}_2 = -\mathbf{a}_1 = (-1, 0, 0)^{\mathrm{T}}$，它位於 x-y 平面上。所以 $\{\mathbf{a}_1, -\mathbf{a}_1\}$ 是其中一個答案。

現在假設 $n \geq 3$，於是 $-\frac{1}{2} \leq \beta < 0$。由於 $\mathbf{a}_2$ 位於 x-y 平面上，沿用二維情況的論證可得

$$\mathbf{a}_2 = \begin{pmatrix} \beta \\ \varepsilon\sqrt{1-\beta^2} \\ 0 \end{pmatrix}, \qquad \varepsilon = \pm 1,$$

而且因為 $\beta \neq \pm 1$，y 分量 $y_2 = \varepsilon\sqrt{1-\beta^2}$ 不等於零。

想一想 對於 $i \geq 3$，這些條件對 $\mathbf{a}_i$ 有甚麼要求？

對 $i \geq 3$，設 $\mathbf{a}_i = (x_i, y_i, z_i)^{\mathrm{T}}$。則

$$\begin{aligned} \mathbf{a}_1 \cdot \mathbf{a}_i = \beta &\implies x_i = \beta, \\ \mathbf{a}_2 \cdot \mathbf{a}_i = \beta &\implies \beta^2 + y_2 y_i = \beta \implies y_i = \frac{\beta(1-\beta)}{y_2}, \\ |\mathbf{a}_i| = 1 &\implies z_i^2 = 1 - \beta^2 - y_i^2. \end{aligned}$$

所以對每個 $i \geq 3$，x_i 和 y_i 都相同，而 z_i 只能是兩個平方根的其中之一。由於各向量互不相同，$\mathbf{a}_3, \ldots, \mathbf{a}_n$ 當中最多只有兩個向量，所以 $n - 2 \leq 2$，即 $n \leq 4$。

$n = 3$。由 (ii)，$\mathbf{a}_3 = -\mathbf{a}_1 - \mathbf{a}_2$，它位於 x-y 平面上。所以這個集合是平面 $z = 0$ 上的一個正則集（regular set），根據二維情況的結果，它就是

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \\ 0 \end{pmatrix},$$

其中 $\mathbf{a}_2$ 可以是後兩者的任何一個。（這與上面的公式吻合：$\beta = -\frac{1}{2}$ 給出 $y_3 = -y_2$ 及 $z_3^2 = 1 - \frac{1}{4} - \frac{3}{4} = 0$。）

$n = 4$。此時 $\beta = -\frac{1}{3}$，所以 $\sqrt{1-\beta^2} = \frac{2\sqrt{2}}{3}$，而 $y_2 = \varepsilon\frac{2\sqrt{2}}{3}$。於是

$$y_i = \frac{(-\frac{1}{3})(\frac{4}{3})}{\varepsilon\frac{2\sqrt{2}}{3}} = -\varepsilon\frac{2}{3\sqrt{2}} = -\varepsilon\frac{\sqrt{2}}{3}, \qquad z_i^2 = 1 - \frac{1}{9} - \frac{2}{9} = \frac{2}{3}.$$

由於 $\mathbf{a}_3 \neq \mathbf{a}_4$，它們必須分別取兩個相反的值 $z = \pm\sqrt{2/3} = \pm\frac{\sqrt{6}}{3}$。由此得出

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} -\frac{1}{3} \\ \varepsilon\frac{2\sqrt{2}}{3} \\ 0 \end{pmatrix}, \quad \mathbf{a}_3, \mathbf{a}_4 = \begin{pmatrix} -\frac{1}{3} \\ -\varepsilon\frac{\sqrt{2}}{3} \\ \pm\frac{\sqrt{6}}{3} \end{pmatrix},$$

對應 $\varepsilon = \pm 1$ 的每個選擇各有一個集合。我們至今只用了涉及 $\mathbf{a}_1$ 和 $\mathbf{a}_2$ 的條件，所以仍須檢查其餘條件。首先，$\mathbf{a}_3 \cdot \mathbf{a}_4 = \frac{1}{9} + \frac{2}{9} - \frac{6}{9} = -\frac{1}{3} = \beta$。其次是向量之和：$x$ 分量給

出 $1-3\cdot\frac{1}{3}=0$，y 分量給出 $\varepsilon\left(\frac{2\sqrt{2}}{3}-\frac{\sqrt{2}}{3}-\frac{\sqrt{2}}{3}\right)=0$，而 z 分量互相抵銷。所以兩個集合都是正則集。從幾何上看，這四個向量指向一個以原點為中心的正四面體（regular tetrahedron）的頂點。

總結。所有正則集為：$\{\mathbf{a}_1,-\mathbf{a}_1\}$（$n=2$）；$x$-$y$ 平面上的等邊三角形（$n=3$）；以及上述兩個正四面體（$n=4$，$\varepsilon=\pm1$，兩者關於 x-z 平面互為鏡像）。不存在 $n\geq 5$ 的正則集。

驗算

在每個集合中，向量之和應為零，而且所有兩兩之間的點積都應相等。對於正四面體，任意兩個向量之間的夾角為 $\cos^{-1}(-\frac{1}{3})\approx 109.47°$，即大家熟悉的甲烷鍵角，而恆等式給出 $(4-1)(-\frac{1}{3})=-1=-\alpha^2$。我們以數值方法檢查了全部五個集合（格拉姆矩陣和形心（centroid）準確至 10^{-12}），並在平面上進行窮舉搜尋，在 n 不超過 11 的範圍內，只在 $n=2$ 和 $n=3$ 時找到正則集。

常見錯誤

- 忘記向量必須*互不相同*：正是這個條件，才把「最多兩個候選向量」變成 n 的上界。
- 在 $n=4$ 的情況中，利用 $\mathbf{a}_3,\mathbf{a}_4$ 與 $\mathbf{a}_1,\mathbf{a}_2$ 的點積求出它們之後，沒有檢查 $\mathbf{a}_3\cdot\mathbf{a}_4=\beta$ 和形心條件。
- 遺漏 $n=2$ 的情況，或遺漏三維情況中正負號 ε 的選擇。

STEP 之外

這些集合就是正單純形（regular simplex）的頂點：在 $(n-1)$ 維空間中的 n 個點，兩兩之間的距離全部相等。由點積組成的矩陣（matrix）$G=(\mathbf{a}_i\cdot\mathbf{a}_j)$（即*格拉姆矩陣*（Gram matrix））等於 $(\alpha^2-\beta)I+\beta J$，其中 J 是所有元素皆為一的矩陣。它的秩（rank）是 $n-1$，而格拉姆矩陣的秩不能超過空間的維數，由此立即得出在 d 維空間中 $n\leq d+1$。Strang 的 *Introduction to Linear Algebra*（線性代數入門教材）涵蓋了秩、點積和格拉姆矩陣。

詳細解答

STEP I 1991，第 6 題

純數學 • 駐點與廣義積分

STEP I 1991，第 6 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s1-q6

本題在 *Beyond STEP* 第 3.1 節找出錯處中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch03/spot-the-flaw.html

關鍵思路

甚麼都不要輕信：每一項計算都要親自重做一遍，並在每個「因此」或「所以」之處問一問，這一步是否真的成立。在 (i) 中，導數（derivative）算錯了；而且即使駐點（stationary point）求對了，其後關於極大值、極小值和全局界的推理仍然有毛病。在 (ii) 中，原函數（antiderivative）是錯的；更嚴重的是，被積函數（integrand）在區間內部趨向無窮，所以微積分基本定理（fundamental theorem of calculus）根本不能使用。

解答

第 (i) 部分

第 1 步：駐點。由商法則（quotient rule），

$$g'(x) = \frac{6x^2(x^4+4) - 4x^3(2x^3+3)}{(x^4+4)^2} = \frac{-2x^6 - 12x^3 + 24x^2}{(x^4+4)^2} = \frac{-2x^2\,(x^4+6x-12)}{(x^4+4)^2}.$$

所以 $g'(x) = 0$ 當且僅當 $x = 0$ 或 $h(x) = 0$，其中 $h(x) = x^4 + 6x - 12$。現在 $h(1) = -5$ 而 $h(-1) = -17$，所以 $x = \pm 1$ 並不是駐點：事實上 $g'(1) = \frac{2}{5}$ 而 $g'(-1) = \frac{34}{25}$。原來的論斷在第一步就已經不成立。

h 有多少個根？由於 $h'(x) = 4x^3 + 6$ 只在 $x_0 = -(3/2)^{1/3}$ 處為零，h 先遞減後遞增，所以它最多穿過零兩次。另外 $h(0) = -12 < 0$，而且當 $x \to \pm\infty$ 時 $h(x) \to +\infty$，所以

它最少穿過零兩次。因此 h 恰好有兩個實根 $\beta < 0 < \alpha$。由於 $h(1) < 0 < h(2) = 16$ 及 $h(-3) = 51 > 0 > h(-2) = -8$，

$$1 < \alpha < 2, \qquad -3 < \beta < -2,$$

而數值上 $\alpha \approx 1.386$，$\beta \approx -2.247$。

第 2 步：駐點處的函數值。$g(0) = \frac{3}{4}$ 是對的，$g(1) = 1$ 也是對的，但 $g(-1) = \frac{1}{5}$，而不是 1（何況 ± 1 這兩點根本都不是駐點）。在 h 的根 t 處，我們有 $t^4 = 12 - 6t$，所以 $t^3 = (12 - 6t)/t$，並且

$$g(t) = \frac{2(12-6t)/t + 3}{16 - 6t} = \frac{24 - 9t}{t(16-6t)} = \frac{3(8-3t)}{2t(8-3t)} = \frac{3}{2t},$$

其中約去了 $8 - 3t \neq 0$（因為 $t = \frac{8}{3}$ 不是 h 的根）。所以駐點處的函數值是 $g(0) = \frac{3}{4}$、$g(\alpha) = \frac{3}{2\alpha} \approx 1.083$ 和 $g(\beta) = \frac{3}{2\beta} \approx -0.668$。

> ? **想一想** 如果一個駐點處的函數值比另一個小，較小的那個就一定是極小值嗎？

第 3 步：駐點的性質。不一定。比較駐點處的函數值，完全不能告訴我們駐點的性質；我們必須考察 g' 在駐點兩側的正負號（或者考察 g''）。由因式分解後的形式可見，當 $x \neq 0$ 時，$g'(x)$ 的正負號與 $-h(x)$ 的正負號相同，因為 $x^2 > 0$ 而 $(x^4+4)^2 > 0$。由於在 (β, α) 上 $h < 0$，而在 $[\beta, \alpha]$ 之外 $h > 0$：

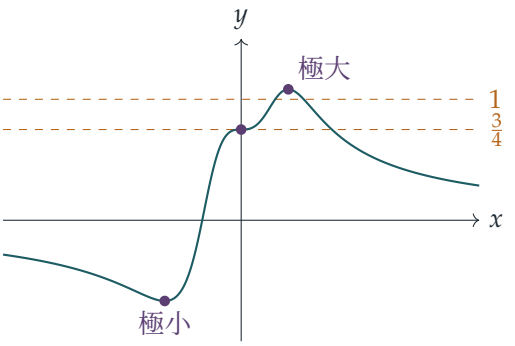
- g 在 $x < \beta$ 時遞減，在 (β, α) 上遞增，在 $x > \alpha$ 時遞減；
- $x = \beta$ 是極小點，而 $x = \alpha$ 是極大點；
- 在 $x = 0$ 處，兩側都有 $g' > 0$，所以 $x = 0$ 是駐拐點（stationary point of inflection），而不是極小點。

第 4 步：全局界。即使局部極大值和極小值都求對了，要得出「對所有 x」都成立的界，還需要更多論據：函數可以在別處超過某個局部極大值，例如當 $x \to \pm\infty$ 時。這裏當 $x \to \pm\infty$ 時 $g(x) \to 0$（分母的次數較高），而且由於 $x^4 + 4 > 0$，g 處處連續。把這些與第 3 步的單調性（monotonicity）結合起來，g 由接近 0 的值下降至 $g(\beta)$，再上升至 $g(\alpha)$，然後回落並趨向 0。因此正確的結論是

$$\frac{3}{2\beta} \leq g(x) \leq \frac{3}{2\alpha} \quad \text{對所有} x, \qquad \text{即約為} -0.668 \leq g(x) \leq 1.083,$$

而且兩個界都可以取到。

結論。$\frac{3}{4} \leq g(x) \leq 1$ 這個結論在兩端都不成立：$g(-1) = \frac{1}{5} < \frac{3}{4}$（其實只要 $2x^3 + 3 < 0$ 就有 $g(x) < 0$），而 $g(\alpha) > 1$ 是因為 $\alpha < \frac{3}{2}$（驗證：$h(\frac{3}{2}) = \frac{81}{16} - 3 > 0$，所以 $\alpha < \frac{3}{2}$，從而 $\frac{3}{2\alpha} > 1$）。



第 (ii) 部分

第 1 步：不定積分。題中給出的答案是錯的：對 $-3(1-x)^{-4}$ 求導得到 $-12(1-x)^{-5}$，而不是 $(1-x)^{-3}$。冪次應該增加 1，由 -3 變為 -2。由於

$$\frac{\mathrm{d}}{\mathrm{d}x}(1-x)^{-2} = -2(1-x)^{-3}\cdot(-1) = 2(1-x)^{-3},$$

正確的結果是

$$\int (1-x)^{-3}\,\mathrm{d}x = \tfrac{1}{2}(1-x)^{-2} + c, \qquad x \neq 1.$$

? **想一想** 用正確的原函數，在 -1 與 3 之間同樣得到 $\frac{1}{2}\cdot\frac{1}{4} - \frac{1}{2}\cdot\frac{1}{4} = 0$。那麼答案終究還是 0 嗎？

第 2 步：定積分。不是。只有當 $F' = f$ 在整個區間上成立，而且 f 在該區間上連續時，計算 $[F(x)]_{-1}^{3}$ 才有根據。這裏 $(1-x)^{-3}$ 在 $x = 1$ 處沒有定義，而這一點位於 $[-1,3]$ 之內，並且函數在該點附近無界，所以這是一個瑕積分（improper integral），必須定義為兩段各自的極限：

$$\int_{-1}^{1-\varepsilon} (1-x)^{-3}\,\mathrm{d}x = \tfrac{1}{2}\left(\varepsilon^{-2} - \tfrac{1}{4}\right) \to +\infty \quad (\varepsilon \to 0^{+}),$$
$$\int_{1+\delta}^{3} (1-x)^{-3}\,\mathrm{d}x = \tfrac{1}{2}\left(\tfrac{1}{4} - \delta^{-2}\right) \to -\infty \quad (\delta \to 0^{+}).$$

兩段都發散（diverge），所以這個積分不存在；「$\infty - \infty$」沒有值。從幾何上看這也合理：被積函數在 $x < 1$ 時為正，在 $x > 1$ 時為負，而兩邊各自圍出的面積都是無窮大。

結論。兩步都是錯的，而且照原樣來說結論並不成立：$\int_{-1}^{3}(1-x)^{-3}\,\mathrm{d}x$ 沒有定義。這個「0」只有在挖去一個對稱的缺口，即 $\varepsilon = \delta$ 時，才算有幾分道理：這時兩段恰好互相抵消，極限是 0（即柯西主值，Cauchy principal value）。但假如取 $\delta = 2\varepsilon$，兩段之和便是 $\frac{1}{2}(\varepsilon^{-2} - \frac{1}{4}\varepsilon^{-2}) \to \infty$，所以答案取決於我們怎樣趨近 $x = 1$，這正是這個積分沒有定義的原因。

驗算

在隨機選取的點上，g 的數值導數與因式分解後的公式吻合；$g'(1) = 0.4 \neq 0$。用二分法（bisection）求得 $\alpha \approx 1.3856$，$\beta \approx -2.2467$，並且 $g(\alpha) = 1.0825 = 3/(2\alpha)$，$g(\beta) = -0.6676 = 3/(2\beta)$；在 $[-100, 100]$ 上對 g 取樣，可確認這兩個值就是全局最大值和最小值。至於 (ii)，在 $[-1, 0.99]$ 上用中點求積法（midpoint quadrature）得到約 4999.87，與 $\frac{1}{2}(10^4 - \frac{1}{4})$ 吻合，而對應的另一段在 $[1.01, 3]$ 上得到 -4999.87。

常見錯誤

- 靠比較駐點處的函數值來判斷駐點的性質。
- 聲稱得到全局界時，忘記考慮 $x \to \pm\infty$ 時的表現。
- 跨過被積函數趨向無窮的點使用 $[F(x)]_a^b$；務必先檢查被積函數在 $[a, b]$ 上是否連續。

STEP 之外

第 (ii) 部分講的是瑕積分：若積分區間包含奇點（singularity），積分便要定義為極限，而且只有當兩個單側極限各自存在時，積分才收斂（converge）。對稱極限，即柯西主值，在複分析（complex analysis）和物理學中確實是很有用的概念，但它是另一回事。第 (i) 部分講的是局部極值與全局極值之間的差距，數學分析藉極值定理（extreme value theorem）把這一點說得精確。Abbott 的 *Understanding Analysis* 對這兩個課題都有仔細的論述。

詳細解答

STEP I 1991，第 7 題

純數學 • 可分離變量微分方程

> **STEP I 1991，第 7 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q7

關鍵思路

兩個變化率都是 $(1+V/V_0)^2$ 的常數倍，所以最自然的變數是 $u = 1+V/V_0$，它滿足 $\frac{\mathrm{d}u}{\mathrm{d}t} = \pm c\,u^2$。分離變數（separating the variables）後可見，倒數 $w = 1/u$ 只是隨時間線性變化：在增長年以速率 a 下降，在衰退年以速率 b 上升。於是整個問題歸結為追蹤一條由直線段組成的圖像，並回答兩個問題：w 會否到達 0（金額變成無限大），又會否到達 1（國庫空了）？

解答

建立模型

以年為單位量度 t，並寫 $V = kV_0$，於是 $u = 1+k$，且 $\mathrm{d}V/\mathrm{d}t = V_0\,\mathrm{d}u/\mathrm{d}t$。在增長年，

$$V_0\frac{\mathrm{d}u}{\mathrm{d}t} = aV_0u^2 \quad\Longrightarrow\quad \frac{1}{u^2}\frac{\mathrm{d}u}{\mathrm{d}t} = a \quad\Longrightarrow\quad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{u}\right) = -a,$$

同理，在衰退年（只要 $V>0$），$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{u}\right) = +b$。因此，設

$$w = \frac{1}{1+V/V_0},$$

則 w 在增長期間以速率 a 線性下降，在衰退期間以速率 b 線性上升。注意 $V \geq 0$ 對應 $0 < w \leq 1$：國庫空了當且僅當 $w=1$，而 $V\to\infty$ 當且僅當 $w\to 0^+$。

不等式的意義

開始時 $w = w_0 = \frac{1}{1+k_0}$，而在第一年內 $w = w_0 - at$。於是：

- $k_0 \geq 0$ 只是說國庫的存款不可能是負數。
- 金額在整個第一年內保持有限，當且僅當 w 直到 $t = 1$ 都保持正值，即 $w_0 - a > 0$，亦即
$$\frac{1}{1+k_0} > a \iff 1 + k_0 < \frac{1}{a} \iff k_0 < a^{-1} - 1.$$
假如這個條件不成立，V 便會在時間 $t = w_0/a \leq 1$，即第一年之內，變成無限大；對任何描述真實國庫的模型來說，這都是荒謬的。

兩個不等式合起來還迫使 $a^{-1} - 1 > 0$，即 $a < 1$；我們會在最後用到這一點。

撐過第二年

第一年結束後 $w_1 = \frac{1}{1+k_1}$（事實上 $w_1 = w_0 - a$）。在衰退年內，只要 $w < 1$，就有 $w = w_1 + bt$。國庫在年底仍未空，當且僅當 w 到 $t = 1$ 時仍未到達 1：

$$\frac{1}{1+k_1} + b < 1 \iff b < \frac{k_1}{1+k_1}.$$

特別地，我們需要 $b < 1$，這時上式可整理成

$$k_1 > \frac{b}{1-b}.$$

反之，若 $b \geq \frac{k_1}{1+k_1}$，存款會在時間 $t = (1 - w_1)/b \leq 1$ 耗盡，並在該年餘下的時間一直維持為零。

第三年的無限增長

> **想一想** 假如國庫在第二年空了，會怎樣？它還有可能在第三年爆破嗎？

若國庫在第二年年底是空的，第三年便由 $w = 1$ 開始，且 $w = 1 - at$；由於 $a < 1$，它在整年內都保持正值。所以空了的國庫不可能在第三年爆破，我們只需考慮存款撐得過第二年的情況。

在這情況下 $w_2 = w_1 + b = \frac{1}{1+k_0} - a + b$，而在第三年 $w = w_2 - at$。金額在第三年內變成無限大，當且僅當 w 在 $t = 1$ 或之前到達 0，即 $w_2 \leq a$：

$$\frac{1}{1+k_0} - a + b \leq a \iff b \leq 2a - \frac{1}{1+k_0}.$$

我們必須檢查這與「但不會更早」是否一致。

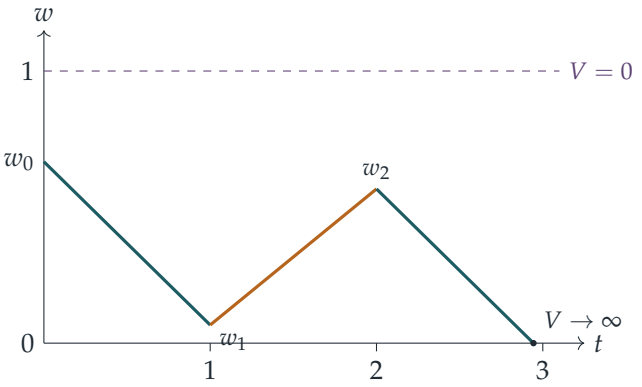
- $k_0 < a^{-1} - 1$ 保證了第一年不會爆破；第二年金額正在減少，也不可能爆破。
- 存款撐得過第二年：若 $w_2 \leq a$，則因 $a < 1$ 而有 $w_2 < 1$，所以存活條件 $w_2 < 1$ 自動成立。

因此，模型預測第三年出現無限增長而不會更早，當且僅當

$$b \leq 2a - \frac{1}{1+k_0}.$$

若不等式嚴格成立，V 會在第三年開始後的時間 $t = w_2/a < 1$ 變成無限大；在等號成立的情況下，$V \to \infty$ 恰好發生在第三年結束之時。（由於 $b > 0$，這只有在 $2a > \frac{1}{1+k_0}$，即 $k_0 > \frac{1}{2a} - 1$ 時才可能發生。）

下圖顯示 $a = 0.6$、$b = 0.5$、$k_0 = 0.5$ 時 w 對 t 的圖像，這是條件成立的一個例子。



驗算

在上面的例子中，$k_1 = 14$，存活條件 $14 > \frac{0.5}{0.5} = 1$ 成立；然後 $k_2 = \frac{13}{17} \approx 0.765$，而第三年的爆破發生在 $t = w_2/a \approx 0.944$，與 $2a - \frac{1}{1+k_0} - b = \frac{1}{30} > 0$ 一致。我們亦以數值方法（四階龍格－庫塔法，Runge–Kutta）對這些微分方程（differential equations）作積分，隨機選取了 300 組 a、b 和 k_0：k_1 和 k_2 的閉合式與數值結果相差約 10^{-11}，而存活準則和爆破準則在每一個情況下都正確預測了數值結果。

常見錯誤

- 忘記國庫可能在第二年空了。在這題中，由於 $a < 1$，這並無影響，但必須說明。
- 把爆破條件弄反：爆破對應的是 $1/(1+k)$ 到達 0，而不是 k 到達 0。
- 把 $\frac{1}{1+k_1} + b < 1$ 整理成關於 k_1 的條件時，忽略了 $b < 1$ 這個要求。

STEP 之外

方程 $\dot{u} = au^2$ 是有限時間爆破（finite-time blow-up）的標準例子：與指數增長不同，它的解可以在有限時間內變成無限大，所以微分方程的解不一定對所有時間都存在。在大學第一門常微分方程（ODE）課程中，解的局部存在性（local existence）與整體存在性（global existence）之分，正是在這裏出現。Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程教科書）在處理非線性一階方程時討論了這一點。

詳細解答

STEP I 1991，第 8 題

純數學 • 代換積分法

STEP I 1991，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s1-q8

關鍵思路

代換（substitution）$y = \pi - x$ 把 $[0, \pi]$ 映射到自身，並且令 $\sin x$ 保持不變，卻把礙事的因子 x 變成 $\pi - y$。把原積分與它反射後的版本相加，x 便完全消失。在第 (ii) 部分，被積函數（integrand）由一個函數及其導數構成，因此只要選對一個代換，便可一步到位。

解答

第 (i) 部分

設 $I = \int_0^{\pi} x\,\mathrm{f}(\sin x)\,\mathrm{d}x$。我們需要一個能把 $[0, \pi]$ 兩端對調、同時令 $\sin x$ 保持不變的代換，故取 $k = \pi$：令 $x = \pi - y$，於是 $\mathrm{d}x = -\mathrm{d}y$，$\sin x = \sin(\pi - y) = \sin y$，而 $x = 0, \pi$ 分別對應 $y = \pi, 0$。因此

$$\begin{aligned} I &= \int_{\pi}^{0} (\pi - y)\,\mathrm{f}(\sin y)\,(-\mathrm{d}y) = \int_0^{\pi} (\pi - y)\,\mathrm{f}(\sin y)\,\mathrm{d}y \\ &= \pi \int_0^{\pi} \mathrm{f}(\sin y)\,\mathrm{d}y - I, \end{aligned}$$

這是因為 y 只是啞變數（dummy variable），所以最後一個積分就是 I 本身。由此可得

$$I = \frac{\pi}{2} \int_0^{\pi} \mathrm{f}(\sin x)\,\mathrm{d}x.$$

? **想一想** 怎樣由 $\frac{\pi}{2}\int_0^{\pi}$ 得出 $\pi \int_0^{\pi/2}$？

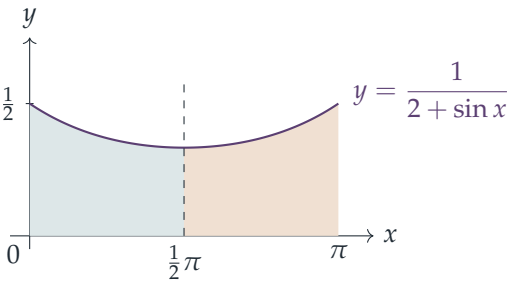
由於 $\sin(\pi - x) = \sin x$，$\mathrm{f}(\sin x)$ 在 $[0, \pi]$ 上的圖像關於 $x = \frac{1}{2}\pi$ 對稱。要嚴格說明這一點，可把積分在 $\frac{1}{2}\pi$ 處分成兩段，並對後半段使用同一個代換 $x = \pi - y$：

$$\int_{\pi/2}^{\pi} \mathrm{f}(\sin x)\,\mathrm{d}x = \int_{\pi/2}^{0} \mathrm{f}(\sin y)\,(-\mathrm{d}y) = \int_{0}^{\pi/2} \mathrm{f}(\sin y)\,\mathrm{d}y.$$

所以 $\int_0^{\pi} \mathrm{f}(\sin x)\,\mathrm{d}x = 2\int_0^{\pi/2} \mathrm{f}(\sin x)\,\mathrm{d}x$，因此

$$\int_0^{\pi} x\,\mathrm{f}(\sin x)\,\mathrm{d}x = \pi \int_0^{\frac{1}{2}\pi} \mathrm{f}(\sin x)\,\mathrm{d}x,$$

這正是所要證明的結果。



現取 $\mathrm{f}(u) = \dfrac{1}{2+u}$，它在 $u \in [-1, 1]$ 上連續（continuous）。於是

$$\int_0^{\pi} \frac{x}{2+\sin x}\,\mathrm{d}x = \pi \int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{2+\sin x}.$$

餘下的積分可用 $t = \tan\frac{1}{2}x$ 處理，此時

$$\sin x = \frac{2t}{1+t^2}, \qquad \mathrm{d}x = \frac{2\,\mathrm{d}t}{1+t^2},$$

而 $x = 0, \frac{1}{2}\pi$ 分別對應 $t = 0, 1$。於是

$$\begin{aligned}
\int_0^{\frac{1}{2}\pi} \frac{\mathrm{d}x}{2+\sin x} &= \int_0^1 \frac{1}{2+\frac{2t}{1+t^2}} \cdot \frac{2\,\mathrm{d}t}{1+t^2} = \int_0^1 \frac{2\,\mathrm{d}t}{2(1+t^2)+2t} = \int_0^1 \frac{\mathrm{d}t}{t^2+t+1} \\
&= \int_0^1 \frac{\mathrm{d}t}{\left(t+\frac{1}{2}\right)^2 + \frac{3}{4}} = \frac{2}{\sqrt{3}}\left[\tan^{-1}\frac{2t+1}{\sqrt{3}}\right]_0^1 \\
&= \frac{2}{\sqrt{3}}\left(\tan^{-1}\sqrt{3} - \tan^{-1}\frac{1}{\sqrt{3}}\right) = \frac{2}{\sqrt{3}}\left(\frac{\pi}{3} - \frac{\pi}{6}\right) = \frac{\pi}{3\sqrt{3}}.
\end{aligned}$$

因此

$$\int_0^{\pi} \frac{x}{2+\sin x}\,\mathrm{d}x = \frac{\pi^2}{3\sqrt{3}} = \frac{\sqrt{3}\,\pi^2}{9}.$$

第 (ii) 部分

？ 想一想 $(\sin^{-1} t)^2$ 的導數是甚麼？

根據鏈式法則（chain rule），

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sin^{-1} t\right)^2 = \frac{2\sin^{-1} t}{\sqrt{1-t^2}},$$

這恰好是與餘弦相乘的那個因子（的兩倍）。所以作代換 $u = (\sin^{-1} t)^2$，其中 $\mathrm{d}u = \dfrac{2\sin^{-1} t}{\sqrt{1-t^2}}\,\mathrm{d}t$。當 $t = 0$ 時，$u = 0$；當 $t = 1$ 時，$\sin^{-1} 1 = \frac{1}{2}\pi$，故 $u = \frac{1}{4}\pi^2$。因此

$$\int_0^1 \frac{(\sin^{-1} t)\cos\left[(\sin^{-1} t)^2\right]}{\sqrt{1-t^2}}\,\mathrm{d}t = \frac{1}{2}\int_0^{\pi^2/4} \cos u\,\mathrm{d}u = \frac{1}{2}\sin\frac{\pi^2}{4}.$$

被積函數在 $t = 1$ 處趨向無窮，所以嚴格來說這是一個瑕積分（improper integral）。代換把這個問題處理得乾淨利落：以 $\theta = \sin^{-1} t$ 表示，該積分就是 $\int_0^{\pi/2} \theta\cos\theta^2\,\mathrm{d}\theta$，其被積函數連續，所以積分收斂（converges），而上面求得的值是正確的。（分兩步來做，先作 $\theta = \sin^{-1} t$，再作 $u = \theta^2$，也是完全可行的另一種做法。）

答案 $\frac{1}{2}\sin\frac{1}{4}\pi^2$ 是精確值，無法再化簡：$\frac{1}{4}\pi^2$ 並不是一個「漂亮」的角，想必這就是考官提醒考生不要動用計算機的原因。

驗算

- 以辛普森法則（Simpson's rule）計算，得 $\int_0^{\pi} x/(2+\sin x)\,\mathrm{d}x \approx 1.899406$，而 $\pi^2/(3\sqrt{3}) \approx 1.899406$。
- 快速檢查一下是否合理：$\frac{1}{3} \le \frac{1}{2+\sin x} \le \frac{1}{2}$，所以該積分介乎 $\frac{1}{3}\cdot\frac{1}{2}\pi^2 \approx 1.64$ 與 $\frac{1}{2}\cdot\frac{1}{2}\pi^2 \approx 2.47$ 之間。
- 第 (i) 部分的恆等式可用 $\mathrm{f}(u) = 1$ 來檢驗：兩邊都等於 $\frac{1}{2}\pi^2$。
- 至於第 (ii) 部分，$\frac{1}{2}\sin(\frac{1}{4}\pi^2) \approx 0.312133$，與 $\int_0^{\pi/2} \theta\cos\theta^2\,\mathrm{d}\theta$ 的數值計算結果吻合。

常見錯誤

- 做到 $I = \frac{\pi}{2}\int_0^{\pi} \mathrm{f}(\sin x)\,\mathrm{d}x$ 便停下來，沒有說明為何可以把積分範圍減半。
- 作代換 $t = \tan\frac{1}{2}x$ 時忘記更改積分上下限（上限變成 1，而不是 ∞）。
- 在第 (ii) 部分漏掉因子 $\frac{1}{2}$：$\mathrm{d}u$ 帶有一個 2。

STEP 之外

第 (i) 部分的技巧，是利用區間的對稱性來計算 $\int x\,g(x)\,\mathrm{d}x$，全程無須求出原函數（antiderivative）。這是一個重要主題的小小例子：要知道一個積分的值，藉助對稱性往往比直接計算省力得多。第 (ii) 部分則觸及瑕積分的理論：變數代換可以把無界的被積函數變成連續函數。Abbott 的 *Understanding Analysis* 嚴格地建立了變數代換定理（change-of-variables theorem），而 Siklos 的 *Advanced Problems in Mathematics* 則收錄了更多這類積分。

詳細解答

STEP I 1991，第 9 題

純數學 • 不等式與調和級數

> **STEP I 1991，第 9 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q9

關鍵思路

在第 (i) 部分，每一條不等式（inequality）都把 x 夾在兩個 k 次方根之間，竅門是找出互相矛盾的一個下界（lower bound）和一個上界（upper bound）。要在不用計算機的情況下比較立方根與五次方根，可把兩邊同時取十五次方。在第 (ii) 部分，我們用（項數）×（最小的項）作為一個長和式的下界。然後把這種形狀、互不相交的區塊串連起來，使調和級數（harmonic series）要多大有多大。

解答

第 (i) 部分

由於 $x > 1$，x 的每一個冪都是正數，而把兩個正數之間的不等式兩邊同時取正數次方，不等號的方向不變。以下會一直用到這一點。

為了引出矛盾，假設 $n \geq 5$。那麼第三條和第五條不等式必然成立：

$$x^3 > 3 \qquad 及 \qquad x^5 < 6.$$

把第一式取五次方，第二式取三次方：

$$x^{15} > 3^5 = 243 \qquad 及 \qquad x^{15} < 6^3 = 216.$$

由此得 $243 < x^{15} < 216$，這是不可能的。所以 $n \geq 5$ 不會發生，因此 $n \leq 4$。

想一想 你怎樣知道要選第三條和第五條不等式？

每一條不等式都說明 $k^{1/k} < x < (k+1)^{1/k}$。下界 $k^{1/k}$ 依次為 1、1.41、1.44、1.41、1.38、...，在 $k = 3$ 時最大。上界 $(k+1)^{1/k}$ 則遞減，而 $6^{1/5} \approx 1.43$ 是第一個跌到 $3^{1/3} \approx$

1.44 以下的。粗略心算一下便能找到這一對，然後十五次方就能嚴格地證明它。（$n=4$ 這個值確實可以出現。例如 $x=1.45$ 給出 $x^2=2.1025$、$x^3\approx 3.05$、$x^4\approx 4.42$。）

第 (ii) 部分

設 $n\geq 2$。和式 $\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n^2}$ 共有 n^2-n 項，每一項都不小於最後一項 $\frac{1}{n^2}$。由於 $n\geq 2$，項數 $n^2-n\geq 2$，所以第一項 $\frac{1}{n+1}$ 嚴格大於 $\frac{1}{n^2}$。因此

$$\frac{1}{n+1}+\cdots+\frac{1}{n^2}>(n^2-n)\cdot\frac{1}{n^2}=1-\frac{1}{n}.$$

兩邊同時加上 $\frac{1}{n}$，

$$\frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{n^2}>1,$$

這正是所要證明的。

使調和和超過 10 每一個區塊 $\frac{1}{m}+\cdots+\frac{1}{m^2}$（其中 $m\geq 2$）的貢獻都大於 1。要同時使用多個區塊，它們必須互不重疊，所以我們讓每個新區塊緊接在上一個區塊結束之後開始。定義

$$a_1=2,\qquad a_{k+1}=a_k^2+1,$$

於是 $a_1=2$, $a_2=5$, $a_3=26$, $a_4=677,\ldots$。第 k 個區塊由 $\frac{1}{a_k}$ 到 $\frac{1}{a_k^2}$，而下一個區塊由 $\frac{1}{a_k^2+1}=\frac{1}{a_{k+1}}$ 開始。所以 $k=1,\ldots,9$ 的各個區塊互不相交而且首尾相接，合起來剛好涵蓋 $\frac{1}{2},\frac{1}{3},\ldots,\frac{1}{a_9^2}$。因此

$$\sum_{n=1}^{a_9^2}\frac{1}{n}=1+\sum_{k=1}^{9}\left(\frac{1}{a_k}+\cdots+\frac{1}{a_k^2}\right)>1+9=10.$$

想一想 a_9^2 是一個龐大無比的數。我們能否改為給出一個明確的 N？

可以，只要給 a_k 定出上界。我們斷言 $a_k<2^{2^k}$。當 $k=1$ 時，由於 $2<4$，斷言成立。若 $a_k<2^{2^k}$，則 $a_k\leq 2^{2^k}-1$，所以

$$a_{k+1}=a_k^2+1\leq\left(2^{2^k}-1\right)^2+1=2^{2^{k+1}}-2^{2^k+1}+2<2^{2^{k+1}}.$$

於是由歸納法（induction）可知 $a_9<2^{512}$ 及 $a_9^2<2^{1024}$。再加上更多正數項只會令和變得更大，所以

$$N=2^{1024}\quad\text{給出}\quad\sum_{n=1}^{N}\frac{1}{n}>10.$$

任何更大的 N 同樣可行。

另一做法：把各項分成區塊 $\frac{1}{2^j+1}+\cdots+\frac{1}{2^{j+1}}$，每個區塊有 2^j 項，而且每一項都不小於 $\frac{1}{2^{j+1}}$，由此可見當 $j \geq 1$ 時每個區塊都大於 $\frac{1}{2}$，當 $j = 0$ 時則等於 $\frac{1}{2}$。所以 $\sum_{n=1}^{2^{18}} \frac{1}{n} > 1 + \frac{18}{2} = 10$，而 $N = 2^{18} = 262\,144$ 是一個小得多的答案。

驗算

- 第 (i) 部分：從數值上看，當 $n = 4$ 時，各條不等式剩下 $x \in (3^{1/3}, 5^{1/4}) \approx (1.442, 1.495)$。加入第五條不等式後必須有 $x < 6^{1/5} \approx 1.431$，於是甚麼也不剩。
- 第 (ii) 部分：用準確分數計算，$\frac{1}{2}+\frac{1}{3}+\frac{1}{4} = \frac{13}{12} > 1$，而且對由 2 到 39 的每一個 n，區塊的和都大於 1。
- 調和和在 $N = 12\,367$ 時首次超過 10，所以任何正確的答案都必須不小於這個數。2^{18}（此時和約為 13.05）和 2^{1024} 都綽綽有餘。

常見錯誤

- 把不等式取次方時，沒有指出兩邊都是正數（這裏由 $x > 1$ 保證）。
- 使用區塊 $[2,4]$、$[4,16]$、$[16,256]$、...，這些區塊共用端點，以致有些項被重複計算了兩次。
- 在第 (ii) 部分寫了 $\geq$ 而不是 $>$，卻沒有檢查至少有一項嚴格大於 $\frac{1}{n^2}$。

STEP 之外

第 (ii) 部分是柯西凝聚（Cauchy condensation）這個想法的一個版本。它證明了調和級數發散（diverge），只是發散得慢得驚人：它的部分和（partial sum）的增長有如 $\ln N + \gamma$，其中 $\gamma \approx 0.5772$ 是歐拉－馬斯刻若尼常數（Euler–Mascheroni constant）。把和式與積分作比較，便可得出這個增長率，也令最小的 N 容易估算（$e^{10-\gamma} \approx 12\,367$）。Abbott 的 *Understanding Analysis*（一本實分析入門書）對凝聚法和級數的收斂（convergence）有仔細的論述。

詳細解答

STEP I 1991，第 10 題

力學 • 懸鏈的平衡

STEP I 1991，第 10 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s1-q10

本題在 *Beyond STEP* 第 11.14 節是拋物線，不是懸鏈線中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch11/bridge-parabola.html

關鍵思路

鏈上每一個接點都在三個力的作用下處於平衡（equilibrium）：相鄰兩節鏈的張力（tension），以及掛在該接點上的重量 w。沿水平方向分解力（resolving），可知每一節鏈的張力都有相同的水平分量。再沿豎直方向分解，可知每經過一個接點，豎直分量恰好減少 w；由於水平間距固定，各節鏈的斜率便構成一個等差數列（arithmetic progression）。斜率成等差數列，正是拋物線（parabola）在離散情況下的標記。

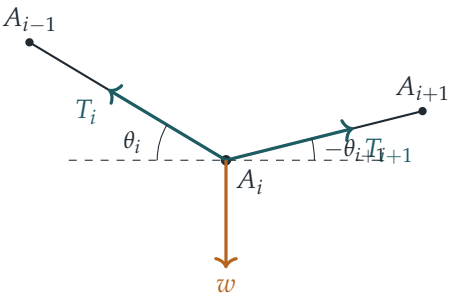
解答

平衡方程

把鏈視為輕、柔軟而且不可伸長（light, flexible and inextensible）的模型，因此每一節鏈 $A_{i-1}A_i$ 的張力都沿着該節鏈的方向。設 θ_i 為 $A_{i-1}A_i$ 與水平線所成的角，沿 x 增加的方向向下量度，所以

$$\tan\theta_i = \frac{y_{i-1} - y_i}{h}.$$

（在橋的右半邊，θ_i 是負數；以下的論證全都不受其正負號影響。）接點 A_i 受三個力作用：沿 A_iA_{i-1} 方向的 T_i、沿 A_iA_{i+1} 方向的 T_{i+1}，以及吊桿向下的拉力 w。



在 A_i 沿水平和豎直方向分解力，對 $1 \le i \le n+1$（其實對所有 $1 \le i \le 2n+1$ 都成立）：

$$T_i \cos\theta_i = T_{i+1}\cos\theta_{i+1},$$
$$T_i \sin\theta_i - T_{i+1}\sin\theta_{i+1} = w.$$

這就是題目要求的方程。

各個比

第一條方程說明 $T_i\cos\theta_i$ 對每一節鏈都相同；把這個共同的水平分量記作 H。它是正數，因為鏈只能拉、不能推。把第二條方程除以 H，並利用 $\tan\theta_i = (y_{i-1}-y_i)/h$，得

$$\frac{H}{h}(y_{i-1}-y_i) - \frac{H}{h}(y_i - y_{i+1}) = w.$$

因此，若以 $V_i = H(y_{i-1}-y_i)/h$ 表示第 i 節鏈張力的豎直分量，則

$$V_i - V_{i+1} = w \quad (1 \le i \le 2n+1), \qquad \text{所以} \qquad V_i = V_1 - (i-1)w \quad (1 \le i \le 2n+2).$$

想一想 現在只剩下一個未知數 V_1。橋的哪一項條件可以把它確定下來？

兩個支點的高度相同。把整條鏈上各節的下降量加起來，

$$0 = y_0 - y_{2n+2} = \sum_{i=1}^{2n+2}(y_{i-1}-y_i) = \frac{h}{H}\sum_{i=1}^{2n+2} V_i$$
$$= \frac{h}{H}\Big[(2n+2)V_1 - w\big(0+1+\cdots+(2n+1)\big)\Big].$$

方括號內的式子是 $(2n+2)V_1 - w(2n+1)(n+1)$，所以 $V_1 = \frac{1}{2}(2n+1)w$。（整體的豎直平衡也指向同一結果：兩個支點平均分擔總負載 $(2n+1)w$。）因此

$$V_i = \tfrac{1}{2}(2n+3-2i)\,w, \qquad \text{即} \qquad y_{i-1}-y_i = \frac{(2n+3-2i)\,wh}{2H}.$$

取 $i = n+1, n, \ldots, 1$，得到的下降量依次是 $wh/(2H)$ 的 $1, 3, \ldots, 2n+1$ 倍：

$$\frac{h}{y_n - y_{n+1}} = \frac{3h}{y_{n-1}-y_n} = \cdots = \frac{(2n+1)h}{y_0 - y_1} = \frac{2H}{w},$$

這就是所要求的一連串等式。（試卷上最後一個分子印作 $(2n+1)y$；它必定是 $(2n+1)h$，這既符合前面各項的規律，也只有這樣，該比才會像其他各個比一樣是無量綱的。）

拋物線

設 $k = wh/(2H)$。從 A_0 起把下降量加起來，

$$y_i = y_0 - k\sum_{j=1}^{i}(2n+3-2j) = y_0 - k\big[(2n+3)i - i(i+1)\big] = y_0 - k\,i\,(2n+2-i).$$

記 $m = n+1$，則 $i(2m-i) = m^2 - (i-m)^2$，所以

$$y_i = (y_0 - km^2) + k(i-m)^2 = y_{n+1} + k(i-n-1)^2,$$

這是因為代入 $i = m$ 可知 $y_0 - km^2 = y_{n+1}$。現在 $x_i = x_0 + ih$，所以 $(i-n-1)h = x_i - x_{n+1}$，於是每一點 A_i（$0 \le i \le 2n+2$）都滿足

$$y = y_{n+1} + \frac{w}{2Hh}\,(x - x_{n+1})^2.$$

這是同一條拋物線，其豎直的對稱軸通過最低點 A_{n+1}，所以 $A_0, A_1, \ldots, A_{2n+2}$ 各點全都在這條拋物線上。兩個端點也包括在內：$i = 0$ 和 $i = 2n+2$ 都給出 y_0，正如所料。

驗算

- 當 $n = 0$ 時，只有正中間的一根吊桿：公式給出 $V_1 = \frac{1}{2}w$，即鏈的每一半各承擔一半負載，這顯然是應有的結果。
- 這個結果與連續的情況吻合：一條水平張力為 H 的纜索，掛上每單位水平長度為 w/h 的均勻負載，會有 $y'' = (w/h)/H$，而我們的拋物線正有 $y'' = w/(Hh)$。
- 我們做了一次數值檢驗：隨機選取 200 座橋（$1 \le n \le 8$，H、w、h 隨機選取），以精確分數把各接點的方程當作線性方程組直接求解，並確認了各個比、公式 $y_i = y_0 - k\,i(2n+2-i)$、拋物線，以及每個接點上完整的力平衡。例如當 $n = 2$、$w = 2$、$H = 5$、$h = 1$ 時，下行一段的下降量依次為 1.0, 0.6, 0.2，而每個比都等於 $2H/w = 5$。

常見錯誤

- 不加說明便以對稱為由斷言 $y_n = y_{n+2}$。這是對的，但它是由兩個支點等高推出來的；上面的求和論證是把它證明出來，而不是假設它成立。
- 忘記張力的水平分量是常數，而嘗試逐一求出每個 T_i。本題從頭到尾都不需要 T_i 本身（它們等於 $\sqrt{H^2 + V_i^2}$）。

- 在鏈上升的那一半弄錯 θ_i 的正負號。全程使用 $\tan\theta_i = (y_{i-1} - y_i)/h$ 便可避免這個問題。

STEP 之外

只受自身重量而懸垂的鏈會形成懸鏈線（catenary）$y = c\cosh(x/c)$；但若鏈所承受的負載是沿水平方向均勻分佈的，例如橋面，則會形成拋物線。本題正是這個事實的離散版本。二階差分方程 $y_{i-1} - 2y_i + y_{i+1} =$ 常數 是 $y'' =$ 常數 的有限差分（finite difference）形式，而微分方程（differential equations）的數值方法也建基於同一想法。變分法（calculus of variations）處理的正是懸鏈線及與之相關的曲線。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）對繩和纜索的靜力學有嚴謹細緻的處理，很值得一讀。

詳細解答

STEP I 1991，第 11 題

力學 · 滑輪與衝量

STEP I 1991，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s1-q11

關鍵思路

每個循環分為兩個階段。重錘撞上樁時會立即停下，繩子隨之鬆弛，於是平衡錘向上飛起，再在重力作用下自由落回原處。接着繩子猛然拉緊，兩個質量分享動量，然後以第一部分求得的恆定加速度一起運動，直到重錘再次落在樁上。恆定加速度令每個階段在時間上都對稱。因此，一個循環中的每個速率，都是上一個循環中對應速率的固定倍數，等比數列（geometric progression）便隨即出現。

解答

恆定加速度

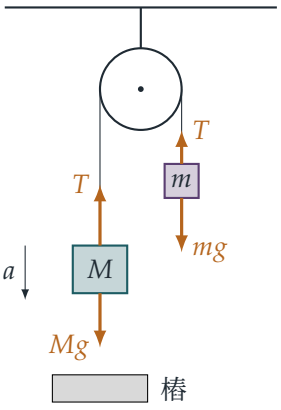
我們把重錘和平衡錘視為質點（particle），並假設繩子輕而不可伸長（light and inextensible），繞過一個光滑的輕滑輪（pulley）。繩子拉緊時，兩個質點以相同的速率 v 運動，重錘向下，平衡錘向上。設它們由靜止開始移動了距離 x 之後，速率為 v。重錘失去了勢能（potential energy）Mgx，平衡錘則增加了 mgx。滑輪光滑，繩子輕而不可伸長，所以沒有能量損失，於是

$$\tfrac{1}{2}(M+m)v^2 = (M-m)gx.$$

對時間求導，並利用 $\dot{x} = v$，

$$(M+m)v\dot{v} = (M-m)gv \quad \Longrightarrow \quad \dot{v} = g\,\frac{M-m}{M+m}$$

其中 $v \neq 0$。這是一個常數，記為 a。用牛頓第二定律（Newton's second law）也可得到同樣的結果：設 T 為張力（tension），則 $Mg - T = Ma$ 及 $T - mg = ma$，兩式相加得 $(M-m)g = (M+m)a$。



從一次碰撞到下一次碰撞

在碰撞前的一刻，重錘以速率 v 向下運動，平衡錘則以相同的速率 v 向上運動。我們分三個階段追蹤這個運動。

*撞擊。*碰撞是完全非彈性（completely inelastic）的，而椿固定不動，所以重錘停下。平衡錘仍以速率 v 向上運動，所以繩子鬆弛。這時張力幫不了平衡錘，而重錘只受重力作用，被壓在椿上，所以重錘留在原處。

*自由飛行。*繩子鬆弛時，平衡錘是一個沿鉛垂方向運動的拋體（projectile）。它上升後再回到繩子重新拉緊的位置，也就是撞擊時它所在的位置。在恆定加速度 g 之下，運動是對稱的，所以它回到該處時以速率 v 向下運動，所需時間為

$$\frac{2v}{g}.$$

*猛拉。*繩子猛然拉緊時，平衡錘以 v 向下運動，重錘則靜止。此後兩個質量必須以共同的速率 u 運動，平衡錘向下，重錘向上。根據題目容許的假設，平衡錘失去的動量（momentum）等於重錘得到的動量：

$$m(v-u) = Mu \quad \Longrightarrow \quad u = \frac{mv}{M+m}.$$

想一想 為甚麼這裏不能直接對整個系統使用動量守恆（conservation of momentum）？

兩個質量沿相反方向運動，而滑輪又施加了衝量（impulse），所以並不存在一個總動量守恆的方向。真正成立的是：由於繩子是輕的、滑輪是光滑的，衝擊張力（impulsive tension）對兩個質量的大小同為 J。因此平衡錘沿繩子失去動量 J，重錘沿繩子得到動量 J。這正是題目容許我們假設的那句話。

*再次一起運動。*現在繩子拉緊，重錘以速率 u 向上運動。根據第一部分，加速度的大小

恆為 a，方向是令較重的質量 M 向下。所以重錘減速至靜止，然後落回樁上。恆定加速度令這段運動同樣對稱：重錘以速率 u 到達樁，所需時間為

$$\frac{2u}{a}.$$

所以在下一次碰撞前的一刻，平衡錘的速率為

$$u = \frac{mv}{M+m},$$

正如所求。

各段時間構成等比數列

記 $k = \dfrac{m}{M+m}$，則 $0 < k < 1$。設 v_0 為第一次撞擊前一刻的速率。根據上一部分，第 $(n+1)$ 次撞擊前一刻的速率為 $v_n = k^n v_0$。把兩個階段相加，從第 $(n+1)$ 次撞擊到下一次撞擊的時間為

$$t_n = \frac{2v_n}{g} + \frac{2kv_n}{a} = \left(\frac{2}{g} + \frac{2k}{a}\right) k^n v_0.$$

這是一個常數乘以 k^n，所以相繼兩次碰撞之間的時間構成等比數列，公比（common ratio）為 $k = m/(M+m)$。

總時間

設 T_0 為從釋放到第一次撞擊的時間。重錘由靜止開始以加速度 a 運動，所以 $v_0 = aT_0$。由於 $0 < k < 1$，這個等比級數（geometric series）收斂（converges），總時間為

$$T_0 + \sum_{n=0}^{\infty} t_n = T_0 + \left(\frac{2}{g} + \frac{2k}{a}\right) \frac{v_0}{1-k}.$$

現在 $1 - k = \dfrac{M}{M+m}$，而 $\dfrac{k}{a} = \dfrac{m}{M+m} \cdot \dfrac{M+m}{g(M-m)} = \dfrac{m}{g(M-m)}$。因此

$$\frac{2}{g} + \frac{2k}{a} = \frac{2}{g}\left(1 + \frac{m}{M-m}\right) = \frac{2M}{g(M-m)},$$

$$\left(\frac{2}{g} + \frac{2k}{a}\right) \frac{v_0}{1-k} = \frac{2M}{g(M-m)} \cdot \frac{M+m}{M} v_0 = \frac{2v_0}{a} = 2T_0.$$

所以碰撞有無限多次，但總共只需有限的時間。重錘最終靜止下來所需的時間為

$$T_0 + 2T_0 = 3T_0,$$

即從開始到第一次撞擊所需時間的三倍。

驗算

- 取極限 $m \to 0$。這時 $k \to 0$ 而 $a \to g$，所以重錘在第一次撞擊後不再被提起。餘下的時間只有平衡錘的飛行時間 $2v_0/g = 2gT_0/g = 2T_0$，總時間為 $3T_0$，與一般公式所預測的一致。
- 我們以隨機的 M、m 和下落高度，用數值方法逐步計算這個運動（例如 $M = 2.403$，$m = 1.571$，故 $k = 0.3953$）。速率之比和時間間隔之比都是 0.3953，而總時間（包括等比級數的尾部）則為 $2.99998\,T_0$。

常見錯誤

- 忘記了自由飛行的階段。撞擊之後平衡錘繼續向上運動，循環中包括它作為拋體所花的時間 $2v/g$。
- 在繩子鬆弛的階段用了 a 而不是 g。繩子一旦鬆弛，平衡錘便只受重力作用。
- 在猛拉的一刻寫「總動量守恆」，卻把兩個速度當作同一方向。滑輪改變了繩子的方向，所以雙方分享的是*沿繩子方向*的動量。

STEP 之外

像這部打樁機或一個彈跳的球那樣，在有限時間內發生無限多個事件，稱為芝諾現象（Zeno behaviour）。這類現象出現在混合系統（hybrid systems）的理論中，當中平滑的運動會被突然的跳躍打斷。繩子的猛拉是一個簡單的衝擊約束（impulsive constraint）。大學一年級的拉格朗日力學（Lagrangian mechanics）會有系統地處理這類約束。想進一步了解衝量、碰撞和受約束運動，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）。

詳細解答

STEP I 1991，第 12 題

力學 • 涉及摩擦力的平衡

> **STEP I 1991，第 12 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q12

關鍵思路

木板 AB 是輕的，而且只在兩端受力，兩端都是光滑鉸鏈（smooth hinge）。所以它施於 CD 的力必定沿 AB 方向。這樣問題就簡化為單一剛體（rigid body）CD，受三個未知的力作用。對 C 取力矩（taking moments）可求出 AB 中的推力（thrust），再分解力（resolving）便得到 C 處的摩擦力（friction）和法向反作用力（normal reaction）。幾何關係（兩塊木板在 B 相接）把 x 與 α、β 聯繫起來，正因如此，l 和 x 才會在答案中消去。

解答

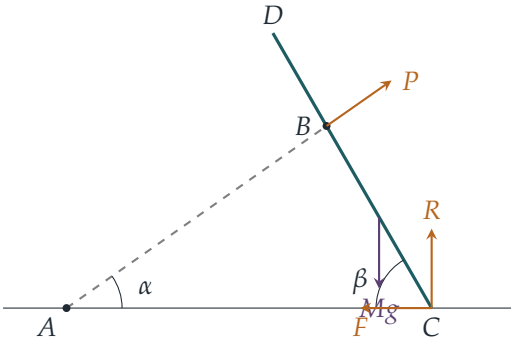
模型

我們把木板視為輕的剛性桿，把油漆工視為質量為 M 的質點（particle），並把 A 和 B 處的鉸鏈視為光滑。地面在 C 處是粗糙的，摩擦系數（coefficient of friction）為 μ。

想一想 為甚麼兩塊木板之間在 B 處的力必定沿 AB 方向？

木板 AB 沒有重量，作用在它上面的力只來自 A 和 B 處的光滑鉸鏈。對 A 取力矩，B 處的力對 A 的力矩為零，所以它沿 BA 方向。於是 A 處的力必定與它大小相等、方向相反。根據牛頓第三定律（Newton's third law），AB 在 B 處以一個沿 AB 方向的力 P 推 CD。我們取 P 的方向為由 A 指向 B（即推力），而下面的計算會確認 $P > 0$。

因此作用在 CD 上的力有：B 處的推力 P，與水平線成仰角 α，方向背離 A；油漆工的重量 Mg，作用點沿木板距 C 為 x；以及 C 處向上的法向反作用力 R 和指向 A 的摩擦力 F。摩擦力必須指向 A，因為推力把 C 推離 A。



幾何關係

點 B 是 AB 的頂端，所以它的高度是 $2l\sin\alpha$。它沿 CD 與 C 的距離又是 $2x$，所以它的高度也是 $2x\sin\beta$。因此

$$x\sin\beta = l\sin\alpha, \qquad \text{即} \qquad x = \frac{l\sin\alpha}{\sin\beta}.$$

水平距離 AC 是 AB 和 BC 的水平投影之和：

$$AC = 2l\cos\alpha + 2x\cos\beta.$$

對 CD 取力矩和分解力

對 C 取力矩。P 的作用線（line of action）就是直線 AB，它與 C 的垂直距離是 $AC\sin\alpha$。重量的作用線與 C 的水平距離是 $x\cos\beta$。兩個力矩的轉向相反，所以

$$P\cdot AC\sin\alpha = Mgx\cos\beta \quad\Longrightarrow\quad P = \frac{Mgx\cos\beta}{(2l\cos\alpha + 2x\cos\beta)\sin\alpha} > 0.$$

沿水平和鉛垂方向分解力：

$$F = P\cos\alpha, \qquad R = Mg - P\sin\alpha.$$

代入 P，

$$R = Mg\left(1 - \frac{x\cos\beta}{2l\cos\alpha + 2x\cos\beta}\right) = Mg\,\frac{2l\cos\alpha + x\cos\beta}{2l\cos\alpha + 2x\cos\beta},$$
$$F = Mg\,\frac{x\cos\beta\cos\alpha}{(2l\cos\alpha + 2x\cos\beta)\sin\alpha}.$$

特別地，$R > 0$，所以 C 端一直與地面保持接觸。兩式相除，

$$\frac{F}{R} = \frac{x\cos\alpha\cos\beta}{\sin\alpha\,(2l\cos\alpha + x\cos\beta)}.$$

現在運用幾何關係 $x = l\sin\alpha/\sin\beta$。於是 $x\cos\beta = l\sin\alpha\cot\beta$，而

$$\frac{F}{R} = \frac{l\sin\alpha\cot\beta\cos\alpha}{\sin\alpha\,(2l\cos\alpha + l\sin\alpha\cot\beta)} = \frac{\cos\alpha\cot\beta}{2\cos\alpha + \sin\alpha\cot\beta} = \frac{\cot\alpha\cot\beta}{2\cot\alpha + \cot\beta},$$

其中最後一步把分子和分母同除以 $\sin\alpha$。

木板 CD 只有在摩擦力足以提供這個力時才處於平衡（equilibrium），即 $F \le \mu R$。由於 $R > 0$，這就是

$$\mu \ge \frac{\cot\alpha\cot\beta}{2\cot\alpha + \cot\beta},$$

正是所求。（木板 AB 在 A 處以鉸鏈固定，所以無論它需要多大的力，都能保持不動。）

最後部分

若 B 與 D 重合，則 $2x = 2l$，所以 $x = l$。由幾何關係得 $\sin\beta = \sin\alpha$，而兩個角都是銳角，所以 $\beta = \alpha$。兩塊木板與地面構成一個等腰三角形，其中 $AC = 4l\cos\alpha$。摩擦條件變成

$$\mu \ge \frac{\cot^2\alpha}{3\cot\alpha} = \tfrac{1}{3}\cot\alpha, \qquad \text{即} \qquad \tan\alpha \ge \frac{1}{3\mu}.$$

所以當且僅當 $\alpha \ge \alpha_0$ 時才可能平衡，其中 $\tan\alpha_0 = 1/(3\mu)$。

想一想 油漆工在哪裏？當 α 變化時，距離朝哪個方向改變？

油漆工站在 CD 的中點，與 C 的距離為 l。以 A 為原點，C 位於 $(4l\cos\alpha, 0)$，而油漆工位於

$$(4l\cos\alpha - l\cos\alpha,\ l\sin\alpha) = (3l\cos\alpha,\ l\sin\alpha).$$

所以油漆工與 A 的距離的平方是

$$d^2 = 9l^2\cos^2\alpha + l^2\sin^2\alpha = l^2\left(1 + 8\cos^2\alpha\right).$$

在 $(0, \frac{\pi}{2})$ 上，這個量隨 α 增大而減小。所以最大距離出現在容許的最小角度 $\alpha = \alpha_0$。這時 $\tan\alpha_0 = 1/(3\mu)$，所以

$$\cos^2\alpha_0 = \frac{1}{1 + \tan^2\alpha_0} = \frac{9\mu^2}{1 + 9\mu^2},$$

而

$$d^2_{\max} = l^2\left(1 + \frac{72\mu^2}{1 + 9\mu^2}\right) = l^2\,\frac{1 + 81\mu^2}{1 + 9\mu^2}.$$

因此油漆工與 A 的最大距離是

$$l\sqrt{\frac{1 + 81\mu^2}{1 + 9\mu^2}},$$

正是所求。這個值是可以達到的，這時摩擦力剛好處於極限狀態（limiting friction）。

驗算

- 最終答案的極限情況：當 $\mu \to 0$ 時，距離趨向 l，對應 $\alpha \to \frac{\pi}{2}$，兩塊木板都豎直。當 $\mu \to \infty$ 時，距離趨向 $3l$，對應兩塊木板平放在地上，油漆工位於 $(3l, 0)$。兩者都合理。
- 若 $\beta \to \frac{\pi}{2}$，即 CD 豎直，則 $\cot\beta \to 0$，所需的 μ 趨向 0。豎直的木板不需要摩擦力，與此吻合。
- 數值驗算：取 $\alpha = 60°$ 和 $\beta = 70°$，直接解 CD 的三條平衡方程，得 $F/R = 0.13837$，與公式相符。取 $\mu = 0.5$，對 α 逐一掃描，得最大距離為 $2.55704\,l$，而公式給出 $l\sqrt{21.25/3.25} = 2.55704\,l$。

常見錯誤

- 假設 B 處的力是水平或鉛垂的。由於 AB 是輕的，而且兩端都以光滑鉸鏈連接，這個力必定沿 AB 方向。
- 在 R 中漏掉推力的鉛垂分量。這個分量會減小法向反作用力，令所需的摩擦力更難提供。
- 在最後部分，對所有 α 求最大值，而不是只對可能平衡的角度求最大值。

STEP 之外

兩端以光滑銷釘連接的輕桿稱為二力構件（two-force member）。辨認出這類構件，是工程靜力學中分析銷接框架（桁架，truss）的第一步，其中每一根這樣的構件只承受純拉力或純推力。最後部分是一個小型的約束優化（constrained optimisation）問題：目標函數是單調的，所以最優值落在可行集（feasible set）的邊界上。這個主題在線性規劃（linear programming）和優化理論中反覆出現。想進一步了解剛體的平衡，可參閱 Kleppner and Kolenkow, *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP I 1991，第 13 題

力學 • 動量守恆

> **STEP I 1991，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q13

關鍵思路

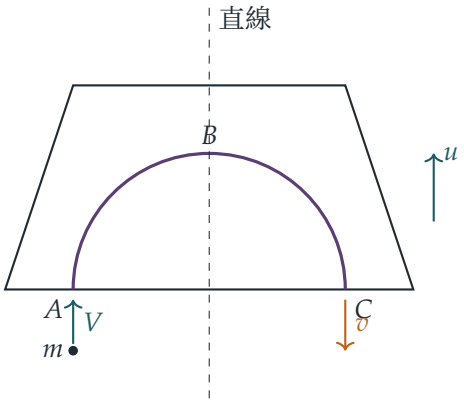
桌面光滑，凹槽也光滑，所以沒有任何外力沿直線方向推動系統，也沒有任何力作功。由此得到兩條守恆定律：沿直線方向的動量（momentum）守恆，以及動能（kinetic energy）守恆。凹槽在出口處的切線與直線平行，所以在出口處兩個速度都沿着直線，這時兩條定律的用法就與一維彈性碰撞（elastic collision）完全一樣。至於固定的薄片（lamina），平均力就是總衝量（impulse）除以所需時間。

關於題目用字，有一點要說明：題目說在 A 和 B 的切線與直線平行，但從圖中可見，與直線平行的其實是兩端 A 和 C 的切線（在 B 處凹槽與直線成直角）。我們採用 A 和 C，這顯然才是題目的原意。

解答

可移動的薄片

把質點（particle）看成點質量，把薄片看成只能沿直線平移的剛體（rigid body），並假設所有接觸面都是光滑的。取該直線為軸，以質點在 A 進入凹槽時的運動方向為正方向。



設質點離開 C 時沿直線的速度為 v（以 V 的方向為正），並設薄片此時沿直線的速度為 u。

想一想 為甚麼質點在 C 的速度必定沿着直線？

相對於薄片，質點沿凹槽運動，所以它在 C 的相對速度（relative velocity）沿着該處的切線，而這條切線與直線平行。薄片本身只沿直線運動。兩者相加，質點在 C 的實際速度同樣沿着直線。因此在出口處，每個物體只需要用一個數來描述。

在運動過程中，系統（質點加薄片）所受的外力有兩個：一是桌面的法向反作用力（normal reaction），方向垂直於桌面；二是令薄片保持在直線上的力，方向與直線垂直（直線是光滑的）。凹槽對質點的力與質點對凹槽的力都是內力，成對互相抵銷。所以沿直線方向的動量守恆：

$$mV = mv + Mu. \tag{1}$$

所有接觸面都是光滑的，而且桌面是水平的，所以沒有作功，動能守恆。（凹槽的法向反作用力確實分別對質點和薄片作功，但這兩個功互相抵銷，因為反作用力垂直於相對速度。）所以

$$\tfrac{1}{2}mV^2 = \tfrac{1}{2}mv^2 + \tfrac{1}{2}Mu^2. \tag{2}$$

把 (1) 和 (2) 改寫成

$$m(V - v) = Mu, \qquad m(V - v)(V + v) = Mu^2.$$

其中一個解是 $v = V$、$u = 0$：這是在 A 處、一切尚未發生時的情況。在 C 處，質點相對於薄片是向後運動的，所以 $v \neq V$ 而且 $u \neq 0$，於是可以把第二條方程除以第一條，得到

$$V + v = u.$$

（這表示質點相對於薄片的速率在兩端都是 V，與彈性碰撞的情況一樣。）把 $u = V + v$

代入 (1)：

$$mV = mv + MV + Mv \quad \Longrightarrow \quad v = -\frac{M-m}{M+m}V, \qquad u = V + v = \frac{2m}{M+m}V.$$

由於 $m < M$，v 是負數：質點從 C 出來時，是沿着來時的方向往回走。所以當質點離開凹槽時，

$$\text{質點的速率} = \frac{M-m}{M+m}V, \qquad \text{薄片的速率} = \frac{2m}{M+m}V,$$

其中薄片沿質點原來的方向運動，質點則沿相反方向運動。

質點確實能夠走完整條凹槽：假如它在某一點相對於薄片靜止，它的速度就會與薄片相同，這時動能會是 $\frac{1}{2}(mV)^2/(M+m) < \frac{1}{2}mV^2$，與能量守恆矛盾。所以它的相對速率永遠不會降至零。

固定不動的薄片

現在薄片不能移動。凹槽對質點的反作用力垂直於質點的速度，所以不作功，質點以恆定速率 V 走過半圓。半圓的長度是 πr，所以質點在凹槽內的時間是

$$T = \frac{\pi r}{V}.$$

質點的動量由向前的 mV 變為向後的 mV，所以薄片給予質點的總衝量是 $2mV$，方向沿直線向後。質點垂直於直線方向的動量在兩端都是零，所以衝量的垂直分量是零。

根據牛頓第三定律（Newton's third law），質點給予薄片的衝量是 $2mV$，方向沿直線向前。薄片保持靜止，所以栓釘（peg）必須提供相反的衝量，而薄片推向栓釘的總衝量就是 $2mV$，方向與質點原來的運動方向相同。（根據上一段，來自直線的任何橫向推力，平均值都是零。）除以 T，栓釘所受的平均力是

$$\frac{2mV}{\pi r/V} = \frac{2mV^2}{\pi r},$$

方向與質點進入凹槽時的方向相同。（題目所說的「average force per unit time」（每單位時間的平均力）正是這個意思：每單位時間的衝量。）

我們也可以直接看出這個結果。質點需要一個指向圓心、大小為 mV^2/r 的力，所以它以 mV^2/r 的力向外推薄片。當連接質點的半徑與 CA 所成的角為 θ 時（其中 θ 以均勻的速率由 0 增至 π），這個推力沿直線方向的分量是 $(mV^2/r)\sin\theta$，而 $\sin\theta$ 在 $[0,\pi]$ 上的平均值是 $2/\pi$。

驗算

- 若 $M \to \infty$，薄片幾乎不動：$u \to 0$，而質點的速率趨向 V，與凹槽固定時的情況一樣。

- 若 $m = M$，質點出來時靜止，而薄片取得它的全部速率 V，就像兩個質量相等的物體之間的彈性碰撞。
- 取 $M = 3$、$m = 1$、$V = 2$：$v = -1$ 而 $u = 1$。動量：$1 \cdot (-1) + 3 \cdot 1 = 2 = mV$。能量：$\frac{1}{2} + \frac{3}{2} = 2 = \frac{1}{2} \cdot 1 \cdot 2^2$。對運動方程作數值模擬（分別用半圓形凹槽和半橢圓形凹槽），所得數值準確至小數點後六位都相同；而對固定薄片的 $(mV^2/r)\sin\theta$ 作數值積分，得到平均力為 $2mV^2/(\pi r)$，例如當 $m = 0.7$、$r = 1.3$、$V = 2.1$ 時是 1.5117。

常見錯誤

- 保留 $v = V$、$u = 0$ 這個根：它滿足兩條方程，但描述的是進入凹槽的一刻，而不是離開的一刻。
- 用了最大的力 mV^2/r，而不是沿直線方向的平均力；橫向分量會互相抵銷，而沿直線方向的分量只在 B 處才達到 mV^2/r。

STEP 之外

第一部分其實是一次隱藏的碰撞：不論凹槽是甚麼形狀，重要的只有守恆量和出口方向，這正是不必知道作用力、單憑守恆定律處理散射（scattering）問題背後的想法。在拉格朗日力學（Lagrangian mechanics）中，沿直線的坐標並不出現在拉格朗日量（Lagrangian）之中，所以它的共軛動量（conjugate momentum）自動守恆；這就是對稱性與守恆定律之間的聯繫（諾特定理，Noether's theorem）的起點。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）正是以這種精神處理動量、衝量和碰撞。

詳細解答

STEP I 1991，第 14 題

概率與統計 • 三角形的計數

> **STEP I 1991，第 14 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q14

關鍵思路

三個長度 $a > b > c$ 能構成非退化三角形（non-degenerate triangle），當且僅當最長的一條嚴格短於其餘兩條之和，即 $b + c > a$；另外兩個三角不等式（triangle inequality）自動成立。因此固定 a 之後，我們對每個 b 數出夾在 $a - b$ 與 b 之間的 c 值有多少個。這些數目每次遞增 2，奇數（或偶數）之和正是由此而來。之後對所有可能的 a 求和，只需用到 $\sum k^2$ 的標準公式。

解答

三角形何時存在？

設 $a > b > c > 0$。非退化三角形須滿足的三個三角不等式是

$$b + c > a, \qquad a + c > b, \qquad a + b > c.$$

後兩個自動成立，因為 a 本身已經大於 b，也大於 c。所以三根棒能構成非退化三角形，當且僅當 $b + c > a$。（若 $b + c = a$，這個「三角形」是扁平的，面積為零，不計算在內。）

最長的棒長度為偶數

固定 a。我們需要整數 b, c 滿足

$$a - b < c < b, \qquad b < a.$$

對於給定的 b，滿足 $a - b + 1 \le c \le b - 1$ 的整數 c 共有

$$(b - 1) - (a - b + 1) + 1 = 2b - a - 1,$$

個，前提是這個數為正；否則一個也沒有。（留意這裏 $c \geq a-b+1 \geq 2$，所以 c 自動是一個真正的棒長。）

現在設 a 為偶數。那麼 $2b-a-1$ 是奇數，而它至少為 1 的充要條件是 $b \geq \frac{a}{2}+1$。因此 b 由 $\frac{a}{2}+1$ 取到 $a-1$，得出

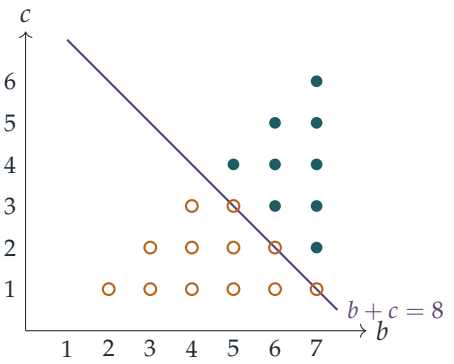
$$\begin{aligned} b = \tfrac{a}{2}+1: &\quad 2b-a-1 = 1, \\ b = \tfrac{a}{2}+2: &\quad 2b-a-1 = 3, \\ &\vdots \\ b = a-1: &\quad 2b-a-1 = a-3. \end{aligned}$$

b 每增加 1，數目便增加 2，所以以這根棒為最長邊的三角形數目是

$$1+3+5+\cdots+(a-3),$$

正如題目所求。（當 $a=2$ 時這個和是空的，這是對的：沒有三角形的最長邊是 2。）首 k 個奇數之和是 k^2，而這裏共有 $\frac{a-2}{2}$ 項，所以這個數目是 $\left(\frac{a-2}{2}\right)^2$。

下圖顯示 $a=8$ 的情況：每一點代表一對 (b,c)，其中 $8>b>c$，而實心點是位於直線 $b+c=8$ 上方的那些點。各列分別有 1、3 和 5 個實心點。



最長的棒長度為奇數

現在設 a 為奇數。那麼 $2b-a-1$ 是偶數，而它為正的充要條件是 $b \geq \frac{a+3}{2}$。當 b 由 $\frac{a+3}{2}$ 取到 $a-1$ 時，數目依次為 $2,4,\ldots,a-3$，所以三角形的數目是

$$2+4+6+\cdots+(a-3) = 2\left(1+2+\cdots+\tfrac{a-3}{2}\right) = \frac{a-3}{2}\cdot\frac{a-1}{2} = \frac{(a-1)(a-3)}{4}.$$

（同樣，$a=3$ 時正確地得出 0。）

三角形的總數

最長的棒可以是 $3,4,\ldots,2N+1$ 中任何一個（更小的值反正只會得出零）。

偶數 $a=2m$，$m=1,\dots,N$：數目是 $(m-1)^2$，所以這部分合共貢獻

$$\sum_{m=1}^{N}(m-1)^2=\sum_{k=1}^{N-1}k^2=\frac{(N-1)N(2N-1)}{6}.$$

奇數 $a=2m+1$，$m=1,\dots,N$：數目是 $\frac{(2m)(2m-2)}{4}=m(m-1)$，所以這部分合共貢獻

$$\begin{aligned}\sum_{m=1}^{N}(m^2-m)&=\frac{N(N+1)(2N+1)}{6}-\frac{N(N+1)}{2}\\&=\frac{N(N+1)(2N-2)}{6}=\frac{(N-1)N(N+1)}{3}.\end{aligned}$$

兩者相加，能構成非退化三角形的選法數目是

$$T_N=\frac{N(N-1)}{6}\big[(2N-1)+2(N+1)\big]=\frac{N(N-1)(4N+1)}{6}.$$

? **想一想** 為甚麼要按 a 的奇偶性（parity）把和式分開，而不是對單一公式求和？

對於給定的 a，當 a 為偶數時數目是 $\frac{(a-2)^2}{4}$，但當 a 為奇數時卻是 $\frac{(a-1)(a-3)}{4}=\frac{(a-2)^2-1}{4}$。這是兩個不同的 a 的多項式（polynomial），所以寫成 $a=2m$ 或 $a=2m+1$ 之後，兩者各自變成 m 的簡單多項式，可以用標準公式求和。

驗證 $N=3$ 的情況

公式給出 $T_3=\frac{3\cdot2\cdot13}{6}=13$。以長度為 $1,\dots,7$ 的棒直接列舉，把三角形寫成 (a,b,c)：

- $a=4$（一個）：$(4,3,2)$；
- $a=5$（兩個）：$(5,4,2)$, $(5,4,3)$；
- $a=6$（四個）：$(6,4,3)$, $(6,5,2)$, $(6,5,3)$, $(6,5,4)$；
- $a=7$（六個）：$(7,5,3)$, $(7,5,4)$, $(7,6,2)$, $(7,6,3)$, $(7,6,4)$, $(7,6,5)$。

合共 $1+2+4+6=13$ 個三角形，與公式吻合（亦與偶數 a 的 $1=1^2, 4=2^2$ 及奇數 a 的 $2=2\cdot1, 6=3\cdot2$ 吻合）。

概率

隨機抽出三根棒而不放回（without replacement），$\binom{2N+1}{3}$ 種無序選法中每一種的可能性均等。由於

$$\binom{2N+1}{3}=\frac{(2N+1)(2N)(2N-1)}{6},$$

構成非退化三角形的概率（probability）是

$$\frac{T_N}{\binom{2N+1}{3}} = \frac{N(N-1)(4N+1)}{(2N+1)\,2N\,(2N-1)} = \frac{(N-1)(4N+1)}{2(4N^2-1)},$$

其中約去了因子 N，並用到 $(2N+1)(2N-1) = 4N^2 - 1$。這就是所求的結果。

驗算

- 當 $N = 3$ 時，概率是 $\frac{2\cdot 13}{2\cdot 35} = \frac{13}{35}$，與 $\binom{7}{3} = 35$ 種選法中有 13 個三角形相符。
- 當 $N \to \infty$ 時，概率趨向 $\frac{4N^2}{8N^2} = \frac{1}{2}$。這與連續的版本一致：在 $[0,1]$ 上取三個獨立的均勻分佈長度，最長者小於其餘兩者之和的機會是 $\frac{1}{2}$。
- 對 N 由 2 至 29 的每個值，以窮舉法（brute force）數遍所有三元組，結果與 T_N 及概率公式完全吻合。

常見錯誤

- 用了 $b + c \geq a$：等號成立時得到的是扁平三角形，題目已把它排除，於是每個數目都會偏大。
- 把三個三角不等式逐一分開檢查，或者忘記了一旦 a 是最長的，就只有 $b + c > a$ 才起作用。
- 把 b 的範圍算錯一位；應以一個小例子檢驗，例如 $a = 6$（三角形為 $(6,4,3)$, $(6,5,2)$, $(6,5,3)$, $(6,5,4)$，即 $1 + 3$）。

STEP 之外

像圖中那樣數算區域內的格點（lattice point），是計算面積的離散版本；而當 N 增大時，這個數目除以 $\binom{2N+1}{3}$ 會趨近單位立方體內的一個體積比。這種由離散均勻分佈（uniform distribution）過渡到連續均勻分佈的做法，以及折棒問題（broken-stick problem）這類「幾何概率」（geometric probability）題目，都是大學一年級的經典課題。想進一步閱讀，Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）是個好去處。

詳細解答

STEP I 1991，第 15 題

概率與統計 • 正態近似

> **STEP I 1991，第 15 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s1-q15

關鍵思路

每個 X_j 的平均值（mean）為 0、方差（variance）為 1，而且各次擲幣互相獨立（independent），所以 S_n 的平均值為 0、方差為 n；當 n 很大時，中心極限定理（central limit theorem）容許我們把 S_n 視為近似正態分佈（normal distribution）。最後一部分看似相仿，卻藏着一個陷阱：每個 Y_k 都包含同一次首擲 X_1，所以各個 Y_k 並不獨立，而這一次共用的擲幣最終支配了整個總分。

解答

每個 X_j 取值 ± 1，概率各為 $\frac{1}{2}$，所以

$$\mathrm{E}(X_j) = 0, \qquad \mathrm{E}(X_j^2) = 1, \qquad \mathrm{Var}(X_j) = 1.$$

各次擲幣互相獨立，所以 $\mathrm{E}(S_n) = 0$ 及 $\mathrm{Var}(S_n) = n$。根據中心極限定理，當 n 很大時，

$$S_n \approx \mathrm{N}(0,\, n).$$

（等價地，若 H 為擲得正面的次數，則 $H \sim \mathrm{B}(n, \frac{1}{2})$，且 $S_n = H - (n - H) = 2H - n$。）

第 (i) 部分

當 $n = 10\,000$ 時，S_n 的標準差（standard deviation）為 $\sqrt{10\,000} = 100$，所以

$$\mathrm{P}(S_n > 100) \approx \mathrm{P}\left(Z > \frac{100}{100}\right) = 1 - \Phi(1) \approx 0.16.$$

如果想再謹慎一點，可對 H 使用連續性校正（continuity correction）：$S_n > 100$ 即 $2H - 10\,000 > 100$，亦即 $H \geq 5051$。由於 H 的平均值為 5000、標準差為 50，

$$\mathrm{P}(H \geq 5051) \approx \mathrm{P}\left(Z > \frac{5050.5 - 5000}{50}\right) = 1 - \Phi(1.01) \approx 0.156.$$

無論用哪一種方法，答案都約為 0.16。

第 (ii) 部分

標準化（standardising）後，

$$\mathrm{P}(S_n > 0.01n) \approx \mathrm{P}\left(Z > \frac{0.01n}{\sqrt{n}}\right) = 1 - \Phi\left(0.01\sqrt{n}\right).$$

這個值隨 n 增大而減小，而且它小於 0.01 當且僅當 $\Phi(0.01\sqrt{n}) > 0.99$。查表得 $\Phi(2.326) = 0.99$，所以我們需要

$$0.01\sqrt{n} > 2.326 \quad \Longleftrightarrow \quad \sqrt{n} > 232.6 \quad \Longleftrightarrow \quad n > 54\,100 \text{(約數)}.$$

所以符合條件的最小 n 大約是 54 000。（若改用 2.33，則得出約 54 300；在這個近似程度下，任何接近 5.4×10^4 的答案都算正確。）

新的計分方法

對於 $2 \leq k \leq n$，首擲與第 k 擲共有四種情況：

首擲	第 k 擲	得分 Y_k	$X_1 + X_k$
H	H	2	$1 + 1 = 2$
T	T	-2	$-1 - 1 = -2$
H	T	0	$1 - 1 = 0$
T	H	0	$-1 + 1 = 0$

最後兩欄在每種情況下都相同，所以 $Y_k = X_1 + X_k$，正如題目所求。

平均值與方差。 由線性性質（linearity），$\mathrm{E}(Y_k) = \mathrm{E}(X_1) + \mathrm{E}(X_k) = 0$。由於 $k \geq 2$，X_1 與 X_k 來自不同的兩次擲幣，因此互相獨立，於是

$$\mathrm{Var}(Y_k) = \mathrm{Var}(X_1) + \mathrm{Var}(X_k) = 2.$$

（直接計算：Y_k 取值 2、-2、0 的概率分別為 $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$，所以 $\mathrm{E}(Y_k^2) = 2$。）

極限。 設 $m = n - 1$ 及 $T = Y_2 + Y_3 + \cdots + Y_n$。那麼

$$T = \sum_{k=2}^{n} (X_1 + X_k) = mX_1 + S', \qquad \text{其中} S' = X_2 + X_3 + \cdots + X_n.$$

想一想 每個 Y_k 的平均值為 0、方差為 2，就像一個經過縮放的 X_j。為甚麼我們不能直接照搬第 (ii) 部分的論證？

因為第 (ii) 部分的論證需要各項互相獨立，而各個 Y_k 並不獨立：它們全都共用 X_1。事實上，對於 $j \neq k$，$\operatorname{Cov}(Y_j, Y_k) = \operatorname{Var}(X_1) = 1$，而 $\operatorname{Var}(T) = m^2 + m$，並不是 $2m$。所以我們把 X_1 的兩個取值分開考慮。留意 S' 是 m 個取值為 ± 1 的項之和，所以 $-m \leq S' \leq m$。

- 若首擲是反面，則 $X_1 = -1$ 且 $T = -m + S' \leq 0$。所以 $T > 0.01m$ 不可能發生。
- 若首擲是正面，則 $X_1 = 1$ 且 $T = m + S'$，所以 $T > 0.01m$ 當且僅當 $S' > -0.99m$。沿用第 (ii) 部分的正態近似，取 $S' \approx \mathrm{N}(0, m)$，

$$\mathrm{P}(S' \leq -0.99m) \approx \Phi\left(-0.99\sqrt{m}\right) \to 0 \quad \text{當} m \to \infty.$$

（其實無須使用近似：根據切比雪夫不等式（Chebyshev's inequality），這個概率至多為 $m/(0.99m)^2 \to 0$。）

由於 S' 與 X_1 獨立，

$$\mathrm{P}(T > 0.01m) = \tfrac{1}{2} \cdot 0 + \tfrac{1}{2}\,\mathrm{P}(S' > -0.99m) \;\to\; \tfrac{1}{2} \quad \text{當} n \to \infty.$$

所以該命題是**錯誤**的：這個概率趨向 $\frac{1}{2}$，而不是 0。用文字來說，當 n 很大時，總分為正基本上就等同於首擲是正面，否則總分便不是正數。

驗算

- 第 (i) 部分：對於 $H \sim \mathrm{B}(10\,000, \frac{1}{2})$，精確的二項概率 $\mathrm{P}(H \geq 5051)$ 為 0.15625，與經連續性校正的數值 0.156 吻合。
- 第 (ii) 部分：$\mathrm{P}(S_n > 0.01n)$ 的精確二項數值在 $n = 50\,000$ 時為 0.0125，在 $n = 54\,000$ 附近約為 0.0100，在 $n = 55\,000$ 時為 0.0094，所以臨界值確實接近 54 000。
- 最後部分：對 $n \leq 14$ 列舉全部 2^n 個結果，所得與 $\frac{1}{2}\,\mathrm{P}(S' > -0.99m)$ 完全一致；在 $n = 14$ 時概率已達 0.49994，而以 $n = 100$ 進行的一次模擬則得出 0.497。

常見錯誤

- 把 $Y_2, \ldots, Y_n$ 當作互相獨立，從而得出 $\operatorname{Var}(T) = 2(n-1)$，令概率看似趨向 0。
- 計算得分的方差時，用了 $\operatorname{Var}(S_n) = n/4$（即正面次數的方差）而不是 n。

STEP 之外

第 (ii) 部分是弱大數定律（weak law of large numbers）的一個定量形式：只要 n 夠大，S_n/n 就極不可能遠離 0。最後一部分則顯示，當各項之間有很強的相依關係時，這條定律會如何失效：這裏每一項都共用首擲，所以平均值 $T/m = X_1 + S'/m$ 趨向 X_1 這個隨機量，而不是趨向平均值 0。相依序列的大數定律（遍歷定理（ergodic theorems）、可交換序列（exchangeable sequences））是大學本科概率論中內容豐富的課題；Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門教材）涵蓋了弱大數定律和切比雪夫不等式。

詳細解答

STEP I 1991，第 16 題

概率與統計 • 條件概率

STEP I 1991，第 16 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s1-q16

關鍵思路

把一切按觀察時刻 $n = 0, 1, 2, \ldots$ 來整理，即每隻爬行動物第一次察看馬路之後的分鐘數。蟾蜍的等候時間服從幾何分佈（geometric distribution），所以牠必定會過路。青蛙則不同：牠的規則迫使牠最遲在 $n = 3$ 時作出決定，所以牠的故事只有四個階段，我們可以逐一列出牠在每個階段安全過路的概率。有了這張清單，與蟾蜍的比較以及條件概率（conditional probability）都只不過是幾個求和而已。

解答

全題假設各次觀察時馬路的狀況互相獨立（independent），每次安全的概率為 0.1，並且蟾蜍與青蛙的行動互相獨立。這是合理的，因為青蛙是「稍後」才過路，時間並不相同。由於過路所需的時間可以忽略，在第 n 次觀察時過路的爬行動物會在時間 n 分鐘到達。

蟾蜍

蟾蜍只會在安全時過路，所以牠永遠不會被輾過。經過第 $0, 1, \ldots, k$ 次觀察後牠仍在等候的概率為 0.9^{k+1}，而這個概率趨向 0。所以牠最終會安全過路，概率為

$$\sum_{k=0}^{\infty} 0.9^k(0.1) = \frac{0.1}{1 - 0.9} = 1.$$

以 T 表示牠過路的時間。那麼 $P(T = k) = 0.9^k(0.1)$，另外為方便稍後使用，$P(T > k) = 0.9^{k+1}$。

青蛙

在第 n 次觀察時，若牠仍在等候：

- 有 0.1 的概率馬路安全，牠安全過路；
- 有 $0.9 \cdot \frac{n}{3}$ 的概率馬路不安全但牠照樣過路，安全到達的概率為 0.2，被輾過的概率為 0.8；
- 有 $0.9\left(1-\frac{n}{3}\right)$ 的概率牠等待下一次觀察。

在 $n=0$ 時牠絕不冒險，而在 $n=3$ 時，不安全也照樣過路的機會是 1，所以若牠之前仍未出發，在 $n=3$ 時必定出發。特別地，$n>3$ 的情況永遠不會出現。

牠在第 n 次觀察時仍在等候的概率為

$$w_0=1,\quad w_1=0.9,\quad w_2=0.9\cdot 0.9\cdot\tfrac{2}{3}=0.54,\quad w_3=0.54\cdot 0.9\cdot\tfrac{1}{3}=0.162.$$

在第 n 次觀察時，牠安全到達的概率為 $w_n\left(0.1+0.9\cdot\frac{n}{3}\cdot 0.2\right)=w_n(0.1+0.06n)$。以 s_n 表示這個概率，

$$s_0=1\times 0.1=0.1,\qquad s_1=0.9\times 0.16=0.144,$$
$$s_2=0.54\times 0.22=0.1188,\qquad s_3=0.162\times 0.28=0.04536.$$

因此

$$P(\text{青蛙安全過路})=0.1+0.144+0.1188+0.04536=\boxed{0.40816=\tfrac{2551}{6250}}.$$

牠被輾過的概率就是餘下的 $1-0.40816=0.59184$。

兩者皆安全，而青蛙較快

設 F 為青蛙安全到達的時間。我們需要兩者都安全到達，並且 $F<T$。蟾蜍必定安全，而由獨立性，

$$P(\text{青蛙安全過路於}n\text{ 且}T>n)=s_n\,P(T>n)=s_n(0.9)^{n+1}.$$

想一想 為甚麼是 $T>n$ 而不是 $T\geq n$？

若兩者在同一個觀察時刻 n 過路，牠們所用的時間相同，青蛙並非嚴格地較快。所以蟾蜍在時間 n 的那次觀察之後必須仍在等候，其概率為 0.9^{n+1}。對青蛙四個可能的時間求和，

$$\begin{aligned}P&=0.1(0.9)+0.144(0.81)+0.1188(0.729)+0.04536(0.6561)\\&=0.09+0.11664+0.0866052+0.029760696\\&=0.323005896\approx\boxed{0.323}.\end{aligned}$$

寫成準確的分數，這就是 $\frac{40375737}{125000000}$。

最後一部分

設 A 為青蛙在第一次觀察後 2 分鐘仍未到達對面這一事件，R 為牠被輾過這一事件。由於過路是瞬間完成的，在 $n = 2$ 那次觀察時安全過路的青蛙，在 2 分鐘時已經到達。所以 A 是「在 $n = 0$、1 或 2 時安全過路」的補事件（complement）：

$$P(A) = 1 - (s_0 + s_1 + s_2) = 1 - 0.3628 = 0.6372.$$

被輾過的青蛙在任何時間都永遠不會到達，所以 $R \subseteq A$，並且 $P(R \cap A) = P(R) = 0.59184$。因此

$$P(R \mid A) = \frac{0.59184}{0.6372} = \frac{59184}{63720} = \boxed{\frac{274}{295} \approx 0.929}.$$

這包括了已經被輾過的青蛙（在 $n = 1$ 或 $n = 2$ 時），以及仍在等候、其後在 $n = 3$ 時被輾過的青蛙，正如「在某個階段」所要求的那樣。

（有些讀者把這個條件理解為：經過 $n = 0$ 和 $n = 1$ 兩次觀察後牠仍未到達，也就是在 $n = 2$ 那次察看之前。這種理解得出 $0.59184/(1 - 0.244) = \frac{137}{175} \approx 0.783$。我們較傾向採用上面的理解，因為恰好在 2 分鐘時到達，也算是在牠開始「之後 2 分鐘」到達。）

驗算

- 青蛙的各種結果必須涵蓋所有情況：在 $n = 1, 2, 3$ 時被輾過的概率分別為 $0.9(0.24)$、$0.54(0.48)$ 和 $0.162(0.72)$，即 $0.216 + 0.2592 + 0.11664 = 0.59184$，恰好等於 $1 - 0.40816$。
- 青蛙安全過路的概率遠低於蟾蜍那必然的 1，對沒有耐性的一方來說理應如此。
- 以分數作準確枚舉，再以蒙地卡羅模擬（Monte Carlo simulation）模擬 10^6 對爬行動物，得出 0.4087、0.3234 和 0.9286，與 0.40816、0.32301 和 0.92881 吻合。

常見錯誤

- 讓青蛙在 $n = 0$ 時冒險：按規則，那時 $n/3 = 0$。
- 忘記青蛙的過程在 $n = 3$ 時終止，於是像處理蟾蜍那樣為牠求一個幾何級數（geometric series）的和。
- 在最後一部分，分母用了 $P(\text{未能安全過路於} n \leq 2)$，分子卻只放進截至 $n = 2$ 為止被輾過的情況。題目問的是在任何階段被輾過，包括 $n = 3$。

STEP 之外

青蛙的旅程是一條小型的馬可夫鏈（*Markov chain*），帶有吸收態（absorbing states）（「安全」和「被輾過」），而蟾蜍的等候時間服從幾何分佈，即無記憶（memoryless）的指數分佈（exponential distribution）的離散版本。馬可夫鏈的吸收概率（absorption probabilities）和首達時間（hitting times）是大學第一門概率課程的核心課題。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率入門教科書）對兩者都有詳細闡述。

詳細解答

STEP II 1991，第 1 題

純數學 • 二次方程與判別式

STEP II 1991，第 1 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s2-q1

關鍵思路

一個二次式是某個完全平方（perfect square）的常數倍，當且僅當它有重根（repeated root），亦即它的判別式（discriminant）等於零。其後每一部分，都是把這一個條件應用於係數取決於 λ 的二次式。要找出令判別式等於零的 λ，便會得到一個關於 λ 的二次方程；至於分解式，則是解兩條以 f 和 g 為未知量的線性方程而得。

解答

形式 $(*)$ 的條件

展開得

$$a(x+k)^2 = ax^2 + 2akx + ak^2.$$

若 $\mathrm{h}(x)$ 具有這個形式，則比較係數（comparing coefficients）可得 $b = 2ak$ 及 $c = ak^2$，所以

$$b^2 = 4a^2k^2 = 4a \cdot ak^2 = 4ac.$$

反過來說，若 $b^2 = 4ac$，取 $k = \dfrac{b}{2a}$（由於 $a \neq 0$，這是可行的）。於是 $2ak = b$ 且 $ak^2 = \dfrac{b^2}{4a} = c$，所以 $\mathrm{h}(x) = a(x+k)^2$。因此所求條件為

$$b^2 = 4ac,$$

即判別式等於零。

λ 的兩個值

我們有

$$\mathrm{f}(x)+\lambda \mathrm{g}(x)=(3+\lambda)x^2+4x-2\lambda.$$

若 $\lambda=-3$，上式變成 $4x+6$。這並不屬於形式 $(*)$，因為該形式要求 x^2 的係數不為零。所以 $3+\lambda\neq 0$，而根據上面的條件，我們需要

$$16=4(3+\lambda)(-2\lambda) \iff \lambda^2+3\lambda+2=0 \iff (\lambda+1)(\lambda+2)=0.$$

所以 $\lambda=-1$ 或 $\lambda=-2$，兩者皆不等於 -3。具體寫出來，就是

$$\mathrm{f}(x)-\mathrm{g}(x)=2x^2+4x+2=2(x+1)^2,$$
$$\mathrm{f}(x)-2\mathrm{g}(x)=x^2+4x+4=(x+2)^2.$$

分解式

? **想一想** 關於 $\mathrm{f}-\mathrm{g}$ 和 $\mathrm{f}-2\mathrm{g}$ 的兩項事實，怎樣能讓我們分別求出 f 和 g？

把上面兩條恆等式看成以 $\mathrm{f}(x)$ 和 $\mathrm{g}(x)$ 為未知量的聯立線性方程（simultaneous linear equations）。從第一式減去第二式，得

$$\mathrm{g}(x)=2(x+1)^2-(x+2)^2,$$

然後 $\mathrm{f}(x)=(\mathrm{f}(x)-2\mathrm{g}(x))+2\mathrm{g}(x)=(x+2)^2+4(x+1)^2-2(x+2)^2$，即

$$\mathrm{f}(x)=4(x+1)^2-(x+2)^2.$$

展開驗算：$4(x^2+2x+1)-(x^2+4x+4)=3x^2+4x$，以及 $2(x^2+2x+1)-(x^2+4x+4)=x^2-2$。所以其中一個選擇是

$$A=4,\quad B=-1,\quad C=2,\quad D=-1,\quad m=1,\quad n=2.$$

把 m 和 n 的角色對調，便得到等價的答案 $m=2$、$n=1$、$A=-1$、$B=4$、$C=-1$、$D=2$。

最後部分

現在

$$\mathrm{f}(x)+\lambda \mathrm{g}(x)=(3+\lambda)x^2+4x+\lambda\alpha.$$

與之前一樣，$\lambda=-3$ 給出線性式 $4x-3\alpha$，它永遠不屬於形式 $(*)$。當 $\lambda\neq -3$ 時，條件 $b^2=4ac$ 變成

$$16=4(3+\lambda)\lambda\alpha \iff \alpha\lambda^2+3\alpha\lambda-4=0. \qquad (\dagger)$$

我們要求滿足 $(\dagger)$ 且 $\lambda\neq -3$ 的 λ 恰好只有一個。

想一想 (†) 會不會有兩個根，而其中一個正是不容許的值 $\lambda = -3$？

不會。把 $\lambda = -3$ 代入 (†) 的左方，得 $9\alpha - 9\alpha - 4 = -4 \neq 0$，所以 $\lambda = -3$ 永遠不是根。因此，符合要求的 λ 值正好就是 (†) 的根。

- 若 $\alpha = 0$，(†) 變成 $-4 = 0$，沒有解。所以 $\alpha \neq 0$。
- 若 $\alpha \neq 0$，(†) 是一個名副其實的關於 λ 的二次方程。它恰好有一個根，當且僅當其判別式等於零：

$$9\alpha^2 + 16\alpha = 0 \iff \alpha(9\alpha + 16) = 0 \iff \alpha = -\tfrac{16}{9},$$

  這裏用到 $\alpha \neq 0$。

因此

$$\alpha = -\frac{16}{9}.$$

重根為 $\lambda = -\frac{3\alpha}{2\alpha} = -\frac{3}{2}$，而事實上

$$f(x) - \tfrac{3}{2}\left(x^2 - \tfrac{16}{9}\right) = \tfrac{3}{2}x^2 + 4x + \tfrac{8}{3} = \tfrac{3}{2}\left(x + \tfrac{4}{3}\right)^2.$$

驗算

代入一個測試值來檢查分解式：當 $x = 1$ 時，$4(2)^2 - (3)^2 = 7 = f(1)$，而 $2(2)^2 - 3^2 = -1 = g(1)$。在最後部分，$\frac{3}{2}(x + \frac{4}{3})^2$ 在 $x = 0$ 時的值是 $\frac{3}{2} \cdot \frac{16}{9} = \frac{8}{3}$，與 $\lambda\alpha = (-\frac{3}{2})(-\frac{16}{9})$ 吻合。以電腦測試所有滿足 $|p| \leq 60$、$q \leq 12$ 的有理數 $\lambda = p/q$，結果在中間部分恰好找到 $\lambda = -1, -2$，而當 $\alpha = -\frac{16}{9}$ 時只找到 $\lambda = -\frac{3}{2}$。

常見錯誤

- 忘記 (∗) 中 $a \neq 0$ 的要求：當 $\lambda = -3$ 時，該式是線性的，你應該說明為何要排除這個值。
- 在最後部分，只寫到「(†) 的判別式等於零」便停下來，沒有處理 $\alpha = 0$ 的情況；這時 (†) 根本不是二次方程，而且沒有根。

STEP 之外

把 f 和 g 寫成兩個線性型的平方的組合，就是把一對二次型（quadratic forms）同時對角化（simultaneous diagonalisation）。令 $f + \lambda g$ 退化（degenerate）的 λ 值，就是對應的對稱矩陣（symmetric matrices）所滿足的方程 $\det(F + \lambda G) = 0$ 的根，這是一個「廣

義特徵值」（generalised eigenvalue）問題。在最後部分，兩個 λ 值重合，這正是這種對角化可能失敗的情況。Strang 的 *Introduction to Linear Algebra*（線性代數入門）正是以這種觀點處理二次型和特徵值。

詳細解答

STEP II 1991，第 2 題

純數學 • 雙曲線的漸近線

STEP II 1991，第 2 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s2-q2

關鍵思路

二次部分 $3y^2 - 10xy + 3x^2$ 可因式分解（factorise）為 $(3y - x)(y - 3x)$，而這就是整題的關鍵：只要把因式分解補全，將一次項也包括在內，曲線便成為「兩個一次式之積 $= 1$」。沿曲線走到極遠處時，其中一個因式必定趨向無限大，另一個因式便必定趨向零；令後者等於零，所得的直線就是一條漸近線（asymptote）。這正是提示所要求的「s 與 t 在極限時的關係」。

解答

切線

把 (†) 對 x 作隱函數微分（implicit differentiation），

$$6y\frac{\mathrm{d}y}{\mathrm{d}x} - 10y - 10x\frac{\mathrm{d}y}{\mathrm{d}x} + 6x + 16\frac{\mathrm{d}y}{\mathrm{d}x} - 16 = 0,$$

所以

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{10y - 6x + 16}{6y - 10x + 16} = \frac{5y - 3x + 8}{3y - 5x + 8}.$$

在 (s, t) 處，斜率(gradient)為 $m = \dfrac{5t - 3s + 8}{3t - 5s + 8}$，而切線(tangent)為 $y = mx + (t - ms)$。把截距（intercept）通分，

$$t - ms = \frac{t(3t - 5s + 8) - s(5t - 3s + 8)}{3t - 5s + 8} = \frac{3t^2 - 10st + 3s^2 + 8t - 8s}{3t - 5s + 8}.$$

由於 (s,t) 在曲線上，$3t^2-10st+3s^2=-16t+16s-15$，所以分子為 $-8t+8s-15=-(8t-8s+15)$。因此切線為

$$y=\left(\frac{5t-3s+8}{3t-5s+8}\right)x-\left(\frac{8t-8s+15}{3t-5s+8}\right),$$

正如所求。（當 $3t-5s+8=0$ 時，切線是鉛垂的，此時應把公式理解為去分母後的形式。）

漸近線

嘗試把 (†) 寫成 $(3y-x+a)(y-3x+b)+k=0$。展開得

$$(3y-x+a)(y-3x+b)=3y^2-10xy+3x^2+(3b+a)y-(b+3a)x+ab,$$

所以我們需要 $3b+a=16$ 及 $b+3a=16$，由此得 $a=b=4$ 及 $ab=16$。因此 (†) 即是

$$(3y-x+4)(y-3x+4)=1.$$

對於點 (s,t)，設

$$u=3t-s+4,\qquad v=t-3s+4,\qquad \text{於是}\qquad uv=1.$$

切線方程中的每一個式子，都是 u、v 與常數的組合。比較 t 和 s 的係數，得

$$\begin{aligned}5t-3s+8&=\tfrac{1}{2}(3u+v),\\3t-5s+8&=\tfrac{1}{2}(u+3v),\\8t-8s+15&=2(u+v)-1=2(u+v)-uv,\end{aligned}$$

其中最後一式用了 $uv=1$。所以切線的斜率和截距為

$$m=\frac{3u+v}{u+3v},\qquad c=\frac{8t-8s+15}{3t-5s+8}=\frac{2(2u+2v-uv)}{u+3v}.$$

想一想 當 $s\to\infty$ 時，u 和 v 會怎樣？

消去 t，得 $u-3v=8s-8$，所以 $u-3v\to\infty$。由於 $uv=1>0$，u 和 v 同號。

- 若兩者皆為正數，$u-3v\to\infty$ 迫使 $u\to\infty$，於是 $v=1/u\to 0$。
- 若兩者皆為負數，$u-3v=3|v|-|u|\to\infty$ 迫使 $|v|\to\infty$，於是 $u=1/v\to 0$。

這就是提示所要的關係：在其中一支上 $t-3s+4\to 0$，在另一支上 $3t-s+4\to 0$。

情況 $v\to 0$，$u\to\infty$。把分子和分母同除以 u，

$$m=\frac{3+v/u}{1+3v/u}\to 3,\qquad c=\frac{2(2+2v/u-v)}{1+3v/u}\to 4,$$

所以切線的極限位置是 $y = 3x - 4$。

情況 $u \to 0$，$v \to \infty$。把分子和分母同除以 v，

$$m = \frac{3u/v + 1}{u/v + 3} \to \tfrac{1}{3}, \qquad c = \frac{2(2u/v + 2 - u)}{u/v + 3} \to \tfrac{4}{3},$$

所以切線的極限位置是 $y = \frac{1}{3}x - \frac{4}{3}$，即 $3y = x - 4$。

因此兩條漸近線是 $y = 3x - 4$ 和 $3y = x - 4$。（若改為令 $s \to -\infty$，所得的仍是這兩條直線，只是兩個情況的角色互換。）留意這兩條正是直線 $v = 0$ 和 $u = 0$：從因式分解後的形式一眼便可看出。

夾角與軸

直線 $y = 3x - 4$ 的斜率為 3，所以它與 x 軸的夾角 θ 滿足 $\tan\theta = 3$。它與 y 軸的夾角便是 $\frac{\pi}{2} - \theta$，而 $\tan(\frac{\pi}{2} - \theta) = \cot\theta = \frac{1}{3}$。直線 $3y = x - 4$ 的斜率為 $\frac{1}{3}$，所以它與 x 軸的夾角是 $\arctan\frac{1}{3}$，與第一條漸近線和 y 軸的夾角相同。

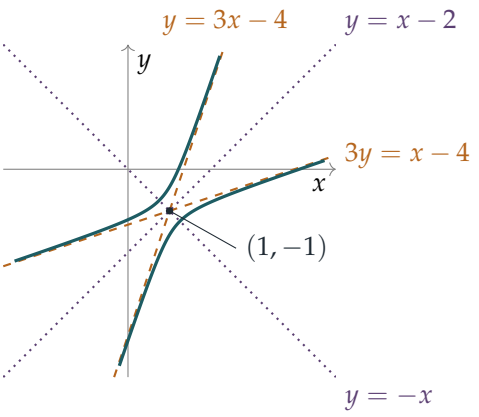
想一想 為甚麼由此可知角平分線是甚麼？

對直線 $y = x$ 作反射（reflection），會把 x 軸和 y 軸互換；而根據剛才所證，它也會把兩條漸近線的方向互換：方向 $(1,3)$ 變成 $(3,1)$。因此兩條漸近線之間的夾角關於 $y = x$ 的方向對稱，亦即被這個方向平分：兩條漸近線的方向與 x 軸分別成 $\arctan 3$ 和 $\arctan\frac{1}{3}$，而這兩個角就是 $\frac{\pi}{4} \pm (\arctan 3 - \frac{\pi}{4})$。另一條角平分線（bisector）與第一條互相垂直。所以兩條角平分線的斜率是 1 和 -1。

雙曲線（hyperbola）的軸，就是通過其中心的兩條漸近線夾角平分線，而中心是兩條漸近線的交點。解 $y = 3x - 4$ 和 $3y = x - 4$：$9x - 12 = x - 4$，所以 $x = 1$，$y = -1$。中心是 $(1, -1)$，而兩條軸是

$$y + 1 = x - 1 \iff y = x - 2, \qquad y + 1 = -(x - 1) \iff y = -x.$$

要分辨哪條是哪條：在 $y = -x$ 上有 $u = v = 4 - 4x$，所以由 $uv = 1$ 得 $x = 1 \pm \frac{1}{4}$，頂點（vertices）是 $(\frac{5}{4}, -\frac{5}{4})$ 和 $(\frac{3}{4}, -\frac{3}{4})$。所以 $y = -x$ 是貫軸（transverse axis）。在 $y = x - 2$ 上有 $u = 2x - 2 = -v$，所以 $uv = -(2x - 2)^2 \le 0$，這條直線（即共軛軸（conjugate axis））永不與曲線相交。



驗算

- 點 $(\frac{5}{4}, -\frac{5}{4})$ 滿足 (†)：$3 \cdot \frac{25}{16} + 10 \cdot \frac{25}{16} + 3 \cdot \frac{25}{16} - 20 - 20 + 15 = 25 - 25 = 0$。
- 以數值檢驗：取 $u = 10^8$（即 $s \approx 1.25 \times 10^7$），切線公式給出斜率 3.0000000 和截距項 3.99999998；取 $u = -10^{-8}$ 則給出 0.33333333 和 1.33333334。
- 中心 $(1, -1)$ 給出 $u = v = 0$，所以它不在曲線上，這就是切線公式永不出現 $0/0$ 的原因。

常見錯誤

- 只從 $3k^2 - 10k + 3 = 0$（設 $t \approx ks$）求出斜率 3 和 $\frac{1}{3}$，卻忘記截距需要用到下一項：$t = 3s - 4 + o(1)$，而不僅是 $t \approx 3s$。
- 把角平分線取為通過原點，而不是通過中心 $(1, -1)$。

STEP 之外

每一條圓錐曲線（conic）$ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$ 都可寫成 $\mathbf{x}^{\mathsf{T}} A \mathbf{x} + \ldots$，其中 $A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$ 是對稱矩陣（symmetric matrix），而曲線的軸沿着 A 的特徵向量（eigenvectors）的方向。這裏 $A = \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}$ 的特徵向量是 $(1, 1)$ 和 $(1, -1)$，對應的特徵值（eigenvalues）-2 和 8 正負相反，曲線因而是雙曲線。像這樣把二次型（quadratic forms）對角化（diagonalising），是大學一年級線性代數（linear algebra）的核心課題；參看 Strang 的 *Introduction to Linear Algebra*（線性代數入門教科書）。

詳細解答

STEP II 1991，第 3 題

純數學 • 對稱聯立方程

> **STEP II 1991，第 3 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q3

關鍵思路

三個相等的式子給出兩條方程，而未知數有三個，所以不能指望求出 x、y、z 本身。我們能做的，是利用當中的循環對稱（cyclic symmetry）。把式子兩兩相減，每個差 $x-y$ 都會變成下一個差 $y-z$ 的倍數；再把三個關係式相乘，幾乎所有東西都會約去。至於第二個結果，把共同值記作 k 並消去分母，便得到三個循環的線性關係式，它們以三種不同方式確定 xyz。

解答

證明 $x^2y^2z^2=1$

令首兩個式子相等，得

$$x-y=\frac{1}{z}-\frac{1}{y}=\frac{y-z}{yz}.$$

同樣地，令第二個與第三個相等，再令第三個與第一個相等，得

$$y-z=\frac{1}{x}-\frac{1}{z}=\frac{z-x}{zx},\qquad z-x=\frac{1}{y}-\frac{1}{x}=\frac{x-y}{xy}.$$

把這三條等式相乘，得

$$(x-y)(y-z)(z-x)=\frac{(y-z)(z-x)(x-y)}{x^2y^2z^2}.$$

由於 x、y 和 z 互不相等（distinct），乘積 $(x-y)(y-z)(z-x)$ 不為零，所以可以把兩邊同時除以它。這樣便剩下 $1=1/(x^2y^2z^2)$，即

$$x^2y^2z^2=1,$$

這正是所要證明的。

> **想一想** 我們究竟在哪裏用到了 x、y、z 互不相等且非零這個條件？

非零是為了令分數 $1/x$、$1/y$、$1/z$ 有意義。互不相等則讓我們可以約去 $(x-y)(y-z)(z-x)$。沒有這個條件，結果便不成立：$x=y=z=2$ 滿足各方程，但 $x^2y^2z^2=64$。

共同值為 ± 1

設 k 為共同值（common value），則

$$x+\frac{1}{y}=k, \qquad y+\frac{1}{z}=k, \qquad z+\frac{1}{x}=k.$$

各式分別乘以 y、z 和 x（全部非零），得

$$xy=ky-1, \qquad yz=kz-1, \qquad zx=kx-1. \tag{$*$}$$

現在把其中第一式乘以 z，並用第二式代換 yz：

$$\begin{aligned} xyz &= k\,yz - z \\ &= k(kz-1)-z \\ &= (k^2-1)z-k. \end{aligned}$$

作代換 $x\to y\to z\to x$ 時，關係式 $(*)$ 保持不變，所以從第二式和第三式出發作同樣的計算，可得

$$xyz=(k^2-1)x-k, \qquad xyz=(k^2-1)y-k.$$

把 xyz 的首兩個表達式相減，得

$$(k^2-1)(z-x)=0.$$

由於 $z\neq x$，必有 $k^2=1$，所以 $x+\frac{1}{y}$ 的值不是 $+1$ 就是 -1，這正是所要證明的。

額外的收穫是：把 $k^2=1$ 代回去，可得 $xyz=-k$。這是得出第一個結果的另一條途徑，因為這樣便有 $x^2y^2z^2=k^2=1$。

驗算

兩個值都確實會出現。當 $k=1$ 時，從任意 $x\neq 0,1$ 出發，取 $y=\frac{1}{1-x}$ 及 $z=\frac{1}{1-y}=\frac{x-1}{x}$。於是 $z+\frac{1}{x}=\frac{x-1}{x}+\frac{1}{x}=1$，所以三個式子都等於 1。例如 $x=2$ 給出 $y=-1$，$z=\frac{1}{2}$：確實有 $2-1=-1+2=\frac{1}{2}+\frac{1}{2}=1$，而且 $xyz=-1=-k$。把 (x,y,z) 換成 $(-x,-y,-z)$，便得到 $k=-1$ 的解，例如 $(-2,1,-\frac{1}{2})$。在這裏三個數永遠互不相等，因為 $x=\frac{1}{1-x}$ 需要 $x^2-x+1=0$，而這條方程沒有實根。我們用精確分數檢驗了約 4000 組這樣的三元組，每次都確認 $x^2y^2z^2=1$ 和 $xyz=-k$。

常見錯誤

- 除以 $(x-y)(y-z)(z-x)$ 時沒有說明它為何不為零。題目給出互不相等的條件，正是為此而設。
- 得出 $xyz = 1$ 而不是 $xyz = \pm 1$ 的結論：事實上 $xyz = -k$，所以當共同值為 1 時，$xyz = -1$。

STEP 之外

映射 $t \mapsto \dfrac{1}{k-t}$ 把 x 變成 y，把 y 變成 z，又把 z 變回 x。它是一個莫比烏斯變換（Möbius transformation），對應的矩陣（matrix）是 $\begin{pmatrix} 0 & 1 \\ -1 & k \end{pmatrix}$，而這類映射的複合正好對應於它們的矩陣相乘。這道題其實是在問：這個映射何時有長度為 3 的循環（cycle）？答案 $k^2 = 1$ 正好就是該矩陣的立方為單位矩陣（identity）的倍數之時。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）有精彩的一章，講述莫比烏斯變換，以及如何按矩陣的跡（trace）把它們分類。

詳細解答

STEP II 1991，第 4 題

純數學 • 餘弦之和

> **STEP II 1991，第 4 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q4

本題在 *Beyond STEP* 第 6.2 節高斯的正十七邊形中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch06/gauss-17gon.html

關鍵思路

由於 $5\phi = 2\pi$ 及 $17\theta = 2\pi$，角的大倍數的餘弦會摺回小倍數的餘弦：當 $n\phi = 2\pi$ 時，$\cos(n-k)\phi = \cos k\phi$。因此，當我們用積化和差公式（product formulae）把餘弦之和平方或相乘時，所有項都會化回同樣的幾個餘弦。在後半部分，我們把八個餘弦 $\cos\theta, \ldots, \cos 8\theta$ 分成兩組，每組四個，求出它們的和與積，從而得到一個以 z 為根（root）的二次方程（quadratic）。

解答

五邊形：$\phi = 2\pi/5$

我們使用積化和差公式 $2\cos A\cos B = \cos(A+B) + \cos(A-B)$，以及 $\cos^2 A = \frac{1}{2}(1+\cos 2A)$。由於 $5\phi = 2\pi$，

$$\cos 3\phi = \cos(2\pi - 2\phi) = \cos 2\phi, \qquad \cos 4\phi = \cos(2\pi - \phi) = \cos\phi.$$

因此

$$\begin{aligned} y^2 &= \cos^2\phi + 2\cos\phi\cos 2\phi + \cos^2 2\phi \\ &= \tfrac{1}{2}(1+\cos 2\phi) + (\cos 3\phi + \cos\phi) + \tfrac{1}{2}(1+\cos 4\phi) \\ &= \tfrac{1}{2}(1+\cos 2\phi) + (\cos 2\phi + \cos\phi) + \tfrac{1}{2}(1+\cos\phi) \\ &= 1 + \tfrac{3}{2}(\cos\phi + \cos 2\phi) = 1 + \tfrac{3}{2}y. \end{aligned}$$

兩邊乘以 2，得 $2y^2 = 3y + 2$，正是所求。

此式可因式分解為 $(2y+1)(y-2) = 0$，所以 $y = -\frac{1}{2}$ 或 $y = 2$。但 $2\phi = \frac{4\pi}{5}$ 介乎 $\frac{\pi}{2}$ 與 π 之間，故 $\cos 2\phi < 0$，而 $\cos\phi \le 1$。於是 $y < 1$，這就排除了 $y = 2$，所以

$$y = \cos\tfrac{2\pi}{5} + \cos\tfrac{4\pi}{5} = -\tfrac{1}{2}.$$

$\theta = 2\pi/17$ 時各餘弦之和

設 $\omega = \cos\theta + \mathrm{i}\sin\theta$。根據棣美弗定理（de Moivre's theorem），$\omega^k = \cos k\theta + \mathrm{i}\sin k\theta$，並且 $\omega^{17} = \cos 2\pi + \mathrm{i}\sin 2\pi = 1$，而 $\omega \ne 1$。把等比級數（geometric series）求和，

$$\sum_{k=0}^{16} \omega^k = \frac{\omega^{17} - 1}{\omega - 1} = 0.$$

取實部（real parts），得 $\displaystyle\sum_{k=0}^{16} \cos k\theta = 0$，正是所求。

z 的值

記 $c_k = \cos k\theta$。由於 $17\theta = 2\pi$，而且餘弦是偶函數（even function），

$$c_{17-k} = c_k \quad \text{及} \quad c_{-k} = c_k,$$

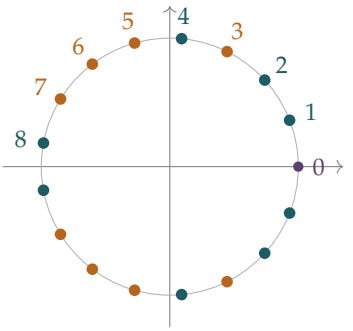
所以每個 c_k 都等於 $c_1, c_2, \ldots, c_8$ 其中之一。特別地，在剛證明的和式中除去 $c_0 = 1$ 後，就得到 $2(c_1 + c_2 + \cdots + c_8) = -1$，所以

$$c_1 + c_2 + \cdots + c_8 = -\tfrac{1}{2}. \qquad (*)$$

z 中的角是 $1, 2, 4, 8$（每個都是前一個的兩倍，並按上述方法化簡：$16 \to 1$）。其餘四個 $3, 5, 6, 7$ 給出與之配對的和

$$w = c_3 + c_5 + c_6 + c_7.$$

下圖顯示單位圓（unit circle）上的十七個點 ω^k：藍綠色的點（及其鏡像）對應 z 中的餘弦，琥珀色的點則對應 w 中的餘弦。



和。 由 $(*)$，$z+w=c_1+c_2+\cdots+c_8=-\frac{1}{2}$。

積。 展開 zw 會得到十六個乘積 c_ac_b，其中 $a\in\{1,2,4,8\}$ 而 $b\in\{3,5,6,7\}$，每個乘積都等於 $\frac{1}{2}(c_{a+b}+c_{a-b})$。把每個下標化到 1 至 8 的範圍內（利用 $c_{-k}=c_k$ 及 $c_{17-k}=c_k$），所得的各對下標如下：

	$b=3$	$b=5$	$b=6$	$b=7$
$a=1$	4, 2	6, 4	7, 5	8, 6
$a=2$	5, 1	7, 3	8, 4	8, 5
$a=4$	7, 1	8, 1	7, 2	6, 3
$a=8$	6, 5	4, 3	3, 2	2, 1

（例如 $a=8$、$b=7$ 給出 c_{15} 和 c_1，而 $c_{15}=c_2$。）

想一想 我們需要知道出現的是哪些餘弦，還是只需知道各出現多少次？

只需知道次數。點算表中各項，下標 $1,2,\ldots,8$ 在 32 個項目中各自恰好出現四次。所以

$$zw=\tfrac{1}{2}\cdot 4\,(c_1+c_2+\cdots+c_8)=2\cdot\left(-\tfrac{1}{2}\right)=-1,$$

這裏再次用到 $(*)$。

二次方程。 由於 $z+w=-\frac{1}{2}$ 及 $zw=-1$，z 和 w 這兩個數是以下方程的根：

$$t^2+\tfrac{1}{2}t-1=0,\qquad 即\qquad 2t^2+t-2=0,$$

所以 $\{z,w\}=\left\{\dfrac{-1+\sqrt{17}}{4},\ \dfrac{-1-\sqrt{17}}{4}\right\}$。第一個根是正數，第二個是負數。

選取正確的根。 我們證明 $z>0$。由於 $\frac{4}{17}<\frac{1}{4}$，θ 和 2θ 都在 $(0,\frac{\pi}{4})$ 內，所以 c_1 和 c_2 都大於 $\frac{1}{\sqrt{2}}$，而 $c_1+c_2>\sqrt{2}>1$。此外 $4\theta=\frac{8\pi}{17}<\frac{\pi}{2}$，故 $c_4>0$，而且 $c_8\geq -1$。因此

$$z=(c_1+c_2)+c_4+c_8>1+0-1=0,$$

所以 z 必定是正根：

$$z=\frac{-1+\sqrt{17}}{4}=-\frac{1-\sqrt{17}}{4}.$$

驗算

以數值計算，$\cos 72^\circ+\cos 144^\circ=0.30902-0.80902=-0.5$。取 $\theta=2\pi/17$，用計算機可得 $z=0.932472+0.739009+0.092268-0.982973=0.780776$，而 $(\sqrt{17}-1)/4=0.780776$；另外 $w=-1.280776=(-1-\sqrt{17})/4$，所以 $z+w=-0.5$ 及 $zw=-1$，與所得結果相符。

常見錯誤

- 在第一部分忘記排除 $y=2$，或在第二部分忘記排除 w 所對應的根：二次方程給出兩個可能的答案，必須另外用正負號的論證來選出其中一個。
- 在 $(*)$ 中把 $k=0$ 也算進去：整個和是 0，但 $c_0=1$，所以 $c_1+\cdots+c_{16}=-1$，除以二才得 $-\frac{1}{2}$。
- 在乘積表中化簡下標時出錯，例如把 c_9 當作新的一項，而不是 c_8。

STEP 之外

這就是高斯（Gauss）的著名計算，它證明了正 17 邊形可以用直尺和圓規作出。按 3 的冪（模 17 的一個生成元（generator））把各餘弦分成層層嵌套、大小依次為 8、4、2 和 1 的組，就得到一連串二次方程，所以 $\cos\frac{2\pi}{17}$ 可以只用平方根寫出；z 和 w 是這一連串方程的第一步，稱為高斯週期（Gaussian periods）。完整的故事屬於伽羅瓦理論（Galois theory）：參閱 Allenby 的 *Rings, Fields and Groups*（環、域與羣）；至於單位根（roots of unity）的算術，可參閱 Davenport 的 *The Higher Arithmetic*（高等算術）。

詳細解答

STEP II 1991，第 5 題

純數學 • 積分歸約公式

STEP II 1991，第 5 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s2-q5

關鍵思路

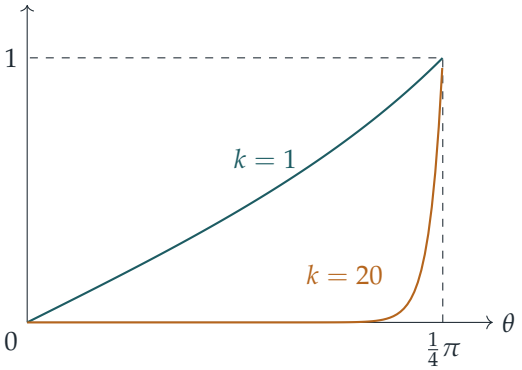
記 $I_k = \int_0^{\pi/4} \tan^k\theta\,\mathrm{d}\theta$。tan 的奇數次冪和偶數次冪都不能直接積分，但把相隔兩次的兩個冪相加就可以：$\tan^k\theta + \tan^{k+2}\theta = \tan^k\theta\sec^2\theta$，而這正是 $\tan^{k+1}\theta/(k+1)$ 的導數（derivative）。由此得到的歸約公式（reduction formula），用歸納法（induction）逐層展開，便得出所求的公式；而 k 很大時的草圖告訴你積分會縮小至零，正是這一點把公式變成無窮級數（infinite series）的和。

解答

草圖

在 $[0, \frac{1}{4}\pi]$ 上，$0 \le \tan\theta \le 1$，且 $\tan 0 = 0$，$\tan\frac{1}{4}\pi = 1$。當 $k = 1$ 時，圖像由 $(0,0)$ 上升至 $(\frac{1}{4}\pi, 1)$，並稍微向上彎曲（起點的斜率為 1，終點的斜率為 $\sec^2\frac{1}{4}\pi = 2$）。

當 $k \gg 1$ 時，$\tan\theta$ 每一個小於 1 的值經過高次乘方後都被壓向 0，而兩端的值 0 和 1 則保持不變。因此 $\tan^k\theta$ 在區間的大部分範圍內都緊貼橫軸，到了非常接近 $\theta = \frac{1}{4}\pi$ 之處才急升至 1。事實上，設 $\theta = \frac{1}{4}\pi - u$，其中 u 很小，則 $\tan\theta \approx 1 - 2u$，所以 $\tan^k\theta \approx (1-2u)^k \approx \mathrm{e}^{-2ku}$：這個尖峯的寬度約為 $1/(2k)$，其下的面積約為 $1/(2k)$。



歸約公式

對 $k \geq 0$，

$$I_k + I_{k+2} = \int_0^{\pi/4} \tan^k\theta\,(1+\tan^2\theta)\,\mathrm{d}\theta = \int_0^{\pi/4} \tan^k\theta\sec^2\theta\,\mathrm{d}\theta = \Big[\frac{\tan^{k+1}\theta}{k+1}\Big]_0^{\pi/4} = \frac{1}{k+1}.$$

我們還需要兩個初始值：

$$I_0 = \tfrac{1}{4}\pi, \qquad I_1 = \int_0^{\pi/4} \frac{\sin\theta}{\cos\theta}\,\mathrm{d}\theta = \Big[-\ln\cos\theta\Big]_0^{\pi/4} = -\ln\tfrac{1}{\sqrt{2}} = \tfrac{1}{2}\ln 2.$$

奇數次冪

設 $J_n = I_{2n+1}$。在歸約公式中取 $k = 2n-1$，得

$$J_n = \frac{1}{2n} - J_{n-1} \qquad (n \geq 1).$$

我們用歸納法證明所求的公式。當 $n=1$ 時：$J_1 = \frac{1}{2} - J_0 = \frac{1}{2} - \frac{1}{2}\ln 2 = (-1)^1\big(\frac{1}{2}\ln 2 - \frac{1}{2}\big)$，這正是 $n=1$ 時的公式。

假設公式對 $n-1 \geq 1$ 成立。則

$$\begin{aligned} J_n &= \frac{1}{2n} - (-1)^{n-1}\Big(\tfrac{1}{2}\ln 2 + \sum_{m=1}^{n-1}\frac{(-1)^m}{2m}\Big) \\ &= (-1)^n\Big(\tfrac{1}{2}\ln 2 + \sum_{m=1}^{n-1}\frac{(-1)^m}{2m}\Big) + (-1)^n\frac{(-1)^n}{2n} \\ &= (-1)^n\Big(\tfrac{1}{2}\ln 2 + \sum_{m=1}^{n}\frac{(-1)^m}{2m}\Big), \end{aligned}$$

其中利用了 $(-1)^n(-1)^n = 1$，把 $\frac{1}{2n}$ 寫成括號內 $m=n$ 的一項。根據歸納法，公式對所有正整數 n 成立，證畢。

第一個級數

> **想一想** 我們怎樣知道 $J_n \to 0$，而不只是「從草圖看來很小」？

我們需要一個確確實實的上界（bound）。在 $[0, \frac{1}{4}\pi]$ 上，$\tan\theta \geq 0$ 且 $\sec^2\theta \geq 1$，所以

$$0 \leq J_n = \int_0^{\pi/4} \tan^{2n+1}\theta \,\mathrm{d}\theta \leq \int_0^{\pi/4} \tan^{2n+1}\theta \sec^2\theta \,\mathrm{d}\theta = \frac{1}{2n+2}.$$

因此當 $n \to \infty$ 時 $J_n \to 0$，這正是 k 很大時的草圖所提示的（而且上界 $1/(2n+2)$ 與尖峯面積約為 $1/(2k)$ 相符）。由於 $|J_n|$ 就是公式中括號的絕對值，

$$\left|\tfrac{1}{2}\ln 2 + \sum_{m=1}^{n} \frac{(-1)^m}{2m}\right| = J_n \leq \frac{1}{2n+2} \to 0,$$

所以部分和（partial sum）$\sum_{m=1}^{n} \frac{(-1)^m}{2m}$ 收斂（converge）至 $-\frac{1}{2}\ln 2$。乘以 -1，得

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2m} = \tfrac{1}{2}\ln 2.$$

第二個級數

現在對偶數次冪照樣做一遍。設 $K_n = I_{2n}$。在歸約公式中取 $k = 2n-2$，得

$$K_n = \frac{1}{2n-1} - K_{n-1} \qquad (n \geq 1), \qquad K_0 = \tfrac{1}{4}\pi.$$

完全相同的歸納法（這次由 $K_1 = 1 - \frac{1}{4}\pi$ 開始）給出

$$K_n = \int_0^{\pi/4} \tan^{2n}\theta \,\mathrm{d}\theta = (-1)^n\left(\tfrac{1}{4}\pi + \sum_{m=1}^{n} \frac{(-1)^m}{2m-1}\right):$$

歸納步驟是 $K_n = \frac{1}{2n-1} - (-1)^{n-1}(\cdots) = (-1)^n(\cdots) + (-1)^n\frac{(-1)^n}{2n-1}$，與之前一樣。同樣的上界給出

$$0 \leq K_n \leq \int_0^{\pi/4} \tan^{2n}\theta\sec^2\theta \,\mathrm{d}\theta = \frac{1}{2n+1} \to 0,$$

所以 $\sum_{m=1}^{n} \frac{(-1)^m}{2m-1} \to -\frac{1}{4}\pi$，即

$$\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{2m-1} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \frac{\pi}{4}.$$

驗算

- 取 $n = 1$：$\int_0^{\pi/4} \tan^3\theta \,\mathrm{d}\theta = \frac{1}{2} - \frac{1}{2}\ln 2 \approx 0.15343$，而對該積分使用森遜法則（Simpson's rule）所得的結果，準確至小數點後十位都相同。我們用這個方法驗算了 $n = 1, \ldots, 11$ 時 I_{2n+1} 和 I_{2n} 的公式。

- 當 k 很大時，數值計算得 $I_{100} \approx 0.0049995$ 及 $I_{1000} \approx 0.00049999996$，非常接近 $1/(2k)$，與草圖的預測一致。
- 兩個級數各取一百萬項，得 $0.346573\ldots$ 和 $0.785398\ldots$，而 $\frac{1}{2}\ln 2 = 0.346574$，$\frac{1}{4}\pi = 0.785398$。誤差大約等於下一項的大小，與上面的上界所提示的一致。

常見錯誤

- 只說「$\tan^{2n+1}\theta \to 0$，所以積分趨向 0」而不給出上界。被積函數（integrand）在 $\theta = \frac{1}{4}\pi$ 處並不趨向 0，所以必須提出理由，例如利用 $\sec^2\theta \geq 1$ 所得的上界。
- 歸納過程中的正負號錯誤：把 $(-1)^n$ 留在括號外，並以直接計算核對 $n = 1$ 的情況。

STEP 之外

第一個級數是交錯調和級數（alternating harmonic series）$1 - \frac{1}{2} + \frac{1}{3} - \cdots = \ln 2$ 的一半，第二個則是 π 的萊布尼茲（Leibniz）級數。兩者都只是條件收斂（conditionally convergent）：重新排列各項可以改變級數的和，這個令人驚訝的事實（黎曼（Riemann）重排定理）在大學第一門分析課程的初段便會遇到。這裏的論證，即寫出一個精確的誤差項再把它夾逼至零，與帶餘項的泰勒（Taylor）定理是同一個想法。Abbott 的 *Understanding Analysis* 對兩者都有仔細的論述。

詳細解答

STEP II 1991，第 6 題

純數學 • 積分的界限

STEP II 1991，第 6 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s2-q6

關鍵思路

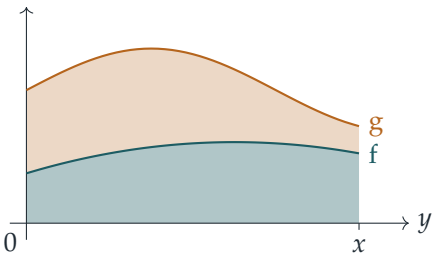
由 0 到 x 的積分（integral）就是面積，而一個較小的非負函數所圍出的面積也較小。只要反覆運用這一個事實，就能把已知的不等式（inequality）變成新的不等式。把 $0 \leq \cos y \leq 1$ 積分，得到 $\sin x$ 的一個界（bound）；再把它積分，得到 $\cos x$ 的一個界；再積分一次，便得到 $\sin x$ 更精確的界。之後每個積分的界，都是把被積函數（integrand）夾在兩個可以準確積分的函數之間而得出的。

解答

比較積分

由於在 $[0, x]$ 上 $\mathrm{f}(y) \geq 0$，f 的圖像位於 y 軸之上或剛好在軸上，所以 $\int_0^x \mathrm{f}(y)\,\mathrm{d}y$ 是圖像與軸之間的面積，必為非負。由於 $\mathrm{f}(y) \leq \mathrm{g}(y)$，f 之下的區域包含在 g 之下的區域之內，所以它的面積不會更大：

$$0 \leq \int_0^x \mathrm{f}(y)\,\mathrm{d}y \leq \int_0^x \mathrm{g}(y)\,\mathrm{d}y.$$



（以分析方式說明：$\mathrm{g} - \mathrm{f} \geq 0$，所以 $\int_0^x (\mathrm{g} - \mathrm{f})\,\mathrm{d}y \geq 0$，而積分是線性的。）

$\sin x$ 與 $\cos x$ 的界

取 $0 \le x \le \frac{1}{2}\pi$。對於 $0 \le y \le x$，我們有 $0 \le \cos y \le 1$，所以在上述結果中取 $\mathrm{f} = \cos$、$\mathrm{g} = 1$，得

$$0 \le \int_0^x \cos y \,\mathrm{d}y \le \int_0^x 1 \,\mathrm{d}y, \qquad \text{即} \qquad 0 \le \sin x \le x.$$

這對 $[0, \frac{1}{2}\pi]$ 內的每一個 x 都成立，所以特別地，當 $0 \le y \le x$ 時有 $0 \le \sin y \le y$。再次應用該結果，這次取 $\mathrm{f} = \sin$、$\mathrm{g}(y) = y$：

$$\int_0^x \sin y \,\mathrm{d}y \le \int_0^x y \,\mathrm{d}y \quad \Longrightarrow \quad 1 - \cos x \le \tfrac{1}{2}x^2 \quad \Longrightarrow \quad \cos x \ge 1 - \tfrac{1}{2}x^2.$$

第一個積分

在 $[0, \frac{1}{10}]$ 上（完全位於 $[0, \frac{1}{2}\pi]$ 之內），我們現在有 $1 - \frac{1}{2}x^2 \le \cos x \le 1$，所以

$$3 - \tfrac{1}{2}x^2 \le 2 + \cos x \le 3.$$

想一想 我們可以直接平方再取倒數嗎？

可以，因為這裏每一個量都是正數：$3 - \frac{1}{2}x^2 \ge 3 - \frac{1}{200} > 0$。對正數而言，平方會保持大小次序，取倒數（reciprocal）則會把次序倒轉。然後乘以 $x \ge 0$，得

$$\frac{x}{9} \le \frac{x}{(2+\cos x)^2} \le \frac{x}{(3 - \frac{1}{2}x^2)^2} \qquad (0 \le x \le \tfrac{1}{10}).$$

三個函數全都非負，所以我們可以由 0 積分到 $\frac{1}{10}$。左邊給出

$$\int_0^{1/10} \frac{x}{9} \,\mathrm{d}x = \frac{1}{18} \cdot \frac{1}{100} = \frac{1}{1800}.$$

至於右邊，分子 x 除正負號外正是 $3 - \frac{1}{2}x^2$ 的導數（derivative），而 $\frac{\mathrm{d}}{\mathrm{d}x}\left(3 - \frac{1}{2}x^2\right)^{-1} = x\left(3 - \frac{1}{2}x^2\right)^{-2}$。所以

$$\int_0^{1/10} \frac{x}{(3 - \frac{1}{2}x^2)^2} \,\mathrm{d}x = \left[\frac{1}{3 - \frac{1}{2}x^2}\right]_0^{1/10} = \frac{200}{599} - \frac{1}{3} = \frac{600 - 599}{1797} = \frac{1}{1797}.$$

因此

$$\frac{1}{1800} \le \int_0^{1/10} \frac{x}{(2+\cos x)^2} \,\mathrm{d}x \le \frac{1}{1797}.$$

$\sin x$ 更精確的界

對於 $0 \le y \le x \le \frac{1}{2}\pi$，我們有 $\cos y \ge 1 - \frac{1}{2}y^2$，即 $0 \le 1 - \cos y \le \frac{1}{2}y^2$（左邊的不等式成立是因為 $\cos y \le 1$）。由 0 積分到 x：

$$\int_0^x (1 - \cos y) \,\mathrm{d}y \le \int_0^x \tfrac{1}{2}y^2 \,\mathrm{d}y \quad \Longrightarrow \quad x - \sin x \le \tfrac{1}{6}x^3 \quad \Longrightarrow \quad \sin x \ge x - \tfrac{1}{6}x^3.$$

第二個積分

在 $[0, \frac{1}{10}]$ 上，我們有 $x - \frac{1}{6}x^3 \le \sin x \le x$，所以

$$1 - \tfrac{1}{6}x^3 \le 1 - x + \sin x \le 1.$$

> **想一想** 為甚麼應該考慮這個組合？

分母中的 $-x$ 正是用來抵消 $\sin x$ 的首項，餘下 1 加上一個 x^3 階的量。下界 $1 - \frac{1}{6}x^3 \ge 1 - \frac{1}{6000}$ 是正數，所以和之前一樣，我們先平方，再取倒數（次序倒轉），然後乘以 $x^2 \ge 0$：

$$x^2 \le \frac{x^2}{(1 - x + \sin x)^2} \le \frac{x^2}{(1 - \frac{1}{6}x^3)^2}.$$

積分後，左邊給出 $\int_0^{1/10} x^2\,\mathrm{d}x = \frac{1}{3}\cdot\frac{1}{1000} = \frac{1}{3000}$。至於右邊，$\frac{\mathrm{d}}{\mathrm{d}x}\left(1 - \frac{1}{6}x^3\right)^{-1} = \frac{1}{2}x^2\left(1 - \frac{1}{6}x^3\right)^{-2}$，所以

$$\int_0^{1/10} \frac{x^2}{(1 - \frac{1}{6}x^3)^2}\,\mathrm{d}x = 2\left[\frac{1}{1 - \frac{1}{6}x^3}\right]_0^{1/10} = 2\left(\frac{6000}{5999} - 1\right) = \frac{2}{5999}.$$

因此

$$\frac{1}{3000} \le \int_0^{1/10} \frac{x^2}{(1 - x + \sin x)^2}\,\mathrm{d}x \le \frac{2}{5999}.$$

驗算

- 用辛普森法則（Simpson's rule）得 $\int_0^{1/10} x(2 + \cos x)^{-2}\,\mathrm{d}x \approx 0.00055648$，介乎 $1/1800 \approx 0.00055556$ 與 $1/1797 \approx 0.00055648$ 之間（上界極為貼近：兩者相差僅約 5×10^{-10}）。
- 同樣地，$\int_0^{1/10} x^2(1 - x + \sin x)^{-2}\,\mathrm{d}x \approx 0.000333389$，介乎 $1/3000 \approx 0.000333333$ 與 $2/5999 \approx 0.000333389$ 之間。
- 在 $[0, \frac{1}{2}\pi]$ 內隨機抽取的 10^5 個點上，三個逐點不等式全部成立，並在 $x = 0$ 時取等號。

常見錯誤

- 沒有檢查正負號就取倒數：只有當 a 與 b 同號時，$a \le b$ 才能推出 $1/a \ge 1/b$。
- 忘記了比較積分的結果需要 f ≥ 0 才能得出左邊的不等式；這裏是由 $\sin y \ge 0$ 和 $1 - \cos y \ge 0$ 提供這個條件。

- 用 $\cos x \leq 1$ 這個界來求積分的上界：$\cos x \leq 1$ 令分母小於 9，得出的是下界 1/1800，而不是上界。

STEP 之外

反覆把 $0 \leq \cos y \leq 1$ 積分，可以得出一整串正負交錯的泰勒（Taylor）界 $x - \frac{1}{6}x^3 \leq \sin x \leq x - \frac{1}{6}x^3 + \frac{1}{120}x^5 \leq \cdots$，這是證明 sin 和 cos 的泰勒級數（Taylor series）收斂（converge）於它們本身的其中一種方法。積分的單調性（monotonicity）以及被積函數的比較，都是實分析（real analysis）的基本工具，用來估計餘項、誤差和極限的界。Abbott 的 *Understanding Analysis*（一本實分析入門書）仔細地建立了積分理論和帶餘項的泰勒定理（Taylor's theorem）。

詳細解答

STEP II 1991，第 7 題

純數學 • 函數方程

> **STEP II 1991，第 7 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q7

關鍵思路

函數方程（functional equation）對*所有* x 和 y 都成立，因此我們可以固定其中一個變量，對另一個變量求導。這樣便把它化為一條微分方程（differential equation），把 g' 在兩個不同點的值聯繫起來。然後巧妙地選取其中一個變量的值，就能確定 g' 在各處的值，最後積分便大功告成。全篇我們都跟題目一樣，假設 g 和 f 可導（differentiable）。

解答

導數關係式

固定 $y>0$，把 $\mathrm{g}(x)+\mathrm{g}(y)=\mathrm{g}(z)$ 對 x 求導。由於 $\mathrm{g}(y)$ 此時是常數，由鏈式法則（chain rule）得

$$\mathrm{g}'(x)=\mathrm{g}'(z)\,\frac{\partial z}{\partial x}.$$

由商法則（quotient rule），

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

這就是第一個形式。至於第二個形式，留意 $z+1$ 可以漂亮地因式分解：

$$z+1=\frac{xy+x+y+1}{x+y+1}=\frac{(x+1)(y+1)}{x+y+1},\qquad\text{所以}\qquad z(z+1)=\frac{xy(x+1)(y+1)}{(x+y+1)^2}.$$

除以 $x(x+1)$ 後恰好得到 $\dfrac{y(y+1)}{(x+y+1)^2}$，所以

$$\mathrm{g}'(x)=\frac{y^2+y}{(x+y+1)^2}\,\mathrm{g}'(z)=\frac{z(z+1)}{x(x+1)}\,\mathrm{g}'(z),$$

正如所求。值得把它寫成以下的對稱形式：

$$x(x+1)\,g'(x) = z(z+1)\,g'(z). \qquad (\dagger)$$

推導 $2g'(1) = (u^2+u)g'(u)$

> **想一想** 我們想得到一個對 $(0,1)$ 內每一個 u 都成立的結論。應該固定哪一個變量？固定在甚麼值？

在 $(\dagger)$ 中代入 $x = 1$；左邊變成 $2g'(1)$，而 $z = \dfrac{y}{y+2}$。當 y 取遍 $(0,\infty)$ 時，$\dfrac{y}{y+2} = 1 - \dfrac{2}{y+2}$ 由 0 連續地遞增並趨向 1，所以它取遍所有滿足 $0 < u < 1$ 的值 u（具體來說，$y = \frac{2u}{1-u}$）。因此

$$2g'(1) = (u^2+u)\,g'(u) \qquad \text{對所有} 0 < u < 1.$$

積分

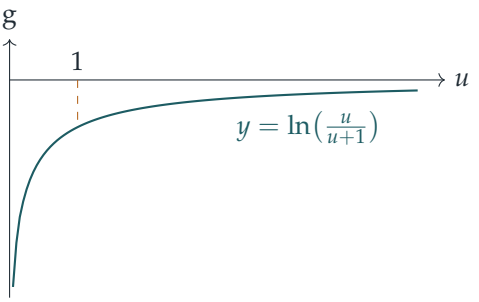
設 $A = 2g'(1)$，這是一個常數。那麼對 $0 < u < 1$，利用部分分式（partial fractions），

$$g'(u) = \frac{A}{u(u+1)} = A\left(\frac{1}{u} - \frac{1}{u+1}\right),$$

再積分（各項皆為正數，所以不需要絕對值符號），

$$g(u) = A\big(\ln u - \ln(u+1)\big) + B = A\ln\left(\frac{u}{u+1}\right) + B.$$

其實這條公式對每一個 $u > 0$ 都成立，而不只是 $u < 1$。對固定的 x，當 y 由 0 增至 ∞ 時，$z = \dfrac{xy}{x+y+1}$ 由 0 遞增並趨向 x，所以我們可以選取 y 使得 $z < 1$。這樣，由 $(\dagger)$ 和上面的結果可得 $x(x+1)g'(x) = z(z+1)g'(z) = A$。因此同樣的積分在整個 $(0,\infty)$ 上都適用，而且只有一個常數 B，因為 $(0,\infty)$ 是一個區間。



（圖中所示的曲線是 $A = 1$、$B = 0$ 的情況：它取負值、遞增，並且當 $u \to \infty$ 時趨向 0。）

驗證 $(*)$ 並求 B

代入 $\mathrm{g}(u) = A\ln\frac{u}{u+1} + B$。$(*)$ 的左邊是

$$\mathrm{g}(x) + \mathrm{g}(y) = A\ln\frac{xy}{(x+1)(y+1)} + 2B.$$

至於右邊，利用先前求得的 $z+1$ 的因式分解，可得

$$\frac{z}{z+1} = \frac{xy}{x+y+1}\cdot\frac{x+y+1}{(x+1)(y+1)} = \frac{xy}{(x+1)(y+1)},$$

所以

$$\mathrm{g}(z) = A\ln\frac{xy}{(x+1)(y+1)} + B.$$

兩邊對所有 $x, y > 0$ 都相等，當且僅當 $2B = B$，即 $B = 0$。所以 $\mathrm{g}(u) = A\ln\dfrac{u}{u+1}$ 滿足 $(*)$。

能否確定 A？

不能。上面的驗證對 A 的每一個值都成立，包括 $A = 0$（零函數）。定義 $A = 2\mathrm{g}'(1)$ 也幫不上忙：對我們的解而言，$\mathrm{g}'(1) = A\left(1 - \frac{1}{2}\right) = \frac{A}{2}$，所以 $A = 2\mathrm{g}'(1)$ 自動成立，並沒有提供任何新資訊。方程 $(*)$ 對 g 而言是線性的，所以一個解的任何倍數仍然是解，而 A 正是這個比例因子。

函數 f

同樣的方法也適用。固定 $y > 0$，把 $\mathrm{f}(x) + \mathrm{f}(y) = \mathrm{f}(xy)$ 對 x 求導：

$$\mathrm{f}'(x) = y\,\mathrm{f}'(xy).$$

> ? **想一想** 在這裏，與代入 $x = 1$ 相對應的做法是甚麼？

代入 $x = 1$，得 $\mathrm{f}'(1) = y\,\mathrm{f}'(y)$ 對每一個 $y > 0$ 都成立，所以設 $C = \mathrm{f}'(1)$，便有

$$\mathrm{f}'(y) = \frac{C}{y} \qquad\Longrightarrow\qquad \mathrm{f}(y) = C\ln y + D \quad (y > 0).$$

代回原式，$\mathrm{f}(x) + \mathrm{f}(y) = C\ln(xy) + 2D$，而 $\mathrm{f}(xy) = C\ln(xy) + D$，所以 $D = 0$。（等價地，在方程中取 $x = y = 1$ 得 $2\mathrm{f}(1) = \mathrm{f}(1)$，所以 $\mathrm{f}(1) = 0$。）因此 $\mathrm{f}(x) = C\ln x$，正如所求，而 C 同樣是任意的。

驗算

► 取 $x = 1$、$y = 2$：則 $z = \frac{1}{2}$，而兩個導數因子分別是 $\frac{6}{16}$ 和 $\frac{\frac{1}{2}\cdot\frac{3}{2}}{2} = \frac{3}{8}$，兩者相符。

- 取 $A=1$：$\mathrm{g}(1)+\mathrm{g}(2)=\ln\frac{1}{2}+\ln\frac{2}{3}=\ln\frac{1}{3}$，而 $\mathrm{g}(\frac{1}{2})=\ln\frac{1/2}{3/2}=\ln\frac{1}{3}$。✓
- 我們用電腦以有限差分（finite differences）在 1000 個隨機點上核實了導數恆等式，並對 A 的若干個值核實了 $\mathrm{g}(x)+\mathrm{g}(y)-\mathrm{g}(z)$ 恰好等於 B（所以只有當 $B=0$ 時才等於零）。

常見錯誤

- 對 x 求導時忘記 $\mathrm{g}(y)$ 是常數，或者漏掉鏈式法則的因子 $\partial z/\partial x$。
- 代入了 $x=1$，卻沒有檢查 z 是否真的取遍整個 $(0,1)$；正是這一步，把單獨一條等式變成對每一個 u 都成立的結論。
- 「因為積分常數無關重要」而假定 B 必定是 0：其實是方程 $(*)$ 迫使 $B=0$，而這一點需要證明。

STEP 之外

第二條方程是柯西（Cauchy）對數型函數方程：代入 $x=\mathrm{e}^{s}$ 便把它化為 $h(s+t)=h(s)+h(t)$。這裏我們假設了可導性；在大學第一門分析課程中，你會看到只要在單獨一點連續（continuity），就已經足以迫使 $h(s)=cs$，而如果完全沒有任何正則性條件，就會出現怪異的非線性解（利用哈默爾基（Hamel basis）構造而成）。第一條方程其實是同一回事的變相：$w=\ln\frac{u}{u+1}$ 把 $(*)$ 變換成加性方程。Abbott 的 *Understanding Analysis*（一本實分析入門書）是閱讀這類連續性論證的好去處。

詳細解答

STEP II 1991，第 8 題

純數學 • 非線性一階常微分方程

> **STEP II 1991，第 8 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q8

關鍵思路

兩條微分方程（differential equation）都是關於 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 的二次方程（quadratic），所以我們可以先用代數方法解出導數（derivative），然後才積分。每條二次方程有兩個根，於是得到兩條候選的一階方程，而 $x = 0$ 處的條件決定了哪一個分支（branch）可取。雙曲恆等式（hyperbolic identities），尤其是 $\cosh^2 x - \sinh^2 x = 1$ 和半角公式（half-angle formulae），可令這些根保持簡潔。

解答

二次方程

由二次公式（quadratic formula），

$$u = \frac{-2\sinh x \pm \sqrt{4\sinh^2 x + 4}}{2} = -\sinh x \pm \sqrt{\sinh^2 x + 1} = -\sinh x \pm \cosh x,$$

當中用到 $\cosh^2 x - \sinh^2 x = 1$ 和 $\cosh x > 0$。由於 $\cosh x - \sinh x = \mathrm{e}^{-x}$ 而 $\cosh x + \sinh x = \mathrm{e}^{x}$，兩個根是

$$u = \mathrm{e}^{-x} \qquad \text{或} \qquad u = -\mathrm{e}^{x}.$$

（驗算：兩根之積是 -1，兩根之和是 $\mathrm{e}^{-x} - \mathrm{e}^{x} = -2\sinh x$，與應有的結果相符。）

第一條微分方程

取 $u = \frac{\mathrm{d}y}{\mathrm{d}x}$，這正是上面的二次方程，所以對每一個 x，要麼 $y' = \mathrm{e}^{-x}$，要麼 $y' = -\mathrm{e}^{x}$。

想一想 y' 會不會在某些 x 值上取其中一個根，而在其他地方取另一個根？

不會。第一個根恆為正，第二個根恆為負，所以兩者永不相遇。導數具有介值性質（intermediate value property）（達布定理（Darboux's theorem）；如果你假設 y' 連續，則用普通的介值定理（intermediate value theorem）即可），所以 y' 不可能由正值跳到負值而不經過 0，但兩個根都不會取這個值。因此 y' 自始至終都停留在同一個根上。

在 $x = 0$ 處，兩個根是 1 和 -1，而我們需要 $y' > 0$，所以 $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{-x}$。積分得 $y = -\mathrm{e}^{-x} + c$，而由 $x = 0$ 時 $y = 0$ 可得 $c = 1$：

$$\boxed{y = 1 - \mathrm{e}^{-x}}$$

第二條微分方程

設 $p = \frac{\mathrm{d}y}{\mathrm{d}x}$，則 $p^2 \sinh x + 2p - \sinh x = 0$。

當 $x = 0$ 時，$\sinh 0 = 0$，方程簡化為 $2p = 0$，所以必定有 $y'(0) = 0$。

當 $x \neq 0$ 時，$\sinh x \neq 0$，由二次公式得

$$p = \frac{-2 \pm \sqrt{4 + 4\sinh^2 x}}{2\sinh x} = \frac{-1 \pm \cosh x}{\sinh x}.$$

利用 $\cosh x = 1 + 2\sinh^2 \frac{x}{2} = 2\cosh^2 \frac{x}{2} - 1$ 和 $\sinh x = 2\sinh \frac{x}{2} \cosh \frac{x}{2}$，兩個根可化簡為

$$\frac{\cosh x - 1}{\sinh x} = \frac{2\sinh^2 \frac{x}{2}}{2\sinh \frac{x}{2} \cosh \frac{x}{2}} = \tanh \frac{x}{2},$$
$$\frac{-1 - \cosh x}{\sinh x} = -\frac{2\cosh^2 \frac{x}{2}}{2\sinh \frac{x}{2} \cosh \frac{x}{2}} = -\coth \frac{x}{2}.$$

選擇分支。第二個根 $-\coth \frac{x}{2}$ 當 $x \to 0$ 時無界，把它積分得 $y = -2\ln\left|\sinh \frac{x}{2}\right| + c$，而無論從哪一側取 $x \to 0$，這個函數都趨向 $+\infty$。所以在原點的任何一側，這個分支都不能給出在 $x = 0$ 處連續且滿足 $y(0) = 0$ 的解。因此在 $x = 0$ 附近（每一側）我們都必須取第一個根，而由於 $\tanh 0 = 0$，這與 $y'(0) = 0$ 吻合。兩個根永不重合（一個與 x 同號，另一個異號，而且當 $x \neq 0$ 時兩者都不為零），所以根據與之前相同的介值論證，對所有 x，解都停留在這個分支上：

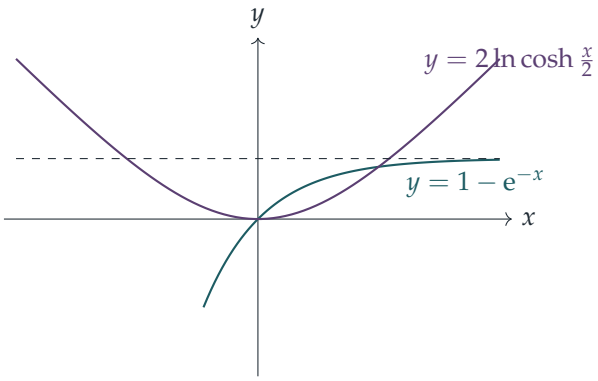
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \tanh \frac{x}{2} \quad \text{對所有} x.$$

由於 $\frac{\mathrm{d}}{\mathrm{d}x} \cosh \frac{x}{2} = \frac{1}{2} \sinh \frac{x}{2}$，我們有 $\tanh \frac{x}{2} = 2 \cdot \frac{\frac{1}{2} \sinh \frac{x}{2}}{\cosh \frac{x}{2}}$，這屬於 $2f'/f$ 的形式，所以積分得 $y = 2\ln \cosh \frac{x}{2} + c$（因為 $\cosh > 0$，無須加上絕對值符號）。由條件 $y(0) = 0$ 得

$c = 0$，所以

$$\boxed{y = 2\ln\cosh\tfrac{x}{2} = \ln\left(\frac{1+\cosh x}{2}\right)}$$

其中第二個形式由 $2\cosh^2\frac{x}{2} = 1 + \cosh x$ 得出。



驗算

- 對於 $y = 1 - e^{-x}$：$y'^2 + 2y'\sinh x = e^{-2x} + e^{-x}(e^x - e^{-x}) = 1$，符合要求，而且 $y'(0) = 1 > 0$。
- 對於 $y = 2\ln\cosh\frac{x}{2}$，設 $t = \tanh\frac{x}{2}$，則 $\sinh x = \frac{2t}{1-t^2}$；於是 $\sinh x\,(t^2-1) + 2t = -2t + 2t = 0$。這個解是偶函數，這是理所當然的，因為方程在 $x \mapsto -x$、$y \mapsto y$ 之下保持不變。
- 數值上，在 $x = 1$ 處兩個解分別是 $1 - e^{-1} \approx 0.6321$ 和 $2\ln\cosh\frac{1}{2} \approx 0.2402$；在許多隨機選取的點代入有限差分導數，兩條方程都成立，誤差在 10^{-6} 以內。

常見錯誤

- 把第二條方程除以 $\sinh x$，卻沒有留意它在 $x = 0$ 處為零，而那裏正是初始條件（initial condition）所在之處；原方程迫使 $y'(0) = 0$。
- 在第二部分保留兩個分支：$-\coth\frac{x}{2}$ 這個分支不可能滿足 $y(0) = 0$，因為它的積分在原點處趨向無窮。

STEP 之外

對 y' 並非線性的方程，例如 $F(x,y,y') = 0$，可以有多個解通過同一點，或者出現沒有任何單一分支能夠描述的「奇異」解（singular solution），所以存在性與唯一性定理（existence and uniqueness theorems，即 Picard–Lindelöf 定理）都是針對已經解成 $y' = f(x,y)$ 形式的方程而敘述的。在這裏兩個根互相分隔，所以每個分支都表現良好；當判別式（discriminant）可以為零時，各分支可以黏合起來，唯一性便不再成立。Boyce 與 DiPrima 的 *Elementary Differential Equations*（基礎微分方程教科書）討論了存在性、唯一性以及這些非線性一階方程。

詳細解答

STEP II 1991，第 9 題

純數學 • 矩陣群

> **STEP II 1991，第 9 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q9

關鍵思路

一切都建基於一個事實：模 5 的整數構成一個域（field）。每個非零剩餘（residue）都有乘法逆元（multiplicative inverse），而非零剩餘之積永不為零。正因如此，對角線上的元素才會規規矩矩，封閉性和逆元幾乎是不費工夫就得到的。至於最後一部分，可解 $M^2 = I$，把階（order）為 1 或 2 的元素明確地找出來，然後檢驗其中兩個相乘之後是否仍留在該集合之內。

解答

以下所有運算均以模 5 進行，並在方便時把 4 寫成 -1。我們把一般的元素寫成

$$M(a,b,c) = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}, \qquad a,c \in \{1,2,3,4\},\ b \in \{0,1,2,3,4\}.$$

G 是一個羣

封閉性（closure）。把兩個元素相乘，得

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\begin{pmatrix} d & e \\ 0 & f \end{pmatrix} = \begin{pmatrix} ad & ae+bf \\ 0 & cf \end{pmatrix}.$$

這是一個上三角矩陣（upper triangular matrix），各元素均為模 5 的數。它的對角線元素 ad 和 cf 都不是零：由於 5 是質數，若 $5 \mid ad$，則 $5 \mid a$ 或 $5 \mid d$，但兩者皆不成立。所以這個積屬於 G。

結合律（associativity）是已知的，因為矩陣乘法符合結合律。

單位元（identity）。$I = M(1,0,1)$ 屬於 G，而且對每個 M 都有 $IM = MI = M$。

逆元（inverses）。每個模 5 的非零剩餘都有逆元（$1^{-1}=1$，$2^{-1}=3$，$3^{-1}=2$，$4^{-1}=4$）。所以，給定 $M(a,b,c)\in G$，我們可以構造

$$N=\begin{pmatrix} a^{-1} & -a^{-1}bc^{-1} \\ 0 & c^{-1} \end{pmatrix},$$

由於 $a^{-1},c^{-1}\neq 0$，它屬於 G。利用上面的乘積公式，

$$MN=\begin{pmatrix} aa^{-1} & -aa^{-1}bc^{-1}+bc^{-1} \\ 0 & cc^{-1} \end{pmatrix}=I, \qquad NM=\begin{pmatrix} a^{-1}a & a^{-1}b-a^{-1}bc^{-1}c \\ 0 & c^{-1}c \end{pmatrix}=I.$$

所以每個元素在 G 中都有逆元，因此 G 是一個羣（group）。

G 的階

a 有 4 個選擇，b 有 5 個，c 有 4 個，而不同的選擇給出不同的矩陣。所以

$$|G|=4\times 5\times 4=80.$$

G 是否可交換？

不是。取 $A=M(1,1,1)$ 和 $B=M(2,0,1)$。則

$$AB=\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}2&0\\0&1\end{pmatrix}=\begin{pmatrix}2&1\\0&1\end{pmatrix}, \qquad BA=\begin{pmatrix}2&0\\0&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}2&2\\0&1\end{pmatrix},$$

所以 $AB\neq BA$，G 不是可交換（commutative）的。

階為 1 或 2 的元素

元素 M 的階為 1 或 2，當且僅當 $M^2=I$。由乘積公式，

$$M(a,b,c)^2=\begin{pmatrix} a^2 & b(a+c) \\ 0 & c^2 \end{pmatrix},$$

所以我們需要 $a^2=1$、$c^2=1$ 以及 $b(a+c)=0$。

$1,2,3,4$ 在模 5 下的平方分別是 $1,4,4,1$，所以 $a^2=1$ 必然得出 $a=\pm 1$，同理 $c=\pm 1$。

- 若 $a=c$，則 $a+c=\pm 2\neq 0$。由於 $\mathbb{Z}_5$ 沒有零因子（zero divisor），$b(a+c)=0$ 必然得出 $b=0$。這給出 I 和 $-I$。
- 若 $c=-a$，則 $a+c=0$，所以 b 可取任何值。這給出 10 個矩陣 $M(1,b,-1)$ 和 $M(-1,b,1)$，其中 $b=0,1,\ldots,4$。

所以，由階為 1 或 2 的元素組成的集合 H 恰好有 12 個成員。

想一想 由自逆元素組成的集合，是否必定封閉？

不是。這個集合包含 I，而且每個元素都是自己的逆元，所以單位元和逆元這兩條公理自動成立。唯一可能不成立的是封閉性，而在非交換羣中它通常都不成立：若 $X^2 = Y^2 = I$，則 $(XY)^2 = XYXY$，除非 X 與 Y 可交換，否則這未必能化簡為 I。

在這裏它確實不成立。取 $X = M(1,1,-1)$ 和 $Y = M(1,0,-1)$，兩者都屬於 H。則

$$XY = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

這個矩陣有 $a = c = 1$ 但 $b = -1 \neq 0$，所以按上面的分類，它不屬於 H。（事實上 $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix}$，所以它的階是 5。）因此 H 並不封閉，它不是 G 的子羣（subgroup）。

如果你認識拉格朗日定理（Lagrange's theorem），還有一個更快捷的論證：子羣的階必須整除羣的階，而 $12 \nmid 80$。

驗算

用電腦把全部 80 個矩陣簡單枚舉一遍，可確認封閉性、逆元公式，以及上述那對矩陣的 $AB \neq BA$。把每個元素平方，找到的正好是預測中 $M^2 = I$ 的 12 個解，而 $XY = M(1,4,1)$ 的階是 5。用手算可驗證 $M(-1,3,1)^2 = I$：非對角線元素是 $3(-1+1) = 0$。

常見錯誤

- 忘記證明逆元屬於 G。它的對角線元素 a^{-1}、c^{-1} 必須不是零。
- 以為 $M^2 = I$ 必然令 M 是對角矩陣，因而漏掉 $b \neq 0$ 的元素 $M(1,b,-1)$ 和 $M(-1,b,1)$。
- 只為 H 檢驗單位元和逆元這兩條公理。不成立的正是封閉性這條公理。

STEP 之外

G 是有限域（finite field）$\mathbb{F}_5$ 上所有可逆 2×2 上三角矩陣組成的羣，是 $\mathrm{GL}_2(\mathbb{F}_5)$ 的一個博雷爾子羣（*Borel subgroup*）。滿足 $a = c = 1$ 的矩陣構成一個階為 5 的正規子羣（normal subgroup），而 G 是由它與對角矩陣以半直積（semidirect product）的方式構成的。有限域上的矩陣羣，是羣論入門課程中豐富的例子來源。想了解更多，可參閱 Allenby, *Rings, Fields and Groups*（環、域與羣）。

詳細解答

STEP II 1991，第 10 題

純數學 • 三維空間中的直線與平面

STEP II 1991，第 10 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s2-q10

關鍵思路

雖然題庫把這題歸入另一類別，但它其實是一道關於三維空間中直線與平面的題目。影子的尖端，就是剛好掠過棒頂的那道太陽光線與地面相交之處。因此，我們把這道光線寫成一條通過太陽與棒頂的直線，把地面寫成一個通過棒底並與棒垂直的平面，然後求兩者的交點。只要把所有位置都用沿棒方向的單位向量（unit vector）表示，一切便變得容易。

解答

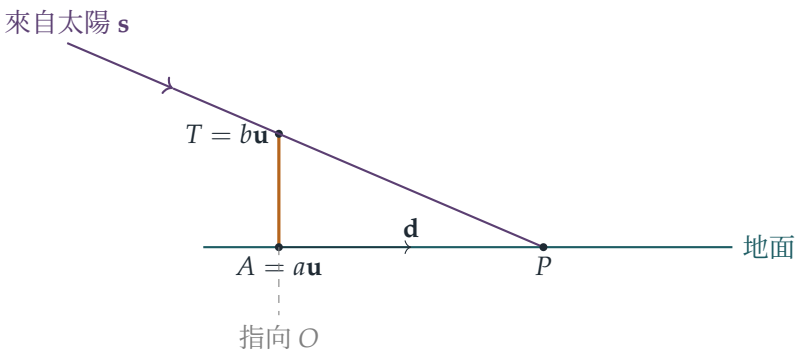
設 O 為地球的中心，並記

$$\mathbf{u} = \begin{pmatrix} 0 \\ \cos\phi \\ \sin\phi \end{pmatrix}, \qquad \mathbf{s} = R\begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}.$$

那麼 $\mathbf{u}$ 是沿棒方向的單位向量，棒底為 $A = a\mathbf{u}$，棒頂為 $T = b\mathbf{u}$，而太陽位於 $\mathbf{s}$。我們唯一需要的點積（dot product）是

$$\mathbf{s}\cdot\mathbf{u} = R\sin\theta\cos\phi.$$

為簡便起見，記 $c = R\cos\phi\sin\theta$。



通過棒頂的光線

到達影子尖端 P 的光線，剛剛經過棒頂。所以 P 位於通過 $\mathbf{s}$ 與 $b\mathbf{u}$ 的直線上：

$$\mathbf{p} = \mathbf{s} + t\,(b\mathbf{u} - \mathbf{s}) = (1-t)\,\mathbf{s} + tb\,\mathbf{u}$$

其中 t 為某個實數。

地面

地面是水平的，即與棒垂直，並且通過棒底 $a\mathbf{u}$。由於 $\mathbf{u}\cdot\mathbf{u} = 1$，地面就是平面 $\mathbf{r}\cdot\mathbf{u} = a$。

求交點

把直線代入平面的方程，並利用 $\mathbf{u}\cdot\mathbf{u} = 1$ 及 $\mathbf{s}\cdot\mathbf{u} = c$，得

$$(1-t)c + tb = a \qquad \Longrightarrow \qquad t = \frac{c-a}{c-b}.$$

我們可以除以 $c-b$，因為太陽正照着這根棒。在棒頂處，太陽位於地平線之上，當且僅當它位於該處切平面（tangent plane）的另一側，即 $\mathbf{s}\cdot\mathbf{u} > b$，所以 $c > b$。

? **想一想** 這個 P 真的在棒後方的地面上，而不是在太陽背後的某處嗎？

由於 $c > b > a$，我們有 $c-a > c-b > 0$，所以 $t > 1$。$t = 0$ 的點是太陽，$t = 1$ 的點是棒頂，因此 $t > 1$ 表示已越過棒頂、背向太陽而行，正如所料。

由 t 的值可得

$$1 - t = \frac{(c-b)-(c-a)}{c-b} = \frac{a-b}{c-b} = -\frac{h}{c-b},$$
$$tb - a = \frac{b(c-a)-a(c-b)}{c-b} = \frac{c(b-a)}{c-b} = \frac{ch}{c-b}.$$

所以由棒底到影子尖端的位移（displacement）為

$$\begin{aligned}\mathbf{d} = \mathbf{p} - a\mathbf{u} &= (1-t)\,\mathbf{s} + (tb - a)\,\mathbf{u} \\ &= \frac{h}{c-b}\left(c\,\mathbf{u} - \mathbf{s}\right) \\ &= \frac{Rh}{c-b}\left[\cos\phi\sin\theta\begin{pmatrix}0\\ \cos\phi\\ \sin\phi\end{pmatrix} - \begin{pmatrix}\cos\theta\\ \sin\theta\\ 0\end{pmatrix}\right] \\ &= \frac{Rh}{c-b}\begin{pmatrix}-\cos\theta\\ -(1-\cos^2\phi)\sin\theta\\ \cos\phi\sin\phi\sin\theta\end{pmatrix}.\end{aligned}$$

由於 $1-\cos^2\phi = \sin^2\phi$ 及 $c - b = R\cos\phi\sin\theta - b$，上式即

$$\mathbf{d} = Rh\,(R\cos\phi\sin\theta - b)^{-1}\begin{pmatrix}-\cos\theta\\ -\sin^2\phi\sin\theta\\ \cos\phi\sin\phi\sin\theta\end{pmatrix},$$

這正是所要求的結果。

中間那一行 $\mathbf{d} = \frac{h}{c-b}(c\,\mathbf{u} - \mathbf{s})$ 值得留意。向量 $c\,\mathbf{u}$ 是 $\mathbf{s}$ 沿棒方向的分量（component），所以 $c\,\mathbf{u} - \mathbf{s}$ 就是太陽位置向量（position vector）的水平部分的負值。影子正正指向與太陽的水平方向相反的一方，與日常經驗吻合。

驗算

- $\mathbf{d}\cdot\mathbf{u} = \frac{Rh}{c-b}\left(-\sin^2\phi\sin\theta\cos\phi + \cos\phi\sin^2\phi\sin\theta\right) = 0$，所以影子確實貼着地面。
- 取 $\phi = 0$，則棒沿 y 軸方向。這時 $\mathbf{d} = \frac{Rh}{R\sin\theta - b}(-\cos\theta, 0, 0)$。當 $\theta = \frac{\pi}{2}$ 時，太陽在正上方，影子消失。當 R 極大時，影長趨向 $h\cot\theta$。這正是大家熟悉的結果：把地面視為平面、太陽仰角為 θ 時的影長。
- 數值方面，我們隨機選取了數千組 R, a, h, θ, ϕ 的值（太陽均在地平線之上），直接以光線與地平面相交的方法求出影子尖端的位置。結果每次都與公式吻合，最大相對誤差約為 10^{-13}。

常見錯誤

- 把地面當作平面 $y = a\cos\phi$，或其他坐標平面。「水平」的意思是與 $\mathbf{u}$ 垂直，即 $\mathbf{r}\cdot\mathbf{u} = a$。

► 一開始就把所有東西按分量展開。把 **u** 和 **s** 保留為向量，直到最後一行才展開，可以省卻大量代數運算。

STEP 之外

在這裏，太陽光是通過同一點（太陽）的一族直線，所以影子是棒在地平面上的中心投影（central projection）。當太陽非常遙遠時，光線幾乎平行，這個投影便變成平行投影（parallel projection）。這正是線性代數（linear algebra）所研究的那類映射，而向量 $c\,\mathbf{u} - \mathbf{s}$ 就是 **s** 在棒上的投影減去 **s** 本身。Strang 的 *Introduction to Linear Algebra*（《線性代數導論》）詳細討論了在直線和平面上的投影。

詳細解答

STEP II 1991，第 11 題

力學 • 拋體與發射點的距離

STEP II 1991，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s2-q11

關鍵思路

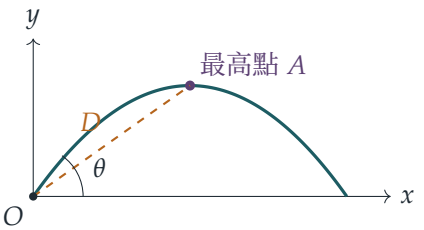
處理砲彈與大砲之間距離的平方：這樣可以避開平方根，並把兩部分都化為多項式（polynomial）問題。在第一部分，最高點距離的平方是 $\sin^2\theta$ 的二次式，所以用配方法（completing the square）（或求導）便可找出最佳角度。在第二部分，砲彈逐漸接近，當且僅當距離的平方正在遞減，因此我們要考察它對時間的導數（derivative）在整個飛行過程中的正負。

解答

模型

我們把砲彈視為只受重力作用而自由運動的質點（particle），不計空氣阻力，從原點 O（砲手所站之處）以速率 u、仰角（angle of elevation）θ 發射，其中 $0 < \theta < \frac{\pi}{2}$。取 x 為水平方向，y 為鉛垂向上。在時間 t，

$$x = ut\cos\theta, \qquad y = ut\sin\theta - \tfrac{1}{2}gt^2.$$



仰角

當 $\dot{y} = u\sin\theta - gt = 0$，即 $t = u\sin\theta/g$ 時，砲彈到達最高點（apex）。此時

$$x_A = \frac{u^2\sin\theta\cos\theta}{g}, \qquad y_A = \frac{u^2\sin^2\theta}{2g}.$$

以 D 表示距離 OA，並設 $s = \sin^2\theta$，於是 $\cos^2\theta = 1 - s$。那麼

$$\begin{aligned} D^2 = x_A^2 + y_A^2 &= \frac{u^4}{g^2}\left(\sin^2\theta\cos^2\theta + \tfrac{1}{4}\sin^4\theta\right) \\ &= \frac{u^4}{g^2}\left(s(1-s) + \tfrac{1}{4}s^2\right) = \frac{u^4}{g^2}\left(s - \tfrac{3}{4}s^2\right). \end{aligned}$$

由於損害隨距離增加而減少，砲手希望 D 愈大愈好，這等同於令 D^2 愈大愈好。配方得

$$s - \tfrac{3}{4}s^2 = \tfrac{1}{3} - \tfrac{3}{4}\left(s - \tfrac{2}{3}\right)^2,$$

它恰好在 $s = \frac{2}{3}$ 時取得最大值 $\frac{1}{3}$。這個值是可以取到的，因為 $s = \sin^2\theta$ 可取 $(0,1)$ 內的每一個值。所以 $\sin^2\theta = \frac{2}{3}$，$\cos^2\theta = \frac{1}{3}$，從而

$$\boxed{\tan\theta = \sqrt{2}, \qquad \theta = \arctan\sqrt{2} \approx 54.7^\circ}$$

此時最高點與大砲的距離為 $D = \dfrac{u^2}{\sqrt{3}\,g}$。

砲彈在下降期間會否更接近砲手？

設 $r^2 = x^2 + y^2$ 為時間 t 時砲彈與大砲距離的平方。砲彈逐漸接近，當且僅當 r^2 正在遞減，所以我們需要知道

$$\tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(r^2) = x\dot{x} + y\dot{y}.$$

的正負。

> **想一想** 為甚麼不直接把最高點的距離與之後幾個點的距離比較一下？

只看幾個點，並不能排除距離在它們之間曾經下跌的可能。導數卻能一次過告訴我們飛行中每一刻的情況。由 $\dot{x} = u\cos\theta$ 及 $\dot{y} = u\sin\theta - gt$，

$$\begin{aligned} x\dot{x} + y\dot{y} &= u^2 t\cos^2\theta + \left(ut\sin\theta - \tfrac{1}{2}gt^2\right)(u\sin\theta - gt) \\ &= u^2 t - \tfrac{3}{2}ugt^2\sin\theta + \tfrac{1}{2}g^2t^3 \\ &= \tfrac{1}{2}t\left[(gt)^2 - 3u\sin\theta\,(gt) + 2u^2\right], \end{aligned}$$

其中用到 $\cos^2\theta + \sin^2\theta = 1$。當 $t > 0$ 時，其正負與方括號內的式子相同，而該式是 gt 的二次式。配方得

$$(gt)^2 - 3u\sin\theta\,(gt) + 2u^2 = \left(gt - \tfrac{3}{2}u\sin\theta\right)^2 + u^2\left(2 - \tfrac{9}{4}\sin^2\theta\right).$$

在所選的角度下 $\sin^2\theta = \frac{2}{3}$，所以常數項為 $u^2\left(2 - \frac{3}{2}\right) = \frac{1}{2}u^2 > 0$。因此方括號內的式子對每個 t 都嚴格為正，故 r^2 在整個飛行過程中嚴格遞增（strictly increasing）。

所以答案是**不會**：砲彈在整個下降過程中一直遠離砲手，而下降路徑上最接近的一點就是最高點本身。這說明砲手只考慮砲彈在最大高度爆炸的捷徑是合理的：對這個角度而言，最高點確實是下降路徑上最接近的一點，距離為 $u^2/(\sqrt{3}\,g)$。

驗算

- 當 $\theta = \frac{\pi}{4}$ 時，最高點位於 $(u^2/2g,\, u^2/4g)$，所以 $D^2 = \frac{5}{16}u^4/g^2$，比我們的最大值 $\frac{1}{3}u^4/g^2$ 略小，正如所料。
- 一般情況的計算顯示，只有當 $\frac{9}{4}\sin^2\theta \geq 2$，即 $\sin\theta \geq \frac{2\sqrt{2}}{3}$ 時，砲彈才開始會有接近的情況，對應的角度約為 70.5°，遠高於 54.7°。（極限 $\theta \to \frac{\pi}{2}$ 的情況合乎常理：垂直向上發射的砲彈會筆直地朝大砲落回來。）
- 數值上，取 $u = g = 1$，對 θ 搜尋，可得最高點距離的最大值 $0.57735 \approx 1/\sqrt{3}$，出現在 $\theta \approx 54.736°$；在該角度下沿軌跡（trajectory）逐步計算，可見與 O 的距離從不減少；而在 $70.5° \pm 0.6°$ 之間，距離由單調（monotonic）變為不單調。

常見錯誤

- 把最高點的水平距離或高度最大化，而不是把它與大砲的直線距離最大化。
- 以為砲彈在水平方向不斷遠離，所以過了最高點後與大砲的距離顯然會增加。其實砲彈同時也在下墜，而當發射角度夠陡時，它確實會回頭接近大砲。

STEP 之外

對 r^2 而非 r 求導，再看 $\mathbf{r}\cdot\dot{\mathbf{r}}$ 的正負，是判斷動點正在接近還是遠離的標準方法，在向量力學（vector mechanics）以至軌道（orbits）的研究中隨處可見。臨界角 $\arcsin(2\sqrt{2}/3)$ 是一個經典結果：仰角低於此值時，拋體（projectile）與發射點的距離永遠遞增。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）闡述了以向量處理運動的方法，令這類論證顯得自然而然。

詳細解答

STEP II 1991，第 12 題

力學 • 受迫彈性振動

> **STEP II 1991，第 12 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q12

本題在 *Beyond STEP* 第 11.1 節共振中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch11/resonance.html

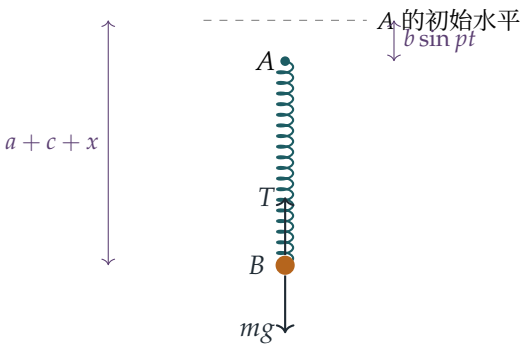
關鍵思路

張力只取決於伸長量（extension），即質點的位置減去移動端的位置，再減去自然長度（natural length）。把它用 x 和 $b\sin pt$ 表示出來，再利用初始平衡狀態把彈性模量（modulus）換成含 c 的式子，運動方程（equation of motion）便自然得出。接下來要處理的是一個受迫簡諧振子（forced simple harmonic oscillator）。求解時把特解（particular solution）加到餘函數（complementary function）上，並謹記「拉緊」的意思是「伸長量永不為負」。

解答

運動方程

把質點視為質量為 m 的一點，繫於一條遵從胡克定律（Hooke's law）、彈性模量為 λ 的輕彈性繩上。所有位置均由 A 的初始位置起向下量度。



在初始平衡狀態下，伸長量為 c，而張力與重量平衡：

$$\frac{\lambda c}{a} = mg, \qquad \text{所以} \qquad \frac{\lambda}{ma} = \frac{g}{c} = n^2.$$

在時間 t，質點位於深度 $a + c + x$，而端點 A 位於深度 $b\sin pt$。只要繩子拉緊，其長度為 $a + c + x - b\sin pt$，所以伸長量為

$$e = c + x - b\sin pt,$$

而張力為 $T = \lambda e/a$。對質點沿向下方向分解力，

$$\begin{aligned} m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} &= mg - \frac{\lambda}{a}(c + x - b\sin pt) \\ &= mg - \frac{\lambda c}{a} - \frac{\lambda}{a}(x - b\sin pt) \\ &= -\frac{\lambda}{a}(x - b\sin pt), \end{aligned}$$

其中用了 $\lambda c/a = mg$。兩邊除以 m，並利用 $\lambda/(ma) = n^2$，得

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -n^2(x - b\sin pt),$$

正是所求。這裏需要 $e \geq 0$：假如繩子鬆弛，張力便會是零，而不是 $\lambda e/a$。

求解 $x(t)$

把方程改寫成 $\ddot{x} + n^2x = n^2b\sin pt$。餘函數是 $C\cos nt + D\sin nt$。至於特積分（particular integral），試用 $x = K\sin pt$：

$$-p^2K + n^2K = n^2b \quad \Longrightarrow \quad K = \frac{n^2b}{n^2 - p^2},$$

這是成立的，因為 $p < n$ 保證 $n^2 - p^2 \neq 0$。所以

$$x = \frac{n^2b}{n^2 - p^2}\sin pt + C\cos nt + D\sin nt.$$

在 $t = 0$ 時，質點位於平衡位置且處於靜止，所以 $x(0) = 0$ 及 $\dot{x}(0) = 0$。由前者得 $C = 0$。由後者得

$$\frac{n^2bp}{n^2 - p^2} + nD = 0 \quad \Longrightarrow \quad D = -\frac{nbp}{n^2 - p^2}.$$

因此

$$x(t) = \frac{bn}{n^2 - p^2}(n \sin pt - p \sin nt),$$

正是所求。

繩子何時保持拉緊

想一想 究竟是甚麼必須保持為正：是 x、張力，還是別的東西？

是伸長量 $e = c + x - b \sin pt$。$x(t)$ 的公式只在 $e \geq 0$ 時才描述這個運動，所以繩子自始至終保持拉緊，當且僅當這個不等式對所有 $t \geq 0$ 都成立。代入 x，

$$\begin{aligned} x - b \sin pt &= b \sin pt \left(\frac{n^2}{n^2 - p^2} - 1 \right) - \frac{bnp}{n^2 - p^2} \sin nt \\ &= \frac{bp}{n^2 - p^2}(p \sin pt - n \sin nt). \end{aligned}$$

所以充分必要條件（necessary and sufficient condition）是

$$c + \frac{bp}{n^2 - p^2}(p \sin pt - n \sin nt) \geq 0 \quad \text{對所有} t \geq 0,$$

又因為 $n^2 - p^2 > 0$，這等價於

$$bp(n \sin nt - p \sin pt) \leq c(n^2 - p^2) \quad \text{對所有} t \geq 0.$$

（如果 e 只在個別孤立的時刻觸及 0，張力在那一瞬間為零，但運動方程仍然成立，繩子亦從未真正鬆弛。這就是不等式不必嚴格的原因。）

現在假設 $pb < (n - p)c$。對任何 t，由於 $\sin nt \leq 1$ 和 $-\sin pt \leq 1$，且 $n, p > 0$，

$$n \sin nt - p \sin pt \leq n + p.$$

取 $b > 0$（即振幅），便有 $bp > 0$，所以

$$bp(n \sin nt - p \sin pt) \leq bp(n + p) < (n - p)c(n + p) = c(n^2 - p^2).$$

條件成立，而且其實是嚴格不等式，所以繩子自始至終保持拉緊。（假如容許 b 為負，只要以 $|b|$ 代替 b，同樣的論證仍然適用，因為 $|n \sin nt - p \sin pt| \leq n + p$。）

驗算

- 取 $t=0$：$x(0)=0$ 及 $\dot{x}(0)=\frac{bn}{n^2-p^2}(np-pn)=0$，與應有的結果相符。
- 當 $p\to 0$ 時，端點 A 幾乎不動，而 $x(t)\approx \frac{bn}{n^2}(npt-pnt)\to 0$。一如所料，甚麼也沒有發生。
- 最小伸長量不可能低於 $c-bp(n+p)/(n^2-p^2)=c-bp/(n-p)$，而此值為正，恰好當 $pb<(n-p)c$。
- 數值上，以隨機選取的 m,λ,a,c,b,p 對完整的胡克定律方程作逐步（RK4）積分，結果與 $x(t)$ 的公式相符，誤差在 4×10^{-10} 之內。取 $n=1$、$p=1/\sqrt{2}$、$c=1$，而 b 只比 $(n-p)c/p$ 高出 5%，則到 $t=20\,000$ 之前，伸長量確實會變成負數（約為 -0.05）。

常見錯誤

- 把 x 本身當作伸長量。上端是會移動的，所以伸長量是 $c+x-b\sin pt$。
- 忘記了 $\dot{x}(0)=0$ 會令解中必須出現 $\sin nt$ 項。沒有這一項，質點起初的速度便會是 $n^2bp/(n^2-p^2)$。
- 把拉緊的條件寫成 $x\geq -c$。只有當 A 固定不動時，這才是對的。

STEP 之外

這是一個無阻尼的受迫振子。它的響應由驅動頻率（driving frequency）p 和固有頻率（natural frequency）n 混合而成，而因子 $1/(n^2-p^2)$ 在 $p\to n$ 時趨向無窮大，這就是共振（resonance）。當 n/p 是無理數時，(nt,pt) 模 2π 的值會稠密地填滿整個範圍，所以 $n\sin nt-p\sin pt$ 可以任意接近 $n+p$。於是充分條件 $pb<(n-p)c$ 基本上也是必要條件（等號仍然容許）。關於受迫振動、共振和阻尼（damping），Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）有詳盡的論述；微分方程方面，則可參看 Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程）。

詳細解答

STEP II 1991，第 13 題

力學 • 剛體轉動

> **STEP II 1991，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q13

關鍵思路

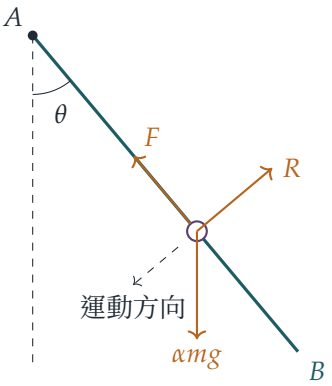
在小環滑動之前，桿和小環如同單一剛體（rigid body）一起繞 A 轉動。因此我們可以求出這個組合體的角速率（angular speed，由能量求得）和角加速度（angular acceleration，由力矩（moments）求得），然後用它們算出桿必須對小環施加甚麼力，才能令小環繼續隨桿運動。只有當小環所需的摩擦力（friction）不超過法向反作用力（normal reaction）的 μ 倍時，小環才會留在原位；這個條件首次不成立的角度，就是開始滑動的角度。

解答

模型

我們把小環視為與 A 相距 x 的質點，把支點視為光滑，並把桿視為剛體。設 θ 為桿與向下鉛垂線之間的夾角，所以釋放時 $\theta = \frac{\pi}{2}$，而桿向下擺動時 θ 遞減。以 $\omega = |\dot{\theta}|$ 表示角速率，以 $\beta = |\ddot{\theta}|$ 表示角加速度的大小。

桿對小環施加兩個力：一個是垂直於桿的法向反作用力 R，以小環運動的方向為正；另一個是沿桿的摩擦力 F，以指向 A 的方向為正。由於小環是套在桿上的，R 的正負號兩種情況都有可能出現。



圖中把 R 畫成與運動方向相反，因為當 $k^2 > xd$ 時結果正是如此，但在代數運算中我們保留一般的正負號。

剛體的運動

小環固定在桿上時，整體繞 A 的轉動慣量（moment of inertia）為 $mk^2 + \alpha mx^2 = m(k^2 + \alpha x^2)$。自釋放起，桿的質心（centre of mass）下降了 $d\cos\theta$，小環下降了 $x\cos\theta$，所以由能量守恆（conservation of energy）可得

$$\tfrac{1}{2}m(k^2 + \alpha x^2)\,\omega^2 = mg(d + \alpha x)\cos\theta \quad\Longrightarrow\quad \omega^2 = \frac{2g(d + \alpha x)\cos\theta}{k^2 + \alpha x^2}.$$

對 A 取力矩（支點的力沒有力矩），兩個重量的總力矩為 $mg(d + \alpha x)\sin\theta$，所以

$$\beta = \frac{g(d + \alpha x)\sin\theta}{k^2 + \alpha x^2}.$$

小環所受的力

小環沿半徑為 x 的圓運動。它的加速度有兩個分量：沿桿指向 A 的 $x\omega^2$，以及垂直於桿、沿運動方向的 $x\beta$。沿桿向外的方向與向下鉛垂線的夾角為 θ，所以重量沿桿向外的分量為 $\alpha mg\cos\theta$，沿運動方向的分量為 $\alpha mg\sin\theta$。

沿桿、指向 A 的方向：$F - \alpha mg\cos\theta = \alpha mx\omega^2$，所以

$$\begin{aligned} F &= \alpha mg\cos\theta\left(1 + \frac{2x(d + \alpha x)}{k^2 + \alpha x^2}\right) \\ &= \frac{\alpha mg\cos\theta\,(k^2 + 3\alpha x^2 + 2xd)}{k^2 + \alpha x^2}. \end{aligned}$$

垂直於桿、沿運動方向：

$$R + \alpha mg\sin\theta = \alpha mx\beta \quad\Longrightarrow\quad R = \alpha mg\sin\theta\left(\frac{x(d + \alpha x)}{k^2 + \alpha x^2} - 1\right) = \frac{\alpha mg\sin\theta\,(xd - k^2)}{k^2 + \alpha x^2}.$$

留意 αx^2 的項在 R 中互相抵消，這就是為甚麼只出現 $k^2 - xd$。

想一想 小環傾向朝哪個方向滑動？R 的正負號又代表甚麼？

當 $0 < \theta < \frac{\pi}{2}$ 時，我們有 $F > 0$：桿必須把小環拉向 A，所以若摩擦力不足，小環便會向外滑，即滑向 B。若 $k^2 > xd$，則 $R < 0$，即桿正在拖住小環。若只有小環自己，它會比桿上這一點下落得更快，所以桿要頂住它。

何時開始滑動？

只要 $|F| \le \mu|R|$，小環便會固定在桿上。當 $k^2 > xd$ 時，F 和 $|R|$ 都含有同一個正因子 $\alpha mg/(k^2 + \alpha x^2)$，所以條件是

$$\cos\theta\,(k^2 + 3\alpha x^2 + 2xd) \le \mu \sin\theta\,(k^2 - xd), \quad \text{即} \quad \mu\tan\theta \ge \frac{3\alpha x^2 + k^2 + 2xd}{k^2 - xd}.$$

釋放時，$\theta = \frac{\pi}{2}$ 而 $F = 0$，所以小環不會立即滑動。當 θ 由 $\frac{\pi}{2}$ 遞減至 0 時，$\mu\tan\theta$ 由 $+\infty$ 平穩地遞減至 0，而右邊是一個固定的正數。所以不等式首次不成立、小環開始向外滑動的時刻，就是在滿足下式的唯一角度：

$$\boxed{\mu\tan\theta = \frac{3\alpha x^2 + k^2 + 2xd}{k^2 - xd}}$$

這正是所要求的結果。小環必定在桿到達鉛垂位置之前滑動。

第 (i) 部分：$k^2 = xd$

現在在任何角度都有 $R = 0$，所以可提供的最大摩擦力 $\mu|R|$ 為零。但只要 $\theta < \frac{\pi}{2}$，便有 $F > 0$。所以不論 μ 的值是多少，桿一釋放，小環便開始沿桿向外滑動。從物理上看，釋放時桿上與 A 相距 $x = k^2/d$ 的那一點，開始下落時的加速度恰好是 g（因為在 $\theta = \frac{\pi}{2}$ 時 $x\beta = g$）。所以小環對桿的兩側都沒有壓力，摩擦力也就無從產生。

第 (ii) 部分：$k^2 < xd$

現在 $R > 0$：桿沿運動方向推動小環，因為桿上這一點在釋放時下落的加速度會大於 g。小環同樣傾向向外滑動，而條件 $F \le \mu R$ 變成

$$\mu\tan\theta \ge \frac{3\alpha x^2 + k^2 + 2xd}{xd - k^2}.$$

按照與之前相同的論證，小環起初保持不動，並在以下情況開始向外滑動：

$$\mu\tan\theta = \frac{3\alpha x^2 + k^2 + 2xd}{xd - k^2},$$

這就是把題目所給結果中的 $k^2 - xd$ 換成 $|k^2 - xd|$。唯一的分別是法向反作用力現在作用在小環的另一側。

驗算

- 當 $k^2 > xd$ 時，隨着 $\mu \to \infty$，滑動角度 $\theta \to 0$：非常粗糙的桿幾乎可以把小環一直留住到最低點。當 $k^2 - xd \to 0$ 時，該角度趨向 $\frac{\pi}{2}$，與第 (i) 部分吻合。
- 長度為 $2a$ 的均勻桿有 $d = a$ 及 $k^2 = \frac{4}{3}a^2$，所以臨界位置 $x = k^2/d = \frac{4}{3}a$ 就是撞擊中心（centre of percussion），即桿釋放時以自由落體方式開始運動的那一點。
- 在數值方面，我們對剛體運動進行積分，用有限差分算出小環的加速度，並就隨機選取的參數值重新求得 F 和 R，誤差在 10^{-5} 之內。例如，當 $k^2 - xd \approx 2.87$ 時，小環首次需要多於 $\mu|R|$ 的摩擦力是在 $\theta \approx 1.3270$，與公式一致。$k^2 < xd$ 的情況亦以同樣方式與 (ii) 的公式吻合。

常見錯誤

- 忘記小環本身對轉動慣量和對重量力矩的貢獻。在滑動之前，小環是剛體的一部分。
- 使用摩擦定律時，以重量的法向分量代替實際的法向反作用力 R，而這裏的法向反作用力還取決於角加速度。
- 在沿桿方向的方程中漏掉向心（centripetal）項 $x\omega^2$，而 $3\alpha x^2$ 中的因子 3 正是由此而來。

STEP 之外

與支點相距 k^2/d 的那一點就是撞擊中心。在該處擊打，支點不會產生衝擊反作用力，這就是板球棒或網球拍有「甜點」的原因。它同時也是與該複擺（compound pendulum）週期相同的單擺的長度。這裏所用的方法，是先求出剛體的運動，再問需要甚麼內力來維持這種運動，這正是拉格朗日力學（Lagrangian mechanics）中分析約束力（constraint forces）的基礎。想進一步了解剛體轉動和複擺，可閱讀 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）。

詳細解答

STEP II 1991，第 14 題

力學 • 相對速度與追蹤

STEP II 1991，第 14 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

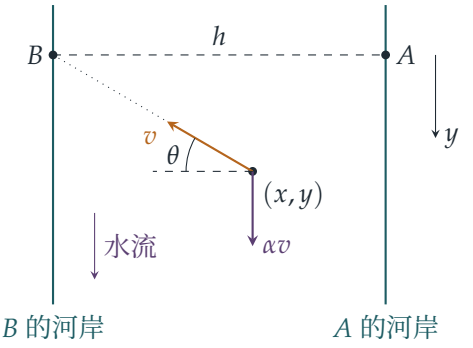
step.maths.org/questions/91-s2-q14

關鍵思路

船相對於地面的速度，等於船相對於水的速度加上水流的速度。「船頭始終指向 B」是一個幾何條件，即 $y = x\tan\theta$；對它求導，便得到同一個速度的第二個表達式。把兩個表達式對照，就得出一條連繫 x 與 θ 的可分離變數（separable）方程。選用代換 $\tan\theta = \sinh(\alpha s)$，是因為它令 $\sec\theta + \tan\theta$ 變成一個指數函數，於是所有式子都化為 s 的簡單指數函數。

解答

我們把船當作質點（particle）來建立模型。取 B 為原點，x 由 B 所在的河岸量向 A，y 則沿下游方向量度，因此 A 位於 $(h, 0)$，而船由該處出發。



以 θ 表示速度

船頭指向 B 所在的河岸，與方向 AB 成 θ 角，並偏向上游（以抵消水流造成的漂移）。因此，以我們的 (x,y) 分量表示，船相對於水的速度（relative velocity）是沿單位向量 $-(\cos\theta,\sin\theta)$ 的 v。再加上水流，即 y 方向上的 αv，便得

$$\begin{pmatrix}\dot{x}\\ \dot{y}\end{pmatrix}=\begin{pmatrix}-v\cos\theta\\ \alpha v-v\sin\theta\end{pmatrix}.$$

指向 B

由位於 (x,y) 的船到 B 的方向是沿 $(-x,-y)$，而這個方向與 AB 成 θ 角，當且僅當 $\tan\theta=y/x$。所以在整個運動過程中，

$$y=x\tan\theta.$$

利用積法則（product rule）和鏈式法則（chain rule）對 t 求導，

$$\dot{y}=\tan\theta\,\frac{\mathrm{d}x}{\mathrm{d}t}+x\sec^2\theta\,\frac{\mathrm{d}\theta}{\mathrm{d}t},$$

所以速度向量 $(\dot{x},\dot{y})$ 具有題目要求的形式。

x 與 θ 的微分方程

現在把兩個表達式對照。由 x 分量得 $\dot{x}=-v\cos\theta$。把它代入 y 分量，

$$\begin{aligned}\tan\theta\,(-v\cos\theta)+x\sec^2\theta\,\dot{\theta}&=\alpha v-v\sin\theta\\ x\sec^2\theta\,\dot{\theta}&=\alpha v,\end{aligned}\qquad(*)$$

這是因為 $\tan\theta\cos\theta=\sin\theta$，兩邊的 $-v\sin\theta$ 項互相抵消。特別地，$\dot{\theta}>0$，所以 θ 由 0 起不斷增加，我們可以把 x 看成 θ 的函數。兩式相除，

$$\frac{\mathrm{d}x}{\mathrm{d}\theta}=\frac{\dot{x}}{\dot{\theta}}=\frac{-v\cos\theta\cdot x\sec^2\theta}{\alpha v}=-\frac{x\sec\theta}{\alpha},\qquad\text{所以}\qquad\alpha\frac{\mathrm{d}x}{\mathrm{d}\theta}=-x\sec\theta,$$

正是所求。

求解 x

這條方程可以分離變數（船到達之前 $x>0$）：

$$\alpha\int\frac{\mathrm{d}x}{x}=-\int\sec\theta\,\mathrm{d}\theta\quad\Longrightarrow\quad\alpha\ln x=-\ln(\sec\theta+\tan\theta)+C,$$

這裏用了標準積分 $\int\sec\theta\,\mathrm{d}\theta=\ln(\sec\theta+\tan\theta)$，它在 $-\frac{\pi}{2}<\theta<\frac{\pi}{2}$ 時成立。開始時船在 A，該處 $x=h$ 而 $\theta=0$，所以 $C=\alpha\ln h$。因此

$$\alpha\ln\frac{x}{h}=-\ln(\sec\theta+\tan\theta),\qquad\text{即}\qquad\left(\frac{x}{h}\right)^{\alpha}=(\sec\theta+\tan\theta)^{-1}.$$

代換 $\tan\theta = \sinh(\alpha s)$

> **想一想** 以 s 表示，$\sec\theta$ 是甚麼？

由於 $\sec^2\theta = 1+\tan^2\theta = 1+\sinh^2(\alpha s) = \cosh^2(\alpha s)$，而 $\sec\theta$ 和 $\cosh(\alpha s)$ 都是正數（因為 $0 \le \theta < \frac{\pi}{2}$），我們得到 $\sec\theta = \cosh(\alpha s)$。於是

$$\sec\theta + \tan\theta = \cosh(\alpha s) + \sinh(\alpha s) = \mathrm{e}^{\alpha s},$$

所以 $(x/h)^\alpha = \mathrm{e}^{-\alpha s}$，兩邊取 α 次方根，得 $x = h\mathrm{e}^{-s}$。

至於變率的方程，把 $\tan\theta = \sinh(\alpha s)$ 對 t 求導：

$$\sec^2\theta\,\dot\theta = \alpha\cosh(\alpha s)\,\dot s.$$

把這個結果和 $x = h\mathrm{e}^{-s}$ 代入 $(*)$，得 $h\mathrm{e}^{-s}\cdot\alpha\cosh(\alpha s)\,\dot s = \alpha v$，再除以 α，

$$h\mathrm{e}^{-s}\cosh(\alpha s)\frac{\mathrm{d}s}{\mathrm{d}t} = v,$$

正是所求。

渡河所需的時間

在 A 處，$\theta = 0$，所以 $s = 0$。由於 $\dot s > 0$，s 不斷增加，而當 $x = h\mathrm{e}^{-s} \to 0$，即 $s \to \infty$ 時，船到達 B。（船確實會到達 B：因為 $\alpha < 1$，所以 $y = x\tan\theta = h\mathrm{e}^{-s}\sinh(\alpha s) \to 0$。）分離變數後，渡河時間為

$$\begin{aligned} T &= \frac{h}{v}\int_0^\infty \mathrm{e}^{-s}\cosh(\alpha s)\,\mathrm{d}s = \frac{h}{2v}\int_0^\infty \left(\mathrm{e}^{-(1-\alpha)s} + \mathrm{e}^{-(1+\alpha)s}\right)\mathrm{d}s \\ &= \frac{h}{2v}\left(\frac{1}{1-\alpha} + \frac{1}{1+\alpha}\right) = \frac{h}{2v}\cdot\frac{2}{1-\alpha^2} = \frac{h}{v(1-\alpha^2)}, \end{aligned}$$

其中兩個積分都收斂（converge），因為 $1-\alpha > 0$。所以渡河所需的時間是 $hv^{-1}(1-\alpha^2)^{-1}$。

驗算

- 沒有水流時（$\alpha = 0$），船筆直駛過對岸，$T = h/v$，與預期相符。當 $\alpha \to 1$ 時，所需時間趨向無限大：船若不比水流快，在 B 附近船頭正對上游時，便永遠無法逆流前進。
- 我們以數值方法逐步計算這個運動（四階龍格—庫塔法（Runge–Kutta），每一步都把船頭對準 B）。取 $h = v = 1$ 和 $\alpha = \frac{1}{2}$，船在 $t \approx 1.33333$ 時到達，與 $1/(1-\frac{1}{4}) = \frac{4}{3}$ 吻合。沿着整條路徑，$(x/h)^\alpha(\sec\theta+\tan\theta)$ 一直等於 1，誤差約為 10^{-13}。

常見錯誤

- 弄錯船頭方向的 y 分量的正負號。船會被沖向下游，所以指向 B 就是要部分指向上游；當 $\theta > 0$ 時，這個分量是 $-v\sin\theta$。
- 忘記說明為何 $\sec\theta = +\cosh(\alpha s)$：這需要 θ 一直小於 $\frac{\pi}{2}$，而這是成立的，因為 θ 是由 B 到一個 $x > 0$ 的點的直線所成的角。

STEP 之外

這是一道經典的追逐曲線（pursuit curve）問題：一個物體不斷把方向對準目標，而它的路徑可由一條關於運動方向的微分方程求得。同樣的想法可以描述狗追趕正在走動的主人，或者導彈追蹤目標，並且自然地引出曳物線（tractrix）一類的曲線。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教材）很適合用來熟習相對速度，以及在運動參考系（moving frames）中建立運動方程。

詳細解答

STEP II 1991，第 15 題

概率與統計 • 期望值的線性性質

> **STEP II 1991，第 15 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q15

關鍵思路

把數對 (n_1, n_2) 看成 $m \times m$ 方格表中的一格，全部 m^2 格出現的機會均等。滿足 $|n_1 - n_2| = k$ 的格子位於與主對角線平行的兩條對角線上，所以很容易數出來。之後的一切都來自期望值的線性性質（*linearity of expectation*）：若干個差之和的期望值，等於各個期望值之和，即使相鄰的項共用同一個變量、並非獨立（independent）亦然。監考員（invigilator）所走的路程也是同一類的和，只是第一段路由一張固定的書桌出發，而不是由隨機的一張出發。

解答

$|n_1 - n_2|$ 的分佈

由於 n_1 與 n_2 獨立，且各自在 $\{1, \ldots, m\}$ 上均勻分佈（uniform），m^2 個有序數對 (n_1, n_2) 中每一個的概率都是 $1/m^2$。

固定 k，其中 $1 \le k \le m-1$。滿足 $n_1 - n_2 = k$ 的數對是 $(k+1, 1), (k+2, 2), \ldots, (m, m-k)$，共 $m-k$ 對。由對稱性，滿足 $n_2 - n_1 = k$ 的數對另有 $m-k$ 對，而因為 $k \neq 0$，這兩組數對並不重疊。因此

$$\mathrm{P}\big(|n_1 - n_2| = k\big) = \frac{2(m-k)}{m^2},$$

正如所求。（當 $k = 0$ 時有 m 對，所以 $\mathrm{P}(|n_1 - n_2| = 0) = 1/m$。作為驗算，$\frac{1}{m} + \sum_{k=1}^{m-1} \frac{2(m-k)}{m^2} = \frac{1}{m} + \frac{m(m-1)}{m^2} = 1$。）

$|n_1 - n_2|$ 的期望值

$k = 0$ 這個值對期望值沒有貢獻，所以

$$\begin{aligned}\mathrm{E}|n_1 - n_2| &= \sum_{k=1}^{m-1} k \cdot \frac{2(m-k)}{m^2} = \frac{2}{m^2}\left(m\sum_{k=1}^{m-1} k - \sum_{k=1}^{m-1} k^2\right)\\ &= \frac{2}{m^2}\left(\frac{m\cdot m(m-1)}{2} - \frac{(m-1)m(2m-1)}{6}\right)\\ &= \frac{2}{m^2}\cdot\frac{m(m-1)\big(3m-(2m-1)\big)}{6} = \frac{(m-1)(m+1)}{3m} = \frac{m^2-1}{3m},\end{aligned}$$

當中用了標準求和公式 $\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$ 和 $\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1)$，並取 $n = m-1$。

$m = 2$ 而有三個整數的情況

當 $m = 2$ 時，共有 $2^3 = 8$ 個機會均等的三元組 (n_1, n_2, n_3)。記 $S = |n_1 - n_2| + |n_2 - n_3|$：

(n_1, n_2, n_3)	111	112	121	122	211	212	221	222
S	0	1	2	1	1	2	1	0

總和是 8，所以 $\mathrm{E}(S) = 8/8 = 1$。題目所述的值是 $2(m^2-1)/(3m) = 2\cdot 3/6 = 1$，兩者吻合。

> ? **想一想** $|n_1 - n_2|$ 和 $|n_2 - n_3|$ 兩項都涉及 n_2。我們還可以直接把它們的期望值相加嗎？

可以。期望值的線性性質 $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$ 完全不需要獨立性。每一對 (n_i, n_{i+1}) 單獨來看，都是兩個獨立而均勻分佈的整數，所以每一項的期望值都是 $(m^2-1)/(3m)$。一共有 $r-1$ 項，所以

$$\mathrm{E}\big(|n_1 - n_2| + |n_2 - n_3| + \cdots + |n_{r-1} - n_r|\big) = \frac{(r-1)(m^2-1)}{3m}.$$

$r = 3$ 的情況，就是我們剛才就 $m = 2$ 驗證過的結果。

監考員

把書桌沿直線依次編號為 $1, 2, \ldots, m$，於是書桌 i 與書桌 j 相距 $d|i-j|$。設 c_j 為第 j 次召喚所來自的書桌的編號，其中 $j = 1, \ldots, N+1$。我們把召喚者設定為互相獨立，且各自在 $\{1, \ldots, m\}$ 上均勻分佈。這就是我們對「m 名考生中任何一人召喚的機會均等」的理解。由於考生從不同時召喚，也不會在另一次召喚進行期間召喚，監考員會逐一處理各

次召喚，而當第 $j+1$ 次召喚出現時，監考員正站在書桌 c_j 旁。（若同一名考生連續召喚兩次，該段路的長度為 0，而公式已經涵蓋這種情況。）

因此所走的總距離是

$$D = d\,|c_1 - 1| + d\sum_{j=1}^{N} |c_{j+1} - c_j|.$$

*第一段路。*監考員由書桌 1 出發，而 c_1 是均勻分佈的，所以

$$\mathrm{E}|c_1 - 1| = \frac{1}{m}\sum_{i=1}^{m}(i-1) = \frac{m-1}{2}.$$

留意這個值比 $(m^2-1)/(3m)$ 大。由這一列書桌的盡頭出發，令第一段路比平均的長。

*其餘 N 段路。*這裏 $c_1, \ldots, c_{N+1}$ 擔當 $n_1, \ldots, n_r$ 的角色，其中 $r = N+1$，所以由上面的結果可得

$$\mathrm{E}\sum_{j=1}^{N} |c_{j+1} - c_j| = \frac{N(m^2-1)}{3m}.$$

*合併。*再用一次線性性質，

$$\begin{aligned}\mathrm{E}(D) &= d\left(\frac{m-1}{2} + \frac{N(m-1)(m+1)}{3m}\right)\\ &= \frac{d(m-1)}{6m}\big(3m + 2N(m+1)\big) = \frac{d(m-1)}{6m}\big[2N(m+1) + 3m\big],\end{aligned}$$

這正是所求的結果。

驗算

- 當 $m = 1$ 時只有一張書桌，所以監考員從不走動。公式給出 0，理應如此。
- 當 $N = 0$ 時只有第一次召喚，公式化簡為 $d(m-1)\cdot 3m/(6m) = \frac{1}{2}d(m-1)$，即第一段路的期望值。
- 把召喚者的每一個可能序列逐一列舉，可確認公式對所有 $m \le 4$ 及 $N \le 3$ 都成立。以 $m = 10$、$N = 20$、$d = 1$ 進行 2×10^5 次模擬，所得平均值為 70.50，而公式給出的是 $\frac{9}{60}(440 + 30) = 70.5$。

常見錯誤

- 把第一段路當作與其餘各段一樣處理。監考員由書桌 1 出發，而不是由隨機的一張書桌出發，所以該段路的期望值是 $\frac{1}{2}(m-1)$，而不是 $(m^2-1)/(3m)$。

- 只數 $n_1 - n_2 = k$ 的情況，忘記了因 $n_2 - n_1 = k$ 而來的因子 2；或者錯誤地把 $k = 0$ 的情況也乘以二。
- 以為和式中各項必須互相獨立，它們的期望值才可以相加。

STEP 之外

監考員的位置是一條馬可夫鏈（*Markov chain*）：他們下一步往哪裏去，只取決於他們現在身在何處。這裏的鏈非常簡單，每一步都把過去完全「忘記」。在較複雜的模型中（例如召喚較可能來自附近的書桌），我們便需要這條鏈的平穩分佈（stationary distribution），才能求出長遠而言每次召喚平均所走的路程。第一個結果的連續版本是：在 $[0,1]$ 上兩個獨立而均勻分佈的點，平均相距 $\frac{1}{3}$，這正是 $(m^2-1)/(3m)$ 除以 m 後的極限。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）涵蓋了期望值的線性性質和馬可夫鏈。

詳細解答

STEP II 1991，第 16 題

概率與統計 • 二項分佈的正態近似

> **STEP II 1991，第 16 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s2-q16

關鍵思路

追蹤水位差額（deficit）：即水位低於壩頂多少。排水令差額穩定上升，而每場驟雨則令它恰好下降 V；溢流就是差額變為負數。在兩天假期的問題中，只有一種降雨情況有可能引致溢流，而且屆時只有一個均勻分佈的時刻起作用。到了暑假，排水已經關掉，於是一切歸結為數算驟雨的次數，而這個次數服從二項分佈（binomial distribution）。

解答

第 (i) 部分：兩天假期

以日為單位，由假期開始的午夜起計算時間，並設 $D(t)$ 為時刻 t 的差額，故 $D(0) = V$。在兩場驟雨之間，D 以速率 u 上升，而每場驟雨令它下降 V。只要 D 一旦變為負數，水壩便會溢流。（我們按題目的原意假設，水庫存水充足，排水在這兩天內可以一直持續。）

設 R_1、R_2 分別為第 1 天和第 2 天下雨的事件，並設下雨的時刻（以一天的分數表示）為 T_1 和 T_2，兩者各自在 $[0,1]$ 上服從均勻分佈（uniform distribution）。各天互相獨立（independent），這一點已隱含在「任何一天下雨的概率為 p」這句話之中。

- **沒有下雨：**差額只會增加；不會溢流。
- **只有一天下雨**，時刻為 s：驟雨剛過後 $D = V + us - V = us \geq 0$，所以不會溢流（概率為零的事件 $s = 0$ 除外）。
- **兩天都下雨：**第一場驟雨過後 $D = uT_1 \geq 0$，這是安全的。第二場驟雨剛過後，即在時刻 $1 + T_2$，

$$D = V + u(1 + T_2) - 2V = u(1 + T_2) - V.$$

留意 T_1 已經消失：無論第一場驟雨何時落下，排水的速率都一樣。

所以，溢流恰好在兩天都下雨而且 $u(1+T_2) < V$ 時發生，即 $T_2 < \frac{V}{u} - 1$。若 $\frac{1}{2}V \le u \le V$，則 $\frac{V}{u} - 1$ 這個數位於 $[0,1]$ 之內，而由於 T_2 服從均勻分佈，

$$\mathrm{P}(\text{溢流}) = p^2\left(\frac{V}{u} - 1\right).$$

令上式等於 Q：

$$\frac{V}{u} = 1 + \frac{Q}{p^2} \quad\Longrightarrow\quad u = \frac{p^2 V}{p^2 + Q}.$$

條件 $Q < p^2$ 正是令這個結果前後一致的原因：由 $0 < Q/p^2 < 1$ 可得 $\frac{1}{2}V < u < V$，與假設相符。（沒有這個條件便無解。溢流的概率永遠不會超過 p^2，因為必須兩天都下雨。）

第 (i) 部分：暑假

現在 $u = 0$，所以差額永不增加。它由 kV 開始，每場驟雨令它下降 V，因此若 N 為 18 天之中下雨的日數，則

$$\text{溢流} \iff NV > kV \iff N > k.$$

各天互相獨立，每天下雨的概率為 $\frac{1}{3}$，所以 $N \sim \mathrm{B}(18, \frac{1}{3})$，其平均值為 6，方差（variance）為 $18 \cdot \frac{1}{3} \cdot \frac{2}{3} = 4$。

> **想一想** 這裏哪一種近似才算「合適」？為甚麼需要連續性修正（continuity correction）？

由於 $np = 6$ 而 $n(1-p) = 12$，正態近似（normal approximation）$N \approx \mathrm{N}(6,4)$ 是合理的。（泊松近似（Poisson approximation）則不合適，因為 p 並不小。）N 只取整數值，所以對整數 k 而言，事件 $N > k$ 就是事件 $N \ge k+1$，我們以 $X > k + \frac{1}{2}$ 來近似它，其中 $X \sim \mathrm{N}(6,4)$。我們需要

$$\mathrm{P}\left(Z > \frac{k + \frac{1}{2} - 6}{2}\right) < \frac{1}{10} \iff \frac{k - 5.5}{2} > 1.2816 \iff k > 8.06.$$

具體來說，$k = 8$ 給出 $\mathrm{P}(Z > 1.25) \approx 0.106$，太大了；而 $k = 9$ 給出 $\mathrm{P}(Z > 1.75) \approx 0.040$，符合要求。由於溢流與否只取決於 $N > k$ 是否成立，介乎 8 與 9 之間的非整數 k 並不比 $k = 9$ 好。所以工程師應把水庫排水至

$$k = 9,$$

即低於壩頂 $9V$。

第 (ii) 部分

現在共有 36 個半天，互相獨立，每個半天下雨的概率為 $\frac{1}{6}$。排水量仍然是零，所以溢流同樣表示 $N > k$，只是現在 $N \sim \mathrm{B}(36, \frac{1}{6})$。平均值仍然是 6，但方差是 $36 \cdot \frac{1}{6} \cdot \frac{5}{6} = 5$，比 4 大。

平均值相同時，方差較大即表示上尾較重：對每個高於平均值的 k，$\mathrm{P}(N > k)$ 都比之前大。所以所需的 k 較高。以 $\mathrm{N}(6, 5)$ 連同連續性修正作近似，我們需要

$$\frac{k - 5.5}{\sqrt{5}} > 1.2816 \iff k > 8.37,$$

而之前則是 $k > 8.06$。（這裏正態近似給出 $\mathrm{P}(N \geq 9) \approx 0.132$，而第 (i) 部分則是 0.106。）所以門檻上升了。若把 k 向上捨入至整數，兩種情況仍然都得出 $k = 9$。不過在 $k = 9$ 時，尾部概率現在約為 0.06 而不是 0.04，所以安全餘裕較小。

驗算

- 以蒙地卡羅模擬（Monte Carlo simulation）模擬兩天假期（400,000 次），取 $p = 0.5$、$Q = 0.1$ 及 $u = \frac{0.25}{0.35} V \approx 0.714V$，得出的溢流頻率為 0.0996。$p$ 和 Q 取其他數值時，結果同樣吻合。
- 極限情況：當 $Q \to p^2$ 時，$u \to \frac{1}{2}V$，這是兩天下雨仍可引致溢流的最慢排水速率。當 $Q \to 0$ 時，$u \to V$，在這個速率下，第二場驟雨永遠不會造成溢流。
- 精確的二項分佈尾部概率印證了正態近似。對 $\mathrm{B}(18, \frac{1}{3})$，$\mathrm{P}(N \geq 9) = 0.1076$ 及 $\mathrm{P}(N \geq 10) = 0.0433$。對 $\mathrm{B}(36, \frac{1}{6})$，$\mathrm{P}(N \geq 9) = 0.1332$ 及 $\mathrm{P}(N \geq 10) = 0.0652$。

常見錯誤

- 以為在兩天假期的問題中，第一場雨的時刻 T_1 有影響。其實它會互相抵消，因為排水速率是恆定的。
- 忘記連續性修正。沒有它，不等式便變成 $k > 8.56$，掩蓋了 N 是整數這個事實。
- 因為平均值相等，便說 (ii) 的答案「一樣」。其實控制尾部的是分佈的離散程度。

STEP 之外

雨水在隨機的時刻到來，而水則以穩定的速率排走，這是應用概率中的一個經典模型：儲存理論（storage theory），由 Moran 為研究水壩而發展出來。它與排隊（queues）和保險破產（insurance ruin）問題關係密切。在第 (ii) 部分，若把一天分割成愈來愈多、

愈來愈小的時段，同時保持平均值不變，則 $\mathrm{B}(n, 6/n)$ 趨向平均值為 6、方差為 6 的泊松分佈（Poisson distribution），而雨水的到來則成為一個泊松過程（Poisson process）。想進一步了解，可參閱 Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）。

詳細解答

STEP III 1991，第 1 題

純數學 • 級數與展開式

STEP III 1991，第 1 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q1

關鍵思路

三個部分都值得先想一想才動手計算。在 (i) 中，部分分式（partial fractions）把和式拆成可以裂項相消（telescope）的幾部分。在 (ii) 中，三次式 $1+x+x^2+x^3$ 是一個等比級數（geometric series），因此可以寫成兩個非常簡單的式子之商，而對數亦隨之分拆成兩個標準級數。在 (iii) 中，先把指數寫成級數；由於它由 x^2 開始，指數級數（exponential series）中只有頭幾項能夠到達 x^4。

解答

第 (i) 部分

我們要找出常數，使得

$$\frac{6}{r(r+1)(r+3)} = \frac{A}{r} + \frac{B}{r+1} + \frac{C}{r+3}.$$

利用掩蓋法（cover-up rule），$A = \dfrac{6}{1\cdot 3} = 2$，$B = \dfrac{6}{(-1)(2)} = -3$ 及 $C = \dfrac{6}{(-3)(-2)} = 1$。（快速檢查一下：$A+B+C=0$，這是必然的，因為當 r 很大時，左方的表現有如 $6/r^3$，並沒有 $1/r$ 項。）

係數 $2, -3, 1$ 之和為零，這提示我們把它們拆成 $2-2$ 和 $-1+1$：

$$\frac{6}{r(r+1)(r+3)} = 2\left(\frac{1}{r} - \frac{1}{r+1}\right) - \left(\frac{1}{r+1} - \frac{1}{r+3}\right).$$

現在每個括號都可以裂項相消。第一個給出

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

在第二個之中，兩項相隔兩步，因此兩端各有兩項留下：

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

把它們合起來，

$$\begin{aligned}\sum_{r=1}^{n}\frac{6}{r(r+1)(r+3)}&=2-\frac{2}{n+1}-\frac{5}{6}+\frac{1}{n+2}+\frac{1}{n+3}\\&=\frac{7}{6}-\frac{2}{n+1}+\frac{1}{n+2}+\frac{1}{n+3}.\end{aligned}$$

第 (ii) 部分

? **想一想** $1+x+x^2+x^3$ 有沒有更簡潔的寫法？

它是公比為 x 的等比級數，所以當 $x\neq 1$ 時

$$1+x+x^2+x^3=\frac{1-x^4}{1-x}.$$

當 $|x|<1$ 時，$1-x^4$ 和 $1-x$ 均為正數，所以我們可以把對數拆開：

$$\ln(1+x+x^2+x^3)=\ln(1-x^4)-\ln(1-x).$$

利用 $\ln(1-y)=-\sum_{k\geq 1}y^k/k$，此式對 $|y|<1$ 成立（而這裏亦有 $|x^4|<1$），

$$-\ln(1-x)=\sum_{k=1}^{\infty}\frac{x^k}{k},\qquad \ln(1-x^4)=-\sum_{m=1}^{\infty}\frac{x^{4m}}{m}.$$

第二個級數只涉及 k 為 4 的倍數的冪 x^k。對於這樣的 $k=4m$，合併後的係數為

$$\frac{1}{4m}-\frac{1}{m}=-\frac{3}{4m}=-\frac{3}{k}.$$

因此

$$\ln(1+x+x^2+x^3)=x+\frac{x^2}{2}+\frac{x^3}{3}-\frac{3x^4}{4}+\frac{x^5}{5}+\cdots,$$

而通項為

$$\frac{x^k}{k}\ \text{若}\ 4\nmid k,\qquad -\frac{3x^k}{k}\ \text{若}\ 4\mid k.$$

每個係數都不是零，所以上面列出的五項就是首五個非零項。

第 (iii) 部分

當 $-1<x\leq 1$ 時，我們有 $\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots$，所以

$$u=x\ln(1+x)=x^2-\frac{x^3}{2}+\frac{x^4}{3}-\cdots.$$

想一想 $e^u = 1 + u + \frac{u^2}{2} + \frac{u^3}{6} + \cdots$ 之中，我們需要多少項？

由於 u 由 x^2 開始，u^k 這一項便由 x^{2k} 開始。所以 u^3 及其後各項對 x^4 或以下的項毫無貢獻，而從 u^2 我們只需要它的首項 x^4。因此

$$\begin{aligned} e^{x\ln(1+x)} &= 1 + \left(x^2 - \frac{x^3}{2} + \frac{x^4}{3}\right) + \frac{1}{2}x^4 + \cdots \\ &= 1 + x^2 - \frac{x^3}{2} + \frac{5x^4}{6} + \cdots . \end{aligned}$$

（當然，$e^{x\ln(1+x)}$ 其實就是 $(1+x)^x$。）

驗算

- 在 (i) 中，$n = 1$ 給出 $\frac{7}{6} - 1 + \frac{1}{3} + \frac{1}{4} = \frac{3}{4}$，與單獨一項 $\frac{6}{1\cdot 2\cdot 4}$ 吻合。令 $n \to \infty$，便得無窮項之和 $\frac{7}{6}$。
- 在 (ii) 中，直接展開至二階：取 $\ln(1+y) \approx y - \frac{y^2}{2}$，其中 $y = x + x^2 + \cdots$，得出 $x + x^2 - \frac{x^2}{2} = x + \frac{x^2}{2}$，與我們的首兩項一致。數值上，在 $x = 0.5$ 時，把級數加至很多項，結果與 $\ln 1.875 \approx 0.6286$ 吻合。
- 在 (iii) 中，在 $x = 0.01$ 時，展開式與 $(1.01)^{0.01}$ 相差約 7.4×10^{-11}，其大小屬 x^5 的量級，正如所料（下一個係數原來是 $-\frac{3}{4}$）。

常見錯誤

- 在 (i) 中，忘記了相隔兩步的裂項和會在兩端各留下兩項，而不是一項。
- 在 (ii) 中，把 x^4 的係數寫成 $\frac{1}{4} - 1$，但其後卻把通項寫成 $(-1)^{k+1}x^k/k$ 或類似的式子；其實規律取決於 k 除以 4 的餘數。
- 在 (iii) 中，漏掉 $\frac{1}{2}u^2$ 這一項：它貢獻了 $\frac{1}{2}x^4$。

STEP 之外

第 (iii) 部分是冪級數（power series）複合的一個小例子，而追蹤每一項中最低次冪這個記帳技巧，正是大學裏處理漸近展開式（asymptotic expansions）的方法。冪級數何時收斂，以及為何在收斂半徑（radius of convergence）之內可以隨意把它們相加、相乘和複合，這些問題都是分析學（analysis）入門課程的內容；Abbott 的 *Understanding Analysis*（分析學入門教材）對此有清晰的論述。

詳細解答

STEP III 1991，第 2 題

純數學 • 幾何中的複數

> **STEP III 1991，第 2 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s3-q2

關鍵思路

一個複數（complex number）可以同時記錄長度之比和一個角。商 $(z_1 - z_2)/(z_1 - z_3)$ 記下了三角形在 P_1 處兩條邊的長度之比（即它的模（modulus)），以及由一條邊轉到另一條邊所轉過的角（即它的輻角（argument)）；而這兩項資料，正是相似三角形（similar triangles）的 SAS 判別法所需要的。只要把條件寫成行列式（determinant），各部分便歸結為把一個行列式因式分解（factorise），或者巧妙地選取一個三角形 $Q_1Q_2Q_3$ 來作比較。

解答

先處理一個印刷錯誤

按題目所印，左邊的分母是 $z_2 - z_3$。這不可能是對的：取 $w_i = z_i$，使兩個三角形完全相同。這時右邊是 $(z_1 - z_2)/(z_1 - z_3)$，由於 $z_2 \neq z_1$，它與左邊並不相等。題目原意的條件，亦即與行列式吻合的條件，是

$$\frac{z_1 - z_2}{z_1 - z_3} = \frac{w_1 - w_2}{w_1 - w_3}. \qquad (*)$$

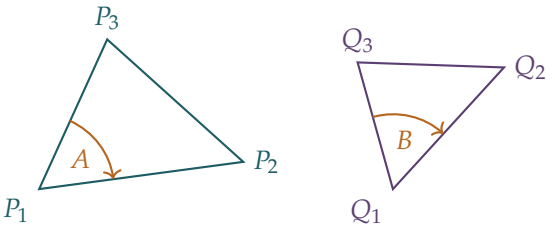
我們要證明的就是這個條件。

相似的條件

設 A 為三角形 $P_1P_2P_3$ 在 P_1 處的角，故 $0 < A < \pi$。以極式（polar form）表示，

$$\frac{z_1 - z_2}{z_1 - z_3} = \frac{z_2 - z_1}{z_3 - z_1} = \frac{P_1P_2}{P_1P_3}\,\mathrm{e}^{\mathrm{i}\theta},$$

其中 θ 是由 $\overrightarrow{P_1P_3}$ 的方向轉到 $\overrightarrow{P_1P_2}$ 的方向所轉過的角。因此，若 $P_1 \to P_2 \to P_3$ 沿逆時針方向排列，則 $\theta = -A$（由 P_1P_3 出發，要順時針轉才到達 P_1P_2）；若沿順時針方向排列，則 $\theta = +A$。簡而言之，這個商的輻角（取值於 $(-\pi, \pi)$）的大小等於 P_1 處的角，而其正負號則告訴我們三角形的繞向（sense）。對各個 w 而言情況相同，其中 B 是 Q_1 處的角。



同向相似 $\Rightarrow (*)$。若兩個三角形相似，且 $P_i \leftrightarrow Q_i$ 互相對應，則 P_1 和 Q_1 處的兩邊之比相同，即 $P_1P_2/P_1P_3 = Q_1Q_2/Q_1Q_3$，而該處的角亦相等，即 $A = B$。由於兩個三角形的繞向相同，兩個輻角 $\mp A$ 和 $\mp B$ 的正負號相同，所以它們相等。模相等，輻角也相等，便得出 $(*)$。

$(*) \Rightarrow$ 同向相似。若 $(*)$ 成立，比較兩邊的模可得 $P_1P_2/P_1P_3 = Q_1Q_2/Q_1Q_3$。比較兩邊的輻角可得 $A = B$，而且正負號相同，所以兩個三角形的繞向相同。兩邊成比例而夾角相等，正是 SAS 條件，故兩個三角形相似，且 $P_i \leftrightarrow Q_i$ 互相對應。（若 $(*)$ 是實數，則兩組三點各自共線（collinear），這兩個退化的「三角形」仍然相似。）

還可以從另一個角度看同一件事：$(*)$ 恰好表示 $w_i = az_i + b$，其中常數 $a = (w_1 - w_2)/(z_1 - z_2) \neq 0$，$b = w_1 - az_1$。映射 $z \mapsto az + b$ 是倍數為 $|a|$ 的放大（enlargement），再加上轉角為 $\arg a$ 的旋轉（rotation）和一個平移（translation），而這樣的映射保持形狀和繞向不變。

行列式形式

去分母後（由於各點互不相同，兩個分母均不為零），$(*)$ 等價於

$$(z_1 - z_2)(w_1 - w_3) - (z_1 - z_3)(w_1 - w_2) = 0.$$

展開後，z_1w_1 項互相抵消，餘下

$$z_1(w_2 - w_3) - z_2(w_1 - w_3) + z_3(w_1 - w_2) = 0,$$

這正是該行列式沿第一行的展開式。所以 $(*)$ 等同於 $\det = 0$。

處理以下各部分時，用另一個方法計算這個行列式會較為方便。從第二列和第三列分別

減去第一列，行列式的值不變；然後沿最底一行展開，這一行現在是 $(1,0,0)$：

$$\det\begin{pmatrix} z_1 & z_2 & z_3 \\ w_1 & w_2 & w_3 \\ 1 & 1 & 1 \end{pmatrix} = \det\begin{pmatrix} z_2-z_1 & z_3-z_1 \\ w_2-w_1 & w_3-w_1 \end{pmatrix}. \tag{†}$$

第 (i) 部分

取 $w_i = z_i^2$，(†) 變成

$$\begin{aligned}(z_2-z_1)(z_3^2-z_1^2)-(z_3-z_1)(z_2^2-z_1^2) &= (z_2-z_1)(z_3-z_1)\big[(z_3+z_1)-(z_2+z_1)\big] \\ &= (z_2-z_1)(z_3-z_1)(z_3-z_2).\end{aligned}$$

由於各 z_i 互不相同，每個因子均不為零。所以行列式永不為零，兩個三角形永不相似（按上述意義）。

第 (ii) 部分

取 $w_i = z_i^3$，並利用 $z_j^3 - z_1^3 = (z_j - z_1)(z_j^2 + z_j z_1 + z_1^2)$：

$$\begin{aligned}(z_2-z_1)(z_3^3-z_1^3)-(z_3-z_1)(z_2^3-z_1^3) &= (z_2-z_1)(z_3-z_1)\big[(z_3^2+z_3z_1)-(z_2^2+z_2z_1)\big] \\ &= (z_2-z_1)(z_3-z_1)(z_3-z_2)(z_1+z_2+z_3),\end{aligned}$$

這是因為 $z_3^2 - z_2^2 + z_1(z_3 - z_2) = (z_3 - z_2)(z_3 + z_2 + z_1)$。首三個因子均不為零，所以行列式為零當且僅當 $z_1 + z_2 + z_3 = 0$。換言之，兩個三角形相似，當且僅當形心(centroid) $\frac{1}{3}(z_1 + z_2 + z_3)$ 位於原點。

第 (iii) 部分

竅門是把三角形與重新標記頂點後的自己作比較。取 Q_1, Q_2, Q_3 為 P_2, P_3, P_1，即

$$w_1 = z_2, \qquad w_2 = z_3, \qquad w_3 = z_1.$$

? **想一想** 「等邊」為甚麼會與這樣的重新標記有關？

若在 $P_1 \leftrightarrow P_2$、$P_2 \leftrightarrow P_3$、$P_3 \leftrightarrow P_1$ 的對應下相似，則各組對應邊必有相同的比 $k > 0$：

$$\frac{P_2P_3}{P_1P_2} = \frac{P_3P_1}{P_2P_3} = \frac{P_1P_2}{P_3P_1} = k.$$

把三者相乘得 $k^3 = 1$，所以 $k = 1$，三邊全部相等。反之，若三角形是等邊 (equilateral) 三角形，則繞其中心旋轉 120°（取適當的方向）會把 $P_1 \to P_2 \to P_3 \to P_1$ 依次對應過去。旋轉保持形狀和繞向不變，所以該三角形與重新標記後的自己同向相似。因此

$$P_1P_2P_3 \text{ 為等邊三角形} \iff \text{它同向相似於} P_2P_3P_1.$$

根據第一部分，右邊成立當且僅當

$$0 = \det\begin{pmatrix} z_1 & z_2 & z_3 \\ z_2 & z_3 & z_1 \\ 1 & 1 & 1 \end{pmatrix} = z_1(z_3 - z_1) - z_2(z_2 - z_1) + z_3(z_2 - z_3),$$

即

$$z_1z_3 - z_1^2 - z_2^2 + z_1z_2 + z_2z_3 - z_3^2 = 0 \iff z_2z_3 + z_3z_1 + z_1z_2 = z_1^2 + z_2^2 + z_3^2,$$

這正是所要證明的結果。

驗算

- 第 (ii) 部分：$z_1, z_2, z_3 = 1, \omega, \omega^2$（其中 $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$）的形心為 0，但它們的立方全都等於 1，並非互不相同。改用 $z = 1,\ \mathrm{i},\ -1 - \mathrm{i}$。形心為 0，而立方分別是 $1,\ -\mathrm{i},\ 2 - 2\mathrm{i}$。按 P_1P_2, P_2P_3, P_3P_1 的次序，兩個三角形的邊長都是 $\sqrt{2}, \sqrt{5}, \sqrt{5}$（在這個例子中它們甚至全等（congruent）），而且兩者都沿逆時針方向排列。
- 第 (iii) 部分：對於 $z = 0, 1, \mathrm{e}^{\mathrm{i}\pi/3}$，兩邊都等於 $\mathrm{e}^{\mathrm{i}\pi/3}$，因為 $1 + \mathrm{e}^{2\mathrm{i}\pi/3} = \mathrm{e}^{\mathrm{i}\pi/3}$。
- 以電腦對 10^4 個隨機三角形作測試，結果證實 $w_i = az_i + b$ 總是滿足 $(*)$，而經反射的三角形 $w_i = a\bar{z}_i + b$ 則從不滿足。測試亦證實了 (†) 的兩個因式分解，以及等邊三角形的判別準則。至於題目所印的條件（分母為 $z_2 - z_3$），在所有真正相似的情況中無一成立。

常見錯誤

- 忽略三角形的繞向。單看模，只能得出「容許反射的相似」。互為鏡像的三角形，其輻角的正負號相反，所以並不滿足 $(*)$。
- 在第 (iii) 部分忘記證明逆命題（converse）：「與重新標記後的自己相似」這個論證必須雙向都成立，而 $k^3 = 1 \Rightarrow k = 1$ 這一步用到了 k 是正實數。

STEP 之外

映射 $z \mapsto az + b$ 就是平面上保持定向（orientation-preserving）的相似變換。它們是最簡單的莫比烏斯變換（Möbius transformations），而四個點的交比（cross-ratio）則是 $(*)$ 的推廣。等邊條件亦可因式分解為 $(z_1 + \omega z_2 + \omega^2 z_3)(z_1 + \omega^2 z_2 + \omega z_3) = 0$，這表示 z_1, z_2, z_3 沿其中一個繞向圍成一個等邊三角形。Needham 的 *Visual Complex Analysis* 把這種以幾何眼光看待複數的思考方式發揮得十分精彩。

詳細解答

STEP III 1991，第 3 題

純數學 • 分段函數與反函數

STEP III 1991，第 3 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q3

關鍵思路

絕對值（modulus）不過是一個函數，其內部的式子變號時，它的表達式就跟著改變。這裏三個絕對值只在 $x = -1$、0 和 1 處變號，因此在它們把 $x < 2$ 分割出來的四個區間上，函數各自都是一個普通的二次函數（quadratic）。一旦求得這四個二次函數，每一個問題（轉向點、拐點、連續性、光滑性、反函數）都只是關於簡單拋物線及其如何相接的問題。

解答

去掉絕對值

把絕對值內的式子因式分解：$x^2 - x = x(x-1)$、$x^2 - 1 = (x-1)(x+1)$ 和 $x^2 + x = x(x+1)$。它們的正負號如下：

區間	$x^2 - x$	$x^2 - 1$	$x^2 + x$	$\mathrm{f}(x)$
$x \le -1$	$+$	$+$	$+$	$x^2 - 4x - 1$
$-1 \le x \le 0$	$+$	$-$	$-$	$3x^2 + 1$
$0 \le x \le 1$	$-$	$-$	$+$	$1 - 5x^2$
$1 \le x < 2$	$+$	$+$	$+$	$x^2 - 4x - 1$

例如，在 $-1 \le x \le 0$ 上，

$$\mathrm{f}(x) = 2(x^2 - x) + (1 - x^2) + 2(x^2 + x) = 3x^2 + 1,$$

而在 $0 \le x \le 1$ 上，$\mathrm{f}(x) = -2(x^2 - x) + (1 - x^2) - 2(x^2 + x) = 1 - 5x^2$。在外側的兩個區間上，三個絕對值內的式子全為正，而 $2(x^2 - x) + (x^2 - 1) - 2(x^2 + x) = x^2 - 4x -$

1。在端點處，其中一個絕對值內的式子為零，所以兩邊的公式在該處都正確。

連續性與可導性

f 由 $|\cdot|$ 和多項式組合而成，這些全都是連續的，所以 f 在 $x < 2$ 上**處處連續（continuous）**。作為驗算，相鄰的公式在接駁處吻合：$f(-1) = 4$、$f(0) = 1$、$f(1) = -4$。

至於可導性（differentiability），可比較接駁處的單側導數（one-sided derivative）。四段的導數分別為 $2x - 4$、$6x$、$-10x$ 和 $2x - 4$。

- 在 $x = -1$：左側為 $2(-1) - 4 = -6$，右側為 $6(-1) = -6$。兩者相同，所以 $f'(-1) = -6$。
- 在 $x = 0$：兩側都得 0，所以 $f'(0) = 0$。
- 在 $x = 1$：左側為 -10，右側為 $2 - 4 = -2$。兩者不同，所以圖像在 $(1, -4)$ 處有一個**尖角（corner）**。

因此 f **並非處處可導**：它恰好只在 $x = 1$ 處不可導。

轉向點與拐點

在每一段上都有 $f' < 0$，只有 $x = 0$ 這一點例外：

- $x < -1$ 和 $1 < x < 2$：由於 $x < 2$，$f'(x) = 2x - 4 < 0$；
- $-1 < x < 0$：$f'(x) = 6x < 0$；
- $0 < x < 1$：$f'(x) = -10x < 0$。

由於 f 連續，而且在每一段上都嚴格遞減，所以它**在整個 $x < 2$ 上嚴格遞減（strictly decreasing）**。特別地，它**沒有極大點或極小點**，連局部的也沒有。

> ? **想一想** $f'(0) = 0$，那麼 $(0, 1)$ 不就是轉向點（turning point）嗎？

不是。左邊一段 $3x^2 + 1$ 在 0 處取得極小值，右邊一段 $1 - 5x^2$ 則在該處取得極大值，但在各自公式適用的區間上，兩者都是遞減的。所以 f 經過 $x = 0$ 時仍然在下降，只是切線是水平的。

現在看凹凸性（concavity）。各段的二階導數為

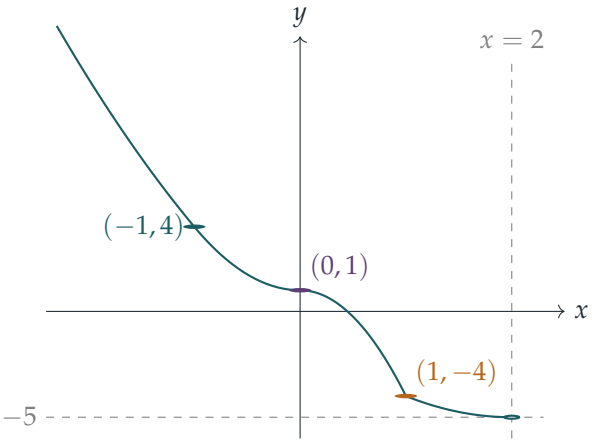
$$f'' = 2\ (x < -1), \qquad 6\ (-1 < x < 0), \qquad -10\ (0 < x < 1), \qquad 2\ (1 < x < 2).$$

- 在 $x = 0$ 處，圖像由下凸（convex）變為上凸（concave），而且在該處有切線，所以 $(0, 1)$ 是一個**駐拐點（stationary point of inflection）**。
- 在 $x = -1$ 處，f'' 由 2 跳到 6，但仍然為正，所以沒有拐點。

- 在 $x = 1$ 處，曲線由上凸變回下凸，但該處是尖角，沒有切線。按照通常的定義（切線穿過曲線），$(1, -4)$ 不是拐點。作答時值得明確指出這一點。

草圖

當 $x \to -\infty$ 時，$f(x) \to +\infty$。當 $x \to 2^-$ 時，$f(x) \to 4 - 8 - 1 = -5$，這個值並不會取到。



曲線一直下降。它在 $(0, 1)$ 處變平，形成一個水平拐點，並在 $(1, -4)$ 處有一個尖角，斜率在該處由 -10 跳到 -2。

反函數

由於 f 在 $x < 2$ 上連續且嚴格遞減，而且在左端 $f \to \infty$，在右端 $f \to -5$，所以它是由 $(-\infty, 2)$ 到 $(-5, \infty)$ 的雙射（bijection）。因此 f^{-1} 定義在 $x > -5$ 上。我們利用 $y = f(x)$，在每一段所覆蓋的值域上分別求反函數（inverse）。

- $x \le -1$，值域 $y \ge 4$：$y = x^2 - 4x - 1 = (x-2)^2 - 5$，所以 $(x-2)^2 = y + 5$。這裏 $x - 2 < 0$，所以取負根：$x = 2 - \sqrt{y+5}$。
- $-1 \le x \le 0$，值域 $1 \le y \le 4$：$x^2 = (y-1)/3$，而 $x \le 0$，所以 $x = -\sqrt{(y-1)/3}$。
- $0 \le x \le 1$，值域 $-4 \le y \le 1$：$x^2 = (1-y)/5$，而 $x \ge 0$，所以 $x = \sqrt{(1-y)/5}$。
- $1 \le x < 2$，值域 $-5 < y \le -4$：同樣有 $(x-2)^2 = y + 5$ 且 $x - 2 < 0$，所以 $x = 2 - \sqrt{y+5}$。

把變量改寫為 x：

$$\mathrm{f}^{-1}(x)=\begin{cases}2-\sqrt{x+5}, & -5<x\le -4,\\ \sqrt{\frac{1-x}{5}}, & -4\le x\le 1,\\ -\sqrt{\frac{x-1}{3}}, & 1\le x\le 4,\\ 2-\sqrt{x+5}, & x\ge 4.\end{cases}$$

各公式在共用的端點處吻合。例如，在 $x=4$ 處兩者都得 -1，在 $x=-4$ 處兩者都得 1。

驗算

- 抽查數值：$\mathrm{f}(\frac{1}{2})=2\cdot\frac{1}{4}+\frac{3}{4}-2\cdot\frac{3}{4}=-\frac{1}{4}$，而 $1-5\cdot\frac{1}{4}=-\frac{1}{4}$。此外 $\mathrm{f}^{-1}(-\frac{1}{4})=\sqrt{1.25/5}=\frac{1}{2}$。
- 直接計算，$\mathrm{f}(-3)=2\cdot 12+8-2\cdot 6=20$。公式給出 $9+12-1=20$，而 $\mathrm{f}^{-1}(20)=2-\sqrt{25}=-3$。
- 數值上，對 10^5 個隨機選取的 $x<2$ 都有 $\mathrm{f}^{-1}(\mathrm{f}(x))=x$；函數在一個細密網格的每一點上都在遞減；而 $x=1$ 處的單側斜率算出來是 -10 和 -2。

常見錯誤

- 因為 $\mathrm{f}'(0)=0$ 就把 $(0,1)$ 稱為極大點或極小點。如果某條拋物線本身的轉向點位於該公式適用區間的邊緣，那個轉向點就無關重要。
- 求反函數時選錯平方根。應由 x（或 $x-2$）在每個區間上的正負號來決定。
- 以為兩個公式的接駁處必定是不可導點。在 $x=-1$ 和 $x=0$ 處，兩邊的斜率是吻合的。

STEP 之外

分段定義的函數（piecewise-defined function）可以連續而不可導，也可以可導而二階導數不連續（這裏 f'' 在 $x=-1$ 和 $x=0$ 處跳躍）。反函數定理（inverse function theorem）指出，嚴格單調（strictly monotonic）的連續函數有連續的反函數，而且在 $\mathrm{f}'\neq 0$ 的地方反函數可導。留意 f^{-1} 在 $x=1$ 處有一條垂直切線，恰好對應 $\mathrm{f}'(0)=0$ 之處。Abbott 的 *Understanding Analysis*（數學分析入門讀本）對這些概念有細緻的闡述。

詳細解答

STEP III 1991，第 4 題

純數學 • 三維空間中的直線與平面

STEP III 1991，第 4 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q4

關鍵思路

同一位觀察者 O_1 的兩條不同視線（line of sight）都經過 O_1，所以它們張成一個平面（plane），而 P 的路徑與這兩條視線都相交。一條直線若與經過 O_1 的兩條直線相交，而本身又不經過 O_1，就必定位於這兩條直線所在的平面上。對 O_2 也是同樣道理，所以路徑就是兩個平面的交線。這樣一來，一個關於四條看似異面（skew）的直線的問題，就化為一個叉積（cross product）和一組聯立方程（simultaneous equations）。

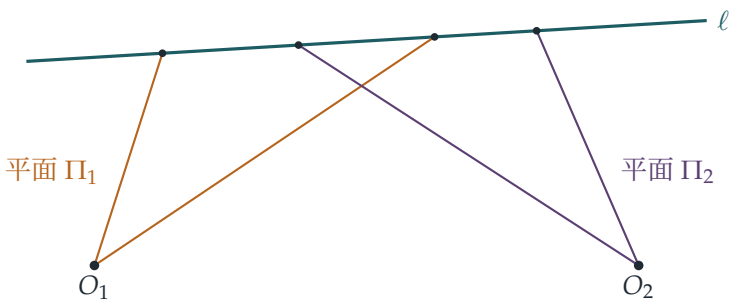
解答

設 P 沿直線 ℓ 運動。四條視線的方向為

$$\mathbf{d}_1 = (2,0,3), \qquad \mathbf{d}_1' = (0,3,4) \qquad （經過 O_1）,$$

$$\mathbf{d}_2 = (-3,1,3), \qquad \mathbf{d}_2' = (-4,2,5) \qquad （經過 O_2）.$$

每條視線都連接一位觀察者與 P 的某個位置，所以 ℓ 與這四條視線全部相交。



（僅為示意圖：兩條琥珀色視線與 ℓ 位於同一平面 Π_1 上，兩條紫紅色視線與 ℓ 則位於另一平面 Π_2 上。）

為何 ℓ 位於經過每位觀察者的一個平面上

首先，ℓ 不可能經過 O_1。假如它經過，那麼每當 $P \neq O_1$ 時，從 O_1 出發的視線就是 ℓ 本身，於是從 O_1 出發不可能有兩條不同的視線。同理，ℓ 也不經過 O_2。

從 O_1 出發的兩條視線是經過 O_1 的不同直線，所以它們只在 O_1 相交，並張成一個平面 Π_1。直線 ℓ 與它們交於兩點，而這兩點互不相同，因為兩條視線唯一的公共點是 O_1，而 ℓ 並不經過 O_1。這兩點都在 Π_1 上，所以整條 ℓ 都在 Π_1 上。同樣地，ℓ 位於平面 Π_2 上，這個平面經過 O_2，並包含 $\mathbf{d}_2$ 和 $\mathbf{d}_2'$。

> ? **想一想** 難道 ℓ 不可以只是與每條視線各交於一點，而不位於任何特定的平面上嗎？

若只有一條視線，的確可以。但這裏有兩個交點落在經過 O_1 的直線上，而兩個相異的點就確定了 ℓ。這兩點都在 Π_1 上，所以 ℓ 別無選擇：它就是經過這兩點的直線，而這條直線位於 Π_1 上。

平面 Π_1

Π_1 的一個法向量（normal）是

$$\mathbf{d}_1 \times \mathbf{d}_1' = (2,0,3) \times (0,3,4) = (0-9,\ 0-8,\ 6-0) = (-9,-8,6),$$

而 Π_1 經過原點，所以

$$\Pi_1: \quad 9x + 8y - 6z = 0.$$

平面 Π_2

Π_2 的一個法向量是

$$\mathbf{d}_2 \times \mathbf{d}_2' = (-3,1,3) \times (-4,2,5) = (5-6,\ -12+15,\ -6+4) = (-1,3,-2),$$

而 Π_2 經過 $O_2(8a,0,0)$，所以 $-(x-8a) + 3y - 2z = 0$，即

$$\Pi_2: \quad x - 3y + 2z = 8a.$$

（快速驗算：$(-3,1,3)\cdot(-1,3,-2) = 3+3-6 = 0$ 及 $(-4,2,5)\cdot(-1,3,-2) = 4+6-10 = 0$。）

我們假設 $a \neq 0$，否則兩位觀察者會重合。於是 O_2 不在 Π_1 上（因為 $9 \cdot 8a \neq 0$），而 O_1 也不在 Π_2 上，這與上文所說的一致。

路徑

法向量 $(-9,-8,6)$ 與 $(-1,3,-2)$ 並不平行，所以 Π_1 與 Π_2 相交於唯一一條直線，而這條直線必定是 ℓ。它的方向與兩個法向量都垂直：

$$(-9,-8,6)\times(-1,3,-2)=(16-18,\ -6-18,\ -27-8)=(-2,-24,-35),$$

所以我們可以取方向為 $(2,24,35)$。要找直線上的一點，可設 $y=0$。兩條方程變成 $9x=6z$ 和 $x+2z=8a$，所以 $x=\frac{2}{3}z$ 及 $\frac{8}{3}z=8a$。由此得 $z=3a$ 和 $x=2a$。所以 $(2a,0,3a)$ 在 ℓ 上。它也在從 O_1 出發的第一條視線上，正如所料。

因此 P 的路徑為

$$\boxed{\frac{x-2a}{2}=\frac{y}{24}=\frac{z-3a}{35}}$$

或等價地，由 $9x+8y-6z=0$ 和 $x-3y+2z=8a$ 這一對方程給出。

最後，ℓ 確實與全部四條視線相交，而且交點都不是觀察者所在的點。把 ℓ 寫成 $(2a+2\lambda,\ 24\lambda,\ 3a+35\lambda)$，它與四條視線的交點依次為

$$(2a,0,3a),\qquad (0,-24a,-32a),\qquad \left(\tfrac{80a}{37},\tfrac{72a}{37},\tfrac{216a}{37}\right),\qquad \left(\tfrac{56a}{25},\tfrac{72a}{25},\tfrac{36a}{5}\right)$$

。

想一想 當 $a>0$ 時，第二個點的 y 和 z 都是負數。這有問題嗎？

題目給出的視線是整條直線，而 $(0,-24a,-32a)$ 滿足 $x=0,\ y/3=z/4$。我們只用到這些直線，並沒有用到觀察者面向哪一方，所以答案成立。如果你把「視線」理解為從觀察者出發的射線，這次觀測就落在 O_1 的背後。那只是數據的一個古怪之處，並非方法有錯。

驗算

- 把點 $(2a,0,3a)$ 和方向 $(2,24,35)$ 代入兩個平面方程：$9(2)+8(24)-6(35)=18+192-210=0$ 及 $2-72+70=0$，所以這個方向位於兩個平面上。
- 取 $a=1$，點 $\left(\frac{80}{37},\frac{72}{37},\frac{216}{37}\right)$ 等於 $O_2+\frac{72}{37}(-3,1,3)$，而且 $\frac{80/37-2}{2}=\frac{3}{37}=\frac{72/37}{24}$。所以它既在視線上，也在路徑上。
- 我們對 a 的若干個值，以精確的有理數運算驗證了全部四個交點。

常見錯誤

- 試圖求出各對視線的交點：不同觀察者在不同時刻的視線，根本不一定相交。

- 忘記 Π_2 經過的是 O_2 而不是原點，結果得出 $x - 3y + 2z = 0$ 而不是 $= 8a$。
- 計算叉積時正負號出錯。務必把法向量與兩個方向向量分別再作點積驗算。

STEP 之外

空間中與四條給定直線都相交的直線有多少條？若四條直線處於一般位置（general position），答案是兩條（在複數範圍內計算；在實數範圍內則可以是零條）。這是枚舉幾何（enumerative geometry）的一個經典結果，而本題的四條直線比較特殊，因為它們兩兩一組地經過兩位觀察者。同一個「視線平面」的想法，也是三角測量（triangulation）和立體視覺中對極幾何（epipolar geometry）的基礎：兩部相機藉此還原三維空間中各點的位置。至於直線、平面與交集的線性代數，可閱讀 Strang 的 *Introduction to Linear Algebra*（線性代數入門教材）。

詳細解答

STEP III 1991，第 5 題

純數學 • 極坐標曲線與常微分方程

STEP III 1991，第 5 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q5

關鍵思路

先把極坐標（polar coordinates）放在一旁：這條方程其實只是一條以 θ 為變量的二階常係數線性常微分方程（linear second-order ODE with constant coefficients）。按慣常方法求解（餘函數加特別積分），用 $\theta = \frac{\pi}{2}$ 處的條件定出常數，然後再回到幾何。由 $r = f(\theta)$ 給出的曲線，若 r 起點和終點皆為 0，而中間一直保持正值，便是一個閉合環圈（closed loop），其面積為 $\frac{1}{2}\int r^2\,\mathrm{d}\theta$。

解答

解微分方程

輔助方程(auxiliary equation)$m^2+4=0$ 的根為 $m=\pm 2\mathrm{i}$，所以餘函數(complementary function）是

$$r = A\cos 2\theta + B\sin 2\theta.$$

至於特別積分（particular integral），$\sin 3\theta$ 並不在餘函數之內，所以我們試用 $r = k\sin 3\theta$。於是 $r'' = -9k\sin 3\theta$，而

$$-9k\sin 3\theta + 4k\sin 3\theta = 5\sin 3\theta \quad\Longrightarrow\quad k = -1.$$

因此通解（general solution）為

$$r = A\cos 2\theta + B\sin 2\theta - \sin 3\theta, \qquad \frac{\mathrm{d}r}{\mathrm{d}\theta} = -2A\sin 2\theta + 2B\cos 2\theta - 3\cos 3\theta.$$

在 $\theta = \frac{\pi}{2}$ 處，有 $\cos 2\theta = -1$、$\sin 2\theta = 0$、$\sin 3\theta = -1$ 及 $\cos 3\theta = 0$。兩個條件變成

$$r = -A + 1 = 1 \quad\Longrightarrow\quad A = 0, \qquad \frac{\mathrm{d}r}{\mathrm{d}\theta} = -2B = -2 \quad\Longrightarrow\quad B = 1.$$

故此

$$r = \sin 2\theta - \sin 3\theta.$$

曲線是一個閉合環圈

要看出 r 在何處為零、其正負號如何，可利用恆等式 $\sin P - \sin Q = 2\cos\frac{P+Q}{2}\sin\frac{P-Q}{2}$ 作因式分解：

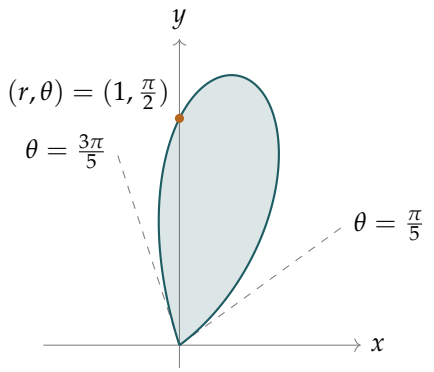
$$r = 2\cos\frac{5\theta}{2}\sin\left(-\frac{\theta}{2}\right) = -2\sin\frac{\theta}{2}\cos\frac{5\theta}{2}.$$

想一想 為甚麼因式分解後的形式能解決「閉合環圈」這個問題？

因為在給定的範圍內，兩個因子的正負號都一目了然。當 $\frac{\pi}{5} \le \theta \le \frac{3\pi}{5}$ 時：

- $\frac{\theta}{2}$ 位於 $\left[\frac{\pi}{10}, \frac{3\pi}{10}\right]$ 之內，所以 $\sin\frac{\theta}{2} > 0$；
- $\frac{5\theta}{2}$ 位於 $\left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$ 之內，所以 $\cos\frac{5\theta}{2} \le 0$，等號只在兩個端點成立。

所以在 $\theta = \frac{\pi}{5}$ 及 $\theta = \frac{3\pi}{5}$ 處 $r = 0$，而對嚴格介乎兩者之間的每一個 θ 都有 $r > 0$。因此曲線沿射線 $\theta = \frac{\pi}{5}$ 離開原點，以正的 r 繞行一周，再沿射線 $\theta = \frac{3\pi}{5}$ 返回原點。其間每一條射線 $\theta =$ 常數 都與曲線恰好相交一次（就是對應該 $r > 0$ 的那一點），所以曲線不會自相交。它是單一個閉合環圈，圍出區域 $0 \le r \le \sin 2\theta - \sin 3\theta$，$\frac{\pi}{5} \le \theta \le \frac{3\pi}{5}$。



面積

所圍的面積為

$$\mathcal{A} = \frac{1}{2}\int_{\pi/5}^{3\pi/5} r^2\,\mathrm{d}\theta = \frac{1}{2}\int_{\pi/5}^{3\pi/5}\left(\sin^2 2\theta - 2\sin 2\theta\sin 3\theta + \sin^2 3\theta\right)\mathrm{d}\theta.$$

利用 $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$ 及 $2\sin P\sin Q = \cos(P-Q) - \cos(P+Q)$，

$$r^2 = 1 - \tfrac{1}{2}\cos 4\theta - \tfrac{1}{2}\cos 6\theta - \cos\theta + \cos 5\theta.$$

記 $D = \sin\frac{2\pi}{5} - \sin\frac{\pi}{5}$。我們在 $\left[\frac{\pi}{5}, \frac{3\pi}{5}\right]$ 上逐項積分，並利用 $\sin(x+2\pi) = \sin x$ 及 $\sin(\pi - x) = \sin x$ 把每個正弦值化為 $\pm\sin\frac{\pi}{5}$ 或 $\pm\sin\frac{2\pi}{5}$：

$$\begin{aligned}
\int 1\,\mathrm{d}\theta &= \tfrac{2\pi}{5},\\
\int -\tfrac{1}{2}\cos 4\theta\,\mathrm{d}\theta &= -\tfrac{1}{8}\left[\sin\tfrac{12\pi}{5} - \sin\tfrac{4\pi}{5}\right] = -\tfrac{1}{8}\left[\sin\tfrac{2\pi}{5} - \sin\tfrac{\pi}{5}\right] = -\tfrac{1}{8}D,\\
\int -\tfrac{1}{2}\cos 6\theta\,\mathrm{d}\theta &= -\tfrac{1}{12}\left[\sin\tfrac{18\pi}{5} - \sin\tfrac{6\pi}{5}\right] = -\tfrac{1}{12}\left[-\sin\tfrac{2\pi}{5} + \sin\tfrac{\pi}{5}\right] = \tfrac{1}{12}D,\\
\int -\cos\theta\,\mathrm{d}\theta &= -\left[\sin\tfrac{3\pi}{5} - \sin\tfrac{\pi}{5}\right] = -D,\\
\int \cos 5\theta\,\mathrm{d}\theta &= \tfrac{1}{5}\left[\sin 3\pi - \sin\pi\right] = 0.
\end{aligned}$$

（第三行中，$\sin\frac{18\pi}{5} = \sin\left(-\frac{2\pi}{5}\right) = -\sin\frac{2\pi}{5}$ 而 $\sin\frac{6\pi}{5} = -\sin\frac{\pi}{5}$。）相加得

$$\int_{\pi/5}^{3\pi/5} r^2\,\mathrm{d}\theta = \frac{2\pi}{5} + \left(-\frac{1}{8} + \frac{1}{12} - 1\right)D = \frac{2\pi}{5} - \frac{25}{24}D,$$

再除以二，

$$\mathcal{A} = \frac{\pi}{5} - \frac{25}{48}D = \frac{\pi}{5} + \frac{25}{48}\left[\sin\frac{\pi}{5} - \sin\frac{2\pi}{5}\right],$$

這正是所求的結果。

驗算

- 在 $\theta = \frac{\pi}{2}$ 處：$r = \sin\pi - \sin\frac{3\pi}{2} = 1$ 而 $r' = 2\cos\pi - 3\cos\frac{3\pi}{2} = -2$，符合要求。
- 以數值計算，$\mathcal{A} \approx 0.6283 - \frac{25}{48}(0.3633) \approx 0.4391$。對 $\frac{1}{2}\int r^2\,\mathrm{d}\theta$ 使用辛普森法則（Simpson's rule），以及對描出的環圈使用鞋帶公式（shoelace formula），兩者都得出 0.43911。由 $\theta = \frac{\pi}{2}$ 出發以數值方法逐步求解微分方程，可重現 $r = \sin 2\theta - \sin 3\theta$，並確認 r 在兩端都回到 0。
- 答案合理：環圈最遠達到 $r \approx 1.22$，所張的角為 $\frac{2\pi}{5}$。半徑為 1.22、圓心角為該角的扇形，面積約為 0.93，而環圈佔去其中略少於一半。

常見錯誤

- 弄錯特別積分的正負號：$(-9+4)k = 5$ 給出 $k = -1$，而不是 $+1$。
- 忘記證明在兩個端點之間 $r > 0$。單憑 r 有兩個零點並不足以構成環圈：r 有可能在中間變號，那樣 $\frac{1}{2}\int r^2\,\mathrm{d}\theta$ 就不是所圍的面積了。
- 化簡 $\sin\frac{12\pi}{5}$ 或 $\sin\frac{18\pi}{5}$ 時粗心大意：兩者各為 $\pm\sin\frac{2\pi}{5}$，而正負號至關重要。

STEP 之外

這裏的方法之所以行得通，是因為方程是線性的：每一個解都是一個特解加上兩個獨立解 $\cos 2\theta$ 和 $\sin 2\theta$ 的組合，而兩個初始條件（initial conditions）恰好選出唯一的一個解。這就是線性常微分方程的存在性與唯一性定理（existence and uniqueness theorem），大學一年級的課程會證明它，並且經常用到。一個相關而又引人入勝的事實是：對於在平方反比力（inverse-square force）作用下的軌道，$u = 1/r$ 滿足 $u'' + u =$ 常數，牛頓（Newton）定律正是這樣推出圓錐曲線（conic section）軌道的。Boyce 與 DiPrima 的 *Elementary Differential Equations*（常微分方程入門教材）對線性理論有詳盡的論述。

詳細解答

STEP III 1991，第 6 題

純數學 • 雙曲線性變換

> **STEP III 1991，第 6 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s3-q6

本題在 *Beyond STEP* 第 7.6 節雙曲旋轉與相對論中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch07/hyperbolic-rotations.html

關鍵思路

這個矩陣(matrix)的表現與旋轉矩陣(rotation matrix)相似，只是以雙曲函數(hyperbolic functions) 取代了圓函數，而恆等式 $\cosh^2 u - \sinh^2 u = 1$ 以及 cosh 和 sinh 的加法公式（addition formulae）就足以完成全部工作。把斜率（gradient）為 $\tanh v$ 的方向寫成向量 $(\cosh v, \sinh v)$，T 對斜率的作用便一目了然：它只是把 u 加到「雙曲角」v 上。至於夾角的條件，我們把兩條直線夾角的正切用 v_1 和 v_2 表示，再看看甚麼必須保持不變。

解答

全篇記 $c = \cosh u$ 及 $s = \sinh u$，於是 $c^2 - s^2 = 1$，而且因為 $u > 0$，所以 $s > 0$。這個變換（transformation）是

$$x' = cx + sy, \qquad y' = sx + cy.$$

雙曲線

對任意一點，

$$\begin{aligned} x'^2 - y'^2 &= (cx + sy)^2 - (sx + cy)^2 \\ &= (c^2 - s^2)x^2 + 2csxy - 2csxy - (c^2 - s^2)y^2 \\ &= x^2 - y^2. \end{aligned}$$

所以 T 保持 $x^2 - y^2$ 不變，而 $x^2 - y^2 = 1$ 上的每一點都被送到 $x^2 - y^2 = 1$ 上的一點。整條曲線都被完全覆蓋，而不僅是被映射到自身之內：把 u 換成 $-u$ 所得的矩陣是 T 的逆矩陣（inverse）（乘開並利用 $c^2 - s^2 = 1$ 即可），而它同樣保持 $x^2 - y^2$ 不變。所以雙曲線（hyperbola）上的每一點都是雙曲線上某一點的像（image）。這條曲線變換為自身。

（事實上每一支都變換為自身：若 $x \geq 1$，則 $|y| < x$，所以 $x' = cx + sy > (c - s)x = \mathrm{e}^{-u}x > 0$。）

通過原點的不變直線

由於 T 是線性(linear)的，它把原點送到原點，並把通過原點、方向為 $\mathbf{d}$ 的直線送到通過原點、方向為 $T\mathbf{d}$ 的直線。所以我們需要 $T\mathbf{d}$ 平行於 $\mathbf{d}$，即 $\mathbf{d}$ 是特徵向量（eigenvector）。試取 $\mathbf{d} = (1, m)$：

$$T\begin{pmatrix}1\\m\end{pmatrix} = \begin{pmatrix}c + sm\\s + cm\end{pmatrix},$$

它平行於 $(1, m)$ 當且僅當 $s + cm = m(c + sm)$，即 $s(1 - m^2) = 0$。由於 $s \neq 0$，故 $m = \pm 1$。（鉛垂方向 $(0, 1)$ 被送到 (s, c)，後者並非鉛垂，所以不會得出新的直線。）這兩條直線是

$$\boxed{y = x \quad \text{及} \quad y = -x}$$

而事實上 $T(1, 1) = \mathrm{e}^{u}(1, 1)$ 及 $T(1, -1) = \mathrm{e}^{-u}(1, -1)$。它們正是雙曲線的漸近線（asymptotes），這與雙曲線映射到自身的結果吻合。

斜率：$v' = v + u$

取一條通過點 $\mathbf{p}$、方向為 $\mathbf{d}$ 的直線，其上的點為 $\mathbf{p} + t\mathbf{d}$。由線性可知它的像是 $T\mathbf{p} + tT\mathbf{d}$：一條通過 $T\mathbf{p}$、方向為 $T\mathbf{d}$ 的直線。所以直線的位置無關重要，只有方向才有影響。

斜率為 $\tanh v$ 的直線，方向為 $(1, \tanh v)$，或者乘以 $\cosh v > 0$ 後，方向為 $(\cosh v, \sinh v)$。於是由加法公式，

$$T\begin{pmatrix}\cosh v\\\sinh v\end{pmatrix} = \begin{pmatrix}\cosh u \cosh v + \sinh u \sinh v\\\sinh u \cosh v + \cosh u \sinh v\end{pmatrix} = \begin{pmatrix}\cosh(u + v)\\\sinh(u + v)\end{pmatrix}$$

因此像直線的斜率為 $\tanh(v + u)$。由於 tanh 是一對一（one-to-one）的，$\tanh v' = \tanh(v + u)$ 給出

$$v' = v + u,$$

這正是所要證明的。（只有斜率嚴格介乎 -1 與 1 之間的直線才可以這樣表示，這就是題目採用 $\tanh v$ 的原因。）

相等的夾角

若兩條直線的斜率為 m_1 和 m_2，且 $1+m_1m_2 \neq 0$，則它們之間的銳角 θ 滿足

$$\tan\theta = \left|\frac{m_1-m_2}{1+m_1m_2}\right|.$$

這裏 $|m_1|, |m_2| < 1$，所以 $1+m_1m_2 > 0$：兩條直線永不互相垂直，$\theta \in [0, \frac{\pi}{2})$，而且 θ 由 $\tan\theta$ 唯一決定。

> **想一想** 這條公式看來很像 $\tanh(v_1-v_2)$，但分母的正負號不對。那麼它實際上化簡成甚麼？

代入 $m_i = \tanh v_i$，並把分子和分母同乘以 $\cosh v_1 \cosh v_2 > 0$：

$$\tan\theta = \frac{|\sinh v_1\cosh v_2 - \cosh v_1 \sinh v_2|}{\cosh v_1\cosh v_2 + \sinh v_1 \sinh v_2} = \frac{|\sinh(v_1-v_2)|}{\cosh(v_1+v_2)}.$$

變換後的直線有 $v_i' = v_i + u$。它們的差不變，但它們的和增加了 $2u$，所以它們之間的夾角 θ' 滿足

$$\tan\theta' = \frac{|\sinh(v_1-v_2)|}{\cosh(v_1+v_2+2u)}.$$

要 ℓ_1 和 ℓ_2 相交，必須有 $v_1 \neq v_2$，所以分子不是零。因此

$$\theta = \theta' \iff \cosh(v_1+v_2) = \cosh(v_1+v_2+2u).$$

由於 $\cosh a = \cosh b$ 當且僅當 $a = \pm b$，這表示要麼 $v_1+v_2 = v_1+v_2+2u$，但因為 $u>0$，這不可能；要麼 $v_1+v_2 = -(v_1+v_2+2u)$。所以充分必要條件（necessary and sufficient condition）是

$$\boxed{v_1+v_2 = -u}$$

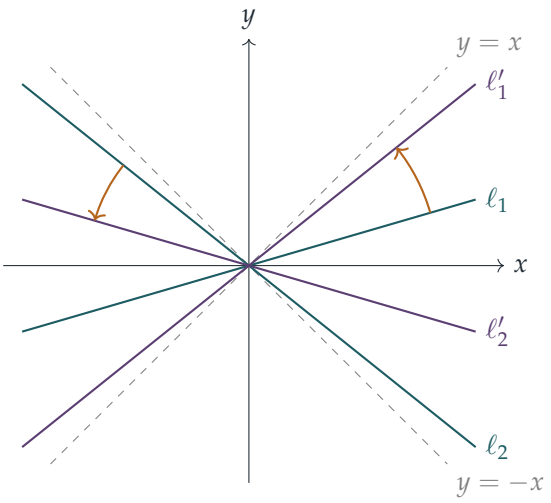
或等價地 $v_1' + v_2' = u$。

兩條直線相交於原點時的圖像

當 $v_1+v_2 = -u$ 時，我們有

$$v_1' = v_1+u = -v_2, \qquad v_2' = v_2+u = -v_1.$$

所以 ℓ_1' 的斜率為 $-\tanh v_2$：它是 ℓ_2 對 x 軸的反射（reflection）。同樣，ℓ_2' 是 ℓ_1 的反射。若兩條直線通過原點，變換後的一對直線就只是原來那一對對 x 軸的鏡像，這就是它們之間的夾角相同的原因。圖中採用 $u=0.8$、$v_1=0.3$、$v_2=-1.1$：ℓ_1, ℓ_2 的斜率約為 0.29 和 -0.80，而 ℓ_1', ℓ_2' 的斜率約為 0.80 和 -0.29。每條直線都轉向 $y=x$，即被 T 拉伸的那條不變直線。



藍綠色的直線是 ℓ_1, ℓ_2，紫紅色的直線是它們的像；琥珀色的箭頭顯示 $\ell_1 \to \ell_1'$ 和 $\ell_2 \to \ell_2'$。紫紅色的一對是藍綠色的一對對 x 軸的反射。

驗算

- 取 $u = 0.8$、$v_1 = 0.3$、$v_2 = -1.1$，ℓ_1 與 ℓ_2 之間的夾角以及它們的像之間的夾角，數值上都是 54.92°。把 v_2 稍作改動，兩個夾角便不再相等。
- 當 $u \to 0$ 時，T 變成恆等變換（identity），條件變成 $v_1 + v_2 = 0$：兩條直線對 x 軸對稱，而恆等映射顯然會保持這種對稱。
- $T(1,1) = \mathrm{e}^u(1,1)$ 很容易用手算驗證，因為 $\cosh u + \sinh u = \mathrm{e}^u$。

常見錯誤

- 只證明了雙曲線映射到自身之內；要用逆變換才能說明沒有遺漏任何一點。
- 把夾角公式寫成 $\tanh(v_1 - v_2)$；$1 + m_1 m_2$ 中的 $+$ 號使分母變成 $\cosh(v_1 + v_2)$，而在 T 之下改變的正是這個和。
- 忘記了 $\cosh a = \cosh b$ 的根 $v_1 + v_2 = v_1 + v_2 + 2u$，又沒有說明它為何不可能。

STEP 之外

這個矩陣是狹義相對論（special relativity）中的洛倫茲變換（Lorentz boost），其中 x 和 y 分別扮演時間和空間的角色（採用光速為 1 的單位）。$x^2 - y^2$ 這個量是不變間隔

（invariant interval），$\tanh v$ 是速度，而 v 是它的快度（rapidity）。$v' = v + u$ 這個結果說明快度可以相加，這其實就是相對論速度加法公式的另一種面貌。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）就是以這種方式建立狹義相對論的。至於以特徵向量的觀點看不變直線，可參閱 Strang 的 *Introduction to Linear Algebra*（線性代數入門）。

詳細解答

STEP III 1991，第 7 題

純數學 • 利用對稱性積分

> **STEP III 1991，第 7 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s3-q7

關鍵思路

$\ln(\sin x)$ 的積分無法透過原函數（antiderivative）求出，但對稱性（symmetry）足以解決一切。代換（substitution）$x \mapsto \frac{1}{2}\pi - x$ 會把 sin 與 cos 互換，而 $x \mapsto \pi - x$ 則說明 $\sin x$ 在 $[0, \pi]$ 上的取值，正是它在 $[0, \frac{1}{2}\pi]$ 上的取值重複兩次。再寫出 $\sin x \cos x = \frac{1}{2}\sin 2x$，便得到一條未知積分必須滿足的方程。第 (iii) 部分是一次分部積分（integration by parts），把 $x \cot x$ 化為 $\ln(\sin x)$，而第 (ii) 部分則用來處理邊界項（boundary term）。

解答

全題記

$$I = \int_0^{\frac{1}{2}\pi} \ln(\sin x)\,\mathrm{d}x.$$

第 (i) 部分

正弦與餘弦的積分相等。代入 $x = \frac{1}{2}\pi - t$，則 $\mathrm{d}x = -\mathrm{d}t$ 且 $\sin x = \cos t$。積分上下限 $0, \frac{1}{2}\pi$ 變成 $\frac{1}{2}\pi, 0$，而負號又把它們調換回來：

$$I = \int_{\frac{1}{2}\pi}^{0} \ln(\cos t)\,(-\mathrm{d}t) = \int_0^{\frac{1}{2}\pi} \ln(\cos t)\,\mathrm{d}t.$$

與 $\sin 2x$ 的關係。把這兩個相等的積分相加，並利用 $\ln a + \ln b = \ln(ab)$（由於在開區間

上 $\sin x, \cos x > 0$，此式成立）：

$$\begin{aligned} 2I &= \int_0^{\frac{1}{2}\pi} \ln(\sin x \cos x)\,\mathrm{d}x = \int_0^{\frac{1}{2}\pi} \ln\left(\tfrac{1}{2}\sin 2x\right)\mathrm{d}x \\ &= \int_0^{\frac{1}{2}\pi} \ln(\sin 2x)\,\mathrm{d}x - \tfrac{1}{2}\pi \ln 2. \end{aligned}$$

兩邊除以二，得 $I = \frac{1}{2}\int_0^{\frac{1}{2}\pi} \ln(\sin 2x)\,\mathrm{d}x - \frac{1}{4}\pi \ln 2$，第一串等式至此證畢。

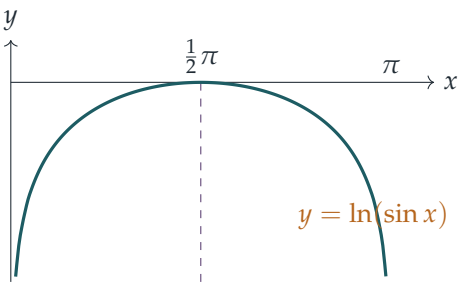
第二串等式。設 $u = 2x$，則 $\mathrm{d}x = \frac{1}{2}\,\mathrm{d}u$，而積分上下限變成 0 和 π：

$$\int_0^{\frac{1}{2}\pi} \ln(\sin 2x)\,\mathrm{d}x = \tfrac{1}{2}\int_0^{\pi} \ln(\sin u)\,\mathrm{d}u.$$

現在把 $[0, \pi]$ 在 $\frac{1}{2}\pi$ 處分成兩段。在後半段代入 $u = \pi - v$。由於 $\sin(\pi - v) = \sin v$，而積分上下限 $\frac{1}{2}\pi, \pi$ 變成 $\frac{1}{2}\pi, 0$，

$$\int_{\frac{1}{2}\pi}^{\pi} \ln(\sin u)\,\mathrm{d}u = \int_0^{\frac{1}{2}\pi} \ln(\sin v)\,\mathrm{d}v = I.$$

下圖顯示了這種對稱性：曲線 $y = \ln(\sin x)$ 關於 $x = \frac{1}{2}\pi$ 對稱。



所以 $\int_0^{\pi} \ln(\sin u)\,\mathrm{d}u = 2I$，因此

$$\int_0^{\frac{1}{2}\pi} \ln(\sin 2x)\,\mathrm{d}x = \tfrac{1}{2}\int_0^{\pi} \ln(\sin x)\,\mathrm{d}x = \tfrac{1}{2}(2I) = I,$$

正如所求。

計算 I。把第二個結果代入第一個結果，得

$$I = \tfrac{1}{2}I - \tfrac{1}{4}\pi \ln 2 \quad \Longrightarrow \quad I = \int_0^{\frac{1}{2}\pi} \ln(\sin x)\,\mathrm{d}x = -\tfrac{1}{2}\pi \ln 2.$$

這一步需要 I 是一個有限的數，這正是題目容許我們假設積分收斂（convergence）的原因。若沒有這個假設，「$I = \frac{1}{2}I + c$」甚麼也證明不了。

第 (ii) 部分

設 $x \geq 1$，並取 $u = \sqrt{x}$。則 $u \geq 1$，所以題目給出的不等式適用：

$$\ln \sqrt{x} < \sqrt{x}, \qquad \text{即} \qquad \tfrac{1}{2} \ln x < \sqrt{x}.$$

第一個極限（limit）。當 $x \geq 1$ 時，我們有 $\ln x \geq 0$，而把 $\ln x < 2\sqrt{x}$ 除以 $x > 0$ 可得

$$0 \leq \frac{\ln x}{x} < \frac{2}{\sqrt{x}}.$$

當 $x \to \infty$ 時，右邊趨向 0，所以根據夾逼原理（squeeze principle，又稱 sandwich principle），$\dfrac{\ln x}{x} \to 0$。

第二個極限。當 $0 < x < 1$ 時，設 $x = 1/t$，則 $t > 1$，且當 $x \to 0^+$ 時 $t \to \infty$。於是根據第一個極限，

$$x \ln x = \frac{1}{t} \ln \frac{1}{t} = -\frac{\ln t}{t} \to 0$$

。所以當 $x \to 0$（只取正值）時，$x \ln x \to 0$。

第 (iii) 部分

首先留意被積函數（integrand）在 0 處並無問題：$x \cot x = \dfrac{x}{\sin x} \cos x \to 1$。由於在 $(0, \frac{1}{2}\pi)$ 上 $\cot x$ 是 $\ln(\sin x)$ 的導數，我們取 $u = x$ 和 $\mathrm{d}v = \cot x \, \mathrm{d}x$，$v = \ln(\sin x)$ 作分部積分。對數在 0 處趨向無窮，所以我們先在 $[\varepsilon, \frac{1}{2}\pi]$ 上計算，然後令 $\varepsilon \to 0^+$：

$$\int_\varepsilon^{\frac{1}{2}\pi} x \cot x \, \mathrm{d}x = \Big[x \ln(\sin x) \Big]_\varepsilon^{\frac{1}{2}\pi} - \int_\varepsilon^{\frac{1}{2}\pi} \ln(\sin x) \, \mathrm{d}x.$$

在上限處，$\ln(\sin \frac{1}{2}\pi) = \ln 1 = 0$。所以邊界項只剩下 $-\varepsilon \ln(\sin \varepsilon)$。

> **想一想** 為甚麼 $\varepsilon \ln(\sin \varepsilon)$ 趨向 0？第 (ii) 部分討論的是 $x \ln x$，而不是 $x \ln(\sin x)$。

記 $s = \sin \varepsilon$，當 $\varepsilon \to 0^+$ 時，它取正值而趨向 0。於是

$$\varepsilon \ln(\sin \varepsilon) = \frac{\varepsilon}{\sin \varepsilon} \cdot s \ln s.$$

第一個因子趨向 1（即標準極限 $\sin \varepsilon / \varepsilon \to 1$），而根據第 (ii) 部分，$s \ln s \to 0$。所以邊界項趨向 0。

令 $\varepsilon \to 0^+$，並利用第 (i) 部分：

$$\int_0^{\frac{1}{2}\pi} x \cot x \, \mathrm{d}x = 0 - I = \tfrac{1}{2}\pi \ln 2.$$

驗算

- 正負號：在該區間上 $\ln(\sin x) \le 0$，所以 I 必為負數，而 $-\frac{1}{2}\pi \ln 2 \approx -1.0888$ 的確是負數。另一方面，在 $(0, \frac{1}{2}\pi)$ 上 $x\cot x > 0$，而我們的答案是正數。
- 大小：$x\cot x$ 在長度為 $\frac{1}{2}\pi \approx 1.571$ 的區間上由 1 下降至 0，所以積分值約為 1.09 是合理的。
- 數值上，以中點求積法（midpoint quadrature）計算可得 $\int_0^{\frac{1}{2}\pi} x\cot x\,\mathrm{d}x \approx 1.0887930$，與 $\frac{1}{2}\pi\ln 2 = 1.0887930\ldots$ 吻合，並且 $\int_0^{\frac{1}{2}\pi} \ln(\sin x)\,\mathrm{d}x \approx -1.08879$。

常見錯誤

- 轉換為 $u = 2x$ 時，漏掉由 $\mathrm{d}x = \frac{1}{2}\,\mathrm{d}u$ 而來的因子 $\frac{1}{2}$，結果得出 $\int_0^{\frac{1}{2}\pi} \ln(\sin 2x)\,\mathrm{d}x = 2I$，之後便無從入手。
- 在第 (iii) 部分，直接寫 $[x\ln(\sin x)]_0^{\frac{1}{2}\pi} = 0$，卻沒有論證在 0 處的極限。第 (ii) 部分正是為了求這個極限而設。

STEP 之外

這裏的兩個積分都是*瑕積分*（improper integral）：它們的被積函數，或所用的原函數，在某個端點處無界。這類積分是以極限來定義的，正如我們利用 ε 所做的那樣。大學第一門分析（analysis）課程會把這一點嚴格化，並給出判別這類極限何時存在的方法。題目中「可假設」那一句的背後，正是這些判別法。同樣的對稱技巧還可以求出許多其他積分，例如 $\int_0^{\pi/2} \ln(\tan x)\,\mathrm{d}x = 0$。Abbott 的 *Understanding Analysis* 講述了支撐這一切的極限與積分理論。

詳細解答

STEP III 1991，第 8 題

純數學 • 歸約公式與求和

STEP III 1991，第 8 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q8

關鍵思路

第 (i) 部分是一條歸約公式（reduction formula），所以分部積分法（integration by parts）自然是第一步。餘下的被積函數（integrand）含有 $\sin x \sin kx$，而 $\cos(k-1)x$ 的複角公式（compound-angle formula）可把它變回同一族的積分。第 (ii) 部分是 $(1+\mathrm{e}^{2\mathrm{i}\theta})^m$ 的二項展開式（binomial expansion）的實部（real part）。第 (iii) 部分則是把 (ii) 積分，再以裂項相消（telescoping）的方法把 (i) 的遞推關係（recurrence）逐層展開而得出。

解答

第 (i) 部分

取 $k \geq 1$，並以 $u = \cos^k x$ 和 $\mathrm{d}v = \cos kx\,\mathrm{d}x$ 作分部積分，於是 $v = \frac{1}{k}\sin kx$ 及 $\mathrm{d}u = -k\cos^{k-1}x \sin x\,\mathrm{d}x$：

$$\begin{aligned} I_k &= \left[\tfrac{1}{k}\cos^k x \sin kx\right]_0^\theta + \int_0^\theta \cos^{k-1}x\,\sin x\,\sin kx\,\mathrm{d}x \\ &= \tfrac{1}{k}\cos^k\theta\sin k\theta + \int_0^\theta \cos^{k-1}x\,\sin x\,\sin kx\,\mathrm{d}x, \end{aligned}$$

因為方括號內的式子在 $x = 0$ 時等於零。

? **想一想** 怎樣把 $\sin x \sin kx$ 改寫成 x 的倍數的餘弦？

把 $(k-1)x = kx - x$ 代入複角公式，得

$$\cos(k-1)x = \cos kx\cos x + \sin kx \sin x,$$

所以

$$\sin x \sin kx = \cos(k-1)x - \cos x \cos kx.$$

因此

$$\begin{aligned}\int_0^\theta \cos^{k-1} x \, \sin x \, \sin kx \, \mathrm{d}x &= \int_0^\theta \cos^{k-1} x \cos(k-1)x \, \mathrm{d}x \\ &\quad - \int_0^\theta \cos^k x \cos kx \, \mathrm{d}x \\ &= I_{k-1} - I_k.\end{aligned}$$

代回原式並乘以 k，

$$kI_k = \cos^k \theta \sin k\theta + kI_{k-1} - kI_k,$$

即

$$2kI_k = kI_{k-1} + \cos^k \theta \sin k\theta,$$

這正是所要證明的結果。

第 (ii) 部分

設 $z = \mathrm{e}^{2\mathrm{i}\theta}$。根據二項式定理（binomial theorem），

$$\begin{aligned}(1+z)^m &= \sum_{r=0}^{m} \binom{m}{r} z^r \\ &= \sum_{r=0}^{m} \binom{m}{r} \big(\cos 2r\theta + \mathrm{i} \sin 2r\theta\big).\end{aligned}$$

現在把「半角」抽出來：

$$1 + \mathrm{e}^{2\mathrm{i}\theta} = \mathrm{e}^{\mathrm{i}\theta}\big(\mathrm{e}^{-\mathrm{i}\theta} + \mathrm{e}^{\mathrm{i}\theta}\big) = 2\cos\theta \, \mathrm{e}^{\mathrm{i}\theta},$$

所以

$$(1+z)^m = 2^m \cos^m \theta \, \mathrm{e}^{\mathrm{i}m\theta} = 2^m \cos^m \theta \, (\cos m\theta + \mathrm{i} \sin m\theta).$$

由於 $2^m \cos^m \theta$ 是實數，比較兩個表達式的實部，得

$$\sum_{r=0}^{m} \binom{m}{r} \cos 2r\theta = 2^m \cos^m \theta \cos m\theta,$$

這就是所求的恆等式（$r = 0$ 和 $r = m$ 兩項分別是 1 和 $\cos 2m\theta$）。

第 (iii) 部分

把第一個表達式記作 S。(ii) 的右邊是 I_m 的被積函數的 2^m 倍，所以我們把 (ii) 由 0 積分至 θ。$r = 0$ 一項積分後得 θ，而當 $r \geq 1$ 時，$\int_0^\theta \cos 2rx \, \mathrm{d}x = \frac{\sin 2r\theta}{2r}$。因此

$$\theta + \sum_{r=1}^{m} \binom{m}{r} \frac{\sin 2r\theta}{2r} = 2^m I_m, \qquad \text{所以} \qquad S = 2^m I_m - \theta.$$

現在利用 (i) 計算 $2^m I_m$。把遞推關係除以 $2k$ 再乘以 2^k，對 $k \geq 1$ 得

$$2^k I_k = 2^{k-1} I_{k-1} + \frac{2^{k-1}}{k} \cos^k \theta \sin k\theta.$$

想一想 為甚麼要乘以 2^k？

若寫成 $I_k = \frac{1}{2} I_{k-1} + \ldots$，遞推關係每一步都帶有因子 $\frac{1}{2}$。改用 $J_k = 2^k I_k$ 便可把它消去：這樣 $J_k - J_{k-1}$ 就是一個簡單的顯式項，而求和時會裂項相消。把 $k = 1, 2, \ldots, m$ 的各條等式相加，中間各項互相抵消：

$$2^m I_m = 2^0 I_0 + \sum_{k=1}^{m} \frac{2^{k-1}}{k} \cos^k \theta \sin k\theta.$$

最後，$I_0 = \int_0^\theta 1 \, \mathrm{d}x = \theta$，所以

$$\begin{aligned} S = 2^m I_m - \theta &= \sum_{k=1}^{m} \frac{1}{k} 2^{k-1} \cos^k \theta \sin k\theta \\ &= \cos\theta \sin\theta + \cos^2\theta \sin 2\theta + \cdots + \frac{1}{m} 2^{m-1} \cos^m \theta \sin m\theta, \end{aligned}$$

這就是所求的結果。

驗算

- $m = 1$：所要證明的是 $\frac{1}{2} \sin 2\theta = \cos\theta \sin\theta$，這正是二倍角公式（double-angle formula）。
- $m = 2$：左邊是 $\sin 2\theta + \frac{1}{4} \sin 4\theta$，而右邊是 $\cos\theta \sin\theta + \cos^2\theta \sin 2\theta = \frac{1}{2} \sin 2\theta + \frac{1}{2} \sin 2\theta (1 + \cos 2\theta)$，兩者相同，因為 $\sin 2\theta \cos 2\theta = \frac{1}{2} \sin 4\theta$。
- 數值上，取 $m = 5$ 和 $\theta = 0.7$，兩個表達式都等於 1.51963（準確至小數點後 5 位）。對隨機選取的 k 和 θ，歸約公式亦與用辛普森法則（Simpson's rule）算出的 I_k 值吻合。

常見錯誤

- 分部積分時弄錯正負號：$\frac{\mathrm{d}}{\mathrm{d}x} \cos^k x = -k \cos^{k-1} x \sin x$，而這個負號會與 $\int v \, \mathrm{d}u$ 前面的負號互相抵消。
- 忘記 (ii) 中 $r = 0$ 的一項。它積分後得 θ，並與裂項求和最底端的 $I_0 = \theta$ 互相抵消。

STEP 之外

把 $\cos^m\theta\cos m\theta$ 寫成 $\cos 2r\theta$ 的組合，其實就是計算一個傅立葉級數（Fourier series）。這裏的級數只有有限項，而逐項積分正是大學一年級數學方法課程中處理傅立葉級數的做法。(i) 中那類歸約公式，也是處理 $\int \cos^n x\,\mathrm{d}x$ 這種形式的積分以及沃利斯乘積（Wallis product）的標準工具。Needham 的 *Visual Complex Analysis* 為 (ii) 中用到的技巧 $1+\mathrm{e}^{2\mathrm{i}\theta}=2\cos\theta\,\mathrm{e}^{\mathrm{i}\theta}$ 提供了一幅幾何圖像。

詳細解答

STEP III 1991，第 9 題

純數學 • 橢圓的弧長

> **STEP III 1991，第 9 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s3-q9

關鍵思路

E_1 是以中心為原點的橢圓（ellipse）方程，而 E_2 是同一條曲線的極坐標（polar）形式，極點位於一個焦點（focus）。把 E_2 化為直角坐標（Cartesian）形式再配方（completing the square），便可看出它的中心和軸在哪裏；比較兩者的半軸（semi-axes）即可求得 k 和 b。至於積分，周界（perimeter）只是一個數，但每一種參數化（parametrisation）都會給出外表不同的弧長（arc length）積分。

解答

坐標軸、k 和 b

以 (x,y) 表示 E_2 的坐標，並設 $r=\sqrt{x^2+y^2}$，$\theta=\theta_2$。於是 E_2 即 $r=\dfrac{k}{1+e\cos\theta}$，亦即 $r+e\,r\cos\theta=k$，所以

$$r=k-ex.$$

兩邊平方（這一步沒有損失任何資料，因為在曲線上 $r=k/(1+e\cos\theta)>0$），

$$x^2+y^2=k^2-2kex+e^2x^2 \quad\Longrightarrow\quad (1-e^2)x^2+2kex+y^2=k^2.$$

對 x 配方：

$$(1-e^2)\left(x+\frac{ke}{1-e^2}\right)^2+y^2=k^2+\frac{k^2e^2}{1-e^2}=\frac{k^2}{1-e^2},$$

所以

$$\frac{\left(x+\frac{ke}{1-e^2}\right)^2}{k^2/(1-e^2)^2}+\frac{y^2}{k^2/(1-e^2)}=1.$$

這是一個橢圓，在 E_2 坐標軸下中心為 $\left(-\frac{ke}{1-e^2},0\right)$，對稱軸與 E_2 的坐標軸平行，半軸為

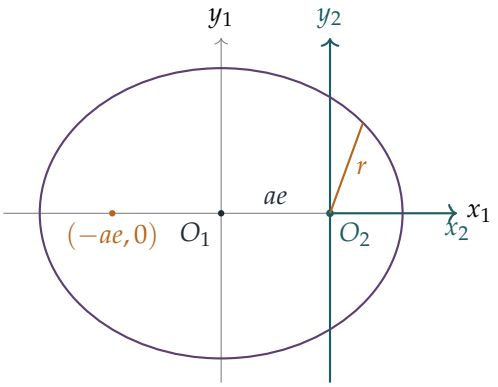
$$A=\frac{k}{1-e^2}\text{（沿 }x\text{ 方向）}, \qquad B=\frac{k}{\sqrt{1-e^2}}\text{（沿 }y\text{ 方向）}.$$

由於 $0<1-e^2<1$，我們有 $A>B$，所以長軸（major axis）位於 E_2 的 x 軸上。對於 E_1，$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 且 $a>b$，所以它的長軸位於 E_1 的 x 軸上。兩條曲線是同一條曲線，所以半長軸相等，半短軸也相等：

$$a=\frac{k}{1-e^2}, \quad \text{即} \quad k=a(1-e^2), \qquad b^2=\frac{k^2}{1-e^2}=\frac{a^2(1-e^2)^2}{1-e^2}=a^2(1-e^2),$$

這正是所需的兩個結果。

中心（即 E_1 的原點）位於 E_2 的 x 軸上 $x=-\frac{ke}{1-e^2}=-ae$ 之處。所以 E_2 的坐標軸就是把 E_1 的坐標軸沿長軸平移距離 ae 而得：E_2 的原點是 E_1 中的點 $(ae,0)$，而由於 $a^2-b^2=a^2e^2$，這一點正是焦點。兩條 x 軸位於同一條直線（長軸）上，兩條 y 軸互相平行；以 E_1 的坐標表示，$x_1=x_2+ae$，$y_1=y_2$。橢圓對它的兩條軸都對稱，所以我們同樣可以把 E_2 的原點取在另一個焦點 $(-ae,0)$，並把 x 軸反向。無論哪一種取法，$\theta_2=0$ 所對應的點都是最接近該焦點的頂點（vertex）。



兩個積分

我們用兩種方法計算橢圓上半部分（$y\geq 0$）的長度。

利用 E_1。當 θ_1 由 0 變到 π 時，動點剛好走遍上半部分一次。由 $\dot{x}=-a\sin\theta_1$，$\dot{y}=b\cos\theta_1$，

$$\dot{x}^2+\dot{y}^2=a^2\sin^2\theta_1+b^2\cos^2\theta_1=a^2-(a^2-b^2)\cos^2\theta_1=a^2\left(1-e^2\cos^2\theta_1\right),$$

其中用到 $a^2-b^2=a^2e^2$。所以半周界為

$$\tfrac{1}{2}P=a\int_0^{\pi}\sqrt{1-e^2\cos^2\theta}\,\mathrm{d}\theta.$$

利用 E_2。這裏由於 $r>0$，$y_2 = r\sin\theta_2 \geq 0$ 當且僅當 $0 \leq \theta_2 \leq \pi$。所以 $\theta_2 \in [0,\pi]$ 同樣剛好描出上半部分一次。在極坐標中，$\left(\frac{\mathrm{d}s}{\mathrm{d}\theta}\right)^2 = r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2$。由 $r = k(1+e\cos\theta)^{-1}$，

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{ke\sin\theta}{(1+e\cos\theta)^2},$$

所以

$$r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 = \frac{k^2\left[(1+e\cos\theta)^2 + e^2\sin^2\theta\right]}{(1+e\cos\theta)^4} = \frac{k^2\left(1+e^2+2e\cos\theta\right)}{(1+e\cos\theta)^4}.$$

因此，利用 $k = a(1-e^2)$，

$$\tfrac{1}{2}P = a(1-e^2)\int_0^\pi \frac{\sqrt{1+e^2+2e\cos\theta}}{(1+e\cos\theta)^2}\,\mathrm{d}\theta.$$

令 $\frac{1}{2}P$ 的兩個表達式相等，再除以 $a>0$，便得到所需的恆等式（identity）。

想一想 為甚麼兩個參數都必須取遍 $[0,\pi]$，而不是其中一個取另一個範圍？

兩個弧長必須是同一段曲線的長度。θ_1 和 θ_2 是不同的角（一個是從中心量度的離心角（eccentric angle），另一個是從焦點量度的極角），但兩者在 $[0,\pi]$ 上都是由右頂點 $(a,0)$ 到左頂點 $(-a,0)$ 掃過上半部分。所以兩個積分量度的是同一段弧。

近似值

由於 $0 \leq e^2\cos^2\theta < 1$，二項級數（binomial series）適用：

$$\sqrt{1-u} = 1 - \tfrac{1}{2}u - \tfrac{1}{8}u^2 + O(u^3), \qquad u = e^2\cos^2\theta.$$

$O(u^3)$ 項是 $O(e^6)$，我們把它略去。利用

$$\int_0^\pi \cos^2\theta\,\mathrm{d}\theta = \frac{\pi}{2}, \qquad \int_0^\pi \cos^4\theta\,\mathrm{d}\theta = \frac{3\pi}{8}$$

（第二個由 $\cos^4\theta = \frac{3}{8} + \frac{1}{2}\cos 2\theta + \frac{1}{8}\cos 4\theta$ 得出），我們得到

$$\begin{aligned}\int_0^\pi \sqrt{1-e^2\cos^2\theta}\,\mathrm{d}\theta &\approx \pi - \frac{e^2}{2}\cdot\frac{\pi}{2} - \frac{e^4}{8}\cdot\frac{3\pi}{8} = \pi\left(1 - \frac{e^2}{4} - \frac{3e^4}{64}\right)\\ &= \tfrac{1}{64}\pi\left(64 - 16e^2 - 3e^4\right).\end{aligned}$$

根據剛證明的恆等式，第二個積分的值相同。至此全題完成。

驗算

► 當 $e=0$ 時，兩個積分都等於 π，而 E_2 變成圓 $r=k$，其中 $k=a$：正確。

- 取 $a = 2$，$b = 1$（所以 $e^2 = \frac{3}{4}$），用一個 200 000 邊形計算，得周界為 9.6884482，與 $2a$ 乘以第一個積分的值吻合。
- 數值上，當 $e = 0.2$ 時兩個積分都是 3.1099371，而近似式給出 3.1099411。兩者之差約為 $-0.063e^6$，與級數的下一項 $-\frac{5\pi}{256}e^6 \approx -0.061e^6$ 相符。

常見錯誤

- 把 E_2 的原點放在中心。分母為 $1 + e\cos\theta$ 的極坐標形式，極點必定在焦點。
- 在極坐標弧長公式中漏掉 $(\mathrm{d}r/\mathrm{d}\theta)^2$ 一項，或者一邊在 $[0, 2\pi]$ 上積分，另一邊卻在 $[0, \pi]$ 上積分。

STEP 之外

積分 $\int_0^{\pi/2} \sqrt{1 - e^2\sin^2\theta}\,\mathrm{d}\theta$ 是第二類完全橢圓積分（complete elliptic integral of the second kind）$E(e)$，所以橢圓的周界是 $4aE(e)$。它不能用初等函數（elementary functions）表示，這正是本題這類級數（以及拉馬努金（Ramanujan）那些準確得驚人的近似公式）之所以重要的原因。由橢圓積分再走下去，便是橢圓函數（elliptic functions）和橢圓曲線（elliptic curves），後者是現代數論的核心課題。想做更多這種風格的題目，可參閱 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）。

詳細解答

STEP III 1991，第 10 題

純數學 • 牛頓恆等式

> **STEP III 1991，第 10 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s3-q10

關鍵思路

這個多項式只有三個非零係數，所以對稱函數（symmetric function）Σ_m 幾乎全部為零，而題目給出的恆等式（牛頓恆等式，Newton's identity）就簡化成一個只有兩項的遞推關係（recurrence）。次數超過 n 之後，這個恆等式不再適用，但有一個更簡單的做法：每個根都滿足方程本身，所以把方程乘以 α^k 再對所有根求和，便可把 S_{n+k} 與較早的冪和（power sum）聯繫起來。

解答

直到 S_{n-1} 的冪和

把 $x^n - qx^{n-1} + r$ 與 $\prod_i (x - \alpha_i) = x^n - \Sigma_1 x^{n-1} + \Sigma_2 x^{n-2} - \cdots + (-1)^n \Sigma_n$ 比較，可得

$$\Sigma_1 = q, \qquad \Sigma_2 = \Sigma_3 = \cdots = \Sigma_{n-1} = 0, \qquad (-1)^n \Sigma_n = r.$$

在題目給出的恆等式中取 $1 \le p \le n-1$。由於 $p \le n-1$，所有含有 $\Sigma_2, \ldots, \Sigma_p$ 的項都等於零。（當 $p = 1$ 時，恆等式就是 $S_1 - \Sigma_1 = 0$。）剩下的是

$$S_p - q\, S_{p-1} = 0 \quad (2 \le p \le n-1), \qquad S_1 = q.$$

所以 $S_1 = q$，而每個冪和都是前一個的 q 倍。對 p 用歸納法（induction），得

$$S_p = q^p \qquad \text{對於} 1 \le p \le n-1,$$

這正是所要證明的結果。另外，由於每個 $\alpha_i^0 = 1$，順帶記下 $S_0 = n$ 會很方便。

S_n、S_{n+1} 和 S_{n+2}

當 $p = n$ 時，恆等式仍然適用，而這次最後一項會出現：

$$S_n - q\,S_{n-1} + (-1)^n n\,\Sigma_n = 0.$$

由於 $(-1)^n\Sigma_n = r$，由此得

$$S_n = q \cdot q^{n-1} - nr = q^n - nr.$$

想一想 題目給出的恆等式到 $p = n$ 就停了。怎樣才能再走下去？

我們利用方程本身。每個根都滿足 $\alpha_i^n = q\alpha_i^{n-1} - r$。乘以 α_i^k（其中 $k \geq 0$）再對 i 求和，得

$$S_{n+k} = q\,S_{n+k-1} - r\,S_k \qquad (k \geq 0). \tag{$*$}$$

取 $k = 0$，這再次給出 $S_n = q^n - nr$，是一個令人安心的交叉驗證。現在

$$\begin{aligned} S_{n+1} &= q\,S_n - r\,S_1 = q^{n+1} - nqr - qr = q^{n+1} - (n+1)qr, \\ S_{n+2} &= q\,S_{n+1} - r\,S_2 = q^{n+2} - (n+1)q^2r - q^2r = q^{n+2} - (n+2)q^2r. \end{aligned}$$

這裏用到了 $S_1 = q$ 和 $S_2 = q^2$；它們成立，是因為 $n \geq 5$ 保證了 $2 \leq n-1$。

一般公式

從 $S_n = q^n - nr$、$S_{n+1} = q^{n+1} - (n+1)qr$、$S_{n+2} = q^{n+2} - (n+2)q^2r$ 這個規律，可以猜想

$$S_{n+m} = q^{n+m} - (n+m)\,q^m r \qquad \text{對於} 0 \leq m \leq n-1.$$

我們對 m 用歸納法來證明。$m = 0$ 的情況就是 $S_n = q^n - nr$，上面已經證明。

假設公式對某個 $m-1$ 成立，其中 $1 \leq m \leq n-1$。在 $(*)$ 中取 $k = m$，得

$$\begin{aligned} S_{n+m} &= q\,S_{n+m-1} - r\,S_m \\ &= q\big(q^{n+m-1} - (n+m-1)q^{m-1}r\big) - r\,q^m \\ &= q^{n+m} - (n+m)\,q^m r, \end{aligned}$$

其中用到了 $S_m = q^m$，這是因為 $1 \leq m \leq n-1$ 才成立。這就是 m 的情況下的公式。根據歸納法，公式對所有 $0 \leq m \leq n-1$ 成立，即對 $m < n$ 成立。

想一想 條件 $m < n$ 究竟在哪裏用到？

只在 $S_m = q^m$ 這一步。當 $m = n$ 時，我們需要的是 S_n，而它等於 $q^n - nr$，不是 q^n，於是多出來的 $-nr$ 會帶進計算之中，規律就此打破。事實上，由 $(*)$ 可得 $S_{2n} = q^{2n} - 2nq^nr + nr^2$。

驗算

- 當 $r = 0$ 時，方程是 $x^{n-1}(x-q) = 0$，根為 q 和 0（重複 $n-1$ 次），所以每個 $S_k = q^k$。我們所有的公式都化為這個結果。
- 當 $q = 0$ 時，各根就是 $-r$ 的 n 次方根。於是對於 $1 \le p \le n-1$ 有 $S_p = 0$，而 $S_n = n \cdot (-r) = -nr$，與 $q^n - nr$ 吻合。
- 數值上，我們對隨機選取的 q, r 和 $5 \le n \le 9$ 求出所有根，再把它們的冪加起來。$S_p = q^p$ 以及 S_{n+m} 的公式（$0 \le m \le n-1$）都吻合至約 10^{-12} 的誤差，而在 $m = n$ 時公式明顯不成立。

常見錯誤

- 忘記恆等式最後一項的因子 p，結果得出 $S_n = q^n - r$ 而不是 $q^n - nr$。
- 弄錯 Σ_n 的正負號：常數項是 $(-1)^n \Sigma_n$，所以不論 n 是奇是偶，都有 $(-1)^n n\Sigma_n = nr$。
- 在 $p > n$ 時套用題目給出的恆等式，而照題目所寫的形式，它在這個範圍並不成立。

STEP 之外

這些就是牛頓恆等式，而 $(*)$ 表明各冪和滿足一個線性遞推關係（linear recurrence），其特徵多項式（characteristic polynomial）正是原來的方程。用矩陣（matrix）的語言來說，S_k 是 A^k 的跡（trace），其中 A 是該多項式的友矩陣（companion matrix）。各根的任何對稱多項式（symmetric polynomial）都可以用 Σ_m 表示，這個事實就是對稱多項式基本定理（fundamental theorem of symmetric polynomials），由此可以通往伽羅瓦理論（Galois theory）。Allenby 的 *Rings, Fields and Groups*（環、域與羣）是了解這套代數如何展開的好去處。

詳細解答

STEP III 1991，第 11 題

力學 • 剛體平衡

STEP III 1991，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

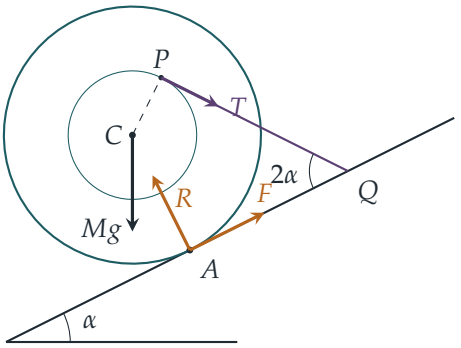
step.maths.org/questions/91-s3-q11

關鍵思路

圓柱體是一個剛體（rigid body），在四個力的作用下處於平衡（equilibrium）：重力、法向反作用力(normal reaction)、接觸點處的摩擦力(friction)，以及繩子的張力(tension)。其中重力和法向反作用力都通過圓柱體的中心，所以對中心取力矩（moments），便可把摩擦力與張力直接聯繫起來。再沿斜面方向和垂直於斜面方向分解力，就得到摩擦力和法向反作用力，餘下的工作交給不等式 $F \leq \mu R$ 即可。在第二部分，圓柱體並不「知道」是甚麼東西拉住繩子，所以張力不變，我們只需考慮質點。

解答

我們把圓柱體看作均勻剛體，其質心（centre of mass）就是它的中心 C；把繩子看作輕而不可伸長（light and inextensible）；把第二個物體看作質點（particle）。設 A 為圓柱體與斜面的接觸點，並設繩子的張力為 T。由於繩子繞在凹槽上，自由的一段 PQ 在 P 點與凹槽相切，所以它的作用線與 C 的垂直距離為 a。由圖可見，Q 在斜面上的位置比 A 更高，而繩子把圓柱體由 P 拉向 Q：既沿斜面向上，也壓向斜面，與斜面成 2α 角。



虛線所示的半徑 CP 長度為 a，並與繩子垂直。我們以 F 表示作用於圓柱體的摩擦力，以沿斜面向上為正。

圓柱體：證明 $\mu \geq \frac{1}{3}\tan\alpha$

對 C 取力矩。重力作用於 C，法向反作用力沿 AC 作用，所以兩者對 C 都沒有力矩。摩擦力的作用點與 C 相距 $2a$，張力的作用線與之垂直距離為 a，兩者使圓柱體朝相反方向轉動，所以

$$F \cdot 2a = T \cdot a, \qquad \text{即} \qquad T = 2F.$$

想一想 為甚麼摩擦力和張力會使圓柱體朝相反方向轉動？

只有這兩個力對 C 有力矩，所以它們的力矩必須互相抵消。單靠繩子在 P 點的拉力，圓柱體會轉動而沿斜面向上滾。A 點處沿斜面向上的摩擦力則使它朝另一方向轉動。假如我們猜錯了摩擦力的方向，下面的方程只會給出 $F < 0$，所以這個正負號並不是額外的假設。

沿斜面方向分解（以沿斜面向上為正）：

$$F + T\cos 2\alpha = Mg\sin\alpha \quad \Longrightarrow \quad F(1+2\cos 2\alpha) = Mg\sin\alpha.$$

垂直於斜面方向分解：張力有一個壓向斜面的分量 $T\sin 2\alpha$，所以

$$R = Mg\cos\alpha + T\sin 2\alpha = Mg\cos\alpha + 2F\sin 2\alpha.$$

代入 $F = Mg\sin\alpha/(1+2\cos 2\alpha)$，

$$\begin{aligned} R(1+2\cos 2\alpha) &= Mg\big(\cos\alpha + 2\cos 2\alpha\cos\alpha + 2\sin 2\alpha\sin\alpha\big) \\ &= Mg\big(\cos\alpha + 2\cos(2\alpha-\alpha)\big) = 3Mg\cos\alpha. \end{aligned}$$

因此

$$F = \frac{Mg\sin\alpha}{1+2\cos 2\alpha}, \qquad R = \frac{3Mg\cos\alpha}{1+2\cos 2\alpha}.$$

圓柱體保持與斜面接觸，所以 $R > 0$。由於 α 是銳角，$\cos\alpha > 0$，於是 $1 + 2\cos 2\alpha > 0$，這表示 $F > 0$ 亦成立。兩式相除，

$$\frac{F}{R} = \frac{\sin\alpha}{3\cos\alpha} = \tfrac{1}{3}\tan\alpha.$$

要保持平衡，摩擦力必須滿足 $F \le \mu R$，所以 $\mu \ge F/R = \frac{1}{3}\tan\alpha$，正是所要證明的結果。

質點：摩擦系數的最小值

繩子留在原來的位置，所以圓柱體的位置完全相同，它的三條方程也不變。特別是，張力仍然是

$$T = 2F = \frac{2Mg\sin\alpha}{1 + 2\cos 2\alpha}.$$

當 $\tan\alpha = \frac{1}{2}$ 時，我們有 $\sin\alpha = \frac{1}{\sqrt{5}}$ 和 $\cos\alpha = \frac{2}{\sqrt{5}}$，所以

$$\cos 2\alpha = \frac{1 - \tan^2\alpha}{1 + \tan^2\alpha} = \frac{3}{5}, \qquad \sin 2\alpha = \frac{2\tan\alpha}{1 + \tan^2\alpha} = \frac{4}{5},$$

而 $1 + 2\cos 2\alpha = \frac{11}{5}$。因此

$$T = \frac{2Mg}{\sqrt{5}} \cdot \frac{5}{11} = \frac{2\sqrt{5}}{11} Mg.$$

（圓柱體同時需要 $\mu \ge \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$。）

現在考慮位於 Q 的質點。繩子把它拉向 P，也就是說，拉力有沿斜面向下的分量 $T\cos 2\alpha$，以及離開斜面的分量 $T\sin 2\alpha$。它的重力 $3Mg$ 同樣有沿斜面向下的分量，所以作用於質點的摩擦力 F' 必定沿斜面向上。設作用於質點的法向反作用力為 R'。

沿斜面方向：

$$F' = 3Mg\sin\alpha + T\cos 2\alpha = \frac{3Mg}{\sqrt{5}} + \frac{3}{5} \cdot \frac{2\sqrt{5}}{11} Mg = \frac{33\sqrt{5}}{55} Mg + \frac{6\sqrt{5}}{55} Mg = \frac{39\sqrt{5}}{55} Mg.$$

垂直於斜面方向：

$$R' = 3Mg\cos\alpha - T\sin 2\alpha = \frac{6Mg}{\sqrt{5}} - \frac{4}{5} \cdot \frac{2\sqrt{5}}{11} Mg = \frac{66\sqrt{5}}{55} Mg - \frac{8\sqrt{5}}{55} Mg = \frac{58\sqrt{5}}{55} Mg.$$

這是正數，所以質點確實留在斜面上。平衡的條件是 $F' \le \mu' R'$，其中 μ' 是質點與斜面之間的摩擦系數（coefficient of friction），所以

$$\mu' \ge \frac{F'}{R'} = \frac{39}{58}.$$

摩擦系數的最小值是 $\dfrac{39}{58}$。

驗算

- 改為對 A 取力矩，得 $T\,a(1+2\cos 2\alpha) = 2a\,Mg\sin\alpha$，張力相同。（由 A 到繩子的垂直距離是 $a + 2a\cos 2\alpha$。）
- 當 $\alpha \to 0$ 時，斜面變成水平，張力和摩擦力都消失，而下界 $\frac{1}{3}\tan\alpha \to 0$，與預期相符。
- 數值驗算：隨機選取 10 000 個小於 60° 的銳角 α，用坐標建立幾何圖形，解出平衡方程，確認 $F/R = \frac{1}{3}\tan\alpha$。當 $\tan\alpha = \frac{1}{2}$ 時，結果是 $T \approx 0.4066\,Mg$、$F' \approx 1.5856\,Mg$、$R' \approx 2.3580\,Mg$，以及 $F'/R' \approx 0.67241 = 39/58$。

常見錯誤

- 忘記張力會把圓柱體壓向斜面，令 R 增大。漏掉 $T\sin 2\alpha$ 會得出錯誤的比值。
- 弄錯繩子對質點的拉力方向。繩子把質點拉向 P，所以它減少了作用於質點的法向反作用力，並加大了沿斜面向下的力。
- 在第二部分從頭重新計算張力。圓柱體的平衡沒有改變，所以 T 也沒有改變。

STEP 之外

剛體靜力學（statics）歸結為兩條向量方程：合力為零，以及對任何一點的總力矩為零。選好這一點是最主要的技巧；本題選的是中心，因為有兩個未知的力通過它。到了大學，這會發展成對力系及其等效「力螺旋」（wrench）的研究。摩擦力問題還會引出平衡如何被破壞的問題：是滑動、滾動，還是翻倒。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）對剛體靜力學和轉動有細緻的論述。

詳細解答

STEP III 1991，第 12 題

力學 • 曲線上的運動

> **STEP III 1991，第 12 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/91-s3-q12

關鍵思路

質點被約束在橢圓上，所以它的位置完全由單一參數 θ 決定。把位置向量（position vector）對時間求導兩次，再把加速度投影到向內的單位法向量（inward unit normal）上：含 $\ddot{\theta}$ 的項會消去，因為它們沿切線方向。至於第二部分，當法向反作用力（normal reaction）降至零時，質點便離開表面；而管是光滑的，所以由能量守恆（energy conservation）可求得每一點的速率。

解答

法向加速度

我們把質點視為在光滑曲線上運動的一點，因此它只受兩個力：向下的重量 mg，以及管所施的法向反作用力 R。

位置為 $\mathbf{r} = (a\cos\theta,\ b\sin\theta)$。利用鏈式法則（chain rule）對時間求導，

$$\dot{x} = -a\sin\theta\,\dot{\theta}, \qquad \dot{y} = b\cos\theta\,\dot{\theta},$$
$$\ddot{x} = -a\sin\theta\,\ddot{\theta} - a\cos\theta\,\dot{\theta}^2, \qquad \ddot{y} = b\cos\theta\,\ddot{\theta} - b\sin\theta\,\dot{\theta}^2.$$

切線方向為 $(-a\sin\theta,\ b\cos\theta)$，所以 $(b\cos\theta,\ a\sin\theta)$ 是一個法線方向（兩者的點積（dot product）為 0）。在 $\theta = 0$ 時，它等於 $(b, 0)$，指向橢圓之外，因此向內的單位法向量為

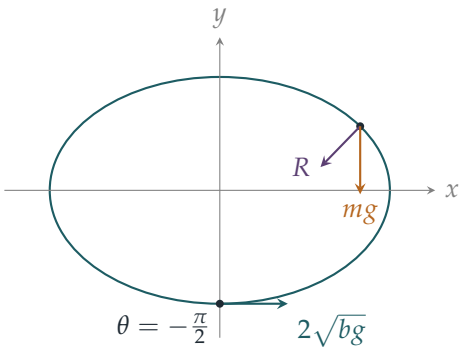
$$\mathbf{n} = -\frac{(b\cos\theta,\ a\sin\theta)}{D}, \qquad D = \sqrt{a^2\sin^2\theta + b^2\cos^2\theta}.$$

於是沿向內法線方向的加速度為

$$\begin{aligned}\ddot{\mathbf{r}}\cdot\mathbf{n} &= -\frac{1}{D}\Big[b\cos\theta\big(-a\sin\theta\,\ddot\theta - a\cos\theta\,\dot\theta^2\big) + a\sin\theta\big(b\cos\theta\,\ddot\theta - b\sin\theta\,\dot\theta^2\big)\Big]\\ &= -\frac{1}{D}\Big[-ab\,\dot\theta^2\big(\cos^2\theta+\sin^2\theta\big)\Big] = \frac{ab\,\dot\theta^2}{\sqrt{a^2\sin^2\theta + b^2\cos^2\theta}},\end{aligned}$$

因為兩個含 $\ddot\theta$ 的項互相抵消。這正是所求的結果。

離開表面



最低點對應 $\theta=-\frac{\pi}{2}$，該處 $y=-b$。管是光滑的，所以反作用力不作功，能量守恆。設質點在參數為 θ 的點上速率為 v，則

$$\tfrac12 mv^2 + mgb\sin\theta = \tfrac12 m(4bg) - mgb \quad\Longrightarrow\quad v^2 = 2gb(1-\sin\theta).$$

由第一部分，$v^2=\dot x^2+\dot y^2 = D^2\dot\theta^2$，所以 $\dot\theta^2 = v^2/D^2$。

現在沿向內法線 $\mathbf{n}$ 的方向分解各力。反作用力 R 沿 $\mathbf{n}$ 作用（質點在內側，所以管向內推，而且只要仍有接觸便有 $R\ge 0$）。重量為 $(0,-mg)$，它沿 $\mathbf{n}$ 的分量是 $(0,-mg)\cdot\mathbf{n} = mga\sin\theta/D$。由牛頓第二定律（Newton's second law）得

$$R + \frac{mga\sin\theta}{D} = \frac{m\,ab\,\dot\theta^2}{D} = \frac{m\,ab\,v^2}{D^3}, \qquad \text{所以} \qquad R = \frac{ma}{D^3}\Big(bv^2 - g\sin\theta\,D^2\Big).$$

想一想 為甚麼重量的分量是 $+mga\sin\theta/D$，而不是帶負號？

在上半部分（$\sin\theta>0$），向內法線有向下的分量，所以重力在那裏有助把質點拉向內側，情況就如在鉛垂圓周的頂部一樣。R 和重量的這個分量共同提供向心加速度（centripetal acceleration）。

現在代入 $a=\frac32 b$ 和 $v^2=2gb(1-\sin\theta)$，並記 $s=\sin\theta$。那麼

$$D^2 = \tfrac94 b^2 s^2 + b^2(1-s^2) = b^2\big(1+\tfrac54 s^2\big),$$

所以

$$bv^2 - g\sin\theta\, D^2 = gb^2\left[2(1-s) - s\left(1+\tfrac{5}{4}s^2\right)\right] = -\tfrac{1}{4}gb^2\left(5s^3+12s-8\right).$$

由於 $ma/D^3 > 0$，R 的正負號與 $-(5s^3+12s-8)$ 的正負號相同。

質點在 R 首次變為零之處離開表面。函數 $p(s) = 5s^3+12s-8$ 嚴格遞增（strictly increasing），因為 $p'(s) = 15s^2+12 > 0$，而且 $p(-1) = -25 < 0$，$p(1) = 9 > 0$。所以 p 在 $(-1,1)$ 內恰有一個根。質點由最低點上升時，s 由 -1 開始遞增，在 s 到達該根之前一直有 $R > 0$。此外，對所有 $s < 1$ 都有 $v^2 = 2gb(1-s) > 0$，所以質點到達該處時仍在沿管向上運動。因此 R 首次變為零、質點離開表面的位置滿足

$$5\sin^3\theta + 12\sin\theta - 8 = 0,$$

正如所求。數值上，$\sin\theta \approx 0.584$，所以質點在略高於軸的水平之處離開。

驗算

- 若 $a = b = r$，法向加速度變成 $r\dot{\theta}^2 = v^2/r$，即熟悉的圓周運動結果。
- 在最低點，$R = m\left(\frac{bv^2}{a^2}+g\right)$，其中 $v^2 = 4bg$，其值為正（該處的曲率半徑（radius of curvature）為 a^2/b），與 $p(-1) = -25 < 0$ 吻合。
- 取 $b = 1$、$a = 1.5$ 和 $g = 9.81$，我們由最低點開始，逐步對以 θ 表示的運動方程作數值積分。反作用力首次變為零時 $\sin\theta \approx 0.58379$，與三次方程的根 0.58377 相符（$\theta \approx 35.7°$）。

常見錯誤

- 把 θ 當作極角（polar angle）。除非 $a = b$，否則點 $(a\cos\theta, b\sin\theta)$ 與 x 軸的夾角並不是 θ，法線也不沿半徑方向。
- 以 v^2/b 或 v^2/a 作為法向加速度。在橢圓上，曲率半徑逐點不同，而這正是第一部分所計算的。
- 弄錯重量法向分量的正負號，結果三次式變成 $5s^3 - 4s + \ldots$，得不到合理的根。

STEP 之外

第一部分是法向加速度公式 v^2/ρ 的特殊情況，其中 ρ 是曲率半徑。事實上，這個橢圓的 $\rho = D^3/(ab)$。把加速度分解為切向和法向兩部分（弗勒內標架（Frenet frame）），是

大學一年級力學和微分幾何（differential geometry）的標準內容。以單一坐標 θ 處理約束，自然引向拉格朗日力學（Lagrangian mechanics）。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門）對約束運動，以及極坐標和內稟坐標（intrinsic coordinates），都有很好的講解。

詳細解答

STEP III 1991，第 13 題

力學 • 拋體與天花板的碰撞

STEP III 1991，第 13 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

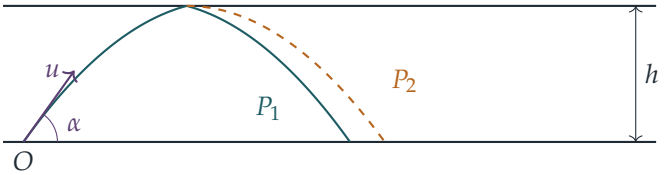
step.maths.org/questions/91-s3-q13

關鍵思路

天花板是光滑的，所以水平速度 $u\cos\alpha$ 始終不變，每個射程都只是 $u\cos\alpha$ 乘以飛行時間。全部工夫都在鉛垂方向的運動上，而能量方程（energy equation）可以立即給出質點到達天花板時的鉛垂速率。經過彈性碰撞（elastic impact）之後，質點下落的過程與上升的過程完全一樣，只是倒過來播放；經過非彈性碰撞（inelastic impact）之後，它就只是從靜止落下。

解答

把兩個質點都視為在重力作用下自由運動的質點（point mass），不計空氣阻力。記 $u=\sqrt{8gh}$、$s=\sin\alpha$ 及 $c=\cos\alpha$。以 O 為原點，取 x 沿水平方向，y 鉛垂向上。



到達天花板

初始鉛垂速度為 us，而在上升期間

$$\dot{y}^2=u^2s^2-2gy=8gh\,s^2-2gy.$$

質點要到達高度 h，必須在該處有 $\dot{y}^2\geq 0$，即 $8ghs^2\geq 2gh$，所以 $s^2\geq\frac{1}{4}$。由於 α 是銳角，$s>0$，故此條件即 $\sin\alpha\geq\frac{1}{2}$，亦即 $\alpha\geq\frac{1}{6}\pi$。（當 $\alpha=\frac{1}{6}\pi$ 時，質點恰好擦過天花

板。）當它確實到達天花板時，它在該處的向上速率為

$$w = \sqrt{8ghs^2 - 2gh} = \sqrt{2gh}\,\sqrt{4s^2 - 1}.$$

P_1 的射程

在上升途中，鉛垂速度以 g 的變率由 us 均勻地減至 w，所以到達天花板所需的時間為

$$t_1 = \frac{us - w}{g}.$$

想一想 P_1 從天花板回落到地面需要多少時間？

碰撞是完全彈性的，而天花板是光滑的，所以鉛垂速度變為向下的 w，水平速度則不變。於是，由天花板向下的鉛垂運動，就是上升時的鉛垂運動在時間上倒過來進行（從高度 h 出發，以速率 w 朝向地面，並以 g 加速）。因此它到達地面時的鉛垂速率為 us，所需時間同樣是 t_1。直接計算亦可：鉛垂速率以 g 的變率由 w 增至 $\sqrt{w^2 + 2gh} = us$。

總飛行時間為 $2t_1$，所以射程（range）為

$$\begin{aligned} R_1 &= uc \cdot \frac{2(us - w)}{g} = \frac{2\sqrt{8gh}\,c}{g}\left(2\sqrt{2gh}\,s - \sqrt{2gh}\sqrt{4s^2 - 1}\right) \\ &= \frac{2\sqrt{8gh}\sqrt{2gh}}{g}\,c\left(2s - \sqrt{4s^2 - 1}\right) = 8h\cos\alpha\left(2\sin\alpha - \sqrt{4\sin^2\alpha - 1}\right), \end{aligned}$$

當中用到 $us = \sqrt{8gh}\,s = 2\sqrt{2gh}\,s$ 及 $\sqrt{8gh}\sqrt{2gh} = 4gh$。這就是所求的射程。

P_2 的射程與差距 D

P_2 沿相同的路徑上升至天花板，所需時間為 t_1。碰撞是完全非彈性的，所以它的鉛垂速度變為零；天花板是光滑的，所以水平速度仍然是 uc。其後它從靜止下落高度 h，所需時間為 $\sqrt{2h/g}$。因此

$$R_2 = uc\left(t_1 + \sqrt{\tfrac{2h}{g}}\right), \qquad R_2 - R_1 = uc\left(\sqrt{\tfrac{2h}{g}} - t_1\right).$$

現在

$$t_1 = \frac{\sqrt{2gh}\,(2s - \sqrt{4s^2 - 1})}{g} = \sqrt{\frac{2h}{g}}\,(2s - \sqrt{4s^2 - 1})$$

而 $uc\sqrt{2h/g} = \sqrt{8gh}\sqrt{2h/g}\,c = 4hc$，所以

$$R_2 - R_1 = 4h\cos\alpha\left(1 - 2\sin\alpha + \sqrt{4\sin^2\alpha - 1}\right).$$

當 $\frac{1}{6}\pi < \alpha < \frac{1}{2}\pi$ 時，此式為正：對於 $s > \frac{1}{2}$，我們有 $(2s-1)^2 = 4s^2 - 1 - 2(2s-1) < 4s^2 - 1$，所以 $2s - 1 < \sqrt{4s^2 - 1}$。因此 P_2 總是落在較遠處（彈性反彈令 P_1 更快地被拋回地面），而

$$D = 4h\cos\alpha\left(1 - 2\sin\alpha + \sqrt{4\sin^2\alpha - 1}\right).$$

求 D 的極大值

設 $f(\alpha) = D/(4h) = c(1-2s) + cr$，其中 $r = \sqrt{4s^2-1}$，定義於 $\frac{1}{6}\pi \le \alpha < \frac{1}{2}\pi$。留意 $f(\frac{1}{6}\pi) = 0$，當 $\alpha \to \frac{1}{2}\pi$ 時 $f \to 0$（因為 $c \to 0$），而在兩者之間 $f > 0$。求導（differentiate），並利用 $\frac{\mathrm{d}r}{\mathrm{d}\alpha} = \frac{4sc}{r}$，得

$$f'(\alpha) = -s(1-2s) - 2c^2 - sr + \frac{4sc^2}{r}.$$

代入 $c^2 = 1-s^2$，再乘以 $r > 0$（當 $\alpha > \frac{1}{6}\pi$）：

$$\begin{aligned} rf'(\alpha) &= (4s^2 - s - 2)\,r - sr^2 + 4s(1-s^2) \\ &= (4s^2 - s - 2)\,r - s(8s^2 - 5). \end{aligned}$$

當 $\alpha = \frac{1}{4}\pi$ 時，$s = \frac{1}{\sqrt{2}}$ 而 $r = 1$，得出 $\left(2 - \frac{1}{\sqrt{2}} - 2\right) - \frac{1}{\sqrt{2}}(4-5) = 0$。所以 $\alpha = \frac{1}{4}\pi$ 是一個駐點（stationary point）。

想一想 我們怎樣知道它就是極大值，而不是別的駐點？

在任何駐點，都有 $(4s^2 - s - 2)r = s(8s^2 - 5)$。兩邊平方並利用 $r^2 = 4s^2 - 1$，展開後因式分解，得

$$(4s^2 - s - 2)^2(4s^2 - 1) - s^2(8s^2 - 5)^2 = -2(2s^2 - 1)(8s^3 - s^2 - 2s - 2),$$

所以每個駐點都滿足 $s^2 = \frac{1}{2}$ 或 $q(s) = 8s^3 - s^2 - 2s - 2 = 0$。在 $[\frac{1}{2}, 1]$ 上，$q'(s) = 24s^2 - 2s - 2 > 0$，所以 q 是遞增的；由於 $q(0.8) = -0.144 < 0 < 0.0985 = q(0.82)$，它在該區間內唯一的根位於 $(0.8, 0.82)$ 之內。但對於這樣的 s，

$$4s^2 - s - 2 < 4(0.82)^2 - 0.8 - 2 < 0, \qquad 8s^2 - 5 > 8(0.8)^2 - 5 > 0,$$

所以左邊 $(4s^2 - s - 2)r$ 為負，右邊 $s(8s^2 - 5)$ 為正：這個根是平方時產生的，並不是駐點。因此 $\alpha = \frac{1}{4}\pi$ 是 $(\frac{1}{6}\pi, \frac{1}{2}\pi)$ 內唯一的駐點。由於 f 是連續的，在區間內部為正，而在兩端為零（或趨向零），它在區間內部必有極大值，而這極大值必定出現在這個駐點。所以當 $\alpha = \frac{1}{4}\pi$ 時 D 最大，且

$$D_{\max} = 4h \cdot \tfrac{1}{\sqrt{2}}\left(1 - \sqrt{2} + 1\right) = 4h(\sqrt{2} - 1).$$

驗算

- 當 $\alpha = \frac{1}{6}\pi$ 時，質點剛好觸及天花板，所以兩個質點的結果應該一致：果然 $D = 0$，而 $R_1 = 8h \cdot \frac{\sqrt{3}}{2} = 4\sqrt{3}\,h$ 正是慣常的射程 $u^2 \sin 2\alpha / g$。
- 取 $h = 1.7\,\mathrm{m}$，在 $\alpha = 1$ 時以數值方法逐步計算兩個運動，得 $R_1 \approx 2.420\,\mathrm{m}$ 及 $D \approx 2.464\,\mathrm{m}$，與公式吻合。對 α 作細密搜尋，得出 $D/(4h)$ 的極大值位於 $\alpha = 0.78540 \approx \frac{1}{4}\pi$，其值為 $0.41421 \approx \sqrt{2} - 1$。

常見錯誤

- 由 $h = ust - \frac{1}{2}gt^2$ 解出 t_1 時取了較大的根；那其實是假如沒有天花板，質點在下落途中經過高度 h 的時間。
- 寫成 $D = R_1 - R_2$，這是負數：非彈性碰撞的質點落在較遠處。
- 得出 $f'(\frac{1}{4}\pi) = 0$ 便停下來，沒有證明這個駐點是極大值。

STEP 之外

對 P_1 所用的時間反演（time reversal）論證，讓我們初次窺見一個深刻的事實：在保守力（conservative force）作用下，牛頓（Newton）定律在時間倒流時形式不變；完全彈性碰撞保持這種對稱性，而非彈性碰撞（會損失能量）則破壞了它。介乎 0 與 1 之間的恢復係數（coefficient of restitution），會給出介乎這裏求得的兩個射程之間的結果。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）正是以這種精神處理碰撞與能量守恆。

詳細解答

STEP III 1991，第 14 題

力學 • 涉及彈簧的圓周運動

STEP III 1991，第 14 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

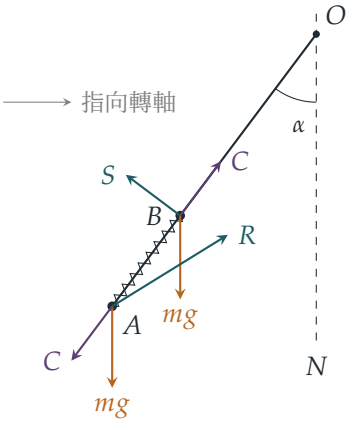
step.maths.org/questions/91-s3-q14

關鍵思路

所有物體都以相同的角速率（angular speed）繞 ON 作水平圓周運動，所以每個質點都有已知的加速度 $\omega^2 r$，方向指向轉軸，此外再無其他加速度。小環套在一根光滑（smooth）的桿上，對它來說最簡潔的方程是沿桿方向的那一條，因為未知的法向反作用力（normal reaction）不會在其中出現。接着，彈簧的力把小環與 A 處的質點聯繫起來，而桿在該處的反作用力的方向就決定了 k。

解答

模型。桿是輕而光滑的，A 和 B 是質量為 m 的質點，彈簧沿着桿位於 B 與 A 之間。設 $\alpha = \sin^{-1}\frac{3}{5}$，於是 $\sin\alpha = \frac{3}{5}$，$\cos\alpha = \frac{4}{5}$，並且 $\omega^2 = \frac{5g}{2a}$。設 $OB = x$，則 $AB = 2a - x$。



這裏 C 是彈簧中的推力（thrust）（我們稍後會發現彈簧是被壓縮的，所以它把 B 推向 O，並把 A 推離 O），S 是光滑桿對 B 的法向反作用力（垂直於桿，位於鉛垂面內，指向有待確定），而 R 是桿對 A 的反作用力。由於 ω 是常數，而且 B 相對於桿靜止，所以沒有橫越鉛垂面的加速度，因此所有的力都在該平面內。

長度 OB

B 沿半徑為 $x\sin\alpha$ 的水平圓周運動，所以它的加速度是 $\omega^2 x\sin\alpha$，方向水平指向 ON。沿桿按 $O\to A$ 的方向（向下並遠離轉軸）分解（resolve）。加速度在這個方向上的分量是 $-\omega^2 x\sin\alpha\cdot\sin\alpha$，因為加速度指向轉軸，而桿與鉛垂線的夾角是 α。

最簡單的做法，是把彈簧作用於 B 的力設為張力（tension）T（把 B 拉向 A），讓正負號自行在計算中顯現。由胡克定律（Hooke's law），

$$T=\frac{kmg\left((2a-x)-a\right)}{a}=\frac{kmg(a-x)}{a}.$$

因為桿是光滑的，反作用力 S 沒有沿桿方向的分量。所以，沿 OA 方向，

$$mg\cos\alpha+\frac{kmg(a-x)}{a}=-m\omega^2 x\sin^2\alpha.$$

代入 $\cos\alpha=\frac{4}{5}$、$\sin^2\alpha=\frac{9}{25}$ 和 $\omega^2=\frac{5g}{2a}$，右邊就是 $-\frac{9gx}{10a}m$。兩邊除以 mg 再乘以 $10a$：

$$\begin{aligned}8a+10k(a-x)&=-9x\\(10k+8)a&=(10k-9)x,\end{aligned}$$

所以

$$OB=x=\frac{(10k+8)a}{10k-9},$$

這正是要證明的結果。留意 $x>a$（答案要為正數，必須有 $10k>9$），所以 $AB=2a-x<a$：彈簧是被壓縮的，推力為

$$C=-T=\frac{kmg(x-a)}{a}=\frac{17k\,mg}{10k-9},$$

其中用到 $x-a=\dfrac{17a}{10k-9}$。

求 k

想一想 為甚麼 A 處的反作用力不像 B 處的那樣垂直於桿？

A 處的質點是固定在桿上的，而不是套在桿上，所以桿可以沿任何方向推它或拉它。光滑這個條件只限制了作用於滑動小環的力。因此我們把 R 寫成分量：R_h 水平指向轉軸，R_v 鉛垂向上。

A 沿半徑為 $2a\sin\alpha = \frac{6a}{5}$ 的圓周運動，所以它的加速度是 $\omega^2 \cdot \frac{6a}{5} = 3g$，方向指向轉軸。推力 C 沿 OA 作用於 A，方向遠離 O，所以它的分量是遠離轉軸的 $C\sin\alpha$ 和向下的 $C\cos\alpha$。因此

$$\text{水平（指向轉軸）：}\quad R_h - \tfrac{3}{5}C = 3mg,$$
$$\text{鉛垂：}\quad R_v - mg - \tfrac{4}{5}C = 0.$$

設 $C = c\,mg$。反作用力與水平線的夾角是 $\tan^{-1}\frac{13}{21}$，所以

$$\frac{R_v}{R_h} = \frac{1+\frac{4}{5}c}{3+\frac{3}{5}c} = \frac{13}{21} \implies 21 + \tfrac{84}{5}c = 39 + \tfrac{39}{5}c \implies 9c = 18, \quad c = 2.$$

（當 $c > 0$ 時，R_h 和 R_v 都是正數，所以反作用力確實指向上方和內側，正負號並無歧義。）於是

$$\frac{17k}{10k-9} = 2 \implies 17k = 20k - 18 \implies k = 6.$$

所以 $k = 6$，由此得 $OB = \frac{68a}{51} = \frac{4}{3}a$，穩妥地介乎 a 與 $2a$ 之間，而推力為 $C = 2mg$。（這時作用於 A 的反作用力是 $R_h = \frac{21}{5}mg$，$R_v = \frac{13}{5}mg$。）

作用於小環的反作用力

? **想一想** 應該沿哪個方向分解，才能以最少的工夫求出 S？

垂直於桿的方向：這樣彈簧的力就不會出現。在鉛垂面內取垂直於 OA 的單位方向，指向上方並遠離轉軸；它與水平線的夾角是 α，所以它的分量是向外的 $\cos\alpha$ 和向上的 $\sin\alpha$。設 S 為作用於 B 的反作用力在這個方向上的分量。重量的貢獻是 $-mg\sin\alpha$，而加速度（$\omega^2 x\sin\alpha$，指向轉軸）的貢獻是 $-\omega^2 x\sin\alpha\cos\alpha$。所以

$$S - mg\sin\alpha = -m\omega^2 x\sin\alpha\cos\alpha.$$

代入 $x = \frac{4}{3}a$：

$$m\omega^2 x\sin\alpha\cos\alpha = m\cdot\frac{5g}{2a}\cdot\frac{4a}{3}\cdot\frac{12}{25} = \frac{8}{5}mg, \qquad \text{所以} \qquad S = \frac{3}{5}mg - \frac{8}{5}mg = -mg.$$

因此桿與小環之間的反作用力的大小是

$$\boxed{mg},$$

而負號告訴我們，它垂直於桿，方向是向下並指向轉軸。

驗算

- 對 B 作完整的向量驗算（水平分量以指向轉軸為正，鉛垂分量以向上為正），取 $C = 2mg$ 沿 AO 作用，以及上面求得的 $S = mg$。水平方向：$\frac{3}{5}(2mg) + \frac{4}{5}(mg) = 2mg$，

而 $m\omega^2 x\sin\alpha = m\cdot\frac{5g}{2a}\cdot\frac{4a}{3}\cdot\frac{3}{5} = 2mg$。鉛垂方向：$\frac{4}{5}(2mg) - \frac{3}{5}mg - mg = 0$。兩者都平衡。

- 當 $k\to\infty$ 時，彈簧變成剛性，而 $OB\to a$，即彈簧的自然長度（natural length），與預期相符。
- 數值方面，取 $m = g = a = 1$，我們以 200 個隨機的 k 值精確地驗證了 OB 的公式，發現夾角只在 $k = 6$ 時才等於 $\tan^{-1}\frac{13}{21}$，並得到 $|S| = 1.0$。由一個小位移開始，逐步計算 B 沿桿的運動，結果它一直留在 $x\approx 1.3333$ 附近。

常見錯誤

- 假設彈簧被拉長，然後對 $x > a$ 感到困惑。應讓張力的正負號由代數運算自行得出。
- 把 A 處的反作用力當作垂直於桿。只有滑動的小環才受這個限制。
- 圓周運動的半徑用了 x 而不是 $x\sin\alpha$。

STEP 之外

在隨桿一起轉動的參考系中，B 受到一個向外的「離心力」（centrifugal force）$m\omega^2 r$，沿桿方向的方程就變成一道普通的平衡問題。把這個觀點再推進一步，就會得到有效勢（effective potential）、科里奧利力（Coriolis force）（在這裏，一旦 B 沿桿移動，它就由桿的反作用力抵銷），以及平衡的穩定性：$x = \frac{4}{3}a$ 是否穩定？Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教材）正是以這個程度討論非慣性參考系（non-inertial frames）和轉動系統。

詳細解答

STEP III 1991，第 15 題

概率與統計 • 樣本中配對的計數

STEP III 1991，第 15 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q15

關鍵思路

每一組 k 張牌出現的機會均等，所以每個概率（probability）都等於符合條件的手牌數目除以 $\binom{2n}{k}$。要數出具有指定對子結構的手牌，可以分階段構造。首先選定哪些種類以完整的一對出現，然後選定另外哪些種類只出現一張，最後決定每個單張種類的兩張牌中，我們手上拿的是哪一張。這個計數一旦完成，題目的每一部分都只是同一條公式的特殊情況。

解答

把 $2n$ 張牌看成全部可以互相區分（同一種類的兩張牌不妨編號為 1 和 2）。這樣，不放回地抽出 k 張牌時，$\binom{2n}{k}$ 個大小為 k 的子集各有相同的概率。當某一種類的兩張牌都被抽中時，我們說手牌中有該種類的一個對子（pair）。

四張牌，沒有對子

沒有對子的手牌包含四個不同的種類，每個種類各一張。選種類有 $\binom{n}{4}$ 種方法，每個種類選哪一張牌有 2 種方法，所以共有 $2^4\binom{n}{4}$ 手牌。由於

$$\binom{2n}{4} = \frac{2n(2n-1)(2n-2)(2n-3)}{24} = \frac{4n(n-1)(2n-1)(2n-3)}{24},$$

我們得到

$$\mathrm{P}(\text{沒有對子}) = \frac{16\cdot n(n-1)(n-2)(n-3)/24}{4n(n-1)(2n-1)(2n-3)/24} = \frac{4(n-2)(n-3)}{(2n-1)(2n-3)},$$

正如所求。（這裏 $n \geq 4$ 保證了四個不同的種類確實存在。）

四張牌，恰好一個對子

先選出以對子出現的種類（n 種方法）。其餘兩張牌必須來自餘下 $n-1$ 個種類中的兩個不同種類，選法有 $\binom{n-1}{2}$ 種，而每個種類選哪一張牌又有 2 種選擇。手牌的數目是

$$n \cdot \frac{(n-1)(n-2)}{2} \cdot 4 = 2n(n-1)(n-2),$$

所以

$$\mathrm{P}(\text{恰好一個對子}) = \frac{24 \cdot 2n(n-1)(n-2)}{4n(n-1)(2n-1)(2n-3)} = \frac{12(n-2)}{(2n-1)(2n-3)}.$$

k 張牌，恰好 r 個對子

假設手牌恰好有 r 個對子。它們佔去 $2r$ 張牌。餘下的 $k-2r$ 張牌全部屬於不同的種類，而且這些種類都不是那 r 個成對的種類，否則對子就會更多。於是我們數：

- 成對的種類：$\binom{n}{r}$ 種方法；
- 只出現一張的 $k-2r$ 個種類，從其餘 $n-r$ 個種類中選出：$\binom{n-r}{k-2r}$ 種方法；
- 每個單張種類的兩張牌中，我們拿的是哪一張：2^{k-2r} 種方法。

每一手恰好有 r 個對子的牌，都恰好由一組這樣的選擇得出。因此

$$\boxed{\mathrm{P}(\text{恰好}r\text{ 個對子}) = \frac{2^{k-2r}\binom{n}{r}\binom{n-r}{k-2r}}{\binom{2n}{k}}}$$

想一想 這個概率會不會等於零？它與四張牌的答案是否吻合？

單張的種類需要滿足 $k-2r \le n-r$，即 $r \ge k-n$。若 $k > n$，任意 k 張牌必定包含至少 $k-n$ 個對子，而 $\binom{n-r}{k-2r} = 0$ 自動處理了這個情況。取 $k=4$，$r=0$ 的情況給出 $2^4\binom{n}{4}/\binom{2n}{4}$，$r=1$ 則給出 $2^2\,n\binom{n-1}{2}/\binom{2n}{4}$，正好就是上面的兩個計數。

k 為偶數，所有牌都成對

設 $k=2m$。手牌由 m 個對子組成，當且僅當它是 m 個完整種類的併集（union）。這樣的手牌有 $\binom{n}{m}$ 手（這就是一般公式取 $r=m$ 的情況）。所以

$$\mathrm{P} = \frac{\binom{n}{m}}{\binom{2n}{2m}} = \frac{n!}{m!\,(n-m)!} \cdot \frac{(2m)!\,(2n-2m)!}{(2n)!}.$$

要化簡這個式子，可把階乘（factorial）中的偶數因子和奇數因子分開歸組：

$$(2j)! = \big(2 \cdot 4 \cdots 2j\big)\big(1 \cdot 3 \cdots (2j-1)\big) = 2^j\, j!\,\big(1 \cdot 3 \cdots (2j-1)\big).$$

記 $O_j = 1\cdot 3\cdots(2j-1)$，並設 $O_0 = 1$。那麼 $(2j)!/j! = 2^j O_j$，所以

$$\mathrm{P} = \frac{2^m O_m \cdot 2^{n-m} O_{n-m}}{2^n O_n} = \frac{O_m\, O_{n-m}}{O_n}.$$

把 O_n 除以 O_{n-m}，剩下的是由 $2n-2m+1 = 2n-k+1$ 到 $2n-1$ 的奇數，所以

$$\mathrm{P}(\tfrac{1}{2}k \text{ 個對子}) = \frac{1\times 3\times 5\times\cdots\times(k-1)}{(2n-1)(2n-3)\cdots(2n-k+1)},$$

正如所求。

另外還有一個簡潔的直接論證。把整副牌洗勻，以最上面的 k 張牌作為手牌。看看最頂的那張牌。它的拍檔出現在其餘 $2n-1$ 個位置中任何一個的機會均等，而其中 $k-1$ 個位置在手牌之內，所以拍檔在手牌中的概率是 $\frac{k-1}{2n-1}$。如果它確實在手牌中，就把這一對拿走。剩下的是 $2n-2$ 張牌的一個隨機排列，而手牌就是其中最上面的 $k-2$ 張。重複這個論證得到 $\frac{k-3}{2n-3}$，如此類推，直到 $\frac{1}{2n-k+1}$。把這些條件概率(conditional probability)相乘，便得到同一個乘積。

驗算

- 四張牌時共有三種情況：0、1 和 2 個對子，而 2 個對子的概率是 $\binom{n}{2}/\binom{2n}{4} = \frac{3}{(2n-1)(2n-3)}$。果然

  $$4(n-2)(n-3) + 12(n-2) + 3 = 4n^2 - 8n + 3 = (2n-1)(2n-3),$$

  所以三個概率之和為 1。
- 取 $n = 5$ 和四張牌，三個概率分別是 $\frac{8}{21}$、$\frac{4}{7}$ 和 $\frac{1}{21}$。用電腦對 $4 \le n \le 8$ 以及由 3 到 $2n-1$ 的每個 k 窮舉所有手牌，結果與一般公式及全部成對的公式均相符。
- 取 $k = 2$，全部成對的公式給出 $\frac{1}{2n-1}$，即第二張牌與第一張牌配成一對的機會。

常見錯誤

- 用 $\binom{2n-2r}{k-2r}$ 從餘下全部 $2n-2r$ 張牌中選出單張的牌。這樣會容許其中兩張屬於同一種類，從而產生額外的對子。
- 遺漏因子 2^{k-2r}。每個只出現一張的種類，拿到的可以是它兩張牌中的任何一張。

STEP 之外

恰好有 r 個對子的手牌數目，是典型的「先選結構，再填內容」式計數，而乘積 $1\cdot 3\cdots(2m-1)$（即雙階乘（double factorial）$(2m-1)!!$）數的是把 $2m$ 件物件分成若

干對的方法數目。雙階乘還會以標準正態分佈（standard normal distribution）的偶數階矩（moment）的身分再次出現。同樣的配對問題，即隨機樣本中有多少個對子得以保存，也見於經典的「失襪」問題和生日類型的問題。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）為這類計數打下良好的基礎。

詳細解答

STEP III 1991，第 16 題

概率與統計 • 聯合分佈與邊際分佈

STEP III 1991，第 16 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q16

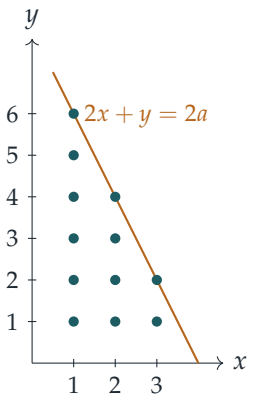
關鍵思路

這裏的一切都是在一個由格點（lattice points）組成的三角形上的有限和。先把區域畫出來：對每個 x，我們要知道哪些 y 是容許的；對每個 y，又有哪些 x 是容許的。一旦求和的上下限定對了，每個邊際分佈（marginal distribution）都是一個等差數列（arithmetic progression）之和。唯一要留神的地方，是對給定的 y，容許的最大 x 是 $\lfloor (2a-y)/2 \rfloor$，而這取決於 y 是偶數還是奇數。

解答

區域

我們需要 $x \geq 1$、$y \geq 1$ 和 $2x + y \leq 2a$。由於 $y \geq 1$，必須有 $2x \leq 2a - 1$，所以 $x \leq a - 1$。對於固定的 x（$1 \leq x \leq a - 1$），容許的值是 $y = 1, 2, \ldots, 2a - 2x$。同樣地，因為 $x \geq 1$，所以 $y \leq 2a - 2$。



圖中所示為 $a = 4$ 的情況：第 x 列有 $2a - 2x$ 個點。

X 的邊際分佈與 c 的值

固定 x，並記 $n = 2a - 2x$。把等差數列 $2x + 1, \ldots, 2x + n$ 求和，

$$\begin{aligned} \mathrm{P}(X = x) &= c \sum_{y=1}^{n} (2x + y) = c \left(2xn + \frac{n(n+1)}{2} \right) = c \, \frac{n}{2} \, (4x + n + 1) \\ &= c \, (a - x)(2a + 2x + 1), \qquad x = 1, 2, \ldots, a - 1, \end{aligned}$$

而在其他情況下 $\mathrm{P}(X = x) = 0$。

所有概率之和必須等於 1。設 $k = a - x$，它由 1 取到 $a - 1$；於是 $2a + 2x + 1 = 4a + 1 - 2k$，並且

$$\begin{aligned} \sum_{k=1}^{a-1} k(4a + 1 - 2k) &= (4a + 1) \frac{(a-1)a}{2} - 2 \cdot \frac{(a-1)a(2a-1)}{6} \\ &= a(a-1) \left(\frac{4a+1}{2} - \frac{2a-1}{3} \right) = \frac{a(a-1)(8a+5)}{6}. \end{aligned}$$

令 c 乘以此和等於 1，便得

$$c = \frac{6}{a(a-1)(8a+5)},$$

正如所求。

Y 的邊際分佈

固定 y，其中 $1 \le y \le 2a - 2$。容許的 x 是 $1, 2, \ldots, N$，其中 N 是滿足 $2N \le 2a - y$ 的最大整數。對這 N 項求和，

$$\sum_{x=1}^{N} (2x + y) = N(N + 1) + yN = N(N + 1 + y).$$

y 為偶數。此時 $2a-y$ 是偶數，$N=\frac{1}{2}(2a-y)$，所以 $N+1+y=\frac{1}{2}(2a+2+y)$，並且

$$\mathrm{P}(Y=y)=c\cdot\frac{(2a-y)(2a+2+y)}{4}=\frac{3(2a-y)(2a+2+y)}{2a(a-1)(8a+5)},$$

其中用了 c 的值。這就是所要求的結果。

> **想一想** 當 y 是奇數時，為甚麼不能用 $N=\frac{1}{2}(2a-y)$？

當 y 是奇數時，$\frac{1}{2}(2a-y)$ 不是整數。滿足 $2x\le 2a-y$ 的最大整數 x 是 $N=\frac{1}{2}(2a-y-1)$：向下取整把那半個捨去了。

y 為奇數。取 $N=\frac{1}{2}(2a-1-y)$，便得 $N+1+y=\frac{1}{2}(2a+1+y)$，所以

$$\mathrm{P}(Y=y)=c\cdot\frac{(2a-1-y)(2a+1+y)}{4}=\frac{3(2a-1-y)(2a+1+y)}{2a(a-1)(8a+5)}.$$

（這兩種情況可以寫成 $3\big((2a+1)^2-(y+1)^2\big)$ 和 $3\big(4a^2-(y+1)^2\big)$ 除以同一個分母。）

Y 的期望值

把 $y\,\mathrm{P}(Y=y)$ 求和，就要把偶數和奇數的 y 分開處理。較簡潔的做法是回到聯合分佈（joint distribution），逐列求和。對於固定的 x，同前面一樣取 $n=2a-2x$，

$$\sum_{y=1}^{n}y(2x+y)=2x\cdot\frac{n(n+1)}{2}+\frac{n(n+1)(2n+1)}{6}=\frac{n(n+1)}{6}\big(6x+2n+1\big).$$

再次設 $k=a-x$，於是 $n=2k$，$6x+2n+1=6a-2k+1$。每列之和變成

$$\frac{2k(2k+1)(6a+1-2k)}{6}=\frac{k(2k+1)(6a+1-2k)}{3}.$$

展開得 $k(2k+1)(6a+1-2k)=(6a+1)(2k^2+k)-4k^3-2k^2$。利用求和至 $a-1$ 的標準公式，

$$\sum_{k=1}^{a-1}(2k^2+k)=\frac{a(a-1)(2a-1)}{3}+\frac{a(a-1)}{2}=\frac{a(a-1)(4a+1)}{6},$$

$$\sum_{k=1}^{a-1}k^3=\frac{a^2(a-1)^2}{4},\qquad\sum_{k=1}^{a-1}k^2=\frac{a(a-1)(2a-1)}{6}.$$

所以

$$\begin{aligned}&\sum_{k=1}^{a-1}k(2k+1)(6a+1-2k)\\&\quad=a(a-1)\left(\frac{(6a+1)(4a+1)}{6}-a(a-1)-\frac{2a-1}{3}\right)\\&\quad=\frac{a(a-1)}{6}\big(24a^2+10a+1-6a^2+6a-4a+2\big)\\&\quad=\frac{a(a-1)(6a^2+4a+1)}{2}.\end{aligned}$$

因此

$$\mathrm{E}(Y) = \frac{c}{3}\cdot\frac{a(a-1)(6a^2+4a+1)}{2} = \frac{6}{a(a-1)(8a+5)}\cdot\frac{a(a-1)(6a^2+4a+1)}{6}$$
$$= \frac{6a^2+4a+1}{8a+5}.$$

驗算

- 當 $a = 2$ 時只有兩個點，$(1,1)$ 和 $(1,2)$，權重分別為 $3c$ 和 $4c$，所以 $c = \frac{1}{7}$。這與 $6/(2\cdot 1\cdot 21)$ 一致。於是 $\mathrm{E}(Y) = \frac{3+8}{7} = \frac{11}{7}$，而公式給出 $\frac{33}{21} = \frac{11}{7}$。
- X 和 Y 的邊際分佈公式（奇偶兩種情況）以及 $\mathrm{E}(Y)$，已對 a 由 2 至 29 的每個值用分數作精確枚舉核對。例如，$a = 3$ 給出 $\mathrm{E}(Y) = \frac{67}{29}$。
- 當 a 很大時，$\mathrm{E}(Y) \approx \frac{3}{4}a$。權重 $2x + y$ 偏重靠近斜邊的點，所以 Y 明顯高於在三角形上取相等權重時所得的 $\frac{2}{3}a$。

常見錯誤

- 讓 x 一直取到 a：$x = a$ 這一點會迫使 $y \le 0$。
- 對奇數 y 使用 $N = \frac{1}{2}(2a - y)$，而它並不是整數。
- 試圖從分兩種情況的邊際分佈求 $\mathrm{E}(Y)$。這樣做是行得通的，但沿聯合分佈的各列求和，就可以完全避開奇偶之分。

STEP 之外

從聯合分佈求邊際分佈，就是把另一個變量求和消去，而難處總在於區域的上下限。對於連續變量，同樣的工作變成在二重積分（double integral）中選取積分限，而交換求和次序（正如我們求 $\mathrm{E}(Y)$ 時所做的）則變成交換積分次序。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）涵蓋了離散和連續變量的聯合分佈與邊際分佈。

Beyond STEP

STEP 是大多數中學生一生中應考過最艱深的數學考試，也是最出色的考試之一：題目簡短、陌生而深刻，當中最精彩的，其實是喬裝了的大學數學小品。

本書精選最有趣的 STEP 試題，由 1987 年首次開考之前的樣本試卷（specimen papers），到 2025 年的試卷，並以每一道題目作為一扇門。每道題目都會仔細解答，沿途提出優秀學生會問的問題，然後帶你穿過這扇門：這個想法從何而來，到了大學會演變成甚麼，以及接下來該讀些甚麼。

本書為正在準備 STEP、有天分又有志投身數學的學生而寫。他們想要的不止是模擬試卷：他們想開始像數學系本科生那樣思考，並一探前方有甚麼正等着自己。

- 共 13 章，涵蓋純數學（pure mathematics）、力學（mechanics）、概率（probability）與統計（statistics）
- 136 道 STEP 歷屆試題，每題均附引導式解答
- 每道題目背後的大學數學
- 附答案的延伸練習，以及延伸閱讀

1991 年詳細解答. 全書各章及
所有詳細解答均可免費閱讀、下載及分享，網址為
beyondstep.keithmcnulty.org

文字內容已貢獻予公有領域（CC0 1.0）。STEP 試題 © Cambridge University Press & Assessment。