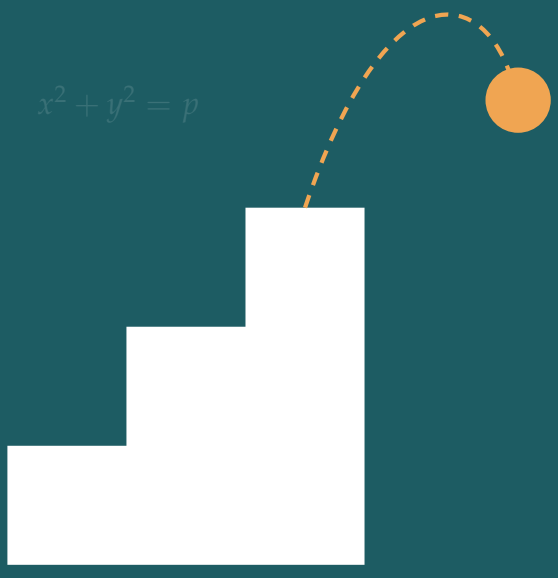
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

1990 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

1990 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

1990 年詳細解答

1990 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 三角函數的最優化 純數學 . . . 3

第 2 題 單位根篩選法 純數學 . . . 6

第 3 題 法線與平行曲線 純數學 . . . 10

第 4 題 等差數列與差分 純數學 . . . 14

第 5 題 部分分式與曲線描繪 純數學 . . . 17

第 6 題 向量與平行四邊形 純數學 . . . 21

第 7 題 一階線性常微分方程 純數學 . . . 25

第 8 題 三角函數無窮乘積 純數學 . . . 28

第 9 題 切線與面積的界限 純數學 . . . 32

第 10 題 摩擦率與阻力 力學 . . . 35

第 11 題 拋體與動量 力學 . . . 38

第 12 題 彈簧與能量 力學 . . . 42

第 13 題 摩擦力與剛體平衡 力學 . . . 46

第 14 題 期望值與最優停止 概率與統計 . . . 50

第 15 題 二項分佈 概率與統計 . . . 53

第 16 題 角分佈與假設檢驗 概率與統計 . . . 57

STEP II

第 1 題 多項式不等式 純數學 . . . 61

第 2 題 托勒密定理 純數學 . . . 65

第 3 題 三次曲線的描繪 純數學 . . . 69

第 4 題 直線排佈與數學歸納法 純數學 . . . 74

第 5 題 複數與圓 純數學 . . . 78

第 6 題 整數不等式與分數 純數學 . . . 82

第 7 題 微分方程 純數學 86
第 8 題 分部積分法 純數學 90
第 9 題 極坐標曲線與弧長 純數學 93
第 10 題 矩陣多項式方程 純數學 97
第 11 題 動量守恆 力學 101
第 12 題 恢復係數與拋體 力學 105
第 13 題 桿的平衡 力學 109
第 14 題 彈性碰撞 力學 112
第 15 題 順序統計量 概率與統計 116
第 16 題 貝葉斯定理 概率與統計 119

STEP III

第 1 題 倍角與多項式的根 純數學 122
第 2 題 三維向量 純數學 126
第 3 題 群公理與共軛 純數學 130
第 4 題 三角級數 純數學 133
第 5 題 萊布尼茲定理 純數學 136
第 6 題 線性變換與橢圓 純數學 140
第 7 題 隱函數曲線描繪 純數學 145
第 8 題 恰當微分方程 純數學 150
第 9 題 雙曲函數 純數學 153
第 10 題 根與漸近近似 純數學 156
第 11 題 剛體轉動 力學 160
第 12 題 穩動楔上的質點 力學 164
第 13 題 球內的運動 力學 168
第 14 題 耦系 力學 172
第 15 題 幾何分佈 概率與統計 176
第 16 題 幾何概率 概率與統計 180

詳細解答

STEP I 1990，第 1 題

純數學 • 三角函數的最優化

> **STEP I 1990，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/90-s1-q1

關鍵思路

窗簾的每個頂點與圓心 D 的距離都是 r，所以半徑 DX 和 DY 把這個四邊形分成三個以 D 為頂點的等腰三角形（isosceles triangle）。每個三角形的面積都是 $\frac{1}{2}r^2\sin(\text{頂角})$，而三個頂角之和為 π。這樣，一道幾何題就變成求幾個正弦之和的最大值，而第一部分只需用和化積公式（sum-to-product formula）便可完成，完全不必動用微積分。

解答

窗簾的面積

由於 A、D、C 共線（collinear），在 D 處的各角滿足

$$\angle ADX + \angle XDY + \angle YDC = \pi, \qquad \text{所以} \qquad \angle YDC = \pi - \theta - \alpha.$$

這裏 $\theta \geq 0$、$\alpha > 0$ 且 $\theta + \alpha \leq \pi$，使得 X 和 Y 在弧上按 A, X, Y, C 的次序排列。

半徑 DX 和 DY 把窗簾 $AXYC$ 分成三角形 ADX、XDY 和 YDC，每個三角形都有兩條長度為 r 的邊相交於 D。對每個三角形使用面積 $= \frac{1}{2}ab\sin C$，得

$$\begin{aligned} \text{面積} &= \tfrac{1}{2}r^2\sin\theta + \tfrac{1}{2}r^2\sin\alpha + \tfrac{1}{2}r^2\sin(\pi - \theta - \alpha) \\ &= \tfrac{1}{2}r^2[\sin\alpha + \sin\theta + \sin(\theta + \alpha)], \end{aligned}$$

因為 $\sin(\pi - u) = \sin u$。

固定 α 時的最佳 θ

當 α 固定時，只有 $\sin\theta + \sin(\theta+\alpha)$ 會變動。根據和化積公式 $\sin P + \sin Q = 2\sin\frac{P+Q}{2}\cos\frac{P-Q}{2}$，

$$\sin\theta + \sin(\theta+\alpha) = 2\cos\tfrac{\alpha}{2}\,\sin\!\left(\theta+\tfrac{\alpha}{2}\right).$$

> **想一想** 為甚麼這樣一寫，最大值就一目了然？

因子 $2\cos\frac{\alpha}{2}$ 是一個固定的正常數，因為 $0 < \frac{\alpha}{2} < \frac{\pi}{2}$。所以面積最大，恰好就是 $\sin(\theta+\frac{\alpha}{2})$ 最大的時候。當 θ 取遍 $[0,\pi-\alpha]$ 時，角 $\theta+\frac{\alpha}{2}$ 取遍 $[\frac{\alpha}{2},\pi-\frac{\alpha}{2}]$，而這個區間包含 $\frac{\pi}{2}$。因此正弦可以達到它的最大可能值 1，此時 $\theta+\frac{\alpha}{2}=\frac{\pi}{2}$，即

$$\boxed{\theta = \frac{\pi}{2} - \frac{\alpha}{2} = \frac{\pi-\alpha}{2}.}$$

這時同樣有 $\angle YDC = \pi-\theta-\alpha = \frac{\pi-\alpha}{2} = \theta$：最佳的窗簾關於垂直的半徑對稱，與直覺相符。對這個 α 而言，最大面積是

$$\tfrac{1}{2}r^2\left[\sin\alpha + 2\cos\tfrac{\alpha}{2}\right].$$

讓 α 變動

現在我們要在 $0<\alpha<\pi$ 的範圍內求 $f(\alpha)=\sin\alpha+2\cos\frac{\alpha}{2}$ 的最大值。（對每個 α，最大面積是 $\frac{1}{2}r^2 f(\alpha)$，所以整體的最大面積就是 $\frac{1}{2}r^2$ 乘以 f 的最大值。）求導數（differentiate），得

$$f'(\alpha) = \cos\alpha - \sin\tfrac{\alpha}{2}.$$

設 $s=\sin\frac{\alpha}{2}$，於是 $0<s<1$，再利用 $\cos\alpha = 1-2\sin^2\frac{\alpha}{2}$：

$$f'(\alpha) = 1 - s - 2s^2 = (1-2s)(1+s).$$

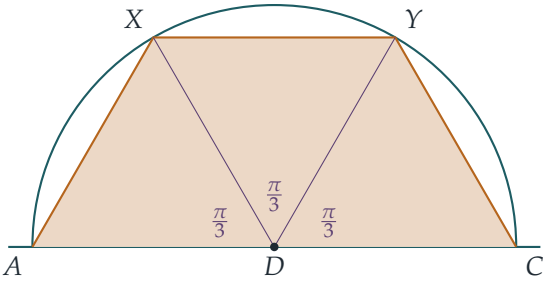
由於 $1+s>0$，f' 的正負號與 $1-2s$ 相同。因為 s 在 $(0,\pi)$ 上隨 α 遞增，f' 在 $s<\frac{1}{2}$ 時為正，在 $s>\frac{1}{2}$ 時為負。所以 f 先遞增後遞減，並在 $\sin\frac{\alpha}{2}=\frac{1}{2}$ 處取得最大值，即 $\alpha=\frac{\pi}{3}$。在該處

$$f\!\left(\tfrac{\pi}{3}\right) = \sin\tfrac{\pi}{3} + 2\cos\tfrac{\pi}{6} = \tfrac{\sqrt{3}}{2} + \sqrt{3} = \tfrac{3\sqrt{3}}{2},$$

所以窗簾的最大可能面積是

$$\boxed{\tfrac{1}{2}r^2\cdot\tfrac{3\sqrt{3}}{2} = \frac{3\sqrt{3}}{4}r^2.}$$

此時 $\alpha=\theta=\angle YDC=\frac{\pi}{3}$，所以三個三角形都是等邊三角形，而窗簾正是內接（inscribed）於整個圓的正六邊形（regular hexagon）的一半。



驗算

- 極限情況：當 $\alpha \to 0$ 時，窗簾變成三角形 AXC，其中 X 位於最高點，面積為 $\frac{1}{2} \cdot 2r \cdot r = r^2$，而確實有 $\frac{1}{2}r^2 f(0) = r^2$。當 $\alpha \to \pi$ 時，窗簾塌縮到直徑上，而 $f(\pi) = 0$。最大值 $\frac{3\sqrt{3}}{4}r^2 \approx 1.299r^2$ 比兩者都大，同時小於半圓的面積 $\frac{\pi}{2}r^2 \approx 1.571r^2$，這是必然的。
- 數值驗證：我們隨機選取 X 和 Y 的位置，把三個三角形的面積公式與鞋帶公式（shoelace formula）互相核對；又對若干個 α 的值，用細密的網格搜尋找出最佳的 θ；並發現取 $r = 1$ 時，對 θ 和 α 同時搜尋所得的最大面積為 $1.299038\ldots = \frac{3\sqrt{3}}{4}$。

常見錯誤

- 遺漏第三個三角形 YDC，或者把它的角寫成 $\pi - \alpha$ 而不是 $\pi - \theta - \alpha$。
- 在第二部分直接取 $\theta = \alpha = \frac{\pi}{3}$ 這條捷徑，卻沒有給出理由。必須用微積分（或對稱性的論證）才能證明最佳的窗簾確實是正六邊形的一半。

STEP 之外

這個答案是一個經典結果的特殊情況：在內接於同一個圓的所有 n 邊形之中，以正多邊形的面積最大。有一個簡潔的證明，用到 sin 在 $[0, \pi]$ 上的凹性（concavity）和琴生不等式（Jensen's inequality）：若圓心角（central angle）$\phi_1, \ldots, \phi_n$ 之和為 2π，則 $\sum \sin \phi_i \leq n \sin \frac{2\pi}{n}$。我們的窗簾就是 $n = 6$ 這個情況的一半。凸性（convexity）和這類不等式是大學一年級分析（analysis）課程中貫穿始終的主題；Abbott 的 *Understanding Analysis* 是正式認識這些概念的好起點，而 Siklos 的 *Advanced Problems in Mathematics* 則收錄了更多同類風格的 STEP 題目。

詳細解答

STEP I 1990，第 2 題

純數學 • 單位根篩選法

> **STEP I 1990，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/90-s1-q2

本題在 *Beyond STEP* 第 5.2 節單位根篩選法中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch05/roots-filter.html

關鍵思路

單位立方根（cube roots of unity）在這裏起着篩選器的作用。當我們用二項式定理（binomial theorem）展開 $(1+\omega^r)^n$，再把 $r=1,2,3$ 的結果相加時，每個二項式係數 $\binom{n}{j}$ 都會乘上 $1+\omega^j+\omega^{2j}$；當 3 整除 j 時這個數是 3，否則是 0。因此這個和剛好只保留我們想要的係數。計算每個 $(1+\omega^r)^n$ 的另一個方法，是把 $1+\omega^r$ 寫成模—輻角形式（modulus–argument form），再用棣美弗定理（de Moivre's theorem）。

解答

單位立方根

由於 $\omega=\mathrm{e}^{2\pi\mathrm{i}/3}$，我們有 $\omega^3=\mathrm{e}^{2\pi\mathrm{i}}=1$，但 $\omega\neq 1$。於是

$$0=\omega^3-1=(\omega-1)(1+\omega+\omega^2),$$

而因為 $\omega-1\neq 0$，可以把它約去，得到 $1+\omega+\omega^2=0$。

由此可知 $1+\omega^2=-\omega$。乘以 $-1=\mathrm{e}^{-\mathrm{i}\pi}$ 相當於旋轉 $-\pi$，所以

$$1+\omega^2=\mathrm{e}^{-\mathrm{i}\pi}\,\mathrm{e}^{2\pi\mathrm{i}/3}=\mathrm{e}^{-\mathrm{i}\pi/3}.$$

因此 $1+\omega^2$ 的模是 1，輻角是 $-\frac{\pi}{3}$。（驗算一下：$1+\omega^2=1+\left(-\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i}\right)=\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i}$。）同理，$1+\omega=-\omega^2=\mathrm{e}^{\mathrm{i}\pi/3}$，即它的共軛複數（complex conjugate）。

恆等式

*方法一：二項式定理。*對 $r = 1, 2, 3$，

$$(1+\omega^r)^n = \sum_{j=0}^{n} \binom{n}{j} \omega^{rj}.$$

把三種情況相加，

$$(1+\omega)^n + (1+\omega^2)^n + (1+\omega^3)^n = \sum_{j=0}^{n} \binom{n}{j} \left(\omega^j + \omega^{2j} + \omega^{3j}\right).$$

> **想一想** 對一般的整數 j，$\omega^j + \omega^{2j} + \omega^{3j}$ 等於甚麼？

由於 $\omega^{3j} = 1$，括號內的式子就是 $1 + \omega^j + \omega^{2j}$；又因為 $\omega^3 = 1$，它只取決於 j 模 3 的值。

- 若 j 是 3 的倍數，則 $\omega^j = 1$，括號內的值是 $1 + 1 + 1 = 3$。
- 若 $j \equiv 1 \pmod 3$，則 $\omega^j = \omega$ 及 $\omega^{2j} = \omega^2$，所以括號內的值是 $1 + \omega + \omega^2 = 0$。
- 若 $j \equiv 2 \pmod 3$，則 $\omega^j = \omega^2$ 及 $\omega^{2j} = \omega^4 = \omega$，所以括號內的值同樣是 0。

因此只有 $3 \mid j$ 的項留下來，而

$$(1+\omega)^n + (1+\omega^2)^n + (1+\omega^3)^n = 3\left[\binom{n}{0} + \binom{n}{3} + \cdots + \binom{n}{k}\right],$$

其中 k 是不超過 n 的最大的 3 的倍數。

*方法二：模與輻角。*由第一部分，$1 + \omega = \mathrm{e}^{\mathrm{i}\pi/3}$ 及 $1 + \omega^2 = \mathrm{e}^{-\mathrm{i}\pi/3}$，而 $1 + \omega^3 = 2$。根據棣美弗定理，

$$\begin{aligned}(1+\omega)^n + (1+\omega^2)^n + (1+\omega^3)^n &= \mathrm{e}^{\mathrm{i}n\pi/3} + \mathrm{e}^{-\mathrm{i}n\pi/3} + 2^n \\ &= 2^n + 2\cos\left(\frac{n\pi}{3}\right).\end{aligned}$$

比較兩種算法的結果，再除以 3，

$$\binom{n}{0} + \binom{n}{3} + \binom{n}{6} + \cdots + \binom{n}{k} = \frac{1}{3}\left(2^n + 2\cos\left(\frac{n\pi}{3}\right)\right),$$

這就是所要證明的結果。

$n = 25$ 的情況

不超過 25 的最大的 3 的倍數是 24，所以第一個和正正就是恆等式取 $n = 25$ 的情況。現在 $\frac{25\pi}{3} = 8\pi + \frac{\pi}{3}$，所以 $\cos\frac{25\pi}{3} = \cos\frac{\pi}{3} = \frac{1}{2}$，而

$$\binom{25}{0} + \binom{25}{3} + \cdots + \binom{25}{24} = \frac{1}{3}\left(2^{25} + 1\right) = \frac{33554433}{3} = 11184811.$$

第二個和

這裏 $n = 24$，而我們要的是 $j \equiv 2 \pmod 3$ 的係數 $\binom{24}{j}$。把所有係數 $\binom{24}{j}$ 按 j 除以 3 的餘數分組：

$$S_0 + S_1 + S_2 = \sum_{j=0}^{24} \binom{24}{j} = 2^{24},$$

其中 S_m 是所有 $j \equiv m \pmod 3$ 的 $\binom{24}{j}$ 之和。

對稱性 $\binom{24}{j} = \binom{24}{24-j}$ 把這些項配成對：若 $j \equiv 2$，則 $24 - j \equiv -2 \equiv 1 \pmod 3$，因為 24 是 3 的倍數。所以 $j \mapsto 24 - j$ 把 S_2 的各項與 S_1 的各項一一對應，於是 $S_1 = S_2$。

取 $n = 24$ 的恆等式可給出 S_0：由於 $\cos 8\pi = 1$，

$$S_0 = \frac{1}{3}\left(2^{24} + 2\right).$$

因此

$$S_2 = \frac{1}{2}\left(2^{24} - S_0\right) = \frac{1}{2} \cdot \frac{2 \cdot 2^{24} - 2}{3} = \frac{2^{24} - 1}{3}.$$

把題目給出的數值減半，得 $2^{24} = 16777216$，所以

$$\binom{24}{2} + \binom{24}{5} + \cdots + \binom{24}{23} = \frac{16777215}{3} = 5592405.$$

驗算

- 小數值的情況：當 $n = 3$ 時，左邊是 $1 + 1 = 2$，右邊是 $\frac{1}{3}(8 + 2\cos\pi) = 2$；當 $n = 4$ 時，左邊是 $1 + 4 = 5$，右邊是 $\frac{1}{3}\left(16 + 2\cos\frac{4\pi}{3}\right) = \frac{1}{3}(16 - 1) = 5$。
- 整除性：$2^{25} + 1$ 和 $2^{24} - 1$ 都必須是 3 的倍數；事實確是如此，因為 $2 \equiv -1 \pmod 3$。
- 我們用電腦直接把二項式係數相加：恆等式對 199 或以下的每個 n 都成立，第一個和是 11184811，第二個和是 5592405。

常見錯誤

- 把 $1 + \omega^2$ 的輻角寫成 $\frac{\pi}{3}$：它位於第四象限，所以輻角主值（principal argument）是 $-\frac{\pi}{3}$（等價地，$\frac{5\pi}{3}$）。
- 忘記了 $r = 3$ 的情況，而 2^n 正是由它而來。
- 在第二個和中，試圖直接套用公式：公式只給出 3 的倍數的項，所以要處理餘數為 2 的項，還需要另一個想法（這裏是對稱性）。

STEP 之外

這個技巧稱為單位根篩選法（roots of unity filter），它是離散傅立葉變換（discrete Fourier transform）的第一個例子：把 $f(\omega^r)$ 在所有 m 次單位根上取平均，就能抽出多項式 f 中屬於模 m 某個指定剩餘類（residue class）的係數。同樣這種在一個對稱羣上取平均的做法，會在有限羣的特徵標理論（character theory）以及伯恩賽德引理（Burnside's lemma）等計數論證中再次出現。想進一步了解為何複數令這類計算變得如此自然，可參閱 Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）。

詳細解答

STEP I 1990，第 3 題

純數學 • 法線與平行曲線

> **STEP I 1990，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/90-s1-q3

關鍵思路

只要把向下的單位法向量（unit normal）明確寫出來，從而以 x 表示 $\mathbf{s}$ 的坐標，其餘一切便水到渠成。第 (i) 部分只需迫使 $\mathbf{s}$ 的 y 坐標等於零。在第 (ii) 部分，題目建議的代換（substitution）令 $\mathbf{s}$ 的坐標變得非常簡單，而斜率（gradient）條件則化為一條一階微分方程（first-order differential equation），這條方程原來恰好是一個全導數（exact derivative）。

解答

準備工作。曲線 $y = \mathrm{f}(x)$ 的切線（tangent）方向為 $(1, \mathrm{f}'(x))$，所以法向量為 $\pm(\mathrm{f}'(x), -1)$。向下的那一個的 y 分量為負，故

$$\mathbf{n} = \frac{1}{N}\begin{pmatrix} \mathrm{f}'(x) \\ -1 \end{pmatrix}, \qquad N = \sqrt{1 + \mathrm{f}'(x)^2},$$

因此

$$\mathbf{s} = \mathbf{r} + \mathrm{p}(x)\mathbf{n} = \begin{pmatrix} x + \dfrac{\mathrm{p}(x)\,\mathrm{f}'(x)}{N} \\ \mathrm{f}(x) - \dfrac{\mathrm{p}(x)}{N} \end{pmatrix}.$$

另外留意，由於 $\mathbf{n}$ 是單位向量，所以 $|\mathbf{r} - \mathbf{s}| = |\mathrm{p}(x)|$。

第 (i) 部分

這裏 $\mathrm{f}(x) = x^k$，$\mathrm{f}'(x) = kx^{k-1}$，而 $N = \sqrt{1 + k^2 x^{2k-2}}$。要使 $\mathbf{s}$ 位於 x 軸上，它的 y 坐標必須為零：

$$x^k - \frac{\mathrm{p}(x)}{N} = 0 \quad \Longrightarrow \quad \mathrm{p}(x) = x^k \sqrt{1 + k^2 x^{2k-2}}.$$

這樣選取之後，$\mathbf{s}$ 的 x 坐標為

$$x + x^k \cdot kx^{k-1} = x + kx^{2k-1},$$

它是連續（continuous）而且遞增的，當 $x \to 0^+$ 時趨向 0，當 $x \to \infty$ 時趨向無窮大。所以 $\mathbf{s}$ 確實掃過整條正 x 軸，而不只是其中一部分。

現在 $|\mathbf{r}|^2 = x^2 + x^{2k}$，而 $|\mathbf{r} - \mathbf{s}|^2 = \mathrm{p}(x)^2 = x^{2k} + k^2 x^{4k-2}$（這裏 $\mathrm{p} > 0$）。所以對每一個 $x > 0$ 都有 $|\mathbf{r}| = |\mathbf{r} - \mathbf{s}|$，當且僅當

$$x^2 = k^2 x^{4k-2} \quad \text{對所有} x > 0.$$

取 $x = 1$ 便必有 $k^2 = 1$，所以 $k = 1$；反過來，$k = 1$ 時兩邊都等於 x^2。因此 $k = 1$。

從幾何角度看：在直線 $y = x$ 上，由 (x, x) 出發的向下法線與 x 軸相交於 $(2x, 0)$，而以 O、(x, x)、$(2x, 0)$ 為頂點的三角形是等腰三角形（isosceles triangle），兩條等邊的長度均為 $\sqrt{2}\,x$。

第 (ii) 部分

現在 $\mathrm{f}(x) = x^2$，$\mathrm{f}'(x) = 2x$，而 $N = \sqrt{1 + 4x^2}$。寫成 $\mathrm{p} = \mathrm{q}N$ 便可消去那些礙事的平方根：

$$\mathbf{s} = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} x + 2x\,\mathrm{q}(x) \\ x^2 - \mathrm{q}(x) \end{pmatrix}.$$

新曲線是以 x 為參數（parameter）給出的，所以它在 $\mathbf{s}$ 的斜率為

$$\frac{\mathrm{d}Y}{\mathrm{d}X} = \frac{\mathrm{d}Y/\mathrm{d}x}{\mathrm{d}X/\mathrm{d}x} = \frac{2x - \mathrm{q}'}{1 + 2\mathrm{q} + 2x\mathrm{q}'}.$$

令它等於 $2x$，即 $y = x^2$ 在 $\mathbf{r}$ 的斜率，然後兩邊乘以分母：

$$\begin{aligned} & 2x - \mathrm{q}' = 2x + 4x\mathrm{q} + 4x^2\mathrm{q}' \\ \Longleftrightarrow \quad & (1 + 4x^2)\,\mathrm{q}' + 4x\,\mathrm{q} = 0. \end{aligned}$$

> **想一想** 我們能否不除以 q（它可能在某處為零）而解出這條方程？

可以：改為除以 $N = \sqrt{1 + 4x^2}$，它永不為零。由於 $N' = 4x/N$，

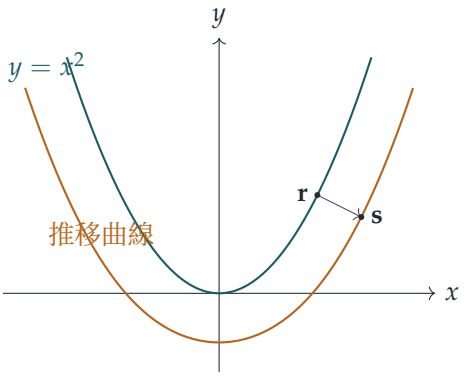
$$\frac{(1 + 4x^2)\,\mathrm{q}' + 4x\,\mathrm{q}}{N} = N\mathrm{q}' + \frac{4x}{N}\,\mathrm{q} = \frac{\mathrm{d}}{\mathrm{d}x}(N\mathrm{q}) = \mathrm{p}'(x).$$

所以條件其實就是對所有 x 都有 $\mathrm{p}'(x) = 0$，因此

$$\mathrm{p}(x) = c, \qquad \text{為一常數}$$

（等價地，$\mathrm{q}(x) = c/\sqrt{1+4x^2}$）。反過來，每一個常數 c 都可行，因為只要新曲線的斜率有定義，上述每一步都是可逆的。

新曲線是一條**平行曲線**（parallel curve）：拋物線（parabola）上每一點都沿着該點的法線被推移一段固定距離 c。下圖所示為 $c = \frac{1}{2}$ 的情況。



有一處細微之處：$\mathrm{d}X/\mathrm{d}x = 1 + 2\mathrm{q} + 2x\mathrm{q}'$，代入 $\mathrm{q} = c/N$ 後可化簡為 $1 + 2c/N^3$。當 $c > -\frac{1}{2}$ 時，它恆為正數，所以推移曲線是一條光滑曲線。若 $c \le -\frac{1}{2}$（即向上推移至少 $\frac{1}{2}$，亦即頂點處的曲率半徑（radius of curvature）），它會在 $N^3 = -2c$ 之處變為零，推移曲線便在該處出現尖點（cusp）。

驗算

- 在 (i) 中取 $k = 2$ 及 $x = 1$：$\mathrm{p}(1) = \sqrt{5}$，而 $\mathbf{s} = (1,1) + \sqrt{5}\cdot\frac{1}{\sqrt{5}}(2,-1) = (3,0)$，一如所求位於 x 軸上；但 $|\mathbf{r}| = \sqrt{2} \neq \sqrt{5}$，所以 $k = 2$ 不符合要求。
- 在 (ii) 中，當 $x = 0$ 時，常數為 c 的推移曲線通過 $(0, -c)$，而根據對稱性，它在該處的斜率為 0，與拋物線頂點的斜率吻合。
- 以數值驗證，取 $c = 1$ 及 $x = 0.8$，用有限差分（finite difference）計算推移曲線的斜率，結果與 $2x = 1.6$ 相符；若改用非常數的 $\mathrm{p} = x^2$，得出的斜率則約為 0.41。

常見錯誤

- 誤用向上的法向量 $(-\mathrm{f}', 1)/N$：$\mathbf{s}$ 中的正負號會全部相反。

- 把新曲線的斜率寫成 $\mathrm{d}Y/\mathrm{d}x$，而不是 $(\mathrm{d}Y/\mathrm{d}x)/(\mathrm{d}X/\mathrm{d}x)$；新曲線的 x 坐標是 X，而不是 x。
- 在 (i) 中，只比較指數（$4k-2=2$）便斷定 $k=1$，而沒有同時檢查係數 $k^2=1$。

STEP 之外

以固定距離推移而成的曲線稱為*平行曲線*或*偏移曲線*（offset curve），它們在實際應用中十分重要：銑牀（milling machine）刀具的中心，正是沿着所切削形狀的平行曲線移動。當偏移距離超過曲率半徑時出現的尖點，都位於*漸屈線*（evolute）之上，即曲率中心的軌跡（locus），它是曲線微分幾何（differential geometry）中的核心對象。Pressley 的 *Elementary Differential Geometry*（初等微分幾何）是一本很好的入門書。

詳細解答

STEP I 1990，第 4 題

純數學 • 等差數列與差分

> **STEP I 1990，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/90-s1-q4

關鍵思路

數列的每一項，都等於首項加上各個差的累計總和。等差數列（arithmetic progression）的差是常數，所以求和很容易。至於雙重等差數列（double arithmetic progression），各個差本身又構成一個等差數列，於是要多做一層求和，二項式係數（binomial coefficient）$\binom{n}{2}$ 和 $\binom{n}{3}$ 便自然出現。最後一部分則化為關於兩個未知差的兩條線性方程。

解答

等差數列

設公差（common difference）為 $d = a_2 - a_1$，則 $a_k = a_1 + (k-1)d$。於是

$$\sum_{k=1}^{n} a_k = na_1 + d\sum_{k=1}^{n}(k-1) = na_1 + \frac{n(n-1)}{2}(a_2 - a_1),$$

其中用到 $0 + 1 + \cdots + (n-1) = \frac{1}{2}n(n-1)$。把 a_1 和 a_2 的項分別合併，得

$$a_1 + a_2 + \cdots + a_n = \frac{n}{2}\big[(3-n)a_1 + (n-1)a_2\big].$$

雙重等差數列

設各個差為 $\delta_k = b_{k+1} - b_k$。它們構成一個等差數列，首項為 $p = \delta_1 = b_2 - b_1$，公差為

$$e = \delta_2 - \delta_1 = (b_3 - b_2) - (b_2 - b_1) = b_3 - 2b_2 + b_1.$$

因此 $\delta_j = p + (j-1)e$，把首 $k-1$ 個差加起來（即第一部分的公式，把 n 換成 $k-1$），得

$$b_k = b_1 + \sum_{j=1}^{k-1} \delta_j = b_1 + (k-1)p + \frac{(k-1)(k-2)}{2}\,e = b_1 + \binom{k-1}{1}p + \binom{k-1}{2}e.$$

現在對 $k = 1, \ldots, n$ 求和。我們需要兩個和。第一個是 $\sum_{k=1}^{n}(k-1) = \binom{n}{2}$，與前面相同。至於第二個，令 $j = k-1$ 由 0 取到 $n-1$，則

$$\begin{aligned}\sum_{j=0}^{n-1} \frac{j(j-1)}{2} &= \frac{1}{2}\left[\frac{(n-1)n(2n-1)}{6} - \frac{(n-1)n}{2}\right] = \frac{n(n-1)}{12}\left[(2n-1)-3\right] \\ &= \frac{n(n-1)(n-2)}{6} = \binom{n}{3}.\end{aligned}$$

所以

$$\sum_{k=1}^{n} b_k = nb_1 + \binom{n}{2}(b_2 - b_1) + \binom{n}{3}(b_3 - 2b_2 + b_1).$$

這已經是一條以 n、b_1、b_2、b_3 表示的公式。如果你想分別看到每個 b_i 的係數，展開後得

$$b_1 + \cdots + b_n = \frac{n(n^2 - 6n + 11)}{6}\,b_1 + \frac{n(n-1)(7-2n)}{6}\,b_2 + \frac{n(n-1)(n-2)}{6}\,b_3.$$

（快速檢驗：$n = 1, 2, 3$ 分別給出 b_1、$b_1 + b_2$ 和 $b_1 + b_2 + b_3$。）

> **想一想** 有沒有更簡潔的方法看出 $\binom{n}{3}$ 從何而來？

有：$\binom{k-1}{2}$ 是從 $k-1$ 件東西中選 2 件的方法數目，而曲棍球棒恆等式（hockey-stick identity）$\sum_{k=1}^{n}\binom{k-1}{2} = \binom{n}{3}$（從 n 件中選 3 件時，先固定最大的一件 k）一步便完成求和。每多一層「差的差」，二項式係數便升高一級，正如積分把冪次升高一次。

階乘數列

現在 $b_1 = 1$，而 p 和 e 如上所定，故 $b_k = 1 + (k-1)p + \frac{1}{2}(k-1)(k-2)e$。所以

$$b_2 = 1 + p, \qquad b_4 = 1 + 3p + 3e, \qquad b_6 = 1 + 5p + 10e.$$

在一個 $c_1 = b_2$、$c_2 = b_4$、$c_3 = b_6$、$c_4 = 220$ 的階乘數列（factorial progression）中，相鄰兩項之差依次為 $1!\,d$、$2!\,d$、$3!\,d$，即 d、$2d$、$6d$：

$$\begin{aligned} b_4 - b_2 &= 2p + 3e = d, \\ b_6 - b_4 &= 2p + 7e = 2d, \\ 220 - b_6 &= 219 - 5p - 10e = 6d. \end{aligned}$$

把第一式代入第二式，得 $2p+7e=4p+6e$，所以 $e=2p$，繼而 $d=8p$。第三條方程變成

$$219-5p-20p=48p \quad\Longrightarrow\quad 73p=219 \quad\Longrightarrow\quad p=3,$$

所以 $e=6$ 及 $d=24$，符合非零的要求。由於每一步都是必然的，這個解是唯一的。

於是 $b_k=1+3(k-1)+3(k-1)(k-2)=3k^2-6k+4$，得出數列 $1,4,13,28,49,76,\ldots$，而 $b_2,b_4,b_6,220=4,28,76,220$ 的差確實是 $24,48,144=1!\cdot 24,\ 2!\cdot 24,\ 3!\cdot 24$。

最後，把 $b_1=1$、$b_2-b_1=3$ 和 $b_3-2b_2+b_1=6$ 代入求和公式：

$$\begin{aligned}1+b_2+\cdots+b_n &= n+3\cdot\frac{n(n-1)}{2}+6\cdot\frac{n(n-1)(n-2)}{6}\\ &= \frac{n}{2}\left[2+3(n-1)+2(n-1)(n-2)\right]\\ &= \frac{n(2n^2-3n+3)}{2}.\end{aligned}$$

驗算

當 $n=1,2,3$ 時，最後的公式給出 1、5 和 18，與 1、$1+4$ 和 $1+4+13$ 吻合。當 $n=6$ 時，公式給出 $3\cdot 57=171$，而直接計算亦得 $1+4+13+28+49+76=171$。此外，我們也用電腦以分數作精確運算，對數百個隨機數列及所有 $n\le 24$ 驗證了兩條一般求和公式。

常見錯誤

- 把階乘數列的差當作 $d,2d,3d$：它們其實是 $1!\,d$、$2!\,d$、$3!\,d$，所以第三個是 $6d$。
- 寫成 $b_k=b_1+kp+\ldots$：b_1 與 b_k 之間只有 $k-1$ 個差。

STEP 之外

這就是有限差分學（calculus of finite differences）：若一個數列的第 m 階差是常數，它便是 n 的 m 次多項式，而它的部分和則是 $m+1$ 次多項式，最宜用 $\binom{n}{0},\binom{n}{1},\binom{n}{2},\ldots$ 這組基來表示（牛頓（Newton）前向差分公式）。把 $\binom{n}{r}$ 求和而得 $\binom{n+1}{r+1}$，正是 $\int x^r\,\mathrm{d}x=\frac{x^{r+1}}{r+1}$ 的離散版本。想多做同類題目，可試試 Siklos 的 *Advanced Problems in Mathematics*（進階數學難題集）中的問題。

詳細解答

STEP I 1990，第 5 題

純數學 • 部分分式與曲線描繪

STEP I 1990，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/90-s1-q5

關鍵思路

在 (i) 中，二次式可以因式分解，所以用部分分式（partial fractions）便可把積分化為兩個對數之差。在 (ii) 中，巨大的冪次看似嚇人，但 $1760 = 8 \times 220$，所以這個函數其實就是 $u^8 - u + q$，其中 $u = x^{220} \geq 0$。圖像的形狀由駐點（stationary points）決定，而根的數目只取決於橫軸相對於局部極大值（local maximum）和兩個極小值的位置。

解答

第 (i) 部分

由於 $6 \times 15 = 90 = 9 \times 10$ 而 $9 + 10 = 19$，這個二次式可分解為：

$$6x^2 + 19x + 15 = (2x+3)(3x+5).$$

設

$$\frac{1}{(2x+3)(3x+5)} = \frac{A}{2x+3} + \frac{B}{3x+5}, \qquad \text{所以} \qquad 1 = A(3x+5) + B(2x+3).$$

代入 $x = -\frac{3}{2}$ 得 $\frac{1}{2}A = 1$，所以 $A = 2$；代入 $x = -\frac{5}{3}$ 得 $-\frac{1}{3}B = 1$，所以 $B = -3$。$2x+3$ 和 $3x+5$ 在 $[1,3]$ 上都是正數，所以無須加上絕對值符號：

$$\int_1^3 \frac{\mathrm{d}x}{6x^2+19x+15} = \int_1^3 \left(\frac{2}{2x+3} - \frac{3}{3x+5}\right)\mathrm{d}x = \Big[\ln(2x+3) - \ln(3x+5)\Big]_1^3$$
$$= \ln\frac{9}{14} - \ln\frac{5}{8} = \ln\frac{72}{70} = \ln\frac{36}{35}.$$

第 (ii) 部分

對稱性與兩端的變化。兩個冪次都是偶數，所以 f 是偶函數（even function），其圖像對稱於 y 軸。當 $|x|$ 很大時，x^{1760} 一項佔主導，所以當 $x \to \pm\infty$ 時 $\mathrm{f}(x) \to +\infty$。此外，$\mathrm{f}(0) = q$ 而 $\mathrm{f}(\pm 1) = q$。

駐點。求導得

$$\mathrm{f}'(x) = 1760x^{1759} - 220x^{219} = 220\,x^{219}\big(8x^{1540} - 1\big).$$

它在 $x = 0$ 處以及 $x^{1540} = \frac{1}{8}$ 之處為零，後者即 $x = \pm a$，其中

$$a = 8^{-1/1540} \approx 0.9987.$$

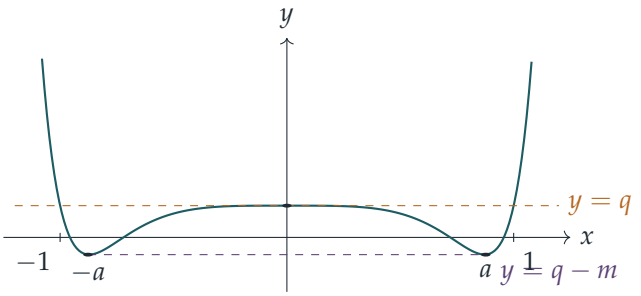
當 $x > 0$ 時，$\mathrm{f}'(x)$ 的正負號與 $8x^{1540} - 1$ 相同：在 $(0, a)$ 上為負，在 $x > a$ 時為正。由於 x^{219} 是奇數次冪，當 $x < 0$ 時正負號剛好相反。所以 f 在 $(-\infty, -a)$ 上遞減，在 $(-a, 0)$ 上遞增，在 $(0, a)$ 上遞減，在 (a, ∞) 上遞增：在 $x = 0$ 處有局部極大值 q，在 $x = \pm a$ 處有兩個相等的極小值。

想一想 怎樣才能求出極小值，而不必跟 $8^{-1/1540}$ 糾纏？

改用 $u = x^{220}$ 來計算。在 $x = \pm a$ 處有 $x^{1540} = u^7 = \frac{1}{8}$，所以 $u = 8^{-1/7}$，而 $u^8 = u \cdot u^7 = \frac{1}{8}u$。因此

$$\mathrm{f}(\pm a) = u^8 - u + q = q - \tfrac{7}{8}u = q - m, \qquad \text{其中} \qquad m = \tfrac{7}{8} \cdot 8^{-1/7} = 7 \cdot 2^{-24/7} \approx 0.650.$$

草圖。圖像呈「W」形：它從高處開始，下降至 $x = -a$ 處的極小值 $q - m$，上升至原點處的局部極大值 q，再下降至 $x = a$ 處的極小值 $q - m$，然後急劇上升。由於冪次極大，真實的圖像極為平坦：除非 $|x|$ 非常接近 1，否則 x^{220} 極小，所以在 $(-1, 1)$ 的大部分範圍內 $\mathrm{f}(x) \approx q$，只在 $x = \pm 1$（圖像在該處回到高度 q）內側不遠處有兩個狹窄的凹陷，而在外側則是近乎垂直的陡壁。下圖把凹陷誇大了，好讓各項特徵清楚可見。



點算根的數目。$\mathrm{f}(x) = 0$ 的根就是圖像與 x 軸相交之處，所以我們把 0 與局部極大值 q 和極小值 $q - m$ 作比較。

- $q > m$：極小值 $q - m$ 為正，所以 $\mathrm{f} > 0$ 處處成立：**沒有根**。
- $q = m$：圖像只在極小值處與橫軸相切：**兩個根**，$x = \pm a$。
- $0 < q < m$：此時 $q - m < 0 < q$。在四個單調區間 $(-\infty, -a)$、$(-a, 0)$、$(0, a)$、(a, ∞) 的每一個上，f 都是連續（continuous）而且嚴格單調（strictly monotonic）的，並且由正值變為負值，或由負值變為正值，所以在每個區間上恰有一個根：**四個根**。
- $q = 0$：$\mathrm{f}(x) = x^{220}(x^{1540} - 1)$，它恰好在 $x = 0$ 和 $x = \pm 1$ 處為零：三**個根**。
- $q < 0$：局部極大值 $\mathrm{f}(0) = q$ 為負，所以在整個 $[-a, a]$ 上 $\mathrm{f} < 0$，而在 (a, ∞) 和 $(-\infty, -a)$ 的每一個上，f 都嚴格單調地由負值變至 $+\infty$：**兩個根**。

所以相異根的數目可能是 0、2、3 和 4，其中

$$\begin{aligned} &0 \text{ 個根：} \quad q > m, \\ &2 \text{ 個根：} \quad q < 0 \text{ 或 } q = m, \\ &3 \text{ 個根：} \quad q = 0, \\ &4 \text{ 個根：} \quad 0 < q < m, \end{aligned} \qquad m = \tfrac{7}{8} \cdot 8^{-1/7} = 7 \cdot 2^{-24/7} \approx 0.650.$$

驗算

對 (i) 的被積函數使用辛普森法則（Simpson's rule），得 0.0281709，與 $\ln \frac{36}{35} = 0.0281709$ 吻合。至於 (ii)，在 $x = 1$ 附近以細密的數值搜尋求 $x^{1760} - x^{220}$ 的極小值，得出在 $x \approx 0.99865$ 處的 -0.650123，與 $-m$ 和 a 一致。取 $q = 1,\ 0.5,\ 0.1,\ 0.001,\ 0,\ -0.01,\ -5$，在細密的格點上點算 f 的變號次數，得出根的數目為 $0, 4, 4, 4, 3, 2, 2$，與上表的預測相符。$q = 0$ 的情況正好用來檢驗草圖是否合理：圖像必須通過原點和 $(\pm 1, 0)$。

常見錯誤

- 遺漏邊界情況 $q = 0$（三個根，因為局部極大值正好落在橫軸上）和 $q = m$（在極小值處有兩個重根）。
- 當 $q < 0$ 時，「憑對稱性」聲稱有兩個根，卻沒有檢查 $-a$ 與 a 之間是否別無變化：正是因為局部極大值 q 為負，才排除了該處有額外的根。
- 把草圖畫成一條彎度平緩的「W」。冪次這麼大，圖像在 $|x| < 1$ 時幾乎平坦地停留在高度 q；即使草圖誇大了凹陷，也值得把這一點說明。

STEP 之外

透過確定駐點處的函數值來點算根的數目，讓人初嘗介值定理（intermediate value theorem）和單調性論證的滋味，而這些內容會在大學第一門分析課程中嚴格建立起來；代換 $u = x^{220}$ 也顯示出 x^n 在 $n \to \infty$ 時的表現：它在 $(-1, 1)$ 上收斂至 0，但並非一致收斂（uniform convergence）。Abbott 的 *Understanding Analysis*（《理解分析》）對連續性和一致收斂的處理既嚴謹又易讀。

詳細解答

STEP I 1990，第 6 題

純數學 • 向量與平行四邊形

STEP I 1990，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/90-s1-q6

關鍵思路

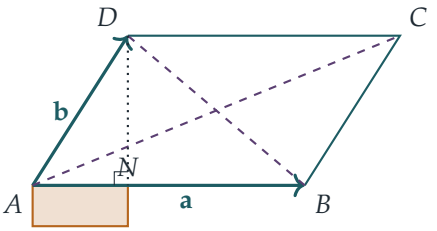
把原點設在 A，並用兩條邊的向量（vector）$\mathbf{a} = \overrightarrow{AB}$ 和 $\mathbf{b} = \overrightarrow{AD}$ 來描述一切。這樣，兩條對角線（diagonal）就是 $\mathbf{a}+\mathbf{b}$ 和 $\mathbf{b}-\mathbf{a}$，而每個長度的平方都是一個向量與自身的點積（dot product）。把這些點積展開，整道題就解決了。唯一用到的幾何事實是：$|\mathbf{u}\cdot\mathbf{v}|$ 等於 $\mathbf{u}$ 的長度乘以 $\mathbf{v}$ 在 $\mathbf{u}$ 上的投影（projection）的長度。

解答

取 A 為原點，設 $\overrightarrow{AB} = \mathbf{a}$ 及 $\overrightarrow{AD} = \mathbf{b}$。由於 $ABCD$ 是平行四邊形（parallelogram），$\overrightarrow{BC} = \overrightarrow{AD} = \mathbf{b}$，所以

$$\overrightarrow{AC} = \mathbf{a}+\mathbf{b}, \qquad \overrightarrow{BD} = \mathbf{b}-\mathbf{a}.$$

對任何向量都有 $|\mathbf{v}|^2 = \mathbf{v}\cdot\mathbf{v}$，而且點積是對稱的，並對加法滿足分配律。記 $a = |\mathbf{a}|$ 及 $b = |\mathbf{b}|$。



圖中，N 是由 D 到直線 AB 的垂足，所以 AN 是 AD 在 AB 上的投影，其長度為 $|\mathbf{a}\cdot\mathbf{b}|/a$。

第 (i) 部分

對邊相等，所以左方等於 $2a^2 + 2b^2$。至於對角線，

$$AC^2 = (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) = a^2 + 2\,\mathbf{a}\cdot\mathbf{b} + b^2,$$
$$BD^2 = (\mathbf{b}-\mathbf{a})\cdot(\mathbf{b}-\mathbf{a}) = a^2 - 2\,\mathbf{a}\cdot\mathbf{b} + b^2.$$

兩式相加，交叉項互相抵消，得 $AC^2 + BD^2 = 2a^2 + 2b^2 = AB^2 + BC^2 + CD^2 + DA^2$，這正是所要證明的。（這個結果通常稱為平行四邊形定律（parallelogram law）。）

第 (ii) 部分

這次改為把同樣的兩式相減，

$$AC^2 - BD^2 = 4\,\mathbf{a}\cdot\mathbf{b}, \qquad \text{所以} \qquad \left|AC^2 - BD^2\right| = 4\,|\mathbf{a}\cdot\mathbf{b}|.$$

現在來解釋 $|\mathbf{a}\cdot\mathbf{b}|$ 的意義。若 θ 是 $\mathbf{a}$ 與 $\mathbf{b}$ 之間的夾角，則 $|\mathbf{a}\cdot\mathbf{b}| = a\cdot b|\cos\theta|$，而 $b|\cos\theta|$ 正好是 AD 在 AB 上的投影的長度 AN。所以 $|\mathbf{a}\cdot\mathbf{b}|$ 就是以 AB 和 AD 在 AB 上的投影為邊的長方形（圖中的陰影部分）的面積。

想一想 題目聲稱這個結果對任何一條邊和它任何一條鄰邊都成立。為甚麼我們不需要分別做八次計算？

平行四邊形的每條邊都是向量 $\pm\mathbf{a}$ 或 $\pm\mathbf{b}$，而與它相鄰的兩條邊，除了正負號之外，就是另一個向量。例如，邊 BC 是 $\mathbf{b}$，它的鄰邊 AB 和 CD 分別是 $\mathbf{a}$ 和 $-\mathbf{a}$。改變正負號並不會改變向量的長度，也不會改變投影的長度，因為投影的長度是 $|\mathbf{u}\cdot\mathbf{v}|/|\mathbf{u}|$。所以無論選哪一條邊和哪一條鄰邊，所得的長方形面積都同樣是 $|\mathbf{a}\cdot\mathbf{b}|$：它或是 $a\times$ (向量$\mathbf{b}$ 投影於$\mathbf{a}$)，或是 $b\times$ (向量$\mathbf{a}$ 投影於$\mathbf{b}$)，而兩者都等於 $|\mathbf{a}\cdot\mathbf{b}|$。因此，在所有情況下，

$$\left|AC^2 - BD^2\right| = 4\times\big(\text{該長方形的面積}\big). \qquad \blacksquare$$

最後部分

第 (ii) 部分來自兩條邊的點積。現在把邊和對角線的角色對調，考慮兩條對角線的點積：

$$\overrightarrow{AC}\cdot\overrightarrow{BD} = (\mathbf{a}+\mathbf{b})\cdot(\mathbf{b}-\mathbf{a}) = b^2 - a^2,$$

因為交叉項 $\mathbf{a}\cdot\mathbf{b}$ 和 $-\mathbf{b}\cdot\mathbf{a}$ 互相抵消。因此

$$\left|AB^2 - AD^2\right| = |\overrightarrow{AC}\cdot\overrightarrow{BD}|.$$

仿照第 (ii) 部分的論證，便得到以下結果。

結果。$\left|AB^2 - AD^2\right|$ 等於以下長方形的面積：它的兩邊分別是任何一條對角線，以及另一條對角線在它上面的投影。

證明。若 φ 是兩條對角線之間的夾角，則 $\left|\overrightarrow{AC} \cdot \overrightarrow{BD}\right| = AC \cdot BD\left|\cos\varphi\right|$。這裏 $BD\left|\cos\varphi\right|$ 是 BD 在 AC 上的投影的長度，而 $AC\left|\cos\varphi\right|$ 是 AC 在 BD 上的投影的長度。所以無論從哪一條對角線出發，面積都相同，而且由上面的式子可知它等於 $\left|AB^2 - AD^2\right|$。■

想讓它看起來更像 (ii)，可設 O 為兩條對角線互相平分的那一點。這樣 $AC = 2OA$，而 BD 在 AC 上的投影是 OB 在 OA 上的投影的兩倍。所以 $\left|AB^2 - AD^2\right|$ 也等於以 OA 和 OB 在 OA 上的投影為邊的長方形面積的 4 倍：半條對角線現在擔當了邊的角色。

由此還可得出兩個漂亮的推論（corollary）。兩條對角線相等，當且僅當 $\mathbf{a} \cdot \mathbf{b} = 0$，亦即當且僅當該平行四邊形是長方形。兩條對角線互相垂直，當且僅當 $a = b$，亦即當且僅當它是菱形（rhombus）。

驗算

取 $A = (0,0)$、$B = (4,0)$、$D = (1,2)$，所以 $C = (5,2)$。這樣 $AC^2 = 29$ 及 $BD^2 = 13$，而 $2(16+5) = 42 = 29+13$，驗證了 (i)。至於 (ii)，$AC^2 - BD^2 = 16$，而 AD 在 AB 上的投影長度為 1，所以長方形的面積是 $4 \times 1 = 4$，而 $16 = 4 \times 4$。至於最後部分，$AB^2 - AD^2 = 16 - 5 = 11$，而 $\overrightarrow{AC} \cdot \overrightarrow{BD} = (5,2) \cdot (-3,2) = -11$。我們還用 10^5 個隨機平行四邊形，就每一種邊和鄰邊的選法測試了全部三個結果；最大的偏差約為 10^{-13}，只是捨入誤差而已。

常見錯誤

- 寫成 $\overrightarrow{BD} = \mathbf{a} - \mathbf{b}$ 或 $\mathbf{a} + \mathbf{b}$。正負號對 BD^2 沒有影響，但把方向寫對可以避免之後出現混亂。
- 忘記取絕對值：當 A 處的角是鈍角時，$\mathbf{a} \cdot \mathbf{b}$ 是負數，但面積不可能是負數，所以題目寫的是 $\left|AC^2 - BD^2\right|$。
- 在最後部分用了整條對角線，卻寫上因子 4。用整條對角線時因子是 1；只有用半條對角線時才會出現因子 4。

STEP 之外

恆等式 $|\mathbf{a}+\mathbf{b}|^2-|\mathbf{a}-\mathbf{b}|^2=4\,\mathbf{a}\cdot\mathbf{b}$ 就是極化恆等式（polarisation identity）：它表明單憑長度就可以還原出點積。它與第 (i) 部分的平行四邊形定律一起，是關於內積空間（inner product space）的關鍵事實。一個範數（norm）來自某個內積，當且僅當它滿足平行四邊形定律，這是約當（Jordan）與馮·諾伊曼（von Neumann）的一條定理。Strang 的 *Introduction to Linear Algebra*（線性代數入門）是個好去處，可以看到點積、投影和正交性（orthogonality）如何在任意維數中發展出來。

詳細解答

STEP I 1990，第 7 題

純數學 • 一階線性常微分方程

> **STEP I 1990，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/90-s1-q7

關鍵思路

設 $y = uv$ 之後，我們有兩個未知函數，但其實只需要一個，所以大可對其中一個多加一個條件。最自然的選擇，是讓 u 消去含 P 的那一項，餘下的微分方程（differential equation）便可直接積分。這其實是喬裝了的積分因子法（integrating-factor method）：$1/u$ 正正就是積分因子。

解答

代換

設 $y = uv$，由乘積法則（product rule）得 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = u\dfrac{\mathrm{d}v}{\mathrm{d}x} + v\dfrac{\mathrm{d}u}{\mathrm{d}x}$，於是方程變成

$$u\frac{\mathrm{d}v}{\mathrm{d}x} + v\left(\frac{\mathrm{d}u}{\mathrm{d}x} + Pu\right) = Q.$$

如果我們選取 u 使它滿足一階方程（first order equation）

$$\frac{\mathrm{d}u}{\mathrm{d}x} + Pu = 0,$$

括號內的式子便等於零，餘下

$$\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{Q}{u},$$

這正是題目所要求的、只以 x、Q 和 u 表示 $\frac{\mathrm{d}v}{\mathrm{d}x}$ 的方程。（要作除法，必須有 $u \neq 0$；$u' + Pu = 0$ 的非零解是 $u = A\mathrm{e}^{-\int P\,\mathrm{d}x}$，其中 $A \neq 0$，這些解永不為零。）

題中的方程

這裏 $P = -\dfrac{2}{x+1}$，$Q = (x+1)^{5/2}$。留意 Q 只在 $x \geq -1$ 時才有定義（作為實數），而 P 在 $x = -1$ 處沒有定義，所以我們在 $x > -1$ 上求解，此時 $x+1 > 0$。

? **想一想** 我們需要 u 的通解（general solution）嗎？還是任何一個解都可以？

任何一個非零解都可以：最終答案中的任意常數，會在積分求 v 時出現。於是我們用分離變數法（separating variables）解

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{2u}{x+1}$$

得 $\ln|u| = 2\ln(x+1) + \text{常數}$，並取

$$u = (x+1)^2.$$

（驗算：$u' = 2(x+1) = \frac{2u}{x+1}$。）於是

$$\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{(x+1)^{5/2}}{(x+1)^2} = (x+1)^{1/2}, \qquad \text{所以} \qquad v = \tfrac{2}{3}(x+1)^{3/2} + C.$$

因此通解是

$$y = uv = \tfrac{2}{3}(x+1)^{7/2} + C(x+1)^2.$$

代入 $y(1) = 0$：

$$0 = \tfrac{2}{3}\cdot 2^{7/2} + 4C \quad \Longrightarrow \quad C = -\tfrac{1}{6}\cdot 2^{7/2} = -\tfrac{1}{6}\cdot 8\sqrt{2} = -\frac{4\sqrt{2}}{3}.$$

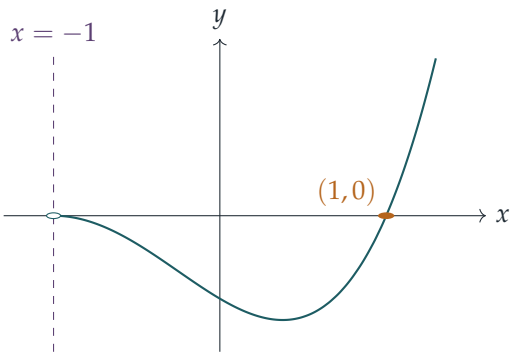
所以

$$\boxed{y = \tfrac{2}{3}(x+1)^{7/2} - \tfrac{4\sqrt{2}}{3}(x+1)^2 = \tfrac{2}{3}(x+1)^2\left[(x+1)^{3/2} - 2\sqrt{2}\right]}$$

解在何處成立

? **想一想** y 的公式在 $x = -1$ 處是有意義的。為甚麼不把 $x = -1$ 包括在內？

這個解在 $x > -1$ 時成立。當 $x < -1$ 時，右邊的 $(x+1)^{5/2}$ 不是實數，所以微分方程本身在那裏毫無意義（我們的 $(x+1)^{7/2}$ 也一樣）。在 $x = -1$ 處，$\frac{2y}{x+1}$ 這一項沒有定義，所以方程在那裏同樣不成立，儘管按我們的公式，當 $x \to -1^+$ 時 $y \to 0$。在整個區間 $x > -1$ 上，由於 $x+1 > 0$，上面每一步都站得住腳，而且解經過題目給定的點 $x = 1$，這一點正在此區間之內。



空心圓圈標示被排除的端點 $x = -1$。

驗算

- 代回原方程：$y' = \frac{7}{3}(x+1)^{5/2} - \frac{8\sqrt{2}}{3}(x+1)$，而 $\frac{2y}{x+1} = \frac{4}{3}(x+1)^{5/2} - \frac{8\sqrt{2}}{3}(x+1)$；兩者之差是 $(x+1)^{5/2}$，符合要求。
- $y(1) = \frac{2}{3} \cdot 4\left(2\sqrt{2} - 2\sqrt{2}\right) = 0$。
- 特殊值：$y(0) = \frac{2-4\sqrt{2}}{3} \approx -1.219$，$y(3) = \frac{256-64\sqrt{2}}{3} \approx 55.16$。由 $y(1) = 0$ 出發，對方程作數值積分（龍格—庫塔法，Runge–Kutta），所得結果與這兩個值吻合至九位有效數字。

常見錯誤

- 弄錯 P 的正負號：這裏 $P = -\frac{2}{x+1}$，所以 $u = (x+1)^2$，而不是 $(x+1)^{-2}$。
- 積分求 v 時漏掉常數，結果沒有餘地去滿足 $y(1) = 0$。
- 把範圍寫成 $x \geq -1$：方程在 $x = -1$ 處沒有定義。

STEP 之外

選取 u 來消去麻煩的一項，正是*參數變易法*（variation of parameters）背後的想法：先解齊次方程（homogeneous equation），然後讓那個「常數」變動起來，以吸收外加項（forcing）。同樣的技巧也適用於二階線性方程；至於解在何處成立這個問題，則發展成存在與唯一性定理（existence and uniqueness theorem），它保證在 P 和 Q 皆連續的任何區間上都有唯一的解。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教科書）對兩者都有論述。

詳細解答

STEP I 1990，第 8 題

純數學 • 三角函數無窮乘積

> **STEP I 1990，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/90-s1-q8

關鍵思路

二倍角公式（double-angle formula）$\sin 2\theta = 2\sin\theta\cos\theta$ 可把每個餘弦化成兩個正弦之比：$\cos\theta = \dfrac{\sin 2\theta}{2\sin\theta}$。把這些比值相乘，幾乎所有項都會互相抵消，情況就如裂項相消的和（telescoping sum）一樣。一旦把有限乘積化成一個簡潔的分數，令 $n \to \infty$ 時便只需用到標準極限（limit）：當 $x \to 0$ 時 $\dfrac{\sin x}{x} \to 1$。

解答

關於 α 的條件

以下每一步都要除以某個 $m \geq 0$ 所對應的 $\sin(\alpha/2^m)$，所以必須確保這些值永不為零。若 $\sin(\alpha/2^m) = 0$，則對某個整數 j 有 $\alpha/2^m = j\pi$，於是 $\alpha = (2^m j)\pi$ 是 π 的整數倍，而這種情況已被排除。因此對所有 $m \geq 0$ 都有 $\sin(\alpha/2^m) \neq 0$。

兩個餘弦的乘積

兩次運用 $\sin 2\theta = 2\sin\theta\cos\theta$，先取 $\theta = \alpha/2$，再取 $\theta = \alpha/4$，得

$$\begin{aligned}\sin\alpha &= 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\\ &= 2\left(2\sin\frac{\alpha}{4}\cos\frac{\alpha}{4}\right)\cos\frac{\alpha}{2} = 4\sin\frac{\alpha}{4}\cos\frac{\alpha}{2}\cos\frac{\alpha}{4}.\end{aligned}$$

由於 $\sin(\alpha/4) \neq 0$，我們可以除以 $4\sin(\alpha/4)$，得出

$$\cos\frac{\alpha}{2}\cos\frac{\alpha}{4} = \frac{\sin\alpha}{4\sin(\alpha/4)},$$

這正是所需的結果。

n 個餘弦的乘積

記 $P_n = \cos\frac{\alpha}{2}\cos\frac{\alpha}{4}\cdots\cos\frac{\alpha}{2^n}$。我們對 n 用歸納法（induction）證明

$$P_n = \frac{\sin\alpha}{2^n\sin(\alpha/2^n)}.$$

當 $n=1$ 時，上式即 $\cos\frac{\alpha}{2} = \dfrac{\sin\alpha}{2\sin(\alpha/2)}$，也就是把二倍角公式除以 $2\sin(\alpha/2)\neq 0$ 所得的結果。（$n=2$ 的情況就是剛才證明的結果。）

現假設公式對某個 $n \geq 1$ 成立。在二倍角公式中取 $\theta = \alpha/2^{n+1}$，得 $\sin\frac{\alpha}{2^n} = 2\sin\frac{\alpha}{2^{n+1}}\cos\frac{\alpha}{2^{n+1}}$。因此

$$\begin{aligned} P_{n+1} = P_n\cos\frac{\alpha}{2^{n+1}} &= \frac{\sin\alpha}{2^n\sin(\alpha/2^n)}\cos\frac{\alpha}{2^{n+1}} \\ &= \frac{\sin\alpha\,\cos(\alpha/2^{n+1})}{2^n\cdot 2\sin(\alpha/2^{n+1})\cos(\alpha/2^{n+1})} = \frac{\sin\alpha}{2^{n+1}\sin(\alpha/2^{n+1})}. \end{aligned}$$

約去 $\cos(\alpha/2^{n+1})$ 是合理的：它是 $\sin(\alpha/2^n)$ 的一個因子，而後者不為零，所以它本身也不為零。這正是 $n+1$ 時的公式，故由歸納法可知，公式對所有正整數 n 都成立。

令 $n\to\infty$

無窮乘積（infinite product）的意思，是部分積（partial products）P_n 的極限。設 $x_n = \alpha/2^n$，則當 $n\to\infty$ 時 $x_n\to 0$，並且

$$2^n\sin\frac{\alpha}{2^n} = \alpha\cdot\frac{\sin x_n}{x_n} \longrightarrow \alpha\cdot 1 = \alpha.$$

> **想一想** 為甚麼可以放心地把 $\lim P_n = \dfrac{\sin\alpha}{\alpha}$ 上下倒轉？

因為這個極限不是零：對我們所考慮的 α，有 $\alpha\neq 0$ 及 $\sin\alpha\neq 0$。所以 $P_n\to\dfrac{\sin\alpha}{\alpha}\neq 0$，無窮乘積存在，而以它除 $\sin\alpha$ 便得

$$\frac{\sin\alpha}{\cos\frac{\alpha}{2}\cos\frac{\alpha}{4}\cos\frac{\alpha}{8}\cdots} = \lim_{n\to\infty}\frac{\sin\alpha}{P_n} = \lim_{n\to\infty}2^n\sin\frac{\alpha}{2^n} = \alpha,$$

這就是所需的公式。

$\pi/2$ 的公式

取 $\alpha = \pi/2$，它不是 π 的整數倍，並留意 $\sin\alpha = 1$。於是

$$\frac{\pi}{2} = \frac{1}{\cos\frac{\pi}{4}\cos\frac{\pi}{8}\cos\frac{\pi}{16}\cdots}.$$

餘下要證明的，是這些餘弦正是那些嵌套平方根（nested square roots）。對 $k \geq 1$，記 $c_k = \cos(\pi/2^{k+1})$。那麼 $c_1 = \cos\frac{\pi}{4} = \sqrt{\frac{1}{2}}$。在半角公式（half-angle formula）$\cos^2\theta = \frac{1}{2} + \frac{1}{2}\cos 2\theta$ 中取 $\theta = \pi/2^{k+2}$，得

$$c_{k+1}^2 = \tfrac{1}{2} + \tfrac{1}{2}c_k, \qquad \text{所以} \qquad c_{k+1} = \sqrt{\tfrac{1}{2} + \tfrac{1}{2}c_k},$$

這裏取正根，是因為 $0 < \pi/2^{k+2} < \pi/2$，所以 $c_{k+1} > 0$。因此

$$c_1 = \sqrt{\tfrac{1}{2}}, \quad c_2 = \sqrt{\tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2}}}, \quad c_3 = \sqrt{\tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2}}}}, \quad \ldots,$$

而每個因子都是把前一個因子代入 $t \mapsto \sqrt{\frac{1}{2} + \frac{1}{2}t}$ 所得。這些正好就是所需乘積中的各個因子，所以

$$\frac{\pi}{2} = \frac{1}{\sqrt{\frac{1}{2}}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}}}\cdots}.$$

驗算

- 取 $\alpha = \pi/2$ 及 $n = 1$，有限乘積的公式給出 $\cos\frac{\pi}{4} = \dfrac{1}{2\sin(\pi/4)} = \dfrac{1}{\sqrt{2}}$，結果正確。
- 嵌套根式（nested radicals）各部分積的倒數依次為 1.4142、1.5307、1.5607、1.5683、1.5702,…；乘到第 60 個因子後得到 1.5707963267949，在機器精度之內與 $\pi/2$ 吻合。
- 對隨機選取的 α 以及不超過 14 的 n，有限乘積的恆等式在數值上同樣成立；而對 $\alpha = 2.5$ 和 $\alpha = -4$ 等數值，$\sin\alpha / P_{60}$ 也能重現 α。

常見錯誤

- 除以 $\sin(\alpha/2^n)$ 而沒有說明它為何不為零。這正是設下條件 $\alpha \neq k\pi$ 的原因。
- 寫出 $1/\lim P_n = \lim(1/P_n)$ 而沒有指出 $\lim P_n \neq 0$。
- 在半角公式一步中，忘記說明為何取正平方根。

STEP 之外

最後的公式就是 1593 年的韋達（Viète）乘積，是歷史上第一個寫下來的 π 的精確無窮表達式。到了大學，你會在分析學（analysis）中正式學習無窮乘積：若一個由非零因子組成的乘積，其部分積趨向一個非零的極限，便稱該乘積收斂（converge），而這正是我們剛才檢查過的一點。歐拉（Euler）的乘積 $\sin\alpha = \alpha\prod_{k\geq 1}\left(1 - \frac{\alpha^2}{k^2\pi^2}\right)$ 則是與它一脈相

承而更為深刻的結果。Abbott 的 *Understanding Analysis* 是一本平易近人的書，可以從中看到這些極限如何得到嚴謹的處理。

詳細解答

STEP I 1990，第 9 題

純數學 • 切線與面積的界限

> **STEP I 1990，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/90-s1-q9

關鍵思路

曲線 $y = 1/x$ 向上彎曲（即它是凸（convex）的），所以每一條切線（tangent）都位於曲線之下。因此，兩條切線下方的兩個梯形（trapezium）都包含在曲線下方由 $x = 1$ 至 $x = b$ 的區域之內，而該區域的面積為 $\ln b$。比較這兩個面積，便得到對數的下界（lower bound）。同樣道理，一個位於曲線之上的長方形則給出上界（upper bound）。

解答

點 C

對於 $y = 1/x$，我們有 $\frac{\mathrm{d}y}{\mathrm{d}x} = -1/x^2$。所以在 $(t, 1/t)$ 的切線為

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

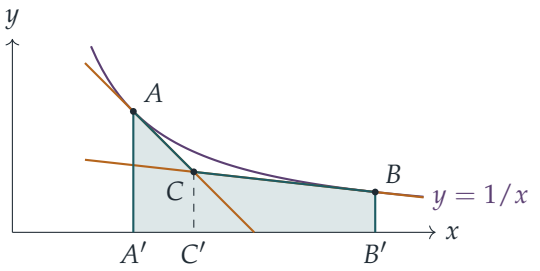
取 $t = 1$，在 A 的切線為 $y = 2 - x$。取 $t = b$，在 B 的切線為 $y = \frac{2}{b} - \frac{x}{b^2}$。在兩線相交之處，

$$2 - x = \frac{2}{b} - \frac{x}{b^2} \quad \Longrightarrow \quad x\left(1 - \frac{1}{b^2}\right) = 2\left(1 - \frac{1}{b}\right).$$

由於 $b > 1$，我們可以除以 $1 - \frac{1}{b} \neq 0$，餘下 $x\left(1 + \frac{1}{b}\right) = 2$。因此

$$C = \left(\frac{2b}{b+1}, \frac{2}{b+1}\right),$$

其中 y 坐標由 $y = 2 - x$ 求得。注意 $1 < \frac{2b}{b+1} < b$：左邊的不等式（inequality）即 $b + 1 < 2b$，右邊的即 $2 < b + 1$。兩者皆因 $b > 1$ 而成立，所以 C 在水平方向上位於 A 與 B 之間。



此圖按 $b = 3$ 繪畫。陰影區域由該兩個四邊形（quadrilateral）組成。

面積之和

每個四邊形都有兩條鉛垂的邊，所以它是梯形，其面積等於闊度乘以兩條鉛垂邊的平均值。

對於 $ACC'A'$，闊度為 $\frac{2b}{b+1} - 1 = \frac{b-1}{b+1}$，而兩條鉛垂邊分別為 1 和 $\frac{2}{b+1}$：

$$\text{面積}(ACC'A') = \frac{1}{2} \cdot \frac{b-1}{b+1}\left(1 + \frac{2}{b+1}\right) = \frac{(b-1)(b+3)}{2(b+1)^2}.$$

對於 $CBB'C'$，闊度為 $b - \frac{2b}{b+1} = \frac{b(b-1)}{b+1}$，而兩條鉛垂邊分別為 $\frac{2}{b+1}$ 和 $\frac{1}{b}$：

$$\text{面積}(CBB'C') = \frac{1}{2} \cdot \frac{b(b-1)}{b+1}\left(\frac{2}{b+1} + \frac{1}{b}\right) = \frac{(b-1)(3b+1)}{2(b+1)^2}.$$

兩式相加，分子合併為 $(b-1)(4b+4) = 4(b-1)(b+1)$，所以

$$\text{面積}(ACC'A') + \text{面積}(CBB'C') = \frac{2(b-1)}{b+1}.$$

不等式

想一想 為甚麼兩個梯形必定包含在曲線之下？

對於 $x > 0$，曲線與在 $(t, 1/t)$ 的切線之間的差距為

$$\frac{1}{x} - \left(\frac{2}{t} - \frac{x}{t^2}\right) = \frac{t^2 - 2tx + x^2}{t^2 x} = \frac{(x-t)^2}{t^2 x} \geq 0.$$

所以對所有 $x > 0$，每條切線都位於曲線上或曲線之下。陰影區域的上邊界，在 $[1, \frac{2b}{b+1}]$ 上是在 A 的切線，在 $[\frac{2b}{b+1}, b]$ 上是在 B 的切線。兩段都位於 $y = 1/x$ 之下，而該區域位於 x 軸之上。所以陰影區域包含在曲線下方由 $x = 1$ 至 $x = b$ 的區域之內，而

$$\frac{2(b-1)}{b+1} \leq \int_1^b \frac{\mathrm{d}x}{x} = \ln b.$$

至於上界，當 $x \geq 1$ 時 $\frac{1}{x} \leq 1$，所以曲線下方的區域位於以 $A'B'$ 為底、高為 1 的長方形之內。因此

$$\ln b = \int_1^b \frac{\mathrm{d}x}{x} \leq \int_1^b 1\,\mathrm{d}x = b - 1.$$

現設 $z > 0$ 並取 $b = 1 + z$，於是 $b > 1$，符合要求。這時 $b - 1 = z$ 及 $b + 1 = 2 + z$，兩個不等式合起來就是

$$\frac{2z}{2+z} \leq \ln(1+z) \leq z,$$

這正是我們要證明的結果。（事實上，當 $z > 0$ 時兩個不等式都是嚴格的，因為曲線與作為邊界的直線只在孤立的點相交。）

驗算

- 當 $b = 1$ 時，C 會是 $(1,1) = A$，而面積之和會是 0，與預期相符。
- 當 $b = 2$ 時：$C = (\frac{4}{3}, \frac{2}{3})$，面積之和為 $\frac{2}{3} \approx 0.667$，僅略小於 $\ln 2 \approx 0.693$，而後者又小於 1。
- 對於細小的 z，$\frac{2z}{2+z} = z - \frac{z^2}{2} + \cdots$ 與 $\ln(1+z) = z - \frac{z^2}{2} + \cdots$ 直至 z^2 項都一致，所以下界在 0 附近非常精確。

我們以精確運算對若干個 b 值驗算了坐標和兩個梯形的面積，並對 10^5 個介乎 10^{-6} 與 10^6 之間的隨機 z 值驗算了該不等式。

常見錯誤

- 忘記檢查 C 是否位於 $x = 1$ 與 $x = b$ 之間。缺少這一步，那些「四邊形」未必是你所想像的形狀。
- 聲稱「從圖中可見」切線位於曲線之下。用 $(x - t)^2$ 作出的那一行計算，才令它成為一個證明。

STEP 之外

用簡單的圖形從上方和下方與函數的圖像作比較，正是黎曼和（Riemann sum）背後的想法，也是大學第一門分析（analysis）課程中嚴格定義積分的方法。凸函數（convex function）的切線位於其圖像之下，這個事實是凸函數理論的起點。由此可推出延森不等式（Jensen's inequality）以及許多經典不等式，包括算術平均—幾何平均不等式（AM–GM）。Abbott 的 *Understanding Analysis*（一本分析入門教科書）正是按這種精神建立積分理論的。

詳細解答

STEP I 1990，第 10 題

力學 • 功率與阻力

STEP I 1990，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/90-s1-q10

關鍵思路

「穩定速率」（steady speed）即是沒有加速度，所以跑手的驅動力（driving force）恰好與各阻力平衡，而功率（power）等於驅動力乘以速率。唯一要留神的地方，是空氣阻力（air resistance）取決於相對於空氣的速率，而這個速率在去程和回程並不相同。至於題目的後半部分，應該考慮總能量（energy）而不是功率：每一程所消耗的能量等於阻力乘以距離，於是固定的能量預算就化為兩個速率之間的一個線性條件。

解答

穩定速率與完成賽事的時間

把跑手看成沿直線運動的質點（particle）。設 v（單位為 $\mathrm{m\,s^{-1}}$）為跑手在某一程的對地速率，D（單位為 N）為跑手產生的向前驅動力。阻力分別為 50 N 和 $2w$ N，兩者分別阻礙相對於地面和相對於空氣的運動。



沿跑道方向分解各力，由於以穩定步速前進時加速度為零，所以

$$D = 50 + 2w.$$

跑手作功的速率等於驅動力乘以其作用點的速率，所以 $300 = Dv$。

去程。風以 $3\,\mathrm{m\,s^{-1}}$ 的速率順着跑手的運動方向吹。若 $v > 3$，跑手相對於空氣以 $w = v - 3$ 向前運動，空氣阻力向後作用。（下文會驗證 $v > 3$。）於是 $D = 50 + 2(v-3) =$

$44+2v$，而

$$v_1(44+2v_1)=300 \iff v_1^2+22v_1-150=0 \iff v_1=-11+\sqrt{271}\approx 5.462,$$

這裏取正根。這個值確實大於 3，所以假設成立。（另一個看法：當 $v=3$ 時沒有空氣阻力，所需功率只是 $3\times 50=150<300\,\mathrm{W}$，所以跑手必定跑得比風快。）

回程。這時跑手逆風而跑，所以 $w=v+3$，而 $D=50+2(v+3)=56+2v$：

$$v_2(56+2v_2)=300 \iff v_2^2+28v_2-150=0 \iff v_2=-14+\sqrt{346}\approx 4.601.$$

時間。忽略起跑和折返時的加速過程，

$$T_0=\frac{5000}{v_1}+\frac{5000}{v_2}\approx 915.4+1086.7=2002.1,$$

所以完成賽事需時 2002 s（準確至最接近的秒），即大約 33 分 22 秒。

另一種策略

現在假設跑手去程以 $x_1\,\mathrm{W}$、回程以 $x_2\,\mathrm{W}$ 的功率作功，穩定速率分別為 v_1 和 v_2。力的平衡關係不變，所以只要 $v_1>3$，

$$x_1=v_1(44+2v_1), \qquad x_2=v_2(56+2v_2).$$

去程所需時間為 $5000/v_1$，所以去程消耗的能量為

$$x_1\cdot\frac{5000}{v_1}=5000(44+2v_1),$$

這正好是阻力乘以距離，理應如此。同樣地，回程消耗 $5000(56+2v_2)$。因此總能量為

$$E=5000\big(100+2(v_1+v_2)\big).$$

想一想 固定的總能量在哪裏派上用場？

跑手消耗的總能量必須與原來的賽事相同，所以 E 是一個固定的數。由於 E 只透過 v_1+v_2 取決於兩個速率，這個和就被確定了：

$$v_1+v_2=\frac{E}{10\,000}-50,$$

無論 x_1 和 x_2 怎樣選取都是如此。這就證明了所求結果。它的值就是原來賽事中的值，

$$S=v_1+v_2=\sqrt{271}+\sqrt{346}-25\approx 10.063.$$

最佳策略。現在完成賽事的時間為

$$T=5000\left(\frac{1}{v_1}+\frac{1}{v_2}\right)=\frac{5000\,(v_1+v_2)}{v_1v_2}=\frac{5000\,S}{v_1v_2}.$$

由於 S 固定，當 $v_1 v_2$ 最大時這個時間最小。根據算術—幾何平均不等式（AM–GM inequality），$v_1 v_2 \le \left(\frac{S}{2}\right)^2$，等號恰好在 $v_1 = v_2 = \frac{S}{2}$ 時成立。所以最佳方案是來回都以相同的速率跑，即 $\frac{S}{2} \approx 5.032\,\mathrm{m\,s^{-1}}$。這仍然比風快，所以去程中 x_1 的公式依然有效。最短時間為

$$T_{\min} = \frac{5000\,S}{S^2/4} = \frac{20\,000}{S} \approx 1987.4\ \mathrm{s}.$$

跑手節省了 $T_0 - T_{\min} \approx 2002.1 - 1987.4 = 14.7\,\mathrm{s}$，即大約 15 秒，正如題目所求。所用功率約為去程 $x_1 \approx 272\,\mathrm{W}$、回程 $x_2 \approx 332\,\mathrm{W}$：跑手在順風時放鬆一點，逆風時則加把勁。

驗算

- 能量是否一致：$300 \times 2002.105 \approx 600\,632\,\mathrm{J}$，而 $5000(100 + 2 \times 10.0632) \approx 600\,632\,\mathrm{J}$ 亦然。
- 若沒有風，兩程都會得出 $v^2 + 25v - 150 = 0$，所以 $v = 5$，兩種方法就會完全一致。正是有風，才有可能節省時間。
- 我們以數值方法驗證過：隨機選取 1000 個 x_1 的值（x_2 由能量預算確定），都有 $v_1 + v_2 = 10.0632$；沒有任何選擇能快過 1987.4 s；節省的時間為 14.66 s。

常見錯誤

- 在空氣阻力一項中用了對地速率，而不是相對於空氣的速率；又或者在其中一程把風的正負號弄反了。
- 以為來回用相同的*功率*才算合理：其實是來回用相同的*速率*才令時間最短。
- 試圖單憑功率方程證明 $v_1 + v_2$ 為常數。你需要用到總能量固定這個條件。

STEP 之外

最後一部分是一個小型的約束優化（constrained optimisation）問題：在一個線性約束條件下，求 $1/v_1 + 1/v_2$ 的最小值。這裏用 AM–GM 不等式就能解決，但一般的工具是大學一年級課程會學到的拉格朗日乘數法（method of Lagrange multipliers），而同樣的想法也支配着單車和賽艇運動中的最佳配速。若想看對功、功率和阻力的嚴謹論述，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）。

詳細解答

STEP I 1990，第 11 題

力學 • 拋體與動量

STEP I 1990，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/90-s1-q11

關鍵思路

整個過程分為三個階段：先是普通的拋體（projectile）運動，直至飛行的最高點；然後是一次瞬時碰撞（接住炮彈）；最後是一次瞬時「爆炸」（推開炮彈）。在兩個瞬時階段中，重力來不及起作用，所以炮彈與超人組成的系統動量守恆（conservation of momentum），而且在每個方向上各自守恆。此後炮彈再次成為拋體，從已知的高度以已知的向下速率出發。

解答

我們把炮彈和超人都視為質點（particle），忽略空氣阻力，並假設炮彈從水平地面發射，最後亦落在水平地面上。取 x 為沿運動方向的水平方向，y 為垂直向上。

直至飛行的最高點

炮彈初速度的分量為

$$u_x = v\cos\tfrac{\pi}{3} = \tfrac{1}{2}v, \qquad u_y = v\sin\tfrac{\pi}{3} = \tfrac{\sqrt{3}}{2}v.$$

在最高點，垂直速度為零，所以所需時間和到達的高度分別是

$$t_1 = \frac{u_y}{g} = \frac{\sqrt{3}\,v}{2g}, \qquad H = \frac{u_y^2}{2g} = \frac{3v^2}{8g}.$$

此刻炮彈正以速率 $\frac{1}{2}v$ 水平運動。

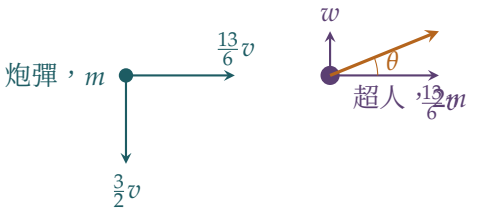
接住炮彈

超人以 $3v$ 的速率水平飛行，從後方接住炮彈，然後兩者一起運動。在（瞬時的）接住過程中水平動量守恆，所以若 V 是兩者的共同速率，則

$$2m(3v) + m\left(\tfrac{1}{2}v\right) = 3mV \quad \Longrightarrow \quad V = \frac{13v}{6}.$$

此刻兩者都沒有垂直速度，所以合起來的垂直動量為零。

推開炮彈



超人現在把炮彈垂直向下推。這一推是內部衝量（impulse），所以兩者的總動量在兩個方向上都保持不變。

水平方向：炮彈保持水平速率 $\frac{13}{6}v$。總水平動量仍然是 $3m \cdot \frac{13}{6}v$，所以超人的水平速率同樣保持為 $\frac{13}{6}v$。（這並不意外：衝量是垂直方向的。）

垂直方向：總垂直動量原本為零，所以若超人以速率 w 向上運動，則

$$2m\,w - m\left(\tfrac{3}{2}v\right) = 0 \quad \Longrightarrow \quad w = \tfrac{3}{4}v.$$

想一想 重力自始至終都在作用，為甚麼我們仍然可以使用動量守恆？

接住和推開都是瞬時發生的。重力是有限的力，所以在長度為零的時間區間內，它的衝量為零；真正起作用的只有超人與炮彈之間巨大而短暫的內力，而這些內力成對地互相抵消。

炮彈餘下的飛行

炮彈現在從高度 $H = \frac{3v^2}{8g}$ 出發，垂直向下的速率為 $\frac{3}{2}v$。取 s 以向下為正，炮彈經過時間 t_2 後到達地面，其中

$$\tfrac{3}{2}v\,t_2 + \tfrac{1}{2}gt_2^2 = \frac{3v^2}{8g}.$$

兩邊乘以 $8g$，得到一個關於 gt_2 的二次方程：

$$4(gt_2)^2 + 12v(gt_2) - 3v^2 = 0 \quad \Longrightarrow \quad gt_2 = \frac{-12v + \sqrt{144v^2 + 48v^2}}{8} = \frac{-12 + 8\sqrt{3}}{8}\,v,$$

由於 $t_2 > 0$，這裏取正根。所以

$$t_2 = \left(\sqrt{3} - \tfrac{3}{2}\right)\frac{v}{g}.$$

（這是正數，因為 $\sqrt{3} \approx 1.732 > 1.5$。）總飛行時間為

$$\begin{aligned} T = t_1 + t_2 &= \frac{\sqrt{3}}{2}\frac{v}{g} + \left(\sqrt{3} - \frac{3}{2}\right)\frac{v}{g} \\ &= \frac{3(\sqrt{3}-1)}{2}\frac{v}{g} \;\approx\; 1.10\,\frac{v}{g}. \end{aligned}$$

超人的運動方向

緊接推開炮彈之後，超人的速度分量為水平方向 $\frac{13}{6}v$ 和向上 $\frac{3}{4}v$。因此他的軌跡起初與水平成仰角 θ，其中

$$\tan\theta = \frac{\frac{3}{4}v}{\frac{13}{6}v} = \frac{9}{26}, \qquad \theta = \arctan\frac{9}{26} \approx 19.09^\circ,$$

所以該角準確至最接近的度數是 19°。

驗算

- 若沒有超人，炮彈會在時間 $2t_1 = \sqrt{3}\,v/g \approx 1.73\,v/g$ 落地。把炮彈向下推理應縮短飛行時間，而 $T \approx 1.10\,v/g$ 的確較小。
- 質量比：炮彈的質量是超人的一半，所以它得到的垂直速率應是超人的兩倍，方向相反。事實正是如此：$\frac{3}{2}v = 2 \times \frac{3}{4}v$。
- 數值驗證：取 $v = 13.7$ 和 $g = 9.81$，以 10^{-6} 為步長逐步計算炮彈從高度 H 下落的過程，得 $t_2 \approx 0.32407\,\mathrm{s}$，與 $(\sqrt{3} - \frac{3}{2})v/g$ 吻合，而 $T \approx 1.5335\,\mathrm{s} = 1.0981\,v/g$。

常見錯誤

- 忘記了接住炮彈這一步：推開之後沒有使用共同速率 $\frac{13}{6}v$，而是用了炮彈本身的水平速率 $\frac{1}{2}v$，或超人的 $3v$。
- 忽略質量，以為超人向上的速率是 $\frac{3}{2}v$。
- 取了二次方程的負根，或把 t_2 加到 $2t_1$ 而不是 t_1 上。

STEP 之外

接住炮彈是一次完全非彈性碰撞（perfectly inelastic collision），而推開炮彈則是它的逆過程，即一次爆炸：同一條守恆定律在兩個方向上都適用，只是動能（kinetic energy）在前者中損失，在後者中得到補充。在大學一年級的課程中，這會發展成衝力（impulsive force）和系統質心（centre of mass）的一般理論：無論系統各部分之間互相推撞得多麼猛烈，質心的運動只取決於外力。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）正是本着這種精神處理動量、衝量和質心。

詳細解答

STEP I 1990，第 12 題

力學 • 彈簧與能量

STEP I 1990，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：

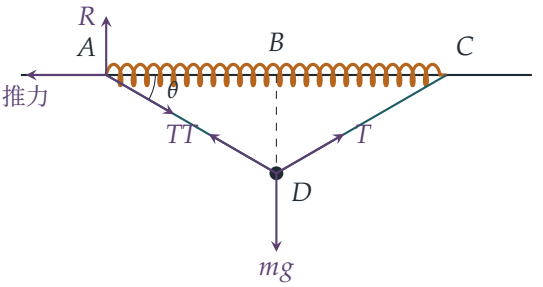
step.maths.org/questions/90-s1-q12

關鍵思路

這裏除了質點之外，所有物件都是輕的，所以兩端 A 和 C 所受的力必須完全平衡：繩子對每一端的水平拉力，等於被壓縮的彈簧的推力。這就給出平衡條件。至於運動，能量是最自然的工具：唯一的動能是質點的動能，所以質點靜止的時刻，正正就是勢能（potential energy）回到初始值的時刻。

解答

模型。彈簧和繩子都是輕的，繩子不可伸長（inextensible），而桿是光滑的，所以桿只能沿垂直方向推彈簧。由於整個系統關於通過 B 的垂直線對稱，質點 D 一直位於 B 的正下方，且 $AD = DC = l$。設 T 為每半段繩子的張力（tension），θ 為每半段繩子與桿所成的角。



平衡

當 $\theta = \pi/6$ 時，AC 的一半長度是 $l\cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}l$，所以

$$AC = \sqrt{3}\,l < 2l,$$

彈簧被壓縮了 $(2-\sqrt{3})l$。根據虎克定律（Hooke's law），它對 A（以及對 C）施加向外的推力（thrust）

$$\frac{\lambda(2-\sqrt{3})l}{2l} = \lambda\Big(1-\frac{\sqrt{3}}{2}\Big).$$

在 D 沿垂直方向分解力：$2T\sin\frac{\pi}{6} = mg$，所以 $T = mg$。

在 A 沿水平方向分解力：

想一想 A 處明明甚麼也沒有，為甚麼可以在那裏列出力的平衡？

彈簧的末端是一個輕的點：質量為零的物體若受到任何淨力，加速度便會是無限大，所以它所受的力必須在每一刻都平衡（而不僅是在平衡狀態下）。桿是光滑的，所以它的反作用力（reaction）R 是垂直的，水平方向就只有繩子和彈簧在起作用：

$$T\cos\frac{\pi}{6} = \lambda\Big(1-\frac{\sqrt{3}}{2}\Big) \quad\Longrightarrow\quad mg\,\frac{\sqrt{3}}{2} = \lambda\,\frac{2-\sqrt{3}}{2},$$

所以

$$mg = \lambda\,\frac{2-\sqrt{3}}{\sqrt{3}} = \lambda\Big(\frac{2}{\sqrt{3}}-1\Big),$$

這正是所要證明的結果。

勢能

設 $AC = 2xl$，則繩子拉緊時 $0 < x \le 1$。AC 的一半是 xl，而 $AD = l$，所以由畢氏定理（Pythagoras' theorem），D 在桿下方的深度為

$$BD = l\sqrt{1-x^2}$$

。彈簧的伸長量是 $2xl - 2l$，所以它的彈性勢能(elastic energy)是 $\frac{\lambda(2l-2xl)^2}{2\cdot 2l} = \lambda l(1-x)^2$。取桿為重力勢能（gravitational potential energy）的零位，

$$V(x) = \lambda l(1-x)^2 - mgl\sqrt{1-x^2}.$$

作為一致性的檢查，

$$V'(x) = -2\lambda l(1-x) + \frac{mglx}{\sqrt{1-x^2}},$$

而在 $x = \frac{\sqrt{3}}{2}$ 處（這時 $\cos\theta = x$，所以 $\theta = \frac{\pi}{6}$），上式為零的充要條件正是 $mg\sqrt{3} = \lambda(2-\sqrt{3})$，也就是我們在上面求得的條件。

從 B 落下

起初質點位於 B，所以繩子沿着桿平放，$AC = 2l$ 而 $x = 1$。於是 $V(1) = 0$，系統由靜止開始運動。

桿是光滑的，繩子不可伸長，所以沒有能量損失；又因為彈簧和繩子都是輕的，唯一的動能（kinetic energy）就是質點的 $\frac{1}{2}mv^2$。因此在整個運動過程中

$$\tfrac{1}{2}mv^2 + V(x) = V(1) = 0.$$

> **想一想** 為甚麼「靜止下來」的意思是 $V(x) = 0$，而不是 $V'(x) = 0$？

質點靜止時 $v = 0$，根據能量守恆（energy conservation），這表示 $V(x) = 0$。條件 $V'(x) = 0$ 描述的是平衡位置，那是質點運動得最快的地方，而不是它停下來的地方。所以 x 滿足

$$\lambda(1-x)^2 = mg\sqrt{1-x^2}, \qquad \text{即} \qquad (1-x)^2 = \left(\frac{2}{\sqrt{3}} - 1\right)\sqrt{1-x^2}.$$

根 $x = 1$ 是起始位置。對於 $x < 1$，我們可以把兩邊除以 $\sqrt{1-x} > 0$ 再平方，得到一個較簡潔的三次方程（cubic）：

$$(1-x)^3 = k^2(1+x), \qquad k = \frac{2}{\sqrt{3}} - 1 \approx 0.15470, \quad k^2 \approx 0.023932.$$

這裏平方是穩妥的，因為 $(1-x)^{3/2} = k\sqrt{1+x}$ 的兩邊都是正數。設 $f(x) = (1-x)^3 - k^2(1+x)$。則

$$f(0.65) = 0.042875 - 0.039488 \approx +0.0034,$$
$$f(0.66) = 0.039304 - 0.039727 \approx -0.0004,$$

所以在 0.65 與 0.66 之間有一個根，而且非常接近 0.66。由於 $f'(x) = -3(1-x)^2 - k^2 < 0$，$f$ 是嚴格遞減（strictly decreasing）的，所以這是唯一滿足 $x < 1$ 的根。質點下一次靜止時 $x \approx 0.66$（更精確地說 $x \approx 0.6589$），這時 D 在桿下方的深度約為 $0.75l$。

驗算

- 在 $\theta = \pi/6$ 處，取 $\lambda = l = m = 1$ 及 $g = \frac{2}{\sqrt{3}} - 1$，$A$ 處的水平拉力和彈簧的推力都等於 0.13397，而 V 在 $x = \frac{\sqrt{3}}{2}$ 處的數值導數為零（且 $V'' > 0$，所以這個平衡是穩定的）。
- 用二分法（bisection）求得根 $x = 0.658861$；另外，就 D 的深度 y，由靜止於 $y = 0$ 開始，以 RK4 逐步計算運動方程 $m\ddot{y} = mg - 2\lambda(1-x)\,y/(lx)$，結果顯示質點首次停下時 $x = 0.658861$，兩者吻合。

- 平衡時的值 $x = \frac{\sqrt{3}}{2} \approx 0.866$ 位於 0.66 與 1 之間，這是必然的：質點圍繞平衡位置振盪。

常見錯誤

- 以為彈簧是被拉長的。這裏 $AC < 2l$，所以彈簧是被壓縮的，它把 A 和 C 向外推開。
- 以 $V'(x) = 0$ 來求靜止位置。這樣求得的是平衡位置，那裏速率最大。
- 忘記了重力勢能在桿的下方是負的，結果得出一條在 $(0, 1)$ 內無解的能量方程。

STEP 之外

「無質量的物體所受的淨力必為零」這個想法是一個約束（constraint），而不是一條運動定律。這類問題用拉格朗日力學（Lagrangian mechanics）處理最為乾淨利落：你只需用單一個坐標（這裏是 x）寫下一個能量函數，約束力根本不會出現。勢能 $V(x)$ 還可以透過 V'' 給出圍繞平衡位置的小振盪（small oscillations）的週期。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）正是以這種精神講解能量方法和小振盪。

詳細解答

STEP I 1990，第 13 題

力學 • 摩擦力與剛體平衡

STEP I 1990，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：

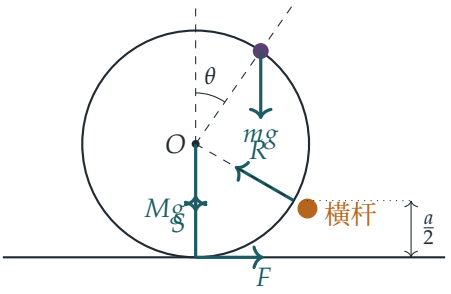
step.maths.org/questions/90-s1-q13

關鍵思路

先處理質點（particle）：它只是一個置於粗糙斜面上的普通質點，而該斜面就是圓柱體在接觸點處的切平面（tangent plane）。然後留意，由於質點處於平衡（equilibrium），它施於圓柱體的總力就只是它的重量，並且通過接觸點鉛垂向下作用。對圓柱體的中心取力矩（moments），除了這個力和來自平面的摩擦力（friction）之外，其餘各力全部消失，這是最快捷的入手方法。

解答

把質點視為質量為 m 的點質量，把圓柱體視為質量為 M 的均勻剛體（rigid body）圓柱，其質心（centre of mass）O 位於軸上。橫杆是光滑的，所以它只施加法向力；平面和圓柱體表面則是粗糙的，摩擦系數（coefficient of friction）為 μ。一切都發生在通過 O 的鉛垂橫截面內，以下我們只考慮這個截面。把橫杆放在右邊。



圖中顯示作用於圓柱體的各個力（質點位置的 mg 會在第二部分解釋）。箭咀並非按比例繪畫。

質點的平衡

在接觸點 P，向外的法線與向上的鉛垂線成 θ 角，所以該處的切平面是一個與水平成 θ 角的斜面。作用於質點的力有：它的重量 mg、沿向外法線方向的法向反作用力（normal reaction）N，以及沿切線、指向斜面上方的摩擦力 f。沿法線方向及垂直於法線方向分解，

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

只要 $f \le \mu N$，摩擦力便足以提供這個力，即

$$\tan\theta \le \mu, \qquad \text{等價於} \qquad \theta \le \arctan\mu.$$

（我們需要 $0 \le \theta < 90^\circ$ 才能使 N 為正，而由於質點位於圓柱體的上半部，這是成立的。）

放手後圓柱體的平衡

> ? **想一想** 質點以一個法向力和一個摩擦力壓在圓柱體上。我們需要分開處理這兩個力嗎？

不需要。根據牛頓第三定律（Newton's third law），圓柱體在 P 點受到 $-N$ 和 $-f$。由於 N 與 f 合起來抵銷質點的重量，它們的和是鉛垂向上的 mg，所以質點施於圓柱體的總力是*鉛垂向下、作用於 P 點*的 mg。它的作用線與 O 的水平距離為 $a\sin\theta$，位於橫杆的一側。

作用於圓柱體的其他力有：

- 它的重量 Mg，作用於 O；
- 橫杆的反作用力 R，垂直於圓柱體表面，因此沿半徑指向 O。橫杆的接觸點高度為 $a/2$，即在 O 之下 $a/2$，所以這條半徑與向下的鉛垂線成 60°。因此 R 有一個背離橫杆的水平分量 $\frac{\sqrt{3}}{2}R$ 和一個向上的分量 $\frac{1}{2}R$；
- 在最低點，有向上的法向反作用力 S 和水平方向的摩擦力 F（取其方向為指向橫杆）。

對 O 取力矩。Mg、R 和 S 的作用線全部通過 O。質點的 mg 的力矩為 $mg\,a\sin\theta$（使圓柱體轉向橫杆），而 F 作用於 O 之下距離 a 處，轉向相反。所以

$$Fa = mg\,a\sin\theta \quad \Longrightarrow \quad F = mg\sin\theta.$$

這是正數，所以摩擦力確實如圖所示指向橫杆。（若沒有橫杆，圓柱體會向橫杆那一側滾動；這就是為甚麼質點必須在那一側，橫杆才會有用。）

水平方向分解：$\frac{\sqrt{3}}{2}R = F$，所以

$$R = \frac{2mg\sin\theta}{\sqrt{3}} > 0,$$

這與橫杆只能推而不能拉相符。

鉛垂方向分解：

$$S = (M+m)g - \tfrac{1}{2}R = (M+m)g - \frac{mg\sin\theta}{\sqrt{3}},$$

由於 $\sin\theta < 1 < \sqrt{3}$ 且 $M > 0$，這是正數，所以圓柱體不會離開平面。

只要 $F \le \mu S$，平面便能提供所需的摩擦力：

$$m\sin\theta \le \mu\left[(M+m) - \frac{m\sin\theta}{\sqrt{3}}\right] \iff m\sin\theta\left(\sqrt{3}+\mu\right) \le \sqrt{3}\,\mu\,(M+m).$$

所以，放手之後，只要 $\tan\theta \le \mu$ 與

$$\boxed{\sin\theta \le \frac{\sqrt{3}\,\mu\,(M+m)}{m\,(\sqrt{3}+\mu)}}$$

同時成立，兩個物體便能保持平衡。（光滑的橫杆處不需要摩擦力，而質點與圓柱體之間的摩擦力已在第一部分處理。）

最後部分

當 $m = 7M$ 及 $\mu = \frac{3}{4}$ 時，質點的條件是 $\tan\theta \le \frac{3}{4}$，即 $\sin\theta \le \frac{3}{5}$。圓柱體的條件是

$$\sin\theta \le \frac{\sqrt{3}\cdot\frac{3}{4}\cdot 8M}{7M\,(\sqrt{3}+\frac{3}{4})} = \frac{24\sqrt{3}}{7(4\sqrt{3}+3)} = \frac{24\sqrt{3}\,(4\sqrt{3}-3)}{7\cdot 39} = \frac{24(4-\sqrt{3})}{91}.$$

哪一個較嚴格？我們有

$$\frac{24(4-\sqrt{3})}{91} < \frac{3}{5} \iff 120(4-\sqrt{3}) < 273 \iff 207 < 120\sqrt{3} \iff 42849 < 43200,$$

這是成立的。所以圓柱體在平面上滑動才是真正起限制作用的條件（僅僅如此），而能保持平衡的最大 θ 值是

$$\theta_{\max} = \arcsin\frac{24(4-\sqrt{3})}{91} \approx 36.7^\circ,$$

僅略小於 $\arctan\frac{3}{4} \approx 36.9^\circ$。

驗算

- 極限情況 $M \to \infty$：圓柱體的條件變得形同虛設，因為非常重的圓柱體總能靠摩擦力固定；只剩下 $\tan\theta \le \mu$。
- 當 $\theta \to 0$ 時，質點位於頂部，它的重量通過 O，而 $F = R = 0$：正如所料，不需要橫杆或摩擦力。

- 數值上，隨機選取 20 000 組 M, m, a, μ, θ，直接由原始的力向量解出三條平衡方程，結果每次都與方框內的條件吻合；當 $m = 7M$、$\mu = 0.75$ 時，對 θ 作精細掃描，得到 $\theta_{\max} \approx 36.737^\circ$，而 $\sin\theta_{\max} \approx 0.59814$。

常見錯誤

- 把橫杆的力當作水平方向。它垂直於圓柱體表面，所以沿半徑方向，在這裏是與水平成 30° 向上。
- 忘記橫杆的力有向上的分量 $\frac{1}{2}R$，它會減小 S，從而減小平面所能提供的摩擦力。
- 做到 $\theta = \arctan\frac{3}{4}$ 便停下來，沒有與圓柱體的條件比較，而後者（稍為）更嚴格。

STEP 之外

這裏用到的技巧，是把質點的接觸力換成一個等於其重量的合力（resultant）。由此可以初步窺見一個想法：作用於剛體的任何力系，都可以化簡為通過某個選定點的單一力再加上一個力偶（couple）。具有多個摩擦接觸點的剛體靜力學（statics）是大學一年級的標準課題，而「哪一個接觸點先滑動」這個問題，會引申到更微妙的情況：摩擦力並不能單靠方程確定。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）對力、力矩和剛體平衡有嚴謹的論述。

詳細解答

STEP I 1990，第 14 題

概率與統計 • 期望值與最優停止

STEP I 1990，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/90-s1-q14

關鍵思路

由於每次抽出的球都會放回，各次抽球互相獨立（independent），所以每次抽到白球、紅球或黑球的概率都固定不變。對 Juggins 而言，我們只需知道連續三次抽球都「存活」的機會，以及一次存活抽球的平均得分。對 Muggins 而言，問題在於取捨：多抽一次，平均可多得少許分數，卻要冒失去已累積的全部分數的風險，而這個風險隨 N 增大而增大。比較「再抽一次的期望得分」與「就此停手的得分」，便得出一個臨界值（threshold）。

解答

在任何一次抽球中，設 W、R、B 分別為抽到白球、紅球、黑球的事件。由於球會放回，每次抽球都與其他各次獨立，且

$$\mathrm{P}(W)=\frac{5}{10}=\frac{1}{2},\qquad \mathrm{P}(R)=\frac{3}{10},\qquad \mathrm{P}(B)=\frac{2}{10}=\frac{1}{5}.$$

Juggins 的策略

Juggins 最終得零分，當且僅當他首三次抽球中出現黑球。否則三次抽球全都不是黑球，他便停手。因此

$$\mathrm{P}(\text{最終得分為0})=1-\mathrm{P}(\text{三次抽球均非黑球})=1-\left(\frac{4}{5}\right)^3=1-\frac{64}{125}=\frac{61}{125}.$$

至於期望得分（expected score），設 S 為他的最終得分。若首三次抽球全都不是黑球（概率為 $\frac{64}{125}$），則 $S=X_1+X_2+X_3$，其中 X_i 是已知第 i 次抽球不是黑球時該次的得分。已知某次抽球不是黑球，則它是白球的概率為 $\frac{1/2}{4/5}=\frac{5}{8}$，是紅球的概率為 $\frac{3/10}{4/5}=\frac{3}{8}$。因此

$$\mathrm{E}(X_i)=1\cdot\frac{5}{8}+2\cdot\frac{3}{8}=\frac{11}{8}.$$

在其餘所有情況下 $S = 0$。所以

$$\mathrm{E}(S) = \tfrac{64}{125} \cdot 3 \cdot \tfrac{11}{8} = \tfrac{264}{125} = 2.112.$$

想一想 為甚麼這裏用 $\frac{5}{8}$ 和 $\frac{3}{8}$，而不是 $\frac{1}{2}$ 和 $\frac{3}{10}$？

我們已經以「存活」為條件（conditioning），即已知三次抽球每次都不是黑球，所以各概率必須重新按比例調整，使其總和為 1。換一個不用條件概率的說法：$\mathrm{E}(S)$ 是對 $2^3 = 8$ 個白／紅序列求（概率）×（得分）的總和。每次抽球的期望貢獻為 $1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{10} = \frac{11}{10}$，而只有當另外兩次抽球也存活時（概率為 $\frac{16}{25}$），這份貢獻才得以保留。由此同樣得出 $3 \cdot \frac{11}{10} \cdot \frac{16}{25} = \frac{264}{125}$。

Muggins 的抉擇

假設 Muggins 現有 N 分，並再抽一次（然後停手）。他的得分會變成 $N+1$、$N+2$ 或 0，所以

$$\mathrm{E}(\text{再抽一次後的得分}) = \tfrac{1}{2}(N+1) + \tfrac{3}{10}(N+2) + \tfrac{1}{5} \cdot 0 = \tfrac{4}{5}N + \tfrac{11}{10}.$$

平均而言，再抽一次勝過停手，當且僅當

$$\tfrac{4}{5}N + \tfrac{11}{10} > N \iff \tfrac{11}{10} > \tfrac{1}{5}N \iff N < 5.5.$$

抽一次球的平均得益是 $\frac{11}{10}$，平均損失則是 $\frac{1}{5}N$，即抽到黑球的機會乘以所押上的分數。由於得分是整數，這提示了以下策略：

得分爲 5 或以下時繼續抽球；一旦達到 6 或以上便立即停手。

想一想 只比較「多抽一次」與「停手」，看來有點短視。在 $N \geq 6$ 時先抽一次，並打算稍後再抽，會不會更有利？

不會。得分一旦達到 6 或以上，其後任何一次抽球都只會令得分仍然是 6 或以上（或以 0 分結束遊戲），所以在往後每一個決策點，「只看一次抽球」的比較仍然會說「停手」。從這樣的得分出發，每多抽一次都會降低最終得分的期望值，所以由此開始的最佳方案就是立即停手。在 6 分以下，多抽一次已經勝過停手，而可以繼續抽下去這個選擇只會錦上添花。因此這個臨界值法則不僅合理，更是最佳策略。

採用這個策略，設 $V(N)$ 為由 N 分開始的最終得分期望值。則當 $N \geq 6$ 時 $V(N) = N$，而當 $N \leq 5$ 時 $V(N) = \frac{1}{2}V(N+1) + \frac{3}{10}V(N+2)$。由上而下逐一計算：

N	6	5	4	3	2	1	0
$V(N)$	6	5.1	4.35	3.705	3.1575	2.690	2.292

所以 Muggins 平均每局約得 2.29 分，而 Juggins 則是 2.112 分。

驗算

- 以電腦精確地列舉了所有最多三次抽球的序列，確認 $\mathrm{P}(0) = \frac{61}{125} = 0.488$ 及 $\mathrm{E}(S) = \frac{264}{125}$。以蒙地卡羅（Monte Carlo）方法模擬 2×10^5 局，得出 0.488 和 2.111。
- 在 $N = 5$ 時，抽一次得 $\frac{4}{5} \cdot 5 + 1.1 = 5.1 > 5$。在 $N = 6$ 時則得 $5.9 < 6$。所以轉捩點在 5 與 6 之間。
- 對所有策略進行價值迭代（value iteration），得出相同的法則。若改為在 ≥ 5 或 ≥ 7 時停手，平均得分較低，分別為 2.261 和 2.265，而在 6 時停手則為 2.292。

常見錯誤

- 把 $\mathrm{E}(S)$ 算成 $3 \times \frac{11}{10}$，忘記了其後抽到黑球會把先前的分數一筆勾銷。
- 把條件平均值 $\frac{11}{8}$ 再乘以無條件概率，以致把存活因子重複計算了一次。
- 把臨界值寫成 $N \leq 5.5$，卻沒有指出 N 是整數，所以法則應是「5 分時抽，6 分時停」。

STEP 之外

Muggins 的問題是一個最優停止（optimal stopping）問題。「當現在停手勝過多走一步時便停手」這個法則稱為一步前瞻（one-step look-ahead）法則；只要像本題一樣，法則指示停手的那些狀態所組成的集合一經進入便無法離開，這個法則就是最優的。同樣的想法可引申至馬可夫決策過程（Markov decision processes）和動態規劃（dynamic programming）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門）涵蓋期望值和馬可夫鏈（Markov chain）的背景知識，而 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）則有許多類似的首步分析（first-step analysis）。

詳細解答

STEP I 1990，第 15 題

概率與統計 • 負二項分佈

STEP I 1990，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/90-s1-q15

關鍵思路

第一部分的做法，是把事件拆成兩個獨立（independent）的部分：最後一次擲幣的結果，以及之前所有擲幣的結果。至於其餘部分，關鍵在於硬幣沒有記憶。每當剛擲出一個正面，等待下一個正面的過程便重新開始，所以我們只需知道一個服從幾何分佈（geometric distribution）的等待時間有多大機會是偶數或奇數。

解答

以下一律記 $q = 1 - p$。各次擲幣互相獨立，每次擲出正面的概率為 p。

第 k 個正面的分佈

設 T_k 為第 k 個正面出現時的擲幣次數。事件 $\{T_k = n\}$ 發生，當且僅當

- 第 n 次擲幣是正面，並且
- 在第 $1, 2, \ldots, n-1$ 次擲幣之中恰好有 $k-1$ 個正面（因而有 $n-k$ 個反面）。

兩個條件缺一不可。假如第 n 次擲幣是反面，第 k 個正面便會在第 n 次擲幣之前或之後出現，而不是在這一次。假如頭 $n-1$ 次擲幣中的正面數目不同，第 n 次擲幣便不會給出第 k 個正面。反過來，兩個條件合起來亦已足夠。

這兩個事件涉及的擲幣次數互不重疊，所以它們是獨立的。第一個事件的概率是 p。至於第二個事件，那 $k-1$ 個正面在頭 $n-1$ 次擲幣中的位置共有 $\binom{n-1}{k-1}$ 種選法。每一種排列都是由 $k-1$ 個正面和 $n-k$ 個反面組成的一個特定序列，根據獨立性，其概率為

$p^{k-1}q^{n-k}$，而不同的排列是互斥（mutually exclusive）的。因此

$$\mathrm{P}(T_k = n) = p \cdot \binom{n-1}{k-1} p^{k-1}q^{n-k} = \binom{n-1}{k-1} p^k (1-p)^{n-k}, \qquad n \geq k,$$

這正是所要證明的結果。

由偶數到偶數

設 $G = T_k - T_{k-1}$ 為第 $(k-1)$ 個正面出現之後還需要再擲的次數（其中 $T_0 = 0$）。第 T_{k-1} 次擲幣之後的各次擲幣，與此前發生的一切互相獨立。所以 G 與 T_{k-1} 獨立，並且服從幾何分佈

$$\mathrm{P}(G = g) = q^{g-1}p, \qquad g = 1, 2, \ldots$$

（即第一部分中 $k = 1$ 的情況）。

> ? **想一想** 如果 T_{k-1} 是偶數，T_k 在甚麼情況下是偶數？

由於 $T_k = T_{k-1} + G$，如果 T_{k-1} 是偶數，則 T_k 是偶數當且僅當 G 是偶數。因為 G 與 T_{k-1} 獨立，以 T_{k-1} 為偶數作為條件（conditioning）並不會改變 G 的分佈。因此

$$\mathrm{P}(T_k \text{ 為偶數} \mid T_{k-1} \text{ 為偶數}) = \mathrm{P}(G \text{ 為偶數}).$$

把等比級數（geometric series）求和，

$$\mathrm{P}(G \text{ 為偶數}) = p\left(q + q^3 + q^5 + \cdots\right) = \frac{pq}{1-q^2} = \frac{pq}{(1-q)(1+q)} = \frac{q}{1+q},$$

$$\mathrm{P}(G \text{ 為奇數}) = p\left(1 + q^2 + q^4 + \cdots\right) = \frac{p}{1-q^2} = \frac{1}{1+q}.$$

所以所求的概率是

$$\frac{q}{1+q} = \frac{1-p}{2-p}.$$

答案與 k 無關。

恰好有兩個正面落在偶數次擲幣

現在一直擲到第三個正面出現為止。設 $G_1 = T_1$、$G_2 = T_2 - T_1$ 和 $G_3 = T_3 - T_2$ 為三段等待時間。根據同樣的無記憶性（memoryless）論證，它們互相獨立，而且各自服從上述的幾何分佈。記

$$e = \mathrm{P}(G_i \text{ 為偶數}) = \frac{q}{1+q}, \qquad o = \mathrm{P}(G_i \text{ 為奇數}) = \frac{1}{1+q}.$$

我們從第 0 次擲幣開始，這是偶數。間隔為偶數時，當前位置的奇偶性（parity）保持不變；間隔為奇數時，奇偶性便會改變。下表列出 (G_1, G_2, G_3) 的八種可能奇偶組合，以及相應的 (T_1, T_2, T_3) 的奇偶性：

G_1	G_2	G_3	T_1	T_2	T_3	偶數次的正面數目
偶	偶	偶	偶	偶	偶	3
偶	偶	奇	偶	偶	奇	**2**
偶	奇	偶	偶	奇	奇	1
偶	奇	奇	偶	奇	偶	**2**
奇	偶	偶	奇	奇	奇	0
奇	偶	奇	奇	奇	偶	1
奇	奇	偶	奇	偶	偶	**2**
奇	奇	奇	奇	偶	奇	1

其中有三種情況恰好有兩個正面落在偶數次擲幣，它們的概率分別是 e^2o、eo^2 和 o^2e。由於各段間隔互相獨立，所求的概率是

$$\begin{aligned} e^2o + 2eo^2 = eo(e+2o) &= \frac{q}{(1+q)^2}\cdot\frac{q+2}{1+q} \\ &= \frac{q(2+q)}{(1+q)^3} = \frac{(1-p)(3-p)}{(2-p)^3}. \end{aligned}$$

驗算

- 驗算第一部分：把 $\mathrm{P}(T_k = n)$ 對 $n \geq k$ 求和，結果應該是 1。這就是負二項級數 $\sum_{n\geq k}\binom{n-1}{k-1}q^{n-k} = (1-q)^{-k}$。
- 兩個奇偶概率 $\frac{q}{1+q}$ 與 $\frac{1}{1+q}$ 之和為 1。當 p 接近 1 時，第一個正面幾乎必定立即出現，即落在奇數次擲幣，所以 $\mathrm{P}(G \text{ 為偶數}) \to 0$，與公式給出的結果一致。
- 對於公平硬幣（$p = \frac{1}{2}$），最後的答案是 $\frac{(1/2)(5/2)}{(3/2)^3} = \frac{10}{27} \approx 0.370$。一次進行了 2×10^5 輪的模擬得出 0.369。直接對三段等待時間求和，所得結果在 $p = 0.3$、0.5 和 0.8 時均與公式吻合，誤差在 10^{-9} 以內。

常見錯誤

- 在第一部分使用 $\binom{n}{k}$。最後一次擲幣必定是正面，所以只有頭 $n-1$ 次擲幣可以自由安排。
- 在第二部分把負二項分佈對偶數的 n 求和。利用無記憶的間隔，便可完全避開這一步。
- 在最後一部分，計算落在偶數位置的正面時，把各個位置當作互相獨立。它們並不獨立，因為每個位置都是由前一個位置推算出來的。

STEP 之外

第 k 個正面出現的時間服從負二項（negative binomial）分佈。把它寫成若干個獨立幾何間隔之和，正是泊松（Poisson）過程的離散版本：泊松過程的到達時間，同樣是由獨立的指數等待時間構成的。最後一部分記錄奇偶性的做法，其實是一個雙狀態的馬可夫鏈（Markov chain），這是大學一年級概率課程中很早便會遇到的課題。Grimmett and Welsh 的 *Probability: An Introduction*（概率入門教科書）涵蓋了以上所有內容。

詳細解答

STEP I 1990，第 16 題

概率與統計 • 三角分佈與假設檢驗

STEP I 1990，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/90-s1-q16

關鍵思路

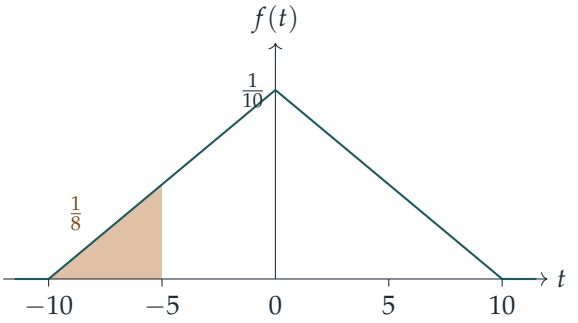
概率密度函數（probability density function）的圖像是一個三角形，所以我們需要的每一個概率都是某個小三角形的面積，而累積分佈函數（cumulative distribution function）會替我們完成所有工作。第二點要留意的是，只有 8 時的那班巴士才有可能在上午 8 時前後到站：7 時的那班巴士在 7.10 之前已經離開，而 9 時的那班巴士不可能在 8.50 之前到達。所以早上「趕上巴士」的真正意思是「比 8 時的那班巴士更早到站」。最後一部分於是變成一個只有一個參數的二項模型（binomial model）和一個顯著性檢定（significance test）。

解答

T 的密度函數

密度由 $t=-10$ 處的 0 線性上升至 $t=0$ 處的峯值，然後以相同的速率下降至 $t=10$ 處的 0，所以它在 $[-10,10]$ 上的形式是 $f(t)=k(10-|t|)$，在其他地方則為零。總面積等於一個底為 20、高為 $10k$ 的三角形的面積，所以 $\frac{1}{2}\cdot 20\cdot 10k=1$，由此得 $k=\frac{1}{100}$：

$$f(t)=\begin{cases}\dfrac{10+t}{100}, & -10\le t\le 0,\\ \dfrac{10-t}{100}, & 0\le t\le 10,\\ 0, & \text{其他情況。}\end{cases}$$



為方便稍後使用，先求左半部分的累積分佈函數：它是一個底為 $a+10$、高為 $f(a)$ 的三角形的面積，

$$\mathrm{P}(T<a)=\tfrac{1}{2}(a+10)\cdot\frac{a+10}{100}=\frac{(a+10)^2}{200},\qquad -10\le a\le 0,$$

而根據關於 $t=0$ 的對稱性，當 $0\le a\le 10$ 時，$\mathrm{P}(T>a)=\dfrac{(10-a)^2}{200}$。圖中陰影部分的面積是 $\mathrm{P}(T<-5)=\frac{25}{200}=\frac{1}{8}$。

在上午 8.10 或之前趕上巴士

我們假設我會在車站一直等到有巴士來為止。預定 8 時到站的巴士在上午 8 時之後 T 分鐘到達，其中 $-10\le T\le 10$；上一班巴士在 7.10 之前已經離開，而下一班不可能在 8.50 之前到達。所以，我能在 8.10 或之前趕上巴士，當且僅當我比 8 時的那班巴士更早到達車站，而這時 8.10 的期限自然會得到滿足。

如果我在上午 8 時之後 a 分鐘到達車站，那麼當 $T>a$ 時我便能趕上這班車（恰好出現 $T=a$ 的機會為零）。三種情況分別給出

$$\begin{aligned}
7{\cdot}55,\ a=-5:&\quad \mathrm{P}(T>-5)=1-\tfrac{25}{200}=\tfrac{7}{8},\\
7{\cdot}58,\ a=-2:&\quad \mathrm{P}(T>-2)=1-\tfrac{64}{200}=\tfrac{17}{25},\\
8{\cdot}02,\ a=2:&\quad \mathrm{P}(T>2)=\tfrac{64}{200}=\tfrac{8}{25}.
\end{aligned}$$

我在 7.55 到達車站的概率是 0.5，在 7.58 到達的概率是 $0.5\times 0.8=0.4$，在 8.02 到達的概率是 $0.5\times 0.2=0.1$。我睡過頭與否和巴士毫無關係，所以根據全概率公式（law of total probability），

$$\begin{aligned}
\mathrm{P}(\text{在 8.10 或之前趕上巴士})&=0.5\cdot\tfrac{7}{8}+0.4\cdot\tfrac{17}{25}+0.1\cdot\tfrac{8}{25}\\
&=0.4375+0.272+0.032=\frac{1483}{2000}=0.7415.
\end{aligned}$$

新時鐘是否不準？

首先假設時鐘是準確的，那麼我每天早上都在 7:55 到達車站。設 T_8 和 T_9 分別為 8 時和 9 時那兩班巴士的到站時間，以整點之後的分鐘數計算；它們互相獨立(independent)，各自的密度都是 f。我在上午 9 時之後仍在等車，當且僅當我錯過了 8 時的巴士（$T_8 < -5$）而且 9 時的巴士遲到（$T_9 > 0$）。所以在任何一個早上，

$$p = \mathrm{P}(T_8 < -5)\,\mathrm{P}(T_9 > 0) = \tfrac{1}{8} \cdot \tfrac{1}{2} = \tfrac{1}{16}.$$

> ? **想一想** 檢定應該朝哪個方向進行：5 次是「太多」還是「太少」？

時鐘調校錯誤，可以令 p 向任何一個方向偏移。假設時鐘快了 d 分鐘。那麼我到達車站的真實時間是 $7.55 - d$，而時鐘顯示 9.00 時的真實時間是 $9.00 - d$，所以等到「上午 9 時」之後的概率變成

$$g(d) = \mathrm{P}(T_8 < -5 - d)\,\mathrm{P}(T_9 > -d).$$

時鐘快了（$d > 0$）會令我更早到站，所以我很少會錯過 8 時的巴士，於是 $g(d) < \frac{1}{16}$；時鐘稍為慢了則會產生相反的效果（例如 $g(-2.5) \approx 0.079$）。所以，合理而不偏頗的做法是採用雙尾檢定（two-sided test）。

設 X 為這 100 個早上之中，按新時鐘計算我等到上午 9 時之後的早上數目。不同巴士的到站時間互相獨立，而我每天都在同一時間到站，所以各個早上是獨立試驗，並且

$$H_0：\text{時鐘準確，}\quad X \sim \mathrm{B}\big(100, \tfrac{1}{16}\big);\qquad H_1：\text{時鐘調校錯誤，因此} p \neq \tfrac{1}{16}.$$

在 H_0 之下，我們預期有 $\mathrm{E}(X) = 6.25$ 次，而 $\mathrm{Var}(X) = 100 \cdot \frac{1}{16} \cdot \frac{15}{16} \approx 5.86$，所以標準差（standard deviation）約為 2.42。觀察所得的 5 只是略低於平均值。準確地計算，

$$\mathrm{P}(X \le 5) = \sum_{k=0}^{5} \binom{100}{k} p^k (1-p)^{100-k} \approx 0.40 \qquad \text{其中} p = \tfrac{1}{16},$$

而採用連續性校正（continuity correction）的正態近似（normal approximation）則給出 $\Phi\big((5.5 - 6.25)/2.42\big) = \Phi(-0.31) \approx 0.38$。這與 0.025（5% 雙尾檢定的下尾）相去甚遠，事實上與任何常用的顯著性水平都相去甚遠。所以我們不拒絕 H_0：**沒有證據顯示新時鐘調校錯誤**。

這並不等於有證據顯示時鐘是準確的。事實上 $g(d) = g(-5 - d)$，因為密度是對稱的（把兩班巴士的角色對調即可），所以一個慢了 5 分鐘的時鐘所給出的 $p = \frac{1}{16}$，與準確的時鐘完全相同。這類數據無論收集多少個早上，都無法把這兩種情況區分開來。

驗算

- 密度的面積為 1，而三角形公式與對 f 作數值積分的結果吻合，例如 $\mathrm{P}(T < -2) = 0.32$。

- 答案 0.7415 介乎最好的情況 $\frac{7}{8}$（每次都在 7.55 到達）與最壞的情況 0.32（每次都在 8.02 到達）之間，這是理所當然的。模擬 10^6 個早上，所得結果為 0.741。
- 模擬 10^6 對 (T_8, T_9)，所得結果為 0.0622，對應 $\frac{1}{16} = 0.0625$；而準確的二項尾部概率 $\mathrm{P}(X \leq 5)$ 的數值為 0.4002。

常見錯誤

- 忘記了要等到 9 時之後，必須先錯過 8 時的巴士：答案 $p = \frac{1}{2}$（只考慮「9 時的巴士遲到」）忽略了第一個因子。
- 以為「在 8.10 或之前趕上巴士」需要另外檢查 $T \leq 10$：8 時的巴士必定在 8.10 或之前到達，所以只有 $T > a$ 才是關鍵。
- 數據顯示次數略為偏少，卻針對「太多」進行單尾檢定（one-tailed test）；時鐘不準可以令 p 向任何一個方向移動。

STEP 之外

三角形密度正是兩個獨立均勻隨機變量（uniform random variables）之和的分佈，這是卷積（convolution）的第一個例子，而大學一年級的概率課程正是用卷積來求和的密度。最後一部分讓我們淺嘗統計學上的可識別性（identifiability）：對稱性 $g(d) = g(-5 - d)$ 意味着兩個不同的參數值會產生完全相同的數據，所以沒有任何檢定能夠把它們區分開來。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）對卷積和隨機變量之和有清晰的講解。

詳細解答

STEP II 1990，第 1 題

純數學 • 多項式不等式

> **STEP II 1990，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/90-s2-q1

關鍵思路

沒有實根（real roots）的四次多項式（quartic），其實就是一個恆為正的四次多項式，所以我們要設法把它拆成若干個我們*確知*為非負（non-negative）的部分：完全平方（perfect square），或者在某個區間（interval）上正負號一目了然的乘積。沒有哪一種拆法能處處適用，但不同的拆法在不同的區間上奏效，合起來便覆蓋整條實數線。第一部分把這些部件直接交給我們；到了最後一部分，就要自己動手去找。

解答

兩個非負的表達式

把四次式因式分解（factorise），並對二次式配方（completing the square），

$$x^4 - 2x^3 + x^2 = x^2\left(x^2 - 2x + 1\right) = x^2(x-1)^2,$$
$$x^2 - 8x + 17 = (x-4)^2 + 1.$$

第一個是兩個平方的乘積，所以對任何實數 x 都 ≥ 0。第二個是一個平方加 1，所以至少是 1，自然也是非負的。

第一個方程

設 $p(x) = x^4 - 2x^3 + x^2 - 8x + 17$。留意 p *並不是*上面兩個表達式之和（那個和含有 $2x^2$，而不是 x^2），所以不能簡單地把兩者相加。我們改為在每個區間上用不同的方式拆開 p，從而證明 $p(x) > 0$ 處處成立。

想一想 題目給了兩個非負的表達式，但它們加起來並不等於 $p(x)$。我們還可以怎樣利用它們？

兩者都可以在實數線的某一段上「借用」：在一個區間上，我們保留四次式 $x^2(x-1)^2$，然後直接處理餘下的 $17-8x$；在另一個區間上，我們保留二次式 $(x-4)^2+1$，然後直接處理餘下的 x^4-2x^3。

情況 $x \le 0$。 這時 $x^4 \ge 0$、$-2x^3 \ge 0$（因為 $x^3 \le 0$）、$x^2 \ge 0$ 以及 $-8x \ge 0$。所以 $p(x)$ 的每一項都是非負的，再加上常數項 17，總和至少是 17。因此 $p(x) > 0$。

情況 $0 < x \le 2$。 寫成

$$p(x) = x^2(x-1)^2 + (17-8x).$$

第一部分 ≥ 0，而由於 $x \le 2$，我們有 $17-8x \ge 17-16 = 1$。所以 $p(x) \ge 1 > 0$。

情況 $x > 2$。 寫成

$$p(x) = x^3(x-2) + \left(x^2 - 8x + 17\right) = x^3(x-2) + (x-4)^2 + 1.$$

由於 $x > 2$，x^3 和 $x-2$ 都是正數，所以 $x^3(x-2) > 0$；而 $(x-4)^2+1 \ge 1$。所以 $p(x) > 1 > 0$。

這三個區間覆蓋了所有實數，而在每個區間上都有 $p(x) > 0$，所以 $p(x) = 0$ 沒有實根。

第二個方程

設 $q(x) = x^4 - x^3 + x^2 - 4x + 4$。我們沿用同樣的方法，不過這次要自己選擇拆法。

想一想 應該在哪裏把實數線分段？又應該把哪些項配成一組？

我們希望每一組在各自的區間上都有明顯的正負號。$4-4x = 4(1-x)$ 這一組在 $x \le 1$ 時是非負的。當 $x \ge 1$ 時，$x^4 - x^3 = x^3(x-1)$ 這一組是非負的，而餘下的 $x^2 - 4x + 4 = (x-2)^2$ 是一個完全平方。所以最自然的分界點是 $x = 1$。

情況 $x \le 1$。 寫成

$$q(x) = \left(x^4 - x^3 + x^2\right) + 4(1-x) = x^2\left[\left(x - \tfrac{1}{2}\right)^2 + \tfrac{3}{4}\right] + 4(1-x).$$

第一項 ≥ 0，而且只在 $x = 0$ 時才等於零（方括號內的值至少是 $\frac{3}{4}$）。第二項因 $x \le 1$ 而 ≥ 0，而且只在 $x = 1$ 時才等於零。兩項不可能同時為零，所以 $q(x) > 0$。

情況 $x \geq 1$。 寫成

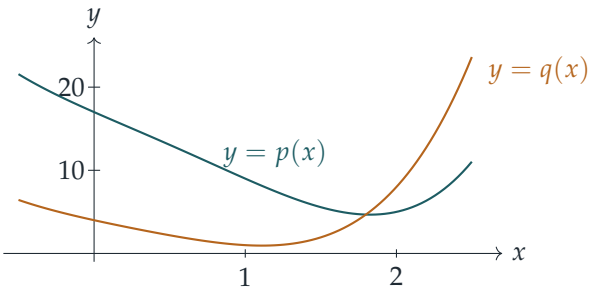
$$q(x) = x^3(x-1) + \left(x^2 - 4x + 4\right) = x^3(x-1) + (x-2)^2.$$

當 $x \geq 1$ 時，x^3 和 $x-1$ 都是非負的，所以第一項 ≥ 0，而且只在 $x = 1$ 時才等於零。第二項是一個平方，只在 $x = 2$ 時才等於零。兩者同樣不可能同時為零，所以 $q(x) > 0$。

所以對所有實數 x 都有 $q(x) > 0$，而 $q(x) = 0$ 沒有實根。

驗算

- 把每一種拆法展開，確保它確實還原成 p 或 q。例如 $x^3(x-1) + (x-2)^2 = x^4 - x^3 + x^2 - 4x + 4$。
- 在 $-5 \leq x \leq 8$ 的範圍內以 10^{-4} 為步長作數值掃描，得出 p 的最小值約為 4.67（在 $x \approx 1.83$ 附近），而 q 的最小值約為 0.94（在 $x \approx 1.11$ 附近）。兩者都明顯是正數，正如下圖所示。



常見錯誤

- 把題目給出的兩個表達式相加，然後聲稱其和就是 $p(x)$：其實並不是，因為 x^2 項對不上。
- 由「非負項加非負項」就斷定 $q(x) > 0$。這樣只能得出 $q(x) \geq 0$；要排除有根的可能，必須證明兩項不可能同時為零。
- 以為 $-2x^3$ 或 $-8x$ 永遠是負數。當 $x \leq 0$ 時，它們非但無害，反而有幫助。

STEP 之外

把多項式（polynomial）寫成平方和（sum of squares）來證明它是正的，是一個很有份量的課題。每一個非負的單變量多項式，都是兩個實係數多項式的平方之和，但希爾伯

特（Hilbert）證明了這個結論在多變量的情況下並不成立，由此引出了他的第十七個問題。時至今日，人們在最優化（optimisation）問題中會利用電腦去搜尋這類分解。如果想學習更多這一類不等式技巧，Siklos 的 *Advanced Problems in Mathematics* 是下一步的好選擇。

詳細解答

STEP II 1990，第 2 題

純數學 • 托勒密定理

> **STEP II 1990，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/90-s2-q2

關鍵思路

第一部分是一條三角恆等式（trigonometric identity），只要把每個正弦之積都化成餘弦之差，它便迎刃而解；然後利用條件 $A+B+C+D=\pi$ 改寫 $\cos(A+C)$ 和 $\cos(A-C)$。第二部分是托勒密定理（Ptolemy's theorem）。圓心 O 就是提示：半徑為 r 的圓上每條弦（chord）的長度都是 $2r\sin$(所對角的一半，而該角的頂點為O)，而圍繞 O 的四個半角之和為 π，這正是第一部分的條件。

解答

恆等式

我們三次使用積化和差公式（product-to-sum formula）

$$\sin X\sin Y=\tfrac{1}{2}\big[\cos(X-Y)-\cos(X+Y)\big]$$

。在左方，

$$\sin(A+B)\sin(A+D)=\tfrac{1}{2}\big[\cos(B-D)-\cos(2A+B+D)\big],$$
$$\sin B\sin D=\tfrac{1}{2}\big[\cos(B-D)-\cos(B+D)\big],$$

所以 $\cos(B-D)$ 兩項互相抵消，得

$$\sin(A+B)\sin(A+D)-\sin B\sin D=\tfrac{1}{2}\big[\cos(B+D)-\cos(2A+B+D)\big].$$

在右方，

$$\sin A\sin C=\tfrac{1}{2}\big[\cos(A-C)-\cos(A+C)\big].$$

現在利用 $C = \pi - A - B - D$。由此可得

$$A + C = \pi - (B + D), \qquad \text{所以} \quad \cos(A + C) = -\cos(B + D),$$
$$A - C = (2A + B + D) - \pi, \qquad \text{所以} \quad \cos(A - C) = -\cos(2A + B + D),$$

當中用到 $\cos(\pi - \theta) = -\cos\theta$ 和 $\cos(\theta - \pi) = -\cos\theta$。因此

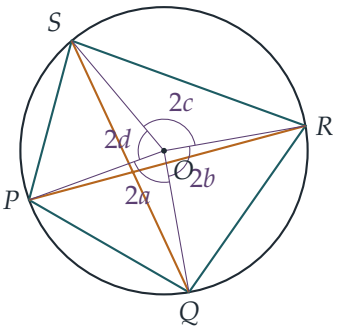
$$\sin A \sin C = \tfrac{1}{2}\big[-\cos(2A + B + D) + \cos(B + D)\big],$$

這與左方相同。恆等式得證。

圓

設圓的半徑為 r。沿圓周由 P 到 Q 到 R 到 S 再回到 P，設弧（arc）PQ、QR、RS、SP 在 O 所對的角分別為 $2a$、$2b$、$2c$、$2d$。由於各點按此次序排列，這四段弧互不重疊，並合成整個圓，所以

$$a, b, c, d > 0, \qquad 2a + 2b + 2c + 2d = 2\pi, \qquad \text{即} \quad a + b + c + d = \pi.$$



想一想 若一條弦所對的弧在圓心所對的角為 2θ，這條弦有多長？

若 $0 < 2\theta \le \pi$，由 O 和弦的兩個端點組成的三角形是等腰（isosceles）三角形，兩腰長 r，頂角為 2θ。由 O 所作的垂線同時平分弦和頂角，所以弦長為 $2r\sin\theta$。若 $\pi < 2\theta < 2\pi$，則三角形內位於 O 的角為 $2\pi - 2\theta$，弦長為 $2r\sin(\pi - \theta) = 2r\sin\theta$。所以在任何情況下

$$\text{弦長} = 2r\sin\theta.$$

把這個結果應用於四條邊，以及兩條對角線（diagonal）（弧 $P \to Q \to R$ 所對的角為 $2(a + b)$，弧 $Q \to R \to S$ 所對的角為 $2(b + c)$），便得

$$PQ = 2r\sin a, \quad QR = 2r\sin b, \quad RS = 2r\sin c, \quad PS = 2r\sin d,$$

$$PR = 2r\sin(a + b), \qquad QS = 2r\sin(b + c).$$

注意 $a+b$ 或 $b+c$ 可能大於 $\frac{\pi}{2}$；這就是我們需要為大於半圓的弧驗證該公式的原因。

除以 $4r^2$ 之後，要證明的結果變成

$$\sin a \sin c + \sin b \sin d = \sin(a+b)\sin(b+c),$$

或等價地，

$$\sin(b+a)\sin(b+c) - \sin a \sin c = \sin b \sin d.$$

這正是第一部分的恆等式，其中 $A=b$、$B=a$、$D=c$ 及 $C=d$；這樣代入是容許的，因為 $b+a+d+c=\pi$。再乘回 $4r^2$，得

$$PQ \times RS + QR \times PS = PR \times QS,$$

這就是所要證明的結果。

驗算

- 內接於半徑為 $\frac{1}{\sqrt{2}}$ 的圓的正方形，邊長為 1，對角線長 $\sqrt{2}$：$1\cdot 1 + 1\cdot 1 = 2 = \sqrt{2}\cdot\sqrt{2}$。
- 對任何邊長為 p 和 q 的長方形，托勒密定理變成 $p^2+q^2 =$ (對角線)^2：畢氏定理（Pythagoras's theorem）是它的特殊情況。對於 3×4 的長方形，$9+16=25=5\times 5$。
- 以 10^5 組隨機的 A, B, D 值（其中 $C=\pi-A-B-D$）對恆等式作數值檢驗，最大誤差約為 5×10^{-15}；另外 10^5 個隨機的圓內接四邊形（cyclic quadrilateral）均滿足最後的結果，相對誤差在 10^{-15} 以內。若各點的次序被打亂（例如沿圓周依次為 P, R, Q, S），等式便不成立，可見次序確實重要。

常見錯誤

- 使用以圓上某點為頂點的圓周角（inscribed angle），卻沒有說明它們對着哪一段弧。按題目的提示改用圓心角，便可避免這個問題。
- 忘記對角線所截的弧可以大於半圓；公式 $2r\sin\theta$ 依然成立，但需要上文那一行的理由來支持。
- 把錯誤的角配對到 A, B, C, D：A 和 C 必須是只在右方出現的兩個角，在這裏即 b 和 d。

STEP 之外

托勒密定理有一個影響深遠的推廣：對平面上任意四點，$PQ \cdot RS + QR \cdot PS \geq PR \cdot QS$，而等號成立當且僅當這些點依次位於同一個圓（或同一條直線）上。有一個巧妙的證明：把各點視為複數（complex numbers）p, q, r, s，留意代數恆等式 $(p-q)(r-s) + (q-r)(p-s) = (p-r)(q-s)$，再應用三角不等式（triangle inequality）；等號成立的情況，正是複平面的幾何派上用場之處。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）是欣賞這類論證的絕佳讀物，書中也介紹了反演（inversion），而反演同樣能簡潔地證明這個結果。

詳細解答

STEP II 1990，第 3 題

純數學 • 三次曲線的描繪

STEP II 1990，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/90-s2-q3

關鍵思路

把這個三次式（cubic）寫成 $f(x) = x(x^2 - 2bx + c^2)$。兩個判別式（discriminant）決定了一切：$f'(x)$ 的判別式告訴我們曲線有沒有轉向點（turning point），而二次因式的判別式則告訴我們曲線除了原點之外，還會不會在別處與 x 軸相交。題目的七種情況，不過是 b 相對於兩個臨界值 $\frac{\sqrt{3}}{2}c$ 和 c 的各種位置。三次曲線一旦畫好，$y^2 = f(x)$ 幾乎是順手可得：只保留 $f(x) \geq 0$ 的部分，取平方根，再對 x 軸作反射。

解答

各種情況都成立的事實

設 $f(x) = x^3 - 2bx^2 + c^2x = x(x^2 - 2bx + c^2)$。則

$$f'(x) = 3x^2 - 4bx + c^2, \qquad f''(x) = 6x - 4b.$$

- 轉向點。f' 的判別式是 $16b^2 - 12c^2 = 4(4b^2 - 3c^2)$。所以若 $2b > c\sqrt{3}$，f' 有兩個相異實零點；若 $2b = c\sqrt{3}$，有一個重零點；若 $2b < c\sqrt{3}$，則沒有零點。有零點時，它們的和是 $\frac{4}{3}b \geq 0$，積是 $\frac{1}{3}c^2 \geq 0$，所以兩個都不是負數：所有轉向點都位於 $x \geq 0$ 的範圍內。
- 根。$f(0) = 0$ 恆成立。二次因式是 $(x - b)^2 + c^2 - b^2$，所以若 $b < c$，它沒有實零點；若 $b = c$，有二重零點 $x = b$；若 $b > c$，則有兩個零點 $b \pm \sqrt{b^2 - c^2}$。
- 負的 x。當 $x < 0$ 時，x^3、$-2bx^2$ 和 c^2x 每一項都 ≤ 0，而且 $x^3 < 0$，所以 $f(x) < 0$。曲線在 y 軸左邊的部分全部位於 x 軸下方。
- 形狀。當 $x \to \pm\infty$ 時 $f(x) \to \pm\infty$；曲線在原點的斜率（gradient）是 $f'(0) = c^2$；唯

一的拐點（point of inflection）在 $x = \frac{2}{3}b$，f'' 在該處變號。

由於只要 $c > 0$ 就有 $c\sqrt{3} < 2c$，所以對固定的 c，各種情況是按 b 由小至大的次序排列的：先是沒有轉向點，然後是駐拐點，再然後是轉向點都在橫軸上方，接着是曲線在一個根處與橫軸相切，最後是三個根。

三次曲線

(i) $2b < c\sqrt{3}$。這時對所有 x 都有 $f'(x) > 0$（它的判別式是負數，首項係數是正數），所以 f 嚴格遞增。此外 $b < \frac{\sqrt{3}}{2}c < c$（注意這裏 $c > 0$），所以原點是唯一的根。曲線以斜率 c^2 上升穿過原點，並在 $x = \frac{2}{3}b$ 處有一個拐點。

(ii) $2b = c\sqrt{3} \neq 0$。這時 $f'(x) = 3(x - \frac{2}{3}b)^2 \geq 0$，所以 f 仍然遞增，但在 $x = \frac{2}{3}b = c/\sqrt{3}$ 處有一瞬間是水平的。這也正是 $f'' = 0$ 的地方，所以它是一個駐拐點（stationary point of inflection）。利用 $c^2 = \frac{4}{3}b^2$，

$$f\left(\tfrac{2}{3}b\right) = \tfrac{2}{3}b\left(\tfrac{4}{9}b^2 - \tfrac{4}{3}b^2 + \tfrac{4}{3}b^2\right) = \tfrac{8}{27}b^3 = \tfrac{\sqrt{3}}{9}c^3 > 0.$$

同樣有 $b < c$，所以原點是唯一的根。

(iii) $c\sqrt{3} < 2b < 2c$。這時 f' 有兩個零點 $x_1 < x_2$，兩個都是正數，於是在 $x_1 = \frac{1}{3}\left(2b - \sqrt{4b^2 - 3c^2}\right)$ 處有極大點（maximum），在 $x_2 = \frac{1}{3}\left(2b + \sqrt{4b^2 - 3c^2}\right)$ 處有極小點（minimum）。

? **想一想** 極小點會不會跌到 x 軸下方？

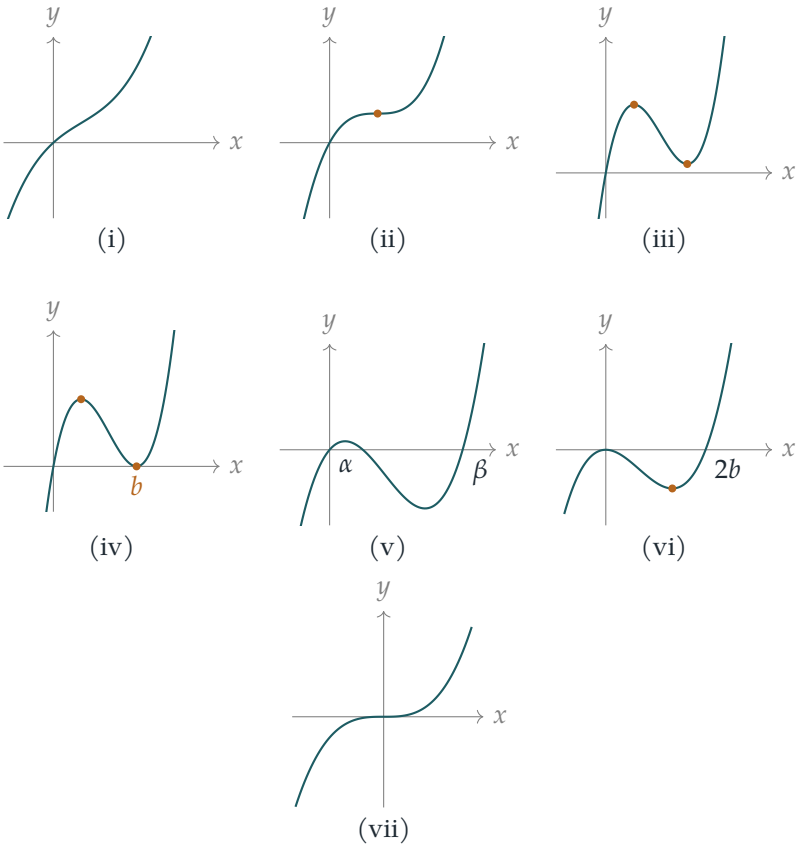
不會。由於 $b < c$，因式 $(x - b)^2 + c^2 - b^2$ 對每一個 x 都是正數，所以對所有 $x > 0$ 都有 $f(x) > 0$。因此兩個轉向點都在橫軸上方，曲線仍然只在原點穿過橫軸。它先上升，稍微回落，然後一直上升。

(iv) $b = c \neq 0$。這時 $f(x) = x(x - b)^2$。$x = b$ 是二重根，表示曲線在該處與橫軸相切；這就是極小點 $(b, 0)$。$f'(x) = (x - b)(3x - b)$ 的另一個零點給出極大點 $\left(\frac{1}{3}b, \frac{4}{27}b^3\right)$。

(v) $b > c > 0$。這時有三個根 $0 < \alpha < \beta$，其中 $\alpha = b - \sqrt{b^2 - c^2}$，$\beta = b + \sqrt{b^2 - c^2}$（因為 $\sqrt{b^2 - c^2} < b$，所以兩個都是正數）。在相鄰兩個根之間曲線必定要轉向，所以極大點在 $(0, \alpha)$ 內，該區間上 $f > 0$；極小點在 (α, β) 內，該區間上 $f < 0$。所以 $f \geq 0$ 恰好在 $[0, \alpha]$ 和 $[\beta, \infty)$ 上成立。

(vi) $c = 0, b \neq 0$。這時 $f(x) = x^2(x - 2b)$，$f'(x) = x(3x - 4b)$。原點是二重根，所以它是極大點，曲線在該處與橫軸相切；另外在 $\left(\frac{4}{3}b, -\frac{32}{27}b^3\right)$ 處有極小點。曲線在 $x = 2b$ 處穿過橫軸。

(vii) $b = c = 0$。這就是 $y = x^3$，在原點有一個駐拐點。



這些草圖取 $c = 1$，在 (i)–(iv) 中 $b = \frac{1}{2}, \frac{\sqrt{3}}{2}, 0.99, 1$，在 (v) 中 $(b, c) = (1, 0.8)$，在 (vi) 中 $b = 0.6$。各小圖的縱向比例並不相同，好讓 (iii) 和 (iv) 中細小的起伏能夠看得見。

曲線 $y^2 = f(x)$

以下三條法則可把每條三次曲線變成曲線 $y^2 = f(x)$：

- 在 $f(x) < 0$ 的地方沒有點，尤其是 $x < 0$ 的地方沒有點；
- 在 $f(x) \geq 0$ 的地方，$y = \pm\sqrt{f(x)}$，所以曲線對 x 軸對稱，並且在 f 最大的地方最高；
- 求導可得 $2y\dfrac{\mathrm{d}y}{\mathrm{d}x} = f'(x)$。

? **想一想** 曲線與 x 軸相交的地方是甚麼樣子的？

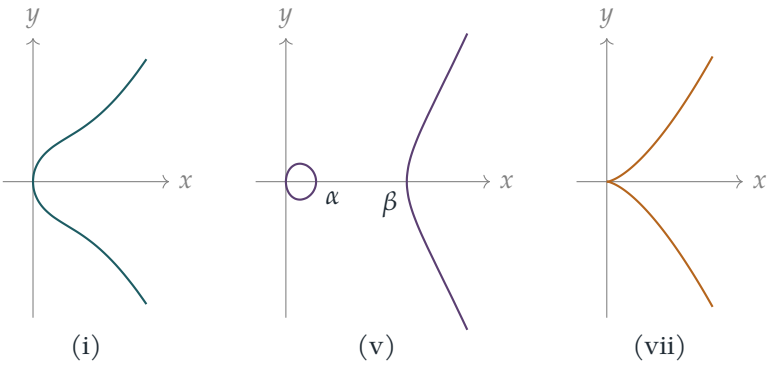
在 f 的單根（simple root）x_0 處，我們有 $f'(x_0) \neq 0$，而當 $y \to 0$ 時 $\dfrac{\mathrm{d}x}{\mathrm{d}y} = \dfrac{2y}{f'(x)} \to 0$。所以該處的切線（tangent）是垂直的，上下兩半平滑地接合。在這樣的根附近 $y^2 \approx$

$f'(x_0)(x - x_0)$，那是一條橫放的拋物線。重根（repeated root）的表現則有所不同，情況 (vii) 就是例子。

情況 (i)。$f(x) \geq 0$ 恰好當 $x \geq 0$ 時成立，而 0 是單根（$f'(0) = c^2 > 0$）。所以曲線是單獨一條無界的分支。它以垂直切線離開原點，在該處看來像 $y^2 \approx c^2 x$；由於 f 遞增，$|y| = \sqrt{f(x)}$ 穩步增加，最終增長得像 $x^{3/2}$。上半部分起初是凹（concave）的，像 $\sqrt{x}$，最後變成凸（convex）的，像 $x^{3/2}$，所以它有一個拐點（由對稱性，下半部分也有）。

情況 (v)。$f(x) \geq 0$ 在 $[0, \alpha]$ 和 $[\beta, \infty)$ 上成立，而三個根都是單根。所以曲線分成互不相連的兩部分：一個封閉的卵形線，經過 $(0,0)$ 和 $(\alpha, 0)$，在這兩點的切線都是垂直的，並且在三次曲線極大點的 x 坐標處最寬（就 y 方向而言）；以及一條無界的分支，從 $(\beta, 0)$ 以垂直切線出發，像情況 (i) 的曲線那樣向右張開。在三次曲線取負值的空隙 $\alpha < x < \beta$ 中，甚麼都沒有。

情況 (vii)。$y^2 = x^3$，所以 $x \geq 0$ 且 $y = \pm x^{3/2}$。這裏當 $x \to 0$ 時 $\mathrm{d}y/\mathrm{d}x = \pm\frac{3}{2}x^{1/2} \to 0$：兩半分別從上方和下方到達原點，並且都與 x 軸相切，形成一個尖角，稱為尖點（cusp）。



(i)　(v)　(vii)

這些圖所用的 b 和 c 的值與相應的三次曲線相同，兩軸的比例相等。

驗算

- 把 $b = c = 0$ 代入任何一個一般性的結論，結果都應該與 (vii) 吻合；例如 (ii) 中的駐點值 $\frac{8}{27}b^3$ 和 (vi) 中的極小值 $-\frac{32}{27}b^3$ 都收縮到原點。
- 取 $c = 1$ 和 $b = 0.99$（情況 (iii)），轉向點大約在 $(0.340, 0.150)$ 和 $(0.980, 0.020)$，正如所說，兩個都在橫軸上方。取 $(b, c) = (1, 0.8)$，根是 0、0.4、1.6，極小值約為 -0.388，在 $x \approx 1.147$ 處取得，位於 $(0.4, 1.6)$ 之內。
- 就情況 (i)、(iii) 和 (v) 隨機選取數千組 (b, c) 作數值測試，結果確認了轉向點和根的數目與位置，以及 $f \geq 0$ 成立的集合。

常見錯誤

- 把情況 (iii) 的極小點畫在橫軸下方：當 $b < c$ 時，三次曲線對所有 $x > 0$ 都取正值。
- 畫 $y^2 = f(x)$ 時，在它於單根處穿過 x 軸的地方畫成一個角，或者畫成水平切線。該處的切線是垂直的。
- 忘記了 $y^2 = f(x)$ 在三次曲線取負值的地方沒有點，所以在情況 (v) 中，卵形線與分支是不相連的。

STEP 之外

若曲線 $y^2 = x^3 + Ax^2 + Bx + C$ 中的三次式有相異的根，這條曲線就是橢圓曲線（elliptic curve），是現代數論的核心對象之一。情況 (v) 有三個實根，給出一個卵形線加一條分支的兩部分圖形；情況 (i) 只有一個實根，給出單獨一部分。當有根重合時，曲線就變成奇異（singular）的：情況 (vii) 有一個尖點，而情況 (iv) 則會給出一條在 $(b, 0)$ 處自相交的曲線。這類曲線上的點竟然可以用一種奇妙的方式「相加」，想初步了解，可參閱 Silverman 與 Tate 的 *Rational Points on Elliptic Curves*（橢圓曲線上的有理點）。

詳細解答

STEP II 1990，第 4 題

純數學 • 直線排佈與數學歸納法

STEP II 1990，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/90-s2-q4

本題在 *Beyond STEP* 第 9.1 節 直線分割平面中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/lazy-caterer.html

關鍵思路

逐條直線地看這幅圖。每加入一條新直線，它與每條舊直線都恰好相交一次，而且交點各不相同，所以它被切成的段數是可以預知的。每一段都有兩個作用：它本身是一條新的線段，同時又把一個原有的區域一分為二。數清楚加入一條新直線所帶來的變化，便得到遞推關係（recurrence），而這些遞推關係很容易求和。

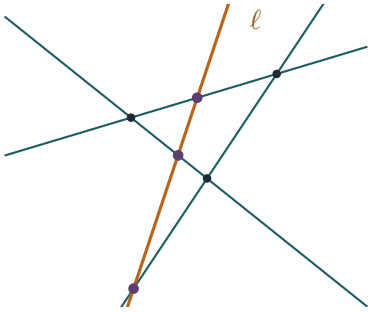
解答

小的情況

在下文中，「線段」（segment）是指某條直線上相鄰兩個交點之間的一段，包括每條直線兩端的兩段無限長部分（射線，ray）；只有一條直線時，整條直線算作一條線段。畫出最初幾種情況，可得：

n	1	2	3	4
線段	1	4	9	16
區域	2	4	7	11

例如，三條直線圍成一個三角形：每條直線被切成一條有限的邊和兩條射線，共得 9 條線段；而區域則有三角形本身、三個與它的邊相接的區域，以及三個與它的頂點相接的區域，合共 7 個。線段的數目令人猜想答案是 n^2。



圖中有三條直線（藍綠色）和第四條直線 ℓ（琥珀色）。ℓ 與舊直線相交於三個紫紅色的點，這三點把它切成四段，每一段各把七個舊區域的其中一個分成兩個，於是得到 11 個區域。

線段的數目

固定 n 條直線中的一條。由於沒有兩條直線平行（parallel），它與其餘 $n-1$ 條直線各相交一次；又由於沒有三條直線共點（concurrent），這 $n-1$ 個交點各不相同（假如有兩個交點重合，便會有三條直線在該處相交）。所以這條直線上有 $n-1$ 個不同的點，把它切成 n 段：$n-2$ 條有限線段和 2 條射線。對 n 條直線逐一這樣做，可得

$$\text{線段的數目} = n \cdot n = n^2.$$

每條線段只被數了一次，因為每條線段都恰好位於一條直線上。

如果你喜歡用歸納法（induction），可設 $S(n)$ 為線段的數目。在 $n-1$ 條處於一般位置（general position）的直線上加入第 n 條直線，新直線被切成 n 段，而它所穿過的 $n-1$ 條舊直線，每條都有一條線段被分成兩條，即每條各多出一條線段。所以

$$S(n) = S(n-1) + n + (n-1) = S(n-1) + 2n - 1,$$

而 $S(1) = 1$，若 $S(n-1) = (n-1)^2$，則 $S(n) = (n-1)^2 + 2n - 1 = n^2$。

區域的數目

設 $R(n)$ 為 n 條處於一般位置的直線所分成的區域數目，所以 $R(1) = 2$。從 $n-1$ 條這樣的直線出發，加入第 n 條直線 ℓ。如上所述，ℓ 與舊直線相交於 $n-1$ 個不同的點，並被切成 n 段。

取其中一段。它的內部不與任何舊直線相交，所以它位於單一個舊區域之內，並且由邊界到邊界（或一直伸延至無窮遠）橫貫該區域。因此它把該區域分成兩個。

想一想 這 n 段之中，會否有兩段位於同一個舊區域內，以致新增的區域少於 n 個？

不會。每個區域都是若干個半平面（half-plane）的交集（即位於每條舊直線的某一固定側），所以它是凸（convex）的。一條直線與一個凸集的交集是單一個區間，所以 ℓ 最多只穿過每個舊區域一次，不同的段位於不同的區域。此外，ℓ 沒有經過的區域則完全不受影響。

所以加入 ℓ 恰好把 n 個舊區域一分為二，其餘的保持不變：

$$R(n) = R(n-1) + n.$$

現在使用歸納法。當 $n = 1$ 時，公式 $\frac{1}{2}(n^2+n+2)$ 給出 2，這是正確的。若 $R(n-1) = \frac{1}{2}\big((n-1)^2+(n-1)+2\big) = \frac{1}{2}(n^2-n+2)$，則

$$R(n) = \tfrac{1}{2}(n^2-n+2) + n = \tfrac{1}{2}(n^2+n+2),$$

所以公式對所有 $n \geq 1$ 成立，證明完畢。（等價地，$R(n) = 1 + (1+2+\cdots+n) = 1 + \frac{1}{2}n(n+1)$，其中開首的 1 是沒有任何直線時的整個平面。）

驗算

- 當 $n = 1, 2, 3, 4$ 時，公式給出 $2, 4, 7, 11$，與所畫的圖吻合。
- 歐拉公式（Euler's formula）提供了一個獨立的驗算方法。把所有無限長的射線在單一個「無窮遠點」（point at infinity）處連接起來。這樣便有 $V = \binom{n}{2} + 1$ 個頂點、$E = n^2$ 條邊和 F 個面，且 $V - E + F = 2$，所以 $F = n^2 - \binom{n}{2} + 1 = \frac{1}{2}(n^2+n+2)$。
- 我們用電腦隨機畫出最多 8 條係數為整數的直線，透過找出交點的位置來數線段，並在每個交點附近及遠處選取樣本點，以這些樣本點「位於每條直線哪一側」的不同組合數目作為區域的數目。例如，8 條直線每次都得出 64 條線段和 37 個區域。

常見錯誤

- 忘記每條直線上的兩條無限長射線，結果得出 $n(n-2)$ 而不是 n^2。
- 說新直線增加 $n-1$ 個區域（每個交點一個），而不是 n 個（每段一個）。
- 沒有用上題設條件：若沒有「無三線共點」這個條件，一條直線上的交點便不一定各不相同，數目亦會減少。

STEP 之外

這是超平面排列（hyperplane arrangement）最簡單的情況。把同樣的論證提高一個維度，可證明 n 個處於一般位置的平面把空間切成 $\frac{1}{6}(n^3+5n+6)$ 塊，因為每個新平面本身都被舊平面切成 $R(n-1)$ 個區域。用歐拉公式所作的驗算，讓我們初次窺見歐拉示性數（Euler characteristic），它是拓撲學（topology）的一個核心概念。若想讀一章講述歐拉公式及其應用的精彩文字，可參閱 Aigner 與 Ziegler 的 *Proofs from THE BOOK*（《數學天書中的證明》）。

詳細解答

STEP II 1990，第 5 題

純數學 • 複數與圓

> **STEP II 1990，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/90-s2-q5

關鍵思路

在圓 $|z| = r$ 上，取共軛（conjugation）就等於做除法：$\bar{z} = r^2/z$。這樣一來，「這個比是純虛數（purely imaginary）」（判斷兩個方向互相垂直的準則）就化成一個簡潔的多項式條件 $lm + pq = 0$。此後，題目餘下的部分只是一路追蹤下去：每條垂直的弦（chord）都以形如 $z \mapsto c/z$ 的公式給出新的點，你只須把這些映射複合（compose）起來，看看結果是甚麼。

解答

垂直的判別準則

設 S 的半徑為 r，則 $|l| = |m| = |p| = |q| = r$，因此

$$\bar{l} = \frac{r^2}{l}, \qquad \bar{m} = \frac{r^2}{m}, \qquad \bar{p} = \frac{r^2}{p}, \qquad \bar{q} = \frac{r^2}{q}.$$

弦 LM 和 PQ 的方向分別是 $l - m$ 和 $p - q$，由於各點互不相同，兩者皆不為零。兩個非零複數所代表的方向互相垂直，當且僅當它們的商是純虛數，所以設

$$w = \frac{l - m}{p - q}.$$

那麼 w 是純虛數當且僅當 $w + \bar{w} = 0$。現在

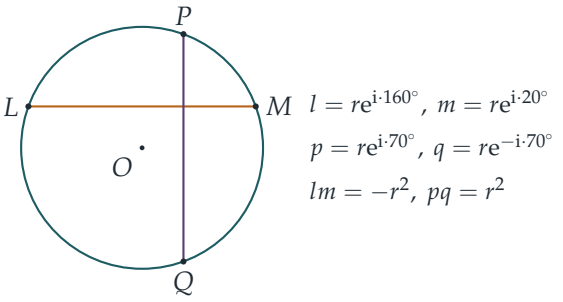
$$\bar{w} = \frac{r^2/l - r^2/m}{r^2/p - r^2/q} = \frac{(m - l)/(lm)}{(q - p)/(pq)} = \frac{l - m}{p - q} \cdot \frac{pq}{lm} = w\,\frac{pq}{lm},$$

所以

$$w + \bar{w} = w\left(1 + \frac{pq}{lm}\right) = \frac{w}{lm}\,(lm + pq).$$

由於 $w \neq 0$ 且 $lm \neq 0$，上式等於零當且僅當 $lm + pq = 0$。所以 $LM \perp PQ$ 當且僅當 $lm + pq = 0$，這正是所要證明的。

（用垂直平分線（perpendicular bisector）的做法，其實是同一個想法換上幾何的外衣：弦 LM 的垂直平分線通過 O，而弦的方向是該平分線方向的 i 倍。）



三個點

以 a_1, a_2, a_3 表示 A_1, A_2, A_3，並以 z 表示 A_1'。根據上述準則，「$A_1'A_2' \perp A_1A_2$」即是 $z\,a_2' = -a_1a_2$，所以

$$a_2' = -\frac{a_1a_2}{z}.$$

這個點確實在 S 上，因為 $|a_2'| = r \cdot r/r = r$；而且它是 S 上唯一符合要求的點。同理

$$a_3' = -\frac{a_2a_3}{a_2'} = \frac{a_3}{a_1}z,$$

$$a_1'' = -\frac{a_3a_1}{a_3'} = -\frac{a_1^2}{z}.$$

因此 $A_1'' = A_1'$ 當且僅當 $z = -a_1^2/z$，即

$$z^2 = -a_1^2 \iff z = \pm \mathrm{i}\, a_1.$$

這兩個值的模（modulus）都是 r，所以都在 S 上；又因為 $a_1 \neq 0$，它們並不相同。所以 A_1' 恰好有兩個位置：就是把 A_1 繞 O 旋轉 $\pm 90^\circ$ 所得的兩點。

想一想 映射 $z \mapsto -a_1^2/z$ 對這個圓做了甚麼？

在 S 上有 $1/z = \bar{z}/r^2$，所以 $-a_1^2/z = -(a_1^2/r^2)\,\bar{z}$，這是先取共軛、再作旋轉（rotation）：也就是把圓沿一條直徑作反射（reflection）。反射恰好只固定鏡線的兩個端點，這就是為何有兩個解。事實上，每一步 $z \mapsto -a_ka_{k+1}/z$ 都是這樣的反射，而這一點解釋了下文的一切。

四個點

繼續同樣的追蹤，只是最後一步改用 a_4：

$$a_2' = -\frac{a_1 a_2}{z}, \qquad a_3' = \frac{a_3}{a_1}\, z,$$
$$a_4' = -\frac{a_3 a_4}{a_3'} = -\frac{a_1 a_4}{z}, \qquad a_1'' = -\frac{a_4 a_1}{a_4'} = z.$$

所以無論 A_1' 取 S 上哪一點，都有 $A_1'' = A_1'$。

n 個點

對 k 用歸納法（induction），這樣追蹤下去可得

$$a_k' = \frac{a_k}{a_1}\, z \quad (k \text{ 為奇數}), \qquad a_k' = -\frac{a_1 a_k}{z} \quad (k \text{ 為偶數}),$$

因為若 a_k' 具有其中一種形式，則 $a_{k+1}' = -a_k a_{k+1}/a_k'$ 便具有另一種形式。最後一步是 $a_1'' = -a_n a_1 / a_n'$：

- 若 n 是偶數，$a_1'' = -a_n a_1 \cdot \dfrac{z}{-a_1 a_n} = z$，所以對每一個 A_1'，A_1'' 都與 A_1' 重合；
- 若 n 是奇數，$a_1'' = -a_n a_1 \cdot \dfrac{a_1}{a_n z} = -\dfrac{a_1^2}{z}$，所以與 $n = 3$ 的情況完全一樣，A_1'' 與 A_1' 重合的情況恰好只有 A_1' 的兩個位置，即 $z = \pm \mathrm{i} a_1$。

用文字來說：偶數次反射合成一個旋轉，而在這裏它原來就是恆等映射（identity）；奇數次反射則合成一個反射，它在圓上有兩個不動點（fixed point）。

驗算

- 圖中的例子：取 $l = r\mathrm{e}^{\mathrm{i}\cdot 160^\circ}$、$m = r\mathrm{e}^{\mathrm{i}\cdot 20^\circ}$ 時，弦 LM 是水平的；取 $p = r\mathrm{e}^{\pm \mathrm{i}\cdot 70^\circ}$ 時，弦 PQ 是垂直的；而確實有 $lm = r^2 \mathrm{e}^{\mathrm{i}\pi} = -r^2 = -pq$。
- 數值驗證：在半徑隨機的圓上隨機取點，並讓 n 由 3 取到 9，沿着垂直的弦逐步追蹤，每一步所得的點都落回 S 上；當 n 是偶數時，最後準確地回到起點；當 n 是奇數時，沿圓周掃描 200 000 個起始位置，恰好找到兩個不動點，位於 $\pm \mathrm{i} a_1$。

常見錯誤

- 以「w 是實數」來檢驗垂直。平行的弦給出實數比；垂直的弦給出的是純虛數比。
- 忘記 $\bar{z} = r^2/z$ 需要 $|z| = r$；正是因為四個點全都在以 O 為圓心的同一個圓上，這個準則才會如此簡潔。

► 在 n 個點的情況中，忘記答案取決於 n 的奇偶性（parity）。

STEP 之外

把映射 $z \mapsto c/z$ 限制在圓上，所得的就是反射，而反射的複合正是研究等距變換羣（isometry group）的起點：圓的對稱構成正交羣（orthogonal group）$O(2)$，當中兩個反射複合起來就是一個旋轉。像 $z \mapsto c/z$ 和 $z \mapsto (az+b)/(cz+d)$ 這樣的映射稱為莫比烏斯變換（Möbius transformations），是複分析（complex analysis）和幾何的核心。Needham 的 *Visual Complex Analysis* 把這兩種觀點都闡述得十分精彩。

詳細解答

STEP II 1990，第 6 題

純數學 • 整數不等式與分數

> **STEP II 1990，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/90-s2-q6

關鍵思路

一切都建基於整數（integer）的一個事實：整數之間的嚴格不等式（strict inequality）可以收緊 1。若 $a < b$，則 $b \geq a+1$；若 $d < c$，則 $d \leq c-1$。這樣一來，(ii) 就變成一行便可完成的估計。其後各部分都是把 (ii) 應用於巧妙選取的整數：先把分數不等式中的分母消去，再把所得的整數不等式代回 (ii)。

解答

以下所有字母均代表正整數（positive integer）。

第 (i) 部分

由於 $b > a \geq 1$ 而 b 是整數，故 $b \geq 2$。兩邊乘以 $c > 0$ 得 $bc \geq 2c$。

由於 $c > 1$ 而 c 是整數，故 $c \geq 2$，所以 $2c = c + c \geq c + 2$。合起來便得 $bc \geq 2c \geq 2 + c$，正是所求。

第 (ii) 部分

因為兩個不等式都是嚴格的，而各數均為整數，所以 $b \geq a+1$ 及 $d \leq c-1$。在前者用上 $c > 0$，在後者用上 $a > 0$，

$$\begin{aligned} bc - ad &\geq (a+1)c - a(c-1) \\ &= ac + c - ac + a \\ &= a + c. \end{aligned}$$

所以 $bc-ad \geq a+c$，而等號成立當且僅當 $b=a+1$ 及 $d=c-1$。

（這與 (i) 所用的「收緊 1」技巧相同；(i) 只是熱身，之後不會再用到。）

第 (iii) 部分

利用 $b,d>0$，把兩個不等式乘開：

$$\frac{a}{b}<p \iff a<bp, \qquad p<\frac{c}{d} \iff dp<c.$$

現在應用 (ii)，把 b 換成 bp，把 d 換成 dp（兩者都是正整數）。所需的假設 $a<bp$ 和 $dp<c$ 正是我們已有的，所以

$$(bp)c-a(dp)\geq a+c, \qquad \text{即} \qquad (bc-ad)p\geq a+c.$$

第 (iv) 部分

消去分母（全部為正數），得到兩個整數不等式：

$$aq<bp \qquad \text{及} \qquad dp<cq.$$

此外 $\frac{a}{b}<\frac{c}{d}$，所以 $ad<bc$，而 $bc-ad$ 是一個正整數；因此我們可以除以它而無須把任何不等號反轉。

> **想一想** 應該由哪四個數來擔當 (ii) 中 a,b,c,d 的角色？

我們有兩個嚴格不等式，而 (ii) 正好需要兩個：「第一個 < 第二個」和「第四個 < 第三個」。配對的自然方法有兩種，每種各給出其中一個下界。

p 的下界。取 $aq<bp$ 作為「$a<b$」那個不等式，取 $dp<cq$ 作為「$d<c$」那個。把第 (ii) 部分應用於 $aq,\ bp,\ cq,\ dp$，得

$$(bp)(cq)-(aq)(dp)\geq aq+cq, \qquad \text{所以} \qquad pq(bc-ad)\geq q(a+c).$$

先除以 $q>0$，再除以 $bc-ad>0$，

$$p\geq \frac{a+c}{bc-ad}.$$

q 的下界。現在把角色對調：取 $dp<cq$ 作為「$a<b$」那個不等式，取 $aq<bp$ 作為「$d<c$」那個。把第 (ii) 部分應用於 $dp,\ cq,\ bp,\ aq$，得

$$(cq)(bp)-(dp)(aq)\geq dp+bp, \qquad \text{所以} \qquad pq(bc-ad)\geq p(b+d).$$

除以 $p>0$，再除以 $bc-ad>0$，

$$q\geq \frac{b+d}{bc-ad}.$$

兩個下界都已證明。（第 (iii) 部分是第一個下界在 $q=1$ 時的特殊情況。）

介乎 8/9 與 9/10 之間的分數

取 $a = 8$、$b = 9$、$c = 9$、$d = 10$。那麼

$$bc - ad = 81 - 80 = 1,$$

所以任何嚴格介乎 8/9 與 9/10 之間的分數 p/q（其中 p, q 為正整數，不論是否最簡分數（lowest terms））都滿足

$$p \geq \frac{8+9}{1} = 17, \qquad q \geq \frac{9+10}{1} = 19.$$

因此，小於 20 的分母必定恰好是 $q = 19$。這時我們需要

$$\frac{8}{9} < \frac{p}{19} < \frac{9}{10} \iff \frac{152}{9} < p < \frac{171}{10} \iff 16.88\ldots < p < 17.1,$$

所以 $p = 17$。而 17/19 的確落在這個區間內，因為 $8 \cdot 19 = 152 < 153 = 9 \cdot 17$ 及 $17 \cdot 10 = 170 < 171 = 9 \cdot 19$。

因此，介乎 8/9 與 9/10 之間而分母小於 20 的分數只有

$$\frac{17}{19}.$$

由於下界 $q \geq 19$ 適用於分數的每一種寫法，而不限於最簡形式，所以像 34/38 這類未約簡的寫法，也不可能以更小的分母混進來。

驗算

- 17/19 是中間分數（mediant）$\frac{8+9}{9+10}$，而 (iv) 中兩個下界都由它達到，這是個好兆頭，說明這些下界是最佳的。
- 小數：$8/9 \approx 0.8889$，$17/19 \approx 0.8947$，$9/10 = 0.9$。
- 我們對不超過 29 的所有數值檢驗了 (i) 和 (ii)，以約 60,000 個隨機例子檢驗了 (iii) 和 (iv)，並用窮舉法（brute force）列出所有介乎 8/9 與 9/10 之間而 $q \leq 19$ 的 p/q：只有 17/19 出現。

常見錯誤

- 在 (iv) 中除以該數之前，忘記說明 $bc - ad > 0$。
- 在最後一部分只用 p 的下界。它容許 q 最大為 19，但單憑它並不能排除更小的分母；真正起作用的是 q 的下界。
- 在 (iv) 中只從 (ii) 得到一個下界：第二個下界需要把兩個不等式的角色對調。

STEP 之外

滿足 $bc - ad = 1$ 的兩個分數 $a/b < c/d$ 稱為法里相鄰分數（Farey neighbours），而本題證明了關於它們的關鍵事實：介乎兩者之間而分母最小的分數，就是它們的中間分數 $\frac{a+c}{b+d}$。不斷插入中間分數，便可建構出法里數列（Farey sequences）和 Stern–Brocot 樹（Stern–Brocot tree）；它們把每個正有理數（rational）恰好列出一次，並與連分數（continued fractions）以及實數的最佳有理逼近（best rational approximations）密切相關。Davenport 的 *The Higher Arithmetic* 是閱讀連分數與逼近理論的絕佳讀物。

詳細解答

STEP II 1990，第 7 題

純數學 • 微分方程

> **STEP II 1990，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/90-s2-q7

關鍵思路

以指數函數作試解（trial solution），可把時滯方程（delay equation）化為關於 p 的普通方程，因為時間平移只是把 e^{pt} 乘以一個常數。方程 $(*)$ 同時含有 p 和 e^{-p}，所以無法精確求解；但全式乘以 e^{p} 後，所有含 p 的部分都歸到同一邊：方程的根就是單一條曲線 $y=(p+1)\mathrm{e}^{p}$ 與水平線 $y=k$ 的交點。於是，數算根的個數就變成描繪一條曲線的問題。

解答

關於 p 的方程

設 $\mathrm{f}(t)=A\mathrm{e}^{pt}$，其中 $A\neq 0$（我們要的是非零解）。則 $\mathrm{f}'(t)=pA\mathrm{e}^{pt}$，而 $\mathrm{f}(t-1)=A\mathrm{e}^{p(t-1)}=\mathrm{e}^{-p}A\mathrm{e}^{pt}$，所以 $(\dagger)$ 變成

$$A\mathrm{e}^{pt}\left(p+1-k\mathrm{e}^{-p}\right)=0 \quad \text{對所有} t.$$

由於 $A\neq 0$ 且 $\mathrm{e}^{pt}>0$，上式對所有 t 成立，當且僅當

$$p+1=k\mathrm{e}^{-p},$$

這正是 $(*)$。因此，這種形式的解存在，恰好是當 p 為 $(*)$ 的實根之時，而此時任何 A 都可以。

有多少個實根？

把 $(*)$ 乘以 $\mathrm{e}^{p}>0$，得到等價的方程

$$g(p)=k, \qquad \text{其中} g(p)=(p+1)\mathrm{e}^{p}.$$

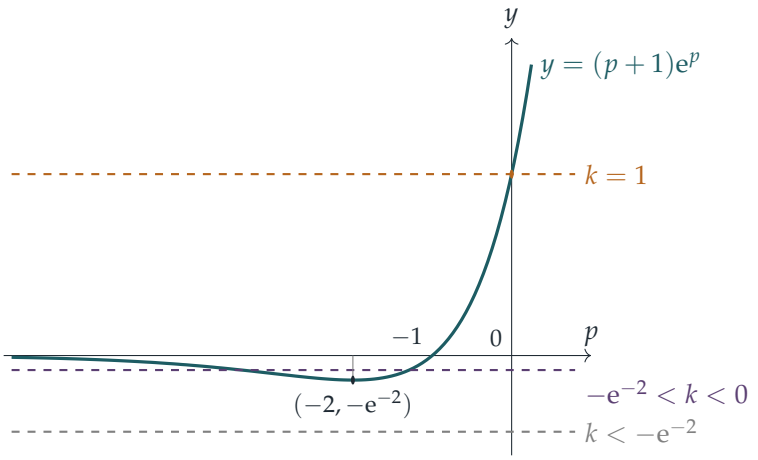
所以我們要數的是曲線 $y = g(p)$ 與水平線 $y = k$ 的交點數目。

> **想一想** $y = (p+1)\mathrm{e}^p$ 的圖像是甚麼模樣？它在哪裏取得最小值？兩端的走勢又如何？

求導得 $g'(p) = \mathrm{e}^p + (p+1)\mathrm{e}^p = (p+2)\mathrm{e}^p$，當 $p < -2$ 時為負，當 $p > -2$ 時為正。所以 g 在 $(-\infty, -2]$ 上遞減，在 $[-2, \infty)$ 上遞增，並有全域最小值（global minimum）

$$g(-2) = -\mathrm{e}^{-2}.$$

在兩端：當 $p \to \infty$ 時 $g(p) \to \infty$；當 $p \to -\infty$ 時 $g(p) \to 0$（指數衰減壓倒線性增長）；當 $p < -1$ 時因子 $p+1$ 為負，所以 g 從下方趨近 0。曲線只在 $p = -1$ 處穿過橫軸，且 $g(0) = 1$。



從圖中讀出結論（上面的單調性（monotonicity）令論證嚴格）：

- $k > 0$：在 $(-\infty, -1]$ 上有 $g \le 0 < k$，而在 $(-1, \infty)$ 上 g 由 0 嚴格遞增至 ∞，所以**恰有一個**根，且它滿足 $p > -1$。
- $-\mathrm{e}^{-2} < k < 0$：在 $(-\infty, -2]$ 上，g 由接近 0 的值嚴格遞減至 $-\mathrm{e}^{-2}$，所以它取值 k 一次；在 $[-2, -1]$ 上它由 $-\mathrm{e}^{-2}$ 嚴格遞增至 0，所以它再取值 k 一次；當 $p > -1$ 時，$g > 0 > k$。因此有**兩個**根，一個在 $(-\infty, -2)$ 內，一個在 $(-2, -1)$ 內。
- $k = -\mathrm{e}^{-2}$：直線與曲線在最低點相切，得出**唯**一的根 $p = -2$。
- $k < -\mathrm{e}^{-2}$：直線位於整條曲線之下，所以**沒有**根。

所以 $(*)$ 的實根數目視乎 k 而為 0、1 或 2，正如題目所求；而形如 $A\mathrm{e}^{pt}$ 的解存在，恰好是當

$$k \ge -\mathrm{e}^{-2}, \quad k \ne 0.$$

（題目已排除 $k = 0$ 這個值；它會給出單一的根 $p = -1$。）

解在甚麼時候衰減？

想一想 哪些 p 會令 $Ae^{pt} \to 0$？不解 $(*)$ 的話，怎樣判斷根相對於 0 的位置？

在 $A \neq 0$ 之下，$Ae^{pt} \to 0$（當 $t \to +\infty$）當且僅當 $p < 0$。我們從來不需要把根明確求出，只需知道它們相對於 0 的位置。

- 若 $-e^{-2} \leq k < 0$，每個根都在 $(-\infty, -1)$ 內（事實上當 $p \geq -1$ 時 $g \geq 0$，所以 $g(p) = k < 0$ 必然導致 $p < -1$）。所有這些解都會衰減。
- 若 $k > 0$，唯一的根位於 g 嚴格遞增的範圍內，所以 $p < 0 \iff g(p) < g(0) \iff k < 1$。當 $k = 1$ 時，根是 $p = 0$，給出常數解 $f = A$，它並不趨向零；當 $k > 1$ 時，根為正數，解會增長。

因此，當 $t \to +\infty$ 時解趨向零，恰好是當

$$-e^{-2} \leq k < 1, \quad k \neq 0.$$

驗算

- $k = 1$：$p = 0$ 滿足 $0 + 1 = 1 \cdot e^{0}$，而 $f = A$ 顯然是 $f' + f - f(t-1) = 0$ 的解。這正好是衰減與增長之間的分界，合乎預期。
- $k = -e^{-2} \approx -0.1353$：$p = -2$ 給出 $-1 = -e^{-2}e^{2}$。數值上，$k = -0.13$ 給出兩個根 $p \approx -2.311$ 和 $p \approx -1.742$，分別在 -2 的兩側，而 $k = -0.14$ 則沒有根。
- 物理意義：沒有反饋（feedback）時（$k \to 0$），系統是 $f' = -f$，衰減率為 $p = -1$，這正是 $|k|$ 很小時根所在的位置。強的正反饋（$k > 1$）會蓋過阻尼（damping），令解增長。

常見錯誤

- 把 $y = p + 1$ 與 $y = ke^{-p}$ 畫在一起比較，卻忽略了當 $k < 0$ 時直線可以與曲線相交兩次，也可以完全不相交；改用 $g(p) = (p+1)e^{p}$，相切（tangency）的條件會較容易找到。
- 遺漏端點 $k = -e^{-2}$；它確實給出一個（重）根 $p = -2$，因而有一個衰減的解。
- 把 $k = 1$ 納入衰減的範圍：$p = 0$ 給出的是常數解，而不是衰減的解。

STEP 之外

方程 (†) 是一個時滯微分方程（delay differential equation），而 $(*)$ 是它的特徵方程（characteristic equation）。與一般的線性常微分方程（ODE）不同，它有無窮多個複數根 p，而通解是由所有這些根構成的，所以要完全穩定，就需要每個複數根的實部都是負數；這是控制理論（control theory）研究的課題。$(p+1)\mathrm{e}^{p} = k$ 的實根是朗伯（Lambert）W 函數的值：$p+1 = W(k\mathrm{e})$，而 $-\mathrm{e}^{-1}$（這裏即 $k\mathrm{e} = -\mathrm{e}^{-1}$）正好是 W 的分支點（branch point）。關於線性常微分方程和特徵方程，可參看 Boyce 與 DiPrima 的 *Elementary Differential Equations*。

詳細解答

STEP II 1990，第 8 題

純數學 • 分部積分法

> **STEP II 1990，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/90-s2-q8

關鍵思路

微分方程（differential equation）的每一項都是，或者幾乎是，一個全導數（exact derivative）。由於 y 和 y' 在 $t = 0$ 與 $t = 1$ 取相同的值，這些全導數的積分都等於零。每一條關係式的做法，都是把方程乘以一個合適的函數，然後反覆使用分部積分法（integration by parts），直至所有邊界項（boundary term）都消失。不妨先列出能令邊界項消失的事實：$x(0) = 0$、$y(0) = y(1)$、$y'(0) = y'(1)$，以及（待我們證明之後）$x(1) = 0$。

解答

第一條恆等式

由於 $x' = y$，被積函數（integrand）$xy = xx'$ 正是 $\frac{1}{2}x^2$ 的導數。所以

$$\int_0^1 x(t)\,y(t)\,\mathrm{d}t = \Big[\tfrac{1}{2}x(t)^2\Big]_0^1 = \tfrac{1}{2}[x(1)]^2,$$

因為 $x(0) = \int_0^0 y(u)\,\mathrm{d}u = 0$。這就是所要求的結果。

證明 $x(1) = 0$

把 $(*)$ 的中間項寫成全導數：

$$(y^2 - 1)\,y' = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\tfrac{1}{3}y^3 - y\Big).$$

現在把 $(*)$ 由 0 到 1 逐項積分：

$$\Big[y'\Big]_0^1 + \Big[\tfrac{1}{3}y^3 - y\Big]_0^1 + \int_0^1 y\,\mathrm{d}t = 0.$$

第一個括號等於 $y'(1)-y'(0)=0$，第二個括號等於零則是因為 $y(1)=y(0)$。根據 x 的定義，餘下的積分就是 $x(1)$。因此 $x(1)=0$。

關係式 $\int y^2=\frac{1}{3}\int y^4$

把 $(*)$ 乘以 x，然後在 $[0,1]$ 上積分。這樣會得到三個積分，我們逐一處理。

最後一項。由第一條恆等式和 $x(1)=0$，可得

$$\int_0^1 x\,y\,\mathrm{d}t=\tfrac{1}{2}[x(1)]^2=0.$$

y'' 項。我們使用分部積分法，對 x 求導（於是 $x'=y$）：

$$\begin{aligned}\int_0^1 x\,y''\,\mathrm{d}t&=\Big[x\,y'\Big]_0^1-\int_0^1 y\,y'\,\mathrm{d}t\\&=x(1)y'(1)-x(0)y'(0)-\Big[\tfrac{1}{2}y^2\Big]_0^1=0,\end{aligned}$$

因為 $x(1)=x(0)=0$ 而且 $y(1)=y(0)$。

> **想一想** 我們應該怎樣對 $x\,(y^2-1)\,y'$ 進行分部積分？應該對哪一個因子積分？

中間項。對因子 $(y^2-1)y'$ 積分（我們已經知道它是 $\frac{1}{3}y^3-y$ 的導數），並對 x 求導：

$$\begin{aligned}\int_0^1 x\,(y^2-1)\,y'\,\mathrm{d}t&=\Big[x\big(\tfrac{1}{3}y^3-y\big)\Big]_0^1-\int_0^1 y\big(\tfrac{1}{3}y^3-y\big)\,\mathrm{d}t\\&=0-\tfrac{1}{3}\int_0^1 y^4\,\mathrm{d}t+\int_0^1 y^2\,\mathrm{d}t.\end{aligned}$$

邊界項消失，是因為 x 在兩個端點都等於零。

把三部分相加，結果為零：

$$\int_0^1 y^2\,\mathrm{d}t-\tfrac{1}{3}\int_0^1 y^4\,\mathrm{d}t=0,\qquad\text{即}\qquad\int_0^1[y(t)]^2\,\mathrm{d}t=\tfrac{1}{3}\int_0^1[y(t)]^4\,\mathrm{d}t,$$

這正是所要求的結果。

關係式 $\int y'^2=\int y^2$

> **想一想** 我們想得到 y'^2 的積分。乘以哪一個函數才會產生它？

對 $y\,y''$ 進行分部積分會產生 $-\int y'^2$，所以這次我們把 $(*)$ 乘以 y 再積分。逐項計算：

$$\int_0^1 y\,y''\,\mathrm{d}t=\Big[y\,y'\Big]_0^1-\int_0^1 y'^2\,\mathrm{d}t=-\int_0^1 y'^2\,\mathrm{d}t,$$

$$\int_0^1 (y^3-y)\,y'\,\mathrm{d}t=\Big[\tfrac{1}{4}y^4-\tfrac{1}{2}y^2\Big]_0^1=0.$$

邊界項 $y(1)y'(1) - y(0)y'(0)$ 等於零，因為 y 和 y' 的值都會重複。第二行等於零，是因為 $y(1) = y(0)$。餘下的就是

$$-\int_0^1 y'^2\,\mathrm{d}t + \int_0^1 y^2\,\mathrm{d}t = 0, \qquad \text{所以} \qquad \int_0^1 [y'(t)]^2\,\mathrm{d}t = \int_0^1 [y(t)]^2\,\mathrm{d}t.$$

驗算

第一條恆等式對任何 y 都成立。取 $y = \cos 3t + t^2 - 0.7e^t$，以數值積分（quadrature）計算可得 $\int_0^1 xy\,\mathrm{d}t \approx 0.338191$，與 $\frac{1}{2}x(1)^2$ 相等。

$(*)$ 並沒有在時間 1 之後重複的非平凡解（non-trivial solution）（見下文）。不過，對於任何在時間 T 之後重複的解，同樣的論證可以一字不改地套用到 $[0, T]$ 上。我們以數值方法對 $(*)$ 積分，直至解落在它的週期軌道（periodic orbit）上，其週期為 $T \approx 6.6633$。在一個週期內，我們求得 $\int y \approx 10^{-9}$、$\int y^2 \approx 13.72222$、$\frac{1}{3}\int y^4 \approx 13.72222$ 和 $\int y'^2 \approx 13.72228$，這些數值在步長誤差範圍內吻合。

常見錯誤

- 忘記第三部分需要用到 $x(1) = 0$。沒有這個條件，$\int xy$ 和各個邊界項都不會消失。
- 在中間項中對 x 積分，而不是對 $(y^2 - 1)y'$ 積分。正是利用了 $x' = y$，積分才能化簡成 y 的冪。
- 說邊界項等於零是「因為 $y(0) = 0$」。題目並沒有給出這個條件。邊界項消失，是因為在 0 和 1 處的值相等。

STEP 之外

方程 $(*)$ 是范德波爾振盪器（van der Pol oscillator），是自持振盪電路的經典模型。它只有一條週期軌道，而你所證明的關係式，就是沿一個循環的能量平衡。這些關係式還揭示了一個意想不到的事實。由於 $\int_0^1 y = 0$，由傅立葉級數（Fourier series）得出的維廷格（Wirtinger）不等式（又稱龐加萊（Poincaré）不等式）給出 $\int_0^1 y'^2 \geq 4\pi^2 \int_0^1 y^2$。結合最後一條關係式，這就迫使 $y \equiv 0$。所以在時間 1 之後重複的解只有平凡解，因為真正的週期約為 6.66。想進一步了解週期軌道和極限環（limit cycle），可參閱 Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）。

詳細解答

STEP II 1990，第 9 題

純數學 • 極坐標曲線與弧長

STEP II 1990，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/90-s2-q9

關鍵思路

兩條曲線都以 θ 表示 r，所以可以逐條射線比較：從原點出發的每條半直線上，拋物線（parabola）恰好有一點，而這一點不是在心臟線（cardioid）之內，就是在其外。這樣立刻就知道拋物線在哪裏進入和離開心臟線。至於長度，心臟線最好保留極坐標（polar）形式（其弧長（arc length）的被積函數（integrand）用半角恆等式（half-angle identity）便可大幅化簡），而拋物線則用直角坐標（Cartesian）形式容易得多。

解答

兩條曲線

對拋物線而言，$r + r\cos\theta = 1$ 即 $r = 1 - x$。兩邊平方，得 $x^2 + y^2 = (1-x)^2$，所以

$$y^2 = 1 - 2x.$$

這是一條焦點（focus）在原點、頂點（vertex）在 $(\frac{1}{2}, 0)$、開口向左的拋物線。原點位於它凹的一側，即 $y^2 < 1 - 2x$ 的一側；我們稱這一側為拋物線的內部。

心臟線 $r = 4(1+\cos\theta)$ 的尖點（cusp）在原點（當 $\theta = \pi$ 時），並在 $\theta = 0$ 時到達 $x = 8$。

兩曲線的交點

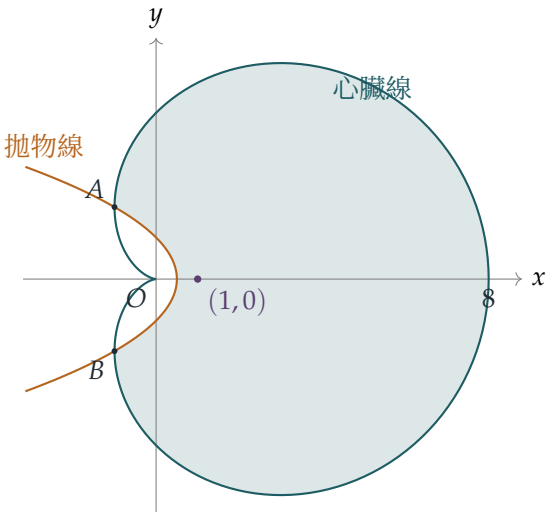
固定一個方向 θ，其中 $-\pi < \theta < \pi$，使得 $1 + \cos\theta > 0$。在這條射線上，拋物線與原點的距離為 $1/(1+\cos\theta)$，心臟線的邊界與原點的距離為 $4(1+\cos\theta)$，而心臟線的內部就是射線上位於邊界之前的那一段。所以拋物線上的點在心臟線之內，當且僅當

$$\frac{1}{1+\cos\theta} < 4(1+\cos\theta) \iff (1+\cos\theta)^2 > \tfrac{1}{4} \iff \cos\theta > -\tfrac{1}{2} \iff |\theta| < \tfrac{2\pi}{3},$$

其中我們乘以正數 $1+\cos\theta$，並取了正平方根。等號在 $\theta = \pm\frac{2\pi}{3}$ 時成立，此時 $r = 2$，得出兩個交點

$$A = (-1, \sqrt{3}), \qquad B = (-1, -\sqrt{3}).$$

所以拋物線在心臟線之內的部分是一段連續不斷的弧，由 B 經頂點 $(\frac{1}{2}, 0)$ 到 A，其餘部分全在心臟線之外。這段弧連接心臟線邊界上的兩點，因此把心臟線的內部分成兩部分：一部分在拋物線之內（包含原點，以心臟線在尖點附近的小環為邊界），另一部分在拋物線之外（以心臟線外側的大弧為邊界）。



（草圖中可見一個巧妙的細節：$x = 4(1+\cos\theta)\cos\theta = (2\cos\theta+1)^2 - 1 \geq -1$，所以 A 和 B 也是心臟線上最左邊的點。）

邊界的長度

$r = 1$、$\theta = 0$ 所表示的點是 $(1,0)$。它在心臟線之內（$1 < 8$），又在拋物線之外（$0^2 > 1 - 2$），所以位於陰影部分。這部分的邊界由心臟線的弧 $-\frac{2\pi}{3} \leq \theta \leq \frac{2\pi}{3}$ 和拋物線由 B 到 A 的弧組成。

心臟線的弧。在極坐標中，$\mathrm{d}s = \sqrt{r^2 + (\mathrm{d}r/\mathrm{d}\theta)^2}\,\mathrm{d}\theta$。這裏 $\mathrm{d}r/\mathrm{d}\theta = -4\sin\theta$，所以

$$r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 = 16\left(1 + 2\cos\theta + \cos^2\theta + \sin^2\theta\right) = 32(1+\cos\theta) = 64\cos^2\tfrac{\theta}{2}.$$

想一想 我們可以直接寫 $\sqrt{64\cos^2(\theta/2)} = 8\cos(\theta/2)$ 嗎？

可以，但全因為在整個範圍內 $\cos\frac{\theta}{2} \geq 0$：當 $|\theta| \leq \frac{2\pi}{3}$ 時，有 $|\frac{\theta}{2}| \leq \frac{\pi}{3}$。利用圖形對 x 軸的對稱性，

$$L_{\text{心臟線}} = 2\int_0^{2\pi/3} 8\cos\tfrac{\theta}{2}\,\mathrm{d}\theta = 32\left[\sin\tfrac{\theta}{2}\right]_0^{2\pi/3} = 32\cdot\tfrac{\sqrt{3}}{2} = 16\sqrt{3}.$$

拋物線的弧。以 y 為參數（parameter）：$x = \frac{1}{2}(1 - y^2)$，所以 $\mathrm{d}x/\mathrm{d}y = -y$，而 $\mathrm{d}s = \sqrt{1+y^2}\,\mathrm{d}y$，其中 y 由 $-\sqrt{3}$ 變到 $\sqrt{3}$。由對稱性，

$$L_{\text{拋物線}} = 2\int_0^{\sqrt{3}} \sqrt{1+y^2}\,\mathrm{d}y.$$

設 $y = \sinh u$，$\mathrm{d}y = \cosh u\,\mathrm{d}u$，被積函數變成 $\cosh^2 u = \frac{1}{2}(1+\cosh 2u)$，所以

$$\int \sqrt{1+y^2}\,\mathrm{d}y = \tfrac{1}{2}u + \tfrac{1}{4}\sinh 2u = \tfrac{1}{2}\left(\operatorname{arsinh} y + y\sqrt{1+y^2}\right) + C.$$

當 $y = \sqrt{3}$ 時，$\sqrt{1+y^2} = 2$，而 $\operatorname{arsinh}\sqrt{3} = \ln\left(\sqrt{3}+2\right)$，所以

$$L_{\text{拋物線}} = \left[\operatorname{arsinh} y + y\sqrt{1+y^2}\right]_0^{\sqrt{3}} = \ln(2+\sqrt{3}) + 2\sqrt{3}.$$

總長度。兩者相加，

$$L = 16\sqrt{3} + 2\sqrt{3} + \ln(2+\sqrt{3}) = 18\sqrt{3} + \ln(2+\sqrt{3}),$$

這正是所要求的結果。

驗算

- 整條心臟線的長度是 $\int_{-\pi}^{\pi} 8\cos\frac{\theta}{2}\,\mathrm{d}\theta = 32$，而我們所用的外側弧長 $16\sqrt{3} \approx 27.71$ 佔了其中大部分，與草圖吻合：尖點附近的小環長度為 $32 - 16\sqrt{3} \approx 4.29$。
- 拋物線的弧長約為 4.78，比由 B 到頂點 $(\frac{1}{2}, 0)$ 再到 A 的兩條直線段的總長度 $2\sqrt{1.5^2+3} \approx 4.58$ 稍長，這是理所當然的。
- 數值上，對兩條曲線的極坐標弧長被積函數用中點法則（midpoint rule）作數值積分（quadrature），所得總和為 32.49387，與 $18\sqrt{3} + \ln(2+\sqrt{3}) \approx 32.49387$ 一致。

常見錯誤

- 選錯了部分：$(1, 0)$ 在拋物線之外（在其頂點 $x = \frac{1}{2}$ 的右方），所以它所在部分的邊界用的是心臟線外側的大弧，而不是尖點附近的小環。
- 用 $8\cos\frac{\theta}{2}$ 一口氣把心臟線由 0 積分到 2π，但這個式子在 $\theta > \pi$ 時變成負數；被積函數應是 $8\left|\cos\frac{\theta}{2}\right|$。

STEP 之外

兩條曲線都是圓錐曲線（conic）或蚶線（limaçon），寫成以焦點或尖點為極點（pole）的形式；而 $r = \ell/(1 + e\cos\theta)$ 正是圓錐曲線在軌道力學（orbital mechanics）中出現的

形式，太陽位於焦點。弧長公式 $\int \sqrt{r^2 + r'^2}\,\mathrm{d}\theta$ 是參數曲線（parametrised curve）長度 $\int |\mathbf{r}'(t)|\,\mathrm{d}t$ 的特殊情況，後者在大學一年級的向量微積分（vector calculus）中會進一步引出曲率（curvature）和線積分（line integral）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）用圓錐曲線的極坐標形式推導出開普勒（Kepler）軌道。

詳細解答

STEP II 1990，第 10 題

純數學 • 矩陣多項式方程

> **STEP II 1990，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/90-s2-q10

關鍵思路

行列式（determinant）把矩陣（matrix）的乘積化為普通數字的乘積，而實數的乘積為零，必定是其中一個因子為零。至於那個三次式，$\mathbf{M}$ 與 $\mathbf{I}$ 可交換（commute），所以 $\mathbf{M}$ 的矩陣多項式的因式分解，與 x 的多項式完全一樣。上三角矩陣（upper triangular matrix）的表現則十分理想：乘積的對角線元素就是各對角線元素的乘積，所以對角線很快便可處理好，只有右上角才需要真正花工夫。

解答

滿足 $\mathbf{AB} = 0$ 的矩陣

由於乘積的行列式等於行列式的乘積，

$$\det \mathbf{A}\ \det \mathbf{B} = \det(\mathbf{AB}) = \det \mathbf{0} = 0.$$

左方是兩個實數的乘積，所以其中至少一個為零：$\det \mathbf{A} = 0$，或 $\det \mathbf{B} = 0$，或兩者皆為零。這就是所需的結果。

若 $\det \mathbf{B} \neq 0$，則 $\mathbf{B}^{-1}$ 存在，把 $\mathbf{AB} = \mathbf{0}$ 右乘 $\mathbf{B}^{-1}$，得

$$\mathbf{A} = \mathbf{ABB}^{-1} = \mathbf{0}\,\mathbf{B}^{-1} = \mathbf{0}.$$

所以 $\mathbf{A}$ 必定是零矩陣（zero matrix）。

不過，兩個行列式皆為零並不足夠。取

$$\mathbf{A} = \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

則 $\det \mathbf{A} = \det \mathbf{B} = 0$，但 $\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \mathbf{0}$。

三次式的因式分解

首先把多項式 $f(x) = x^3 + 2x^2 - 5x - 6$ 因式分解。試代 6 的細小因數：

$$\begin{aligned} f(-1) &= -1 + 2 + 5 - 6 = 0, \\ f(2) &= 8 + 8 - 10 - 6 = 0, \\ f(-3) &= -27 + 18 + 15 - 6 = 0, \end{aligned}$$

由於首項係數為 1，所以 $f(x) = (x+1)(x-2)(x+3)$。

現在展開 $(\mathbf{M} + p\mathbf{I})(\mathbf{M} + q\mathbf{I})(\mathbf{M} + r\mathbf{I})$。因為 $\mathbf{I}$ 與 $\mathbf{M}$ 可交換（而且 $\mathbf{I}^2 = \mathbf{I}$），展開的結果與數字的情況完全相同：

$$\mathbf{M}^3 + (p+q+r)\mathbf{M}^2 + (pq+qr+rp)\mathbf{M} + pqr\,\mathbf{I}.$$

與上面的多項式比較，我們可取

$$p = 1, \qquad q = -2, \qquad r = 3$$

（次序不拘），於是對任何方陣 $\mathbf{M}$，都有 $\mathbf{M}^3 + 2\mathbf{M}^2 - 5\mathbf{M} - 6\mathbf{I} = (\mathbf{M} + \mathbf{I})(\mathbf{M} - 2\mathbf{I})(\mathbf{M} + 3\mathbf{I})$。

> **想一想** 我們現在可否斷定三個因子之中必有一個是零矩陣？

不可以。第一部分已經說明，矩陣的乘積可以為零，而沒有任何一個因子為零；我們只知道至少有一個因子是奇異（singular）的。所以我們必須動手計算。

上三角矩陣的解

設 $\mathbf{M} = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$。每個因子都是上三角矩陣，而且右上角的元素同樣是 c：

$$\mathbf{M} + k\mathbf{I} = \begin{pmatrix} a+k & c \\ 0 & b+k \end{pmatrix}.$$

對於三個對角線為 (d_i, e_i)、右上角元素全為 c 的上三角矩陣，乘開後得

$$\begin{aligned} &\begin{pmatrix} d_1 & c \\ 0 & e_1 \end{pmatrix} \begin{pmatrix} d_2 & c \\ 0 & e_2 \end{pmatrix} \begin{pmatrix} d_3 & c \\ 0 & e_3 \end{pmatrix} \\ &\quad = \begin{pmatrix} d_1 d_2 d_3 & c\,(d_1 d_2 + d_1 e_3 + e_2 e_3) \\ 0 & e_1 e_2 e_3 \end{pmatrix}. \end{aligned}$$

取 $d_1, d_2, d_3 = a+1, a-2, a+3$ 及 $e_1, e_2, e_3 = b+1, b-2, b+3$，方程便化為三個條件：

$$\begin{aligned}(a+1)(a-2)(a+3) &= 0,\\ (b+1)(b-2)(b+3) &= 0,\\ c\left[(a+1)(a-2)+(a+1)(b+3)+(b-2)(b+3)\right] &= 0.\end{aligned}$$

頭兩個條件表示 $a, b \in \{-1, 2, -3\}$。把第三個條件中的括號展開，

$$(a^2-a-2)+(ab+3a+b+3)+(b^2+b-6) = a^2+ab+b^2+2(a+b)-5.$$

> **想一想** 要看出這個括號何時為零，有沒有比逐一試驗全部九對數值更聰明的方法？

有。若 $a \neq b$，括號等於

$$\frac{a^3-b^3}{a-b}+2\frac{a^2-b^2}{a-b}-5\frac{a-b}{a-b} = \frac{f(a)-f(b)}{a-b},$$

由於 $f(a) = f(b) = 0$，此式為 0。所以當 $a \neq b$ 時，無論 c 是甚麼，第三個條件都成立。

若 $a = b$，括號等於 $3a^2+4a-5 = f'(a)$。在 $a = -1, 2, -3$ 時，其值分別為 -6、15 和 10（全部不為零，因為 f 有三個相異的根），所以第三個條件迫使 $c = 0$。

因此，解為：

- $\mathbf{M} = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$，其中 a 和 b 是 $\{-1, 2, -3\}$ 中兩個不同的數，而 c 是任意實數（共六族）；
- 三個純量矩陣（scalar matrix）$-\mathbf{I}$、$2\mathbf{I}$ 和 $-3\mathbf{I}$（即 $a = b$ 的情況，此時必有 $c = 0$）。

驗算

取 $a = -1$、$b = 2$、$c = 5$。則 $\mathbf{M}+\mathbf{I} = \begin{pmatrix} 0 & 5 \\ 0 & 3 \end{pmatrix}$、$\mathbf{M}-2\mathbf{I} = \begin{pmatrix} -3 & 5 \\ 0 & 0 \end{pmatrix}$ 及 $\mathbf{M}+3\mathbf{I} = \begin{pmatrix} 2 & 5 \\ 0 & 5 \end{pmatrix}$。頭兩個的乘積已經是 $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$，所以無論 c 取何值，整個乘積都是零。另一方面，$\mathbf{M} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ 給出 $\mathbf{M}^3+2\mathbf{M}^2-5\mathbf{M}-6\mathbf{I} = \begin{pmatrix} 0 & 15 \\ 0 & 0 \end{pmatrix} \neq \mathbf{0}$，與 $f'(2) = 15$ 吻合。以電腦搜尋大量有理數 a, b, c，所得的解恰好就是上述各解；因式分解亦已用隨機的 3×3 整數矩陣驗證。

常見錯誤

- 由 $(\mathbf{M}+\mathbf{I})(\mathbf{M}-2\mathbf{I})(\mathbf{M}+3\mathbf{I})=\mathbf{0}$ 斷定 $\mathbf{M}$ 必定是 $-\mathbf{I}$、$2\mathbf{I}$ 或 $-3\mathbf{I}$。第一部分正是為了提醒你不要犯這個錯。
- 遺漏 $a=b$ 的情況，或以為在這情況下 c 同樣可以任意取值。
- 為第一部分舉例時，不慎寫出一個 $\mathbf{AB}=\mathbf{0}$ 的例子；要檢查你的乘積確實不是零。

STEP 之外

這個答案讓我們初次窺見*最小多項式*（minimal polynomial）：一個矩陣滿足某個具相異根的多項式，當且僅當它可對角化（diagonalisable），而且其特徵值（eigenvalue）都在這些根之中。我們那些 $a\neq b$ 的解，對任何 c 都可對角化；而 $\begin{pmatrix}a & c\\ 0 & a\end{pmatrix}$ 在 $c\neq 0$ 時則不可對角化，這正是它不成立的原因。右上角出現的均差（divided difference）$\big(f(a)-f(b)\big)/(a-b)$，在 $a=b$ 時變成 $f'(a)$，一般的矩陣函數正是這樣定義的。Strang 的 *Introduction to Linear Algebra*（線性代數入門）對特徵值和對角化有清晰的講解。

詳細解答

STEP II 1990，第 11 題

力學 • 角動量守恆

STEP II 1990，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/90-s2-q11

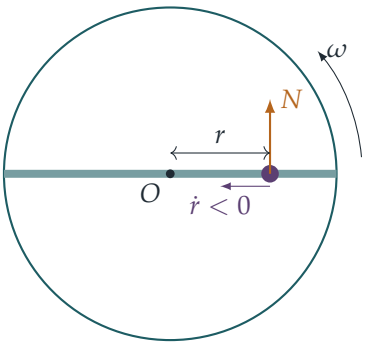
關鍵思路

圓盤與質點合起來是一個系統：對垂直軸沒有外力矩（torque），亦沒有外力作功，所以有兩個量守恆：對該軸的角動量（angular momentum）和動能（kinetic energy）。由這兩條方程可直接求得 ω 和 $\mathrm{d}r/\mathrm{d}t$。把兩個變率相除便可消去時間，得出一條描述路徑的可分離變數方程（separable equation），即 r 作為 θ 的函數。

解答

模型

把質點視為置於光滑凹槽內的點質量，並假設轉軸光滑（軸承無摩擦）。從上方俯視，質點位於凹槽上距離 r 處，沿凹槽方向的速度為 $\dot{r}$，垂直於凹槽方向的速度為 $r\omega$，其中 $\omega = \dot{\theta}$。



質點所受的水平力只有凹槽的反作用力 N，方向垂直於凹槽（凹槽光滑，故沿凹槽方向

沒有分量）。圓盤則受到 $-N$。

角動量

重力和垂直方向的反作用力均與轉軸平行，而光滑的轉軸不施加力矩，所以對轉軸的總角動量是常數。圓盤的貢獻是 $mk^2\omega$，質點的貢獻是 $m\,r\cdot r\omega = mr^2\omega$。與 $t=0$ 時比較：

$$mk^2\omega + mr^2\omega = mk^2\Omega + ma^2\Omega \quad\Longrightarrow\quad \omega = \frac{\Omega(k^2+a^2)}{k^2+r^2}.$$

能量

想一想 質點橫向移動時，反作用力 N 正好橫向推它。這樣不會改變動能嗎？

它會改變質點的動能，但不會改變總動能。N 對質點作功的功率為 $N\cdot r\omega$，而作用於圓盤上大小相等、方向相反的力，作用點同樣在距離 r 處，產生力矩 $-Nr$，故作功的功率為 $-Nr\omega$。兩者互相抵消，所以總動能守恆：

$$\tfrac{1}{2}mk^2\omega^2 + \tfrac{1}{2}m\big(\dot r^2 + r^2\omega^2\big) = \tfrac{1}{2}mk^2\Omega^2 + \tfrac{1}{2}m\big(V^2 + a^2\Omega^2\big).$$

因此

$$\dot r^2 = V^2 + (k^2+a^2)\Omega^2 - (k^2+r^2)\omega^2.$$

利用已知條件 $V^2 = \Omega^2 a^2(k^2+a^2)/k^2$，

$$V^2 + (k^2+a^2)\Omega^2 = \Omega^2(k^2+a^2)\,\frac{a^2+k^2}{k^2} = \frac{\Omega^2(k^2+a^2)^2}{k^2},$$

而由角動量的結果可得 $(k^2+r^2)\omega^2 = \Omega^2(k^2+a^2)^2/(k^2+r^2)$。於是

$$\dot r^2 = \Omega^2(k^2+a^2)^2\left(\frac{1}{k^2} - \frac{1}{k^2+r^2}\right) = \frac{\Omega^2(k^2+a^2)^2\,r^2}{k^2(k^2+r^2)}.$$

質點向轉軸移動時，$\dot r < 0$，所以取負平方根（其中 $r>0$）：

$$\frac{\mathrm{d}r}{\mathrm{d}t} = -\frac{\Omega r(k^2+a^2)}{k(k^2+r^2)^{1/2}},$$

這正是所求。

路徑

由於 $\omega = \mathrm{d}\theta/\mathrm{d}t > 0$，由鏈式法則（chain rule）得

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{\mathrm{d}r/\mathrm{d}t}{\mathrm{d}\theta/\mathrm{d}t} = -\frac{\Omega r(k^2+a^2)}{k(k^2+r^2)^{1/2}}\cdot\frac{k^2+r^2}{\Omega(k^2+a^2)} = -\frac{r(k^2+r^2)^{1/2}}{k},$$

即 $k\,\mathrm{d}r/\mathrm{d}\theta = -r(k^2+r^2)^{1/2}$。留意 Ω 和 a 已經消失。

現在設 $u = 1/r$，則 $\mathrm{d}r/\mathrm{d}\theta = -u^{-2}\,\mathrm{d}u/\mathrm{d}\theta$，而 $r(k^2+r^2)^{1/2} = u^{-2}(1+k^2u^2)^{1/2}$（因為 $u>0$）。方程變成

$$k\frac{\mathrm{d}u}{\mathrm{d}\theta} = \sqrt{1+k^2u^2}.$$

這個平方根提示我們使用雙曲代換（hyperbolic substitution）：設 $ku = \sinh\phi$。那麼 $k\,\mathrm{d}u/\mathrm{d}\theta = \cosh\phi\,\mathrm{d}\phi/\mathrm{d}\theta$，且 $\sqrt{1+\sinh^2\phi} = \cosh\phi$，所以

$$\frac{\mathrm{d}\phi}{\mathrm{d}\theta} = 1 \quad\Longrightarrow\quad \phi = \theta + \alpha.$$

當 $\theta = 0$ 時 $r = a$，所以 $\sinh\alpha = ku = k/a$。因此 $k/r = \sinh(\theta+\alpha)$，即

$$r\sinh(\theta+\alpha) = k, \qquad \sinh\alpha = \frac{k}{a}.$$

質點永遠不會到達轉軸

> **想一想** 當 $\theta\to\infty$ 時，$r = k/\sinh(\theta+\alpha)$ 趨向 0。圓盤有可能在有限時間內轉過無限大的角度嗎？

不可能。由於 $r^2 \geq 0$，$\omega = \Omega(k^2+a^2)/(k^2+r^2) \leq \Omega(k^2+a^2)/k^2$，所以到時間 t 為止，圓盤轉過的角度最多為 $\theta \leq \Omega(k^2+a^2)t/k^2$，這是有限的。對任何有限的 θ，$\sinh(\theta+\alpha)$ 都是有限的，所以 $r = k/\sinh(\theta+\alpha) > 0$。此外，對每個 $r>0$，$\dot r$ 都嚴格為負，所以質點永遠不會停下來再折返：這些公式對所有 t 都成立，而 r 不斷減小並趨向 0，卻永遠不會到達。

（另一做法是直接從 $\mathrm{d}r/\mathrm{d}t$ 的方程入手：由 a 走到 r 所需的時間為 $\frac{k}{\Omega(k^2+a^2)}\int_r^a \frac{\sqrt{k^2+x^2}}{x}\,\mathrm{d}x$，而在 $x=0$ 附近，被積函數的表現有如 k/x，其積分在 $r\to 0$ 時以對數方式發散（diverge）。）

驗算

- 在 $t=0$ 時，$r=a$ 給出 $\omega = \Omega$ 和 $\dot r = -\Omega a\sqrt{k^2+a^2}/k = -V$，與初始數據吻合；而 $\theta = 0$ 給出 $r = k/\sinh\alpha = a$。
- 當 θ 很大時，$r \approx 2ke^{-(\theta+\alpha)}$：質點以指數方式螺旋向內，與永遠不會到達的結論一致。
- 數值上，對隨機選取的 k, a, Ω 逐步計算運動方程（$\ddot r = r\omega^2$，角動量固定），所得的 $\dot r$、能量和 $r\sinh(\theta+\alpha) = k$ 均吻合至約 10^{-12}。取 $k = a = \Omega = 1$ 時，到達 $r = 10^{-8}$ 所需的時間約為 9.32，增長方式有如 $\frac{1}{2}\ln(1/r)$。

常見錯誤

- 以為質點本身的能量守恆：凹槽的反作用力對質點作功，守恆的只是總能量。
- 在質點的動能或角動量中漏掉橫向速度 $r\omega$。
- 只憑 θ 有限時 $r > 0$ 便斷定質點永遠不會到達轉軸，卻沒有檢查 θ 在有限時間內是否保持有限。

STEP 之外

關於 V 的特殊條件，令質點恰好落在永遠向 $r = 0$ 緩緩靠近的軌跡上：在旋轉參考系（rotating frame）中，$r = 0$ 是一個不穩定平衡點（unstable equilibrium），而這條軌跡就是它的*穩定流形*（*stable manifold*）。V 只要稍大一點，質點便會穿過中心；稍小一點，質點便會折返。這類想法，連同耦合物體的守恆定律，可見於 Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*。

詳細解答

STEP II 1990，第 12 題

力學 • 恢復係數與拋體

STEP II 1990，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：

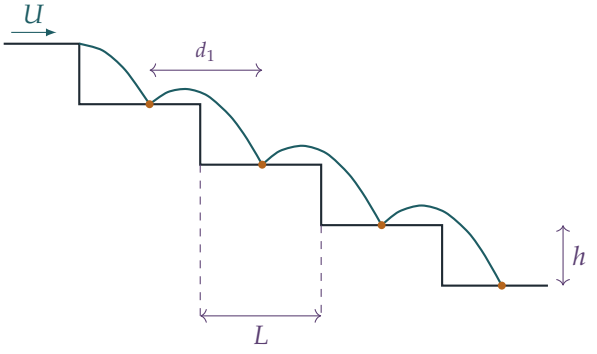
step.maths.org/questions/90-s2-q12

關鍵思路

樓梯是光滑（smooth）的，所以水平速度永遠不變：它始終是 U，而每一段水平距離都只是 U 乘以一段飛行時間。真正要花工夫的是垂直方向的運動：每次反彈令垂直速率乘以 e，而每下降 h 則令其平方增加 $2gh$。由此可得每次反彈時垂直速率的平方所滿足的一個簡單遞推關係（recurrence）。至於最後一部分，想想反彈點必須落在哪裏：在一道無窮無盡的樓梯上，每一級恰好落一次。

解答

我們把質點（particle）看成一個點，在兩次反彈之間只受重力作用；又把樓梯看成光滑的水平面，每一級比上一級低 h，相鄰梯級邊緣的水平距離為 L。



相鄰兩次反彈之間的距離

水平運動。樓梯是光滑的，所以梯級每次施加的衝量（impulse）都純粹沿垂直方向，而在兩次反彈之間唯一的力就是重力。因此水平速度任何時候都是 U，一段歷時 T 的飛行所走的水平距離為 UT。

垂直運動。設 w_n 為質點第 n 次反彈時撞上梯級的（向下）速率。質點離開最高一級時沒有垂直速度，然後下跌 h，所以

$$w_1^2 = 2gh.$$

在第 n 次反彈時，由於梯級固定不動，根據牛頓恢復定律（Newton's law of restitution），質點以速率 ew_n 向上彈起。之後它必須落在低 h 的下一級上，所以由 $v^2 = u^2 + 2as$（取向下為正，位移為 $+h$）得：

$$w_{n+1}^2 = e^2 w_n^2 + 2gh.$$

由 $w_1^2 = 2gh$ 出發，反覆應用這關係，

$$\begin{aligned} w_n^2 &= 2gh\left(1 + e^2 + e^4 + \cdots + e^{2(n-1)}\right) \\ &= 2gh\,\frac{1 - e^{2n}}{1 - e^2} = 2gh\,\alpha_n^2, \end{aligned}$$

這裏用了等比級數（geometric series）求和（因 $e^2 \neq 1$ 而成立）。所以 $w_n = \sqrt{2gh}\,\alpha_n$ 對每個 $n \geq 1$ 都成立；用歸納法（induction）簡單驗證一下便可使論證嚴謹，因為 $e^2 \cdot 2gh\,\alpha_n^2 + 2gh = 2gh\big(e^2\alpha_n^2 + 1\big)$，而 $e^2\alpha_n^2 + 1 = \dfrac{e^2 - e^{2n+2} + 1 - e^2}{1 - e^2} = \alpha_{n+1}^2$。

想一想 由第 n 次反彈到第 $(n+1)$ 次反彈，飛行時間有多長？

垂直速度以恆定的變率 g（向下）改變。它開始時是向上的 ew_n，結束時是向下的 w_{n+1}，所以總共改變了 $ew_n + w_{n+1}$，飛行時間為

$$T_n = \frac{ew_n + w_{n+1}}{g} = \frac{\sqrt{2gh}}{g}\left(e\alpha_n + \alpha_{n+1}\right) = \left(\frac{2h}{g}\right)^{\frac{1}{2}}\left(e\alpha_n + \alpha_{n+1}\right).$$

（這樣便不必解一條關於 T_n 的二次方程：我們早已知道末速率。）乘以水平速率 U，

$$d_n = U\left(\frac{2h}{g}\right)^{\frac{1}{2}}\left(e\alpha_n + \alpha_{n+1}\right),$$

正是所求。

一道很長的樓梯

由於 $0 \leq e < 1$，我們有 $e^{2n} \to 0$，所以當 $n \to \infty$ 時 $\alpha_n \to \dfrac{1}{\sqrt{1 - e^2}}$，因此

$$d_n \to D = U\left(\frac{2h}{g}\right)^{\frac{1}{2}}\frac{1 + e}{\sqrt{1 - e^2}} = U\left(\frac{2h}{g}\right)^{\frac{1}{2}}\left(\frac{1 + e}{1 - e}\right)^{\frac{1}{2}},$$

其中用了 $1-e^2=(1-e)(1+e)$。

想一想 為甚麼這個極限距離必須恰好等於 L？

由最高一級的邊緣起沿水平方向量度 x。下面第一級佔據 $0\le x\le L$，一般而言，最高一級以下的第 n 級佔據 $(n-1)L\le x\le nL$。質點在每一級上都反彈一次，所以它第 n 次反彈的位置（設為 x_n）滿足

$$(n-1)L\le x_n\le nL,\qquad \text{所以}\qquad \frac{x_n}{n}\to L\text{，當}n\to\infty.$$

另一方面，$x_n=x_1+d_1+d_2+\cdots+d_{n-1}$，而 $d_k\to D$。若 $D>L$，則由某一項起，每個 d_k 都比 L 大至少一個固定的量，於是質點每反彈一次，便相對梯級邊緣多走一段固定距離，最終會越過某一級而不落在其上；若 $D<L$，質點則逐步落後，最終會在同一級上落兩次。（嚴格來說，各項趨向 D 時，其平均值 x_n/n 亦趨向 D。）無論是哪種情況，只要樓梯夠長，質點都不會每級恰好反彈一次。所以對於很大的 N，我們需要 $D=L$：

$$U\left(\frac{2h}{g}\right)^{\frac12}\left(\frac{1+e}{1-e}\right)^{\frac12}=L\quad\Longrightarrow\quad U=\left(\frac{L^2g}{2h}\right)^{\frac12}\left(\frac{1-e}{1+e}\right)^{\frac12},$$

正是所求。

驗算

- 特殊值。$\alpha_1=1$，所以 $d_1=U\sqrt{2h/g}\,(e+\sqrt{1+e^2})$。當 $e=0$ 時，質點每次反彈後垂直方向完全停下，所以每段飛行都只是下跌 h 的自由下落，需時 $\sqrt{2h/g}$：果然 $d_n=U\sqrt{2h/g}$，而最後的條件變成 $U\sqrt{2h/g}=L$，完全符合預期。
- 量綱。$\sqrt{2h/g}$ 是時間，所以 d_n 是長度；$\sqrt{L^2g/h}$ 是速率。
- 數值檢驗。對隨機選取的 U、h 和 e（取 $g=9.8$）以數值方法逐步模擬這運動，所得的每個 d_n 與公式相差都在 10^{-6} 以內。取 $e=\frac12$、$L=0.3$ 和 $h=0.2$，公式給出 $U\approx0.857$，而其後的反彈位置在 2000 級樓梯上都依次落在相繼的梯級上。

常見錯誤

- 把每次反彈都當成由靜止下跌 h，或由上一次反彈的高度下跌：質點每次反彈後都會上升，所以它的垂直速率會一級一級地累積起來。
- 寫成 $w_{n+1}=ew_n+\sqrt{2gh}$：相加的是速率的平方，而不是速率本身。

STEP 之外

最後一部分其實是一個關於平均值的命題：若 $d_n \to D$，則其累計平均值亦滿足 $\frac{1}{n}\sum d_k \to D$（切薩羅定理，Cesàro's theorem），這是實分析（real analysis）入門階段的一道標準習題。還要留意，關於 U 的條件是必要（necessary）條件，而非充分（sufficient）條件：最初幾次反彈的距離比 L 短，而若 e 接近 1，累積起來的差額可以超過起步時領先的 x_1，那麼任何發射速率都行不通。想看數列極限的嚴謹處理，可參閱 Abbott 的 *Understanding Analysis*；至於碰撞的力學，可參閱 Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*。

詳細解答

STEP II 1990，第 13 題

力學 • 桿的平衡

STEP II 1990，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/90-s2-q13

關鍵思路

由於釘子光滑（smooth）而繩子為輕繩（light），繩上各處的張力（tension）都相同，所以桿受到兩個相等的拉力，分別沿 PO 和沿 QO 方向。它們的合力（resultant）平分角 POQ，而要達致平衡（equilibrium），這個合力必須與重量沿同一條鉛垂線互相抵銷。因此，過 O 的鉛垂線就是角 POQ 的角平分線，並且通過 G，餘下的工作交給角平分線定理（angle bisector theorem）便可。至於那個不等式，其實只是說三角形 OPQ 是一個真正的三角形。

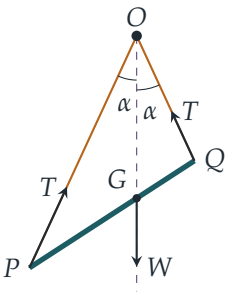
解答

求 OP 和 OQ

把桿模型化為一枝剛性桿，其重心（centre of gravity）為 G，其中 $PG = a + d$ 及 $GQ = a - d$；我們取 $0 \leq d < a$，使得 G 位於桿上，在中點靠 Q 的一側（或正好在中點）。繩子是不可伸長（inextensible）的輕繩，釘子細小而光滑，所以 OP 和 OQ 兩段繩的張力 T 相同，並且

$$OP + OQ = 2b.$$

作用在桿上的力只有三個：作用於 G 的重量 W、在 P 處沿 PO 方向的張力 T，以及在 Q 處沿 QO 方向的張力 T。



兩個張力的作用線相交於 O，所以（對 O 取力矩（moments））重量的作用線也必須通過 O：過 O 的鉛垂線通過 G。此外，兩個張力的大小相等，所以它們的合力指向兩個單位向量之和的方向，這兩個單位向量一個沿 PO，一個沿 QO，而這正是兩者夾角的平分線方向。這個合力必須是鉛垂的，才能與 W 平衡。因此鉛垂線 OG 平分角 POQ；把每一半記作 α。

想一想 知道 OG 平分角 POQ，能告訴我們關於各長度的甚麼資料？

在三角形 OPQ 中應用角平分線定理，由 O 引出的角平分線把 PQ 按鄰邊之比分割：

$$\frac{OP}{OQ} = \frac{PG}{GQ} = \frac{a+d}{a-d}.$$

（直接的看法：三角形 OPG 和 OQG 有共同的由 O 引出的高，所以它們的面積之比為 $PG : GQ$；而它們的面積又分別是 $\frac{1}{2} OP \cdot OG \sin\alpha$ 和 $\frac{1}{2} OQ \cdot OG \sin\alpha$。）所以對某個 $k > 0$，有 $OP = k(a+d)$ 及 $OQ = k(a-d)$，而由 $OP + OQ = 2b$ 得 $2ka = 2b$，故 $k = b/a$：

$$OP = \frac{b(a+d)}{a}, \qquad OQ = \frac{b(a-d)}{a}.$$

不等式

設 $\theta = \angle POQ = 2\alpha$。由於桿並非鉛垂，$O$ 不在直線 PQ 上，所以 OPQ 是一個真正的三角形，且 $0 < \theta < \pi$。以 $PQ = 2a$ 代入餘弦定理（cosine rule），得

$$\begin{aligned} 4a^2 &= OP^2 + OQ^2 - 2\,OP \cdot OQ \cos\theta \\ &= \frac{b^2}{a^2}\Big[(a+d)^2 + (a-d)^2 - 2(a^2-d^2)\cos\theta\Big] \\ &= \frac{2b^2}{a^2}\Big[(a^2+d^2) - (a^2-d^2)\cos\theta\Big]. \end{aligned}$$

現在，因為 $\theta \neq 0$，所以 $\cos\theta < 1$；又因為 $d < a$，所以 $a^2 - d^2 > 0$。於是

$$4a^2 > \frac{2b^2}{a^2}\Big[(a^2+d^2) - (a^2-d^2)\Big] = \frac{4b^2d^2}{a^2},$$

亦即 $b^2d^2 < a^4$。取正平方根（其中 $b > 0$，$d \geq 0$），得 $bd < a^2$，所以

$$d < \frac{a^2}{b},$$

這正是所要證明的結果。

換一個等價的說法，這就是三角不等式（triangle inequality）$OP - OQ < PQ$：這裏 $OP - OQ = 2bd/a$，而 $2bd/a < 2a$ 正是同一個條件。（另一個不等式 $\cos\theta > -1$ 給出 $OP + OQ > PQ$，即 $b > a$，這只是說繩子的長度足以繞過釘子而已。）

驗算

- 若 $d = 0$（均勻桿），則 $OP = OQ = b$，桿水平懸掛，正如對稱性所要求的；不等式 $0 < a^2/b$ 自動成立。
- 當 $d \to a^2/b$ 時，$OP - OQ \to 2a = PQ$，三角形被壓扁，桿變成鉛垂：超過這個值之後，桿唯一的懸掛方式就是鉛垂懸掛。
- 取 $a = 1$，$b = 2$，$d = 0.3$：$OP = 2.6$，$OQ = 1.4$，$\cos\theta \approx 0.648$。把這個三角形擺放成 OG 為鉛垂，兩個張力的水平分量互相抵銷，而對 O 的力矩在機器精度之內平衡。對繩子的所有位置逐一掃描，確認了這是唯一的非鉛垂平衡位置，並且一旦 $d \geq a^2/b$ 便不存在這樣的位置。

常見錯誤

- 假設兩段繩的張力不同。繩子繞過光滑釘子時，兩段的張力相等，而正是這一點令 OG 成為角平分線。
- 把角平分線的比例弄反了：較長的一段繩 OP 連接的是離 G 較遠的一端。
- 以為這個平衡是穩定的。事實上它是不穩定的：當 $d = 0$ 時，水平的桿的 G 在 $\sqrt{b^2 - a^2}$ 的距離處位於 O 之下，而鉛垂懸掛時則在其下方 b 的距離處，位置更低。題目只問及平衡。

STEP 之外

相對於桿而言，O 的軌跡（locus）是一個以 P 和 Q 為焦點（foci）的橢圓，而平衡位置就是 G 的高度取駐值之處。這就是虛功原理（principle of virtual work），由此可以通往拉格朗日力學（Lagrangian mechanics）。它也解釋了為甚麼相等的張力會平分夾角：這正是橢圓的反射性質。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教科書）是進一步學習靜力學和穩定性的好選擇。

詳細解答

STEP II 1990，第 14 題

力學 • 彈性碰撞

STEP II 1990，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/90-s2-q14

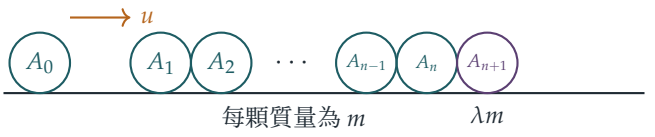
本題在 *Beyond STEP* 第 10.1 節牛頓擺中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch10/newtons-cradle.html

關鍵思路

我們不知道衝量（impulse）怎樣沿着這列彈珠傳遞，所以無法逐次追蹤每一下碰撞。我們能做的，是用三個條件去檢驗每個提出的最終運動：動量（momentum）守恆、動能（kinetic energy）守恆，以及彈珠永遠不會互相穿過。第三個條件很容易被遺忘，但每一部分的結論正是由它決定的。若一顆彈珠的正後方緊貼着一顆靜止的彈珠，它就不能向後運動；而一顆彈珠的速度也不能比它正前方的彈珠快。

解答

我們把彈珠看成在一條光滑直線（即槽）上的質點（particle），速度以 A_0 原來的運動方向為正。碰撞之前，總動量是 mu，總動能是 $\frac{1}{2}mu^2$，兩者都守恆。



由於彈珠不能重疊，任何可能的最終運動都必須符合以下兩點。

- 若 A_k 靜止，而 A_{k+1} 在它的正前方並與它接觸，則 A_{k+1} 的速度不能是負數。
- 若 A_k 與 A_{k+1} 互相接觸，則 A_k 的速度不能大於 A_{k+1} 的速度。

第 (i) 部分

假設只有 A_n 和 A_{n+1} 運動，速度分別為 v 和 w。兩式同除以 m，得

$$v + \lambda w = u,$$
$$v^2 + \lambda w^2 = u^2.$$

把 $v = u - \lambda w$ 代入能量方程，得

$$u^2 - 2\lambda uw + \lambda^2 w^2 + \lambda w^2 = u^2 \quad \Longrightarrow \quad \lambda w\big((1+\lambda)w - 2u\big) = 0.$$

根 $w = 0$ 給出 $v = u$。這樣 A_n 就會向前撞進靜止的 A_{n+1}，這是不可能的。另一個根是

$$w = \frac{2u}{1+\lambda}, \qquad v = u - \lambda w = \frac{(1-\lambda)u}{1+\lambda}.$$

若 $\lambda < 1$：則 $v > 0$，所以 A_n 離開靜止的 A_{n-1}。此外

$$w - v = \frac{(1+\lambda)u}{1+\lambda} = u > 0,$$

所以 A_{n+1} 離開 A_n。兩條守恆定律都成立，而且沒有彈珠重疊，因此這個運動是可能的，其中 $A_0, \ldots, A_{n-1}$ 全部靜止。

若 $\lambda > 1$：則 $v < 0$，所以 A_n 必須向後運動。但 A_{n-1} 就在它的正後方，與它接觸而且靜止。因此，唯一同時滿足兩條守恆定律的運動在物理上並不可能，所以當 $\lambda > 1$ 時，這類運動並不存在。

第 (ii) 部分

假設只有 A_{n-1}、A_n 和 A_{n+1} 運動，速度分別為 a、b 和 c。由於 $n \geq 3$，彈珠 A_{n-2} 存在，它靜止不動，並從後面與 A_{n-1} 接觸。所以 $a \geq 0$。由於三顆運動中的彈珠不能互相超越，$a \leq b \leq c$。因此 a、b 和 c 全部非負。

> ? **想一想** 兩條守恆定律都給出關於 a、b、c 的方程。怎樣可以不解方程而比較它們？

我們把動量方程平方，再減去能量方程。由守恆得 $a + b + \lambda c = u$ 和 $a^2 + b^2 + \lambda c^2 = u^2$，所以

$$(a + b + \lambda c)^2 - (a^2 + b^2 + \lambda c^2) = u^2 - u^2 = 0.$$

展開左方，得

$$\lambda(\lambda - 1)c^2 + 2ab + 2\lambda c(a + b) = 0.$$

當 $\lambda > 1$ 且 $a, b, c \geq 0$ 時，左方每一項都非負，所以每一項都必須是零。特別地，$\lambda(\lambda - 1)c^2 = 0$，由此得 $c = 0$。於是 $0 \leq a \leq b \leq c = 0$，三顆彈珠全部靜止。由於 $u > 0$，這與動量方程矛盾。因此當 $\lambda > 1$ 時，這樣的運動並不存在。

第 (iii) 部分

(ii) 的論證適用於任何一組運動中的彈珠：若所有速度都非負，而 A_{n+1} 運動，則動能會太小。所以當 $\lambda > 1$ 時，必定有某顆彈珠向後運動。除了 A_0 之外，每顆彈珠的正後方都有另一顆彈珠，所以 A_0 自然是最有可能反彈的一顆。

想一想 我們應該試哪兩顆彈珠？

試試 A_0 和 A_{n+1}，速度分別為 p 和 q，而 $A_1, \ldots, A_n$ 保持靜止。所得的兩條方程與 (i) 中的相同，只是以 p, q 代替 v, w，而非平凡解（non-trivial solution）是

$$p = \frac{(1-\lambda)u}{1+\lambda} = -\frac{(\lambda-1)u}{\lambda+1}, \qquad q = \frac{2u}{1+\lambda}.$$

與 (i) 一樣，動量和能量都守恆。這個運動在物理上也是可能的：

- $p < 0$，所以 A_0 向後運動，離開 A_1，而 A_0 後面沒有任何東西。
- $q > 0$，所以 A_{n+1} 離開靜止的 A_n。

所以其中一個可能的運動是：A_0 以速率 $\dfrac{(\lambda-1)u}{\lambda+1}$ 反彈，A_{n+1} 以速率 $\dfrac{2u}{\lambda+1}$ 向前運動，而 $A_1, \ldots, A_n$ 保持靜止。

如果我們把這次碰撞看成相鄰彈珠之間一連串獨立的碰撞，結果也是如此。每對質量相等的彈珠會交換速度，所以速率 u 一直傳遞到 A_n。當 A_n 撞上 A_{n+1} 時，它以速度 p 反彈，然後這個向後的速度再沿着這列彈珠傳回 A_0。

驗算

- 當 $\lambda = 1$ 時，所有彈珠完全相同，兩個答案都變成 A_{n+1} 以速率 u 向前運動，而其餘全部靜止，這就是大家熟悉的「牛頓擺」（Newton's cradle）。
- 取 $\lambda = 2$ 和 $u = 3$，第 (iii) 部分給出 $p = -1$ 和 $q = 2$。動量：$-1 + 2 \times 2 = 3$。能量：$1 + 2 \times 4 = 9 = 3^2$。
- 我們就 $n = 3$ 和 $n = 6$ 以數值方法逐步計算了這一連串兩兩之間的彈性碰撞（elastic collision）。當 $\lambda = 2$ 時，最終速度是 $-\frac{1}{3}u$ 和 $\frac{2}{3}u$；當 $\lambda = 4$ 時，則是 $-0.6u$ 和 $0.4u$，與預測完全相符。當 $\lambda = \frac{1}{2}$ 時，模擬結束時 A_n 和 A_{n+1} 分別以 $\frac{1}{3}u$ 和 $\frac{4}{3}u$ 運動，與 (i) 吻合。

常見錯誤

- 解了兩條守恆方程便停下來。每一部分都取決於所得的解在物理上是否可能，即取決於各速度的正負，以及彈珠會否互相穿過。
- 在 (i) 中保留 $w = 0$、$v = u$ 這個根。這個根只是描述碰撞之前的運動。
- 忘記了在 (ii) 中，正是 $n \geq 3$ 保證了 A_{n-1} 後面有一顆靜止的 A_{n-2}。

STEP 之外

對於一列互相接觸的彈珠，單憑兩條守恆定律並不能確定結果。這道題真正要問的是：在眾多滿足守恆量的解之中，哪些在物理上是可能的。要在它們之間作出選擇，就需要一個更仔細的接觸模型，例如一連串兩兩之間的碰撞，或把彈珠當作剛硬的彈簧。把彈珠當作彈簧來建立模型，就會引向質點鏈中的波動理論。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）就是以這種精神講解碰撞和守恆定律的。

詳細解答

STEP II 1990，第 15 題

概率與統計 • 順序統計量

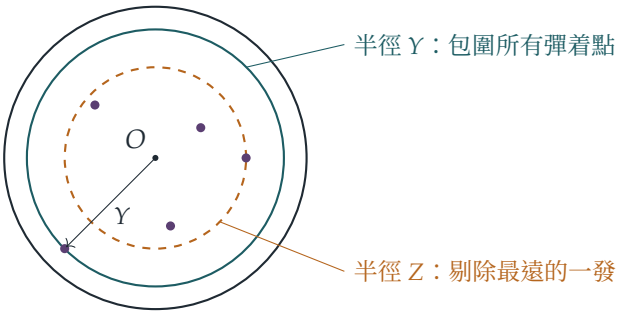
STEP II 1990，第 15 題

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/90-s2-q15

關鍵思路

以 O 為圓心、能包圍一組彈着點的最小圓，其半徑等於最遠一發與圓心的距離，所以 Y 是若干個獨立（independent）距離的最大值。最大值不超過 y，當且僅當每一發都落在距離 y 之內，於是累積分佈函數（cumulative distribution function）便化為一個乘積。剔除最遠的一發之後，我們要求的半徑是第二大的距離，而它的分佈函數可以透過數算有多少發可以落在給定的圓之外而求得。



解答

設 $R_1, \ldots, R_n$ 為 n 發射擊的彈着點與 O 的距離。它們互相獨立（因為各發是獨立射出的），而且每一個的分佈函數都是

$$P(R_i \le t) = t^2, \qquad 0 \le t \le 1.$$

Y 的分佈

以 O 為圓心、包含所有彈着點的最小圓，半徑為 $Y = \max(R_1, \ldots, R_n)$。對於 $0 \le y \le 1$，

$$P(Y \le y) = P(R_1 \le y, \ldots, R_n \le y) = \prod_{i=1}^{n} P(R_i \le y) = (y^2)^n = y^{2n},$$

其中中間一步用到獨立性。求導後，Y 的概率密度函數（probability density function）為

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} y^{2n} = 2ny^{2n-1}, \qquad 0 \le y \le 1,$$

正如題目所求。（它在 $[0,1]$ 之外為零，而且積分為 1，理應如此。）

圓的面積為 πY^2，所以

$$E(\pi Y^2) = \pi \int_0^1 y^2 \cdot 2ny^{2n-1}\,\mathrm{d}y = 2n\pi \int_0^1 y^{2n+1}\,\mathrm{d}y = \frac{2n\pi}{2n+2} = \frac{n\pi}{n+1}.$$

剔除最遠的一發

我們需要 $n \ge 2$，這樣才會有彈着點剩下。剔除最遠的一發之後，以 O 為圓心的最小包圍圓的半徑為 Z，即 $R_1, \ldots, R_n$ 之中第二大的一個。

> **想一想** 怎樣用「有多少發落在半徑為 z 的圓之外」來描述事件 $Z \le z$？

第二大的距離不超過 z，當且僅當最多只有一發與 O 的距離大於 z。每一發各自獨立地以概率 z^2 落在距離 z 之內，所以落在半徑為 z 的圓之外的發數服從二項分佈（binomial distribution），$\mathrm{B}(n, 1-z^2)$。因此，對於 $0 \le z \le 1$，

$$\begin{aligned} P(Z \le z) &= P(\text{沒有一發在圓外}) + P(\text{恰好一發在圓外}) \\ &= z^{2n} + n(1-z^2)z^{2(n-1)} \\ &= nz^{2n-2} - (n-1)z^{2n}. \end{aligned}$$

求導便得 Z 的密度函數：

$$f_Z(z) = 2n(n-1)z^{2n-3} - 2n(n-1)z^{2n-1} = 2n(n-1)z^{2n-3}(1-z^2), \qquad 0 \le z \le 1.$$

因此新圓的面積的期望值（expected value）為

$$\begin{aligned} E(\pi Z^2) &= 2n(n-1)\pi \int_0^1 \left(z^{2n-1} - z^{2n+1}\right)\mathrm{d}z \\ &= 2n(n-1)\pi \left(\frac{1}{2n} - \frac{1}{2n+2}\right) \\ &= 2n(n-1)\pi \cdot \frac{2}{2n(2n+2)} = \left(\frac{n-1}{n+1}\right)\pi, \end{aligned}$$

這正是所求的結果。

驗算

- 當 $n = 1$ 時，第一個答案是 $\pi/2$：單獨一發的 R^2 在 $[0,1]$ 上服從均勻分佈（uniform distribution）（因為 $P(R^2 \le u) = u$），所以面積 πR^2 的期望值是 $\pi/2$。
- 當 $n = 2$ 時，第二個答案是 $\pi/3$，即兩發之中較近一發所對應的面積期望值；連同第一個答案 $2\pi/3$，兩者的平均是 $\pi/2$，即單獨一發時的值，這是必然的。
- 以 2×10^5 次試驗及 $n = 5$ 進行模擬（simulation），得到的面積期望值為 2.618 和 2.095，而理論值為 $5\pi/6 \approx 2.618$ 和 $4\pi/6 \approx 2.094$。

常見錯誤

- 寫成 $P(Z \le z) = z^{2(n-1)}$，彷彿餘下的彈着點是一個全新的、大小為 $n-1$ 的樣本；被剔除的一發並不是隨機選出的，所以留下來的彈着點會偏向 O。
- 忘記了因子 n，它代表落在圓外的是哪一發。
- 先求 $E(Y)$ 再把它平方：面積的期望值是 $\pi E(Y^2)$，而不是 $\pi\big(E(Y)\big)^2$。

STEP 之外

這些就是順序統計量（order statistics）。由於 $U_i = R_i^2$ 在 $[0,1]$ 上服從均勻分佈，這些面積其實就是 n 個均勻隨機變量的順序統計量，而第 k 小的一個（取自 n 個獨立的均勻隨機變量）服從貝塔分佈（beta distribution），平均值為 $k/(n+1)$；本題的兩個答案分別是 $k = n$ 和 $k = n-1$ 的情況。一個簡潔的圖像是：n 個隨機點把 $[0,1]$ 分割成 $n+1$ 段期望長度相等的間隙。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教科書）涵蓋了這類變換和順序統計量。

詳細解答

STEP II 1990，第 16 題

概率與統計 • 貝葉斯定理

STEP II 1990，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/90-s2-q16

關鍵思路

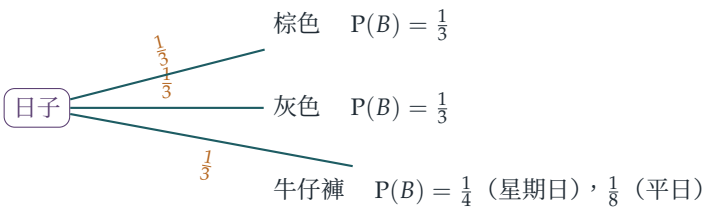
這其實是喬裝了的貝葉斯定理（Bayes' theorem）。照片告訴我們當天繫的是黑色領帶，而我們要求的是已知這一點時，那天是某一天的概率。所以我們先順着推：對每一類日子，用一幅按褲子分支的樹形圖（tree diagram），求出繫黑色領帶的概率。然後再把條件概率（conditional probability）倒轉過來。星期一的規則只改動樹形圖的其中一條分支，所以第二部分幾乎不用多花工夫。

解答

設 S 為照片攝於星期日的事件，B 為我繫着黑色領帶的事件。照片出自七天中任何一天的可能性均等，所以在看照片之前，$\mathrm{P}(S) = \frac{1}{7}$。除了（稍後的）星期一之外，每一天穿每條褲子的概率都是 $\frac{1}{3}$。

各類日子繫黑色領帶的機會

- **棕色褲子**。領巾已被排除，所以領帶是從黑色、波點和軍團領帶三者中隨機選出：$\mathrm{P}(B) = \frac{1}{3}$。
- **灰色褲子**。波點領帶已被排除，所以是從黑色領帶、軍團領帶和領巾中選：$\mathrm{P}(B) = \frac{1}{3}$。
- **星期日穿牛仔褲**。一定會繫領帶，從全部四條中選：$\mathrm{P}(B) = \frac{1}{4}$。
- **平日穿牛仔褲**。有 $\frac{1}{2}$ 的概率穿運動衫而不繫領帶。否則就是四條領帶的其中一條：$\mathrm{P}(B) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$。



沿各分支相加（即按褲子劃分的全概率公式（law of total probability）），

$$\mathrm{P}(B \mid \text{星期日}) = \tfrac{1}{3}\left(\tfrac{1}{3}+\tfrac{1}{3}+\tfrac{1}{4}\right) = \tfrac{1}{3}\cdot\tfrac{11}{12} = \tfrac{11}{36} = \tfrac{22}{72},$$
$$\mathrm{P}(B \mid \text{平日}) = \tfrac{1}{3}\left(\tfrac{1}{3}+\tfrac{1}{3}+\tfrac{1}{8}\right) = \tfrac{1}{3}\cdot\tfrac{19}{24} = \tfrac{19}{72}.$$

第一個答案

根據貝葉斯定理，星期日的先驗概率（prior）為 $\frac{1}{7}$，六個平日合計為 $\frac{6}{7}$，

$$\mathrm{P}(S \mid B) = \frac{\frac{1}{7}\cdot\frac{22}{72}}{\frac{1}{7}\cdot\frac{22}{72}+\frac{6}{7}\cdot\frac{19}{72}} = \frac{22}{22+6\times 19} = \frac{22}{136} = \frac{11}{68},$$

正如所求。（公因數 $\frac{1}{7\times 72}$ 會約去，這就是為甚麼值得把所有分數都通分成以 72 為分母。）

> **想一想** 為甚麼答案不直接是 $\mathrm{P}(B \mid \text{星期日})$，或者 $\frac{1}{7}$？

$\mathrm{P}(B \mid \text{星期日}) = \frac{11}{36}$ 答的是另一條問題：它是已知日子時繫這條領帶的機會，而不是已知領帶時那天是星期日的機會。而 $\frac{1}{7}$ 則忽略了證據。星期日繫黑色領帶的可能性稍高一點（$\frac{22}{72}$ 對 $\frac{19}{72}$），因為那天沒有不繫領帶這個選擇，所以這張照片把星期日的概率由 $\frac{1}{7} = \frac{11}{77}$ 輕輕推高至 $\frac{11}{68}$。

加入星期一的規則

現在星期一必定穿牛仔褲，而星期一仍然是平日，所以運動衫這個選擇依然適用：

$$\mathrm{P}(B \mid \text{星期一}) = 1\cdot\tfrac{1}{2}\cdot\tfrac{1}{4} = \tfrac{1}{8} = \tfrac{9}{72}.$$

星期日和其餘五個平日不變。每一天的先驗概率仍然是 $\frac{1}{7}$，可以約去，所以

$$\mathrm{P}(S \mid B) = \frac{\frac{22}{72}}{\frac{22}{72}+\frac{9}{72}+5\cdot\frac{19}{72}} = \frac{22}{22+9+95} = \frac{22}{126} = \frac{11}{63}.$$

這比 $\frac{11}{68}$ 稍大一點。這是合理的：星期一繫黑色領帶的可能性降低了（由 $\frac{19}{72}$ 變成 $\frac{1}{8}$），所以一張繫黑色領帶的照片會令星期一的可能性下降，而稍為更傾向星期日。

驗算

- 每一天各條領帶的概率，連同「不繫領帶」，加起來等於 1。例如，平日不繫領帶的機會是 $\frac{1}{3}\cdot\frac{1}{2}=\frac{1}{6}$。
- 模擬 4×10^5 個隨機日子，得出繫黑色領帶的日子當中星期日所佔的比例約為 0.161（準確值 $\frac{11}{68}\approx0.1618$），加入星期一的規則後則約為 0.174（準確值 $\frac{11}{63}\approx0.1746$）。

常見錯誤

- 忘記了星期一穿牛仔褲時，仍然有一半機會穿運動衫：星期一是平日。星期一若用 $\frac{1}{4}$ 而不是 $\frac{1}{8}$，會得出 $\frac{22}{135}$。
- 穿棕色或灰色褲子時從全部四條領帶中選。這兩條褲子各自只剩下三條領帶可選。

STEP 之外

把「已知假設時證據的概率」轉化為「已知證據時假設的概率」，正是貝葉斯推斷（Bayesian inference）的核心。有一個簡潔的說法：後驗賠率（posterior odds）等於先驗賠率乘以似然比（likelihood ratio）。這裏星期日對平日的先驗賠率是 1 : 6，似然比是 22 : 19，於是得出 22 : 114，即 11 : 57。大學課程會由此發展出貝葉斯統計學（Bayesian statistics），當中整個分佈都以同樣的方式更新。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門）有一章講條件概率，寫得十分出色，並附有許多像本題這樣的例子。

詳細解答

STEP III 1990，第 1 題

純數學 • 倍角與多項式的根

STEP III 1990，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/90-s3-q1

關鍵思路

九等於 3×3，所以只要把正切的三倍角公式（triple-angle formula）用兩次，就可以得到 $\tan 9\theta$。這比展開 $(\cos\theta + \mathrm{i}\sin\theta)^9$ 省事得多。因式分解後的形式還告訴我們，在 $\theta = \pi/9, 2\pi/9, 4\pi/9$ 時是哪一個因式等於零：這些角正好是令 $\tan^2 3\theta = 3$ 的角。由此得出一個以 $u = \tan^2\theta$ 為未知數的三次方程（cubic），餘下的工作就交給韋達公式（Vieta's formulas，即多項式的根與係數之間的關係）。

解答

$\tan 9\theta$ 的公式

設 $t = \tan\theta$ 及 $T = \tan 3\theta$。三倍角公式可由 $\tan(2\theta + \theta)$ 推出，亦可由棣美弗定理（de Moivre's theorem）取 $n = 3$ 得出，它給出

$$T = \frac{3t - t^3}{1 - 3t^2} = \frac{t(t^2 - 3)}{3t^2 - 1}, \qquad \tan 9\theta = \frac{3T - T^3}{1 - 3T^2} = \frac{T(3 - T^2)}{1 - 3T^2}.$$

現在把 $3 - T^2$ 和 $1 - 3T^2$ 都寫成以 $(1 - 3t^2)^2$ 為公分母的形式：

$$\begin{aligned}(1 - 3t^2)^2(3 - T^2) &= 3(1 - 3t^2)^2 - t^2(3 - t^2)^2 \\ &= (3 - 18t^2 + 27t^4) - (9t^2 - 6t^4 + t^6) \\ &= -(t^6 - 33t^4 + 27t^2 - 3),\end{aligned}$$

$$\begin{aligned}(1 - 3t^2)^2(1 - 3T^2) &= (1 - 3t^2)^2 - 3t^2(3 - t^2)^2 \\ &= (1 - 6t^2 + 9t^4) - (27t^2 - 18t^4 + 3t^6) \\ &= -(3t^6 - 27t^4 + 33t^2 - 1).\end{aligned}$$

兩式相除時，因子 $-(1-3t^2)^{-2}$ 互相約去，所以

$$\tan 9\theta = \frac{t(t^2-3)}{3t^2-1}\cdot\frac{t^6-33t^4+27t^2-3}{3t^6-27t^4+33t^2-1},$$

這就是所求的結果。

若改用棣美弗定理，則 $\tan 9\theta = \operatorname{Im}(1+\mathrm{i}t)^9 / \operatorname{Re}(1+\mathrm{i}t)^9$。這樣便要再驗證

$$9t-84t^3+126t^5-36t^7+t^9 \quad 和 \quad 1-36t^2+126t^4-84t^6+9t^8$$

確實可以分解成上面的分子和分母。它們的確可以，但經由 $\tan 3\theta$ 這條路，因式分解是現成的，不費工夫。

三次方程

設 $P(t) = t^6-33t^4+27t^2-3$。上面第一個計算表明

$$3-\tan^2 3\theta = -\frac{P(t)}{(1-3t^2)^2} \qquad 只要 3t^2 \neq 1.$$

現在取 $\theta = \pi/9,\ 2\pi/9,\ 4\pi/9$。這時 $3\theta = \pi/3,\ 2\pi/3,\ 4\pi/3$，而每一個情況都有 $\tan 3\theta = \pm\sqrt{3}$，所以 $\tan^2 3\theta = 3$。每一個情況下 3θ 都不是 $\pi/6$ 的奇數倍，所以 $3t^2 \neq 1$（否則 $\tan 3\theta$ 便沒有定義）。因此，對這三個角都有 $P(\tan\theta) = 0$。

在 $(0, \pi/2)$ 之內，這些正好是令 $\tan 9\theta = 0$ 的原因在於因式 P，而不在於 t 或 t^2-3 的角。事實上，$t \neq 0$，而且因為 $\theta \neq \pi/3$，所以 $t^2 \neq 3$。

> **想一想** 我們怎樣知道這三個數就是該三次方程的全部根，而且沒有重根？

令 $u = t^2$。$\tan^2(\pi/9)$、$\tan^2(2\pi/9)$、$\tan^2(4\pi/9)$ 每一個都是

$$u^3-33u^2+27u-3 = 0.$$

的根。這三個值互不相同，因為 tan 在 $(0, \pi/2)$ 上取正值且嚴格遞增（strictly increasing），而 $\pi/9 < 2\pi/9 < 4\pi/9 < \pi/2$。三次方程最多有三個根，所以這三個數正好就是它的全部根。這就是所求的整數係數三次方程。

乘積

根據韋達公式，$u^3-33u^2+27u-3$ 的根之積是 3：

$$\tan^2\left(\tfrac{\pi}{9}\right)\tan^2\left(\tfrac{2\pi}{9}\right)\tan^2\left(\tfrac{4\pi}{9}\right) = 3.$$

三個角都在 $(0, \pi/2)$ 之內，所以每個正切值都是正數，我們取正平方根：

$$\tan\left(\tfrac{\pi}{9}\right)\tan\left(\tfrac{2\pi}{9}\right)\tan\left(\tfrac{4\pi}{9}\right) = \sqrt{3}.$$

六次方之和

設三個根為 u_1, u_2, u_3。由該三次方程可知

$$e_1 = u_1 + u_2 + u_3 = 33, \qquad e_2 = u_1u_2 + u_2u_3 + u_3u_1 = 27, \qquad e_3 = u_1u_2u_3 = 3.$$

我們要求的是 $u_1^3 + u_2^3 + u_3^3$，也就是三個正切值的六次方之和。

> **想一想** 不把根求出來，怎樣由 e_1, e_2, e_3 得到立方和？

每個根都滿足 $u_i^3 = 33u_i^2 - 27u_i + 3$。把這三條等式相加，得

$$\sum u_i^3 = 33\sum u_i^2 - 27\sum u_i + 9.$$

這裏 $\sum u_i = 33$，而

$$\sum u_i^2 = e_1^2 - 2e_2 = 1089 - 54 = 1035.$$

因此

$$\begin{aligned}\sum u_i^3 &= 33 \times 1035 - 27 \times 33 + 9\\ &= 34155 - 891 + 9\\ &= 33273.\end{aligned}$$

也就是說，

$$\tan^6\left(\tfrac{\pi}{9}\right) + \tan^6\left(\tfrac{2\pi}{9}\right) + \tan^6\left(\tfrac{4\pi}{9}\right) = 33273,$$

正如所求。等價地，$\sum u_i^3 = e_1^3 - 3e_1e_2 + 3e_3 = 35937 - 2673 + 9$。

驗算

- 以數值計算，$\tan^2(\pi/9) \approx 0.1325$、$\tan^2(2\pi/9) \approx 0.7041$ 及 $\tan^2(4\pi/9) \approx 32.163$。它們的和是 33.000，積是 3.000。
- 三個正切值的積是 $1.7320508\ldots = \sqrt{3}$。
- 上述三個值的立方和是 33273.0。它幾乎全部來自 $\tan^6(4\pi/9) \approx 33272.4$。
- 隨機取數千個 θ 的值，$\tan 9\theta$ 的公式與計算機的結果吻合至約十二位有效數字。
- 簡單情況：取 $t = 1$，即 $\theta = \pi/4$。公式給出 $\frac{1\cdot(-2)(-8)}{2\cdot 8} = 1 = \tan(9\pi/4)$。

常見錯誤

- 未先檢查 $3t^2 \neq 1$（或分母不為零），就斷定那個六次因式等於零。

- 把乘積取為 $-\sqrt{3}$。每個角都是銳角，所以每個正切值都是正數。
- 寫成 $\sum u_i^3 = e_1^3 - 3e_1e_2$ 而漏掉 $+3e_3$ 一項，結果得出 33264。

STEP 之外

三次多項式 $u^3 - 33u^2 + 27u - 3$ 沒有有理根，所以它在 Q 上不可約（irreducible）。它又有三個實根。伽羅華理論（Galois theory）表明，這類三次多項式的根不能單靠實數根式表示出來（即所謂不可約情形（casus irreducibilis））。正因如此，像 $\tan^2(\pi/9)$ 這樣的三角表達式才是寫出這些根的自然方式。至於根的冪和，一般由牛頓恆等式（Newton's identities）支配。想了解棣美弗定理的幾何一面，可讀 Needham 的 *Visual Complex Analysis*（以圖像方式講解複分析）。代數方面，可讀 Allenby 的 *Rings, Fields and Groups*（環、域與羣的入門書）。

詳細解答

STEP III 1990，第 2 題

純數學 • 三維向量

STEP III 1990，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/90-s3-q2

關鍵思路

四點全部位於同一條曲線 $t \mapsto (t^3, t^2, t)$ 上，這條曲線稱為三次撓線（twisted cubic）。兩條相交的直線必定位於同一平面上，所以第 (i) 部分歸結為證明 O、A、B、C 永不共面（coplanar），而這是一個行列式（determinant），它可以像范德蒙行列式（Vandermonde determinant）那樣因式分解。在第 (ii) 部分，條件 $ab = 1$ 令純量積（scalar product）和長度都化為單一個量 $a^2 + 1/a^2$ 的函數。

解答

首先，「相異」告訴我們甚麼？參數為 t 的點，其第三個坐標就是 t 本身，所以不同的點有不同的參數。因此 a、b、c 互不相同，而由於 O 是 $t = 0$ 所對應的點，a、b、c 均不為零。

第 (i) 部分

假設直線 OA 與 BC 確實相交。通過同一點的兩條直線位於同一平面上（若兩線剛好重合，任何包含它們的平面皆可）。於是 O、A、B、C 全部位於同一個通過原點的平面上，而位置向量（position vector）$\mathbf{a} = \overrightarrow{OA}$、$\mathbf{b} = \overrightarrow{OB}$、$\mathbf{c} = \overrightarrow{OC}$ 就會線性相關（linearly dependent）。這種情況恰好在以下條件成立時出現：

$$D = \begin{vmatrix} a^3 & a^2 & a \\ b^3 & b^2 & b \\ c^3 & c^2 & c \end{vmatrix} = 0.$$

從第一行抽出因子 a，第二行抽出 b，第三行抽出 c，

$$D = abc\begin{vmatrix} a^2 & a & 1 \\ b^2 & b & 1 \\ c^2 & c & 1 \end{vmatrix}.$$

把首兩行各自減去第三行，然後沿最後一列展開：

$$D = abc\begin{vmatrix} (a-c)(a+c) & a-c & 0 \\ (b-c)(b+c) & b-c & 0 \\ c^2 & c & 1 \end{vmatrix} = abc\,(a-c)(b-c)\begin{vmatrix} a+c & 1 \\ b+c & 1 \end{vmatrix}$$
$$= abc\,(a-b)(a-c)(b-c).$$

每個因子都不為零：$a, b, c \neq 0$，而且 a、b、c 互不相同。所以 $D \neq 0$，四點不共面，因此直線 OA 與 BC 不可能相交。（兩線亦不平行，所以它們是異面直線（skew lines）。）

想一想 為甚麼只需證明四點不共面便已足夠？

因為相交的直線必定張成一個平面。假如 OA 與 BC 相交於點 P，那麼通過 P 並包含這兩條直線的平面就會包含 O、A、B 和 C。我們已證明這樣的平面並不存在。

第 (ii) 部分

設 $b = 1/a$。那麼 $a \neq 0$，並且 $a \neq \pm 1$，否則 $b = a$，A 與 B 便會重合。純量積為

$$\mathbf{a}\cdot\mathbf{b} = (ab)^3 + (ab)^2 + ab = 3,$$

而長度的平方為

$$|\mathbf{a}|^2 = a^2(a^4 + a^2 + 1), \qquad |\mathbf{b}|^2 = \frac{1}{a^6} + \frac{1}{a^4} + \frac{1}{a^2} = \frac{a^4 + a^2 + 1}{a^6}.$$

兩式相乘，得 $|\mathbf{a}|^2|\mathbf{b}|^2 = (a^4 + a^2 + 1)^2/a^4$，所以 $|\mathbf{a}||\mathbf{b}| = (a^4 + a^2 + 1)/a^2$，這是正數。記 $\theta = \angle AOB$，則

$$\cos\theta = \frac{3a^2}{a^4 + a^2 + 1} = \frac{3}{u+1}, \qquad \text{其中} u = a^2 + \frac{1}{a^2}.$$

哪些值可以出現。根據算術幾何平均不等式（AM–GM）（或由於 $a^2 + a^{-2} - 2 = (a - a^{-1})^2$），$u \geq 2$，等號僅在 $a^2 = 1$ 時成立，而這個情況已被排除。所以 $u > 2$，因此

$$0 < \cos\theta = \frac{3}{u+1} < \frac{3}{3} = 1,$$

由此得 $0 < \theta < \frac{1}{2}\pi$。其他角度都不可能出現。

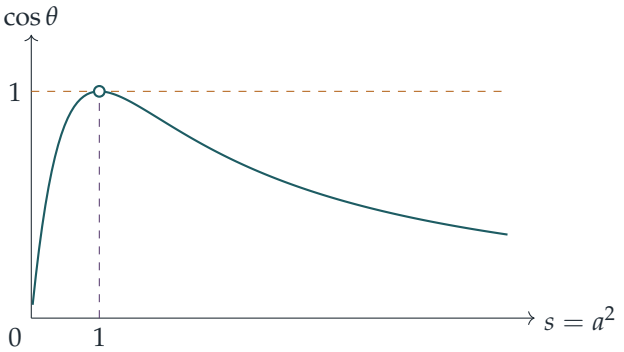
每個這樣的值都會出現。任取滿足 $0 < \theta < \frac{1}{2}\pi$ 的 θ，於是 $0 < \cos\theta < 1$。我們需要 $u = 3/\cos\theta - 1$，而這個值大於 2。方程 $a^2 + 1/a^2 = u$ 是以 a^2 為未知數的二次方程 $(a^2)^2 - u\,a^2 + 1 = 0$，它有一個根

$$a^2 = \frac{u + \sqrt{u^2 - 4}}{2} > 1,$$

由於 $u > 2$，這個根是實數。取 a 為其正平方根，並取 $b = 1/a$。那麼 $a > 1 > b > 0$，所以 $ab = 1$，$A \neq B$，兩點都不是 O，並且 $\cos\angle AOB = 3/(u+1) = \cos\theta$。因此 $\angle AOB = \theta$，因為兩個角都在 $[0, \pi]$ 之內，而餘弦函數在這個區間上是一對一（one-to-one）的。

所以 $\angle AOB$ 可取 $0 < \angle AOB < \frac{1}{2}\pi$ 內的每一個值，而不能取其他值，正如題目所求。

下圖顯示 $\cos\theta$ 對 $s = a^2$ 的圖像。曲線在被排除的值 $s = 1$ 處升至 1，並在兩端趨向 0，所以它的值恰好填滿 $(0, 1)$。



驗算

- 試取 $a = 2$、$b = \frac{1}{2}$：$A = (8, 4, 2)$ 和 $B = (\frac{1}{8}, \frac{1}{4}, \frac{1}{2})$ 給出 $\mathbf{a} \cdot \mathbf{b} = 1 + 1 + 1 = 3$，而公式給出 $\cos\theta = 12/21 = 4/7$，所以 $\theta \approx 55.2°$，確實在區間之內。
- 當 $a \to 1$ 時，點 A 與 B 合而為一，$\theta \to 0$。當 $a \to \infty$ 時，A 幾乎沿著 x 軸遠去，而 B 幾乎沿著 z 軸趨近 O，所以 $\theta \to \frac{1}{2}\pi$。這正是區間的兩端。
- D 的因式分解已用數千組隨機有理數 a、b、c 作精確驗證，角度公式及其值域亦已用數值方法驗證。

常見錯誤

- 忘記說明為何 a、b、c 互不相同且不為零。正是這些事實令 $D \neq 0$。

- 沒有排除 $a = \pm 1$，單憑 $\cos\theta < 1$ 就下結論。嚴格不等式 $u > 2$ 才是排除 $\theta = 0$ 的關鍵。
- 只證明「沒有其他值」。題目還要求你證明區間內每一個值都能取到，這需要明確給出一個 a。

STEP 之外

曲線 (t^3, t^2, t) 就是三次撓線，是代數幾何（algebraic geometry）中的標準例子。第 (i) 部分所說的，就是這條曲線上任意四點都不共面，而你計算的行列式是一個范德蒙行列式，它在多項式插值（polynomial interpolation）和線性獨立（linear independence）的論證中會再次出現。Strang 的 *Introduction to Linear Algebra*（線性代數入門）對行列式和線性獨立有清晰的講解。

詳細解答

STEP III 1990，第 3 題

純數學 • 群公理與共軛

> **STEP III 1990，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/90-s3-q3

關鍵思路

三個部分都只需用到羣公理（group axioms），而每一部分中有用的一步，都是把一個元素改寫成以另一個元素表示。階（order）為 2 的元素是它自己的逆元（inverse），這使 (i) 變成一行計算。在 (ii) 和 (iii) 中，所用的工具是共軛（*conjugation*）：$x \mapsto c^{-1}xc$ 保持乘積，所以它把冪映到冪。

全文中我們把 $x * y$ 寫作 xy，並以 e 表示單位元（identity）。

解答

第 (i) 部分

元素 x 的階為 2，當且僅當 $x \neq e$ 且 $x^2 = e$；第二個條件即是說 $x^{-1} = x$。所以

$$a^{-1} = a, \qquad b^{-1} = b, \qquad (ab)^{-1} = ab.$$

對任何羣元素，都有 $(ab)^{-1} = b^{-1}a^{-1}$，因為 $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = e$。因此

$$ab = (ab)^{-1} = b^{-1}a^{-1} = ba,$$

所以 a 與 b 可交換（commute）。

第 (ii) 部分

首先留意 dc 是 cd 的一個共軛元：

$$c^{-1}(cd)\,c = (c^{-1}c)\,dc = dc.$$

想一想 為甚麼寫成 $dc = c^{-1}(cd)c$ 對處理階有幫助？

因為共軛可以穿過冪。對任何 x 和任何正整數 n，

$$(c^{-1}xc)^n = (c^{-1}xc)(c^{-1}xc)\cdots(c^{-1}xc) = c^{-1}x^nc,$$

因為中間每一對相鄰的 $c\,c^{-1}$ 都互相抵消。取 $x = cd$，便得 $(dc)^n = c^{-1}(cd)^nc$。於是

$$(dc)^n = e \iff c^{-1}(cd)^nc = e \iff (cd)^n = cec^{-1} = e,$$

其中第二步是左乘 c 並右乘 c^{-1}（反方向則乘以 c^{-1} 和 c）。所以滿足 $(cd)^n = e$ 的正整數 n，正好就是滿足 $(dc)^n = e$ 的那些正整數。階就是這樣的 n 中最小的一個，所以 cd 與 dc 的階相同。若不存在這樣的 n，兩者的階都是無限，所以命題在這個情況下同樣成立。

第 (iii) 部分

已知 $c^{-1}bc = b^r$。對正整數 s，用 (ii) 中逐項相消（telescoping）的論證，可得

$$c^{-1}b^sc = (c^{-1}bc)^s = (b^r)^s = b^{sr}.$$

當 $s = 0$ 時同樣成立（兩邊都是 e）。當 s 為負數時，設 $s = -t$，其中 $t > 0$；則 $c^{-1}b^sc = \left(c^{-1}b^tc\right)^{-1}$，因為 $(c^{-1}b^tc)(c^{-1}b^{-t}c) = e$。所以

$$c^{-1}b^sc = \left(b^{tr}\right)^{-1} = b^{-tr} = b^{sr}.$$

因此對每個整數 s，都有 $c^{-1}b^sc = b^{sr}$。

至於第二個結果，我們對 $n \geq 1$ 使用歸納法（induction），並同時處理所有整數 s。$n = 1$ 的情況就是剛才所證明的結果。假設 $c^{-n}b^sc^n = b^{sr^n}$ 對所有整數 s 成立。則

$$\begin{aligned} c^{-(n+1)}\,b^s\,c^{n+1} &= c^{-1}\left(c^{-n}b^sc^n\right)c \\ &= c^{-1}\,b^{sr^n}\,c \\ &= b^{(sr^n)r} = b^{sr^{n+1}}, \end{aligned}$$

其中最後一行用了第一個結果，並以整數 sr^n 代替 s。歸納法至此完成，所以 $c^{-n}b^sc^n = b^{sr^n}$ 對所有正整數 n 成立。

想一想 為甚麼我們要特地為負數 s 證明第一個結果？

若 r 是負數（例如在二面體羣（dihedral group）中 $c^{-1}bc = b^{-1}$），則即使 s 是正數，sr^n 也可能是負數，而歸納步驟正是以 sr^n 代替 s 來應用第一個結果。為所有整數 s 證明第一個結果，歸納法便不論 r 的正負都行得通。

驗算

用小的羣來測試。在正方形的對稱羣中，沿兩條對角線的反射及它們的乘積（旋轉半周）的階都是 2，而這兩個反射確實可交換，與 (i) 的預測一致。在 S_3 中，取 $c = (12)$ 和 $d = (13)$：則 cd 和 dc 就是那兩個 3-循環（3-cycle），階都是 3。至於 (iii)，在二面體羣中，反射 c 與旋轉 b 滿足 $c^{-1}bc = b^{-1}$，所以 $r = -1$，而公式預測 $c^{-n}b^s c^n = b^{(-1)^n s}$，這是對的：以同一個反射共軛兩次，等於甚麼也沒有做。我們亦用電腦驗證了全部三個命題，窮舉了 S_4 中所有合適的元素（就 (i) 而言還包括 S_5），以及模 13 之下由映射 $x \mapsto ux + v$ 所組成的羣。

常見錯誤

- 寫成 $(ab)^{-1} = a^{-1}b^{-1}$。次序是會倒轉的；正是這個倒轉令 (i) 得以成立。
- 在 (ii) 中，只證明 $(cd)^n = e \Rightarrow (dc)^n = e$。兩個方向都要證明（或利用對稱性：把 c 和 d 的角色互換），而階為無限的情況亦應提及。
- 在 (iii) 中，以三點省略號代替歸納法，為第二個結果寫出一個所謂的「證明」。

STEP 之外

映射 $x \mapsto c^{-1}xc$ 是這個羣的一個*自同構*（*automorphism*）：它是由 G 到自身的雙射（bijection），並且保持運算，所以它保持元素的階，並把冪映到冪。第 (iii) 部分是*半直積*（*semidirect products*）中的關鍵關係：由 b 和 c 生成並滿足 $c^{-1}bc = b^r$ 的羣，包括二面體羣（$r = -1$）以及模一個質數之下由映射 $x \mapsto ux + v$ 所組成的羣。Allenby 的 *Rings, Fields and Groups*（環、域與羣）對共軛及這些例子有仔細的闡述。

詳細解答

STEP III 1990，第 4 題

純數學 • 三角級數

STEP III 1990，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/90-s3-q4

關鍵思路

兩個級數（series）都是餘弦之和，而且角度按等差遞增，所以有兩件順手的工具。對於單純的和，可乘以 $2\sin\beta$ 並使用積化和差公式（product-to-sum formula）：每一項都拆成兩個正弦之差，整個和便會裂項相消（telescope）。對於二項式的和，以及最後的恆等式，可把 $\cos\phi$ 看成 $\mathrm{e}^{\mathrm{i}\phi}$ 的實部（real part）。這樣，二項式定理（binomial theorem）或等比級數（geometric series）公式便會代勞，最後取實部或虛部（imaginary part）就得出答案。

解答

第一個級數

把這個和記作 S。根據積化和差公式 $2\sin A\cos B = \sin(B+A) - \sin(B-A)$，

$$2\sin\beta\cos(\alpha+2r\beta) = \sin\big(\alpha+(2r+1)\beta\big) - \sin\big(\alpha+(2r-1)\beta\big).$$

把 $r = 0, 1, \ldots, n$ 的各式相加，和便裂項相消：對應 r 的第一個正弦，與對應 $r+1$ 的第二個正弦互相抵消。因此只剩下兩項：

$$\begin{aligned} 2\sin\beta\, S &= \sin\big(\alpha+(2n+1)\beta\big) - \sin(\alpha-\beta) \\ &= 2\cos(\alpha+n\beta)\sin\big((n+1)\beta\big), \end{aligned}$$

這裏用了 $\sin P - \sin Q = 2\cos\frac{P+Q}{2}\sin\frac{P-Q}{2}$，其中 $\frac{P+Q}{2} = \alpha+n\beta$ 及 $\frac{P-Q}{2} = (n+1)\beta$。由於 $\sin\beta \neq 0$，我們可以除以它，得

$$\boxed{\sum_{r=0}^{n}\cos(\alpha+2r\beta) = \frac{\sin(n+1)\beta\,\cos(\alpha+n\beta)}{\sin\beta}}$$

留意 $\alpha + n\beta$ 正是首尾兩個角的平均值，這可用來粗略核對答案的形式是否合理。

二項式級數

把這個和記作 T。由於 $\cos\phi = \operatorname{Re} \mathrm{e}^{\mathrm{i}\phi}$，而二項式係數（binomial coefficient）都是實數，

$$T = \operatorname{Re} \sum_{r=0}^{n} \binom{n}{r} \mathrm{e}^{\mathrm{i}\alpha} \left(\mathrm{e}^{2\mathrm{i}\beta}\right)^r = \operatorname{Re}\left[\mathrm{e}^{\mathrm{i}\alpha}\left(1 + \mathrm{e}^{2\mathrm{i}\beta}\right)^n\right]$$

這是根據二項式定理得出的。現在抽出「半角」因子：

$$1 + \mathrm{e}^{2\mathrm{i}\beta} = \mathrm{e}^{\mathrm{i}\beta}\left(\mathrm{e}^{-\mathrm{i}\beta} + \mathrm{e}^{\mathrm{i}\beta}\right) = 2\cos\beta\, \mathrm{e}^{\mathrm{i}\beta}.$$

因此 $\mathrm{e}^{\mathrm{i}\alpha}(1 + \mathrm{e}^{2\mathrm{i}\beta})^n = 2^n \cos^n\beta\, \mathrm{e}^{\mathrm{i}(\alpha+n\beta)}$，而由於 $2^n\cos^n\beta$ 是實數，

$$\boxed{\sum_{r=0}^{n} \binom{n}{r} \cos(\alpha + 2r\beta) = 2^n \cos^n\beta\, \cos(\alpha + n\beta)}$$

（這一個其實並不需要 $\sin\beta \neq 0$。）

最後的恆等式

式中出現正割，表示我們需要 $\cos\theta \neq 0$，這是題目隱含的條件。第 r 項是 $\cos r\theta\, \sec^r\theta = \operatorname{Re}\left(\mathrm{e}^{\mathrm{i}\theta}\sec\theta\right)^r$，因為 $\sec^r\theta$ 是實數。於是設

$$w = \mathrm{e}^{\mathrm{i}\theta}\sec\theta = \frac{\cos\theta + \mathrm{i}\sin\theta}{\cos\theta} = 1 + \mathrm{i}\tan\theta.$$

左方就是 $\operatorname{Re}(1 + w + \cdots + w^n)$。

想一想 為甚麼在這裏可以放心使用等比級數公式？w 又為甚麼有幫助？

這條公式需要 $w \neq 1$，而 $w - 1 = \mathrm{i}\tan\theta \neq 0$ 正正是因為 $\sin\theta \neq 0$。題目的條件就是用在這裏。引入 w 的好處在於 $w - 1$ 是純虛數（purely imaginary），所以除以它很容易，而 w^{n+1} 則可由指數形式 $w = \sec\theta\, \mathrm{e}^{\mathrm{i}\theta}$ 輕易算出。設 $t = \tan\theta$，

$$1 + w + \cdots + w^n = \frac{w^{n+1} - 1}{w - 1} = \frac{\sec^{n+1}\theta\, \mathrm{e}^{\mathrm{i}(n+1)\theta} - 1}{\mathrm{i}t}.$$

對任意複數 z 及實數 $t \neq 0$，我們有 $\operatorname{Re}\dfrac{z}{\mathrm{i}t} = \operatorname{Re}\dfrac{-\mathrm{i}z}{t} = \dfrac{\operatorname{Im} z}{t}$。分子中的 -1 虛部為零，所以

$$\begin{aligned}\sum_{r=0}^{n} \cos r\theta\, \sec^r\theta &= \frac{\sec^{n+1}\theta\, \sin(n+1)\theta}{\tan\theta} = \sec^{n+1}\theta\, \sin(n+1)\theta \cdot \frac{\cos\theta}{\sin\theta} \\ &= \frac{\sin(n+1)\theta\, \sec^n\theta}{\sin\theta},\end{aligned}$$

這就是所求的結果。

與第一部分的聯繫。把全式乘以 $\cos^n\theta$，恆等式便變成

$$\sum_{r=0}^{n}\cos r\theta\cos^{n-r}\theta=\frac{\sin(n+1)\theta}{\sin\theta}.$$

你也可以對 n 用歸納法（induction）證明這個形式：若 C_n 為左方，則 $C_n=\cos\theta\,C_{n-1}+\cos n\theta$，而 $\sin n\theta\cos\theta+\cos n\theta\sin\theta=\sin(n+1)\theta$ 便完成歸納步驟。

驗算

- $n=0$：三條公式都化為各自唯一的首項（$\cos\alpha$、$\cos\alpha$ 和 1）。
- 在第一個和中，$\beta\to 0$ 時 $\frac{\sin(n+1)\beta}{\sin\beta}\to n+1$，所以 $S\to(n+1)\cos\alpha$，這與 $n+1$ 個相等的項之和吻合。在第二個和中，$\beta=0$ 給出 $2^n\cos\alpha$，因為 $\sum\binom{n}{r}=2^n$。
- $\theta=\pi/4$，$n=2$：左方是 $1+\frac{1}{\sqrt{2}}\cdot\sqrt{2}+0=2$，右方是 $\sin\frac{3\pi}{4}\cdot 2/\sin\frac{\pi}{4}=2$。
- 數值上，隨機選取 20 000 組 α、β、θ 及 $n\le 15$，三條公式與直接求和的結果都吻合，相對誤差在 10^{-9} 以內。

常見錯誤

- 忘記第一個級數共有 $n+1$ 項（由 $r=0$ 至 n），結果得出 $\sin n\beta$ 而不是 $\sin(n+1)\beta$。
- 把 $1+\mathrm{e}^{2\mathrm{i}\beta}$ 寫成 $2\cos\beta$，漏掉因子 $\mathrm{e}^{\mathrm{i}\beta}$，令答案中的角度出現偏移。
- 在最後一部分，把分子和分母分別取實部。你必須先把分母化為實數。

STEP 之外

把 cos 寫成指數的實部，三角函數的和便化為等比或二項式的和。這正是傅立葉分析（Fourier analysis）的起點：第一個和本質上就是狄利克雷核（Dirichlet kernel）$\sin\big((n+\frac{1}{2})x\big)/\sin\frac{x}{2}$，它決定了傅立葉級數（Fourier series）如何收斂。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）以優美的幾何觀點，解釋了為甚麼乘以 $\mathrm{e}^{\mathrm{i}\theta}$ 就是旋轉，以及為甚麼這類和最終能收攏成簡潔的公式。

詳細解答

STEP III 1990，第 5 題

純數學 • 萊布尼茲定理

> **STEP III 1990，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/90-s3-q5

關鍵思路

帕斯卡法則（Pascal's rule）只不過是含階乘（factorial）的分數運算，把兩項通分即可。萊布尼茲公式（Leibniz's formula）給出乘積的 n 階導數（derivative），其證明是對 n 用數學歸納法（induction）：把 n 階的公式再求導一次，每個混合項都會出現兩次，而帕斯卡法則正好把每一對合併起來。至於最後一部分，直接把 $y = \sin^{-1} x$ 求導 n 次是行不通的。我們改為找出 y 所滿足的一條簡單二階方程，再對它應用萊布尼茲公式。

（試卷上印的是 $y = \sin^{-1} y$。這顯然是 $y = \sin^{-1} x$ 的誤印，我們採用後者。）

解答

帕斯卡法則

回想 $\binom{n}{r} = \dfrac{n!}{r!\,(n-r)!}$，其中 $0! = 1$。當 $1 \le r \le n$ 時，$r-1$ 和 $n-r$ 均為非負數，所以下面每個階乘都有意義。由於 $r! = r\,(r-1)!$ 及 $(n-r+1)! = (n-r+1)\,(n-r)!$，公分母是 $r!\,(n-r+1)!$：

$$\begin{aligned}\binom{n}{r} + \binom{n}{r-1} &= \frac{n!}{r!\,(n-r)!} + \frac{n!}{(r-1)!\,(n-r+1)!} \\ &= \frac{n!\,(n-r+1)}{r!\,(n-r+1)!} + \frac{n!\,r}{r!\,(n-r+1)!} \\ &= \frac{n!\,(n+1)}{r!\,(n+1-r)!} = \frac{(n+1)!}{r!\,(n+1-r)!} = \binom{n+1}{r}.\end{aligned}$$

萊布尼茲公式

我們假設 u 和 v 可按需要求導任意多次。用求和記號表示，要證明的是

$$(uv)^{(n)} = \sum_{r=0}^{n} \binom{n}{r} u^{(n-r)} v^{(r)},$$

其中 $z^{(0)}$ 表示 z 本身。我們對 $n \geq 1$ 用數學歸納法證明。

基礎步驟（base case） $n = 1$。右邊是 $u^{(1)}v + uv^{(1)}$，由乘法法則（product rule）可知它就是 $(uv)'$。

歸納步驟（inductive step）。假設公式對某個 $n \geq 1$ 成立。把它求導，對每一項使用乘法法則：

$$(uv)^{(n+1)} = \sum_{r=0}^{n} \binom{n}{r} u^{(n+1-r)} v^{(r)} + \sum_{r=0}^{n} \binom{n}{r} u^{(n-r)} v^{(r+1)}.$$

在第二個和式中設 $s = r+1$，它便成為 $\sum_{s=1}^{n+1} \binom{n}{s-1} u^{(n+1-s)} v^{(s)}$。現在兩個和式都由形如 $u^{(n+1-r)}v^{(r)}$ 的項組成。把第一個和式的 $r = 0$ 項和第二個和式的 $s = n+1$ 項分拆出來：

$$\begin{aligned}(uv)^{(n+1)} &= u^{(n+1)}v + \sum_{r=1}^{n} \left[\binom{n}{r} + \binom{n}{r-1} \right] u^{(n+1-r)} v^{(r)} + uv^{(n+1)} \\ &= u^{(n+1)}v + \sum_{r=1}^{n} \binom{n+1}{r} u^{(n+1-r)} v^{(r)} + uv^{(n+1)},\end{aligned}$$

這裏用了帕斯卡法則，因為 $1 \leq r \leq n$，所以法則適用。由於 $\binom{n+1}{0} = \binom{n+1}{n+1} = 1$，上式即

$$(uv)^{(n+1)} = \sum_{r=0}^{n+1} \binom{n+1}{r} u^{(n+1-r)} v^{(r)},$$

這正是 $n+1$ 時的公式。根據數學歸納法，公式對所有 $n \geq 1$ 成立。

$\sin^{-1} x$ 的導數

設 $y = \sin^{-1} x$，其中 $-1 < x < 1$。則

$$y^{(1)} = \frac{1}{\sqrt{1-x^2}}, \qquad \text{所以} \qquad (1-x^2)\left(y^{(1)}\right)^2 = 1.$$

把此式對 x 求導，

$$(1-x^2) \cdot 2y^{(1)}y^{(2)} - 2x\left(y^{(1)}\right)^2 = 0.$$

由於在 $(-1,1)$ 上 $y^{(1)} > 0$，我們可以除以 $2y^{(1)}$：

$$(1-x^2)\, y^{(2)} - x\, y^{(1)} = 0. \qquad (*)$$

這就是 $n = 0$ 時所要求的結果。

想一想 為甚麼先把 $y^{(1)}$ 平方再求導，而不直接把 $(1-x^2)^{-1/2}$ 求導？

直接把 $(1-x^2)^{-1/2}$ 求導，得 $y^{(2)} = x(1-x^2)^{-3/2}$，再乘以 $1-x^2$ 同樣得到 $(*)$。兩種做法都可行。重點是要得出一條方程，其係數是 x 的多項式（polynomial），而且對 y 的各階導數而言是線性（linear）的，因為這正是萊布尼茲公式處理得乾淨俐落的情況：多項式係數求導幾次便變成零。

現在把 $(*)$ 求導 n 次，其中 $n \geq 1$，對每個乘積應用萊布尼茲公式。對第一個乘積，取 $u = y^{(2)}$ 和 $v = 1-x^2$。則 $v^{(1)} = -2x$，$v^{(2)} = -2$，而所有更高階的導數均為零，所以只剩下三項：

$$\left((1-x^2)\,y^{(2)}\right)^{(n)} = (1-x^2)\,y^{(n+2)} + \binom{n}{1}(-2x)\,y^{(n+1)} + \binom{n}{2}(-2)\,y^{(n)}.$$

（當 $n=1$ 時最後一項不存在，這與 $\binom{1}{2}=0$ 吻合。）對第二個乘積，取 $u = y^{(1)}$ 和 $v = x$，所以只剩下兩項：

$$\left(x\,y^{(1)}\right)^{(n)} = x\,y^{(n+1)} + n\,y^{(n)}.$$

兩式相減，$(*)$ 的 n 階導數是

$$(1-x^2)\,y^{(n+2)} - 2nx\,y^{(n+1)} - n(n-1)\,y^{(n)} - x\,y^{(n+1)} - n\,y^{(n)} = 0.$$

合併同類項，$-2nx - x = -(2n+1)x$ 及 $-n(n-1) - n = -n^2$，所以

$$(1-x^2)\,y^{(n+2)} - (2n+1)x\,y^{(n+1)} - n^2 y^{(n)} = 0,$$

這正是所要求的結果。它對所有 $n \geq 0$ 及 $(-1,1)$ 內的所有 x 成立。

驗算

- 在萊布尼茲公式中取 $n=1$：得到乘法法則。取 $n=2$：$(uv)'' = u''v + 2u'v' + uv''$，你可以把 $u'v + uv'$ 求導來核實。
- 在 $x = 0$ 處，最後的結果給出 $y^{(n+2)}(0) = n^2 y^{(n)}(0)$。配合 $y^{(1)}(0) = 1$，可得 $y^{(3)}(0) = 1$ 和 $y^{(5)}(0) = 9$，所以麥克勞林級數（Maclaurin series）的開首是 $x + \frac{x^3}{6} + \frac{9x^5}{120} = x + \frac{x^3}{6} + \frac{3x^5}{40}$，這正是 $\sin^{-1} x$ 的已知級數。
- 我們作了數值驗算：對所有 $n < 60$ 檢驗帕斯卡法則；用 300 對隨機多項式精確檢驗萊布尼茲公式；並利用 $(1-x^2)^{-1/2}$ 的泰勒係數（Taylor coefficients），在 $(-1,1)$ 內的六個點上對 n 直至 27 檢驗最後的恆等式。相對殘差小於 10^{-15}。

常見錯誤

- 在歸納步驟中把第二個和式重新編排指標後，弄錯二項式係數（binomial coefficient）。應先把 $r = 0$ 和 $r = n + 1$ 這兩個首尾項分開寫出，然後才應用帕斯卡法則。
- 試圖直接把 $\sin^{-1} x$ 求導 n 次，而不先找出方程 $(*)$。
- 把 $1 - x^2$ 求導時，漏掉因子 $\binom{n}{2}$ 或負號。

STEP 之外

先找出一條係數為多項式的線性微分方程（differential equation），再把它求導 n 次，這個技巧是求泰勒係數遞推關係（recurrence）的標準方法。它直接通往常微分方程（ODE）的冪級數解法（power-series solutions），勒讓德（Legendre）函數和貝塞爾（Bessel）函數正是由此而來。萊布尼茲公式也是二項式定理（binomial theorem）在微積分中的版本：若 $u = \mathrm{e}^{ax}$ 而 $v = \mathrm{e}^{bx}$，它便化為 $(a+b)^n = \sum \binom{n}{r} a^{n-r} b^r$。想進一步了解級數解法，可參閱 Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程）。

詳細解答

STEP III 1990，第 6 題

純數學 • 線性變換與橢圓

> **STEP III 1990，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/90-s3-q6

本題在 *Beyond STEP* 第 7.2 節只伸縮，不轉向中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch07/eigenvectors.html

關鍵思路

那兩個特殊的向量是矩陣的特徵向量（eigenvector），它們道出了整個幾何圖像：T 沿一條直線按某個倍數拉伸，又沿與之垂直的另一條直線按另一個倍數拉伸，於是單位圓變成一個橢圓（ellipse），其兩軸就落在這兩條直線上。代數方面，要求 E 的方程，最簡潔的做法是先求 T 的逆，把 x 和 y 用 X 和 Y 表示，再代入 $x^2 + y^2 = 1$。面積按行列式（determinant）的倍數縮放；又因為 T 把直線映射為直線，所以 E 的切線（tangent）就是 C 的切線的像。

解答

設 $M = \frac{2}{5}\begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$，則 T 就是左乘 M。

兩個特殊方向

計算可得

$$M\begin{pmatrix}1\\2\end{pmatrix} = \frac{2}{5}\begin{pmatrix}9-4\\-2+12\end{pmatrix} = \frac{2}{5}\begin{pmatrix}5\\10\end{pmatrix} = 2\begin{pmatrix}1\\2\end{pmatrix},$$
$$M\begin{pmatrix}2\\-1\end{pmatrix} = \frac{2}{5}\begin{pmatrix}18+2\\-4-6\end{pmatrix} = \frac{2}{5}\begin{pmatrix}20\\-10\end{pmatrix} = 4\begin{pmatrix}2\\-1\end{pmatrix}.$$

所以 $\begin{pmatrix}1\\2\end{pmatrix}$ 方向不變，並乘以 2；而 $\begin{pmatrix}2\\-1\end{pmatrix}$ 方向不變，並乘以 4，正如題目所求。留意這兩個方向互相垂直，因為 $1\cdot 2 + 2\cdot(-1) = 0$。

E 的方程

M 的行列式為

$$\det M = \left(\tfrac{2}{5}\right)^2 (9\cdot 6 - (-2)(-2)) = \tfrac{4}{25}\cdot 50 = 8,$$

其值非零，故 T 可逆（invertible），且

$$M^{-1} = \frac{1}{8}\cdot\frac{2}{5}\begin{pmatrix}6&2\\2&9\end{pmatrix} = \frac{1}{20}\begin{pmatrix}6&2\\2&9\end{pmatrix}.$$

因此，點 (X,Y) 是 (x,y) 的像，當且僅當

$$x = \frac{3X+Y}{10}, \qquad y = \frac{2X+9Y}{20}.$$

點 (X,Y) 在 E 上，當且僅當它的原像在 C 上，即當且僅當 $x^2+y^2=1$。代入得

$$\begin{aligned}\frac{(3X+Y)^2}{100} + \frac{(2X+9Y)^2}{400} = 1 &\iff 4(9X^2+6XY+Y^2) + (4X^2+36XY+81Y^2) = 400\\ &\iff 40X^2+60XY+85Y^2 = 400\\ &\iff 8X^2+12XY+17Y^2 = 80,\end{aligned}$$

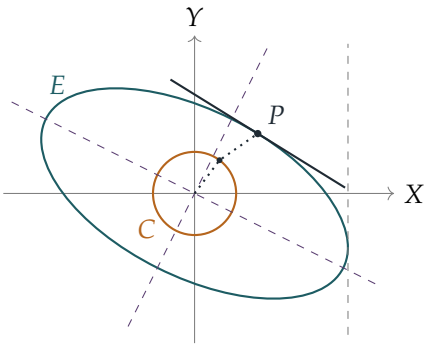
這就是所求的 E 的方程。

面積

線性映射（linear map）會把每一個面積乘以 $|\det M|$。C 所圍成的圓盤面積為 π，所以 E 所圍成的區域面積為

$$8\pi.$$

這與幾何圖像吻合：E 是一個橢圓，其半軸長為 2（沿 $\begin{pmatrix}1\\2\end{pmatrix}$ 方向）和 4（沿 $\begin{pmatrix}2\\-1\end{pmatrix}$ 方向），而 $\pi\cdot 2\cdot 4 = 8\pi$。



圖中顯示 C、橢圓 E 及其兩軸（虛線）、與 E 相切的鉛垂線 $X = 2\sqrt{17/5}$，以及在 P 的切線，該點是 $\frac{1}{5}(3,4)$ 的像。

X 的最大值

把 E 的方程看成 Y 的二次方程：

$$17Y^2 + 12XY + (8X^2 - 80) = 0.$$

某個 X 值會在 E 上出現，當且僅當此方程有實根 Y，即當且僅當其判別式(discriminant)非負：

$$\begin{aligned} 144X^2 - 4 \cdot 17(8X^2 - 80) \geq 0 &\iff 144X^2 - 544X^2 + 5440 \geq 0 \\ &\iff 400X^2 \leq 5440 \\ &\iff X^2 \leq \tfrac{68}{5}. \end{aligned}$$

所以 E 上每一點都滿足 $X \leq \sqrt{68/5} = 2\sqrt{17/5}$，而且等號可以取到（此時判別式為零，得出唯一的一點，其 $Y = -\frac{6X}{17}$）。因此 X 在 E 上的最大值為 $2\sqrt{17/5}$，正如題目所求。

想一想 有沒有辦法不用判別式，直接從圓看出這個結果？

有。在 E 上有 $X = \frac{2}{5}(9x - 2y)$，其中 $x^2 + y^2 = 1$。設 $(x, y) = (\cos\theta, \sin\theta)$，則表達式 $9\cos\theta - 2\sin\theta$ 的最大值為 $\sqrt{81 + 4} = \sqrt{85}$（即 $R\cos(\theta + \alpha)$ 的形式）。所以最大的 X 是 $\frac{2}{5}\sqrt{85} = 2\sqrt{85/25} = 2\sqrt{17/5}$，與上面的結果一致。

切線

C 上的點 $\frac{1}{5}(3,4)$ 被映射到

$$P = \frac{2}{5}\begin{pmatrix} 9 \cdot \frac{3}{5} - 2 \cdot \frac{4}{5} \\ -2 \cdot \frac{3}{5} + 6 \cdot \frac{4}{5} \end{pmatrix} = \frac{2}{5}\begin{pmatrix} \frac{19}{5} \\ \frac{18}{5} \end{pmatrix} = \begin{pmatrix} \frac{38}{25} \\ \frac{36}{25} \end{pmatrix}.$$

C 在 $\frac{1}{5}(3,4)$ 的切線垂直於該處的半徑，所以它的方程是

$$3x + 4y = 5.$$

由於 T 是線性而且可逆的，它把直線映射為直線，並且把一條與 C 恰好交於一點的直線，映射為一條與 E 恰好交於一點的直線。所以 E 在 P 的切線就是這條直線的像。代入 $x = \frac{3X+Y}{10}$ 和 $y = \frac{2X+9Y}{20}$：

$$\begin{aligned}\frac{3(3X+Y)}{10} + \frac{4(2X+9Y)}{20} = 5 &\iff 6(3X+Y) + 4(2X+9Y) = 100\\ &\iff 26X + 42Y = 100,\end{aligned}$$

所以 E 在 P 的切線是

$$13X + 21Y = 50.$$

驗算一下：$13 \cdot \frac{38}{25} + 21 \cdot \frac{36}{25} = \frac{494+756}{25} = 50$，所以 P 在這條直線上。另一個做法是對 $8X^2 + 12XY + 17Y^2 = 80$ 作隱函數微分（implicit differentiation），得出在 P 的法線（normal）方向為 $(16X + 12Y,\ 12X + 34Y) = \frac{1}{25}(1040, 1680)$，它與 $(13,21)$ 平行，於是得出同一條直線。

驗算

- $(1,0)$ 的像是 $\frac{2}{5}(9,-2)$，而 $\frac{4}{25}\left(8 \cdot 81 - 12 \cdot 18 + 17 \cdot 4\right) = \frac{4}{25} \cdot 500 = 80$，所以它在 E 上。
- 兩軸的端點 $\pm 2 \cdot \frac{1}{\sqrt{5}}(1,2)$ 和 $\pm 4 \cdot \frac{1}{\sqrt{5}}(2,-1)$ 都滿足 E 的方程；例如 $\frac{4}{5}(8 + 24 + 68) = 80$。
- 數值上，在 C 上取樣 200 000 個點，得出最大值 $X \approx 3.687818$，與 $2\sqrt{17/5} \approx 3.687818$ 相符；而用鞋帶公式（shoelace formula）計算像多邊形的面積，得 $25.13274 \approx 8\pi$。

常見錯誤

- 代入 $x^2 + y^2 = 1$ 時用了 M 而不是 M^{-1}：要求像曲線的方程，需要把舊坐標用新坐標表示。
- 計算行列式時忘記了因子 $\frac{2}{5}$：它貢獻的是 $\left(\frac{2}{5}\right)^2$，而不是 $\frac{2}{5}$。
- 求出了像點 P，卻用 C 在 $\frac{1}{5}(3,4)$ 的斜率來寫 E 的切線：線性映射保持相切關係，但不保持斜率。

STEP 之外

這裏的矩陣是對稱（symmetric）的，這正是它的特徵向量互相垂直的原因：這是譜定理（spectral theorem）的二維情形，該定理指出每個實對稱矩陣都可以用一個旋轉來對角化（diagonalise）。它解釋了為甚麼任何二次型（quadratic form）$ax^2 + 2hxy + by^2$ 所描述的圓錐曲線（conic）都有互相垂直的主軸，而該二次型的矩陣的特徵值（eigenvalue）可以讓你立刻知道它的形狀。Strang 的 *Introduction to Linear Algebra*（線性代數入門）以大量幾何圖像闡述特徵值、對稱矩陣和二次型。

詳細解答

STEP III 1990，第 7 題

純數學 • 隱函數曲線描繪

STEP III 1990，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/90-s3-q7

關鍵思路

隱函數微分法（implicit differentiation）做兩次，便得到每一點的 $\mathrm{d}y/\mathrm{d}x$ 和 $\mathrm{d}^2y/\mathrm{d}x^2$，而這些正是泰勒級數（Taylor series）的係數。至於不等式，竅門在於留意到這條方程只是 y 的二次方程：點 (x,y) 要存在，這個二次方程必須有實根，所以它的判別式（discriminant）非負。把二次方程直接解出來，便可看清整條曲線：兩條分支（branch），在判別式等於零之處相接。

解答

以下一律記 $F(x,y)=xy^2+x^3+a^2y-a^3$。代入可知在 P、Q 和 R 均有 $F=0$；例如在 R 得 $a\cdot a^2+a^3-a^3-a^3=0$。

在 P、Q 和 R 的泰勒級數

把 y 視為 x 的函數，將 $F(x,y)=0$ 對 x 求導，

$$y^2+2xyy'+3x^2+a^2y'=0,\qquad \text{所以}\qquad y'=-\frac{y^2+3x^2}{2xy+a^2}.$$

把第一條方程再求導一次（$2xyy'$ 一項給出 $2yy'+2xy'^2+2xyy''$）：

$$4yy'+2xy'^2+6x+(2xy+a^2)\,y''=0.$$

在每一點，分母 $2xy+a^2$ 都不等於零，所以這些方程確定了 y' 和 y''，繼而有 $y=y_0+y'h+\frac{1}{2}y''h^2+\cdots$。

在 $P(0,a)$： $y' = -a^2/a^2 = -1$。於是 $4a(-1) + a^2 y'' = 0$，所以 $y'' = 4/a$，而

$$y = a - h + \frac{2}{a}h^2 + \cdots.$$

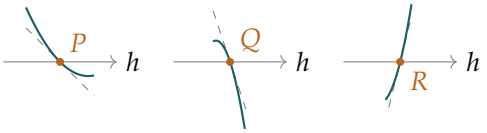
在 $Q(a,0)$： $y' = -3a^2/a^2 = -3$。於是 $0 + 2a(9) + 6a + a^2 y'' = 0$，所以 $y'' = -24/a$，而

$$y = -3h - \frac{12}{a}h^2 + \cdots.$$

在 $R(a,-a)$： 這裏 $2xy + a^2 = -a^2$，所以 $y' = -4a^2/(-a^2) = 4$。於是 $4(-a)(4) + 2a(16) + 6a - a^2 y'' = 0$，即 $22a = a^2 y''$，所以 $y'' = 22/a$，而

$$y = -a + 4h + \frac{11}{a}h^2 + \cdots.$$

因此在 P 附近，曲線以斜率（gradient）-1 下降，並且向上凹（concave up，即彎向切線上方）；在 Q 附近，曲線以斜率 -3 急降，並且向下凹（concave down）；在 R 附近，曲線以斜率 4 上升，並且向上凹。



（取 $a = 1$ 時的局部形狀；虛線為切線。）

不等式

把 C 的方程寫成 y 的二次方程：

$$x\,y^2 + a^2\,y + (x^3 - a^3) = 0.$$

若 $x \neq 0$ 而 (x,y) 在 C 上，這個二次方程便有實根 y，所以它的判別式非負：

$$a^4 - 4x(x^3 - a^3) \geq 0 \quad \Longleftrightarrow \quad 4x^4 - 4a^3 x - a^4 \leq 0.$$

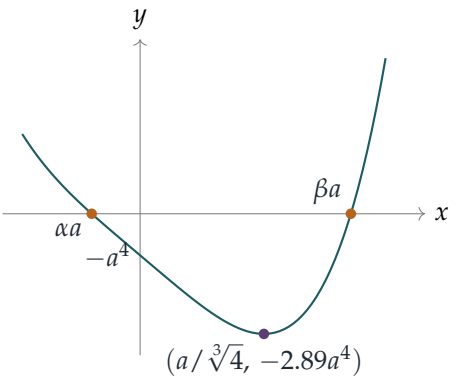
若 $x = 0$，左邊等於 $-a^4 < 0$，所以不等式在該處同樣成立。結果得證。

$y = 4x^4 - 4a^3 x - a^4$ 的圖像

試題原文印的是 $4x^4 - 4a^3 - a^4$，這肯定是誤植，應為我們剛剛遇到的四次多項式（quartic）$g(x) = 4x^4 - 4a^3 x - a^4$。我們有 $g(0) = -a^4$ 及 $g'(x) = 16x^3 - 4a^3$，後者只在 $x_0 = a/\sqrt[3]{4} \approx 0.63a$ 處等於零，g 在該處取得極小值。利用 $x_0^3 = a^3/4$，

$$g(x_0) = a^3 x_0 - 4a^3 x_0 - a^4 = -a^4\left(1 + 3 \cdot 4^{-1/3}\right) \approx -2.89a^4.$$

由於當 $x \to \pm\infty$ 時 $g \to +\infty$，而 g 先遞減後遞增，所以它恰好有兩個零點 $x = \alpha a$ 和 $x = \beta a$，其中 $\alpha \approx -0.246$ 和 $\beta \approx 1.072$ 是 $4t^4 - 4t - 1 = 0$ 的根（它們沒有簡潔的閉合形式，我們也用不着）。此外 $g''(x) = 48x^2 \geq 0$，所以圖像是凸（convex）的。



（以 $a = 1$ 繪畫。）所以 C 上每一點都滿足 $\alpha a \leq x \leq \beta a$：曲線位於一條垂直的帶狀區域之內。

曲線 C

當 $x \neq 0$ 時，解二次方程得出兩條分支

$$y_\pm = \frac{-a^2 \pm \sqrt{D}}{2x}, \qquad D = a^4 + 4a^3x - 4x^4,$$

它們恰好在該帶狀區域上取實數值，並在 $D = 0$ 之處相接，即 $x = \alpha a$ 和 $x = \beta a$。在那裏 $y = -a^2/(2x)$，而切線是垂直的，因為那正是 y' 的分母 $2xy + a^2$ 等於零的地方。由此得到 $(\beta a, -0.47a)$ 和 $(\alpha a, 2.03a)$ 兩點。

? **想一想** 當 $x \to 0$ 時，兩條分支會怎樣？

把分子和分母同乘以 $\sqrt{D} + a^2$，並利用 $D - a^4 = 4x(a^3 - x^3)$，得

$$y_+ = \frac{2(a^3 - x^3)}{a^2 + \sqrt{D}},$$

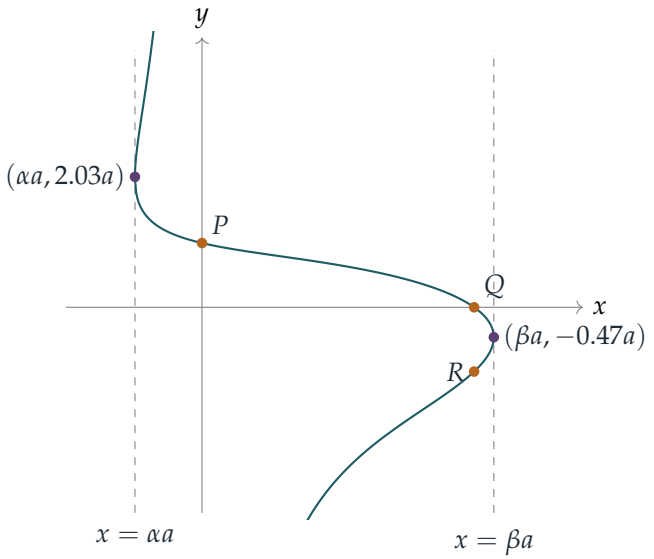
它在 $x = 0$ 處完全沒有問題，並且在該處等於 a。所以 y_+ 是一條通過 P 的光滑分支。另一條分支在 x 很小時滿足 $y_- \approx -a^2/x$，所以當 $x \to 0^+$ 時 $y_- \to -\infty$，當 $x \to 0^-$ 時 $y_- \to +\infty$。這就是沿 y 軸的漸近線（asymptote）。

再多幾項事實，有助確定草圖的形狀。

- 曲線沒有水平切線：$y' = 0$ 需要 $y^2 + 3x^2 = 0$，即 $(0,0)$，而這一點不在 C 上。

- 曲線只在 $x^3 = a^3$ 之處與 $y = 0$ 相交，即 Q；並且只在 P 與 $x = 0$ 相交。
- 在 $x = a$ 處，$D = a^4$，所以 $y_+ = 0$（點 Q）而 $y_- = -a$（點 R）。

綜合以上各點，C 是一條連通的曲線。它在 y 軸左邊不遠處沿 y_- 從 $+\infty$ 下降，到達 $(\alpha a, 2.03a)$ 處的垂直切線。然後它沿 y_+ 折回，經過 P（斜率 -1），再下降經過 Q（斜率 -3），到達 $(\beta a, -0.47a)$ 處的垂直切線，接着沿 y_- 折回，經過 R（斜率 4），在 y 軸右邊不遠處下降至 $-\infty$。曲線在 P 向上凹，在 Q 向下凹，所以兩點之間有一個拐點（point of inflection）。



（以 $a = 1$ 繪畫；虛線是上文求得的帶狀區域的邊界。）

驗算

- 取 $a = 1.7$，我們在每一點附近的 $x = x_0 + h$ 處解出二次方程的 y。每當 h 減半，各個兩項泰勒級數的誤差便縮小約 8 倍，這正是我們預期的 h^3 表現。
- 隨機取 10^5 個 x 值，二次方程有實根的情況，恰好就是 $4x^4 - 4a^3x - a^4 \leq 0$ 的情況。
- 用二分法（bisection）得 $\alpha \approx -0.2463$ 和 $\beta \approx 1.0724$。在 $x = \pm 0.001$ 處（取 $a = 1.7$），一個根接近 a，另一個根接近 $-a^2/x \approx \mp 2890$。

常見錯誤

- 漏掉因子 $\frac{1}{2}$：h^2 的係數是 $\frac{1}{2}y''$，不是 y''。

- 為了使用判別式而除以 x，卻沒有另外檢查 $x = 0$ 的情況。
- 沒有察覺 y_+ 光滑地通過 P。寫成 $(-a^2 + \sqrt{D})/(2x)$ 時，它在 $x = 0$ 處看似奇異，其實不然。

STEP 之外

這一題中，我們利用隱式定義的函數的泰勒展開式，找出曲線在某一點附近的表現。保證在 $\partial F/\partial y \neq 0$ 之處 y 是 x 的光滑函數的定理，就是**隱函數定理**（implicit function theorem），它是多元分析（multivariable analysis）的一個重要結果。帶狀區域兩端的垂直切線，正好出現在 $\partial F/\partial y = 2xy + a^2$ 等於零的點。這條曲線是平面上的三次曲線（cubic），而研究這類曲線會引領我們走向代數幾何（algebraic geometry）和橢圓曲線（elliptic curve）。想做更多這種風格的題目，可試試 Siklos 的 *Advanced Problems in Mathematics*（進階數學問題集）。

詳細解答

STEP III 1990，第 8 題

純數學 • 恰當微分方程

STEP III 1990，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/90-s3-q8

關鍵思路

展開 $\dfrac{\mathrm{d}}{\mathrm{d}x}(py' + qy)$，再與 $Py'' + Qy' + Ry$ 比較：這樣便完全確定了 p 和 q，而條件 $P'' - Q' + R = 0$ 正好令最後一個係數吻合。至於題中的微分方程（differential equation），這個條件成立，所以積分一次就能把二階方程直接化為一階線性方程，再用積分因子（integrating factor）便可完成。

解答

恰當性的判別準則

對任意函數 p、q 和 y，由乘積法則（product rule）得

$$\frac{\mathrm{d}}{\mathrm{d}x}(py' + qy) = py'' + (p' + q)y' + q'y. \qquad (*)$$

僅當。假設 p 和 q（兩者均與 y 無關）使 $Py'' + Qy' + Ry$ 對每一個 y 都等於 $(*)$。取 $y = 1$，得 $R = q'$。取 $y = x$，得 $Q + Rx = (p' + q) + q'x$，所以 $Q = p' + q$。取 $y = \frac{1}{2}x^2$，得 $P + Qx + \frac{1}{2}Rx^2 = p + (p' + q)x + \frac{1}{2}q'x^2$，所以 $P = p$。因此

$$P'' - Q' + R = p'' - (p'' + q') + q' = 0.$$

若。假設 $P'' - Q' + R = 0$。取

$$p = P, \qquad q = Q - P'.$$

那麼 $p' + q = P' + Q - P' = Q$，而利用該條件可得 $q' = Q' - P'' = R$。由 $(*)$，對每一個 y 都有 $\dfrac{\mathrm{d}}{\mathrm{d}x}(py' + qy) = Py'' + Qy' + Ry$，這正是所要證明的。

以上的推導同時說明 p 和 q 必定是 P 和 $Q-P'$，下面我們就用這個方法把它們求出來。

微分方程

這裏 $P=x-x^4$、$Q=1-7x^3$ 及 $R=-9x^2$，所以

$$P''-Q'+R=-12x^2-(-21x^2)-9x^2=0.$$

判別準則適用，其中 $p=P=x-x^4$ 及 $q=Q-P'=(1-7x^3)-(1-4x^3)=-3x^3$。方程變成

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(x-x^4)y'-3x^3y\right]=(x^3+3x)\mathrm{e}^x.$$

反覆運用分部積分法（integration by parts），得 $\int x^3\mathrm{e}^x\,\mathrm{d}x=(x^3-3x^2+6x-6)\mathrm{e}^x$ 及 $\int 3x\mathrm{e}^x\,\mathrm{d}x=(3x-3)\mathrm{e}^x$，所以

$$(x-x^4)y'-3x^3y=(x^3-3x^2+9x-9)\mathrm{e}^x+C.$$

當 $x=2$ 時，$y=2\mathrm{e}^2$ 及 $y'=0$，所以左邊是 $-3\cdot 8\cdot 2\mathrm{e}^2=-48\mathrm{e}^2$，右邊是 $5\mathrm{e}^2+C$。因此 $C=-53\mathrm{e}^2$。

> **想一想** 左邊的 $(x-x^4)y'-3x^3y$ 並不是 $(x-x^4)y$ 的導數。那麼第二次積分該怎樣做？

抽出因子 x：左邊是 $x\left[(1-x^3)y'-3x^2y\right]=x\dfrac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^3)y\right]$。（這就是常用的積分因子：$y'-\frac{3x^2}{1-x^3}y$ 的積分因子是 $\exp\int\frac{-3x^2}{1-x^3}\,\mathrm{d}x=1-x^3$。）由於解位於 $x=2$ 附近，可以除以 x，於是

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^3)y\right]=(x^2-3x+9)\mathrm{e}^x-\frac{9\mathrm{e}^x}{x}-\frac{53\mathrm{e}^2}{x}.$$

再用分部積分法，得 $\int(x^2-3x+9)\mathrm{e}^x\,\mathrm{d}x=(x^2-5x+14)\mathrm{e}^x$。$\mathrm{e}^x/x$ 這一項沒有初等原函數（elementary antiderivative），所以我們把它保留為由起點 $x=2$ 開始的積分：

$$(1-x^3)y=(x^2-5x+14)\mathrm{e}^x-53\mathrm{e}^2\ln\frac{x}{2}-9\int_2^x\frac{\mathrm{e}^t}{t}\,\mathrm{d}t+D.$$

當 $x=2$ 時，左邊是 $-7\cdot 2\mathrm{e}^2=-14\mathrm{e}^2$，右邊是 $8\mathrm{e}^2+D$，所以 $D=-22\mathrm{e}^2$。因此

$$\boxed{y=\frac{1}{1-x^3}\left[(x^2-5x+14)\mathrm{e}^x-22\mathrm{e}^2-53\mathrm{e}^2\ln\frac{x}{2}-9\int_2^x\frac{\mathrm{e}^t}{t}\,\mathrm{d}t\right]}$$

適用於 $x>1$。y'' 的係數 $x-x^4$ 在 $x=1$ 處為零，所以經過 $x=2$ 的解定義在 $x>1$ 上，而不一定能延伸至 $x=1$ 以外。

驗算

- 當 $x = 2$ 時，方括號內的值是 $8e^2 - 22e^2 = -14e^2$，除以 $1 - 8 = -7$ 得 $y = 2e^2$，符合要求。
- 由 $x = 2$ 開始以數值方法逐步求解原方程（四階龍格－庫塔法，Runge–Kutta），得 $y(3) \approx 7.90306$，與公式的結果吻合，其中的積分以辛普森法則（Simpson's rule）計算。兩者在 $x = 1.2$、1.5 和 4 處亦一致。
- 展開 $\frac{\mathrm{d}}{\mathrm{d}x}\left[(x - x^4)y' - 3x^3y\right]$，可還原為 $(x - x^4)y'' + (1 - 7x^3)y' - 9x^2y$。

常見錯誤

- 第一次積分後漏掉常數 C。它在這裏並不是零，而且對數項正是由它產生的。
- 硬要把 $\int e^x/x\,\mathrm{d}x$ 寫成閉合形式（closed form）。這是做不到的，所以應把它保留為下限是 2 的定積分，這樣代入初始條件也方便得多。

STEP 之外

若表達式 $Py'' + Qy' + Ry$ 是某個函數的全導數，便稱它為恰當（exact）的。表達式 $P'' - Q' + R$ 就是伴隨（adjoint）方程 $(Pz)'' - (Qz)' + Rz = 0$ 的左邊在 $z = 1$ 時的值。一般而言，伴隨方程的任何一個解 z 都是積分因子，能使 $z(Py'' + Qy' + Ry)$ 成為恰當的。由伴隨方程再往前走，便是自伴隨（self-adjoint）問題，即施圖姆－劉維爾（Sturm–Liouville）問題，以及格林函數（Green's function）。Boyce 與 DiPrima 所著的 *Elementary Differential Equations*（一本初等微分方程教科書）在二階線性方程的習題中，涵蓋了恰當方程和伴隨方程。

詳細解答

STEP III 1990，第 9 題

純數學 • 雙曲函數

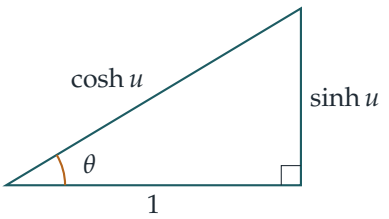
STEP III 1990，第 9 題

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/90-s3-q9

關鍵思路

關係式 $\tan\theta = \sinh u$ 等同於 $\sec^2\theta = 1 + \sinh^2 u = \cosh^2 u$，因此兩個變量的三角函數（trigonometric functions）與雙曲函數（hyperbolic functions）完全對應：$\cos\theta = \operatorname{sech} u$、$\sin\theta = \tanh u$、$\sec\theta = \cosh u$。畫一個邊長為 1、$\sinh u$ 和 $\cosh u$ 的直角三角形，這些關係便一目了然。有了這本「字典」，每一部分都只需先做一次簡短的求導，再把結果翻譯回來。



解答

在整個解答中，由於 $0 \le \theta < \frac{1}{2}\pi$，我們有 $\tan\theta \ge 0$，所以 $\sinh u \ge 0$ 且 $u \ge 0$。此外 $\cos\theta > 0$ 且 $\cosh u > 0$。

第 (i) 部分

利用 $1 + \tan^2\theta = \sec^2\theta$ 和 $1 + \sinh^2 u = \cosh^2 u$，

$$\sec^2\theta = 1 + \tan^2\theta = 1 + \sinh^2 u = \cosh^2 u.$$

$\sec\theta$ 和 $\cosh u$ 均為正數，所以 $\sec\theta = \cosh u$，即 $\cos\theta = \operatorname{sech} u = v$，這正是所要證明的結果。

值得順帶記下與之配對的結果：$\sin\theta = \tan\theta\cos\theta = \sinh u \operatorname{sech} u = \tanh u$。

第 (ii) 部分

把 $\tan\theta = \sinh u$ 對 u 求導（利用鏈式法則（chain rule）作隱函數求導（implicit differentiation））：

$$\sec^2\theta\,\frac{\mathrm{d}\theta}{\mathrm{d}u} = \cosh u \quad\Longrightarrow\quad \frac{\mathrm{d}\theta}{\mathrm{d}u} = \cosh u\cos^2\theta = \cosh u\operatorname{sech}^2 u = \operatorname{sech} u = v,$$

其中 $\cos\theta$ 的表達式來自第 (i) 部分。

第 (iii) 部分

由於 $\frac{\mathrm{d}}{\mathrm{d}u}\cosh u = \sinh u$，

$$\frac{\mathrm{d}v}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}u}\big(\cosh u\big)^{-1} = -\frac{\sinh u}{\cosh^2 u} = -\operatorname{sech} u\tanh u.$$

查一查字典，$\operatorname{sech} u\tanh u = \cos\theta\sin\theta = \frac{1}{2}\sin 2\theta$，所以

$$\sin 2\theta = -2\,\frac{\mathrm{d}v}{\mathrm{d}u}.$$

至於二階導數（second derivative），可運用積法則（product rule），並配合 $\frac{\mathrm{d}}{\mathrm{d}u}\tanh u = \operatorname{sech}^2 u$：

$$\begin{aligned}\frac{\mathrm{d}^2v}{\mathrm{d}u^2} &= -\Big[\big(-\operatorname{sech} u\tanh u\big)\tanh u + \operatorname{sech} u\cdot\operatorname{sech}^2 u\Big]\\ &= \operatorname{sech} u\big(\tanh^2 u - \operatorname{sech}^2 u\big).\end{aligned}$$

乘以 $-\cosh u$ 便可消去 $\operatorname{sech} u$：

$$-\cosh u\,\frac{\mathrm{d}^2v}{\mathrm{d}u^2} = \operatorname{sech}^2 u - \tanh^2 u = \cos^2\theta - \sin^2\theta = \cos 2\theta,$$

這就是第二個結果。

第 (iv) 部分

現在把 u 和 v 看成 θ 的函數。由第 (i) 部分，$v = \cos\theta$ 且 $\cosh u = \sec\theta$；而第 (ii) 部分給出 $\frac{\mathrm{d}u}{\mathrm{d}\theta} = 1\big/\frac{\mathrm{d}\theta}{\mathrm{d}u} = 1/v$（因為 $v > 0$，此式成立）。所以

$$\frac{\mathrm{d}u}{\mathrm{d}\theta} = \sec\theta, \qquad \frac{\mathrm{d}^2u}{\mathrm{d}\theta^2} = \sec\theta\tan\theta,$$
$$\frac{\mathrm{d}v}{\mathrm{d}\theta} = -\sin\theta, \qquad \frac{\mathrm{d}^2v}{\mathrm{d}\theta^2} = -\cos\theta.$$

想一想 在埋頭苦算之前，先想一想：第 (iv) 部分的恆等式，照試卷上所印的樣子看來完全正確嗎？

並不正確。按試卷所印，第一項含有 $\frac{\mathrm{d}^2u}{\mathrm{d}\theta^2}$，於是左方為

$$\begin{aligned}\sec\theta\cdot\sec\theta\tan\theta-\sin\theta\sec\theta\tan\theta+\sec^2\theta &= \sec^2\theta\tan\theta+(\sec^2\theta-\tan^2\theta)\\ &= \sec^2\theta\tan\theta+1\ \geq 1,\end{aligned}$$

它永遠不會等於零。題目原意的恆等式（identity），第一項應為 $\frac{\mathrm{d}^2v}{\mathrm{d}\theta^2}$（把 v 誤印成 u）：

$$\frac{\mathrm{d}u}{\mathrm{d}\theta}\frac{\mathrm{d}^2v}{\mathrm{d}\theta^2}+\frac{\mathrm{d}v}{\mathrm{d}\theta}\frac{\mathrm{d}^2u}{\mathrm{d}\theta^2}+\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2=0.$$

代入上述四個導數，

$$\sec\theta(-\cos\theta)+(-\sin\theta)\sec\theta\tan\theta+\sec^2\theta=-1-\tan^2\theta+\sec^2\theta=0,$$

因為 $\sec^2\theta=1+\tan^2\theta$。這就證明了修正後的結果。

還有一個更簡潔的看法：根據積法則，首兩項合起來就是 $\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\frac{\mathrm{d}v}{\mathrm{d}\theta}\right)$，而 $\frac{\mathrm{d}u}{\mathrm{d}\theta}\frac{\mathrm{d}v}{\mathrm{d}\theta}=-\sec\theta\sin\theta=-\tan\theta$，其導數為 $-\sec^2\theta=-\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2$。

驗算

取 $u=\ln 2$，則 $\sinh u=\frac{3}{4}$、$\cosh u=\frac{5}{4}$ 且 $\tan\theta=\frac{3}{4}$：由 3–4–5 三角形可得 $\cos\theta=\frac{4}{5}=\operatorname{sech}u$ 及 $\sin 2\theta=\frac{24}{25}$，而 $-2\,\mathrm{d}v/\mathrm{d}u=2\cdot\frac{4}{5}\cdot\frac{3}{5}=\frac{24}{25}$。數值方面，在 2000 個隨機點上以有限差分（finite difference）計算的導數，與 (i)–(iii) 及修正後的 (iv) 吻合，誤差約為 10^{-7}，並確認試卷所印的 (iv) 左方等於 $1+\sec^2\theta\tan\theta$。

常見錯誤

- 寫出 $\cos\theta=\pm\operatorname{sech}u$ 卻沒有說明為何應取正號：決定正負號的是範圍 $0\leq\theta<\frac{1}{2}\pi$。
- 忘記 $\frac{\mathrm{d}u}{\mathrm{d}\theta}$ 只不過是 $\frac{\mathrm{d}\theta}{\mathrm{d}u}$ 的倒數，反而從頭對 $\operatorname{arsinh}(\tan\theta)$ 求導。

STEP 之外

函數 $\theta(u)=\arctan(\sinh u)$ 稱為*古德曼函數*（Gudermannian function），它毋須借助複數（complex numbers）便把圓函數與雙曲函數聯繫起來；其反函數（inverse function）$u=\ln(\sec\theta+\tan\theta)=\int_0^\theta\sec t\,\mathrm{d}t$ 正好是麥卡托（Mercator）地圖上的縱坐標。當你寫下 $\cosh(\mathrm{i}x)=\cos x$ 時，同樣的配對也會出現；Needham 的 *Visual Complex Analysis*（《可視化複分析》）是從幾何角度欣賞這些聯繫的絕佳讀物。

詳細解答

STEP III 1990，第 10 題

純數學 • 根與漸近近似

STEP III 1990，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/90-s3-q10

關鍵思路

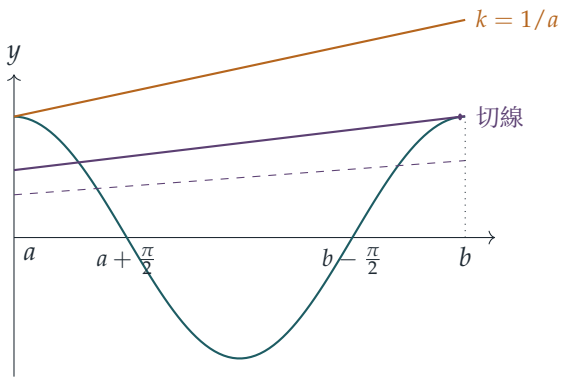
在題目所給的區間內，$\sin x$ 剛好走完一個完整週期，起點和終點都在峯頂，而直線 $y = kx$ 則恆為正，並緩緩上升。根只可能出現在 $\sin x > 0$ 的地方，而這樣的地方分為兩段：一段是 $\sin x$ 下降的部分（最多相交一次），另一段是 $\sin x$ 上升而且是凹函數（concave）的部分（最多相交兩次）。值得留意的臨界情況，是直線剛好與第二個拱形相切之時；把切點寫成在區間右端點左方一小段距離 δ 處，相切條件便化為題目所給的方程。

解答

以下一律記

$$a = (4n+1)\frac{\pi}{2}, \qquad b = (4n+5)\frac{\pi}{2} = a + 2\pi,$$

於是區間就是 $a < x < b$。由於 $\sin a = \sin b = 1$，$\sin x$ 的圖像由峯頂出發，在 $x = a + \frac{\pi}{2}$ 處降至 0，其後一直為負，直至 $x = a + \frac{3\pi}{2} = b - \frac{\pi}{2}$，然後再回升至 b 處的峯頂。



上圖（按 $n = 1$ 的比例繪畫，原點遠在左方圖外）以藍綠色顯示 $\sin x$，以琥珀色顯示通過 $(a, 1)$ 的直線，以紫紅色實線顯示切線（tangent），另以虛線顯示一條斜率較小、與曲線相交三次的直線。

根的可能數目

由於 $k > 0$ 且 $x > 0$，故 $kx > 0$，所以每個根都位於 $\sin x > 0$ 之處：不是在 $I_1 = (a, a + \frac{\pi}{2})$ 內，就是在 $I_2 = (b - \frac{\pi}{2}, b)$ 內。

在 I_1 上。這裏 $\sin x$ 嚴格遞減（strictly decreasing），而 kx 嚴格遞增（strictly increasing），所以 $kx - \sin x$ 嚴格遞增，最多只有一個根。在右端點處 $kx - \sin x > 0$；在 $x = a$ 處其值為 $ka - 1$。因此，若 $ka < 1$，則 I_1 內恰有一個根；若 $ka \geq 1$，則沒有根（此時在整個區間上 $kx - \sin x > 0$）。

在 I_2 上。設 $f(x) = \sin x - kx$。這裏 $f''(x) = -\sin x < 0$，所以 f 是凹函數，在 I_2 上最多有兩個零點。此外，f 在左端點處 $f < 0$，而 $f(b) = 1 - kb$。

- 若 $kb < 1$，則 f 由負變正，而凹函數在上升途中只能變號一次，所以恰有一個根。當 $kb = 1$ 時結論相同：此時 $f(b) = 0$ 且 $f'(b) = \cos b - k = -k < 0$，所以在 b 的左側附近 $f > 0$，於是 I_2 內有一個根（b 本身這個零點不計在內）。
- 若 $kb > 1$，則 f 在兩端均有 $f < 0$，所以根的數目是 0、1 或 2，分別對應 f 在 I_2 上的最大值為負、為零或為正。最大值為 0 即相切的情況：直線與拱形在某點 $t \in I_2$ 相切，設此時的斜率為 k_t。當 k 由很大的值逐漸減小時，直線隨之下降，所以當 $k > k_t$ 時沒有根，當 $k = k_t$ 時有一個根，而當 $\frac{1}{b} < k < k_t$ 時有兩個根。

在切點處 $k_t t = \sin t < 1$ 且 $t > a$，所以 $k_t < 1/a$；又因為相切的情況滿足 $kb > 1$，故 $k_t > 1/b$。把兩段的結果合起來：

k 的範圍	I_1 內的根數	I_2 內的根數	總數
$k \leq 1/b$	1	1	2
$1/b < k < k_t$	1	2	3
$k = k_t$	1	1	2
$k_t < k < 1/a$	1	0	1
$k \geq 1/a$	0	0	0

所以方程在該區間內可以有 0、1、2 或 3 個根，正如題目所求。

恰有一個根

由上表可見，恰有一個根的充要條件是 $k_t < k < \dfrac{1}{a} = \dfrac{2}{(4n+1)\pi}$。餘下的工作是描述 k_t。

想一想 怎樣把「直線與曲線相切」寫成方程？

直線 $y = kx$ 與 $y = \sin x$ 在 $x = t$ 處相切，即兩個圖像在該處相交，而且斜率相等：

$$\sin t = k_t t, \qquad \cos t = k_t.$$

點 t 位於 I_2 內，所以可寫 $t = b - \delta$，其中 $0 < \delta < \frac{\pi}{2}$。由於 $\sin b = 1$ 及 $\cos b = 0$，

$$\sin t = \sin(b-\delta) = \cos\delta, \qquad \cos t = \cos(b-\delta) = \sin\delta.$$

所以由斜率條件得 $k_t = \sin\delta$，而相交條件則變成

$$\cos\delta = \left((4n+5)\frac{\pi}{2} - \delta\right)\sin\delta.$$

因此，恰有一個根的條件是

$$\sin\delta < k < \frac{2}{(4n+1)\pi},$$

其中 δ 如上所述，這正是我們要證明的結果。

n 很大時的近似值

把關於 δ 的方程除以 $\sin\delta$（由於 $0 < \delta < \frac{\pi}{2}$，它不等於零）：

$$\cot\delta = b - \delta, \qquad \text{即} \qquad \tan\delta = \frac{1}{b-\delta}.$$

由於 $\delta < \frac{\pi}{2}$，右邊小於 $1/(b - \frac{\pi}{2})$，而當 n 很大時這個值很小。所以 δ 很小，$\tan\delta \approx \delta$，且 $b - \delta \approx b$，由此得

$$\delta \approx \frac{1}{b} = \frac{2}{(4n+5)\pi}.$$

要求下一級的近似值，可寫 $\delta = \dfrac{1}{b} + \dfrac{c}{b^3}$，並在兩邊保留至 b^{-3} 的項。左邊是 $\tan\delta = \delta + \frac{1}{3}\delta^3 + \cdots = \dfrac{1}{b} + \dfrac{c+\frac{1}{3}}{b^3} + \cdots$。右邊是

$$\frac{1}{b-\delta} = \frac{1}{b}\left(1 - \frac{\delta}{b}\right)^{-1} = \frac{1}{b}\left(1 + \frac{\delta}{b} + \cdots\right) = \frac{1}{b} + \frac{1}{b^3} + \cdots,$$

其中在修正項裏用了 $\delta \approx 1/b$。比較 b^{-3} 的項，得 $c + \frac{1}{3} = 1$，所以 $c = \frac{2}{3}$，於是

$$\delta \approx \frac{1}{b} + \frac{2}{3b^3} = \frac{2}{(4n+5)\pi} + \frac{16}{3(4n+5)^3\pi^3}.$$

（等價的做法是：由 $\delta = 1/b$ 出發，把迭代（iteration）$\delta \mapsto \arctan\dfrac{1}{b-\delta}$ 進行一步，得到 $\arctan\dfrac{b}{b^2-1}$，它的展開式與上式相同。）由此可知，k 的區間的下端是 $\sin\delta \approx \dfrac{1}{b} + \dfrac{1}{2b^3}$。

驗算

- 當 $n = 1$ 時，以數值方法解 $\cos\delta = (\frac{9\pi}{2} - \delta)\sin\delta$ 得 $\delta = 0.070973$，而 $1/b = 0.070736$ 及 $1/b + 2/(3b^3) = 0.070971$：改良後的近似值準確了約 150 倍。當 $n = 10$ 時，誤差降至約 5×10^{-10}。
- 取 $n = 1, 2, 3, 5$，並從表中每一行各取一個 k 值，在細密的格點上數算 $kx - \sin x$ 的變號次數，所得的根數 $2, 3, 1, 0$ 與預測完全吻合；而在 $\sin\delta$ 與 $2/((4n+1)\pi)$ 之間隨機選取 k，則每次都得到一個根。

常見錯誤

- 忘記了相切的情況，或者把 $1/b$、k_t 和 $1/a$ 的大小次序弄錯，以致根的數目數錯。
- 在第二級近似中只用 $\tan\delta \approx \delta$，結果得到 $\delta \approx b/(b^2 - 1) \approx 1/b + 1/b^3$。$\tan\delta$ 中的 $\frac{1}{3}\delta^3$ 項與修正項屬同一階，必須保留。

STEP 之外

當方程含有一個很小或很大的參數時，逐步求出其根越來越精確的近似值，這就是*漸近分析*（asymptotic analysis）：把一個以 $1/b$ 的冪寫成的級數代入方程，然後比較係數，做法與這裏完全一樣。把同樣的想法應用於 $\tan x = x$，便可描述該方程的大根，這些根會在熱傳導和振動問題中出現。如想進一步了解這類展開式背後的分析學，可參閱 Bartle 與 Sherbert 的 *Introduction to Real Analysis*（實分析入門教材）；Siklos 的 *Advanced Problems in Mathematics* 則有更多 STEP 風格的練習。

詳細解答

STEP III 1990，第 11 題

力學 • 剛體轉動

STEP III 1990，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：

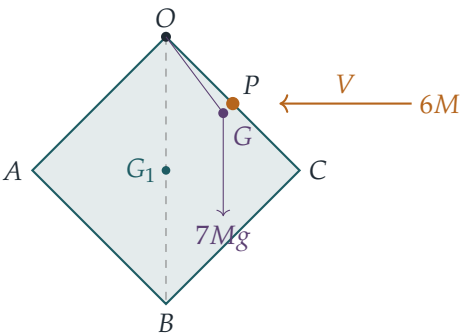
step.maths.org/questions/90-s3-q11

關鍵思路

碰撞在一瞬間發生，其間位於 O 的轉軸會施加一個很大的衝力（impulsive force），所以線動量（linear momentum）並不守恆。不過這個衝量（impulse）作用於 O，因此對 O 沒有力矩（moment）：在整個碰撞過程中，對轉軸的角動量（angular momentum）守恆。碰撞之後，整個系統是一個繞固定軸轉動的剛體（rigid body），只有重力作功，於是能量守恆（conservation of energy）便可告訴我們它能否越過最高點。

解答

我們把薄片（lamina）視為均勻的剛性正方形，把質點視為在碰撞一刻黏附在薄片上的點質量，並假設轉軸光滑。取 O 為原點。當 OB 鉛垂向下時，正方形的中心 G_1 位於深度 $\sqrt{2}a$ 處（即對角線 $2\sqrt{2}a$ 的一半），而 A 和 C 分別位於 OB 兩側、水平距離為 $\sqrt{2}a$ 之處，深度同樣是 $\sqrt{2}a$。因此 OC 的中點 P 與 O 的水平距離為 $a/\sqrt{2}$，深度為 $a/\sqrt{2}$，並且 $OP = a$。



（圖中所示的質點從右方飛來；稍後我們會看到，從左方飛來也會得到相同的答案。）

對轉軸的轉動慣量

對於邊長為 $2a$ 的均勻正方形薄片，它對通過中心並垂直於薄片的軸的轉動慣量（moment of inertia）是 $\frac{1}{12}M\big((2a)^2+(2a)^2\big)=\frac{2}{3}Ma^2$（根據垂直軸定理（perpendicular axis theorem），把薄片平面內兩條軸的轉動慣量相加，每條均為 $\frac{1}{3}Ma^2$）。由於 $OG_1=\sqrt{2}\,a$，根據平行軸定理（parallel axis theorem），

$$I_{\text{薄片}}=\tfrac{2}{3}Ma^2+M\big(\sqrt{2}\,a\big)^2=\tfrac{8}{3}Ma^2.$$

質點最後停在 P，與 O 相距 a，所以合成後的物體有

$$I=\tfrac{8}{3}Ma^2+6Ma^2=\tfrac{26}{3}Ma^2.$$

碰撞

想一想 為甚麼處理這次碰撞時不能使用線動量守恆？

因為轉軸是固定的：要令 O 保持不動，轉軸必須在碰撞期間向薄片施加一個衝量，而這個衝量是未知的。不過它的作用線通過 O，所以它對 O 的力矩為零，而重力在一瞬間內不會產生衝量。因此對轉軸的角動量守恆。

在碰撞前的一刻，質點對 O 的角動量等於它的動量 $6MV$ 乘以由 O 至其（水平）運動路線的垂直距離，這個距離就是 P 的深度，即 $a/\sqrt{2}$。薄片此時靜止。碰撞剛結束時，整個物體以角速率（angular speed）ω 轉動，所以

$$\tfrac{26}{3}Ma^2\,\omega=6MV\cdot\frac{a}{\sqrt{2}}\qquad\Longrightarrow\qquad\omega=\frac{9V}{13\sqrt{2}\,a}.$$

於是緊接碰撞之後的動能（kinetic energy）為

$$\tfrac{1}{2}I\omega^2=\tfrac{1}{2}\cdot\tfrac{26}{3}Ma^2\cdot\frac{81V^2}{338a^2}=\frac{27}{26}MV^2.$$

作完整的旋轉

合成物體的質心（centre of mass）G（質量為 $7M$）相對於 O 的水平位移和深度分別是

$$\frac{M\cdot 0+6M\cdot\frac{a}{\sqrt{2}}}{7M}=\frac{3\sqrt{2}}{7}a,\qquad\frac{M\cdot\sqrt{2}\,a+6M\cdot\frac{a}{\sqrt{2}}}{7M}=\frac{4\sqrt{2}}{7}a,$$

所以 $OG=\frac{\sqrt{2}}{7}a\sqrt{3^2+4^2}=\frac{5\sqrt{2}}{7}a$。

這個物體是剛體，並且釘在 O 上，所以（與繫在繩上的質點不同）它不可能「掉離」所走的圓形軌道；它能作完整的旋轉，當且僅當它在勢能（potential energy）最高的位置

仍有正的動能，而這個位置就是 G 在 O 正上方的時候。由起初在 O 之下深度 $\frac{4\sqrt{2}}{7}a$ 處，到在 O 之上高度 $\frac{5\sqrt{2}}{7}a$ 處，G 上升了

$$\frac{4\sqrt{2}}{7}a + \frac{5\sqrt{2}}{7}a = \frac{9\sqrt{2}}{7}a,$$

無論物體向哪一個方向轉動都是如此，所以質點從哪一邊飛來並無影響。碰撞之後只有重力作功（光滑的轉軸不作功），所以根據能量守恆，我們需要

$$\frac{27}{26}MV^2 > 7Mg \cdot \frac{9\sqrt{2}}{7}a = 9\sqrt{2}\,Mga,$$

即

$$\boxed{V^2 > \frac{26\sqrt{2}}{3}ga.}$$

質量 M 被約去了，這是必然的，因為題目中每一個質量都是 M 的倍數。（若剛好取等號，物體只會緩緩趨近最高位置而永遠不會越過它，所以這是嚴格不等式。）

驗算

- 轉動慣量、質心（它與向下鉛垂線所成角的 tan 值為 $\frac{3}{4}$）以及動能係數 $\frac{27}{26}$ 均已用數值方法核實，其中第一項是在傾斜的正方形上以網格積分求得的。
- 取 $a = 1.3$，$g = 9.81$，以數值方法逐步求解運動方程 $I\ddot{\theta} = -7Mg\,x_G$：當 V^2 為 $\frac{26\sqrt{2}}{3}ga \approx 15.6$ 的 98% 時，物體會擺回來；而在 102% 時，它會轉完一整圈，無論質點從哪一邊飛來都是如此。
- 對能量作合理性檢查：質點原有的 $3MV^2$ 動能之中，只有 $\frac{27}{26}MV^2$ 在碰撞後保留下來，這正符合完全非彈性碰撞（perfectly inelastic collision）應有的情況。

常見錯誤

- 在碰撞過程中使用線動量守恆：轉軸會施加衝量，所以線動量並不守恆。
- 以距離 $OP = a$ 作為質點動量的力臂，而沒有使用到其運動路線的垂直距離 $a/\sqrt{2}$。
- 忘記質點黏上之後質心已經偏離 OB，所以升至最高點的高度並不單單是原來深度的兩倍。

STEP 之外

這是剛體繞固定軸轉動的力學：由於轉軸上有衝擊反作用力，在碰撞中守恆的量是角動量，而不是線動量。一個精彩的延伸課題是撞擊中心（centre of percussion）：在這一

點上施加打擊，轉軸完全不會產生衝擊反作用力，這正是板球棒有「甜蜜點」的原因。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對剛體轉動和角衝量有細緻的論述。

詳細解答

STEP III 1990，第 12 題

力學 • 移動楔上的質點

> **STEP III 1990，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/90-s3-q12

關鍵思路

質點下滑期間有兩個量守恆：水平動量（horizontal momentum）（桌面光滑，所以沒有水平方向的外力）和能量（所有表面都是光滑的）。這兩條定律，加上質點相對於楔塊始終沿斜面 AC 運動這一事實，就確定了所有速度。之後的反彈和飛行只是拋體運動（projectile motion），而楔塊則以恆定速率滑行。

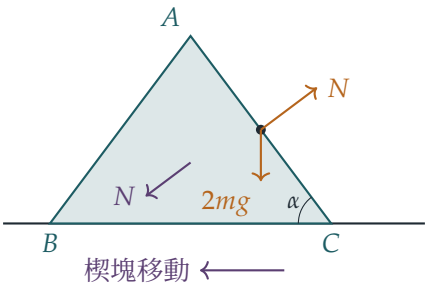
解答

建立模型

把質點視為質量為 $2m$ 的點質量（point mass），並把楔塊（wedge）視為質量為 m 的剛體（rigid body）。所有接觸面都是光滑的。在通過 A、B、C 的鉛垂截面上考慮問題。取 B 為原點，$C = (6a, 0)$，$A = (3a, 4a)$。A 的高度是 $4a$，因為三角形是等腰三角形，$AB = AC = 5a$，而底邊的一半是 $3a$。設 α 為角 ACB。則

$$\sin\alpha = \frac{4}{5}, \qquad \cos\alpha = \frac{3}{5}.$$

質點沿 AC 下滑，向 x 軸正方向運動，而楔塊則向 x 軸負方向後退。



設 u 為質點相對於楔塊的速率，方向沿 AC 向下。設 V 為楔塊的速度，其方向是水平的。質點的速度等於楔塊的速度加上質點相對於楔塊的速度：

$$\mathbf{v}_P = \big(V + u\cos\alpha,\ -u\sin\alpha\big) = \big(V + \tfrac{3}{5}u,\ -\tfrac{4}{5}u\big).$$

下滑期間的動量與能量

外力只有重力和桌面的鉛垂反作用力，所以水平動量守恆。它起初為零，因此

$$2m\big(V + \tfrac{3}{5}u\big) + mV = 0 \quad\Longrightarrow\quad V = -\tfrac{2}{5}u.$$

所以質點的速度是 $\big(\frac{1}{5}u,\ -\frac{4}{5}u\big)$，楔塊的速度是 $-\frac{2}{5}u$。

想一想 為甚麼不能單獨對質點使用能量守恆？

作用於質點的法向反作用力（normal reaction）垂直於斜面 AC，但斜面本身正在移動。因此質點的實際速度並不沿 AC 方向，而 N 對質點作功。只有整個系統的能量才守恆，因為這一對法向力所作的總功為零。

質點到達 AC 底部時，已下降了高度 $4a$。對整個系統應用能量守恆，得

$$\begin{aligned}\tfrac{1}{2}(2m)\Big(\tfrac{1}{25} + \tfrac{16}{25}\Big)u^2 + \tfrac{1}{2}m\cdot\tfrac{4}{25}u^2 &= 2mg\cdot 4a\\ \tfrac{19}{25}u^2 &= 8ga\\ u^2 &= \frac{200ga}{19}.\end{aligned}$$

質點自始至終都留在斜面上。按加速度分解各力，可知法向反作用力是常數 $N = \frac{10}{19}mg > 0$。

碰撞時各物體的位置

動量始終為零，所以系統的質心（centre of mass）沒有水平運動。設質點到達桌面時，質點的水平位移為 x_P，楔塊的水平位移為 x_W。則 $2m\,x_P + m\,x_W = 0$。相對於楔塊，質

點由 A 走到 C，水平距離為 $3a$，所以 $x_P - x_W = 3a$。因此

$$x_P = a, \qquad x_W = -2a.$$

在該瞬間，質點位於 $x = 3a + a = 4a$，而頂角 C 位於 $6a - 2a = 4a$。兩者在同一位置，這是必然的，因為質點剛好到達 C。

反彈與飛行

桌面固定而且光滑，而碰撞是完全彈性（perfectly elastic）的。所以質點的水平速度不變，鉛垂速度則反向。它以速度 $\left(\frac{1}{5}u,\ \frac{4}{5}u\right)$ 離開桌面。它返回桌面所需的飛行時間為

$$T = \frac{2 \cdot \frac{4}{5}u}{g} = \frac{8u}{5g}.$$

飛行期間，它向右水平移動了距離 $\frac{1}{5}u \cdot T = \dfrac{8u^2}{25g}$。

質點離開楔塊之後，楔塊不再受水平力作用，所以它繼續以速率 $\frac{2}{5}u$ 向左運動。在時間 T 內，它移動了 $\frac{2}{5}u \cdot T = \dfrac{16u^2}{25g}$。質點向右運動而楔塊向左運動，所以質點不可能再與楔塊相遇。

D 與 E 之間的距離是這兩段距離之和，因為兩者都由 $x = 4a$ 出發，並向相反方向運動：

$$DE = \frac{8u^2}{25g} + \frac{16u^2}{25g} = \frac{24u^2}{25g} = \frac{24}{25} \cdot \frac{200a}{19} = \frac{192}{19}\,a,$$

正如所求。

驗算

- 碰撞時的總動能：$\frac{19}{25}mu^2 = \frac{19}{25} \cdot \frac{200}{19}mga = 8mga$。這等於所失去的位能。
- 動量：$2m \cdot \frac{1}{5}u + m \cdot \left(-\frac{2}{5}u\right) = 0$。
- 取 $a = g = m = 1$，以數值方法逐步計算運動，得 $DE \approx 10.1053$，而 $\frac{192}{19} = 10.10526\ldots$。

常見錯誤

- 把質點的速度當作沿 AC 方向。那只是它*相對於*移動中的楔塊的速度。
- 單獨對質點寫出 $\frac{1}{2}(2m)u^2 = 2mg \cdot 4a$。這忽略了楔塊的動能和 N 所作的功。
- 由 C 的原來位置量度 DE，而不是由碰撞點 $x = 4a$ 量度；在那一刻，質點和 C 都位於該處。

STEP 之外

在光滑桌面上滑動的楔塊，是帶有約束（constraints）的系統的經典例子。拉格朗日力學（Lagrangian mechanics）處理這類系統的方法，是選取適當的坐標（這裏是楔塊的位置和質點沿斜面下滑的距離），使約束力根本無須出現。這樣，水平動量守恆便源於問題在水平平移下的對稱性，這是諾特定理（Noether's theorem）的一個例子。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）涵蓋了運動參考系，以及多體系統中的動量守恆。

詳細解答

STEP III 1990，第 13 題

力學 • 球內的運動

> **STEP III 1990，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/90-s3-q13

關鍵思路

整個過程分為兩個階段：先是在球面上作圓周運動（circular motion），直至法向反作用力（normal reaction）消失為止；然後是作為拋體（projectile）的自由飛行。離開球面的條件決定了離開一刻的速率，並可用 θ 表示；其後只需依照提示去做：把拋體重返球面的條件 $x^2 + y^2 = a^2$ 寫成關於 t 的方程。由於離開條件令其中幾項互相抵銷，這條方程會化簡得非常漂亮。

解答

模型

我們把 P 視為質點（particle），並把球面（半徑 a）視為光滑（smooth），因此 P 與球面接觸時，所受的力只有它的重量 mg 和指向 O 的法向反作用力 R。運動在通過 O 的一個鉛垂平面內進行；我們取 O 為原點，x 軸為水平方向，y 軸鉛垂向上，並設 P 沿球面的右邊向上爬升。

P 若要離開球面，就必定是在高於 O 的水平之處離開：在該水平以下，重力沿 PO 的分量背向 O，所以當 P 仍在圓上運動時，R 在那裏不可能為零。因此 $0 < \theta < \frac{\pi}{2}$。

離開球面

設 P 的速率為 v 時，OP 與向上的鉛垂線所成的角為 θ，於是 P 位於 $a\cos\theta$ 的高度（相對於 O 而言）。沿指向 O 的方向分解，重量在該方向的分量為 $mg\cos\theta$，所以

$$R + mg\cos\theta = \frac{mv^2}{a}.$$

P 在 $R = 0$ 時離開球面，所以在該刻

$$v^2 = ag\cos\theta.$$

（由能量守恆（conservation of energy）可知，拋射速率必須是 u，其中 $u^2 = v^2 + 2ga(1+\cos\theta) = ag(2+3\cos\theta)$，所以只要適當選取 u，$(0, \frac{\pi}{2})$ 內的每一個 θ 都可以達到。隨着 P 向上爬升，v 減小，R 亦減小，所以 R 正是在這一點首次變為零。）

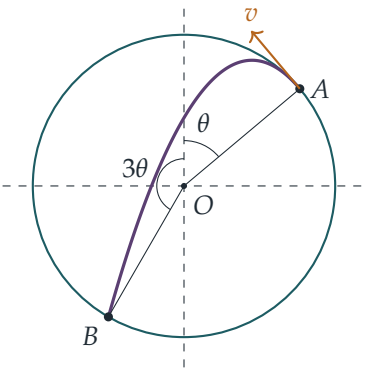
自由飛行

把離開的一刻記為 $t = 0$，此時 P 位於 $(a\sin\theta, a\cos\theta)$，其速度沿切線（tangent）方向，與 OP 垂直，並指向上方、越過球頂：

$$\text{速度} = v(-\cos\theta,\ \sin\theta).$$

此後只有重力作用，所以

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



P 在 A 離開球面，並且（如下文所證）在 B 再次撞上球面。

P 位於球面上的條件是 $x^2 + y^2 = a^2$。展開後，關於 t 的常數項和一次項分別是

$$a^2\sin^2\theta + a^2\cos^2\theta = a^2 \qquad \text{及} \qquad -2avt\sin\theta\cos\theta + 2avt\cos\theta\sin\theta = 0,$$

所以

$$\begin{aligned} x^2 + y^2 - a^2 &= v^2t^2\cos^2\theta + v^2t^2\sin^2\theta - gt^2\left(a\cos\theta + vt\sin\theta\right) + \tfrac{1}{4}g^2t^4 \\ &= \left(v^2 - ag\cos\theta\right)t^2 - gv\sin\theta\, t^3 + \tfrac{1}{4}g^2t^4 \\ &= gt^3\left(\tfrac{1}{4}gt - v\sin\theta\right), \end{aligned}$$

最後一步用了 $v^2 = ag\cos\theta$。

想一想 為甚麼 $t = 0$ 是三重根（triple root），而不只是單根？

在離開的一刻，單靠重力就剛好提供了沿圓運動所需、指向 O 的加速度 v^2/a。因此拋物線起初不但與球面有相同的方向，而且有相同的曲率（curvature）：它先緊貼着球面，然後才向球內偏離。這就是為甚麼除了 t 項之外，連 t^2 項也抵銷了。

當 $0 < t < 4v\sin\theta/g$ 時，上式為負，所以 P 嚴格位於球內，而它下一次與球面相遇的時刻是

$$T = \frac{4v\sin\theta}{g}.$$

落點位置

在 $t = T$ 時，利用 $vT\sin\theta = 4v^2\sin^2\theta/g = 4a\cos\theta\sin^2\theta$ 和 $\frac{1}{2}gT^2 = 8v^2\sin^2\theta/g = 8a\cos\theta\sin^2\theta$，可得

$$y = a\cos\theta + 4a\cos\theta\sin^2\theta - 8a\cos\theta\sin^2\theta = a\cos\theta\left(1 - 4\sin^2\theta\right).$$

（這就是 $a\cos 3\theta$，因為 $\cos 3\theta = 4\cos^3\theta - 3\cos\theta = \cos\theta(1 - 4\sin^2\theta)$；同樣地 $x = -a\sin 3\theta$，所以撞擊點 B 位於另一邊、與向上的鉛垂線成 3θ 角之處，正如圖中所示。）

P 撞在低於 O 的水平之處，當且僅當 $y < 0$。由於 $0 < \theta < \frac{\pi}{2}$，故 $\cos\theta > 0$，所以

$$y < 0 \iff 1 - 4\sin^2\theta < 0 \iff \sin\theta > \tfrac{1}{2} \iff \theta > \tfrac{\pi}{6},$$

最後一步成立，是因為 sin 在 $(0, \frac{\pi}{2})$ 上遞增。當 $\theta = \frac{\pi}{6}$ 恰好成立時，P 撞上球面的位置恰好與 O 同一水平。所以 θ 必須超過的角是

$$\boxed{\theta = \frac{\pi}{6} = 30^\circ}$$

而每一個 θ 只要滿足 $30^\circ < \theta < 90^\circ$，撞擊點確實都低於 O 的水平。

驗算

- $\theta = 60^\circ$：公式給出 $y = a\cos 180^\circ = -a$，所以 P 落回最低點，即拋射點。若取 $\theta = 45^\circ$，它落在與向上的鉛垂線成 135° 之處；兩者都有 $y < 0$，與預期相符。
- 當 $\theta \to 0$ 時，$T \to 0$ 且 $B \to A$：一個在球頂僅僅離開球面的質點，會立即落回球面上。
- 以數值模擬（逐步計算球面上的運動直至 $R = 0$，然後計算拋體運動直至它與球面相遇）檢驗 $\theta = 10^\circ, 25^\circ, 29^\circ, 30.5^\circ, 31^\circ, 45^\circ, 60^\circ, 80^\circ$ 各情況，結果與 $T = 4v\sin\theta/g$ 和 $y = a\cos 3\theta$ 吻合至小數點後四位；例如 $\theta = 31^\circ$ 給出 $y/a = -0.0523$。

常見錯誤

- 弄錯速度的方向：速度與 OP 垂直，並且仍然向上、越過球頂，其分量為 $(-v\cos\theta, v\sin\theta)$，而不是 $(v\sin\theta, v\cos\theta)$。
- 忘記先代入 $v^2 = ag\cos\theta$ 才去解 $x^2 + y^2 = a^2$，結果得到一條棘手的四次方程（quartic），而不是簡潔的因式分解（factorisation）。
- 把 $1 - 4\sin^2\theta < 0$ 兩邊相除時，沒有指出 $\cos\theta > 0$（這是因為 P 在高於 O 的水平之處離開）。

STEP 之外

$t = 0$ 處的三重根，其實是關於密切（osculation）的一個陳述：拋物線與圓在離開點有相同的切線和相同的曲率，所以兩者有二階切觸。曲率和密切圓（osculating circle）是曲線的微分幾何（differential geometry）中的核心概念；在力學中，每當我們把加速度分解為切向和法向兩部分時，它們都會再次出現。Kleppner 與 Kolenkow 所著的 *An Introduction to Mechanics*（力學入門教科書）對極坐標（polar coordinates）下的運動，以及這類受約束的運動，都有細緻的論述。

詳細解答

STEP III 1990，第 14 題

力學 • 力系

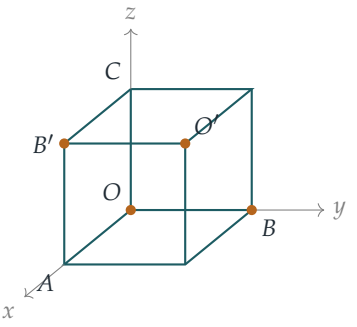
> **STEP III 1990，第 14 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/90-s3-q14

關鍵思路

作用於剛體（rigid body）上的力系，歸根究底只由兩個向量決定：合力（resultant）$\mathbf{R} = \sum \mathbf{F}_i$，以及對某一選定點的力矩（moment）$\mathbf{G}_P = \sum(\mathbf{r}_i - \mathbf{p}) \times \mathbf{F}_i$。把基點由 O 移到 P，力矩按一條簡單的規則改變：$\mathbf{G}_P = \mathbf{G}_O - \mathbf{p} \times \mathbf{R}$。本題每一部分，都是把這條規則用於不同的 P。

解答

把正方體視為邊長為 a 的剛體，於是 $A = (a, 0, 0)$，$B = (0, a, 0)$，$C = (0, 0, a)$，而與 B 相對的頂點是 $B' = (a, 0, a)$。



求 λ、μ、ν

作用於 O 的力對 O 沒有力矩，所以只需考慮其餘三個力。利用 $(x,y,z)\times(p,q,r) = (yr-zq,\ zp-xr,\ xq-yp)$：

$$\overrightarrow{OO'}\times\begin{pmatrix}\lambda\\0\\1\end{pmatrix} = a\begin{pmatrix}1\\1\\1\end{pmatrix}\times\begin{pmatrix}\lambda\\0\\1\end{pmatrix} = a\begin{pmatrix}1\\\lambda-1\\-\lambda\end{pmatrix},$$

$$\overrightarrow{OB}\times\begin{pmatrix}-1\\0\\2\end{pmatrix} = a\begin{pmatrix}0\\1\\0\end{pmatrix}\times\begin{pmatrix}-1\\0\\2\end{pmatrix} = a\begin{pmatrix}2\\0\\1\end{pmatrix},$$

$$\overrightarrow{OB'}\times\begin{pmatrix}1\\\mu\\\nu\end{pmatrix} = a\begin{pmatrix}1\\0\\1\end{pmatrix}\times\begin{pmatrix}1\\\mu\\\nu\end{pmatrix} = a\begin{pmatrix}-\mu\\1-\nu\\\mu\end{pmatrix}.$$

相加後，對 O 的總力矩為 $a\,(3-\mu,\ \lambda-\nu,\ 1-\lambda+\mu)$。令每個分量為零，得 $\mu=3$，繼而 $\lambda=1+\mu=4$，再得 $\nu=\lambda=4$：

$$\lambda=4,\qquad \mu=3,\qquad \nu=4.$$

以上結果完全不取決於 α,β,γ，而合力為

$$\mathbf{R}=\begin{pmatrix}\alpha+\lambda-1+1\\\beta+\mu\\\gamma+1+2+\nu\end{pmatrix}=\begin{pmatrix}\alpha+4\\\beta+3\\\gamma+7\end{pmatrix},\qquad \mathbf{G}_O=\mathbf{0}.$$

第 (i) 部分

當 $\alpha=\beta=\gamma=0$ 時，$\mathbf{R}=(4,3,7)$。兩個力系等效（equivalent），是指它們有相同的合力，並且對某一點（因而對每一點）有相同的力矩。所以作用於 B 的單一力必定就是 $\mathbf{R}$ 本身，而力偶（couple）必須等於原力系對 B 的力矩：

$$\mathbf{G}_B=\mathbf{G}_O-\overrightarrow{OB}\times\mathbf{R}=\mathbf{0}-a\begin{pmatrix}0\\1\\0\end{pmatrix}\times\begin{pmatrix}4\\3\\7\end{pmatrix}=-a\begin{pmatrix}7\\0\\-4\end{pmatrix}=a\begin{pmatrix}-7\\0\\4\end{pmatrix}.$$

? **想一想** 為甚麼對 B 的力矩等於 $\mathbf{G}_O-\overrightarrow{OB}\times\mathbf{R}$？

對每一個力，$(\mathbf{r}_i-\mathbf{b})\times\mathbf{F}_i=\mathbf{r}_i\times\mathbf{F}_i-\mathbf{b}\times\mathbf{F}_i$；對 i 求和便得 $\mathbf{G}_O-\mathbf{b}\times\mathbf{R}$。

因此，等效力系是一個作用於 B 的力 $(4,3,7)$，連同一個力矩為 $a(-7,0,4)$ 的力偶。驗算一下：作用於 B 的力加上這個力偶，對 O 的力矩等於 $a(7,0,-4)+a(-7,0,4)=\mathbf{0}$，正如所料。

第 (ii) 部分

現在 $\mathbf{R} = (2+4,\ 3+3,\ 2+7) = (6,6,9) = 3(2,2,3)$，而 $\mathbf{G}_O = \mathbf{0}$ 依然成立。在位置向量（position vector）為 $\mathbf{r} = (x,y,z)$ 的點 P，

$$\mathbf{G}_P = \mathbf{G}_O - \mathbf{r}\times\mathbf{R} = -\mathbf{r}\times\mathbf{R}.$$

由於 $\mathbf{R}\neq\mathbf{0}$，上式為零當且僅當 $\mathbf{r}$ 與 $\mathbf{R}$ 平行（或 $\mathbf{r}=\mathbf{0}$）。所以所求軌跡（locus）是通過 O、方向為 $(2,2,3)$ 的直線：

$$\frac{x}{2} = \frac{y}{2} = \frac{z}{3}.$$

從物理上看，整個力系等效於沿這條直線作用的單一力 $(6,6,9)$。

當正方體平移（不帶轉動）時，每一點的位移都相同，所以各力所作的功（work done）是 $\mathbf{R}\cdot\mathbf{s}$，其中 $\mathbf{s}$ 是位移。若沿該直線、朝 $(2,2,3)$ 的方向移動距離 d，則 $\mathbf{s} = \dfrac{d}{\sqrt{17}}(2,2,3)$，而

$$W = (6,6,9)\cdot\frac{d}{\sqrt{17}}(2,2,3) = \frac{51d}{\sqrt{17}} = 3\sqrt{17}\,d.$$

題目沒有為這段距離命名；每單位距離所作的功是 $3\sqrt{17}\approx 12.37$ 個單位。由於 $\mathbf{R}$ 沿着這條直線，這其實就是 $|\mathbf{R}|\,d = \sqrt{36+36+81}\,d = \sqrt{153}\,d$，數值相同。（若沿直線反方向移動，則得 $-3\sqrt{17}\,d$。）

驗算

- 取 $\lambda = 4$、$\mu = 3$、$\nu = 4$，三個對 O 的力矩分別是 $a(1,3,-4)$、$a(2,0,1)$ 和 $a(-3,-3,3)$，其和為零。
- 在 (ii) 中，取 $P=(2,2,3)$：$\mathbf{r}\times\mathbf{R} = (2,2,3)\times(6,6,9) = \mathbf{0}$。在 $P=(1,0,0)$，力矩為 $-(0,-9,6)\neq\mathbf{0}$。
- 以電腦用精確分數運算，核實了每一個力矩，檢查了 (i) 中的等效力系對 20 個隨機點的力矩均與原力系相同，並確認每單位距離所作的功為 $3\sqrt{17}\approx 12.369$。

常見錯誤

- 把 $\mathbf{r}\times\mathbf{F}$ 的次序弄錯：次序顛倒，所有正負號都會反轉。
- 把 (i) 中的力偶當作 $+\overrightarrow{OB}\times\mathbf{R}$。把力由 O 移到 B，所需的補償力偶帶相反的符號。
- 逐一計算作用於不同頂點的各個力所作的功再相加。在純平移中，只有合力才起作用。

STEP 之外

作用於剛體上的任何力系，都可以化簡為沿某條直線的單一力，連同一個繞同一條直線的力偶。這稱為*力螺旋*（wrench），其螺距（pitch）為 $\mathbf{R}\cdot\mathbf{G}/|\mathbf{R}|^2$。在 (ii) 中螺距為零，所以力螺旋只是一個單一力，而零力矩軌跡就是它的作用線（line of action），即中心軸（central axis）。這些概念通往剛體動力學（rigid-body dynamics），以及機械人學中的旋量理論（screw theory）。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）以恰當的程度講解力矩和剛體。

詳細解答

STEP III 1990，第 15 題

概率與統計 • 幾何分佈

STEP III 1990，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/90-s3-q15

關鍵思路

每次投擲出現 A 的概率（probability）是 $\frac{1}{4}$，出現 B 的概率是 $\frac{3}{4}$，而且各次投擲互相獨立（independent）。因此事件 $\{M = m,\ N = n\}$ 就是某一串特定的投擲結果，其概率是一個乘積。這個乘積可以拆成只與 m 有關的因子乘以只與 n 有關的因子，而這一點幾乎包辦了所有工作：邊際分佈（marginal distribution）是幾何分佈（geometric distribution），兩個變量互相獨立，其後每一個概率都是一個等比級數（geometric series）。

解答

設單次投擲中 $p = \mathrm{P}(A) = \frac{3}{12} = \frac{1}{4}$，$q = \mathrm{P}(B) = \frac{9}{12} = \frac{3}{4}$。各次投擲互相獨立。

事件 $\{M = m,\ N = n\}$ 發生，當且僅當最初的 $m + n + 1$ 次投擲為

$$\underbrace{A\,A\cdots A}_{m}\ \underbrace{B\,B\cdots B}_{n}\ A,$$

所以根據獨立性，對於 $m \geq 0$ 及 $n \geq 1$，

$$\mathrm{P}(M = m,\ N = n) = \left(\tfrac{1}{4}\right)^m \left(\tfrac{3}{4}\right)^n \tfrac{1}{4} = \left(\tfrac{1}{4}\right)^{m+1} \left(\tfrac{3}{4}\right)^n.$$

（這串投擲結果以概率 1 終會出現一個 B，然後再出現一個 A，所以這些概率之和為 1；下面的邊際分佈可以印證這一點。）

第 (i) 部分

求 M 的邊際分佈時，對 $n \geq 1$ 求和，並利用 $\sum_{n\geq 1}\left(\frac{3}{4}\right)^n = \dfrac{3/4}{1-3/4} = 3$：

$$\mathrm{P}(M=m) = \left(\tfrac{1}{4}\right)^{m+1}\cdot 3 = \tfrac{3}{4}\left(\tfrac{1}{4}\right)^{m}, \qquad m = 0,1,2,\ldots$$

求 N 的邊際分佈時，對 $m \geq 0$ 求和，並利用 $\sum_{m\geq 0}\left(\frac{1}{4}\right)^{m+1} = \dfrac{1/4}{1-1/4} = \frac{1}{3}$：

$$\mathrm{P}(N=n) = \tfrac{1}{3}\left(\tfrac{3}{4}\right)^{n} = \tfrac{1}{4}\left(\tfrac{3}{4}\right)^{n-1}, \qquad n = 1,2,3,\ldots$$

兩個結果都可以直接理解：M 是第一個 B 出現之前 A 的個數，而 N 則是一連串 B 開始之後，等待第一個 A 出現所需的次數。兩者各自的總和都是 1。

第 (ii) 部分

我們利用 $\sum_{k\geq 1} kx^{k-1} = \dfrac{1}{(1-x)^2}$，其中 $|x|<1$（把等比級數求導即得）。於是

$$\mathrm{E}(M) = \tfrac{3}{4}\sum_{m\geq 1} m\left(\tfrac{1}{4}\right)^{m} = \tfrac{3}{4}\cdot\tfrac{1}{4}\cdot\frac{1}{(3/4)^2} = \tfrac{1}{3},$$

$$\mathrm{E}(N) = \tfrac{1}{4}\sum_{n\geq 1} n\left(\tfrac{3}{4}\right)^{n-1} = \tfrac{1}{4}\cdot\frac{1}{(1/4)^2} = 4.$$

獨立性

把兩個邊際分佈相乘，

$$\mathrm{P}(M=m)\,\mathrm{P}(N=n) = \tfrac{3}{4}\left(\tfrac{1}{4}\right)^{m}\cdot\tfrac{1}{4}\left(\tfrac{3}{4}\right)^{n-1} = \left(\tfrac{1}{4}\right)^{m+1}\left(\tfrac{3}{4}\right)^{n} = \mathrm{P}(M=m,\ N=n)$$

對所有 $m \geq 0$、$n \geq 1$ 都成立。所以 M 與 N 確實互相獨立。

> **想一想** 一連串 A 的長度，難道不會影響一連串 B 的長度嗎？為甚麼不會？

一連串 A 在第一個 B 出現時結束，而由該次投擲起，整個過程重新開始：餘下的投擲與之前發生的一切互相獨立。一連串 B 的長度只取決於這些較後的投擲，所以知道 M 的值，並不能告訴我們任何關於 N 的資料。聯合概率（joint probability）可以分解為兩個因子，正是這個事實的代數形式。

涉及 N 與 M 的概率

對於整數 $k \geq 1$，把幾何分佈的尾部求和：

$$\mathrm{P}(N>k) = \sum_{n\geq k+1}\tfrac{1}{4}\left(\tfrac{3}{4}\right)^{n-1} = \tfrac{1}{4}\cdot\frac{(3/4)^k}{1-3/4} = \left(\tfrac{3}{4}\right)^{k}.$$

這條公式同時給出 $\mathrm{P}(N>0)=1$，由於 $N\geq 1$，這是正確的。

求 $\mathrm{P}(N>M)$ 時，以 M 的值為條件，並利用獨立性：

$$\begin{aligned}\mathrm{P}(N>M) &= \sum_{m\geq 0}\mathrm{P}(M=m)\,\mathrm{P}(N>m) = \sum_{m\geq 0}\tfrac{3}{4}\left(\tfrac{1}{4}\right)^m\left(\tfrac{3}{4}\right)^m \\ &= \tfrac{3}{4}\sum_{m\geq 0}\left(\tfrac{3}{16}\right)^m = \tfrac{3}{4}\cdot\tfrac{16}{13} = \tfrac{12}{13}.\end{aligned}$$

至於 $\mathrm{P}(N=M)$，由於 $N\geq 1$，兩者的共同值必須至少為 1：

$$\begin{aligned}\mathrm{P}(N=M) &= \sum_{m\geq 1}\tfrac{3}{4}\left(\tfrac{1}{4}\right)^m\cdot\tfrac{1}{4}\left(\tfrac{3}{4}\right)^{m-1} = \tfrac{3}{64}\sum_{m\geq 1}\left(\tfrac{3}{16}\right)^{m-1} \\ &= \tfrac{3}{64}\cdot\tfrac{16}{13} = \tfrac{3}{52}.\end{aligned}$$

最後，$N>M$、$N=M$、$N<M$ 這三個事件互斥（exclusive）且窮盡（exhaustive）所有可能，所以

$$\mathrm{P}(N<M) = 1-\tfrac{12}{13}-\tfrac{3}{52} = \tfrac{52-48-3}{52} = \tfrac{1}{52},$$

正是所求。

驗算

- 兩個平均值都符合直覺：A 很少出現，所以開頭的一連串 A 很短（平均為 $\frac{1}{3}$），而我們平均要等 $1/p=4$ 次投擲，才有一個 A 出現來結束一連串 B。
- 直接計算，$\mathrm{P}(M>n)=\left(\frac{1}{4}\right)^{n+1}$，所以 $\mathrm{P}(N<M)=\sum_{n\geq 1}\frac{1}{4}\left(\frac{3}{4}\right)^{n-1}\left(\frac{1}{4}\right)^{n+1} = \frac{1}{64}\sum_{n\geq 1}\left(\frac{3}{16}\right)^{n-1} = \frac{1}{64}\cdot\frac{16}{13} = \frac{1}{52}$，與用減法所得的結果一致。$\mathrm{P}(N>k)=\left(\frac{3}{4}\right)^k$ 的直接解釋是：第一個 B 之後，接着的 k 次投擲必須全部是 B。
- 把聯合分佈精確求和，得 $\mathrm{P}(N>M)=0.92308$，$\mathrm{P}(N=M)=0.05769$，$\mathrm{P}(N<M)=0.01923$；模擬 200 000 組投擲，所得結果為 0.923、0.058、0.019，樣本平均值則為 0.334 和 3.99。

常見錯誤

- 在 $\mathrm{P}(N=M)$ 的求和中包括了 $m=0$：N 永遠不可能是 0。
- 忘記 M 可以是 0（第一次投擲可能就是 B），以致 M 的邊際分佈移了位，算出錯誤的平均值。

STEP 之外

這裏兩段連串的長度互相獨立，是因為一連串獨立試驗會在任何只由過去決定的時刻「重新開始」，例如第一個 B 出現的時刻。這就是強馬可夫性質（strong Markov property），它是更新理論（renewal theory）和馬可夫鏈（Markov chain）分析的基礎，而這些都是大學一、二年級的課題。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教科書）以恰到好處的程度講解幾何分佈、獨立性和馬可夫鏈。

詳細解答

STEP III 1990，第 16 題

概率與統計 • 幾何概率

STEP III 1990，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/90-s3-q16

本題在 *Beyond STEP* 第 12.21 節折斷一根棒中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/broken-stick.html

關鍵思路

三段總長固定的線段，恰好在每一段都小於總長的一半時，才可以組成三角形。三角形一旦存在，它是鈍角三角形（obtuse triangle）的充要條件，就是最長邊的平方大於另外兩邊的平方和（餘弦定理（cosine rule））。兩部分都可化為關於隨機變量（random variable）的條件，而所求的概率在 (i) 中變成區間的長度，在 (ii) 中變成區域的面積。

解答

第 (i) 部分

三段的長度分別是 X、$\frac{1}{2}(1-X)$ 和 $\frac{1}{2}(1-X)$，總長為 1。三個正數長度可以組成三角形，當且僅當每一個都小於另外兩個之和，即小於周界的一半，在這裏就是 $\frac{1}{2}$。只要 $X>0$，兩段相等的線段都滿足 $\frac{1}{2}(1-X)<\frac{1}{2}$，所以唯一的條件是 $X<\frac{1}{2}$。設 T 為可以組成三角形這個事件，則

$$\mathrm{P}(T)=\mathrm{P}\left(0<X<\tfrac{1}{2}\right)=\tfrac{1}{2}.$$

一個三角形是鈍角三角形，當且僅當最長邊的平方大於另外兩邊的平方和。哪一條邊最長，則取決於 X。

- 若 $X\le\frac{1}{3}$，則 $X\le\frac{1}{2}(1-X)$，所以兩段相等的線段之一是最長邊。這時要檢驗的條件是 $\frac{1}{4}(1-X)^2>\frac{1}{4}(1-X)^2+X^2$，即 $X^2<0$，這是不可能發生的。這個三角形是等

腰三角形（isosceles triangle），底邊不比兩條等邊長，所以它永遠不是鈍角三角形。

- 若 $\frac{1}{3} < X < \frac{1}{2}$，最長邊是 X，而三角形是鈍角三角形，當且僅當

$$X^2 > 2 \cdot \tfrac{1}{4}(1-X)^2 \iff X > \frac{1-X}{\sqrt{2}},$$

這裏取正平方根，因為兩邊都是正數。整理後得 $X(\sqrt{2}+1) > 1$，即 $X > \dfrac{1}{\sqrt{2}+1} = \sqrt{2}-1 \approx 0.414$，這個值確實落在 $\left(\frac{1}{3}, \frac{1}{2}\right)$ 之內。

設 O 為三角形是鈍角三角形這個事件，則 $O \cap T$ 就是事件 $\sqrt{2}-1 < X < \frac{1}{2}$。由於 X 服從均勻分佈（uniform distribution），

$$\mathrm{P}(O \mid T) = \frac{\mathrm{P}(\sqrt{2}-1 < X < \frac{1}{2})}{\mathrm{P}(T)} = \frac{\frac{1}{2} - (\sqrt{2}-1)}{\frac{1}{2}} = 1 - 2\sqrt{2} + 2 = 3 - 2\sqrt{2},$$

正如所求。

第 (ii) 部分

這個構作總是可行的。 以 $(1,0)$、$(1,1)$、$(0,1)$ 為頂點的三角形 R，就是滿足 $x \le 1$、$y \le 1$ 和 $x + y \ge 1$ 的點所組成的集合。題目給出的三條邊 x、y、$2-x-y$ 總長為 2，它們可以組成一個三角形（可能是退化（degenerate）的），當且僅當三者皆非負，而且沒有一個超過周界的一半，即 1：

- $0 \le x \le 1$ 和 $0 \le y \le 1$ 在 R 上成立；
- $2 - x - y \le 1$ 其實就是 $x + y \ge 1$，而 $2 - x - y \ge 0$ 可由 $x \le 1$、$y \le 1$ 推出。

所以 R 上的每一點都給出一個三角形。只有在 R 的邊上，三角形才是退化的，而這些邊的概率為零。

$\angle ABC$ 何時是鈍角？ 位於 B 的角夾在 BA 與 BC 之間，所以它的對邊是 $CA = y$。根據餘弦定理，$\angle ABC$ 是鈍角，當且僅當 $CA^2 > AB^2 + BC^2$：

$$y^2 > x^2 + (2-x-y)^2.$$

展開 $(2-x-y)^2 = (2-x)^2 - 2(2-x)y + y^2$，$y^2$ 項互相抵消，條件變成

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

由於 $x \le 1$，我們有 $2 - x > 0$，所以可以把兩邊除以 $2(2-x)$ 而不必改變不等號的方向：

$$y > \frac{x^2 - 2x + 2}{2 - x},$$

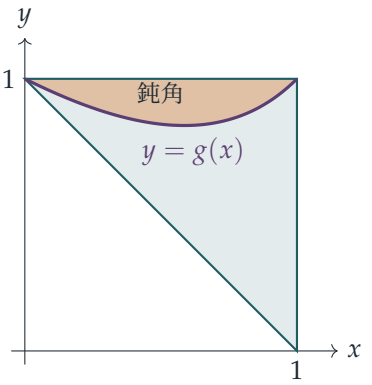
正如所求。把右邊記作 $g(x)$。

想一想 曲線 $y = g(x)$ 真的一直留在三角形 R 之內嗎？

我們需要它在 $0 \le x \le 1$ 時位於直線 $y = 1 - x$ 與 $y = 1$ 之間，這樣鈍角區域就正好是 R 在曲線上方的部分。事實上，

$$1 - g(x) = \frac{(2-x) - (x^2 - 2x + 2)}{2 - x} = \frac{x(1-x)}{2-x} \ge 0, \qquad g(x) - (1-x) = \frac{x}{2-x} \ge 0,$$

而且 $g(0) = g(1) = 1$。所以曲線由 $(0,1)$ 走到 $(1,1)$，中途降到 $y = 1$ 之下，但從不低於斜邊 $x + y = 1$。



概率。 三角形 R 的面積是 $\frac{1}{2}$，所以聯合密度（joint density）在 R 上等於 2，於是

$$\mathrm{P}(\angle ABC \text{ 為鈍角}) = 2 \times \text{面積}\{(x,y) : g(x) < y \le 1,\ 0 \le x \le 1\} = 2\int_0^1 \big(1 - g(x)\big)\,\mathrm{d}x.$$

要對 g 積分，可代入 $u = 2 - x$，於是 $x = 2 - u$，而

$$x^2 - 2x + 2 = (2-u)^2 - 2(2-u) + 2 = u^2 - 2u + 2, \qquad g = \frac{u^2 - 2u + 2}{u} = u - 2 + \frac{2}{u}.$$

當 x 由 0 變到 1 時，u 由 2 變到 1，而 $\mathrm{d}x = -\mathrm{d}u$，所以

$$\begin{aligned}\int_0^1 g(x)\,\mathrm{d}x &= \int_1^2 \left(u - 2 + \frac{2}{u}\right)\mathrm{d}u = \left[\tfrac{1}{2}u^2 - 2u + 2\ln u\right]_1^2 \\ &= (2 - 4 + 2\ln 2) - \left(\tfrac{1}{2} - 2\right) = 2\ln 2 - \tfrac{1}{2}.\end{aligned}$$

因此

$$\mathrm{P}(\angle ABC \text{ 為鈍角}) = 2\left(1 - 2\ln 2 + \tfrac{1}{2}\right) = 3 - 4\ln 2,$$

正如所求。

驗算

- 數值上，$3-2\sqrt{2}\approx 0.1716$，$3-4\ln 2\approx 0.2274$。以 10^6 次試驗進行蒙地卡羅模擬（Monte Carlo simulation），結果約為 0.1719 和 0.2278。
- 兩個答案都在 0 與 1 之間。在 (ii) 中，曲線在兩端都碰到 $y=1$：在 $x=0$ 時邊 BC 消失，在 $x=1$ 時邊 AB 消失，所以兩者都是退化情況，正如所料。
- 以辛普森法則（Simpson's rule）計算 $\int_0^1 g(x)\,\mathrm{d}x$，得 0.886294，與 $2\ln 2-\frac{1}{2}$ 吻合。

常見錯誤

- 在 (i) 中忘記取條件：出現鈍角三角形的無條件概率是 $\frac{3}{2}-\sqrt{2}$，只有所求答案的一半。
- 在 (ii) 中，餘弦定理用錯了邊。位於 B 的角的對邊是 $CA=y$，而不是 AB。
- 忘記 R 上的密度是 2 而不是 1，結果答案只有應有的一半。

STEP 之外

把概率化為長度或面積，就是幾何概率（geometric probability），而隨機折斷一根棒正是其中一個經典問題：在兩個獨立的均勻隨機點把棒折斷，可以組成三角形的機會是 $\frac{1}{4}$。第 (ii) 部分讓你初次接觸連續聯合分佈（continuous joint distribution），當中的概率是聯合密度在某個區域上的二重積分（double integral）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教科書）有講解聯合密度，並收錄了好幾道這類題目。

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

1990 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。