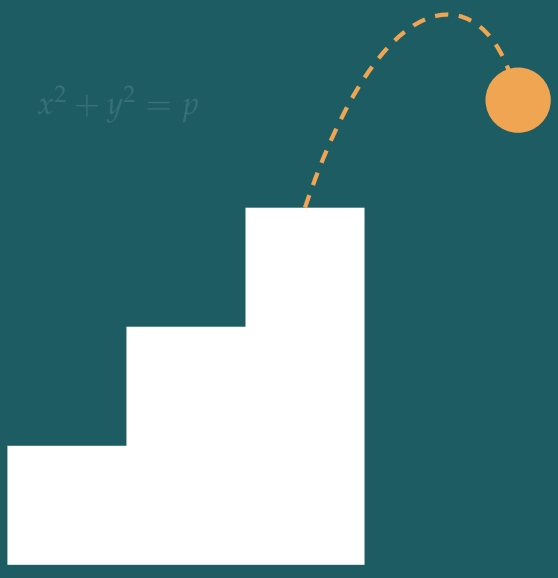
STEP 備試 · 大學數學

Beyond STEP

STEP I, II 及 III

1989 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

1989 年詳細解答

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

1989 年詳細解答

1989 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 組合區域的面積 純數學 3

第 2 題 黎曼和與極限 純數學 6

第 3 題 向量與純量積 純數學 10

第 4 題 拉姆齊理論 純數學 14

第 5 題 二項式係數恆等式 純數學 17

第 6 題 法線與微分方程 純數學 20

第 7 題 曲線描繪與最優化 純數學 23

第 8 題 倍角公式 純數學 27

第 9 題 圖像的變換 純數學 31

第 10 題 與速度相關的阻力 力學 36

第 11 題 拋體射程的最優化 力學 39

第 12 題 彈性繩與能量 力學 42

第 13 題 張及摩擦力的梯子平衡 力學 46

第 14 題 條件分佈 概率與統計 49

第 15 題 差佈的比較 概率與統計 53

第 16 題 變和博弈 概率與統計 56

STEP II

第 1 題 三次方程的三角解法 純數學 60

第 2 題 冪級數的係數 純數學 64

第 3 題 複雙曲函數映射 純數學 67

第 4 題 有理函數的描繪 純數學 71

第 5 題 極坐標曲線與反射 純數學 76

第 6 題 反函數的積分 純數學 80

第 7 題 二階微分方程 純數學 84
第 8 題 變換群 純數學 87
第 9 題 矩陣指數與方程組 純數學 91
第 10 題 線與平面 純數學 94
第 11 題 連接質點與衝量 力學 98
第 12 題 剛體平衡 力學 102
第 13 題 轉動慣量 力學 106
第 14 題 圓周運動 力學 111
第 15 題 幾何概率 概率與統計 114
第 16 題 項分佈 概率與統計 118

STEP III

第 1 題 旋轉曲面 純數學 122
第 2 題 三維空間中的平面 純數學 126
第 3 題 矩陣群 純數學 129
第 4 題 曲線描繪與根 純數學 133
第 5 題 切比雪夫多項式 純數學 138
第 6 題 圓錐曲線的極坐標方程 純數學 142
第 7 題 錯切矩陣 純數學 146
第 8 題 耦合微分方程 純數學 150
第 9 題 冪級數求和 純數學 153
第 10 題 黎曼和 純數學 156
第 11 題 球體的斜碰撞 力學 160
第 12 題 旋轉參考系 力學 164
第 13 題 簡諧運動 力學 168
第 14 題 飛輪與衝量 力學 172
第 15 題 順位數的分佈 概率與統計 175
第 16 題 概率與假設檢驗 概率與統計 179

詳細解答

STEP I 1989，第 1 題

純數學 • 組合區域的面積

> **STEP I 1989，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/89-s1-q1

關鍵思路

這個區域由半圓（semicircle）組成。先取一個大的半圓盤（half-disc），再減去和加上較小的半圓盤，就可以得到它。所以我們只需要知道每個半徑，尤其是下方半圓 CDE 的半徑，而這可以由直線 AG 上的各段長度求得。把半圓換成拋物線（parabola），方法完全不變，因為拋物線拱（parabolic arch）有一條簡單的面積公式，可以取代 $\frac{1}{2}\pi a^2$ 的角色。

解答

半圓

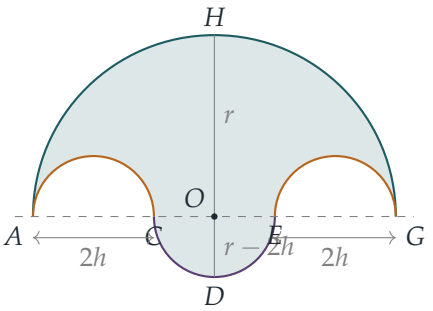
從 A 起沿直線量度。由於 ABC 的半徑是 h，所以 $AC = 2h$。同理 $EG = 2h$，而 $AG = 2r$，因為 AHG 的半徑是 r。於是

$$CE = 2r - 4h,$$

由於 $2h < r$，這個值是正數。所以半圓 CDE 的半徑是 $r - 2h$。它的圓心是 CE 的中點，而根據對稱性，這也是 AG 的中點，所以它與大半圓有共同的圓心 O。

從圖中可見，ABC 和 EFG 與 H 位於 AG 的同一側，並且在大半圓之內，而 CDE 則位於另一側。因此所圍成的區域是

$$(\text{半圓盤}AHG) \ - \ (\text{半圓盤}ABC \text{ 和}EFG) \ + \ (\text{半圓盤}CDE).$$



半徑為 a 的半圓盤面積是 $\frac{1}{2}\pi a^2$，所以所圍成的面積是

$$\begin{aligned}\tfrac{1}{2}\pi r^2 - 2\cdot\tfrac{1}{2}\pi h^2 + \tfrac{1}{2}\pi(r-2h)^2 &= \tfrac{1}{2}\pi(r^2 - 2h^2 + r^2 - 4rh + 4h^2)\\ &= \tfrac{1}{2}\pi(2r^2 - 4rh + 2h^2)\\ &= \pi(r-h)^2.\end{aligned}$$

? 想一想 HD 有多長？

點 H 是大半圓的最高點，所以它位於通過 O 而垂直於 AG 的直線上，與 O 相距 r。點 D 是 CDE 的最低點，所以它也在同一條垂直線上，位於直線的另一側，與 O 相距 $r-2h$。所以

$$HD = r + (r-2h) = 2(r-h),$$

而以 HD 為直徑的圓，半徑是 $r-h$，面積是 $\pi(r-h)^2$。這正是我們求得的面積，符合題目要求。

拋物線

每個半徑為 a 的半圓，都換成一個拋物線拱：它建在同一條長度為 $2a$ 的弦（chord）上，頂點（vertex）在弦的中點上方，高度為 a。把原點設在弦的中點。拋物線關於通過頂點的軸對稱，所以它的形式是 $y = a - kx^2$。它通過 $(\pm a, 0)$，所以 $k = 1/a$。拱與弦之間的面積是

$$\int_{-a}^{a}\left(a - \frac{x^2}{a}\right)\mathrm{d}x = 2a^2 - \frac{2a^3}{3a} = \frac{4a^2}{3}.$$

（這就是阿基米德（Archimedes）的結果：外接矩形面積的三分之二，即 $\frac{2}{3}\cdot 2a\cdot a$。）

每個拱仍然位於 AG 的同一側，與原來的半圓一樣。我們還應該檢查每個小拱是否仍然在大拱之下。兩個拱都從 A 出發，斜率（gradient）同為 2，所以在 A 附近這一點並不明顯。在距離 A 為 x 之處，其中 $0 \le x \le 2h$，大拱的高度是 $x(2r-x)/r$，而拱 ABC 的高度是 $x(2h-x)/h$。兩者之差是

$$\frac{x(2r-x)}{r} - \frac{x(2h-x)}{h} = x^2\left(\frac{1}{h} - \frac{1}{r}\right) \ge 0,$$

所以 ABC 確實在 AHG 之內，而根據對稱性，EFG 也是如此。因此同樣的加減分解依然適用，只需以 $\frac{4}{3}a^2$ 代替 $\frac{1}{2}\pi a^2$：

$$\begin{aligned}\tfrac{4}{3}r^2 - 2\cdot\tfrac{4}{3}h^2 + \tfrac{4}{3}(r-2h)^2 &= \tfrac{4}{3}(2r^2 - 4rh + 2h^2)\\ &= \tfrac{8}{3}(r-h)^2.\end{aligned}$$

所以換成拋物線後，所圍成的面積是 $\frac{8}{3}(r-h)^2$。

驗算

- 兩個答案都是一個常數乘以 $(r-h)^2$。這是必然的，因為兩個計算的結構完全相同：只有單個拱的係數改變了，由 $\frac{\pi}{2}$ 變成 $\frac{4}{3}$。比值 $\frac{8/3}{\pi} \approx 0.849$ 小於 1，這是合理的，因為弦和高度都相同的拋物線拱位於半圓之內。
- 極限情況 $h \to 0$：小拱消失，而 CDE 變成一個半徑為 r 的完整下半圓，合起來是一個面積為 πr^2 的完整圓盤。公式 $\pi(r-h)^2$ 與此吻合。
- 數值上，取 $r = 3$ 和 $h = 1$（即試卷圖中的比例），面積分別是 $4\pi \approx 12.566$ 和 $\frac{32}{3} \approx 10.667$。對描出的邊界以鞋帶公式（shoelace formula）計算，取八組隨機的 (r,h)，結果與兩條公式都吻合，誤差在離散化誤差範圍之內。

常見錯誤

- 把 CDE 的半徑當作 $r-h$ 或 $r-4h$。把 AG 上的各段長度畫出來：$CE = 2r - 4h$，所以它的半徑是 $r-2h$。
- 把上方的小半圓加上去，而不是減去。它們是從大半圓盤中切走的部分，所以其面積要減掉。

STEP 之外

第一部分的圖形是*鞋匠刀形*（arbelos，即「鞋匠的刀」），是阿基米德研究過的一種形狀。它的經典形式是一個大半圓，內有兩個較小的半圓，而它的面積同樣等於以某條精心選取的弦為直徑的圓的面積。阿基米德還發現了拋物線弓形的 $\frac{2}{3}$ 法則，所用的極限論證已經很接近積分。在大學裏，從他的「窮竭法」（method of exhaustion）走到黎曼積分（Riemann integral）的那一步，是第一門分析課程的核心內容。想了解這一步，Abbott 的 *Understanding Analysis*（分析學入門教材）是很好的選擇。

詳細解答

STEP I 1989，第 2 題

純數學 • 黎曼和與極限

STEP I 1989，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/89-s1-q2

關鍵思路

這個古怪的式子，其實只是一些長方形面積之和的偽裝。如果直接寫出 n 個長方形的總面積，會得到一個含有 $k\ln k$ 的和；那些階乘（factorial）的對數，不過是把同一個和重新分組而已。只要把兩者對上，極限（limit）就是那個積分，而第一部分正好讓我們把它算出來。

解答

不定積分

用分部積分法（integration by parts），取 $u=\ln x$ 和 $\frac{\mathrm{d}v}{\mathrm{d}x}=x$，於是 $\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{1}{x}$ 而 $v=\frac{1}{2}x^2$，

$$\begin{aligned}\int x\ln x\,\mathrm{d}x &= \tfrac{1}{2}x^2\ln x-\int \tfrac{1}{2}x^2\cdot\frac{1}{x}\,\mathrm{d}x\\ &= \tfrac{1}{2}x^2\ln x-\tfrac{1}{4}x^2+C.\end{aligned}$$

積分的值

由於 $\ln(1/x)=-\ln x$，被積函數（integrand）是 $-x\ln x$。被積函數在 $x=0$ 處沒有定義，所以我們先在 $[\varepsilon,1]$ 上積分，再令 $\varepsilon\to 0^+$：

$$\int_\varepsilon^1 x\ln(1/x)\,\mathrm{d}x=\left[\tfrac{1}{4}x^2-\tfrac{1}{2}x^2\ln x\right]_\varepsilon^1=\tfrac{1}{4}-\tfrac{1}{4}\varepsilon^2+\tfrac{1}{2}\varepsilon\,(\varepsilon\ln\varepsilon).$$

當 $\varepsilon\to 0$ 時，$\varepsilon^2\to 0$ 且 $\varepsilon\ln\varepsilon\to 0$（即題目給出的事實），所以

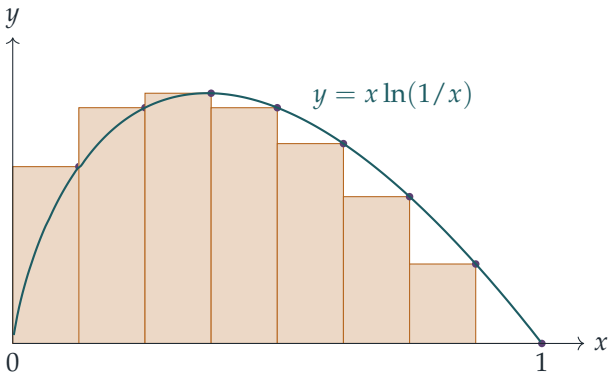
$$\int_0^1 x\ln(1/x)\,\mathrm{d}x=\frac{1}{4}.$$

同一個事實也說明，當 $x \to 0^+$ 時 $f(x) = x\ln(1/x)$ 趨向 0，所以只要定義 $f(0) = 0$，f 就在整個 $[0,1]$ 上連續（continuous）。曲線由 0 上升，在 $x = \frac{1}{\mathrm{e}}$ 處（該處 $f'(x) = -\ln x - 1 = 0$）達到極大值 $\frac{1}{\mathrm{e}}$，然後在 $x = 1$ 處回落到 0。

長方形

把 $[0,1]$ 分成 n 條闊度為 $\frac{1}{n}$ 的長條。第 k 條是 $\left[\frac{k-1}{n}, \frac{k}{n}\right]$，其上的長方形的右上角落在曲線上，所以它的高是

$$f\left(\frac{k}{n}\right) = \frac{k}{n}\ln\frac{n}{k} = \frac{k}{n}(\ln n - \ln k).$$



因此，這些長方形的總面積是

$$\begin{aligned} S_n = \sum_{k=1}^{n} \frac{1}{n}\cdot\frac{k}{n}(\ln n - \ln k) &= \frac{\ln n}{n^2}\sum_{k=1}^{n} k - \frac{1}{n^2}\sum_{k=1}^{n} k\ln k \\ &= \frac{\ln n}{n^2}\cdot\frac{n(n+1)}{2} - \frac{1}{n^2}\sum_{k=1}^{n} k\ln k \\ &= \frac{1}{2}\left(1+\frac{1}{n}\right)\ln n - \frac{1}{n^2}\sum_{k=1}^{n} k\ln k. \end{aligned}$$

第一項已經與我們想要的式子吻合。

？ 想一想 那些階乘是從哪裏來的？

對於 $0 \le m \le n-1$，

$$\ln\frac{n!}{m!} = \ln\big((m+1)(m+2)\cdots n\big) = \sum_{k=m+1}^{n} \ln k.$$

現在把 $m = 0, 1, \ldots, n-1$ 的這些式子加起來，並數一數每個 $\ln k$ 出現了多少次。$\ln k$ 這一項恰好在 $m+1 \le k$ 時出現，即 $m = 0, 1, \ldots, k-1$ 的時候：一共 k 次。因此

$$\ln\frac{n!}{0!} + \ln\frac{n!}{1!} + \cdots + \ln\frac{n!}{(n-1)!} = \sum_{k=1}^{n} k\ln k.$$

（例如取 $n=3$：$\ln 6+\ln 6+\ln 3=\ln 1+2\ln 2+3\ln 3$，兩邊都等於 $\ln 108$。）所以題目中的式子正正就是 S_n，即 n 個長方形的面積。

極限

當 $n\to\infty$ 時，長方形與曲線貼合得越來越緊密，而由於 f 在 $[0,1]$ 上連續，長方形的總面積趨向曲線下的面積。因此

$$\frac{1}{2}\left(1+\frac{1}{n}\right)\ln n-\frac{1}{n^2}\left[\ln\frac{n!}{0!}+\ln\frac{n!}{1!}+\cdots+\ln\frac{n!}{(n-1)!}\right]=S_n\to\int_0^1 x\ln(1/x)\,\mathrm{d}x=\frac{1}{4},$$

這正是所要證明的結果。

如果你想要比「長方形貼合得越來越緊密」更紮實的說法，這裏有一個上界。在每一條長條上，長方形的面積與曲線下的面積相差至多為 $\frac{1}{n}$ 乘以 f 在該長條上的變化幅度。把所有長條上的變化幅度加起來：f 上升了 $\frac{1}{\mathrm{e}}$，又回落了 $\frac{1}{\mathrm{e}}$，所以總和至多為 $\frac{2}{\mathrm{e}}$。因此 $\left|S_n-\frac{1}{4}\right|\le\frac{2}{\mathrm{e}n}\to 0$。

驗算

- 對第一部分的答案求導：$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{2}x^2\ln x-\frac{1}{4}x^2\right)=x\ln x+\frac{1}{2}x-\frac{1}{2}x=x\ln x$。
- 取 $n=2$ 時，式子的值是 $\frac{3}{4}\ln 2-\frac{1}{4}(\ln 2+\ln 2)=\frac{1}{4}\ln 2\approx 0.173$，這正是兩個長方形的面積 $\frac{1}{2}\cdot\frac{1}{2}\ln 2+\frac{1}{2}\cdot 0$。
- 數值上，式子的值在 $n=10$ 時約為 0.24559，在 $n=100$ 時約為 0.249937，在 $n=1000$ 時約為 0.2499992，逐步逼近 $\frac{1}{4}$。

常見錯誤

- 弄丟正負號：被積函數是 $x\ln(1/x)=-x\ln x$，所以第一部分求得的原函數（antiderivative）必須取負，面積是 $+\frac{1}{4}$，而不是 $-\frac{1}{4}$。
- 把 $\frac{1}{2}n(n+1)$ 誤寫成 $\sum k=\frac{1}{2}n^2$：這樣括號內的 $\frac{1}{n}$ 就會不見了。
- 忘記這個積分在 0 處是瑕積分（improper integral）：處理下限時需要用到 $x\ln x\to 0$ 這個事實，而題目已經給出。

STEP 之外

這個長方形面積之和是一個黎曼和（Riemann sum）；對於閉區間上的每一個連續函數，黎曼和都收斂到積分，這是大學分析課程中最早遇到的重要定理之一。我們所用的上

界，考慮的是 f 升降的總幅度，這正是有界變差（bounded variation）函數的想法。反過來看，這道題也提供了關於 $\sum k \ln k$ 的資訊；這個量與超階乘（hyperfactorial）$1^1\,2^2 \cdots n^n$ 密切相關，而後者的增長速度可以用類似 $n!$ 的斯特靈公式（Stirling's formula）的思路更精確地求出。Abbott 的 *Understanding Analysis* 仔細地建立了黎曼積分的理論。

詳細解答

STEP I 1989，第 3 題

純數學 • 向量與純量積

> **STEP I 1989，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/89-s1-q3

關鍵思路

以 O 為原點（origin），把一切都用兩個向量（vector）$\mathbf{a}$ 和 $\mathbf{b}$ 表示。這樣，「垂直（perpendicular）」就變成「純量積（scalar product）為零」，而由於 $\mathbf{a}\cdot\mathbf{a}=\mathbf{b}\cdot\mathbf{b}=1$，這個純量積只取決於 x、OY 和唯一的未知量 $\mathbf{a}\cdot\mathbf{b}$。利用 $OX+OY=1$，便只剩下一條聯繫 x 與 $\mathbf{a}\cdot\mathbf{b}$ 的方程。至於上下界，則取決於 x 可以在多大的範圍內變動。

解答

條件

取 O 為原點，並記 $k=\mathbf{a}\cdot\mathbf{b}$。由於 C 位於由 A 到 B 的三分之一處，而 D 位於三分之二處，

$$\overrightarrow{OC}=\mathbf{a}+\tfrac{1}{3}(\mathbf{b}-\mathbf{a})=\tfrac{1}{3}(2\mathbf{a}+\mathbf{b}),\qquad \overrightarrow{OD}=\tfrac{1}{3}(\mathbf{a}+2\mathbf{b}).$$

由於 X 在 OA 上，且 $OX=x$ 及 $|\mathbf{a}|=1$，故有 $\overrightarrow{OX}=x\mathbf{a}$。同理 $\overrightarrow{OY}=y\mathbf{b}$，其中 $y=OY=1-x$。所以

$$3\overrightarrow{CY}=(3y-1)\mathbf{b}-2\mathbf{a},\qquad 3\overrightarrow{DX}=(3x-1)\mathbf{a}-2\mathbf{b}.$$

兩條直線垂直，當且僅當純量積為零。展開並利用 $\mathbf{a}\cdot\mathbf{a}=\mathbf{b}\cdot\mathbf{b}=1$，得

$$\begin{aligned}9\overrightarrow{CY}\cdot\overrightarrow{DX}&=(3x-1)(3y-1)k-2(3y-1)-2(3x-1)+4k\\&=k\big[(3x-1)(3y-1)+4\big]-2(3x+3y-2).\end{aligned}$$

現在利用 $x+y=1$。於是 $3x+3y-2=1$，並且

$$(3x-1)(3y-1)=9xy-3(x+y)+1=9xy-2.$$

所以垂直的條件變成 $k(9xy+2)=2$。代入 $y=1-x$，即

$$\boxed{\mathbf{a}\cdot\mathbf{b}=\frac{2}{9x(1-x)+2}}$$

（分母至少為 2，所以除以它並無問題）。

上下界

X 位於線段 OA 上，而該線段的長度為 1，所以 $0\le x\le 1$。（這時 $y=1-x$ 也在 $[0,1]$ 之內，所以 Y 自然在 OB 上。）在這個區間上

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

所以 $0\le x(1-x)\le\frac{1}{4}$，其中極大值在 $x=\frac{1}{2}$ 取得，極小值在 $x=0$ 或 1 取得。因此

$$2\le 9x(1-x)+2\le\tfrac{9}{4}+2=\tfrac{17}{4}.$$

取倒數會把不等式（inequality）的方向倒轉（各項皆為正數），再乘以 2 便得

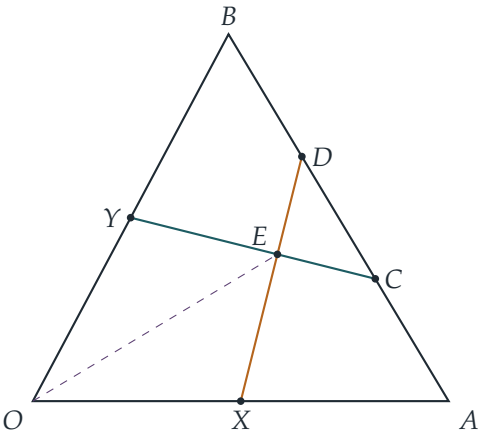
$$\frac{8}{17}\le\mathbf{a}\cdot\mathbf{b}\le 1,$$

正如所求。（取值 1 需要 $\mathbf{a}=\mathbf{b}$，這並不構成真正的三角形，所以實際上 $\mathbf{a}\cdot\mathbf{b}<1$；題目所述的不等式仍然成立。）

最大的角與 OE

> ? **想一想** 為甚麼最大的角對應的是 $\mathbf{a}\cdot\mathbf{b}$ 的最小值？

由於 $|\mathbf{a}|=|\mathbf{b}|=1$，$\mathbf{a}\cdot\mathbf{b}=\cos A\hat{O}B$，而餘弦在 $[0,\pi]$ 上是遞減的。所以當 $\mathbf{a}\cdot\mathbf{b}=\frac{8}{17}$ 時，即當 $x=\frac{1}{2}$（因而 $y=\frac{1}{2}$）時，這個角最大。這時 X 和 Y 分別是 OA 和 OB 的中點，而這個角為 $\arccos\frac{8}{17}\approx 61.9°$。



現在，把 **a** 和 **b** 對調，整個圖形是對稱的：這樣對調會交換 $A \leftrightarrow B$、$C \leftrightarrow D$ 和 $X \leftrightarrow Y$，因而交換直線 CY 和 DX。所以它們的交點 E 在對調之下保持不變，亦即它位於對稱軸（line of symmetry）上：對某個 t 有 $\overrightarrow{OE} = t(\mathbf{a}+\mathbf{b})$。

DX 上的一般點為

$$\overrightarrow{OX} + s\,\overrightarrow{XD} = \tfrac{1}{2}\mathbf{a} + s\left(\tfrac{1}{3}\mathbf{a} + \tfrac{2}{3}\mathbf{b} - \tfrac{1}{2}\mathbf{a}\right) = \left(\tfrac{1}{2} - \tfrac{s}{6}\right)\mathbf{a} + \tfrac{2s}{3}\mathbf{b}.$$

由於 **a** 和 **b** 不平行，我們可以比較係數：當 $\frac{1}{2} - \frac{s}{6} = \frac{2s}{3}$ 時，這是 $\mathbf{a}+\mathbf{b}$ 的倍數，由此得 $s = \frac{3}{5}$ 及 $t = \frac{2}{5}$。所以 $\overrightarrow{OE} = \frac{2}{5}(\mathbf{a}+\mathbf{b})$。最後

$$|\mathbf{a}+\mathbf{b}|^2 = \mathbf{a}\cdot\mathbf{a} + 2\,\mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{b} = 2 + \tfrac{16}{17} = \tfrac{50}{17},$$

所以

$$OE = \frac{2}{5}\sqrt{\frac{50}{17}} = \frac{2}{5}\cdot\frac{5\sqrt{2}}{\sqrt{17}} = \frac{2\sqrt{2}}{\sqrt{17}} = \boxed{\frac{2\sqrt{34}}{17}} \approx 0.686.$$

驗算

- 用坐標表示，取 $\mathbf{a} = (1,0)$ 及 $\mathbf{b} = (\frac{8}{17}, \frac{15}{17})$。於是 $X = (\frac{1}{2}, 0)$，$D = (\frac{11}{17}, \frac{10}{17})$，所以 $\overrightarrow{XD} = (\frac{5}{34}, \frac{10}{17})$。另外 $Y = (\frac{4}{17}, \frac{15}{34})$ 及 $C = (\frac{14}{17}, \frac{5}{17})$，所以 $\overrightarrow{CY} = (-\frac{10}{17}, \frac{5}{34})$。它們的純量積為 $-\frac{50}{578} + \frac{50}{578} = 0$，正如所料。
- $E = \frac{2}{5}(\frac{25}{17}, \frac{15}{17}) = (\frac{10}{17}, \frac{6}{17})$，而 $|OE| = \frac{\sqrt{136}}{17} = \frac{2\sqrt{34}}{17} \approx 0.6860$。這與直接以數值方法求兩條直線交點所得的結果吻合。
- $s = \frac{3}{5}$ 這個值在 $[0,1]$ 之內，所以 E 確實位於 X 與 D 之間，在三角形之內。

常見錯誤

- 忘記了 $\mathbf{a}\cdot\mathbf{a} = 1$（以致在條件中留下 $|\mathbf{a}|^2$ 的項），或者把兩個三等分點（trisection points）的位置調轉了。
- 以為最大的角對應最大的 $\mathbf{a}\cdot\mathbf{b}$：角愈大，餘弦愈小。
- 寫成 $\overrightarrow{OX} = x$ 而不是 $x\mathbf{a}$。點 X 是以向量表示的；x 只是它的長度。

STEP 之外

這裏的方法，是把每一點寫成兩個基向量（basis vectors）的組合，並把幾何問題化為純量積；這正是學習內積（inner product）和格拉姆矩陣（Gram matrix）的起點。一組向量的構形，在不計旋轉的意義下，完全由它的純量積表決定，而在這裏該表只有 1、1 和

$\mathbf{a} \cdot \mathbf{b}$。這就是整道題目可以歸結為一個數的原因。Strang 的 *Introduction to Linear Algebra*（線性代數入門）對這些想法有詳細的闡述。

詳細解答

STEP I 1989，第 4 題

純數學 • 拉姆齊理論

> **STEP I 1989，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/89-s1-q4

本題在 *Beyond STEP* 第 9.2 節小型版的拉姆齊定理中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch09/ramsey.html

關鍵思路

先把幾何放在一旁：題目對各點所設的條件，只是保證任意三點都構成一個真正的三角形，所以這純粹是一個為連接六點的 15 條邊着色的問題。按照提示，先看其中一點。它連出五條邊，而顏色只有兩種，所以根據鴿巢原理（pigeonhole principle），這些邊之中至少有三條同色。這三條邊另一端的三點之間的邊，要麼與最初那一點合成一個單色三角形，要麼自己構成一個單色三角形。

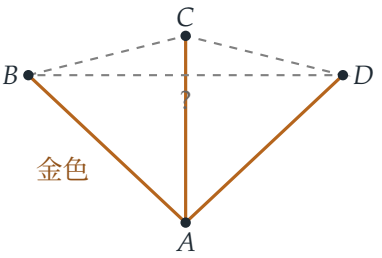
解答

六點

由於任意三點都不共線（collinear），任意三點都是一個真正的三角形的頂點，而它的三條邊都在我們所畫的線段之中。（任意四點不共面（coplanar）這個條件，在論證中用不着。）

考慮由 A 連出的五條線段 AB, AC, AD, AE, AF。每條線段不是金色就是銀色。假如金色的最多兩條，銀色的也最多兩條，線段總數便最多只有四條，這與事實不符。所以其中至少有三條同色。把 $B, \ldots, F$ 各點重新標記，並在有需要時把兩種顏色的名稱對調，我們可以假設

$$AB,\ AC,\ AD \text{ 全是金色。}$$



現在看看連接另一端三點的三條線段 BC、CD 和 DB。

> **想一想** 無論 BC、CD 和 DB 是甚麼顏色，為甚麼以下兩種情況之中必有一種會給出單色三角形？

只有兩種可能。

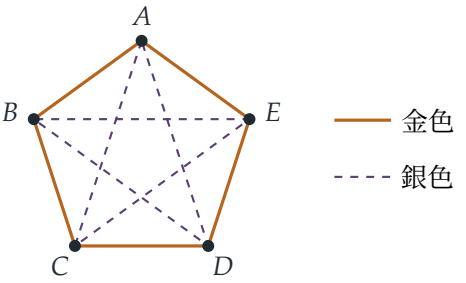
- BC, CD, DB 之中至少有一條是金色。設它是 BC（其餘情況只是把字母換一換，道理相同）。那麼 AB、AC 和 BC 全是金色，所以三角形 ABC 的邊都是同一種顏色。
- BC, CD, DB 之中沒有一條是金色。那麼三條都是銀色，三角形 BCD 的邊都是同一種顏色。

無論哪一種情況，都有一個各邊同色的三角形，這正是我們要證明的。

五點

若只有五點，每一點只連出四條邊，可以分成兩條一種顏色、兩條另一種顏色，於是上述論證不再成立。要說明結論確實不成立，我們需要找出一種沒有單色三角形的着色方法。

把 A, B, C, D, E 五點依次放在一個五邊形的頂點上。把五條邊 AB, BC, CD, DE, EA 塗成金色，五條對角線 AC, BD, CE, DA, EB（即五角星）塗成銀色。



任何三角形都用上五點之中的三點。在五邊形的任意三個頂點之中，至少有兩個並不相鄰（若三個頂點兩兩相鄰，5-圈之內便要有一個三角形），所以每個三角形都有一條銀色

的邊。此外，任意三個頂點之中也至少有兩個是相鄰的：假如三個頂點兩兩不相鄰，那麼沿五邊形走一圈，每兩個相繼的選定頂點之間都至少隔着一個未被選中的頂點，這就需要 $3+3=6$ 點。所以每個三角形也有一條金色的邊。因此沒有任何三角形是單色的。

> **想一想** 正五邊形是平的，所以它有四個頂點共面。這不是違反了關於一般位置的規定嗎？

假如各點真的在同一平面上，那確實違反了規定；但着色只取決於哪兩點相連，而與各點的位置無關。把各頂點稍稍提離平面，而且各自升高不同的幅度（例如把第 k 個頂點升到高度 $0.3k^2$，其中 $k=0,1,2,3,4$）。這樣便沒有三點共線，也沒有四點共面，而從上方看下去，圖形仍然是這裏所畫的五邊形和五角星，着色相同，也沒有單色三角形。

驗算

六點的邊只有 $2^{15}=32768$ 種着色方法，所以可以用電腦逐一檢驗：每一種都有單色三角形，而且事實上總有至少兩個。至於五點，把五邊形着色中全部 $\binom{5}{3}=10$ 個三角形逐一檢查，可以確認每個都有兩種顏色。再用行列式（determinant）檢驗，可以確認上面升高了的五邊形沒有四點共面。

常見錯誤

- 只說「由對稱性」而沒有交代假設了甚麼：要明確說明，你是從 A 連出的邊之中選出佔多數顏色的三條，並把那種顏色稱為金色。
- 只證明了 A 有三條同色的邊便停下來；第二步，即考察 BC、CD、DB，才是這個證明的核心。
- 處理五點時，只畫一個草圖而沒有檢查每一個三角形；像「由 A 連出的星形塗金色，其餘全塗銀色」這樣的着色是行不通的。

STEP 之外

這是拉姆齊理論（Ramsey theory）的第一個例子：拉姆齊數（Ramsey number）$R(3,3)$ 是使得完全圖（complete graph）K_n 的邊的每一種雙色着色都含有單色三角形的最小的 n，而本題正好證明了 $R(3,3)=6$。拉姆齊定理（Ramsey's theorem）指出，無論要求的單色完全子圖有多大，這樣的數都存在，但它們極難計算：就連 $R(5,5)$ 也仍然未知。Wilson 的 *Introduction to Graph Theory*（圖論入門）是一本很出色的圖論（graph theory）入門書，當中也講到拉姆齊定理。

詳細解答

STEP I 1989，第 5 題

純數學 • 二項式係數恆等式

STEP I 1989，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/89-s1-q5

關鍵思路

二項展開式（binomial expansion）是關於 x 的恆等式（identity），所以無論代入甚麼數都成立。取 $x = 1$ 會把所有係數（coefficient）加起來，取 $x = -1$ 則以正負相間的符號把它們相加；把兩者結合，就能把偶數項與奇數項的係數分開。至於 (ii) 中的加權和，$\binom{n}{r}$ 前面的因子 r 正好是對 x^r 求導時產生的，所以我們先對展開式求導（differentiate），然後再代入。

解答

對正整數 n，

$$(1+x)^n = \sum_{r=0}^{n} \binom{n}{r} x^r = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n,$$

其中 $\binom{n}{r} = \dfrac{n!}{r!\,(n-r)!}$。這對所有實數 x 都成立。

第 (i) 部分

設 $E = \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{n}$ 為偶數項係數之和，$O = \binom{n}{1} + \binom{n}{3} + \cdots + \binom{n}{n-1}$ 為奇數項係數之和。由於 n 是偶數，最後一項 $\binom{n}{n}$ 屬於 E，而 $\binom{n}{n-1}$ 屬於 O，所以 E 與 O 合起來恰好包含每個係數一次。

在展開式中代入 $x = 1$，得

$$E + O = 2^n.$$

代入 $x = -1$，每個偶數次冪 $(-1)^r$ 都是 $+1$，每個奇數次冪都是 -1，而由於 $n \geq 1$，左

方是 $0^n = 0$。所以

$$E - O = 0.$$

因此 $E = O$，於是 $2E = 2^n$，故

$$E = O = 2^{n-1},$$

這就是所求的結果。

第 (ii) 部分

把展開式兩邊對 x 求導。常數項消失，而 x^r 變成 rx^{r-1}：

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

現在代入 $x = 1$：

$$n2^{n-1} = \binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n},$$

正如所求。這對所有正整數 n 都成立，不論奇偶。

想一想 有沒有不用微積分（calculus）也能看出第 (ii) 部分的方法？

有。由於對 $1 \le r \le n$ 有 $r\binom{n}{r} = r \cdot \dfrac{n!}{r!\,(n-r)!} = n \cdot \dfrac{(n-1)!}{(r-1)!\,(n-r)!} = n\binom{n-1}{r-1}$，在 $(1+x)^{n-1}$ 的展開式中代入 $x = 1$，便知這個和等於 $n\sum_{r=1}^{n}\binom{n-1}{r-1} = n \cdot 2^{n-1}$。用文字來說：從 n 個人中選出一個人數不限的委員會，再從委員會中選出一位主席，這等同於先選主席（n 種方法），再從其餘 $n-1$ 個人中任選一個子集（subset）（2^{n-1} 種方法）。

最後的和

把這個和拆成兩部分：

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

第一個和中 $r = 0$ 的一項是零，所以第一個和正是 (ii) 中的和，即 $n2^{n-1}$。第二個和是 $(1+x)^n$ 的展開式在 $x = -1$ 時的值，即對所有正整數 n 都有 $(1-1)^n = 0$。（這就是 (i) 中 $E - O = 0$ 的那一步；它只需要 $n \ge 1$，不需要 n 是偶數。）因此

$$\boxed{\sum_{r=0}^{n}(r + (-1)^r)\binom{n}{r} = n2^{n-1}.}$$

驗算

- 當 $n = 2$：係數為 $1, 2, 1$，所以 $E = 1 + 1 = 2 = O$，而 $2^1 = 2$。(ii) 中的和是 $2 + 2 \cdot 1 = 4 = 2 \cdot 2^1$。
- 當 $n = 3$，最後的和是 $(0+1)\cdot 1 + (1-1)\cdot 3 + (2+1)\cdot 3 + (3-1)\cdot 1 = 1 + 0 + 9 + 2 = 12 = 3 \cdot 2^2$。
- 我們對 1 至 29 的每個 n 都精確地驗證了全部三個恆等式；例如 $n = 6$ 給出 $192 = 6 \cdot 2^5$。

常見錯誤

- 忘記了 $\sum(-1)^r\binom{n}{r} = 0$ 對奇數 n 同樣成立，因而以為最後的答案取決於 n 的奇偶性（parity）。它只在 $n = 0$ 時才不成立，而這裏 n 是正數。
- 在 (ii) 中先代入 $x = 1$ 才求導。代入必須放在最後：一旦 x 成了一個數，就再沒有甚麼可以求導了。

STEP 之外

把多項式（polynomial）看成儲存係數的倉庫，然後透過代入、求導或相乘來提取資訊，這就是*生成函數*（generating function）的想法，它是組合數學（combinatorics）和概率論中最強大的工具之一。若代入的不是 ± 1 而是複數單位根（roots of unity），便能以同樣的方式抽出每隔三項或每隔四項的係數；而對概率生成函數（probability generating function）求導，就得到分佈的平均值（mean），正如在這裏它給出了 $\sum r\binom{n}{r}$ 一樣。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）展示了生成函數的實際應用。

詳細解答

STEP I 1989，第 6 題

純數學 • 法線與微分方程

STEP I 1989，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/89-s1-q6

關鍵思路

在 P 點的法線（normal）斜率（gradient）為 $-1/\mathrm{f}'(x)$，所以要求 Q，只須把 $X=0$ 代入法線方程。於是 PQ^2 就是水平距離的平方加上垂直距離的平方，而題目給出的長度便化成一條只含 $\mathrm{f}'(x)$ 的方程。唯一要留神的是最後開平方的一步：正負號的取捨由極小點（minimum point）決定。

解答

Q 的 y 坐標

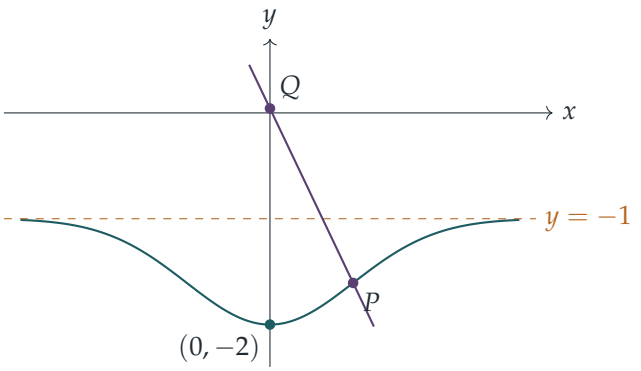
在 $P=(x,\mathrm{f}(x))$ 處，切線（tangent）的斜率為 $\mathrm{f}'(x)$，所以（只要 $\mathrm{f}'(x)\neq 0$）法線的斜率為 $-1/\mathrm{f}'(x)$。以 (X,Y) 表示法線上動點的坐標，法線方程為

$$Y-\mathrm{f}(x)=-\frac{1}{\mathrm{f}'(x)}\,(X-x).$$

代入 $X=0$，法線與 y 軸相交於

$$Y=\mathrm{f}(x)+\frac{x}{\mathrm{f}'(x)},\qquad \text{所以}\qquad Q=\left(0,\ \mathrm{f}(x)+\frac{x}{\mathrm{f}'(x)}\right).$$

（若 $\mathrm{f}'(x)=0$，法線是鉛垂線，除非 $x=0$，否則它永不與 y 軸相交。）



微分方程（differential equation）

P 與 Q 的水平距離為 x，垂直距離為 $x/\mathrm{f}'(x)$，所以

$$PQ^2 = x^2 + \frac{x^2}{\big(\mathrm{f}'(x)\big)^2}.$$

令上式等於 $\mathrm{e}^{x^2} + x^2$，兩邊的 x^2 項互相抵消，得到

$$\frac{x^2}{\big(\mathrm{f}'(x)\big)^2} = \mathrm{e}^{x^2}, \qquad \text{即} \qquad \boxed{\big(\mathrm{f}'(x)\big)^2 = x^2\,\mathrm{e}^{-x^2}.}$$

這等價於 $\mathrm{f}'(x) = \pm x\,\mathrm{e}^{-x^2/2}$。

極小點位於 $(0,-2)$ 的曲線

想一想 應該取哪一個正負號？正負號又會否在曲線中途改變？

由於 f′ 連續（continuous）（曲線有極小點，所以我們假設它可微（differentiable）），而 $\mathrm{f}'(x)^2 = x^2\mathrm{e}^{-x^2}$ 只在 $x = 0$ 時為零，因此 f′ 在 $x > 0$ 上或在 $x < 0$ 上都不可能變號：要變號就必須經過零。要使 $(0,-2)$ 成為極小點，f 必須在緊靠 0 的左方遞減，在緊靠其右方遞增，所以當 $x < 0$ 時 $\mathrm{f}'(x) < 0$，當 $x > 0$ 時 $\mathrm{f}'(x) > 0$。這正好與 x 的正負號相同，因此在兩邊都有

$$\mathrm{f}'(x) = x\,\mathrm{e}^{-x^2/2}.$$

（若全程取負號，則在 $x = 0$ 處得到極大點；若兩邊取不同的正負號，則 f′ 的正負號保持不變，即得到拐點（point of inflection）。）

進行積分，由於 $\frac{\mathrm{d}}{\mathrm{d}x}\big(\mathrm{e}^{-x^2/2}\big) = -x\,\mathrm{e}^{-x^2/2}$，

$$\mathrm{f}(x) = -\mathrm{e}^{-x^2/2} + c.$$

由條件 $\mathrm{f}(0) = -2$ 得 $-1 + c = -2$，所以 $c = -1$，曲線為

$$\boxed{y = -1 - \mathrm{e}^{-x^2/2}.}$$

它在 $x = 0$ 處取得極小值 -2，並在兩邊向漸近線（asymptote）$y = -1$ 上升，一如略圖所示。

驗算

- 在 $x = 1$ 處：$\mathrm{f}(1) = -1 - \mathrm{e}^{-1/2} \approx -1.6065$，而 $\mathrm{f}'(1) = \mathrm{e}^{-1/2} \approx 0.6065$，所以 $Q \approx (0,\, 0.0422)$，且 $PQ \approx \sqrt{1 + 1.6487^2} \approx 1.9283$，等於 $\sqrt{\mathrm{e} + 1}$，符合要求。
- 電腦在 2000 個隨機點上，以法線與 y 軸的交點重新計算 Q，所得的 PQ 與 $\sqrt{\mathrm{e}^{x^2} + x^2}$ 相符，（相對）誤差在 10^{-15} 以內。

常見錯誤

- 誤用切線的斜率 $\mathrm{f}'(x)$，而不是法線的斜率 $-1/\mathrm{f}'(x)$，結果得出 Q 位於 $\mathrm{f}(x) - x\mathrm{f}'(x)$。
- 開平方時直接寫 $\mathrm{f}'(x) = x\mathrm{e}^{-x^2/2}$，卻沒有說明為何捨去負號。

STEP 之外

在這一題，一個對曲線上每一點都成立的幾何條件，化成了一條一階微分方程，但方程中的 f' 以平方形式出現，所以要靠額外的資料選定分支，才能確定唯一的解。這類沒有就 y' 解出的方程，正是常微分方程（ODE）的唯一性定理（uniqueness theorem）開始需要小心處理的地方，並由此引申到包絡線（envelope）和奇解（singular solution）。Boyce 與 DiPrima 的 *Elementary Differential Equations* 涵蓋一階方程及其存在性與唯一性理論。

詳細解答

STEP I 1989，第 7 題

純數學 • 曲線描繪與最優化

STEP I 1989，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/89-s1-q7

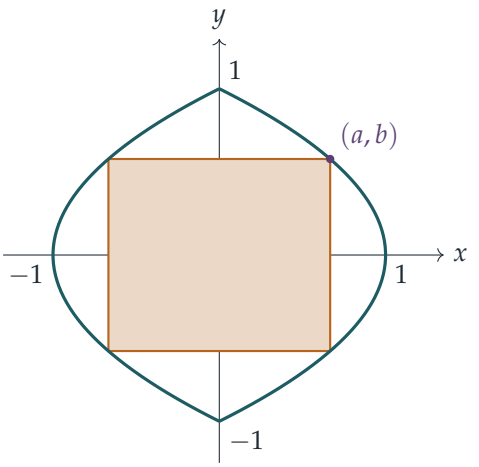
關鍵思路

這條曲線對兩條坐標軸都對稱，所以最大的長方形應以原點為中心，並有一個頂點位於第一象限的曲線上。這樣一來，原本要在兩個變量中搜尋的問題，就變成求一個單變量函數的最大值。對於一般的曲線，只要把 $|y|$ 解出來，寫成 $1-|x|$ 的冪，同樣的想法依然適用。

解答

草圖

由於 $y^2 \geq 0$，必須有 $1-|x| \geq 0$，故 $-1 \leq x \leq 1$。當 $x \geq 0$ 時，曲線是 $x = 1-y^2$，即一條開口向左、頂點（vertex）為 $(1,0)$ 的拋物線（parabola）；當 $x \leq 0$ 時，曲線是它的鏡像 $x = y^2 - 1$。兩段弧在 $(0,\pm 1)$ 相接，並在該處形成尖角（斜率 $\mathrm{d}y/\mathrm{d}x = \mp\frac{1}{2}$ 與 $\pm\frac{1}{2}$ 並不相同），又在 $(\pm 1, 0)$ 垂直地穿過 x 軸。這條曲線是一個閉合的環，圍住區域 $|x| + y^2 \leq 1$，形狀像一個圓角的菱形。



最大的長方形

設長方形 $[x_1, x_2] \times [y_1, y_2]$ 位於區域 $|x| + y^2 \leq 1$ 之內。令 $A = \max(|x_1|, |x_2|)$ 及 $B = \max(|y_1|, |y_2|)$。各頂點為 (x_i, y_j)，其中 i, j 可獨立選取，所以必有一個頂點同時滿足 $|x| = A$ 和 $|y| = B$。該頂點位於區域之內，故 $A + B^2 \leq 1$。於是長方形 $[-A, A] \times [-B, B]$ 也位於區域之內，因為當 $|x| \leq A$、$|y| \leq B$ 時，

$$|x| + y^2 \leq A + B^2 \leq 1.$$

它的闊度 $2A \geq x_2 - x_1$，高度 $2B \geq y_2 - y_1$，所以它的面積不會比原來的小。因此我們可以只考慮以原點為中心、頂點為 $(\pm a, \pm b)$ 的長方形；再把頂點向外推，直至碰到曲線為止，便有 $a = 1 - b^2$，其中 $0 \leq b \leq 1$。

? **想一想** 為甚麼只需考慮以原點為中心的長方形？

這正是上面的論證：把每條邊換成它到坐標軸的兩個距離中較大的一個，長方形仍然留在曲線之內（由對稱性可知），而且絕不會縮小。

面積為

$$S(b) = 4ab = 4b(1 - b^2) = 4b - 4b^3, \qquad S'(b) = 4 - 12b^2.$$

$S'(b) = 0$ 於 $b = 1/\sqrt{3}$，在此之前 $S' > 0$，在此之後 $S' < 0$，而且 $S(0) = S(1) = 0$，所以這就是最大值。此時 $a = 1 - \frac{1}{3} = \frac{2}{3}$，而

$$S_{\max} = 4 \cdot \frac{2}{3} \cdot \frac{1}{\sqrt{3}} = \frac{8}{3\sqrt{3}} = \frac{8}{\sqrt{27}},$$

正如所求。

一般的曲線

現在考慮 $y^{2m} = (1-|x|)^n$。左邊為非負數，而由於 n 是奇數，$(1-|x|)^n$ 與 $1-|x|$ 同號。所以同樣有 $|x| \le 1$，取（正的）$2m$ 次方根，得

$$|y| = (1-|x|)^k, \qquad k = \frac{n}{2m} > 0.$$

（假如 n 是偶數，則任何 x 都可以取，曲線會伸延至無窮遠，因而不存在最大的長方形：這就是 n 必須是奇數的原因。）這條曲線同樣是一個經過 $(\pm1, 0)$ 和 $(0, \pm1)$ 的閉合環，對兩條坐標軸都對稱，圍住區域 $|y| \le (1-|x|)^k$。由於 $(1-|x|)^k$ 隨 $|x|$ 增加而遞減，與之前相同的論證說明，我們可以取頂點為 $(\pm a, \pm b)$ 的長方形，其中 $b = (1-a)^k$，$0 \le a \le 1$。它的面積為

$$S(a) = 4a(1-a)^k.$$

求導（differentiate），得

$$S'(a) = 4(1-a)^{k-1}\big[(1-a) - ka\big] = 4(1-a)^{k-1}\big[1-(1+k)a\big],$$

當 $0 < a < \frac{1}{1+k}$ 時為正，當 $\frac{1}{1+k} < a < 1$ 時為負。由於 $S(0) = S(1) = 0$，最大值出現於 $a = \frac{1}{1+k}$，此時 $1-a = \frac{k}{1+k}$：

$$S_{\max} = \frac{4}{1+k}\left(\frac{k}{1+k}\right)^k = \frac{4k^k}{(1+k)^{k+1}}.$$

代入 $k = \frac{n}{2m}$ 及 $1+k = \frac{n+2m}{2m}$，$2m$ 的各個冪合併為 $(2m)^{-k} \cdot (2m)^{k+1} = 2m$，所以

$$S_{\max} = \frac{8m\, n^{n/(2m)}}{(n+2m)^{(n+2m)/(2m)}} = 8m\left(\frac{n^n}{(n+2m)^{n+2m}}\right)^{1/(2m)}.$$

驗算

- 取 $m = n = 1$，得 $8\left(\frac{1}{27}\right)^{1/2} = 8/\sqrt{27} \approx 1.5396$，與第一部分的結果一致。
- 對 $(m,n) = (1,3), (2,1), (2,3), (3,5)$ 在 $a \in [0,1]$ 上作精細的數值搜尋，結果與公式吻合；例如 $(m,n) = (2,1)$ 給出 $16/5^{5/4} \approx 2.1400$。隨機選取的偏離中心的長方形，面積從未超過這個值。
- 當 n 固定而 $m \to \infty$ 時，$k \to 0$，曲線逐漸填滿正方形 $|x| \le 1$、$|y| \le 1$；公式的值趨向 4，正是該正方形的面積。

常見錯誤

- 漏掉因子 4：頂點 (a,b) 所對應的長方形，闊度是 $2a$，高度是 $2b$。
- 忽略 n 為奇數的重要性：若 n 是偶數，曲線便不是有界的。

STEP 之外

求 $a(1-a)^k$ 的最大值，其實是一道喬裝了的加權算術—幾何平均不等式（weighted AM–GM）問題：把 $a\cdot\left(\frac{1-a}{k}\right)^k$ 寫成 $1+k$ 個和為常數的因子之積，無須微積分便可求得最大值。曲線 $|x|^p+|y|^q=1$ 與 ℓ^p 範數（norm）的單位球（unit ball）是近親，後者會在大學一年級的分析和線性代數中出現。想進一步了解這類不等式，可參閱 Siklos 的 *Advanced Problems in Mathematics*；至於最優化（optimisation）背後的分析，則可參閱 Abbott 的 *Understanding Analysis*。

詳細解答

STEP I 1989，第 8 題

純數學 • 倍角公式

> **STEP I 1989，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/89-s1-q8

關鍵思路

棣莫弗定理（De Moivre's theorem）把 $\cos n\theta$ 化為 $(\cos\theta + \mathrm{i}\sin\theta)^n$ 的實部（real part），再把 $\sin^2\theta$ 換成 $1-\cos^2\theta$，便得到一個以 $\cos\theta$ 為變量的多項式（polynomial）。至於最後一部分，那個六次多項式（sextic）看來像是剛求得的兩個多項式的某種組合，於是我們着手找出這個組合：原來它正是 $\cos 6\theta - \cos 4\theta$ 的一半，而解 $\cos 6\theta = \cos 4\theta$ 十分容易。

解答

以下一律記 $c = \cos\theta$ 及 $s = \sin\theta$，故 $s^2 = 1 - c^2$。

第 (i) 部分

由棣莫弗定理，$\cos 4\theta + \mathrm{i}\sin 4\theta = (c + \mathrm{i}s)^4$。以二項式定理（binomial theorem）展開，並只保留實數項（即 $\mathrm{i}s$ 的冪次為偶數的各項），

$$\begin{aligned}\cos 4\theta &= c^4 - \tbinom{4}{2}c^2s^2 + s^4 \\ &= c^4 - 6c^2(1-c^2) + (1-c^2)^2 \\ &= c^4 - 6c^2 + 6c^4 + 1 - 2c^2 + c^4 \\ &= 8c^4 - 8c^2 + 1,\end{aligned}$$

這正是所求的恆等式（identity）。

第 (ii) 部分

同樣地，取 $(c+\mathrm{i}s)^6$ 的實部，得

$$\cos 6\theta = c^6 - 15c^4s^2 + 15c^2s^4 - s^6,$$

其中用到 $\binom{6}{2} = \binom{6}{4} = 15$ 以及 $\mathrm{i}^2 = -1$、$\mathrm{i}^4 = 1$、$\mathrm{i}^6 = -1$。設 $u = c^2$，則 $s^2 = 1-u$：

$$\begin{aligned}\cos 6\theta &= u^3 - 15u^2(1-u) + 15u(1-u)^2 - (1-u)^3 \\ &= u^3 - 15u^2 + 15u^3 + 15u - 30u^2 + 15u^3 - 1 + 3u - 3u^2 + u^3 \\ &= 32u^3 - 48u^2 + 18u - 1 \\ &= 32c^6 - 48c^4 + 18c^2 - 1,\end{aligned}$$

證畢。

方程

把左方記為 $p(x) = 16x^6 - 28x^4 + 13x^2 - 1$。

> **想一想** 這個多項式並不是那兩個恆等式中的任何一個，它們又怎能派上用場？

比較係數（coefficient）。首項係數 16 是 32 的一半，常數項 -1 是 -2 的一半，這提示我們把 (ii) 減去 (i)：

$$\cos 6\theta - \cos 4\theta = 32c^6 - 56c^4 + 26c^2 - 2 = 2\,p(c).$$

因此，若 $x = \cos\theta$，則

$$p(x) = 0 \iff \cos 6\theta = \cos 4\theta.$$

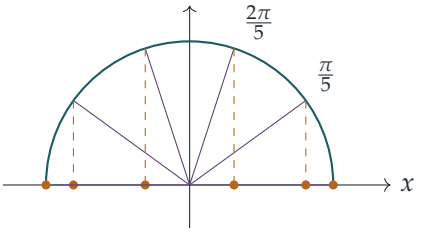
而 $\cos A = \cos B$ 當且僅當存在整數 k 使得 $A = \pm B + 2k\pi$。在這裏即是

$$6\theta = 4\theta + 2k\pi \;\Rightarrow\; \theta = k\pi, \qquad \text{或} \qquad 6\theta = -4\theta + 2k\pi \;\Rightarrow\; \theta = \frac{k\pi}{5}.$$

第一組解已包含在第二組之內，所以解為 $\theta = k\pi/5$，而對應的 $x = \cos\theta$ 的值是

$$\cos 0 = 1, \quad \cos\tfrac{\pi}{5}, \quad \cos\tfrac{2\pi}{5}, \quad \cos\tfrac{3\pi}{5} = -\cos\tfrac{2\pi}{5}, \quad \cos\tfrac{4\pi}{5} = -\cos\tfrac{\pi}{5}, \quad \cos\pi = -1.$$

k 取其他值時只會重複以上各數。這是六個互不相同的數（餘弦在 $[0,\pi]$ 上嚴格遞減），而 p 的次數（degree）為 6，所以它們就是 p 的全部根（root），不論實根還是複根。特別要留意，代換 $x = \cos\theta$ 只能涵蓋 $-1 \le x \le 1$，但我們並沒有遺漏這個區間以外的根，因為已經再沒有其他根可找了。



這些還不算是可接受的答案，因為題目不容許使用數值近似，我們應該給出這些餘弦的準確值。由於 $x = 1$ 是根，$x = -1$ 也是根（p 是偶函數），所以 $x^2 - 1$ 是一個因式（factor）。把它除去（或比較係數），

$$p(x) = (x^2 - 1)(16x^4 - 12x^2 + 1).$$

這個四次因式是 x^2 的二次式：

$$x^2 = \frac{12 \pm \sqrt{144 - 64}}{32} = \frac{12 \pm 4\sqrt{5}}{32} = \frac{3 \pm \sqrt{5}}{8}.$$

注意

$$\left(\frac{1+\sqrt{5}}{4}\right)^2 = \frac{6+2\sqrt{5}}{16} = \frac{3+\sqrt{5}}{8}, \qquad \left(\frac{\sqrt{5}-1}{4}\right)^2 = \frac{6-2\sqrt{5}}{16} = \frac{3-\sqrt{5}}{8}.$$

x^2 的兩個值都是正數，所以各給出兩個實數 x。因此方程的實根就是以下六個數：

$$x = \pm 1, \qquad x = \pm\frac{1+\sqrt{5}}{4}, \qquad x = \pm\frac{\sqrt{5}-1}{4}.$$

額外的收穫是：按大小與上面的列表對照（$\cos\frac{\pi}{5} > \cos\frac{2\pi}{5} > 0$），便得到兩個著名的值 $\cos\frac{\pi}{5} = \frac{1+\sqrt{5}}{4}$ 和 $\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}$。

驗算

- 當 $\theta = 0$ 時，兩個恆等式都給出 $1 = 1$；當 $\theta = \frac{\pi}{2}$ 時，(i) 給出 $1 = 1$，(ii) 給出 $-1 = -1$。
- p 的各係數之和為 $16 - 28 + 13 - 1 = 0$，印證了 $x = 1$ 是根。
- 以數值計算，$\frac{1+\sqrt{5}}{4} \approx 0.80902 = \cos 36°$ 及 $\frac{\sqrt{5}-1}{4} \approx 0.30902 = \cos 72°$，而 p 在這兩點的值都是零（在捨入誤差範圍內）。

常見錯誤

- 遺漏了由 $6\theta = -4\theta + 2k\pi$ 得出的解，結果只找到 $x = \pm 1$。
- 把答案寫成 $\cos\frac{\pi}{5}$ 之類，或寫成小數：題目要求的是準確的根式（surd）形式。
- 沒有解釋為何不存在 $|x| > 1$ 的根：數一數次數便可解決。

STEP 之外

把 $\cos n\theta$ 用 $\cos\theta$ 表示的多項式，就是切比雪夫多項式（Chebyshev polynomial）T_n，所以 (i) 和 (ii) 說的是 $T_4(x) = 8x^4 - 8x^2 + 1$ 和 $T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$。切比雪夫多項式在數值分析（numerical analysis）中佔有核心地位，它們的根正是多項式插值（polynomial interpolation）的最佳取點。$\cos\frac{\pi}{5}$ 中出現 $\sqrt{5}$，反映了正五邊形的幾何與黃金比（golden ratio）。Needham 的 *Visual Complex Analysis*（《複分析：可視化方法》）是從幾何角度認識棣莫弗定理和單位根（roots of unity）的絕佳讀物。

詳細解答

STEP I 1989，第 9 題

純數學 • 圖像的變換

> **STEP I 1989，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/89-s1-q9

關鍵思路

只要畫出基本的三次曲線（cubic），其餘每一個圖像都只是同一條曲線經坐標變換後的樣子。代換（substitution）$X = ax + b$, $Y = cy + d$ 不過是在每個方向上先作伸縮（stretch）（或反射（reflection）），再作平移（translation），而且它會把轉向點（turning point）變到轉向點。因此，全部工夫就在於找出原來三次曲線的兩個轉向點，並追蹤它們去了哪裏。

解答

基本圖像

寫成 $y = \frac{1}{8}(x^3 - 12x) = \frac{1}{8}x(x^2 - 12)$。於是

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3x^2 - 12}{8} = \frac{3(x-2)(x+2)}{8}, \qquad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{6x}{8} = \frac{3x}{4}.$$

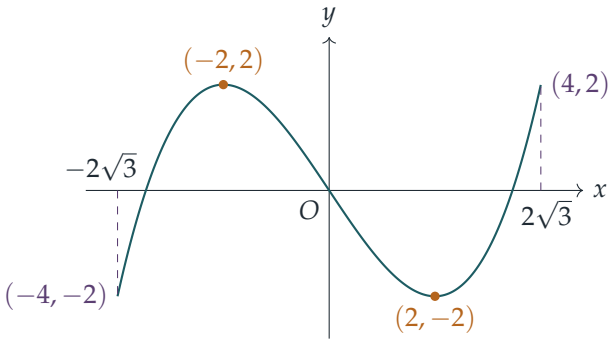
斜率（gradient）在 $x = \pm 2$ 處為零。

- 在 $x = -2$：$y = \frac{1}{8}(-8 + 24) = 2$，而 $y'' = -\frac{3}{2} < 0$，所以 $(-2, 2)$ 是局部極大點（local maximum）。
- 在 $x = 2$：$y = \frac{1}{8}(8 - 24) = -2$，而 $y'' = \frac{3}{2} > 0$，所以 $(2, -2)$ 是局部極小點（local minimum）。

其他有用的特徵：

- 曲線與坐標軸相交於 $x(x^2 - 12) = 0$ 之處，即 $x = 0$ 及 $x = \pm 2\sqrt{3} \approx \pm 3.46$。

- 這個函數是奇函數（odd function），所以圖像繞原點有 2 階旋轉對稱（rotational symmetry），而原點亦是拐點（point of inflection）；該處的斜率為 $-\frac{3}{2}$。
- 在兩端，$y(4) = \frac{1}{8}(64 - 48) = 2$ 及 $y(-4) = -2$。所以曲線恰好在 $x = 4$ 處回升到極大點的高度。



為何轉向點會變到轉向點

想一想 當我們變換坐標時，怎樣知道新的轉向點就是舊轉向點的像？

若 $X = ax + b$ 及 $Y = cy + d$，其中 $a, c \neq 0$，則由鏈式法則（chain rule）可得

$$\frac{\mathrm{d}Y}{\mathrm{d}X} = \frac{\mathrm{d}Y/\mathrm{d}x}{\mathrm{d}X/\mathrm{d}x} = \frac{c}{a}\frac{\mathrm{d}y}{\mathrm{d}x},$$

它恰好在 $\mathrm{d}y/\mathrm{d}x$ 為零之處為零。所以新的轉向點就是 $(-2, 2)$ 和 $(2, -2)$ 的像。（若 $c/a < 0$，圖形會被反射，極大點與極小點的角色互換。）從幾何上看：

- $X = \frac{1}{2}x$ 把圖像沿水平方向壓縮，壓縮因子為 2；
- $Y = \frac{1}{2}y$ 把圖像沿垂直方向壓縮，壓縮因子為 2；
- 再把 1 加到 X 或 Y 上，則把圖像向右或向上平移 1 個單位。

四個變換後的圖像

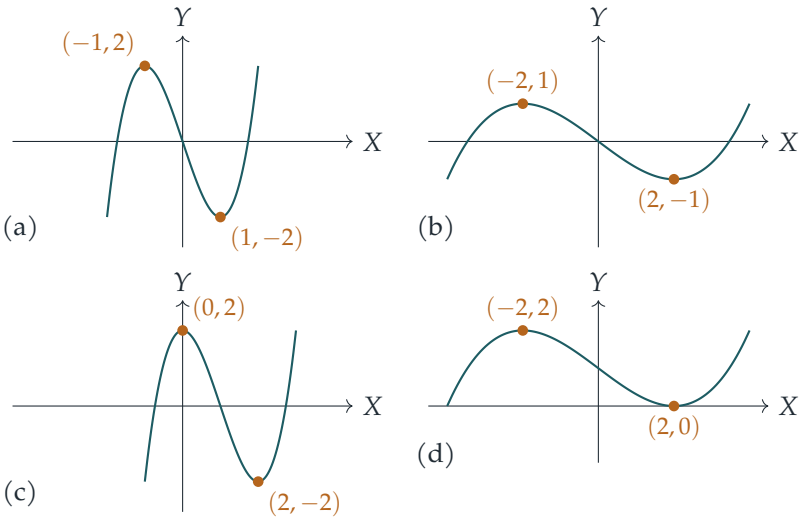
把每個映射分別應用於 $(-2, 2)$、$(2, -2)$ 以及範圍 $-4 \leq x \leq 4$：

(a) $X = \frac{1}{2}x$, $Y = y$：極大點 $(-1, 2)$，極小點 $(1, -2)$，範圍為 $-2 \leq X \leq 2$。曲線是 $Y = X^3 - 3X$。

(b) $X = x$, $Y = \frac{1}{2}y$：極大點 $(-2, 1)$，極小點 $(2, -1)$，範圍為 $-4 \leq X \leq 4$。曲線是 $16Y = X^3 - 12X$。

(c) $X = \frac{1}{2}x + 1, Y = y$：極大點 $(0,2)$，極小點 $(2,-2)$，範圍為 $-1 \le X \le 3$。這是把圖像 (a) 向右移 1 個單位，拐點在 $(1,0)$。

(d) $X = x, Y = \frac{1}{2}y + 1$：極大點 $(-2,2)$，極小點 $(2,0)$，範圍為 $-4 \le X \le 4$。這是把圖像 (b) 向上移 1 個單位，拐點在 $(0,1)$；它在極小點處與 X 軸相切。



最後部分

我們需要 $a, c \neq 0$（否則圖像會塌縮成一條直線），這樣新圖像的轉向點就是

$$(-2a + b,\ 2c + d) \quad 及 \quad (2a + b,\ -2c + d).$$

這兩點必須是 $(0,0)$ 和 $(1,1)$，次序不拘，由此得到兩組線性方程組。

情況 1：$(-2,2) \mapsto (0,0)$ 及 $(2,-2) \mapsto (1,1)$。於是 $-2a + b = 0$ 及 $2a + b = 1$，所以 $b = \frac{1}{2}, a = \frac{1}{4}$；又 $2c + d = 0, -2c + d = 1$，所以 $d = \frac{1}{2}, c = -\frac{1}{4}$。

情況 2：$(-2,2) \mapsto (1,1)$ 及 $(2,-2) \mapsto (0,0)$。同理可得 $a = -\frac{1}{4}, b = \frac{1}{2}, c = \frac{1}{4}, d = \frac{1}{2}$。

所以其中一個答案是

$$a = \tfrac{1}{4}, \quad b = \tfrac{1}{2}, \quad c = -\tfrac{1}{4}, \quad d = \tfrac{1}{2},$$

另一個答案是 $a = -\frac{1}{4}, b = \frac{1}{2}, c = \frac{1}{4}, d = \frac{1}{2}$。

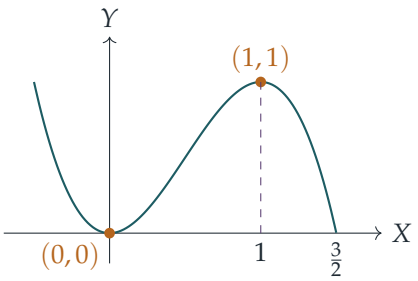
新曲線的方程值得求出來。在情況 1 中，$x = 4X - 2$ 及 $y = 2 - 4Y$，所以

$$\begin{aligned} 8(2 - 4Y) &= (4X - 2)^3 - 12(4X - 2) \\ &= 64X^3 - 96X^2 + 48X - 8 - 48X + 24 \\ &= 64X^3 - 96X^2 + 16, \end{aligned}$$

化簡後得

$$Y = 3X^2 - 2X^3 = X^2(3 - 2X).$$

這條曲線有 $Y' = 6X(1 - X)$，所以它的轉向點確實是位於 $(0,0)$ 的極小點和位於 $(1,1)$ 的極大點。情況 2 給出完全相同的曲線：由於這個三次函數是奇函數，把 (x,y) 換成 $(-x,-y)$ 並不會改變原來的圖像，而這正是兩種情況之間唯一的分別。



當 $-4 \le x \le 4$ 時，這條曲線由 $(-\frac{1}{2}, 1)$ 延伸至 $(\frac{3}{2}, 0)$。

驗算

- 把 $X = 0$ 和 $X = 1$ 代入 $Y = 3X^2 - 2X^3$，分別得 0 和 1，而 $Y' = 6X(1 - X)$ 在這兩處都等於零。
- 每個變換後的圖像都保留了原圖像的一個特徵：曲線在範圍的一端回到某個轉向點的高度。例如在 (d) 中，$Y = 2$ 於 $X = 4$，而 $Y = 0$ 於 $X = -4$。
- 數值上，在 $x = 1$ 處我們有 $y = -\frac{11}{8}$；在情況 1 中這給出 $X = 0.75$, $Y = 0.84375$，而 $3(0.75)^2 - 2(0.75)^3 = 0.84375$。

常見錯誤

- 弄錯伸縮的方向：$X = \frac{1}{2}x$ 令圖像變窄，而不是變闊，因為位於 $x = 2$ 的點會移到 $X = 1$。
- 忘記把定義域（domain）也一併變換：圖像 (c) 的範圍是 $-1 \le X \le 3$，而不是 $-4 \le X \le 4$。
- 最後部分只給出一個解，而沒有留意到轉向點還有另一種對應次序（給出任何一個已經足夠，但第一種情況中的 $c < 0$ 很容易被忽略）。

STEP 之外

映射 $X = ax + b, Y = cy + d$ 是仿射變換(affine transformation)，而最後部分說明，任何有兩個轉向點的三次曲線，都可以用其中一個這樣的變換移到標準曲線 $Y = 3X^2 - 2X^3$ 上；這就是電腦圖像學中所用的「smoothstep」函數。像這樣按某個變換羣（group of transformations）來為曲線分類，是貫穿幾何與線性代數（linear algebra）的一個主題。至於這個故事中與矩陣有關的一面，可參閱 Strang 的 *Introduction to Linear Algebra*（線性代數入門）。

詳細解答

STEP I 1989，第 10 題

力學 • 與速度相關的阻力

> **STEP I 1989，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/89-s1-q10

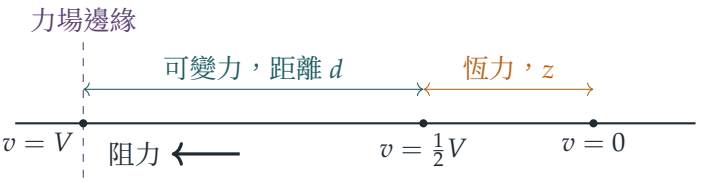
關鍵思路

阻力取決於速率 v，而非時間，而我們要求的是距離。這正是提示我們把加速度（acceleration）寫成 $v\dfrac{\mathrm{d}v}{\mathrm{d}x}$ 的信號：這樣一來，牛頓第二定律（Newton's second law）便化為一條直接聯繫 v 與 x 的可分離變數方程（separable equation）。第二部分是勻減速問題，所以唯一要做的，就是求出轉變一刻的力，再讓它在 $\pm 10\%$ 的範圍內變動。

解答

模型

我們把太空船視為質量為 M、沿直線運動的質點（particle）。設 x 為進入力場後所行的距離，v 為速率，且在 $x = 0$ 時 $v = V$。唯一的力是阻力（resistance），其方向向後，因此速率一直遞減。



速率減半所需的距離

取運動方向為正，並把加速度寫成 $v\dfrac{\mathrm{d}v}{\mathrm{d}x}$，由牛頓第二定律得

$$M v \frac{\mathrm{d}v}{\mathrm{d}x} = -\frac{M\omega\,(v^2 + V^2)}{v} \qquad \Longrightarrow \qquad \frac{v^2}{v^2 + V^2}\,\mathrm{d}v = -\omega\,\mathrm{d}x.$$

想一想 怎樣求 $\dfrac{v^2}{v^2+V^2}$ 的積分？

把分子寫成 $(v^2+V^2)-V^2$，於是

$$\frac{v^2}{v^2+V^2} = 1 - \frac{V^2}{v^2+V^2}, \qquad \int \frac{v^2}{v^2+V^2}\,\mathrm{d}v = v - V\arctan\frac{v}{V} + C,$$

這裏用了標準積分（standard integral）$\int \frac{\mathrm{d}v}{v^2+V^2} = \frac{1}{V}\arctan\frac{v}{V}$。

設 d 為速率降至 $\frac{1}{2}V$ 時所行的距離。由 $(x,v)=(0,V)$ 積分至 $(d,\frac{1}{2}V)$，

$$\begin{aligned}\omega d &= \int_{V/2}^{V} \frac{v^2}{v^2+V^2}\,\mathrm{d}v = \Big[v - V\arctan\frac{v}{V}\Big]_{V/2}^{V} \\ &= \frac{V}{2} - V\Big(\arctan 1 - \arctan\tfrac{1}{2}\Big) = V\Big(\frac{1}{2} - \frac{\pi}{4} + \arctan\frac{1}{2}\Big).\end{aligned}$$

因此所求的距離為

$$d = \frac{V}{\omega}\Big(\frac{1}{2} - \frac{\pi}{4} + \arctan\frac{1}{2}\Big) = \frac{V}{\omega}\Big(\frac{1}{2} - \arctan\frac{1}{3}\Big) \approx 0.178\,\frac{V}{\omega}.$$

第二個形式用了 $\arctan 1 - \arctan\frac{1}{2} = \arctan\dfrac{1-\frac{1}{2}}{1+\frac{1}{2}} = \arctan\frac{1}{3}$，這來自正切的差角公式（tangent subtraction formula）（兩個角都在 $(0,\frac{\pi}{2})$ 之內，所以沒有歧義）。兩個形式都是完整的答案。

恆力

在轉變前的一刻，$v=\frac{1}{2}V$，所以阻力的大小為

$$\frac{M\omega\left(\frac{1}{4}V^2+V^2\right)}{\frac{1}{2}V} = \frac{5}{2}M\omega V.$$

新的恆力 F 與此值相差不超過 10%，所以

$$\frac{9}{10}\cdot\frac{5}{2}M\omega V \le F \le \frac{11}{10}\cdot\frac{5}{2}M\omega V, \qquad \text{即} \qquad \frac{9}{4}M\omega V \le F \le \frac{11}{4}M\omega V.$$

在恆力作用下，減速度（deceleration）$a = F/M$ 是常數，所以由 $v^2 = u^2 - 2as$，取 $u=\frac{1}{2}V$ 及末速率為 0，得

$$z = \frac{(V/2)^2}{2a} = \frac{MV^2}{8F}.$$

這個量隨 F 增大而減小，所以最大的力對應最小的 z，反之亦然：

$$\frac{MV^2}{8\cdot\frac{11}{4}M\omega V} \le z \le \frac{MV^2}{8\cdot\frac{9}{4}M\omega V}, \qquad \text{亦即} \qquad \boxed{\frac{V}{22\,\omega} \le z \le \frac{V}{18\,\omega}}$$

（若把「相差不超過 10%」理解為嚴格的範圍，則不等式為嚴格不等式）。假如力完全不變，則 $z = V/(20\omega)$，位於這個範圍的中間。

驗算

- 量綱（dimensions）：ω 的單位是 s^{-1}（因為 $M\omega v$ 必須是力），所以 V/ω 是長度，正如兩個答案所要求的。
- 係數 $\frac{1}{2} - \arctan\frac{1}{3} \approx 0.1782$ 是正數，且小於 $\frac{1}{2}$。這是應該的：太空船的速率介乎 $\frac{1}{2}V$ 與 V 之間，而減速度至少為 $2\omega V$，所以這段行程最多歷時 $1/(4\omega)$，所行距離小於 $V/(4\omega)$。
- 取 $V = \omega = 1$，以數值方法逐步計算 $\dot{v} = -\omega(v^2 + V^2)/v$，得出 $v = \frac{1}{2}$ 時的距離為 0.17825；再逐步計算恆力階段，當力為原來的 110%、100% 及 90% 時，分別得 $z = 0.0455$、0.0500、0.0556，與 $\frac{1}{22}, \frac{1}{20}, \frac{1}{18}$ 吻合。

常見錯誤

- 用了 $\dfrac{\mathrm{d}v}{\mathrm{d}t}$ 而不是 $v\dfrac{\mathrm{d}v}{\mathrm{d}x}$：這樣求得的是時間而不是距離，其後還要多做一次積分。
- 把上下限配反了：力愈大，停下來所需的距離愈短，所以 1.1 倍的力對應的是下限 $V/(22\omega)$。

STEP 之外

寫出 $a = v\,\mathrm{d}v/\mathrm{d}x$，是初次窺見一個更宏大的想法：對於只取決於位置或速率的一維力，我們可以用另一個變數取代時間，從而降低方程的階（order）；而當力只取決於位置時，這就引出能量守恆（energy conservation）以及描述運動的相平面（phase plane）圖像。第二部分是一道關於敏感度（sensitivity）的小練習：數據中的不確定性（這裏是力的 $\pm 10\%$）怎樣傳遞到答案之中（z 大約相差 $\mp 10\%$）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）正是以這種精神處理與速度有關的力。

詳細解答

STEP I 1989，第 11 題

力學 • 拋體射程的最優化

STEP I 1989，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/89-s1-q11

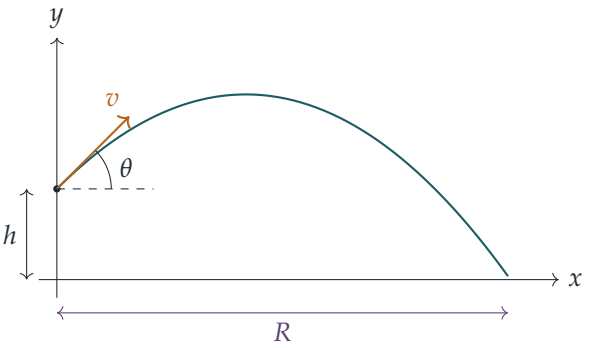
關鍵思路

飛行時間來自一條關於 t 的二次方程（quadratic equation）：鉛球的高度回到零時便落地。取正根，再乘以水平速率，便得到 R；餘下的工作只是把平方根整理成題目要求的形式。至於第二部分，「θ 稍為減小，R 會否增大？」其實是問 $\mathrm{d}R/\mathrm{d}\theta$ 在 $\theta = 45°$ 時的正負號，而對稱關係 $\sin 45° = \cos 45°$ 令這個導數（derivative）化簡得十分俐落。

解答

模型

我們把鉛球視為只受重力作用的質點（particle），從點 $(0, h)$ 以速率 v、與水平成仰角 θ 擲出，其中 $0 < \theta < \frac{\pi}{2}$ 及 $h > 0$。取 x 軸為水平方向並指向投擲的方向，y 軸由地面起豎直向上。



射程

經過時間 t 後，

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

鉛球在 $y = 0$ 時着地，即

$$\tfrac{1}{2}gt^2 - (v\sin\theta)\,t - h = 0 \quad\Longrightarrow\quad t = \frac{v\sin\theta \pm \sqrt{v^2\sin^2\theta + 2gh}}{g}.$$

由於 $h > 0$，兩根之積 $-2h/g$ 為負數，所以恰好有一個根是正數；取負號得到的是負根（即擲出之前的時刻），因此我們取正號。於是

$$\begin{aligned} R = v\cos\theta\, t &= \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gh}\right) \\ &= \frac{v^2\sin\theta\cos\theta}{g}\left(1 + \sqrt{1 + \frac{2gh}{v^2\sin^2\theta}}\right) \qquad \text{（提出 } v\sin\theta > 0\text{）} \\ &= \frac{v^2}{2g}\sin 2\theta\left(1 + \sqrt{1 + \frac{2hg}{v^2\sin^2\theta}}\right), \end{aligned}$$

其中用到 $2\sin\theta\cos\theta = \sin 2\theta$。這就是所要求的結果。

? **想一想** 為甚麼可以把 $v\sin\theta$ 以 $v^2\sin^2\theta$ 的形式放進平方根之內？

因為 $\sqrt{a^2 b} = a\sqrt{b}$ 只在 $a \geq 0$ 時成立。這裏 $0 < \theta < \frac{\pi}{2}$，所以 $v\sin\theta > 0$，這一步是成立的。（若是向下擲出，這一步便不成立，題目所給的公式也會出錯。）

45° 是最佳角度嗎？

求導時，用未化簡的形式較為方便。寫成

$$R(\theta) = \frac{v}{g}\cos\theta\,(v\sin\theta + S), \qquad S = \sqrt{v^2\sin^2\theta + 2gh},$$

於是 $\dfrac{\mathrm{d}S}{\mathrm{d}\theta} = \dfrac{v^2\sin\theta\cos\theta}{S}$。由積法則（product rule），

$$\frac{\mathrm{d}R}{\mathrm{d}\theta} = \frac{v}{g}\left[-\sin\theta\,(v\sin\theta + S) + \cos\theta\left(v\cos\theta + \frac{v^2\sin\theta\cos\theta}{S}\right)\right].$$

在 $\theta = 45°$ 時，設 $\sin\theta = \cos\theta = c = \frac{1}{\sqrt{2}}$。$-vc^2$ 與 $+vc^2$ 兩項互相抵銷，餘下

$$\begin{aligned} \left.\frac{\mathrm{d}R}{\mathrm{d}\theta}\right|_{\theta=45°} &= \frac{vc}{g}\left(-S + \frac{v^2c^2}{S}\right) = \frac{vc}{gS}\left(v^2c^2 - S^2\right) \\ &= \frac{vc}{gS}\left(\tfrac{1}{2}v^2 - \tfrac{1}{2}v^2 - 2gh\right) = -\frac{2vhc}{S} = -\frac{\sqrt{2}\,vh}{\sqrt{v^2 + 4gh}}, \end{aligned}$$

其中最後一步用到 $S = \sqrt{\frac{1}{2}v^2 + 2gh} = \frac{1}{\sqrt{2}}\sqrt{v^2 + 4gh}$。

由於 $v > 0$ 及 $h > 0$，這個導數嚴格為負。所以在 45° 附近，R 是 θ 的嚴格遞減函數（strictly decreasing function），而當減幅 $\delta\theta > 0$ 足夠小時，便有 $R(45° - \delta\theta) \approx R(45°) + \frac{\sqrt{2}\,vh}{\sqrt{v^2+4gh}}\,\delta\theta > R(45°)$。**答案是會的：$\theta$ 稍為減小，R 便會增大。**

從物理角度看：鉛球從高處擲出，留在空中的時間比在平地上擲出時稍長，而這段額外的時間用來作水平移動更為划算，所以擲得平一點會有好處。只有當 $h = 0$ 時，導數才等於零，45° 才是最佳角度。

驗算

- 當 $h = 0$ 時，公式給出 $R = \frac{v^2}{g}\sin 2\theta$，即熟悉的平地射程，而在 45° 的導數為 0，正如所料。
- 單位：v^2/g 是長度，而 $2hg/v^2$ 是無量綱（dimensionless）的量。
- 數值上，取 $g = 9.8$、$v = 13\,\mathrm{m\,s^{-1}}$ 及 $h = 2\,\mathrm{m}$：$R(45°) \approx 19.05\,\mathrm{m}$，但 $R(40°) \approx 19.10\,\mathrm{m}$。把運動按時間逐步推算，所得結果與 R 的公式吻合；而對許多隨機選取的 v 和 h，在 45° 以有限差分（finite difference）計算的導數亦與 $-\sqrt{2}\,vh/\sqrt{v^2+4gh}$ 相符。

常見錯誤

- 解 t 的二次方程時取錯了根，或忘記說明為何捨去負根。
- 推論說「$\sin 2\theta$ 在 45° 時最大，所以 R 也是」。括號內的式子同樣取決於 θ（它隨 θ 增大而減小），所以這個論證並不完整。

STEP 之外

一般地令 $\mathrm{d}R/\mathrm{d}\theta = 0$，可證明從高度 h 擲出的最佳角度滿足 $\tan\theta = v/\sqrt{v^2 + 2gh}$，這個角度略小於 45°（以上述數值計算約為 42°）；現實中的鉛球運動員出手角度還要更低，因為角度較平時他們可以推得更用力。像這樣對一族軌跡求最優，便會引出這族曲線的*包絡線*（envelope），即「安全拋物線」（parabola of safety），它劃出了鉛球所能到達的每一點。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門教材）對拋體運動（projectile motion）及其各種變化有細緻的論述。

詳細解答

STEP I 1989，第 12 題

力學 • 彈性繩與能量

STEP I 1989，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/89-s1-q12

關鍵思路

一切都建基於正四面體（regular tetrahedron）的幾何：它的質心（centre of mass）與頂點、與面、以及與通過頂點的軸之間的距離各是多少。一旦求得這三個長度，第一部分只不過是虎克定律（Hooke's law）的應用。至於第二部分，對稱性確保運動是純粹的鉛垂平移。四面體開始和結束時都處於靜止，所以我們根本不需要運動方程（equation of motion）：只要考慮重力勢能與彈性勢能（elastic potential energy）之間的能量平衡，便可解決問題。

解答

四面體的幾何

我們把各桿視為均勻剛桿，把各弦線視為符合虎克定律的輕彈性弦線。每根桿的質量可視為集中在其中點。每個頂點恰好屬於六根桿中的三根，所以六個中點的平均等於四個頂點的平均。因此質心 G 就是四面體的形心（centroid）。

面 BCD 是邊長為 $2a$ 的等邊三角形，所以它的中心 N 與 B、C、D 每一點的距離都是

$$\rho = \frac{2a}{\sqrt{3}}$$

。四面體的高是 $AN = \sqrt{(2a)^2 - \rho^2} = \sqrt{\frac{8}{3}}\,a = \frac{2\sqrt{6}}{3}a$。形心位於 AN 上，在由該面往上四分之一的位置。所以

$$GN = r = \frac{\sqrt{6}}{6}\,a, \qquad GA = GB = GC = GD = R = \frac{\sqrt{6}}{2}\,a.$$

（驗算：$R^2 = \rho^2 + r^2 = \frac{4}{3}a^2 + \frac{1}{6}a^2 = \frac{3}{2}a^2$。）

張力

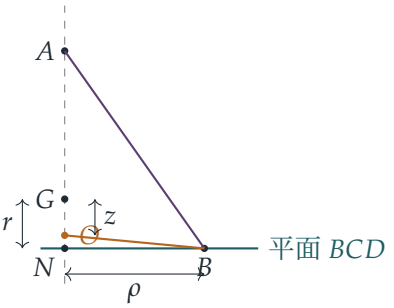
當 X 位於 G 時，每條弦線的長度為 $R = \frac{\sqrt{6}}{2}a$，伸長量（extension）為 $R - a = \frac{\sqrt{6}-2}{2}a > 0$，所以四條弦線全部拉緊。根據虎克定律，每條弦線的張力（tension）為

$$T = \frac{\lambda(R-a)}{a} = \frac{\sqrt{6}-2}{2}\lambda \approx 0.225\,\lambda.$$

XA 斷裂之後

斷裂之前，四個相等的張力沿着指向 G 的四個方向拉動各頂點。由於 G 是形心，由 G 指向各頂點的單位向量之和為零，所以各弦線不產生淨力，桌面承受全部重量 Mg。

當 XA 斷裂時，餘下三條弦線的合力大小為 T，方向鉛垂向上，即等於 XA 原有拉力的負值。在水平方向上，由於繞鉛垂線 AN 有三重對稱性，三個拉力互相抵銷。每個拉力都位於通過該線的鉛垂平面內，所以它們不產生扭轉力偶（couple），而它們的傾側力偶亦因對稱而互相抵銷。因此四面體的運動是純粹的鉛垂平移，AN 始終留在通過 O 的鉛垂線上。它之所以離開桌面，是因為 $T > Mg$，這一點我們會在最後確認。



設 z 為 G 高於 O 的高度。那麼 B 與通過 O 的鉛垂線的水平距離為 ρ，相對於 O 的高度為 $z - r$，所以三條弦線每條的長度都是

$$L(z) = \sqrt{\rho^2 + (z-r)^2}.$$

開始時 $z = 0$，而 $L = \sqrt{\rho^2 + r^2} = R$。在最高點，$BCD$ 與 O 同高，所以 $z = r$，而 $L = \rho = \frac{2}{\sqrt{3}}a$。當 $0 \le z \le r$ 時，L 一直介乎 ρ 與 R 之間，兩者都大於 a，所以弦線自始至終保持拉緊。

想一想 怎樣才可以不必解一條棘手的非線性運動方程，而求出最大高度？

桌面光滑，弦線為輕弦線，所以機械能（mechanical energy）守恆。由於四面體立即離開桌面，法向反作用力（normal reaction）不作功。四面體在開始時和在最高點時都處於靜止，所以動能不必考慮。三條弦線所失去的彈性能量 $3 \cdot \frac{\lambda(L-a)^2}{2a}$，必定等於所獲

得的重力勢能 Mgr：

$$Mg\,r = \frac{3\lambda}{2a}\Big[(R-a)^2 - (\rho - a)^2\Big].$$

把方括號內的式子因式分解，並利用 $R^2 - \rho^2 = \frac{3}{2}a^2 - \frac{4}{3}a^2 = \frac{1}{6}a^2$，

$$\begin{aligned}(R-a)^2 - (\rho-a)^2 &= R^2 - \rho^2 - 2a(R-\rho)\\ &= a^2\Big(\tfrac{1}{6} - \sqrt{6} + \tfrac{4}{\sqrt{3}}\Big).\end{aligned}$$

由於 $\dfrac{3}{2ar} = \dfrac{3}{2a}\cdot\dfrac{6}{\sqrt{6}\,a} = \dfrac{3\sqrt{6}}{2a^2}$，由此可得

$$\begin{aligned}\frac{Mg}{\lambda} &= \frac{3\sqrt{6}}{2}\Big(\tfrac{1}{6} - \sqrt{6} + \tfrac{4}{\sqrt{3}}\Big) = \frac{\sqrt{6}}{4} - 9 + 6\sqrt{2}\\ &\approx 0.6124 - 9 + 8.4853 = 0.0977,\end{aligned}$$

準確至 2 位有效數字即為 0.098，正如題目所求。

最後，$Mg \approx 0.098\lambda < 0.225\lambda \approx T$，所以剛斷裂後的向上淨力 $T - Mg$ 為正數，四面體確實會升離桌面，與題目所述一致。

驗算

- 取 $a = 1$，我們以坐標表示四個頂點，確認全部六條棱的長度都是 2，四個頂點與形心的距離都是 $\frac{\sqrt{6}}{2} \approx 1.2247$，並重新得出 $T \approx 0.2247\lambda$。
- 取 $Mg/\lambda = 6\sqrt{2} + \frac{\sqrt{6}}{4} - 9 \approx 0.09765$，我們由 $z = 0$ 處的靜止狀態出發，以數值方法逐步計算運動方程 $M\ddot{z} = 3T(L)\,\frac{r-z}{L} - Mg$。速度首次回到零時 $z \approx 0.40825$，與 $r = \frac{\sqrt{6}}{6}$ 吻合。

常見錯誤

- 把 R 當作形心到面的距離。形心位於由頂點往下 $\frac{3}{4}$ 的位置，所以 $R = 3r$。
- 忘記弦線在最高點時仍然處於拉伸狀態。每條弦線在該處的能量是 $\frac{\lambda(\rho-a)^2}{2a}$，而不是零。
- 計算伸長量 $L - a$ 時，把 L 取作由 O 到面的中心 N 的距離，而不是由 O 到頂點的距離。

STEP 之外

這裏所用的形心技巧（六根桿的中點的平均等於四個頂點的平均）是以下一般事實的一個小例子：一個系統的質心，是其各部分質心按質量加權的平均。把運動視為純粹平移，靠的是對稱性論證；這類論證在剛體動力學（rigid-body dynamics）中會變得有系統，因為物體的運動可分解為質心的平移和繞質心的轉動。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門）對這兩個想法都有細緻的闡述。

詳細解答

STEP I 1989，第 13 題

力學 • 涉及摩擦力的梯子平衡

STEP I 1989，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：

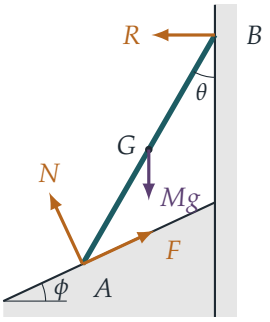
step.maths.org/questions/89-s1-q13

關鍵思路

這是一道三個力再加摩擦力的平衡（equilibrium）問題，竅門在於選對方向。以梯腳為支點取力矩（moments），可同時消去 N 和 F，立刻得出牆的反作用力。接着，我們不沿水平和鉛垂方向分解，而是*沿垂直於斜面和平行於斜面的方向*分解。這樣得出的兩條方程各自只含未知量 N 和 F 的其中一個，題目給出的公式便直接得出。

解答

我們把梯子模型化為一根長度為 $2a$ 的均勻桿（uniform rod）AB，其質心（centre of mass）G 位於中點。梯腳 A 置於粗糙斜面上，梯頂 B 靠在光滑的牆上，因此牆的反作用力 R 沿水平方向，指向遠離牆的一方。在 A 點有垂直於斜面的法向反作用力（normal reaction）N 和沿斜面的摩擦力（friction）F。梯子若要滑動，梯腳會沿斜面向下滑，遠離牆壁，所以我們取 F 的方向為沿斜面向上、指向牆壁。（稍後求得的 F 的公式為正數，印證了這個選擇。）



牆的反作用力

以 A 為支點取力矩，可消去 N 和 F。梯子與鉛垂線的夾角為 θ，所以 B 比 A 高 $2a\cos\theta$，而 G 與 A 的水平距離為 $a\sin\theta$。作用於 B 的水平力 R 的力矩為 $R\cdot 2a\cos\theta$，重量的力矩為 $Mg\cdot a\sin\theta$，轉向相反，因此

$$2aR\cos\theta = Mga\sin\theta \qquad\Longrightarrow\qquad R = \tfrac{1}{2}Mg\tan\theta.$$

沿垂直於斜面和平行於斜面的方向分解

想一想 應該沿哪些方向分解，才能令每條方程只含一個未知量？

沿垂直於斜面的方向和沿斜面的方向分解（resolve）。F 在垂直於斜面的方向沒有分量，N 在沿斜面的方向沒有分量。我們需要求出另外兩個力的分量。

- 重量 Mg 有壓向斜面的分量 $Mg\cos\phi$ 和沿斜面向下的分量 $Mg\sin\phi$。
- 斜面向着牆壁上升，所以它向上的法線朝遠離牆的一方傾斜，而水平力 R（同樣指向遠離牆的一方）沿該法線的分量為 $R\sin\phi$，即指向斜面之外。沿斜面方向，R 指向遠離牆的一方，亦即下坡方向，分量為 $R\cos\phi$。

垂直於斜面：

$$N + R\sin\phi = Mg\cos\phi,$$
$$\text{所以}\quad N = Mg\cos\phi - \tfrac{1}{2}Mg\tan\theta\sin\phi = Mg\left(\cos\phi - \frac{\tan\theta\sin\phi}{2}\right).$$

沿斜面（以沿斜面向上為正）：

$$F = Mg\sin\phi + R\cos\phi \quad\Longrightarrow\quad F = Mg\left(\sin\phi + \frac{\tan\theta\cos\phi}{2}\right).$$

這就是題目要求的兩個結果。

最大的角

梯腳不滑動的條件是：接觸得以維持，即 $N > 0$；以及摩擦力不超過其上限，即 $F \le \mu N$，其中 $\mu = 2$。（這裏 $F > 0$，所以這是唯一的摩擦條件。）當 $\phi = 45^\circ$ 時，$\sin\phi = \cos\phi = \frac{1}{\sqrt{2}}$，記 $t = \tan\theta$，則

$$F = \frac{Mg}{\sqrt{2}}\left(1 + \frac{t}{2}\right), \qquad N = \frac{Mg}{\sqrt{2}}\left(1 - \frac{t}{2}\right).$$

所以 $N > 0$ 要求 $t < 2$，而 $F \le 2N$ 則變成

$$1 + \frac{t}{2} \le 2 - t \quad\Longleftrightarrow\quad \frac{3t}{2} \le 1 \quad\Longleftrightarrow\quad \tan\theta \le \frac{2}{3}.$$

這比 $t < 2$ 更嚴格，所以它才是真正起限制作用的條件。由於 $\tan\theta$ 在 $0 < \theta < 90^\circ$ 上遞增，梯子能夠保持平衡的最大角為

$$\theta = \arctan\tfrac{2}{3} \approx 33.7^\circ,$$

而在這個角度下，摩擦力達到極限值（limiting friction）。

驗算

- 當 $\phi = 0$（水平地面）時，公式給出 $N = Mg$ 和 $F = \frac{1}{2}Mg\tan\theta = R$，正是梯子置於水平地面時為人熟悉的結果。
- 當 $\theta = 0$（梯子鉛垂）時，牆不起任何作用，而 $N = Mg\cos\phi$，$F = Mg\sin\phi$，與質點靜止於斜面上的情況完全一樣。
- 數值上，對隨機選取的 θ 和 ϕ 直接解水平和鉛垂方向的力方程，所得結果與兩條公式吻合，誤差在 10^{-13} 以內；而當 $\phi = 45^\circ$ 時，比值 F/N 在 $\theta = 33.690^\circ$ 處剛好越過 2。

常見錯誤

- 弄錯 $R\sin\phi$ 一項的正負號。由於斜面向着牆壁上升，R 有一部分把梯腳拉離斜面，令 N 減小。
- 沿水平和鉛垂方向分解，然後在解所得的聯立方程時出錯。這個做法行得通，但沿斜面和垂直於斜面的方向分解要快得多。

STEP 之外

留意斜面越陡，梯子可以傾斜的程度就越小，因為牆的反作用力一方面令 N 減小，另一方面又令 F 增大。這類問題是靜力學（statics）的起點：我們選擇合適的坐標軸和取矩點，把各個未知量分離開來。同樣的想法也是分析桁架（trusses）和框架結構（frameworks）的基礎，更廣泛而言也是分析穩定性的基礎。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對剛體平衡和摩擦有嚴謹細緻的論述。

詳細解答

STEP I 1989，第 14 題

概率與統計 • 條件分佈

STEP I 1989，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/89-s1-q14

關鍵思路

一切都建基於一組概率：種子落在每一條闊一里格的地帶內的機會。種子當且僅當落在園林或農地時才會長成樹，所以樹的數目服從二項分佈（binomial distribution）。樹生長之處的距離 Y，其實就是在「落在陸地上」這個條件下的 L。它的密度函數（density）就是把 f 只保留陸地地帶的部分，再按比例放大，使總面積等於 1。最後一部分是假設檢驗（hypothesis test）：對一個二項計數而言，農地上有 23 棵樹是否多得出奇？

解答

各地帶的概率

L 的累積分佈函數（cumulative distribution function）在 $0 \le x \le 4$ 時為

$$\mathrm{F}(x) = \int_0^x \left(\tfrac{1}{2} - \tfrac{1}{8}t\right)\mathrm{d}t = \frac{x}{2} - \frac{x^2}{16},$$

所以 $\mathrm{F}(1) = \frac{7}{16}$、$\mathrm{F}(2) = \frac{3}{4}$、$\mathrm{F}(3) = \frac{15}{16}$ 及 $\mathrm{F}(4) = 1$。因此

$$\mathrm{P}(\text{園林}) = \mathrm{P}(0 \le L \le 1) = \tfrac{7}{16}, \qquad \mathrm{P}(\text{湖}) = \tfrac{3}{4} - \tfrac{7}{16} = \tfrac{5}{16},$$
$$\mathrm{P}(\text{農地}) = \mathrm{P}(2 \le L \le 3) = \tfrac{15}{16} - \tfrac{3}{4} = \tfrac{3}{16}, \qquad \mathrm{P}(\text{海}) = \tfrac{1}{16}.$$

（由於 L 是連續的，各邊界的概率為零，所以邊界屬於哪一條地帶並無影響。）一顆種子長成樹的概率為

$$p = \tfrac{7}{16} + \tfrac{3}{16} = \tfrac{5}{8}.$$

設 T 為一年內長出的樹的數目。由於 96 顆種子各自獨立地被風吹走，每顆長成樹的概率為 $\frac{5}{8}$，故 $T \sim \mathrm{B}(96, \frac{5}{8})$，而每年新長出的發燒樹數目的平均值為

$$\mathrm{E}(T) = 96 \times \tfrac{5}{8} = 60.$$

第 (i) 部分

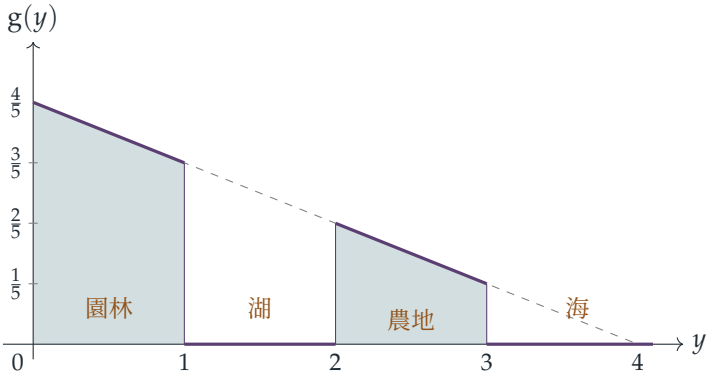
Y 的分佈，就是已知種子落在陸地上時 L 的分佈。以 $A = [0,1] \cup [2,3]$ 表示陸地。對於 $y \in A$，

$$\mathrm{P}(y \le Y \le y + \delta y) = \frac{\mathrm{P}(y \le L \le y + \delta y)}{\mathrm{P}(L \in A)} \approx \frac{\mathrm{f}(y)\,\delta y}{5/8},$$

而 Y 永遠不會取 A 以外的值。所以 Y 的密度函數為

$$\mathrm{g}(y) = \begin{cases} \frac{8}{5}\left(\frac{1}{2} - \frac{1}{8}y\right) = \dfrac{4-y}{5} & \text{當} 0 \le y \le 1 \text{ 或} 2 \le y \le 3, \\ 0 & \text{其他情況。} \end{cases}$$

驗算一下，它的面積是 $\frac{8}{5}\left(\frac{7}{16} + \frac{3}{16}\right) = 1$。圖像是同一條下降直線上的兩段：在園林上由 $\frac{4}{5}$ 降至 $\frac{3}{5}$，在湖上為零，然後在農地上由 $\frac{2}{5}$ 降至 $\frac{1}{5}$。



虛線是 $\frac{8}{5}\mathrm{f}$，畫出來以作比較；Y 的密度函數只在陸地上才與它重合。

至於平均值，$\int y(4-y)\,\mathrm{d}y = 2y^2 - \frac{1}{3}y^3$，所以

$$\begin{aligned} \mathrm{E}(Y) &= \frac{1}{5}\left[2y^2 - \tfrac{1}{3}y^3\right]_0^1 + \frac{1}{5}\left[2y^2 - \tfrac{1}{3}y^3\right]_2^3 \\ &= \frac{1}{5}\left(\frac{5}{3}\right) + \frac{1}{5}\left(9 - \frac{16}{3}\right) = \frac{1}{5}\cdot\frac{5}{3} + \frac{1}{5}\cdot\frac{11}{3} = \frac{16}{15}. \end{aligned}$$

所以新樹平均生長在城堡以南 $\frac{16}{15}$ 里格處。這個位置剛過了園林，落在湖中，但這並無不妥：平均值不一定是變量可以取到的值。

第 (ii) 部分

設 X 為一年內在農地長出的樹的數目。若有依從傳統，則放出的種子有 96 顆，每顆獨立地以概率 $\frac{3}{16}$ 落在農地，所以

$$X \sim \mathrm{B}\left(96, \tfrac{3}{16}\right), \qquad \mathrm{E}(X) = 18, \qquad \mathrm{Var}(X) = 96 \cdot \tfrac{3}{16} \cdot \tfrac{13}{16} = \tfrac{117}{8} \approx 14.6,$$

由此得標準差（standard deviation）約為 3.82。

> **想一想** 我們究竟在檢驗甚麼？朝哪個方向檢驗？

零假設（null hypothesis）是傳統得到依從：種子有 96 顆，所以 $X \sim \mathrm{B}\left(96, \frac{3}{16}\right)$。農地上的樹比平常多，最自然的懷疑是放出的種子多於 96 顆，所以我們採用單尾檢驗（one-tailed test），問在零假設下出現 23 棵或以上的可能性有多大。

由於 $np = 18$ 和 $n(1-p) = 78$ 都夠大，我們採用正態近似（normal approximation）$X \approx \mathrm{N}(18, 14.625)$，並作連續性修正（continuity correction）：

$$\mathrm{P}(X \geq 23) \approx \mathrm{P}\left(Z \geq \frac{22.5 - 18}{3.824}\right) = \mathrm{P}(Z \geq 1.18) \approx 0.12.$$

（二項分佈的精確值是 0.121。）所以在正常情況下，農地上有 23 棵或以上的樹，大約每八年便有一年出現。這遠遠未達 5% 水平（甚至 10% 水平）的顯著性。這個觀察結果與傳統得到依從的說法相當吻合，單憑這項證據，國王的懷疑並無充分理由。

作為比較，即使在 5% 水平下，數目也要達到至少 25（精確值 $\mathrm{P}(X \geq 25) \approx 0.049$），國王才算有合理的根據。假如他同時掌握園林的樹木數目，便可以檢驗樹的總數 $T \sim \mathrm{B}\left(96, \frac{5}{8}\right)$，這對於判斷放出了多少顆種子，會是好得多的證據。

驗算

- 四個地帶的概率 $\frac{7}{16}, \frac{5}{16}, \frac{3}{16}, \frac{1}{16}$ 相加得 1，而 Y 的密度函數的總面積為 1。
- 用中點法則作數值積分得 $\mathrm{E}(Y) = 1.06667$，與 $\frac{16}{15}$ 吻合。模擬 400 000 顆種子（取 $L = 4 - 4\sqrt{1-U}$，其中 U 在 $[0,1]$ 上均勻分佈），結果有 62.5% 落在陸地上，而這些種子的平均落點距離為 1.069。
- 把二項概率逐項精確相加，得 $\mathrm{P}(X \geq 23) = 0.1213$，與正態估計值 0.120 相近。

常見錯誤

- 直接把 f 當作 Y 的密度函數。它在陸地上的面積只有 $\frac{5}{8}$，所以必須除以 $\frac{5}{8}$。
- 計算了 $\mathrm{E}(L\,;\, L \in A) = \int_A y\,\mathrm{f}(y)\,\mathrm{d}y = \frac{2}{3}$，卻忘記除以 $\mathrm{P}(L \in A)$。

- 檢驗 $\mathrm{P}(X = 23)$ 而非尾部概率 $\mathrm{P}(X \geq 23)$。X 取任何一個個別數值的可能性本來就不大。

STEP 之外

第 (i) 部分是條件分佈（conditional distribution）的一個例子：把隨機變量限制在某個事件上，再重新歸一化。密度函數形式的貝葉斯定理（Bayes' theorem），以及統計學中隨處可見的截尾分佈（truncated distribution），背後都是同一種運算。第 (ii) 部分讓你初嘗假設檢驗和 p 值，大學一年級的課程會用奈曼－皮爾遜（Neyman–Pearson）框架把這些概念嚴格化。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門教材）講解條件分佈時，附有大量像本題這樣帶故事情境的例題。

詳細解答

STEP I 1989，第 15 題

概率與統計 • 分佈的比較

STEP I 1989，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/89-s1-q15

關鍵思路

以 $12\,\mathrm{km\,h^{-1}}$ 的速率行駛，每公里剛好需時 5 分鐘，所以每條路線的行程時間都是一個固定的分鐘數，再加上某個隨機變量（random variable）的簡單倍數：這個變量分別服從均勻分佈（uniform distribution）、泊松分佈（Poisson distribution）或二項分佈（binomial distribution）。於是，每一個關於時間的問題，都變成關於這一個變量的問題。中間的一問則是貝葉斯定理（Bayes' theorem）的應用，其中三條路線被選中的可能性相等。

解答

行程時間

設 T_A、T_B、T_D 為三條路線的行程時間，以分鐘計。由於 $12\,\mathrm{km\,h^{-1}}$ 即每 5 分鐘行 1 km：

- Angle Avenue：$T_A = 25 + A$，其中 $A \sim \mathrm{U}(0, 10)$。
- Bend Boulevard：$T_B = 20 + 3B$，其中 $B \sim \mathrm{Po}(4)$。
- Detour Drive：設 X 為繞道的次數。假設五處道路工程互相獨立（independent），每處導致繞道的概率均為 $\frac{4}{5}$，則 $X \sim \mathrm{B}(5, \frac{4}{5})$，路程為 $2 + X$ km，而 $T_D = 10 + 5X$。

至少 27 分鐘

Angle Avenue。$T_A \geq 27 \iff A \geq 2$，所以

$$\mathrm{P}(T_A \geq 27) = \frac{10 - 2}{10} = \frac{4}{5}.$$

Bend Boulevard。$T_B \geq 27 \iff 3B \geq 7 \iff B \geq 3$，因為 B 是整數。所以

$$\mathrm{P}(T_B \geq 27) = 1 - \mathrm{e}^{-4}\left(1 + 4 + \frac{4^2}{2!}\right) = 1 - 13\mathrm{e}^{-4} \approx 0.7619.$$

Detour Drive。$T_D \geq 27 \iff 5X \geq 17 \iff X \geq 4$。所以

$$\mathrm{P}(T_D \geq 27) = 5\left(\frac{4}{5}\right)^4\frac{1}{5} + \left(\frac{4}{5}\right)^5 = \left(\frac{4}{5}\right)^4\left(1 + \frac{4}{5}\right) = \frac{256}{625}\cdot\frac{9}{5} = \frac{2304}{3125} = 0.73728.$$

已知行程較長，走的是哪條路線？

設 R 為所選的路線，其中 $\mathrm{P}(R = A) = \mathrm{P}(R = B) = \mathrm{P}(R = D) = \frac{1}{3}$，並設 L 為行程時間至少 27 分鐘這一事件。根據貝葉斯定理，各個 $\frac{1}{3}$ 因子互相約去：

$$\begin{aligned}\mathrm{P}(R = B \mid L) &= \frac{\mathrm{P}(L \mid B)}{\mathrm{P}(L \mid A) + \mathrm{P}(L \mid B) + \mathrm{P}(L \mid D)} = \frac{1 - 13\mathrm{e}^{-4}}{\frac{4}{5} + 1 - 13\mathrm{e}^{-4} + \frac{2304}{3125}} \\ &= \frac{1 - 13\mathrm{e}^{-4}}{\frac{7929}{3125} - 13\mathrm{e}^{-4}} \approx \frac{0.7619}{2.2992} \approx 0.331.\end{aligned}$$

這個數值只比 $\frac{1}{3}$ 略高，十分合理：三條路線令行程達 27 分鐘或以上的可能性幾乎相同。

第 (i) 部分

利用 $\mathrm{E}(A) = 5$、$\mathrm{E}(B) = 4$ 和 $\mathrm{E}(X) = 5 \times \frac{4}{5} = 4$，以及期望值的線性性質（linearity of expectation），

$$\mathrm{E}(T_A) = 25 + 5 = 30, \qquad \mathrm{E}(T_B) = 20 + 3 \times 4 = 32, \qquad \mathrm{E}(T_D) = 10 + 5 \times 4 = 30.$$

? **想一想** 有兩條路線打成平手。這是計錯了嗎？

不是。Angle Avenue 和 Detour Drive 的平均時間都剛好是 30 分鐘，而 Bend Boulevard 的平均時間是 32 分鐘。所以要令平均時間盡量短，我應該選用 Angle Avenue 或 Detour Drive 其中任何一條：兩者同樣好，而且都勝過 Bend Boulevard。

第 (ii) 部分

現在比較各條路線的 $\mathrm{P}(T < 32)$。

- $T_A < 32 \iff A < 7$，所以 $\mathrm{P}(T_A < 32) = \frac{7}{10} = 0.7$。
- $T_B < 32 \iff 3B < 12 \iff B \leq 3$，所以

$$\mathrm{P}(T_B < 32) = \mathrm{e}^{-4}\left(1 + 4 + 8 + \frac{32}{3}\right) = \frac{71}{3}\mathrm{e}^{-4} \approx 0.4335.$$

- $T_D < 32 \iff 5X < 22 \iff X \le 4$，所以

$$\mathrm{P}(T_D < 32) = 1 - \left(\frac{4}{5}\right)^5 = 1 - \frac{1024}{3125} = \frac{2101}{3125} \approx 0.6723.$$

由於 $0.7 > 0.6723 > 0.4335$，我應該每次都選用 **Angle Avenue**，這樣有 70% 的日子可以在 32 分鐘內到達。

驗算

- 模擬 400 000 次行程，得出 $\mathrm{P}(T \ge 27) \approx 0.801,\ 0.761,\ 0.740$，平均時間 $\approx 30.0,\ 32.0,\ 30.0$ 分鐘，$\mathrm{P}(T < 32) \approx 0.700,\ 0.435,\ 0.670$，以及 $\mathrm{P}(R = B \mid L) \approx 0.331$。全部與精確值吻合。
- 檢查邊界情況是否合理：在 Detour Drive 上，繞道 3 次需時 25 分鐘，繞道 4 次則需時 30 分鐘，所以「至少 27 分鐘」確實等同於「至少繞道 4 次」。

常見錯誤

- 處理整數條件時不夠嚴謹：$3B \ge 7$ 表示 $B \ge 3$，而不是 $B \ge 2$；$3B < 12$ 表示 $B \le 3$，而不是 $B \le 4$。
- 忘記 Detour Drive 本身有 2 km 的基本路程，或者忘記其他路線上每公里五分鐘的騎車時間。
- 在 (i) 中聲稱只有一條路線勝出。其實兩個平均值完全相等。

STEP 之外

第 (i) 和第 (ii) 部分說明，「最好」取決於我們要優化的是甚麼：平均值只是概括一個分佈的其中一個數字，兩條平均值相同的路線，其離散程度和尾部概率可以相差很遠。在決策理論（decision theory）中，這就是如何選擇*損失函數*（loss function）的問題。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教科書）涵蓋了本題用到的各種分佈和貝葉斯定理。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）則收錄了許多這一類的題目。

詳細解答

STEP I 1989，第 16 題

概率與統計 • 零和博弈

> **STEP I 1989，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/89-s1-q16

本題在 *Beyond STEP* 第 12.14 節混合策略中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/mixed-strategies.html

關鍵思路

一切都源自一個小小的表：九對猜測各自對應的 B 的贏款。只要把這個表寫下來，每個期望值（expected value）都不過是表中各項的加權和。至於最後一部分，把總和按 A 的概率 a_j 分組，「無論 A 怎樣做」便化為對三個係數的條件，而要令三者皆為非負，只需解一個小型的線性不等式組。

解答

設 $w(i,j)$ 為在 A 猜 i、B 猜 j 的一局中 B 的贏款（以英鎊計）。按規則逐一讀出：

- 若 $|i-j|=2$，說 1 的一方付給說 3 的一方兩英鎊，故 $w(1,3)=2$ 及 $w(3,1)=-2$；
- 若 $|i-j|=1$，說較大數字的一方付出一英鎊，故 $w(2,1)=w(3,2)=1$ 及 $w(1,2)=w(2,3)=-1$；
- $w(i,i)=0$。

	B 說 1	B 說 2	B 說 3
A 說 1	0	-1	2
A 說 2	1	0	-1
A 說 3	-2	1	0

留意 $w(j,i) = -w(i,j)$：規則對兩名玩家一視同仁，所以把兩人的猜測對調，付款的一方也隨之對調。

第 (i) 部分

若 A 總是說 N，B 應沿着表的第 N 行看，並總是選取最大的一項。

- $N = 1$：B 總是說 3，每局贏 £2。
- $N = 2$：B 總是說 1，每局贏 £1。
- $N = 3$：B 總是說 2，每局贏 £1。

在每種情況下，這個選擇都是唯一的最佳選擇，其他任何回應不是一無所獲，就是輸錢。

第 (ii) 部分

在任何一局中，A 的猜測與 B 的猜測互相獨立（independent）（A 自行隨機選取，而 B 則暗中決定），所以 A 說 i 而 B 說 j 的概率是 $\frac{1}{3}b_j$。因此 B 的期望贏款是

$$\begin{aligned}\mathrm{E}(W) &= \sum_{j=1}^{3} b_j \cdot \tfrac{1}{3}\big(w(1,j) + w(2,j) + w(3,j)\big) \\ &= \tfrac{1}{3}\big(b_1(0+1-2) + b_2(-1+0+1) + b_3(2-1+0)\big) \\ &= \tfrac{1}{3}\,(b_3 - b_1).\end{aligned}$$

由於 $b_3 \le 1$ 及 $b_1 \ge 0$，此值至多為 $\frac{1}{3}$，而等號恰好在 $b_3 = 1$ 時成立（這時必有 $b_1 = b_2 = 0$）。所以 B 應該總是猜 3，這樣每局的期望贏款是 £$\frac{1}{3}$。面對均勻隨機選取的 A，猜 3 會贏 £2、輸 £1 或打和，概率各為 $\frac{1}{3}$。

第 (iii) 部分

設 A 以概率 a_i 猜 i，B 以概率 b_j 猜 j，兩者獨立，則

$$\mathrm{E}(W) = \sum_{i=1}^{3}\sum_{j=1}^{3} a_i b_j\, w(i,j).$$

按每個 a_i 合併各項（每次取表中的一行）：

$$\mathrm{E}(W) = \underbrace{(2b_3 - b_2)}_{X}\, a_1 + \underbrace{(b_1 - b_3)}_{Y}\, a_2 + \underbrace{(b_2 - 2b_1)}_{Z}\, a_3.$$

我們希望 $X \ge 0$、$Y \ge 0$ 及 $Z \ge 0$，即

$$b_2 \le 2b_3, \qquad b_3 \le b_1, \qquad 2b_1 \le b_2.$$

把這些不等式串連起來，得 $b_2 \le 2b_3 \le 2b_1 \le b_2$，所以全部都必須是等式：$b_3 = b_1$ 及 $b_2 = 2b_1$。連同 $b_1 + b_2 + b_3 = 1$，便得

$$b_1 = \tfrac{1}{4}, \qquad b_2 = \tfrac{1}{2}, \qquad b_3 = \tfrac{1}{4},$$

這時 $X = Y = Z = 0$，當然都是非負的。所以這樣的選擇確實存在（而且事實上是唯一的一個）。

想一想 B 能否做得更好，令 X、Y 和 Z 全部嚴格為正？

不能。對任何 b_j，

$$b_1X + b_2Y + b_3Z = b_1(2b_3 - b_2) + b_2(b_1 - b_3) + b_3(b_2 - 2b_1) = 0,$$

這不過是說，B 與另一個自己對賽時不賺不蝕（這個表是反對稱（antisymmetric）的）。以總和為 1 的非負權重對 X, Y, Z 取加權和，若三者皆為正數，結果不可能是零。這就是為甚麼上面的不等式只能以等號成立。

現在來作出推論。取 $b = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$，則對任何 a_1, a_2, a_3 的選擇，都有

$$\mathrm{E}(W) = 0 \cdot a_1 + 0 \cdot a_2 + 0 \cdot a_3 = 0$$

；更一般地，只要 $X, Y, Z \ge 0$，便有 $\mathrm{E}(W) = Xa_1 + Ya_2 + Za_3 \ge 0$，因為每個 $a_i \ge 0$。B 在不同局的贏款是獨立同分佈（independent and identically distributed）的（兩名玩家的猜測在各局之間互相獨立），每局的平均值至少為 0。根據大數定律（law of large numbers），經過許多局之後，B 每局的平均贏款會趨向這個平均值，所以長遠而言，無論 A 怎樣做，B 都不會輸錢。 □

驗算

- 第 (ii) 部分是第 (iii) 部分在 $a_1 = a_2 = a_3 = \frac{1}{3}$ 時的特殊情況：$\frac{1}{3}(X + Y + Z) = \frac{1}{3}(b_3 - b_1)$，與之前所得相符。
- 第 (i) 部分是 a 把全部權重放在 N 上的情況：當 $N = 1$ 時，$\mathrm{E}(W) = X = 2b_3 - b_2$，在 $b_3 = 1$ 時取得最大值，即 2。
- 模擬 2×10^5 局，B 採用 $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$，A 採用 $(0.6, 0.1, 0.3)$，所得的平均贏款約為每局 -0.004，與 0 吻合。

常見錯誤

- 把正負號的規定弄反：相差 1 時，付款的是猜較大數字的一方，但相差 2 時，付款的卻是猜較小數字的一方。應先把表列出，再檢查它是否反對稱。

- 從第 (ii) 部分推斷 B 應該混合不同的猜測。面對一個固定的混合策略（mixed strategy），最佳回應總是一個純策略（pure strategy）（這裏是總是猜 3）；只有當 B 必須防範 A 的每一個策略時，混合才有好處，正如第 (iii) 部分那樣。

STEP 之外

這是一個二人零和博弈（two-player zero-sum game），而第 (iii) 部分是馮·諾伊曼（von Neumann）極小極大定理（minimax theorem）的一個小型例子：B 有一個混合策略，無論 A 怎樣做，都能保證他至少取得博弈的值；又因為這個博弈是對稱的，所以該值為 0。尋找這樣的策略是一個線性規劃（linear programming）問題，它把概率與最優化及經濟學連繫起來。概率方面，可參閱 Grimmett and Welsh, *Probability: An Introduction*（概率入門教材）；至於博弈本身，Binmore 的 *Playing for Real*（博弈論入門讀物）是很好的入門書。

詳細解答

STEP II 1989，第 1 題

純數學 • 三次方程的三角解法

> **STEP II 1989，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/89-s2-q1

關鍵思路

這條恆等式（identity）說明，當 $z = \cos\theta$ 時，$4z^3 - 3z$ 其實就是喬裝了的 $\cos 3\theta$。因此，只要能把三次方程（cubic）化成 $4z^3 - 3z = k$ 的形式，而且 $|k| \leq 1$，解這條方程就等同於解 $\cos 3\theta = k$，而後者有三個容易求得的解。兩個代換（substitution）各司其職：平移 $y = x + a$ 用來消去平方項，縮放 $z = by$ 則把餘下兩個係數（coefficient）之比調整為 $4 : 3$。

解答

恆等式

利用複角公式（addition formulae）及 $\sin^2\theta = 1 - \cos^2\theta$，

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

正是所求。

化簡三次方程

設 $y = x + a$，則 $x = y - a$，而 $24(y-a)^3 - 72(y-a)^2 + \cdots$ 中 y^2 的係數是 $-72a - 72$。要消去這一項，須取 $a = -1$，即 $x = y + 1$。於是

$$\begin{aligned} &24(y+1)^3 - 72(y+1)^2 + 66(y+1) - 19 \\ &\quad = (24y^3 + 72y^2 + 72y + 24) - (72y^2 + 144y + 72) \\ &\qquad + (66y + 66) - 19 \\ &\qquad\qquad = 24y^3 - 6y - 1, \end{aligned}$$

所以 $(*)$ 變成

$$24y^3 - 6y = 1.$$

現在代入 $y = z/b$：

$$\frac{24}{b^3}z^3 - \frac{6}{b}z = 1.$$

想一想 應該怎樣選取 b，才能令左方成為 $4z^3 - 3z$ 的倍數？

我們需要 z^3 與 z 的係數之比為 $4 : -3$，所以

$$\frac{24/b^3}{6/b} = \frac{4}{3} \quad\Longrightarrow\quad \frac{4}{b^2} = \frac{4}{3} \quad\Longrightarrow\quad b^2 = 3.$$

取 $b = \sqrt{3}$。這時方程是 $\dfrac{8}{\sqrt{3}}z^3 - \dfrac{6}{\sqrt{3}}z = 1$，兩邊乘以 $\frac{\sqrt{3}}{2}$，得

$$4z^3 - 3z = \frac{\sqrt{3}}{2}.$$

所以 $a = -1$、$b = \sqrt{3}$ 及 $k = \frac{\sqrt{3}}{2}$，其中 $z = \sqrt{3}\,(x-1)$。（若改取 $b = -\sqrt{3}$，則 $k = -\frac{\sqrt{3}}{2}$，最後得到的根（root）完全相同。）

求根

由於 $|k| \le 1$，我們可以寫 $z = \cos\theta$，而根據恆等式，方程變成

$$\cos 3\theta = \frac{\sqrt{3}}{2} = \cos 30^\circ.$$

所以 $3\theta = 30^\circ,\ 390^\circ,\ 750^\circ$，從而 $\theta = 10^\circ,\ 130^\circ,\ 250^\circ$。這三個角給出三個不同的 $\cos\theta$ 值，而三次方程最多只有三個根，所以這就是全部的根：

$$z = \cos 10^\circ,\quad \cos 130^\circ,\quad \cos 250^\circ.$$

把代換還原，$x = 1 + \dfrac{z}{\sqrt{3}}$，所以

$$x_1 = 1 + \frac{\cos 10^\circ}{\sqrt{3}} = 1.5686\ldots \approx 1.57,$$
$$x_2 = 1 + \frac{\cos 130^\circ}{\sqrt{3}} = 0.6289\ldots \approx 0.629,$$
$$x_3 = 1 + \frac{\cos 250^\circ}{\sqrt{3}} = 0.8025\ldots \approx 0.803.$$

並非每條三次方程都適用

對於一般的首一（monic）三次方程 $x^3 + \alpha x^2 + \beta x + \gamma = 0$，平移的選擇別無他法：代入 $x = y - a$ 後，y^2 的係數是 $\alpha - 3a$，所以 $a = \frac{\alpha}{3}$，由此得到 $y^3 + py + q = 0$，其中 $p = \beta - \frac{\alpha^2}{3}$。代入 $y = z/b$ 並乘以 $4b^3$，得

$$4z^3 + 4pb^2 z + 4qb^3 = 0,$$

要與 $4z^3 - 3z$ 吻合，須有 $4pb^2 = -3$，即 $b^2 = -\dfrac{3}{4p}$。只有當 $p < 0$ 時，這才有實數解。

一個反例（counterexample）是

$$x^3 + 3x - 4 = 0.$$

這裏 $\alpha = 0$，所以 $a = 0$ 而 $p = 3 > 0$；我們需要 $b^2 = -\frac{1}{4}$，但沒有實數 b 能滿足這個條件。所以這條方程根本無法化成 $4z^3 - 3z = k$ 的形式。（方程本身確實有根：它可分解為 $(x-1)(x^2 + x + 4) = 0$，而當中二次式的判別式（discriminant）為 $-15 < 0$，所以 $x = 1$ 是它唯一的實根。失效的是這個方法，而不是方程。）

即使 $p < 0$，這個方法也可能在下一步失效。對於 $x^3 - 3x - 4 = 0$，代換 $x = 2z$ 給出 $4z^3 - 3z = 2$；但 $|\cos 3\theta| \le 1$，所以沒有實數 θ 能使 $\cos 3\theta = 2$。

驗算

- 三個根之和應為 $\frac{72}{24} = 3$，三個根之積應為 $\frac{19}{24}$。事實上，各 cos 項之和為零（$\cos 10^\circ + \cos 130^\circ + \cos 250^\circ = 0$），而數值上 $1.5686 \times 0.6289 \times 0.8025 = 0.7917 \approx \frac{19}{24}$。
- 把 $x = 1.5686$ 代入 $24x^3 - 72x^2 + 66x - 19$，在計算機的準確度之內結果為零；其餘兩個根亦然。

常見錯誤

- 求得 $\theta = 10^\circ$ 便停下來：另外兩個根來自 $3\theta = 390^\circ$ 和 750°。

- 弄錯 a 的正負號：$y = x + a$ 而 $x = y + 1$，即 $a = -1$，不是 $+1$。
- 舉反例時，選了一條 $p < 0$ 而且 $|k| \leq 1$ 的三次方程，但這個方法其實完全應付得來。在斷言方法失效之前，先檢查 $b^2 > 0$ 和 $|k| \leq 1$ 這兩個條件。

STEP 之外

這個方法恰好在三次方程有三個實根時適用，這正是所謂「不可約情形」（casus irreducibilis）：儘管每個答案都是實數，卡爾達諾（Cardano）公式卻迫使你繞道複數的立方根。當 $|k| > 1$ 時，把 cos 換成 cosh，同樣的想法依然可行，因為 $\cosh 3u = 4\cosh^3 u - 3\cosh u$；對於 $x^3 - 3x - 4 = 0$，這給出實根 $x = 2\cosh\left(\frac{1}{3}\operatorname{arcosh} 2\right) \approx 2.196$。至於哪些多項式方程可以用根式求解（solved by radicals）這個更深入的問題，則通往伽羅瓦理論（Galois theory）；Allenby 的 *Rings, Fields and Groups* 是很好的入門讀物，而 Needham 的 *Visual Complex Analysis* 則展示了立方根的幾何。

詳細解答

STEP II 1989，第 2 題

純數學 • 冪級數的係數

> **STEP II 1989，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/89-s2-q2

關鍵思路

一條把函數在 x 處的值與同一函數在 $2x$ 處的值聯繫起來的三角恆等式（trigonometric identity），會化為冪級數（power series）係數之間的關係。把 x 換成 $2x$，x^n 的係數就乘以 2^n，所以只要把兩邊都寫成冪級數，便可以比較係數（compare coefficients），直接讀出所求的關係。有用的竅門是先把整條式乘以 x，使每一項都是在 0 處確實具有冪級數的函數。

解答

以下一律假設 x 很小而且不等於零，因此下面出現的每個級數都收斂（converge）（所涉及的函數在例如 $0 < |x| < \frac{\pi}{4}$ 的範圍內性質都良好）。兩個收斂的冪級數若在 0 附近的一個區間上相等，它們的係數就相等，所以我們可以比較 x^n 的係數。

a_{n-1} 與 b_n 之間的關係

首先驗證 $(*)$：

$$\cot x - \tan x = \frac{\cos^2 x - \sin^2 x}{\sin x \cos x} = \frac{\cos 2x}{\frac{1}{2}\sin 2x} = 2\cot 2x.$$

乘以 x：

$$x\cot x - x\tan x = (2x)\cot(2x).$$

現在把每一項寫成級數。

▶ $x\cot x = 1 + \sum_{n\geq 1} b_n x^n$。

- $x\tan x = \sum_{m\geq 0} a_m x^{m+1} = \sum_{n\geq 1} a_{n-1}x^n$，其中取 $n = m+1$。
- 由於只要 x 很小，$2x$ 也很小，我們可以在 $x\cot x$ 的級數中把 x 換成 $2x$：

$$(2x)\cot(2x) = 1 + \sum_{n\geq 1} b_n (2x)^n = 1 + \sum_{n\geq 1} 2^n b_n x^n.$$

因此

$$1 + \sum_{n\geq 1} (b_n - a_{n-1})x^n = 1 + \sum_{n\geq 1} 2^n b_n x^n.$$

常數項相等，而對 $n \geq 1$ 比較 x^n 的係數，得 $b_n - a_{n-1} = 2^n b_n$，即

$$a_{n-1} = (1-2^n)\, b_n \qquad (n \geq 1),$$

這正是所要證明的結果。

想一想 為甚麼要先乘以 x 才展開，而不直接處理 $\cot x$ 本身？

$\cot x$ 在 0 處趨向無窮大，所以在該處沒有冪級數；題目只給了我們 $x\cot x$ 的級數。把整條恆等式乘以 x 之後，每一項都成為「具有冪級數的函數」的形式，而右邊的餘切項變成 $(2x)\cot(2x)$，也就是同一個函數在 $2x$ 處的值。這正是我們需要的形式。

係數 c_n

與之配對的恆等式是把差換成和：

$$\cot x + \tan x = \frac{\cos^2 x + \sin^2 x}{\sin x \cos x} = \frac{1}{\frac{1}{2}\sin 2x} = 2\operatorname{cosec} 2x.$$

再次乘以 x，

$$x\cot x + x\tan x = (2x)\operatorname{cosec}(2x).$$

利用與先前相同的展開式，以及把 x 換成 $2x$ 後的 $x\operatorname{cosec} x$ 的級數，

$$1 + \sum_{n\geq 1} (b_n + a_{n-1})x^n = 1 + \sum_{n\geq 1} 2^n c_n x^n,$$

所以對 $n \geq 1$ 有

$$2^n c_n = b_n + a_{n-1}.$$

現在用第一個結果消去 b_n。由於 $n \geq 1$，$1 - 2^n \neq 0$，所以 $b_n = \dfrac{a_{n-1}}{1-2^n}$，於是

$$2^n c_n = a_{n-1}\left(\frac{1}{1-2^n} + 1\right) = a_{n-1}\frac{2-2^n}{1-2^n} = a_{n-1}\frac{2\,(2^{n-1}-1)}{2^n - 1}.$$

除以 2^n，並利用 $2/2^n = 1/2^{n-1}$，

$$c_n = \frac{2^{n-1}-1}{2^n-1}\cdot\frac{1}{2^{n-1}}\,a_{n-1} \qquad (n \geq 1),$$

這就是所求的結果。

驗算

由 $\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots$ 可得 $a_1 = 1$，$a_3 = \frac{1}{3}$。第一個結果預測 $b_2 = a_1/(1-4) = -\frac{1}{3}$ 和 $b_4 = a_3/(1-16) = -\frac{1}{45}$，與 $x\cot x = 1 - \frac{1}{3}x^2 - \frac{1}{45}x^4 - \cdots$ 吻合。第二個結果預測 $c_2 = \frac{1}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$ 和 $c_4 = \frac{7}{15}\cdot\frac{1}{8}\cdot\frac{1}{3} = \frac{7}{360}$，與 $x\operatorname{cosec} x = 1 + \frac{1}{6}x^2 + \frac{7}{360}x^4 + \cdots$ 吻合。另外留意 $n = 1$ 的情況：$a_0 = 0$ 而 $b_1 = c_1 = 0$，與兩條公式都一致。我們把 sin 和 cos 的級數相除，以分數作精確計算，對所有 $n \le 23$ 驗證了這兩個關係，並用 $\tan 0.3 \approx 0.309336$ 和 $0.3\operatorname{cosec} 0.3 \approx 1.015159$ 核對了這些級數。

常見錯誤

- 忘記下標的平移：$x\tan x$ 中 x^n 的係數是 a_{n-1}，而不是 a_n。
- 在 $x\cot x$ 中以 $2x$ 代替 x 時，卻寫成 $2x\cot x$ 或 $x\cot 2x$；恆等式中的那一項必須恰好是 $(2x)\cot(2x)$。
- 除以 $1 - 2^n \neq 0$ 之前，沒有指出這一點。

STEP 之外

這裏的係數，除了相差一些簡單的因子之外，就是伯努利數（Bernoulli numbers）：$x\cot x = \sum_{k\ge 0}(-1)^k \frac{2^{2k}B_{2k}}{(2k)!}x^{2k}$，而題目中的關係則以伯努利數表示出正切和餘割的級數。兩個冪級數若在 0 附近相等，便可以令它們的係數相等，這個事實（冪級數的恆等定理（identity theorem））會在大學第一個分析課程中證明。想進一步了解，可參閱 Abbott 的 *Understanding Analysis*（分析學入門教材）中關於冪級數的部分，以及 Needham 的 *Visual Complex Analysis*（以幾何直觀講解複分析的著作），這些展開式在書中會從複函數自然地產生。

詳細解答

STEP II 1989，第 3 題

純數學 • 複雙曲函數映射

> **STEP II 1989，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/89-s2-q3

關鍵思路

右邊是 $2\sinh(x+\mathrm{i}y)$，所以整道題目其實是在研究映射（map）$u+\mathrm{i}v=\sinh(x+\mathrm{i}y)$。把和的 sinh 展開，便可分成實部（real part）和虛部（imaginary part），從而明確寫出 u 和 v。固定 x 或 y 之後，剩下的是一條參數曲線（parametric curve），再利用恆等式 $\cos^2+\sin^2=1$ 和 $\cosh^2-\sinh^2=1$ 便可消去參數。

解答

求 u 和 v

兩邊除以 2，得 $u+\mathrm{i}v=\sinh(x+\mathrm{i}y)$。利用 $\sinh(\mathrm{i}y)=\mathrm{i}\sin y$ 和 $\cosh(\mathrm{i}y)=\cos y$，

$$\begin{aligned}\sinh(x+\mathrm{i}y)&=\sinh x\cosh(\mathrm{i}y)+\cosh x\sinh(\mathrm{i}y)\\&=\sinh x\cos y+\mathrm{i}\cosh x\sin y.\end{aligned}$$

（如果你喜歡，也可以直接展開 $\frac{1}{2}\left(\mathrm{e}^{x}(\cos y+\mathrm{i}\sin y)-\mathrm{e}^{-x}(\cos y-\mathrm{i}\sin y)\right)$；結果完全一樣。）由於 x、y、u、v 全是實數，比較實部和虛部可得

$$u=\sinh x\cos y,\qquad v=\cosh x\sin y.$$

直線 $x=a$

固定 $x=a$ 而讓 y 任意變化，則 $u=\sinh a\cos y$，$v=\cosh a\sin y$。由於 $a>0$，$\sinh a$ 和 $\cosh a$ 都是正數，所以可以相除：

$$\left(\frac{u}{\sinh a}\right)^2+\left(\frac{v}{\cosh a}\right)^2=\cos^2 y+\sin^2 y=1.$$

反過來，這個橢圓（ellipse）上的每一點都可以寫成 $(\sinh a\cos y, \cosh a\sin y)$ 的形式（對某個 y），因為這正是橢圓的標準參數表示（parametrisation）。所以直線 $x = a$ 對應整個橢圓（y 每走過一個長度為 2π 的區間，橢圓便被描畫一次），其半軸（semi-axes）沿 u 軸方向為 $\sinh a$，沿 v 軸方向為 $\cosh a$。

直線 $y = b$

現在固定 $y = b$ 而讓 x 任意變化：$u = \sinh x\cos b$，$v = \cosh x\sin b$。題目給出 $0 < \sin b < 1$，所以 $\sin b \neq 0$，而且 $\cos^2 b = 1 - \sin^2 b > 0$，所以 $\cos b \neq 0$ 亦成立。相除後得

$$\left(\frac{v}{\sin b}\right)^2 - \left(\frac{u}{\cos b}\right)^2 = \cosh^2 x - \sinh^2 x = 1.$$

這是一條雙曲線（hyperbola），其軸沿 v 軸，頂點（vertices）位於 $(0, \pm\sin b)$。

> **想一想** 為甚麼只有一支？這一支又是否全部都被覆蓋？

由於 $\cosh x > 0$ 且 $\sin b > 0$，必有 $v > 0$：只能到達上面的一支。而且這一支上的每一點都能到達：任取上支上的一點 (u, v)，選 $x = \sinh^{-1}(u/\cos b)$，使得 $\sinh x\cos b = u$；這時 $\cosh x\sin b$ 就是雙曲線方程對於這個 u 的正根，亦即 v。所以直線 $y = b$ 恰好對應其中一支，即 $v > 0$ 的那一支。

交點與直角

點 $x = a$、$y = b$ 同時在兩條直線上，所以它的像（image）同時在兩條曲線上：

$$(u, v) = (\sinh a\cos b,\ \cosh a\sin b).$$

要找出兩條曲線在該點的方向，可把每個參數表示對各自的參數求導（differentiate）。沿橢圓（$x = a$，參數為 y），在 $y = b$ 處：

$$\mathbf{t}_1 = \left(\frac{\mathrm{d}u}{\mathrm{d}y}, \frac{\mathrm{d}v}{\mathrm{d}y}\right) = (-\sinh a\sin b,\ \cosh a\cos b).$$

沿雙曲線（$y = b$，參數為 x），在 $x = a$ 處：

$$\mathbf{t}_2 = \left(\frac{\mathrm{d}u}{\mathrm{d}x}, \frac{\mathrm{d}v}{\mathrm{d}x}\right) = (\cosh a\cos b,\ \sinh a\sin b).$$

兩個向量都不是零向量，因為它們各有一個分量是 $\cosh a\cos b \neq 0$，所以兩者都是真正的切向量（tangent vectors）。它們的純量積（scalar product）是

$$\mathbf{t}_1 \cdot \mathbf{t}_2 = -\sinh a\sin b\cosh a\cos b + \cosh a\cos b\sinh a\sin b = 0,$$

所以橢圓與雙曲線的這一支在該點正交，正如題目所求。

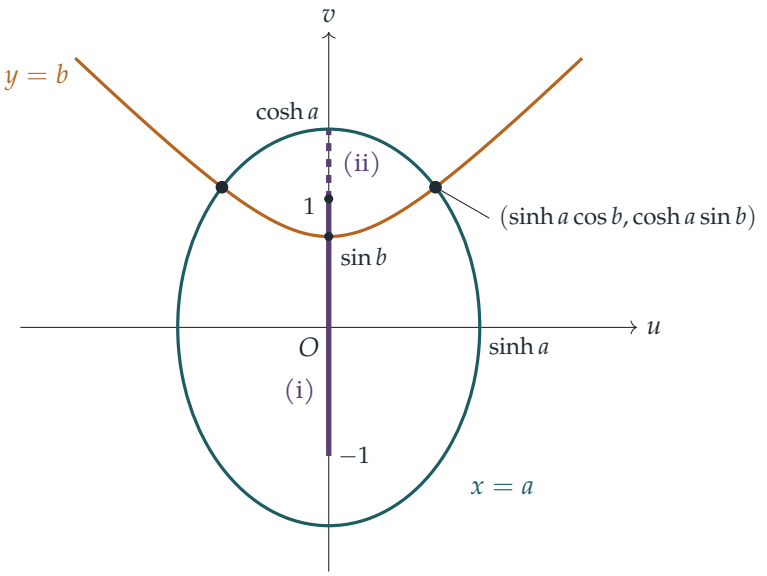
（由於兩條曲線都具有 $u \mapsto -u$ 的對稱性，它們在 $(-\sinh a\cos b, \cosh a\sin b)$ 亦正交；這一點是 $x = -a$、$y = b$ 的像，也同樣是 $x = a$、$y = \pi - b$ 的像。）

略圖

(i) $x = 0$ 給出 $u = 0$，$v = \sin y$：即 v 軸上由 $(0, -1)$ 至 $(0, 1)$ 的線段。留意它是橢圓在 $a \to 0$ 時的極限，因為 $\sinh a \to 0$ 而 $\cosh a \to 1$。

(ii) $y = \frac{1}{2}\pi$，$0 \le x \le a$ 給出 $u = 0$，$v = \cosh x$，其值由 1 遞增至 $\cosh a$：即 v 軸上由 $(0, 1)$ 向上至橢圓頂端 $(0, \cosh a)$ 的線段。

雙曲線這一支的頂點在 $(0, \sin b)$，位於線段 (i) 上。略圖中取 $a = 1$，$b = \frac{1}{4}\pi$。



線段 (i) 以實線表示，線段 (ii) 以虛線表示；兩者合起來沿 v 軸由 -1 一直向上伸延至橢圓頂端。假如改為取 $y = \frac{1}{2}\pi$ 而讓 x 取遍所有實數，便會得到 v 軸上的射線 $v \ge 1$，而且重複覆蓋兩次；這就是題目排除了的退化情況 $\sin b = 1$。

驗算

- 兩條圓錐曲線(conics)的焦點(foci)都在 $(0, \pm 1)$：橢圓方面是因為 $\cosh^2 a - \sinh^2 a = 1$，雙曲線方面是因為 $\sin^2 b + \cos^2 b = 1$。共焦（confocal）的橢圓和雙曲線必定正交，與我們的結果吻合。
- 數值上，取 $a = 1$，$b = \frac{1}{4}\pi$：$\sinh(1 + \frac{1}{4}\pi\mathrm{i}) \approx 0.8310 + 1.0911\mathrm{i}$，與 $(\sinh 1 \cos \frac{1}{4}\pi,\ \cosh 1 \sin \frac{1}{4}\pi)$ 相符。我們亦在許多隨機選取的 (a, b) 上驗證了兩條曲線的方程，以及純量積為零。

常見錯誤

- 把 $\sinh(\mathrm{i}y) = \sin y$ 寫成等式，而不是 $\mathrm{i}\sin y$，結果把 u 和 v 的角色對調了。
- 只說「一條雙曲線」，卻沒有解釋為甚麼只出現 $v > 0$ 的那一支。
- 除以 $\cos b \neq 0$ 之前忘記檢查這個條件；條件 $\sin b < 1$ 正是為此而設。

STEP 之外

映射 $z \mapsto \sinh z$ 是保角（conformal）的：凡是導數不為零的地方，它都保持角度不變，所以由直線 $x =$ 常數 和 $y =$ 常數 組成的垂直網格，必定變成由曲線組成的垂直網格，在這裏就是共焦的橢圓和雙曲線。這就是為甚麼在平板或狹縫周圍求解拉普拉斯方程（Laplace's equation）時，會出現這類「橢圓坐標」（elliptic coordinates）。Needham 的 *Visual Complex Analysis*（《可視化複分析》）把保角映射講解得十分精彩，書中有許多正是這一類的圖。

詳細解答

STEP II 1989，第 4 題

純數學 • 有理函數的描繪

> **STEP II 1989，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/89-s2-q4

關鍵思路

由線性因式相乘或相除而成的式子，最容易用取對數求導（logarithmic differentiation）的方法處理，因為對數會把它變成一個和。至於值域（range），不妨把問題倒過來問：對哪些 z，$\mathrm{f}(x) = z$ 有實數解？這條方程是 x 的二次方程（quadratic），所以答案來自判別式（discriminant）。之後的草圖，就由零點（zeros）、極點（poles）、$|x|$ 很大時的表現，以及是否有某些 z 值取不到來決定。

關於印出來的題目，要說明一點。條件 $a+d \neq c+d$ 只是說 $a \neq c$，這是我們早已知道的。實際用到的條件是 $a+b \neq c+d$，以下我們一直假設這個條件成立。

解答

導數

在 $\mathrm{f}(x) \neq 0$ 的地方，即 $x \neq a, b$（以及 $x \neq c, d$）時，我們可以對絕對值取對數（logarithms of moduli）：

$$\ln|\mathrm{f}(x)| = \ln|x-a| + \ln|x-b| - \ln|x-c| - \ln|x-d|.$$

求導，並利用 $\frac{\mathrm{d}}{\mathrm{d}x}\ln|x-k| = \frac{1}{x-k}$，得

$$\frac{\mathrm{f}'(x)}{\mathrm{f}(x)} = \frac{1}{x-a} + \frac{1}{x-b} - \frac{1}{x-c} - \frac{1}{x-d}.$$

現在乘以 x。當 $x \neq 0$ 時，每一項變成

$$\frac{x}{x-k} = \frac{1}{1-k/x} = \left(1 - \frac{k}{x}\right)^{-1},$$

於是得到所求的恆等式

$$\frac{x\,\mathrm{f}'(x)}{\mathrm{f}(x)} = \left(1-\frac{a}{x}\right)^{-1} + \left(1-\frac{b}{x}\right)^{-1} - \left(1-\frac{c}{x}\right)^{-1} - \left(1-\frac{d}{x}\right)^{-1}.$$

$|x|$ 很大時

當 $|x|$ 遠大於 $|k|$ 時，由二項式展開（binomial expansion）得 $(1-k/x)^{-1} = 1 + \frac{k}{x} + O\left(\frac{1}{x^2}\right)$。四個首項 1 互相抵消（$1+1-1-1=0$），餘下

$$\frac{x\,\mathrm{f}'(x)}{\mathrm{f}(x)} = \frac{a+b-c-d}{x} + O\left(\frac{1}{x^2}\right), \qquad \text{所以} \qquad \mathrm{f}'(x) = \frac{\mathrm{f}(x)}{x^2}\left[(a+b-c-d) + O\left(\tfrac{1}{x}\right)\right].$$

由於 $a+b-c-d \neq 0$，只要 $|x|$ 足夠大，方括號內的式子便與 $a+b-c-d$ 同號。另外，當 $|x| \to \infty$ 時 $\mathrm{f}(x) \to 1$，所以在那裏 $\mathrm{f}(x) > 0$，而且 $x^2 > 0$。因此，當 $|x|$ 很大時，在實數線的兩端，$\mathrm{f}'(x)$ 都與 $a+b-c-d$ 同號。

曲線位於漸近線（asymptote）$y=1$ 的哪一側，也值得記下來。直接計算，

$$\mathrm{f}(x) - 1 = \frac{(c+d-a-b)x + ab - cd}{(x-c)(x-d)} \approx \frac{c+d-a-b}{x} \quad \text{對於很大的} |x|,$$

而由分子可見，$\mathrm{f}(x) = 1$ 恰好只在一點成立，即 $x_0 = \dfrac{cd-ab}{c+d-a-b}$。

值域的條件：一個方法

把 $\mathrm{f}(x) = z$ 改寫成 $(x-a)(x-b) = z(x-c)(x-d)$，這就是二次方程

$$(1-z)x^2 - \big(a+b-(c+d)z\big)x + (ab - cdz) = 0.$$

- 這條方程的根不可能是 $x=c$ 或 $x=d$，因為左邊的 $(x-a)(x-b)$ 在那裏不等於零。所以二次方程的實根恰好就是 $\mathrm{f}(x) = z$ 的實數解。
- 當 $z \neq 1$ 時，方程有實根，當且僅當判別式 $\Delta(z) = \big(a+b-(c+d)z\big)^2 - 4(1-z)(ab-cdz)$ 滿足 ≥ 0。把它展開，乘以 $(c-d)^2 > 0$，再按 z 的冪次整理。設 $P = (a-c)(b-d)$ 和 $Q = (a-d)(b-c)$，有用的恆等式是 $P - Q = (a-b)(c-d)$，所以常數項 $(a-b)^2(c-d)^2$ 等於 $(P+Q)^2 - 4PQ$。於是題目給出的不等式正好就是 $(c-d)^2\Delta(z) \geq 0$。
- 當 $z=1$ 時，方程是線性的，x 的係數為 $c+d-a-b \neq 0$，所以有一個根；而不等式亦成立，因為兩邊相差 $(c-d)^2(a+b-c-d)^2 > 0$。

草圖

記 $G(z) = \left[(c-d)^2 z + P + Q\right]^2 - 4PQ$。當 z 變化時，平方項取遍 $[0, \infty)$ 內的每一個值，所以 G 的最小值是 $-4PQ$。因此：

$$\text{函數f 的值域是整個}\mathbb{R} \iff PQ \leq 0.$$

若 $PQ > 0$，則使 $G(z) < 0$ 的 z 值構成一個開區間，這些就是取不到的值。

> **想一想** 沒有實際的數值，我們怎樣知道 PQ 的正負？

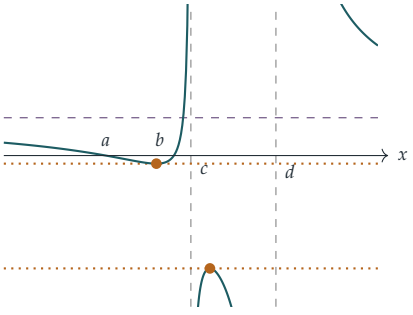
重新組合：$PQ = (a-c)(a-d) \cdot (b-c)(b-d)$。由於 $a < c < d$，第一對是兩個負數的積，所以是正數。於是 PQ 與 $(b-c)(b-d)$ 同號，而後者為負，恰好就是 b 位於兩個極點 c 與 d 之間的時候。一切都取決於 b 相對於 $c < d$ 的位置，以及 $a + b - c - d$ 的正負：

1. $a < b < c < d$：這時 $a + b < c + d$ 自動成立。
2. $a < c < b < d$：同樣有 $a + b < c + d$，因為 $a < c$ 而且 $b < d$。
3. $a < c < d < b$，並且 $a + b < c + d$。
4. $a < c < d < b$，並且 $a + b > c + d$。

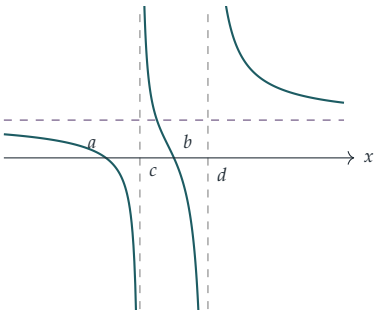
所以只有在情況 2，值域才是整個 $\mathbb{R}$。在情況 1、3 和 4，都有一個區間的值取不到。

還有一件工具，對找出轉向點（turning points）很有幫助。若 f 在 x_1 有駐點（stationary point），而 $\mathrm{f}(x_1) = z_1$，則 x_1 是 $(x-a)(x-b) - z_1(x-c)(x-d) = 0$ 的重根（repeated root）。（當 $z_1 = 1$ 時這是線性方程，所以不可能發生。）有重根即是 $G(z_1) = 0$。所以駐點處的函數值恰好就是取不到的區間的兩個端點。在情況 2，$G > 0$ 恆成立，所以完全沒有駐點。

f 在每個極點兩側附近的正負，可以靠數一數負因式的個數得出。在每幅圖中，垂直虛線是極點 $x = c$、$x = d$，水平虛線是漸近線 $y = 1$，琥珀色的點線標出取不到的區間的兩個端點，也就是駐點（琥珀色圓點）處的函數值。各曲線是按特定數值（列於每幅圖下方）繪畫的，但在同一情況下，無論怎樣選取數值，形狀都是一樣的。



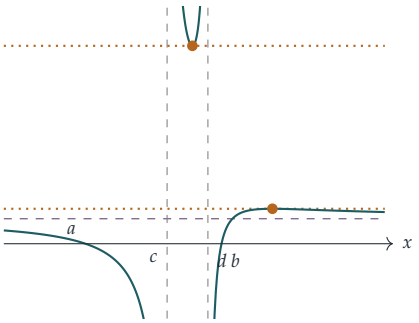
情況 1：$a, b, c, d = -2, 0, \frac{1}{2}, 3$



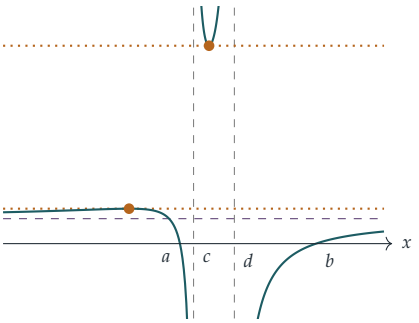
情況 2：$a, b, c, d = -1, 1, 0, 2$

情況 1（$a < b < c < d$）。 這裏在左邊遠處 $\mathrm{f} < 1$，在右邊遠處 $\mathrm{f} > 1$。在 $x < c$ 上，曲線由略低於 1 的位置進來，在 a 處降至 0，在 a 與 b 之間下探至一個負的極小值，然後上升穿過 b，在 c 處趨向 $+\infty$。在兩個極點之間，分母的兩個因式異號，而分子為正，所以在兩端都有 $\mathrm{f} \to -\infty$，中間有一個負的局部極大值（local maximum）。在 d 的右邊，曲線由 $+\infty$ 下降至 1。(c, d) 上的局部極大值低於 (a, b) 上的局部極小值（local minimum），嚴格介乎兩者之間的值都取不到。值域不是整個 $\mathbb{R}$。

情況 2（$a < c < b < d$）。 這裏沒有駐點。曲線由略低於 1 的位置下降，穿過 a，在 c 處趨向 $-\infty$；然後由 $+\infty$ 穿過 b，在 d 處趨向 $-\infty$；再由 $+\infty$ 下降至 1。單是中間的一支已經取遍所有實數值。值域是整個 $\mathbb{R}$。



情況 3：$a, b, c, d = -3, 2, 0, \frac{3}{2}$



情況 4：$a, b, c, d = -\frac{1}{2}, \frac{9}{2}, 0, \frac{3}{2}$

情況 3（$a < c < d < b$，$a + b < c + d$）。 在 $x < c$ 上，曲線由略低於 1 的位置下降，穿過 a，趨向 $-\infty$。在兩個極點之間，分子為負，所以在兩端都有 $\mathrm{f} \to +\infty$，中間有一個局部極小值 m。在 d 的右邊，曲線由 $-\infty$ 上升穿過 b，在 x_0 處穿過 $y = 1$，到達略高於 1 的局部極大值 M，然後從上方回落至 1。(M, m) 內的值都取不到，其中 $1 < M < m$。值域不是整個 $\mathbb{R}$。

情況 4（$a < c < d < b$，$a + b > c + d$）。 這是把情況 3 的外側表現對調。在左邊遠處，f 略高於 1 並且遞增。它到達局部極大值 $M > 1$，在 $x_0 < a$ 處穿過 $y = 1$，再下降穿過 a，在 c 處趨向 $-\infty$。中間的一支與之前一樣呈 U 形，極小值為 m。在 d 的右邊，曲線由 $-\infty$ 上升穿過 b，從下方趨向 1。(M, m) 同樣取不到。值域不是整個 $\mathbb{R}$。

驗算

- 恆等式：隨機選取 a, b, c, d, x，用有限差分（finite difference）估算 $x\mathrm{f}'(x)/\mathrm{f}(x)$，結果與右邊相符，誤差約為 10^{-9}。

- 情況 1，取 $a,b,c,d=-2,0,\frac{1}{2},3$：這裏 $P=\frac{15}{2}$，$Q=\frac{5}{2}$，$(c-d)^2=\frac{25}{4}$，而 $G(z)=0$ 給出 $z=-\frac{8}{5}\pm\frac{4}{5}\sqrt{3}$，約為 -0.2144 和 -2.9856。這正好就是解 $\mathrm{f}'(x)=0$ 所得的駐點處的函數值（在 $x\approx-0.515$ 和 $x\approx1.060$）。在 4×10^5 個點上對 f 取樣，沒有一個值嚴格落在兩者之間。
- 情況 3 和 4（圖中所用的數值）：取不到的區間是 $(1.4007, 7.9327)$，同樣與駐點處的函數值吻合。在情況 2，$\mathrm{f}'(x)=0$ 沒有實數解。

常見錯誤

- 忘記了 $\mathrm{f}(x)=z$ 只在 $z\neq1$ 時才是二次方程：$z=1$ 這個值需要另外交代一句。
- 以為曲線在兩端都從同一側趨近 $y=1$。事實並非如此：$\mathrm{f}-1\approx(c+d-a-b)/x$ 的正負隨 x 而改變。

STEP 之外

把「f 的值域是甚麼？」變成「對哪些 z，$\mathrm{f}(x)=z$ 有解？」，就是原像（preimage）的想法；而觀察到駐點處的函數值正是兩個解合而為一的地方，則是臨界值（critical values）理論的起點。像這樣的 2 次有理函數（rational function）是一個二對一的映射，它在複數上的表現要簡潔得多：在複數上，每個值（按重數（multiplicity）計算）都恰好取到兩次。Needham 的 *Visual Complex Analysis* 把這種幾何觀點發揮得十分精彩。

詳細解答

STEP II 1989，第 5 題

純數學 • 極坐標曲線與反射

> **STEP II 1989，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/89-s2-q5

關鍵思路

第 (i) 部分只不過是把 $x = r\cos\theta$, $y = r\sin\theta$ 代入 $\mathrm{d}y/\mathrm{d}x = (\mathrm{d}y/\mathrm{d}\theta)/(\mathrm{d}x/\mathrm{d}\theta)$。至於鏡面，以 O 為中心的極坐標（polar coordinates）是最自然的選擇，因為反射光線總是沿着半徑回到極點（pole）。這樣，反射定律（law of reflection）就能單憑 θ 告訴我們每一點上法線（normal）的方向，而第 (i) 部分則把這一點化為關於 $r(\theta)$ 的一階微分方程（differential equation），而且這個方程可以分離變量。

解答

第 (i) 部分

極坐標為 (r,θ) 的點，其笛卡兒坐標（Cartesian coordinates）為 $x = r\cos\theta, y = r\sin\theta$。沿着曲線 $r = r(\theta)$，兩者都是 θ 的函數，由積法則（product rule）得

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta, \qquad \frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}\theta}\sin\theta + r\cos\theta.$$

由鏈式法則（chain rule，用於以參數表示的曲線），凡 $\mathrm{d}x/\mathrm{d}\theta \neq 0$ 之處，

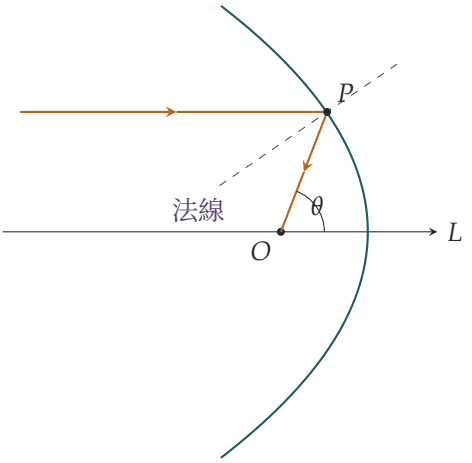
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{\dfrac{\mathrm{d}r}{\mathrm{d}\theta}\sin\theta + r\cos\theta}{\dfrac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta} = \frac{\dfrac{\mathrm{d}r}{\mathrm{d}\theta}\tan\theta + r}{\dfrac{\mathrm{d}r}{\mathrm{d}\theta} - r\tan\theta},$$

其中最後一步把分子和分母同時除以 $\cos\theta$，只要 $\cos\theta \neq 0$ 便可以這樣做。這就是所求的公式。

第 (ii) 部分

化為平面曲線問題。任取一個包含 L 的平面 Π，並在 Π 內取鏡面上的一點 P。經過 P 且平行於 L 的光線位於 Π 內，O 亦然；因此反射光線 PO 也位於 Π 內。P 處的法線平分 Π 內兩條直線的夾角，所以它也位於 Π 內。設鏡面與 Π 相交所得的曲線為 C；它在 P 處的切線（tangent）既在 Π 內，又在鏡面的切平面（tangent plane）內，所以與法線垂直。於是，在平面 Π 內，C 在 P 處的法線平分入射光線與反射光線的夾角：問題化為一個二維問題。

坐標。在 Π 內採用極坐標，以 O 為極點，始線（initial line）沿着 L，並選定方向使入射光線沿 $\theta = 0$ 的方向前進（如圖所示，光線來自 L 上鏡面開口以外的那一側）。設 $P = (r, \theta)$ 在 C 上。



法線的方向。從 P 出發，入射光線沿方向 $(-1, 0)$ 往回走，而反射光線沿方向 $(-\cos\theta, -\sin\theta)$ 射向 O。兩者都是單位向量（unit vector），所以它們的和指向兩者夾角的平分線（bisector）方向，亦即法線方向：

$$(-1-\cos\theta, -\sin\theta) = -2\cos\tfrac{\theta}{2}\left(\cos\tfrac{\theta}{2}, \sin\tfrac{\theta}{2}\right),$$

這裏用了 $1+\cos\theta = 2\cos^2\frac{\theta}{2}$ 和 $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$。所以 P 處的法線與 L 成 $\frac{\theta}{2}$ 角，切線與法線垂直，而 C 在 P 處的斜率（gradient）為

$$m = -\cot\tfrac{\theta}{2}.$$

微分方程。記 $r' = \mathrm{d}r/\mathrm{d}\theta$, $t = \tan\theta$ 及 $\tau = \tan\frac{\theta}{2}$，於是 $t = \dfrac{2\tau}{1-\tau^2}$ 而 $m = -\dfrac{1}{\tau}$。由第 (i) 部分，

$$\frac{r't + r}{r' - rt} = m \quad\Longrightarrow\quad r'(t-m) = -r(1+mt) \quad\Longrightarrow\quad \frac{r'}{r} = -\frac{1+mt}{t-m}.$$

現在

$$1+mt = 1-\frac{2}{1-\tau^2} = -\frac{1+\tau^2}{1-\tau^2}, \qquad t-m = \frac{2\tau}{1-\tau^2}+\frac{1}{\tau} = \frac{1+\tau^2}{\tau(1-\tau^2)},$$

因此

$$\frac{1}{r}\frac{\mathrm{d}r}{\mathrm{d}\theta} = \tan\tfrac{\theta}{2}.$$

? **想一想** 第 (i) 部分在 $\cos\theta = 0$ 之處失效。這會否留下漏洞？

不會。方程 $r'/r = \tan\frac{\theta}{2}$ 只涉及 θ 在 $-\pi < \theta < \pi$ 上的連續函數。除了孤立的值 $\theta = \pm\frac{\pi}{2}$（以及切線垂直的任何孤立點）可能例外，它在每一點都成立，所以由 r' 的連續性，它在這些點上同樣成立。（另一個做法是：半徑與切線之間的夾角 ψ 滿足 $\tan\psi = r/r'$，而由幾何關係可直接得出 $\psi = \frac{\pi}{2} - \frac{\theta}{2}$，並無任何例外點。）

求解。兩邊積分，並注意當 $-\pi < \theta < \pi$ 時 $\cos\frac{\theta}{2} > 0$，

$$\ln r = \int \tan\tfrac{\theta}{2}\,\mathrm{d}\theta = -2\ln\cos\tfrac{\theta}{2} + \text{常數}, \qquad \text{所以} \qquad r = \frac{k}{\cos^2\frac{\theta}{2}} = \frac{2k}{1+\cos\theta}$$

其中 $k > 0$ 為常數（因為 $r > 0$）。令 $\ell = 2k$。於是 $r + r\cos\theta = \ell$，即 $r = \ell - x$。兩邊平方，並利用 $r^2 = x^2 + y^2$，

$$x^2 + y^2 = (\ell - x)^2 \quad \Longleftrightarrow \quad y^2 = \ell^2 - 2\ell x.$$

這是一條以 L 為軸、以 O 為焦點（focus）的拋物線（parabola）（它就是 $Y^2 = 4aX$，其中 $X = \frac{\ell}{2} - x$, $a = \frac{\ell}{2}$，所以焦點位於 $x = \frac{\ell}{2} - a = 0$）。由於 Π 是任意一個包含 L 的平面，鏡面與每個這樣的平面都相交於一條拋物線，正如所求。（事實上，只有當鏡面是單一個旋轉拋物面（paraboloid of revolution）時，常數 ℓ 才會在每個平面內都相同；本題並不需要這一點。）

驗算

- 當 $\theta = 0$ 時得 $r = \frac{\ell}{2}$：頂點位於 O 與準線（directrix）$x = \ell$ 的正中間，這正是以 O 為焦點的拋物線應有的性質。
- 當 $\ell = 2$ 而 $\theta = \frac{\pi}{2}$ 時，該點為 $P = (0,2)$，$y^2 = 4 - 4x$ 在該處的斜率為 $-1 = -\cot\frac{\pi}{4}$，而把方向 $(1,0)$ 對這條切線作反射便得 $(0,-1)$，筆直向下射到 O。
- 數值驗證：第 (i) 部分的斜率公式在 2000 條隨機曲線上與有限差分（finite differences）的結果吻合（準確至 10^{-8}）；以龍格—庫塔法（Runge–Kutta）求解 $r' = r\tan\frac{\theta}{2}$，由 $r(0) = 1.35$ 出發，得 $r(2.5) = 13.5776380$，與 $2.7/(1+\cos 2.5)$ 相符；又在 1000 個隨機點上，把平行於 L 的光線經 $y^2 = \ell^2 - 2\ell x$ 反射後，光線都通過 O。

常見錯誤

- 取錯了角平分線：沿兩條光線背離 P 方向的兩個單位向量，其和給出法線；其差則給出切線。
- 忘記解釋為何問題是平面問題：經過 L 的平面包含入射光線和 O，因而也包含反射光線和法線。
- 寫成 $\ln r = -2\ln\cos\frac{\theta}{2}$ 而漏掉常數，這樣便失去了整族拋物線。

STEP 之外

方程 $r(1+\cos\theta) = \ell$ 是圓錐曲線（conic）的極坐標形式，其焦點位於極點，離心率（eccentricity）為 1；把 1 換成 e 便得到橢圓和雙曲線，而在極坐標下求解牛頓（Newton）方程時，行星軌道正是以這種形式出現的。這個反射論證也正是衛星碟形天線和汽車車頭燈採用拋物面的原因。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）就是以這種方法推導出圓錐曲線軌道的。

詳細解答

STEP II 1989，第 6 題

純數學 • 反函數的積分

> **STEP II 1989，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/89-s2-q6

關鍵思路

右邊有一個 g 對 y 的積分，所以最自然的做法是把左邊的積分改寫成以 $y = \mathrm{f}(x)$ 表示，即作代換（substitution）$x = \mathrm{g}(y)$。新的被積函數（integrand）含有 $y\,\mathrm{g}'(y)$，一看便知要用分部積分法（integration by parts），而邊界項正好給出那兩個長方形。第二條公式的做法完全一樣，只是把 x 換成 $x^2/2$，把面積換成繞 y 軸的旋轉體體積（volume of revolution）。

解答

第一條公式

由於在 $[a, b]$ 上 $\mathrm{f}'(x) > 0$，f 在該區間上嚴格遞增（strictly increasing），所以它是由 $[a, b]$ 到 $[\alpha, \beta]$ 的雙射（bijection），反函數（inverse function）$\mathrm{g}\colon [\alpha, \beta] \to [a, b]$ 存在。此外 g 可微（differentiable），且

$$\mathrm{g}'(y) = \frac{1}{\mathrm{f}'(\mathrm{g}(y))},$$

這個式子有意義，因為 f' 永不為零。

代入 $x = \mathrm{g}(y)$，於是 $\mathrm{f}(x) = y$，$\mathrm{d}x = \mathrm{g}'(y)\,\mathrm{d}y$。當 $x = a$ 時，$y = \mathrm{f}(a) = \alpha$；當 $x = b$ 時，$y = \beta$。因此

$$\int_a^b \mathrm{f}(x)\,\mathrm{d}x = \int_\alpha^\beta y\,\mathrm{g}'(y)\,\mathrm{d}y.$$

現在用分部積分法，取 $u = y$ 和 $\mathrm{d}v = \mathrm{g}'(y)\,\mathrm{d}y$，所以 $v = \mathrm{g}(y)$：

$$\begin{aligned}\int_\alpha^\beta y\,\mathrm{g}'(y)\,\mathrm{d}y &= \Big[y\,\mathrm{g}(y)\Big]_\alpha^\beta - \int_\alpha^\beta \mathrm{g}(y)\,\mathrm{d}y \\ &= \beta\,\mathrm{g}(\beta) - \alpha\,\mathrm{g}(\alpha) - \int_\alpha^\beta \mathrm{g}(y)\,\mathrm{d}y \\ &= b\beta - a\alpha - \int_\alpha^\beta \mathrm{g}(y)\,\mathrm{d}y,\end{aligned}$$

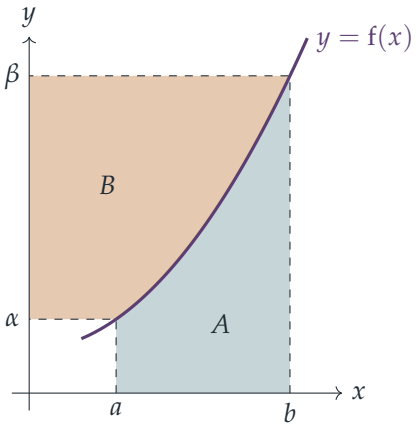
因為 $\mathrm{g}(\beta) = b$ 及 $\mathrm{g}(\alpha) = a$。這就是所求的公式。

幾何意義

取 $a > 0$ 及 $\alpha > 0$。由於 f 遞增，$0 < a < b$ 且 $0 < \alpha < \beta$，所以曲線 $y = \mathrm{f}(x)$ 由 (a, α) 到 (b, β) 的一段弧位於第一象限（first quadrant）。把公式寫成

$$\underbrace{\int_a^b \mathrm{f}(x)\,\mathrm{d}x}_{\text{面積}A} + \underbrace{\int_\alpha^\beta \mathrm{g}(y)\,\mathrm{d}y}_{\text{面積}B} = b\beta - a\alpha.$$

這裏 A 是 $a \le x \le b$ 時曲線與 x 軸之間的面積，B 是 $\alpha \le y \le \beta$ 時曲線與 y 軸之間的面積（把圖形橫過來看，曲線就是 $x = \mathrm{g}(y)$）。兩者合起來，恰好填滿長方形 $[0, b] \times [0, \beta]$ 除去較小的長方形 $[0, a] \times [0, \alpha]$ 後餘下的部分，而這個 L 形區域的面積是 $b\beta - a\alpha$。



第二條公式

用同一個代換 $x = \mathrm{g}(y)$，$\mathrm{d}x = \mathrm{g}'(y)\,\mathrm{d}y$：

$$2\pi \int_a^b x\,\mathrm{f}(x)\,\mathrm{d}x = 2\pi \int_\alpha^\beta y\,\mathrm{g}(y)\,\mathrm{g}'(y)\,\mathrm{d}y.$$

想一想 這一次應該對哪一個因子積分？

留意根據鏈式法則（chain rule），$\mathrm{g}(y)\,\mathrm{g}'(y)$ 是 $\frac{1}{2}[\mathrm{g}(y)]^2$ 的導數。所以取 $u = y$ 和 $v = \frac{1}{2}[\mathrm{g}(y)]^2$ 作分部積分：

$$\begin{aligned} 2\pi \int_\alpha^\beta y\,\mathrm{g}(y)\,\mathrm{g}'(y)\,\mathrm{d}y &= 2\pi \left(\left[\tfrac{1}{2} y\,[\mathrm{g}(y)]^2\right]_\alpha^\beta - \int_\alpha^\beta \tfrac{1}{2}[\mathrm{g}(y)]^2\,\mathrm{d}y \right) \\ &= \pi(\beta b^2 - \alpha a^2) - \pi \int_\alpha^\beta [\mathrm{g}(y)]^2\,\mathrm{d}y, \end{aligned}$$

這裏同樣用到 $\mathrm{g}(\beta) = b$ 及 $\mathrm{g}(\alpha) = a$。這就是第二條公式。

第二條公式的幾何意義

同樣取 $a > 0$、$\alpha > 0$，把上圖繞 y 軸旋轉一整周。

- 區域 A 掃出一個立體，用圓柱殼法（method of cylindrical shells）計算，其體積為 $2\pi \int_a^b x\,\mathrm{f}(x)\,\mathrm{d}x$：與旋轉軸相距 x、高 $\mathrm{f}(x)$、闊 $\mathrm{d}x$ 的窄條，旋轉後成為一個體積約為 $2\pi x\,\mathrm{f}(x)\,\mathrm{d}x$ 的圓柱殼。
- 區域 B 掃出一個立體，用圓盤法（discs）計算，其體積為 $\pi \int_\alpha^\beta [\mathrm{g}(y)]^2\,\mathrm{d}y$：位於高度 y 的窄條，旋轉後成為一個半徑為 $\mathrm{g}(y)$ 的圓盤。
- 兩個立體合起來，恰好填滿一個半徑為 b、高為 β 的圓柱體，再從其底部挖去一個半徑為 a、高為 α 的同軸圓柱體，其體積為 $\pi b^2\beta - \pi a^2\alpha$。

所以這條公式是說：圓柱殼體積加圓盤體積，等於兩個圓柱體的體積之差。條件 $a > 0$ 和 $\alpha > 0$ 確保兩個區域都在旋轉軸的同一側，並且在 x 軸上方，這樣兩個積分才真正是體積，而兩個立體也不會重疊。

驗算

試取 $[1,2]$ 上的 $\mathrm{f}(x) = x^2$，於是 $\alpha = 1$，$\beta = 4$，$\mathrm{g}(y) = \sqrt{y}$。第一條公式左邊是 $\frac{7}{3}$，右邊是 $8 - 1 - \frac{2}{3}(8-1) = \frac{7}{3}$。第二條公式左邊是 $2\pi \cdot \frac{15}{4} = \frac{15\pi}{2}$，右邊是 $\pi(16-1) - \pi \cdot \frac{15}{2} = \frac{15\pi}{2}$。我們亦用辛普森法則（Simpson's rule）對兩條公式作了數值驗算，其中 g 以二分法（bisection）求得；所驗算的幾個函數包括 $[\frac{1}{2}, 2]$ 上的 $x^3 + x$，這時第一條公式兩邊都等於 5.859375。

常見錯誤

- 忘記更改積分上下限：代入 $x = \mathrm{g}(y)$ 之後，上下限變成 α 和 β，而不是 a 和 b。

- 沒有說明 g 為何存在和可微。條件 $f' > 0$ 正是為此而設。
- 把區域 B 說成位於曲線之下。其實它位於曲線與 y 軸之間。

STEP 之外

f 與 f^{-1} 的圖像互為關於 $y = x$ 的反射，而第一條公式正是這個事實的積分版本。它是楊氏不等式（Young's inequality）背後的想法：若 f 遞增、$f(0) = 0$ 且 $a, b \geq 0$，則 $ab \leq \int_0^a f + \int_0^b f^{-1}$；凸分析（convex analysis）和哈密頓力學（Hamiltonian mechanics）中所用的勒讓德變換（Legendre transform），背後也是同一個想法。arcsin 和 ln 等反函數的積分，同樣可以靠它求出，無需任何巧妙的代換。想看代換與反函數的嚴謹處理，可參閱 Abbott 的 *Understanding Analysis*（一本實分析入門書）。

詳細解答

STEP II 1989，第 7 題

純數學 • 二階微分方程

> **STEP II 1989，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/89-s2-q7

關鍵思路

$x\,y''$ 與 $-b\,y'$ 這兩項放在一起看似棘手，但只要更換自變量（independent variable），設 $t = x^{\alpha}$，y'' 便會產生一個一階導數項，而這一項可以恰好與 $-b\,y'$ 抵消。適當選取 α 做到這一點之後，餘下的就是關於 t 的簡諧運動（simple harmonic motion）。此後，當 $x \to 0$ 時的表現，就取決於 $\mathrm{d}y/\mathrm{d}x$ 中的因子 x^{b} 如何變化，而這又取決於 b 的正負。

解答

更換變量

設 $t = x^{\alpha}$，其中 α 待定，並把 y 視為 t 的函數。記 $\dot{y} = \mathrm{d}y/\mathrm{d}t$，由鏈式法則（chain rule）得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \alpha x^{\alpha-1}\dot{y},$$
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \alpha(\alpha-1)x^{\alpha-2}\dot{y} + \alpha^2 x^{2\alpha-2}\ddot{y},$$

其中第二行用了積法則（product rule）以及 $\frac{\mathrm{d}}{\mathrm{d}x}\dot{y} = \alpha x^{\alpha-1}\ddot{y}$。代入後得

$$x\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - b\frac{\mathrm{d}y}{\mathrm{d}x} = \alpha^2 x^{2\alpha-1}\ddot{y} + \alpha(\alpha-1-b)\,x^{\alpha-1}\dot{y}.$$

若取 $\alpha = b+1$，$\dot{y}$ 項便會消失；由於 $b > -1$，這個值不等於零。此時 $2\alpha - 1 = 2b+1$，於是方程變成

$$(b+1)^2 x^{2b+1}\ddot{y} + x^{2b+1}y = 0 \quad\Longrightarrow\quad \ddot{y} + \frac{y}{(b+1)^2} = 0,$$

這裏兩邊除以了 $x^{2b+1} \neq 0$（因為 $x > 0$）。這是關於 t 的簡諧運動，角頻率（angular frequency）為 $1/(b+1)$，所以

$$\boxed{y = A\cos\left(\frac{x^{b+1}}{b+1}\right) + B\sin\left(\frac{x^{b+1}}{b+1}\right)}$$

其中 A、B 為任意常數。二階方程配上兩個任意常數：這就是通解（general solution）。

當 $x \to 0$ 時的表現

記 $u = x^{b+1}/(b+1)$。由於 $b+1 > 0$，當 $x \to 0^+$ 時 $u \to 0^+$，所以

$$y \to A\cos 0 + B\sin 0 = A.$$

因此 $y \to 1$ 迫使 $A = 1$。利用 $\mathrm{d}u/\mathrm{d}x = x^b$ 求導，得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x^b\left(-\sin u + B\cos u\right) \qquad (A = 1).$$

情況 $b > 0$。此時 $x^b \to 0$，而括號內的式子趨向 B，是一個有限值。所以無論 B 取甚麼值，都有 $\mathrm{d}y/\mathrm{d}x \to 0$。以下這族函數

$$y = \cos\left(\frac{x^{b+1}}{b+1}\right) + B\sin\left(\frac{x^{b+1}}{b+1}\right), \qquad B \in \mathbb{R},$$

中的每一個都滿足兩個條件。解是存在的，但並不唯一，正如題目所要求證明的。

> ? **想一想** 對於一般的二階方程，給定 y 的一個值和 y' 的一個值，就能確定解。為甚麼在這裏卻行不通？

唯一性定理（uniqueness theorem）要求方程可寫成 $y'' = f(x, y, y')$ 的形式，而且 f 的性質要夠好。兩邊除以 x 得 $y'' = (b/x)\,y' - x^{2b}y$，而係數 b/x 在 $x = 0$ 處趨向無窮。$x = 0$ 這一點是奇點（singular），所以通常的保證在該處並不適用。

情況 $b = 0$。此時 $x^b = 1$，而 $\mathrm{d}y/\mathrm{d}x \to B$。條件迫使 $B = 0$，得出唯一解 $y = \cos x$。（這其實就是 $y'' + y = 0$。）

情況 $-1 < b < 0$。此時 $x^b \to \infty$。若 $B \neq 0$，則 $Bx^b\cos u \to \pm\infty$，而另一項很小：由於 $\sin u \leq u$，

$$\left|x^b \sin u\right| \leq x^b \cdot \frac{x^{b+1}}{b+1} = \frac{x^{2b+1}}{b+1}.$$

更準確地說，$\sin u/u \to 1$，所以 $x^b \sin u \sim x^{2b+1}/(b+1)$。因此必須有 $B = 0$，這時

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -x^b \sin u \sim -\frac{x^{2b+1}}{b+1}.$$

這個式子趨向 0，當且僅當 $2b+1>0$，即 $b>-\frac{1}{2}$。當 $b=-\frac{1}{2}$ 時它趨向 -2，而當 $b<-\frac{1}{2}$ 時它無界。所以當 $-\frac{1}{2}<b<0$ 時恰好有一個解 $y=\cos\left(x^{b+1}/(b+1)\right)$，而當 $-1<b\le-\frac{1}{2}$ 時則沒有解。

結論

這些條件能確定唯一解，恰好是在

$$-\tfrac{1}{2}<b\le 0,$$

的時候，而該解為 $y=\cos\left(x^{b+1}/(b+1)\right)$。



驗算

- $b=0$ 時方程為 $y''+y=0$，其解 $A\cos x+B\sin x$ 與公式吻合。
- 數值上，通解在捨入誤差範圍內滿足方程（在隨機選取的 x, b, A, B 處用有限差分檢驗）。取 $b=-\frac{1}{4}$, $B=0$，我們在 $x=10^{-9}$ 處得到 $y'\approx -4.2\times10^{-5}$，趨向 0。取 $b=-\frac{1}{2}$ 時，該值保持在 -2；而取 $b=1, B=5$ 時，在 $x=10^{-9}$ 處的值為 5×10^{-9}。

常見錯誤

- 做到「$b\le0$ 時解唯一」就停下來：當 $-1<b\le-\frac{1}{2}$ 時，即使是餘弦解也有 $y'\not\to 0$，所以根本沒有解。
- 把 $\alpha x^{\alpha-1}\dot{y}$ 對 x 求導時忘記用積法則。多出來的那一項正是用來抵消 $-b\,y'$ 的。

STEP 之外

$x=0$ 這一點是這個方程的正則奇點（regular singular point）。在這類點附近，弗羅貝尼烏斯法（Frobenius method）會尋找形如 $x^r\sum c_n x^n$ 的解，而這裏的指標方程（indicial equation）的根為 $r=0$ 和 $r=b+1$。在奇點處哪些條件能確定一個解，取決於這些指數，正如我們剛才所見。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教材）涵蓋了級數解和正則奇點。

詳細解答

STEP II 1989，第 8 題

純數學 • 變換群

> **STEP II 1989，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/89-s2-q8

關鍵思路

一切都建基於一條恆等式（identity）：$\Omega^3 = -1$，所以 Ω 滿足第一行的二次方程，我們因而可以把 Ω 的任何次冪化為 Ω 的線性式。一旦知道 R 和 T 怎樣調動這四個點，餘下的點算工夫便只與這四個點有關，而與公式無關：這個羣（group）必須能把 z_0 送到四個點中的每一個，而它本身已包含三個保持 z_0 不動的不同映射。至於最後一部分，不妨問問 R 和 T 屬於哪一類置換（permutation）。

解答

恆等式

由於 $\Omega = e^{i\pi/3}$，我們有 $\Omega^3 = e^{i\pi} = -1$，所以

$$0 = \Omega^3 + 1 = (\Omega + 1)(\Omega^2 - \Omega + 1).$$

因為 $\Omega \neq -1$，第二個因式必為零：$\Omega^2 - \Omega + 1 = 0$，這正是所求。

我們會以下列形式運用這條式：

$$\Omega^2 = \Omega - 1, \qquad \Omega^3 = -1, \qquad \Omega^4 = -\Omega.$$

另外留意 $2\Omega - 1 = 2\cos\frac{\pi}{3} - 1 + 2i\sin\frac{\pi}{3} = i\sqrt{3} \neq 0$。

R 和 T 置換這四個點

先看 R：乘以 $\Omega^2 = e^{2i\pi/3}$ 即是繞 0 旋轉 120°：

$$R(0) = 0, \qquad R(1) = \Omega^2, \qquad R(\Omega^2) = \Omega^4 = -\Omega, \qquad R(-\Omega) = -\Omega^3 = 1.$$

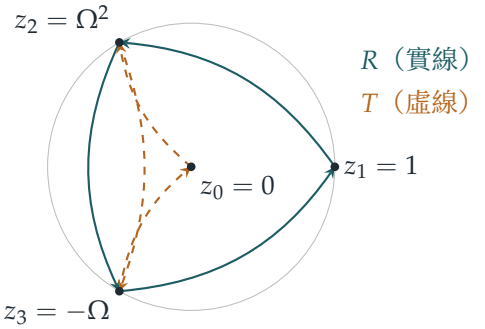
所以 R 保持 z_0 不動，並把 $z_1 \to z_2 \to z_3 \to z_1$。

再看 T：把各點逐一代入，並利用 $\Omega^2 = \Omega - 1$ 化簡：

$$\begin{aligned} T(0) &= \frac{\Omega^2}{1} = z_2, \\ T(1) &= \frac{\Omega + \Omega^2}{2\Omega^2 + 1} = \frac{2\Omega - 1}{2\Omega - 1} = 1 = z_1, \\ T(\Omega^2) &= \frac{\Omega^3 + \Omega^2}{2\Omega^4 + 1} = \frac{\Omega - 2}{1 - 2\Omega} = -\Omega = z_3, \\ T(-\Omega) &= \frac{-\Omega^2 + \Omega^2}{-2\Omega^3 + 1} = \frac{0}{3} = 0 = z_0. \end{aligned}$$

在第三行中，$\Omega^3 + \Omega^2 = -1 + (\Omega - 1) = \Omega - 2$，而最後一步成立，是因為 $-\Omega(1 - 2\Omega) = -\Omega + 2\Omega^2 = -\Omega + 2\Omega - 2 = \Omega - 2$。所有分母均不為零（它們分別是 1、$\mathrm{i}\sqrt{3}$、$-\mathrm{i}\sqrt{3}$ 和 3）。

所以 T 保持 z_1 不動，並把 $z_0 \to z_2 \to z_3 \to z_0$。$R$ 和 T 都置換這四個點；兩者各是一個 3-循環（cycle）。



這個羣的階至少為 12

設 G 為包含 R 和 T 的最小變換羣：即由 R、T 及其逆變換經複合（composition）所能得到的一切變換。G 的每個元素都把集合 $\{z_0, z_1, z_2, z_3\}$ 映射到它自身之上，因為 R 和 T 如此（它們的逆變換亦然，因為那是這個有限集合上的雙射（bijection）的逆）。

三個保持 z_0 不動的映射。集合 $H = \{I, R, R^2\}$ 包含於 G 之內，其中每個元素都保持 z_0 不動，而且三者是不同的變換（分別是旋轉 $0°$、$120°$、$240°$）。

把 z_0 送到每一個點的映射。利用上面求得的循環，

$$g_0 = I, \qquad g_1 = R^2T, \qquad g_2 = T, \qquad g_3 = RT$$

滿足 $g_k(z_0) = z_k$：例如 $RT(z_0) = R(z_2) = z_3$ 以及 $R^2T(z_0) = R^2(z_2) = R(z_3) = z_1$。

現在考慮十二個乘積 $g_k h$，其中 $k \in \{0,1,2,3\}$ 而 $h \in H$。由於 $h(z_0) = z_0$，我們得到 $g_k h(z_0) = z_k$。

想一想 這十二個乘積當中，會否有兩個其實是同一個變換？

不會。若 $g_k h = g_l h'$，把兩邊同時作用於 z_0：我們得到 $z_k = z_l$，所以 $k = l$。然後在左邊複合上 g_k^{-1}，便得 $h = h'$。因此這十二個乘積是 G 中互不相同的元素，於是 $|G| \geq 12$。（用羣論的語言來說：z_0 的軌道（orbit）大小為 4，而它的穩定子（stabiliser）大小至少為 3，所以根據軌道－穩定子定理（orbit–stabiliser theorem），$|G| \geq 4 \times 3$。）

最後一部分

有的：這四個點的 24 個置換當中，恰好有一半永遠無法產生。

想一想 作為置換，R 和 T 有甚麼共通之處？

兩者各是一個 3-循環，而 3-循環是偶置換（even permutation）：例如 $(z_1\, z_2\, z_3) = (z_1\, z_3)(z_1\, z_2)$，是兩個對換（transposition）的乘積。它們的逆也是 3-循環。偶置換的複合仍是偶置換，所以由 R 和 T 產生的每一個點置換都是偶置換。四件物件只有 12 個偶置換，所以其餘 12 個，即奇置換，永遠不會產生。例如，把 z_0 和 z_1 對調而保持 z_2 和 z_3 不動，是單獨一個對換，因而是奇置換，不可能得到。

事實上，結合上一部分，這就完全確定了 G。R 和 T 都是莫比烏斯變換（Möbius map）（T 有 $ad - bc = \Omega - 2\Omega^4 = 3\Omega \neq 0$），而保持三個點不動的莫比烏斯變換必為恆等變換，所以 G 的每個元素都由它怎樣置換這四個點所決定。因此 $|G| \leq 12$，再配合上一部分，$|G| = 12$：G 在這些點上的作用，正是全體偶置換所成的羣 A_4。

驗算

- 以數值計算，$|\Omega^2 - \Omega + 1| \approx 2 \times 10^{-16}$，而在四個點上計算 R 和 T 的值，所得的正是上述兩個 3-循環。
- $R^3 = I$ 顯然成立；而 $T^3 = I$ 亦成立，正如 3-循環所應有的性質（已在一個隨機點上檢驗，準確至 10^{-16}）。
- 用電腦生成 R 和 T 的所有複合，恰好得到 12 個不同的映射，全部誘導出偶置換，其中 z_0 被移到全部四個點，並有 3 個映射保持它不動。

常見錯誤

- 論證說「G 含有階（order）為 3 的元素，所以 $|G| \geq 12$」。單憑階為 3 的元素，只能得出 3 整除 $|G|$；因子 4 是來自在四個點上的可遷性（transitivity）。
- 沒有化簡 Ω 的冪，以致 $T(\Omega^2) = -\Omega$ 看似不成立。應先把一切化為 $a + b\Omega$ 的形式。

STEP 之外

$|G| =$ (軌道大小) $\times$ (穩定子大小) 這個點算結果就是軌道－穩定子定理，而全體偶置換構成交錯羣（alternating group）A_4，它也是正四面體的旋轉羣。這裏出現這個現象，背後有個很漂亮的原因：莫比烏斯變換保持四個點的交比（cross-ratio）不變，而這四個點的交比是 $e^{-i\pi/3}$，正是那個特殊的「等非調和」（equianharmonic）值，對這個值而言，恰好只有偶置換保持交比不變。關於羣和羣作用（group action），可參閱 Allenby 的 *Rings, Fields and Groups*；關於莫比烏斯變換和黎曼球面（Riemann sphere），可參閱 Needham 的 *Visual Complex Analysis*。

詳細解答

STEP II 1989，第 9 題

純數學 • 矩陣指數與方程組

STEP II 1989，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/89-s2-q9

關鍵思路

一切都繫於一個小小的計算：$\mathbf{A}^2$ 原來就是單位矩陣（identity matrix）。於是 $\mathbf{A}$ 的每個偶次冪都是 $\mathbf{I}$，每個奇次冪都是 $\mathbf{A}$，所以該級數可拆成指數級數（exponential series）的偶數項部分和奇數項部分，也就是 $\cosh t$ 和 $\sinh t$。至於微分方程（differential equation），我們仿照純量（scalar）情況下的積分因子（integrating factor）法，由矩陣 $\mathbf{F}$ 擔當 $\mathrm{e}^{\int a\,\mathrm{d}t}$ 的角色。

解答

矩陣 F

乘開可得

$$\mathbf{A}^2 = \begin{pmatrix} -3 & -1 \\ 8 & 3 \end{pmatrix}\begin{pmatrix} -3 & -1 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} 9-8 & 3-3 \\ -24+24 & -8+9 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}.$$

因此對所有 $k \geq 0$，$\mathbf{A}^{2k} = (\mathbf{A}^2)^k = \mathbf{I}$ 及 $\mathbf{A}^{2k+1} = \mathbf{A}^{2k}\mathbf{A} = \mathbf{A}$。寫成 $\mathbf{I} = \frac{t^0}{0!}\mathbf{A}^0$，並把級數按 n 為偶數和奇數拆開（矩陣級數的每個元素都是絕對收斂（absolutely convergent）純量級數的重排，所以這樣做是容許的），

$$\begin{aligned}\mathbf{F} &= \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!}\mathbf{A}^{2k} + \sum_{k=0}^{\infty} \frac{t^{2k+1}}{(2k+1)!}\mathbf{A}^{2k+1} \\ &= \mathbf{I}\sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!} + \mathbf{A}\sum_{k=0}^{\infty} \frac{t^{2k+1}}{(2k+1)!} = \mathbf{I}\cosh t + \mathbf{A}\sinh t,\end{aligned}$$

當中用到麥克勞林級數（Maclaurin series）$\cosh t = \sum t^{2k}/(2k)!$ 和 $\sinh t = \sum t^{2k+1}/(2k+1)!$。這正是所求的形式。

逆矩陣

設 $\mathbf{G} = \mathbf{I}\cosh t - \mathbf{A}\sinh t$。由於 $\mathbf{I}$ 與 $\mathbf{A}$ 可交換（commute），我們可以像處理普通數字一樣乘開：

$$\mathbf{FG} = \mathbf{I}\cosh^2 t - \mathbf{A}^2\sinh^2 t + (\mathbf{A} - \mathbf{A})\cosh t \sinh t = \mathbf{I}(\cosh^2 t - \sinh^2 t) = \mathbf{I},$$

當中用到 $\mathbf{A}^2 = \mathbf{I}$ 和 $\cosh^2 t - \sinh^2 t = 1$。同樣的計算給出 $\mathbf{GF} = \mathbf{I}$，所以 $\mathbf{F}^{-1} = \mathbf{I}\cosh t - \mathbf{A}\sinh t$。

導數

逐個元素求導（矩陣 $\mathbf{I}$ 和 $\mathbf{A}$ 都是常數矩陣），

$$\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}t} = \mathbf{I}\sinh t + \mathbf{A}\cosh t, \qquad \text{而} \qquad \mathbf{FA} = \mathbf{A}\cosh t + \mathbf{A}^2\sinh t = \mathbf{A}\cosh t + \mathbf{I}\sinh t.$$

兩者相同，所以 $\dfrac{\mathrm{d}\mathbf{F}}{\mathrm{d}t} = \mathbf{FA}$。另外留意 $\mathbf{FA} = \mathbf{AF}$，因為 $\mathbf{F}$ 是由 $\mathbf{I}$ 和 $\mathbf{A}$ 組合而成的。

微分方程

在純量方程 $r' + ar = 0$ 中，我們乘以 e^{at}，它的導數是 $\mathrm{e}^{at}a$，於是左方恰好變成某個式子的導數。矩陣 $\mathbf{F}$ 正有這個性質，即 $\mathbf{F}' = \mathbf{FA}$，所以我們在方程的左邊乘上 $\mathbf{F}$：

$$\mathbf{F}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} + \mathbf{FAr} = \mathbf{0}.$$

根據積法則（product rule）（只要保持各因子的次序，它對矩陣乘積同樣成立），

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{Fr}) = \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}t}\mathbf{r} + \mathbf{F}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{FAr} + \mathbf{F}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t},$$

這恰好就是左方。所以方程說的是

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{Fr}) = \mathbf{0},$$

而 $\mathbf{Fr}$ 是一個常向量 $\mathbf{c}$。

> ? **想一想** 為甚麼要在左邊乘而不是在右邊乘？又怎樣求出 $\mathbf{c}$？

乘積 $\mathbf{rF}$ 根本沒有意義（$\mathbf{r}$ 是一個 2×1 的列向量），而且積法則得出的是 $\mathbf{FAr}$，只有當 $\mathbf{F}$ 在左邊時，它才與 $\mathbf{F}(\mathbf{Ar})$ 這一項吻合。至於常向量，代入 $t = 0$：這時 $\cosh 0 = 1$ 而 $\sinh 0 = 0$，所以 $\mathbf{F} = \mathbf{I}$，並且

$$\mathbf{c} = \mathbf{I}\,\mathbf{r}(0) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

現在在 $\mathbf{Fr} = \mathbf{c}$ 的左邊乘上前面已求得的 $\mathbf{F}^{-1}$：

$$\mathbf{r} = \mathbf{F}^{-1}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \cosh t\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \sinh t\begin{pmatrix}-3 & -1\\ 8 & 3\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \cosh t\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \sinh t\begin{pmatrix}-3\alpha-\beta\\ 8\alpha+3\beta\end{pmatrix}.$$

讀出各分量，

$$\begin{aligned} x(t) &= \alpha\cosh t + (3\alpha+\beta)\sinh t,\\ y(t) &= \beta\cosh t - (8\alpha+3\beta)\sinh t, \end{aligned}$$

正如所求。

驗算

- 在 $t = 0$ 時，公式給出 $x = \alpha$、$y = \beta$，理應如此。
- 直接代入：$x' = \alpha\sinh t + (3\alpha+\beta)\cosh t$，而 $-(-3x - y) = 3x + y = (3\alpha + \beta)\cosh t + \alpha\sinh t$（用到 $9\alpha + 3\beta - 8\alpha - 3\beta = \alpha$）。所以 $x' = -$(第一行，取自$\mathbf{Ar}$)，y 的情況亦類似。
- 數值上，把級數加到第 60 項，結果與 $\mathbf{I}\cosh t + \mathbf{A}\sinh t$ 吻合；而取 $\alpha = 1$、$\beta = 0$ 時，以龍格—庫塔(Runge–Kutta)法解這個方程組，得出 $x(2) \approx 14.6428$，與 $\cosh 2 + 3\sinh 2$ 相符。

常見錯誤

- 在右邊乘上 $\mathbf{F}$，又或者把 $\mathbf{AF}$ 和 $\mathbf{FA}$ 隨意互換，卻沒有說明為何在這裏可以這樣做（它們可交換，是因為 $\mathbf{F}$ 是 $\mathbf{I}$ 和 $\mathbf{A}$ 的組合）。
- 拆開級數時漏掉 $\mathbf{I}$ 這一項（即 $n = 0$ 的一項），結果失去了 $\cosh t$ 的常數項 1。
- 把 $-\mathbf{A}$ 乘到 $(\alpha, \beta)^{\mathrm{T}}$ 上時弄錯正負號。

STEP 之外

矩陣 $\mathbf{F}$ 就是矩陣指數（matrix exponential）$\mathrm{e}^{t\mathbf{A}}$，而這裏的方法說明 $\mathbf{r}' = \mathbf{Mr}$ 的解總是 $\mathbf{r} = \mathrm{e}^{t\mathbf{M}}\mathbf{r}(0)$；本題中 $\mathbf{M} = -\mathbf{A}$，所以答案才會涉及 $\mathbf{F}^{-1} = \mathrm{e}^{-t\mathbf{A}}$。一般而言，計算 $\mathrm{e}^{t\mathbf{M}}$ 的方法是把 $\mathbf{M}$ 對角化（diagonalising）：我們的 $\mathbf{A}$ 的特徵值（eigenvalues）是 ± 1（這正是 $\mathbf{A}^2 = \mathbf{I}$ 的含意），$\mathrm{e}^{\pm t}$ 以至 cosh 和 sinh 就是由此而來。Strang 的 *Introduction to Linear Algebra*（線性代數入門）有一章把矩陣指數和微分方程組講得很好，而 Boyce 與 DiPrima 的 *Elementary Differential Equations*（初等微分方程）則詳細討論線性方程組。

詳細解答

STEP II 1989，第 10 題

純數學 • 直線與平面

> **STEP II 1989，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/89-s2-q10

關鍵思路

空間中的兩條直線不交於唯一一點，有三種情況：兩者是同一條直線；兩者平行；或者兩者是*異面直線*（skew lines），分處不同平面而互不相遇。因此「恰好交於一點」需要兩個條件：兩個方向不平行，而且兩條直線位於同一平面上。向量積（cross product）可以同時處理這兩個條件。至於那兩個平面，其交線的方向就是兩個法向量（normal）的向量積，而最後一部分其實只是第一部分換了個樣子再出現一次。

解答

第一部分：兩條直線何時恰好交於一點

兩條直線相交之處滿足 $\mathbf{a}+\lambda\mathbf{u}=\mathbf{b}+\mu\mathbf{v}$，即

$$\lambda\mathbf{u}-\mu\mathbf{v}=\mathbf{b}-\mathbf{a}.$$

要它們是真正的直線，需要 $\mathbf{u}\neq\mathbf{0}$ 和 $\mathbf{v}\neq\mathbf{0}$。於是：

- 若 $\mathbf{u}$ 與 $\mathbf{v}$ 平行（$\mathbf{u}\times\mathbf{v}=\mathbf{0}$），兩條直線要麼平行而相異（沒有公共點），要麼完全重合（有無限多個公共點）。無論哪一種，都不是恰好一點。
- 若 $\mathbf{u}\times\mathbf{v}\neq\mathbf{0}$，則 $\mathbf{u}$ 與 $\mathbf{v}$ 張成（span）一個平面，而方程有解當且僅當 $\mathbf{b}-\mathbf{a}$ 位於該平面上，即與 $\mathbf{u}\times\mathbf{v}$ 垂直。解若存在則必唯一，因為 $\mathbf{u}$ 與 $\mathbf{v}$ 線性無關（independent）。

所以條件是

$$\mathbf{u}\neq\mathbf{0},\quad \mathbf{v}\neq\mathbf{0},\quad \mathbf{u}\times\mathbf{v}\neq\mathbf{0},\quad (\mathbf{b}-\mathbf{a})\cdot(\mathbf{u}\times\mathbf{v})=0.$$

（頭兩個條件已由第三個推出，但仍值得寫明。）若最後一個條件不成立，兩條直線便是

異面直線。

第二部分：兩個平面

兩個法向量是 $\mathbf{n}_1 = (1,1,b)$ 和 $\mathbf{n}_2 = (b,b,1)$，而

$$\mathbf{n}_1 \times \mathbf{n}_2 = \left(1-b^2,\ b^2-1,\ 0\right) = (1-b^2)(1,-1,0).$$

兩個平面不交於一條直線，當且僅當它們的法向量平行；在這裏即 $1-b^2 = 0$，所以 $b = 1$ 或 $b = -1$。

- $b = 1$：兩條方程都變成 $x+y+z = 3$。
- $b = -1$：兩條方程是 $x+y-z = 1$ 和 $-x-y+z = -1$，同樣是同一個平面。

所以在這兩種情況下，兩個平面都是重合，而不是平行而分開。

現在取 $b \neq \pm 1$。交線沿 $\mathbf{n}_1 \times \mathbf{n}_2$ 的方向，所以可以取 $\mathbf{u} = (1,-1,0)$。要找出線上的點，留意 x 和 y 只以 $s = x+y$ 的形式出現：

$$s + bz = b+2, \qquad bs + z = 2b+1.$$

從第二式減去第一式的 b 倍，得 $(1-b^2)z = 1-b^2$，所以 $z = 1$（由於 $b^2 \neq 1$，可以相除），繼而 $s = 2$。因此交線是 $x+y = 2$，$z = 1$，與 b 無關。

我們需要在這條直線上找一點 $\mathbf{a}$ 使得 $\mathbf{a}\cdot\mathbf{u} = 0$，即 $x-y = 0$。連同 $x+y = 2$，得 $x = y = 1$，所以

$$\mathbf{r} = \begin{pmatrix}1\\1\\1\end{pmatrix} + \lambda\begin{pmatrix}1\\-1\\0\end{pmatrix}, \qquad b \neq \pm 1.$$

點 $\mathbf{a}$ 就是由原點到這條直線的垂足。

最後部分

坐標同時滿足兩條方程 $(*)$ 的點，當 $b \neq \pm 1$ 時組成上述直線，當 $b = \pm 1$ 時則組成整個平面。我們把這兩種情況分開處理。

> **想一想** 當 $b = \pm 1$ 時，同時滿足兩條方程的點組成甚麼集合？

是整個平面，而不是空集，因為我們已發現兩個平面重合。在這些情況下，答案因此完全不同。

情況 $b \neq \pm 1$。我們需要題目所給的直線與直線 $\mathbf{r} = (1,1,1) + \lambda(1,-1,0)$ 恰好交於一點。應用第一部分，取 $\mathbf{a} = (1,1,1)$，$\mathbf{u} = (1,-1,0)$，$\mathbf{b} = (0,0,c)$，$\mathbf{v} = (1,d,0)$。於是

$$\mathbf{u}\times\mathbf{v} = (0,\,0,\,d+1), \qquad (\mathbf{b}-\mathbf{a})\cdot(\mathbf{u}\times\mathbf{v}) = (-1,-1,c-1)\cdot(0,0,d+1) = (c-1)(d+1).$$

我們需要 $\mathbf{u}\times\mathbf{v}\neq\mathbf{0}$，所以 $d\neq -1$，於是 $(c-1)(d+1)=0$ 迫使 $c=1$。所以條件是 $c=1$ 和 $d\neq -1$。（直接驗證：題目所給的直線上 $z=c$，所以只有當 $c=1$ 時它才與 $z=1$ 相遇，而此時 $x+y=\mu(1+d)=2$ 在 $d\neq -1$ 時有唯一解 $\mu=2/(1+d)$。）

情況 $b=\pm 1$。現在我們需要題目所給的直線與一個平面恰好交於一點。點 $(\mu,d\mu,c)$ 位於 $x+y+z=3$ 上（對應 $b=1$）當 $\mu(1+d)=3-c$，位於 $x+y-z=1$ 上（對應 $b=-1$）當 $\mu(1+d)=1+c$。每條方程對 μ 恰有一個解，當且僅當 $1+d\neq 0$；若 $d=-1$，直線與平面平行，方程或無解，或有無限多個解。這裏 c 可取任何值。

答案。恰有一個這樣的點，當且僅當

$$d\neq -1 \quad 及 \quad \big(c=1 \text{ 或 } b=\pm 1\big),$$

可以簡潔地寫成 $d\neq -1$ 及 $(b^2-1)(c-1)=0$。

驗算

- 把 $(1+\lambda,1-\lambda,1)$ 代入 $(*)$：$2+b=b+2$ 及 $2b+1=2b+1$，對所有 b 和 λ 都成立。
- 取 $b=2$，$c=1$，$d=3$，交點對應 $\mu=\frac{1}{2}$，即 $(\frac{1}{2},\frac{3}{2},1)$：$\frac{1}{2}+\frac{3}{2}+2=4=b+2$ 及 $1+3+1=5=2b+1$。
- 用電腦以精確算術測試了 $\{-3,\ldots,3,\frac{1}{2},-\frac{5}{3}\}$ 中所有的 b,c,d（共 729 個情況），直接求解 μ，結果解的數目為一，恰好就是上述條件成立之時。

常見錯誤

- 只因兩條直線不平行，就說它們交於一點。在三維空間中，它們可以是異面直線。
- 以為當 $b=\pm 1$ 時兩個平面平行而分開，從而斷定完全沒有這樣的點。檢查一下右邊的常數：兩個平面是重合的。
- 除以 $1-b^2$ 而沒有說明它為何不等於零。

STEP 之外

這其實是一個小型的線性方程組（system of linear equations），只是換了個樣子：三個未知數的兩條方程，當係數矩陣的秩（rank）為 2 時，解組成一條直線；當秩為 1 時，解組成一個平面（或者無解）。加入直線的參數式後，它就變成一個關於 μ 的超定方程組（overdetermined system），而「恰有一個解」就是一個關於秩與相容性（consistency）的

命題。Strang 的 *Introduction to Linear Algebra*（《線性代數導論》）正是由這裏出現的行圖像（平面相交）和列圖像（向量的組合）出發，建立起這一切。

詳細解答

STEP II 1989，第 11 題

力學 • 連接質點與衝量

> **STEP II 1989，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/89-s2-q11

關鍵思路

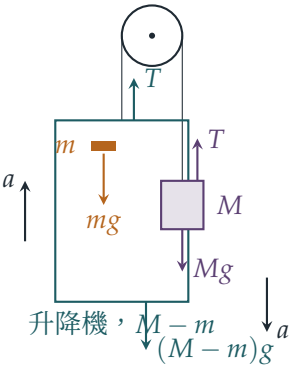
升降機與平衡錘一開始互相平衡，所以升降機的質量 M 必定已包括瓷磚在內。瓷磚一旦脫落，升降機便比平衡錘輕，於是向上加速，而瓷磚則自由下墜，地板因而升上來迎向瓷磚。處理碰撞時，把升降機、纜索和平衡錘視為沿纜索運動的單一系統。這個系統的「沿纜索的動量」只會因重力而改變，而碰撞產生的衝量（impulse）則屬於系統的內力。

解答

建立模型

我們把升降機、平衡錘和瓷磚都當作質點（particle），並假設纜索輕而不可伸長（light and inextensible），繞過一個光滑而輕的滑輪（pulley）。導軌是光滑的，所以不會施加任何垂直方向的力。瓷磚掉下之前，系統處於靜止狀態，兩邊的質量分別為 M 和 M。因此升降機的質量 M 已包括瓷磚。（否則升降機一邊的重量是 $(M+m)g$，另一邊只有 Mg，系統不可能靜止。）

瓷磚脫落之後，升降機的質量為 $M-m$，平衡錘的質量為 M，而瓷磚是一個自由質點。



到達地板所需的時間

設升降機以加速度 a 向上加速，則平衡錘以加速度 a 向下加速。設 T 為纜索的張力（tension）。由牛頓第二定律（Newton's second law）可得

$$\begin{aligned} &\text{升降機（向上）：} && T-(M-m)g=(M-m)a,\\ &\text{平衡錘（向下）：} && Mg-T=Ma. \end{aligned}$$

把兩式相加便可消去 T：

$$a=\frac{mg}{2M-m},\qquad T=M(g-a)=\frac{2M(M-m)g}{2M-m}>0,$$

所以纜索一直保持拉緊。

想一想 為甚麼不能直接說瓷磚在重力作用下墜落了距離 h？

因為地板也在移動。瓷磚以 g 向下加速，地板以 a 向上加速，兩者都由靜止開始。因此兩者之間的距離以恆定的相對加速度（relative acceleration）收窄：

$$g+a=g+\frac{mg}{2M-m}=\frac{2Mg}{2M-m}.$$

這段距離最初是 h，所以瓷磚在 $\frac{1}{2}(g+a)t^2=h$ 時着地，由此可得

$$t^2=\frac{2h}{g+a}=\frac{(2M-m)h}{Mg},\qquad t=\sqrt{\frac{(2M-m)h}{Mg}}.$$

這並不是題目所印的式子，題目中出現的是 $2(M-m)$ 而不是 $2M-m$。兩者只在 $m=0$ 時才一致，而我們的結果會在下文以數值方法逐步模擬運動來驗證。題目所印的式子，需要升降機以 $mg/(M-m)$ 加速才會成立。如果忘記了平衡錘也要被加速，並把張力當作維持在 Mg，就會得到這個結果。我們相信題目所印的公式是原題的誤植或疏忽。題目其餘部分完全不依賴這個式子。

碰撞

在撞擊前的一刻，升降機向上運動，平衡錘向下運動，兩者的速率都是 $u = at$。瓷磚則以速率 $v = gt$ 向下運動。於是

$$(2M - m)u = (2M - m) \cdot \frac{mg}{2M - m}\, t = mgt = mv.$$

現在考慮碰撞本身。取「正方向」為升降機向上／平衡錘向下，即纜索帶動兩者一起運動的方向。撞擊歷時極短，所以有限的力（各物體的重量）所產生的衝量可以忽略。僅有的衝量是瓷磚與地板之間的衝量（它是「升降機連瓷磚」這個整體的內力），以及纜索中的衝擊張力（impulsive tension）J。設 w 為碰撞後的共同速率（升降機連瓷磚，質量為 M，向上；平衡錘向下）。碰撞是完全非彈性的（perfectly inelastic），所以瓷磚最終隨地板一起運動。於是

$$\begin{aligned} \text{升降機連瓷磚：} \quad & Mw - \big[(M - m)u - mv\big] = J, \\ \text{平衡錘：} \quad & Mw - Mu = -J. \end{aligned}$$

把兩式相加，

$$2Mw = (M - m)u + Mu - mv = (2M - m)u - mv = 0,$$

所以 $w = 0$：升降機停了下來。（另外 $J = Mu > 0$，所以纜索確實是被猛然拉緊而不是變鬆，這與我們的假設一致。）撞擊之後，兩邊的質量再次是 M 和 M，互相平衡，所以一切保持靜止。

> **想一想** 為甚麼在撞擊前的一刻，總的「沿纜索的動量」是零？

撞擊之前，作用在這三個物體上而又沿運動方向的外力只有重力，按上面所定的正方向計算，其淨效果為 $Mg - (M - m)g - mg = 0$。（滑輪的反作用力不作功，而纜索屬於系統內部。）所以這個動量自始至終保持在初始值 0。碰撞同樣不能改變它。

能量損失

系統由靜止開始，最後也歸於靜止，所以損失的機械能（mechanical energy）等於開始與結束之間重力勢能（gravitational potential energy）的減少量。在墜落期間，設升降機上升了 s_1（因此平衡錘下降了 s_1），並設瓷磚下降了 s_2。瓷磚走完了原來相距的 h，所以 $s_1 + s_2 = h$。勢能的變化為

$$(M - m)g\, s_1 - Mg\, s_1 - mg\, s_2 = -mg(s_1 + s_2) = -mgh.$$

碰撞之後再沒有任何東西移動，所以損失的能量是 mgh，正如題目所求。撞擊前一刻所有的動能（kinetic energy）都在碰撞中消失。留意這個論證完全沒有用到 t 的值。

驗算

- 若 $m \to 0$，則 $a \to 0$ 且 $t \to \sqrt{2h/g}$，即自由落體落在靜止的地板上。兩條公式都通過這個測試，所以它無法分辨兩者。
- 取 $M = 5$、$m = 1$、$h = 1$ 及 $g = 9.8$，以數值方法逐步模擬運動，得 $t \approx 0.4286\,\text{s}$，與 $\sqrt{9/49} \approx 0.4286$ 吻合。題目所印的公式則給出 $\sqrt{8/49} \approx 0.404\,\text{s}$。
- 在同一個模擬中，撞擊前一刻的動能等於 mgh，準確至五位有效數字，而碰撞後的速率為零，誤差只在捨入範圍之內。

常見錯誤

- 把升降機的質量當作 M 另加瓷磚：這樣系統一開始就不可能靜止。
- 忘記了平衡錘的慣性（inertia），以致算出升降機的加速度為 $mg/(M-m)$。
- 在碰撞中只對升降機和瓷磚使用線動量守恆（conservation of linear momentum），忽略了纜索中的衝擊張力。

STEP 之外

這裏所用的「沿纜索的動量」是一種廣義動量（generalised momentum）：對於只有一個自由度（degree of freedom）的系統，其坐標就是已繞過滑輪的纜索長度。拉格朗日力學（Lagrangian mechanics）把這套做法系統化。凡是力在其上沒有淨效果的坐標，都對應一個守恆的廣義動量，而衝擊約束（impulsive constraint）也可用同樣的方法處理。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學導論）仔細討論了以繩和滑輪連接的系統，以及碰撞中的衝量和能量。

詳細解答

STEP II 1989，第 12 題

力學 • 剛體平衡

> **STEP II 1989，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/89-s2-q12

關鍵思路

作用在薄片（lamina）上的力只有三個：它的重量和兩根釘子的反作用力（reaction）。光滑的釘子只會沿垂直於所接觸的邊的方向施力，而在長方形中，這就表示每個反作用力都平行於另一條邊。一個物體若在三個互不平行的力作用下保持平衡（equilibrium），這三個力的作用線（line of action）必須交於一點。因此，兩個反作用力的交點必定位於通過中心的鉛垂線上。兩個結果都可由這一事實，加上兩根釘子等高這個條件推導出來。

解答

建立模型

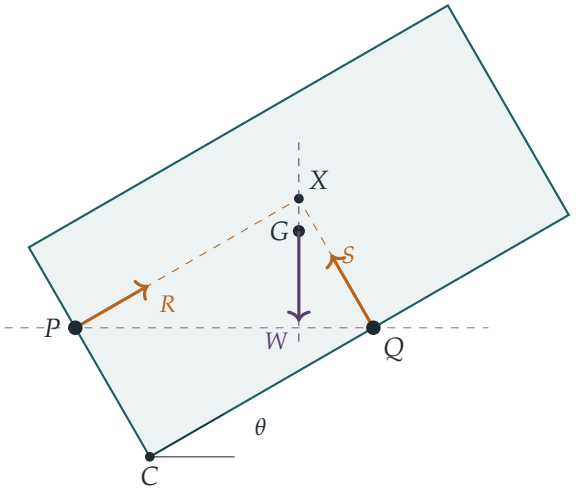
我們把薄片看成一個重量為 W、中心為 G 的均勻剛性長方形，並把兩根釘子看成位於同一高度的光滑點 P 和 Q。設 C 為最低的頂點。通過 C 而長度為 $2a$ 的邊擱在 P 上，其中 $CP = p$；通過 C 而長度為 $2b$ 的邊擱在 Q 上，其中 $CQ = q$。$2b$ 邊與水平線成銳角 θ，所以（與它垂直的）$2a$ 邊與鉛垂線成角 θ。

最方便的做法是沿兩條邊設立坐標：由 C 出發，沿 $2b$ 邊量度 u，沿 $2a$ 邊量度 v。於是

$$C = (0,0), \qquad P = (0,p), \qquad Q = (q,0), \qquad G = (b,a).$$

點 (u,v) 相對於 C 的水平位移和鉛垂位移為

$$x = u\cos\theta - v\sin\theta, \qquad y = u\sin\theta + v\cos\theta.$$



兩個反作用力交於通過 G 的鉛垂線上

P 處的反作用力 R 垂直於 $2a$ 邊，所以它平行於 $2b$ 邊：它的作用線是 $v = p$。同理，Q 處的反作用力 S 沿 $u = q$ 作用。這兩條直線交於 $X = (q, p)$，即長方形 $CPXQ$ 的第四個頂點。

對 X 取力矩（moment），兩個反作用力的力矩均為零，所以重量的力矩也必須為零：通過 G 的鉛垂線通過 X。換句話說，X 和 G 的水平坐標相同：

$$q\cos\theta - p\sin\theta = b\cos\theta - a\sin\theta, \qquad \text{即} \qquad (q-b)\cos\theta = (p-a)\sin\theta. \qquad (1)$$

（分解各力可得 $R = W\sin\theta$ 和 $S = W\cos\theta$，兩者皆為正數，所以釘子確實能夠提供這些力。）

第一個結果

由於 $0 < \theta < \frac{\pi}{2}$，$\cos\theta$ 和 $\sin\theta$ 都是正數，所以由 (1) 可知 $q-b$ 與 $p-a$ 同號。若 $p < a$，則 (1) 的右邊為負數，所以 $q - b < 0$，即 $q < b$。這正是所求的結果：若該頂點與 $2a$ 邊上的釘子相距少於 a，則它與另一根釘子相距少於 b。

? **想一想** 為甚麼問題還未完結？d 在哪裏出現？

方程 (1) 只用到力矩條件。它還未涉及 d，而且 p 和 q 並非互相獨立：兩根釘子等高，並且相距 d。現在我們把這兩個事實用上。

與 d 的關係

兩根釘子位於同一高度，所以 P 和 Q 的 y 坐標相等：

$$p\cos\theta = q\sin\theta, \qquad \text{所以} \qquad p = q\tan\theta. \tag{2}$$

它們的水平距離為 d：

$$d = x_Q - x_P = q\cos\theta + p\sin\theta = q\cos\theta + \frac{q\sin^2\theta}{\cos\theta} = \frac{q}{\cos\theta}. \tag{3}$$

（等價地說，三角形 CPQ 以 C 為直角，斜邊（hypotenuse）$PQ = d$ 是水平的，所以 $q = d\cos\theta$ 及 $p = d\sin\theta$。）

現在把 $q = d\cos\theta$ 和 $p = d\sin\theta$ 代入 (1) 的第一個形式的左邊：

$$\begin{aligned} b\cos\theta - a\sin\theta &= q\cos\theta - p\sin\theta \\ &= d\cos^2\theta - d\sin^2\theta \\ &= d\cos 2\theta, \end{aligned}$$

這就是所求的結果。

驗算

- 正方形的情況 $a = b$、$\theta = \frac{\pi}{4}$：兩邊都等於零。根據對稱性，對角線 CG 是鉛垂的，所以結果吻合；而且這對任何 d 都成立，正方形對稱地放置時理應如此。
- 取 $a = 1$、$b = 2$、$\theta = \frac{\pi}{6}$，平衡時需要 $q = d\cos\theta$ 及 $d = (b\cos\theta - a\sin\theta)/\cos 2\theta \approx 2.4641$。由此得 $p \approx 1.2321 > a$ 及 $q \approx 2.1340 > b$，與 (1) 吻合。以 $R = W\sin\theta$ 和 $S = W\cos\theta$ 對 G 取力矩，結果為零。
- 數值上，超過 10^5 個隨機的平衡位置（隨機選取 a、b、θ，而釘子落在邊上）全部滿足該恆等式，誤差在 10^{-14} 以內，而且從不違反「$p < a \Rightarrow q < b$」。

常見錯誤

- 把釘子的反作用力當作鉛垂。光滑接觸的施力方向垂直於*接觸面*（這裏是薄片的邊），而不是垂直於地面。
- 混淆了哪一條邊與水平線成角 θ：$2a$ 邊是與*鉛垂線*成角 θ。

STEP 之外

三力原理是剛體（rigid body）靜力學一般理論的特殊情況：一個力系處於平衡，當且僅當它的合力以及它對任意一點的合力矩同時為零。到了大學，這會發展成力螺旋（wrench）和虛功（virtual work）的語言。若用虛功的方法，可以在薄片與兩根釘子保持接觸的約束下，把 G 的高度減至最小，從而求出這個平衡位置。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）對剛體靜力學和力矩（torque）有嚴謹的論述。

詳細解答

STEP II 1989，第 13 題

力學 • 轉動慣量

STEP II 1989，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/89-s2-q13

本題在 *Beyond STEP* 第 11.7 節桿中的辛普森法則中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch11/simpsons-rod.html

關鍵思路

一切都建基於轉動慣量（moment of inertia）的兩個事實。平行軸定理（parallel axis theorem）指出，對任何一條軸的轉動慣量，都由質量、質心（centre of mass）以及對通過質心的平行軸的轉動慣量所決定。此外，轉動慣量具有可加性（additive），所以我們可以把一個物體切成若干小塊，再把每一塊換成等效的東西。薄片（lamina）就是一疊幼桿，所以只要知道怎樣替換一根桿，就能替換整塊薄片。

解答

對每一條軸均等效

設 ℓ 為任意一條軸，並設 ℓ_O 為通過 O 的平行軸，與 ℓ 相距 d。物體與質點（particle）系統的質量同為 m，質心同為 O，所以可以對兩者分別應用平行軸定理：

$$I_\ell^{\text{物體}} = I_{\ell_O}^{\text{物體}} + md^2, \qquad I_\ell^{\text{質點}} = I_{\ell_O}^{\text{質點}} + md^2.$$

由於 ℓ_O 通過 O，根據假設 $I_{\ell_O}^{\text{物體}} = I_{\ell_O}^{\text{質點}}$。因此兩式的右邊相等，於是對每一條軸 ℓ 都有 $I_\ell^{\text{物體}} = I_\ell^{\text{質點}}$，這正是所要證明的。

均勻桿

把均勻桿（uniform rod）沿 x 軸放置，由 $x=-a$ 伸展至 $x=a$，於是它的質心 O 就是原點。在 $x=-a$、0、a 處分別放置質量為 p、q、r 的質點。要令總質量為 m 且質心位於 O，我們需要

$$p+q+r=m, \qquad -pa+ra=0,$$

所以 $r=p$ 及 $q=m-2p$。

現在取任意一條通過 O、與桿成角 θ 的軸。桿上位於 x 的一點與這條軸的距離為 $|x|\sin\theta$。因此，對任何一條通過 O 的軸，轉動慣量都是 $\sin^2\theta$ 乘以 $\sum(\text{質量})\,x^2$，我們只需令這一個和相等便可。對於桿，每個微元的質量為 $\frac{m}{2a}\,\mathrm{d}x$，所以

$$\int_{-a}^{a}\frac{m}{2a}\,x^2\,\mathrm{d}x=\frac{m}{2a}\cdot\frac{2a^3}{3}=\tfrac{1}{3}ma^2,$$

而對於質點，這個和是 $pa^2+q\cdot 0+pa^2=2pa^2$。當 $p=\frac{1}{6}m$ 時，兩者對每個 θ 都相等。所以這根桿動力等效（dynamically equivalent）於

$$\tfrac{1}{6}m \text{ 於每一端}, \qquad \tfrac{2}{3}m \text{ 於中點}.$$

根據第一部分，兩者對任何一條軸的轉動慣量都相同，而不限於通過 O 的軸。

長方形薄片

設長方形的中心為 O，邊長為 $2a$（平行於 x 軸）及 $2b$，於是它每單位面積的質量為 $M/(4ab)$。把它切成平行於 x 軸的幼條。位於高度 y、闊度為 δy 的幼條是一根長度為 $2a$、質量為 $\delta M=\frac{M}{2b}\,\delta y$ 的均勻桿。

> **想一想** 幼條的中心並不在 O。為甚麼我們仍然可以把每條幼條換成它的三個質點？

第一部分的用處正在於此。每條幼條與它的三個質點對每一條軸的轉動慣量都相同，而不只是對通過幼條本身中心的軸。轉動慣量可以相加，所以對任何一條軸，薄片的轉動慣量都與全部替換質點所組成的系統相同。這個系統的總質量同樣是 M，質心同樣是 O，因為每條幼條的替換質點與該幼條有相同的質量和質心。

位於高度 y 的幼條換成：在 $(\pm a,y)$ 兩點各放 $\frac{1}{6}\delta M=\frac{M}{12b}\,\delta y$，以及在 $(0,y)$ 放 $\frac{2}{3}\delta M=\frac{M}{3b}\,\delta y$。當 y 由 $-b$ 變到 b 時，這些質點填滿三根長度為 $2b$ 的均勻桿：

- 沿兩邊 $x=\pm a$ 各有一根，每根的質量為 $\frac{M}{12b}\cdot 2b=\frac{1}{6}M$；
- 沿中線 $x=0$ 有一根，質量為 $\frac{M}{3b}\cdot 2b=\frac{2}{3}M$。

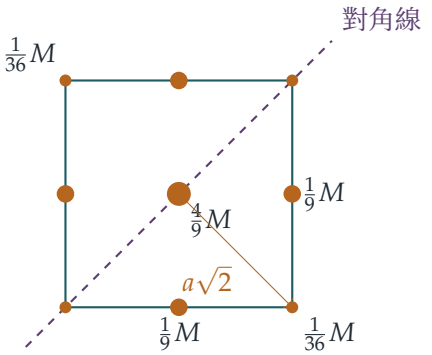
現在以同樣的可加性論證，對這三根桿逐一應用均勻桿的結果：

- 每根邊桿（$\frac{1}{6}M$）變成：在兩端（即長方形的角）各有 $\frac{1}{6}\cdot\frac{1}{6}M=\frac{1}{36}M$，以及在中點（即長度為 $2b$ 的邊的中點）有 $\frac{2}{3}\cdot\frac{1}{6}M=\frac{1}{9}M$；
- 中央的桿（$\frac{2}{3}M$）變成：在兩端（即長度為 $2a$ 的邊的中點）各有 $\frac{1}{6}\cdot\frac{2}{3}M=\frac{1}{9}M$，以及在中心有 $\frac{2}{3}\cdot\frac{2}{3}M=\frac{4}{9}M$。

所以每個角有 $\frac{1}{36}M$，每條邊的中點有 $\frac{1}{9}M$，中心有 $\frac{4}{9}M$。總和為 $\frac{4}{36}M+\frac{4}{9}M+\frac{4}{9}M=M$，而由對稱性可知質心為 O。這就是所求的等效系統。

正方形對其對角線

對於邊長為 $2a$ 的正方形，把四個角放在 $(\pm a,\pm a)$，並取對角線 $y=x$。點 (x_0,y_0) 與這條直線的距離為 $|x_0-y_0|/\sqrt{2}$。



- 角 (a,a)、$(-a,-a)$ 和中心都在對角線上，沒有貢獻。
- 角 $(a,-a)$、$(-a,a)$ 與對角線的距離為 $2a/\sqrt{2}=a\sqrt{2}$。
- 每個中點 $(\pm a,0)$、$(0,\pm a)$ 與對角線的距離為 $a/\sqrt{2}$。

因此

$$I_{\text{對角線}}=2\cdot\tfrac{1}{36}M\cdot 2a^2+4\cdot\tfrac{1}{9}M\cdot\tfrac{1}{2}a^2=\tfrac{1}{9}Ma^2+\tfrac{2}{9}Ma^2=\tfrac{1}{3}Ma^2.$$

非均勻桿

設桿的長度為 $2a$，沿 x 軸由 $x=0$ 伸展至 $x=2a$，線密度（line density）為 kx。那麼

$$m=\int_0^{2a}kx\,\mathrm{d}x=2ka^2,\qquad\text{所以}\qquad k=\frac{m}{2a^2}.$$

它的質心位於

$$\bar{x}=\frac{1}{m}\int_0^{2a}kx^2\,\mathrm{d}x=\frac{k}{m}\cdot\frac{8a^3}{3}=\frac{4a}{3},$$

並且

$$\int_0^{2a} kx^3\,\mathrm{d}x = \frac{k\,(2a)^4}{4} = 4ka^4 = 2ma^2.$$

與均勻桿的情況一樣，所有點都在同一條直線上，所以對一條通過質心、與桿成角 θ 的軸，轉動慣量為 $\sin^2\theta\sum(\text{質量})(x-\bar{x})^2$。若質量與質心都相同，則 $\sum(\text{質量})(x-\bar{x})^2 = \sum(\text{質量})x^2 - m\bar{x}^2$，所以只需令 $\sum(\text{質量})\,x^2$ 相等便已足夠，而這正是對通過端點 $x=0$ 的垂直軸的轉動慣量。

設質點為：$x=0$ 處的 p_0、$x=a$ 處的 p_1 及 $x=2a$ 處的 p_2。令質量、一階矩（first moment）和二階矩（second moment）分別相等，得出方程

$$\begin{aligned} p_0 + p_1 + p_2 &= m, \\ p_1 a + 2p_2 a &= \tfrac{4}{3}ma, \\ p_1 a^2 + 4p_2 a^2 &= 2ma^2. \end{aligned}$$

把第三式（除以 a^2）減去第二式（除以 a），得 $2p_2 = \frac{2}{3}m$，所以 $p_2 = \frac{1}{3}m$。於是 $p_1 = \frac{4}{3}m - \frac{2}{3}m = \frac{2}{3}m$，最後 $p_0 = m - \frac{2}{3}m - \frac{1}{3}m = 0$。

位於較輕一端的質點質量為零，所以只需要兩個質點：中點的 $\frac{2}{3}m$ 和較重一端的 $\frac{1}{3}m$。

驗算

- 邊長為 $2a$ 的正方形薄片，對一條通過中心、位於薄片平面內且平行於一邊的軸，其轉動慣量的標準結果是 $\frac{1}{3}Ma^2$。根據垂直軸定理（perpendicular axis theorem）和對稱性，每一條通過中心而位於平面內的軸都給出相同的值，所以對角線也必然如此。結果吻合。
- 對於非均勻桿，兩個質點給出質量 m、一階矩 $\frac{2}{3}m\cdot a + \frac{1}{3}m\cdot 2a = \frac{4}{3}ma$（所以質心位於 $\frac{4}{3}a$）以及二階矩 $\frac{2}{3}ma^2 + \frac{4}{3}ma^2 = 2ma^2$，全部符合要求。
- 在數值上，我們用細密的點質量網格，把每個物體與它的質點系統對空間中 100–200 條隨機的軸（不通過中心）作比較。均勻桿和非均勻桿的誤差小於 10^{-7}，長方形的誤差為 4×10^{-5}（網格誤差），而正方形對其對角線在 $M=a=1$ 時給出 0.33332。

常見錯誤

- 把每條幼條換成它的三個質點，卻沒有說明為甚麼可以這樣做。幼條的中心並不在 O，所以你需要第一部分關於「任何軸」的結果。
- 在最後一部分，只施加質量和質心兩個條件。兩條方程不足以確定三個質量：正是轉動慣量的條件迫使 $p_0 = 0$。

STEP 之外

對於在空間中運動的剛體（rigid body），在動力學上起決定作用的是它的質量、質心和慣性張量（inertia tensor）。慣性張量是一個對稱的 3×3 矩陣，可給出對每一條通過質心的軸的轉動慣量。兩個物體若這三者都相同，在相同的力作用下運動便完全一樣，這就是「動力等效」這個名稱貼切的原因。用幾個點質量去配合一個分佈的頭幾階矩，也是數值積分（numerical integration）背後的想法。均勻桿的權重 $\frac{1}{6}, \frac{2}{3}, \frac{1}{6}$ 正好就是辛普森法則（Simpson's rule），而若除了權重之外連位置也一併選取，便會引出高斯求積法（Gaussian quadrature）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics* 詳細講述了慣性張量和剛體運動。

詳細解答

STEP II 1989，第 14 題

力學 • 圓周運動

STEP II 1989，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：

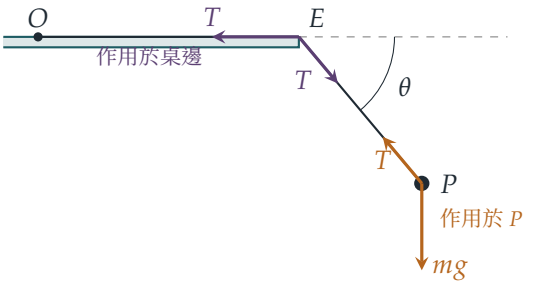
step.maths.org/questions/89-s2-q14

關鍵思路

質點一經釋放，繩子留在桌面上的那一段便保持不動：它在固定端與桌邊之間被拉緊。因此桌邊相當於一個固定的支點，質點繞着它沿半徑為 a 的圓運動，情況與由水平位置釋放的單擺（simple pendulum）完全相同。由能量可求得速率，由徑向運動方程（radial equation of motion）可求得張力（tension）。桌邊是光滑的，所以繩子在桌邊兩側的張力相同，而作用於桌邊的力，就是沿兩段繩子拉扯的兩個相等張力的向量和。

解答

模型。 繩子是輕而不可伸長的（light and inextensible），桌面光滑而且很薄，質點是一個點質量（point mass）。設 O 為繩子的固定端，E 為繩子跨過桌邊的那一點，於是 $OE = l - a$，而懸垂部分 EP 的長度為 a。設 θ 為懸垂部分從水平方向（即它起初所指、背離桌子的方向）向下轉過的角。



由於 O 固定，而繩子一直保持拉緊（下文會驗證 $T > 0$），OE 這一段留在桌面上靜止不動。因此 P 沿以 E 為圓心、半徑為 a 的圓運動。

速率。 桌面光滑，而張力始終與 P 的運動方向垂直，所以兩者都不作功，機械能守恆（conservation of mechanical energy）。當繩子轉過角 θ 時，P 下降了高度 $a\sin\theta$，故由靜止開始，

$$\tfrac{1}{2}mv^2 = mga\sin\theta \qquad \Longrightarrow \qquad v^2 = 2ga\sin\theta.$$

張力。 沿 PE 向圓心 E 的方向分解：張力沿 PE 作用，而重量沿 PE 的分量為 $mg\sin\theta$，方向背離 E（當 $0<\theta<\pi$ 時，P 位於 E 的水平以下）。所以

$$T - mg\sin\theta = \frac{mv^2}{a} = 2mg\sin\theta, \qquad \text{由此得} \qquad \boxed{T = 3mg\sin\theta.}$$

當 $0<\theta<\pi$ 時這個值為正，所以繩子自始至終保持拉緊。在 $\theta=\pi$ 時有 $v=0$：質點剛好到達桌子底面、與 E 同高的位置，然後往回擺。因此運動發生在 $0\le\theta\le\pi$ 的範圍內。

作用於桌邊的力。 桌邊光滑，所以繩子在桌邊兩側的張力都是 T。繩子以兩個大小為 T 的拉力壓在桌邊上：一個沿 EO，水平指向固定端；另一個沿 EP。取 x 軸為水平背離桌子的方向，y 軸為鉛垂向上。沿 EO 和 EP 的單位向量（unit vector）分別是 $(-1,0)$ 和 $(\cos\theta,-\sin\theta)$，所以作用於桌邊的力為

$$\mathbf{F} = T\,(\cos\theta - 1,\ -\sin\theta).$$

它的大小為

$$|\mathbf{F}| = T\sqrt{(1-\cos\theta)^2+\sin^2\theta} = T\sqrt{2-2\cos\theta} = 2T\sin\tfrac{\theta}{2}$$

這裏用了 $1-\cos\theta = 2\sin^2\frac{\theta}{2}$，以及當 $0\le\theta\le\pi$ 時 $\sin\frac{\theta}{2}\ge 0$。（從幾何上看，兩個夾角為 $\pi-\theta$ 的相等的力 T，合力為 $2T\cos\frac{\pi-\theta}{2}$，結果相同。）代入 T 和 $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$，

$$|\mathbf{F}| = 6mg\sin\theta\sin\tfrac{\theta}{2} = 12mg\,s^2\sqrt{1-s^2}, \qquad \text{其中} s=\sin\tfrac{\theta}{2}\in[0,1].$$

想一想 怎樣才能乾淨俐落地求 $s^2\sqrt{1-s^2}$ 的最大值？

由於它是非負的，只需求它的平方的最大值，即 $h(u) = u^2(1-u)$，其中 $u = s^2 \in [0,1]$。於是 $h'(u) = 2u-3u^2 = u(2-3u)$，當 $0<u<\frac{2}{3}$ 時為正，當 $\frac{2}{3}<u\le 1$ 時為負。所以 h 在 $[0,1]$ 上的最大值出現在 $u=\frac{2}{3}$，此時 $h = \frac{4}{9}\cdot\frac{1}{3} = \frac{4}{27}$。因此 $s^2\sqrt{1-s^2}$ 的最大值為 $\sqrt{4/27} = \dfrac{2}{3\sqrt{3}}$，而

$$|\mathbf{F}|_{\max} = 12mg\cdot\frac{2}{3\sqrt{3}} = \frac{8mg}{\sqrt{3}},$$

正如所求。它出現在 $\sin^2\frac{\theta}{2}=\frac{2}{3}$ 時，即 $\cos\theta = 1-2\sin^2\frac{\theta}{2} = -\frac{1}{3}$，所以 $\theta\approx 109.5^\circ$：這時質點已經越過桌邊的正下方，正往桌子底下擺回去。

驗算

- 在 $\theta = \frac{\pi}{2}$（繩子鉛垂）時，速率為 $\sqrt{2ga}$，而 $T = 3mg$，這正是由水平位置釋放的單擺的熟悉結果。這時作用於桌邊的力為 $T\sqrt{2} = 3\sqrt{2}\,mg \approx 4.24mg$，略小於 $8mg/\sqrt{3} \approx 4.62mg$，與最大值出現在鉛垂位置之後這一點吻合。
- 在 $\theta = 0$ 和 $\theta = \pi$ 時，質點在與桌邊同高的位置瞬時靜止，所以 $T = 0$，桌邊不受任何力，與公式所示一致。
- 我們對運動方程 $\ddot{\theta} = (g/a)\cos\theta$ 作了數值積分：由徑向方程求得的 T 與 $3mg\sin\theta$ 相符，誤差在 10^{-12} 以內；而作用於桌邊的力的最大值出現在 $\theta \approx 1.9106$（即 $\arccos(-\frac{1}{3})$），其值為 $4.6188\,mg$，與 $8/\sqrt{3}$ 相符。

常見錯誤

- 把作用於桌邊的力當作只有 T，或只取懸垂繩子拉力的鉛垂部分。桌面上的那段繩子也會水平地拉扯桌邊，兩個拉力必須以向量方式相加。
- 求最大值時只考慮到 $\theta \le \frac{\pi}{2}$ 為止。質點會繼續擺到桌子底下，而最大值正是出現在這個較後的階段。

STEP 之外

這裏的張力是一種*約束力*（constraint force）：它不作功，其大小恰好足以令質點保持在圓周上。在拉格朗日力學（Lagrangian mechanics）中，這類力會從運動方程中完全消失，事後才再求回來（就像我們利用徑向方程所做的那樣），或者藉拉格朗日乘數（Lagrange multipliers）求得。Kleppner 與 Kolenkow 合著的 *An Introduction to Mechanics*（力學入門）在這個程度上對圓周運動、約束和能量方法有細緻的論述。

詳細解答

STEP II 1989，第 15 題

概率與統計 • 幾何概率

STEP II 1989，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/89-s2-q15

關鍵思路

「在周界上隨機」選取的一點，是按長度均勻分佈（uniform distribution）的，所以一切都由總長度 $\pi+2$ 出發。三角形總有一個頂點在圓心 O，所以它的面積是 $\frac{1}{2}ab\sin C$，其中 a 和 b 是兩點與 O 的距離，C 是在 O 處的夾角。這樣，每個條件情況都化為一個簡單的期望值（expectation），再用全期望定律（law of total expectation）把它們合併起來。

解答

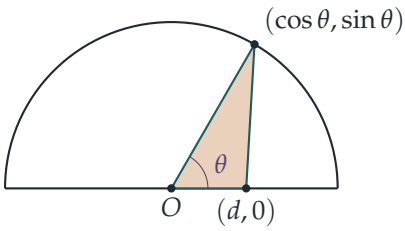
把 O 放在原點，直徑沿 x 軸由 $(-1,0)$ 到 $(1,0)$，圓弧位於 $y\geq 0$。周界的長度是 $2+\pi$：直徑佔 2，圓弧佔 π。由於每一點都沿周界均勻分佈，

$$\mathrm{P}(\text{一點落在直徑上})=\frac{2}{\pi+2},\qquad \mathrm{P}(\text{一點落在圓弧上})=\frac{\pi}{\pi+2}.$$

已知該點落在圓弧上時，它是 $(\cos\theta,\sin\theta)$，其中 θ 在 $[0,\pi]$ 上均勻分佈。已知該點落在直徑上時，它是 $(d,0)$，其中 d 在 $[-1,1]$ 上均勻分佈。

兩點互相獨立（independent），而「恰好一點在直徑上」可以是第一點在直徑上，也可以是第二點在直徑上，所以

$$\mathrm{P}(\text{恰好一點在直徑上})=2\cdot\frac{2}{\pi+2}\cdot\frac{\pi}{\pi+2}=\frac{4\pi}{(\pi+2)^2}.$$



第 (i) 部分

設其中一點是直徑上的 $(d,0)$，另一點是圓弧上的 $(\cos\theta,\sin\theta)$。由 O 出發的兩條邊長度分別是 $|d|$ 和 1，它們之間的夾角是 θ（若 $d<0$，夾角是 $\pi-\theta$，其正弦值相同）。所以

$$A=\tfrac{1}{2}\,|d|\sin\theta.$$

在這個情況下，d 和 θ 互相獨立，所以乘積的期望值等於期望值的乘積。由於 d 在 $[-1,1]$ 上均勻分佈，$|d|$ 在 $[0,1]$ 上均勻分佈，故 $\mathrm{E}|d|=\frac{1}{2}$。另外，

$$\mathrm{E}(\sin\theta)=\frac{1}{\pi}\int_0^{\pi}\sin\theta\,\mathrm{d}\theta=\frac{2}{\pi}.$$

因此

$$\mathrm{E}(A\mid\text{恰好一點在直徑上})=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{2}{\pi}=\frac{1}{2\pi},$$

正是所求。

第 (ii) 部分

現在兩點都在圓弧上，角度分別是 θ_1 和 θ_2，兩者互相獨立，並在 $[0,\pi]$ 上均勻分佈。兩點與 O 的距離都是 1，而它們在 O 處的夾角是 $X=|\theta_1-\theta_2|$，位於 $[0,\pi]$ 之內。所以

$$A=\tfrac{1}{2}\sin X.$$

? **想一想** 為甚麼我們可以使用題目給出的 X 的密度函數？

角 θ 量度的是沿單位半圓的弧長，所以在圓弧上均勻地選取兩點，等同於在一條長度為 π 的線段上均勻地選取兩點。角 X 正是兩點之間的距離，所以 X 在 $[0,\pi]$ 上的密度函數（density）是 $2(\pi-x)/\pi^2$。於是

$$\mathrm{E}(\sin X)=\frac{2}{\pi^2}\int_0^{\pi}(\pi-x)\sin x\,\mathrm{d}x.$$

把積分拆開，$\int_0^{\pi}\pi\sin x\,\mathrm{d}x=2\pi$；再取 $u=x$ 作分部積分（integration by parts），

$$\int_0^{\pi}x\sin x\,\mathrm{d}x=\Big[-x\cos x\Big]_0^{\pi}+\int_0^{\pi}\cos x\,\mathrm{d}x=\pi+0=\pi.$$

所以 $\int_0^\pi (\pi - x)\sin x\,\mathrm{d}x = 2\pi - \pi = \pi$，由此得 $\mathrm{E}(\sin X) = \dfrac{2}{\pi}$，以及

$$\mathrm{E}(A \mid \text{兩點都不在直徑上}) = \frac{1}{2}\cdot\frac{2}{\pi} = \frac{1}{\pi},$$

正是所求。

A 的期望值

還有第三種情況：若兩點都落在直徑上，這個「三角形」是退化（degenerate）的（三點都在 x 軸上），所以 $A = 0$。兩點都不在直徑上的概率是 $\left(\frac{\pi}{\pi+2}\right)^2$。根據全期望定律，

$$\begin{aligned}\mathrm{E}(A) &= \frac{4\pi}{(\pi+2)^2}\cdot\frac{1}{2\pi} + \frac{\pi^2}{(\pi+2)^2}\cdot\frac{1}{\pi} + \frac{4}{(\pi+2)^2}\cdot 0\\ &= \frac{2+\pi}{(\pi+2)^2}\\ &= \frac{1}{2+\pi},\end{aligned}$$

這就是我們要證明的結果。

驗算

- 三種情況的概率 $\frac{4}{(\pi+2)^2}$、$\frac{4\pi}{(\pi+2)^2}$ 和 $\frac{\pi^2}{(\pi+2)^2}$ 相加得 $\frac{(2+\pi)^2}{(\pi+2)^2} = 1$。
- $\mathrm{E}(A) \le \frac{1}{2}$，即不超過最大可能的面積（兩條半徑互相垂直），這是必然的；事實上 $\frac{1}{2+\pi} \approx 0.1945$。
- 以蒙地卡羅模擬（Monte Carlo simulation）抽取 2×10^6 對點，得出 P(恰好一點) $\approx$ 0.4757（準確值為 0.4754），條件平均值為 0.1591 和 0.3184（準確值為 $\frac{1}{2\pi} \approx 0.1592$ 和 $\frac{1}{\pi} \approx 0.3183$），以及 $\mathrm{E}(A) \approx 0.1944$。

常見錯誤

- 在 P(恰好一點) 中漏掉因子 2：兩點之中任何一點都可以是落在直徑上的那一點。
- 把「在直徑上」的概率當作 $\frac{1}{2}$：選點是按長度均勻分佈，而不是按落在邊界的哪一段來平均分配。
- 在第 (ii) 部分使用 $\sin\theta_1\sin\theta_2$ 或類似的式子，而不是兩條半徑之間夾角的正弦。

STEP 之外

把期望值按各種情況的劃分（partition）拆開，即 $\mathrm{E}(A) = \sum_i \mathrm{E}(A \mid B_i)\,\mathrm{P}(B_i)$，就是全期望定律，它是大學一年級概率論中最常用的工具之一；其一般形式 $\mathrm{E}(A) = \mathrm{E}\big(\mathrm{E}(A \mid Y)\big)$ 更是無處不在，由分支過程（branching processes）到馬可夫鏈（Markov chains）都會用到。第 (ii) 部分引用的 X 的密度函數，本身就是一道不錯的練習：它來自正方形 $[0, \pi]^2$ 內一條帶狀區域的面積。這類題目屬於幾何概率（geometric probability）；Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）以合適的程度講解條件期望（conditional expectation）和聯合密度函數（joint densities）。

詳細解答

STEP II 1989，第 16 題

概率與統計 • 二項分佈

> **STEP II 1989，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/89-s2-q16

關鍵思路

把整批產品分成表現不同的兩部分：樣本中的 n 件小配件，以及從未被抽查的 N 件。每一部分的次品數目都服從二項分佈（binomial distribution），而且兩個數目互相獨立（independent）。一切取決於樣本：如果樣本中有兩件或以上次品，整批最終會完美無瑕，否則一切維持不變。因此，若一批產品最後有 $k \geq 1$ 件次品，樣本中必定只有 0 或 1 件次品，其餘的次品則在另外 N 件之中。

解答

記 $q = 1 - p$。設 X 為樣本中次品的數目，Y 為該批其餘 N 件中次品的數目。每件小配件各自獨立地以概率（probability）p 為次品，所以

$$X \sim \mathrm{B}(n, p), \qquad Y \sim \mathrm{B}(N, p),$$

而且 X 與 Y 獨立，因為它們所數算的是兩組互不相交的小配件中的次品。

設 K 為經過品質管制後該批產品中次品的數目。若 $X \geq 2$，所有次品都會被修正，於是 $K = 0$。若 $X \leq 1$，則甚麼都不會修正，於是 $K = X + Y$。所以

$$K = \begin{cases} X + Y, & X \leq 1, \\ 0, & X \geq 2. \end{cases}$$

K 的分佈

取 $1 \le k \le N$。由於 $k \ge 1$，事件 $K = k$ 不可能來自 $X \ge 2$ 的情況，所以它恰好在 $X = 0,\ Y = k$ 或 $X = 1,\ Y = k-1$ 時發生。這兩個事件互斥，而由獨立性可得

$$\begin{aligned}\mathrm{P}(K=k) &= \mathrm{P}(X=0)\,\mathrm{P}(Y=k) + \mathrm{P}(X=1)\,\mathrm{P}(Y=k-1)\\ &= q^n\binom{N}{k}p^k q^{N-k} + npq^{n-1}\binom{N}{k-1}p^{k-1}q^{N-k+1}\\ &= \left[\binom{N}{k} + n\binom{N}{k-1}\right]p^k q^{N+n-k}.\end{aligned}$$

兩項中 p 的冪次相同，q 的冪次也相同，所以才能合併。現在把二項式係數（binomial coefficient）通分，公分母為 $k!\,(N-k+1)!$：

$$\binom{N}{k} = \frac{N!\,(N-k+1)}{k!\,(N-k+1)!}, \qquad n\binom{N}{k-1} = \frac{n\,N!\,k}{k!\,(N-k+1)!}.$$

相加後，分子為 $N!\,(N-k+1+nk) = N!\,[N+1+k(n-1)]$，所以

$$\mathrm{P}(K=k) = \frac{[N+1+k(n-1)]\,N!}{(N-k+1)!\,k!}\,p^k(1-p)^{N+n-k},$$

正如所求。

> **想一想** 當 $k = 0$ 時會怎樣？

這條公式只在 $k \ge 1$ 時成立。當 $k = 0$ 時，它給出 $\frac{(N+1)N!}{(N+1)!}q^{N+n} = q^{N+n}$，這是完全沒有次品的概率。但只要 $X \ge 2$，$K = 0$ 同樣會發生，所以實際上

$$\mathrm{P}(K=0) = q^{N+n} + \mathrm{P}(X \ge 2) = q^{N+n} + 1 - q^n - npq^{n-1}.$$

題目中的「$k \le N$」應理解為 $1 \le k \le N$。

$k = N+1$ 的情況

要使 $K = N+1$，需要 $X \le 1$ 且 $X + Y = N+1$。由於 $Y \le N$，這就迫使 $X = 1$ 且 $Y = N$，所以

$$\mathrm{P}(K=N+1) = npq^{n-1}\cdot p^N = np^{N+1}q^{n-1}.$$

把 $k = N+1$ 代入公式，得到係數

$$\frac{[N+1+(N+1)(n-1)]\,N!}{0!\,(N+1)!} = \frac{(N+1)\,n\,N!}{(N+1)!} = n,$$

而 q 的冪為 $q^{N+n-(N+1)} = q^{n-1}$。所以公式給出 $np^{N+1}q^{n-1}$，結果正確。（當 $k \ge N+2$ 時概率為零，因為 $K \le N+1$。）

次品數目的期望值

我們可以對 $k\,\mathrm{P}(K=k)$ 求和，但利用以 X 和 Y 表示 K 的寫法會快得多。以 $\mathbf{1}_{\{X\le 1\}}$ 表示事件 $X\le 1$ 的指示變量（indicator），則

$$K=(X+Y)\,\mathbf{1}_{\{X\le 1\}}=X\,\mathbf{1}_{\{X\le 1\}}+Y\,\mathbf{1}_{\{X\le 1\}}.$$

對於第一項，$X\,\mathbf{1}_{\{X\le 1\}}$ 在 $X=1$ 時等於 1，其餘情況等於 0，所以它的期望值（expectation）是 $\mathrm{P}(X=1)=npq^{n-1}$。對於第二項，$Y$ 與 X 獨立，所以

$$\mathrm{E}\big[Y\,\mathbf{1}_{\{X\le 1\}}\big]=\mathrm{E}[Y]\,\mathrm{P}(X\le 1)=Np\left(q^n+npq^{n-1}\right).$$

相加，並提出因子 pq^{n-1}，

$$\begin{aligned}\mathrm{E}[K]&=npq^{n-1}+Np\,q^{n-1}(q+np)\\&=pq^{n-1}\left[n+N(1-p)+Nnp\right]\\&=\left[N+n+pN(n-1)\right]p(1-p)^{n-1},\end{aligned}$$

正如所求。

驗算

- 各概率之和應為 1。取 $n=2$、$N=1$：$\mathrm{P}(K=1)=(2+1)pq^2=3pq^2$，$\mathrm{P}(K=2)=2p^2q$，以及 $\mathrm{P}(K=0)=q^3+p^2$，三者相加得 $q^3+3pq^2+3p^2q+p^3=1$（用到 $p^2=p^2(p+q)$）。
- 若 n 極大，樣本中有次品的批次幾乎全部都會被修正，期望值趨向 0，這與因子 $(1-p)^{n-1}$ 吻合。
- 我們以精確分數驗算了每一條公式，範圍是 $2\le n\le 6$、$0\le N\le 6$ 以及若干個 p 值，並且模擬了整個程序：取 $n=4$、$N=6$、$p=0.3$ 時，公式給出 $\mathrm{E}[K]\approx 1.585$，而模擬 200 000 批所得的結果是 1.581。

常見錯誤

- 把「兩件或以上」的規則套用在整批的次品總數上。次品會否被修正，只由樣本中的數目 X 決定。
- 忘記 $K=0$ 還有來自已修正批次的額外貢獻，所以公式在 $k=0$ 時不成立。
- 以硬算求和的方法計算 $\mathrm{E}[K]$。這樣做是可行的，但用指示變量把 K 拆開，再利用獨立性，會簡短得多。

STEP 之外

把一個隨機變量（random variable）寫成若干部分各自乘以指示變量之和，再運用期望值的線性性質（linearity of expectation），是大學一年級概率論中最有用的技巧之一。當只需要平均值時，這個方法可以免卻求出整個分佈的工夫。這道題目也是品質管制中驗收抽樣（acceptance sampling）的一個小型模型，當中我們要選擇 n，在檢測成本與漏網的次品之間取得平衡。想進一步了解，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門），或 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論導論），後者全書貫徹使用指示變量。

詳細解答

STEP III 1989，第 1 題

純數學 • 旋轉曲面

STEP III 1989，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/89-s3-q1

關鍵思路

球面是半圓的旋轉曲面（surface of revolution），所以球帶（zone）的面積是一個標準的旋轉曲面積分。出人意表的是，被積函數（integrand）原來是常數，這正是答案只取決於兩平面之間距離的原因。在第二部分，曲面分成一個圓錐（由 OP 生成）和一個球帶（由弧 PR 生成）。這時第一個結果便擔起了最吃力的工作，而相切的條件則提供了我們所需的直角三角形。

關於題目印出的方程，要說明一下：若右邊是 c^2，圓便會通過 O，而常數 a 則從未出現。題目原意的曲線顯然是

$$(x-c)^2+y^2=a^2,$$

即圓心為 $C=(c,0)$、半徑為 a 的圓。由於 $c>a$，原點在圓外，所以由 O 出發的切線（tangent）存在。我們就這條曲線來解題。

解答

球帶的面積

取一個半徑為 r、球心在原點的球面，並選取 x 軸與兩個平面垂直，這兩個平面是 $x=h_1$ 和 $x=h_2$，其中 $-r\le h_1<h_2\le r$。球面由上半圓 $y=\sqrt{r^2-x^2}$ 繞 x 軸旋轉 2π 而成，所以球帶的面積是

$$S=\int_{h_1}^{h_2}2\pi y\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x.$$

這裏 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -\dfrac{x}{\sqrt{r^2-x^2}}$，所以當 $|x| < r$ 時

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

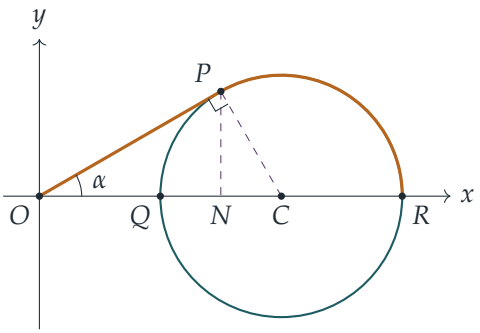
? **想一想** 在 $x = \pm r$ 處，$\mathrm{d}y/\mathrm{d}x$ 是無窮大，這時會怎樣？

導數（derivative）在兩極處趨向無窮大，但乘積 $y\sqrt{1+(y')^2}$ 在整個 $(-r, r)$ 上都等於 r，所以它是有界的。若其中一個平面與球面相切（$h_1 = -r$ 或 $h_2 = r$），這個積分雖屬反常積分（improper integral），卻無傷大雅：先積分至 $h_2 - \varepsilon$，再令 $\varepsilon \to 0$ 即可。（另一個做法是以 $x = r\cos\theta$、$y = r\sin\theta$ 作參數化（parametrise），這樣便完全避開了這個問題。）因此

$$S = \int_{h_1}^{h_2} 2\pi r\,\mathrm{d}x = 2\pi r\,(h_2 - h_1),$$

也就是 $2\pi \times$ (半徑) $\times$ (兩平面之間的距離)，正是所求。

曲面



設 $C = (c, 0)$ 為圓心，N 為由 P 至 x 軸的垂足。那麼 $Q = (c-a, 0)$，$R = (c+a, 0)$。弧 PR 是指不經過 Q 的那一段（圖中以琥珀色標示）：由 P 到 R 的另一段弧會經過 Q 並繞過圓的底部，把它旋轉便會掃出整個球面。

切線與過切點的半徑互相垂直，所以角 OPC 是直角。在三角形 OPC 中，$OC = c$ 而 $CP = a$，所以由畢氏定理（Pythagoras' theorem）得

$$OP = \sqrt{c^2-a^2}.$$

設 α 為角 POC。那麼 $\sin\alpha = \frac{a}{c}$ 而 $\cos\alpha = \frac{\sqrt{c^2-a^2}}{c}$，由此得

$$PN = OP\sin\alpha = \frac{a\sqrt{c^2-a^2}}{c}, \qquad ON = OP\cos\alpha = \frac{c^2-a^2}{c}.$$

圓錐。把 OP 繞 x 軸旋轉，得到一個斜高（slant height）為 $l = OP$、底半徑為 PN 的圓錐。它的曲面面積是 $\pi \times$ (底半徑) $\times$ (斜高)，即

$$\pi \cdot \frac{a\sqrt{c^2 - a^2}}{c} \cdot \sqrt{c^2 - a^2} = \frac{\pi a(c^2 - a^2)}{c}.$$

球帶。把弧 PR 旋轉，得到球面（球心 C，半徑 a）上夾在兩個平面之間的部分：一個是過 N 且垂直於軸的平面，另一個是在 R 處的切平面（tangent plane）。這兩個平面相距

$$NR = (c + a) - \frac{c^2 - a^2}{c} = \frac{ac + a^2}{c} = \frac{a(c + a)}{c}$$

，所以由第一個結果，球帶的面積是

$$2\pi a \cdot \frac{a(c + a)}{c} = \frac{2\pi a^2(c + a)}{c}.$$

總和。把兩部分相加，

$$\begin{aligned} \text{面積} &= \frac{\pi a}{c}\left[(c^2 - a^2) + 2a(c + a)\right] \\ &= \frac{\pi a}{c}\left(c^2 + 2ac + a^2\right) \\ &= \frac{\pi a(c + a)^2}{c}. \end{aligned}$$

驗算

- 取 $a = 1$，$c = 2$。那麼 $P = \left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right)$，圓錐的面積是 $\frac{3\pi}{2}$，球帶的面積是 3π，所以總面積是 $\frac{9\pi}{2} \approx 14.137$。以 20 000 條短線段逼近曲線，再把所得各個圓台（frustum）的面積相加，得 14.13717。
- 當 $a \to c$ 時，P 趨近 O，圓錐隨之消失。曲面變成半徑為 c 的整個球面，而果然 $\pi c(2c)^2/c = 4\pi c^2$。
- 當 a 固定而 $c \to \infty$ 時，面積的增長與 πac 相若，這正是一個斜高約為 c、底半徑約為 a 的細長圓錐的面積。這與我們的預期相符。

常見錯誤

- 以 OC 或 ON 代替 OP 作為圓錐的斜高。
- 由 C 或 Q 起量度球帶的厚度，而不是由 N（即由 P 所作垂線的垂足）起量度。
- 使用公式 $2\pi rh$ 時取 $r = c$ 而非 $r = a$：被截的球面是以 C 為球心、半徑為 a 的那一個，而不是任何以 O 為球心的球面。

STEP 之外

球帶的結果就是阿基米德（Archimedes）定理。把球面沿水平方向投影到外接圓柱上，面積保持不變，這就是蘭伯特（Lambert）等積地圖投影背後的想法。它也解釋了為甚麼在球面上均勻地選取一點時，該點的高度服從均勻分佈（uniform distribution）。旋轉曲面的面積，是大學一年級向量微積分（vector calculus）課程中曲面積分（surface integral）$\iint |\mathbf{r}_u \times \mathbf{r}_v| \, \mathrm{d}u \, \mathrm{d}v$ 的第一個例子。若想從優美的幾何角度認識曲率（curvature）與曲面，可參閱 Needham 的 *Visual Differential Geometry and Forms*（以圖像講解微分幾何與微分形式）。

詳細解答

STEP III 1989，第 2 題

純數學 • 三維空間中的平面

> **STEP III 1989，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/89-s3-q2

關鍵思路

引入坐標：東、北和上。這樣每個平面都是一個圖像 $z = px + qy + r$，而三個點就能確定一個平面。平面有多陡，只取決於 z 沿水平方向變化得多快，所以所求的角可由 p 和 q 得出。第二部分有一條巧妙的捷徑：礦層在地面以下的深度同樣是位置的線性函數（linear function），而礦層恰好在深度為零之處露出地面。

解答

建立坐標

取 A 為原點，x 軸指向東，y 軸指向北，z 軸鉛垂向上（所有長度以米為單位）。於是

$$A = (0,0,0), \qquad B = (0,100,-5), \qquad C = (60,100,-10),$$

因為 B 比 A 低 5 m，而 C 又比 B 再低 5 m。鑽孔（bore-hole）是鉛垂的，所以鑽孔碰到礦層（seam）的點就在這些點的正下方：

$$A' = (0,0,-95), \qquad B' = (0,100,-50), \qquad C' = (60,100,-86).$$

礦層的傾角

礦層並非鉛垂，因為 A', B', C' 位於三個在平面圖（plan）上不共線（collinear）的點的上方。所以我們可以把它寫成 $z = px + qy + r$。

- 在 A'：$r = -95$。
- 由 A' 到 B'，y 增加 100，z 增加 45，所以 $100q = 45$，即 $q = \frac{9}{20}$。

► 由 B' 到 C'，x 增加 60，z 的變化為 -36，所以 $60p = -36$，即 $p = -\frac{3}{5}$。

因此礦層為

$$z = -\tfrac{3}{5}x + \tfrac{9}{20}y - 95,$$

它的一個法向量（normal vector）是 $\mathbf{n} = \left(\frac{3}{5}, -\frac{9}{20}, 1\right)$，或者乘以 20 後得 $(12, -9, 20)$。

兩個平面之間的夾角等於它們的法線之間的夾角。水平面的法向量是 $\mathbf{k} = (0,0,1)$，所以礦層與水平面的夾角 θ 滿足

$$\cos\theta = \frac{|(12,-9,20)\cdot(0,0,1)|}{\sqrt{12^2+9^2+20^2}} = \frac{20}{\sqrt{144+81+400}} = \frac{20}{\sqrt{625}} = \frac{20}{25} = \frac{4}{5}.$$

故此礦層與水平面的傾角為 $\cos^{-1}\left(\frac{4}{5}\right)$，正如題目所求。

（另一個看法：沿最大斜率線（line of greatest slope），水平方向每走一米，z 的變化為 $\sqrt{p^2+q^2} = \sqrt{\frac{9}{25} + \frac{81}{400}} = \frac{3}{4}$。所以 $\tan\theta = \frac{3}{4}$，是一個 3–4–5 三角形，於是 $\cos\theta = \frac{4}{5}$。）

礦層在何處露出地面

地面是一個通過原點的平面。與之前的做法一樣，由 A 向北走到 B，地面每 100 下降 5；由 B 向東走到 C，地面每 60 下降 5，所以

$$z = -\tfrac{1}{12}x - \tfrac{1}{20}y.$$

? 想一想 我們真的需要聯立求解兩個平面方程嗎？

不盡然。設 $d(x,y)$ 為在水平位置 (x,y) 處礦層在地面以下的深度。它是兩個線性函數之差，所以本身也是線性的：

$$d(x,y) = \left(-\tfrac{1}{12}x - \tfrac{1}{20}y\right) - \left(-\tfrac{3}{5}x + \tfrac{9}{20}y - 95\right) = \tfrac{31}{60}x - \tfrac{1}{2}y + 95.$$

作為驗算，$d(0,0) = 95$、$d(0,100) = 45$ 及 $d(60,100) = 45 + 31 = 76$，正是三個鑽孔的深度。（其實我們可以直接由這些數據寫出 d：向北每走 100 m，深度減少 50；向東每走 60 m，深度增加 31。）

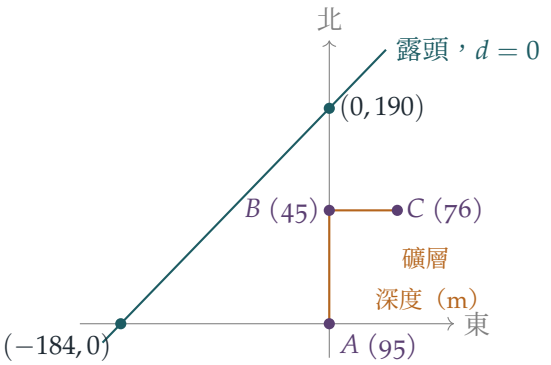
礦層在 $d = 0$ 之處露出地面，在平面圖上這是直線

$$\tfrac{31}{60}x - \tfrac{1}{2}y + 95 = 0, \qquad \text{即} \qquad y = \tfrac{31}{30}x + 190.$$

沿這條直線，每向東移動 30 m，就同時向北移動 31 m。所以這條直線與正北方向成角 φ，其中

$$\tan\varphi = \frac{\text{東}}{\text{北}} = \frac{30}{31}, \qquad \varphi = \tan^{-1}\left(\tfrac{30}{31}\right) \approx 44.06^\circ.$$

因此這條直線的方位角（bearing）是 $\tan^{-1}\left(\frac{30}{31}\right) \approx 044^\circ$。若從相反方向來讀，則是 224°；一條直線有這兩個方位角，用哪一個來描述它都可以。這條直線的走向略偏於東北方向以西，在平面圖上它通過 A 正北 190 m 處。



驗算

- 礦層的法向量 $(12, -9, 20)$ 的長度恰好是 25，所以 $\cos\theta = \frac{20}{25}$，是一個整齊的 3–4–5 答案，與題目所預告的一致。
- 深度公式 d 重現了全部三個鑽孔的深度，所以兩個平面都是對的。
- 數值上，$\tan^{-1}(30/31) = 44.0608°$。在直線 $y = \frac{31}{30}x + 190$ 上隨機取點，兩個平面給出的高度相同，只差捨入誤差。

常見錯誤

- 把每個鑽孔由孔口起計的深度直接拿來用，彷彿各孔口處於同一水平。B 和 C 分別比 A 低 5 m 和 10 m，所以在該兩處礦層位於 A 的水平以下 50 m 和 86 m，而不是 45 m 和 76 m。
- 給出礦層內一條水平線的方位角（即 $-\frac{3}{5}x + \frac{9}{20}y$ 為常數之處，方位角為 $\tan^{-1}\frac{3}{4} \approx 037°$），而不是露頭（outcrop）的方位角。由於地面也是傾斜的，兩者並不相同。
- 把 $\tan\varphi =$ 東/北 與北/東混淆。方位角是由北起順時針方向量度的。

STEP 之外

這是地質學家的「走向與傾角」（strike and dip）問題，所用的工具屬於大學一年級的向量幾何：以 $\mathbf{n}\cdot\mathbf{r} = c$ 表示平面、透過法線求平面之間的夾角，以及交線。深度這一招是一個普遍原理的例子：兩個仿射函數（affine function）之差仍是仿射函數，所以兩個不平行平面的交集就是某一個線性函數的零點集（zero set）。Strang 的 *Introduction to Linear Algebra*（線性代數入門）把這幅圖像融入線性方程的幾何之中。

詳細解答

STEP III 1989，第 3 題

純數學 • 矩陣群

STEP III 1989，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/89-s3-q3

關鍵思路

$\mathbf{M}$ 是轉角為 $2\pi/m$ 的旋轉（rotation），所以 $\mathbf{M}^m = \mathbf{I}$，而它各次冪之和是一個等比級數（geometric series）：把這個和乘以 $\mathbf{I}-\mathbf{M}$，各項便會逐一相消（telescope）。至於遞推關係（recurrence），只要寫出 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$，規律便一目了然。之後的羣（group）全靠一條恆等式：把其中一個表達式代入另一個，結果只是把下標相加，所以這些元素的運算方式就像模（modulo）m 的整數一樣。

解答

M 的各次冪之和

$\mathbf{M}$ 是繞原點逆時針旋轉 $\theta = 2\pi/m$ 的矩陣。連續施行 k 次，就是旋轉 $k\theta$，所以

$$\mathbf{M}^k = \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix}, \qquad \text{特別地} \qquad \mathbf{M}^m = \begin{pmatrix} \cos 2\pi & -\sin 2\pi \\ \sin 2\pi & \cos 2\pi \end{pmatrix} = \mathbf{I}.$$

（由於 $\mathbf{M}^{k+1} = \mathbf{M}\mathbf{M}^k$，$\mathbf{M}^k$ 的公式可用歸納法（induction）由複角公式（compound-angle formulae）推出。）

設 $\mathbf{S} = \mathbf{I} + \mathbf{M} + \cdots + \mathbf{M}^{m-1}$。這些矩陣全是 $\mathbf{M}$ 的冪，所以它們可交換（commute），而乘積中各項逐一相消：

$$(\mathbf{I}-\mathbf{M})\mathbf{S} = \mathbf{I} - \mathbf{M}^m = \mathbf{I} - \mathbf{I} = \mathbf{O}.$$

? **想一想** 我們可以直接「約去」因子 $\mathbf{I}-\mathbf{M}$ 嗎？

只有當 $\mathbf{I}-\mathbf{M}$ 可逆（invertible）時才可以：非零矩陣的乘積可以是 $\mathbf{O}$。這裏

$$\det(\mathbf{I}-\mathbf{M}) = (1-\cos\theta)^2 + \sin^2\theta = 2-2\cos\theta,$$

而由於 $m>1$，我們有 $0<\theta\le\pi$，所以 $\cos\theta\neq 1$，行列式（determinant）不等於零。在 $(\mathbf{I}-\mathbf{M})\mathbf{S}=\mathbf{O}$ 的左邊乘以 $(\mathbf{I}-\mathbf{M})^{-1}$，便得 $\mathbf{S}=\mathbf{O}$，這正是所求的結果。

$\mathbf{X}_k$ 的公式

算出最初幾項，

$$\mathbf{X}_1 = \mathbf{P}\mathbf{X}_0+\mathbf{Q},\quad \mathbf{X}_2 = \mathbf{P}^2\mathbf{X}_0+\mathbf{P}\mathbf{Q}+\mathbf{Q},\quad \mathbf{X}_3 = \mathbf{P}^3\mathbf{X}_0+(\mathbf{P}^2+\mathbf{P}+\mathbf{I})\mathbf{Q},$$

由此可猜想，對 $k\ge 0$，

$$\mathbf{X}_k = \mathbf{P}^k\mathbf{X}_0+\mathbf{S}_k\mathbf{Q},\qquad \text{其中}\quad \mathbf{S}_k = \mathbf{I}+\mathbf{P}+\cdots+\mathbf{P}^{k-1}\quad(\mathbf{S}_0=\mathbf{O}).$$

當 $k=0$ 時，這式子就是 $\mathbf{X}_0=\mathbf{X}_0$，顯然成立。假設它對某個 $k\ge 0$ 成立，則

$$\begin{aligned}\mathbf{X}_{k+1} = \mathbf{P}\mathbf{X}_k+\mathbf{Q} &= \mathbf{P}^{k+1}\mathbf{X}_0+(\mathbf{P}+\mathbf{P}^2+\cdots+\mathbf{P}^k)\mathbf{Q}+\mathbf{Q}\\ &= \mathbf{P}^{k+1}\mathbf{X}_0+\mathbf{S}_{k+1}\mathbf{Q},\end{aligned}$$

所以它對 $k+1$ 也成立，按歸納法，它對所有 $k\ge 0$ 都成立。要保持因子的次序：$\mathbf{P}$ 自始至終都是從左邊乘上去的，而 $\mathbf{P}$ 與 $\mathbf{Q}$ 不一定可交換。

羣

把 $\mathbf{X}_i$ 看成表達式 $\mathbf{P}^i\mathbf{X}_0+\mathbf{S}_i\mathbf{Q}$，即一條把任意矩陣 $\mathbf{X}_0$ 變成另一個矩陣的規則。以 $\mathbf{X}_j$ 代替 $\mathbf{X}_0$，得

$$\begin{aligned}\mathbf{X}_i * \mathbf{X}_j &= \mathbf{P}^i\left(\mathbf{P}^j\mathbf{X}_0+\mathbf{S}_j\mathbf{Q}\right)+\mathbf{S}_i\mathbf{Q}\\ &= \mathbf{P}^{i+j}\mathbf{X}_0+\left(\mathbf{P}^i\mathbf{S}_j+\mathbf{S}_i\right)\mathbf{Q}.\end{aligned}$$

現在 $\mathbf{P}^i\mathbf{S}_j = \mathbf{P}^i+\cdots+\mathbf{P}^{i+j-1}$，再加上 $\mathbf{S}_i=\mathbf{I}+\cdots+\mathbf{P}^{i-1}$，剛好就是 $\mathbf{S}_{i+j}$。所以，對任意 $\mathbf{P}$，

$$\mathbf{X}_i * \mathbf{X}_j = \mathbf{X}_{i+j}.$$

（這並不令人意外：先按遞推關係走 j 步，再多走 i 步，就是走了 $i+j$ 步。）

現在取 $\mathbf{P}=\mathbf{M}$。由第一部分，$\mathbf{M}^m=\mathbf{I}$ 且 $\mathbf{S}_m=\mathbf{O}$，所以

$$\mathbf{X}_m = \mathbf{X}_0,$$

也就是說，$\mathbf{X}_m$ 是令每個矩陣保持不變的表達式。於是對任意 k，$\mathbf{X}_{k+m}=\mathbf{X}_k*\mathbf{X}_m=\mathbf{X}_k$，所以這列表達式以週期 m 重複出現，而

$$G=\{\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3,\ldots\}=\{\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_m\}$$

是一個有限集。我們逐一檢驗羣公理。

- 封閉性（*closure*）。$\mathbf{X}_i * \mathbf{X}_j = \mathbf{X}_{i+j}$，而 $i+j \geq 1$，所以它屬於 G。
- 結合律（*associativity*）。$(\mathbf{X}_i * \mathbf{X}_j) * \mathbf{X}_k$ 與 $\mathbf{X}_i * (\mathbf{X}_j * \mathbf{X}_k)$ 都等於 $\mathbf{X}_{i+j+k}$。（更一般地說，代入就是函數的複合，而函數的複合必定滿足結合律。）
- 單位元（*identity*）。$\mathbf{X}_m * \mathbf{X}_j = \mathbf{X}_{m+j} = \mathbf{X}_j$ 且 $\mathbf{X}_j * \mathbf{X}_m = \mathbf{X}_j$，所以 $\mathbf{X}_m$ 是單位元。
- 逆元（*inverses*）。對 $1 \leq i < m$，$\mathbf{X}_i * \mathbf{X}_{m-i} = \mathbf{X}_{m-i} * \mathbf{X}_i = \mathbf{X}_m$，而 $\mathbf{X}_m$ 的逆元就是它自己。每個逆元都屬於 G。

所以 G 在運算 $*$ 下是一個有限羣。

事實上，它剛好有 m 個元素：$\mathbf{X}_1, \ldots, \mathbf{X}_m$ 是互不相同的表達式，因為它們的係數 $\mathbf{M}^1, \ldots, \mathbf{M}^m$ 是轉角各不相同的旋轉，轉角都在 $(0, 2\pi]$ 之內。由於每個元素都是 $\mathbf{X}_1$ 的「冪」，G 就是階（order）為 m 的循環羣（cyclic group）。

驗算

- 當 $m = 2$ 時，$\mathbf{M} = -\mathbf{I}$，和為 $\mathbf{I} + (-\mathbf{I}) = \mathbf{O}$。當 $m = 4$ 時，$\mathbf{M}$ 是四分之一圈的旋轉，而 $\mathbf{I} + \mathbf{M} + \mathbf{M}^2 + \mathbf{M}^3 = \mathbf{I} + \mathbf{M} - \mathbf{I} - \mathbf{M} = \mathbf{O}$。
- 以數值計算，對由 2 至 39 的每個 m，各次冪之和都是 $\mathbf{O}$（在捨入誤差之內）；對隨機選取的 $\mathbf{P}$、$\mathbf{Q}$、$\mathbf{X}_0$，$\mathbf{X}_k$ 的公式與遞推關係吻合；而當 $\mathbf{P} = \mathbf{M}$，$\mathbf{Q}$ 和 $\mathbf{X}_0$ 隨機選取時，按遞推關係走 m 步後剛好回到 $\mathbf{X}_0$。

常見錯誤

- 沒有檢查 $\mathbf{I} - \mathbf{M}$ 是否可逆便把它約去。
- 寫成 $\mathbf{Q}\mathbf{S}_k$ 或 $\mathbf{X}_0\mathbf{P}^k$：矩陣乘法不可交換，而遞推關係中 $\mathbf{P}$ 是放在左邊的。
- 把 $\mathbf{X}_0$ 當作單位元：按題目所列，$\mathbf{X}_0$ 並不在 G 之內；單位元是 $\mathbf{X}_m$，它之所以等於 $\mathbf{X}_0$，全因 $\mathbf{S}_m = \mathbf{O}$。

STEP 之外

每個 $\mathbf{X}_i$ 都是一個仿射映射（*affine map*）$\mathbf{X} \mapsto \mathbf{A}\mathbf{X} + \mathbf{B}$，而 $\mathbf{A}$ 可逆的仿射映射在複合運算下構成一個羣；這裏我們找到的是一個階為 m 的循環子羣（cyclic subgroup），它由單獨一個映射生成，而這個映射的 m 次冪是單位元。$\mathbf{I} + \mathbf{M} + \cdots + \mathbf{M}^{m-1} = \mathbf{O}$ 這個事實，正是「m 次單位根（roots of unity）之和為零」的矩陣版本；而背後的一般想法，即

一個羣以變換的方式作用於一個集合，是大學第一門羣論（group theory）課程的核心。Allenby 的 *Rings, Fields and Groups* 對此有詳細闡述。

詳細解答

STEP III 1989，第 4 題

純數學 • 曲線描繪與根

STEP III 1989，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/89-s3-q4

關鍵思路

把分式拆成多項式加一個真分式（proper fraction），斜漸近線（oblique asymptote）便立刻出現，求導數（derivative）也變得容易。結果發現導數處處大於 2，所以三段分支各自由 $-\infty$ 一直上升到 $+\infty$。題目後半部的每一條方程，其實都是「曲線 = 某條較簡單的曲線」的變相寫法。兩者之差在每段分支上的嚴格單調性（strict monotonicity），便能準確數出交點的數目。

解答

曲線

由於 $2x(x^2-5) = 2x(x^2-4) - 2x$，

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

這個函數是奇函數（odd function），所以曲線關於原點有 2 階旋轉對稱（rotational symmetry）。曲線只在 $x(x^2-5) = 0$ 之處與坐標軸相交，即 $x = 0$ 及 $x = \pm\sqrt{5}$。

漸近線。分母在 $x = \pm 2$ 時為零（此時分子為 $\mp 4 \neq 0$），所以 $x = 2$ 和 $x = -2$ 是鉛垂漸近線（vertical asymptote）。當 $|x| \to \infty$ 時，$\frac{2x}{x^2-4} \to 0$，所以 $y = 2x$ 是斜漸近線。此外，$y - 2x = -\frac{2x}{x^2-4}$ 在 x 為很大的正數時是負的，在 x 為很大的負數時是正的。因此曲線在右方從下方趨近 $y = 2x$，在左方則從上方趨近。

斜率。對拆開後的式子求導，

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2 - \frac{2(x^2-4) - 2x\cdot 2x}{(x^2-4)^2} = 2 + \frac{2(x^2+4)}{(x^2-4)^2} > 2.$$

所以曲線沒有駐點（stationary point），而 y 在 $x<-2$、$-2<x<2$ 和 $x>2$ 這三個區間上各自嚴格遞增。再求導一次，得 $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2}=-\dfrac{4x(x^2+12)}{(x^2-4)^3}$，所以唯一的拐點（point of inflection）是原點。

原點處的切線。在 $x=0$ 處斜率為 $2+\frac{2\cdot4}{16}=\frac{5}{2}$，所以該處的切線（tangent）是

$$y=\tfrac{5}{2}x.$$

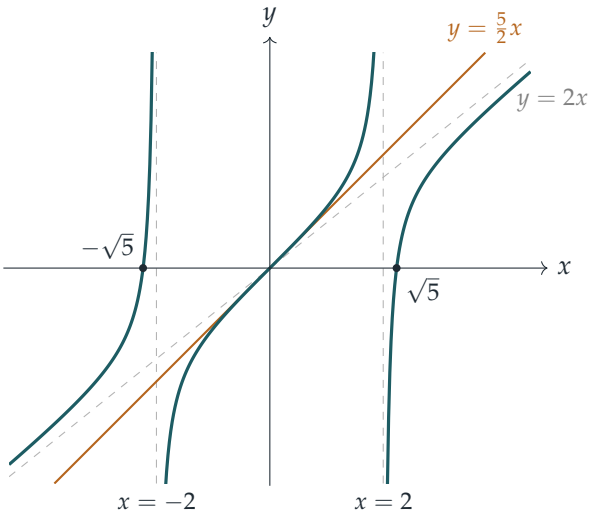
在 $x=\pm2$ 附近的性態。在 $x=2$ 附近，$-\frac{2x}{x^2-4}$ 這一項起主導作用。當 $x\to2^+$ 時它趨向 $-\infty$，當 $x\to2^-$ 時趨向 $+\infty$；由對稱性，當 $x\to-2^-$ 時它趨向 $+\infty$，當 $x\to-2^+$ 時趨向 $-\infty$。因此：

- 當 $x<-2$ 時，y 由 $-\infty$ 上升至 $+\infty$，並在 $-\sqrt{5}$ 處穿過橫軸；
- 當 $-2<x<2$ 時，y 由 $-\infty$ 上升至 $+\infty$，並經過原點；
- 當 $x>2$ 時，y 由 $-\infty$ 上升至 $+\infty$，並在 $\sqrt{5}$ 處穿過橫軸。

還有一個事實稍後會用得着。當 $0<x<2$ 時，

$$y-\tfrac{5}{2}x=\frac{x\left[4(x^2-5)-5(x^2-4)\right]}{2(x^2-4)}=\frac{-x^3}{2(x^2-4)}>0,$$

所以在這一段上，曲線位於原點處切線的上方（由對稱性，當 $-2<x<0$ 時則位於切線下方）。



第 (i) 部分

$x = 0$（左邊為 0，右邊為 8）和 $x = \pm 2$（左邊為 -16，右邊為 0）都不是根。所以我們可以除以 $2x(x^2-4)$，方程變成

$$\frac{2x(x^2-5)}{x^2-4} = \frac{5x-2}{2x} = \frac{5}{2} - \frac{1}{x}.$$

因此，方程的根就是曲線與雙曲線（hyperbola）$H\colon y = \frac{5}{2} - \frac{1}{x}$ 各交點的 x-坐標。我們逐段分支來看。

- $x < -2$：這裏 H 的斜率為 $\frac{1}{x^2} < \frac{1}{4}$，而曲線的斜率 > 2。所以兩者之差（曲線 $-$ H）嚴格遞增。當 $x \to -\infty$ 時它趨向 $-\infty$，當 $x \to -2^-$ 時趨向 $+\infty$，因為 H 一直介乎 $\frac{5}{2}$ 與 3 之間。恰有一個根。
- $-2 < x < 0$：曲線上 $y < 0$（三個因子 x、x^2-5、x^2-4 全是負的），但 H 上 $y > \frac{5}{2}$。沒有根。
- $0 < x < 2$：由上面的事實，$y > \frac{5}{2}x$。另外

  $$\tfrac{5}{2}x - \left(\tfrac{5}{2} - \tfrac{1}{x}\right) = \frac{5x^2-5x+2}{2x} > 0,$$

  因為 $5x^2-5x+2$ 的判別式（discriminant）為 $25-40 < 0$。所以曲線嚴格位於 H 的上方。沒有根。
- $x > 2$：與第一種情況一樣，兩者之差嚴格遞增（斜率 $> 2 - \frac{1}{4}$）。它由 $-\infty$ 走到 $+\infty$，因為 H 一直介乎 2 與 $\frac{5}{2}$ 之間。恰有一個根。

所以方程 (i) 恰有**兩**個實根，一個小於 -2，一個大於 2。

想一想 (ii) 的方程兩邊都有平方，怎樣才能從曲線上把它讀出來？

除以 $(x^2-4)^2$。左邊變成 y^2，所以方程說的是 $y^2 = x^2+1$。

第 (ii) 部分

同樣，$x = \pm 2$ 不是根（左邊為 16，右邊為 0）。所以方程就是 $y^2 = x^2+1$，其中 y 是曲線的縱坐標（ordinate）。方程的根就是曲線與等軸雙曲線（rectangular hyperbola）$y^2 - x^2 = 1$ 的兩支 $y = \pm\sqrt{x^2+1}$ 的交點。

上支。設 $f(x) = y - \sqrt{x^2+1}$。則

$$f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} - \frac{x}{\sqrt{x^2+1}} > 2 - 1 > 0,$$

所以 f 在三個區間上各自嚴格遞增。在每個區間上，y 由 $-\infty$ 走到 $+\infty$，而 $\sqrt{x^2+1}$ 保持有限（它的增長只與 $|x|$ 相若，而當 $x \to +\infty$ 時對手是 $2x$；當 $x \to -\infty$ 時，f 的兩項都是負的）。所以 f 由 $-\infty$ 走到 $+\infty$，每個區間上恰有一個根：合共三個。

下支。映射 $(x, y) \mapsto (-x, -y)$ 把曲線映到自身，並把下支映到上支。所以曲線與下支也恰有三個交點。

這兩組根不會重疊：對同一個 x，曲線只有一個 y，它不可能既正又負。所以方程 (ii) 恰有**六**個實根。（作為驗算，它是一條 6 次多項式方程，即 $3x^6 - 33x^4 + 92x^2 - 16 = 0$，所以它的根全是實數。）

? **想一想** (iii) 的變量是 z 而不是 x，而且 z 可能是負數。怎樣才能把曲線用上？

當 $z > 0$ 時，我們令 $z = x^2$，其中 $x > 0$。至於 $z \le 0$ 的情況，則另外直接處理。

第 (iii) 部分

左邊永不為負，所以任何根都必須滿足 $z + 1 \ge 0$，即 $z \ge -1$。另外 $z = 0$ 不是根（左邊為 0，右邊為 16），$z = 4$ 也不是（左邊為 64，右邊為 0）。

滿足 $z > 0$ 的根。令 $z = x^2$，其中 $x > 0$，使每個正根 z 恰好對應一個 $x > 0$。方程變成 $4x^4(x^2-5)^2 = (x^2-4)^2(x^2+1)$。除以 $x^2(x^2-4)^2$ 得

$$y^2 = 1 + \frac{1}{x^2}, \qquad \text{即} \qquad y = \pm s(x), \quad s(x) = \sqrt{1 + x^{-2}}.$$

這裏 s 由 $+\infty$（當 $x \to 0^+$）遞減至 1（當 $x \to \infty$），而 $s'(x) = -\dfrac{1}{x^3\sqrt{1+x^{-2}}}$，所以 $|s'(x)| < \frac{1}{x^3}$。

- $y = s(x)$：函數 $y - s$ 的導數為 $y' - s' > 2$。在 $(0, 2)$ 上，它由 $-\infty$（因為 $y \to 0$ 而 $s \to \infty$）走到 $+\infty$。在 $(2, \infty)$ 上，它同樣由 $-\infty$ 走到 $+\infty$。每個區間上各有一個根，所以共有兩個根。
- 在 $(0, 2)$ 上的 $y = -s(x)$：曲線在那裏滿足 $y > 0 > -s$。沒有根。
- 在 $(2, \infty)$ 上的 $y = -s(x)$：函數 $y + s$ 的導數為 $y' + s' > 2 - \frac{1}{8} > 0$，並由 $-\infty$ 走到 $+\infty$。有一個根。

所以恰有三個正根 z。

滿足 $-1 \le z < 0$ 的根。寫 $z = -t$，其中 $0 < t \le 1$。方程變成

$$4t^2(t+5)^2 = (t+4)^2(1-t).$$

在 $(0, 1]$ 上，左邊嚴格遞增，由 0（當 $t \to 0$）增至 144。右邊嚴格遞減，由 16 減至 0。兩者之差嚴格遞增並且變號，所以恰有一個根。

因此方程 (iii) 恰有**四**個實根：三個正根，另一個在 $(-1,0)$ 內。展開後，它是四次方程（quartic）$4z^4-41z^3+107z^2-8z-16=0$，所以它的四個根全是實數。

驗算

- 在 $x=1000$ 處，$y-2x\approx -0.002$，印證了曲線是從下方趨近 $y=2x$。
- 用數值計算，(i) 的根為 $x\approx -2.143$ 和 2.401，一如所料，外側兩段分支上各有一個。
- (ii) 的根為 $x\approx \pm 0.431,\ \pm 2.155,\ \pm 2.484$，對雙曲線的每一支而言，曲線的每段分支上各有一個。
- (iii) 的根為 $z\approx -0.332,\ 0.473,\ 4.799,\ 5.310$。後三個給出 $x=\sqrt{z}\approx 0.688,\ 2.191,\ 2.304$，曲線上相應的值為 $y\approx 1.77,\ -1.10,\ 1.09$，與分情況討論的結果吻合。

常見錯誤

- 沒有先檢查 $x=0$ 和 $x=\pm 2$ 不是根，就除以 x 或 x^2-4。
- 在 (ii) 中，只數與 $y=+\sqrt{x^2+1}$ 的交點而忘記了下支，結果得出 3 而不是 6。
- 在 (iii) 中，代入 $z=x^2$ 而忘記了 z 可以是負數，因而漏掉 $(-1,0)$ 內的根。

STEP 之外

這裏的計數論證，建基於介值定理（intermediate value theorem，即變號的連續函數必有零點），再加上嚴格單調性（零點最多只有一個）。把這兩個事實嚴格地建立起來，是大學分析課程最先要做的工作之一。Abbott 的 *Understanding Analysis* 對兩者都有細緻而易讀的論述。笛卡兒（Descartes）符號法則（rule of signs）和施圖姆（Sturm）序列是數算多項式實根數目的代數方法，與這裏所用的幾何方法對照起來，相當有趣。

詳細解答

STEP III 1989，第 5 題

純數學 • 切比雪夫多項式

> **STEP III 1989，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/89-s3-q5

關鍵思路

設 $x = \cosh\theta$。這樣 $e^{\pm\theta}$ 正好就是 $x \pm \sqrt{x^2-1}$，而 $y = \cosh n\theta$ 是它們的 n 次冪的平均值。用二項式定理（binomial theorem）展開這些冪，$\sqrt{x^2-1}$ 的奇次冪會互相抵消，餘下的就是 x 的多項式（polynomial）。只要把這個多項式寫成整齊的形式，每個係數（coefficient）不過是幾個二項式係數（binomial coefficient）之和。

解答

閉合形式

設 $\theta = \cosh^{-1} x$，則 $\theta \geq 0$ 且 $x = \cosh\theta$。於是 $\sinh\theta \geq 0$ 而 $\sinh^2\theta = x^2 - 1$，所以 $\sinh\theta = \sqrt{x^2-1}$。因此

$$e^{\theta} = \cosh\theta + \sinh\theta = x + \sqrt{x^2-1}, \qquad e^{-\theta} = \cosh\theta - \sinh\theta = x - \sqrt{x^2-1}.$$

故此

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

這正是所要證明的結果。

為何 y 是奇多項式

記 $s = \sqrt{x^2-1}$。由二項式定理，

$$(x+s)^n + (x-s)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} s^j \left(1 + (-1)^j\right) = 2 \sum_{j \text{ 為偶數}} \binom{n}{j} x^{n-j} s^j.$$

s 的奇次冪互相抵消，而每個偶次冪 $s^{2m} = (x^2-1)^m$ 都是 x 的多項式。所以

$$y = \sum_{m=0}^{\lfloor n/2 \rfloor} \binom{n}{2m} x^{n-2m}(x^2-1)^m.$$

現在取 $n = 2k+1$，則 m 由 0 取到 k。每一項都是奇次冪 $x^{2k+1-2m}$ 乘以一個 x^2 的多項式，所以只含 x 的奇次冪，其中最高的是 x^{2k+1}。因此 $y = a_0x + a_1x^3 + \cdots + a_kx^{2k+1}$。

要讀出各係數，設 $j = k-m$（即 $j = 0,\ldots,k$），並利用 $\binom{2k+1}{2k-2j} = \binom{2k+1}{2j+1}$：

$$y = \sum_{j=0}^{k} \binom{2k+1}{2j+1} x^{2j+1}(x^2-1)^{k-j}.$$

指標為 j 的那一項由 x^{2j+1} 開始。所以 x^{2r+1} 只來自 $j \le r$ 的各項，而在其中第 j 項，我們需要 $(x^2-1)^{k-j}$ 中的 $(x^2)^{r-j}$ 項，即 $\binom{k-j}{r-j}(-1)^{k-r}$。因此

$$a_r = (-1)^{k-r} \sum_{j=0}^{r} \binom{2k+1}{2j+1}\binom{k-j}{r-j}. \qquad (*)$$

a_0

當 $r = 0$ 時，只有 $j = 0$ 有貢獻：

$$a_0 = (-1)^k \binom{2k+1}{1} = (-1)^k(2k+1).$$

第 (i) 部分

當 $r = 1$ 時，$j = 0, 1$ 有貢獻：

$$\begin{aligned} a_1 &= (-1)^{k-1}\left[(2k+1)\binom{k}{1} + \binom{2k+1}{3}\right] = (-1)^{k-1}(2k+1)\left[k + \frac{2k(2k-1)}{6}\right] \\ &= (-1)^{k-1}(2k+1)\,k\,\frac{3+2k-1}{3} = (-1)^{k-1}\,\frac{2k(k+1)(2k+1)}{3}, \end{aligned}$$

這正是所要證明的結果。

第 (ii) 部分

這裏需要 $k \ge 2$，x^5 才會出現。當 $r = 2$ 時，$j = 0, 1, 2$ 有貢獻：

$$a_2 = (-1)^k\left[(2k+1)\binom{k}{2} + \binom{2k+1}{3}\binom{k-1}{1} + \binom{2k+1}{5}\right].$$

每一項都有因子 $(2k+1)k(k-1)$；把它抽出，

$$\begin{aligned} a_2 &= (-1)^k(2k+1)k(k-1)\left[\frac{1}{2}+\frac{2k-1}{3}+\frac{(2k-1)(2k-3)}{30}\right] \\ &= (-1)^k(2k+1)k(k-1)\,\frac{15+(20k-10)+(4k^2-8k+3)}{30} \\ &= (-1)^k(2k+1)k(k-1)\,\frac{4(k+1)(k+2)}{30} = (-1)^k\,\frac{2(k-1)k(k+1)(k+2)(2k+1)}{15}. \end{aligned}$$

處理第三項時，我們用了 $\binom{2k+1}{5} = \frac{(2k+1)(2k)(2k-1)(2k-2)(2k-3)}{120} = \frac{(2k+1)k(k-1)(2k-1)(2k-3)}{30}$。

> **想一想** 我們的答案有一個因子 $(k+1)$，而 (ii) 中印出的公式卻沒有。哪一個才對？

用 $k=2$ 來檢驗。這時 $y=\cosh(5\theta)$，其中 $x=\cosh\theta$，而 $\cosh 5\theta = 16x^5-20x^3+5x$，所以 $a_2 = 16$。我們的公式給出 $2\cdot1\cdot2\cdot3\cdot4\cdot5/15 = 16$。印出的公式則給出 $2\cdot1\cdot2\cdot4\cdot5/15 = \frac{16}{3}$，這不可能正確，因為各個 a_r 都是整數。所以試卷有印刷錯誤：漏了因子 $(k+1)$。正確的結果就是上面所示的那一個。

係數之和

在多項式中代入 $x=1$。由於恆等式對 $x\geq 1$ 成立，而 $x=1$ 在容許範圍之內，

$$\sum_{r=0}^{k} a_r = y(1) = \cosh\bigl(n\cosh^{-1}1\bigr) = \cosh 0 = 1.$$

驗算

- $k=1$：$\cosh 3\theta = 4x^3-3x$。公式給出 $a_0=-3$ 和 $a_1 = 2\cdot1\cdot2\cdot3/3 = 4$，而 $-3+4=1$。
- $k=3$：$\cosh 7\theta = 64x^7-112x^5+56x^3-7x$。公式給出 $a_0=-7$、$a_1=56$ 和 $a_2 = -2\cdot2\cdot3\cdot4\cdot5\cdot7/15 = -112$，而各係數之和為 1。
- 我們用電腦把 $(*)$ 與遞推關係（recurrence）$T_{m+1}=2xT_m-T_{m-1}$ 對照，檢驗了所有 $k\leq 14$ 的情況，並在隨機選取的 $x\geq 1$ 處把多項式與 $\cosh(n\cosh^{-1}x)$ 作比較。

常見錯誤

- 忘記了 x^{2r+1} 的係數來自展開式中的好幾項，而不只是一項。
- 正負號出錯：a_r 的每一項貢獻都帶有相同的符號 $(-1)^{k-r}$，所以應先把它抽出。

- 硬要把代數運算湊成 (ii) 中印出的形式。只要檢驗一個小的情況，例如 $k = 2$，便知道那是印錯了的。

STEP 之外

多項式 y 就是切比雪夫多項式（Chebyshev polynomial）$T_n(x)$，通常以 $T_n(\cos\phi) = \cos n\phi$ 來定義。我們的雙曲函數（hyperbolic function）版本是同一個恆等式在 $x \geq 1$ 時的情形，而兩者一致，是因為多項式由它在任何一個區間上的值完全確定。切比雪夫多項式在逼近論（approximation theory）中佔有核心地位，因為在所有 n 次首一多項式（monic polynomial）之中，$2^{1-n}T_n$ 在 $[-1, 1]$ 上的最大絕對值是最小的。在任何數值分析（numerical analysis）的入門課程中，它們都會再次出現。如果想多做一些這類代數運算的 STEP 式練習，可以試試 Siklos 的 *Advanced Problems in Mathematics*（STEP 進階數學題集）。

詳細解答

STEP III 1989，第 6 題

純數學 • 圓錐曲線的極坐標方程

STEP III 1989，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/89-s3-q6

關鍵思路

在極坐標（polar coordinates）中，一條不經過原點的直線的 $1/r$ 等於 $\cos\theta$ 與 $\sin\theta$ 的組合，而把它寫成 $A\cos\theta + B\cos(\theta-\gamma)$，就可以與圓錐曲線（conic）本身的 $\cos\theta$ 項完全配對。只要 $e\cos\theta$ 的部分一致，弦（chord）的方程便只需處理單獨一個餘弦項，其餘的就交給關於 $\theta=\gamma$ 的對稱性。讓弦收縮成一點便得到切線（tangent），而有了切線方程，平分線（bisector）性質幾乎是自動得出的。

解答

全文之中，(r,θ) 為極坐標，其中 $x = r\cos\theta$，$y = r\sin\theta$，並且容許 $r<0$（點 (r,θ) 與 $(-r,\theta+\pi)$ 是同一點）。圓錐曲線的焦點（focus）S 位於原點。

直線的極坐標形式

由於 $lx+my=1$ 不可能經過原點，直線上 $r\neq 0$，兩邊除以 r 得

$$\frac{1}{r} = l\cos\theta + m\sin\theta.$$

把題目建議的形式展開，

$$A\cos\theta + B\cos(\theta-\gamma) = (A + B\cos\gamma)\cos\theta + B\sin\gamma\,\sin\theta.$$

若 $B\sin\gamma = m$ 及 $A + B\cos\gamma = l$，兩式便對所有 θ 都相等。由於 $\sin\gamma\neq 0$，可取

$$B = \frac{m}{\sin\gamma}, \qquad A = l - m\cot\gamma,$$

這就是所求的極坐標方程。反過來說，任何方程 $1/r = A\cos\theta + B\cos(\theta-\gamma)$，只要 A、B 不會令兩個係數同時為零，都是一條直線 $lx+my=1$，其中 $l = A + B\cos\gamma$，

$m = B\sin\gamma$。留意這個形式在 $(r,\theta) \to (-r,\theta+\pi)$ 之下保持不變，所以無論 r 取哪個符號，它描述的都是整條直線。

弦 PQ

設弦的方程具有 $a/r = A\cos\theta + B\cos(\theta-\gamma)$ 的形式。在 P 和 Q 處有 $\theta = \gamma \mp \delta$ 及 $a/r = 1 + e\cos\theta$，所以我們需要

$$A\cos(\gamma\mp\delta) + B\cos(\mp\delta) = 1 + e\cos(\gamma\mp\delta).$$

因為對兩點而言都有 $\cos(\mp\delta) = \cos\delta$，取 $A = e$，$B\cos\delta = 1$ 便可同時滿足兩條方程，而 $\cos\delta \neq 0$ 保證可以解出 B。兩個相異的點確定一條直線，所以弦 PQ 為

$$\boxed{\frac{a}{r} = e\cos\theta + \sec\delta\,\cos(\theta-\gamma).}$$

（若 $\cos\delta = 0$，則 P 和 Q 位於 $\theta = \gamma \mp \frac{\pi}{2}$，而 PQ 是一條焦弦（focal chord）；經過原點的直線沒有 $1/r = \ldots$ 這種形式的方程，這就是該情況被排除的原因。）

想一想 我們怎樣知道這條弦只在 P 和 Q 與圓錐曲線相交？

對每個 θ，圓錐曲線上恰有一點，滿足 $a/r = 1 + e\cos\theta$。弦經過該點，當且僅當 $1 + e\cos\theta = e\cos\theta + \sec\delta\cos(\theta-\gamma)$，即 $\cos(\theta-\gamma) = \cos\delta$，所以 $\theta = \gamma \pm \delta$。$e\cos\theta$ 項互相抵消，這正是選取 $A = e$ 的用意所在。

在 $\theta = \gamma$ 處的切線

設 $\theta = \gamma$ 所對應的點為 G，該點的切線就是當 P 和 Q 向 G 靠攏時，亦即 $\delta \to 0$ 時，弦 PQ 的極限。弦的方程連續地依賴於 δ，而 $\sec\delta \to 1$，所以切線為

$$\boxed{\frac{a}{r} = e\cos\theta + \cos(\theta-\gamma).}$$

作為印證，這條直線與圓錐曲線相交於 $\cos(\theta-\gamma) = 1$ 之處，即只在 G 相交，切線理應如此。更具決定性的驗證是：切線的直角坐標（cartesian）形式為 $(e+\cos\gamma)x + (\sin\gamma)y = a$，其法向量（normal vector）為 $(e+\cos\gamma,\ \sin\gamma)$。圓錐曲線為 $r = a - ex$；沿曲線求微分得 $(x\,\mathrm{d}x + y\,\mathrm{d}y)/r = -e\,\mathrm{d}x$，即 $(e + x/r)\,\mathrm{d}x + (y/r)\,\mathrm{d}y = 0$。在 G 處，$x/r = \cos\gamma$ 及 $y/r = \sin\gamma$，所以曲線在該處的法向量正是同一個向量。

最後一部分

把焦點 S 置於原點，圓錐曲線便是 $a/r = 1 + e\cos\theta$，並設 L 和 M 分別為 $\theta = \alpha$ 和 $\theta = \beta$ 所對應的點。記

$$\gamma = \tfrac{1}{2}(\alpha+\beta), \qquad \delta = \tfrac{1}{2}(\beta-\alpha),$$

於是 L 和 M 位於 $\theta = \gamma - \delta$ 和 $\gamma + \delta$，與上文的 P 和 Q 完全一樣。根據上一個結果，在 L 和 M 處的切線分別為

$$\frac{a}{r} = e\cos\theta + \cos(\theta - \alpha), \qquad \frac{a}{r} = e\cos\theta + \cos(\theta - \beta).$$

在 T 處兩式同時成立，所以 $\cos(\theta - \alpha) = \cos(\theta - \beta)$。因此，要麼 $\theta - \alpha = \theta - \beta + 2k\pi$，這會令 $L = M$；要麼

$$\theta - \alpha = -(\theta - \beta) + 2k\pi \quad \Longrightarrow \quad \theta = \tfrac{1}{2}(\alpha + \beta) + k\pi = \gamma + k\pi.$$

所以 T 位於經過 S、方向為 $\theta = \gamma$ 的直線上，且

$$r_T = \frac{a}{e\cos\gamma + \cos\delta}$$

（兩條切線平行、T 不存在的情況，恰好是 $e\cos\gamma + \cos\delta = 0$ 之時）。

射線 SL 和 SM 分別指向 $\alpha = \gamma - \delta$ 和 $\beta = \gamma + \delta$ 的方向（這裏取通常的情況，即在 L 和 M 處 $r > 0$，橢圓（ellipse）或拋物線（parabola）便是如此）。直線 $\theta = \gamma$ 與兩條射線所成的角都是 δ，所以直線 ST 平分角 LSM。（若 $r_T < 0$，則 T 位於反方向的射線上，ST 平分的是優角（reflex angle）LSM；角 LST 與角 MST 仍然相等。）

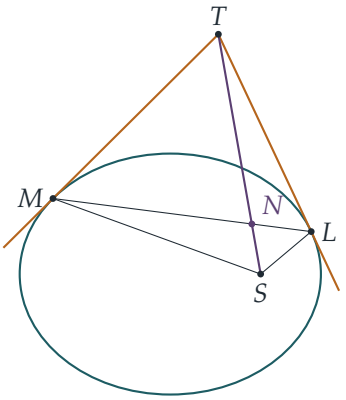
要找出 ST 與 LM 的交點，把該點稱為 N。弦 LM 為 $a/r = e\cos\theta + \sec\delta\cos(\theta - \gamma)$，而在 ST 上有 $\theta = \gamma$，所以

$$\frac{a}{SN} = e\cos\gamma + \sec\delta \quad \Longrightarrow \quad SN = \frac{a\cos\delta}{1 + e\cos\gamma\cos\delta}.$$

利用 $2\cos\gamma\cos\delta = \cos(\gamma - \delta) + \cos(\gamma + \delta) = \cos\alpha + \cos\beta$，上式即

$$\boxed{N \text{ 位於} \theta = \tfrac{1}{2}(\alpha + \beta), \quad r = \frac{2a\cos\frac{1}{2}(\beta - \alpha)}{2 + e\cos\alpha + e\cos\beta}.}$$

當 LM 是焦弦（$\cos\delta = 0$）時，這給出 $N = S$，正如所料。



驗算

- 試試圓的情況（$e = 0$，半徑為 a）：切線變成 $a/r = \cos(\theta - \gamma)$，即 $x\cos\gamma + y\sin\gamma = a$，正是 $x^2 + y^2 = a^2$ 為人熟悉的切線。
- 三角形 LSM 中的角平分線定理（angle bisector theorem）指出 $LN : NM = SL : SM$。取 $a = 1.5$，$e = 0.6$，$\alpha = 40°$，$\beta = 160°$（即圖中的情況），得 $SN \approx 0.791$，而該比例經數值驗算後成立。
- 我們用電腦在數以千計隨機選取的圓錐曲線和點上，測試了弦與切線的方程、ST 的方向、兩角相等的性質，以及 SN 的公式。

常見錯誤

- 解 $\cos(\theta - \alpha) = \cos(\theta - \beta)$ 時只取 $+$ 號。這個分支得不出任何結果（它需要 $\alpha = \beta$），有用的解來自 $-$ 號。
- 沒有善用 A 的選擇自由：取 $A = e$，才能令 $e\cos\theta$ 項在與圓錐曲線求交點時互相抵消。

STEP 之外

ST 平分 $\angle LSM$ 這個結果，是圓錐曲線一個經典的焦點性質，而直線 LM 則是 T 關於該圓錐曲線的*極線*（*polar*）。極點與極線的對偶（pole–polar duality）是通往射影幾何（projective geometry）的大門；在射影幾何中，所有非退化的圓錐曲線都是等價的，相切與關聯都變成純粹的代數關係。Brannan、Esplen 與 Gray 合著的 *Geometry*（《幾何學》）是一本很好的大學一年級程度讀物，講述圓錐曲線與射影幾何。

詳細解答

STEP III 1989，第 7 題

純數學 • 錯切矩陣

> **STEP III 1989，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/89-s3-q7

關鍵思路

用兩個單位向量（unit vector）來處理，一切都最簡單：$\mathbf{u}$ 沿直線 $L\colon y = x\tan\alpha$ 的方向，而 $\mathbf{n}$ 與它垂直。這個錯切（shear）把每一點沿平行於 $\mathbf{u}$ 的方向移動，移動量與該點到 L 的有向距離（signed distance）成正比，所以 $\mathbf{M} = \mathbf{I} + k\,\mathbf{u}\mathbf{n}^{\mathsf{T}}$。至於第二部分，一個矩陣與 $\mathbf{I} + k\mathbf{A}$ 可交換（commute），當且僅當它與 $\mathbf{A}$ 可交換，這樣便可撇除一切枝節。

解答

記 $s = \sin\alpha$ 及 $c = \cos\alpha$。由於 α 是銳角，故 $s > 0$ 且 $c > 0$。設

$$\mathbf{u} = \begin{pmatrix} c \\ s \end{pmatrix}, \qquad \mathbf{n} = \begin{pmatrix} -s \\ c \end{pmatrix},$$

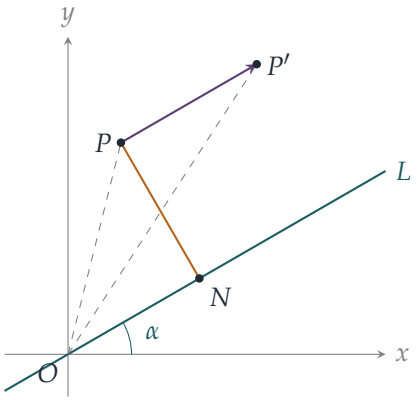
則 $\mathbf{u}$ 是沿 L 方向、與 x 軸成角 α 的單位向量，而 $\mathbf{n}$ 是指向 L 左上方一側的單位法向量（unit normal）。

求 M

對於 $P = (x, y)$，設

$$d = \mathbf{n}\cdot\overrightarrow{OP} = y\cos\alpha - x\sin\alpha,$$

即由 P 到 L 的有向距離。於是 $PN = |d|$，而 $d > 0$ 當且僅當 P 位於 L 的上方。



取 P 在 L 的上方，故 $d > 0$。由 (i)，$\overrightarrow{PP'}$ 沿 $\mathbf{u}$ 的方向，所以 $\overrightarrow{PP'} = \lambda\mathbf{u}$，其中 $\lambda > 0$。條件 (ii) 與此相符：二維叉積（cross product）為

$$\overrightarrow{OP} \times \overrightarrow{OP'} = \overrightarrow{OP} \times (\overrightarrow{OP} + \lambda\mathbf{u}) = \lambda\,(xs - yc) = -\lambda d < 0,$$

所以由 OP 轉到 OP' 是順時針方向，符合要求。條件 (iii) 給出 $\lambda = k\,PN = kd$。（要這些條件同時成立，必須有 $k > 0$。）因此，對於 L 上方的每一點 P，

$$\overrightarrow{OP'} = \overrightarrow{OP} + kd\,\mathbf{u} = \overrightarrow{OP} + k\,\mathbf{u}\,(\mathbf{n} \cdot \overrightarrow{OP}).$$

想一想 我們只用了 L 上方的點。為甚麼這樣已足以完全確定 $\mathbf{M}$？

T 是線性（linear）的，而映射 $\mathbf{v} \mapsto \mathbf{v} + k\,\mathbf{u}(\mathbf{n} \cdot \mathbf{v})$ 也是線性的。兩者在 L 上方的開半平面（open half-plane）上一致，而這個半平面包含兩個線性獨立（linearly independent）的向量（例如 $\mathbf{n}$ 和 $\mathbf{n} + \mathbf{u}$）。兩個線性映射若在一組基（basis）上一致，便完全相等。所以對於所有向量，

$$\mathbf{M} = \mathbf{I} + k\,\mathbf{u}\mathbf{n}^{\mathsf{T}} = \mathbf{I} + k\begin{pmatrix} c \\ s \end{pmatrix}\begin{pmatrix} -s & c \end{pmatrix} = \mathbf{I} + k\begin{pmatrix} -sc & c^2 \\ -s^2 & sc \end{pmatrix}.$$

亦即

$$\mathbf{M} = \begin{pmatrix} 1 - k\sin\alpha\cos\alpha & k\cos^2\alpha \\ -k\sin^2\alpha & 1 + k\sin\alpha\cos\alpha \end{pmatrix},$$

正如所求。作為合理性檢查，由 $\mathbf{n} \cdot \mathbf{u} = 0$ 得 $\mathbf{M}\mathbf{u} = \mathbf{u}$，所以 L 上每一點都固定不動。此外 $\det\mathbf{M} = 1 - k^2s^2c^2 + k^2s^2c^2 = 1$，正是錯切應有的行列式（determinant）值。

可交換的條件

設 $\mathbf{A} = \mathbf{u}\mathbf{n}^{\mathsf{T}} = \begin{pmatrix} -sc & c^2 \\ -s^2 & sc \end{pmatrix}$ 及 $\mathbf{X} = \begin{pmatrix} p & q \\ r & t \end{pmatrix}$。由於 $\mathbf{M} = \mathbf{I} + k\mathbf{A}$，而 $\mathbf{I}$ 與任何矩陣都可交換，

$$\mathbf{XM} - \mathbf{MX} = k(\mathbf{XA} - \mathbf{AX}).$$

因為 $k \neq 0$，$\mathbf{X}$ 與 $\mathbf{M}$ 可交換，當且僅當它與 $\mathbf{A}$ 可交換。乘開得

$$\mathbf{XA} = \begin{pmatrix} -psc - qs^2 & pc^2 + qsc \\ -rsc - ts^2 & rc^2 + tsc \end{pmatrix}, \qquad \mathbf{AX} = \begin{pmatrix} -psc + rc^2 & -qsc + tc^2 \\ -ps^2 + rsc & -qs^2 + tsc \end{pmatrix},$$

所以

$$\mathbf{XA} - \mathbf{AX} = \begin{pmatrix} -(qs^2 + rc^2) & -c^2(t-p) + 2qsc \\ -s^2(t-p) - 2rsc & qs^2 + rc^2 \end{pmatrix}.$$

這個矩陣為零，當且僅當

$$\begin{aligned} &\text{(a)} \quad qs^2 + rc^2 = 0, \\ &\text{(b)} \quad c^2(t-p) = 2qsc, \\ &\text{(c)} \quad s^2(t-p) = -2rsc. \end{aligned}$$

由於 $c \neq 0$，(b) 即 $t - p = 2q\tan\alpha$。由於 $s \neq 0$，(c) 即 $t - p = -2r\cot\alpha$。

> ? **想一想** 這裏有三條方程，但題目所述的條件只包含兩條。(a) 到哪裏去了？

它可由另外兩條推出。若 $2q\tan\alpha = -2r\cot\alpha$，兩邊乘以 $\frac{1}{2}sc$ 便得 $qs^2 = -rc^2$，這正是 (a)。所以

$$\mathbf{XM} = \mathbf{MX} \iff t - p = 2q\tan\alpha = -2r\cot\alpha.$$

兩個方向都成立：以上每一步都是等價關係。這就是所求的充分必要條件（necessary and sufficient condition）。

驗算

- 取 $\alpha = \frac{\pi}{4}$、$k = 2$，得 $\mathbf{M} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$。點 $(0,1)$ 位於 $y = x$ 上方，距離為 $\frac{1}{\sqrt{2}}$，它被移到 $(1,2)$。這是沿 $45°$ 方向移動了 $\sqrt{2} = 2 \times \frac{1}{\sqrt{2}}$，而且轉向是順時針的。
- $\mathbf{X} = \mathbf{M}$ 本身滿足這個條件：$t - p = 2ksc$，$2q\tan\alpha = 2kc^2 \cdot \frac{s}{c} = 2ksc$，以及 $-2r\cot\alpha = 2ks^2 \cdot \frac{c}{s} = 2ksc$。$\mathbf{I}$ 也滿足這個條件，三個式子都等於零。
- 以數值方法檢驗：隨機選取 α、k 和 L 上方的 P，$\mathbf{M}$ 都重現了 (i)–(iii)。滿足條件的隨機矩陣與 $\mathbf{M}$ 可交換，誤差在 10^{-8} 之內；不滿足條件的隨機矩陣則不可交換。

常見錯誤

- 對每一點都使用不帶正負號的距離 PN。位移是有向距離的 k 倍。L 下方的點向相反方向移動，正因如此 T 才是線性的。
- 就可交換的條件只證明了「若」（或只證明了「僅若」）的方向，或忘記說明 $(1,1)$ 元素並沒有帶來新的條件。

STEP 之外

這裏與 $\mathbf{M}$ 可交換的矩陣，恰好就是形如 $a\mathbf{I}+b\mathbf{A}$ 的矩陣：請驗證 $t-p=2q\tan\alpha=-2r\cot\alpha$ 是一個雙參數族。這是一個一般事實的特例。對於一個不是 $\mathbf{I}$ 的倍數的 2×2 矩陣，與它可交換的矩陣恰好就是 $a\mathbf{I}+b\mathbf{M}$。對於 $n\times n$ 矩陣，當最小多項式（minimal polynomial）等於特徵多項式（characteristic polynomial）時，相應的結論（可交換的矩陣都是 $\mathbf{M}$ 的多項式）成立。Strang 的 *Introduction to Linear Algebra*（線性代數入門教材）闡述了秩一矩陣（rank-one matrix）$\mathbf{u}\mathbf{n}^{\mathsf{T}}$ 以及這背後的思想。

詳細解答

STEP III 1989，第 8 題

純數學 • 耦合微分方程

STEP III 1989，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/89-s3-q8

關鍵思路

第一條方程把 y 用 x 和 $\mathrm{d}x/\mathrm{d}t$ 表示出來。因此只要把它再微分一次，出現的 $\mathrm{d}y/\mathrm{d}t$ 就可以用第二條方程代換，y 便完全消失。剩下的是一條二階常係數線性方程（second-order linear equation with constant coefficients）。它的輔助方程（auxiliary equation）有重根（repeated root），而其中一個強迫項（forcing term）正好與這個根吻合。這種共振（resonance）才是本題真正的考驗。

解答

消去 y

以字母上方的點表示對 t 的導數。把 $\dot{x} = 4(x - y)$ 微分，得

$$\ddot{x} = 4\dot{x} - 4\dot{y} = 4\dot{x} - 4x + 48\left(\mathrm{e}^{2t} + \mathrm{e}^{-2t}\right),$$

其中 $\dot{y}$ 用了第二條方程代入。因此

$$\boxed{\ddot{x} - 4\dot{x} + 4x = 48\left(\mathrm{e}^{2t} + \mathrm{e}^{-2t}\right),}$$

式中不含 y。

求解 x

餘函數（*complementary function*）。輔助方程 $m^2 - 4m + 4 = (m-2)^2 = 0$ 有重根 $m = 2$，所以餘函數是

$$x_{\mathrm{c}} = (A + Bt)\,\mathrm{e}^{2t}.$$

特別積分（*particular integral*）。由於方程是線性的，我們可以把兩個強迫項逐一處理。

對於 $48\mathrm{e}^{-2t}$，試用 $x = D\mathrm{e}^{-2t}$。代入後得 $D(4+8+4)\mathrm{e}^{-2t} = 48\mathrm{e}^{-2t}$，所以 $D = 3$。

想一想 對於 $48\mathrm{e}^{2t}$ 這一項，為甚麼 $C\mathrm{e}^{2t}$ 行不通，甚至連 $Ct\,\mathrm{e}^{2t}$ 也行不通？

e^{2t} 和 $t\,\mathrm{e}^{2t}$ 都已經在餘函數之中，所以方程左邊會把它們變成 0。因為 2 是二重根，我們必須升到 $t^2\mathrm{e}^{2t}$。試用 $x = Ct^2\mathrm{e}^{2t}$。一個簡潔的代入方法是寫成 $x = u\,\mathrm{e}^{2t}$。這樣

$$\dot{x} = (\dot{u} + 2u)\mathrm{e}^{2t}, \qquad \ddot{x} = (\ddot{u} + 4\dot{u} + 4u)\mathrm{e}^{2t},$$

所以

$$\ddot{x} - 4\dot{x} + 4x = \ddot{u}\,\mathrm{e}^{2t}.$$

取 $u = Ct^2$，上式便是 $2C\mathrm{e}^{2t}$，所以 $2C = 48$，即 $C = 24$。（這個代換同時說明了為甚麼餘函數是 $(A + Bt)\mathrm{e}^{2t}$：我們需要 $\ddot{u} = 0$。）

因此通解（general solution）是

$$x = (A + Bt)\,\mathrm{e}^{2t} + 24t^2\mathrm{e}^{2t} + 3\mathrm{e}^{-2t}.$$

初始條件（*initial conditions*）。題目只給出 $x(0) = 0$ 和 $y(0) = 0$。二階方程還需要 $\dot{x}(0)$，而原來的第一條方程正好提供了它：$\dot{x}(0) = 4\big(x(0) - y(0)\big) = 0$。

由 $x(0) = 0$ 得 $A + 3 = 0$，所以 $A = -3$。微分得

$$\dot{x} = \big(B + 2A + 2Bt + 48t + 48t^2\big)\mathrm{e}^{2t} - 6\mathrm{e}^{-2t},$$

所以 $\dot{x}(0) = B + 2A - 6 = 0$，得 $B = 12$。因此

$$\boxed{x = \left(24t^2 + 12t - 3\right)\mathrm{e}^{2t} + 3\mathrm{e}^{-2t}.}$$

求 y

不必再解另一條微分方程，因為第一條方程給出 $y = x - \frac{1}{4}\dot{x}$。代入 $A = -3$ 和 $B = 12$，

$$\dot{x} = \left(48t^2 + 72t + 6\right)\mathrm{e}^{2t} - 6\mathrm{e}^{-2t},$$

所以

$$\begin{aligned} y &= \left(24t^2 + 12t - 3\right)\mathrm{e}^{2t} + 3\mathrm{e}^{-2t} - \left(12t^2 + 18t + \tfrac{3}{2}\right)\mathrm{e}^{2t} + \tfrac{3}{2}\mathrm{e}^{-2t} \\ &= \left(12t^2 - 6t - \tfrac{9}{2}\right)\mathrm{e}^{2t} + \tfrac{9}{2}\mathrm{e}^{-2t}. \end{aligned}$$

即

$$\boxed{y = \tfrac{3}{2}\left(8t^2 - 4t - 3\right)\mathrm{e}^{2t} + \tfrac{9}{2}\mathrm{e}^{-2t}.}$$

驗算

- 當 $t = 0$ 時：$y(0) = -\frac{9}{2} + \frac{9}{2} = 0$。這個條件從未直接用上，所以這是一項真正的驗算。
- 第二條方程我們只用來消去 y，它也應該成立。微分得

$$\dot{y} = \left(24t^2 + 12t - 15\right)\mathrm{e}^{2t} - 9\mathrm{e}^{-2t},$$

  而

$$x - 12\left(\mathrm{e}^{2t} + \mathrm{e}^{-2t}\right) = \left(24t^2 + 12t - 15\right)\mathrm{e}^{2t} - 9\mathrm{e}^{-2t}.$$

  兩者相符。
- 數值方面，用龍格－庫塔法（Runge–Kutta method）由 $(0,0)$ 出發逐步推算原來的方程組，得 $x(1) \approx 244.2449$ 和 $y(1) \approx 11.6926$。公式給出 $33\mathrm{e}^2 + 3\mathrm{e}^{-2}$ 和 $\frac{3}{2}\mathrm{e}^2 + \frac{9}{2}\mathrm{e}^{-2}$，數值與上述結果相同。

常見錯誤

- 對共振項試用 $Ct\,\mathrm{e}^{2t}$。根 2 是重根，所以我們需要 $t^2\mathrm{e}^{2t}$。
- 看不出第二個初始條件從何而來。$\dot{x}(0) = 0$ 這個值隱藏在第一條方程之中，再配合 $x(0) = y(0)$ 便可得出。
- 另外為 y 解一條二階方程。這樣會引入兩個新常數，之後還要把它們與 x 配對。利用 $y = x - \frac{1}{4}\dot{x}$ 便可避免這個麻煩。

STEP 之外

消去 y 是解線性方程組（linear system）的其中一種方法。大學的做法是把它寫成 $\dot{\mathbf{v}} = M\mathbf{v} + \mathbf{f}(t)$，其中 $M = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}$。這裏 M 有重特徵值（eigenvalue）2，但只有一個特徵向量（eigenvector），所以不能對角化（diagonalise），而 $t\,\mathrm{e}^{2t}$ 項正是由此而來。若爾當標準形（Jordan normal form）和矩陣指數（matrix exponential）e^{Mt} 可以把這一點說得精確。Boyce 與 DiPrima 的 *Elementary Differential Equations*（微分方程入門教科書）涵蓋線性方程組，包括重特徵值的情況。

詳細解答

STEP III 1989，第 9 題

純數學 • 冪級數求和

> **STEP III 1989，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/89-s3-q9

關鍵思路

每個級數(series)都是一個熟悉的冪級數(power series)$\sum a_r x^r$ 在某個特定的 x 處的取值，所以要做的就是認出它是哪一個。幾何級數（geometric series）$\frac{1}{1-x} = 1 + x + x^2 + \cdots$ 是頭兩個級數的源頭。對它求導，會帶出一個因子 r；對它積分，則會除以 r。至於第三個，奇數的連乘積除以階乘，正是指數為負分數的二項展開式（binomial expansion）所產生的形式。

解答

以下我們一直會用到：當 $|x| < 1$ 時，

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots = \sum_{r=1}^{\infty} x^{r-1}.$$

題目說明不必理會收斂（convergence）的問題，所以我們可以放心逐項求導、逐項積分。(我們用到的三個 x 值全都穩穩落在收斂區間（interval of convergence）之內，所以這樣做其實是合理的。)

第 (i) 部分

通項是 $r\left(\frac{1}{2}\right)^{r-1}$，所以我們要求的是 $\sum rx^{r-1}$ 在 $x = \frac{1}{2}$ 處的值。把幾何級數逐項求導，

$$\sum_{r=1}^{\infty} rx^{r-1} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right) = \frac{1}{(1-x)^2}.$$

代入 $x = \frac{1}{2}$：

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

第 (ii) 部分

通項是 $\frac{1}{r}\left(\frac{1}{2}\right)^{r-1}$。求導可以抵銷除以 r 這一步，所以這次我們改為積分。把幾何級數由 0 積分至 x，

$$\sum_{r=1}^{\infty} \frac{x^r}{r} = \int_0^x \frac{\mathrm{d}t}{1-t} = -\ln(1-x).$$

這裏的冪次高了一次，所以我們除以 x（其中 $x \neq 0$）：

$$\sum_{r=1}^{\infty} \frac{1}{r} x^{r-1} = -\frac{\ln(1-x)}{x}.$$

代入 $x = \frac{1}{2}$：

$$1 + \frac{1}{2} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2^2} + \cdots = -\frac{\ln \frac{1}{2}}{\frac{1}{2}} = 2\ln 2.$$

第 (iii) 部分

設

$$c_k = \frac{1 \times 3 \times \cdots \times (2k-1)}{k!}, \qquad c_0 = 1.$$

? **想一想** 你以前在哪裏見過奇數的連乘積除以 $k!$？

就在 $(1+y)^{-1/2}$ 的二項展開式裏。y^k 的係數是

$$\binom{-\frac{1}{2}}{k} = \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{2k-1}{2}\right)}{k!} = \frac{(-1)^k}{2^k} \cdot \frac{1 \times 3 \times \cdots \times (2k-1)}{k!} = \left(-\frac{1}{2}\right)^k c_k.$$

因此，取 $y = -2x$，正負號和 2 的冪便互相抵銷：

$$(1-2x)^{-1/2} = \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} (-2x)^k = \sum_{k=0}^{\infty} c_k x^k = 1 + x + \frac{1 \times 3}{2!}x^2 + \frac{1 \times 3 \times 5}{3!}x^3 + \cdots,$$

此式在 $|2x| < 1$ 時成立。

現在仔細看看題目給出的級數。它的通項是 $c_k \cdot \frac{1}{3^{k-1}} = 3c_k\left(\frac{1}{3}\right)^k$，而且由 $k = 2$ 開始（首項是 $\frac{1\times 3}{2!} \times \frac{1}{3}$）。所以它等於上面的級數在 $x = \frac{1}{3}$ 處的值，去掉 $k = 0$ 和 $k = 1$ 兩項之後，

再乘以 3：

$$\begin{aligned}\sum_{k=2}^{\infty} c_k \frac{1}{3^{k-1}} &= 3\left(\sum_{k=0}^{\infty} c_k \left(\tfrac{1}{3}\right)^k - c_0 - c_1 \cdot \tfrac{1}{3}\right)\\ &= 3\left(\left(1-\tfrac{2}{3}\right)^{-1/2} - 1 - \tfrac{1}{3}\right)\\ &= 3\left(\sqrt{3} - \tfrac{4}{3}\right)\\ &= 3\sqrt{3} - 4.\end{aligned}$$

這裏 $\left(\frac{1}{3}\right)^{-1/2} = \sqrt{3}$，而 $x = \frac{1}{3}$ 滿足 $|2x| < 1$。

驗算

- 第 (i) 部分：部分和（partial sums）1, 2, 2.75, 3.25, 3.5625, … 逐步向 4 靠近，而把 200 項加起來，在機器精度之內得到 4。
- 第 (ii) 部分：$2\ln 2 \approx 1.3863$，而部分和 1, 1.25, 1.3333, 1.3646, … 與此吻合。
- 第 (iii) 部分：各項為 0.5, 0.2778, 0.1620, …，把 400 項加起來得到 1.196152…，與 $3\sqrt{3} - 4 \approx 1.1962$ 相符。再做一個簡單的合理性檢查：答案必須大於首項 0.5，而且是正數；這就排除了忘記去掉 $k = 1$ 一項的情況（那樣會得到 $3\sqrt{3} - 3 \approx 2.196$，由於各項大致按幾何級數遞減，這個值太大了）。

常見錯誤

- 在 (ii) 中，忘記除以 x：$-\ln(1-x)$ 在 $x = \frac{1}{2}$ 處的值是 $\ln 2$，只有答案的一半。
- 在 (iii) 中，沒有留意到級數由 $k = 2$ 開始，或者漏掉了由 $\frac{1}{3^{k-1}} = 3 \cdot \frac{1}{3^k}$ 而來的額外因子 3。

STEP 之外

冪級數在其收斂半徑（radius of convergence）之內可以逐項求導和逐項積分，原因是冪級數在嚴格位於該半徑之內的每一個閉區間上都一致收斂（converges uniformly）。這是大學一年級分析學（analysis）的標準結果，也正因如此，「不必考慮收斂問題」這句話在這裏是穩妥的。指數並非整數的二項級數（binomial series），正是這一類冪級數。Abbott 的 *Understanding Analysis*（分析學入門教材）對一致收斂和冪級數有清晰的講解。

詳細解答

STEP III 1989，第 10 題

純數學 • 黎曼和

> **STEP III 1989，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/89-s3-q10

本題在 *Beyond STEP* 第 4.1 節 *甚麼是積分？* 中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch04/riemann.html

關鍵思路

連續整數之積（product of consecutive integers）是 r^5 的「下降冪」（falling power）版本，而且可以用裂項相消（telescoping）求和，與 r^5 的積分如出一轍：兩個相鄰的六項乘積之差，等於一個五項乘積的六倍。第 (i) 和第 (ii) 部分分別給出一個略大於和一個略小於 $\sum r^5$ 的乘積之和。第 (iii) 部分再把 x^5 的兩個黎曼和（Riemann sum）夾在這兩個界之間，而兩個界趨向同一個極限。

解答

第 (i) 部分

設 $P(r) = r(r+1)(r+2)(r+3)(r+4)(r+5)$。那麼 $P(r)$ 與 $P(r-1) = (r-1)r(r+1)(r+2)(r+3)(r+4)$ 有五個公共因子 $r(r+1)\cdots(r+4)$，所以

$$\begin{aligned} P(r) - P(r-1) &= r(r+1)(r+2)(r+3)(r+4)\big[(r+5) - (r-1)\big] \\ &= 6\,r(r+1)(r+2)(r+3)(r+4). \end{aligned}$$

由 $r = 1$ 加至 n，右邊裂項相消：

$$\begin{aligned} \sum_{r=1}^{n} r(r+1)(r+2)(r+3)(r+4) &= \tfrac{1}{6}\sum_{r=1}^{n}\big[P(r) - P(r-1)\big] = \tfrac{1}{6}\big[P(n) - P(0)\big] \\ &= \tfrac{1}{6}\,n(n+1)(n+2)(n+3)(n+4)(n+5), \end{aligned}$$

因為 $P(0) = 0$。（對 n 使用歸納法（induction）同樣可行；歸納步驟用的正是同一個因式分解。）

至於推論：若 $r \geq 1$，則對 $k = 1, 2, 3, 4$，每個因子都滿足 $r + k > r > 0$，所以

$$r^5 < r(r+1)(r+2)(r+3)(r+4) \qquad (r \geq 1).$$

把 $r = 1, \ldots, n$ 的這些嚴格不等式（strict inequality）相加，便得 $\sum_{r=1}^{n} r^5 < \frac{1}{6}n(n+1)(n+2)(n+3)(n+4)(n+5)$，正是所求。

第 (ii) 部分

現在我們需要一個位於 r^5 之下的乘積，所以把各因子向下平移：改用 $(r-4)(r-3)(r-2)(r-1)r$。設 $Q(r) = (r-4)(r-3)(r-2)(r-1)r(r+1)$，同樣的因式分解給出

$$\begin{aligned} Q(r) - Q(r-1) &= (r-4)(r-3)(r-2)(r-1)r\big[(r+1) - (r-5)\big] \\ &= 6\,(r-4)(r-3)(r-2)(r-1)r. \end{aligned}$$

由 $r = 0$ 加至 $n-1$：

$$\begin{aligned} \sum_{r=0}^{n-1} (r-4)(r-3)(r-2)(r-1)r &= \tfrac{1}{6}\big[Q(n-1) - Q(-1)\big] \\ &= \tfrac{1}{6}\,(n-5)(n-4)(n-3)(n-2)(n-1)n, \end{aligned}$$

因為 $Q(-1) = (-5)(-4)(-3)(-2)(-1)\cdot 0 = 0$。

? **想一想** 對和式中的每一個 r，$(r-4)(r-3)(r-2)(r-1)r < r^5$ 是否都成立？

不盡然，而這正是條件 $n > 1$ 發揮作用的地方。我們逐項比較：

- $r = 0$：兩邊都是 0，所以只有等式成立。
- $r = 1, 2, 3, 4$：乘積中有一個因子為零，所以乘積是 $0 < r^5$。
- $r \geq 5$：對 $k = 1, 2, 3, 4$，每個因子都滿足 $0 < r - k < r$，所以乘積 $< r^5$。

因此對所有 $r \geq 0$，都有 $(r-4)(r-3)(r-2)(r-1)r \leq r^5$，而且對每個 $r \geq 1$，不等式都是嚴格的。若 $n > 1$，由 $r = 0$ 至 $n-1$ 的和包含 $r = 1$ 這一項，所以最少有一個不等式是嚴格的，於是

$$\sum_{r=0}^{n-1} r^5 > \tfrac{1}{6}(n-5)(n-4)(n-3)(n-2)(n-1)n.$$

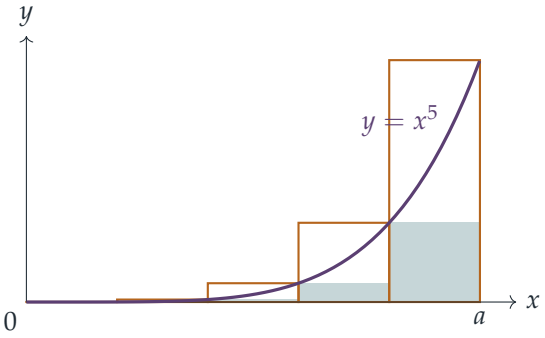
（當 $n = 1$ 時，兩邊都是 0，這就是需要該條件的原因。）

第 (iii) 部分

取 $\mathrm{f}(x) = x^5$，這是一個遞增函數（increasing function）。定義中的兩個和是

$$L_n = \sum_{r=0}^{n-1} \frac{a}{n}\left(\frac{ra}{n}\right)^5 = \frac{a^6}{n^6}\sum_{r=0}^{n-1} r^5, \qquad U_n = \sum_{r=1}^{n} \frac{a}{n}\left(\frac{ra}{n}\right)^5 = \frac{a^6}{n^6}\sum_{r=1}^{n} r^5.$$

當 $a > 0$ 時，它們分別是曲線下方和上方各長方形的面積之和：



（塗色部分：L_n；框線部分：U_n；圖中 $n = 5$。）不過，以下的代數運算並不需要 $a > 0$：對任何 $a \neq 0$ 都有 $a^6 > 0$，而 $a = 0$ 的情況顯然成立，因為每個和都是 0。

對 L_n 使用第 (ii) 部分（其中 $n > 1$），對 U_n 使用第 (i) 部分，

$$L_n > \frac{a^6}{6n^6}(n-5)(n-4)(n-3)(n-2)(n-1)n = \frac{a^6}{6}\prod_{k=1}^{5}\left(1-\frac{k}{n}\right),$$

$$U_n < \frac{a^6}{6n^6}n(n+1)(n+2)(n+3)(n+4)(n+5) = \frac{a^6}{6}\prod_{k=1}^{5}\left(1+\frac{k}{n}\right).$$

此外 $U_n - L_n = \dfrac{a^6}{n^6}\left(n^5 - 0^5\right) = \dfrac{a^6}{n} > 0$，所以綜合起來

$$\frac{a^6}{6}\prod_{k=1}^{5}\left(1-\frac{k}{n}\right) < L_n < U_n < \frac{a^6}{6}\prod_{k=1}^{5}\left(1+\frac{k}{n}\right).$$

當 $n \to \infty$ 時，每個括號都趨向 1，所以外側的兩個界都趨向 $\frac{1}{6}a^6$。根據夾逼定理（squeeze theorem，又稱 sandwich theorem），L_n 和 U_n 都收斂（converge），而且收斂至同一個極限 $\frac{1}{6}a^6$。所以定義中的兩個極限都存在而且相等，因此

$$\int_0^a x^5\,\mathrm{d}x = \tfrac{1}{6}a^6.$$

驗算

- 第 (i) 部分的小數值情況：當 $n = 1$ 時，$1\cdot 2\cdot 3\cdot 4\cdot 5 = 120 = \frac{1}{6}\cdot 720$；當 $n = 2$ 時，$120 + 720 = 840 = \frac{1}{6}\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7$。

- 第 (i) 和第 (ii) 部分的恆等式（identity），以及兩個不等式，對 200 以內的每個 n 都準確成立（已用電腦核對）。
- 取 $a = 1$ 和 $n = 1000$：$L_n \approx 0.166167$，$U_n \approx 0.167167$，分別位於 $\frac{1}{6} \approx 0.166667$ 的兩側，而且 $U_n - L_n = 0.001 = a^6/n$。

常見錯誤

- 聲稱對所有 $r \geq 0$ 都有 $(r-4)\cdots r < r^5$：在 $r = 0$ 時其實是等式，所以嚴格不等式要靠 $r = 1$ 這一項，亦即需要 $n > 1$。
- 只證明 L_n 和 U_n 有界（bounded），或只證明 $U_n - L_n \to 0$。定義要求兩個極限都*存在*；夾逼定理則一併給出存在性和極限值。

STEP 之外

第 (iii) 部分是黎曼（Riemann）積分（或稱達布（Darboux）積分）的一次親手實踐：若一個函數的下和（lower sum）與上和（upper sum）可以夾在一起，該函數便是可積（integrable）的；對單調函數（monotonic function）而言，兩者之差恆為 $(b-a)\,|\mathrm{f}(b) - \mathrm{f}(a)|/n$，所以閉區間（closed interval）上的每個單調函數都可積。第 (i) 和第 (ii) 部分的乘積分別是「上升階乘冪」（rising factorial power）和「下降階乘冪」（falling factorial power），它們令有限差分演算（calculus of finite differences）看起來與普通的微積分一模一樣。Abbott 的 *Understanding Analysis*（一本實分析入門書）正是以這種方式建立積分理論。

詳細解答

STEP III 1989，第 11 題

力學 • 球體的斜碰撞

STEP III 1989，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/89-s3-q11

關鍵思路

首先確定碰撞一刻的幾何關係：它決定了連心線（line of centres）的方向，其餘一切都由此而來。然後套用光滑（smooth）球體斜碰撞（oblique impact）而其中一個球受約束（constraint）時的標準方法。兩球之間的衝量（impulse）沿連心線作用。牛頓恢復定律（Newton's law of restitution）適用於沿該直線的速度分量。至於動量守恆（conservation of momentum），由於繩子或支點（pivot）本身也提供衝量，這裏並不成立；我們改用不受約束影響的那一條方程：有繩子時用 A 的水平動量，有桿時用對 O 的角動量（angular momentum）。

解答

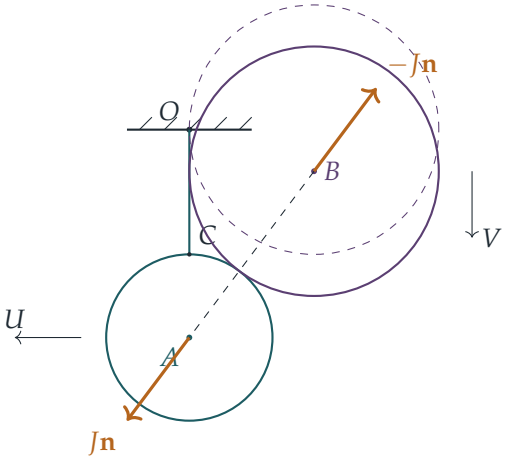
碰撞時的幾何關係

我們把兩個球都視為光滑而均勻（uniform），並把繩子視為輕而不可伸長（light and inextensible）。碰撞前繩子垂直懸掛。由於 C 位於直線 OA 上，我們有 $OA = 3a + 2a = 5a$，而 A 在 O 的正下方。

球 B 起初的位置使 $OB = 3a$ 為水平，然後垂直下落。設它下落的距離為 h。碰撞時兩球相切，所以 $AB = 2a + 3a = 5a$。A 與 B 的水平距離是 $3a$，所以兩者的垂直距離是 $\sqrt{25a^2 - 9a^2} = 4a$。因此 B 在 A 之上 $4a$ 處，亦即在 O 的水平線之下 a 處，故 $h = a$。

所以連心線構成一個 3–4–5 三角形：由 B 到 A，它沿水平方向（朝向 OA）走 $3a$，再向下走 $4a$。設 $\mathbf{n}$ 為由 B 指向 A 的單位向量（unit vector），並取 x 軸為水平，方向由 OA 指向 B。那麼

$$\mathbf{n} = \left(-\tfrac{3}{5}, -\tfrac{4}{5}\right).$$



虛線的球標示 B 的起始位置，實線的球則標示它在碰撞時的位置。

有繩子時的碰撞

設 J 為兩球之間衝量的大小，所以 A 受到 $J\mathbf{n}$，而 B 受到 $-J\mathbf{n}$。兩球都是光滑的，所以 $J\mathbf{n}$ 沿連心線作用，並通過 A。繩子的衝量沿 OC 作用，同樣通過 A。所以沒有任何東西令球 A 自轉，A 的運動就如同一個質點（particle），在以 O 為圓心、半徑為 $5a$ 的圓上運動。緊接碰撞之後繩子仍然垂直，所以 A 沿水平方向運動，速率為 U，方向為 x 的負方向。

*球 A，水平方向。*繩子不提供水平衝量，所以

$$3mU = \tfrac{3}{5}J, \qquad \text{即} \qquad J = 5mU.$$

（在垂直方向，繩子提供一個向上的衝量 $\frac{4}{5}J$，抵消了 $J\mathbf{n}$ 向下的部分。這個量是正數，所以繩子確實保持拉緊。）

*球 B。*碰撞前 $\mathbf{v}_B = (0, -V)$，它沿 $\mathbf{n}$ 的分量是 $\frac{4}{5}V$。衝量 $-J\mathbf{n}$ 只改變這個分量，並使它減少 $J/(25m)$：

$$\mathbf{v}'_B \cdot \mathbf{n} = \tfrac{4}{5}V - \frac{J}{25m} = \tfrac{4}{5}V - \tfrac{1}{5}U.$$

想一想 為甚麼我們不能直接沿連心線運用動量守恆？

繩子對 A 的衝量是垂直的，而它沿 $\mathbf{n}$ 有一個分量。所以沿 $\mathbf{n}$ 的總動量會改變。唯一沒有外來衝量作用的方向是水平方向，而那正是我們對 A 所用的方程。

*恢復定律。*沿 $\mathbf{n}$ 方向，靠近速率（speed of approach）是 $\frac{4}{5}V - 0$，因為 A 起初靜止。分離速率（speed of separation）是 $\mathbf{v}'_A \cdot \mathbf{n} - \mathbf{v}'_B \cdot \mathbf{n}$，其中 $\mathbf{v}'_A \cdot \mathbf{n} = (-U)(-\frac{3}{5}) = \frac{3}{5}U$。所以

$$\tfrac{3}{5}U - \left(\tfrac{4}{5}V - \tfrac{1}{5}U\right) = \tfrac{3}{4} \cdot \tfrac{4}{5}V \quad \Longrightarrow \quad \tfrac{4}{5}U = \tfrac{7}{5}V \quad \Longrightarrow \quad U = \tfrac{7}{4}V.$$

有桿時的碰撞

現在球與桿組成單一個剛體（rigid body），繞 O 轉動。支點可以提供任何方向的衝量，所以水平動量不再守恆。但支點的衝量作用於 O，所以我們對 O 取矩（take moments）。根據平行軸定理（parallel axis theorem），它對 O 的轉動慣量（moment of inertia）是

$$I = \tfrac{2}{5}(3m)(2a)^2 + 3m(5a)^2 = \tfrac{24}{5}ma^2 + 75ma^2 = \tfrac{399}{5}ma^2.$$

衝量 $J\mathbf{n}$ 的作用線通過 A，而它位於 O 之下 $5a$ 的深度。只有它的水平分量 $\frac{3}{5}J$ 對 O 有矩，而該矩為 $5a \cdot \frac{3}{5}J = 3aJ$。若剛體在緊接碰撞之後的角速率（angular speed）是 ω，則 A 以速率 $U' = 5a\omega$ 沿水平方向運動，並且

$$I\omega = 3aJ \quad\Longrightarrow\quad \tfrac{399}{5}ma^2 \cdot \frac{U'}{5a} = 3aJ \quad\Longrightarrow\quad J = \tfrac{133}{25}\,mU'.$$

關於 B 的方程和恢復定律的方程與之前完全相同，因為它們只涉及沿 $\mathbf{n}$ 的分量。所以

$$\tfrac{3}{5}U' - \left(\tfrac{4}{5}V - \frac{J}{25m}\right) = \tfrac{3}{5}V \quad\Longrightarrow\quad \left(\tfrac{3}{5} + \tfrac{133}{625}\right)U' = \tfrac{7}{5}V \quad\Longrightarrow\quad \tfrac{508}{625}U' = \tfrac{7}{5}V,$$

所以 $U' = \frac{875}{508}V$。把這個結果與 $U = \frac{7}{4}V$ 比較，

$$\frac{U'}{U} = \frac{875}{508} \cdot \frac{4}{7} = \frac{500}{508} = \frac{125}{127},$$

所以 A 在緊接碰撞之後的速率是 $125U/127$，正如題目所求。

答案比之前稍小，這是合理的。有繩子時，A 抵抗水平運動的表現就像一個質量為 $3m$ 的質點。有桿時，球還必須開始轉動，而它在 A 處的等效水平質量變成 $I/(5a)^2 = \frac{399}{125}m$，比 $3m$ 略大。

驗算

- 能量：取 $m = a = V = 1$，碰撞前的動能（kinetic energy）是 12.5。碰撞後，有繩子時是 11.625，有桿時約為 11.58。兩者都較小，在 $e < 1$ 時必然如此。
- 彈性極限情況：取 $e = 1$ 並用繩子，同樣的步驟給出 $\frac{4}{5}U = \frac{8}{5}V$，所以 $U = 2V$，而動能恰好守恆。這確認了正負號無誤。
- 用電腦以精確分數求解兩次碰撞，取 $a = m = V = 1$。結果是 $U = 7/4$ 和 $U' = 875/508$，得出比值 $125/127$。電腦亦核對了 B 的切向速度（tangential velocity）不變，以及 B 在接觸前下落的距離為 a。

常見錯誤

- 沿連心線運用總動量守恆。繩子或支點會提供外來衝量，而該衝量在這個方向上有分量。

- 忘記桿上的球會轉動。它對 O 的轉動慣量除了 $3m(5a)^2$ 之外，還需要 $\frac{2}{5}(3m)(2a)^2$ 這一項。
- 把連心線當作成 45° 角，或當作垂直。相切的條件決定了它構成一個 3–4–5 三角形。

STEP 之外

受約束的碰撞讓我們初次接觸剛體的衝擊運動（impulsive motion of rigid bodies）。一般的法則是：固定支點上的衝量，不會出現在對該支點的角動量方程之中。由此可引伸出撞擊中心（centre of percussion）的概念：敲擊一個裝有支點的物體時，若擊中這一點，支點便完全感受不到衝量，這就是板球棒有「甜蜜點」的原因。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教科書）正是以這個程度講解角動量和剛體衝量。

詳細解答

STEP III 1989，第 12 題

力學 · 旋轉參考系

STEP III 1989，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/89-s3-q12

關鍵思路

平面是光滑（smooth）的，而桿是輕（light）桿並可繞樞軸自由轉動，所以作用於 P 的水平力只能沿桿的方向。因此，P 的加速度在垂直於 AP 方向上的分量必定為零。只要求出這個分量，令它等於零，便得到關於 ϕ 的單擺方程（pendulum equation），再用慣常的「乘以 $\dot\phi$ 再積分」技巧，便可求出 $\dot\phi$。

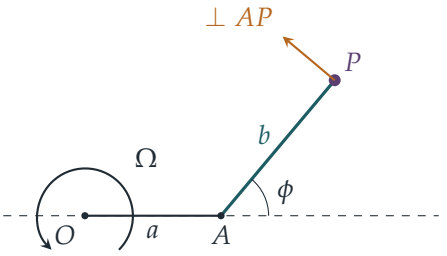
解答

模型與設定

把 P 看成光滑平面上的一個質點（particle），把桿看成輕而剛硬的桿，可繞 A 自由轉動。我們在一個固定的慣性參考系（inertial frame）中考慮問題。用複數表示平面上的位置，以 O 為原點，實軸取為 $t=0$ 時 OA 的方向，並取 $\Omega>0$ 為逆時針方向。由於 A 固定在平面上，它沿半徑為 a 的圓運動；而 AP 與 OA 的延長線成角 ϕ，該延長線本身已轉過 Ωt。因此，記 $\theta=\Omega t+\phi$，便有

$$z_A = a\,\mathrm{e}^{\mathrm{i}\Omega t}, \qquad z_P = a\,\mathrm{e}^{\mathrm{i}\Omega t} + b\,\mathrm{e}^{\mathrm{i}\theta}.$$

這裏 $\mathrm{e}^{\mathrm{i}\theta}$ 是沿 AP 方向的單位向量，而 $\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}$ 是垂直於 AP、指向 ϕ 增加方向的單位向量。



加速度的垂直分量

求導，其中 $\dot\theta = \Omega + \dot\phi$ 及 $\ddot\theta = \ddot\phi$（因為 Ω 是常數），得

$$\dot z_P = \mathrm{i}a\Omega\,\mathrm{e}^{\mathrm{i}\Omega t} + \mathrm{i}b(\Omega+\dot\phi)\,\mathrm{e}^{\mathrm{i}\theta},$$
$$\ddot z_P = -a\Omega^2\,\mathrm{e}^{\mathrm{i}\Omega t} + b\big(\mathrm{i}\ddot\phi - (\Omega+\dot\phi)^2\big)\mathrm{e}^{\mathrm{i}\theta}.$$

要把加速度分解為沿 AP 和垂直於它的分量，可寫 $\mathrm{e}^{\mathrm{i}\Omega t} = \mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\phi} = \mathrm{e}^{\mathrm{i}\theta}(\cos\phi - \mathrm{i}\sin\phi)$。於是

$$\ddot z_P = \big[-a\Omega^2\cos\phi - b(\Omega+\dot\phi)^2\big]\mathrm{e}^{\mathrm{i}\theta} + \big[\,b\ddot\phi + a\Omega^2\sin\phi\,\big]\,\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}.$$

所以，P 的加速度在垂直於 AP 方向（指向 ϕ 增加的一方）的分量是

$$\boxed{\,b\,\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} + a\Omega^2\sin\phi\,}$$

（第一項是我們熟悉的 $b\ddot\phi$，來自繞 A 的圓周運動。第二項來自 A 本身指向 O 的向心加速度（centripetal acceleration）$a\Omega^2$，再分解到垂直於桿的方向。）

求 $\mathrm{d}\phi/\mathrm{d}t$

? **想一想** 為甚麼加速度的這個分量必定為零？

作用於 P 的力有：它的重量和法向反作用力（normal reaction）（兩者都是鉛垂的，並且互相抵消）；沒有摩擦力（平面是光滑的）；以及桿所施的力。一根可繞 A 自由轉動的輕桿，只能沿自身的長度方向推或拉。所以在垂直於 AP 的方向上沒有水平力，由牛頓第二定律（Newton's second law）得

$$b\ddot\phi + a\Omega^2\sin\phi = 0.$$

這就是單擺方程。乘以 $\dot\phi$ 再積分，得

$$\tfrac{1}{2}b\dot\phi^2 = a\Omega^2\cos\phi + C.$$

初始條件。開始時 O, A, P 在同一直線上。由於 $b > a$，P 不可能位於 O 與 A 之間，我們取 P 在 A 之外的一方，所以在 $t = 0$ 時 $\phi = 0$。這時兩個指數函數的值都是 1，而

$$\dot{z}_P(0) = \mathrm{i}\big[(a+b)\Omega + b\dot{\phi}(0)\big],$$

它垂直於 OP（即實軸）。留意 $(a+b)\Omega$ 正是平面上位於 P 正下方那一點的速率，而 $b\dot{\phi}(0)$ 則是 P 相對於平面的速率。題目印出的速率是 $(a+2\sqrt{ab})\Omega$，但這看來是 $(a+b+2\sqrt{ab})\Omega = (\sqrt{a}+\sqrt{b})^2\Omega$ 的誤印，因為後者才能得出題目所述的結果（見文末的備註）。採用這個值，P 相對於平面的速率是 $2\sqrt{ab}\,\Omega$，所以

$$\dot{\phi}(0) = 2\Omega\sqrt{a/b}.$$

把 $\phi = 0$ 代入能量方程，得 $\frac{1}{2}b \cdot 4a\Omega^2/b = a\Omega^2 + C$，所以 $C = a\Omega^2$，而

$$\tfrac{1}{2}b\dot{\phi}^2 = a\Omega^2(1+\cos\phi) = 2a\Omega^2\cos^2\tfrac{\phi}{2}, \qquad \text{所以} \qquad \dot{\phi}^2 = \frac{4a\Omega^2}{b}\cos^2\tfrac{\phi}{2}.$$

開始時 $\dot{\phi} > 0$，而 $\dot{\phi}$ 只有經過零才能改變正負號，這需要 $\cos\frac{\phi}{2} = 0$，即 $\phi = \pm\pi$。所以只要 $-\pi < \phi < \pi$，我們便取正根：

$$\boxed{\frac{\mathrm{d}\phi}{\mathrm{d}t} = 2\Omega\sqrt{\frac{a}{b}}\cos\frac{\phi}{2}}$$

P 永不越過 OA

? **想一想** ϕ 難道不可以慢慢爬升到 π，然後越過它嗎？

只有當 ϕ 是 π 的倍數時，P 才會在直線 OA 上（當 $\phi = \pi$ 時，它位於 O 之外、距離為 $b-a$ 的地方）。由於當 $0 \le \phi < \pi$ 時 $\dot{\phi} > 0$，ϕ 由 0 開始遞增，而回到這條直線上的唯一途徑就是到達 $\phi = \pi$。我們解出這個方程，從而證明這種情況永不發生。分離變數（separating the variables），並設 $k = \sqrt{a/b}$，得

$$\int_0^{\phi} \sec\tfrac{u}{2}\,\mathrm{d}u = 2k\Omega t \quad\Longrightarrow\quad \ln\Big(\sec\tfrac{\phi}{2} + \tan\tfrac{\phi}{2}\Big) = k\Omega t.$$

設 $s = \sin\frac{\phi}{2}$，則 $\sec\frac{\phi}{2} + \tan\frac{\phi}{2} = \frac{1+s}{\sqrt{1-s^2}} = \sqrt{\frac{1+s}{1-s}}$，所以左邊等於 $\frac{1}{2}\ln\frac{1+s}{1-s} = \operatorname{artanh} s$。因此

$$\sin\frac{\phi}{2} = \tanh\Big(\Omega t\sqrt{\tfrac{a}{b}}\Big).$$

對於每個 $t > 0$，右邊都嚴格介乎 0 與 1 之間，所以對所有 $t > 0$ 都有 $0 < \phi < \pi$。於是 P 始終嚴格地留在直線 OA 的同一側。當 $t \to \infty$ 時，它趨近這條直線（趨近 O 之外的那一點），但永遠不會到達，所以 P 永不越過 OA。 ■

（不解方程的做法：$\phi = \pi, \dot{\phi} = 0$ 是單擺方程的一個平衡點（equilibrium），而在它附近，$\psi = \pi - \phi$ 滿足 $\dot{\psi} \approx -k\Omega\psi$，所以 ψ 以指數形式衰減，永不為零。）

關於印出的速率的備註。如果我們按字面把速率取為 $(a+2\sqrt{ab})\Omega$，則 $\dot\phi(0) = (2k-1)\Omega$，而同樣的積分給出 $\dot\phi^2 = \Omega^2\left[(2k-1)^2 - 4k^2\sin^2\frac{\phi}{2}\right]$。這並不是一個簡潔的式子。這時 P 會繞 $\phi = 0$ 來回擺動（每擺一次都越過 OA 的延長線），如果 $b < 16a$；又或者轉足一整圈，越過 OA，如果 $b > 16a$。所以印出的數值不可能是出題者的原意。

驗算

- 令 $a = 0$，A 便落在 O 上，這時 $\dot\phi \equiv 0$：P 只是隨着平面以速率 $b\Omega$ 一起轉動。這是對的，因為初始速率變成了 $b\Omega$。
- 單位：$\Omega\sqrt{a/b}$ 是一個變率，所以它的單位是 s^{-1}，正如 $\dot\phi$ 應有的單位。
- 我們以數值方法（RK4）逐步求解 $b\ddot\phi = -a\Omega^2\sin\phi$，取 $a = 1, b = 3, \Omega = 1.3$。在 $t = 2$ 時，得出 $\phi \approx 2.2644$ 和 $\dot\phi \approx 0.6375$，與 $2\arcsin\tanh(\Omega t\sqrt{a/b})$ 和 $2\Omega\sqrt{a/b}\cos\frac{\phi}{2}$ 相符，準確至 10 位有效數字。在 $t = 10$ 時，$\phi \approx 3.1394 < \pi$。

常見錯誤

- 忘記了參考系本身對 P 的初始速度所貢獻的 $(a+b)\Omega$，把題目給出的速率當成相對於平面的速率。
- 把 P 的加速度只寫成繞 A 作圓周運動的加速度。樞軸 A 本身也在加速，$a\Omega^2\sin\phi$ 這一項正是由此而來。
- 把 $a\Omega^2\sin\phi$ 的正負號弄錯。如果符號錯了，$\phi = 0$ 便會是不穩定的，這不符合物理實際，因為「離心」拉力是背向 O 的。

STEP 之外

從旋轉的平面上看，P 是一個單擺，懸掛在大小為 $a\Omega^2$、方向沿 OA 延長線的均勻「重力」之中。科里奧利力（Coriolis force）沿桿的方向作用，所以它不做功，可以不予考慮。題目的初始條件令運動恰好落在 分界線（*separatrix*）上：這是來回擺動與越過頂點之間的臨界情況，此時單擺需要無限長的時間才到達倒立位置。旋轉參考系（rotating frames）、虛擬力（fictitious forces）以及這類相平面（phase plane）圖像，可參閱 Kleppner and Kolenkow, *An Introduction to Mechanics*（一本大學力學入門教科書）。

詳細解答

STEP III 1989，第 13 題

力學 • 簡諧運動

> **STEP III 1989，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/89-s3-q13

關鍵思路

每條弦線都由 O（在桌面之下）出發，向上穿過所屬的小孔，再沿桌面連到 P。桌下的一段長度恆為 a，正好等於自然長度（natural length），所以每條弦線的伸長量（extension）就是 P 與該小孔之間的距離。這樣每個張力（tension）都與一個向量成正比；用位置向量（position vector）來表示，五個力只需一行便可加起來，因為五個小孔對稱排列，它們的位置向量之和為零。

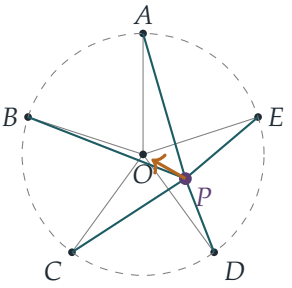
解答

建立模型

把 P 視為光滑桌面上的質點（particle），弦線視為輕彈性弦線（light elastic string），小孔視為光滑，因此每條弦線上各處的張力相同。取桌面上的 O 為原點，設 $\mathbf{r}$ 為 P 的位置向量，並設 $\mathbf{a}_1, \ldots, \mathbf{a}_5$ 為 $A, \ldots, E$ 的位置向量，每個的長度均為 a。

穿過第 i 個小孔的弦線由兩段組成：一段是由 O（在桌面之下，位於中心的正下方）到小孔的垂直部分，長度為 a；另一段在桌面上，長度為 $|\mathbf{a}_i - \mathbf{r}|$。弦線的總長度是 $a + |\mathbf{a}_i - \mathbf{r}|$，所以伸長量是 $|\mathbf{a}_i - \mathbf{r}|$。這個量永不為負，因此無論 P 在哪裏，每條弦線都是拉緊的（或當 P 位於小孔時剛好鬆弛）：不會有弦線鬆掉而令方程改變。根據虎克定律（Hooke's law），張力為 $\frac{\lambda}{a}|\mathbf{a}_i - \mathbf{r}|$，而在桌面上它把 P 拉向小孔，方向為單位向量 $\frac{\mathbf{a}_i - \mathbf{r}}{|\mathbf{a}_i - \mathbf{r}|}$。所以第 i 條弦線作用於 P 的力是

$$\frac{\lambda}{a}|\mathbf{a}_i - \mathbf{r}| \cdot \frac{\mathbf{a}_i - \mathbf{r}}{|\mathbf{a}_i - \mathbf{r}|} = \frac{\lambda}{a}(\mathbf{a}_i - \mathbf{r}).$$



藍綠色的線是從上方俯視時桌面上的五條弦線；琥珀色箭頭表示作用於 P 的合力（resultant force），它正正指向 O。

第一部分

作用於 P 的垂直方向的力（重量與法向反作用力）互相平衡。把五個水平方向的力相加，

$$\mathbf{F} = \frac{\lambda}{a}\sum_{i=1}^{5}(\mathbf{a}_i - \mathbf{r}) = \frac{\lambda}{a}\Big(\sum_{i=1}^{5}\mathbf{a}_i\Big) - \frac{5\lambda}{a}\mathbf{r}.$$

? **想一想** 為甚麼 $\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4 + \mathbf{a}_5 = \mathbf{0}$？

把整個圖形繞 O 旋轉 72°，五個小孔只是互相置換，所以總和 $\mathbf{s} = \sum \mathbf{a}_i$ 保持不變。但旋轉 72° 後仍然不變的向量只有零向量，所以 $\mathbf{s} = \mathbf{0}$。（換句話說，五次單位根（fifth roots of unity）之和為零。）

因此由牛頓第二定律（Newton's second law）可得

$$m\ddot{\mathbf{r}} = -\frac{5\lambda}{a}\mathbf{r}, \qquad \text{即} \qquad \ddot{\mathbf{r}} = -\omega^2\mathbf{r} \quad \text{其中} \quad \omega^2 = \frac{5\lambda}{am}.$$

P 由 F 從靜止開始運動，所以 $\mathbf{r}(0) = \overrightarrow{OF}$ 及 $\dot{\mathbf{r}}(0) = \mathbf{0}$。解為 $\mathbf{r} = \cos(\omega t)\,\overrightarrow{OF}$，所以 P 沿直線 FO 來回運動（去到 F 關於 O 的反射點再折返），而它相對於 O 的位移滿足簡諧運動（simple harmonic motion）的方程。週期（period）為

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{am}{5\lambda}},$$

這正是所需的結果。（注意這對任何起點 F 都成立，並不限於小位移：力的規律是完全線性的。）

最後部分

現在弦線 PAO 的彈性模量（modulus）為 $k\lambda$，所以它作用於 P 的力是 $\frac{k\lambda}{a}(\mathbf{a}_1 - \mathbf{r})$。利用 $\mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4 + \mathbf{a}_5 = -\mathbf{a}_1$，

$$\begin{aligned}\mathbf{F} &= \frac{k\lambda}{a}(\mathbf{a}_1 - \mathbf{r}) + \frac{\lambda}{a}\sum_{i=2}^{5}(\mathbf{a}_i - \mathbf{r}) \\ &= \frac{k\lambda}{a}(\mathbf{a}_1 - \mathbf{r}) + \frac{\lambda}{a}(-\mathbf{a}_1 - 4\mathbf{r}) \\ &= \frac{\lambda}{a}\big[(k-1)\mathbf{a}_1 - (k+4)\mathbf{r}\big] = -\frac{(k+4)\lambda}{a}\,(\mathbf{r} - \mathbf{r}_0)\,, \qquad \mathbf{r}_0 = \frac{k-1}{k+4}\,\mathbf{a}_1.\end{aligned}$$

所以平衡位置（equilibrium position）是 OA 上的點 $\mathbf{r}_0$（該處 $\mathbf{F} = \mathbf{0}$），並且

$$\ddot{\mathbf{r}} = -\frac{(k+4)\lambda}{am}(\mathbf{r} - \mathbf{r}_0).$$

若把 P 由 $\mathbf{r}_0$ 沿 OA 移開，然後從靜止釋放，則 $\mathbf{r} - \mathbf{r}_0$ 起初與 $\mathbf{a}_1$ 平行，加速度又恆與 $\mathbf{r} - \mathbf{r}_0$ 平行，而初速度為零。所以與之前完全一樣，$\mathbf{r} - \mathbf{r}_0 = \cos(\omega' t)(\mathbf{r}(0) - \mathbf{r}_0)$：$P$ 沿 OA 作直線運動，以 $\mathbf{r}_0$ 為中心作簡諧運動，角頻率（angular frequency）為 ω'，其中

$$\omega'^2 = \frac{(k+4)\lambda}{am}.$$

週期為 $T/2$，當且僅當 $\omega' = 2\omega$，即 $\omega'^2 = 4\omega^2$：

$$\frac{(k+4)\lambda}{am} = \frac{20\lambda}{am} \quad\Longrightarrow\quad k + 4 = 20 \quad\Longrightarrow\quad k = 16.$$

當 $k = 16$ 時，平衡位置是 $\mathbf{r}_0 = \frac{15}{20}\mathbf{a}_1$，即由 O 到 A 的四分之三處。（要令模型成立，振幅（amplitude）必須小於 $\frac{1}{4}a$，使 P 不會到達 A 處的小孔。）

驗算

- 代入 $k = 1$ 便回到第一部分：$\mathbf{r}_0 = \mathbf{0}$ 及 $\omega'^2 = 5\lambda/(am)$。
- 單位：λ 是一個力，所以 $\lambda/(am)$ 的單位是 $\mathrm{N/(m\,kg)} = \mathrm{s^{-2}}$，而 T 是時間。
- 我們在 1000 個隨機位置把五個張力向量相加（取 $a = 1.3$、$m = 0.7$、$\lambda = 2.1$），結果恰好是 $-5\lambda\mathbf{r}/a$；而當 $k = 16$ 時恰好是 $-20\lambda(\mathbf{r} - \mathbf{r}_0)/a$。由 $(0.4, -0.3)$ 出發，以數值方法逐步計算運動，P 在時間 $T \approx 1.85$ 後回到起點，並一直留在通過 O 的直線上；當 $k = 16$ 時由 $\mathbf{r}_0 + 0.1a\,\hat{\mathbf{a}}_1$ 出發，則在 $T/2$ 後回到起點。

常見錯誤

- 把伸長量當作 $|\mathbf{a}_i - \mathbf{r}| - a$ 或 $|\mathbf{r}|$：自然長度 a 剛好全部用於桌下的那一段。

- 用角度和三角學來分解五個張力。這樣做是可以的，但用向量的話，各項一眼便可看出互相抵消。
- 假設新的振動仍以 O 為中心：較硬的弦線會把平衡位置拉向 A。

STEP 之外

二維空間中的力 $\mathbf{F} = -c\,\mathbf{r}$ 給出各向同性諧振子（isotropic harmonic oscillator）：一般的運動（初速度帶有橫向分量時）是以平衡點為中心的橢圓，而且每條軌道的週期都相同。它與平方反比定律（inverse-square law）是僅有的兩種令每條有界軌道都閉合的有心力（central force），這就是伯特蘭定理（Bertrand's theorem）。Kleppner 與 Kolenkow 的 *An Introduction to Mechanics*（力學入門教材）正是以這種向量方式處理振子和有心力。

詳細解答

STEP III 1989，第 14 題

力學 • 滑輪與衝量

> **STEP III 1989，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/89-s3-q14

關鍵思路

第 (i) 部分：把圓盤切成許多薄圓環，每個圓環上的質量全都與圓心相距同一距離，因此質量和轉動慣量（moment of inertia）都只是對 r 的單一積分。第 (ii) 部分：整個系統只有一個自由度（degree of freedom），即 D 的深度，所以一條運動方程（equation of motion）便可求出加速度。當 B 突然被截停時，繩的約束（constraint）立刻告訴我們 D 此後可以運動得多快，而衝量（impulse）正是它的動量（momentum）的改變。

解答

第 (i) 部分

（第 (i) 部分所指的軸是通過圓心 O 的軸；試卷上的「C」是印刷錯誤。）設與 O 相距 r 處的密度為 $\rho = kr$，其中 k 為常數。半徑 r 與 $r + \mathrm{d}r$ 之間的薄圓環面積為 $2\pi r\,\mathrm{d}r$，故質量為 $2\pi k r^2\,\mathrm{d}r$，而且環上每一點與軸的距離都是 r。因此

$$m = \int_0^a 2\pi k r^2\,\mathrm{d}r = \frac{2\pi k a^3}{3}, \qquad I = \int_0^a r^2 \cdot 2\pi k r^2\,\mathrm{d}r = \frac{2\pi k a^5}{5}.$$

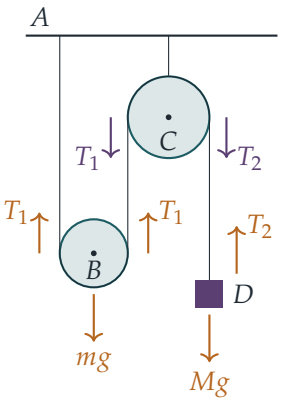
兩式相除，得 $I/m = \frac{3}{5}a^2$，所以

$$I = \tfrac{3}{5}ma^2,$$

這正是所要證明的結果。（它比均勻圓盤的 $\frac{1}{2}ma^2$ 大，這是合理的：較多的質量集中在邊緣附近。）

第 (ii) 部分：B 被截停前的運動

模型：B 是質量為 m 的光滑滑輪（pulley），C 是 (i) 中那種圓盤，質量為 m、半徑為 a，所以它繞其固定軸的轉動慣量是 $\frac{3}{5}ma^2$。繩是輕而不可伸長的，並且在 C 上不打滑。由於 B 是光滑的，AB 和 BC 兩段繩的張力（tension）T_1 相同。把 CD 段繩的張力記為 T_2；由於 C 是粗糙的，T_2 不一定等於 T_1。



設 D 以加速度 f 向下運動。支承 B 的兩段繩各自縮短的長度，是 CD 伸長量的一半，所以 B 以加速度 $\frac{1}{2}f$ 上升；又因為繩在 C 上不打滑，C 的角加速度（angular acceleration）是 f/a。運動方程為：

$$\begin{aligned} D\text{ (向下)}: &\quad Mg - T_2 = Mf, \\ B\text{ (向上)}: &\quad 2T_1 - mg = \tfrac{1}{2}mf, \\ C\text{ (對其軸取矩)}: &\quad (T_2 - T_1)a = \tfrac{3}{5}ma^2 \cdot \frac{f}{a}. \end{aligned}$$

把 $T_2 = Mg - Mf$ 和 $T_1 = \frac{1}{2}mg + \frac{1}{4}mf$ 代入最後一條方程，

$$Mg - Mf - \tfrac{1}{2}mg - \tfrac{1}{4}mf = \tfrac{3}{5}mf \quad\Longrightarrow\quad f\left(M + \tfrac{17}{20}m\right) = g\left(M - \tfrac{1}{2}m\right),$$

所以

$$f = \frac{10(2M - m)g}{20M + 17m}.$$

這是常數，而且正因為 $m < 2M$ 才是正數：題目給出這個條件，原因就在這裏。它說明 D 確實會由靜止開始下降。因此 t 秒後，D 向下運動的速率為

$$v = ft = \frac{10(2M - m)gt}{20M + 17m},$$

B 以 $\frac{1}{2}v$ 上升，而 C 以角速率 v/a 轉動。

（用能量驗算可得相同的 f：當 D 下降了 x 時，動能（kinetic energy）為 $\frac{1}{2}\left(M + \frac{1}{4}m + \frac{3}{5}m\right)\dot{x}^2$，而失去的勢能（potential energy）為 $(M - \frac{1}{2}m)gx$。）

第 (ii) 部分：衝量

> **想一想** B 被截停之後，D 還能運動嗎？

只要繩保持拉緊，D 的速率總是 B 的速率的兩倍，因為 A 是固定的，而 AB 和 BC 的長度每次都改變相同的量。一旦 B 靜止，這兩段的長度都固定了，於是再沒有繩可以繞過 C，D 也必須停下來，C 亦然。繩確實保持拉緊：D 正在遠離 C，所以它把 CD 拉緊；而仍在轉動的 C 則從 BC 一側把繩往上扯，所以那一段也被拉緊。（事實上，算出來的衝擊張力（impulsive tension）在 CD 中是 Mv，在 BC 中是 $(M+\frac{3}{5}m)v$，兩者皆為正數，印證了這一點。）

所以 D 的速度由向下的 v 降至零，而（繩）作用於 D 的衝量等於它的動量改變：

$$\text{衝量施於} D = Mv = \frac{10M(2M-m)gt}{20M+17m} \quad \text{，方向鉛垂向上。}$$

驗算

- 若 $m \to 0$（輕滑輪），則 $f \to g$，衝量為 Mgt：D 只是自由下落，一如所料。
- 若 $m = 2M$，衝量為零：系統處於平衡狀態，根本不會開始運動。
- 取 $M = 3$、$m = 2.5$、$g = 9.8$、$t = 1.3$，以數值方法逐步模擬運動，得 $Mv \approx 13.0507$，與公式吻合，而所得的動能亦與失去的勢能相符。用數值積分算出的轉動慣量比值是 0.6000。

常見錯誤

- 把 C 兩邊的張力當作相同。滑輪 C 有質量而且粗糙，正是兩邊張力之差令它轉動。
- 忘記 B 的速率只有 D 的一半；這使有效質量中出現的是 $\frac{1}{4}m$，而不是 m。
- 只給出衝量的大小而不說明方向：它向上作用於 D。

STEP 之外

用單一坐標把動能寫成 $\frac{1}{2}M_{\text{有效}}\dot{x}^2$，正是拉格朗日力學（Lagrangian mechanics）的起點；在那裏，像滑輪幾何這樣的約束從一開始就已內置，張力根本不會出現。像截停 B 這類突然施加的約束，也可用同樣的方法，以衝量方程處理。Kleppner and Kolenkow, *An Introduction to Mechanics*（力學入門教科書）以這種風格講述剛體轉動（rigid-body rotation）和衝量。

詳細解答

STEP III 1989，第 15 題

概率與統計 • 中位數的分佈

> **STEP III 1989，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/89-s3-q15

關鍵思路

$2n+1$ 個數值的中位數（median）落在 z 附近，當且僅當其中一個數值落在 z 處的一個極小區間內，另有 n 個數值在它之下、n 個在它之上。第 (i) 及第 (ii) 部分給出了「全部在下」和「全部在上」這兩部分的概率，再數一數有多少種方法選定哪些數值擔當哪個角色，便得到密度。之後，利用密度的積分等於 1 這個事實，不費吹灰之力就能求出該積分；把同一技巧用於以 $n+1$ 代替 n 的情況，便得到方差。

解答

在整個解答中，X 在 $[-c,c]$ 上的密度為 $\frac{1}{2c}$，所以對於 $-c \le k \le c$，

$$\mathrm{P}(X > k) = \frac{c-k}{2c}, \qquad \mathrm{P}(X < k) = \frac{c+k}{2c}.$$

第 (i) 部分

這 n 個數值互相獨立（independent），所以把概率相乘：

$$\mathrm{P}(\text{全部 } n \text{ 個數值均大於} k) = \left(\frac{c-k}{2c}\right)^n.$$

第 (ii) 部分

同理，

$$\mathrm{P}(\text{全部 } n \text{ 個數值均小於} k) = \left(\frac{c+k}{2c}\right)^n.$$

中位數的密度

設 $X_1, \ldots, X_{2n+1}$ 為樣本，Z 為其中位數，即第 $(n+1)$ 小的數值。由於 X 是連續的，出現相同數值的概率為 0，所以我們不予理會。每個 X_i 都在 $[-c, c]$ 內，所以 Z 也在其中。

取 $(-c, c)$ 內的 z 和一個小的 $\delta z > 0$。略去 $(\delta z)^2$ 階的項後，事件 $z < Z \le z + \delta z$ 發生，當且僅當

- 其中一個數值落在 $(z, z+\delta z]$ 內，
- 其餘數值中有 n 個小於 z，並且
- 餘下的 n 個數值大於 $z + \delta z$。

（有兩個或以上數值落在這個小區間內的情況，其概率為 $(\delta z)^2$ 階，在取極限時無關重要。）

想一想 把這 $2n+1$ 個數值分配到這三個角色，共有多少種方法？

落在小區間內的數值有 $2n+1$ 種選擇，然後從其餘數值中選出 n 個放在 z 之下，有 $\binom{2n}{n}$ 種方法。於是共有

$$(2n+1)\binom{2n}{n} = \frac{(2n+1)!}{(n!)^2}$$

種分配方法。這些分配互斥（mutually exclusive），而且每一種的概率相同。對於一種固定的分配，由獨立性以及第 (i) 和第 (ii) 部分（取 $k = z$），其概率為

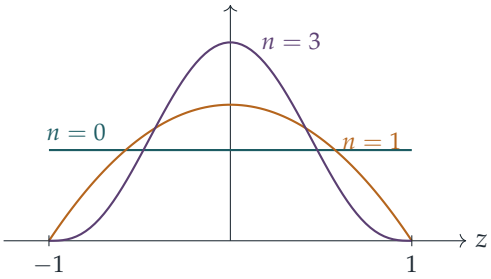
$$\frac{\delta z}{2c}\left(\frac{c+z}{2c}\right)^n \left(\frac{c-z-\delta z}{2c}\right)^n.$$

對所有分配求和，除以 δz，再令 $\delta z \to 0$，便得 Z 的密度為

$$\begin{aligned} g(z) &= \frac{(2n+1)!}{(n!)^2} \cdot \frac{1}{2c} \cdot \frac{(c+z)^n (c-z)^n}{(2c)^{2n}} \\ &= \frac{(2n+1)!}{(n!)^2 (2c)^{2n+1}} (c^2 - z^2)^n, \qquad -c \le z \le c, \end{aligned}$$

其中用到 $(c+z)(c-z) = c^2 - z^2$。這就是所要求的結果。

下圖顯示 $c = 1$ 時的密度。當 $n = 0$ 時，「中位數」只是單獨一個數值，服從均勻分佈（uniform distribution）。隨着 n 增大，中位數愈來愈集中在 0 附近。



積分

概率密度（probability density）在其取值範圍上的積分等於 1，所以

$$\frac{(2n+1)!}{(n!)^2(2c)^{2n+1}}\int_{-c}^{c}(c^2-z^2)^n\,\mathrm{d}z=1,$$

因此

$$\int_{-c}^{c}(c^2-z^2)^n\,\mathrm{d}z=\frac{(n!)^2(2c)^{2n+1}}{(2n+1)!}.$$

這對每個整數 $n\geq 0$ 都成立，稍後我們就會用到這一點。

平均值與方差

密度 g 是 z 的偶函數（even function），所以 $zg(z)$ 是奇函數，它在對稱區間 $[-c,c]$ 上的積分為零：

$$\mathrm{E}(Z)=0.$$

因此 $\mathrm{var}(Z)=\mathrm{E}(Z^2)$。記 $K_n=\dfrac{(2n+1)!}{(n!)^2(2c)^{2n+1}}$ 及 $I_n=\displaystyle\int_{-c}^{c}(c^2-z^2)^n\,\mathrm{d}z=\frac{1}{K_n}$。

想一想 怎樣才能不展開而求出 $z^2(c^2-z^2)^n$ 的積分？

寫成 $z^2=c^2-(c^2-z^2)$。於是

$$\begin{aligned}\mathrm{E}(Z^2)&=K_n\int_{-c}^{c}\big(c^2-(c^2-z^2)\big)(c^2-z^2)^n\,\mathrm{d}z\\&=K_n\big(c^2I_n-I_{n+1}\big)=c^2-\frac{K_n}{K_{n+1}},\end{aligned}$$

其中用到 $K_nI_n=1$ 和 $I_{n+1}=1/K_{n+1}$，後者就是以 $n+1$ 代替 n 的積分結果。現在

$$\frac{K_n}{K_{n+1}}=\frac{(2n+1)!}{(2n+3)!}\cdot\frac{((n+1)!)^2}{(n!)^2}\cdot\frac{(2c)^{2n+3}}{(2c)^{2n+1}}=\frac{(n+1)^2\cdot 4c^2}{(2n+2)(2n+3)}=\frac{2(n+1)c^2}{2n+3}.$$

因此

$$\mathrm{var}(Z)=c^2\left(1-\frac{2n+2}{2n+3}\right)=\frac{c^2}{2n+3}.$$

驗算

- 當 $n = 0$ 時，中位數就是 X 的一個數值。密度變成 $\frac{1}{2c}$，積分為 $2c$，方差為 $\frac{c^2}{3}$，正是 $[-c, c]$ 上均勻分佈的方差。
- 當 $n = 1$ 及 $c = 1$ 時，積分公式給出 $\int_{-1}^{1}(1 - z^2)\,\mathrm{d}z = \frac{1 \cdot 8}{6} = \frac{4}{3}$，結果正確。
- 以 10^5 個大小為 5 的樣本（$n = 2$）並取 $c = 1.7$ 進行模擬，得到中位數的方差為 0.415，而 $c^2/7 \approx 0.413$。

常見錯誤

- 忘記了選擇中間數值所帶來的因子 $2n + 1$，只用了 $\binom{2n}{n}$。
- 漏掉中間數值本身的密度所帶來的因子 $\frac{1}{2c}$，結果得到的冪是 $(2c)^{2n}$ 而不是 $(2c)^{2n+1}$。
- 用二項式展開 $(c^2 - z^2)^n$ 來求 $\mathrm{E}(Z^2)$。這樣做是可行的，但所得的和很難化簡。

STEP 之外

中位數是樣本的順序統計量（order statistics）之一。同樣的計數論證可以給出任何連續分佈的 m 個數值中第 r 小者的密度，即 $\frac{m!}{(r-1)!\,(m-r)!}F^{r-1}(1 - F)^{m-r}f$。對於均勻分佈，這些都是貝塔分佈（beta distributions），而你所推導出的積分正是貝塔函數 $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ 的一個特例。方差 $c^2/(2n + 3)$ 以 $1/n$ 的速度縮小，這是邁向樣本中位數的漸近正態性（asymptotic normality）的第一步。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（概率論入門）對順序統計量和貝塔分佈有很好的講解。

詳細解答

STEP III 1989，第 16 題

概率與統計 • 概率與假設檢驗

STEP III 1989，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/89-s3-q16

關鍵思路

三個「是／否」特徵把人口分成八格，而那八項陳述不過是關於各格比例的線性方程。先填綠眼的各格（那裏的條件最具體），再填淡褐眼的各格。之後的統計部分全靠一個數，即一個魯里坦尼亞人是窄頭的概率，並問在這個數值之下，觀察所得是否令人意外。

解答

試卷的印刷本有三處筆誤。我們把 (iv) 中的「arrow-headed」讀作 *narrow-headed*（窄頭），把 (vi) 中的「fair-headed」讀作 *fair-haired*（金髮），把 (viii) 中的「hazel-haired」讀作 *hazel-eyed*（淡褐眼）。這是唯一合理的讀法，因為三個特徵分別是頭髮（金或深色）、眼睛（綠或淡褐）和頭形（窄或闊）。

窄頭者所佔的比例

隨機選出一個魯里坦尼亞人。以 F、D 表示金髮和深色頭髮，以 G、H 表示綠眼和淡褐眼，以 N、B 表示窄頭和闊頭。已知 $\mathrm{P}(F) = \frac{1}{4}$ 及 $\mathrm{P}(G) = \frac{1}{5}$。

綠眼的各格。由 (v)，$\mathrm{P}(N \mid G) = \mathrm{P}(B \mid G) = \frac{1}{2}$，所以

$$\mathrm{P}(G \cap N) = \mathrm{P}(G \cap B) = \frac{1}{10}.$$

由 (vi)，綠眼的闊頭者平均分為金髮和深色頭髮兩半：

$$\mathrm{P}(F \cap G \cap B) = \mathrm{P}(D \cap G \cap B) = \frac{1}{20}.$$

由 (iv)，$\mathrm{P}(F \cap G \cap N) = 0$，所以 $\mathrm{P}(D \cap G \cap N) = \frac{1}{10}$。

淡褐眼的各格。設

$$a = \mathrm{P}(F \cap H \cap B), \quad b = \mathrm{P}(F \cap H \cap N), \quad c = \mathrm{P}(D \cap H \cap B), \quad d = \mathrm{P}(D \cap H \cap N).$$

- 金髮：綠眼的金髮者佔 $0 + \frac{1}{20}$，所以 $a + b = \frac{1}{4} - \frac{1}{20} = \frac{1}{5}$。
- 淡褐眼：$a + b + c + d = \frac{4}{5}$，所以 $c + d = \frac{3}{5}$。
- 陳述 (vii)：$\mathrm{P}(B \cap D) = \frac{1}{20} + c = \frac{1}{2}$，所以 $c = \frac{9}{20}$，因此 $d = \frac{3}{20}$。
- 陳述 (viii)：$\mathrm{P}(F \cap B \mid H) = \mathrm{P}(D \cap N \mid H)$；兩邊乘以 $\mathrm{P}(H)$ 得 $a = d = \frac{3}{20}$，於是 $b = \frac{1}{20}$。

全部八個比例，以二十分之一為單位，列出如下：

	$G \cap N$	$G \cap B$	$H \cap N$	$H \cap B$
F	0	1	1	3
D	2	1	3	9

它們全部非負，而且加起來是 20，所以這個模型並無矛盾。把窄頭的兩欄相加，

$$\mathrm{P}(N) = \frac{0+2+1+3}{20} = \frac{3}{10}.$$

所以按照這個模型，有 30% 的魯里坦尼亞人是窄頭的。

六個闊頭的相識

假如我這六個相識是一個隨機樣本（random sample），而且彼此獨立（independent），那麼六人全是闊頭的機會是

$$\left(\tfrac{7}{10}\right)^6 = 0.1176\ldots \approx 0.12.$$

想一想 一件概率為 12% 的事件，算得上是否定模型的有力證據嗎？

算不上。一件大約每 8 次就有 1 次發生的事並不罕見；在 5% 這類慣用的顯著性水平（significance level）下，它都不顯著。這些數據還有另一個弱點：一個人的相識並不是隨機樣本。他們可能來自同一家庭、同一地區或同一社交圈子，所以未必彼此獨立，也未必能代表整個人口。這項觀察頂多只提供極輕微的懷疑理由。

一個 200 人的隨機樣本

設 X 為一個 200 人的隨機樣本中窄頭者的數目。人口很大，所以各次抽取實際上是獨立的，而在這個模型之下，$X \sim \mathrm{B}(200, \frac{3}{10})$，且

$$\mathrm{E}(X) = 60, \qquad \mathrm{Var}(X) = 200 \times 0.3 \times 0.7 = 42, \qquad \sqrt{42} \approx 6.481.$$

由於 np 和 $n(1-p)$ 都很大，我們採用正態近似（normal approximation）$X \approx \mathrm{N}(60, 42)$，並加上連續性校正（continuity correction）：

$$\mathrm{P}(X \leq 50) \approx \Phi\left(\frac{50.5-60}{\sqrt{42}}\right) = \Phi(-1.466) \approx 0.071.$$

所以在這個模型之下，窄頭者不多於 50 人的情況大約有 7% 的機會出現。

想一想 我們現在應該否定這個模型嗎？

在 5% 的水平下不應該：以 $p = 0.3$ 對 $p < 0.3$ 作單尾檢驗（one-tailed test），所得的 p 值約為 0.07。不過，這個結果確實令人對模型有些懷疑。樣本比例 $\frac{50}{200} = 0.25$ 比預測值低了約 1.5 個標準差（standard deviation）。它的方向也與我那些相識的情況一致，他們當中的窄頭者同樣比預期少。在相信模型所給的 30% 這個數字之前，多收集一些數據才是明智的做法。（假如檢驗是雙尾的，p 值便約為 0.14，證據就更弱了。）

驗算

- 八個條件每一個都可以直接從表中讀出，例如 (vii)：$1 + 9 = 10$ 個二十分之一 $= \frac{1}{2}$。
- 用電腦核對，可確認各格的比例是唯一解，$\mathrm{P}(N) = \frac{3}{10}$，以及 $(0.7)^6 = 0.117649$。精確的二項概率 $\mathrm{P}(X \leq 50) = 0.0695$ 與正態近似值 0.0713 相當接近。

常見錯誤

- 把 (v) 或 (vi) 讀成關於整個人口的陳述，而不是條件（conditional）陳述：在綠眼的一組之內「同樣可能」（as likely as），意思是把 $\frac{1}{5}$ 平分成兩半。
- 忘記連續性校正。用 50 而不用 50.5 會得到 $\Phi(-1.543) \approx 0.061$。
- 把相識當作隨機樣本而不加任何說明。

STEP 之外

最後一部分是一個假設檢驗（hypothesis test），而那兩項證據引出一個很自然的問題：幾項薄弱的證據應該怎樣合併起來？大學的統計學以似然（likelihood）來回答這個問題（對貝葉斯學派（Bayesian）而言，還有後驗概率（posterior probability）），它們衡量 p 的每一個值對全部數據合起來解釋得有多好。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）提供所需的概率背景，而 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）則對貝葉斯觀點有一段淺白易讀的介紹。

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

1989 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。